The Bioscan 10(4): 1929-1933, 2015 (Supplement on Genetics and Plant Breeding) www.thebioscan.in
AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES

SEQUENCE AND STRUCTURAL ANALYSIS OF PATHOGENESIS RELATED PROTEIN CLASS I CHITINASE FROM SEED COAT OF *GLYCINE MAX* [L.] MERR

MANVENDRA SINGH* AND A. K. RAI
Department of Genetics and Plant Breeding,
Institute of Agricultural Sciences, Banaras Hindu University, Varanasi - 221 005, Utter Pradesh, INDIA
e-mail: manvendra.jhansi76@gmail.com

KEYWORDS
Chitinase
Pathogenesis related
Protein
Homology modeling
Conserved domain

Received on :
29.04.2015

Accepted on :
15.09.2015

***Corresponding author**

ABSTRACT
Chitinase are enzymes that catalyze the hydrolysis of β-1,4-N-acetylglucosamine linkage present in chitin. Chitin is a major component of fungal cell walls, chitinases play a role in plant defense against pathogens. Proteomic studies carried on 32 kDa class I chitinase of soybean seed coat, the FASTA sequences for enzyme were taken from NCBI. The sequence with known PDB structure was taken for multiple sequence alignment. The sequences were compared with BLOSUM 62 matrix. The sequence similarity among these chitinases varied between 30 to 70 %, with maximum similarity exists between soybean seed coat chitinase and rice chitinase.Homology modeling at SWISS- MODEL gave 3 models with different scores. The model 1developed using jack bean chitinase has highest homology. Two conserved domain ChtBD1 [cd00035], Chitin binding domain, involved in recognition or binding of chitin subunits and Glycoside hitinase_glyco_hydro_19[cd00325], Glycoside hydrolase family 19 chitinase domain were detected in the sequence using Conserved domain detection tool at NCBI. Soybean seed coats are particularly rich in defense related proteins and peptides, although there are abundant proteins that have yet to be identified.In this study revealed complex structure and more refined mechanism of the chitinase catalysis.

INTRODUCTION

Plants produce an array of proteins when explored to biological stress. Most of the proteins expressed during the pathogenesis, play important roles in the defences of plants against bacterial and fungal pathogens (Sajeesh *et al.,* 2014) (Chavan and Suryawanshi, 2014). In recent years, the antimicrobial potential of some members of pathogenes-related (PR) proteins has been reported (Van Loon *et al.,* 2006). Chitinase are one of them which extensively distributed among plants, fungi, bacteria and viruses. In higher plants, chitinase are used as defense against plant pathogen (Koga *et al.,* 1999). These enzymes are found at low levels in healthy plants; however, their expression is increased during pathogen attack. The production of chitinase elicits other plant responses including the synthesis of antifungal phytoalexins (Gooday, 1999) (Ingle *et al.,* 2014). The antifungal activity of chitinase and β-1, 3-glucase cause rapid lysis of fungal hyphal tips and germinating spores. The enzymes are an effective tool for the complete degradation of mycelial or conidial walls of phytopathogenic fungi (Hakala *et al.,* 1993).

Chitinases are glycosyl hydrolases that catalyse the hydrolytic cleavage of the β-1, 4-glycoside bond present in biopolymers of *N*-acetylglucosamine (Collinge *et al.*, 1993). The main substrate of chitinases is chitin, an insoluble homopolymer of β-1, 4-linked *N*-acetylglucosamine (GlcNAc) residues which is the second most abundant polymer in nature after cellulose(Ornum, 1992) and serves a structural role in fungal cell walls and arthropod cuticles including those of insects, nematodes and crustaceans (Kramer, 2005). Depending on the organism of origin, these enzymes have different functions. Bacteria produce chitinases to meet nutritional needs. In animals and plants, chitinases mainly play a role in the defense against pathogen attacks(Patil*et al.*, 2000).

Seed chitinase in soybean, little attention has been paid to the physiological and biochemical basis underlying its defense mechanisms in response to pathogen and herbivore attack (Gomez *et al.*, 2002), (Vega-sanchez*et al.*, 2005). A recent investigation on the proteomics of seed filling in soybean showed that >600 proteins are expressed during five key stages of seed development. However, most of them, including 7% involved in plant defense, have not been purified (Hajduch *et al.*, 2005). In soybean seed coat, a 32-kDa class I chitinase have been identified (Gijzen *et al.*, 2001).

In present study attempt has been made to determine class I chitinase from seed coat of soybean, homologous search and sequence alignment followed by structural modeling to identify the conserve domain. This information is useful few simultaneous additions of conserved domains are suggested to develop pathogen resistant agriculturally important plant as important feature for plant defense.

MATERIALS AND METHODS

Sequence retrieval and alignment

Sequence of soyabean seed coat class I chitinase in FASTA format was retrieved from protein database of NCBI. The

sequence ID is gi|12698917|gb|AAK01734.1|AF335589_1 chitinase class I of *Glycine max*. The sequence was analyzed for its biochemical properties such as pI, extinction coefficient and other features using ProtPram from ExPASy proteomic server (http://web.expasy.org/protparam/). The sequence obtained was compared with non-redundant protein database using BLAST from NCBI (http://blast.ncbi.nlm.nih.gov/Blast). Construction of a multiple sequence alignment aims at arranging residues with inferred common evolutionary origin or structural/functional equivalence in the same column position for a set of sequences. The high homology sequences obtained after BLAST analysis was submitted CLUSTAL W (ExPASy Proteomic server) for multiple sequence alignment. The alignment file obtained was analyzed using JALVIEW 3.0.Progressive methods assemble a multiple alignment by making a series of pair wise alignments of sequences or pre-aligned groups. The order of these pair wise alignments is guided by a tree or dendrogram so that similar sequences tend to be aligned before divergent sequences. Using scoring functions based on general residue substitution models, classic progressive methods such as ClustalW (Thompson *et al*., 1994).

Comparative modelling and molecular visualization

The sequence was submitted to SWISS-MODEL in FASTA format for structural determination. The PDB model with significant score was selected for structural analysis. The PDB file obtained after homology modelling was visualized using PYMOL and RASMOL. The PDB file was submitted to PDBSUM for the detailed structural analysis.The program PROCHECK was used for checking the stereochemistry and quality of the model(Laskows ki *et al*., 1997). These are due to interesting properties of the structure or possible errors in interpretation, and require further investigation during rebuilding step. The PDB file was submitted to DALI server for structural comparison. PDBsum does contain some functional annotation. Data from the Gene Ontology annotations for the corresponding UniProt sequence are provided where available as is functional annotation from the UniProt Knowledgebase (The Gene Ontology Consortium, 2000).

Structure-function relationship

The complementary analysis for protein possible function was analyzed using PROFUNC server. The program compares the structure with other protein and based on fold similarity can elucidate the functional features such as active site and ligand binding site. 3D structure of a protein opens up the possibility of ascertaining its function from an analysis of that structure. Many methods have been developed for predicting protein function from structure (Laskowski *et al.,* 2005).

RESULTS AND DISCUSSION

Sequence retrieval and analysis

The linear amino acid sequence of chitinase class I (Glycine max) derived protein database with protein Id >gi|12698917|gb|AAK01734.1|AF335589_1 contains 320 amino acid residues (Fig. 1). The theoretical pI and extinction coefficient obtained from Protparam are 7.40 and 60360 M^{-1} cm^{-1} respectively. Total number of positively charged residues (Asp &Glu) and negatively charged residue (Lys and Arg) are 25 and 26 respectively.

Secondary structure

The secondary structure of the protein is inferred from the

```
        10         20         30         40         50         60
MKNMKLCSVM LCLSLAFLLG ATAEQCGTQA GGALCPNRLC CSKFGWCGDT DSYCGEGCQS
        70         80         90        100        110        120
QCKSATPSTP TPTTPSSGGD ISRLISSSLF DQMLKYRNDG RCSGHGFYRY DAFIAAAGSF
       130        140        150        160        170        180
NGFGTTGDDN TRKKEIAAFL AQTSHETTGG WASAPDGPYA WGYCFINEQN QATYCDGGNW
       190        200        210        220        230        240
PCAAGKKYYG RGPIQLTHNY NYGQAGKALG LDLINNPDLV ATDATVSFKT ALWFWMTAQG
       250        260        270        280        290        300
NKPSSHDVIT GRWTPSSADS SAGRAPGYGV ITNIINGGLE CGHGQDNRVQ DRIGFYRRYC
       310        320
QMMGISPGDN LDCNNQRPFA
```

Figure 1: The primary sequence obtained from protein sequence database

Table 1: Models generated from SWISS model

Models	Residue range	Based on template	Resolution (Å)	Sequence identity (%)	E-value
1	80-320	1dxjA (Jack bean)	1.80	69.84	$1.69e^{-95}$
2	24-319	2dkvA(*Oryza sativa*)	2.00	66.22	$1.93e^{-102}$
3	24-63	1q9bA (Hevein)	1.50	69.05	$1.07e^{-8}$

Table 2: Cleft and the accessibility in the tertiary structure

Gap region	Volume	Accessible vertices	Buried vertices	Ave. depth
1	3225.66	66.11%	10.23%	1
2	494.86	66.61%	6.26%	8
3	386.44	61.70%	4.32%	7
4	377.16	59.93%	6.43%	2
5	280.55	64.44%	8.81%	3
6	356.06	65.96%	6.46%	9
7	266.20	54.30%	4.75%	5
8	282.66	54.05%	4.05%	4
9	241.73	59.67%	3.77%	10
10	198.70	72.63%	5.35%	6

Ramachandran angles and the intra-molecular hydrogen bonding of peptide backbone.The 12 alpha helices form most of the protein structure as seen in the topology diagram. The secondary structure of seed coat class Ichitinase consists of 12 helices, 22 beta turns and 3 gamma turns. It contains two right handed spiral disulfide bond Cys 102- Cys 164 and Cys 281- Cys 313 (Fig. 2). Disulfide bridges are very well conserved structural features in extracellular and other class I chitinases synthesized at the rough endoplasmic reticulum in the plant.

Sequence alignment

Thesequence with known PDB structure was taken for multiple sequence alignment. The sequence were compared with

Table 3: The r. m. s. deviation in Cá positions of seed coat class Ichitinaseand homologous protein structures of DALI databse

Homologous Protein (source)	PDB Code	Z-score	Aligned residues	RMSD (Å)	% identity
Barley	1cns	40	238	0.7	69
Jackfruit	1dxj	39.3	237	0.9	70
Brassica juncea	2z37	39.7	240	1.1	62
Rice	3iwr	43.9	279	0.5	70

Table 4: Ligand binding template analysis

Matched PDB entry identical/similar	Matched residues	template residues	RMSD	Similarity score	Matched residues
1cns Crystal structure of chitinase at 1.91a resolution	Glu146, Glu168, Thr197	Glu67, Glu89, Ser120	0.19Å	487.50	24/4

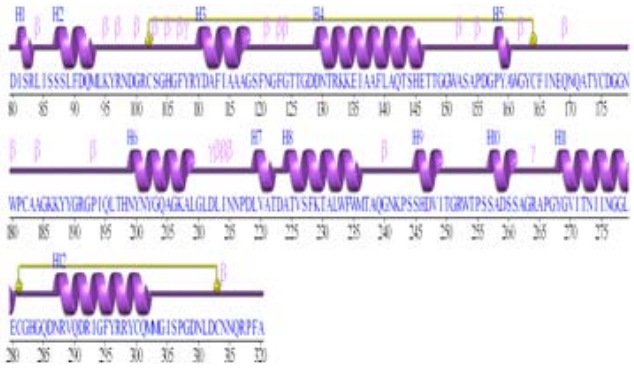

Figure 2: Secondary structural elements and connectivity of the chitinase, The alpha helices strand is shown as helices. The yellow line indicates disulfide bonds

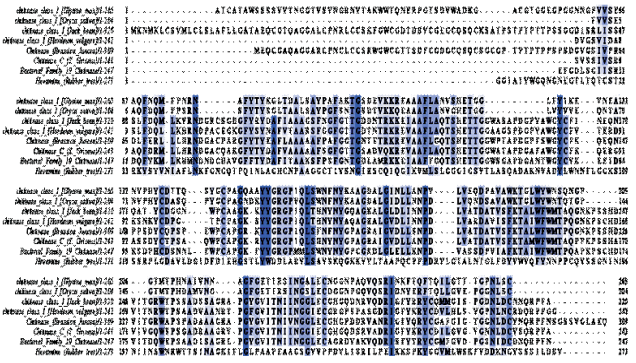

Figure 3: Multiple sequence alignment of soybean seed coat chitinase with other class I realted chitinases using CLUSTAL W. The Boxes dark grey, light grey and pale grey corresponds to the conserved residues, substitution by a similar type of amino acid and substitution by a non-similar type of amino acid residues, respectively.

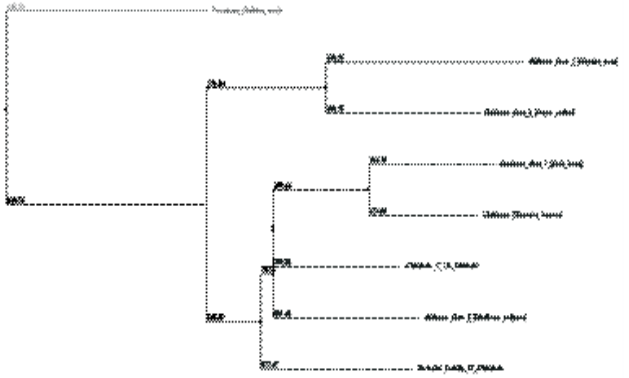

Figure 4: Neighbor joining tree using BLOSUM62 scores for the residue pair at each aligned position to measure of similarity between each pair of sequences in the alignment

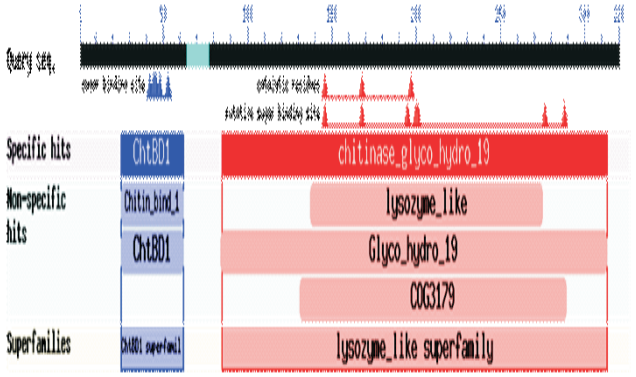

Figure 5: Conserved domain detection of the chitinase class I *Glycine max*

BLOSUM 62 matrix and coloured based on conservation of residues (Fig. 3). The sequence similarity among these chitinases varied between 30 to 70 %, with maximum similarity exists between soyabean seed coat chitinase and rice chitinase as shown in the neighbor joining tree(Fig. 4).

Comparative Modeling and molecular visualization

The FASTA sequence submitted to automated mode modeling, gave 3 models with different scores. The model 1developed using jack bean chitinase has highest homology and was used for further analysis. The 1-79 residues were omitted for the model building. The residue range coverage was 80-320 amino acids (Table 1).

Conserved Domain Detection

Two conserved domain ChtBD1 [cd00035], Chitin binding domain, involved in recognition or binding of chitin subunits and Glycoside chitinase_glyco_hydro_19[cd00325]. Glycoside hydrolase family 19 chitinasedomain were detected in the sequence using Conserved domain detection tool at NCBI (Fig. 5). Through a sequence comparison with homologous plant chitinases as well as a structural comparison

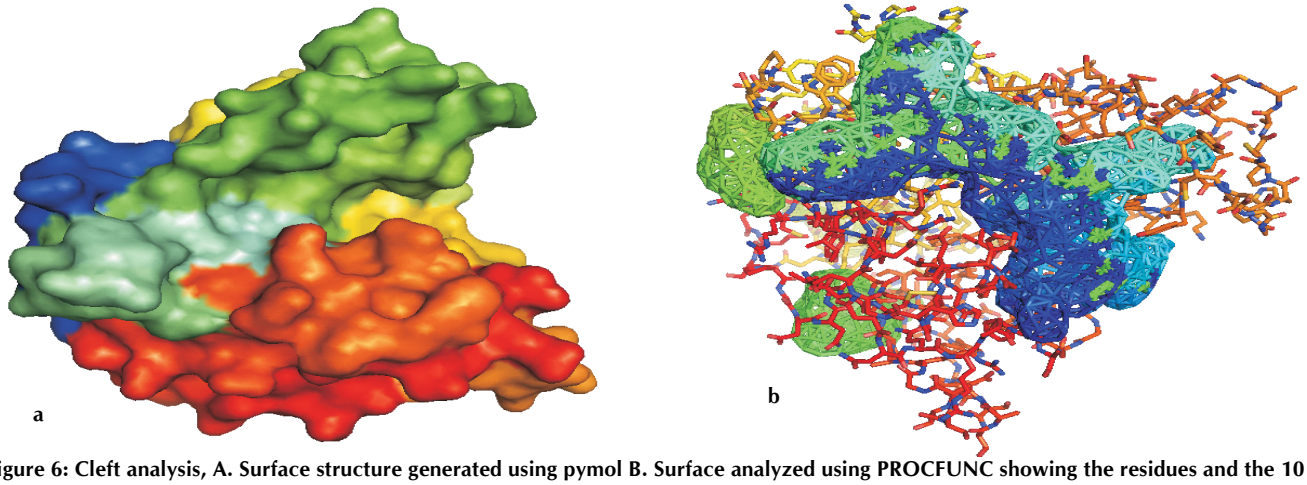

Figure 6: Cleft analysis, A. Surface structure generated using pymol B. Surface analyzed using PROCFUNC showing the residues and the 10 cleft, colored in different color

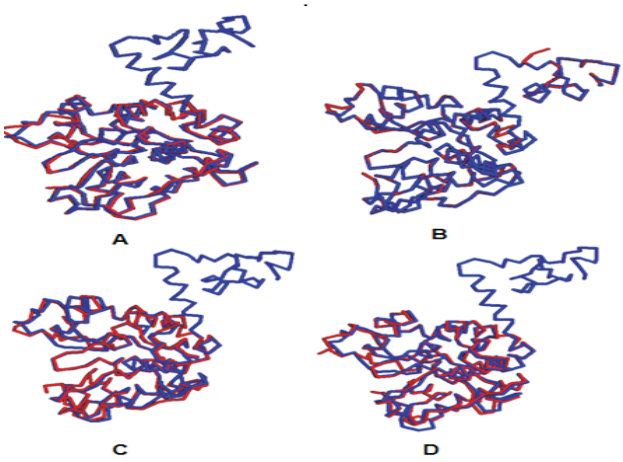

Figure 7: Structural aligment (Cα backbone) of soybean class I chitinase (red) with blue. A. barley B. rice C. jackbean D. *B.juncea*. Soybean class I chitinase is structurally similar to most of the known class I chitinase from plant.

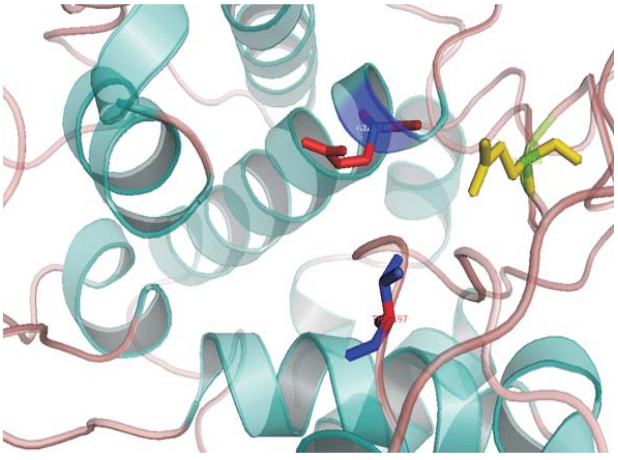

Figure 8: The proposed ligand binding site Glu146, Glu168, Thr197 shown in green is surrounded by number of aromatic residues (grey) and thus may provide the hydrophobic environment for chitin binding. The image was generated using PYMOL.

with the active sites of other glycosidases, key catalytic residues have been identified and the active site has been located in the three-dimensional structure of the soybean seed coat chitinase.

The soybean seed coat chitinase is an α-helical globular domain with approximate dimensions of 42 Å X 52 Å X 52 Å. This domain exhibits 70 and 35% amino acid sequence identities to the available crystal structures of plant chitinases and streptomyceschitinases, respectively. The α-helix rich fold of this domain was, as expected, very similar to that of rice chitinase. The RMS deviations from barley chitinase and rice and barley chitinase are 0.5 Å (279 atoms) and 0.7 Å (238 atoms), respectively, with overlapping of the Cá atoms of the helices, indicating that the spatial position and orientation of the helices are strictly conserved.

Cleft analysis and accessible and buried region

Accessibility of each residue and its relative accessibility, defined as percentage of its accessibility compared to the accessibility of that residue type in an extended Ala-X-Ala tripeptide, were computed for the subunit using NACCESS program. 33 residues were found to have relative accessibility values less than 5 % and constitute the hydrophobic core of the molecule. The Cleft analysis shows that most of the residues are easily accessible to solvent (Fig. 6). There are 10 gap regions in the protein with different degree of accessibility to the surface (Table 2). Thus protein is globular which is required for accessibility and binding of different kind of ligand.

Structural comparison with chitinase using DALI data base

Structural comparison of classI chitinase from seed coat of soybean with the DALI database of protein structures revealed number of homologous protein structures. Three of them (rice, barley, jack bean and *Hordeumvulagre*) belong to class I family (Fig. 7). When classI chitinase from seed coat of soybean and rice structures were superposed the Cα positions had r. m. s. deviation of only 0.50 Å, indicating that the two proteins possessed similar folds (Table 3).

Ligand binding template analysis

Ligand binding template results using Profunc shows a single

PDB match with 0.19Å root mean square deviation in their active site (Table 4). The residue Glu146, Glu168 and Thr197 showed a closed match with Glu67, Glu89 and Ser120 of barley chitinase. The active site of soybean class I chitinase contains polar residue for binding of chitin.The template matching is shown in Fig. 8. The residues surrounding the matched residues are aromatic and thus they may provide the ideal hydrophobic condition for binding of chitin.

In our investigation on the relevance of plant defense-related proteins, a novel enzyme having remarkable properties was purified by combining several chromatographic procedures in previous reports (Gijzen*et al*., 2001). Such an economic purification procedure combined with the easy availability of the seed makes large-scale preparation of the enzyme possible allowing a broad study of its various aspects and hence probable applications. Soybean seed coats are particularly rich in defense related proteins and peptides, although there are abundant proteins that have yet to be identified. Besides the biological role of providing defense and protection of the seed until germination occurs, seed coat tissues affect the overall quality and value of soybean food and feed products. Thus, characterization of seed coat constituents and their corresponding genes is important from a biological, nutritional, and economic standpoint in a widely grown crop species such as soybean. In addition to their role in plant defense, class I chitinasesare emerging as a distinct group of panallergens causing cross sensitization to different foods and materials in susceptible persons. Sensitization is usually limited to raw or uncooked foods, since IgE mediated recognition of the chitin binding domain is lost upon heat denaturation. Several different allergenic proteins have been identified from soybeans including those that cause food and inhalant allergies, but chitinases have not been included among these to date. The finding that a class I chitinase is an abundant component of the soluble protein fraction from seed coats indicates that this protein should be considered as a potential determinant of allergenicity to raw or uncooked soybean products. Structure prediction of chitinase from soyabean seed coat has shown that it belonged to Class I family and shared close structural and sequence similarity with rice and jackbean class I chitinase. The mechanism of chitin hydrolysis is still not clear. In conclusion, the three-dimensional structure of Class I from soyabean seed coat provides structure for another member of Class I chitinase and chitinase in general. The structure could account for the specificity and affinity of the chitinase for its ligands. The structure determined with bound ligand may be helpful in evaluating the exact specificity of the purified chitinase.

REFERENCES

Chavan, S. S. and Suryawanshi, A. P. 2014. *In vivo* evaluation of fungicides, botanicals andbioagents against *C. truncatum* causing anthracnose/pod blight of soybean. *The Bioscan*. **9(4):** 1581-1585.

Collinge, D. B., Kragh, K. M., Mikkelsen, J. D., Nielsen, K. K., Rasmussen, U. and Vad, K.1993. Plant chitinases.*The Plant J.* **3:** 31-40.

Gijzen, M., Kuflu, K., Qutob, D. and Chernys, J. T. 2001. A class I chitinase from soybean seed coat. *J. Exp. Bot.* **52:** 2283-9.

Gomez, L., Allona, I., Casado, R. and Aragoncillo, C. 2002. Seed chitinases. *Seed Sciences Research*. **12:** 217-230.

Gooday, G. W. 1999. 'Aggressive and Defensive role of Chitinase.' *(Birkhauser, Verlag: Basel/Switzerland).*

Hajduch, M., Ganapathy, A., Stein, J. W. and Thelen, J. 2005. A systematic proteomic study of seed filling in soybean: establishment of high-resolution two-dimensional reference maps, expression profiles, and an interactive proteome database. *Plant Physiol.***137:** 1397-419.

Hakala, B. E., White, C. and Recklies, A. D. 1993. Human cartilage gp-39, a major secretory product of articular chondrocytes and synovial cells, is a mammalian member of a chitinase protein family. *J. Biol. Chem.* **268:** 25803-25810.

Ingle, Y. V., Patil, C. U., Thakur, K. D. and Ingle, K. 2014. Effect of fungicides and plant resistance activator on colletotrichum leaf spot of soybean.*The Bioscan*. **9(3):** 1187-1190.

Koga, D., Mitsutomi, M., Kono, M. and Matsumiya, M. 1999. Biochemistry of chitinase. (BirkhauserVeerlag: Basel/ Switzerland).

Kramer, K.J. and Muthukrishnan, S. 2005. Chitin metabolism in insects. In: Gilbert L.I., Latrou K., and Gill S., eds., Comprehensive Molecular Insect Science. Vol. 4, Biochemistry and Molecular Biology, Chapter 3, Elsevier Press, Oxford, UK, pp. 111-144.

Laskowski, R. A., Watson, J. D. and Thornton, J. M. 2005. ProFunc: a server for predicting protein function from 3D structure. *Nucleic Acids Res.* **33:** W89-W93.

Laskowski, R. A., Hutchinson, E. G., Michie, A. D., Wallace, A. C., Jones, M. L. and Thornton, J. M. 1997. PDBsum: aweb-based database of summaries and analyses of all PDB structures. *Trends Biochem. Sci.* **22:** 488-490.

Ornum, J. V.1992. Shrimp waste-must it be wasted. *Infofish Int.* **6:** 48-52.

Patil, S. R., Ghormade, V. and Deshpande, M. V. 2006. Chitinolytic enzymes: an exploration. *Enzyme Microb Technol.* **26:** 473-483.

Sajeesh, P. K., rao, M. S. L. and Jahagirdar, S. 2014. Molecular detection, transmission and histopathological studies of seed-borne fungal infection of soybean (*Glycine max* (l.) merill). **9(1):** 247-251.

The Gene Ontology Consortium. 2000. Gene Ontology: tool for the unification of biology. *Nature Genet.* **25:** 25-29.

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

Van Loon, L. C., Rep, M. and Pieterse, C. M. J. 2006. Significance of InducibleDefense-related Proteins inInfected Plants. *Annual Review of Phytopathology*. **44:** 135-62.

Vega-Sa´nchez, M. E., Redinbaugh, M. G., Costanzo, S. and Dorrance, A. E. 2005. Spatial and temporal expression analysis of defense-related genes in soybean cultivars with different levels of partial resistance to *Phytophthorasojae*. *Physiol Mol. Plant Pathol.* **66:**175-82.

INSTRUCTION TO AUTHORS

The Bioscan

An International Quarterly Journal of Life Science

THE JOURNAL

The Bioscan is an international quarterly journal of life sciences with international editorial board. The journal is online and details can be seen (downloaded from the site. www.thebioscan.in). For any query e-mail at m_psinha@yahoo.com & dr.mp.sinha@gmail.com can be used.

AIM & SCOPE

The journal aims to publish original peerly reviewed/ refereed research papers/reviews on all aspects of life sciences.

SUBMISSION OF MANUSCRIPT

Only original research papers are considered for publication. The authors may be asked to declare that the manuscript has not been submitted to any other journal for consideration at the same time. Two hard copies of manuscript and one soft copy, complete in all respects should be submitted. The soft copy can also be sent by e-mail as an attachment file for quick processing of the paper.

FORMAT OF MANUSCRIPT

All manuscripts must be written in English and should be typed double-spaced with wide margins on all sides of good quality A4 paper.

First page of the paper should be headed with the title page, (in capital, font size 16), the names of the authors (in capitals, font size 12) and full address of the institution where the work was carried out including e-mail address. A short running title should be given at the end of the title page and 3-5 key words or phrases for indexing.

The main portion of the paper should be divided into Abstract, Introduction, Materials and Methods, Results, Discussion (or result and discussion together), Acknowledgements (if any) References and legends.

Abstract should be limited to 200 words and convey the main points of the paper-outline, results and conclusion or the significance of the results.

Introduction should give the reasons for doing the work. Detailed review of the literature is not necessary. The introduction should preferably conclude with a final paragraph stating concisely and clearly the aims and objectives of your investigation.

Materials and Methods should include a brief technical description of the methodology adopted while a detailed description is required if the methods are new.

Results should contain observations on experiment done illustrated by tables and figures. Use well known statistical tests in preference to obscure ones.

Discussion must not recapitulate results but should relate the author's experiments to other work on the subject and give their conclusions.

All tables and figures must be cited sequentially in the text. Figures should be abbreviated to Fig., except in the beginning of a sentence when the word Figure should be written out in full.

The figures should be drawn on a good quality tracing/ white paper with black ink with the legends provided on a separate sheet. Photographs should be black and white on a glossy sheet with sufficient contrast.

References should be kept to a minimum and listed in alphabetical order. Personal communication and unpublished data should not be included in the reference list. Unpublished papers accepted for publication may be included in the list by designating the journal followed by "in press" in parentheses in the reference list. The list of reference at the end of the text should be in the following format.

1. **Witkamp, M. and Olson, J. S. 1963.** Breakdown of confined and non-confined Oak Litter.*Oikos*. **14**:138-147.
2. **Odum, E.P. 1971.** *Fundamentals of Ecology*. W. B. Sauder Co. Publ. Philadelphia.p.28.
3. **Macfadyen, A.1963.** The contribution of microfauna to total soil metabolism. In:*Soil organism*, J. Doeksen and J. Van Der Drift (Eds). North Holland Publ. Comp., pp 3-16.

References in the text should be quoted by the **author's name and year** in parenthesis and presented in year order. When there are more than two authors the reference should be quoted as: first author followed by *et al.,* throughout the text. Where more than one paper with the same senior author has appeared in on year the references should

Cont. ..... P. 1956