Plant Omics Journal POJ

POJ 16(01):1-10 (2024) ISSN:1836-3644
doi: 10.21475/POJ.16.01.24.pne02

Genome-wide identification and characterization of *Inositol Phosphokinase (IPK)* gene family in wheat (*Triticum aestivum* L.)

Saira Ibrahim[1], Tayyaba Andleeb[2], Muhammad Ramzan Khan[3], Muhammad Uzair[3]

[1]Department of Biochemistry, Faculty of Biological Sciences, Quaid-i-Azam University, Islamabad, Pakistan
[2]Department of Plant Sciences, Faculty of Biological Sciences, Quaid-i-Azam University, Islamabad, Pakistan
[3]National Institute for Genomics and Advanced Biotechnology (NIGAB), National Agricultural Research Centre (NARC), Islamabad, Pakistan

***Correspondence: sibrahim@bs.qau.edu.pk**

Submitted: 06/03/2024

Revised: 24/05/2024

Accepted: 05/06/2024

Abstract

Inositol phosphate kinases *(IPKs)* play vital roles in the synthesis and regulation of cellular levels of inositol polyphosphates, which perform vital functions in eukaryotic cells as second messengers. Due to the vital biological roles of the inositol phosphates, the kinases involved in their synthesis have gained great attention. The *IPK* gene family has not been extensively studied in bread wheat (*Triticum aestivum*). In the present study, we reported a genome-wide identification, phylogenetic analysis and expression patterns of the *IPK* gene family in wheat. Gene structure, genome distribution, motif conservation, and gene ontology enrichment analysis were carried out systematically. A total of 24 inositol phosphate kinase *(IPK)* genes were identified in the wheat genome that belonged to 8 homoeologous groups. The *TaIPK* genes were distributed on chromosomes 1, 2, 3, 4, 5 and 7 in all the three sub-genomes A, B, and D. Based on phylogenetic analysis, these genes were classified into four subfamilies. The subfamilies were defined based on conserved domains, motifs, chromosome locations, and gene structures. Eight pairs of paralogous *IPK*s were identified based on the phylogenetic relationships among wheat *IPK*s. Gene ontology enrichment analysis revealed the significant role of *IPK* genes in different biological and molecular processes in addition to their role in inositol phosphate signaling. At different developmental stages, all the 24 *TaIPK*s exhibited diverse expression patterns in different tissues showing diversity in their biological functions. The current study identified and explored the properties of 24 *IPK* genes in the wheat genome at diverse levels, thus providing a strong foundation for researchers to understand the family of these kinases in wheat.

Keywords: Expression patterns, Gene structure, Inositol polyphosphates, Inositol phosphate kinases, Phylogenetic analysis.
Abbreviations: IPK_Inositol phosphokinase; IPMK_inositol phosphate multikinases; ITPK_ inositol tris/tetrakisphosphate; IP5-2K_ Inositol- pentakisphosphate 2 kinase; MEME_multiple em for motif elicitation; PPIP5K_ diphosphoinositol-pentakisphosphate kinase

Introduction

Inositol phosphate kinases *(IPKs)* regulate the cellular levels of inositol polyphosphates which perform various vital biological roles in eukaryotic cells as second messengers (Hatch and York, 2010; Shears, 2017). The kinases involved in the synthesis of inositol polyphosphates have gained great attention. Most of these enzymes are the members of *IPK* superfamily. Which consist of inositol 3-kinases (IP3K) and inositol phosphate multikinases ((IPMK) subgroups. With very low sequence conservation, these superfamily members share several strictly conserved signature motifs with each other and are expected to adopt the same overall fold (Shears and Wang 2019). Substrate selectivity of these enzymes is expected due to differences in their sequence. Among these subgroups, *IPMK*s generally act as 6/3/5-kinases showing catalytic activities towards more substrates (Laha et al., 2021).

InsP6 (Inositol 1, 2, 3, 4, 5, 6-hexakisphosphate) recognized as phytic acid, plays a vital role in cell biology. For instance, IP6 is involved in different processes such as DNA repair, export of mRNA, maintaining basal resistance against plant pathogens, chromatin remodeling and apoptosis (Steger et al., 2003; Agarwal et al., 2009). In addition, phytic acid is involved in the formation of pyrophosphate inositols IP7 and IP8 (Wilson et al., 2013). Phytic acid has gained considerable commercial interest due to its vital roles in cell biology (Bloot et al., 2023).

In plant seeds, IP6 serves as the major phosphorus reserve, and acts as a source of minerals, phosphorus and inositol during germination and development. Concerns arise from the fact that IP6 can cause harmful effects on the environment as well as on human health. Grain-based foods provide high levels of phytic acid that acts as a strong chelator of important ions like Zn^{+2} and Fe^{+2}, thus reducing the bioavailability of these essential ions in diet that contribute to malnutrition or hidden hunger (Brouns, 2021). In the biosynthesis of phytic acid, the enzymes catalyzing the sequential phosphorylation reactions are the members of *IPK* family of kinases which include *inositol tris/tetrakisphosphate (ITPK) kinase*, *inositol-pentakisphosphate 2 kinase (IPK1)* and *inositol phosphate multikinase (IPK2)* (Raboy, 2009; Rasmussen et al., 2010; Sharma et al., 2021).

In plants, phytic acid also known as phytate can be produced by lipid-independent or lipid-dependent pathways. *ITPK*s are involved in the lipid- independent pathway of PA synthesis and are primarily the members of the ATP-grasp proteins family. Plant *ITPK1* might even have evolved to synthesize IP_7 by using IP6 as a substrate (Laha et al., 2019). *IPK1* (*Inositol-pentakisphosphate 2 kinase)* is the key enzyme that catalyzes the last step of phytate synthesis (Gonzalez et al. 2010). This enzyme is a potential target for producing low phytic acid (lpa) crops. Inositol phosphate multikinase *(IPK2)* genes have been described to be involved in the lipid-dependent pathway of phytate synthesis (Punjabi et al., 2018). Multiple *IPK* superfamily genes have been identified in different cereals like maize and rice (Shi et al., 2003; Suzuki et al., 2007). Cereals are rich sources of micronutrients and mineral ions, but the bioavailability of micronutrients is restricted due to the presence of an anti-nutrient phytate (Gupta et al., 2015). In recent years, the role of *IPK* family members has been a research topic of great interest because these enzymes are involved in the sequential phosphorylation steps of phytate synthesis in cereal crops and may be strong candidates for targeting the production of biofortified (lpa) crops. The hexaploid wheat (*Triticum aestivum L*.), being a source of 20% of the food calories and 55% of the carbohydrates consumed globally, is one of the essential food crops. (Shewry and Hey 2015). Among the subgroups of the *IPK* superfamily, *IP3K*s are not represented in yeast and plant genomes (Laha et al., 2021). Bhati et al., (2014) identified the *IPK* family members (*TaIPK1, TaIPK2* and *TaITPK1-4*) in wheat based on the conserved domains and conserved amino acid sequences. However, a detailed analysis of the wheat *IPK* family members known as inositol phosphate kinases (designated *TaIPK*s) has not been performed so far. Given all these facts, the goal of the present study was to identify and understand the *IPK* family genes in wheat.

Since the *T. aestivum* genome has been sequenced, genome-wide analysis of different genes has become more feasible. In the current study, we analyzed 24 *TaIPK* genes for their conserved motifs, domains, structure, chromosomal locations within the *T. aestivum* genome and phylogenetic relationships. Moreover, the expression patterns of all 24 *TaIPK* genes were also analyzed *in silico*. This study provides a strong foundation for understanding the *IPK* family members in wheat.

Results

Identification of IPK family genes in T. aestivum

After eliminating different transcripts of the same gene, a total sum of 24 protein sequences (3 TaPPIP5K1, 3 TaIPK1, 3TaIPMK and 15 TaITPK) expressing the primary transcript were identified. One highly conserved domain of the four *IPK* subfamilies was present in all 24 TaIPKs which was also confirmed through the online Conserved Domains database. To differentiate these genes, the present study named these genes based on chromosomal distribution and subfamily branch as *TaPPIP5K1* (chr3) on A, B, and D sub-genomes, *TaIPK1* (chr2) on A, B, and D sub-genomes, *TaIPMK* (chr7) on A, B, and D sub-genomes and *TaITPK1-4* and *TaITPK6* (chr1, 4, 5, 7) on A, B, and D sub-genomes. All the 24 TaIPKs protein lengths ranged from 287 amino acids to 1036 amino acids (Table 1), displaying a wide distribution of IPK lengths. The *TaPPIP5K* protein lengths (1036 amino acids) were generally the longest and the TaIPMK lengths ranging from 310 to 356 amino acids were the shortest among the four sub families of IPKs. The predicted Mw (molecular weights) of the 24 TaIPKs ranged from 30.61Kdal (TaIPMK-7D) to 117.8Kdal (TaPPIP5K1-3A, B, and D), and the predicted pI (isoelectric points) of the 24 TaIPKs ranged from 4.7 (TaITPK4-7B) to 8.80 (TaIPK1-2B).

Phylogenetic analysis, gene structure and genome distribution of TaIPKs

By using MEGA X software, a phylogenetic tree was constructed to examine the phylogenetic relations between the *TaIPKs* based on the *IPK* nucleotide sequences alignment (Fig.1, Supplementary File1). To show the structural classification of the *TaIPK*s, the NJ phylogenetic trees were constructed. As displayed in Fig. 1, the NJ phylogenetic tree was clearly divided into four groups, indicated by four different colors (*TaITPK*s, blue; *TaIPMK*s, green; *TaIP5-2K*s, red; TaPPIP5Ks purple), which were per our identification results. The *TaPPIP5K1* branch had a simple structure, and a pair of paralogous *TaPPIP5K1*s (*TaPPIP5K1-3B* and *3D*) was detected. *TaIPK1s* branch structure was similar to that of the *TaPPIP5K1*s. Similarly the branch structure of IPMK genes (TaIPMK2-7A, B, and D) was similar having one pair of paralogous genes. The *TaITPK*s branch contained two major groups. Two bifurcations were present in the first group, separating Ta*ITPK2-4A, B,* and *D* from a second subgroup composed of Ta*ITPK3-4A, B,* and *D*. There were also two bifurcations present in the second group, separating Ta*ITPK1-1A, B,* and *D* from a second subgroup composed of Ta*ITPK4-7A, B,* and *D*. A third group containing *TaITPK6-5A*, *B,* and *D* was positioned outside the two other groups. Twenty-four *TaIPK*s were distributed on chr1, 2, 3, 4, 5 and 7 in all the three A, B, and D sub-genomes. The chr4 and chr7 contained two members of *IPK*s each, while other chromosomes contained only one member (Fig. 2). Coding sequences and the genomic DNA sequences of the *TaIPK*s were used to predict the *TaIPK*s gene structure by using Gene Structure Display Server GSDS2.0 (http://gsds.cbi.pku.edu.cn/) (Fig. 3). All *TaIPK*s contained 1-30 introns. *TaIPMK-7A*, *B,* and *D* genes were composed of only one exon. Similarly, some *TaITPK*s *(TaITPK1-1A*, *B*, *D)* and *(TaITPK4-7A*, *B*, *D)* were also composed of only one exon.

Table 1. Information on the *IPK* family genes in wheat.

Gene name	Gene locus	Transcript ID	Gene ID	Protein length (AA)	Mol.wt. (Kdal)	No. of introns	Coding exons	PI	Sub-family
TaIPK1-2A	728405886-728410445	TraesCS2A02G497700	123190668	454	50.28	6	7	7.04	IP5-2K
TaIPK1-2B	719966007-719970727	TraesCS2B02G525900	123046991	576	62.8	7	8	8.80	IP5-2K
TaIPK1-2D	593510756-593514283	TraesCS2D02G612600LC	123054827	469	50.78	6	7	7.15	IP5-2K
TaIPMK-7A	620634772-620637025	TraesCS7A02G427500	123148742	362	38.95	1	1	7.61	IPK
TaIPMK-7B	581794516-581796383	TraesCS7B02G327600	123161229	287	30.66	1	1	6.42	IPK
TaIPMK-7D	538987871-538989945	TraesCS7D02G419800	123167046	287	30.61	1	1	6.60	IPK
TaPPIP5K1-3A	561090306-561119817	TraesCS3A02G319100	123062235	1036	117.8	30	31	6.5	PPIP5K
TaPPIP5K1-3B	556487647-556519303	TraesCS3B02G347000	123070985	1036	117.8	30	31	6.3	PPIP5K
TaPPIP5K1-3D	426845398-426880464	TraesCS3D02G312200	123079366	1036	117.8	30	31	6.3	PPIP5K
TaITPK1-1A	351349023-351350790	TraesCS1A02G194100	123050359	356	39.12	0	1	5.78	ITPK1
TaITPK1-1B	378837882-378838943	TraesCS1B02G209100	123125975	353	38.93	0	1	5.68	ITPK1
TaITPK1-1D	279103037-279104802	TraesCS1D02G198000	123181474	341	38.82	0	1	5.68	ITPK1
TaITPK2-4A	55912521-55919370	TraesCS4A02G059500	123085097	310	34.21	9	10	5.33	ITPK1
TaITPK2-4B	491016693-491022782	TraesCS4B02G235100	123092886	356	39.22	9	10	5.65	ITPK1
TaITPK2-4D	398429576-398436307	TraesCS4D02G236500	123098166	349	38.81	9	10	5.44	ITPK1
TaITPK3-4A	570466220-570471987	TraesCS4A02G257700	123087024	350	38.72	9	10	7.30	ITPK1
TaITPK3-4B	46615028-46621497	TraesCS4B02G056800	123090936	351	38.62	9	10	7.30	ITPK1
TaITPK3-4D	32230928-32237040	TraesCS4D02G057100	123095966	352	38.66	9	10	7.3	ITPK1
TaITPK4-7A	113485607-113486853	TraesCS7A02G158400	123154092	342	36.73	0	1	4.81	ITPK1
TaITPK4-7B	67818474-67819722	TraesCS7B02G063100	123159722	341	36.63	0	1	4.7	ITPK1
TaITPK4-7D	108618713-108620426	TraesCS7D02G159400	123166273	341	36.68	0	1	4.71	ITPK1
TaITPK6-5A	514153380-514157363	TraesCS5A02G305400	123104438	502	55.44	11	12	5.04	ITPK1
TaITPK6-5B	489809723-489813745	TraesCS5B02G305900	123112736	503	55.51	11	12	5.1	ITPK1
TaITPK6-5D	408318334-408322320	TraesCS5D02G312400	123122212	500	55.08	11	12	5.02	ITPK1

AA, amino acid; Mol.wt, molecular weight (Kdal); PI, isoelectric point.

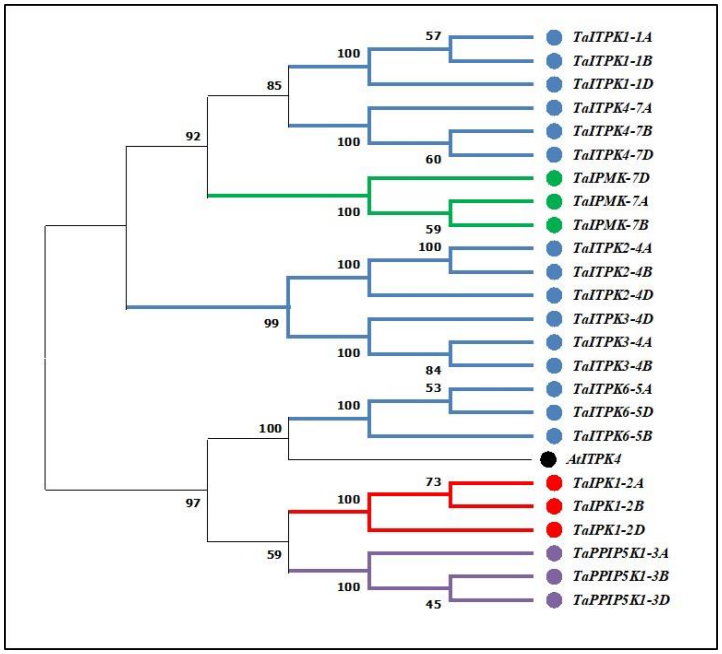

Fig 1. Phylogenetic tree showing the relationship of the *IPK* gene family members in Wheat. 1,000 replicates were used for the Bootstrap test, and the replication frequency is presented next to the branches. Arabidopsis sequence (*AtITPK4*) was used as an out-group. Four subfamilies are represented in different colors.

Conserved motifs of TaIPKs

To identify the putative motifs of the IPK family in wheat, the MEME tool was used to analyze the 24 TaIPKs which is available at the following web site (http://meme-suite.org/tools/meme). Eight conserved motifs were identified and represented as 1-8 (Fig. 4). Motifs 1, 2, and 5 were found in all TaITPK proteins. The main motif 3 specifying the ATP binding domain was found in all TaIPK proteins. Similarly motif 6 was found in all TaIPK proteins. These findings were consistent with the point that they were the members of the same gene family.

Gene ontology analysis of TaIPK genes

Gene ontology analysis was conducted for the *in silico* functional prediction of *TaIPK* genes. The results showed two types of functions that is involvement in Biological Processes

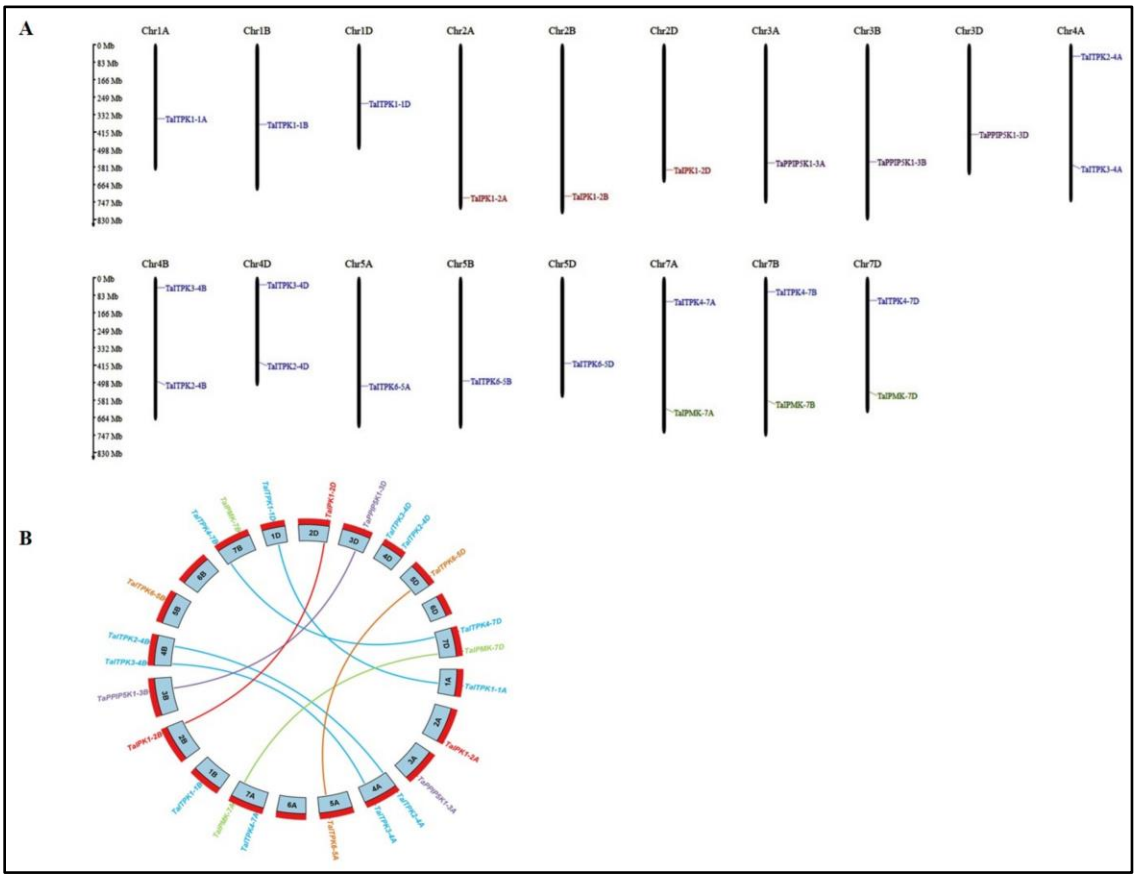

Fig 2. (A) The distribution of 24 *IPK*s in the wheat genome. (B) Physical map visualized by Circos (TBtool) showing eight pairs of paralogous *IPK* genes in the wheat genome. Genes from different sub families are shown in different colors.

(BPs) and Molecular Functions (MFs) (Fig. 5). The molecular processes showed that *IPK* genes have catalytic activity while the biological processes suggested the involvement of *IPK* genes in various metabolic activities and their role in the biosynthesis of various organic compounds. Thus, clearly suggesting the significant role of *IPK* genes in different biological and molecular processes.

Gene expression patterns in different tissues under different developmental stages

Due to the association of gene expression and biological function, we examined the expression patterns of different *IPK* genes. Wheat expression database (http://www.wheat-expression.com/) was used to download RNA-seq data (Fig. 6). It was found that *IPK* genes were broadly expressed in the vegetative (root, shoot, and leaf) and reproductive (spike and grain) tissues showing that *IPK*s work in different tissues and have diversity in biological functions. We observed that some *IPK*s had very low expression levels in the vegetative tissues. *TaIPMK-B, TaITPK2-B, TaITPK4-D*, and *TaITPK6-D* expression levels in shoots/leaves at the vegetative stage were very low. Similarly, expression levels of *TaIPK1-D*, *TaITPK2- B*, *TaITPK4-A*, *B, D*, and *TaPPIP5K1-A, B*, and *D* were very low in roots at the vegetative stage. *TaITPK2-D* gene was observed to be highly expressed in shoots/leaves and spikes at both vegetative and reproductive stages. Some genes like *TaIPK1-A* and *TaITPK6-A, B*, and *D* showed very high expression in grains.

On comparison of the patterns of gene expression in sub-genomes A, B, and D, we observed that the expression levels in sub-genomes were not always the same; for instance, *TaIPK1* in sub-genomes A and D was highly expressed in grain but *TaIPK1* in sub genome B showed low expression in grain. Similarly, *TaITPK2* in sub-genome D showed very high expression levels in shoots/leaves and spikes at both vegetative and reproductive stages but *TaITPK2* in sub-genomes A and B showed very low expression in these tissues. *TaITPK4* in sub-genomes A and B was highly expressed in shoots/leaves at the vegetative stage but the expression was very low in sub-genome D. For some genes like *TaITPK2* and *TaITPK4* expression level in roots at both vegetative and reproductive stages and in spike at a vegetative stage has not been reported yet but these genes show expression in other tissues at different stages (Fig. 6).

Evolutionary relationship of the IPKs in wheat, rice, maize, and Arabidopsis

For observing the evolutionary relationships of the IPK members from wheat, rice, maize, and Arabidopsis, a NJ phylogenetic tree was generated based on the 57 IPK protein sequences (Fig. 7, Supplementary File 2). The NJ phylogenetic tree indicated that all 57 IPK members could be categorized into four groups differentiated by four different colors (purple, red, blue, and green), which were in accordance with the subfamily classification of the IPKs. The ITPK subgroup was the largest, possessing 29 genes. IPMK subgroup contained 10 genes, while the PPIP5K and IP5-2K subgroups contained the same number of genes i.e. 9 genes in each group.

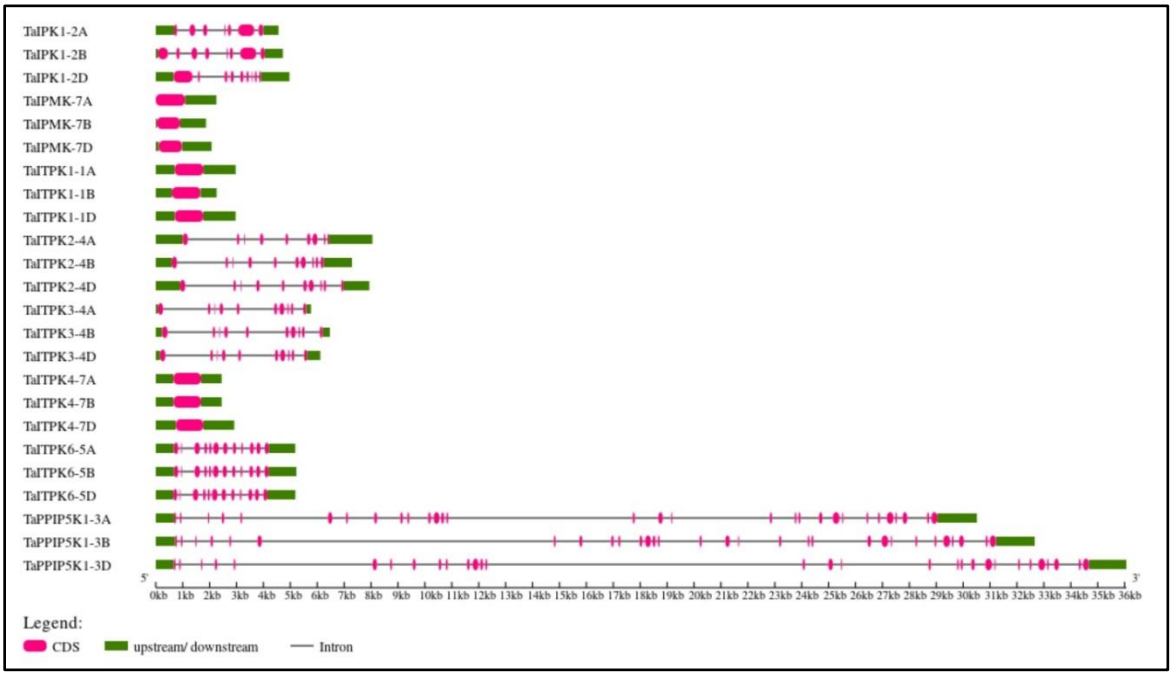

Fig 3. Gene structures of *TaIPKs*. The black line represents intron, the pink box represents CDS and the green box represents the upstream/downstream region. The *TaIPKs* lengths were shown in proportion according to their length.

Discussion

The inositol phosphate kinase (IPK) family of enzymes is involved in the production of inositol phosphates (IPs), a group of signaling molecules involved in vital functions in cellular biology (Tsui and York, 2010). Inositol phosphate signaling is involved in the regulation of ion channels, vesicle transport, DNA repair, and editing (Gosein and Miller, 2013). Among these signaling molecules, Phytic acid (IP6) is the main anti-nutrient found in cereal crops that leads to hidden hunger by reducing the bioavailability of micronutrients in cereal seeds (Perera et al., 2018; Gibson et al., 2018). Wheat, the most important staple food is cultivated all over the world for its seed. Wheat provides carbohydrates (almost 55%) and 20% of total food calories and about 2 billion people (approximately 36%) of the total population consume wheat grain as an essential staple food all over the world (Khalid et al., 2023). However, wheat foods contain very low bioavailable micronutrients because these are rich in anti-nutrients, specifically phytic acid (Balk et al., 2019).

The *IPK*s are present in most eukaryotes and the *IPK* gene family has been extensively studied in cereal crops due to the involvement of these enzymes in the biosynthesis of phytic acid (PA). Bhati et al., (2014) revealed the properties of members of the *IPK* family including conserved domains, conserved amino acid sequences and diverse biological roles. Which were helpful to our research on the *IPK* genes in wheat. A total of 24 *IPK* genes were recognized in common wheat (*Triticum aestivum*). The 24 *IPK*s revealed diversity in several features, including theoretical isoelectric point, molecular weight, gene structure, protein length and expression pattern, suggesting a biological functions diversity among *IPK*s. Based on phylogenetic analysis our 24 *TaIPK* genes were divided into four groups (PPIP5K, IP5-2K, ITPK1 and IPK superfamily) (Fig.1) which was in accordance with the sub-family classification of IPKs (Laha et al., 2021). The hexaploid wheat genome was created because of the hybridization of 3 diploid sub genomes A, B, and D (Marcussen et al., 2014). For each *T. aestivum* gene there should be three homoeologous genes one from each sub-genome (Sharma et al., 2016). The 24*TaIPK* genes in this study were distributed on chromosomes 1, 2, 3, 4, 5 and 7 on all the three sub-genomes A, B, and D, showing the possibility of the absence of deletion of *TaIPK* genes in the course of the evolutionary process of *T. aestivum*. The maximum number of *TaIPKs* were found on chromosomes 4 and 7 (Fig. 2).

Gene organization plays a very important role in the multigene family evolution. In this study, there was great variation in number of introns ranging from 1-30, whereas some *TaITPK*s were composed of only one exon (Fig. 3). This indicates the difference in the tendencies of these genes for deletion or attainment of introns. The same intron distributions have been observed in other cereal crops like rice and maize indicating the evolutionary conservation between different cereal crops. The functionally important parts of proteins are represented by conserved motifs. Motif analysis of *TaIPKs* revealed the conservation of motifs displaying the kinase domain and ATP binding domain with the motifs that were found in all the TaIPK proteins (Fig. 4).

Generally, conversed regions can be used to describe and classify the proteins, and these conserved regions play important roles in their interaction with other molecules. Previous studies have shown that all IPK proteins are recognized by ATP binding domains and inositol phosphate binding (IP) domains (Bhati et al., 2014). Three consensus motif sequences of 13 amino acids “GEGAANLILSYTG”, 9 amino acids “RYKMHQHLK” and 13 amino acids “LDDHDIEGAIHLY” were present in all the genes grouped as TaIPK1 (Supplementary Figure 1 A). Based on the recognized inositol phosphate (IP) binding domain of IPMK in plants, in the N-terminus of the TaIPMK conserved region the similar 21-amino

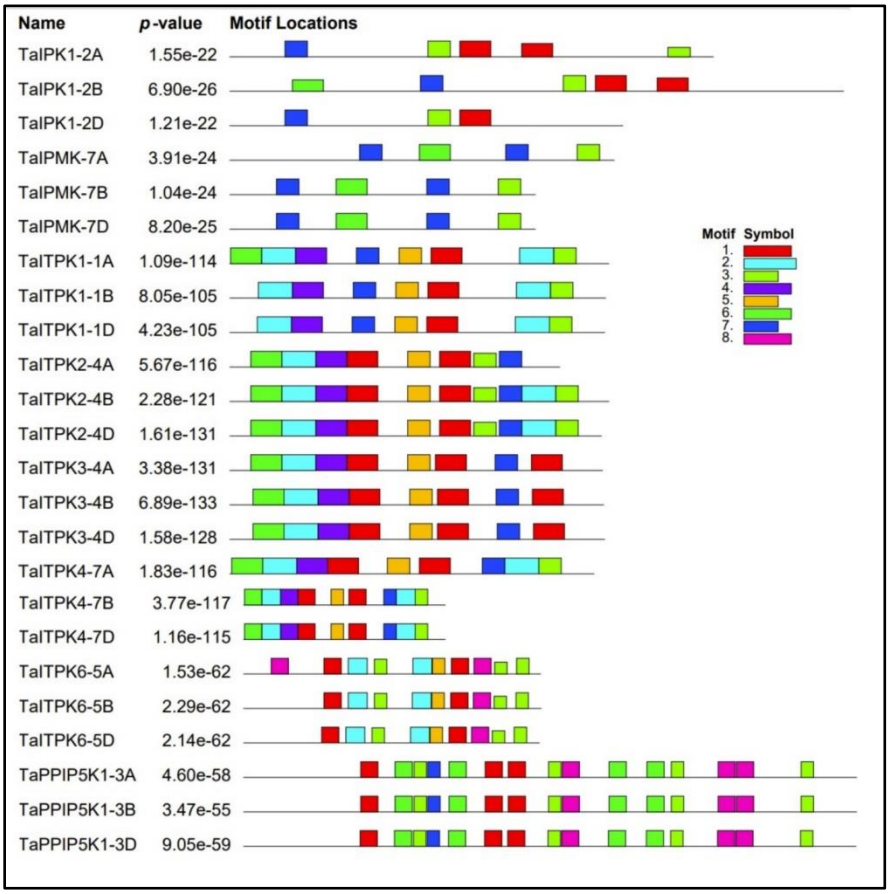

Fig 4. Conserved motifs distribution in wheat IPK gene family proteins identified by MEME. Four groups are differentiated by the motif distribution.

-acid sequence D-L-L-K-G… -G-A-C-T-W was considered to be the inositol phosphate (IP) binding site. Similarly in the C-terminus of the TaIPMK conserved region, the 9- amino acid sequence "VKLVDFAHV was considered to interact with the ATP molecule i.e., ATP binding site (Supplementary Figure 1B). All ITPKs have a similar domain structure, comprising a variable N-terminal end and a conserved catalytic C-terminal end, which is involved in the regulation and targeting of the enzymes (Erneux et al., 2016). All the TaITPK group proteins showed the conserved region in the C-terminus region containing ATP binding sites (E-I-L-A-R…..-D-M-I-R) and Inositol phosphate binding sites (P-L-V-L…-E-R-L-L) (Supplementary Figure 1D). The PPIP5K group has not yet been reported in wheat, in this study TaPPIP5K1 genes were identified based on conserved domains and motif structures. The N-terminal end of these proteins consists of the core catalytic domain containing the inositol phosphate binding site and ATP binding site (Supplementary Figure 1C). As shown in the previous studies, PPIP5K1 proteins are recognized by the PPIP5K1-specific ARK motif, which was found in all TaPPIP5K1 proteins (Randall et al., 2020). So the TaIPK genes within the same group had a similar motif construction. The consistency in motif structure, number of introns, and phylogeny shows that these genes were correctly classified.

Previous studies on the function and expression of *IPK* genes have directed that the IPK family members play vital roles in phytic acid (PA) synthesis in plants through both lipid-dependent and lipid-independent pathways. Most *ITPK* genes are involved in the lipid- independent pathway of PA synthesis while *Inositol phosphate multikinase (IPK2)* genes are reported to be involved in a lipid-dependent pathway (Laha et al., 2019); Punjabi et al., 2018). The resultant compound phytic acid plays a vital role in cell biology, on the other hand, it is a main anti-nutritional factor that reduces the bioavailable minerals in cereal grains. Various approaches have been used to target enzymes involved in the synthesis of phytic acid for enhancing nutritional value. ZFN (zinc finger nuclease) and TALEN (transcription activator-like effector nuclease) approaches were used to knockdown the *ZmIPK1* gene and phytic acid content was reduced in maize (Shukla et al., 2009; Liang et al., 2014). Similarly, many studies have reported that silencing of *AtIPK1*, *OsIPK1*, and *OsITPK1* resulted in a significant decrease in Phytic acid content (Stevenson-Paulik et al., 2005; Ali et al., 2013; Karmakar et al., 2020). Wheat *IPK1* (*TaIPK1*) gene down regulation by RNAi technology resulted in a significant increase in Fe (iron) and Zn (zinc) content by the reduction of phytic acid in wheat grains (Aggarwal et al., 2018). In our previous study, we also disrupted the *TaIPK1* gene through latest genome editing technology CRISPR-Cas9 to increase iron and zinc accumulation in wheat grains through a significant decrease in phytic acid content (Ibrahim et al., 2022). The current study analyzed the expression pattern of *TaIPK*s in different tissues at both vegetative and reproductive stages (Fig. 6). It was found that *TaIPK*s were broadly expressed at different developmental stages showing diversity in their biological functions. Some genes such as *TaIPMK-B, TaITPK2-B,*

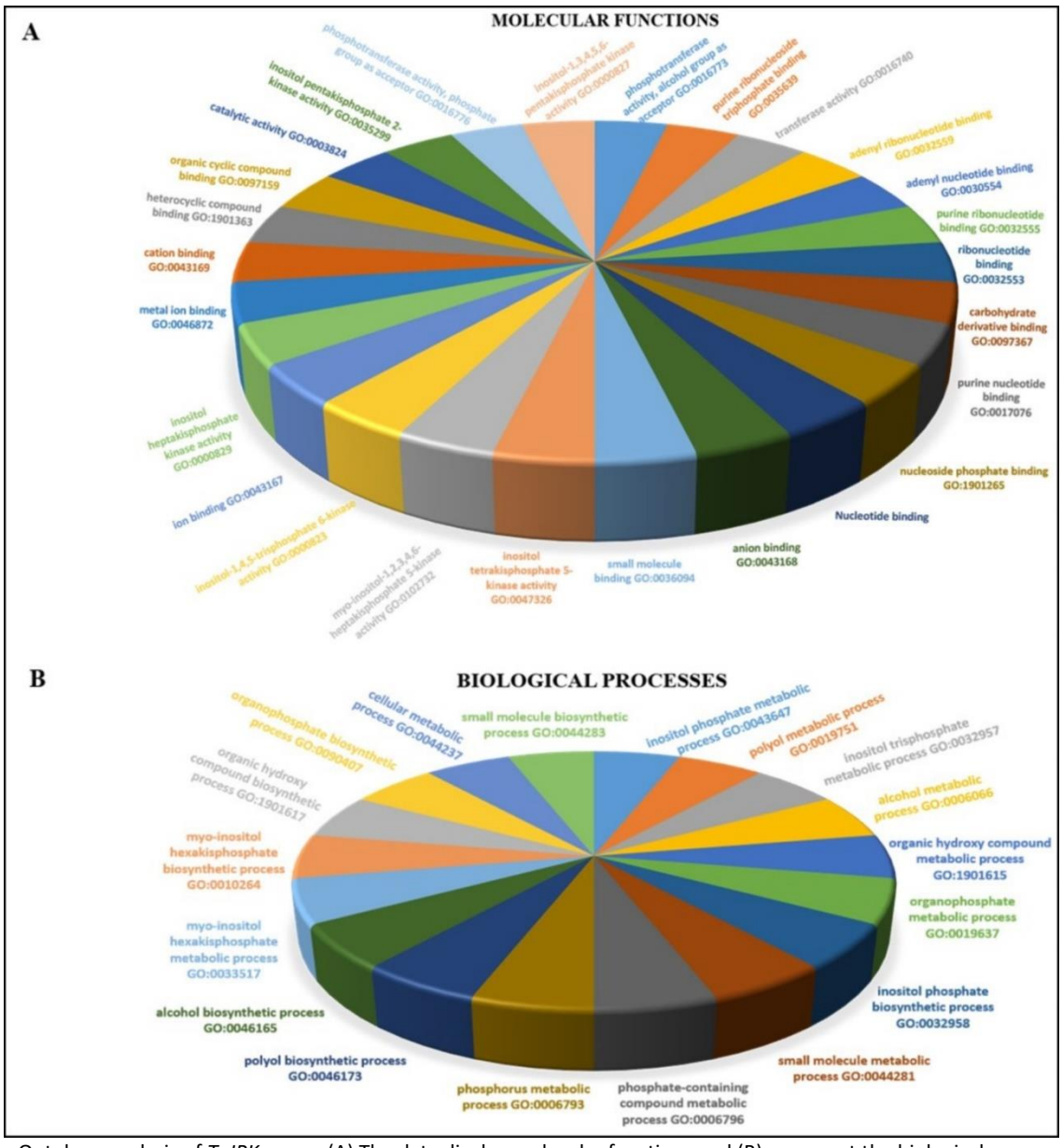

Fig 5. Gene Ontology analysis of *TaIPK* genes. (A) The data display molecular functions and (B) represent the biological processes.

TaITPK4-D, *TaITPK6-D, TaIPK1-D*, *TaITPK2- B*, *TaITPK4-A*, *B*, and *D* and *TaPPIP5K1-A, B*, and *D* had very low expression levels in the vegetative tissues showing that they might be involved in reproductive growth. The *TaITPK2-D* gene was observed to have very high expression in both vegetative and reproductive tissues, indicating that this gene might perform multiple functions in the developmental process. However, some genes like *TaIPK1-A* and *TaITPK6-A, B,* and *D* showed very high expression in grains indicating their participation in the floral organ's development. For further functional prediction of TaIPK genes, we performed the gene ontology (GO) enrichment analysis. *In silico* analysis indicated that in addition to the role in inositol phosphate signaling, *TaIPK* genes also showed many molecular functions (MFs) thus indicating their involvement in various biological processes (BPs) (Fig. 5).

Phylogenetic analysis suggested that the evolution trajectories are similar to family species (*Arabidopsis thaliana*, *Oryza sativa* and *Zea mays*), thus suggesting a single ancestor for the *IPK* gene family. Further evolutionary relationships and sequence homology were indicated by the closed clustering of homoeologous sequences. The NJ tree indicated high conservation of *IPK* genes among *T. aestivum*, *A. thaliana*, *O. sativa*, and *Z. mays.* (Fig. 7). The phylogenetic tree indicated that *IPK* genes could be categorized into four different groups depicting the sub-family classification of IPKs.

Materials and Methods

Identification and classification of IPK family genes in T. aestivum

Bhati et al., (2014) defined the conserved regions of the *TaIPK* family members. This study described 24 *IPK* family members in wheat by using those conserved regions. The candidate genes were considered to be *TaIPK*s. MEME program was used to identify the TaIPKs conserved motifs (https://meme-suite.org/meme/tools/meme) (Bailey et al., 2009) and based

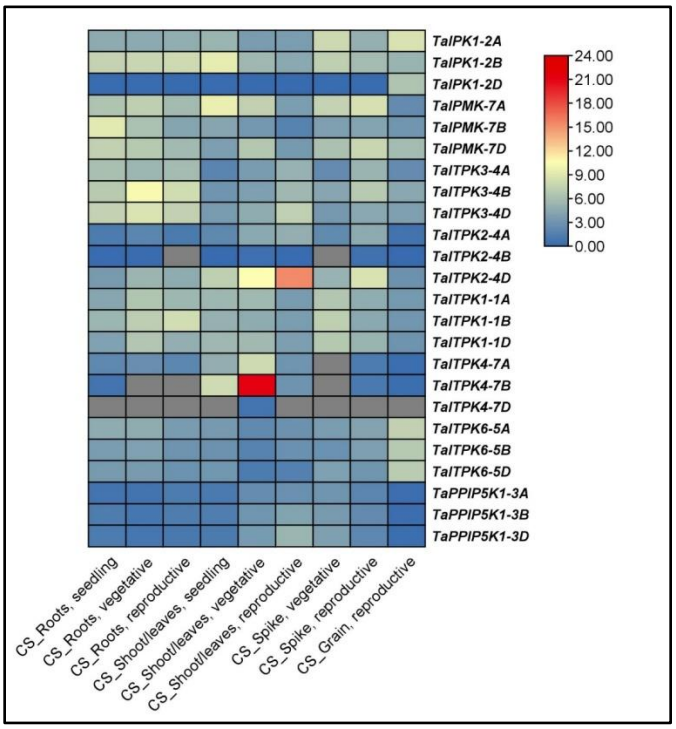

Fig 6. A heat map displaying the *in silico* expression pattern of *IPK* genes was created using TB tool. CS= Chinese Spring vegetative (root, shoot and leaf) and reproductive (spike and grain) tissues in the absence of any stress treatment or disease. The red color indicates higher expression and the green color indicates lower expression of the transcripts.

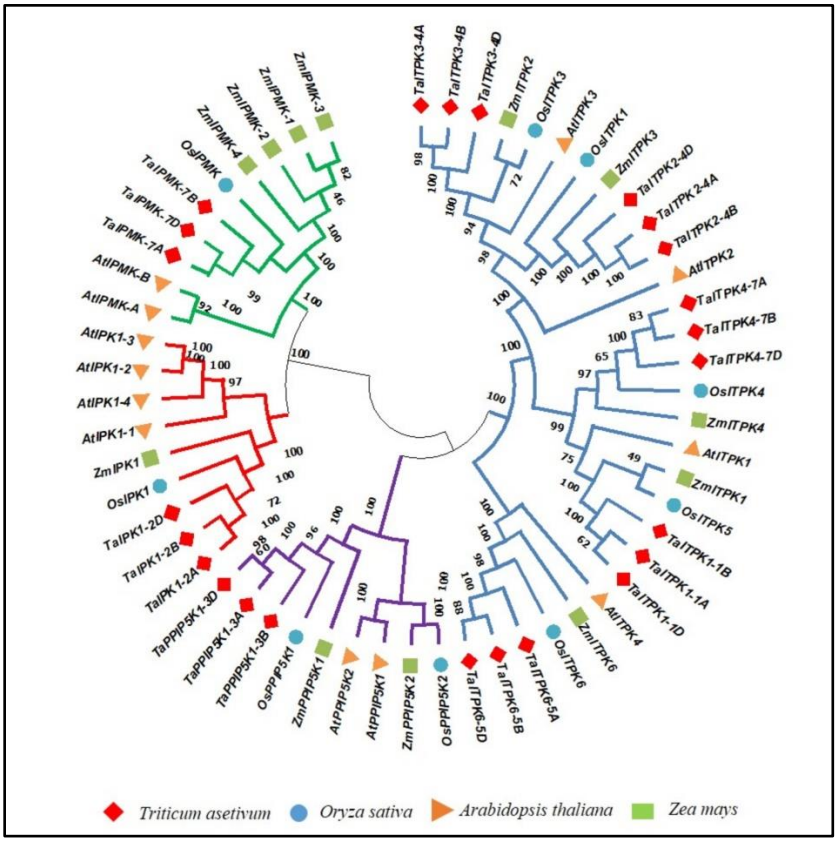

Fig 7. Phylogenetic analysis of IPK proteins from wheat, rice, maize, and Arabidopsis using neighbor-joining method. For the Bootstrap test 1,000 replicates were used, and the replication frequency is presented next to the branches. Four subfamilies are represented in different colors.

on conserved motifs these genes were classified into four subfamilies (PPIP5K, IP5-2K, ITPK1, and IPK superfamily).

TaIPKs gene structure and genome distribution

The Gene Structure Display Server 2.0 software was used to analyze the gene exons and introns distribution which is available at the following web link (https://gsds.gao-lab.org/) by following (Hu et al., 2015). The *IPK* genes locations on the genome were collected from the Ensemble Plant database (https://plants.ensembl.org). MG2C_v2.1 tool was used to display their genome distribution (http://mg2c.iask.in/mg2c_v2.1).

TaIPKs sequence alignment and phylogenetic analysis

MEGA X software was used to analyze the alignments of the IPK protein sequences. (https://www.megasoftware.net/). GeneDoc 2.7.000 software was used to mark the conserved domains (Nicholas et al., 1997). The NJ (neighbor-joining) method of MEGA X software was utilized for the construction of a phylogenetic tree of IPK family proteins (Saitou and Nei, 1987).

GO (Gene Ontology) enrichment analysis

The gene ontology (GO) analysis of *TaIPK* genes was performed through the online tool gProfiler which is available at the following web link (https://biit.cs.ut.ee/gprofiler/gost) by using default parameter values (Raudvere et al., 2019).

In silico expression analysis of IPK genes in different wheat tissues

Wheat expression database (http://www.wheat-expression.com/) was used to download RNA-seq data for the analysis of *IPK* genes expression in different tissues under different developmental (vegetative and reproductive) stages. Heat map showing the expression levels was generated using TB tool (Chen et al., 2018).

Phylogenetic analysis of IPK proteins between T. Aestivum, O. sativa, Z. mays, and A. thaliana

The *O. sativa, Z. mays and A. thaliana* IPK gene family sequences and the resultant protein sequences were retrieved from the Ensemble Plant database. Arabidopsis IPK proteins were designated as AtIPK1s, AtIPMKs, AtPPIP5Ks and AtITPKs. Similarly rice and maize IPK proteins were designated as OsIPK1s, OsIPMKs, OsPPIP5Ks, OsITPKs and ZmIPK1s, ZmIPMKs, ZmPPIP5Ks and ZmIPMKs respectively. For observing the evolutionary relationships, a Neighbor-Joining (NJ) phylogenetic tree was generated based on the 57 IPK protein sequences from wheat, rice, maize, and Arabidopsis.

Conclusions

In this study, the *IPK* gene family comprising 24 members was identified in the *T. aestivum* genome. These genes were classified into four subgroups which is consistent with the sub-family classification of *IPK*s. Chromosomal locations, distributions of conserved motifs and intron/exon structures of *IPK* family members in wheat were also determined. The results of the current study showed that *IPK* genes are highly conserved among wheat, Arabidopsis, and other cereals like rice and maize. The *IPK* genes showed different patterns of expression in different tissues at different developmental stages showing diversity in their biological functions. Since most of these genes are involved in phytic acid (PA) synthesis, this detailed study provides a foundation for researchers to understand these genes and target the good candidate genes to attain low phytate (lpa) crops by exploiting the latest biotechnological tools like genome editing. Therefore, it is clearing the way toward nutritional improvement of common wheat (*T. aestivum*) and other staple food crops.

Acknowledgements

We acknowledge the research facilities provided by the Quaid-e-Azam University and National Institute for Genomics and Advanced Biotechnology (NIGAB), Pakistan. We also thank the NCBI, Ensembl plants, and various other tools for the availability of data and sequences.

Author's contribution

MRK and SI designed the study. SI and TA performed the experiments, data collection, data analysis, and writing of the original draft. MRK and MU provided technical expertise to improve and help with revision of the article. MRK supervised the research. All authors review and edit the manuscript.

References

Agarwal R, Mumtaz H, Ali N (2009) Role of inositol polyphosphates in programmed cell death. Mol Cell Biochem. 328:155-165.

Aggarwal S, Kumar A, Bhati KK, Kaur G, Shukla V, Tiwari S, Pandey AK (2018) RNAi-mediated downregulation of inositol pentakisphosphate kinase (IPK1) in wheat grains decreases phytic acid levels and increases Fe and Zn accumulation. Front Plant Sci. *9*:259.

Ali N, Paul S, Gayen D, Sarkar SN, Datta K, Datta SK (2013a). Development of low phytate rice by RNAi mediated seed-specific silencing of inositol1,3,4,5,6-pentakisphosphate2-kinasegene(IPPK1). PLoS One. 8(7):e68161.

Balk J, Connorton JM, Wan Y, Lovegrove A, Moore KL, Uauy C, Shewry PR (2019) Improving wheat as a source of iron and zinc for global nutrition. Nutr Bull. 44(1):53-59.

Bhati KK, Aggarwal S, Sharma S, Mantri S, Singh SP, Bhalla S, Pandey AK (2014) Differential expression of structural genes for the late phase of phytic acid biosynthesis in developing seeds of wheat (Triticum aestivum L.). Plant Sci. 224:74-85.

Bloot AP, Kalschne DL, Amaral JA, Baraldi IJ, Canan C (2023) A review of phytic acid sources, obtention, and applications. Food Rev Int. 39(1):73-92.

Brouns F (2021) Phytic acid and whole grains for health controversy. Nutrients. 14(1):25.

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009) MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 37:W202-W208

Chen C, Chen H, He Y, Xia R (2018) TB tools, a toolkit for biologists integrating various biological data handling tools with a user-friendly interface. BioRxiv. 10(1101):289660

Erneux C, Ghosh S, Koenig S (2016) Inositol (1, 4, 5) P3 3-kinase isoenzymes: catalytic properties and importance of targeting

to F-actin to understand function. Adv Biol Regul. 1(60):135-143.

Gonzalez B, Banos-Sanz JI, Villate M, Brearley CA, Sanz-Aparicio J (2010) Inositol 1,3,4,5,6-pentakisphosphate 2-kinase is a distant IPK member with a singular inositide binding site for axial 2-OH recognition, Proc Natl Acad Sci U.S.A. 107:9608-9613.

Gosein V, & Miller, GJ (2013) Roles of phosphate recognition in inositol 1,3,4,5,6-pentakisphosphate 2-kinase (ipk1) substrate binding and activation. J Biol Chem. 288:26908-26913.

Gupta, Raj K, Shivraj SG, Nand KS (2015) "Reduction of phytic acid and enhancement of bioavailable micronutrients in food grains." J Food Sci Technol. 52(2):676-684.

Hatch AJ, York JD (2010) Snapshot: inositol phosphates. Cell. 143 (6):1030

Hu B, Jin J, Guo A, Zhang H, Luo J, Gao G (2015) GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics. 31(8):1296-1297.

Ibrahim S, Saleem B, Rehman N, Zafar SA, Naeem MK, Khan MR (2022) CRISPR/Cas9 mediated disruption of Inositol Pentakisphosphate 2-Kinase 1 (TaIPK1) reduces phytic acid and improves iron and zinc accumulation in wheat grains. J Adv Res. 37:33-41.

Karmakar A, Bhattacharya S, Sengupta S, Ali N, Sarkar SN, Datta K, Datta SK (2020) RNAi-mediated silencing of ITPK gene reduces phytic acid content, alters transcripts of phytic acid biosynthetic genes, and modulates mineral distribution in rice seeds. Rice Sci. 27(4):315-328.

Khalid A, Hameed A, Tahir MF (2023). Wheat quality: A review on chemical composition, nutritional attributes, grain anatomy, types, classification, and function of seed storage proteins in bread making quality. Front Nutr. 10:1053196.

Laha D, Parvin N, Hofer A, Giehl RF, Fernandez-Rebollo N, von Wirén N, Saiardi A, Jessen HJ, Schaaf G (2019) Arabidopsis ITPK1 and ITPK2 have an evolutionarily conserved phytic acid kinase activity. ACS chem biol. 14(10):2127-2133.

Laha D, Portela-Torres P, Desfougères Y, Saiardi A (2021). Inositol phosphate kinases in the eukaryote landscape. Adv Biol Regul. 79:100782.

Marcussen T, Sandve SR, Heier L, Spannagl M, Pfeifer, M, Jakobsen KS et al (2014) International Wheat Genome Sequencing Consortium. Ancient Hybridizations Among the Ancestral Genomes of Bread Wheat. Science. 345(6194):1250092.

Nicholas KB, Nicholas HB, Deerfield DW, Nicholas HBJ, Nicholas K, Nicholas HJ, Nicholas KR, Nicholas HBJ, Nicholas A, Deerfield DW, Nicholas H, Gauch H (1997) GeneDoc: a tool for editing and annotating multiple sequence alignments. EMBnet news.

Perera I, Seneweera S, Hirotsu N (2018). Manipulating the phytic acid content of rice grain toward improving micronutrient bioavailability. Rice. 11:1-13.

Punjabi M, Bharadvaja N, Sachdev A, Krishnan V (2018) Molecular characterization, modeling, and docking analysis of late phytic acid biosynthesis pathway gene, inositol polyphosphate 6-/3-/5-kinase, a potential candidate for developing low phytate crops. Biotech. 8:1-20.

Raboy V (2009) Approaches and challenges to engineering see phytate and total phosphorus, Plant Sci. 79:281-296.

Randall TA, Gu C, Li X, Wang H, Shears SB (2020) A two-way switch for inositol pyrophosphate signaling: Evolutionary history and biological significance of a unique, bifunctional kinase/phosphatase. Adv Biol Regul. 75:100674.

Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, et al (2019). g:Profiler: a Web Server for Functional Enrichment Analysis and Conversions of Gene Lists (2019 Update). Nucleic acids Res. 47:W191-W198.

Saitou N, Nei M (1987) The Neighbor-Joining Method-a New Method for Reconstructing Phylogenetic Trees. Mol Biol Evol. 4(4):406-425.

Sharma S, AnandKumar LHD, Tyagi A, Muthumilarasan M, Kumar K, Gaikwad K (2021). An insight into phytic acid biosynthesis and its reduction strategies to improve mineral bioavailability. The Nucleus. 1-13.

Shears SB, Baughman BM, Gu C, Nair VS, Wang H (2017) The significance of the 1-kinase/1-phosphatase activities of the PPIP5K family Adv Biol Regul. 63:98-106

Shears, SB, Wang H (2019) Inositol phosphate kinases: Expanding the biological significance of the universal core of the protein kinase fold. Adv Biol Regul. 71:118-127.

Shewry, Peter R, Sandra JH (2015) The contribution of wheat to human diet and health. Food Energy Secur. 4(3):178-202.

Shi J, Wang H, Wu Y, Hazebroek J, Meeley RB, Ertl DS (2003) The maize low-phytic acid mutant lpa2 is caused by mutation in an inositol phosphate kinase gene. Plant Physiol. 131:507-515

Shukla VK, Doyon Y, Miller JC, DeKelver RC, Moehle EA, Worden SE, Mitchell JC, Arnold NL, Gopalan S, Meng X, Choi VM (2009) Precise genome modification in the crop species Zea mays using zinc-finger nucleases. Nature. 459(7245):437-441.

Steger DJ, Haswell ES, Miller AL, Wente SR, O'Shea EK (2003) Regulation of chromatin remodeling by inositol polyphosphates. Science. 299:114-116.

Stevenson-Paulik J, Bastidas RJ, Chiou ST, Frye RA, York JD (2005) Generation of phytate-free seeds in Arabidopsis through disruption of inositol polyphosphatekinases. Proc Natl Acad Sci U.S.A.102:12612-12617.

Suzuki M, Tanaka K, Kuwano M, Yoshida KT (2007) Expression pattern of inositol phosphate-related enzymes in rice (Oryza sativa L.): implications for the phytic acid biosynthetic pathway. Gene. 405:55-64.

Tsui MM, York, JD (2010) Roles of inositol phosphates and inositol pyrophosphates in development, cell signaling and nuclear processes. Adv in enzyme regul. 50(1):324-337.

Wilson MS, Livermore TM, Saiardi A (2013) Inositol pyrophosphates: Between signaling and metabolism. Biochem J. 452:369-379.