Legume Research, 42(4) 2019: 461-466
Print ISSN:0250-5371 / Online ISSN:0976-0571

AGRICULTURAL RESEARCH COMMUNICATION CENTRE
www.arccjournals.com

Genome-wide identification and characterization of the *Phenylalanine Ammonia-lyase* (*PAL*) gene family in *Medicago truncatula*

Wei Ren[1,2,3], Yingzhe Wang[2], Ankai Xu[2] and Yunge Zhao*[1]

Institute of Soil and Water Conservation, Chinese Academy of Sciences and Ministry of Water Resources, 26 Xinong Road, Yangling-712 100, Shaanxi, China.
Received: 18-05-2018 Accepted: 26-03-2019 DOI: 10.18805/LR-431

ABSTRACT

Phenylalanine ammonia-lyase (PAL) catalyzes the rate-limiting step of phenylpropanoid biosynthesis in plants and supplies precursors for a variety of secondary metabolites, such as flavonoids, lignins and stilbenes. The first draft of the full *Medicago truncatula* genome assembly has been released. However it is observed that, the *PAL* gene family from *Medicago truncatula* (*MtPAL* genes) has not been characterized in detail. In this study, a comprehensive analysis of the *Medicago truncatula PAL* gene family is presented, including chromosomal locations, phylogenetic analyses, gene structures, three-dimensional (3D) structures and expression patterns. Six *Medicago truncatula PAL* genes that encode PAL proteins were identified in the *Medicago truncatula* genome. It was shown that *MtPAL* genes are distributed on four chromosomes. Dynamic expression patterns of *MtPAL* genes were observed in different tissues and abiotic stresses, suggesting that *MtPAL* genes may play important roles in the regulation of development and stress responses in *Medicago truncatula*.

Key words: Gene family, *Medicago truncatula*, PAL, Stress, Tissue specific.

INTRODUCTION

Phenylalanine ammonia-lyase (PAL; EC 4.3.1.5) is an important enzyme that catalyzes the key step in the phenylpropanoid pathway in plants by the nonoxidative deamination of L-phenylalanine to cinnamic acid, which then produces many secondary metabolic products, including lignins, flavonoids, and many other important phenolics in plants, especially stilbenes in grape (Vogt, 2010). Due to the irreplaceable role of PAL in phenylpropanoid biosynthesis, *PAL* genes have been extensively studied. There are typically between three and seven *PAL* genes identified in many plant genomes that form a small gene family. In detail, there are four *PAL* genes in *Arabidopsis thaliana*, three *PAL* genes in *Coffea canephora* (Lepelley *et al.*, 2012), seven *PAL* genes in *Cucumis sativus* (Shang *et al.*, 2012), five *PAL* genes in *Populus trichocarpa*, three *PAL* genes in *Salvia miltiorrhiza* (Hou *et al.*, 2013), and three *PAL* genes in *Scutellaria baicalensis*. This has been summarized as described previously (Hou *et al.*, 2013). The high copy number of the PAL gene family was also discovered in tomato. While the general function of PAL is fully clear in many plants, the functional differentiation of the individual *PAL* genes is not well understood.

PAL has been reported widely in biology, including plants, fungi, viruses and algae, and PAL activity has been detected in many plants. To date, there have been no reports for animal PALs (Hyun *et al.*, 2011). PAL was first isolated from *Hordeum vulgare* and could convert L-phenylalanine to cinnamic acid and ammonia. Soon, PAL was shown to play important roles in animals and markedly inhibit the cell division of human leukemic and murine L5178Y lymphoblasts *in vitro* by depriving these cells of L-phenylalanine. Additionally, PAL benefits phenylketonuric (PKU) patients. Considering the substantial commercial and medical potential of PAL, many studies have been conducted to resolve problems associated with the enzyme activity and stability of PALs in yeast. Continuous research has been carried out on *PAL* genes involved in light response, tissue-specific expression, and stress response among many plant species (Lin *et al.*, 2018). However, an overall analysis for the *PAL* gene family has not been available for *Medicago truncatula* (Karayilanli and Ayhan, 2016; Yin *et al.*, 2018). Recently, a working draft of the *M. truncatula* genome has been published (Young *et al.*, 2011). To generate a comprehensive understanding of the *M. truncatula PAL* gene family, we searched the assembly of the *M. truncatula* genome for *PAL* sequences. In this study, we identified six *PAL* genes by a genome-wide search of the *M. truncatula* genome. The intron/exon structures of *PAL* genes were constructed, and the molecular weights (Mw) and isoelectric

*Corresponding author's e-mail: 454898044@qq.com
[1]Institute of Soil and Water Conservation,Chinese Academy of Sciences and Ministry of Water Resources, 26 Xinong Road, Yangling-712 100, Shaanxi, China.
[2]Jilin Academy of Agricultural Sciences, 1363 Shengtai Street, Changchun-130 124, Jilin, China.
[3]University of Chinese Academy of Sciences, 19 Yuquan Road, Beijing-100 049, China.

MATERIALS AND METHODS

Plant materials: *M. truncatula* (Jemalong) A17 was grown in an incubator with the following conditions: 24°C, 16 hours, light/ 18°C, 8 hours, night. Different tissues were sampled at the flower stage (e.g., root, stem, leaf, and flower). Samples were quickly frozen in liquid nitrogen and stored at -80°C for RNA isolation.

Putative grape *PAL* gene identification and characterization: The hidden Markov model (HMM) domain Lyase_aromatic (PF00221) in the Pfam database was used to search against the Ref Seq database to find candidate PAL sequences on the HMMER website. All the candidate PAL sequences of *M. truncatula* were submitted to the NCBI Conserved Domains Search website tool to confirm whether they contained the phe_am_lyase domain. *M. truncatula* ESTs were searched on the NCBI database. The MW and pI of *M. truncatula* PALs were computed using the online tool Compute pI/Mw (http://web.expasy.org/compute_pi/). The exon/intron structures of the *MtPAL genes* were drawn using FancyGene.

Phylogenetic analysis: Sequence alignments of PAL proteins from *Arabidopsis*, *Oryza sativa* and *M. truncatula* were performed using ClustalX2 conducted on MEGA 5 software. Then, the sequence alignments were used to construct the phylogenetic tree. Specifically, the maximum likelihood method was used to generate the phylogenetic tree, and a bootstrap analysis of 1,000 replicates was used to evaluate the conûdence level of monophyletic groups. Lasergene 7.0 (DNAstar) was used to display the sequence distances of the PAL proteins. PAL protein sequences were aligned by the ClustalW method and then display the sequence distances in MegAlign, which is a composition of Lasergene 7.0.

Chromosomal localization and subcellular localization prediction: The chromosomal distribution of putative *MtPAL genes* was analyzed in the *M. truncatula* Genome Database (http://www.medicagogenome.org/). The subcellular locations of *MtPAL genes* were predicted with WoLF PSORT (http://www.genscript.com/psort/wolf_psort.html) and further confirmed by the TargetP 1.1 online server (http://www.cbs.dtu.dk/services/TargetP). All the parameters were used by default.

Conserved motif analysis and three-dimensional (3D) structure prediction: The conserved protein motifs were investigated with the multiple expectation maximization for the motif elicitation (MEME) version 4.9.1 online program. The analytical parameters were set as described before. Three-dimensional (3D) structure prediction of the *Mt*PALs was carried out on the Phyre web server before visualization in POLYVIEW-3D.

Microarray data analysis: The corresponding DNA sequences of putative *MtPAL* genes were retrieved based on their chromosomal locations. Then, the DNA sequences were used to find the corresponding probe set IDs of putative *MtPAL* genes in PLEXdb. Expression data were extracted based on the IDs of the probe set. Raw microarray datasets were normalized by using the Z-score method.

qRT-PCR and statistical analysis: Total RNA was extracted using the method described previously. Then, the RT-PCR procedure was the same as the previously described method. Gene-specific primers were designed for *MtPAL* genes (Additional file 1).

RESULTS AND DISCUSSION

Identification and chromosomal localization of *MtPAL* genes: To identify members of the PAL gene family in *M. truncatula*, phe_am_lyase domain file (Lyase_aromatic; pfam00221) was used as a query for multiple searches against the Ref Seq database using web-based HMMER with the cut-offs significance E-values being e^{-5}. After the HMMER search, six protein sequences were identified and described in *M. truncatula*. Table 1 shows the six genes with their

Table 1: *PAL* gene family in *M. truncatula.*

Gene Name[a]	Accession	ORF length(bp)[b]	Deduced polypeptide[c]			Chromosome[d]	Sublocation[e]
			Length (AA)	MW(KDa)	PI		(WoLF/TargetP)
MtPAL1	Medtr1g064090.1	2139	712	77.37	5.93	Chr1	Chloroplast/-
MtPAL2	Medtr1g094780.1	2169	722	78.55	5.80	Chr1	Endoplasmic reticulum /-
MtPAL3	Medtr2g094960.1	2103	700	77.84	6.27	Chr2	Cytoplasmic/-
MtPAL4	Medtr5g098720.1	2127	708	77.88	6.23	Chr5	Chloroplast/-
MtPAL5	Medtr7g101395.1	2175	724	78.86	6.00	Chr7	Chloroplast/-
MtPAL6	Medtr7g101425.1	2175	724	78.83	6.03	Chr7	Chloroplast/-

[a]Systematic designation given to *M.truncatula PALs* in this work.
[b]Length of open reading frame in base pairs.
[c]The number of amino acids, molecular weight (kilo Daltons) and isoelectric point (pI) of the deduced polypeptides.
[d]Location of *Medicago truncatula* PALs on the chromosome.
[e]The subcellular location result of *M.truncatula* PALs by WoLF and TargetP.

names, accession numbers, putative subcellular localization, amino acid numbers, Mw, pI and chromosomal location. For the nomenclature of *MtPALs*, we followed the order of the *M. truncatula* PAL sequences on the chromosome. These six *MtPALs* were mapped on 4 out of the 8 *M. truncatula* chromosomes. *MtPAL1* and *MtPAL2* were present on chromosome 1; *MtPAL3* on chromosome 2; *MtPAL4* on chromosome 5 and *MtPAL5* and *MtPAL6* on chromosome 7 (Fig 1). The full-length coding sequences of the *MtPAL* genes ranged from 2103 bp (*MtPAL3*) to 2175 bp (*MtPAL5/6*) with deduced proteins of 700 to 724 amino acids. The isoelectric point of *MtPALs* changed very little from 5.80 to 6.27.

Structural and phylogenetic analysis of *Mt*PALs: The intron/exon structures of *MtPAL* genes were determined by matching cDNA sequences with genomic sequences. This sequence analysis revealed that only one intron was found in all the *MtPAL* genes (Fig 2A).

Phylogenetic analysis and sequence distances were performed and the results showed that most of the *Mt*PAL protein sequences are highly conserved. The most divergence pair was *Mt*PAL3 and *Mt*PAL4, with an identity of 67.1, while the highest identity pair was *Mt*PAL5 and *Mt*PAL6, with an identity of 98.8 (Fig 2B).

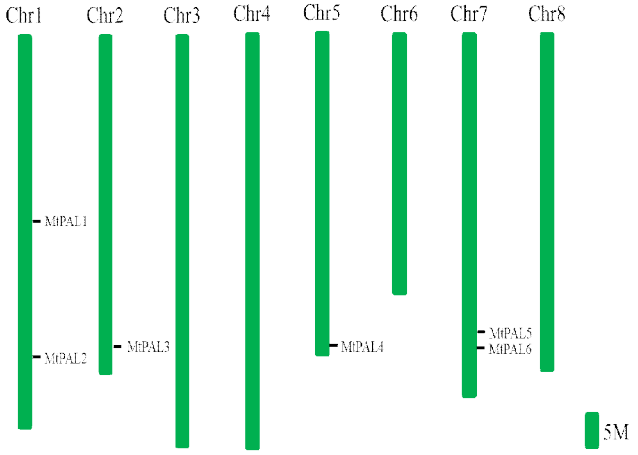


Fig 1: Chromosomal localization of *MtPAL* genes on the *Medicago truncatula* genome. Each black bold line represents one chromosome, and they were drawn in proportion. Bar = 5 Mb.

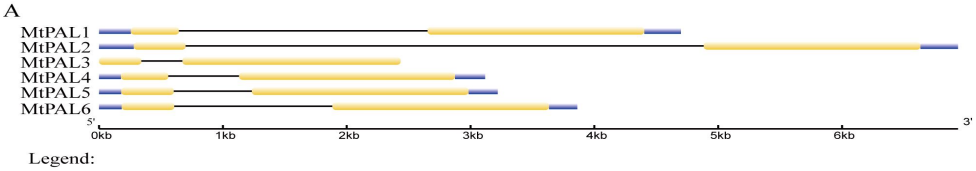
The annotated PALs in *Arabidopsis thaliana* (*At*) and *Oryza sativa* (*Os*) enabled us to determine the phylogenetic relationship of PAL proteins (Additional file 2.). A phylogenetic tree constructed using the protein sequences of 6 *Mt*PALs, 9 *Os*PALs and 4 *At*PALs depicted that all 19 PAL proteins were divided into two clades and *PAL* proteins of *Arabidopsis* and *Medicago* fell into one clade (Fig 3). Analysis of the predicted 3D structure showed that *Mt*PAL proteins had very similar 3D structures and a representative catalytic site was clearly found (Fig 4).

Analysis of the conserved PAL motifs: The conserved motif structure of *Mt*PAL proteins was analyzed by submitting protein sequences to the MEME software. Based on the results of the MEME analysis, we found that all the *Mt*PAL proteins possessed the same conserved motif and were plotted with a colored box in order (Fig 5).

The identified genes were confirmed to be PAL by high sequence similarity between *Mt*PALs and well-characterized PALs from *Arabidopsis thaliana* (Raes *et al.*, 2003). Phylogenetic analysis of *Mt*PALs, *Os*PALs and *At*PALs showed that PALs clustered into two distinct clades, and PALs from *Arabidopsis* and *Medicago* can be divided into one clade. The cluster results showed that the PAL and phenylpropanoid biosynthesis pathways could exist before monocot and dicot differentiation. Gene duplication was supposed to be a mechanism for the adaptation of organisms to the environment. The highly conserved *Mt*PAL protein sequences indicated functional redundancy of the gene family members to adapt to different stresses.

Analysis of the expression patterns of *MtPAL* genes using Affymetrix arrays: In our study, publicly available grape microarray datasets were utilized to study the responses of

A



B

MtPAL1	MtPAL2	MtPAL3	MtPAL4	MtPAL5	MtPAL6	
***	78.5	74.8	72.7	89.2	89	MtPAL1
	***	72.4	68.6	77.6	77.1	MtPAL2
		***	67.1	74.7	74.7	MtPAL3
			***	74	73.4	MtPAL4
				***	98.8	MtPAL5
					***	MtPAL6

Fig 2: Exon-intron structures and sequence divergence of *MtPAL* genes.
A: Coding exons, represented by colored boxes, were drawn to scale. Dashed lines connecting two exons represent introns; B: Divergence and identification between every two protein sequences are present on the bottom left and top right, respectively. The numbers in the boxes are percentages.

MtPAL genes to abiotic and biotic stress conditions and tissue-specific expression patterns. The *MtPAL3* probe was not available in the microarray datasets and the microarray expression analysis for *MtPAL3* was not included in our study. Tissue expression profile analysis indicated that *Mt*PAL1/4/5/6 had the highest expression value in seed that had germinated in 36 days, while *Mt*PAL2 was a root-specific expression gene (Fig 6A). The dynamic change of *MtPAL* genes under drought stress showed that they were induced at short time points (e.g., 2 d and 3 d) while exhibiting low levels of expression at long time points (e.g., 7 d and 10 d). Interestingly, they resumed high expression levels after re-watering (Fig 6B). *MtPAL* genes could be induced by salinity stress, and they were more sensitive to hydroponic stress (Fig 6C).

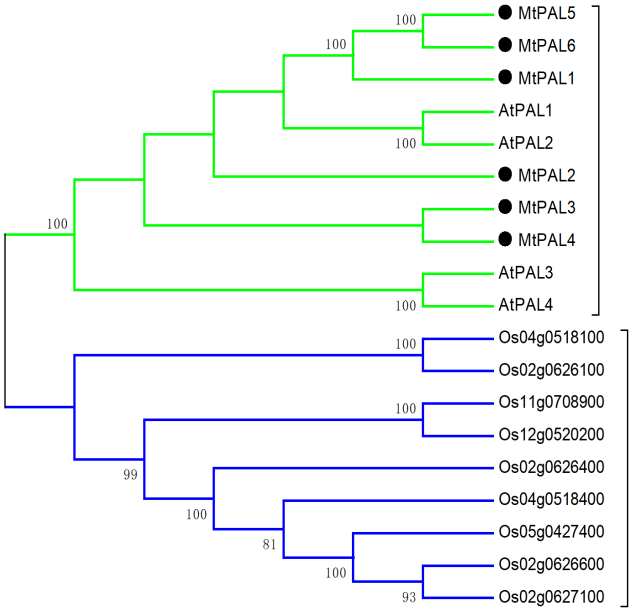


Fig 3: Phylogenetic analysis annotated PALs in *Arabidopsis*, *Oryza sativa* and *Medicago truncatula*. Numbers above or below branches of the tree indicate bootstrap values. Protein sequences are available in Additional file 2.

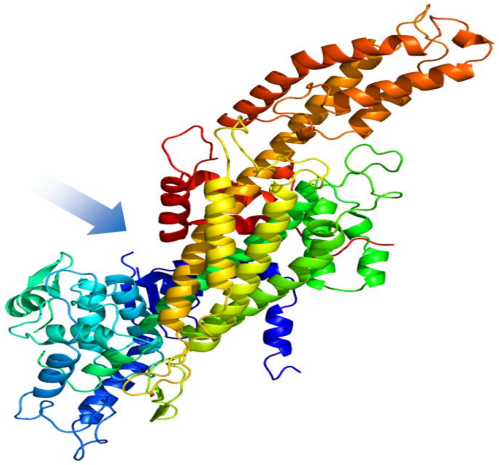


Fig 4: 3D structures of MtPALs. The arrows and loop indicate different structures compared with other PALs.

Expression patterns of *MtPAL* genes in different tissues: To investigate tissue-specific *MtPAL* genes, samples of roots, stems, leaves and flowers were collected for qRT-PCR analysis. *MtPAL1* was highly expressed in the stem and expressed at low levels in the root. However, *MtPAL2* was significantly expressed in the roots and stems and *MtPAL3* showed flower-specific expression. *MtPAL4* was expressed at very low levels in all the tissues. *MtPAL5* and *MtPAL6* had the same expression trends with high levels of expression in the stem and leaf and low levels in the root and flower (Fig 7).

It has been shown that the *PAL* gene family plays an important role in plant growth, development and stress response in *Arabidopsis* (Huang *et al.*, 2010). As the *PAL* gene provides linkages between primary and secondary metabolism, lignin, flavonoids and phenolic compounds were determined by the enzyme activity of PAL (Xu *et al.*, 2008). The different expression profiles of *MtPAL* genes in different tissues indicated functional differentiation of the gene family. The regulatory region of the *MtPAL* genes could differ in cis-acting elements. We analyzed the tissue expression patterns of *MtPAL* genes (Fig 6A). *MtPAL2* was found to be root-specific. In addition, we found that *MtPAL2* was down regulated as the seed geminated, while the other *PAL* genes

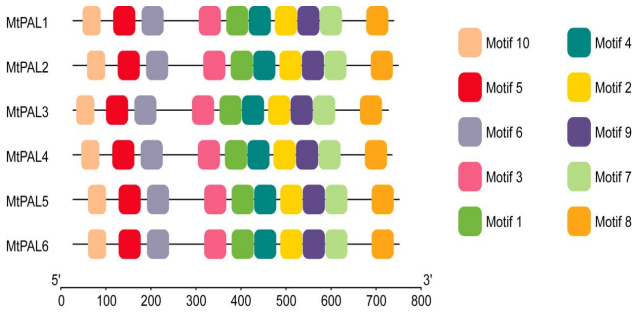


Fig 5: MEME software characterized 10 conserved motifs of MtPALs. Different motifs numbered 1 to 10 are shown with different color boxes.

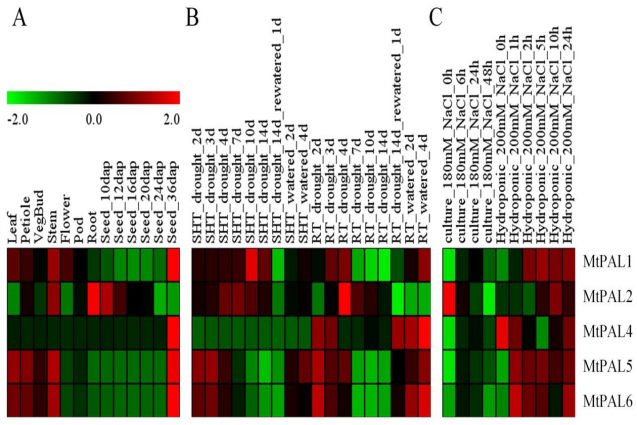


Fig 6: Hierarchical clustering of *MtPAL* genes. Details of the experimental conditions are listed in Additional file 3. The changes in *MtPAL* gene expression levels are shown in color as the upper scale.

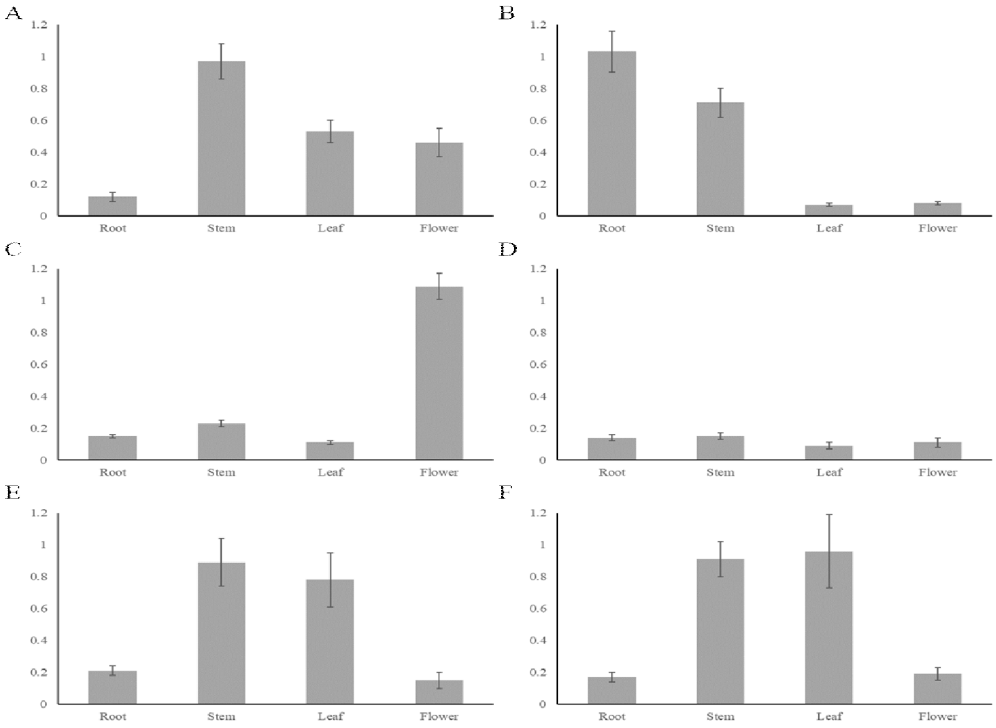


Fig 7: Expression patterns of *MtPAL* genes in different tissues of *Medicago truncatula*. (A) *MtPAL1*, (B) *MtPAL2*, (C) *MtPAL3*, (D) *MtPAL4*, (E) *MtPAL5*, (F) *MtPAL6*. The experimental data were generated from three biological repeats.

were up regulated. This phenomenon suggests that *MtPAL* genes play different roles in *Medicago* seed germination (Moawed, 2016; Kumar *et al.*, 2017). The qRT-PCR results for *MtPAL* genes showed similar expression patterns as microarray data analysis indicated that microarray data are still useful for gene quantity.

The SlPAL5 gene from tomato was proved to respond to NaCl and mannitol stress and other abiotic stresses (Guo and Wang, 2009). *PAL* was also involved in heat stress in banana fruit (Chen *et al.*, 2008). In the present study, we found that *MtPAL* genes were up regulated or down regulated by NaCl and drought treatment. The expression of *MtPAL* genes in the response to drought treatment rose first and then fell, which was the same as that of *SlPAL5*. The results suggested that different plants are subject to the same model of regulation. However, the mechanism is still unknown. In addition to abiotic stresses, *PAL* plays an important role in biotic stresses. *PAL* gene knockdown lines of *Brachypodium* enhanced its susceptibility to fungal pathogen (Cass *et al.*, 2015). Moreover, the *PAL* gene is hypothesized to be SA-dependent, as SA is an important pathogen-related signal (Kim and Hwang, 2014).

CONCLUSION

PAL is the key enzyme for secondary metabolites. In the present study, we systematically performed a genome-wide search of grape *PAL* genes by BLAST analysis of the assembly genome of *M. truncatula*. A total of six *MtPAL* genes were identified. The conserved motifs and features of *MtPAL* genes were analyzed. As an important gene family for plant secondary metabolites, it is valuable to understand the molecular and biological functions of PAL. Our findings provide a comprehensive understanding of *Medicago PAL* genes. Our results suggest that the *MtPAL* genes are involved in abiotic stress responses and play different roles in different developmental stages. Thus, further analysis of the roles of *MtPAL* genes could improve stress tolerance.

ACKNOWLEDGEMENT

This work was supported by earmarked funds for the Modern Agro-industry Technology Research System (Grant No. CARS-35).

REFERENCES

Cass, C.L., Peraldi, A., Dowd, P.F., Mottiar, Y., Santoro, N., Karlen, S.D., Bukhman, Y.V., *et al*. (2015). Effects of Phenylalanine Ammonia Lyase (PAL) knockdown on cell wall composition, biomass digestibility and biotic and abiotic stress responses in Brachypodium. *Journal of Experimental Botany*, **66**: 4317-4335.

Chen, J., He, L., Jiang, Y., Wang, Y., Joyce, D.C., Ji, Z. and Lu, W. (2008). Role of phenylalanine ammonia-lyase in heat pretreatment-induced chilling tolerance in banana fruit. *Physiologia Plantarum*, **132**:318-328.

Guo, J., and Wang, M.H. (2009). Characterization of the phenylalanine ammonia-lyase gene (SlPAL5) from tomato (*Solanum lycopersicum* L.). *Molecular Biology Reports*, **36**:1579-1585.

Hou, X., Shao, F., Ma, Y. and Lu, S. (2013). The phenylalanine ammonia-lyase gene family in Salvia miltiorrhiza: genome-wide characterization, molecular cloning and expression analysis. *Molecular Biology Reports*, **40**:4301-4310.

Huang, J., Gu, M., Lai, Z., Fan, B., Shi, K., Zhou, Y.H., Yu, J.Q., Chen, Z. (2010). Functional analysis of the arabidopsis PAL gene family in plant growth, development, and response to environmental stress. *Plant Physiology*, **153**:1526-1538.

Hyun, M.W., Yun, Y.H., Kim, J.Y., Kim, S.H. (2011). Fungal and plant phenylalanine ammonia-lyase. *Mycobiology*, **39**:257-265.

Karayilanli, E. and Ayhan, V. (2016). Investigation of feed value of alfalfa (*Medicago sativa* L.) harvested at different maturity stages. *Legume Research*, **39:**237-247.

Kim, D.S. and Hwang, B.K. (2014). An important role of the pepper phenylalanine ammonia-lyase gene (PAL1) in salicylic acid-dependent signalling of the defence response to microbial pathogens. *Journal of Experimental Botany*, **65**:2295-2306.

Kumar, P.R., Channakeshava, B.C., Siddaraju, R. (2017). Effect dates of sowing and cutting time on seed yield and quality of alfalfa (medicago sativa l.) cv. rl-88. *Indian Journal of Agricultural Research*, **51**:493-497.

Lepelley, M., Mahesh, V., Mccarthy, J., Rigoreau, M., Crouzillat, D., Chabrillange, N., Kochko, A., Campa, C. (2012). Characterization, high-resolution mapping and differential expression of three homologous PAL genes in Coffea canephora Pierre (Rubiaceae). *Planta*, **236**:313-326.

Lin, W., Liu, A., Weng, C., Li, H., Sun, S., Song, A., Zhu, H. (2018).Cloning and characterization of a novel phenylalanine ammonia-lyase gene from inonotus baumii. *Enzyme and Microbial Technology*, **112**:1-7.

Moawed, M.M.(2016). Evaluation of morphological and anatomical characters for discrimination and verification of some medicago sativa (l.) cultivars. *Indian Journal of Agricultural Research*, **50**:183-192.

Raes, J., Rohde, A., Christensen, J.H., Vande P.Y., Boerjan, W.(2003). Genome-Wide characterization of the lignification toolbox in Arabidopsis. *Plant Physiology*, **133**:1051-1071.

Shang, Q.M., Li, L., Dong, C.J. (2012). Multiple tandem duplication of the phenylalanine ammonia-lyase genes in *Cucumis sativus* L. *Planta*, **236**:1093-1105.

Vogt, T. (2010). Phenylpropanoid Biosynthesis. *Molecular Plant*, **3**:2-20.

Xu, F., Cai, R., Cheng, S., Du, H., Wang, Y. (2008). Molecular cloning, characterization and expression of phenylalanine ammonia-lyase gene from Ginkgo biloba. *African journal of Biotechnology*, **7**:721-729.

Yin, S., Wang, Y., Nan, Z. (2018). Genetic diversity studies of alfalfa germplasm (Medicago sativa L. subsp. sativa) of United States origin using microsatellite analysis. *Legume Research*, **41**:202-207.

Young, N.D., Oldroyd, G.E., Geurts, R., Cannon, S.B., Udvardi, M.K., Benedito, V.A., *et al.* (2011). The Medicago genome provides insight into the evolution of rhizobial symbioses. *Nature*, **480**:520-524.