http://communications.science.ankara.edu.tr

Commun.Fac.Sci.Univ.Ank.Ser. C Biology
Volume 31, Number 1, Pages 1-25 (2022)
ISSN 1303-6025 E-ISSN 2651-3749
DOI: 10.53447/communc.1056557

COMMUNICATIONS SERIES C

Research Article; Received: January 11, 2022; Accepted: February 24, 2022

DETAILED CHARACTERIZATION OF LECTIN GENES IN COMMON BEAN USING BIOINFORMATIC TOOLS

AYBÜKE OKAY[1,2], E. SÜMER ARAS[1], İLKER BÜYÜK[1]

[1]Department of Biology, Faculty of Science, Ankara University, Ankara, Turkey
[2]Department of Vaccine Technology, Vaccine Institute, Hacettepe University, Ankara, Turkey

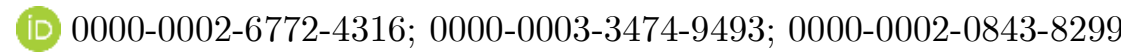
ABSTRACT. Because of differences in molecular structure, biochemical properties, and carbohydrate binding specificity, lectins are considered a complex and heterogeneous group of proteins found in all organisms. Plant lectins are important proteins in terms of their benefits in cancer treatments, biomedical applications, and many medical uses due to their numerous biological roles such as intercellular interactions, defense mechanisms formation, immunomodulation, and anticarcinogenic activity. Despite the discovery of significant amounts of lectin proteins in different plant species, many questions about their potential biological role remain unanswered in *P. vulgaris* L. In this study, using bioinformatics tools, 52 *Pvul-LEC* genes were identified in the *P. vulgaris* genome and these genes were clustered into three subgroups based on phylogenetic analysis. The majority of Pvul-LEC proteins in the same subfamily of phylogenetic tree shared similar motifs and gene structures. Eight pairs of segmental duplications were discovered based on genome wide duplication analysis. Pvul-LEC proteins' three-dimensional structure and functions were also predicted. Simultaneously, gene expression levels of *Pvul-LEC* genes against drought and salt stress in leaf tissues were evaluated based on publicly available RNAseq data. As a result, it is anticipated that the data obtained in the current study will be beneficial to literature and following studies related to lectin genes.

1. INTRODUCTION

Lectins are proteins that bind reversibly to carbohydrate epitopes on polysaccharides, glycoproteins and glycolipids. They can be found in animals, plants, fungi, bacteria and viruses [1]. Because of variances in molecular structure, biochemical characteristics and carbohydrate binding selectivity, plant lectins are thought to constitute a complex and heterogeneous group of proteins.

Keywords. Lectin, in-silico analysis, bioinformatics, common bean, RNAseq

okay.a13@hotmail.com; aras@science.ankara.edu.tr; ilker.buyuk@ankara.edu.tr; -Corresponding author;


0000-0002-6772-4316; 0000-0003-3474-9493; 0000-0002-0843-8299

© 2022 Ankara University
Communications Faculty of Sciences University of Ankara Series C: Biology

Plants contain lectins in a variety of parts, including seeds, leaves, bark, roots, tubers and fruits [2]. Many plant lectins linked to cell motility, cell-cell interaction, embryogenesis and organ formation have been identified as secretory proteins, meaning they enter the secretory system and accumulate in vacuoles, cell walls, or intercellular spaces [3-6].

Plants use lectins to protect themselves from insects and fungus, as well as to transport and store sugar [7]. To defend themselves against invading diseases, plants have evolved an innate immune system [8]. One of the major hormones regulating the pathogen-induced immune response in plants is salicylic acid (SA). The activation of a local sensitivity reaction that inhibits the biotrophic pathogen's proliferation is typical of this response [9]. According to the previous studies, lectin genes play a crucial role in SA biosynthesis. Using *Pseudomonas syringae* pv. avirulence strains on tomato, Armijo et al. (2013) found that increased expression levels of *SAI-LLP* genes triggered cells to avoid leaf death, suggesting that lectin genes are components of SA-mediated defensive pathways [10]. Luo et al. (2017) showed that *LecRK* genes have a role in plant immunological signaling and SA buildup [11]. Furthermore, some lectins are critical for nitrogen fixation in the atmosphere [7]. Lectins operate as immunomodulatory molecules in addition to having antinutritional effects [12-14]. They can adhere to the surface of epithelial cells in the digestive system and generate harmful responses in intestinal permeability because this category of proteins is always resistant to destruction by temperature and digestive enzymes [15, 16].

Plant lectins have also been reported to have biological and medical applications, including the isolation of glycoconjugates from cells, microorganism recognition, monitoring of changes in carbohydrate expression on living cells, mitogenic simulation, anti-proliferative effects, anti-tumor, and drug targeting to the gastrointestinal tract [17-20]. Lectins are one of the most studied protein groups because to their extraordinary biological roles [18]. They are found in legumes, which are an essential food source for both humans and animals. *Phaseolus vulgaris*, a legume, is grown for direct human consumption in a number of countries. Protein (22-27 percent of seed weight) and carbs (39-47 percent of seed weight) are abundant in *P. vulgaris* L., making it an excellent food. Although lectins have been discovered in large quantities in *P. vulgaris*, many issues about their biological function remain unanswered [21, 22].

Plant lectins, specifically *P. vulgaris* lectins, have been found to inhibit leukemia viruses, acquired immunodeficiency syndrome (AIDS) virus, and coronaviruses potently and selectively [23, 24]. As a result, *P. vulgaris* lectins exceptional anti-HIV effectiveness has become a major research topic in recent years. Total parenteral nutrition (TPN) is a valuable medical treatment option that satisfies all of the body's nutritional requirements. Because they can bind to diverse areas of the gastrointestinal tract and induce epithelial proliferation, *P. vulgaris* lectins have recently been proposed as medicines to prevent mucosal atrophy and related

difficulties when patients receive TPN [25]. Furthermore, in recent years, lectin-mediated drug delivery methods have spawned fascinating fields of research in terms of lectin binding to receptor-like structures of the cell membrane and epithelial cells and initiating active transport of medicinal components [26].

Consumption of *P. vulgaris* lectins is critical due to multiple health hazards and major pharmacological effects in humans, however not all of the *Pvul-LEC* genes have been found and characterized so far. For this reason, this study will be used to strategically accompany future studies by discovering *Pvul-LEC* genes utilizing bioinformatics methods.

2. MATERIALS AND METHODS

2.1. Identification of lectin proteins in *P. vulgaris* genome

P. vulgaris lectin family sequences were obtained from NCBI (http https://www.ncbi.nlm.nih.gov/) and Pfam databases [27]. Putative *P. vulgaris* lectin proteins were used for query in blastp (NCBI) for characterization of hypothetical proteins. The physicochemical properties of lectin proteins were calculated using ProtParam Tool (http:web.expasy.org/protparam) and detection of domains was performed using HMMER (http:www.ebi.ac.uk/Tools/hmmer/).

2.2. Structure and physical locations of lectin genes and conserved motifs

'Gene Structure Display Server v2.0' (GSDS, http:// gsds.gao-lab.org) was used to illustrate the exon–intron structure of *Pvul-LEC* genes [28]. The lectin genes have been mapped on *P. vulgaris* chromosomes using the MapChart program [29]. Multiple Expectation Maximizations for Motif Elicitation tool (EM) was used to categorize additional conserved motifs for Pvul-LEC proteins (MEME 4.11.1; http://meme-suite.org/) [30].

2.3. Phylogenetic analysis and sequence alignment

The ClustalW was used to align various sequences of lectin proteins [31]. For the creation of phylogenetic trees with a bootstrap value of 1000 replicates (MEGA11), the Neighbor-joining (NJ) method was used, and the tree was constructed using an Interactive Life Tree (iTOL; http://itol.embl.de/index.shtml) [32].

2.4. Promoter analysis of lectin genes

The PlantCARE database was used to do a cis element scan on the 5′ upstream

regions (2 kb of DNA sequence from each *Pvul-LEC* gene) using the Phytozome database v11 (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/).

2.5. In-silico prediction of miRNA targets in lectin genes

MiRBase v21.0. has been used to download all known miRNA plant sequences (http://www.mirbase.org). The default miRNA prediction parameters were applied using the psRNA Target Server (http://plantgrn.noble.org/psRNATarget) [33]. BLASTX was used to search in-silico predicted miRNA targets against common bean Expressed Sequenced Tags (ESTs) in the NCBI database with a $1e^{-10}$ threshold.

2.6. Detection of gene duplication events and prediction of synonymous and non-synonymous substitution rates

The Plant Genome Duplication Database service (http://chibba.agtec.uga.edu/duplication/index/locus) was used to examine duplicated gene pairs with a display range of 2000 kb. The amino acid sequences of duplicated lectin genes were predicted using the CLUSTALW program. To estimate synonymous (Ks) and non-synonymous (Ka) substitution rates, the PAML (PAL2NAL) CODEML program (http://www.bork.embl.de/pal2nal) was used [34]. The following formula was used to compute the duplication period (Million Years Ago, Mya) and divergence of each lectin gene: $T = Ks/2\lambda$ ($\lambda = 6.56E^{-9}$) [35].

2.7. Identification of expression level of lectin genes through transcriptome data

Illumina RNA-seq data was collected from the Sequence Read Archive (SRA) to measure the *Pvul-LEC* gene expression levels. For this reason, the accession numbers SRR957667 (control leaf) and SRR957668 (salt-treated leaf) were used as defined by Büyük et al. (2016) [36, 37]. The heat maps of hierarchical clustering were eventually built using the CIMMiner algorithm (http://discover.nci.nih.gov/cimminer).

2.8. Homology modeling of lectin proteins

BLASTP (with default parameters) was used to search all lectin proteins against the Protein Data Bank (PDB) to identify the best template(s) with identical sequence and three-dimensional structure [38]. The data was loaded into Phyre2 (Protein Homology/AnalogY Recognition Engine; http://www.sbg.bio.ic.ac.uk/phyre2), which used homology modeling in 'intense' mode to predict protein structure [39].

3. RESULTS AND DISCUSSION

3.1. Determination of *Pvul-LEC* genes in the genome of *P. vulgaris*

The analyses resulted in the discovery of 52 lectin genes encoded in the *P. vulgaris* genome. The order of chromosome placement was taken into account when naming the genes, and all lectin genes were categorized into two groups based on their domain types: Legume lectin domain (Lectin legB) and L-Type Lectin. According to this, starting from *Pvul-LLEC-1* to *Pvul-LLEC-30*, and starting from *Pvul-BLEC-1* to *Pvul-BLEC-22*, the naming process was carried out using the Latin name of the bean plant (*Phaseolus vulgaris*).

According to the studies, lectin legB plays a critical role in biotic and abiotic stress responses. For example, Jiang et al. (2010) discovered that lectin *legB* genes revealed a wide range of expression profiles, especially under salt, cold and drought stress and that they may operate as a regulator of environmental conditions. With their work on *Arabidopsis*, it was also hypothesized that lectin legB and the kinase domains, a structurally conserved protein domain that carries out the phosphorylation process common to all living things, may have duplicated to protect against stressors during the evolutionary process [40].

L-type lectins are mostly present in the seeds of leguminous plants, and they are delivered to the vacuole where they are condensed by specific vesicles after being generated during seed development a few weeks after blooming. L-type lectins remain stable during the drying process of the seeds and can do so indefinitely until the seeds germinate. They are typically referred to as storage proteins because they are one of several types of proteins that are kept in high concentrations in seeds [41, 42]. L-type lectins, which have structural motifs seen in glycan-binding proteins, are also found in the bark of some legume trees and to a lesser extent in other vegetal tissues of legumes (GBPs). Most of their structures have been determined, and numerous L-type lectins have been reported to be employed in a variety of biological and analytical techniques [43]. L-type lectins also play a part in plant-nitrogen-fixing bacteria symbiosis, such as Rhizobium-legumes symbiosis, however the exact role of lectins in this system is unknown. However, investigations demonstrate that the seed lectin is a lipoxygenase with the activity required to start the plant defense pathway [7, 44].

So far, lectin genes have been identified in several plant species. The number of lectin genes (n=52) in *P. vulgaris* in this study is higher than the number identified in *Cucumis sativus* (n=46) [45] however less than the number identified in *Arabidopsis thaliana* (n=72) [46], *Setaria italica* (n=120), *Zea mays* (n=126), *Oryza sativa* (n=153), *Sorghum bicolor* (n=153), *Brachypodium distachyon* (n=204) and *Saccharum spontaneum* (n=429) [47].

Protein length (aa), the isoelectric point (pI), molecular weight (Da), instability index, aliphatic index and GRAVY value were calculated for all identified Pvul-LEC proteins. Accordingly, the lengths of Pvul-LEC proteins ranged from 259 to 699 amino acids. Additionaly, 10 Pvul-LEC proteins were found to have basic and 42 Pvul-LEC proteins have acidic properties (Table 1). The molecular weights of Pvul-LEC proteins were between 26.96 to 78.50 kDa and 5 Pvul-LEC proteins were found to be unstable according to instability index scores. The most of the Pvul-LEC proteins were found to be hydrophilic according to the GRAVY values which ranged between -0.294 to 0.229 (Table 1).

The aliphatic index value, defined as the relative volume occupied by the aliphatic side chains (alanine, valine, isoleucine and leucine) can be considered to be a positive factor for enhancing global protein thermostability [48]. These values ranged from 73.71 to 96.06 for Pvul-LECs, suggesting the strong thermostability of these proteins.

All *Pvul-LEC* genes were found to be unevenly distributed on 10 out 11 chromosomes of *P. vulgaris* with the exception of chromosome 9. chromosome 1 contained the most *Pvul-LEC* genes (eleven), whereas chromosome 11 contained only one (Figure 1). Surprisingly, chromosome 4 contained only *Pvul-BLEC* genes while chromosome 5 and 6 contained only *Pvul-LLEC* genes. Dang et al. (2016) discovered that the lectin gene was found in all chromosomes (7 chromosomes) of *C. sativus* and at different rates, similar to *Pvul-LEC* genes [49]. In another study on the mungbean plant, Singh et al. (2021) classified lectin genes into three types: G-type, L-type, and C-type, and 59 of 73 *VrLecRLK* were randomly distributed on 11 chromosomes. Furthermore, the remaining *VrLecRLK* genes in the study were discovered to be unattached to any specific chromosome [50].

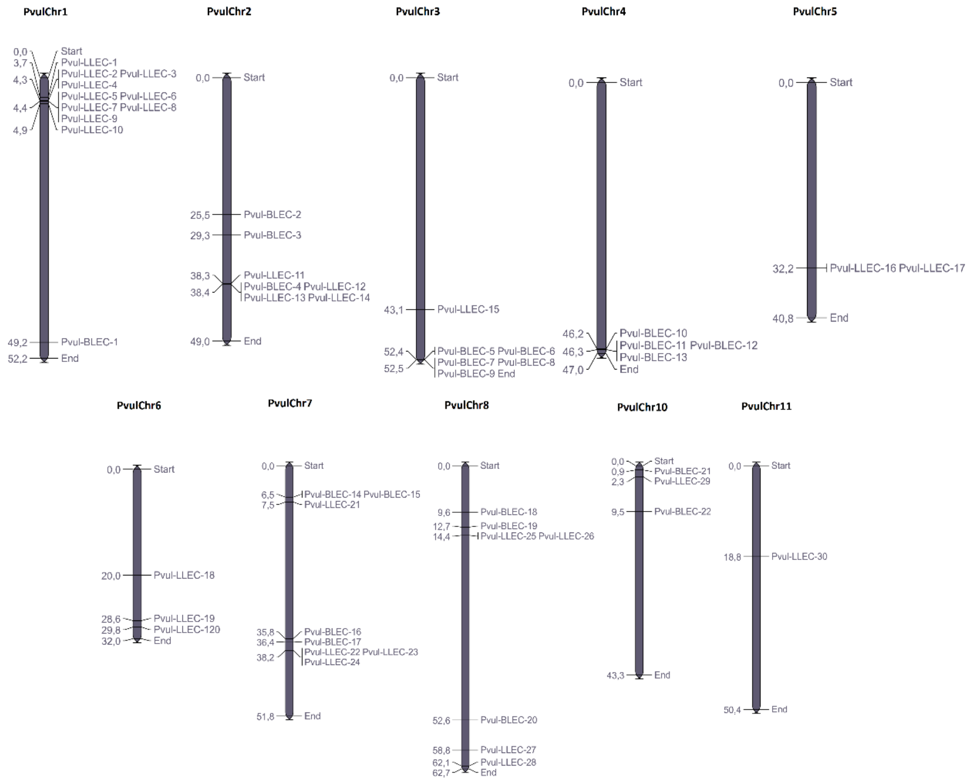


FIGURE 1. Distribution of *Pvul-LEC* genes on *P. vulgaris* chromosomes

TABLE 1. Detailed descriptive information for Pvul-LLEC proteins

ID	*Phaseolus vulgaris* Genomic Database Identifier	Physical position on *P. vulgaris* genome			Protein length (aa)	pI	Molecular weight (kDa)	Instability index	Aliphatic index	GRAVY	NCBI Accession No.
		Chr.	Start position (bp)	End Position (bp)							
Pvul-LLEC-1	Phvul.001G045400.1.p	1	3.677.580	3.680.026	667	6.04	75.27	32.38	87.06	-0.176	XP_007161136.1
Pvul-LLEC-2	Phvul.001G040800.1.p	1	4.327.634	4.329.496	620	5.68	68.75	27.71	90.05	-0.085	XP_007161076.1
Pvul-LLEC-3	Phvul.001G040700.1.p	1	4.336.106	4.338.606	672	5.77	74.99	28.89	83.11	-0.189	XP_007161075.1
Pvul-LLEC-4	Phvul.001G040600.1.p	1	4.339.156	4.340.979	607	6.01	67.9	31.81	94.43	-0.064	XP_007161074.1
Pvul-LLEC-5	Phvul.001G040500.1.p	1	4.359.201	4.361.618	661	5.73	73.92	25.07	89.05	-0.14	XP_007161073.1
Pvul-LLEC-6	Phvul.001G040400.1.p	1	4.370.066	4.372.118	636	5.62	70.87	34.63	91.89	-0.142	XP_007161072.1
Pvul-LLEC-7	Phvul.001G040300.1.p	1	4.379.952	4.382.583	664	5.63	73.98	25.57	87.91	-0.161	XP_007161071.1
Pvul-LLEC-8	Phvul.001G040100.1.p	1	4.394.207	4.396.266	636	5.62	71.1	34.65	90.66	-0.154	XP_007161068.1
Pvul-LLEC-9	Phvul.001G040000.1.p	1	4.399.520	4.401.949	666	5.83	74.15	25.7	88.11	-0.129	XP_007161067.1
Pvul-LLEC-10	Phvul.001G234200.1.p	1	4.880.986	48.811.980	664	8.68	73.85	33.83	92.3	-0.104	XP_007163432.1
Pvul-LLEC-11	Phvul.002G214900.1.p	2	38.345.248	38.347.335	695	5.85	77.3	32.79	87.8	-0.128	XP_007159169.1
Pvul-LLEC-12	Phvul.002G215200.1.p	2	38.371.310	38.373.253	647	7.31	72.85	30.71	85.09	-0.229	XP_007159173.1
Pvul-LLEC-13	Phvul.002G215300.1.p	2	38.383.003	38.384.994	663	7.32	74.17	31.47	90.89	-0.201	XP_007159174.1
Pvul-LLEC-14	Phvul.002G215400.1.p	2	38.393.780	38.395.786	668	6.75	75.53	36.9	91.44	-0.147	XP_007159175.1
Pvul-LLEC-15	Phvul.003G204500.1.p	3	43.061.296	43.063.308	670	5.11	73.7	42.72	86.1	-0.074	XP_007155476.1
Pvul-LLEC-16	Phvul.005G103200.1.p	5	32.195.113	32.197.596	691	5.99	77.7	41.04	86.43	-0.115	XP_007149844.1
Pvul-LLEC-17	Phvul.005G103300.1.p	5	32.198.665	32.200.731	688	6.34	77.07	34.86	91.89	-0.078	XP_007149845.1
Pvul-LLEC-18	Phvul.006G087700.1.p	6	19.974.245	19.976.281	678	5.75	72.71	36.68	85.27	-0.04	XP_007146994.1
Pvul-LLEC-19	Phvul.006G185000.1.p	6	28.613.191	28.616.354	692	5.51	78.5	36.34	89.22	-0.143	XP_007148157.1
Pvul-LLEC-20	Phvul.006G200800.1.p	6	29.750.251	29.752.428	639	6.96	72.04	36.62	88.17	-0.155	XP_007148350.1
Pvul-LLEC-21	Phvul.007G078200.1.p	7	7.489.146	7.491.182	678	6.24	75.42	41.56	84.1	-0.087	XP_007143518.1
Pvul-LLEC-22	Phvul.007G260300.1.p	7	38.169.156	38.172.057	670	6.25	74.86	36.79	84.51	-0.263	XP_007145693.1
Pvul-LLEC-23	Phvul.007G260400.2.p	7	38.176.633	38.178.637	450	7.61	50.66	39.48	84.6	-0.294	XP_007145694.1
Pvul-LLEC-24	Phvul.007G260500.1.p	7	38.182.876	38.185.141	657	6.43	72.94	37.51	88.23	-0.212	XP_007145695.1
Pvul-LLEC-25	Phvul.008G117700.1.p	8	14.407.039	14.409.057	672	7.01	74.2	34.28	99.33	-0.066	XP_007140498.1
Pvul-LLEC-26	Phvul.008G117800.1.p	8	14.436.075	14.438.400	668	7.32	73.97	33.83	94.03	-0.115	XP_007140499.1
Pvul-LLEC-27	Phvul.008G239600.1.p	8	58.823.983	58.826.861	699	5.59	74.93	33.42	87.78	-0.018	XP_007141948.1
Pvul-LLEC-28	Phvul.008G279300.1.p	8	62.071.477	62.073.771	662	6.64	72.8	38.34	86.56	-0.003	XP_007142422.1
Pvul-LLEC-29	Phvul.010G015800.1.p	10	2.345.298	2.347.358	686	5.74	75.22	36.4	93.76	-0.068	XP_007134058.1
Pvul-LLEC-30	Phvul.011G119200.1.p	11	18.774.057	18.775.652	531	5.98	59.06	34.97	94.11	-0.007	XP_007132721.1

Table 2. Detailed descriptive information for Pvul-BLEC proteins

ID	*Phaseolus vulgaris* Genomic Database Identifier	Physical position on *P. vulgaris* genome			Protein length (aa)	pI	Molecular weight (Da)	Instability index	Aliphatic index	GRAVY	NCBI Accession No.
		Chr.	Start position (bp)	End Position (bp)							
Pvul-BLEC-1	Phvul.001G239200.1.p	1	49.244.673	49.246.082	345	5.28	37.51	28.22	89.77	-0.126	XP_007163496.1
Pvul-BLEC-2	Phvul.002G119900.1.p	2	25.542.601	25.545.814	693	6.16	76.67	33.71	90.71	-0.127	XP_007158049.1
Pvul-BLEC-3	Phvul.002G144200.1.p	2	29.269.216	29.270.007	263	4.72	28.25	34.77	87.07	0.102	XP_007158339.1
Pvul-BLEC-4	Phvul.002G215000.1.p	2	38.350.955	38.353.003	664	6.01	73.69	35.44	88.09	-0.133	XP_007159169.1
Pvul-BLEC-5	Phvul.003G286700.1.p	3	52.442.001	52.442.780	259	4.87	27.66	32.2	88.84	0.214	XP_007156448.1
Pvul-BLEC-6	Phvul.003G286800.1.p	3	52.444.624	52.445.439	271	4.93	29.7	30.56	87.05	0.049	XP_007156449.1
Pvul-BLEC-7	Phvul.003G286900.1.p	3	52.450.066	52.450.845	259	6.21	27.85	36.21	93.4	0.223	XP_007156450.1
Pvul-BLEC-8	Phvul.003G287000.1.p	3	52.452.303	52.453.082	259	4.69	27.72	32.79	88.84	0.229	XP_007156451.1
Pvul-BLEC-9	Phvul.003G287101.1.p	3	52.459.994	52.462.334	449	6.35	49.81	36.53	96.06	-0.032	XP_007156452.1
Pvul-BLEC-10	Phvul.004G158000.1.p	4	46.239.426	46.240.567	277	4.73	29.97	24.31	89.03	-0.02	XP_007152769.1
Pvul-BLEC-11	Phvul.004G158100.1.p	4	46.258.583	46.259.432	244	5.03	26.96	33.15	83.85	-0.22	XP_007152770.1
Pvul-BLEC-12	Phvul.004G158200.1.p	4	46.261.905	46.262.865	279	5.15	30.26	21.46	95.05	0.031	XP_007152771.1
Pvul-BLEC-13	Phvul.004G158300.1.p	4	46.267.391	46.268.336	272	4.92	29.55	26.45	89.56	-0.041	XP_007152772.1
Pvul-BLEC-14	Phvul.007G070100.1.p	7	6.474.389	6.475.216	275	5.37	30.64	31.26	73.71	-0.29	XP_007143413.1
Pvul-BLEC-15	Phvul.007G070300.1.p	7	6.481.954	6.483.109	299	5.3	32.91	27.95	79.57	-0.223	XP_007143415.1
Pvul-BLEC-16	Phvul.007G234400.1.p	7	35.840.909	35.843.746	643	7.96	70.63	41.18	84.07	-0.189	XP_007145382.1
Pvul-BLEC-17	Phvul.007G239800.1.p	7	36.355.865	36.356.814	274	5.53	29.36	34.33	91.5	0.039	XP_007145445.1
Pvul-BLEC-18	Phvul.008G094500.1.p	8	9.626.760	9.628.781	673	5.25	74.54	34.81	88.78	-0.156	XP_007140223.1
Pvul-BLEC-19	Phvul.008G111500.1.p	8	12.658.909	12.659.979	356	9.28	39.36	34.84	82.33	0.057	XP_007140433.1
Pvul-BLEC-20	Phvul.008G188700.1.p	8	52.631.783	52.634.655	677	7.28	75.82	38.29	85.64	-0.18	XP_007141356.1
Pvul-BLEC-21	Phvul.010G006300.1.p	10	923.471	924.954	348	9.2	38.09	32.36	85.34	0.063	XP_007133954.1
Pvul-BLEC-22	Phvul.010G060800.1.p	10	9.503.176	9.505.775	662	5.47	73.7	45.83	89.08	-0.096	XP_007134605.1

In order to expand and diversify, genes engage in a dynamic replication process which is referred to as gene duplication. Specialized genes or new gene functions can arise as a result of these evolutionary processes. For this reason, the location of *Pvul-LEC* genes on *P. vulgaris* chromosomes was investigated, as well as their duplications. While no tandem duplication events were discovered, 8 pairs of segmental duplication events were detected in *Pvul-LEC* genes. The duplicated genes homologous (Ks) and non-homologous (Ka) exchange rates were calculated (Table 3). Accordingly, the calculated mean Ka/Ks ratios were found to be 0.12 and hence this is a purifying selection since Ka / Ks <1 [51]. In addition, it was determined that the average time of duplications was 76.62 MYA (Table 3).

TABLE 3. Information regarding segmental duplication events between *Pvul-LEC* genes (Ka: non-synonymous divergence value; Ks: synonymous divergence value; Ka/Ks: rate of change; MYA: Million years ago)

Gene IDs	Gene IDs	Ka	Ks	Ka/Ks	MYA
Pvul-BLEC-3	*Pvul-BLEC-5*	46.1974	0.4735	0.0102	355,36
Pvul-LLEC-10	*Pvul-LLEC-25*	4.0418	0.2910	0.0720	31,09
Pvul-LLEC-16	*Pvul-LLEC-19*	3.0242	0.3787	0.1252	23,26
Pvul-LLEC-16	*Pvul-LLEC-30*	2.6967	0.3044	0.1129	20,74
Pvul-BLEC-18	*Pvul-LLEC-29*	11.3547	0.6572	0.0579	87,34
Pvul-LLEC-18	*Pvul-LLEC-27*	1.1548	0.1411	0.1222	88,83
Pvul-BLEC-19	*Pvul-BLEC-21*	0.7361	0.2274	0.3089	5,66
Pvul-LLEC-25	*Pvul-LLEC-29*	0.6383	0.1574	0.2465	4,91

A comparison was made between the 52 *Pvul-LEC* genes identified in the *P. vulgaris* genome and the lectin genes of *A. thaliana* and *G. max.* (Supplementary Table 1). As a result, 11 orthologous gene pairs were found to be shared between *Phaseolus vulgaris* and *Arabidopsis thaliana*, and 8 orthologous pairs were shared between *Phaseolus vulgaris* and *Glycine max*. The average Ka/Ks ratios were 0.07 and 0.028, respectively (Supplementary table 1).

When orthologous lectin genes between *P. vulgaris* and *G. max* were analyzed, it was discovered that approximately 88% of them were *Pvul-BLEC* genes. This situation has increased the possibility that the lectin *legB* genes were duplicated in the evolutionary process in parallel with the findings of Jiang et al. (2010) [40]. The number of *Pvul-LLEC* genes was compared to the total number of protein-coding genes in different genomes. Accordingly, *G. max, A. thaliana, T. aestivum* and *O. sativa* had rates of 60/56044 (0.107%), 42/27416 (0.153%), 84/99386 (0.085%), and 72/22273 (0.323%), respectively. In this study, the

number of protein-coding genes in *P. vulgaris* was determined to be 85,167, while the number of L-type genes was discovered to be 30, yielding a 30/85167 (0.035%) ratio. Various *LLEC* gene growth rates in different genomes, as well as changes in genome size, could explain differences in *Pvul-LLEC* gene copy counts. Furthermore, it was discovered that only *A. thaliana*, which included *Pvul-LEC* genes, increased the L-type genes to a larger extent than the other types (55-57).

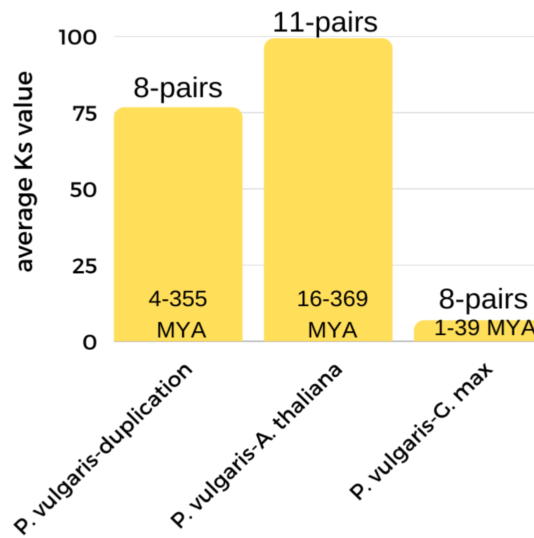


FIGURE 2. Time of duplication and divergence (MYA) based on synonymous substitituon rate (Ks).

3.2. Structure of *Pvul-LEC* genes, phylogenetic tree analysis, determination of preserved motifs, homology modeling and promoter analysis

The structure of *Pvul-LEC* genes was investigated using exon and intron regions. Exon and intron regions play a significant role in modulating gene expression and regulating genes differently [52, 53]. There were no introns discovered in 42 of the *Pvul-LEC* genes. The *Pvul-BLEC-16* gene was discovered to be the only one with multiple intron regions among all the genes examined. Furthermore, this gene has been found to have the most recent orthologous relationship with Glyma.20G072500, a *Glycine max LEC* gene (MYA=1.66). In a study on *Vigna radiata* (mung bean) [50], Singh et al. (2021) discovered that 27 *VrLecRLK* genes (out of 73 total) do not have introns. In soybean, Liu et al. (2018) discovered that nearly 70% of 60 L-type *GmLec* genes do not have intron regions [54]. The presence or lack of introns in many genomes has evolutionary

advantages for a gene's protein output by changing numerous characteristics, such as reverse transcriptase activity or mRNA stability [55].

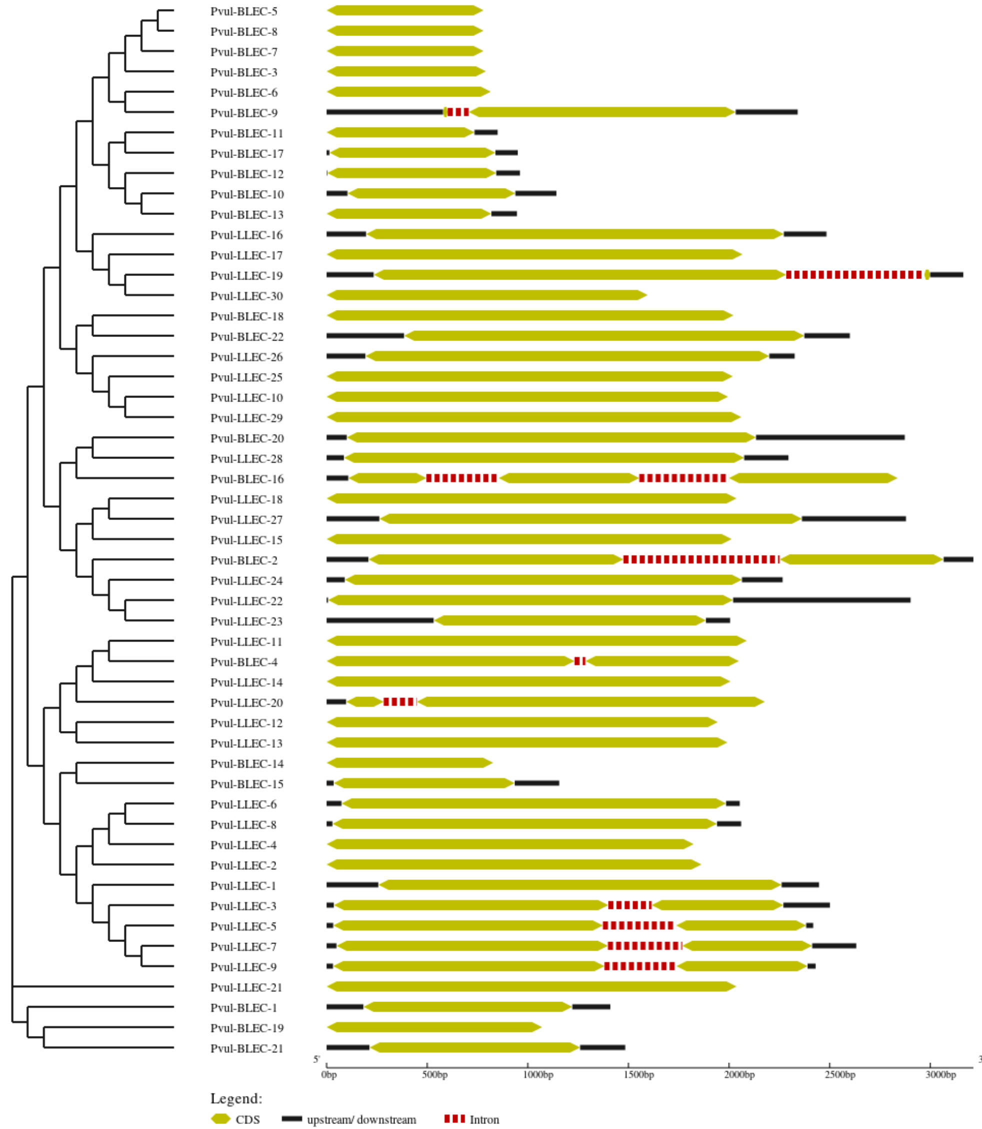


FIGURE 3. Gene structures of lectin family members from *P. vulgaris* with clustering based on NJ phylogenetic tree. Introns are presented by lines. UTR and CDS are indicated by filled black and green boxes, respectively.

A phylogenetic tree was drawn using the neighbor joining method using the evolutionary processes of different organisms (*P. vulgaris, A. thaliana* and *G. max*) to examine the kinship for lectin genes. The examined phylogenetic tree with 196 members was found to be divided into three major groups.

According to this, the smallest group was 'A' with 3 genes and the largest group was 'D' with 106 genes (D1-30/D2-76). Group 'B' and 'C' was found to contain 45 and 42 genes, respectively. The all lectin genes in group 'B' belong to *Pvul-BLEC* family.

Groups C and D contained all of the intron-containing genes. Genes with identical motif content weree clustered in the same groups of the phylogenetic tree; *Pvul-BLEC-5, Pvul-BLEC-8, Pvul-BLEC-11, Pvul-BLEC-13, Pvul-BLEC-15* and *Pvul-BLEC-17* in group B, *Pvul-LLEC-1, Pvul-LLEC-2, Pvul-LLEC-5, Pvul-LLEC-6, Pvul-LLEC-7, Pvul-LLEC-8, Pvul-LLEC-9* and *Pvul-LLEC-14* were in group *C,* while *Pvul-LLEC-17* and *Pvul-LLEC-19* were in group D2.

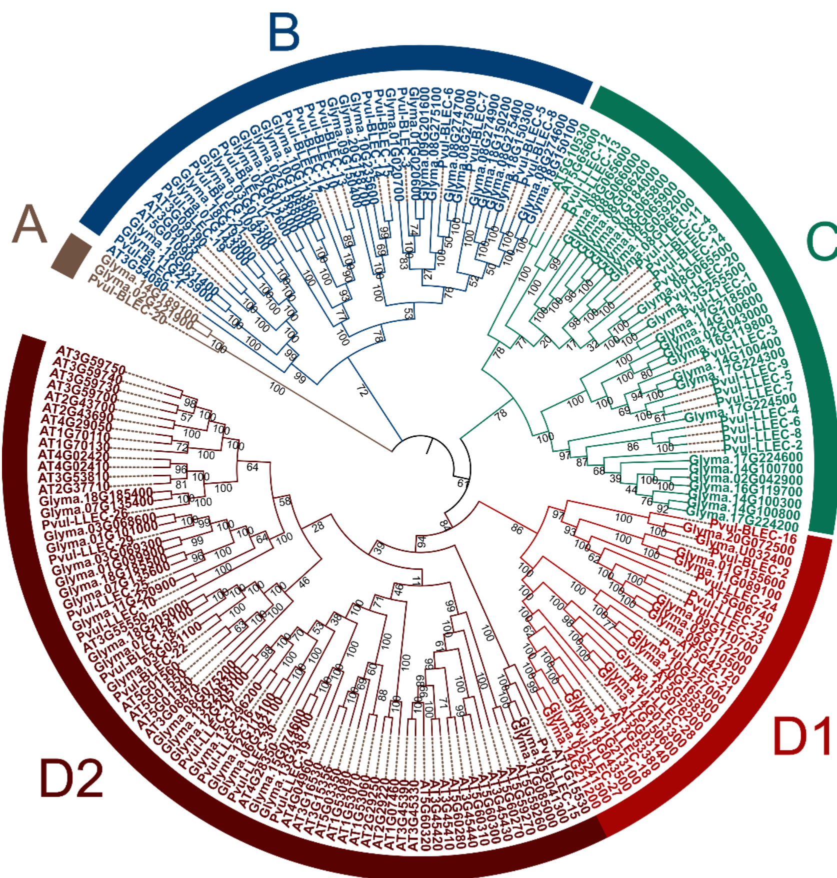


FIGURE 4. Phylogenetic analyses of lectin proteins from three plant species. The phylogenetic tree was constructed using the NJ method. The identifier names of lectin proteins of *Phaseolus vulgaris* L., *Arabidopsis thaliana* and *Glycine max* start with 'Pvul', 'AT' and 'Glyma', respectively.

In genome investigations, the detection of conserved motifs is critical for protein identification and categorization, as well as the identification of functional areas and particular binding sites [56]. In this context, we evaluated the conserved motif structures of lectin proteins in the *P. vulgaris* genome, and the results were reported in Supplementary Table 2. Accordingly, twenty conserved motif patterns were found in 52 Pvul-LEC proteins. As a result, all *Pvul-LLEC* genes were found to contain Motifs 1, 2, 3, 5, and 9, while all *Pvul-BLEC* genes contained Motif 14.

In this study, homology was determined using modeling of Pvul-LEC proteins with a confidence level of 90% or higher in order to obtain information about the protein structures. As a result, all of the proteins were found to be 90% or more reliable belong to the Pvul-BLEC family. Legume lectin loops have a jelly roll subtopology, also known as a superfold [57]. This fold is thought to have had multiple evolutionary origins rather than a single common ancestor [58]. The fact that approximately 88 percent of the existing orthologous genes located between *Pvul-LEC* and *G. max* are *Pvul-BLEC* genes supports the findings.

The hydrogen bond pattern between adjacent strands is broken in two places, resulting in the formation of two four-stranded sheets. Because adjacent strips appear in different layers, both layers are completely anti-parallel, with the exception of the fourth and fifth strands, which appear in the same layer. This results in a structure with a single hairpin and arched - junctions [59]. When Pvul-BLEC proteins were examined, anti-parallel – structures could be found in all of them. In addition to these structures, almost all Pvul-BLEC proteins had alpha structures (Figure 5).

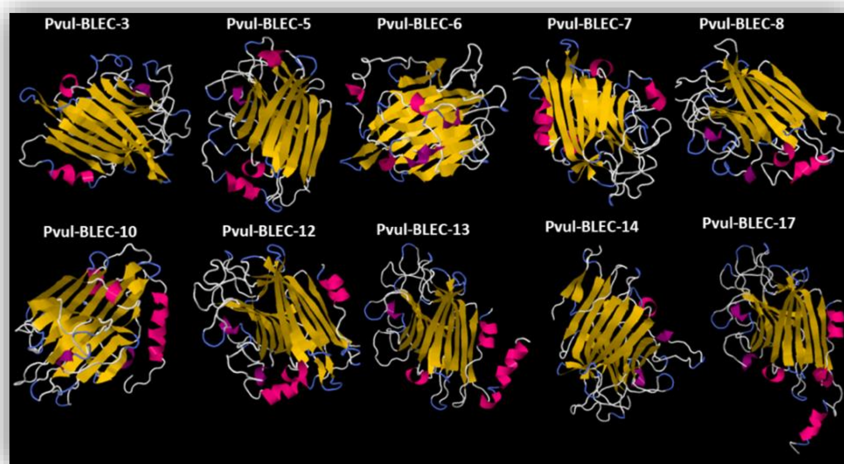


FIGURE 5. Predicted 3D models of common bean Pvul-BLEC proteins. Models were generated by using Phyre2 server. The secondary structure elements: α-helices (pink), β-sheets (yellow), and coils (blue-white) are indicated for the predicted 3D structures of Pvul-BLECs.

Cis-acting regulatory sequence elements are involved in the transcriptional regulation of gene activities, which control many biological processes, particularly abiotic stress responses, hormones, and development [60]. In this study, in silico promoter analysis was done in order to better understand the regulatory mechanisms of *Pvul-LEC* genes. The discovered cis-acting elements were classified into following eight categories: development, environmental stress, hormone, light, promoter, site binding, biotic stress, and others (Supplementary Table 3).

Plants require a variety of signaling chemicals, including SA (Salicyclic acid), JA (Jasmonic acid), ABA (Abscisic acid), and ET (Ethylene) to respond to stress. MYCs, which are the major regulators of the JA signaling pathway and are categorized in the environmental stress category, are found in roughly 55% (29 genes) of *Pvul-LEC* genes. MYB, which is categorized in the biotic stress category and is known to have a role in the regulation of auxins as well as drought, salt, and abscisic acid, was discovered to be present in around 53% of *Pvul-LEC* genes (28 genes). In the promoters of 20, 19, and 11 *Pvul-LEC* genes, ABA-sensitive elements (ABREs), MeJA-responsive elements (CGTCA-pattern), and SA-responsive elements (TCA-elements) were found respectively. Additionaly, *Pvul-LLEC-12* and *Pvul-BLEC-6* genes, which both have ABRE, CGTCA motif, TCA-element and WUN-motifs, were found to reveal almost same expression patterns according to the heatmaps drawn in this study (Figure 6 and 7).

3.3. Detection of miRNAs targeting *Pvul-LEC* genes

Plants exposed to abiotic stress factors develop response mechanisms at the transcriptome level. MiRNAs, in addition to TF in restoring metabolism, are involved in the regulation of gene expression [61, 62]. Under abiotic stress conditions, plants can increase or decrease the efficiency of miRNAs or synthesize new miRNAs [62). MiRNAs and associated *Pvul-LEC* genes were identified in this study (Supplementary Table 4). A total of 8226 targets were examined, and it was discovered that the most targeted gene was *Pvul-LLEC-14* among all *Pvul-LLEC* genes, and the most targeted gene was *Pvul-BLEC-4* among all *Pvul-BLEC* genes. It was worth noting that both genes were clustered together in the phylogenetic tree's group C (Figure 4).

MiRNAs targeting *Pvul-LLEC-14* gene, which was most targeted *Pvul-LLEC* gene, were studied based on the targeting frequency. Among these, miR169 is a miRNA which involves in early flowering. Additionaly, in a study to grow resistant tobacco varieties, Jianyu et al. (2019) discovered that miRNA169 played a role in response to low temperature stress. In the same study, LTR elements were detected in the promoter region of the genes targeted by miR169 in tobacco [63]. Similarly, an LTR element was also found in the promoter region of the *Pvul-LLEC-14* gene. Based on this information, it is believed that the

miR169 and LTR element may interact to reveal stress response in *P. vulgaris* under stress conditions.

MiR395, is an another miRNA targeted *Pvul-LLEC-14* gene, has been identified as a regulatory component of the sulfate assimilation network [64]. The expression of miR395 was also found to change in response to environmental stimuli [65]. MiR5241 was found to be significantly increased in a study by Cao et al (2018) that examined the responses of miRNAs to salt and alkali stress in *Medicago truncatula* [66].

At the same time, the frequency of miRNA targeting of the *Pvul-BLEC-4* gene was investigated. Plants have a set of defensive responses to prevent pathogen invasion. Some of these defense mechanisms are built into the plant, while others can be triggered by pathogen detection. According to our findings, miR160 is involved in the local defense mechanism, and Natarjan et al. (2018) discovered an increased susceptibility to infection in potato associated with miR160 regulation [67].

We hypothesize that the interaction of *Pvul-BLEC-4* - miR160 - SA may have an antagonistic effect against infections because lectin genes are known to play a central role in SA biosynthesis. According to the literature, miR172 confers plant tolerance, particularly against salt stress and water deficiency, but increases ABA sensitivity. In a study on Arabidopsis, it was discovered that overexpression of miR172 kept the plant alive [68].

3.4. Tissue-specific mRNA levels of the *Pvul-LEC* genes

The expression data of *Pvul-LEC* genes in various tissues (flower buds, flowers, leaves, stem 10, young pods, stem 19, young trifoliats, root 10, root 19, green mature pods and nodules) was obtained from the Phytosome v12 database and a heat map was plotted (Figure 6). Accordingly, *Pvul-BLEC-1, Pvul-LLEC-15, Pvul-LLEC-16, Pvul-LLEC-18* and *Pvul-LLEC-27* genes were found to be highly expressed in all tissues in comparison to the other *Pvul-LEC* genes (Figure 6). Genes with similar expression levels were surprisingly found to be located in the same phylogenetic groups (Figure 4 and Figure 6). For example, *Pvul-LLEC-15, Pvul-LLEC-16, Pvul-LLEC-18* and *Pvul-LLEC-27* genes with similar expression levels were found to be located in D1-D2 groups according to the phylogenetic tree.

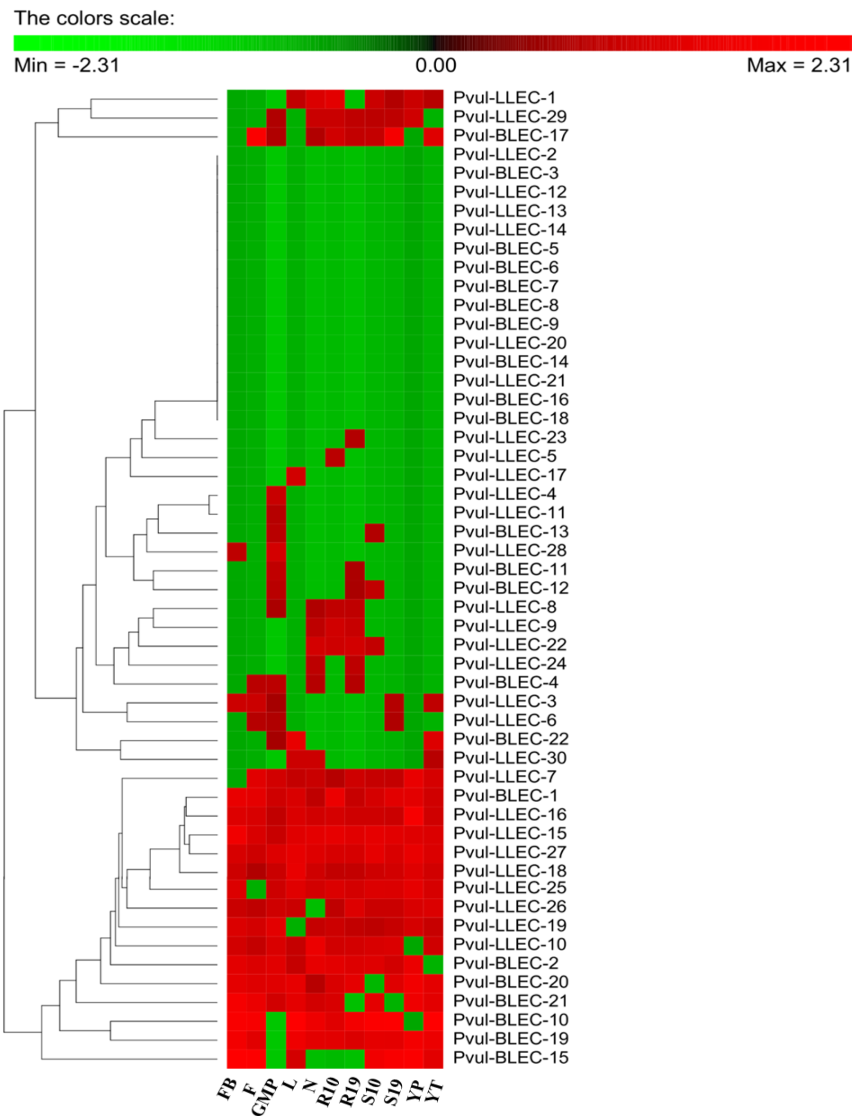


FIGURE 6. Heat map image for tissue specific mRNA levels of *Pvul-LEC* genes – from left to right; FB: Flower Buds, F: Flowers, GMP: Green Mature Pods, L: Leaves, N: Nodules, R10: Root 10, R19: Root 19, S10: Stem 10, S19: Stem 19, YP: Young Pods, YT: Young Trifoliates Tempo-spatial expression patterns of *Pvul-LEC* genes in different organs in *P. vulgaris*. The log10-transformation of the average of expression values were used to generate the heat map with PermutMatrix software. Red and green in the color scale indicate high and low transcript expression, respectively. Root 19: 19 days after planting; Root 10: 10 days after planting; Stem 19: 19 days after planting; Stem 10: 10 days after planting.

3.5. Responses of *Pvul-LEC* genes to salt stress and drought stress through RNAseq analysis

Plants are grown best in environments that are ideal for them. Exposure to an unexpected condition may result in diseases or physiological changes that affect the development and survival of the organism [69]. The leading environmental factors threatening agricultural production are salinity and drought [70]. In this regard, it is critical to understand plant defense mechanisms against these stress conditions. Accordingly, the heat map analysis was done in order to examine the responses of *Pvul-LEC* genes against salt and drought stress using RNAseq data obtained from GenBank (Figure 7). Upon salt stress treatment, the expression levels of *Pvul-BLEC-9, Pvul-LLEC-9, Pvul-LLEC-19, Pvul-LLEC-22,* and *Pvul-LLEC-29* were decreased, whereas *Pvul-BLEC-12, Pvul-BLEC-15*, and *Pvul-LLEC-23* expression levels were found to be increased in leaves. On the other hand, the expression levels of *Pvul-BLEC-1, Pvul-BLEC-3, Pvul-LLEC-9, Pvul-LLEC-17, Pvul-LLEC-18, Pvul-LLEC-19, Pvul-LLEC-27* and *Pvul-LLEC-28* genes were found to be increased whereas *Pvul-BLEC-11* expression level was decreased under drought stress conditions.

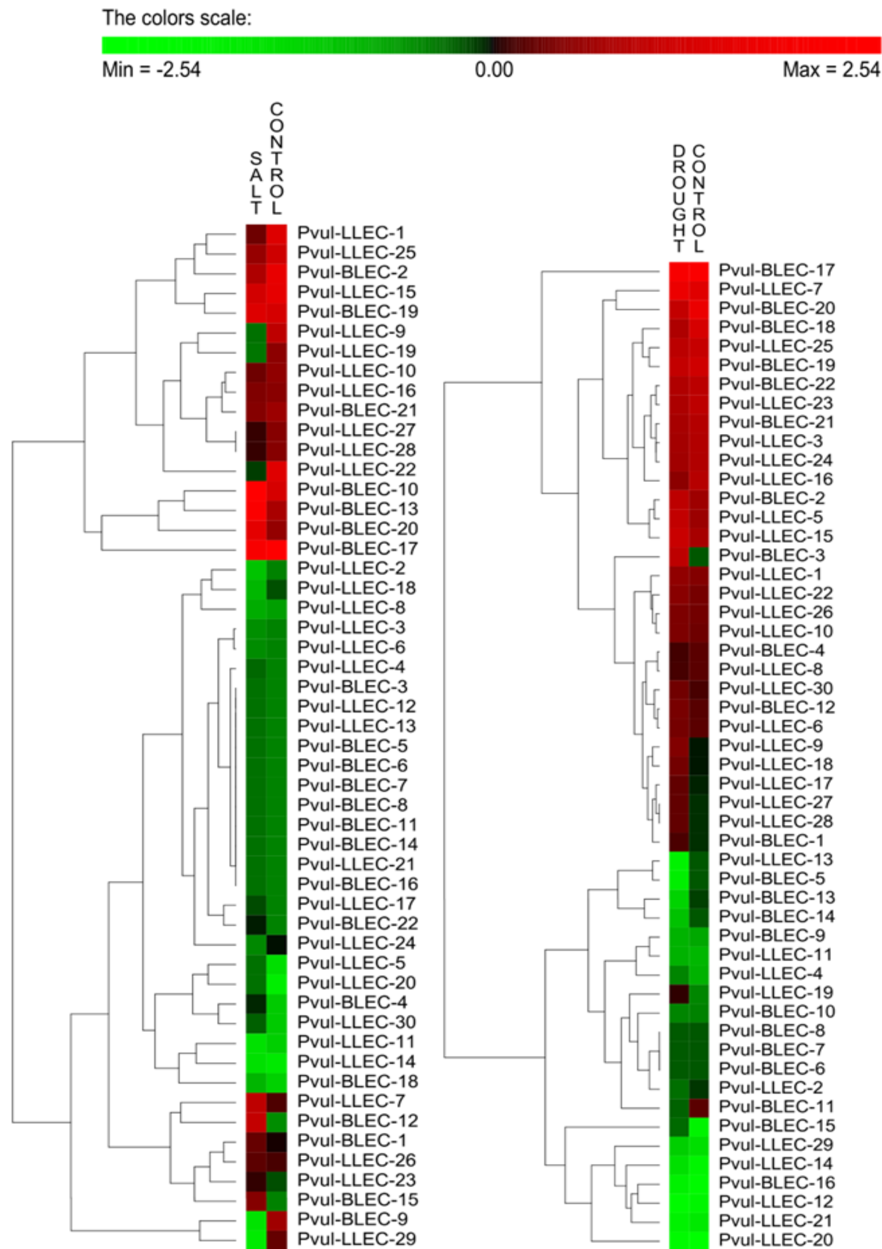


FIGURE 7. Heatmaps of *Pvul-LEC* genes differentially expressed under salt-control / drought- control stress conditions derived from RNAseq analysis. The log10-transformation of the average of expression values were used to generate the heat map with PermutMatrix software. Red and green in the color scale indicate high and low transcript expression, respectively.

4. CONCLUSIONS

The lectin gene family encoded in the *P. vulgaris* genome was identified and characterized for the first time in this study. Bioinformatics analyses were used to determine the properties of 52 newly identified *Pvul-LEC* genes. RNAseq data was also used to assess the expression levels of *Pvul-LEC* genes under salt and drought stress conditions. Since *LEC* genes were first discovered in *P. vulgaris*, we believe that our findings will open up new avenues for future research on *LEC* genes in other plant species.

Author Contribution Statements AO- data collection, management and manuscript writing. SA- project development, manuscript editing. IB- project development, data analysis, manuscript writing and manuscript editing. All authors have read and approved the manuscript.

Declaration of Competing Interests The authors declare no conflict of interest.

REFERENCES

[1] Sharon, N., Lis, H., History of lectins: from hemagglutinins to biological recognition molecules, *Glycobiology*, 14 (2004), 53R-62R. https://doi.org/10.1093/glycob/cwh122.

[2] Nareddy, P.K., Bobbili, K.B., Swamy, M.J., Purification, physico-chemical characterization and thermodynamics of chitooligosaccharide binding to cucumber (*Cucumis sativus*) phloem lectin, *International Journal of Biological Macromolecules*, 95 (2017), 910-919. https://doi.org/10.1016/j.ijbiomac.2016.10.078.

[3] Ogawa, H., Date, K., The "white kidney bean incident" in Japan, Lectins, Springer, (2014), 39-45.

[4] Vasconcelos, I.M., Oliveira, J.T.A., Antinutritional properties of plant lectins, *Toxicon*, 44 (2004), 385-403. https://doi.org/10.1016/j.toxicon.2004.05.005.

[5] Van Holle, S., Van Damme, E.J., Signaling through plant lectins: modulation of plant immunity and beyond, *Biochemical Society Transactions,* 46 (2018), 217-233. https://doi.org/10.1042/BST20170371.

[6] Brown, G.D., Willment, J.A., Whitehead, L., C-type lectins in immunity and homeostasis, *Nature Reviews Immunology*, 18 (2018), 374-389. https://doi.org/10.1038/s41577-018-0004-8

[7] Rüdiger, H., Gabius, H.-J., Plant lectins: occurrence, biochemistry, functions and applications, *Glycoconjugate journal*, 18 (2001), 589-613. https://doi.org/10.1023/A:1020687518999.

[8] Bittel, P., Robatzek, S., Microbe-associated molecular patterns (MAMPs) probe plant immunity, *Current Opinion in Plant Biology*, 10 (2007), 335-341. https://doi.org/10.1016/j.pbi.2007.04.021.

[9] Vlot, A.C., Dempsey, M. A., Klessig, D.F., Salicylic acid, a multifaceted hormone to combat disease, *Annual Review of Phytopathology*, 47 (2009), 177-206. https://doi.org/10.1146/annurev.phyto.050908.135202.

[10] Armijo, G., Salinas, P., Monteoliva, M.I., Seguel, A., García, C., Villarroel-Candia, E., Song W., van der Krol, A.R., Álvarez, M.E., Holuigue, L., A salicylic acid–induced lectin-like protein plays a positive role in the effector-triggered immunity response of *Arabidopsis thaliana* to *Pseudomonas syringae* Avr-Rpm1, Molecular *Plant-Microbe Interactions*, 26 (2013), 1395-1406. https://doi.org/10.1094/MPMI-02-13-0044-R.

[11] Luo, X., Xu, N., Huang, J., Gao, F., Zou,H., Boudsocq, M., Coaker, G., Liu, J., A lectin receptor-like kinase mediates pattern-triggered salicylic acid signaling, *Plant Physiology*, 174 (2017), 2501-2514. https://doi.org/10.1104/pp.17.00404.

[12] He, S., Shi, J., Walid, E. , Zhang, H., Ma, Y., Xue, S.J., Reverse micellar extraction of lectin from black turtle bean (*Phaseolus vulgaris*): Optimisation of extraction conditions by response surface methodology, *Food Chemistry*, 166 (2015), 93-100. https://doi.org/10.1016/j.foodchem.2014.05.156.

[13] Mazalovska, M., Kouokam, J.C., Lectins as promising therapeutics for the prevention and treatment of HIV and other potential coinfections, *Biomed Research International*, (2018). https://doi.org/10.1155/2018/3750646.

[14] Fitches, E., Woodhouse, S.D., Edwards, J.P., Gatehouse, J.A., In vitro and in vivo binding of snowdrop (*Galanthus nivalis* agglutinin; GNA) and jackbean (Canavalia ensiformis; Con A) lectins within tomato moth (Lacanobia oleracea) larvae; mechanisms of insecticidal action, *Journal of Insect Physiology*, 47 (2001), 777-787. https://doi.org/10.1016/S0022-1910(01)00068-3

[15] Miyake, K., Tanaka,T., McNeil, P.L., Lectin-based food poisoning: a new mechanism of protein toxicity, *PloS One*, 2 (2007), e687. https://doi.org/10.1371/journal.pone.0000687

[16] Menard, S., Cerf-Bensussan, N., Heyman, M., Multiple facets of intestinal permeability and epithelial handling of dietary antigens, *Mucosal Immunology*, 3 (2010), 247-259. https://doi.org/10.1038/mi.2010.5

[17] Lehr, C.-M., Bioadhesion technologies for the delivery of peptide and protein drugs to the gastrointestinal tract, *Critical Reviews In Therapeutic Drug Carrier Systems*, 11 (1994), 119-160.

[18] Zhang, J., Shi, J., Ilic, S., Xue, Jun., Kakuda, Y., Biological properties and characterization of lectin from red kidney bean (*Phaseolus vulgaris*), *Food Reviews International*, 25 (2008), 12-27. https://doi.org/10.1080/87559120802458115

[19] Gabius, H.-J., André, S., Jiménez-Barbero, J., Romero, A., Solís, D., From lectin structure to functional glycomics: principles of the sugar code, *Trends in biochemical sciences*, 36 (2011), 298-313. https://doi.org/10.1016/j.tibs.2011.01.005

[20] De Hoff, P.L., Brill, L.M., Hirsch, A.M., Plant lectins: the ties that bind in root symbiosis and plant defense, *Molecular Genetics and Genomics*, 282 (2009), 1-15. https://doi.org/10.1007/s00438-009-0460-8

[21] Van der Poel, A., Effect of processing on antinutritional factors and protein nutritional value of dry beans (*Phaseolus vulgaris* L.). A review, *Animal Feed Science and Technology*, 29 (1990), 179-208.

[22] Büyük, İ., Okay, A. Aksoy, T., Aras, S., The NIN-like Protein (NLP) Family in Common Bean: Genome-Wide Identificationn Evolution and Expression Analysis, *Communications Faculty of Sciences University of Ankara Series C Biology*, 30 (2021), 58-84. https://dergipark.org.tr/en/pub/communc/issue/58411/869501.

[23] Hirsch, M., Phillips, S., Solnik, C., Black, P., Schwartz, R., Carpenter, C., Activation of leukemia viruses by graft-versus-host and mixed lymphocyte reactions in vitro, *Proceedings of the National Academy of Sciences*, 69 (1972), 1069-1072.

[24] Koyanagi, Y., Miles, S., Mitsuyasu, R.T., Merrill, J.E., Vinters, H.V., Chen, I., Dual infection of the central nervous system by AIDS viruses with distinct cellular tropisms, *Science*, 236 (1987), 819-822.

[25] Lajolo, F.M., Genovese M.I., Nutritional significance of lectins and enzyme inhibitors from legumes, *Journal of Agricultural and Food Chemistry*, 50 (2002), 6592-6598. https://doi.org/10.1021/jf020191k.

[26] Woodley, J.F., Lectins for gastrointestinal targeting–15 years on, *Journal of Drug Targeting*, 7 (1999), 325-333. https://doi.org/10.3109/10611869909085515.

[27] Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes,R.D., Fazo, J., Mitros,T., Dirks, W., Hellsten,U., Putnam, N., Phytozome: a comparative platform for green plant genomics, *Nucleic Acids Research*, 40 (2012), D1178-D1186. https://doi.org/10.1093/nar/gkr944.

[28] Guo, A., Zhu, Q., Chen, X., Luo, J., GSDS: a gene structure display server, *Yi chuan= Hereditas*, 29 (2007), 1023-1026. https://doi.org/10.1360/yc-007-1023

[29] Voorrips, R., MapChart: software for the graphical presentation of linkage maps and QTLs, *Journal of Heredity*, 93 (2002), 77-78. https://doi.org/10.1093/jhered/93.1.77

[30] Bailey, T.L., Williams, N., Misleh. C., Li,W.W., MEME: discovering and analyzing DNA and protein sequence motifs, *Nucleic Acids Research*, 34 (2006), W369-W373. https://doi.org/10.1093/nar/gkl198

[31] Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, *Molecular Biology and Evolution*, 28 (2011), 2731-2739. https://doi.org/10.1093/molbev/msr121

[32] Letunic, I., Bork, P., Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy, *Nucleic Acids Research*, 39 (2011), W475-W478. https://doi.org/10.1093/nar/gkr201

[33] Zhang Y., miRU: an automated plant miRNA target prediction server, Nucleic acids research, 33 (2005), W701-W704. https://doi.org/10.1093/nar/gki383

[34] Suyama, M., Torrents, D., Bork, P., PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments, *Nucleic Acids Research*, 34 (2006), W609-W612. https://doi.org/10.1093/nar/gkl315

[35] Yang, Z., Nielsen, R., Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models, *Molecular Biology and Evolution*, 17 (2000), 32-43. https://doi.org/10.1093/oxfordjournals.molbev.a026236

[36] Hiz, M.C., Canher, B., Niron, H., Turet, M., Transcriptome analysis of salt tolerant common bean (*Phaseolus vulgaris* L.) under saline conditions, *PloS One*, 9 (2014), e92598. https://doi.org/10.1371/journal.pone.0092598

[37] Büyük, İ., Inal, B., Ilhan, E., Tanriseven, M., Aras, S., Erayman, M., Genome-wide identification of salinity responsive HSP70 s in common bean, *Molecular Biology Reports*, 43 (2016), 1251-1266. https://doi.org/10.1007/s11033-016-4057-0

[38] Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., The protein data bank, *Nucleic Acids Research*, 28 (2000), 235-242. https://doi.org/10.1093/nar/28.1.235

[39] Kelley, L.A., Sternberg, M.J., Protein structure prediction on the Web: a case study using the Phyre server, *Nature Protocols*, 4 (2009), 363. https://doi.org/10.1038/nprot.2009.2.
[40] Jiang, S.-Y., Ma, Z., Ramachandran, S., Evolutionary history and stress regulation of the lectin superfamily in higher plants, *BMC Evolutionary Biology*, 10 (2010), 1-24. https://doi.org/10.1186/1471-2148-10-79.
[41] Kouno, T., Watanabe, N., Sakai, N., Nakamura, T., Nabeshima, Y., Morita, M., Mizuguchi, M., Aizawa, T., Demura, M., Imanaka, T., The structure of *Physarum polycephalum* hemagglutinin I suggests a minimal carbohydrate recognition domain of legume lectin fold, *Journal of Molecular Biology*, 405 (2011), 560-569. https://doi.org/10.1016/j.jmb.2010.11.024
[42] Bouwmeester,K., Govers, F., *Arabidopsis* L-type lectin receptor kinases: phylogeny, classification, and expression profiles, *Journal of Experimental Botany*, 60 (2009), 4383-4396. https://doi.org/10.1093/jxb/erp277
[43] Lis, H., Sharon, N., Lectins: carbohydrate-specific proteins that mediate cellular recognition, *Chemical Reviews*, 98 (1998), 637-674. https://doi.org/10.1021/cr940413g .
[44] Gegg, C.V., Roberts, D.D., Segel,I.H., Etzler, M.E., Characterization of the adenine binding sites of two Dolichos biflorus lectins, *Biochemistry*, 31 (1992), 6938-6942. https://doi.org/10.1021/bi00145a011
[45] Lv, D., Wang, G., Xiong, L.-R., Sun, J.-X., Chen, Y., Guo, C.-L., Yu, Y., He, H.-L., Cai, R., Pan, J.-S., Genome-wide identification and characterization of lectin receptor-like kinase gene family in cucumber and expression profiling analysis under different treatments, *Genes*, 11 (2020), 1032. https://doi.org/10.3390/genes11091032
[46] Vaid, N., Pandey, P.K., Tuteja, N., Genome-wide analysis of lectin receptor-like kinase family from Arabidopsis and rice, *Plant Molecular Biology*, 80 (2012), 365-388. https://doi.org/10.1007/s11103-012-9952-8
[47] Wang, Z., Ren, H., Xu, F., Lu, G., Cheng, W., Que, Y., Xu, L., Genome-Wide Characterization of Lectin Receptor Kinases in *Saccharum spontaneum* L. and Their Responses to Stagonospora tainanensis Infection, *Plants*, 10 (2021), 322.
[48] Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M.R., Appel R.D., Bairoch, A., Protein identification and analysis tools on the ExPASy server. *The proteomics protocols handbook*, (2005), 571-607. https://doi.org/10.1385/1-59259-890-0:571
[49] Dang, L., Van Damme, E.J., Genome-wide identification and domain organization of lectin domains in cucumber, *Plant Physiology and Biochemistry*, 108 (2016), 165-176. https://doi.org/10.1016/j.plaphy.2016.07.009
[50] Singh, P., Mishra A.K., Singh C.M., Genome-wide identification and characterization of Lectin receptor-like kinase (LecRLK) genes in mungbean (*Vigna radiata* L. Wilczek), Journal of Applied Genetics, (2021), 1-12. https://doi.org/10.1007/s13353-021-00613-8
[51] Wang, Y., Wang, X., Paterson, A.H., Genome and gene duplications and gene expression divergence: a view from plants, *Annals of the New York Academy of Sciences*, 1256 (2012), 1-14. https://doi.org/10.1111/j.1749-6632.2011.06384.x
[52] Le Hir, H., Nott, A., Moore, M.J., How introns influence and enhance eukaryotic gene expression, *Trends in Biochemical Sciences*, 28 (2003), 215-220. https://doi.org/10.1016/S0968-0004(03)00052-5

[53] Rose, A., Intron-mediated regulation of gene expression, *Nuclear pre-mRNA Processing in Plants*, (2008), 277-290. https://doi.org/10.1007/978-3-540-76776-3_15

[54] Liu, P.-L., Huang, Y., Shi, P.-H.,Yu,M., Xie, J.-B., Xie, L., Duplication and diversification of lectin receptor-like kinases (LecRLK) genes in soybean, Scientific Reports, 8 (2018), 1-14. https://doi.org/10.1038/s41598-018-24266-6

[55] Carmel, L., Wolf, Y.I., Rogozin, I.B., Koonin, E.V., Three distinct modes of intron dynamics in the evolution of eukaryotes, *Genome Research*, 17 (2007), 1034-1044. http://www.genome.org/cgi/doi/10.1101/gr.6438607

[56] Fierro-Monti, I., Mathews, M.B., Proteins binding to duplexed RNA: one motif, multiple functions, *Trends in Biochemical Sciences*, 25 (2000), 241-246. https://doi.org/10.1016/S0968-0004(00)01580-2

[57] Richardson, J.S., Protein anatomy, Adv. *Protein Chem*, 34 (1981), 339. https://doi.org/10.1016/S0065-3233(08)60520-3

[58] Orengo, C.A., Jones, D.T., Thornton, J.M., Protein superfamilles and domain superfolds, *Nature*, 372 (1994), 631-634. https://doi.org/10.1038/372631a0

[59] Williams, A., Westhead, D., Sequence relationships in the legume lectin fold and other jelly rolls, *Protein Engineering*, 15 (2002), 771-774. https://doi.org/10.1093/protein/15.10.771

[60] Yamaguchi-Shinozaki, K., Shinozaki, K., Organization of cis-acting regulatory elements in osmotic-and cold-stress-responsive promoters, *Trends in Plant Science*, 10 (2005), 88-94. https://doi.org/10.1016/j.tplants.2004.12.012

[61] Lu, X.-Y., Huang, X.-L., Plant miRNAs and abiotic stress responses, *Biochemical and Biophysical Research Communications*, 368 (2008), 458-462. https://doi.org/10.1016/j.bbrc.2008.02.007

[62] Khraiwesh, B., Zhu, J.-K., Zhu, J., Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants, *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, 1819 (2012), 137-148. https://doi.org/10.1016/j.bbagrm.2011.05.001

[63] Tang, F., Chu, L., Shu, W., He, X., Wang, L., Lu, M., Selection and validation of reference genes for quantitative expression analysis of miRNAs and mRNAs in Poplar, *Plant Methods,* 41 (2019), 28-37. https://doi.org/10.1186/s13007-019-0420-1

[64] Matthewman, C.A., Kawashima, C.G., Húska, D., Csorba, T., Dalmay, T., Kopriva, S., miR395 is a general component of the sulfate assimilation regulatory network in *Arabidopsis*, *FEBS Letters*, 586 (2012), 3242-3248. https://doi.org/10.1016/j.febslet.2012.06.044

[65] Jones-Rhoades, M.W., Bartel, D.P., Computational identification of plant microRNAs and their targets, including a stress-induced miRNA, *Molecular Cell*, 14 (2004), 787-799. https://doi.org/10.1016/j.molcel.2004.05.027

[66] Cao, C., Long, R., Zhang, T., Kang, J., Wang, Z., Wang, P., Sun, H., Yu, J., Yang, Q., Genome-wide identification of microRNAs in response to salt/alkali stress in *Medicago truncatula* through high-throughput sequencing, *International Journal of Molecular Sciences*, 19 (2018), 4076. https://doi.org/10.3390/ijms19124076

[67] Natarajan, B., Kalsi, H.S., Godbole, P., Malankar, N., Thiagarayaselvam,A., Siddappa, S., Thulasiram, H.V., Chakrabarti, S.K., Banerjee, A.K., MiRNA160 is associated with local defense and systemic acquired resistance against *Phytophthora infestans* infection in potato, *Journal of Experimental Botany*, 69 (2018), 2023-2036. https://doi.org/10.1093/jxb/ery025

[68] Li, W., Wang, T., Zhang, Y., Li, Y., Overexpression of soybean miR172c confers tolerance to water deficit and salt stress, but increases ABA sensitivity in transgenic *Arabidopsis thaliana*, *Journal of Experimental Botany*, 67 (2016), 175-194. https://doi.org/10.1093/jxb/erv450

[69] Shao, H.-B., Chu, L.-Y., Jaleel, C.A., Zhao, C.-X., Water-deficit stress-induced anatomical changes in higher plants, *Comptes Rendus Biologies*, 331 (2008), 215-225. https://doi.org/10.1016/j.crvi.2008.01.002

[70] Zapata, P.J., Serrano, M., Pretel, M.T., Botella, M.A., Changes in free polyamine concentration induced by salt stress in seedlings of different species, *Plant Growth Regulation*, 56 (2008), 167-177. https://doi.org/10.1007/s10725-008-9298-z