# Novel allelic variant of *Lpa1* gene associated with a significant reduction in seed phytic acid content in rice (*Oryza sativa* L.)

**D.S. Kishor[1], Choonseok Lee[1], Dongryung Lee[1], Jelli Venkatesh[1], Jeonghwan Seo[1], Joong Hyoun Chin[2], Zhuo Jin[1], Soon-Kwan Hong[3], Jin-Kwan Ham[4], Hee-Jong Koh[1*]**

[1] Department of Plant Science, Plant Genomics and Breeding Institute, and Research Institute of Agriculture and Life Science, Seoul National University, Seoul, South Korea

[2] Graduate School of Integrated Bioindustry, Sejong University, Republic of Korea

[3] Division of Biotechnology, Kangwon National University, Republic of Korea

[4] Gangwon provincial Agricultural Research & Extension Services, Republic of Korea

[*] Corresponding author

Email: heejkoh@snu.ac.kr (HJK)

**Data availability statement:** All relevant data are within the paper and in the NCBI Genbank database (accession number: MH707666).

## Abstract

In plants, *myo*-inositol-1,2,3,4,5,6-hexa*kis*phosphate ($InsP_6$), also known as phytic acid (PA), is a major component of organic phosphorus (P), and accounts for up to 85% of the total

P in seeds. In rice (*Oryza sativa* L.), PA mainly accumulates in rice bran, and chelates mineral cations, resulting in mineral deficiencies among brown rice consumers. Therefore, considerable efforts have been focused on the development of low PA (LPA) rice cultivars. In this study, we performed genetic and molecular analyses of *OsLpa*1, a major PA biosynthesis gene, in Sanggol, a low PA mutant variety developed via chemical mutagenesis of Ilpum rice cultivar. Genetic segregation and sequencing analyses revealed that a recessive allele, *lpa1-3*, at the *OsLpa*1 locus (Os02g0819400) was responsible for a significant reduction in seed PA content in Sanggol. The *lpa1-3* gene harboured a point mutation (C623T) in the fourth exon of the predicted coding region, resulting in threonine (Thr) to isoleucine (Ile) amino acid substitution at position 208 (Thr208Ile). Three-dimensional analysis of Lpa1 protein structure indicated that *myo*-inositol 3-monophosphate [Ins(3)$P_1$] kinase binds to the active site of Lpa1, with ATP as a cofactor for catalysis. Furthermore, the presence of Thr208 in the loop adjacent to the entry site of the binding pocket suggests that Thr208Ile substitution is involved in regulating enzyme activity via phosphorylation. Therefore, we propose that Thr208Ile substitution in *lpa1-3* reduces Lpa1 enzyme activity in Sanggol, resulting in reduced PA biosynthesis.

# Introduction

In most cereal crops, *myo*-inositol-1,2,3,4,5,6-hexa*kis*phosphate ($InsP_6$), also known as phytic acid (PA), is considered a major source of phosphorus (P) available in the form of phytate, and accounts for 65–85% of the total P in seeds [1]. Monogastric animals poorly digest PA, as they lack the phytase enzyme, which is responsible for the release of phosphate residues [2]. PA is an efficient chelator of mineral cations, such as zinc ($Zn^{2+}$), iron ($Fe^{2+}$), magnesium ($Mg^{2+}$), potassium ($K^{2+}$), and calcium ($Ca^{2+}$), in the nutritional tract. Because of these attributes,

PA is considered as an antinutrient [3, 4]. Hence, there is a need to develop low PA (LPA) crop cultivars to maximize the nutritional benefits of grains.

Mutants associated with the LPA phenotype have been identified in several crop plants including maize (*Zea mays*) [5, 6], barley (*Hordeum vulgare*) [7], soyabean (*Glycine max*) [8], rice (*Oryza sativa*) [9], and wheat (*Triticum aestivum*) [10]. Although, LPA mutants are identified primarily on the basis of percentage reduction of PA and high inorganic P ($P_i$) content in seeds [5, 11], some mutants show a significant accumulation of *myo*-inositol and inositol phosphate [Ins(1,3,4)$P_3$ 5-/6] intermediates in seeds [12, 13].

Previously, the LPA phenotype of seeds has been associated with reduced agronomic performance of mutant crop plants in the field [5, 14]. It is important to understand the genetic and molecular bases of reduced agronomic performance of LPA mutants for effective utilization in breeding programs. In addition, studies show that climate change and elevated carbon dioxide ($CO_2$) levels negatively affect micronutrient bioavailability and total P in grains [15, 16]. Therefore, developing crop cultivars with increased micronutrient bioavailability in seeds and greater adaptability to environmental variations, by reducing the PA content in grains, is an important priority of breeding programs.

PA is biosynthesized via two different routes: lipid dependent and lipid independent [3, 17]. The lipid dependent pathway operates in all plant organs, whereas the lipid independent pathway is predominant only in seeds [13, 17, 18]. In the first step of PA biosynthesis, D-glucose-6-phosphate is converted to *myo*-inositol 3-monophosphate [Ins(3)$P_1$] by *myo*-inositol 3-phosphate synthase (MIPS) [19]. This is followed by the sequential phosphorylation of specific inositol to $InsP_6$ through the action of various inositol phosphate kinases (S1 Fig). However, enzymes involved in lipid independent PA biosynthesis, from Ins(3)$P_1$ seem to be complicated and are not well understood [3]. Nevertheless, PA biosynthetic genes encoding other *myo*-inositol enzyme and inositol phosphate kinases are well documented in major plants

[12, 13, 20, 21]. Additionally, biochemical and functional analyses of PA biosynthetic genes encoding Ins monophosphate kinase could address the missing steps in the lipid independent pathway.

In rice, several mutants with low seed PA content have been reported [14, 21-27]. Genetic studies of LPA mutants have shown that a single recessive gene is responsible for the LPA phenotype in rice and other crop plants [21, 22, 27, 28]. The first *lpa* gene encoding inositol 1,3,4-triskisphophate 5/6-kinase (ITPK5/6) was identified in maize, and designated as *Lpa2*. Subsequently, *myo*-inositol kinase gene *Lpa3*, and multidrug resistance protein (MRP) ATP binding cassette (ABC) transporter gene *Lpa1* were identified [12, 13, 29]. In addition, reduction of PA content in *Arabidopsis atipk2β* mutant indicates the inositol 1,4,5-tris-phosphate (IPK2) kinase of lipid dependent pathway is also active the seeds [20]. In rice, *OsLpa*1 gene has been associated with the reduction in seed PA content and increase in seed $P_i$ content, with little change in the total P content in seeds [22, 30]. *OsLpa*1 have homology with one gene, Os09g0572200 (*OsLpa*1 paralog) within the rice genome, suggests possible overlapping or redundant functions [22].

Genetic studies in rice have confirmed that a mutation in the *OsLpa*1 locus generates the LPA phenotype in seeds. Molecular characterization of LPA mutants has previously revealed three alleles of the *OsLpa*1 locus, including KBNT *lpa 1-1*, DR1331-2, and Os-lpa-XQZ-1, responsible for the low PA phenotype of seeds [22, 30]. In the present study, we report a novel allele of *OsLpa*1, *OsLpa1-3*, responsible for a significant reduction in the seed PA content in a new LPA mutant rice cultivar Sanggol developed in the Republic of Korea [31]. Sequence analysis of *OsLpa*1-3 revealed a point mutation in the gene coding sequence. Our data suggest that this mutation is responsible for the LPA phenotype of Sanggol mutant.

# Material and methods

# Plant material

The low PA mutant rice cultivar Sanggol derived from a *japonica* rice cultivar Ilpum mutagenized with *N*-methyl-*N*-nitrosourea (MNU) [31]. Ilpum was used as the wild type in comparing phenotypic data. Sanggol was crossed with Ilpum to develop $F_2$ population. Segregation analysis was performed using the $F_2$ population. Both parent cultivars and $F_2$ individuals were grown in experimental fields of Seoul National University, Republic of Korea.

# Agronomic trait analysis

To characterize agronomic traits, 15 phenotypic observations were recorded during various stages of plant growth, according to the Standard Evaluation System (SES) for rice, 2014. Yield data was obtained from "3.6 m X 3.6 m" plot size. All agronomic data were analyzed using the Student's *t*-test in SPSS 16.0 (https://www.ibm.com/analytics/spss-statistics-software) to determine significant differences among Sanggol, and Ilpum cultivars.

# Analysis of $P_i$ and PA content in seeds

Concentrations of $P_i$ and PA in seeds were examined using $P^{31}$ nuclear magnetic resonance ($P^{31}$ NMR) spectroscopy [32], with slight modifications.

## Sample preparation

Fine powdered samples (1 g dry weight) of brown rice were thoroughly mixed with 10 mL of 2.4% HCl in 14 mL Falcon tubes. Samples were incubated at room temperature for 16 h on an HB-201SF shaker (HANBAEK Scientific Co) at 220 rpm, and subsequently centrifuged at 1,500 × *g* (combi 514R, Hanil science Inc.) at 10°C for 20 min. Crude extracts were transferred to a new 14 mL Falcon tube containing 1 g NaCl, and incubated at 25°C for 40 min on a shaker at 220 rpm to dissolve NaCl. Samples were allowed to settle at 4°C for 60 min, and then centrifuged at 1,500 × *g* at 10°C for 20 min.

## $^{31}$P NMR

For $^{31}P$ NMR spectroscopy, samples were prepared by mixing 450 µL of NaCl treated acid extract with 450 µl of buffer containing 0.11mM EDTA-disodium salt and 0.75 mM NaOH, 40 mg NaOH, and 100 µL $D_2O$ in 1.5 mL microtubes. Sample and standard peaks were obtained on a 600 MHz spectrometer using Advance 600 $^{31}P$ NMR system (Bruker, Germany). PA sodium salt and 85% phosphoric acid were used as external standards for peak identification and further analysis [33, 34]. For internal calibration, 1 mM of phenylphosphonic acid was included in 100 µL $D_2O$ during NMR measurements. All standards were purchased from Sigma-Aldrich, USA.

To determine significant differences in seed PA and $P_i$ contents among parents and $F_2$ individuals, data were analyzed using the Student's *t*-test in SPSS 16.0 (https://www.ibm.com/analytics/spss-statistics-software).

## Expression analysis of PA biosynthetic genes

Genes involved in PA biosynthesis and transport were identified from the RAB-DB and from recent studies [25, 35, 36]. The rice microarray database, RiceX-Pro, shows different expression patterns of most of the PA biosynthetic genes in various tissues and organs [37]. To confirm the expression pattern of PA biosynthetic genes, spikelets were harvested from the wild cultivar Ilpum at 5 days after flowering (DAF), and total RNA was extracted using RNAiso Plus (Takara Bio, Japan). The extracted RNA samples were treated with RNase-free recombinant DNase I (Takara Bio, Japan) to eliminate genomic DNA contamination, and first-strand cDNA was synthesized using M-MLV reverse transcriptase (Promega, USA). The PA biosynthetic genes (200–550bp) were amplified from cDNA samples by reverse transcription polymerase chain reaction (RT-PCR) using gene-specific primers (Table 1) with the following conditions: initial denaturation at 95°C for 2 min, followed by 32 cycles of denaturation at 95°C for 20 s, annealing at 58°C for 40 s, and extension at 72°C for 1 min, and a final extension at 72°C for 5 min. The *Actin* gene was used as an internal control.

146

147 **Table 1. RT-PCR primers used to amplify PA biosynthetic genes.**

| Gene ID | Primer name | Sequence (5'→3') |
|---|---|---|
| Os03g0192700 | *OsRINO1F* | AGTGGACAAGGTGGTGGTGT |
| | *OsRINO1R* | ATCACCACCAATCAGGCAGT |
| Os10g0369900 | *OsRINO2F* | GAAGAGCAAGGTGGACAAGG |
| | *OsRINO2R* | CATCTTGGTCTGCCCACTCT |
| Os03g0587000 | *OsIMP1F* | GTGGATTTGGTGACGGAGAC |
| | *OsIMP1R* | ATCGAGACGCAAACAAAAGG |
| Os02g0169900 | *OsIMP2F* | CCTCTTCACACCGCAGGAAT |
| | *OsIMP2R* | CTGGATGACGCCGAGGAG |
| Os07g0507300 | *OsMIK1F* | TCTACTGGGACGGTGGAGAG |
| | *OsMIK1R* | TAGCCGCTTCTTGGAGTGAT |
| Os02g0819400 | *OsLPA1F* | TATGTGGGACTAGCGGATGC |
| | *OsLPA1R* | GAGCAACTGCAACAGGGTCT |
| Os09g0572200 | *OsLPA1-P F* | CGGCTGATGTTCCACCTAAT |
| | *OsLPA1-P R* | TTGACGCTTTCTCAATGTGC |
| Os10g0103800 | *OsITPK1F* | ACAAGGAGTGGCAGCAAGTT |
| | *OsITPK1R* | CAACCAAGGGCAACGTTAGT |
| Os03g0230500 | *OsITPK2F* | TCTGGTCCTCCAGGAATTTG |
| | *OsITPK2R* | CCAGTCTTCCACGAAGCTCT |
| Os03g0726200 | *OsITPK3F* | AGGGAGGAACACCCAGAAGT |
| | *OsITPK3R* | ACCAGAGGCTTTGCCACTAA |
| Os02g0466400 | *OsITPK4F* | ACATGCGCCTCGTCTACC |
| | *OsITPK4R* | GTTGGAGATGTTGGCGAAG |
| Os10g0576100 | *OsITPK5F* | CCAGCTCCTCAAAGTCTGCT |
| | *OsITPK5R* | TTTGTCCATGCTCCTTCTCA |
| Os09g0518700 | *OsITPK6F* | GCAAAACGAGGTGCAAGATA |
| | *OsITPK6R* | GCTTGATTGCATCCCAGAAT |
| Os04g0661200 | *OsIPK1F* | CAACCGGCACCAAACTGTAT |
| | *OsIPK1R* | CAGAATCAGCTCCAGCATCA |
| Os02g0596100 | *OsGLE1F* | AGACCGCGTCTTGTCTGC |
| | *OsGLE1R* | GTCGAGCTCGGTGAGGAC |
| Os02g0523800 | *OsIPK2F* | CTCTTCTACAAGCCCCTCCA |
| | *OsIPK2R* | GAGGCACTTGGCGACGTA |
| Os03g0142800 | *OsMRP13F* | GCTTATTGCATTGGGTAGGG |
| | *OsMRP13R* | TTACCCGAAGCTCTGATGCT |
| Os03g0836000 | *Actin 4* | AGGCAGTCAGTCAGATCACGA |
| | *Actin 5* | GAGACATTCAATGCACCAGCA |

148

149 # Sequence analysis

Genomic DNA and cDNA were isolated from young leaves and spikelets of Sanggol low PA mutant cultivar, respectively. Fragments of size 300bp-2000bp were amplified from the coding region and untranslated region (UTR) of 16 genes of lipid dependent and independent pathways using gene-specific primers designed with prime3 (http://bioinfo.ut.ee/primer3-0.4.0/) (S1 Table). The PCR products were purified using the DNA Purification Kit (Inclone, Korea), and analyzed with an ABI Prism 3730 XL DNA Analyzer (PE Applied Biosystems, USA). In addition, sequences of all 16 genes in Ilpum was downloaded from the crop molecular breeding lab server (http://nature.snu.ac.kr/rice/). Sequences were aligned using the Codon Code Aligner software (Codon Code Corporation, USA).

## Candidate gene analysis

To confirm nucleotide polymorphisms in the candidate genes, validation primers were designed using Primer3 for cDNA sequencing (Table 2). The PCR products were purified using the DNA Purification Kit (Inclone, Korea), and analyzed with an ABI Prism 3730 XL DNA Analyzer (PE Applied Biosystems, USA). Sequences were aligned using the Codon Code Aligner software (Codon Code Corporation, USA). Simultaneously, BLAST search was performed using the predicted amino acid sequences of the candidate genes in the NCBI database (https://blast.ncbi.nlm.nih.gov/Blast.cgi), and deleterious amino acid substitutions were predicted using Provean web server with proven scores [38].

**Table 2. Primers used for validating cDNA sequences and genotyping the $F_2$ population.**

| Analysis | Primer name | Sequence (5'→3') | Amplicon size |
|---|---|---|---|
| cDNA validation | *Lpa1-3 F* | GCCATGCCTTCAAGATTAGC | 1,186 bp |
| | *Lpa1-3 R* | TGAAACATTCCCTTGGAACC | |
| Genotyping | *Lpa1-3-1F* | AGCATTCGCCTGCATGATCG | Homozygous wild-type (192 bp), homozygous mutant (174 bp), heterozygous (both 192 bp and 174 bp) |
| | *Lpa1-3-1R* | CGCTTACCGAACAATGAATG | |

## Expression analysis of *Lpa* and lipid dependent PA biosynthesis genes in Sanggol and Ilpum cultivars.

Total RNA was extracted from the leaves at 15 days after germination (DAG) to analyze the expression of *Lpa* and lipid dependent pathway genes, and 5 DAF from spikelets to analyze the expression of *OsLpa*1 in Sanggol and Ilpum cultivars. For the expression analysis of *OsLpa*1 paralog and *OsIpk*2 genes, total RNA was extracted only from spikelets at 5 DAF. RNA extraction was performed as described above. The extracted RNA was subjected to RT-PCR using gene-specific primers (Table 3). The *Actin* gene was used as an internal control.

183

184 **Table 3. RT-PCR primers used to amplify *Lpa* and *Ipk2* genes.**

| Gene ID | Transcript | Primer name | Primer sequence (5'→3') | Amplicon size |
|---|---|---|---|---|
| Os02g0819400 | *OsLpa*1.1 | *lpa1.1 F* | TATGTGGGACTAGCGGATGC | 192 bp |
| | | *1pa1.1 R* | GAGCAACTGCAACAGGGTCT | |
| | *OsLpa*1.2 | *lpa1.2 F* | TATGTGGGACTAGCGGATGC | 445 bp |
| | | *lpa1.2 R* | GAGCAACTGCAACAGGGTCT | |
| | *OsLpa*1.3 | *lpa1.3 F* | ATCTTTCGGGATCAGTGCAT | 200 bp |
| | | *lpa1.3 R* | TGGCAGCATGTTTCTCCTATC | |
| Os09g0572200 | *OsLpa*1 paralog | *OsLPA2F* | CGGCTGATGTTCCACCTAAT | 236 bp |
| | | *OsLPA2R* | TTGACGCTTTCTCAATGTGC | |
| Os02g0523800 | *OsIpk*2-0 | *OsIPK2F* | CTCTTCTACAAGCCCCTCCA | 291 bp |
| | | *OsIPK2R* | GAGGCACTTGGCGACGTA | |
| Os03g0836000 | *OsActin* | *Actin 4* | AGGCAGTCAGTCAGATCACGA | 194 bp |
| | | *Actin 5* | GAGACATTCAATGCACCAGCA | |

185

186

## Derived cleaved amplified polymorphic sequence (dCAPS) analysis

Genomic DNAs were isolated from all 96 $F_2$ plants derived from cross between Sanggol and Ilpum, and subjected to dCAPS analysis. The $F_2$ genotyping primers (Table 2) were designed using dCAPS 2.0 (http://helix.wustl.edu/dcaps/) to validate a single nucleotide substitution (C to T) in the *OsLpa*1 gene in Sanggol cultivar, which generates a *Taq*I restriction site (TCGA) in the amplified PCR product. PCR was performed using the following conditions: initial denaturation at 95°C for 2 min, followed by 32 cycles of denaturation at 95°C for 20 s, annealing at 58°C for 40 s, and extension at 72°C for 30 s, and a final extension at 72°C for 1 min. The amplified PCR product was digested with *Taq*I restriction endonuclease (Promega, USA), and separated on 3% agarose gel.

## Multiple sequence alignment and phylogenetic analysis

Amino acid sequences of the Lpa superfamily were obtained from the NCBI protein database (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins), and subjected to multiple sequence alignment using the Clustal Omega program (https://www.ebi.ac.uk/Tools/msa/clustalo/). Multiple sequence alignment editing, visualization, and analysis was performed using Jalview 2.10.4 (http://www.jalview.org/). The Lpa and other superfamily proteins obtained from the NCBI protein database were used for phylogenetic analysis. Neighbour-joining tree was constructed using MEGA 7 [39] with 1,000 bootstrap replicates.

## Biocomputational analysis

A three-dimensional (3D) model of Lpa1 protein was produced under the intensive mode of the Phyre2 server [40] (www.sbg.bio.ic.ac.uk/phyre2/html/). The ligand and cofactor were downloaded from the PubChem database (https://pubchem.ncbi.nlm.nih.gov/)

211 for protein ligand analysis. Furthermore, auto docking and 3D model were analyzed using
212 the CLC drug discovery workbench 4.0 (QIAGEN, Denmark). Putative phosphorylation sites
213 were predicted with the GPS 3.0 server (http://gps.biocuckoo.org/) using high cut-off values
214 ranging from 1.36 to 17.72.
215

# 216 Results

## 217 Agronomic characterization of Sanggol low PA mutant and Ilpum
## 218 cultivars

219 Analysis of agronomic traits demonstrated a significant reduction in the plant height
220 (cm), number of productive tillers, culm length (cm), first intermodal length (cm), 1,000-grain
221 weight (g), number of spikelets per panicle, number of panicles per plant, and yield components
222 of Sanggol compared with Ilpum (Table 4 and Fig 1A). By contrast, the number of days to 50%
223 flowering was significantly higher in Sanggol than in the Ilpum, indicating delayed flowering
224 in the mutant cultivar. In addition, Sanggol exhibited significantly higher percentage of chalky
225 grains compared with the wild cultivar. However, no significant differences were observed
226 between the two cultivars in morphological characteristics, such as secondary internodal length,
227 grain length, grain width, panicle length, and spikelet fertility (Fig 1B and Fig 1C). Overall,
228 these data indicate that Sanggol low PA mutant shows poor agronomic performance with
229 respect to the flowering time, yield, and yield components compared with the Ilpum.
230

**Table 4. Agronomic traits of the wild-type cultivar Ilpum and low PA mutant cultivar Sanggol.**

| Traits[a] | PH (cm) | NPT | DFF | CL (cm) | FIL (cm) | SIL (cm) | GL (mm) | GW (mm) | KGW (g) | GC (%) | NSP | NPP | PL (cm) | SPF (%) | YPP (g) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Wild | 97 | 22 | 116.7 | 75 | 40.5 | 15 | 5.01 | 2.97 | 20.93 | 43.62 | 187 | 12 | 22 | 97 | 760 |
| Mutant | 77 | 13 | 122.5 | 54 | 35.5 | 14.83 | 4.81 | 3.03 | 18.9 | 54.66 | 165 | 8 | 23 | 96 | 470 |
| Significance[b] | ** | ** | ** | ** | ** | NS | NS | NS | ** | ** | ** | * | NS | NS | ** |

[a]Traits: PH, plant height; NPT, number of productive tillers; DFF, days to 50% flowering; CL, culm length; FIL, first internodal length; SIL, second internodal length; GL, grain length; GW, grain width; KGW, 1,000 grain weight; GC, grain chalkiness; NSP, number of spikelets per panicle; NPP, number of panicles per plant; PL, panicle length; SPF, spikelet fertility; YPP, yield per plot

[b]Asterisks indicate the level of significance (*, $P < 0.05$; **, $P < 0.01$). NS, non-significant.

238 **Fig 1. Phenotypic comparison between Sanggol mutant (*lpa*) and wild-type (WT) cultivar**
239 **Ilpum.** (A) Whole plant. (B) Spikelet. (C) Mature grain.

240

## 241 Determination of PA and $P_i$ content in Sanggol and Ilpum seeds

242 To quantify PA and $P_i$ content in seeds, brown rice extracts of Sanggol and Ilpum were
243 analyzed via $^{31}P$ NMR spectroscopy. Results showed that PA contents were significantly
244 reduced (49% reduction), and $P_i$ content was significantly increased in the seeds of Sanggol
245 compared with Ilpum (Table 5). The $^{31}P$ NMR analysis showed peaks analogous to standard
246 (Fig 2A) for $P_i$ and PA peak identification. Similarly, $P_i$ and PA analogous peaks were observed
247 for wild (WT) (Fig 2B), and mutant (*lpa*) (Fig 2C) types.

248 Additionally, PA and $P_i$ amounts were also quantified among 96 $F_2$ individuals using
249 $^{31}P$ NMR spectroscopy. Segregation analysis revealed that 77 $F_2$ plants showed the wild-type
250 phenotype, whereas 19 $F_2$ plants showed the mutant phenotype (Table 6), and the phenotype
251 segregation fitted a 3:1 ratio, suggesting that a single recessive allele controls the low PA in
252 the seeds of Sanggol mutant cultivar.

253

254 **Table 5. Seed PA and $P_i$ content in Sanggol and Ilpum cultivars.**

| Cultivar | PA P (mg $g^{-1}$) | $P_i$ (mg $g^{-1}$) | Total P (mg $g^{-1}$) |
|---|---|---|---|
| Ilpum | 5.7 ± 0.34 | 0.1 ± 0.04 | 5.85 ± 0.34 |
| Sanggol | 2.9 ± 0.69* | 1.8 ± 0.10** | 5.21 ± 0.62 |

255

256 Data represent mean ± standard error ($n$ = 3).
257 Asterisks indicate the level of significance (*, $P < 0.05$; **, $P < 0.01$) between Sanggol and Ilpum.
258

259 **Fig. 2. $^{31}P$ NMR spectrum of standard, Ilpum (WT), and Sanggol mutant (*lpa*)** (A)
260 Reference standard peaks. (B) Wild 'WT' (C) *lpa* 'Mutant'.

**Table 6. Segregation and co-segregation analysis of seed PA content among 96 $F_2$ individuals derived from a cross between Sanggol and Ilpum cultivars.**

| Cross | No. of $F_2$ plants | PA phenotype | | | | | dCAPS genotyping | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | High PA | Low PA | Expected | χ2 | *P*-value† | W* | M* | H* | Expected | χ2 | *P*-value† |
| Sanggol /Ilpum | 96 | 77 | 19 | 3:1 | 1.38 | 0.23 | 26 | 19 | 51 | 1:2:1 | 1.39 | 0.49 |

†Not significant ($P > 0.05$).

*Wild: homozygous wild-type, H: heterozygous, M: homozygous mutant.

## Expression of PA biosynthetic gene and sequence analysis

To identify the gene responsible for reduced PA content in seeds, the candidate gene approach was followed. In rice, PA biosynthesis and accumulation begins after flowering [42, 43], and continues until 25 DAF during seed development [44]. Therefore, we extracted total RNA from 'Ilpum' spikelets at 5 DAF, and subjected it to RT-PCR analysis. Results showed that 15 genes in the PA biosynthesis pathway were expressed at 5 DAF (Fig 3). Further, we amplified and sequenced 16 genes involved in PA biosynthesis from sanggol and Ilpum cultivars (S1 Table).

**Fig 3. Semi-quantitative RT-PCR analysis of PA biosynthetic genes at 5 DAF in the Ilpum cultivar.**

Sequence analysis of PA biosynthetic genes revealed a single nucleotide polymorphism (SNP) in the *OsLpa*1 gene of Sanggol *lpa* mutant (Fig 4A); none of the other PA biosynthetic genes showed mutations in Sanggol *lpa* mutant. Previously, the *OsLpa*1 locus has been mapped to chromosome 2 [11], and narrowed down to a region less than 150 kb using microsatellite and sequence tagged site markers [45]. Further, the *OsLpa*1 has been characterized in *lpa* mutants of rice [22, 30]. The *OsLpa*1 gene encodes three expressed splice variants in rice [22, 35]. Sequence analysis of the *OsLpa*1 locus (position +1 to 2,058 bp; Genbank accession number: MH707666) showed a SNP (C623T) in the fourth exon of the largest splice variant, designated as *OsLpa*1-3.1, in Sanggol *lpa* mutant. Additionally, another SNPs (C53T) was identified in the first exon of the small splice variants, *OsLpa*1-3.2 and *OsLpa*1-3.3 (S2 Fig). Further, sequence analysis of *OsLpa*1-3.1 cDNA confirmed the presence of *lpa1*-3 allele in Sanggol mutant (Fig 4B).

**Fig 4. Gene structure of *OsLpa*1 according to Kim et al. [22] and Zhao et al. [30].** (A) Structure of the novel splice variant *OsLpa*1-3.1 carrying the C623T mutation in Sanggol low PA mutant cultivar. Empty boxes represent 5′UTR and 3′UTR, black box represents the coding region, and lines between boxes indicate introns. ATG (start codon) and TGA (stop codon) are shown for each splicing variant. (B) cDNA validation of *OsLpa*1-3.1 showing *Lpa*1-3 allele in the Sanggol mutant cultivar.

To determine the expression of *OsLpa*1 splice variants in mutant and wild-type cultivars, we performed RT-PCR analysis of *OsLpa*1 gene at 15 DAG using total RNA isolated from leaves and spikelets at 5 DAF. Expression analysis revealed that both *OsLpa*1-3.1 and *OsLpa*1-3.2 were expressed at 15 DAG, with slightly different expression patterns, whereas *OsLpa*1-3.3 showed no expression at 15 DAG in both cultivars (Fig 5A), indicating that *OsLpa*1-3.1 and *OsLpa*1-3.2 play an important role in seedling growth. At 5 DAF, *OsLpa*1-3.1 showed strong expression in both Sanggol *lpa* mutant and wild cultivar Ilpum; however, *OsLpa*1-3.3 exhibited low expression in both cultivars, and *OsLpa*1-3.2 exhibited no expression in either cultivar, suggesting *OsLpa*1-3.1 as a candidate transcript responsible for the seed low PA phenotype of Sanggol mutant. Protein analysis of Lpa1 amino acid sequence predicted deleterious amino acid substitution changes threonine (Thr) to isoleucine (Ile) in *OsLpa*3.1 (Thr208Ile), with a -5.715 proven score. Similarly, deleterious amino acid substitution changes were observed in *OsLpa*3.2 and *OsLpa*3.3 (Thr18Ile), with -5.482 proven scores.

**Fig 5. RT-PCR analysis of the *Lpa* gene family in Sanggol and Ilpum.** (A) Expression of *OsLpa*1 gene at 15 DAG and 5 DAF. (B) Expression of *OsLpa*1 gene paralog at 5 DAF.

317

318 Additionally, expression of the *OsLpa*1 paralog, reported previously by Kim et al. [22],
319 was investigated at 5 DAF in Sanggol *lpa* mutant and wild cultivars using RT-PCR. The
320 *OsLpa*1 paralog exhibited strong expression in both Sanggol *lpa* mutant and wild cultivars (Fig
321 5B), suggesting that sequence variation in the coding region of *OsLpa1* was responsible for the
322 low PA content of Sanggol seeds. In addition, reduction of PA content in *Arabidopsis atipk2β*
323 mutant indicates the IPK2 kinase of lipid dependent pathway is active the seeds [20]. We also
324 ruled out the possibility for seed PA biosynthesis similar to *Arabidopsis* in Sanggol low PA
325 mutant cultivar. However, our RT-PCR results showed no expression of *OsIpk2*, a key PA
326 biosynthesis gene in the lipid dependent pathway (data not shown), suggesting that the lipid
327 dependent pathway is not active in the Sanggol or Ilpum cultivar.

328 Next, we performed multiple sequence alignment of Lpa1 amino acid sequences of
329 Sanggol and other major plant species. Results revealed an amino acid substitution in the
330 conserved kinase domain in Sanggol (Fig 6A), thus showing the impact of a SNP in gene
331 coding sequence. The kinase domain of Lpa1 shows weak homology with that of 2-
332 phosphoglycerate kinase (2-PGK) found in hyperthermophilic methanogens [22]. However,
333 there is structural similarity among the substrates and products of 2-PGK and Lpa1 [46].

334 Phylogenetic analysis revealed a strong relationship among the kinase proteins in the
335 glycolysis and PA biosynthesis pathways. This suggests that Lpa proteins encoding Ins(3)$P_1$
336 kinase are classified into the Lpa clade (Fig 6B).

337

338 **Fig 6. Multiple sequence alignment and phylogenetic analysis of Lpa protein.** (A) Multiple
339 sequence alignment of Lpa1 protein from Sanggol (s-mtLpa1) and other major plant species.
340 The P-loop kinase domain of 2-PGK found in *Methanothermus fervidus* is marked with a red
341 line. P-loop is an ATP/GTP binding site motifs (Walker A and Walker B) are indicated using

342 pink and black arrowheads, respectively. The yellow box shows a single amino acid
343 substitution in the conserved kinase domain in Sanggol cultivar. (B) Phylogenetic relationship
344 among protein families of various kinases, including 2-PGK, serine threonine protein kinase
345 (S/T), fructokinase (FK), *myo*- inositol kinase 1 (MIK), inositol 1,3,4,5,6 penta*kis*phosphate 2-
346 kinase (IPK1), inositol 1,4,5-tris-phosphate kinase (IPK2), hexokinase 1 (HXK), and
347 hexokinase 2, with the Lpa protein clade. Phylogenetic tree was constructed using amino acid
348 sequences of Lpa and other kinase protein families from selected species using the neighbour-
349 joining method with 1,000 bootstrap replicates. Protein homologs are indicated with an 'H'.
350 ob, *Oryza brachyantha*; Os-J, *Oryza sativa* L. *japonica*; Osi, *Oryza sativa* L. *indica*; Agt/Ast,
351 *Aegilops tauschii*; sb, *Sorghum bicolor*; zm, *Zea Mays*; Si, *Setaria italica*; Ao, *Asparagus*
352 *officinalis*; Ga, *Gossypium arboretum*; Cc, *Cajanus cajan*; Gm, *Glycine max*; Cm, *Cucurbita*
353 *maxima*; ta, *Nicotiana tabacum*; Bn, *Brassica napus*; At, *Arabidopsis thaliana*; Ac, *Ananas*
354 *cosmos*; Ztm, *Zotera marina*; Ci, *Chrysanthemum indicum*; Mf, *methanothermus fervidus*; R-
355 570, *Saccharum*; Al, *Arabidopsis lyrata*.
356

## 357 Co-segregation analysis of low PA phenotype with *lpa1-3* allele

358 A dCAPS assay was developed to determine the co-segregation of *lpa*1-3 allele with
359 the low PA phenotype (Fig 7). A pair of dCAPS markers amplified a 192 bp PCR product.
360 Digestion of this PCR product with *Taq*I yielded a 174 bp fragment in Sanggol, but an uncut
361 fragment (192 bp) in Ilpum. Genotyping the $F_2$ individuals using this dCAPS marker showed a
362 segregation ratio, which was consistent with the expected ratio of 1:2:1 (Table 3). In addition,
363 the dCAPS marker genotype co-segregated with the low PA phenotype in the $F_2$ population.
364 Statistical analysis using Student's *t*-test revealed significant differences in the seed PA (S3
365 Fig) and $P_i$ (S4 Fig) contents of Ilpum, Sanggol, and $F_2$ individuals.
366

**Fig 7. dCAPS analysis of *lp1*-3 allele in the $F_2$ population derived from a cross between Sanggol and Ilpum.** P1, 'Ilpum'; P2, 'Sanggol'; W, homozygous wild type; M, homozygous mutant; H, heterozygous.

# Biocomputational analysis

Structural analysis of Lpa1 using molecular docking of ligand and cofactors showed that Ins(3)$P_1$ kinase was able to bind to the active site of Lpa1 protein, with ATP as a cofactor for catalysis (Fig 8A). Detailed view of the 3D protein model showed that Thr residue at amino acid position 208 was located in the kinase loop (Fig 8B) on the outer surface of the protein, adjacent to the entry site of the binding pocket, thus indicating its potential involvement in regulating the enzyme activity of Lpa1 protein. In addition, GPS 3.0 predicted Thr208 residue as a putative phosphorylation site, with a score of 9.66 above the cut-off value of 8.31. In previous studies, biochemical characterization of the regulatory mechanisms of various other metabolic enzymes has shown that amino acid substitutions are responsible for the reduction in enzyme activity of mutant proteins compared with wild-type proteins [47, 48]. Altogether, our results suggest that Thr208Ile amino acid substitution regulates the enzyme activity of Lpa1 protein via phosphorylation in Sanggol mutant cultivar.

**Fig 8. Three-dimensional (3D) model of Lpa1 protein structure.** (A) 3D model of Lpa1 protein showing Ins(3)$P_1$ (ligand) binding site and ATP (cofactor). (B) Thr208Ile substitution is indicated with an arrow, and the P loop containing Walker A and Walker B motifs is shown.

# Discussion

To date, several genes controlling PA biosynthesis have been reported in major crop plants [13, 20, 25, 28, 49-51]. The biosynthesis of PA proceeds via two major routes: a lipid dependent pathway, which operates in all plant tissues, and lipid independent pathway, which operates predominantly in seeds [3, 17]. In rice, molecular characterization of genes encoding MIPS, MIK, Lpa1, ITPK5/6, and IPK1 has revealed association with the low PA phenotype [21-25]. The first step of PA biosynthesis involves the conversion of D-glucose-6-phosphate to Ins(3)$P_1$ by MIPS [19], which is followed by a series of phosphorylation steps, leading to the formation of $InsP_6$ (S1 Fig). However, biochemical pathways leading to the conversion of Ins(3)$P_1$ to $InsP_4$, and the enzymes involved are very complex and not yet fully understood in plants [3].

Understanding the genetic basis of low PA phenotype is important for developing cultivars with low PA content in seeds. Therefore, we obtained the low PA mutant cultivar 'Sanggol' from Kangwon National University, Republic of Korea [31, 46]. In this study, Sanggol showed relatively poor agronomic performance compared with the wild cultivar Ilpum (Table 4). These results are in agreement with previous studies showing superior agronomic performance of wild cultivars compared with the low PA mutants [5, 14]. Edwards et al. [41] report an association between *Lpa*1 locus and grain chalkiness in rice. Similarly, the Sanggol showed high percentage of chalky grains compared with Ilpum, indicating that the low PA phenotype interacts with grain chalkiness. Thus, results of this study and previous studies suggest that the *lpa* allele plays an important role in determining the yield potential and seed quality of rice.

Phenotypic analysis using $P^{31}$ NMR spectroscopy showed a significant reduction in PA content and an increase in $P_i$ content in Sanggol seeds (Table 5). Expression analysis of PA biosynthetic genes in spikelets of the wild-type cultivar Ilpum at 5 DAF indicated that 15 genes from the lipid independent pathway were possibly responsible for the low PA content in

Sanggol (Fig 3). Our data showed that a point mutation in the *OsLpa*1 locus was associated with low PA content in Sanggol seeds. Previous studies have also shown that rice low PA mutants exhibit a reduction in seed PA content because of SNPs [25, 30]. Candidate gene sequencing (Fig 4) and co-segregation analysis (Fig 7 and Table 6) confirmed that a new single recessive allele of *Lpa1*, designated as *lpa1-3*, was responsible for the low PA phenotype of Sanggol *lpa* mutant because of a C/T SNP located at nucleotide position 623 in *OsLpa*1, resulting in a single amino acid substitution (Thr208Ile). In a previous study, the *japonica* mutant 'KBNT *lpa1-1*' exhibited a 28% reduction in seed PA content because of a SNP (C/G to T/A), resulting in a nonsense mutation at amino acid position 409 whereas the DR1331-2 (*lpa1-2)* mutant showed a 48% reduction in seed PA content because of a single nucleotide deletion (T/A) at position 313, causing a frame shift mutation [22]. In addition, molecular characterization of the *indica* mutant 'Os-lpa-XQZ-1' shows the deletion of a 1,475 bp fragment in *lpa1-1*, resulting in a 38% reduction in seed PA content [30].

The *OsLpa*1 gene encodes three splice variants, all of which are expressed in seeds, suggesting that these variants play different roles in rice seed development [22, 35]. However, RT-PCR analysis of *OsLpa*1 locus revealed that *OsLpa*1-3.1 expression exhibited both vegetative and seed specificity, which indicates a major role of *OsLpa*1-3.1 in PA biosynthesis; however, *OsLpa*1-3.2 and *OsLpa*1-3.3 showed significant and dynamic changes at 15 DAG and 5 DAF, respectively (Fig 5A), suggesting that these variants play important roles in seedling growth and seed development, respectively. This finding is consistent with a previous study in rice [52]. Additionally, we investigated the expression of *OsLpa*1 paralog (Os09g0572200) and a IPK2 kinase is specific for the lipid independent pathway, *OsIpk*2 (Os02g0523800), in spikelets at 5 DAF, to provide an alternative explanation for the low level of PA in Sanggol mutant seeds. However, expression analysis suggests that the *OsLpa*1 paralog gene involved in PA biosynthesis in Sanggol (Fig 5B).

440 According to a previous study, *OsLpa*1 shows a weak homology to 2-PGK found in
441 *Methanothermus fervidus* [22]. 2,3-bisphosphoglycerate (2,3-BPG), derived from 2-PGK, is a
442 strong inhibitor of inositol polyphosphate 5-phosphatases [53]; thus, removing this inhibition
443 may degrade inositol polyphosphate intermediates, causing a reduction in seed PA content in
444 low PA mutants [22, 54]. Based on the structural similarity among substrates and products of
445 *OsLpa*1 and 2-PGK, it is possible that Lpa1 protein functions as a kinase [3]. Additionally, our
446 results revealed a single amino acid substitution (Thr208Ile) in the kinase domain of *Lpa1* in
447 Sanggol. The *Lpa1* gene encodes Ins(3)$P_1$ kinase, which phylogenetically groups with the Lpa
448 clade. From the molecular docking analysis, it is evident that Ins(3)$P_1$ binds to the Lpa1 protein,
449 with ATP as a cofactor for catalysis (Fig 8A). Overall, these results suggest that Lpa1 protein
450 functions as a kinase, and is probably involved in the conversion of Ins(3)$P_1$ to *myo*-inositol
451 3,4-bisphosphate [Ins (3,4) $P_2$].

452 In *Arabidopsis*, aspartic acid (Asp) to alanine (Ala) substitutions at amino acid positions
453 98 and 100 (Asp98Ala and Asp100Ala) in two genes encoding inositol polyphosphate kinases
454 result in inactive enzymes and LPA phenotypes [55]. Similarly, analysis of phosphorylation
455 deficient mutants in yeast and human shows decreased MIPS activity compared with wild- type
456 because of amino acid substitutions at phosphorylation sites [48]. Several studies of various
457 kinases and other metabolic enzymes show reduced enzyme activity of the mutant protein
458 because of Thr and other amino acid substitutions at phosphorylation sites [47, 56-60].
459 Therefore, we speculate that a point mutation (C623T) causing Thr208Ile amino acid
460 substitution in the loop adjacent to the entry site of the binding pocket of OsLpa1 is responsible
461 for the altered enzyme activity of OsLpa1-3.1, resulting in reduced PA biosynthesis in Sanggol
462 mutant seeds. Additionally, enzyme activity analysis is necessary to confirm the association of
463 Thr208Ile substitution with the reduction in seed PA content in Sanggol low PA mutant cultivar.

464 Previous findings suggest that climate change and elevated $CO_2$ levels negatively affect
465 micronutrient bioavailability and total P in grains [15, 16]. However, rising $CO_2$ levels are
466 likely to increase the yield of rice crop because of the stimulation of photosynthetic rate [61].
467 A 1.2% increase in seed PA content under elevated $CO_2$ conditions has been reported in rice
468 [62]. In addition, crop plants need more P under elevated $CO_2$ levels [63]. Soil P also exhibits
469 a positive correlation with seed PA content in rice [64]. Limited information is available on
470 how climate change affects seed PA content, and further studies are needed to avoid nutrient
471 deficiencies. Reducing the seed PA content and increasing P uptake by crop plants should be a
472 major focus of future crop breeding programs. The results of Sanggol mutant reported in this
473 study will facilitate the development of new low PA lines with increased seed micronutrient
474 bioavailability, high P uptake, better nutrition, and enhanced agronomic performance, despite
475 elevated $CO_2$ levels, using marker assisted introgression of the *lpa1-3* allele into elite rice
476 varieties.
477

478 # Funding

479 This work was supported by a Grant from the Next-Generation BioGreen 21 Program (Plant
480 Molecular Breeding Center number PJ013165), Rural Development Administration, Republic
481 of Korea. The funder had no role in study design, data collection and analysis, decision to
482 publish, or preparation of the manuscript.
483

484 # Competing interests

485 The authors have declared that no competing interests exist.
486

# Acknowledgments

This work was supported by a Grant from the Next-Generation BioGreen 21 Program (Plant Molecular Breeding Center number PJ013165), Rural Development Administration, Republic of Korea.

# Author Contributions

**Conceptualization:** Kishor Doddanakatte Shivaramegowda, Hee-Jong Koh

**Data curation:** Kishor Doddanakatte Shivaramegowda, Hee-Jong Koh

**Formal analysis:** Kishor Doddanakatte Shivaramegowda

**Funding acquisition:** Hee-Jong Koh

**Methodology:** Kishor Doddanakatte Shivaramegowda, Choonseok Lee, Hee-Jong Koh

**Project administration:** Hee-Jong Koh

**Resources:** Hee-Jong Koh, Soon-Kwan Hong, Jin-Kwan Ham

**Software:** Kishor Doddanakatte Shivaramegowda, Dongryung Lee, Jeonghwan Seo

**Supervision:** Hee-Jong Koh

**Validation:** Choonseok Lee

**Visualization:** Kishor Doddanakatte Shivaramegowda, Choonseok Lee

**Writing ± original draft:** Kishor Doddanakatte Shivaramegowda

**Writing, review & editing:** Kishor Doddanakatte Shivaramegowda, Choonseok Lee, Jelli Venkatesh, Zhuo Jin, Dongryung Lee, Jeonghwan Seo, Joong Hyoun Chin, Hee-Jong Koh

# References

1. Raboy V. Accumulation and Storage of Phosphate and Minerals. In: Larkins BA, Vasil IK, editors. Cellular and Molecular Biology of Plant Seed Development. Dordrecht:

Springer Netherlands; 1997. p. 441-77. doi:https://doi.org/10.1007/978-94-015-8909-3_12.

2. Marounek M, Skřivan M, Rosero O, Rop O. Intestinal and total tract phytate digestibility and phytase activity in the digestive tract of hens fed a wheat-maize-soyabean diet. J. Anim. Feed Sci. 2010;19(3):430–439.
3. Raboy V. Approaches and challenges to engineering seed phytate and total phosphorus. Plant Science. 2009;177(4):281-96. doi: https://doi.org/10.1016/j.plantsci.2009.06.012.
4. Perera I, Seneweera S, Hirotsu N. Manipulating the Phytic Acid Content of Rice Grain Toward Improving Micronutrient Bioavailability. Rice. 2018;11(1):4. doi: 10.1186/s12284-018-0200-y.
5. Raboy V, Gerbasi PF, Young KA, Stoneberg SD, Pickett SG, Bauman AT, et al. Origin and Seed Phenotype of Maize low phytic acid 1-1 and low phytic acid 2-1. Plant Physiology. 2000;124(1):355-68.PMID: PMC59149.
6. Pilu R, Panzeri D, Gavazzi G, Rasmussen SK, Consonni G, Nielsen E. Phenotypic, genetic and molecular characterization of a maize low phytic acid mutant (lpa241). TAG Theoretical and applied genetics Theoretische und angewandte Genetik. 2003;107(6):980-7. Epub 2003/10/03. doi:10.1007/s00122-003-1316-y. PMID: 14523526.
7. Dorsch JA, Cook A, Young KA, Anderson JM, Bauman AT, Volkmann CJ, et al. Seed phosphorus and inositol phosphate phenotype of barley low phytic acid genotypes. Phytochemistry. 2003;62(5):691-706. Epub 2003/03/07. PMID: 12620321.
8. Hitz WD, Carlson TJ, Kerr PS, Sebastian SA. Biochemical and Molecular Characterization of a Mutation That Confers a Decreased Raffinosaccharide and Phytic Acid Phenotype on Soybean Seeds. Plant Physiology. 2002;128(2):650-60. doi: 10.1104/pp.010585. PMID: PMC148927.

9. Liu QL, Xu XH, Ren XL, Fu HW, Wu DX, Shu QY. Generation and characterization of low phytic acid germplasm in rice (Oryza sativa L.). TAG Theoretical and applied genetics Theoretische und angewandte Genetik. 2007;114(5):803-14. Epub 2007/01/16. doi: 10.1007/s00122-006-0478-9. PMID: 17219209.
10. Guttieri M, Bowen D, Dorsch JA, Raboy V, Souza E. Erratum: Identification and characterization of a low phytic acid wheat. Crop Science.2004;44(2):418-424.
11. Larson SR, Rutger JN, Young KA, Raboy V. Isolation and Genetic Mapping of a Non-Lethal Rice (Oryza sativa L.) low phytic acid 1 Mutation This work was supported by the USDA-ARS and by the USDA's National Research Initiative Competitive Grants Program Award no. 97-35300-4421. Crop Science. 2000;40(5):1397-405. doi: 10.2135/cropsci2000.4051397x.
12. Shi J, Wang H, Wu Y, Hazebroek J, Meeley RB, Ertl DS. The maize low-phytic acid mutant lpa2 is caused by mutation in an inositol phosphate kinase gene. Plant Physiol. 2003;131(2):507-15. Epub 2003/02/15. doi: 10.1104/pp.014258. PMID: 12586875.
13. Shi J, Wang H, Hazebroek J, Ertl DS, Harp T. The maize low-phytic acid 3 encodes a myo-inositol kinase that plays a role in phytic acid biosynthesis in developing seeds. The Plant journal : for cell and molecular biology. 2005;42(5):708-19. doi: 10.1111/j.1365-313X.2005.02412.x. PMID: 15918884.
14. Zhao H-J, Liu Q-L, Fu H-W, Xu X-H, Wu D-X, Shu Q-Y. Effect of non-lethal low phytic acid mutations on grain yield and seed viability in rice. Field Crops Research. 2008; 108: 206-211.
15. Seneweera S, Blakeney A, Milham P, S. Basra A, Barlow E, Conroy J. Influence of rising atmospheric CO2 and phosphorus nutrition on the grain yield and quality of rice (Oryza sativa cv. Jarrah). Cereal Chemistry.1996. 73(2): 239-43.

16. Dietterich LH, Zanobetti A, Kloog I, Huybers P, Leakey ADB, Bloom AJ, et al. Impacts of elevated atmospheric CO2 on nutrient content of important food crops. Scientific Data. 2015;2:150036. doi: 10.1038/sdata.2015.36.

17. Sparvoli F, Cominelli E. Seed Biofortification and Phytic Acid Reduction: A Conflict of Interest for the Plant?. Plants. 2015;4(4):728. doi:10.3390/plants4040728.

18. Suzuki M, Tanaka K, Kuwano M, Yoshida K. Expression pattern of inositol phosphate-related enzymes in rice (Oryza sativa L.): Implications for the phytic acid biosynthetic pathway. Gene.2007. 405(1-2):55-64.

19. Mitsuhashi N, Kondo M, Nakaune S, Ohnishi M, Hayashi M, Hara-Nishimura I, et al. Localization of myo-inositol-1-phosphate synthase to the endosperm in developing seeds of Arabidopsis. Journal of Experimental Botany. 2008;59(11):3069-76. doi: 10.1093/jxb/ern161. PMID: PMC2504351.

20. Stevenson-Paulik J, Bastidas RJ, Chiou S-T, Frye RA, York JD. Generation of phytate-free seeds in Arabidopsis through disruption of inositol polyphosphate kinases. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(35):12612-7. doi: 10.1073/pnas.0504172102.

21. Kim SI, Andaya CB, Newman JW, Goyal SS, Tai TH. Isolation and characterization of a low phytic acid rice mutant reveals a mutation in the rice orthologue of maize MIK. Theoretical and Applied Genetics. 2008;117(8):1291. doi: 10.1007/s00122-008-0863-7.

22. Kim SI, Andaya CB, Goyal SS, Tai TH. The rice OsLpa1 gene encodes a novel protein involved in phytic acid metabolism. Theoretical and Applied Genetics. 2008;117(5):769-79. doi: 10.1007/s00122-008-0818-z.

23. Ali N, Paul S, Gayen D, Sarkar SN, Datta K, Datta SK. Development of Low Phytate Rice by RNAi Mediated Seed-Specific Silencing of Inositol 1,3,4,5,6-

585 Pentakisphosphate 2-Kinase Gene (IPK1). PLOS ONE. 2013;8(7):e68161. doi:
586 10.1371/journal.pone.0068161.
587 24. Ali N, Paul S, Gayen D, Sarkar SN, Datta SK, Datta K. RNAi mediated down regulation
588 of myo-inositol-3-phosphate synthase to generate low phytate rice. Rice. 2013;6:12. doi:
589 10.1186/1939-8433-6-12. PMID: PMC4883737.
590 25. Kim S-I, Tai TH. Identification of novel rice low phytic acid mutations via TILLING
591 by sequencing. Molecular Breeding. 2014;34(4):1717-29. doi: 10.1007/s11032-014-
592 0127-y.
593 26. Xu XH, Zhao HJ, Liu QL, Frank T, Engel KH, An G, et al. Mutations of the multi-drug
594 resistance-associated protein ABC transporter gene 5 result in reduction of phytic acid
595 in rice seeds. TAG Theoretical and applied genetics Theoretische und angewandte
596 Genetik. 2009;119(1):75-83. Epub 2009/04/17. doi: 10.1007/s00122-009-1018-
597 1.PMID: 19370321.
598 27. Zhao HJ, Cui HR, Xu XH, Tan YY, Fu JJ, Liu GZ, et al. Characterization of OsMIK in
599 a rice mutant with reduced phytate content reveals an insertion of a rearranged
600 retrotransposon. TAG Theoretical and applied genetics Theoretische und angewandte
601 Genetik. 2013;126(12):3009-20. Epub 2013/09/18. doi: 10.1007/s00122-013-2189-3.
602 PMID: 24042572.
603 28. Kim S-I, Tai TH. Genetic analysis of two OsLpa1-like genes in Arabidopsis reveals
604 that only one is required for wild-type seed phytic acid levels. Planta.
605 2010;232(5):1241-50. doi: 10.1007/s00425-010-1243-5.
606 29. Shi J, Wang H, Schellin K, Li B, Faller M, Stoop JM, et al. Embryo-specific silencing
607 of a transporter reduces phytic acid content of maize and soybean seeds. Nature
608 Biotechnology. 2007;25:930.

609 30. Zhao H-J, Liu Q-L, Ren X-L, Wu D-X, Shu Q-Y. Gene identification and allele-specific
610 marker development for two allelic low phytic acid mutations in rice (Oryza sativa L.).
611 Molecular Breeding. 2008;22(4):603-12. doi: 10.1007/s11032-008-9202-6.
612 31. Li C-Y, Park D-S, Won S-R, Hong S-K, Ham J-K, Choi J-K, et al. Food chemical
613 properties of low-phytate rice cultivar, Sang-gol. Journal of Cereal Science.
614 2008;47(2):262-5. doi: https://doi.org/10.1016/j.jcs.2007.04.012.
615 32. Gao Y, Shang C, Maroof S, Biyashev R, Grabau E, Kwanyuen P, Burton JW, Buss GR.
616 A Modified Colorimetric Method for Phytic Acid Analysis in Soybean. Crop
617 Science. 2007;47:1797–1803. doi: 10.2135/cropsci2007.03.0122.
618 33. Cade-Menun BJ, Carter MR, James DC, Liu CW. Phosphorus Forms and Chemistry in
619 the Soil Profile under Long-Term Conservation Tillage: A Phosphorus-31 Nuclear
620 Magnetic Resonance Study All rights reserved. No part of this periodical may be
621 reproduced or transmitted in any form or by any means, electronic or mechanical,
622 including photocopying, recording, or any information storage and retrieval system,
623 without permission in writing from the publisher. Journal of Environmental Quality.
624 2010;39:1647-56. doi: 10.2134/jeq2009.0491.
625 34. Nanganoa LT, Njukeng JN. Phosphorus Speciation by 31P NMR Spectroscopy in Leaf
626 Litters and Crop Residues from Para Rubber, Cocoa, Oil Palm, and Banana Plantations
627 in the Humid Forest Zone of Cameroon. Journal of Applied Chemistry. 2018;2018:8.
628 doi: 10.1155/2018/6290236.
629 35. Sakai H, Lee SS, Tanaka T, Numa H, Kim J, Kawahara Y, et al. Rice Annotation Project
630 Database (RAP-DB): An Integrative and Interactive Database for Rice Genomics. Plant
631 and Cell Physiology. 2013;54(2):e6-e. doi: 10.1093/pcp/pcs183.
632 36. Lee H-S, Lee D-H, Cho HK, Kim SH, Auh JH, Pai H-S. InsP<sub>6</sub>-Sensitive
633 Variants of the Gle1 mRNA Export Factor Rescue Growth and Fertility Defects of the

634 ipk1 Low-Phytic-Acid Mutation in Arabidopsis. The Plant Cell.
635 2015;27(2):417-31. doi: 10.1105/tpc.114.132134.
636 37. Sato Y, Takehisa H, Kamatsuki K, Minami H, Namiki N, Ikawa H, et al. RiceXPro
637 Version 3.0: expanding the informatics resource for rice transcriptome. Nucleic Acids
638 Research. 2013;41(D1):D1206-D13. doi: 10.1093/nar/gks1125.
639 38. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the Functional Effect of
640 Amino Acid Substitutions and Indels. PLOS ONE. 2012;7(10):e46688. doi:
641 10.1371/journal.pone.0046688.38.
642 39. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis
643 Version 7.0 for Bigger Datasets. Molecular Biology and Evolution. 2016;33(7):1870-
644 4. doi: 10.1093/molbev/msw054.
645 40. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal
646 for protein modeling, prediction and analysis. Nature Protocols. 2015;10:845. doi:
647 10.1038/nprot.2015.053.
648 41. Edwards JD, Jackson AK, McClung AM. Genetic architecture of grain chalk in rice
649 and interactions with a low phytic acid locus. Field Crops Research. 2017;205:116-23.
650 doi: https://doi.org/10.1016/j.fcr.2017.01.015.
651 42. Yoshida KT, Wada T, Koyama H, Mizobuchi-Fukuoka R, Naito S. Temporal and
652 Spatial Patterns of Accumulation of the Transcript of Myo-Inositol-1-
653 Phosphate Synthase and Phytin-Containing Particles during Seed Development in Rice.
654 Plant Physiology. 1999;119(1):65-72. doi: 10.1104/pp.119.1.65.
655 43. Ogawa M, Tanaka K, Kasai Z. Accumulation of phosphorus, magnesium and potassium
656 in developing rice grains: followed by electron microprobe X-ray analysis focusing on
657 the aleurone layer. Plant and Cell Physiology. 1979;20(1):19-27. doi:
658 10.1093/oxfordjournals.pcp.a075791.

659 44. Iwai T, Takahashi M, Oda K, Terada Y, Yoshida KT. Dynamic Changes in the
660 Distribution of Minerals in Relation to Phytic Acid Accumulation during Rice Seed
661 Development. Plant Physiology. 2012;160(4):2007-14. doi: 10.1104/pp.112.206573.
662 45. Andaya CB, Tai TH. Fine mapping of the rice low phytic acid (Lpa1) locus. Theoretical
663 and Applied Genetics. 2005;111(3):489-95. doi: 10.1007/s00122-005-2038-0.
664 46. Li C-Y, Li H, Choi J-K, Kim H-W, Hong S-K, Xu Z, et al. Physicochemical properties
665 of low-phytate rice cultivar, Sang-gol. Journal of the Korean Society for Applied
666 Biological Chemistry. 2012;55(1):101-4. doi: 10.1007/s13765-012-0017-0.
667 47. Foltyn VN, Zehl M, Dikopoltsev E, Jensen ON, Wolosker H. Phosphorylation of mouse
668 serine racemase regulates D-serine synthesis. FEBS letters. 2010;584(13):2937-41.
669 Epub 2010/05/25. doi: 10.1016/j.febslet.2010.05.022. PMID: 20493854.
670 48. Deranieh RM, He Q, Caruso JA, Greenberg ML. Phosphorylation Regulates myo-
671 Inositol-3-phosphate Synthase: A NOVEL REGULATORY MECHANISM OF
672 INOSITOL BIOSYNTHESIS. The Journal of biological chemistry.
673 2013;288(37):26822-33. doi: 10.1074/jbc.M113.479121. PMID: PMC3772229.
674 49. Yuan F-J, Zhao H-J, Ren X-L, Zhu S-L, Fu X-J, Shu Q-Y. Generation and
675 characterization of two novel low phytate mutations in soybean (Glycine max L. Merr.).
676 Theoretical and Applied Genetics. 2007;115(7):945-57. doi: 10.1007/s00122-007-
677 0621-2.
678 50. Panzeri D, Cassani E, Doria E, Tagliabue G, Forti L, Campion B, et al. A defective
679 ABC transporter of the MRP family, responsible for the bean lpa1 mutation, affects the
680 regulation of the phytic acid pathway, reduces seed myo-inositol and alters ABA
681 sensitivity. The New phytologist. 2011;191(1):70-83. Epub 2011/03/15. doi:
682 10.1111/j.1469-8137.2011.03666.x. PMID: 21395595.

683 51. Ye H, Zhang X-Q, Broughton S, Westcott S, Wu D, Lance R, et al. A nonsense mutation
684 in a putative sulphate transporter gene results in low phytic acid in barley. Functional
685 & Integrative Genomics. 2011;11(1):103-10. doi: 10.1007/s10142-011-0209-4.
686 52. Lu H-p, Pang W-q, Li W-x, Tan Y-y, Wang Q, Zhao H-j, et al. Tissue-specific
687 expression, developmentally and spatially regulated alternative splicing, and protein
688 subcellular localization of OsLpa1 in rice. Journal of Zhejiang University Science B.
689 2016;17(2):100-9. doi: 10.1631/jzus.B1500205. PMID: PMC4757580.
690 53. Downes CP, Mussat MC, Michell RH. The inositol trisphosphate phosphomonoesterase
691 of the human erythrocyte membrane. Biochemical Journal. 1982;203(1):169-77. PMID:
692 PMC1158207.
693 54. Rasmussen Søren K, Ingvardsen Christina R, Torp Anna M. Mutations in genes
694 controlling the biosynthesis and accumulation of inositol phosphates in seeds.
695 Biochemical Society Transactions. 2010;38(2):689.
696 55. Stevenson-Paulik J, Odom AR, York JD. Molecular and biochemical characterization
697 of two plant inositol polyphosphate 6-/3-/5-kinases. The Journal of biological chemistry.
698 2002;277(45):42711-8. Epub 2002/09/13. doi: 10.1074/jbc.M209112200. PMID:
699 12226109.
700 56. Banerjee DK, Carrasquillo EA, Hughey P, Schutzbach JS, Martinez JA, Baksi K. In
701 vitro phosphorylation by cAMP-dependent protein kinase up-regulates recombinant
702 Saccharomyces cerevisiae mannosylphosphodolichol synthase. The Journal of
703 biological chemistry. 2005;280(6):4174-81. Epub 2004/11/19. doi:
704 10.1074/jbc.M406962200. PMID: 15548536.
705 57. Smal C, Vertommen D, Bertrand L, Ntamashimikiro S, Rider MH, Van Den Neste E,
706 et al. Identification of in vivo phosphorylation sites on human deoxycytidine kinase.
707 Role of Ser-74 in the control of enzyme activity. The Journal of biological chemistry.

2006;281(8):4887-93. Epub 2005/12/20. doi: 10.1074/jbc.M512129200. PMID: 16361699.

58. G SY, S KR, Kachhap S, Thakur M, Roy A, Singh B, et al. Eukaryotic-type serine/threonine kinase mediated phosphorylation at Thr(169) perturbs mycobacterial guanylate kinase activity. Bioscience reports. 2017;37(6). Epub 2017/10/01. doi: 10.1042/bsr20171048. PMID: 28963370.
59. Greif DM, Kou R, Michel T. Site-specific dephosphorylation of endothelial nitric oxide synthase by protein phosphatase 2A: evidence for crosstalk between phosphorylation sites. Biochemistry. 2002;41(52):15845-53. PMID: 12501214.
60. Choi H-S, Su W-M, Morgan JM, Han G-S, Xu Z, Karanasios E, et al. Phosphorylation of Phosphatidate Phosphatase Regulates Its Membrane Association and Physiological Functions in Saccharomyces cerevisiae: IDENTIFICATION OF SER(602), THR(723), AND SER(744) AS THE SITES PHOSPHORYLATED BY CDC28 (CDK1)-ENCODED CYCLIN-DEPENDENT KINASE. The Journal of biological chemistry. 2011;286(2):1486-98. doi: 10.1074/jbc.M110.155598. PMID: PMC3020757.
61. Hasegawa T, Sakai H, Tokida T, Nakamura H, Zhu C, Usui Y, Yoshimoto M, Fukuoka M, Wakatsuki H, Katayanagi N, Matsunami T, Kaneta Y, Sato T, Takakai F, Sameshima R, Okada M, Mae T, Makino A. Rice cultivar responses to elevated CO 2 at two free-air CO 2 enrichment (FACE) sites in Japan. Funct Plant Biol. 2013;7:148–159. doi: 10.1071/FP12357.
62. Myers SS, Zanobetti A, Kloog I, Huybers P, Leakey AD, Bloom AJ, et al. Increasing CO2 threatens human nutrition. Nature. 2014;510(7503):139-42. Epub 2014/05/09. doi: 10.1038/nature13179. PMID: 24805231.
63. Rakshit R, Patra A, Pal D, Kumar DM, Singh R. Effect of Elevated CO2 and Temperature on Nitrogen Dynamics and Microbial Activity During Wheat (Triticum

733 aestivum L.) Growth on a Subtropical Inceptisol in India. Journal of Agronomy and
734 Crop Science. 2012;198(6):452-465.
735 64. Rose T, Kretzschmar T, Liu L, Lancaster G, Wissuwa M. Phosphorus Deficiency Alters
736 Nutrient Accumulation Patterns and Grain Nutritional Quality in Rice. Agronomy.
737 2016;6(4):52. doi:10.3390/agronomy6040052.

738

739

# Supporting information

**S1 Fig. Schematic representation of the phytic acid (PA) biosynthetic pathway.** Glu6p, glucose-6-phosphate; Ins, *myo*-inositol; PtdIns, phosphatidyl inositol; MIPS, *myo*-inositol-3-phosphate synthase; IMP, *myo*-inositol monophosphatase; MIK, *myo*-inositol kinase; LPA1, low phytic acid 1; ITPK, inositol 1,3,4-triphosphate 5/6-kinase; IPK1, inositol 1,3,4,5,6 penta*kis*phosphate 2-kinase; PIS, phosphatidyl inositol phosphate synthase; PI4K, phosphatidyl inositol 4-kinase; PIP5K, phosphatidyl inositol 4 phosphate 5-kinase; PLC, phospholipase C; IPK2, inositol 1,4,5-tris-phosphate kinase; MRP, multidrug resistance protein.

**S2 Fig. Structure of *OsLpa*1-3.2 and *OsLpa*1-3.3 splice variants.** Mutations (C53T) in the first exon of *OsLpa*1-3.2 and *OsLpa*1-3.3 are indicated with red lines. Empty boxes represent 5′ and 3′ untranslated regions (UTRs), black box represents the coding region, and lines between boxes indicate introns. ATG (start codon) and TGA (stop codon) are shown.

**S3 Fig. Statistical analysis of seed PA content among $F_2$ plants derived from a cross between the mutant cultivar Sanggol and wild-type cultivar Ilpum.** Data were analyzed using the Student's *t*-test. M, mutant; W, wild-type.

**S4 Fig. Statistical analysis of inorganic phosphorous ($P_i$) content in seeds of $F_2$ plants.** Data were analyzed using the Student's *t*-test. M, mutant; W, wild-type.

**S1 Table. Primers used to sequence 16 PA biosynthetic genes in the lipid dependent and independent pathways.**

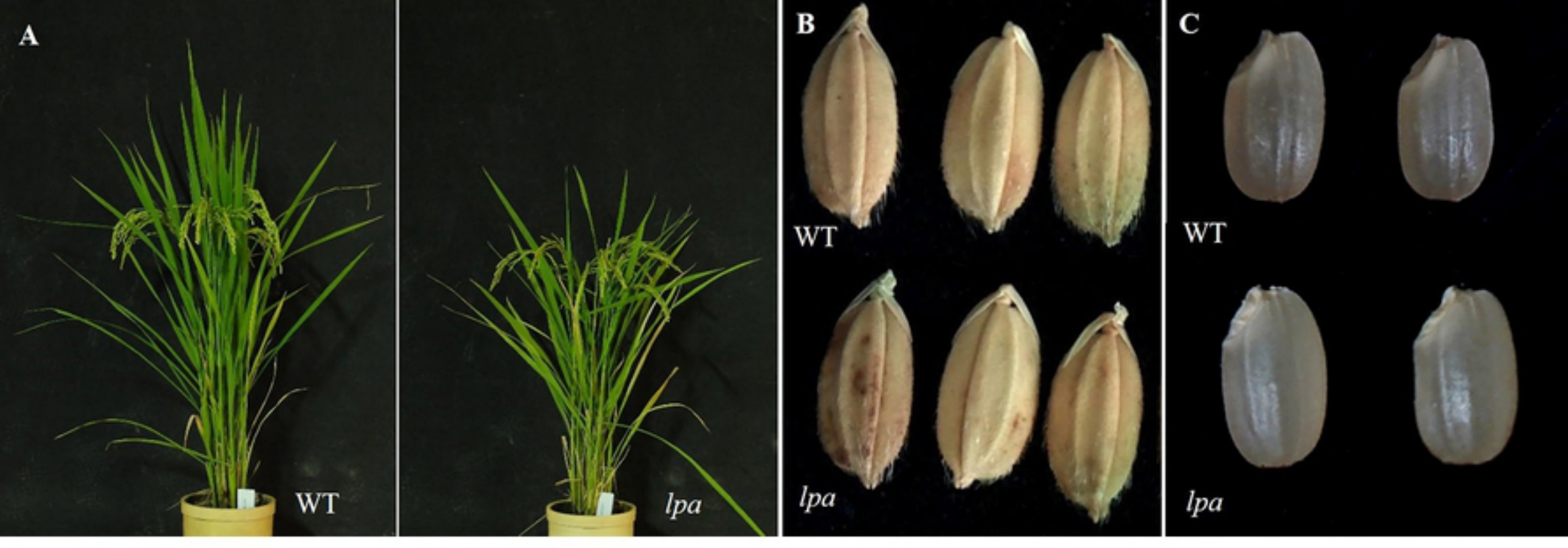
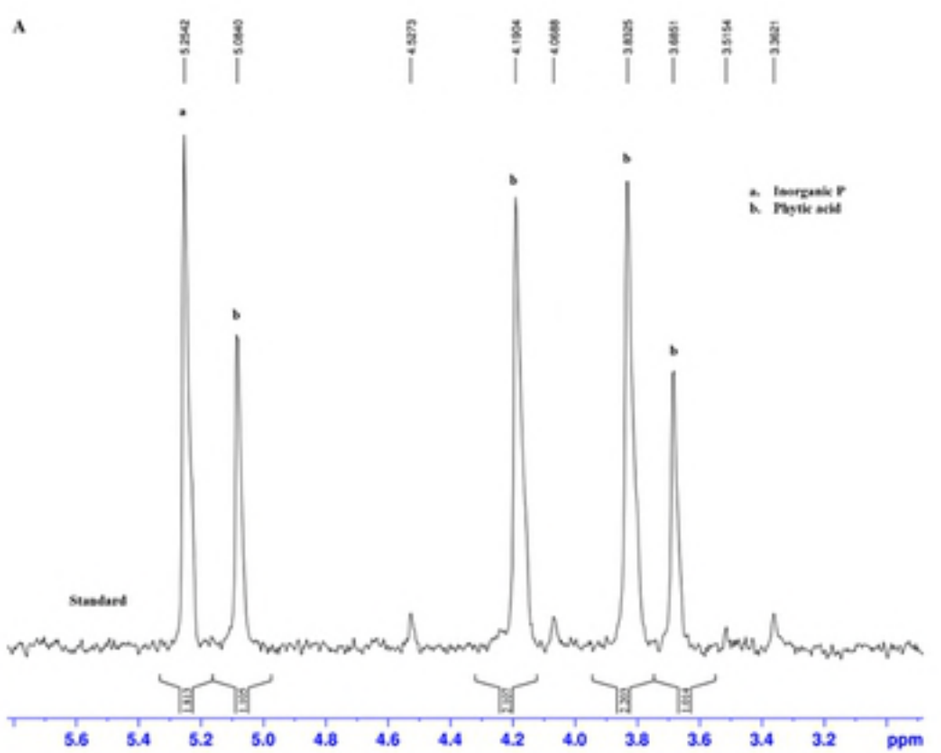
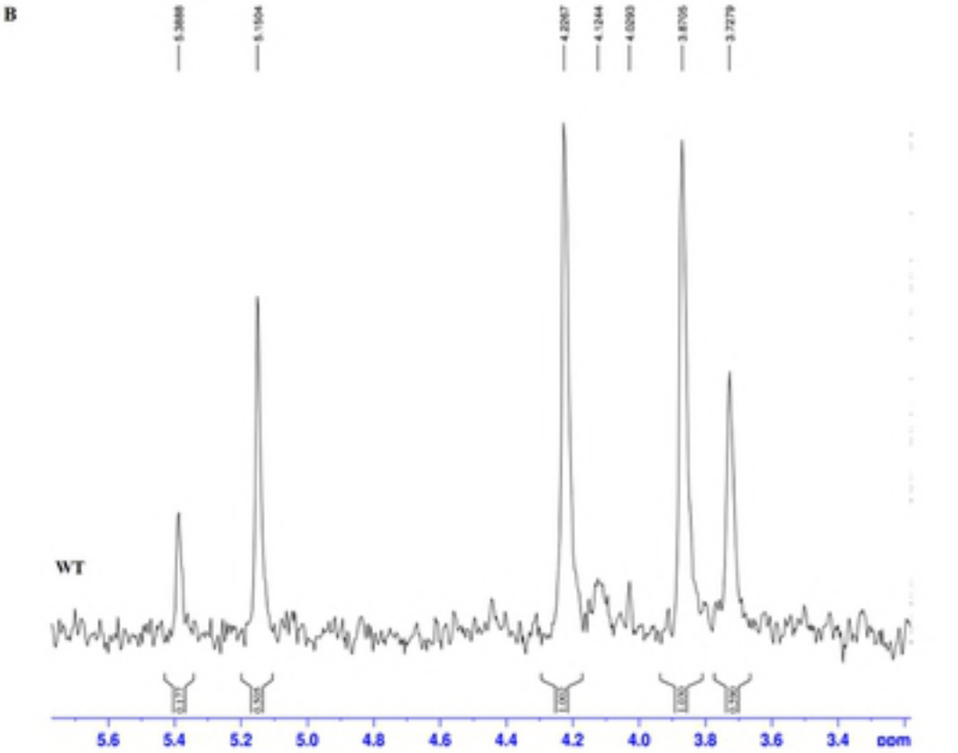



Figure1


A
5.2542
5.0840
4.5273
4.1904
4.0588
3.8325
3.6851
3.5154
3.3621
a
b
a. Inorganic P
b. Phytic acid
Standard
1.813
1.000
4.107
2.202
1.014
5.6 5.4 5.2 5.0 4.8 4.6 4.4 4.2 4.0 3.8 3.6 3.4 3.2 ppm


B
5.3888
5.1504
4.2267
4.1244
4.0293
3.8705
3.7279
WT
0.137
0.505
7.000
1.026
0.592
5.6 5.4 5.2 5.0 4.8 4.6 4.4 4.2 4.0 3.8 3.6 3.4 ppm

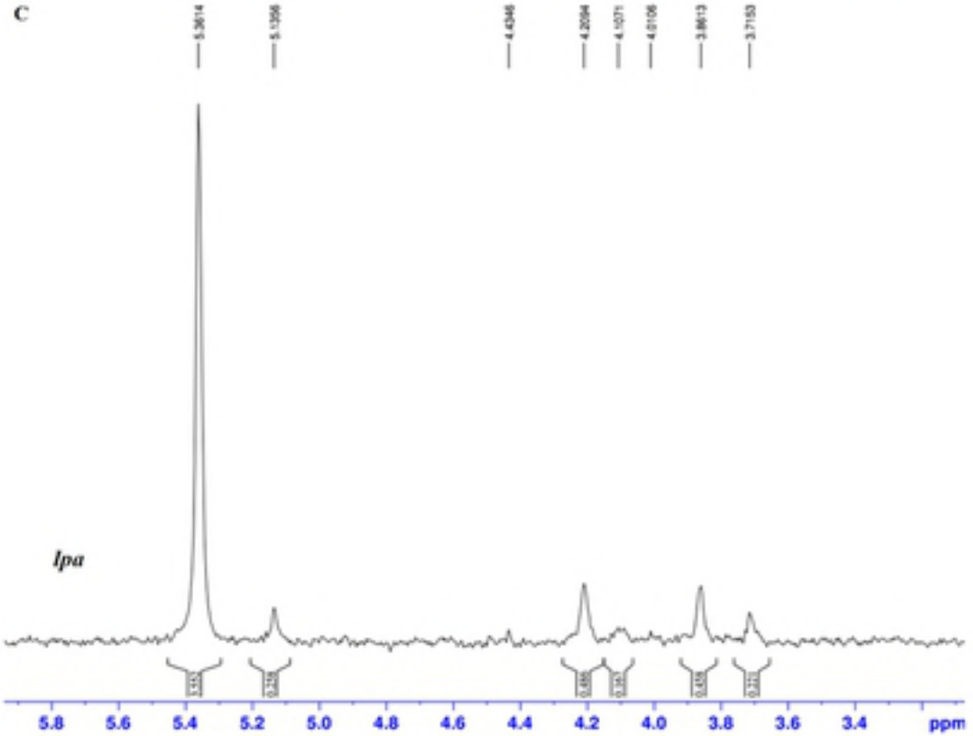

C
5.3614
5.1356
4.4248
4.2094
4.1071
4.0106
3.8613
3.7153
Ipa
3.451
0.250
0.695
0.381
0.458
0.231
5.8 5.6 5.4 5.2 5.0 4.8 4.6 4.4 4.2 4.0 3.8 3.6 3.4 ppm

Figure2

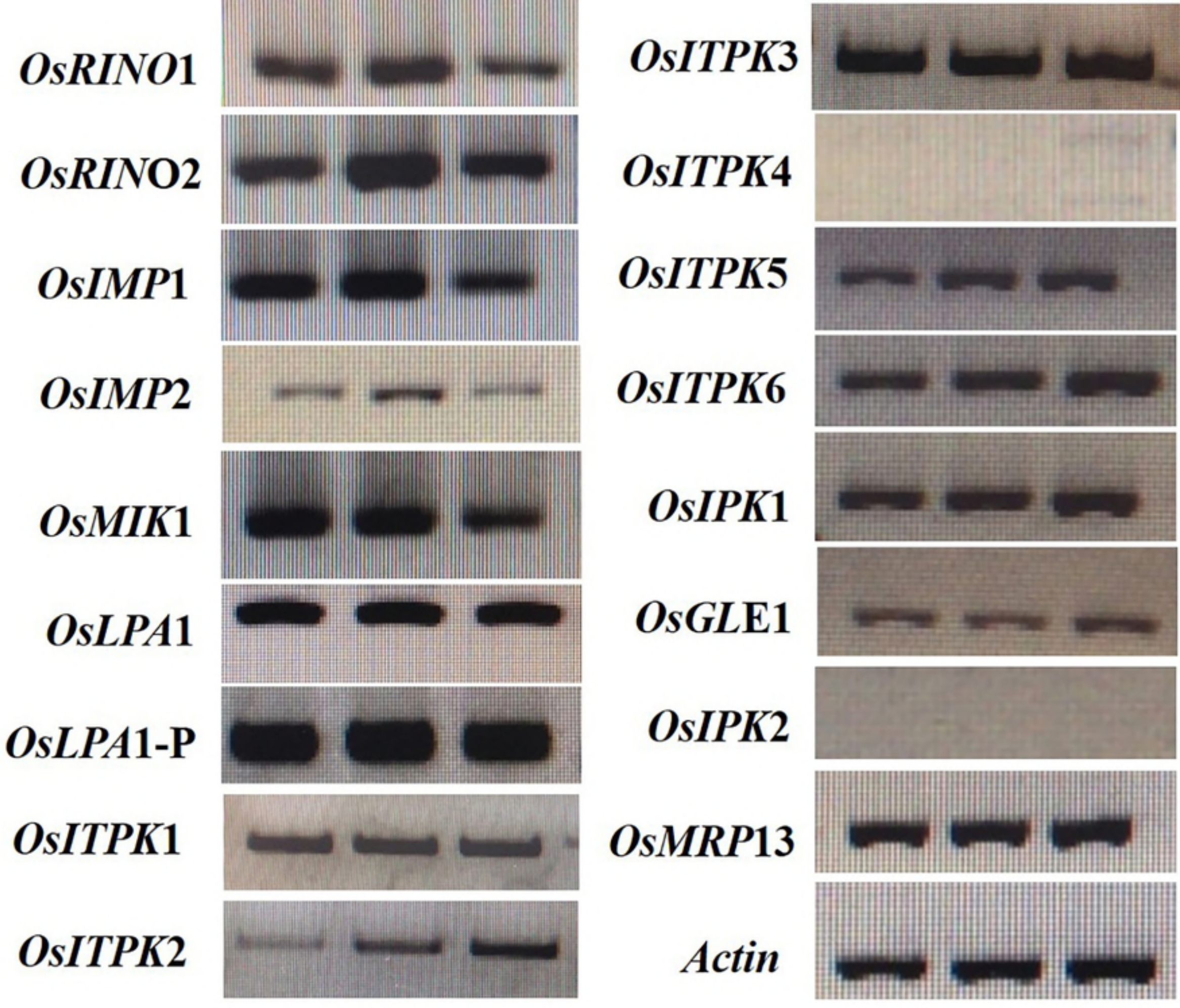

OsRINO1
OsRINO2
OsIMP1
OsIMP2
OsMIK1
OsLPA1
OsLPA1-P
OsITPK1
OsITPK2
OsITPK3
OsITPK4
OsITPK5
OsITPK6
OsIPK1
OsGLE1
OsIPK2
OsMRP13
Actin

Figure3

A

WT 623 ACC
*lpa* 623 ATC

***OsLpa*1-3.1**

ATG

TGA

B

C T A T G C A T C A A T C T A T C A T G C A G G C G A

50 60

*Lpa*1-3

Figure4

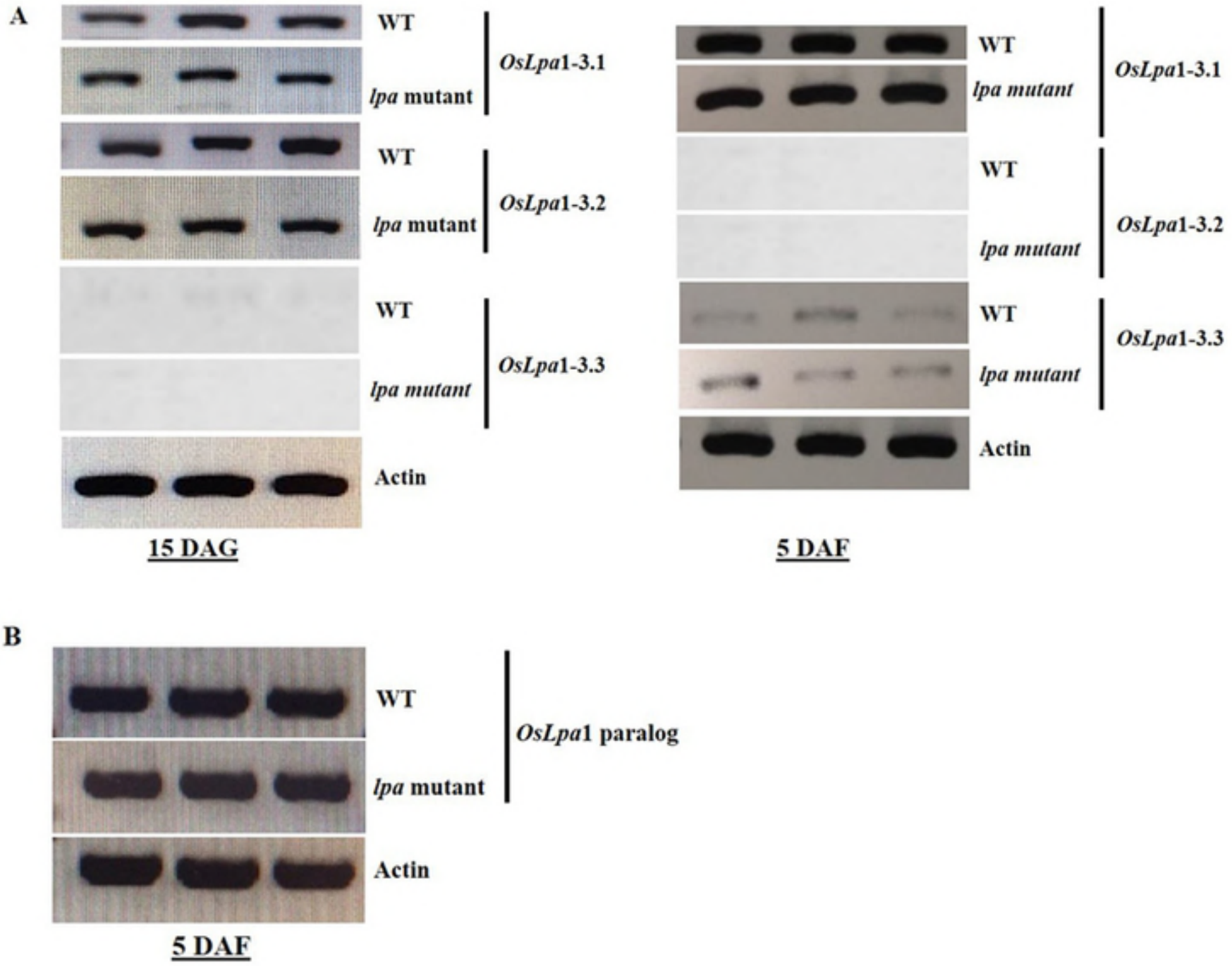

A
WT
lpa mutant
OsLpa1-3.1
WT
lpa mutant
OsLpa1-3.2
WT
lpa mutant
OsLpa1-3.3
Actin
15 DAG
WT
lpa mutant
OsLpa1-3.1
WT
lpa mutant
OsLpa1-3.2
WT
lpa mutant
OsLpa1-3.3
Actin
5 DAF
B
WT
lpa mutant
OsLpa1 paralog
Actin
5 DAF

Figure5

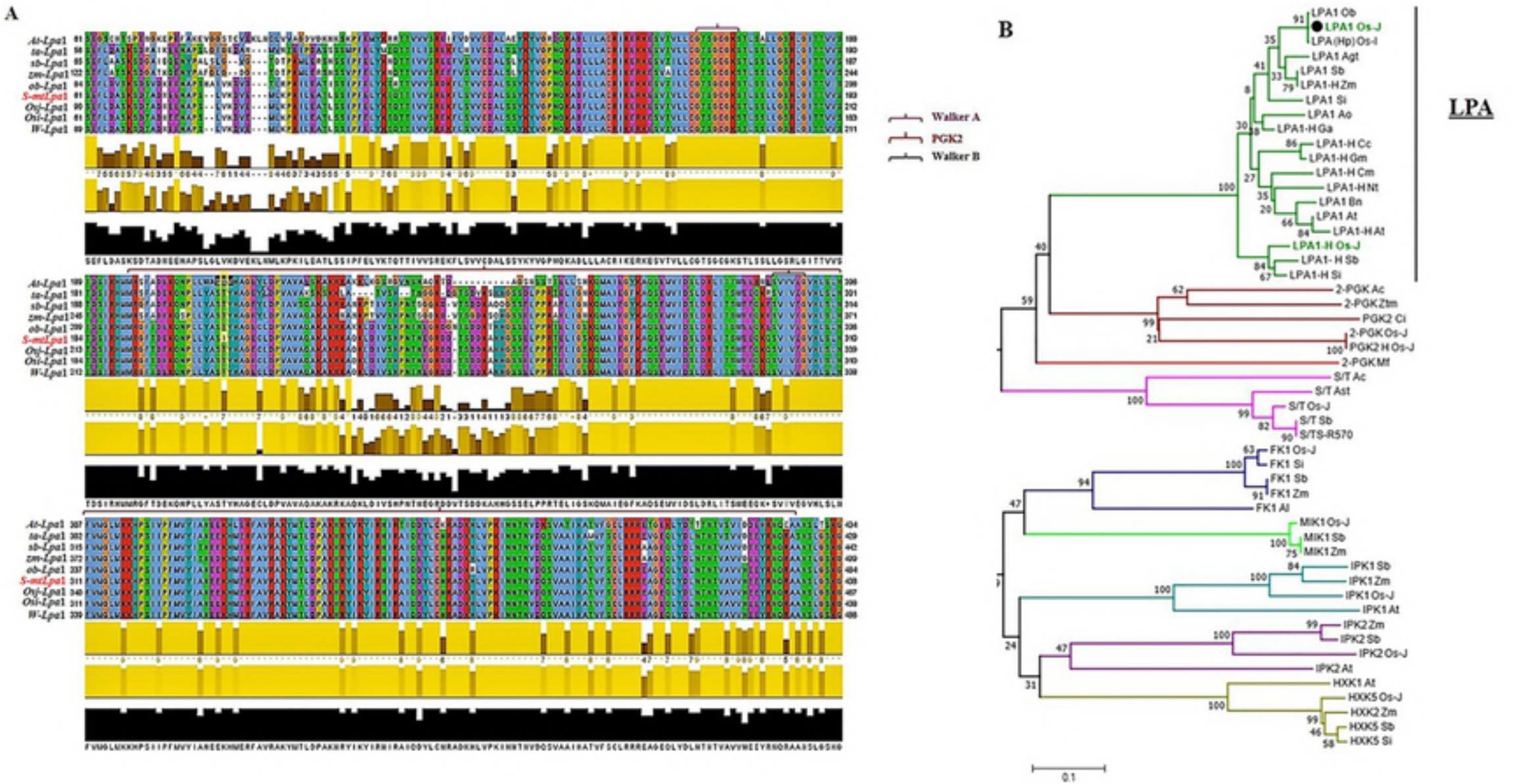

A
Walker A
PGK2
Walker B
B
LPA
LPA1 Ob
LPA1 Os-J
LPA(Hp) Os-I
LPA1 Agt
LPA1 Sb
LPA1-H Zm
LPA1 Si
LPA1 Ao
LPA1-H Ga
LPA1-H Cc
LPA1-H Gm
LPA1-H Cm
LPA1-H Nt
LPA1 Bn
LPA1 At
LPA1-H At
LPA1-H Os-J
LPA1- H Sb
LPA1-H Si
2-PGK Ac
2-PGK Ztm
PGK2 Ci
2-PGK Os-J
PGK2 H Os-J
2-PGK Mf
S/T Ac
S/T Ast
S/T Os-J
S/T Sb
S/TS-R570
FK1 Os-J
FK1 Si
FK1 Sb
FK1 Zm
FK1 Al
MIK1 Os-J
MIK1 Sb
MIK1 Zm
IPK1 Sb
IPK1 Zm
IPK1 Os-J
IPK1 At
IPK2 Zm
IPK2 Sb
IPK2 Os-J
IPK2 At
HXK1 At
HXK5 Os-J
HXK2 Zm
HXK5 Sb
HXK5 Si
0.1

Figure6


Figure7

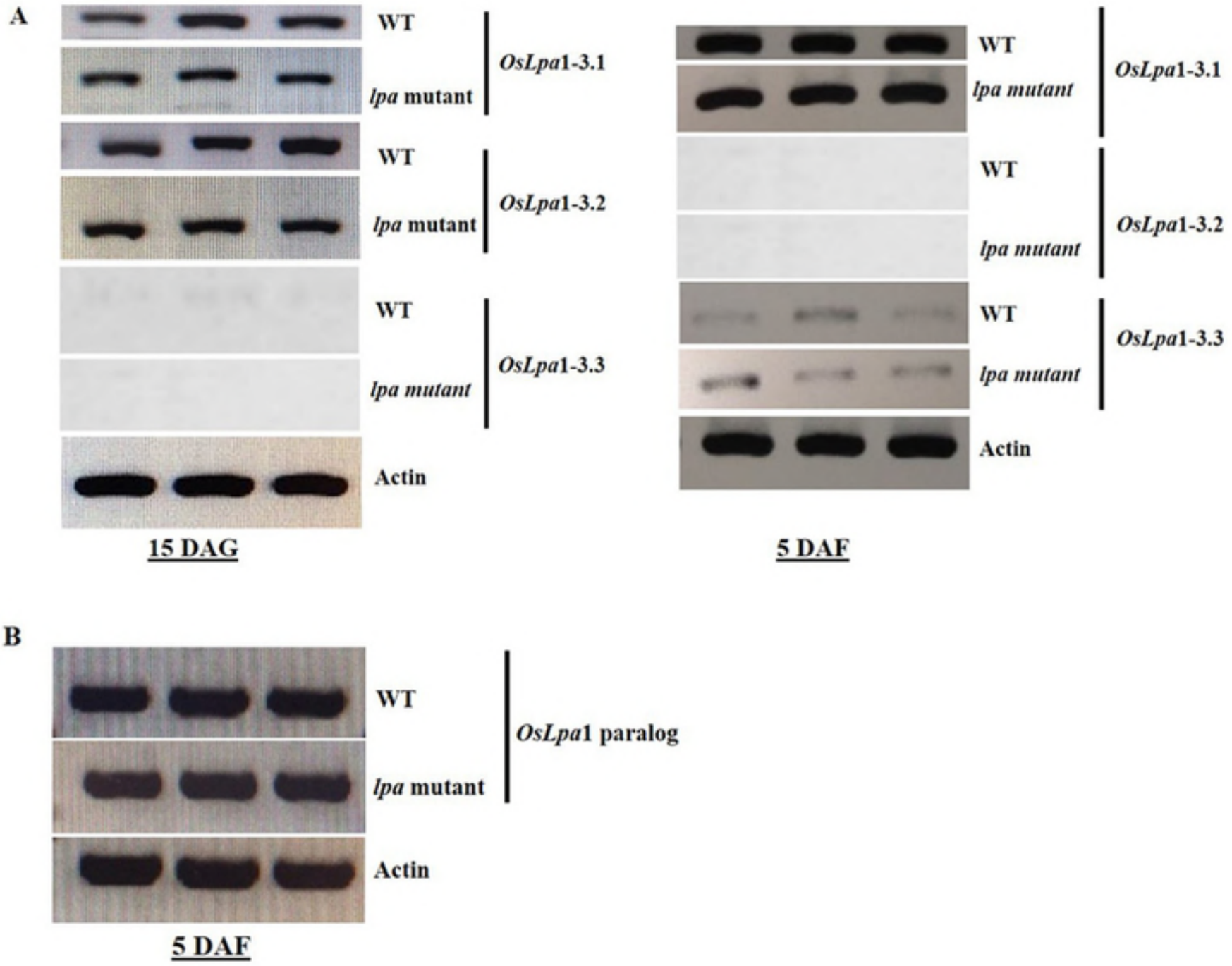

A
WT
lpa mutant
OsLpa1-3.1
WT
lpa mutant
OsLpa1-3.2
WT
lpa mutant
OsLpa1-3.3
Actin
15 DAG
WT
lpa mutant
OsLpa1-3.1
WT
lpa mutant
OsLpa1-3.2
WT
lpa mutant
OsLpa1-3.3
Actin
5 DAF
B
WT
lpa mutant
OsLpa1 paralog
Actin
5 DAF

Figure8