Expression and *in silico* characterization of Phenylalanine ammonium lyase against karnal bunt (*Tilletia indica*) in wheat (*Triticum aestivum*)

Shalini Purwar[1]*, Shanthy Sundaram[1], Sukrat Sinha[1], Ankit Gupta[1], Neha Dobriyall[1] & Anil Kumar[2]

[1]Center for Biotechnology, Nehru Science Center, University of Allahabad, Allahabad, India, [2]Department of Molecular Biology & Genetic Engineering, GBPUA & T, Pantnagar, India; Shalini Purwar - Email: purwarshalini@gmail.com; Phone: 09473850596; *Corresponding author

Abstract:
To investigate the lignifications process and its physiological significance under Karnal Bunt (KB), the changes in enzymes responsible for lignifications likes, phenylalanine ammonia lyase (PAL), were determined in resistant (HD-29) and susceptible genotype (WH-542) of wheat during different developmental stages. The PAL gene was cloned and sequenced. The expression of PAL gene was measured by means of semi-quantitative RT-PCR. The enzyme was expressed constitutively in both the susceptible and resistant genotype. However, the activity was higher in all the developmental stages of resistant genotype, indicating that this genotype has a significant higher basal level of these enzymes as compared to the susceptible line and could be used as marker(s) to define KB resistance. The activity of PAL was significantly higher in WSv stage (Z=16). Structural comparisons based on alignments of all the protein sequences using the clustal W program and searches for conserved motifs using the MEME program have revealed broad conservation of main motifs characteristic of the plant PAL. MSA and phylogenetic analyses of different plants PAL demonstrate that all PAL cluster divided in to two main cluster. The PAL also possesses a specific consensus sequences [GS]-[STG]-[LIVM]-[STG]-[SAC]-S-G-[DH]-L-x-[PN]-L-[SA]-x(2,3)-[SAGVTL]. The pathway might be associated with the enhancement of structural defense barrier due to lignifications of cell wall as evident from the enhanced synthesis of lignin in all the stages of resistant genotype. Our results clearly indicate the possible role of enzymes of Phenyl propanoid pathway metabolism provides genotype and stage dependant structural barrier resistance in wheat against KB.

Key words: Phenyl propanoid pathway, lignifications, RT PCR, KB, PAL

Background:
The lignifications (a phenolic compound lignin), is a well known plant response to fungal attack. Lignin deposition is correlated with an increased activity enzyme from phenylpropanoid pathway. Phenylalanine ammonia-lyase (PAL; EC 4.3.1.5) catalyzes the initial step in the biosynthesis of phenylpropanoids [1]. PAL enzyme a member of the ammonia lyase family that catalyzes a reaction converting L phenylalanine to ammonia and trans-cinnamic acid. It has a molecular mass in the range of 270-330 kDa.

Phenylpropanoids are essential for growth, development, and survival of vascular plants [2]. The activity of PAL is induced dramatically in response to various stimuli such as tissue wounding, pathogenic attack, light, low temperatures, and hormones [3]. PAL is the first and committed step in the phenyl propanoid pathway and is therefore involved in the biosynthesis of the polyphenol compounds (such as flavonoids, phenylpropanoids, and lignin in plants) through general phenylpropanoid metabolism involves a sequence of enzyme-regulated reactions [4].

Karnal bunt (KB) caused by *Tilletia indica* (Syn. *Neovossia indica*), a semi-biotroph fungus. It is a seed borne disease which typically causes partial conversion of individual kernels into sori filled with fetid teliospores thus affecting yield and quality [5]. The pathogen affects wheat spikes and proliferates within seeds, leading to the impairment of grain quality, seed vigor and reduction in market acceptability of the product. Since the pathogen disseminates through seeds, it also necessitates quarantine regulations to restrict the spread of the disease to unaffected areas. Elucidation of the Molecular basis of resistance and stage specific immunity against KB is needed to be examined for analyzing the role of enzymes of phenylpropanoid metabolism which provides structural barriers in conferring the resistance against KB in wheat. Fungi, which are serious plant pathogens, are known to cause serious yield losses in economically important agricultural crops. Recombinant DNA based approach could be utilized to develop fungal resistant crop varieties, provided that the cellular and molecular basis of fungal phyto-pathogenesis is delineated. In the present study attempts were made to elucidate possible role of PAL in impart stage specific and differential immunity as well as their role in providing structural barrier against KB pathogen has been demonstrated by comparing the basal levels of enzyme activity that are involved in tissue lignifications in developing wheat spikes and *In silico* characterization of PAL protein.
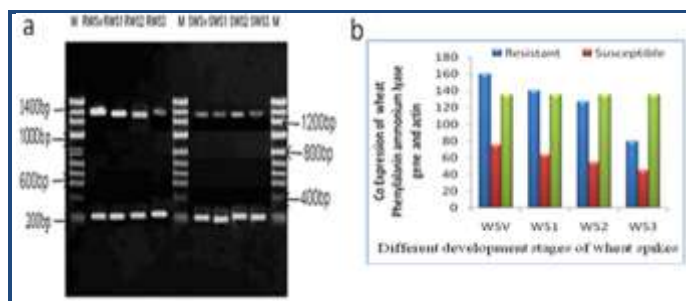

Figure 1: Semi quantitative & differential expression analysis of multi-domain wheat PAL at different stages (Sv, S1, S2 and S3) of developing wheat spikes in Resistant and Susceptible genotypes. **(a)** RT PCR based amplification followed by **(b)** densitometry analysis of wheat PAL transcripts in Resistant and Susceptible genotype.

Methodology:
Plant Material
In the present study two parent genotypes of bread wheat (*Triticum aestivum)*, one highly resistant HD-29 (designated designated as PR) and another highly susceptible WH-542 (designated as PS) based on their pathogenicity testing under field conditions against KB, were used in the present investigation. The seeds of these genotypes were collected from the Crop Research Centre, G.B. Pant University of Agriculture and Technology, Pantnagar (UK) [6].

Isolation of total RNA
Total RNAs were extracted from the different stages of the wheat spikes of resistant and susceptible genotypes using Plant Easy RNA isolation kit and RNA preparations were purified using the same kit according to the manufacturer's protocol (QIAGEN, Germany). Equal quantities of total RNA isolated from each of three plants randomly selected from each

treatment were bulked and subjected for RT PCR analysis. Total RNA (10 μg) was separated on 1.2% agarose gel containing formaldehyde for assessing the quality of isolated RNA [6].

Designing of primers
The different sets of primers were designed (by using BI tool DNASTAR) for PCR amplification of PAL gene by RT-PCR based on available gene sequences in NCBI database i.e., JQ005112 and AY423548 for co-amplification of wheat Actin gene. The details of forward and reverse primers, their Tm and %GC contents are given in **Table 1 (see supplementary material).**

Molecular cloning and Expression analysis of Phenylalanine ammonia lyase gene
For RT-PCR, QIAGEN One Step RT-PCR kit was used. The QIAGEN One Step RT-PCR Enzyme mix contains a specifically formulated enzyme blend for both reverse transcription and PCR amplification. PCR amplification of wheat PAL genes was done using the specific primers designed as above. The PCR amplicons were ligated in pGEMT-EASY cloning vector (Promega). Nucleotide sequence of independent clones were determined with the dye terminator kit (ABI Prism, Perkin Elmer, NJ) and analyzed on Applied Biosystems 370. RT-PCR was carried out to check the expression of wheat PAL mRNA transcripts. Actin gene was selected as endogenous internal standard [6]. Thermal cycler was programmed as follows reverse transcription at 50^0C for 30 min and subsequent 35 cycles of 94^0C denaturation for 30 s 55^0C annealing for 30 s and 2 min extension at 72^0C. Densitometry analyses of expressed genes (PAL and actin) were done with the help of Gene Profiler software, Alpha Innotech Corporation USA.

In silico sequence analysis
The Genomic sequence was translated to protein sequence using translation tool (http://ca.expasy.org/tools/dna.html). The homology search of the PAL protein was done through BLAST search tool of NCBI (http://www.ncbi.nlmnih.gov). Sequences from putative protein (Acc. No. *Triticum aestivum* (AEX59143), and 24 different plants **(Table 1)** were aligned using Clustal W (www.ebi.ac.uk/Tools/msa/clustalw2), and phylogenetic tree was constructed using UPGMA method by MEGA version 4.0.02 (www.megasoftware.net) [7]. Domain and family analysis of amino acid sequence was done using CDD tool of NCBI (http://www.ncbi.nlm.nih.gov/cdd) and MEME (meme.nbcr.net) [8]. Protein variability cheek by Protein Variability Server http://imed.med.ucm.es/PVS.

Results & Discussion:
Molecular cloning and Expression of analysis of Phenylalanine ammonia lyase gene
Total RNA was isolated from immature spike lets of wheat genotypes HD-29 and WH- 542, PAL were amplified using the same PCR conditions in both resistant and susceptible genotype. The primer sets designed from highly conserved domains were efficient in reproducible amplification of expected sizes 1.5 kb of amplicons. Amplification of PAL gene gave product size of 1500 bp **(Figure 1).** The sequence was submitted to NCBI databank and is available online at NCBI website. Stage dependent co-amplification of the PAL and actin mRNA transcripts expression was studied in both the genotypes **(Figure 1a)**. Semi-quantitative RT-PCR, and

BIOINFORMATION

Densitometry (relative values) clearly showed PAL mRNA transcripts were significantly higher (P\0.001) in resistant genotype compared to susceptible **(Figure 1a–b)**. In HD29, PAL gene expression utmost at Sv stage, its expression gradually drop off upto S3 stage (slight grain formation) in resistant genotype. The PAL gene was maximally expressed at Sv-stage (P\0.001) of wheat spike (i.e., vegetative leaf stages) accumulates high level of lignin and phytoalexin and then degree of expression was declined gradually. In susceptible genotype PAL mRNAs express in all the stages although the intensity of each amplicon was much lower than that of resistance genotype **(Figure 1a-b).** However, these results indicate that higher expression of PAL seemed to impart resistance against KB. However, this work provides possible evidences in favor of the role of PAL in defining the KB resistance. The expression of PAL occurring from Sv to S2 stages a highly prone stage to KB infection, and the intensity of the detected bands in RT PCR, showed that PAL seemed to be expressed more in resistant than in susceptible genotype. Interestingly, expression and localization of *pal mRNA* was observed in Sv tissue, suggesting their involvement in flavonoid biosynthesis as well (Mahroug *et al.*, 2005). Increased PAL activity has been observed in response to *E. graminis* attack in wheat, barley and oat **[9]**. In barley, induction of both PAL and peroxidase transcripts correlates temporally with *E. graminis* attack **[9]**. Karnal bunt (KB) a seed borne disease in which proliferation of mycelia of *T. indica* in endosperm of wheat grains completely converted into sori filled with teliospores through necrotic cell death. In Poplar, one PAL gene product was associated with formation of condensed tannins while another was associated with lignin production **[10]**. Spatio-temporal organization of secondary metabolites in aerial organs of *Cataranthus roseus* revealed that the first enzymes of the phenylpropanoid pathway (PAL) are preferentially localized in lignifying tissues, illustrating their involvement in the lignin biosynthesis pathway **[11]**. In the family Gramineae, active lignifications appears to be of special importance in induced resistance mechanisms and Inhibition of Lignification Breaks Hypersensitive Resistance of Wheat to Stem Rust **[12].**

In silico Analysis of PAL Gene

We performed comparative studies of PAL gene of *Triticum aestivum* (AEX59143.1) with twenty four plant species **(Table 1)** that represent two major classes. The selection of sequences was random belongs to very diverse families of dicotyledonous like fruits (*Vitis vinifera*), vegetable (*Daucus carota, Lactuca sativa, Solanum tuberosum*), floweres (*Hibiscus cannabinus, Solenostemon scutellarioides),* trees *(Eucalyptus robusta)*, from Monocotyledons**,** cereals (*Triticum aestivum, Triticum urartu, Sorghum bicolor, Zea mays***,** *Oryza sativa, Hordeum vulgar, Aegilops tauschii*), trees (*Bambusa oldhamii, Musa acuminate, Phyllostachys edulis, Pleioblastus maculosoides, Phyllostachy parvifolia*) flower (*Lilium*) and model dicotyledon and monocotyledon species (*Arabidopsis thaliana* and *Brachypodium distachyon*) **(Table 1)**.

Multiple sequence alignment and Evolutionary relationships of taxa

In silico studies revealed that all the resistant and susceptible PAL gene clones translated to protein sequences by using ExPASy Software. The partial cDNA clone consists of 1,514 nucleotides, and encoded protein 504 amino acid with molecular mass approximate 15 KB. Multiple alignment of

amino acid analysis showed, 83.34% homology with in monocotyledon and 70% homology between monocotyledon and dicotyledon PAL mRNA and proteins. Comparison of PAL amino acid sequence among the monocotyledon plants, highest homology with Poaceae plant then other monocotyledon PAL proteins. *T. aestivum* PAL (AAG02280.1), followed by 74.96 % identity to a *Triticun urartu* (EMS63197.1), 67.64%, 63.70 % , 62.94%, 62.08%, 61.76%, 62.22%, 62.08% *to Brachypodium distachyon , Zea mays , Hordeum vulgar, Aegilops tauschii, Sorghaum bicolor, Pleioblastus maculosoides, Phyllostachy parvifolia*, respectively. 56. 74% homology with *Musa acuminate*. Phylogenetic analysis was performed to evaluate the evolutionary relationships among the 24 PAL sequences from 24 taxa selected for this analysis. Trees were estimated from the multiple sequence alignment using clustal W (www.ebi.ac.uk/Tools/msa/clustalw2) **(Figure 2).**

One dicotyledon specific clade was placed just above the out-group branches in the PAL gene tree. A clade with the remaining genes split into another monocotyledon species. A second clade containing only monocotyledon PAL genes divided into two subclade, one subclade 2a strictly contain Poaceae while in subclade2b along with Poaceae remaining gene split to another monocolyledon familes like *Musaceae and Liliaceae* **(Figure 2)**. This result suggests that PAL genes were transfer from dicotyledonous plants to monocotyledonous. The high bootstrap values and posterior probability evidence provided strong support for the organization of the angiosperms genes into these three distinct clusters. Physical position of PAL gene on chromosome also varied in monocot and dicotelyledon like *vitis vinifere* the PAL gene position on chromosomes No. 16 (NC_012022.3), in *Sorgham bicolar* present on chromosome no. 4 (NC_012873), in *Brachypodium distachyon* on chromosome 4. Because complete genome sequences are not yet available for wheat. It was not clear the position of *Triticum* PAL genes on chromosomes.
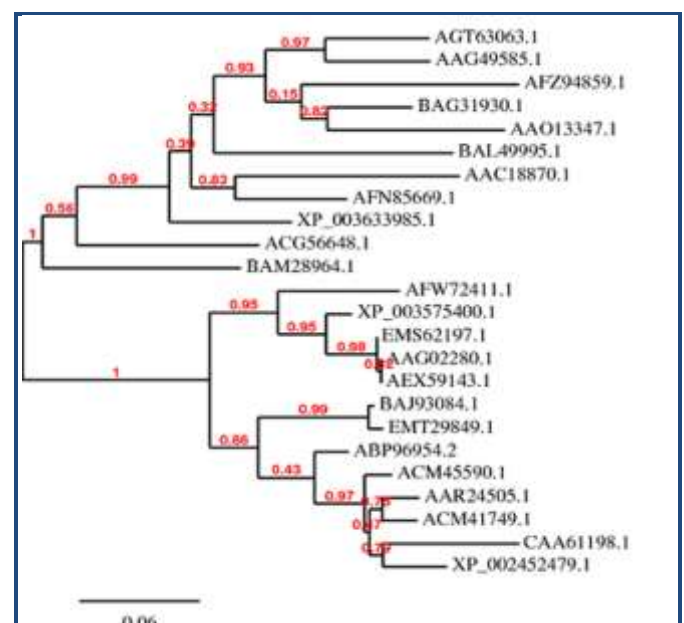

Figure 2: A Dendrogram analysis for establishing the similarity index between different PAL protein sequences available in NCBI databases by UPGMA (MEGA 4. 1 version) with *Triticum aestivum* PAL (AEX59143), were as follows: *Triticum aestivum*

(AAG02280.1), *Triticum urartu* (EMS63197.1), *Brachypodium distachyon* (XP_003575400), *Zea mays* (AFW72411.1), *Aegilops tauschii* (EMT29849.1), *Hordeum vulgar* (BAJ93084.1), *Phyllostachys edulis* (ABP96954.2), *Sorghaum bicolor* (XP_002452479.1), *Pleioblastus maculosoides* (ACM45590.1), *Phyllostachy parvifolia* (ACM41749.1), *Bambusa oldhamii* (AAR24505.1), *Oryza sativa* (CAA61198.1), *Lilium* (BAM28964.1), *Musa acuminate* (ACG56648.1), *Arabidopsis thaliana* (AAC18870.1), *Lactuca sativa* (AAO13347.1), *Daucus carota* (BAG31930.1), *Vitis vinifera* (XP_003633985.1), *Ipomoea* (AAG49585.1), *Hibiscus cannabinus* (AFN85669.1), *Solenostemon scutellarioides* (AFZ94859.1), *Eucalyptus robusta* (BAL49995.1), *Solanum tuberosum* (AGT63063.1).
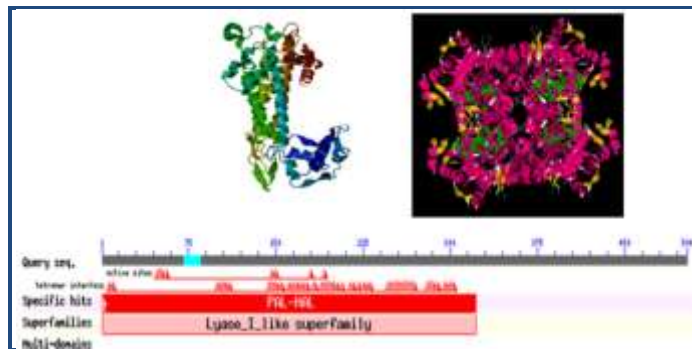

Figure 3: Functional domain analysis of representative member of PAL.
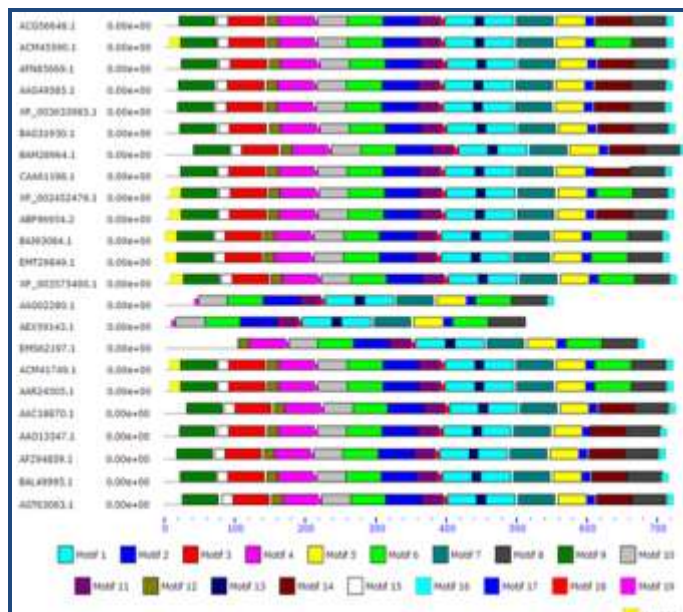

Figure 4: Schematic diagram of motif distribution of PAL genes.MEME4.6.0 was applied to show that different sub groups were distinguished by the motif distribution.

Conserved Domains and Protein Classification

The CDD search [13] with the deduced protein sequence of PAL identifies is multidomain protein. It belongs to a member of the Lyase class I_like superfamily of enzymes that catalyze beta-elimination reactions. Unique identifier for domain model PSSM ID of PAL is 176460.3 D structure of PAL (PSSM View cd00332) showes four active sites of the homotetrameric enzyme are each formed by residues from three different subunits and are active as homotetramers [13]. *In silico* analysis

of amino acid sequences employing Pfam tool [14] revealed existence of one functional domains PDOC00424 (**Figure 3**). Conserved Signature sequence of PAL shows [GS]- [STG]- [LIVM]- [STG]- [SAC]-S-G- [DH] -L-x-[PN] -L-[SA] -x(2,3) - [SAGVTL] (PROSITE PS00488). Multiple Expectation-Maximization for Motif Elicitation (MEME) [8] is a suite of tools for motif discovery and searching. Twenty different motifs (subdomain) between 6 and 50 residues were detected and distributed by MEME software **(Figure 4)**. Most of the group and subgroup sharing same motifs. Many motifs are almost conserved and found in every group and subgroup. Motif possesses well conserved signature sequences. These conserved motifs could be the essential elements determining the PAL family's common molecular function among different plant species. Motif 14 absent all the monocotyledon plants. Motifs 3, 9, 15, 18 absent only in *Triticum species like T. aestivum and T. ururtus*. Motif 12 absent in *Triticum aestivum*. Motif 6 occurs twice in monocotyledon while only once found in dicotyledon plants. Duplication or absence of Motifs are either, substitution, accumulation of mutation or subjected to rearrangements. It is not necessary that changes their activity, because that do not have a direct impact on the active site contain altered residues. Duplication events have been an important theme in the evolution of the PAL gene family. At least five distinct duplication events can be identified in the PAL gene tree, with the oldest event following the divergence of *Physcomitrella*. Duplication events in the ancestral lineage, as well at the tip of the gymnosperm branch, suggest potential sources of functional variability [10].
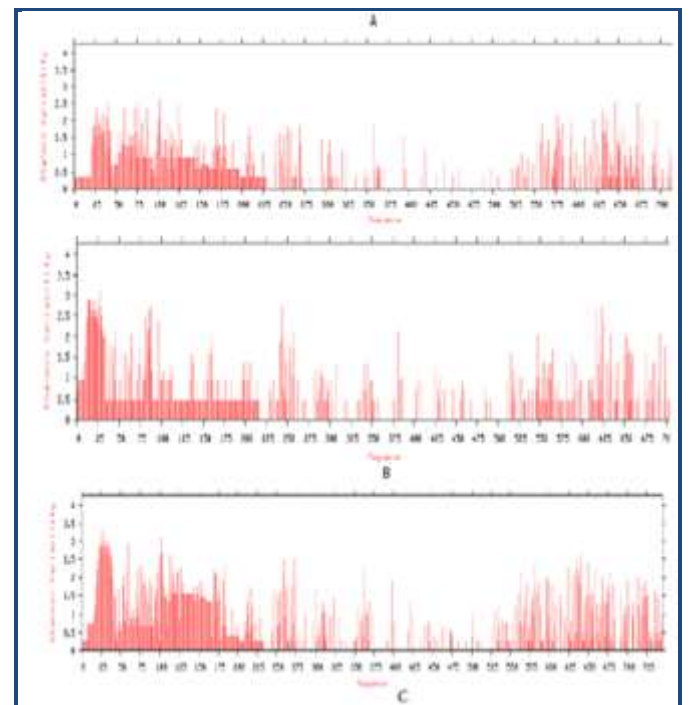

Figure 5: Shinon Protein Variability of **a)** with in monocotyledonous plants **b)** with in Dicotyledonous plants **c)** both monocotyledon and Dicotyledonous plant.

Protein variability analysis of PAL

Prediction of putative protein variability sites was performed using PVR server [15]. The analysis protein variability sites location was compared between cloned *Triticum aestivum* PAL

with other monocotyledon and dicotyledon PAL protein database individually by calculating Shannon entropy (H) **[16]**. In monocotyledon plants shows low variability level (H>2) when compared to clones *Triticum aestivum* **(Figure 5a)**. In monocotyledons plants sequence N terminal and intermediate sequences shows highly conserved sequence H<1, some intermediate sequence shows conserve sequence H<2. Protein variability is very high towards the N terminal in dicotyledon plants as the Shannon entropy H >2.0 compared to 'C' terminal. Sequence from amino acids 30 to 200 shows highly conserver as H<1 while intermediate sequence shows H< 2 that is conserved **(Figure 5b)**. Conserved regions with in monocotyledons and dicotyledons PAL proteins regions between 220-244, 275-299, 325 - 354, 370-397, 399-420, 422-530, and 535-556 **Table 2 (see supplementary material)**. Shannon entropy result shows variability was found to be more at dicotyledons protein compared to monocotyledons plants. It indicates that during evolution rate of mutation is very high compared to monocotyledon plant.

Conclusions:

Phenylalanine ammonia lyase, is a primary control point for the phenylpropanoid pathway which belongs to the lyase class I super-family of enzymes. This study is the most extensive stage specific molecular expression and in silico phylogenomic studies so far for the PAL gene, particularly with respect to karnal bunt resistant (HD 29) and susceptible (WH 542) genotype of *Triticum aestivum*. The partial sequence of cloned PAL genes available in the GenBank. Differential expression patterns of RT PCR suggest that the PAL gene products may be responsible for providing biosynthetic precursors to different phenylpropanoid branch pathways under different developmental conditions or in response to various external stimuli. The PAL transcript was particularly abundant in developing spikes. The phylogenetic tree constructed using PAL gene sequences from 24 species including dicotyledonous and monocotyledon shows a 70 % homology between them, which may suggest different functional regulation. The lignin content of higher plants has long been recognized as an important factor in the resistance response against potential pathogens. However, induced lignifications have been proposed as an important mechanism of disease resistance of wheat against a variety of fungal pathogens. Several lines of evidence point to a role for cellular lignification in the hypersensitive resistance response of wheat to the karnal bunt fungus, *Tilletia indica*.

References:

[1] Ritter H & Schulz GE, *Plant Cell* 2004 **16**: 3426 [PMID: 15548745]
[2] Junli Huang *et al. Plant Physiol.* 2010 **153:** 1526 [PMID: 20566705]
[3] Marcos JF *et al. J Exp Bot*. 2005 **56:** 2183 [PMID: 15983011]
[4] Adriano R *et al. Genetics and Molecular Biology* 2007 **30:** 819
[5] Sekhon KS *et al. Bull Grain Technol*. 1980 **18:** 208
[6] Purwar S *et al. Mol Biol Rep*. 2010 **37**: 1377 [PMID: 19347606]
[7] Tamura K *et al. Mol Biol Evol.* 2007 **24:** 1596 [PMID: 17488738]
[8] Timothy LB, *Current Protocols in Bioinformatics.* 2002 DOI: 10.1002/0471250953.bi0204s00
[9] Zhang L *et al. Physiological and Molecular Plant Pathology* 1997 **51**: 15
[10] Bagal UR *et al. BMC Genomics* 2012 **13**: S1 [PMID: 22759610]
[11] Mahroug S *et al. Planta* 2006 **223**: 1191 [PMID: 16322983]
[12] Moerschbacher BM *et al. Plant Physiol*. 1990 **93:** 465 [PMID: 16667489]
[13] Marchler-Bauer *et al. Nucleic Acids Res*. **39:** 225 [PMID: 21109532]
[14] Finn RD *et al. Nucleic Acids Res.* 2008 **36:** D281 [PMID: 18039703]
[15] Garcia-Boronat M *et al. Nucleic Acids Res.* 2008 **36:** W35 [PMID: 18442995]
[16] Shannon CE, *The Bell system Technical Journal* 1948 **27**: 379

BIOINFORMATION

Supplementary material:

Table 1: Selected Monocotyledonous and Dicotyledonous plant with Accession Number

S. No.	Plants Name	Accession No.
1.	Triticum aestivum	AAG02280.1
2.	Triticum aestivum	AEX59143.1
3.	Triticum urartu	EMS63197.1
4.	Brachypodium distachyon	XP_003575400
5.	Zea mays	AFW72411.1
6.	Aegilops tauschii	EMT29849.1
7.	Hordeum vulgar	BAJ93084.1
8.	Phyllostachys edulis	ABP96954.2
9.	Sorghaum bicolor	XP_002452479.1
10.	Pleioblastus maculosoides	ACM45590.1
11.	Phyllostachy parvifolia	ACM41749.1
12.	Bambusa oldhamii	AAR24505.1
13.	Oryza sativa	CAA61198.1
14.	Lilium	BAM28964.1
15.	Musa acuminate	ACG56648.1
16.	Arabidopsis thaliana	AAC18870.1
17.	Lactuca sativa	AAO13347.1
18.	Daucus carota	BAG31930.1
19.	Vitis vinifera	XP_003633985.1
20.	Ipomoea	AAG49585.1
21.	Hibiscus cannabinus	AFN85669.1
22.	Solenostemon scutellarioides	AFZ94859.1
23.	Eucalyptus robusta	BAL49995.1
24.	Solanum tuberosum	AGT63063.1

Table 2: Conserved novel motifs regions analysis of PAL proteins.

N	Start	End	Sequence
1	220	244	--------------DLVPLSYIAGL
2	275	299	GFFELQPKEGLAMVNGTAVGSGLAS
3	325	354	GKPEFTDHLTHKLKHHPGQIEAAAIMEHIL
4	370	397	DPLMKPKQDRYALRTSPQWLGPQIEVIR
5	399	420	ATKSIEREINSVNDNPLIDVSR
6	422	530	KAIHGGNFQGTPIGVSMDNTRLAIAAIGKLMFAQFSELVNDFYNNGLPSNLSGGRNPSLDYG FKGAEIAMASYCSELQFLGNPVTNHVSAEQHNQDVNSLGLISSRKT
7	535	556	DILKIMSSTFLVALCQAIDLRH