Indian Journal of Biotechnology
Vol 16, July 2017, pp 456-464

Comparative analyses of pathogenesis-related protein-10 (PR10) in plants

Ibrahim Ilker Ozyigit[1], Recep Vatansever[1], Ertugrul Filiz[2] *

[1]Marmara University, Faculty of Science and Arts, Department of Biology, 34722, Goztepe, Istanbul, Turkey, [2]Duzce University, Department of Crop and Animal Production, Cilimli Vocational School, 81750, Cilimli, Duzce, Turkey

Received 3 March 2015; revised 1 September 2015; accepted 8 September 2015

In the present study, we have comparatively analyzed *PR10* genes and proteins from 28 plant species in order to understand the relationship (conservation or divergence) between different PR10s in various plant species. In analyzed species, PR10 proteins were found to be small (157-166 aa long and 14.3-18.2 kDa weight) and acidic (4.69-6.17) in nature. Besides, PR10 sequences had highly conserved GxGGxG motif (P-loop motif) structure at various positions. These positional variations in glycine (Gly) residues may become the result of substitution, deletions and insertions occurred during the course of PR10 evolution. In general, primary sequences of PR10s in various plant species may have a well conserved structure. Digital expression data of tomato and maize showed that expression of *PR10* genes may significantly increase in plant parts (root, lateral root and root tips) where it is more open to the mechanical perturbation and pathogenic attack, supporting the involvement of PR10 in plant defense. In phylogenetic tree, a clear monocot/polycot and dicot separation were observed. This separation could have been arising from the well conserved structure of *PR10* genes of monocots and polycots than dicots. All modeled species contained the same number of β-strands (seven) but α-helices varied between 2 and 4 depending on species. The results of this study will provide a theoretical reference regarding the primary, secondary and tertiary structure of PR10s in various plant species and will support the future studies that aiming to characterize the pathogenesis-related (PR) proteins.

Key words: *In silico*, pathogenesis-related protein 10, 3D structure, phylogenetic, polycot, divergence

Introduction

Pathogenic infection causes to induce various genes at the infection site and/or in other parts of the plant, which subsequently leads to developing a hypersensitive reaction (HR) and systemic acquired resistance (SAR)[1,2]. Pathogenesis-related (*PR*) genes are some of the first genes that overly expressed in different parts of the plants during an infection or under some stress-related conditions. Products of these genes comprise the very diverse group of proteins, including chitinases, glucanases, endoproteinases, peroxidases, proteinase inhibitors, osmotins, defensins, thionins and lipid transfer proteins (LTPs)[3]. The PR proteins are grouped into 17 families based on their primary structure, serological relationships and biological activities[4,5]. PR1, PR2, PR3 and PR4 have been characterized as β-1,3-glucanases, PR6 as proteinase inhibitors, PR9 as chitinases and PR11 as peroxidases, PR12 as defensins, PR13 as thionins, PR14 as lipid transfer proteins, and PR15 as oxalate oxidases, but precise biological functions of PR5, PR10, PR16 and PR17 remain unclear[6,7]. PR10 proteins, which was first identified as a major white birch pollen allergen (Bet_v_1)[8], are small (15–19 kDa) acidic intra-cellular proteins and comprise the largest family of PRs with more than its 100 members[9,10]. They are induced by pathogenic attack[11,12], extreme abiotic factors[4] or even by normal physiological changes in many plant species. Methyl jasmonate (MeJA), salicylic acid (SA), gibberellic acid (GA3), hydrogen peroxide (H_2O_2), sodium chloride (NaCl) and ethylene are reported to induce the transcription of *PR10* genes in different plant species[13,14]. In addition, PR10 proteins have been demonstrated to exhibit ribonuclease activity and to have cytokinin-binding ability in various plant species[15,16].

Although PR10 proteins are mainly known to be involved in plant defense under biotic/abiotic stress conditions, the overall biological function of many PR10 members still awaits to be elucidated. In the present study, we have comparatively analyzed the *PR10* genes and proteins in 28 plant species in order to understand the evolutionary (conservation or divergence) relationship between these genes/proteins in various plants. The results of this study will provide a theoretical reference for the future studies

*Author for correspondence
ertugrulfiliz@gmail.com

of structure-function relationships that characterize the pathogenesis-related (PR) proteins. In this context, we have analyzed exon/intron organization of *PR10* genes, determined the conserved motifs in PR10 protein sequences, evaluated the digital expression profile, constructed phylogenetic tree, and predicted 3D models.

Materials and Methods

Sequence Retrieval, Domain and Interaction Partner Analysis of PR10s

The PR10 protein sequences were retrieved from the protein database of NCBI (http://www.ncbi.nlm.nih.gov/protein/). The query against the plant protein database resulted in 28 plant species, including 20 dicots, 6 monocots and 2 polycots after the partial, redundant and ambiguous sequences were removed (Table 1). Then, domain analysis was performed by Pfam server for retrieved sequences in order to confirm Bet_v_1 (PF00407) domain structure (http://pfam.sanger.ac.uk)[17]. The interaction partners of PR10 proteins were predicted using STRING 9.1 server (http://string-db.org/)[18].

Physicochemical and Conserved Motif Analysis of PR10s

Sequence length, molecular weight, and theoretical isoelectric point (*pI*) of proteins were analyzed by Expasy's ProtParam server (http://web.expasy.org/protparam/)[19]. The conserved motif analysis was performed by MEME (Multiple Em for Motif Elicitation) tool[20] (http://meme.nbcr.net/meme/cgi-bin/meme.cgi) with following parameters: distribution of motifs, 0 or 1 per sequence; maximum number of motifs to find 5; minimum width of motif, 6; maximum width of motif, 50. The sub-cellular localizations of PR10s were predicted by CELLO v.2.5 (subCELlular LOcalization predictor)[21] (http://cello.life.nctu.edu.tw/).

Table 1 — Physicochemical properties and predicted sub-cellular localization of PR10s in 28 plant species

Species name	Groups	Access. No. (NCBI)	Length (aa)	MW (kDa)	*pI*	Predicted subcellular localization
Pisum sativum	Dicot	AAA90954	158	16.75	4.94	Cytoplasm
Gossypium barbadense	Dicot	ACM17134	159	17.31	4.95	Cytoplasm/Chloroplast
Capsicum annuum	Dicot	ACB30364	166	18.25	5.95	Cytoplasm
Arachis hypogaea	Dicot	AAU81922	150	16.22	5.34	Cytoplasm
Nicotiana tabacum	Dicot	AEY11296	160	17.77	5.56	Cytoplasm
Prunus domestica	Dicot	ABW99634	160	17.66	5.79	Cytoplasm
Prunus persica	Dicot	ABW99628	160	17.65	5.79	Cytoplasm
Capsicum baccatum	Dicot	ABC74797	159	17.24	5.21	Cytoplasm
Lilium regale	Monocot	AHC68874	157	16.84	5.39	Chloroplast/Cytoplasm
Vicia faba	Dicot	AFD29283	133	14.34	5.53	Cytoplasm
Crocus sativus	Monocot	ADL09408	162	17.40	6.08	Chloroplast/Cytoplasm
Triticum aestivum	Monocot	ACG68733	160	17.06	5.19	Cytoplasm
Vitis pseudoreticulata	Dicot	ABC86747	159	17.47	6.07	Cytoplasm
Pinus pinaster	Polycot	ADJ53040	160	17.80	6.17	Cytoplasm
Sorghum bicolor	Monocot	AAW83209	160	16.95	5.19	Chloroplast
Capsicum chinense	Dicot	CAI51309	159	17.31	4.81	Cytoplasm
Lupinus albus	Dicot	CAA03926	158	17.92	4.83	Cytoplasm
Tanacetum cinerariifolium	Dicot	AEL17175	157	17.18	5.79	Cytoplasm
Pinus monticola	Polycot	AAL50007	161	17.80	5.34	Cytoplasm
Pisum fulvum	Dicot	AAB07447	158	16.72	4.94	Cytoplasm
Solanum lycopersicum	Dicot	AHC08074	160	17.37	5.44	Cytoplasm
Ziziphus jujuba	Dicot	AGL07712	160	17.51	5.40	Cytoplasm
Medicago truncatula	Dicot	XP_003594834	158	16.81	4.76	Cytoplasm
Fragaria chiloensis	Dicot	ADN05762	157	17.27	5.52	Cytoplasm
Glycine max	Dicot	XP_006582821	158	16.77	4.69	Cytoplasm
Elaeis guineensis	Monocot	AEB96227	160	17.06	5.19	Cytoplasm
Rheum australe	Dicot	ACH63224	160	17.81	5.87	Cytoplasm
Oryza sativa	Monocot	BAD03969	160	16.90	4.88	Cytoplasm

Gene Structure and Expression Profile Analysis of PR10s

Exon/intron organization of *PR10* genes were analyzed by GSDS 2.0 (Gene Structure Display Server) (http://gsds.cbi.pku.edu.cn/)[22]. Digital expression profile of *PR10* genes were obtained from Genevestigator platform[23] (https://www.genevestigator.com/gv/plant.jsp).

Phylogenetic Analysis of PR10s

PR10 protein sequences were aligned by ClustalW[24] and phylogenetic tree was constructed by MEGA version 5.1[25] with following parameters; maximum likelihood (ML) method, Poisson correction, pairwise deletion and 1000 replicates bootstrap value.

3D Modeling of PR10s

Binding sites of PR10 proteins were predicted by 3DLigandSite server (http://www.sbg.bio.ic.ac.uk/3dligandsite/)[26], and visualized by PyMOL[27]. Structural evaluation and stereochemical analyses were done with Rampage Ramachandran plot analysis (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php)[28].

Results and Discussion

Sequence and Physico-chemical Analysis of PR10 Proteins

PR10 protein sequences were confirmed to contain Bet_v_1 (PF00407) domain structure, and found to have a close molecular weight (14.3-18.2 kDa but mainly 17-18 kDa) and amino acid length (157-166 aa), with a *pI* value of 4.69-6.17. The sub-cellular localization of PR10 proteins was predicted as cytoplasmic (Table 1). PR10s are reported to be small (17.1-18.4 kDa) acidic intra-cellular proteins[10]. This complies with the physico-chemical properties of PR10 proteins retrieved for this study.

The most conserved five motifs of PR10 proteins were detected. Motif 1 was observed in 24 of 28 species; motif 2 was observed in 25 of 28 species; motif 3 was observed in 26 of 28 species; motif 4 was observed in 27 of 28 species and motif 5 was present in all species. Only motif 1 was found to be related with the Bet_v_I domain structure (Fig. 1 and Table 2). P-loop structure with an amino acid sequence GXGGXG in birch pollen Betv1 resembles

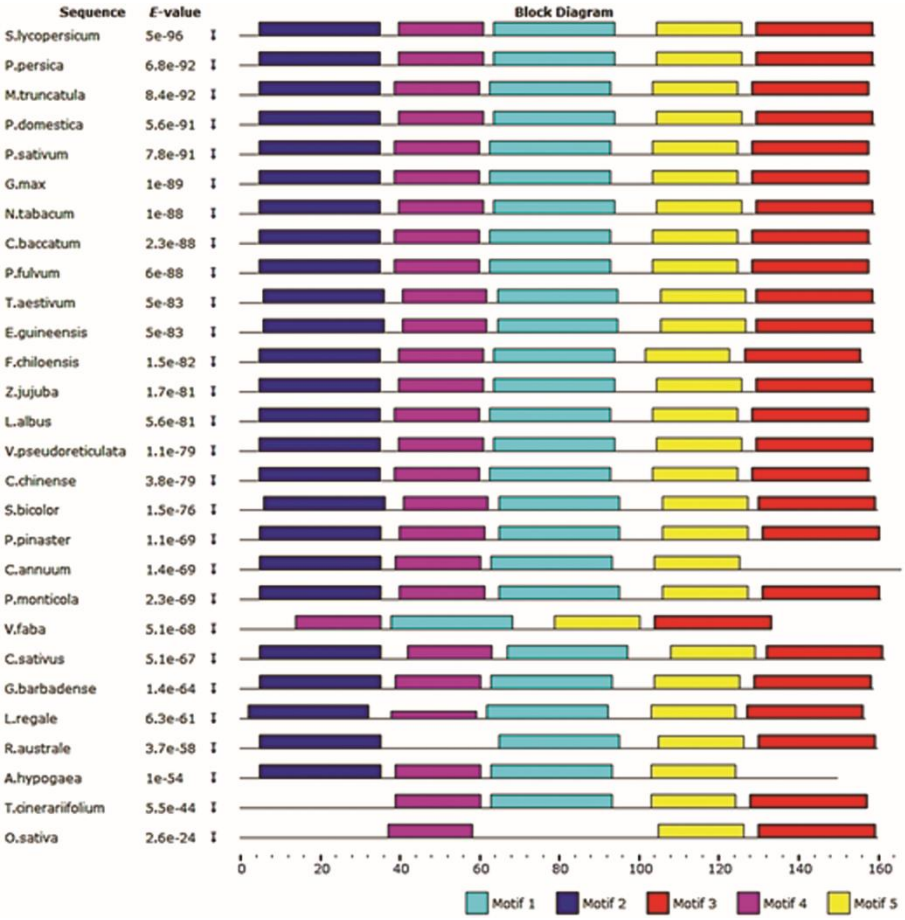

Fig. 1 — The schematic diagram of conserved motif analysis in PR10 proteins. Each motif was represented using boxes with different colors: Motif 1, cyan; Motif 2, blue; Motif 3, red; Motif 4, purple; and Motif 5, yellow.

Table 2 — The most conserved five motifs of PR10s in 28 plant species. Red residues in motif 4 indicate the phosphate-binding loop (P-loop; GXGGXG) (Saraste et al. 1990)

Motif number	Width Sequence	Protein Sequences	Pfam domain
1	30	FKYMKHRIDFIDEENCVCNYSLIEGDGLGD	Bet v I family
2	30	WEHEITSPVAPARLWKALVMDWHNLWPKLW	Not found
3	29	EEQHKQGKERADGLFKAIEQYCLANPDYY	Not found
4	21	SIEIVQ**GNGGPG**TIRQMNFVE	Not found
5	21	EARPNGGCICKWTCHYHTKGD	Not found

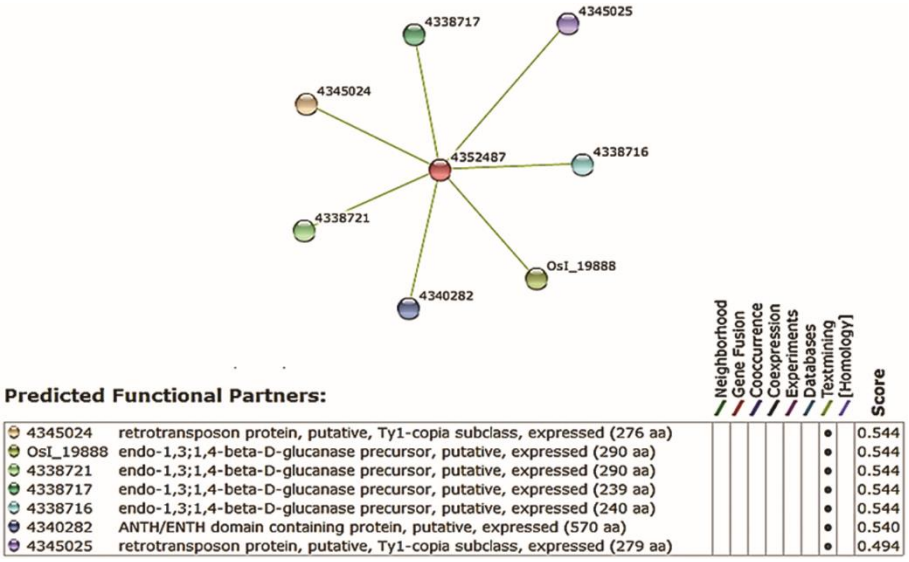

Fig. 2 — Predicted interaction partners of rice PR10 protein

a motif found in many nucleotide-binding proteins[29]. We found a similar motif with GNGGPG pattern in all species, except for *Rheum australe*. Overall, variations at different motif structures might be caused by deletion, insertion etc., causing to the formation of new catalytic sites in PR10 family in terms of improving the survivability under the various biotic and abiotic stress conditions.

Putative retrotransposon protein (Ty1-copia subclass), endo-1,3; 1,4-β-D-glucanase precursor and ANTH/ENTH domain containing protein were predicted to be main interaction partners of rice PR10 protein (Fig. 2). Retrotransposons are mobile genetic elements affecting genome structure, function and evolution in eukaryotic organisms[30]. 1,3;1,4-β-glucanase type I (E.C. 3.2.1.73; glycosyl hydrolase family 17) evolved specifically in the grasses that might control growth responses in plants[31]. Many ANTH Sla2/HIP1 and ENTH Ent1-2/Epsin1, 2, 3 domains play role as lipid binding domains, functioning in phosphoinositide enrichment at the target membranes, membrane trafficking, and in metabolism and cytokinesis[32]. Interaction partners of rice PR10 proteins implicate that *PR10* could be involved in a diverse metabolic pathways in plants.

Exon/intron Structure and Expression Profiles of *PR10* Genes

Gene structure (exon/intron) analysis provides valuable information for understanding of genetic and evolutionary relationship between inter-and intra-species[33,34]. Exon-intron analysis was performed for six selected species, including *G. max*, *M. truncatula*, *O. sativa*, *P. persica*, *S. lycopersicum* and *S. bicolor* that have complete genome sequenced, thereby genomic DNA and CDS were available for exon-intron deduction. The analysis in six selected species showed that *PR10* genes have similar exon/intron patterns (two exons) without any monocot/dicot separation (Fig. 3). In order to see how exons are related to each other, we obtained the nucleotide sequences of these potential two exons from the JBrowser of Phytozome and aligned by ClustalW (Fig. 4). Notably, some nucleotide residues in alignment were found to be highly conserved in both

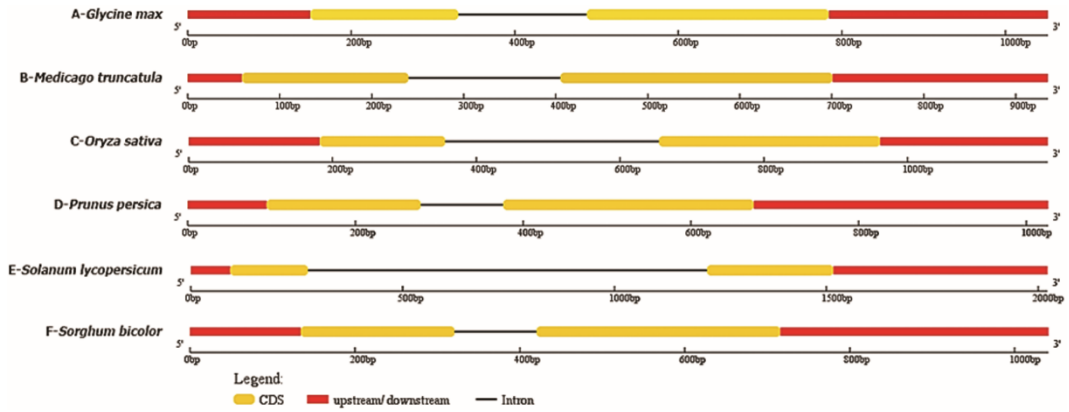

Fig. 3 — Exon/intron and upstream/downstream organization of *PR10* genes in selected six species. Exon, intron, and upstream/ downstream regions were indicated as filled yellow boxes, lines, and red boxes, respectively.

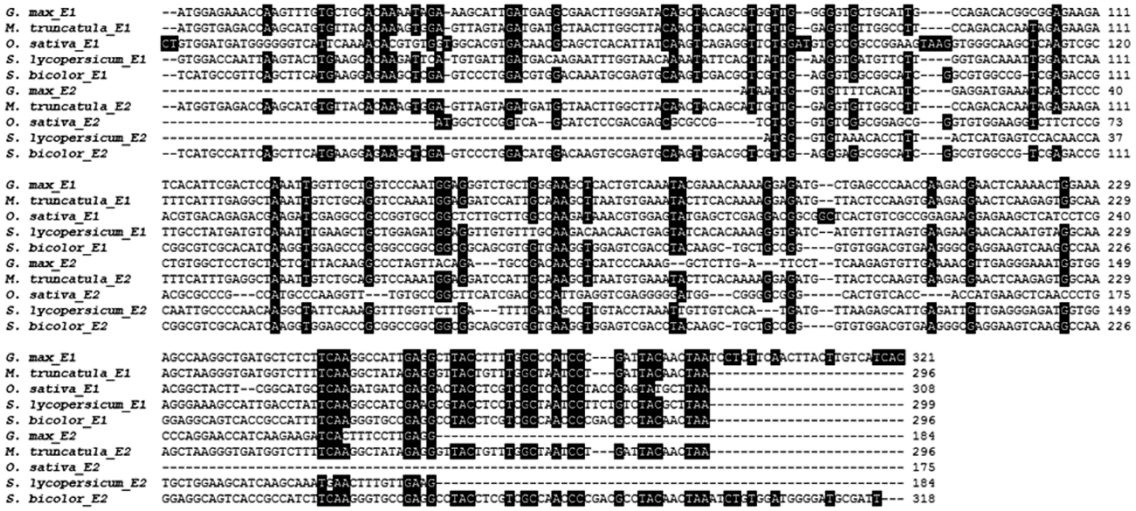

Fig. 4 — Alignment of exons in six selected species, except for *P. prunus persica*. Two potential exons were predicted by GSDS server for selected species. Nucleotide sequence of these two exons were obtained from the JBrowser of Phytozome and aligned by ClustalW. Black shaded residues show the highly conserved nucleotides. E1 and E2 tags after species name indicate the exon 1 and exon 2, respectively.

exons, indicating that these nucleotide residues may encode the some functional residues of the proteins. Overall, this shows that *PR10* genes might have been well preserved along the divergence of monocots and dicots during *PR10* gene evolution in plants.

Digital gene expression profile of *PR10* genes were analyzed for selected two species (tomato and maize) at developmental and anatomical part levels (Fig. 5). Tomato and maize represent for dicots and monocots, respectively. Tomato and maize genes were evaluated within seven developmental stages. Tomato gene (*Solyc09g090980.2*) was highly expressed in the early stage of flowering and fruit ripening (Fig. 5A). In maize, expression level of *GRMZM2G075283* and *GRMZM2G112524* genes relatively increased in seed germination period and *GRMZM2G112524* was found to be highly expressed during seed development period (Fig. 5B). The reason of increased gene expression level in seed germination, flowering, fruit ripening and seed development could be related with increased plant hormones because plant PR10 proteins are known as more properly linked with general phytohormone-binding proteins (PhBP)[35-37]. At anatomical level, tomato gene (*Solyc09g090980.2*) was observed to mainly express in lateral root, root, exocarp (skin), pericarp and fruit, and the highest expression levels were found in lateral root and exocarp (skin) (Fig. 5C). In maize, *GRMZM2G075283* gene was detected to be highly expressed in glume while *GRMZM2G112524* gene was highly expressed in caryopsis (Fig. 5D). Also, *GRMZM2G075283* and *GRMZM2G112524* genes were observed to moderately be expressed in roots and lateral roots. This shows that *PR10* gene

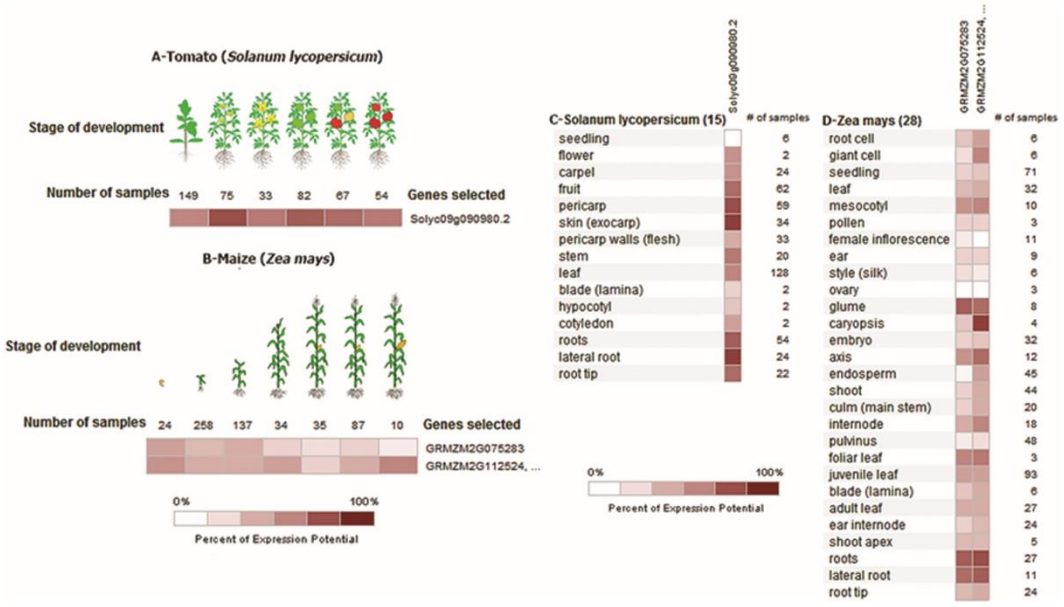

Fig. 5 — The digital expression profile analysis of tomato and maize *PR10* genes by using gene investigator search engine. Heat map shows the expression profiles in different developmental stages (5A and 5B, respectively) of tomato and maize at different anatomical parts (5C and 5D, respectively).

expression significantly increases in the parts (root, lateral root and root tips) of plant where it is more open to the mechanical perturbation and pathogenic attack.

Phylogenetic Analysis of PR10s

Phylogenetic tree was constructed for PR10 protein sequences of 28 species. First, constructed phylogenetic tree was divided into two main groups, namely as A and B. Then, group A and B were further subdivided into subgroups A1 and A2, and B1 and B2, respectively (Fig. 6). All monocots and polycots were observed to be clustered in group B (B1 and B2) with 67% bootstrap value while all dicots were in group A (A1 and A2) with 61% bootstrap value. In dicot group, *R. australe* (A2) was separated from other dicots and in monocot group, *O. sativa* (B1) was diverged from other monocots. A distinction between monocots/polycots (67% bootstrap value) and dicots (61% bootstrap value) could be explained with the well conserved structure of *PR10* genes in monocots and polycots.

3D Structure Prediction of PR10s

G. max, *M. truncatula*, *O. sativa*, *P. persica*, *S. lycopersicum*, and *S. bicolor* plants selected for 3D modeling (Fig. 7). Selection based on the phylogenetic clustering of these species (Fig. 6). *O. sativa* and *S. bicolor* selected as representatives of monocots while *G. max*, *M. truncatula*, *P. persica* and *S. lycopersicum* selected as representatives of dicot species. In models, seven β-strands and three α-helices in *G. max*, *S. lycopersicum*, and *S. bicolor*; seven β-strands and four α-helices in *M. truncatula* and *O. sativa*; and seven β-strands and two α-helices in *P. persica* were predicted. Notably, all six models contained the same number of β-strands (seven). In 3D models, phosphate-binding loop (P-loop; GXGGXG), which is a strongly conserved region of nucleotide-binding proteins, were identified. P-loop is located at positions Gly46-48-49-51 in *G. max* and *M. truncatula*; Gly44-46-47-49 in *O. sativa*; Gly47-49-50-52 in *P. persica* and *S. lycopersicum* and Gly48-50-51-53 in *S. bicolor*. Positional variations in glycine (Gly) residues in P-loop motif suggest that mutations, deletions and insertions may cause the variations within *PR10* genes during the course of evolution.

Protein-ligand binding sites have a vital importance in understanding the functional diversity among species[38]. Thus, predicted binding sites were identified in PR10 proteins. These are; 23Phe-24Val-27Ala-39Ile-57Ile-59Phe-70His-82Tyr-84Tyr-101Tyr-144Leu in *P. persica*; 23Leu-27Ala-69His-81Tyr-83Tyr-89Ala-90Ala-100Phe-139Lys-140Ala-143Leu in *G. max*; 35Ile-83Tyr-136Gly-138Ala-139Lys-140Gly-141Asp-142Gly-143Leu-144Phe in *M. truncatula*; 24Ala-25Val-28Trp-56Arg-58Phe-60Phe-69Met-71Glu-141Val-144Ile in *S. bicolor*; 23Leu-24Val-27Phe-55Lys-57Met-59Phe-65Ile-68Leu-70His-84Tyr-91Val-101Tyr-141Ala-144Leu in *S. lycopersicum*

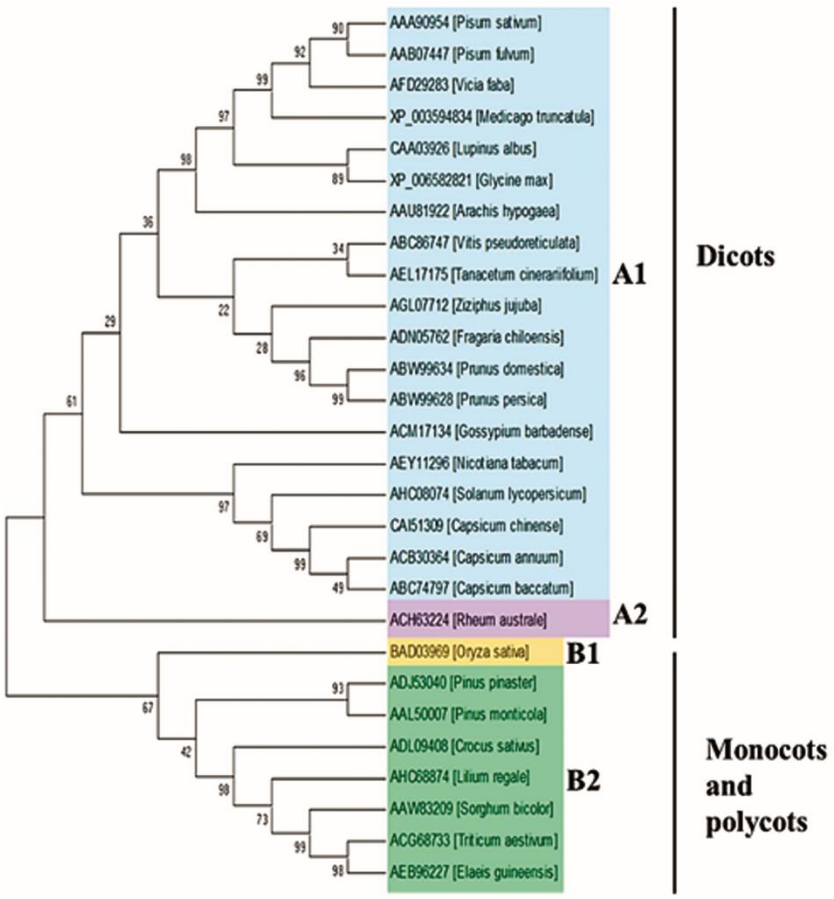

Fig. 6 — Phylogenetic analyses of PR10 proteins in 28 plant species. Sequence alignment was done by ClustalW and phylogenetic tree was generated by using MEGA 5.1 with maximum likelihood (ML) method for 1.000 replicates bootstrap value.

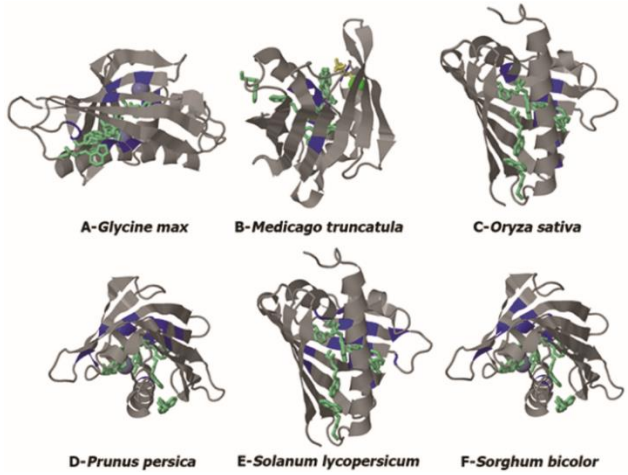

Fig. 7 — The predicted 3D structure and putative active site of PR10 proteins generated by 3DLigandSite server. The green color indicates the predicted active site position of PR10s.

and 23Phe-24Ser-26Ala-54Met-66Phe-89Ala-92Lys-141Tyr-144Met in *O. sativa* (Fig. 8). Generally, in proteins, solvent-exposed residues show variational tendencies due to mutations[39], and amino acid residues at such binding sites represent higher degree of variations[40]. Therefore, residues at predicted binding sites in PR10 proteins might be connected with deleterious or insertional mutations etc., leading to the formation of new catalytic sites with different functions in diverse metabolic pathways.

Model validation was done by Ramachandran plot analysis. It was found that 97.4%, 98.1%, 91.8%, 92.4%, 98.1% and 87.8% of residues were in favored region; 2.6%, 1.9%, 5.1%, 7.6%, 1.9% and 8.3% of residues were in allowed region and 0.0%, 0.0%, 3.2%, 0.0%, 0.0%, and 3.8% of residues were in outlier region in *G. max*, *M. truncatula*,, *O. sativa*, *P. persica*, *S. lycopersicum*, and *S. bicolor*, respectively. This indicates that our 3D models were fairly in good quality.

Overall, *PR* genes are highly expressed in different anatomical parts of the plants under various biotic/abiotic stress conditions. Products of these genes comprise the very diverse group of proteins with various metabolic functions. However, we still have limited knowledge about many

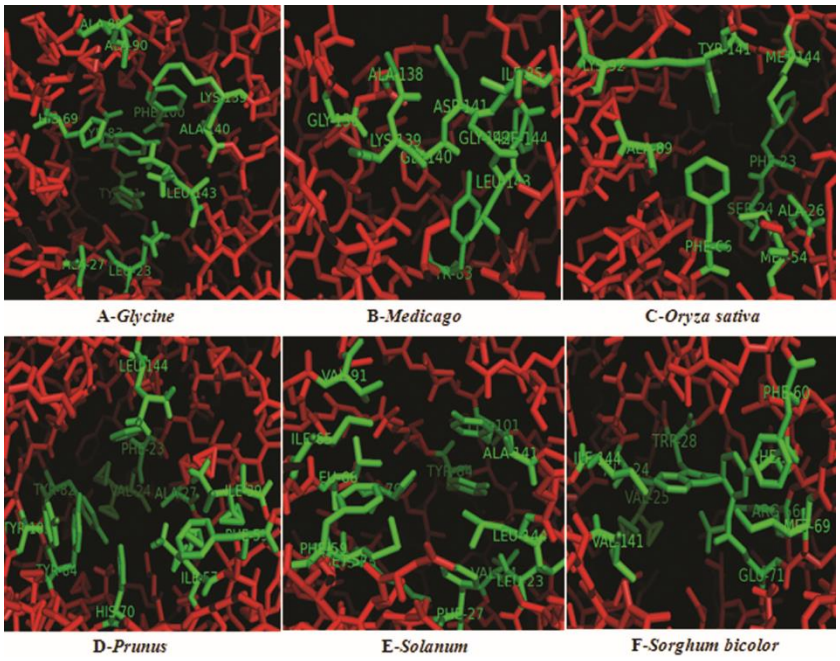

Fig. 8 — Ligand binding sites of PR10 proteins visualized by PyMOL. Green color shows the predicted ligand binding residues in PR10 proteins.

PR proteins of different plant species. Therefore, to have an insight about the structural and functional roles of PR proteins, we have comparatively analyzed the *PR10* genes and proteins from 28 different plant species. Analyses revealed that PR10 proteins are structurally well conserved and functionally involve in plant defense. However, to better understand the physiological response of PR10 proteins in plant defense, further genome-wide analyses are required with various stress perturbations in various plant species. We believe that the results of this study will provide a theoretical reference regarding the primary, secondary and tertiary structure of PR10s in various plant species and will support the future studies that aims to characterize the pathogenesis-related (PR) proteins.

References

1 Jackson A O & Taylor C B, Plant-microbe interactions: life and death at the interface, *Plant Cell*, 8 (1996) 1651-1668.

2 Durrant W E & Dong X, Systemic acquired resistance, *Annu Rev Phytopathol*, 42 (2004) 185-209.

3 Jain S, Kumar D, Jain M, Chaudhary P, Deswal R *et al*, Ectopic overexpression of a salt stress-induced pathogenesis-related class 10 protein (*PR10*) gene from peanut (*Arachis hypogaea* L.) affords broad spectrum abiotic stress tolerance in transgenic tobacco, *Plant Cell Tiss Org*, 109 (2012) 19-31.

4 Liu J J & Ekramoddoullah A K M, The family 10 of plant pathogenesis-related proteins: their structure, regulation, and function in response to biotic and abiotic stresses, *Physiol Mol Plant Pathol*, 68 (2006) 3-13.

5 Vidhyasekaran P, *Fungal pathogenesis in plants and crops*, 2nd edn (Boca Raton: CRC press), 2008.

6 Edreva A, Pathogenesis-related proteins: research progress in the last 15 years, *Gen Appl Plant Physiol*, 31 (2005) 105-124.

7 Lee O R, Pulla R K, Kim Y J, Balusamy S R D & Yang D C, Expression and stress tolerance of *PR10* genes from *Panax ginseng* CA Meyer, *Mol Biol Rep*, 39 (2012) 2365-2374.

8 Breiteneder H, Pettenburger K, Bito A, Valenta R, Kraft D *et al*, The gene coding for the major birch pollen allergen Betv1, is highly homologous to a pea disease resistance response gene, *Embo J*, 8 (1989) 1935-1938.

9 Fernandes H, Pasternak O, Bujacz G, Bujacz A, Sikorski M M *et al*, *Lupinus luteus* pathogenesis-related protein as a reservoir for cytokinin, *J Mol Biol*, 378 (2008) 1040-1051.

10 Lebel S, Schellenbaum P, Walter B & Maillot P, Characterization of the *Vitis vinifera PR10* multigene family, *BMC Plant Biol*, 10 (2010) 184-197.

11 McGee J D, Hamer J E & Hodges T K, Characterization of a PR-10 pathogenesis-related gene family induced in rice during infection with *Magnaporthe grisea*, *Mol Plant Microb*, 14 (2001) 877-886.

12 Tanaka N, Che F S, Watanabe N, Fujiwara S, Takayama S *et al*, Flagellin from an incompatible strain of *Acidovorax avena*e mediates H_2O_2 generation accompanying hypersensitive cell death and expression of *PAL*, *Cht-1*, and *PBZ1*, but not of LOX in rice, *Mol Plant Microb*, 16 (2003) 422-428.

13 Agrawal G K, Rakwa R & Jwa N S, Differential induction of three pathogenesis-related genes, *PR10*, *PR1b* and *PR5* by the ethylene generator ethephon under light and dark in rice (*Oryza sativa* L.) seedlings, *J Plant Physiol*, 158 (2001) 133-137.

14 Xiong L & Yang Y, Disease resistance and abiotic stress tolerance in rice are inversely modulated by an abscisic acid inducible mitogen-activated protein kinase, *Plant Cell*, 15 (2003) 745-759.

15 Swoboda I, Hoffmann-Sommergruber K, O'Ríordáin G, Scheiner O, Heberle-Bors E *et al*, Bet_v_1 proteins, the major birch pollen allergens and members of a family of conserved pathogenesis-related proteins, show ribonuclease activity *in vitro*, *Physiol Plantarum*, 96 (1996) 433-438.

16 Fujimoto Y, Nagata R, Fukasawa H, Yano K, Azuma M *et al*, Purification and cDNA cloning of cytokinin-specific binding protein from mung bean (*Vigna radiata*), *Eur J Biochem*, 258 (1998) 794-802.

17 Sonnhammer E L, Eddy S R & Durbin R, Pfam: a comprehensive database of protein domain families based on seed alignments, *Proteins*, 28 (1997) 405-420.

18 Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M *et al*, STRING v9. 1: protein-protein interaction networks, with increased coverage and integration, *Nucleic Acids Res*, 41 (2013), 808-815.

19 Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins M R *et al*, Protein identification and analysis tools on the ExPASy server, In *The Proteomics Protocols Handbook*, Edited by John M. Walker (Humana Press, New York) 2005, 571-607.

20 Timothy L, Mikael B B, Buske F A, Frith M, Grant C E *et al*, MEME SUITE: tools for motif discovery and searching, *Nucleic Acids Res*, 37 (2009) 202-208.

21 Yu C S, Chen Y C, Lu C H & Hwang J K, Prediction of protein subcellular localization, *Proteins*, 64 (2006) 643-651.

22 Guo A Y, Zhu Q H, Chen X & Luo J C, GSDS: a gene structure display server, *Yi Chuan*, 29 (2007) 1023-1026.

23 Hruz T, Laule O, Szabo G, Wessendorp F, Bleuler S *et al*, Genevestigator v3: a reference expression database for the meta analysis of transcriptomes, *Adv Bioinformatics*, 2008.

24 Thompson J D, Higgins D G & Gibson T J, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res*, 22 (1994) 4673-4680.

25 Tamura K, Peterson D, Peterson N, Stecher G, Nei M *et al*, MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, *Mol Biol Evol*, 28 (2011) 2731-2739.

26 Wass M N, Kelley L A & Sternberg M J, 3DLigandSite: predicting ligand-binding sites using similar structures, *Nucleic Acids Res*, 38 (2010) 469-473.

27 DeLano W L, The PyMOL molecular graphics system, 2002.

28 Lovell S C, Davis I W, Arendall W B, Bakker P I W, Word J M *et al*, Structure validation by Cα geometry: φ, ψ and Cβ deviation, *Proteins*, 50 (2003) 437-450.

29 Gajhede M, Osmark P, Poulsen F M, Ipsen H, Larsen J N *et al*, X-ray and NMR structure of Bet v 1, the origin of birch pollen allergy, *Nat Struct Biol*, 3 (1996) 1040-1045.

30 Weber B, Wenke T, Frömmel U, Schmidt T & Heitkam T, The Ty1-copiafamilies SALIRE and Cotzilla populating the *Beta vulgaris* genome show remarkable differences in abundance, chromosomal distribution, and age, *Chromosome Res*, 18 (2010) 247-263.

31 Thomas B R, Inouhe M, Simmons C R & Nevins D J, Endo-1,3;1,4-β-glucanase from coleoptiles of rice and maize: role in the regulation of plant growth, *Int J Biol Macromol*, 27 (2000) 145-149.

32 De Craene J O, Ripp R, Lecompte O, Thompson J D, Poch O *et al*, Evolutionary analysis of the ENTH/ANTH/VHS protein superfamily reveals a coevolution between membrane trafficking and metabolism, *BMC Genomics*, 13 (2012) 297.

33 Paquette S M, Bak S & Feyereisen R, Intron-exon organization and phylogeny in a large superfamily, the paralogous cytochrome *P450* genes of *Arabidopsis thaliana*, *DNA Cell Biol*, 19 (2000) 307-317.

34 Hu L F & Liu S Q, Genome-wide identification and phylogenetic analysis of the *ERF* gene family in cucumbers, *Genet Mol Biol*, 34 (2011) 624-633.

35 Buttner M & Singh K B, *Arabidopsis thaliana* ethylene-responsive element binding protein (AtEBP), an ethylene-inducible, GCC box DNA-binding protein interacts with an ocs element binding protein, *Proc Natl Acad Sci USA*, 94 (1997) 5961-5966.

36 Goldberg-Moeller R, Shalom L, Shlizerman L, Samuels S, Zur N *et al*, Effects of gibberellin treatment during flowering induction period on global gene expression and the transcription of flowering-control genes in *Citrus* buds, *Plant Sci*, 198 (2013) 46-57.

37 Ruszkowski M, Sliwiak J, Ciesielska A, Barciszewski J, Sikorski M *et al*, Specific binding of gibberellic acid by cytokinin-specific binding proteins: a new aspect of plant hormone-binding proteins with the PR-10 fold, *Acta Crystallogr*, 70 (2014) 2032-2041.

38 Khazanov N A & Carlson H A, Exploring the composition of protein-ligand binding sites on a large scale, *Plos Comput Biol*, 9 (2013) e1003321.

39 Bowie J U, Reidhaar-Olson J F, Lim W A & Sauer R T, Deciphering the message in protein sequences: tolerance to amino acid substitutions, *Science*, 247 (1990) 1306-1310.

40 Magliery T J & Regan L, Sequence variation in ligand binding sites in proteins, *BMC Bioinformatics*, 6 (2005) 240.