Article

Not peer-reviewed version

Unraveling the Genomic Landscape and Stress Responses of Cowpea Thaumatin-Like Proteins: A Comprehensive Analysis

Carolline de Jesús-Pires , José Ribamar Costa Ferreira-Neto * , Roberta Lane de Oliveira-Silva , Jéssica Barboza da Silva , Manassés Daniel da Silva , Antonio Felix da Costa , Ana Maria Benko-Iseppon *

Posted Date: 9 September 2024

doi: 10.20944/preprints202409.0657.v1

Keywords: PR-5 gene family; osmotin-like; TLPs; Vigna unguiculata; zeamatin

 doi:10.20944/preprints202409.0657.v1

Article

Unraveling the Genomic Landscape and Stress Responses of Cowpea Thaumatin-Like Proteins: A Comprehensive Analysis

Carolline de Jesús-Pires [1], José Ribamar Costa Ferreira-Neto [1,*], Roberta Lane de Oliveira Silva [1], Jessica Barboza da Silva [1], Manassés Daniel da Silva [1], Antônio Félix da Costa [2] and Ana Maria Benko-Iseppon [1,*]

[1] Laboratory of Plant Genetics and Biotechnology, Center of Biosciences, Genetics Department, Federal University of Pernambuco, Av. Prof. Moraes Rego, 1235, Recife 50670-901, PE, Brazil
[2] Pernambuco Agronomic Institute, Av. Gen. San Martin, 1371 - Bongi, Recife, 50761-000, PE, Brazil
* Correspondence: ana.iseppon@ufpe.br (A.M. Benko-Iseppon); joseribamar.ferreiraneto@ufpe.br (José Ribamar Costa Ferreira-Neto)

Abstract: Cowpea (*Vigna unguiculata* (L.) Walp.) is an important legume cultivated mainly in regions with limited water availability across the African and American continents. Its productivity is significantly affected by environmental stresses. Thaumatin-like proteins (TLPs), which belong to the PR-5 (Pathogenesis-Related 5) protein family, are known to be responsive to both biotic and abiotic stresses. However, their role remains controversial, with some TLPs associated with plant defense (particularly against fungal infections) and others associated with drought response. In this study, we evaluated the structural diversity and gene expression of TLPs in cowpea (VuTLPs) under different stress conditions, including biotic [mechanical injury followed by inoculation with Cowpea Aphid-borne Mosaic Virus (CABMV) or Cowpea Severe Mosaic Virus (CPSMV)] and abiotic (root dehydration). Genomic anchoring of VuTLPs revealed 34 loci encoding these proteins. Neighbor-joining analysis clustered the VuTLPs into three distinct groups. We identified 15 segmental duplication and six tandem duplication gene pairs, with the majority of VuTLP genes found to be under purifying selection. Promoter analysis associated VuTLPs with bHLH, Dof-type, and MYB-related transcription factors, supporting their diverse roles. Diversity in VuTLP function was also observed in their expression profiles under various stress conditions. Gene expression data showed that most VuTLPs are recruited within the first minutes after stress imposition, except in the root dehydration assay, where most transcripts were induced 150 minutes post-stress, a finding also validated by qPCR. Moreover, the data suggest that VuTLPs exhibit functional specialization depending on the stress condition, highlighting their diverse roles and biotechnological potential.

Keywords: PR-5 gene family; osmotin-like; TLPs; *Vigna unguiculata*; zeamatin

1. Introduction

In the course of the evolution, plants have developed a series of components to fight pathogens. Among several protein families involved in plant defense, we stand out thaumatin-like proteins (TLPs). TLPs are members of the Pathogenesis-related 5 (PR-5) gene family. This protein group received this name because of their structural similarity to proteins called thaumatins, isolated from the fruit of *Thaumatococcus daniellii* Benth. (Maranthaceae family) (van der Wel and Loeve, 1972). The mentioned group share the Thaumatin conserved domain (Pfam: PF00314) that covers almost 95% of the entire mature peptide (Petre et al. 2011). About their potential biological functions, TLPs have shown antifungal activity (Li et al. 2023; Wang et al. 2023; Yang et al. 2024; Liu et al. 2021), acting on fungal membrane permeabilization (Batalia et al. 1996). There is evidence that some TLPs representatives act on invading fungi through their binding activities or hydrolysis of β-1,3-glucans (Grenier et al. 1999; Trudel et al. 1998) or, still, through inhibition of enzymes such as xylanases (Fierens et al. 2007).

 doi:10.20944/preprints202409.0657.v1

Hormones associated with plant immunity can also induce TLPs-encoding genes. For example, in tobacco seedlings, the induction of TLPs genes occurred in response to the application of ethylene (ET), methyl jasmonate (methylJA) and salicylic acid (SA) (Xu et al. 1994). A similar fact was also reported for transgenic rice, where the rice blast fungus elicitor, besides the hormones SA and methylJA, induced the β-glucuronidase (GUS) reporter gene driven by the Rtlp1 (Rice thaumatin-like protein 1) promoter (pRtlp1GUS) (Hiroyuki et al. 2008). TLPs expression are also associated with abiotic responses, as drought (Zhao et al. 2024; Muoki et al. 2021; Jung et al. 2005), wounding (Ruperti et al. 2002), and freezing (Liu et al. 2023; Hon et al. 1995). There are, additionally, reports that TLPs were induced after infection by virus or bacteria (Breiteneder and Radauer, 2004); however, in these situations, their action mechanism remains indeterminate. Regarding viruses, the work of Kim et al. (2005) observed that a tobacco (*Nicotiana tabacum*), TLP1, interacted with *Cucumber mosaic virus* proteins.

Due to their involvement in the response to pathogens and biotechnological potential, TLPs have been analyzed in a large number of species, such as *Arabidopsis thaliana*, rice (*Oryza sativa*), *Populus* spp., maize (*Zea mays*), *Physcomitrella patens*, *Chlamydomonas* spp., and wheat (*Triticum aestivum*) (Muoki et al. 2021; Sharma et al. 2022; Cao et al. 2016; Wang et al. 2010), among others. About legumes, information on this protein class is still quite limited (Feng et al. 2024). The vast majority of these studies, however, does not explicitly focus on global analysis of TLPs; only quote the some TLPs participation in the global transcriptional response, or focus on point TLPs. The only report found in the literature addressing the structural and functional genomics of TLPs was performed by Petre et al. (2011). These authors identified 42 validated TLP-loci in the genome *Populus trichocarpa* 'Nisqually-1' (version 2.0). In addition, it was observed that most TLPs were responsive to abiotic stresses (high ozone content, UV-B rays, drought, and high copper content) besides being also responsive to biotic stress (infection by the fungus *Melampsora laricis-populina*).

Cowpea is an important legume. Its social role is important: such crop is a relevant source of proteins and minerals for millions of people in sub-Saharan Africa and other developing regions (Boukar et al., 2016). Cowpea demonstrates exceptional phenotypic plasticity and strong physiological performance under a wide array of biotic and abiotic stresses. Given its importance and adaptive capacity, considerable efforts in omics research have been focused on this species. Recently, its reference genome was released (Lonardi et al., 2019), providing a vital resource for elucidating the crop's physiological advantages under stress conditions, identifying genes with biotechnological potential, and serving as a key platform for comparative genomics within legumes. In Brazil, which contributes 12% of global cowpea production, the Cowpea Genomics Consortium (CpGC) was established. The CpGC manages sequenced genomes from multiple accessions and cultivars of this species, which are in various stages of assembly. In addition, it holds comprehensive transcriptomes (RNA-Seq) from cultivars demonstrating tolerance or resistance, which have been subjected to root dehydration or combined stressors (mechanical injury followed by viral inoculation with either Cowpea Aphid-borne Mosaic Virus or Cowpea Mosaic Severe Virus).

Thus, given the importance, multifunctionality, the responsiveness of plant TLPs under unfavorable conditions and the availability of CpGC data, the present work brings structural characterization and gene expression data of cowpea TLPs expressed under different stress types (mechanical injury, CABMV or CPSMV inoculation and root dehydration). The present work uncovered relevant information to the molecular physiology of cowpea under unfavorable conditions and provide data about the performance of TLPs in a legume, an issue still little addressed.

2. Materials and Methods

2.1. Biological Material, Experimental Design and Stress Application

- Root dehydration assay

Seeds of *Vigna ungiculata* cv. Pingo de Ouro (considered tolerant to water deficit and drought; Rodrigues *et al*., 2017; Bastos *et al*., 2011) were treated with 0.05 % (w/v) Thiram (tetramethylthiuram disulfide) and germinated during 2 days at 25° C ± 1 °C and 65% ± 5% of temperature and relative humidity, respectively. Seedlings were transferred to a hydroponic system (Rodrigues *et al*., 2012)

 doi:10.20944/preprints202409.0657.v1

(**Supplementary Figure 1A**) with aerated pH 6.6 balanced nutrient solution (Hoagland and Arnon, 1950) in a randomized block experimental design, with three biological replicates (**Supplementary Figure 1B**). Each biological replicate was composed of two individuals. Plantlets were placed in supports in such a way that the roots of the seedlings were completely immersed in the solution (**Supplementary Figure 1A**). Plantlets were grown for three weeks (estágio de desenvolvimento V3) in greenhouse under a natural photoperiod of approximately 13/11 h light/dark cycle, temperature of 30 ± 5 °C and 60 ± 10 % relative humidity (RH). After this period, root dehydration treatment was initiated by withdrawing the nutrient solution from treated plants (**Supplementary Figure 1A)**. Roots were collected after 25 min (RD25) and 150 min (RD150) after solution removing (**Supplementary Figure 1B**). The tissue was immediately frozen in liquid nitrogen and stored at -80 °C until RNA extraction. For each treatment, the respective control plants (Cont.25' and Cont.150'; **Supplementary Figure 1B**) were maintained in the nutrient solution and subsequently collected.

Throughout this manuscript, the expression contrast for 'RD25 vs Cont.25' was referred to as T25; similarly, the contrast for 'RD150 vs Cont.150' was referred to as T150.

❖ Mechanical injury and virus inoculation assays

The experiments involving mechanical injury and inoculation by CABMV (Cowpea Aphid-borne Mosaic Virus) or CPSMV (Cowpea Severe Mosaic Virus) were conducted under controlled conditions in a greenhouse at the Instituto Agronômico de Pernambuco (IPA; Recife, Pernambuco, Brazil) (**Supplementary Figure 2A**). The CABMV assay utilized the resistant genotype IT85F-2687 (Rocha et al., 1996; Oliveira et al., 2011), while the CPSMV assay employed the resistant genotype BR-14 Mulato (Cardoso et al., 1990).

The experimental procedures for both assays were conducted separately but followed similar protocols. Both accessions were sown and grown for three weeks, reaching the V3 development stage, under natural photoperiods and temperatures ranging from 28 to 32°C (**Supplementary Figure 2A**). The leaves of the youngest trifoliate were mechanically injured using carborundum (silicon carbide) to facilitate viral entry into the plant tissue, followed by the application of the viral inoculum (Supplementary Figure 2A).

Two post-injury/inoculation time points were implemented for each assay (Supplementary Figure 2B): 60 minutes and 16 hours. Each treatment had its corresponding absolute control or mock (**Supplementary Figure 2B**). The tissues were immediately frozen in liquid nitrogen and stored at -80°C until RNA extraction.

The experimental design was factorial, based on cultivar and post-inoculation period, with three biological replicates (BRs) for each control and treatment (**Supplementary Figure 2B**). Each BR consisted of five plants. All treatments were carried out in isolated areas to avoid cross-contamination through volatile compounds released by plants.

Differential gene expression in the two virus-related assays was induced by a combination of stress factors: mechanical injury and viral inoculation. Plant viruses cannot initiate infection independently - they require a vector or specific agricultural practices to breach the plant cell wall. According to Barna and Király (2004), plant viruses lack specific cellular receptors, unlike bacteriophages and animal viruses. Therefore, the combination of 'mechanical injury and viral inoculation' was employed to mimic natural infection processes.

Throughout this study, the expression contrasts for the mechanical injury (MI) and viral inoculation (CABMV or CPSMV) assays were defined as follows:

- MI+CABMV(60min) vs. control: MI_CABMV60'
- MI+CABMV(16h) vs. control: MI_CABMV16h
- MI+CPSMV(60min) vs. control: MI_CPSMV60'
- MI+CPSMV(16h) vs. control: MI_CPSMV16h

2.2. RNA Extraction and cDNA Synthesis

Total RNA was isolated using the "SV Total RNA Isolation System" kit (Promega, US) following manufacturer's protocol. Agarose gel (1.5 %) and Agilent 2100 Bioanalyzer (Agilent Technologies,

 doi:10.20944/preprints202409.0657.v1

EUA) evaluated both concentration and quality of total extracted RNA. Only samples with RNA integrity number (RIN) ≥ 8.0 were sequenced.

A "RNAm TruSeq® Stranded LT-Set A" kit (RS-122-2101) (Illumina, San Diego, CA, EUA) was employed in messenger RNA purification and cDNA library construction according to the manufacturer's instructions. Paired-end reads, with 100 pb in length were generated via the Illumina HiSeq 2500 system, using the following kits: "HiSeq® Rapid PE Cluster Kit v2" (PE-402-4002); "SBS Kit v2" (200 Cycle; FC-402-4021); and "TruSeq® Stranded mRNA LT - Set A" (RS-122-2101). All sequencing steps were performed at the Center for Functional Genomics, University of São Paulo (São Paulo, Brazil).

2.3. RNA-Seq Libraries Assembly and Differential Expression Analysis

The 12 RNA-Seq libraries sequenced for the root dehydration assay were combined with the 12 libraries from the "mechanical injury and viral inoculation by Cowpea Aphid-borne Mosaic Virus (CABMV)" experiment and the 12 libraries from the "mechanical injury and inoculation by Cowpea Severe Mosaic Virus (CPSMV)" experiment. This combined assembly facilitated the generation of longer and more robust transcripts. A total of 72 replicates (12 biological replicates for root dehydration, 12 for CABMV, and 12 for CPSMV, each with two technical replicates) were processed through the pipeline, briefly outlined below.

Initially, the Trimmomatic v0.36 package was used to remove adapters and low-quality sequences (Bolger et al., 2014). Reads were trimmed from the 3' end to ensure a Phred score of at least 30. Illumina sequencing adapters were removed, and only reads with a minimum length of 32 bp were retained. Subsequently, de novo transcriptome assembly was performed using the Trinity software (Grabherr et al., 2011), as described by Haas et al. (2013). The quality of the assembled transcriptomes was assessed by evaluating the N50 of the resulting libraries, along with the count and distribution of contig lengths.

Notably, differential expression analysis was conducted independently for each assay (root dehydration; mechanical injury + CABMV inoculation; mechanical injury + CPSMV inoculation). This analysis was performed using the edgeR tool (Robinson et al., 2010) implemented within the Bioconductor package (Huber et al., 2015). Transcripts with Log2FC values between -1 and 1, $p < 0.05$, and FDR < 0.05 were considered differentially expressed.

Hierarchical clustering of VuTLPs was carried out with CLUSTER 3.0 (Eisen et al. 1998), and the *heatmap* was visualized using TreeView (Page, 1996). In addition, Venn diagrams were also generated using software Venny (Oliveros, 2007).

2.4. Cowpea TLPs (VuTLPs) in Silico Characterization

The in silico analysis was carried out using reads of cowpea RNA-Seq libraries under CABMV or CPSMV inoculation or under root dehydration (**Supplementary Table S1**). VuTLPs were predicted based on sequence homology searches, using thaumatin-domain with sequences previously characterized (see Jesús-Pires et al. 2019). An outline of the annotation strategy is presented in Fig. 1. For the identification of VuTLPs candidates, the alignments were carried out against CpGC database using tBLASTn (Altschul et al. 1990) (**Figure 1**) with a cut-off value of $1e^{-05}$. VuTLP candidates were annotated against NCBI and UniProt and analyzed for their score, e-values, sequence size, and presence of Conserved Thaumatin Domains (CTDs), as shown in **Supplementary Table S2**.

All sequences were screened for conserved motifs, with the aid of rps-BLAST on CD-search tool (Altschul et al. 1990) (**Figure 1**). Only orthologs presenting the expected features (CTDs and motifs) were translated into proteins and considered for subsequent evaluation (**Figure 1**).

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 September 2024 doi:10.20944/preprints202409.0657.v1

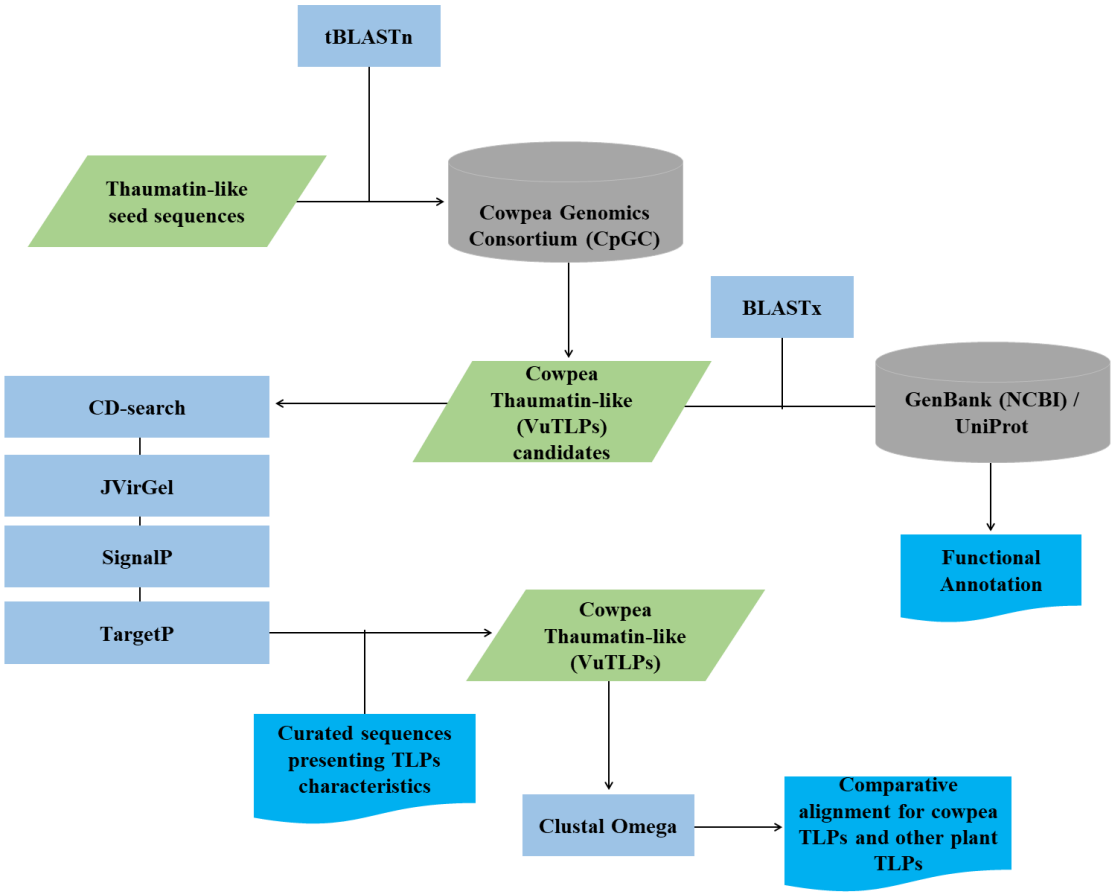

Figure 1. Workflow used for in silico identification and prediction of VuTLPs applied to the cowpea transcriptome.

Aiming to enrich the characterization of predicted VuTLPs, the identified amino acid sequences were further characterized including the following analyzes (**Figure 1**):

1. Determination of the putative isoelectric point and molecular weight using ProtParam (https://web.expasy.org/protparam/; (Gasteiger et al. 2005); **Figure 1**);
2. Prediction of signal peptide for each cowpea TLP candidate with SignalP 4.1 Server (Petersen et al. 2011; http://www.cbs.dtu.dk/services/ SignalP/);
3. Prediction of the subcellular localization with TargetP 1.1 Server (Emanuelsson et al. 2007; http://www.cbs.dtu.dk/services/TargetP).

Additionally, we used Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/; Sievers et al. 2011) to generate a multiple sequence alignment of full-length VuTLPs compared with previously described TLPs from other plant species, aiming at the localization and comparison of CTDs and residues among TLP sequences. For this step, some homolog proteins denominated "osmotin" and "zeamatin" were also employed (**Supplementary Figure 3**; **Supplementary Material SM1**), considering literature evidences (Singh et al. 1989; Roberts and Selitrennikoff, 1990; Skadsen et al. 2000) that such proteins present similarities with TLPs.

2.5. Distribution of VuTLP-Candidates in V. unguiculata Genome

To determine the genomic distribution of VuTLPs, cowpea candidates were aligned against *V. unguiculata* genome V12 available at Phytozome Database (https://phytozome.jgi.doe.gov/) aiming to map these sequences in virtual chromosomes through the BLASTp tool. This step aimed to infer on the distribution, relative position, and abundance of TLPs-coding loci. We considered the best-hit (e-value cut off < $1e^{-10}$) to allow the identification of VuTLPs along the virtual chromosomes. Afterward,

 doi:10.20944/preprints202409.0657.v1

the identified mapping positions were used to build a virtual ideogram (**Figure 2**) regarding the TLPs distribution on *V. unguiculata* pseudochromosomes (n = 11).

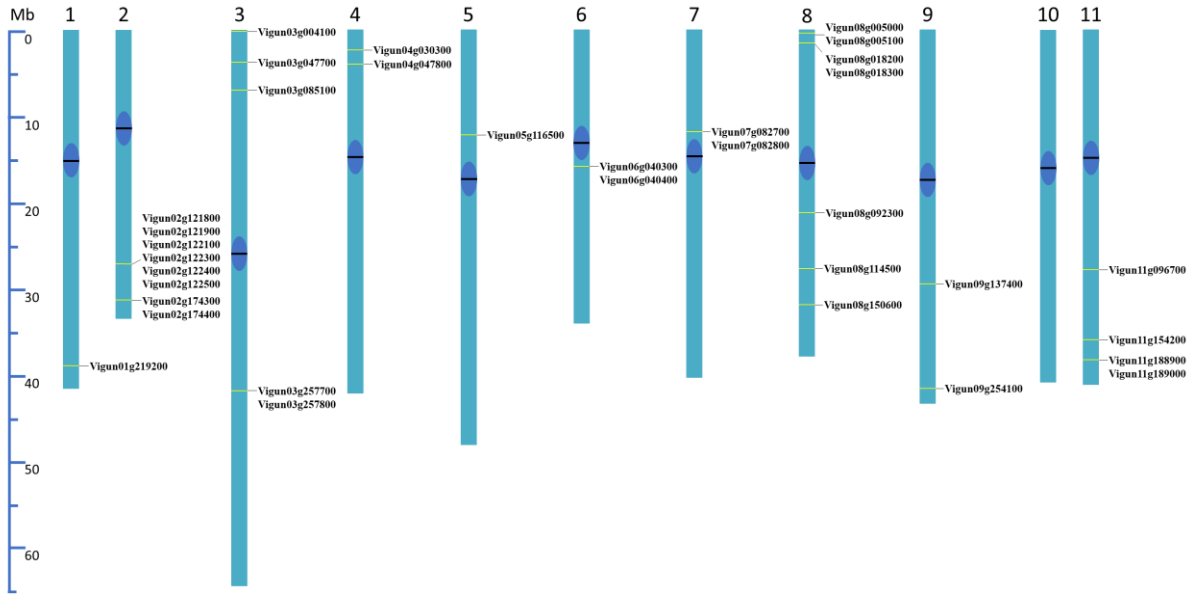

Figure 2. Distribution of cowpea thaumatin-like genes in *Vigna unguiculata* virtual chromosomes (Chr1 to Chr11). Positions of TLP mapped genes or clusters are indicated in yellow. Centromeres are indicated with black bars and dark blue spheres.

2.6. NJ Analysis and VuTLPs Gene Features

For the Neighbor-Joining (NJ) analysis, used 34 TLP-coding genes. The multiple alignment was submitted to manual edition and elimination of not aligned extremities and uninformative autapomorphies. The NJ analysis was carried out using the MEGA 7 (Kumar et al. 2016). Statistical support of the branches was evaluated by a bootstrap analysis with 2,000 replicates.

The gene structure was examined in terms of exon-intron organization of each VuTLP. The display of exon-intron was performed using GSDS 2.0 server (Hu et al. 2015). The VuTLPs gene features também foram determinadas via Genestats script.

The following parameters were scrutinized:

1. transcript sequence length; 2. number of exons; 3. total exon sequence length; 4. number of introns; 5. total intron sequence length; 6. number of CDS chunks; 7. total CDS sequence length; 8. number of 5' UTR sequences; 9. total 5' UTR sequence length; 10. number of 3' UTR sequences; 11. total 3' UTR sequence length.

2.7. Analysis of VuTLPs Duplication

Multiple Collinearity Scan toolkit (MCScanX; Wang *et al*., 2012) package - downstream analysis mode: 'duplicate_gene_classifier' - was applied to classify the origins of the duplicate VuTLPs genes of the cowpea genome, into:

1. Whole genome /segmental (i.e., collinear genes in collinear blocks);
2. Tandem (adjacent loci in a genome region);
3. Proximal (gene pairs in nearby chromosomal region but not adjacent);
4. Dispersed (other modes than segmental, tandem and proximal) duplications.

The following procedure was used by MCScanX to assign the duplication mechanisms:

a) All genes were initially classified as 'singletons' (i.e., without duplicates in the genome) and assigned gene ranks according to their order of appearance along chromosomes;
b) BLASTp results were evaluated and the genes with BLASTp hits to other genes were re-labeled as 'dispersed duplicates';
c) In any BLASTp hit, the two genes were re-labeled as 'proximal duplicates' if they had a difference of gene rank < 20 (configurable);

 doi:10.20944/preprints202409.0657.v1

d) In any BLASTp hit, the two genes were re-labeled as 'tandem duplicates' if they had a difference of gene rank = 1;
e) MCScanX was executed. The anchor genes in collinear blocks were re-labeled as 'WGD/segmental'.
f) So, if a gene appeared in multiple BLASTp hits, it was assigned a unique class according to the order of priority: whole genome /segmental > tandem > proximal > dispersed.

2.8. Ratio of Synonymous (Ks) and Non-Synonymous (Ka) Substitutions per Site for Tandem Duplicated Genes

The ratio of Ka to Ks (Ka/Ks) was used to determine the selection pressure among tandem and segmental duplication events. ClustalW 2.0 (Larkin *et al*., 2007) software first aligned the full-length coding sequences of tandemly and segmentaly duplicated VuTLPs genes. Then, the Ka and Ks rates were calculated with the standard genetic code table by the Nei–Gojobori method (Jukes-Cantor model) implemented on MEGA 7 (Kumar *et al*., 2016).

The Ka/Ks ratio is an indicator of the selection history on genes or gene regions:

- if its value is lower than '1', the duplicated gene pairs may have evolved from purifying selection (also called as negative selection; conserves the amino acid sequence);
- If Ka/Ks equal '1', means neutral selection (that had no constraint for sequence divergence);
- while Ka/Ks greater than '1' means positive selection (that led to different peptides).

2.9. Promoter Analysis

Cowpea promoter regions (up to 1.0 kb upstream of TLP-coding genes) were downloaded from Phytozome database v8.0 (https://phytozome.jgi.doe.gov). The motifs (candidate cis-regulatory elements) in each promoter were revealed through the MEME v5.0.3 software (Bailey and Elkan, 1994) (http://meme-suite.org/tools/meme). The software reports an e-value for each identified motif and gives an estimate of the number of expected motifs found by chance if the input sequences were shuffled. In the present simulation, an e-value < 10^{-2} was adopted as the cut-off for characterization of *bona fide* cis-regulatory element candidates. The maximum number of motifs searched in the present analysis was 10. The motifs analyzed exhibited between six and 50 nucleotides in length.

After MEME software analysis, the Tomtom v4.11.2 software (Gupta et al. 2007) (http://meme-suite.org/tools/tomtom) was used with the JASPAR database (file: JASPAR2018_CORE_plants_non-redundant) to annotate putative cis-regulatory elements. JASPAR searches one or more queries (candidate cis-regulatory element) against annotated motifs ranked by p-value (cut-off < 10^{-2}). The q-value (false discovery rate; cut-off < 10^{-2}) of each match is also reported. In this work, the presented identities of the cis-regulatory element candidates were associated with the target motif that exhibited the most significant p-value.

2.10. qPCR Analysis

For validation of the RNA-Seq data of the different experiments, RT-qPCR analyses were performed with treatments and control for each time sampled. For each time point three biological replicates, each with three technical replicates, were analyzed to ensure statistical reliability. The qPCR reactions were performed on the CFX96 Touch Real-Time PCR Detection System (Bio-Rad), using SYBR Green detection. The PCR program was adjusted to 95 °C for 2 min, followed by 35 cycles of 95 °C for 15 s, 60 °C for 1 min, and 72 °C for 15 s. After amplification, dissociation curves were produced (65–95 °C at a heating rate of 0.5 °C / sec and acquiring fluorescence data every 0.3 °C) to confirm the specificity of the PCR products.

The amplification efficiency ($E = 10^{(-1/\text{slope of the standard curve})}$ for all primer pairs was

determined from a 4-point standard curve generated by serial dilutions of cDNA (10-fold each) in technical triplicates. *β-tubulin-2*, *F-box* and *UBQ10* were used as reference genes for normalization of MI_CABMV and MI_CPSMV assays while *Actin* and *UE21D* were used as reference genes for normalization of radicular dehydration assay (Amorim et al. 2018). Primers for VuTLPs candidates

 doi:10.20944/preprints202409.0657.v1

were designed based on the *V. unguiculata* transcriptome using the Primer3 tool (http://bioinfo.ut.ee/primer3-0.4.0/) under program´s default settings (**Supplementary Table S3)**.

The Rest2009 (Pfaffl et al. 2009) software package (standard mode) was used for relative expression analysis of target transcripts. Hypothesis testing ($p < 0.05$) was used to determine whether the differences in target transcripts expression between the control and treatment conditions were significant.

3. Results

3.1. In Silico Identification and Characterization of VuTLPs

The tBLASTn search allowed the identification of 126 putative VuTLPs (Suppl. Tab. S2) expressed under biotic and abiotic stresses. From these 33 presented no ORF associated with TLP, while 32 presented complete TLP conserved domain (CTD) and 61 presented incomplete CTDs. Among candidate VuTLPs with complete CTD, 30 were expressed in cowpea RNA-Seq libraries generated under radicular dehydration, while 28 and 27 VuTLPs were expressed under CABMV and CPSMV inoculation, respectively. Considering the total amount (abiotic + biotic stresses), 30 non-redundant VuTLPs were recognized (**Supplementary Material SM2**).

All 30 predicted VuTLPs presented CTD in at least one searched database (NCBI CDD: cd09218 and cl02511, Pfam: pfam00314, SMART ID: smart00205; **Supplementary Table S2**). Five highly conserved amino acids (REDDD motif), in an acidic cleft, were detected in most VuTLPs. However, the alignment of the referred candidates with other TLPs from *A. thaliana*, *Nicotiana tabacum*, *Zea mays*, and *Cicer arietinum* revealed that some VuTLPs did not contain all expected amino acids in REDDD motif. For this group, some VuTLPs exhibited the following substitutions (**Supplementary Figure 3**): the glutamic acid (E) by a glutamine (Q) in the transcript Vu54351|c0_g1_i1 and in the transcripts Vu85277|c0_g1_i1 and Vu28453|c0_g1_i1 the glutamic acid (E) was replaced by aspartic acid (D) of the same functional group; and the first aspartic acid (D) by a glycine (G) in the transcript Vu54351|c0_g1_i1. Other modifications in the mentioned motif were also perceived. For instance, in the transcripts Vu85277|c0_g1_i1 and Vu28453|c0_g1_i1 (**Supplementary Figure 3**) the second aspartic acid (D) was replaced by an asparagine (N).

Regarding the 16-cysteine conservation, just three cowpea sequences (Vu58024|c0_g1_i1, Vu49480|c1_g2_i2 and Vu49480|c1_g2_i1; **Supplementary Figure 3**) lacked the first cysteine; whereas the sixth cysteine of transcript Vu83475|c0_g1_i1 was replaced by a serine (S). In turn, all other VuTLPs presented the 16 residues. Additionally, expressed VuTLPs exhibited a molecular weight between 18.97 and 34.93 kDa and an isoelectric point varying from 4.23 to 9.05 (**Supplementary Table S2**).

The SignalP prediction, in turn, revealed that 20 VuTLPs present hydrophobic signal peptide sequences.

3.2. Mapping of VuTLPs in V. unguiculata Genome and Analysis of Gene Duplication

The approach revealed the distribution of the VuTLPs in 34 loci located along the chromosomes, with exception chromosome ten (**Figure 2**).

A clustering of VuTLPs loci was observed, being more pronounced on chromosomes two and eight (**Figure 2**; **Supplementary Table S4**). For example, the long arm of *V. unguiculata* chromosome 2 presents six and two clustered TLP coding genes (**Figure 2**). Such a distribution is indicative of evolutionary processes through tandem duplication. Besides the observed clustering in most chromosomes, the position of VuTLP coding genes was mainly subterminal or intercalary, while a pericentric position of VuTLPs is observed only in chromosomes 6 and 7 (**Figure 2**).

Gene duplications, including segmental and tandem duplication, have been considered as one of the main forces in the evolution and expansion of a gene Family (Lynch and Conery, 2000). The analysis of the expansion of TLP-coding genes revealed that the 34 TLP-coding genes have expanded through gene duplication processes (**Supplementary Table S5**). In this study, we found four types of gene duplications: dispersed, proximal, tandem, and WGD or segmental duplication by MCScanX program. Our analysis provides evidence that whole-genome duplication (WGD) or segmental

 doi:10.20944/preprints202409.0657.v1

duplication has been the primary mechanism driving the expansion of VuTLPs in the cowpea genome, resulting in a total of 15 gene pairs (**Supplementary Table S6**). Moreover, VuTLP genes were observed to cluster within six tandem duplication regions across four of the 11 chromosomes (**Supplementary Table S7**). Finally, we identified five dispersed duplication events and one proximal duplication event contributing to the expansion of these genes (**Supplementary Table S5**).

3.3. Ratio of Synonymous and Non-Synonymous Substitutions for Tandem Duplicated Genes

The forces that drive natural selection can be inferred from the types of nucleotide substitutions in coding sequence of the genes. An informative parameter of the evolution of genes under selection has been the ratio of the rate of nonsynonymous substitution (Ka; that causes an amino acid change) to the rate of synonymous substitution (Ks; that not causes an amino acid change). Os valores de Ka/Ks variaram de 0,00 a 1,36 for tandem duplication (**Supplementary Table S7)** e de 0,00 a 0,31 for segmental duplication (**Supplementary Table S6**). Sixteen pairs (16 duets) presented a Ka/Ks ratio of less than 1 (Supplementary Tables S6-S7), suggesting that these genes have been subjected to purifying selection. Em outro contexto, observou-se que cinco duetos apresentaram Ka/Ks > 1 (**Supplementary Table S6-S7)**, implying that positive selection played an important role in the evolution desses duetos.

3.4. NJ Analysis and Exon-Intron Organization

A Neighbor-Joining tree was constructed using the protein sequences of all the 34 VuTLP genes. The cowpea VuTLP family was divided into three groups (**Supplementary Figure 4**). The Group III colored in purple (58.82%) contained the most members, followed by Group II colored in pink (27.27%) and the least represented group was the I colored in green (5.88%) with only two VuTLPs (VuTLP16 and VuTLP32). To understand the formation of gene groups, several characteristics were analyzed, including the presence of a signal peptide, molecular weight, isoelectric point, presence of REDDD residues, and subcellular localization (Table 1). Group I comprised genes with molecular weights ranging from 24.16 to 25.18 kDa and isoelectric points from 6.80 to 7.32. Notably, in the gene VuTLP32, alterations were observed in the REDDD residues, where E was replaced by Q, and the first D by G. Group II included genes with molecular weights ranging from 21.48 to 29.20 kDa and isoelectric points from 4.04 to 8.76. In the gene VuTLP12, a change in the REDDD residues was noted, with the second D replaced by N. Group III consisted of genes with molecular weights from 13.75 to 108.18 kDa and isoelectric points from 4.00 to 8.19. In the gene VuTLP6, the REDDD residue showed an alteration where R was replaced by Q. Additionally, in genes VuTLP7, VuTLP5, and VuTLP18, the first D residue was substituted by Y, and in VuTLP18, the second D was replaced by Q. Moreover, in this group, two genes (VuTLP5 and VuTLP18) were localized to the plasma membrane.

Besides, the most of thirty-four VuTLP genes shown in gene structure analysis had three exons, while VuTLP7, VuTLP31, VuTLP20, VuTLP21, VuTLP29, VuTLP1, VuTLP17, VuTLP15, VuTLP24, VuTLP25 and VuTLP16 had two and VuTLP4, VuTLP19, VuTLP9, VuTLP11, VuTLP8, VuTLP10 and VuTLP32 had one exon. Four genes (VuTLP11, VuTLP8, VuTLP10 and VuTLP32) had no introns (**Supplementary Figure 4**).

Table 1.

Gene	Loci	Duplication pattern	Signal peptide	MW (kDa)	pI	REDDD motif	Subcelular location
VuTLP1	Vigun01g219200	WGD or Segmental	YES	23.27	7.32	R E D D D	Extracelular
VuTLP2	Vigun02g121800	WGD or Segmental	YES	30.75	4.35	R E D D D	Extracelular
VuTLP3	Vigun02g121900	WGD or Segmental	YES	28.48	4.05	R E D D D	Extracelular

 doi:10.20944/preprints202409.0657.v1

VuTLP4	Vigun02g122100	Proximal	YES	21.57	5.66	R	E	D	D	D	Extracelular
VuTLP5	Vigun02g122300	Tandem	NO	31.03	7.31	R	E	Y	E	D	Membrana Plasmática
VuTLP6	Vigun02g122400	Tandem	YES	28.11	8.19	Q	E	D	E	D	Extracelular
VuTLP7	Vigun02g122500	Tandem	YES	21.36	6.86	R	E	Y	E	D	Extracelular
VuTLP8	Vigun02g174300	WGD or Segmental	YES	21.56	4.04	R	E	D	D	D	Extracelular
VuTLP9	Vigun02g174400	Tandem	YES	24.06	6.74	R	E	D	D	D	Extracelular
VuTLP10	Vigun03g004100	WGD or Segmental	YES	23.80	8.76	R	E	D	D	D	Extracelular
VuTLP11	Vigun03g047700	WGD or Segmental	YES	21.48	4.29	R	E	D	D	D	Extracelular
VuTLP12	Vigun03g085100	Dispersed	YES	28.55	6.06	R	E	D	N	D	Extracelular
VuTLP13	Vigun03g257700	WGD or Segmental	YES	28.72	4.19	R	E	D	D	D	Extracelular
VuTLP14	Vigun03g257800	WGD or Segmental	YES	27.88	5.71	R	E	D	D	D	Extracelular
VuTLP15	Vigun04g030300	WGD or Segmental	YES	24.89	8.40	R	E	D	D	D	Extracelular
VuTLP16	Vigun04g047800	Dispersed	YES	25.18	7.32	R	E	D	D	D	Extracelular
VuTLP17	Vigun05g116500	WGD or Segmental	YES	23.47	7.32	R	E	D	D	D	Extracelular
VuTLP18	Vigun06g040300	Tandem	YES	108.18	5.22	R	E	Y	Q	-	Membrana Plasmática
VuTLP19	Vigun06g040400	Tandem	NO	13.75	5.46	R	E	D	D	-	Extracelular
VuTLP20	Vigun07g082700	Tandem	NO	27.34	4.77	R	E	D	D	D	Extracelular
VuTLP21	Vigun07g082800	Tandem	NO	27.34	4.77	R	E	D	D	D	Extracelular
VuTLP22	Vigun08g005000	WGD or Segmental	YES	30.35	4.21	R	E	D	D	D	Extracelular
VuTLP23	Vigun08g005100	Tandem	YES	30.42	4.24	R	E	D	D	D	Extracelular
VuTLP24	Vigun08g018200	Tandem	NO	29.20	4.65	R	E	D	D	D	Extracelular
VuTLP25	Vigun08g018300	Tandem	YES	22.77	4.30	R	E	D	D	D	Extracelular
VuTLP26	Vigun08g092300	Dispersed	YES	28.05	8.36	R	E	D	D	D	Extracelular
VuTLP27	Vigun08g114500	WGD or Segmental	NO	33.67	4.93	R	E	D	D	D	Extracelular
VuTLP28	Vigun08g150600	WGD or Segmental	YES	23.48	7.52	R	E	D	D	D	Extracelular
VuTLP29	Vigun09g137400	WGD or Segmental	YES	29.89	7.32	R	E	D	D	D	Extracelular
VuTLP30	Vigun09g254100	WGD or Segmental	YES	31.53	4.86	R	E	D	D	D	Extracelular
VuTLP31	Vigun11g096700	Dispersed	YES	23.26	4.11	R	E	D	D	D	Extracelular
VuTLP32	Vigun11g154200	Dispersed	YES	24.16	6.80	R	Q	G	D	D	Extracelular
VuTLP33	Vigun11g188900	WGD or Segmental	YES	33.65	4.00	R	E	D	D	D	Extracelular

 doi:10.20944/preprints202409.0657.v1

VuTLP34	Vigun11g189000	WGD or Segmental	YES	30.44	8.15	R	E	D	D	D	Extracelular

3.5. Promoter VuTLPs

The 34 cowpea TLP-coding genes had their predicted promoters (1 kb) analyzed regarding the presence and identity of candidate cis-regulatory elements (CCREs), via the MEME and Tomtom tools. Four of the identified promoter motifs were within the stipulated cut-off (e-value < 10^{-2}, **Figure 3**; **Supplementary Figure 5A**, **5B**, **5C** e **5D**), being considered as *bona fide* CCREs, (see **Supplementary Figure 5A**, **5B**, **5C** e **5D**). The six remaining motifs (**Figure 3**) were associated to TFs [Dof-type (dark green and pink boxes), C2H2 (light green), G2-like (orange), and CPP (light blue), and one without specific annotation] but did not achieve the specified parameters, being excluded from the discussion.

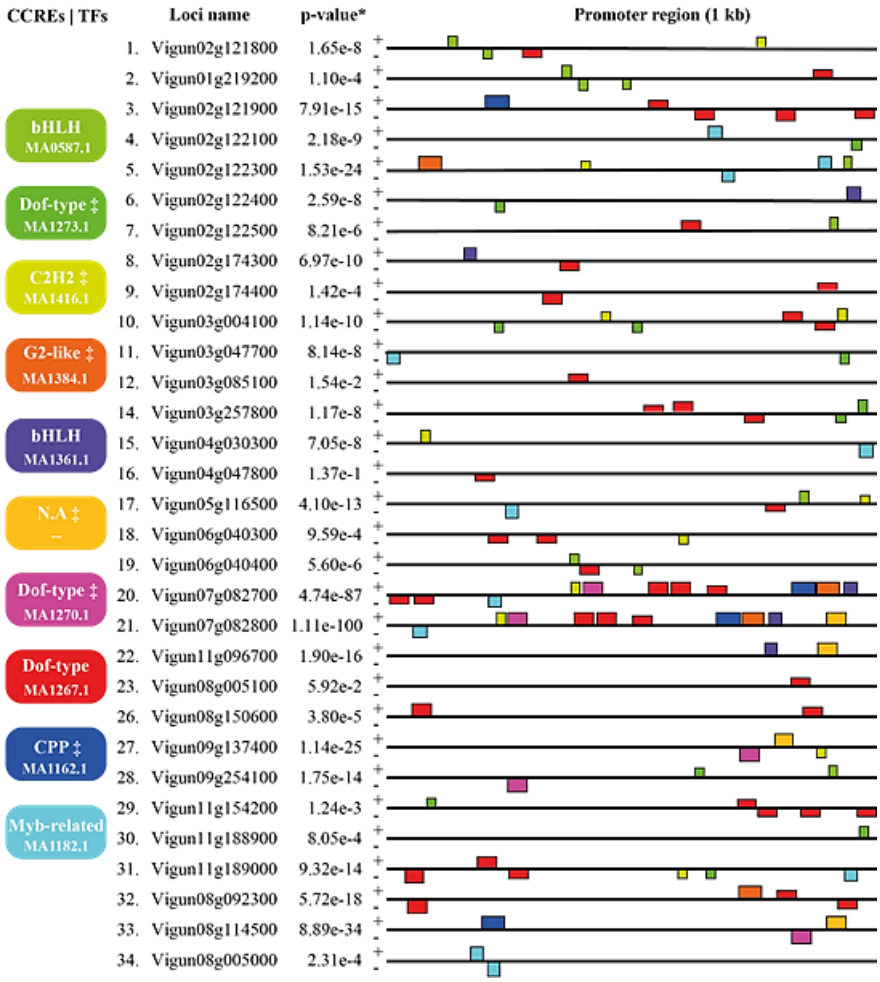

Figure 3. Motif analysis of cis-regulatory elements (CCREs) detected in promoter regions of thaumatin-like protein coding genes. Promoter content and distribution represented by colored rectangles and squares. Colored boxes (CCREs | TFs section) bring information on transcription factors (TFs) associated with the identified motifs, as well as their respective JASPAR ID Matrix. "+" and "-" signals represent sense and antisense strands of the promoter regions analyzed. *MEME combined match p-value. ‡ Motifs with statistical significance below of the adopted cut-off (e-value < 10^{-2}; data not shown). N.A (not annotated motif).

The identified *bona fide* CCREs are associated with the three TF families (**Figure 3**): bHLH (two motifs; light green and blue boxes, **Supplementary Figure 5A** and **5C**), MYB-related (light blue, **Figure 3**; **Supplementary Figure 5B**) and Dof-type (red, **Figure 3**; **Supplementary Figure 5D**). CEMR associated with a Dof-type transcription factor (JASPAR ID MA1267.1; **Figure 3**; **Supplementary Figure 5D**) was the most abundant promoter, observed in 42 sites, followed by MYB-related (JASPAR

 doi:10.20944/preprints202409.0657.v1

ID MA1182.1; **Figure 3**; **Supplementary Figure 5B**) and bHLH (JASPAR ID MA0587.1; **Figure 3**; **Supplementary Figure 5C**) each anchored in 11 sites. Finally, bHLH (JASPAR ID 1361.1; **Figure 3**; **Supplementary Figure 5A**) anchored in five sites.

3.6. TLP Content and Expression in Cowpea Transcriptomes under Different Stress Types

Considering all available RNA-Seq libraries for biotic and abiotic stresses (**Supplementary Table S8**), 30 VuTLPs were identified. A partir daqui iremos dividir os ensaios para melhor analisar a regulação das VuTLPs.

3.6.1. Mechanical Injury (MI) and Virus Inoculation Assays

Twenty-eight VuTLPs-coding transcripts were expressed in the MI_CABMV RNA-Seq libraries (**Supplementary Table S8**; **Supplementary Figure 6A**). Of these, six (Vu3908|c0_g1_i1, Vu93958|c3_g1_i3, Vu85277|c0_g1_i1, Vu95346|c0_g1_i1, Vu58024|c0_g1_i1 and Vu83475|c0_g1_i1) were up-regulated exclusively at 60 min (**Supplementary Figure 6B**), indicating that VuTLPs are preferably recruited during the first hours after the MI_CABMV treatments. In addition, only one VuTLP-encoding transcript (Vu44758|c0_g1_i1) showed up-regulation at 16 h. Besides, the transcript Vu9781|c2_g1_i3 encoding VuTLP was up-regulated at 60 min and repressed at 16 h treatment, whereas the others presented constitutive expression or were not expressed in the comparisons.

Regarding the MI_CPSMV assay, twenty-seven VuTLPs-encoding transcripts were expressed (**Supplementary Table S8**; **Supplementary Figure 7A**). Of these, five were up-regulated exclusively at 60 min (Vu95899|c0_g1_i1, Vu133945|c0_g1_i1, Vu9781|c2_g1_i3, Vu9763|c0_g1_i1 and Vu95346|c0_g1_i1) and three at 16 h (Vu49480|c1_g2_i2, Vu88791|c0_g1_i2 and Vu44758|c0_g1_i1) (**Supplementary Figure 7B**). In addition, a VuTLP-encoding (Vu83475|c0_g1_i1) transcript was up-regulated in both treatments (60 min and 16 h; **Supplementary Figure 7A**). These results also point out that potential VuTLPs were recruited in the first hour after MI_CPSMV treatment.

When comparing the transcriptional orchestration in response to the two MI_viral inoculation assays, it was observed that four potential VuTLPs (Vu9781|c2_g1_i3, Vu95346|c0_g1_i1, Vu44758|c0_g1_i1 and Vu83475|c0_g1_i1) were up-regulated in response to both assays (**Supplementary Figure 8A**). Por sua vez, four (Vu3908|c0_g1_i1, Vu93958|c3_g1_i3, Vu85277|c0_g1_i1 and Vu58024|c0_g1_i1) and five (Vu95899|c0_g1_i1, Vu133945|c0_g1_i1, Vu9763|c0_g1_i1, Vu49480|c1_g2_i2 and Vu88791|c0_g1_i2) VuTLPs-encoding transcripts were induced exclusively in response to MI_CABMV and MI_CPSMV, respectively (**Supplementary Figure 8A**).

3.6.2. Root Dehydration Assay

Thirty potential VuTLPs-encoding transcripts were expressed in cowpea under root dehydration (**Supplementary Table S8**; **Supplementary Figure 9A**). Of these, Vu73314|c0_g1_i1 and Vu49480|c1_g2_i1 were up- and repressed at 25 min, respectively. Additionally, the transcript Vu94809|c0_g2_i1 encoding VuTLP was induced at 25 min and repressed at 150 min treatment. In addition, the transcript Vu88791|c0_g1_i2 was induced in both treatments (25 min and 150 min). Besides, eight (Vu88791|c0_g1_i1, Vu8524|c1_g3_i1, Vu93958|c3_g1_i3, Vu93958|c3_g1_i2, Vu49480|c1_g3_i1, Vu9781|c2_g1_i3, Vu149510|c0_g2_i1 and Vu83475|c0_g1_i1) and three (Vu54351|c0_g1_i1, Vu3908|c0_g1_i1 and Vu79818|c1_g2_i8) transcripts encoding for VuTLPs were induced and repressed exclusively at 150 min treatment, respectively. These results indicate that some TLPs are recruited in the first minutes after stress imposition, even though a higher number of transcripts were detected at 150 min (**Supplementary Figure 9B**).

3.6.3. MI_CABMV vs. MI_CPSMV vs. Root Dehydration Assays

A comparison between the transcriptional orchestration in response to both MI_viruses (in leaves) and root dehydration (roots) revealed interesting insights. There were 20 potential VuTLPs expressed jointly in the assays (MI_CABMV and MI_CPSMV and root dehydration, **Supplementary**

 doi:10.20944/preprints202409.0657.v1

Table S8; **Supplementary Figure 8B**), thus indicating a crosstalk role for TLPs under different stress types and tissues.

Qualitative analyses showed a heterogeneous orchestration of VuTLPs. Only two VuTLPs candidates (Vu9781|c2_g1_i3 and Vu83475|c0_g1_i1) were up-regulated in the three assays (**Supplementary Table S8**). Exclusive responses were also observed under different treatments. For example, three (Vu3908|c0_g1_i1, Vu85277|c0_g1_i1 and Vu58024|c0_g1_i1), four (Vu95899|c0_g1_i1, Vu133945|c0_g1_i1, Vu49480|c1_g2_i2 and Vu9763|c0_g1_i1) and seven (Vu88791|c0_g1_i1, Vu8524|c1_g3_i1, Vu94809|c0_g2_i1, Vu93958|c3_g1_i2, Vu49480|c1_g3_i1, Vu73314|c0_g1_i1 and Vu149510|c0_g2_i1) VuTLPs-encoding transcripts were up-regulated exclusively under MI_CABMV and MI_CPSMV inoculation and root dehydration, respectively. Thus, there is no overall conservation in VuTLPs transcriptional orchestration towards different analyzed assays.

3.7. RNA-Seq Data Validation by qPCR

3.7.1. MI_CABMV and MI_CPSMV Assays

The selection of target transcripts (Vu44758|c0_g1_i1, Vu9781|c2_g1_i3, Vu83475|c0_g1_i1, Vu93958|c3_g1_i3 and Vu58024|c0_g1_i1) was based on their regulation in the RNA-Seq libraries (**Supplementary Table S8**) for MI_CABMV experiment. Considering the functional primer pairs, todos apresentaram eficiência dentro dos padrões aceitáveis (**Supplementary Figure 10**). Relative expression analysis showed that Vu9781|c2_g1_i3, Vu83475|c0_g1_i1, Vu93958|c3_g1_i3 and Vu58024|c0_g1_i1 were up-regulated (UR) at 60 min (**Supplementary Figure 11**) while Vu44758|c0_g1_i1 was UR at 16 h (**Supplementary Table S9** and **Supplementary Figure 11**), confirming RNA-Seq data. However, qPCR results have shown that Vu9781|c2_g1_i3 was not modulated at 16 h (not significant; ns), differing from RNA-Seq results where it was down-regulated (**Supplementary Table S9** and **Supplementary Figure 11**).

For MI_CPSMV treatment, three differentially expressed target transcripts (Vu9781|c2_g1_i3, Vu88791|c0_g1_i2 and Vu44758|c0_g1_i1) were evaluated, todos os pares de primers analisados nesta etapa apresentaram eficiência dentro dos padrões aceitáveis; **Supplementary Figure 12**). Vu9781|c2_g1_i3 was UR at 60 min and Vu88791|c0_g1_i2 and Vu44758|c0_g1_i1 were UR at 16 h (**Supplementary Table S9** and **Supplementary Figure 13**). All these results confirmed RNA-Seq expression data.

3.7.2. Root Dehydration

For the mentioned assay, six differentially expressed target transcripts (Vu88791|c0_g1_i2, Vu9781|c2_g1_i3, Vu3908|c0_g1_i1, Vu54351|c0_g1_i1, Vu79818|c1_g2_i8 and Vu49480|c1_g3_i1) were analyzed by qPCR, all with acceptable efficiency values (93.00 to 107.70%; **Supplementary Figure 14**). Relative expression analysis showed that Vu88791|c0_g1_i2, Vu9781|c2_g1_i3 and Vu49480|c1_g3_i1 were up-regulated at 150 min and transcripts Vu3908|c0_g1_i1 and Vu54351|c0_g1_i1 were down-regulated at 150 min (**Supplementary Table S9**; **Supplementary Figure 15**). All these results confirmed RNA-Seq expression data.

4. Discussion

Using TLPs sequences of others species as seed sequences, cowpea VuTLPs candidates were searched in the RNA-Seq libraries generated under radicular dehydration or CABMV or CPSMV inoculation. Such conditions were chosen to exemplify the participation of TLPs in response to both, biotic and abiotic stresses, since TLP participation in response to abiotic stress is still incipient when compared to its biotic counterpart. Additionally, RNA-Seq reads have up to 400 bp, depending on the sequencing technology used (Wang et al. 2009), allowing direct VuTLP structural characterization.

Most of the identified VuTLPs presented the REDDD motif, in an acidic cleft. This configuration is responsible for the reported antifungal activity of TLPs in plants (Cao et al. 2016; Ghosh and Chakrabarti, 2008). However, the alignment of the referred candidates with other TLPs from *A.*

 doi:10.20944/preprints202409.0657.v1

thaliana, *N. tabacum*, *Z. mays*, and *C. arietinum* revealed that some VuTLPs did not contain all expected amino acids in REDDD motif. It is still not clear whether these small differences have a significant impact on the substrate selectivity or protein function (Petre et al. 2011). Petre et al. (2011) found that the acidic cleft is the most conserved region among the eukaryotic TLPs. For example, in sTLPs (small TLPs), the most of REDDD amino acids are conserved, but wheat sTLPs sequences with resolved structure have revealed no acidic clefts nor any particular conserved region which may be linked to the xylanase inhibitor function reported for such proteins (Petre et al. 2011, Fierens et al. 2007).

Regarding the 16-cysteine conservation, with the exception of most VuTLPs, all other VuTLPs presented the 16 residues that form eight disulfide bonds necessary for correct folding and to ensure a high level of thermostability and pH constancy (Liu et al. 2012). Most TLPs presented molecular weights ranging from 21 to 26 kDa (Liu, Sturrock and Ekramoddoullah, 2010) and isoelectric point ranging between 3.4 and 12.0 (Velazhahan, Datta Muthukrishnan, 1999), having this last one influence on the electrostatic potential at the molecular surface (Tachi et al. 2009). Thus, our results show that most of the VuTLPs presented values within the expected standard for both characteristics, except for some VuTLPs that presented specificities regarding their molecular weight.

The SignalP prediction revealed that most VuTLPs present hydrophobic signal peptide sequences, indicating that they are predominantly secreted, a result confirmed by TargetP prediction. TLPs are located mainly in two subcellular compartments; apoplastic space or vacuole because of an N-terminal signal peptide that directs the mature protein to a secretory pathway and in some instances, a C-terminal polypeptide that directs them to a vacuolar compartment (Melcher et al. 1993). Previous studies have used GFP-gene approach associated with a transient expression system to study the localization of these proteins. Most of them point to a location in extracellular spaces (Singh et al. 2013; Zhang et al. 2017; Yang et al. 2024), corroborating with the in silico prediction.

To infer on the distribution, relative position, and abundance of VuTLPs, we anchored candidate transcripts against the cowpea genome, recently available (Lonardi et al. 2019). The amount of VuTLPs is similar to those found in other leguminous plant genomes. Specifically, *Ammopiptanthus nanus* possesses 31 TLP genes, *Vitis vinifera* has 32, *Medicago truncatula* has 49, *Trifolium pratense* has 23, and *Lupinus albus* has 29 (Liu et al., 2023).

The results of gene duplication analysis suggested the existence of gene replication events in cowpea genome. VuTLP genes were found to cluster into six tandem duplication event regions on 4 of 11 chromosomes, implying that they may originate from the recent gene duplication events. Moreover, segmental duplications were common in the cowpea TLP family. Fifteen segmental duplication gene pairs were identified in cowpea genome, compared to 3, 2, 7 and 4 in *A. thaliana*, rice, poplar and maize, suggesting that a majority of the VuTLP genes were generated by segmental replications (Cao et al. 2016). Cao et al. (2016) suggested that tandem and segmental duplications are the main factors leading to the enlargement and diversity of TLP family in plants.

Yang et al. (2024) identifed six segmentally duplicated TLP gene pairs in the *Panax notoginseng* genome. Gene duplication analysis revealed thirteen gene duplication events in the watermelon genome, including one tandem and twelve segmental duplications (Ram et al. 2022). Liu et al. (2020b) identifed twelve segmentally duplicated and seven tandemly duplicated gene pairs that have experienced purifying selection in the melon genome.

Natural selection analysis is widely used to estimate how natural pressure and evolutionary forces affect duplicated genes and corresponding proteins (Ellegren, 2008). In cowpea, a majority of TLP genes were shown to be driven by purifying selection, but few sites of TLP sequences were shown to be driven by positive selection. However, very similar results were found in the study on the evolution of TLPs in *A. thaliana*, rice and maize, which revealed that many TLP sequences were under purifying selection pressure and only on a few sites of TLP sequences may have experienced positive selection during the evolutionary process (Cao et al. 2016). Study on Poplar TLP genes suggested that four TLP kinases and ten TLPs have undergone some positive selection, indicating that plant TLPs may have experienced a diversified natural selection processes (Liu, Sturrock and Ekramoddoullah, 2010). Some of the exposed amino acids of TLPs are under positive selection, the amino acids that form the acidic cleft are always under purifying selection, which indicated that

 doi:10.20944/preprints202409.0657.v1

conservation of the acidic cleft are crucial and could be important for the antifungal activity of TLPs (Misas-Villamil and van der Hoorn, 2008).

In the current study, 34 cowpea VuTLP genes were analyzed into three groups through the NJ tree. In all obtained groups, a clear distinction emerged regarding the VuTLP genes when analyzing the REDDD motif associated with duplication events. Group III, which contained the highest number of clustered genes, displayed a consistent pattern of residue substitution in the REDDD motif, where the first D was replaced by Y. In contrast, genes in the other groups exhibited different substitution patterns. When comparing the conservation and substitution patterns of the REDDD residues with the duplication events, it was observed that residues maintaining conservation were expanded through segmental duplication and were under purifying selection. This indicates that these residues, and consequently the antifungal activity associated with this motif, were preserved. Conversely, residues that underwent alterations or substitutions expanded through tandem duplication and were subject to positive selection, suggesting a diversification of these residues that could lead to the emergence of new functions. Analyzing eukaryotic TLPs, Liu, Sturrock and Ekramoddoullah (2010) observed the formation of nine distinct groups showing, what he called TLP superfamily, as highly diversified. A possible explanation for the formation of these groups can be attributed to the structural diversity presented by TLPs. According to Petre et al. (2011), such diversity may influence the biological and biochemical functions, and differences in the topology around the cleft could determine the specificity of TLPs to their target ligands (Min et al. 2004). This may be true also considering the VuTLPs since their diversity and position in different positions of the NJ tree is possibly due to the presence of orthologs and paralogs, with structural and functional specificities. Zhao et al. (2024) identified four distinct groupings within the TLP gene Family and highlighted the presence of distinct subfamilies within each group indicating the close evolutionary associations between certain TLP genes from faba bean and other legume species.

Over time plants acquired new weapons to promote resistance over time (Burdon and Thrall 2009; Benko-Iseppon et al. 2010) while pathogens also continued to evolve, bringing pressure on the structure and diversity of some associated gene families, increasing the size of some protein families (Kohler et al. 2008), including PR-genes. Such an increase represents a major significance for the functional diversity via sub- or neo-functionalization of paralogs (Paterson et al. 2010). In this scenario, the natural selection of genes with new functions under environmental pressure probably played a significant role also in the evolution of TLPs (Liu, Sturrock and Ekramoddoullah, 2010), as may also be suggested for cowpea.

In addition, regarding the gene structure of cowpea TLPs, genes with three, two and one exon were verified, besides to four genes that did not present introns. The organization of introns and exons of TLPs have been reported to be in the range of one to ten exons (Liu et al., 2010; Petre et al., 2011; Cao et al., 2016). Out of the 10 TLP genes identified in the faba bean transcriptome, two genes exhibited three exons, five genes consisted of only one exon, and three genes displayed two exons (Zhao et al. 2024). Yan et al. (2017) found that most of the grape TLP genes had less than four exons and four genes also had no introns.

The 34 cowpea TLP coding genes had their predicted promoters (1 kb) analyzed in regard to the presence and identity of CCREs. Considering CCREs, a MEME motif is a sequence pattern that repeatedly occurs in one or more sequences in the input group. MEME can be used to discover novel patterns because discoveries are based only on the input sequences, not on any prior knowledge, such as databases of known motifs (Bailey and Elkan, 1994). Discovered *bona fide* CCREs can be used to identify potential physiological or adaptive processes in which TLPs participate, associating transcription factors and their intrinsic biological processes. This action adds value to the molecular dynamics of cowpea TLPs, especially in association with the considered transcriptomic data.

Concerning the promoters analyzed, CCREs associated with bHLH, Dof-type, and MYB-related TFs attended the stipulated statistical parameters. Also considering literature data, these TFs present an intimate association with plant stress response and other physiological processes. bHLH members can act as transcriptional activator or repressor and play essential roles in metabolic and developmental processes (Feller et al. 2011). Some bHLH have also been reported to be modulated

 doi:10.20944/preprints202409.0657.v1

under cold, drought, and salt stress (Wang et al. 2018; Mao et al. 2017). Dof-type transcription factors, in turn, participate widely in plant development and abiotic stress response (Li et al. 2016; Wang et al. 2017). In tobacco, the *Sar8.2b* gene can be activated by the Dof-type TF, which is related to systemic acquired resistance (Song and Goodman, 2002). Finally, there are reports on MYB-related TFs acting as retrotransposon regulators and controlling defense-related genes, besides being induced by wounding and other elicitors, in tobacco plants (Sugimoto, 2000). Such in silico data on promoters associated with TLPs in RNA-Seq libraries under different stresses (biotic and abiotic) indicate a plurality of roles for TLPs, also reinforcing their biotechnological potential.

The RNA-Seq data showed that potential VuTLPs were induced in the first hours in response to the MI_CABMV or MI_CPSMV. There are previous reports in the literature about the interaction between viral proteins and TLPs. Kim et al. (2005) observed that a *Nicotiana tabacum* (NtTLP1) TLP specifically interacted with CMV (*Cucumber mosaic virus*) - 'movement-related protein', 'movement protein' and 'coat protein' - during a yeast two-hybrid screen experiment. However, the determination of the possible role of TLP-virus interaction requires deepened studies.

When comparing the transcriptional orchestration in response to MI_viral, the results indicate a functional specialization of VuTLPs in cowpea submitted to MI_CABMV and MI_CPSMV, since only specific representatives are induced, in detriment of others, in response to inoculation by a given virus / injury.

The results for radicular dehydration assay indicate that some TLPs are recruited in the first minutes after stress imposition, even though a higher number of transcripts were detected at 150 min. Previous studies did not address the global transcriptional orchestration of TLP proteins and referred to individual case studies. An example is given by Jung et al. (2005) that reported a new TLP isoform isolated from carrot, regarding a drought-specific, ABA-independent, non-organ-specific, and non-developmental-stage-specific protein.

The comparison between the orchestration in response to both viruses (in leaves) and radicular dehydration (roots) revealed potential VuTLPs expressed in both assays, indicating a crosstalk role for TLPs under different stress types and tissues. Besides, there is no overall conservation in VuTLPs orchestration towards different viral isolates and radicular dehydration, indicating specific TLP associated with the response of each virus type and radicular dehydration.

Signalization response in plants submitted to abiotic and biotic stressors can induce separate and overlapping sets of genes, leading to the expression of distinct as well as common components (Zhu et al. 2014; Mellacheruvu et al. 2016). These separate pathways show nodal points where they converge and crosstalk to optimize the various defense responses (Xiao et al. 2013), resulting in shared stress mitigation strategy by combined morpho-physiological processes and molecular responses (Pandey, Ramegowda, Senthil-Kumar, 2015). The identification of crosstalk between biotic and abiotic signaling pathways has been crucial for envisaging and strengthening our understanding of the regulation of plants response against combined stresses. Genes such as those coding for TLPs that impart abiotic as well as biotic stress tolerance are highlighted in some studies in order to understand the mechanism underlying tolerance to dual stresses.

Singh et al. (2013) analyzed an *Arachis diogoi* (AdTLP) TLP induced under fungal infection (*Phaeoisariopsis personata*). The cDNA encoding AdTLP was cloned using RACE-PCR and used to transform tobacco plants. Overexpression of AdTLP resulted in increased resistance to pathogenic fungi and tolerance to abiotic stresses (high salinity and oxidative stress). In addition, transgenic plants also exhibited a higher level of transcription of genes *PR1a*, *PI-I* and *PI-II* compared to *WT*. Such genes are associated with plant pathogen defense mechanisms. Chowdhury, Basu and Kundu (2017) showed that transgenic lines of sesame overexpressing an osmotin-like protein (SindOLP) presented tolerance against abiotic stresses (drought and high salinity) and resistance against the fungus *Macrophomina phaseolina*. Overexpression of SindOLP resulted in the induction of three genes [*superoxide dismutase* (*SiSOD*), *cysteine protease inhibitor* (*SiCysPI*), and *glutathione-S-transferase* (*SiGST*)] which encode enzymes for the elimination of reactive oxygen species (ROS), indicating that SindOLP participates in the ROS regulation, which is common to both stresses addressed.

 doi:10.20944/preprints202409.0657.v1

Considering the results of qPCR, it is evident that TLPs are recruited in the first minutes after MI_CABMV or MI_CPSMV treatment. Even though both treatments had the injury in common, different TLPs were recruited after each virus inoculation, indicating some specificity that deserves additional evaluations.

Some recently published papers have shown the induction of TLPs after virus inoculation. Madroñero et al. (2018) analyzed changes in the papaya transcriptome in response to the inoculation with PMeV complex (Papaya Meleira Virus, PMeV, and Papaya Meleira Virus 2, PMeV2). To evaluate the effects of SA signaling on PMeV complex load, *Carica papaya* seedlings were exposed to this hormone and inoculated with both viruses. The SA-treated plants showed increased *PR1, PR5* and *CHIA* levels confirming SA signaling activation. The same plants presented reduced PMeV complex load, suggesting that SA signaling plays a role in *C. papaya* tolerance to PSD before flowering.

Kappagantu et al. (2017) studied the effect of the host response of hop plants (*Humulus lupulus*) to Hop stunt viroid (HSVd) on the development of hop powdery mildew (*Podospheara macularis*; HPM). Transcriptome analysis followed by qPCR analysis showed that transcript levels of PR-genes such as PR-1, chitinase, and TLPs were induced in leaves infected with HPM alone. The response of these genes to HPM was significantly down-regulated in leaves with HSVd-HPM mixed infection. These results confirm that HSVd alters host metabolism, physiology, and plant defense responses.

The results of qPCR show that – despite being traditionally responsive to biotic stresses – TLPs have been reported as recruited under abiotic stresses (Zhao et al. 2024; Muoki et al. 2021; Jung et al. 2005; Ruperti et al. 2002; Petre et al. 2011). In the present work, a higher number of VuTLPs were induced at 150 min after root dehydration. Zhao et al. (2024) identified and characterized 10 TLP genes from the faba bean transcriptome. Two genes (VfTLP4-3 and VfTLP5) displayed high expression levels in the drought-tolerant cultivar when exposed to drought stress. In turn, Misra et al. (2016) identified and characterized a new basil (*Ocimum basilicum*) TLP (ObTLP1) from ESTs recovered after MeJA treatment. The ectopic expression of ObTLP1 in *A. thaliana* provided fungal resistance (*Scleretonia sclerotiorum* and *B. cinerea*) and tolerance to abiotic stresses (drought and high salinity).

Despite the evidence of the present work and previous studies, the mechanism of action of such proteins against abiotic stress remain largely unknown, although investigations have confirmed the roles of several TLP genes in cold temperature, salt and drought stresses (Jung et al. 2005; Muoki et al. 2012; Singh et al. 2013; Misra et al. 2016; Zhao et al. 2024). Zhao et al. (2024) showed ectopically expression of two genes from the faba bean (VfTLP4-3 and VfTLP5) in tobacco increased drought tolerance. This enhancement in drought tolerance was accompanied by changes in the activity of peroxidase (POD) and superoxide dismutase (SOD), as well as the levels of malondialdehyde (MDA) and proline content. VfTLP4-3 appeared to primarily regulate chitin metabolism, while VfTLP5 was associated with the regulation of proline synthesis. The work of Chowdhury, Basu and Kundu (2017) with sesame, indicated a role in the induction of three genes [*superoxide dismutase* (*SiSOD*), *cysteine protease inhibitor* (*SiCysPI*), *glutathione-S-transferase* (*SiGST*)] which encode enzymes for the elimination of reactive oxygen species (ROS). These results suggest that SindOLP participates in the regulation of the ROS pathway, which is common to both types of stresses.

5. Conclusion

The present study describes the structural diversity and gene expression profile of thaumatin-like proteins (VuTLPs) in cowpea under biotic and abiotic stress conditions. Identification of 34 loci encoding VuTLPs along with the detection of both segmental and tandem duplication events underlines the complexity of this gene family in cowpea and its evolutionary dynamics. The prevalence of purifying selection across VuTLPs suggests that these proteins have conserved functions important for plant adaptation to stress. The obtained results showed that different stressors provoke distinct expression profiles in VuTLPs and that the most rapid induction generally occurs within minutes following the onset of stress. Delayed expression in response to root dehydration underlines the possibility of functional specialization associated with certain types of unfavorable conditions. This is further supported by the analysis of promoters linking VuTLPs with

 doi:10.20944/preprints202409.0657.v1

various transcription factors, including bHLH, Dof-type, and MYB-related factors, indicating multifaceted participation in response mechanisms. Collectively, these analyses provide an overview of VuTLPs and suggest them as potential targets for genetic engineering strategies in cowpea. Towards this, future research studies should henceforth be directed to the elucidation of VuTLP-mediated molecular pathways and interaction with other stress-related proteins for fullest biotechnological applications in crop improvement.

Conflicts of interest: The authors declare that they have no conflict of interest.

Funding: This work was supported by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico; grant number: 406657/2023-8, 406048/2022-3) and CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior; grant number: 88887.814478/2023-00).

Acknowledgments: The authors acknowledge the Fundação de Amparo à Pesquisa do Estado de Pernambuco (FACEPE), Conselho Nacional de Desenvolvimento Científico Tecnológico (CNPq), and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for fellowships and financial support. The authors acknowledge the National Laboratory for Scientific Computing (LNCC), for providing high-performance computing resources through the Santos Dumont supercomputer, used to obtain the research results reported in this paper.

Supplementary Materials: Supplementary Material SM1: Thaumatin-like proteins candidates predicted from cowpea (VuTLPs) RNA-Seq libraries under CABMV or CPSMV inoculation and submitted to root dehydration and other TLPs from different species used in alignment construction. **Supplementary Material SM2:** Non-redundant VuTLPs sequences identified in RNA-Seq library under biotic and abiotic stresses. **Supplementary Figure 1.** Implemented procedures and experimental design for root dehydration assay. **A.** Growth conditions and stress application in tanks with hydroponic solution in a controlled environment prior to the application of root dehydration treatment (left) and by withdrawing the nutrient solution from the roots of treated plants (right); **B.** Experimental design with three replicates (A, B, C). Treatment times: 25 and 150 minutes after removing of hydroponic solution (RD25′ and RD150′, respectively) and their respective non stressed controls [Cont.25′ and Cont.150′, respectively] **Abbreviations**: RD (Root Dehydration); Cont. (Control) " ′ " (minutes). **Supplementary Figure 2.** Implemented procedures and experimental design for the 'mechanical injury and CABMV inoculation' and 'mechanical injury and CPSMV inoculation' assays. **A.** Growth conditions and stress application in a controlled environment (right) and the procedures for stress application (left); **B.** Experimental design with three replicates (A, B, C). Treatment times: 60 min and 16 h stresses application ('MI+Vir60′ and MI+Vir16h′, respectively) and their respective non stressed controls [Cont.60′ and Cont.16h, respectively] **Abbreviations**: MI (Mechanical Injury); Vir (Virus: CABMV or CPSMV); Cont. (Control); " ′ " (minutes). **Supplementary Figure 3.** Alignment of amino-acid sequences of *Vigna unguiculata* thaumatin-like proteins (TLPs) together with TLPs from other species. Sequence comparison was carried out with Clustal Omega software with standard parameters. Colored bars in the sequence show conserved residues and conserved motifs. Black boxes indicate conserved domain regions; Pink diamonds indicate the cysteine residues that form the eight disulfide bonds; Yellow triangles indicate the position of REDDD motif and yellow bars in the base show the level of conservation for each position. **Supplementary Figure 4**. Genome wide organization of cowpea TLP (*VuTLP*) genes. (**a**) Phenetic tree based on the protein sequences of 34 *VuTLP* genes. Phenetic tree was constructed using the neighbor-joining method with MEGA7. Three groups were formed and colored in green, pink and purple. Bootstrap values at the nodes from 2000 replicates were used to assess the robustness of the tree. The scale is in amino acid substitutions per site. (**b**) Exon-intron structure of *VuTLP* genes: yellow indicates coding sequence (CDS), blue indicates untranslated 5′- and 3′- regions, black indicates introns. **Supplementary Figure 5**. Candidate motifs of cis-regulatory candidates statistically validated, presenting their consensus sequence, e-value and number of anchoring sites (MEME software output underlined in blue) together with the annotation of the associated transcription factor and its Matrix JASPAR ID (in red) according to the Tomtom software output. **Supplementary Figure 6. (A)** Heat map illustrating the transcriptional regulation of 28 VuTLP-encoding transcripts after inoculation with *Cowpea Aphid-borne Mosaic Virus* (CABMV). **(B)** Venn diagram of VuTLP-candidates in each treatment / comparison. Libraries: IT60E: IT85F-2687, 60 minutes, stressed library; IT60C: IT85F-2687, 60 minutes, control library; IT16E: IT85F-2687, 16 hours, stressed library; IT16C: IT85F-2687, 16 hours, control library. **Supplementary Figure 7. (A)** Heat-map illustrating the transcriptional regulation of 27

 doi:10.20944/preprints202409.0657.v1

VuTLP-encoding transcripts after inoculation with *Cowpea Severe Mosaic Virus* (CPSMV). **(B)** Venn diagram of induced VuTLP-candidates in each treatment / comparison. Libraries: BR60E: BR-14 Mulato, 60 minutes, stressed library; BR60C: BR-14 Mulato, 60 minutes, control library; BR16E: BR-14 Mulato, 16 hours, stressed library; BR16C: BR-14 Mulato, 16 hours, control library. **Supplementary Figure 8. (A)** Venn diagram with all VuTLP-candidates induced in each biotic assay (CABMV and CPSMV inoculation). **(B)** Venn diagram with all VuTLP-candidates induced in each assay (CABMV or CPSMV inoculation and radicular dehydration). **Supplementary Figure 9. (A)** Heat-map illustrating the transcriptional regulation of 30 VuTLP-encoding transcripts after radicular dehydration stress. **(B)** Venn diagram of induced VuTLP-candidates in each treatment / comparison. Libraries: PO25E: Pingo de Ouro, 25 minutes, stressed library; PO25C: Pingo de Ouro, 25 minutes, control library; PO150E: Pingo de Ouro, 150 minutes, stressed library; PO150C: Pingo de Ouro, 150 minutes, control library. **Supplementary Figure 10.** Dilution curve with the points used for the efficiency analysis of TLP-target transcripts in the CABMV experiment: **(A)** Vu44758|c0_g1_i1; **(B)** Vu9781_c2_g1_i3; **(C)** Vu83475|c0_g1_i1; **(D)** Vu93958_c3_g1_i3; **(E)** Vu58024_c0_g1_i1. **Supplementary Figure 11**. Relative expression of TLP-target transcripts at 60 minutes (60 min) and 16 hours (16 h) after CABMV inoculation, as compared with not stressed control. **Supplementary Figure 12.** Dilution curve with the points used for the efficiency analysis of TLP-target transcripts in the CPSMV experiment: **(A)** Vu9781_c2_g1_i3; **(B)** Vu88791_c0_g1_i2; **(C)** Vu44758|c0_g1_i1. **Supplementary Figure 13.** Relative expression of TLP-target transcripts at 60 minutes (60 min) and 16 hours (16 h) after CPSMV inoculation, as compared with not stressed control. **Supplementary Figure 14.** Dilution curve with the points used for the efficiency analysis of TLP-target transcripts in the root dehydration experiment: **(A)** Vu88791_c0_g1_i2; **(B)** Vu9781_c2_g1_i3; **(C)** Vu3908_c0_g1_i1; **(D)** Vu54351_c0_g1_i1; **(E)** Vu79818_c1_g2_i8; **(F)** Vu49480_c1_g3_i1. **Supplementary Figure 15**. Relative expression of TLP-target transcripts at 25 minutes after root dehydration, as compared with not stressed control. **(A)** Vu88791|c0_g1_i1 (blue) and Vu149510|c0_g2_i1 (green); **(B)** Vu49480|c1_g3_i1 (red). **Supplementary Table S1.** Cowpea libraries generated by the Cowpea Genomics Consortium (CpGC) including treatments, stress types, number of sequenced transcripts and sequencing methods used. **Supplementary Table S2.** Annotation of TLPs predicted in Vigna unguiculata (VuTLPs) in RNA-Seq libraries under CABMV and CPSMV inoculation and under radicular dehydration and respective controls. The identification of 126 TLPs in cowpea was based on curated protein sequences from UniProt via tBLASTn. Additionally, data from BLASTx against GenBank, Conserved Domain Evaluation, ProtParam and TargetP were used to predict thaumatin features. Legend: nuc - nucleotide; aa - amino acid; q. start - query start; q.end - query end; s. start - subject start; s.end - subject end; PSSM-ID - position-specific scoring matrix identification; pI - isoelectric point; MW - molecular weight.

Data Availability Statement: The datasets generated for this study can be found in the NCBI BioProject (https://www.ncbi.nlm.nih.gov/bioproject/), with the following accession numbers: Root dehydration | BioSample ID: SAMN14051116; Mechanical injury + CABMV | BioSample ID: SAMN15763372; Mechanical injury + CPSMV | BioSample ID: SAMN15774013.

Abbreviations

BLAST, Basic Local Alignment Search Tool; CABMV, Cowpea Aphid-born Mosaic Virus; CpGC, Cowpea Genomics Consortium; CPSMV, Cowpea Severe Mosaic Virus; CTD, Conserved Thaumatin Domain; ESTs, Expressed Sequence Tags; ET, Ethylene; FPKM, Fragments Per Kilobase Million; IPA, Instituto Agronômico de Pernambuco; JA, Jasmonic acid; MEGA, Molecular Evolutionary Genetic Analysis; MethylJA, Methyl Jasmonic Acid; NCBI, National Center for Biotechnology Information; PR, Pathogenesis-related; RNA-Seq, RNA Sequencing; SA, Salicylic Acid; SuperSAGE, Serial Analysis of Gene Expression; TLPs, Thaumatin-like proteins; VuTLPs, Thaumatin-like proteins of *Vigna unguiculata.*

References

1. Feller, K. MacHemer, E.L. Braun, E. Grotewold, Evolutionary and comparative analysis of MYB and bHLH plant transcription factors, Plant J. 66 (2011) 94–116.
2. Kohler, C. Rinaldi, S. Duplessis, M. Baucher, D. Geelen, F. Duchaussoy, B.C. Meyers, W. Boerjan, F. Martin, Genome-wide identification of NBS resistance genes in *Populus trichocarpa*, Plant Mol. Biol. 66 (2008) 619–636.
3. A.H. Paterson, M. Freeling, H. Tang, X. Wang, Insights from the comparison of plant genome sequences. *Annu. Rev. Plant. Biol.* 61 (2010) 349–372.
4. A.M. Benko-Iseppon, S.L. Galdino, T. Calsa, E.A. Kido, A. Tossi, L.C. Belarmino, S. Crovella, Overview on plant antimicrobial peptides., Curr. Protein Pept. Sci. 11 (2010) 181–188.

5. Hu, J. Jin, A.Y. Guo, H. Zhang, J. Luo, G. Gao, GSDS 2.0: an upgraded gene feature visualization server, Bioinformatics 31 (2015) 1296–1297.
6. Petre, I. Major, N. Rouhier, S. Duplessis, Genome-wide analysis of eukaryote thaumatin-like proteins (TLPs) with an emphasis on poplar., BMC Plant Biol. 11 (2011) 33.
7. Ruperti, L. Cattivelli, S. Pagni, A. Ramina, Ethylene-responsive genes are differentially regulated during abscission, organ senescence and wounding in peach (*Prunus persica*), J. Exp. Bot. 53 (2002) 429–437.
8. Liu, X. He, W. Li, C. Chen, F. Ge, Molecular cloning of a thaumatin-like protein gene from *Pyrus pyrifolia* and overexpression of this gene in tobacco increased resistance to pathogenic fungi, Plant Cell. Tissue Organ Cult. 111 (2012) 29–39.
9. Fierens, S. Rombouts, K. Gebruers, H. Goesaert, K. Brijs, J. Beaugrand, G. Volckaert, S. Van Campenhout, P. Proost, C.M. Courtin, J.A. Delcour, TLXI, a novel type of xylanase inhibitor from wheat (*Triticum aestivum*) belonging to the thaumatin family, Biochem. J. 403 (2007) 583–591.
10. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel, A. Bairoch; Protein Identification and Analysis Tools on the ExPASy Server;
(In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press (2005). pp. 571-607
11. E.A. Kido, J.R.C. Ferreira-Neto, R.L.O. Silva, L.C. Belarmino, N.M. Soares-Cavalcanti, V. Pandolfi, M.D. Silva, A.L. Nepomuceno, A.M. Benko-Iseppon, Expression dynamics and genome distribution of osmoprotectants in soybean: identifying important components to face abiotic stress, BMC Bioinformatics 14 (2013) S7.
12. Sievers, A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J.D. Thompson, D.G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, Mol. Syst. Biol. 7 (2011).
13. Song, R.M. Goodman, Cloning and identification of the promoter of the tobacco *Sar8.2b* gene, a gene involved in systemic acquired resistance, Gene. 290 (2002) 115–124.
14. Breiteneder, C. Radauer, A classification of plant food allergens, J. Allergy Clin. Immunol. 113 (2004) 821–830.
15. Ellegren, Comparative genomics and the study of evolution by natural selection, Mol. Ecol. 17 (2008) 4586–4596, https://doi.org/10.1111/j.1365-294X.2008.03954.x.
16. H. Li, W. Huang, Z.W. Liu, Y.X. Wang, J. Zhuang, Transcriptome-based analysis of Dof family transcription factors and their responses to abiotic stress in tea plant (*Camellia sinensis*), Int. J. Genomics. 2016 (2016).
17. H. Tachi, K. Fukuda-Yamada, T. Kojima, M. Shiraiwa, H. Takahara, Molecular characterization of a novel soybean gene encoding a neutral PR-5 protein induced by high-salt stress, Plant Physiol. Biochem. 47 (2009) 73–79.
18. H. van der Wel, K. Loeve, Isolation and characterization of thaumatin I and II, the sweet-tasting proteins from *Thaumatococcus daniellii* Benth, Eur. J. Biochem. 31 (1972) 221–225.
19. H. Wang, S. Zhao, Y. Gao, J. Yang, Characterization of dof transcription factors and their responses to osmotic stress in poplar (*Populus trichocarpa*), PLoS One. 12 (2017) 1–19.
20. Cao, Y. Lv, Z. Hou, X. Li, L. Ding, Expansion and evolution of thaumatin-like protein (TLP) gene family in six plants, Plant Growth Regul. 79 (2016) 299–307.
21. Grenier, C. Potvin, J. Trudel, A. Asselin, Some thaumatin-like proteins hydrolyse polymeric β-1,3-glucans, Plant J. 19 (1999) 473–480.
22. J. Madroñero, S.P. Rodrigues, T.F.S. Antunes, P.M.V. Abreu, J.A. Ventura, A.A.R. Fernandes, P.M.B. Fernandes, Transcriptome analysis provides insights into the delayed sticky disease symptoms in Carica papaya, Plant Cell Rep. 37 (2018) 967–980.
23. J. Trudel, J. Grenier, C. Potvin, a Asselin, Several thaumatin-like proteins bind to beta-1,3-glucans., Plant Physiol. 118 (1998) 1431–1438.
24. J. Xiao, H. Cheng, X. Li, J. Xiao, C. Xu, S. Wang, Rice WRKY13 Regulates Cross Talk between Abiotic and Biotic Stress Signaling Pathways by Selective Binding to Different cis-Elements, Plant Physiol. 163 (2013) 1868–1882.
25. J.C. Misas-Villamil, R.A.L. van der Hoorn, Enzyme-inhibitor interactions at the plant-pathogen interface, Curr. Opin. Plant Biol. 11 (2008) 380–388, https://doi.org/10.1016/j.pbi.2008.04.007.
26. J.C. Oliveros VENNY: An interactive tool for comparing lists with Venn diagrams. Available: http://bioinfogp.cnb.csic.es/tools/venny/ index.html. (2007) Accessed 17 December 2018.
27. J.J. Burdon, P.H. Thrall, Coevolution of plants and their pathogens in natural habitats, Science 324 (2009) 755–756.
28. J.J. Liu, R. Sturrock, A.K.M. Ekramoddoullah, The superfamily of thaumatin-like proteins: its origin, evolution, and expression towards biological function, Plant. Cell Rep. 29 (2010) 419–436.
29. Hiroyuki, R. Terauchi, Regulation of expression of rice thaumatin-like protein: inducibility by elicitor requires promoter W-box elements, Plant Cell Rep. 27 (2008) 1521–1528.

30. K. Mao, Q. Dong, C. Li, C. Liu, F. Ma, Genome Wide Identification and Characterization of Apple bHLH Transcription Factors and Expression Analysis in Response to Drought and Salt Stress, Front. Plant Sci. 8 (2017).
31. K. Min, S.C. Ha, P.M. Hasegawa, R.A. Bressan, D.J. Yun, K.K. Kim, Crystal structure of osmotin, a plant antifungal protein, Proteins 54 (2004) 170–173.
32. K. Sugimoto, MYB-Related Transcription Factor NtMYB2 Induced by Wounding and Elicitors is a Regulator of the Tobacco Retrotransposon Tto1 and Defense-Related Genes, Plant Cell Online. 12 (2000) 2511–2528.
33. L.L.B. Amorim, J.R.C. Ferreira-Neto, J.P. Bezerra-Neto, V. Pandolfi, F.T. Araújo, M.K. Silva Matos, M.G. Santos, E.A. Kido, A.M. Benko-Iseppon, Cowpea and abiotic stresses: Identification of reference genes for transcriptional profiling by qPCR, Plant Methods. 14 (2018).
34. L.S. Melcher, M.B. Sela-Buurlage, S.A. Vloemans, C.P. Woloshuk, J.S. van Roekel, J. Pen, P.J. van den Elzen, B.J. Cornelissen, Extracellular targeting of the vacuolar tobacco proteins AP24, chitinase and b-1,3-glucanase in transgenic plants, Plant Mol. Biol. 21 (1993) 583–593.
35. Hayashi, S. Shiro, H. Kanamori, S. Mori-Hosokawa, H. Sasaki-Yamagata, T. Sayama, M. Nishioka, M. Takahashi, M. Ishimoto, Y. Katayose, A. Kaga, K. Harada, H. Kouchi, Y. Saeki, Y. Umehara, A thaumatin-like protein, Rj4, controls nodule symbiotic specificity in soybean, Plant Cell Physiol. 55 (2014) 1679–1689.
36. Kappagantu, J.M. Bullock, M.E. Nelson, K.C. Eastwell, *Hop stunt viroid* : Effect on Host (*Humulus lupulus*) Transcriptome and Its Interactions With Hop Powdery Mildew (*Podospheara macularis*), Mol. Plant-Microbe Interact. 30 (2017) 842–851.
37. M. Lynch, J.S. Conery, The evolutionary fate and consequences of duplicate genes, Science 290 (2000) 1151–1155, https://doi.org/10.1126/science.290.5494.1151.
38. M.A. Batalia, A.F. Monzingo, S. Ernst, W. Roberts, J.D. Robertus, The crystal structure of the antifungal protein zeamatin, a member of the thaumatin-like, PR-5 protein family, Nat. Struct. Biol. 3 (1996) 19–23.
39. M.B. Eisen, P.T. Spellman, P.O. Brown, D. Botstein. Cluster analysis and display of genome-wide expression patterns, Proc Natl Acad Sci USA. 95 (1998) 14863–14868.
40. M.D. Robinson, D.J. McCarthy, G.K. Smyth, edgeR: A Bioconductor package for differential expression analysis of digital gene expression data, Bioinformatics. 26 (2009) 139–140.
41. M.J. Kim, B.K. Ham, H.R. Kim, I.J. Lee, Y.J. Kim, K.H. Ryu, Y.I. Park, K.H. Paek, In vitro and in planta interaction evidence between *Nicotiana tabacum* thaumatin-like protein 1 (TLP1) and Cucumber mosaic virus proteins, Plant Mol. Biol. 59 (2005) 981–994.
42. M.W. Pfaffl, G.W. Horgan, L. Dempfle, REST 2009 Software User Guide: For gene expression analysis using real-time Sample & Assay Technologies QIAGEN Sample and Assay Technologies, (2009) 1–28.
43. N.K. Singh, D.E. Nelson, D. Kuhn, P.M. Hasegawa, R.A. Bressan, Molecular cloning of osmotin and regulation of its expression by ABA and adaptation to low water potential, Plant. Physiol. 90 (1989) 1096–1101.
44. N.K. Singh, K.R.R. Kumar, D. Kumar, P. Shukla, P.B. Kirti, Characterization of a pathogen induced thaumatin-like protein gene AdTLP from *Arachis diogoi*, a wild peanut, PLoS One. 8 (2013) 1–18.
45. Emanuelsson, S. Brunak, G. von Heijne, H. Nielsen, Locating proteins in the cell using TargetP, SignalP and related tools, Nat. Protoc. 2 (2007) 953–971.
46. Pandey, V. Ramegowda, M. Senthil-Kumar, Shared and unique responses of plants to multiple individual stresses and stress combinations: physiological and molecular mechanisms, Front. Plant Sci. 6 (2015) 1–14.
47. Wang, L. Su, H. Gao, X. Jiang, X. Wu, Y. Li, Q. Zhang, Y. Wang, F. Ren, Genome-Wide Characterization of bHLH Genes in Grape and Analysis of their Potential Relevance to Abiotic Stress Tolerance and Secondary Metabolite Biosynthesis, Front. Plant Sci. 9 (2018) 1–15.
48. Ghosh, C. Chakrabarti, Crystal structure analysis of NP24-I: A thaumatin-like protein, Planta. 228 (2008) 883–890.
49. Velazhahan, S.K. Datta, S. Muthukrishnan, Pathogenesis-related proteins in plants, in: S.K. Datta, S. Muthukrishnan (Eds.), CRC press Boca Raton, Florida, United States, 1999, pp. 107-129.
50. R.C. Misra, Sandeep, M. Kamthan, S. Kumar, S. Ghosh, A thaumatin-like protein of *Ocimum basilicum* confers tolerance to fungal pathogen and abiotic stress in transgenic *Arabidopsis*., Sci. Rep. 6 (2016) 25340.
51. R.D.M. Page, TREEVIEW: An application to display phylogenetic trees on personal computers. Comp Appl Biosci. 12 (1996) 357–358.
52. R.W. Skadsen, P. Sathish, H.F. Kaeppler, Expression of thaumatin-like permatin PR-5 genes switches from the ovary wall to the aleurone in developing barley and oat seeds, Plant Sci. 156 (2000) 11–22.
53. Chowdhury, A. Basu, S. Kundu, Overexpression of a New Osmotin-Like Protein Gene (SindOLP) Confers Tolerance against Biotic and Abiotic Stresses in Sesame, Front. Plant Sci. 8 (2017) 410.
54. Gupta, J.A. Stamatoyannopoulos, T.L. Bailey, W.S. Noble, Quantifying similarity between motifs, Genome Biol. 8 (2007).
55. S. Kumar, G. Stecher, K. Tamura, MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets., Mol. Biol. Evol. 33 (2016) msw054.

 doi:10.20944/preprints202409.0657.v1

56. S. Lonardi, M. Muñoz-amatriaín, Q. Liang, S. Shu, I. Steve, The genome of cowpea (*Vigna unguiculata* [L .] Walp.), (2019) 1–49.
57. S. Mellacheruvu, S. Tamirisa, D.R. Vudem, V.R. Khareedu, Pigeonpea Hybrid-Proline-Rich Protein (CcHyPRP) Confers Biotic and Abiotic Stress Tolerance in Transgenic Rice, Front. Plant Sci. 6 (2016) 1–13.
58. S. Singh, S. Chand, N.K. Singh, T.R. Sharma, Genome-wide distribution, organisation and functional characterization of disease resistance and defence response genes across rice species, PLoS One 10 (2015) 1–29.
59. S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, T. Pennsylvania, U. Park, Basic Local Alignment Search Tool, J Mol Biol. 215 (1990) 403–410.
60. L. Bailey, C. Elkan, Fitting a mixture model by expectation maximization to discover motifs in biopolymers, Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology, pp. 28-36, AAAI Press, Menlo Park, California, 1994.
61. T.N. Petersen, S. Brunak, G. von Heijne, H. Nielsen, SignalP 4.0: discriminating signal peptides from transmembrane regions, Nat. Methods 8 (2011) 785–786.
62. W.C. Hon, M. Griffith, A. Mlynarz, Y.C. Kwok, D.S.C. Yang, Antifreeze proteins in winter rye are similar to pathogenesis-related proteins, Plant Physiol. 109 (1995) 879–889.
63. W.K. Roberts, C.P. Selitrennikoff, Zeamatin, an antifungal protein from maize with membrane-permeabilizing activity, J. Gen. Microbiol. 136 (1990) 1771–1778.
64. X. Wang, C. Tang, L. Deng, G. Cai, X. Liu, B. Liu, Q. Han, H. Buchenauer, G. Wei, D. Han, L. Huang, Z. Kang, Characterization of a pathogenesis-related thaumatin-like protein gene TaPR5 from wheat induced by stripe rust fungus, Physiol. Plant. 139 (2010) 27–38.
65. X. Yan, H. Qiao, X. Zhang, C. Guo, M. Wang, Y. Wang, X. Wang, Analysis of the grape (Vitis vinifera L.) thaumatin-like protein (TLP) gene family and demonstration that TLP29 contributes to disease resistance, Sci. Rep. 7 (2017) 1–14.
66. X. Zhu, L. Qi, X. Liu, S. Cai, H. Xu, R. Huang, J. Li, X. Wei, Z. Zhang, The Wheat Ethylene Response Factor Transcription Factor Pathogen-induced ERF1 Mediates Host Responses to Both the Necrotrophic Pathogen Rhizoctonia cerealis and Freezing Stresses, Plant Physiol. 164 (2014) 1499–1514.
67. Y. Xu, F.L. Chang, D. Liu, M.L. Narasimhan, K.G. Raghothama, P.M. Hasegawa, R.A. Bressan, Plant defense genes are synergistically induced by ethylene and methyl jasmonate, Plant. Cell 6 (1994) 1077–1085.
68. Y. Yang, H. Guan, F. Wei, Z. Li, S. Yang, J. Huang, Genome-wide identification of thaumatin-like protein family genes in Panax notoginseng and analysis of their responses to Fusarium solani infection, Genet Resour Crop Evol (2024) 71, 2267–2279.
69. Y. Zhang, H. Yan, X. Wei, J. Zhang, H. Wang, D. Liu, Expression analysis and functional characterization of a pathogen-induced thaumatin-like gene in wheat conferring enhanced resistance to *Puccinia triticina*, J Plant Interactions 12 (2017) 332–339.
70. Y.C. Jung, H.J. Lee, S.S. Yum, W.Y. Soh, D.Y. Cho, C.K. Auh, T.K. Lee, H.C. Soh, Y.S. Kim, S.C. Lee, Drought-inducible - but ABA-independent - thaumatin-like protein from carrot (*Daucus carota* L.), Plant Cell Rep. 24 (2005) 366–373.
71. Z. Wang, M. Gerstein, M. Snyder, RNA-Seq: a revolutionary tool for transcriptomics, Nature Rev. Genet. 10 (2009) 57–63.