Indian J. Genet., 77(3): 371-378 (2017)
DOI: 10.5958/0975-6906.2017.00050.5

Molecular characterization of inositol pentakisphosphate 2-kinase (*GmIPk1*) from soybean and its expression pattern in the developing seeds

Nabaneeta Basak[1], Veda Krishnan, Vanita Pandey[2], Mansi Punjabi, Alkesh Hada, Ashish Marathe, Monica Jolly and Archana Sachdev*

Division of Biochemistry, ICAR-Indian Agricultural Research Institute, New Delhi 110 012; [2]Quality and Basic Sciences Division, ICAR-Indian Institute of Wheat and Barley Research, Karnal 132 001

(Received: November 2016; Revised: July 2017; Accepted: July 2017)

Abstract

The principal storage compound of seeds, phytic acid, poses several problems for agriculture due to its antinutritional effects making the development of low phytate crops a prime priority in research. Soybean, with its high nutritive value, wide acceptability and varied usage, but possessing high levels of phytate can be a major target of low phytate research. To this effect, Inositol penta*kis*phosphate 2-kinase (IPK1) can prove to be critical player given its key role in catalyzing the terminal step of phytate biosynthesis. In this paper, we report the expression profile of *GmIPK1* across different tissues and seed development stages, cloning of 1371 bp nucleotide sequence representing the full length cDNA of *GmIPK1* and *in silico* analysis for characterization. Amino acid sequence of *GmIPK1* showed much similarity with that of *Phaseolus vulgaris* and *Cicer arietinum*. It also showed the presence of the characteristic Ins_P5_2-kinase domain required for catalytic activity. Thus this study provides an initial idea on the molecular characterization of *Gm*IPK1 enzyme for its further in detailed study and subsequent utilization in generating low phyate soybean with enhanced mineral bioavailability.

Key words: *GmIPK1*, low phytate, expression profiling, cloning, molecular modeling

Introduction

One of the most important legume crops in the world, soybean is well known for its high content of good quality seed protein for food and feed purposes, and also for its capability for providing oil needed for industry (Harada and Xia 2004). Soybean seeds also contain sufficient amount of phosphorus required for human health and optimal livestock production, in the form of inositol hexaphosphate otherwise known as phytate (Raboy 2007). In addition to sequestering inorganic phosphate, phytate may also chelate divalent cations thereby decreasing their availability (Raboy 2009) besides contributing to water pollution by eutrophication as the phosphate-rich waste discharges into water bodies. Hence a reduction in phytic acid content of grains could have far reaching nutritional and environmental consequences.

Manipulation of the key enzymes of phytic acid biosynthesis pathway can be a probable approach for generating low phytate soybean. Careful examination of the phytic acid biosynthetic pathway reveals the terminal enzyme Inositol penta*kis*phosphate 2-kinase (IPK1) to be one the most important enzyme controlling the metabolic flux of phytate generation. Thus engineering of this enzyme may help to achieve seed specific reduction in levels of phytate synthesis while at the same time avoiding depletion of intermediates of the pathway for other metabolic activities. This enzyme has hitherto been studied in other organisms including *Arabidopsis* (Stevenson-Paulik et al. 2005; Sweetman et al. 2006), maize (Sun et al. 2007), yeast (York et al. 1999), human (Verbsky et al. 2002), and soybean (Stiles 2007). Evidence suggesting the potential of downregulation of the enzyme has been well documented. The *Saccharomyces cerevisiae* IPK1Δ mutant was almost incapable of producing $InsP_6$ (York et al. 1999) while the *Arabidopsis* T-DNA

*Corresponding author's e-mail: arcs_bio@yahoo.com
[1]Present address : Crop Physiology and Biochemistry Division, ICAR-National Rice Research Institute, Cuttack 753 006, Odisha
Published by the Indian Society of Genetics & Plant Breeding, F2, First Floor, NASC Complex, PB#11312, IARI, New Delhi 110 012
Online management by indianjournals.com; http://epubs.icar.org.in/journal/index.php/IJGPB

insertion mutant showed a reduction of 83% in seed phytate levels (Stevenson-Paulik et al. 2005). Recently, RNAi mechanism has also been successively adopted to silence the *IPK1* gene in rice (Ali et al. 2013) to obtain a significant decrease of 69% in phytate content. IPK1 in soybean has also been implicated to have a critical role in phytate biosynthesis (Stiles et al. 2008), nevertheless characterization and information about the fundamental features of this enzyme is still lacking. In this paper, we report the spatiotemporal expression analysis, cloning of *GmIPK1* from soybean variety Pusa-9712 and its characterizations for unravelling the structural complexity with an ultimate aim to specifically down regulate the enzyme for generation of low phytate soybean.

Materials and methods

Differential gene expression by Quantitative Real-Time PCR

The steady state expression levels of *GmIPK1* transcripts were estimated in various tissues *viz*., roots, stems and leaves of 20-day-old seedlings, flowers and developing seeds of soybean, *Glycine max* L. var. Pusa 9712, collected from the Division of Genetics, IARI, New Delhi. As IPK1 is the terminal enzyme in phytic acid biosynthesis, for spatial expression studies seeds in later stages of development (25-30 DAF) were taken whereas to determine the expression pattern of the enzyme during seed development as would be expected for phytic acid biosynthesis, temporal expression studies were carried out at increasing seed sizes representing the stages of development (0-2, 2-4, 4-6, 6-8, 8-10 and 10-12 mm seed size based on length as measured from apical end to basal end of the seed). Total RNA was isolated from the samples using the TRIZOL reagent (Invitrogen, USA) and treated with RNase-free DNase I (Thermo Scientific, USA) as per manufacturer's instructions. All RNA samples were quality assessed on agarose gels and further quantified with Nanodrop 2000 (Thermo Scientific, USA) to achieve uniform concentration for cDNA synthesis. 1 µg of total RNA was reverse transcribed for cDNA synthesis using the RevertAid^{TM} H Minus First strand cDNA Synthesis Kit (Thermo Scientific, USA) according to manufacturer's protocol. Primers were designed from the exonic regions of previously reported complete genomic sequences of *IPK1*, GenBank (accession number ACUP01008083.1) to amplify a fragment of approximally 200 bp from the conserved region of the gene of interest (q*IPK1*F 5'-ACG CGT CGA CTT TTG ATC TTG TTC CTG TG-3' and *qIPK1*R 5'- CCA TCG ATG GTA AAA GAA GGT GAG GAT CCA GC- 3'). The housekeeping gene PEP carboxylase (q*PEPCo*F 5'- CAT GCA CCA AAG GGT GTT TT- 3' and q*PEPCo*R 5'- TTT TGC GGC AGC TAT CTC TC- 3') was used as an internal control for the normalization of cDNA. Real time expression analysis of the reference and target gene was performed in PikoReal 96 Real Time PCR System (Thermo Scientific, USA) using 1µl of template. The experiment was conducted according to standard protocol using the DyNAmo Flash SYBER Green qPCR Kit (Thermo Scientific, USA). All three biological samples were run in triplicates under the following conditions: 1 cycle of 3 min at 94^{o}C followed by 35 cycles of 95^{o}C for 30 s and 30 s at 60^{o}C with a final extension time of 30 s at 72^{o}C. The PCR products were also subjected to melt-curve analysis. A standard fluorescence threshold was set to a ΔRn of 0.5 on the log fluorescence scale to determine the fractional cycle number (Ct value). Relative expression of target genes was calculated using the efficiency calibrated 2^[–$\Delta\Delta$C(T)] method (Livak and Schmittgen, 2001). In case of spatial expression analysis root tissue was considered as the calibrator, whereas for temporal analysis, expression in the representative 0-2 mm sized seeds was taken as calibrator. The resulting PCR products were also analyzed by agarose gel electrophoresis to determine the specificity.

Cloning of GmIPK1 from Glycine max Pusa 9712 cultivar

PCR reactions were performed using the cDNA synthesized as previously described, with *GmIPK1*F [5'-ATG GCA TTG ACT TTG AAA GAG GAA GAC GC-3'] and *GmIPK1*R [5'- TCA ATA TGC AGC AT A GAT -3'] primers in a final volume of 25 µl containing 20 ng of cDNA, 2.5 µl of 10 X Taq buffer, 2 mM $MgCl_2$, 1.0 µl of 10 mM dNTP mix, 200 nM of each primer and 0.5 U of Taq DNA polymerase (Thermo Scientific, USA) to amplify the *GmIPK1* gene sequence. The 1371 bp amplicon was purified and subsequently cloned in the pGEM®- T Easy vector (Promega, USA). The recombinant plasmids were confirmed by restriction analysis and sequenced using universal primers (SP6 and T7) on an automated sequencer (ABI 3730xl DNA Analyzer, USA). The nucleotide sequence data was submitted to the GenBank (Accession number KF359958.1).

In silico characterization of GmIPK1

The related *GmIPK1* amino acid sequences from other plants were identified via Basic Local Alignment Search Tool (BLASTP) implemented on databases of the National Center for Biotechnology Information (http://blast.ncbi.nlm.nih.gov/blastp) using the previously deduced *GmIPK1* amino acid sequence as a query.

The amino acid compositions of the selected IPK1 sequences were computed using PEPSTATS (http://emboss.bioinformatics.nl/cgi-bin/emboss/pepstats/) analysis tool (Rice et al. 2000). The ProtParam tool (http://web.expasy.org/protparam/) of ExPASy was used to compute the physiochemical characterization of the selected 10 IPK1 sequences (Gasteiger et al. 2005).

The amino acid sequences were fed into Clustal Omega version 2.0 for multiple alignments (http://www.ebi.ac.uk/Tools/msa/clustalomega2/) (Goujon et al. 2010). The phylogenetic analysis were carried out using the neighbor-joining (N-J) method with 1000 bootstrap replicates implemented in MEGA (Molecular Evolutionary Genetic Analysis) version 6 program (Tamura et al. 2013), to evaluate the degree of support for the particular grouping pattern in the constructed cladogram. Domain and family analysis of *GmIPK1* was done using CDD (http://www.ncbi.nlm.nih.gov/cdd) tool.

Results and discussion

Expression profiles of GmIPK1

Low phytate trait in plant seeds offers important nutritional and environmental benefits for food and feed uses. Controlling the level of phytate requires its specific targeted silencing of genes involved in the phytate pathway without affecting other critical aspects of inositol metabolism in different plant tissues. Thus the spatiotemporal expression pattern of *GmIPK1* needs to be fully understood to exploit it for achieving the desired level of gene silencing in the target tissue that is seed.

An initial expression profiling of *GmIPK1* was carried out by Stiles (2007) reporting spatial gene expression pattern in limited number of tissues such as leaves, roots and seeds; Yuan et al. (2012) also surveyed *GmIPK1* expression analysis in root, stem, leaf, flower and seeds. Thus to gain greater understanding of the steady state mRNA level of *GmIPK1*, initially, the transcript accumulation pattern was determined across different tissues (root, stem, leaf, flower and seed). The transcript levels were normalized relative to the internal control gene, PEP

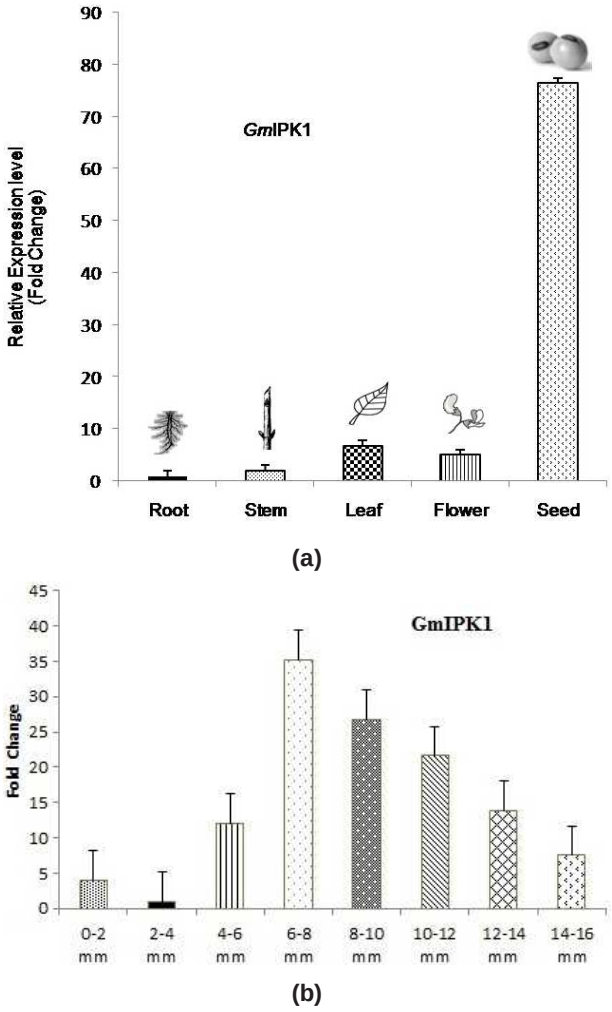


Fig. 1. (a) Transcript levels of *GmIPK1* in different tissues of Pusa 9712 cultivar. The root tissue was taken as calibrator. (b) Transcript levels of *GmIPK1* across developing seed stages of Pusa 9712 cultivar. The 0-2mm seed size was taken as calibrator

carboxylase. The highest level of expression (more than 70 fold change) was observed in seeds (Fig. 1a) supporting its role in seed phytic acid biosynthesis. The fold change expression in other tissues studied were less than fivefold. The reduced transcript levels in other soybean tissues could be due to the fact that occurrence of phytic acid has been reported in every tissues (Singh et al. 2013) where it plays diverse roles ranging from cellular metabolism to tackling various oxidative stresses (Graf and Eaton 1990). In addition, such diverged expression profiles have also been seen

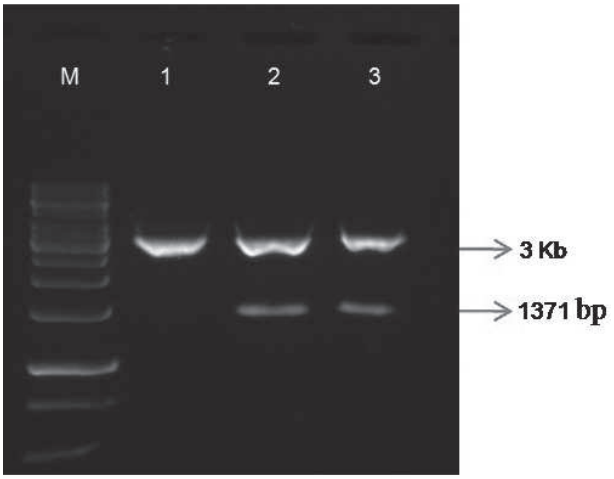
for other enzymes of this distinct pathway thereby justifying their role in coordinating inositol metabolism with cellular growth and other signaling mechanisms (Krishnan et al. 2015; Majumder et al. 1997). The expression profile obtained in the present study is in line with Stiles (2007) and Yuan et al. (2012) where the highest expression was also reported in seeds confirming *GmIPK1* to be seed specific. In addition to tissue specific expression analysis, we also performed temporal expression analysis in developing seeds at 8 different stages of seed development. The expression of *GmIPK1* showed a preferential expression in the later stages of developing seeds particularly from 6-8 mm of seed size showed maximum expression of 35 fold change; and the expression was present up 14-16 mm of seed developmental stages with decreasing trends (Fig. 1b). The semi q-PCR analysis of Yuan et al., (2012) also reported the similar expression pattern of *GmIPK1*. The expression analysis thus confirms *GmIPK1* to be the ideal candidate for designing efficient silencing strategies for reduction of seed phytate levels.

Cloning and sequencing of GmIPK1

Custom designed primers synthesized from the sequence data of the *GmIPK1* cDNA available in NCBI (EU033956) were used to amplify the 1371 bp open reading frame (ORF) encoding a single polypeptide of 456 amino acid residues from the seeds of Pusa 9712, a popular Indian soybean genotype. The amplicon was cloned in pGEM®-T Easy vector, sequenced and submitted in NCBI as KF359958.1 (Fig. 2). *GmIPK1* sequence similarity search using BLASTP analysis retrieved homology to protein sequences of *Phaseolus vulgaris* (XP_007160729.1), *Cicer arietinum* (XP_004499187.1) and *Arabidopsis thaliana* (NP_173629.2) with 85%, 75% and 54% of identity, respectively. Table 1 shows the selected 10 IPK1 sequences retrieved from NCBI BLASTP using *GmIPK1* as the query sequence**.**

Physico-chemical and structural features

The primary structural analysis of the selected IPK1 homologs was done by computing different parameters using ExPASy ProtParam and PEPSTATS tools and tabulated as Table 2. The analysis revealed Leucine (Leu) as the most abundant amino acid for all the IPK1s under study; Methionine (Met) was the least abundant. The predicted average molecular weight (MW) calculated was 50452.32 g mol^{-1} (Table 2). The computed average pI of the IPK1 sequences was 6.46 indicating that this enzyme is likely to precipitate in



(a)

|KF359958.1| Glycine max cultivar Pusa 9712 inositol pentakisphosphate 2-
plete cds
AAGAGGAAGACGCAGCTGACTGGGTTTACAGAGGGGAAGGAGCAGTTAATCTTG
ATCCTCACTTTCTTTTATTGGGAAAGTGGTCCGCATACGTAAGGCTCCAAGGAA
AGTGTGAGTGTGAGGAATAGCATAGCTTTGACTCCACATGAACGTGTTCTCTGG
TTATCTCCTCTTCAGACAAGGAAATAGTTGGTCAACTATATGTGCAGCATGTTA
TTCCAACTCTGTTGATGCTGGGATGCACGTCCTGGTGACCAGGGAATTCCTTGA
GTTTCTGGTCAACGTCCTGCGTGGCGAGTTGAAGCTGCCAGGGTTGATGCACAT
TCATGTCAGATCATTCTCTCTTTGCTTATGGCAGTCAAGGATCTAGCCTCTGCT
GCCCAAATGCGGATTTCTTCCTCTTTCAAGATTCATATCTGAAGTAAATGCTAT
CGATTTGAAATGCACCAAACTCTGAAATTGCTTCAAGGAGAGATATCGCAACTA
TTGATCTGTTCTCTGGATCCAAGGAAAGAATTCTGAAAGCTATTAAAAGTCTCC
CAATTTTCGCGTATTTTTGAATGGCTCTCTCATACTTGGAGGACTGGGAGGTGT
GTTTGTATTGCTAAAGCATTTGAAGATGAACTTAAGTCAATCATTCGAGCTGAT
ATAACTTATCTACTCTTGTTACTGAGGCTTTGCAAAAATCAGGAGTCCTTGATA
GAAGCTTGATAATATTGACGTAGAAGGGGTCATCCATGCATATTATAACATTAC
GTGTGTAAGGAATTGAGTGAAGAACAGGCAAAAATATATACCCCTTTGCATTCA
GTTTGAGAATTGTAAAGGACTACCTGATAGCAACAACTGCAAAAGACTGCAGTT
ACCAAGGGATGAGGAGGATTCTGGATCTGTGTACGATAATGTATATCTGGACTC
GATTATAAGGTGTATTTTATCGACCTTGATTTAAAGCGCTTAAGTAAAGTGGAA
ATAAGAAGATAGTGAACTGCTACAGACAAATCATCAAAATGGATCAAGGAAGAA
GAAGGCATCTAATGCTGCATATTGA

(b)

Fig. 2. Cloning of *GmIPK1* cDNA in pGEMT easy vector system. (a) PCR amplification of 1371 bp fragment isolated from *Glycine max* cv. Pusa 9712 using gene specific primers and separated on 1% agarose gel along with 1 kb molecular weight marker; (b) IPK1 Sequence submitted in NCBI (KF359958.1)

Table 1. IPK1 sequences retrieved from NCBI database (http://www.ncbi.nlm.nih.gov)

Organism	Accession no.	Nucleotide specificity
Glycine max	AGZ03654.1	ATP
Cicer arietinum	XP_004499187.1	ATP
Phaseolus vulgaris	XP_007160729.1	ATP
Glycine soja	KHN45871.1	ATP
Vitis vinifera	XP_010664715.1	ATP
Ricinus communis	DAA64567.1	ATP
Brassica napus	CDY52373.1	ATP
Malus domestica	XP_008347364.1	ATP
Nicotiana sylvestris	XP_009792964.1	ATP
Arabidopsis thaliana	NP_173629.2	ATP

Table 2. Parameters of the selected IPK1 sequences (from plants) calculated using the ProtParam program: molecular weight (MW) (g/mol); isoelectric point (pI); extinction coefficient (EC) (M^{-1} cm^{-1}); instability index (Ii); aliphatic index (Ai); grand average hydropathy (GRAVY); number of negative residues (-R); number of positive residues (+R)

Organism	Seq	MW length	pI	EC	Ii	Ai	GRAVY	-R	+R
Glycine max	456	51132.5	6.19	41830	34.66	99.39	-0.186	60	56
Cicer arietinum	454	50984.1	6.17	38850	37.61	90.42	-0.331	62	57
Phaseolus vulgaris	489	54911.2	8.13	40340	36.10	95.48	-0.198	63	66
Glycine soja	420	47227.4	6.61	34380	32.07	92.88	-0.146	53	51
Vitis vinifera	453	50935.3	6.71	48820	33.98	85.87	-0.233	54	52
Ricinus communis	450	50901.4	6.52	38850	32.74	94.87	-0.259	61	58
Brassica napus	450	50482.5	6.08	41370	43.25	90.73	-0.368	64	60
Malus domestica	430	47614.8	6.57	30370	34.42	87.74	-0.243	54	52
Nicotiana sylvestris	449	50167.2	5.83	40340	28.53	88.98	-0.215	54	46
Arabidopsis thaliana	441	50166.8	5.83	35870	33.43	93.22	-0.269	66	57

acidic buffers (Table 2); this information will be of much use in the development of buffer systems for the purification of recombinant IPK1 proteins by isoelectric focussing. The low pI values (< 6) observed may be attributed to the dominance of surface OH species (Satoshi and Makoto, 2005) which might play a role in bond formation. The *in vivo* half-life ($T_{1/2}$) of the selected kinases was calculated via Instability Index (Ii). Previous studies reported that proteins having Ii > 40 have a $T_{1/2}$ of less than 5 hours, while those having Ii < 40 have a longer $T_{1/2}$ of 16 hours (Rogers et al. 1986). Our study showed that with the exception of *Brassica napus,* the other IPK1 proteins including *GmIPK1* was stable. Although the positional or interaction effects of the adjacent residues have not been taken into consideration for analysing the stability, yet the study provides some indication about the physical state of the protein. GRAVY indices of IPK1 sequences ranged from –0.146 to –0.368 (Table 2) reflecting the hydrophobicity of the amino acids (Kyte and Doolittle, 1982). The present study thus predicted that all IPK1 enzymes tend to be hydrophilic in nature, which endorses its multifaceted role in cell signaling. Aliphatic Index (Ai), another parameter that measures the relative volume occupied by the aliphatic side chains of the amino acids like Alanine (Ala), Valine (Val), Leucine (Leu) and Isoleucine (Ile), serves as a measure of thermo stability of proteins (Gupta et al. 2012). In our study, the Ai values of all the kinases under consideration ranged from 85.87 to 99.39. IPK1 sequences from fruit bearing species [*Vitis vinifera* (85.87), *Malus domestica* (87.74)] are presumed to be less thermostable than other horticultural and legumes crops as high Ai is an indicator of protein stability over a wide temperature range. The extinction coefficient (EC), which indicates the light absorbed by the protein at a wavelength of 280 nm, was found to be in a range of 30370 to 48820 M^{-1} cm^{-1} with respect to its concentration of Cysteine (Cys), Tryptophan (Trp), and Tyrosine (Tyr).

Sequence alignment and domain analysis

Multiple alignment of the selected amino acid sequences of the homologs of IPK1 from 10 different plants were performed using Clustal Omega program. The conserved amino acids are shown in Fig. 3. The analysis revealed most of the amino acids in the Ins_P5_2-kinase domain is well conserved across IPK1 homologs implicating their function in conferring the enzymes its kinase activity. The amino acids playing a major catalytic role may be Arginine (Arg), Valine (Val), Glycine (Gly), Glutamic acid (Glu) and Leucine (Leu) which can be confirmed further by other *in silico* studies. Domain analysis reveals that IPK1s belong to Inositol pentakisphosphate 2 kinase family possessing a major Ins_P5_2-kinase domain required for its catalytic function (Fig. 4).

Phylogenetic tree

Phylogenies organize the knowledge of biodiversity and structural similarities into distinct clades (Baum 2008). Also, phylogenetic analysis provides a useful

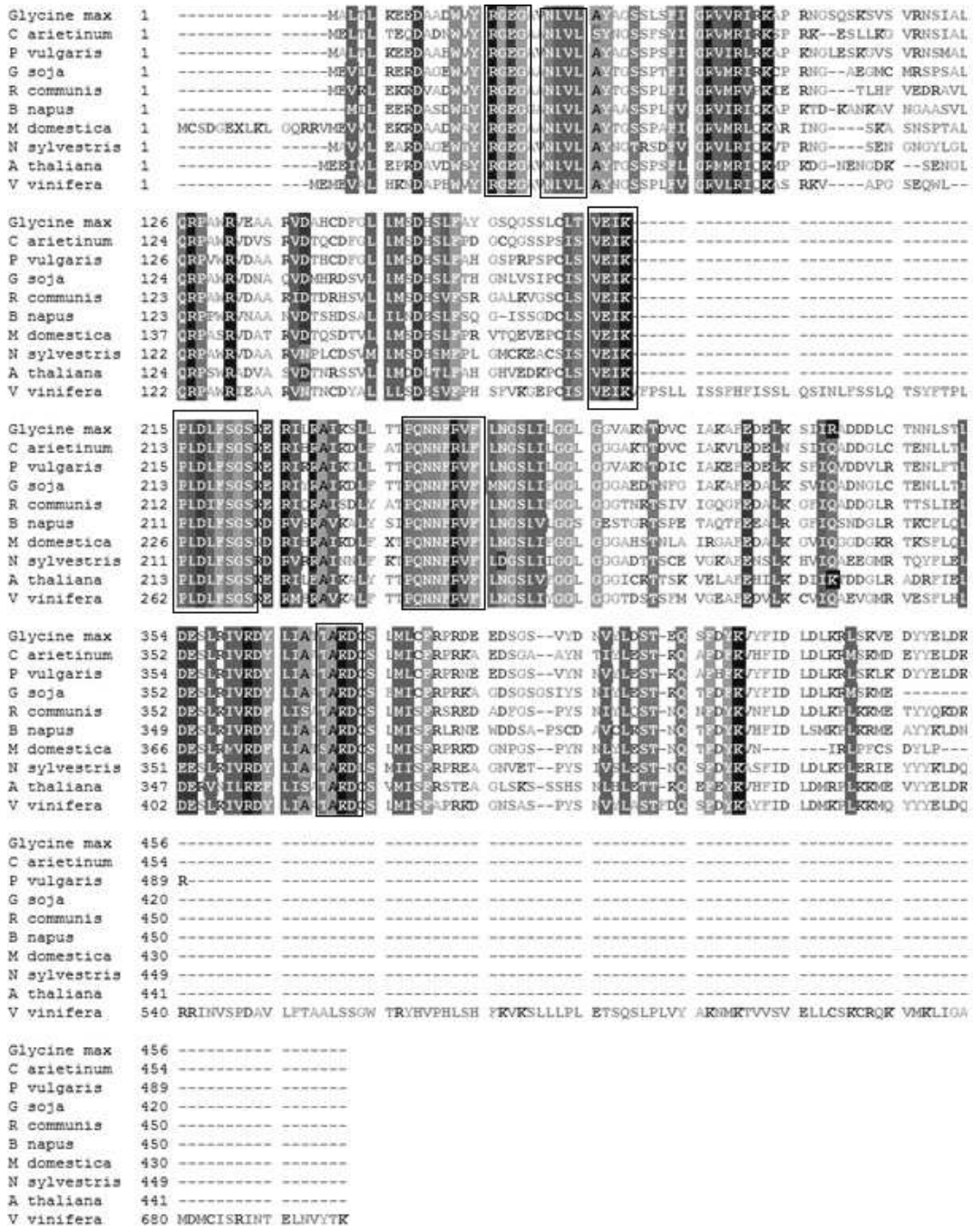


Fig. 3. Significantly conserved sequences as analyzed using Clustal Omega

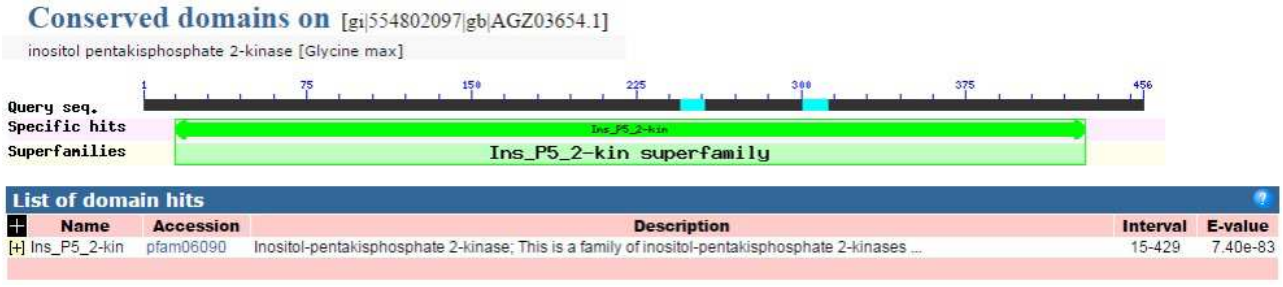


Fig. 4. Domain analysis of *Gm*IPK1

framework to understand the evolutionary relationships of group of homologous proteins. Differential expression of phytic acid biosynthetic pathway have previously been carried out in wheat (Bhati et al. 2014) and *Arabidopsis* (Sweetman et al. 2006) but the similarity study had been carried out not considering a large group of representative organisms. In this

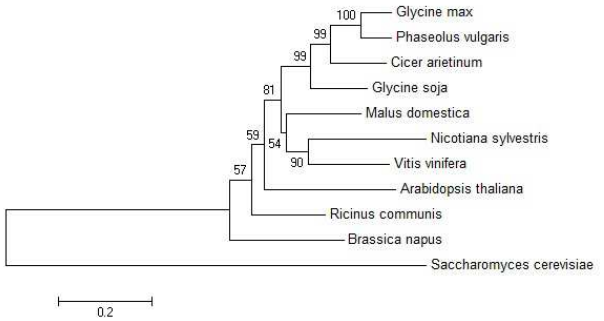


Fig. 5. Cladogram of the selected 10 inositol kinases based on their amino acid sequences as outlined in Table 1 using Mega 6.0 software. The tree was inferred based on the neighbor joining algorithm. *Saccharomyces cerevisiae* was considered as the outgroup

study, we have constructed a neighbor-joining phylogenetic tree of IPK1 homologs of 10 different plants using (Fig. 5) MEGA software. Yeast IPK1 sequence was used as outgroup. Our study showed the presence of a distinct clade among the selected IPK1 homologs. Interestingly, the IPK1s of plants belonging to pulse crop formed a separate subgroup in the clade and consist of members from soybean, *Phaseolus*, chickpea and IPK1 from apple, grapes and *Nicotiana* formed another subgroup in a clade of phylogeny. IPK1s from Arabidopsis, *Ricinus* and *Brassica* were present separately in the phylogeny. Thus, phylogeny of homologs of IPK1 illustrated a high degree of divergence of legume IPK1s and with that of other plants and only there was conservation of catalytic domains in the IPK1 homologs.

Authors' contribution

Conceptualization of research (NB, AS, VK); Designing of the experiments (NB, AS, VK); Contribution of experimental materials (AS, MJ); Execution of field/ lab experiments and data collection (NB, VP, MP, AM, AH); Analysis of data and interpretation (NB, VK, AS); Preparation of manuscript (NB, VK, AS, MJ).

Declaration

The authors declare no conflict of interest.

Acknowledgements

Financial support provided by Department of Science and Technology, Government of India, in the form of INSPIRE fellowship (IF110180) to N.B. and funding by National Agricultural Science Fund (RNAi 20-11), Indian Council of Agricultural Research to A.S. is duly acknowledged.

References

Ali N., Paul S., Gayen D., Sarkar S.N., Datta K., Datta S. K. 2013. Development of Low Phytate Rice by RNAi Mediated Seed-Specific Silencing of Inositol 1,3,4,5,6-pentakisphosphate 2-Kinase Gene (*IPK1*). PLosOne, doi.org/10.1371/journal.pone.006816.

Baum D. 2008. Reading a phylogenetic tree: the meaning of monophyletic groups. Nature Education, **1**(1): 190-196.

Bhati K. K., Aggarwal S., Sharma S., Mantri S., Singh S. P., Bhalla S., Kaur J., Tiwari S., Roy J. K., Tuli R. and Pandey A. K. 2014. Differential expression of structural genes for the late phase of phase of phytic acid biosynthesis in developing seeds of wheat (*Triticum aestivum* L.). Plant Science, **224**: 74-85.

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M. R., Appel R. D. and Bairoch A. 2005. Protein identification and analysis tools on the ExPASy *Server. In:* The Proteomics Protocols Handbook (ed. J. Walker J.), Humana Press, New York, USA: 571-607.

Goujon M., McWilliam H., Li W., Valentin F., Squizzato S., Paern J. and Lopez R. 2010. A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res., **38**: 695-699.

Graf E. and Eaton J. W. 1990. Antioxidant functions of phytic acid. Free Radic Biol. Med., **8**(1): 61-69.

Gupta S. K., Rai A. K., Kanwar S. S. and Sharma T. R. 2012. Comparative analysis of zinc finger proteins involved in plant disease resistance. PLoS One, **7**(8): e42578 doi: 10.1371/journal.pone.0042578.

Harada, K. and Xia Z. 2004. Soybean genomics: efforts to reveal the complex genome. Breeding Science, **54**: 215-224.

Krishnan V., Hada A., Basak N., Pandey V., Marathe A., Punjabi M., Jolly M., Vinutha T., Manickavasagam M. and Sachdev A. 2015. RNA interference-based gene silencing: An efficient tool for developing designer crops.J of Science, Technology and Management, **8**(1): 11-18.

Kyte J. and Doolittle R. F. 1982. A simple method for displaying the hydropathic character of a protein. J.

Mol. Biol., **157**(1): 105-132.

Livak K. J. and Schmittgen T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2[-Delta Delta C (T)] method. Methods, **25**: 402-408.

Majumder A. L., Johnson M. D. and Henry S. A. 1997. 1L-*myo*-Inositol-1-phosphate synthase. Biochimica et Biophysica Acta, **1348**: 245-256.

Raboy V. 2009. Approaches and challenges to engineering seed phytate and total phosphorus. Plant Sci., **177**: 281-296.

Raboy V. 2007. The ABC s of low phytate crops. Nature Biotechnology, **25**(8): 874-875.

Rice P., Longden I. and Bleasby A. 2000. Emboss: The European molecular biology open software suite. Trends Genet., **16**(6): 276-277.

Rogers S., Wells R. and Rechsteiner M. 1986. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. Science, **234**: 364-368.

Satoshi T. and Makoto F. 2005. Role of surface OH groups in surface chemical properties of metal oxide films. Mater Sci. Engr., **119**: 265-267.

Singh P., Punjabi M., Jolly M., Rai R. D. and Sachdev A. 2013.Characterization and expression of codon optimized soybean *phytase* gene in *E. coli*. Indian J Biochem Biophys., **50**: 537-547.

Stevenson-Paulik J., Bastidas R. J., Chiou S. T., Frye R. A. and York J. D. 2005. Generation of phytate-free seeds in Arabidopsis through disruption of inositol polyphosphate kinases. PNAS, **102**: 12612-12617.

Stiles A. R. 2007. Identification and Characterization of Late Pathway Enzymes in Phytic Acid Biosynthesis in *Glycine max.* Ph. D. Thesis. Virginia Polytechnic Institute and State University, USA.

Sun Y., Thompson M., Lin G., Butler H., Gao Z., Thornburgh S., Yau K., Smith D. A. and Shukla V. K. 2007. Inositol 1,3,4,5,6-pentakisphosphate 2-kinase from maize (*Zea mays* L.): molecular and biochemical characterization. Plant Physiol., **144**: 1278-1291.

Sweetman D., Johnson S., Caddick S. E., Hanke D. E. and Brearley C. A. 2006. Characterization of an Arabidopsis inositol 1,3,4,5,6-pentakisphosphate 2-kinase (*At*IPK1). Biochem J., **394**: 95-103.

Tamura K., Stecher G., Peterson D., Filipski A. and Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol. Biol. Evol., **30**(12): 2725-2729.

Verbsky J. W., Wilson M. P., Kisseleva M. V., Majerus P. W. and Wente S. R. 2002. The synthesis of inositol hexakisphosphate- Characterization of human inositol 1,3,4,5,6-pentakisphosphate 2-kinase. J. Biological Chemistry, **277**: 31857-31862.

York J. D., Odom A. R., Murphy R., Ives E. B. and Wente S. R. 1999. A phospholipase C-dependent inositol polyphosphate kinase pathway required for efficient messenger RNA export. Science, **285**(5424): 96-100.