Molecular Biology Reports (2020) 47:401–422
https://doi.org/10.1007/s11033-019-05144-4

ORIGINAL ARTICLE

Novel *cis-acting* regulatory elements in wild *Oryza* species impart improved rice bran quality by lowering the expression of phospholipase D alpha1 enzyme (OsPLDα1)

Amandeep Kaur[1,2] · Kumari Neelam[1] · Ai Kitazumi[2,4] · Karminderbir Kaur[1] · Priti Sharma[1] · Gurjit Singh Mangat[3] · Benildo G. de los Reyes[2,4] · Darshan Singh Brar[1] · Kuldeep Singh[1,5]

Received: 15 April 2019 / Accepted: 16 October 2019 / Published online: 22 October 2019

Abstract

Rice bran oil is good quality edible oil, rich in antioxidants and comprised typically of oleic-linoleic type fatty acids. However, presence of a highly lipolytic enzyme Phospholipase D alpha1 (OsPLDα1) increases free fatty acid content in the oil which further leads to stale flavor and rancidity of the oil, making it unfit for human consumption. In this study, we compared the upstream regions of *OsPLDα1* orthologs across 34 accessions representing 5 wild *Oryza* species and 8 cultivars, to uncover sequence variations and identify *cis*-elements involved in differential transcription of orthologs. Alignment of the upstream sequences to the Nipponbare *OsPLDα1* reference sequence revealed the presence of 39 SNPs. Phylogenetic analysis showed that all the selected cultivars and wild species accessions are closely related to the reference except for three accessions of *O. rufipogon* (IRGC89224, IRGC104425, and IRGC105902). Furthermore, using exon-specific qRT-PCR, *OsPLDα1* expression patterns in immature grains indicated significant differences in transcript abundance between the wild species accessions. In comparison to the control, lowest gene expression was observed in IRGC89224 accession (0.20-fold) followed by IRGC105902 (0.26-fold) and IRGC104425 (0.41-fold) accessions. In-silico analysis of the *OsPLDα1* promoter revealed that the copy number variations of CGCGBOXAT, GT1CONSENSUS, IBOXCORE, NODCON2GM, OSE2ROOTNODULE, SURECOREATSULTR11, and SORLIP1AT *cis*-elements play an important role in the transcriptional activities of orthologous genes. Owing to the presence of ARFAT and SEBF elements only in the IRGC89224 accession, which had the lowest gene expression as well, these putative upstream regulatory sequences have been identified as novel *cis*-elements which may act as repressors in regulating the *OsPLDα1* gene expression. The accessions identified with low *OsPLDα1* expressions could be further deployed as potential donors of ideal *OsPLDα1* allele for transfer of the desired trait into elite rice cultivars.

Keywords *Oryza* · Phospholipase D · Promoter mining · qRT-PCR · *cis*-Elements

Introduction

Rice (*Oryza sativa* L.) is the most important staple food for a large part of the world's human population and most of it is consumed after milling. Rice bran, which is a by-product after milling, is rich in oil (12–23%), proteins (14–16%), vitamins and crude fiber (8–10%). Rice bran oil (RBO) is typically an oleic–linoleic-type fatty acid, and its physico–chemical properties qualify it for good quality edible oil [1]. Moreover, the RBO has been considered superior in comparison to other vegetable oils because it is a rich source of ω-3 and ω-6 fatty acids, and high level of unsaponifiables [2]. In addition, consumption of RBO leads to hypocholestrolemic effects due to the presence of primarily phytosterols, tocols (tocopherols

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11033-019-05144-4) contains supplementary material, which is available to authorized users.

✉ Kuldeep Singh
kuldeep.singh4@icar.gov.in

1 School of Agricultural Biotechnology, Punjab Agricultural University, Ludhiana, Punjab, India

2 School of Biology and Ecology, University of Maine, Orono, ME, USA

3 Department of Plant Breeding and Genetics, Punjab Agricultural University, Punjab, India

4 Department of Plant and Soil Science, Texas Tech University, Lubbock, TX, USA

5 ICAR- National Bureau of Plant Genetic Resources, New Delhi, India

and tocotrienols), γ-Oryzanol, and triterpene alcohol in its unsaponifiable fraction [3, 4]. Although rice bran could be a valuable source of nutrition, its use is severely limited by the rapid degradation of rice bran oil and more rapid development of hydrolytic rancidity in comparison to other vegetable oils. Immediately after the completion of rice milling, triacylglycerols (TAGs) in rice bran rapidly decompose, and the level of free fatty acids (FFAs) increases. As a consequence, the bran becomes unsuitable for human consumption or for the production of edible oil with acceptable quality [5–7].

TAGs are the primary reserve lipids inside the phospholipid membrane-bounded oil bodies of rice bran. It has been reported that phospholipid-degrading enzymes viz. phospholipases, lipid-oxidizing enzymes, and acyl hydrolases, are important contributors to membrane degradation [8, 9]. Of all these enzymes, phospholipase D (PLD) initiates an early step in the membrane degradation and seed deterioration by breaking down phospholipids into phosphatidic acid [10, 11]. The treatment of rice bran fraction with PLD enzyme leads to disintegration of the rice bran oil bodies and cause decline in the levels of phosphatidylcholine, followed by the decomposition of TAGs into FFAs [12–14]. Further, interaction of the produced FFAs with endosperm starches reduces the edibility of the rice. Moreover, action of lipoxygenases (LOXs) on FFAs, having a 1, 4-pentadiene structure, such as linoleic and linolenic acids, and their subsequent oxidation and decomposition, convert them into low molecular-weight volatile products (creating a stale flavor) [15–17]. Thus, PLD is the enzyme which initiates lipid decomposition followed by deterioration of the rice grain and rice bran fractions.

A survey of the rice genome database indicated the presence of 17 PLD genes in the genome including eight isoforms of *OsPLDα*, two isoforms of *OsPLDβ*, three isoforms of *OsPLDγ*, two isoforms of *OsPLD ξ,* one isoform of *OsPLD κ,* and one isoform of *OsPLD φ* [18]. Among these PLD isoforms, *OsPLDα1* has been found to be responsible for rice bran oil rancidity [19]. This gene has been cloned by Suzuki (2011) and is 6.28-kb in size including the promoter region, and is located on the antisense strand of rice chromosome 1 [20].

Although several methods are being used to suppress the *OsPLDα1* activity in rice bran but all these bring only partial inactivation; are associated with negative effects on the nutritional value of rice bran; and add to the cost of oil production and time stringency for treatment [21–23]. Therefore, a cost effective alternative is required to reduce the susceptibility of rice bran to hydrolytic rancidity. Although decline in the activity of this enzyme is must for maintaining the rice bran quality, breeding for reduced activity of this enzyme in rice seeds has not been attempted so far. The use of breeding techniques by transferring the desired allele could be effective in increasing the shelf life of rice bran against lipid hydrolysis if genetic variations exist between cultivars for this trait.

The problem of reduced gene pool of cultivated germplasm is particularly relevant in self-pollinated crops such as rice, where the level of genetic variation among cultivars can be lower than 5% of the total variation in natural populations. The wild relatives of rice constitute a major gene pool for rice improvement and have been used as sources of agronomically important genes particularly those involved in tolerance to biotic and abiotic stresses [24–26]. However, there is no report in literature focusing on the evaluation of wild rice germplasm for variation in OsPLDα1 enzyme activity and the allelic variation that underpins such phenotype.

In plants, a large number of transcription factors are known to control the expression of genes encoding metabolic enzymes by binding to target regulatory sequences usually within the upstream regions [27]. In this study, we performed a parallel analysis of variation in the upstream sequences of *OsPLDα1* orthologs and paralogs and their transcriptional activity during endosperm development across a panel of *Oryza* species and cultivars, in order to identify nucleotide variation in transcription factor binding motif (TBMF), number/frequency, and location of regulatory elements binding sites in promoter regions of allelic variants. The aim of present study was to dissect the molecular mechanism of *OsPLDα1* gene expression and to identify efficient promoters to be used in genetic engineering to improve the rice bran quality in elite lines.

Materials and methods

Plant material

A total of 26 representative accessions belonging to 4 wild *Oryza* species viz. *O. barthii* (n = 2)*, O. nivara* (n = 4)*, O. rufipogon* (n = 14)*, O. longistaminata* (n = 1), and African cultivated rice *O. glaberrima* (n = 5) were chosen for the study. The germplasm accessions were originally procured either from the International Rice Research Institute (IRRI), Philippines or from National Rice Research Institute (NRRI), Cuttack and being actively maintained at Punjab Agricultural University (PAU), Ludhiana. Selection of the wild species accessions for this study was done on the basis of country of their origin. The list of all the wild species accessions and their country of origin is given in the Table 1. The cultivars viz. Punjab Rice 114 (PR 114), PAU201, Nagina 22 (N22), IR64, Pusa 44, Minghui 63, Feng-Ai-Zhan, and Kitake were used as positive checks for the sequence analysis.

Table 1 List of wild *Oryza* species accessions selected for *OsPLDα1* promoter analysis

Species	Accession No.[a]	Country of origin
O. glaberrima	IRGC100854	Congo
O. glaberrima	IRGC101800	Senegal
O. glaberrima	IRGC102489	Liberia
O. glaberrima	IRGC102512	Liberia
O. glaberrima	IRGC102925	Burkina Faso
O. barthii	IRGC100117	Mali
O. barthii	IRGC106294	Chad
O. nivara	CR100400	India
O. nivara	CR100429	India
O. nivara	IRGC92713	Cambodia
O. nivara	IRGC105880	Bangladesh
O. rufipogon	IRGC80610	India
O. rufipogon	CR 100472A	India
O. rufipogon	CR100013	India
O. rufipogon	IRGC81976	Indonesia
O. rufipogon	IRGC83823	Vietnam
O. rufipogon	IRGC89224	Cambodia
O. rufipogon	IRGC103308	Taiwan
O. rufipogon	IRGC104308	Myanmar
O. rufipogon	IRGC104425	Thailand
O. rufipogon	IRGC105569	Cambodia
O. rufipogon	IRGC105902	Bangladesh
O. rufipogon	IRGC106162	Laos
O. rufipogon	IRGC106336	Cambodia
O. rufipogon	IRGC113652	Vietnam
O. longistaminata	IRGC105206	Ethiopia

[a]IRGC signifies the accession numbers of International Rice Germplasm Collection, Philippines; CR signifies the accession numbers of Central Rice Research Institute, Cuttack, India

Primer designing, DNA extraction and PCR amplification

The 1.6-kb promoter sequence (1.2-kb upstream and 0.4-kb downstream of the Translation Start Codon) of the *OsPLDα1* gene from *Oryza sativa japonica* cv. 'Nipponbare' was retrieved from NCBI (http://www.ncbi.nlm.nih.gov/nuccore/AB571657.1) and used as a reference for the present study. Three overlapping oligonucleotide primer pairs were designed (Supplementary Table S1) from this sequence using PerlPrimer (http://perlprimer.sourceforge.net/) software and re-checked by BLAST search to ensure that they matched uniquely with the expected positions in the rice genome. Leaf tissue from each selected wild species accession and cultivar was collected in the field and stored in − 80 °C deep freezer. The collected leaves were pulverized in liquid nitrogen using a sterilized and pre-chilled mortar and pestle to reduce the DNase activity. Further, the DNA extraction was conducted according to the improved CTAB (Cetyltrimethylammonium bromide) method modified from Saghai-Maroof et al. [28]. Quantity and quality of DNA from each sample was accessed on 0.8% agarose gel, and DNA samples were then diluted with 1X TE buffer and stored at − 20 °C for further use. PCR was performed in thermocycler (*Eppendorf* and *Biometra*) using 0.3 µl Phusion® high fidelity DNA polymerase, 3 µl of genomic DNA (20 ng/µl), 6 µl of 5× HF buffer, 6 µl of dNTPs (1 mM), 3 µl of primer (5 µM), and Nuclease Free Water to make a total volume of 30 µl. Cycling conditions were an initial denaturation of 94 °C for 5 min followed by 35 cycles of 1 min at 94 °C, 1 min at 55 °C, 1 min at 72 °C; concluding with a final extension of 72 °C for 5 min. PCR products were analysed by electrophoresis in ethidium bromide stained 1.0% agarose gels. 1 kb plus ladder (Thermo Scientific Generular) was used to estimate PCR fragment size. The amplified fragments were excised and purified using Wizard® SV PCR Clean-Up System (Promega, Inc.) as per manufacturer's protocol.

Sequencing of the PCR products and SNP detection

Sequencing was performed, in two replicates for each sample, on ABI Sequencer 3730xl using BigDye Terminator V3.1 cycle sequencing reaction kit at the School of Agricultural Biotechnology, Punjab Agricultural University, Ludhiana, India. The obtained *OsPLDα1* promoter sequences were extracted from chromatogram files using CHROMAS Lite 2.1.1 (http://technelysium.com.au/). Contigs were generated using *DNA Baser v4.23.0* (http://www.dnabaser.com/) software which also helps in automatic detection and trimming of low quality regions of the sequences. Alignment of the sequenced promoters against reference was carried out using ClustalX 2.1.1. Based on the alignment candidate SNPs (Single Nucleotide Polymorphisms) and InDels (Insertion-Deletions) were predicted. The candidate SNPs were then manually curated by analyzing and comparing chromatogram files with ClustalX alignment files. The bases which had Q-value below 20 were not considered as actual SNPs.

Phylogenetic analysis

A phylogenetic tree was generated using MEGA6 software (Tamura et al. 2013) using the multiple alignment file obtained earlier using ClustalX 2.1.1 [28]. The evolutionary distances were computed using the Maximum Composite Likelihood method with 1000 bootstrap (Tamura et al. 2004) and are in the units of the number of base substitutions per site [29].

In-silico analysis of the *OsPLDα1* promoters

Cis-elements in the *OsPLDα1* promoter regions of Nipponbare, wild species accessions and cultivated rice were identified using PLACE database (http://www.dna.affrc.go.jp/PLACE/).

OsPLDα1 gene expression analysis

To determine *OsPLDα1* gene expression levels and assess their association with *cis*-elements in *OsPLDα1* orthologs, total RNA was extracted 7 days after flowering from immature grains of selected wild species accessions. At that stage, plants have small grain size which leads to low amount of isolated RNA, if collected from the individual plant. Maintenance of wild species on a large area to obtain the required quantity of rice bran is highly tedious and time consuming. Moreover, cultivation of wild species on a large area may lead to the spreading of wild seeds into the broader area and can contaminate the cultivated rice land as these species have the weedy nature. Hence, only five plants per accession have been planted in the field each year to maintain the seed. Due to these limitations, from each selected wild species accession, nine panicles (3 panicles each from three different plants belonging to that accession) were collected after 7 days of flowering and harvested into liquid nitrogen to store at − 80 °C until RNA isolation. Further, same amount of tissue (immature grains) was taken from all the nine panicles and the collected tissue was weighed to get the equal amount (100 mg) for RNA extraction using the TRIzol® reagent (Thermo Fisher Scientific) according to the manufacturer's protocol. Therefore, each experimental replicate in this study represents RNA from three biological replicates.

The 12 wild species accessions, for expression study at *OsPLDα1* locus, were selected from the entire set of 26 wild species accessions on the basis of phylogenetic tree generated using their promoter sequences. From each phylogenetic group, one accession was selected to represent that group. In addition, maximum accessions were selected from *O. rufipogon* accessions as there was maximum variation obtained in the promoter region of this species. Moreover, RNA samples were collected from Punjab Agricultural University, Ludhiana, India and expression analysis was performed using lab facilities at the University of Maine, USA. Hence, a few wild species accessions were not included in the expression study as they remained in the vegetative state or flowering occurred very late after the sample collection, owing to their photoperiod sensitive nature.

The quantity of RNA in the samples was determined by using NanoDrop® ND-1000. RNA samples were reverse transcribed to obtain first stand cDNAs using an iScript cDNA kit (Bio-Rad laboratories, CA, USA). Total cDNA yields from the equimoler amounts of total RNA from each selected wild species accession were diluted to 1/10, and used as template for qRT-PCR performed with three replicates. Real-time PCR was performed in MyiQ™ thermal cycler (Bio-Rad Laboratories, CA, USA) using the iQ™ SYBR® Green Supermix (Bio-Rad) according to the manufacturers protocol. The thermal profile of the real-time system was one step at 95 °C for 30 s, followed by 40 cycles at 95 °C for 5 s (denaturation) and at 60 °C for 30 s (annealing and extension) followed by an added dissociation pattern.

OsPLDα1 genomic loci and transcript sequences (Locus ID *Os01g0172400*), were obtained from the RAP data base (http://rapdb.dna.affrc.go.jp/viewer/gbrowse/). Four exon-specific qRT-PCR primer pairs (PLDE1, PLDE2.1, PLDE2.2, and PLDE3) (Supplementary Table S2) were designed from the coding sequence of *OsPLDα1* gene using Primer-BLAST tool (http://www.ncbi.nlm.nih.gov/tools/primer-blast/). In addition, a primer pair (sense primer 5′-CATGAAGATCAAGGTTGTCGCTCC-3′ and antisense primer 5′- CCAGATTCTTCATACTCAGCCCTTG-3′) was also designed for *actin* gene (Locus ID *Os10g0510000*). Each primer was dissolved in 1X TE buffer (stock solution) to have master stock of 100 µM.

Transcript abundance was normalized relative to the baseline constitutive expression of a reference gene i.e., *actin-2* (*ACT2*). Relative expression levels were calculated using the $2^{-\Delta C_T}$ method [30]. According to this method, data have been normalized against the selected internal control (*ACT2*) to get ΔC_T values.

$$\Delta C_T = \left[C_T(\text{gene of interest}) - C_T(\text{internal control})\right] (\text{SampleA})$$

Calculated ΔC_T values of each of the three replicates of a sample were then log transformed into $2^{-\Delta C_T}$ value. Mean of the $2^{-\Delta C_T}$ values was used to measure relative expression of different samples and to generate the heat map. Standard deviation for each sample was also calculated using the $2^{-\Delta C_T}$ value of the replicates of a sample.

Results

Detection of SNPs in *OsPLDα1* promoter

The 1.6-kb upstream region of the *OsPLDα1* promoter obtained from selected wild species accessions and cultivars was used for SNP/InDel detection. Multiple alignments of *OsPLDα1* promoter sequences to the Nipponbare resulted in detection of 39 SNPs (Table 2). Of the detected SNPs, 32 were identified in the 1.2 kb region upstream of the translation start site. These SNPs represent base transitions (22) which were more frequent than base transversion (17) mutations. Interestingly, all the accessions of *O. glaberrima* were found to be sharing similar mutations in *OsPLDα1*

Table 2 The nucleotide variations observed in the *OsPLDα1* promoter sequence of wild *Oryza* species accessions and *Oryza* cultivars in comparison to the Nipponbare *OsPLDα1* promoter

S. No.	Position[a]	Alleles	Wild species accessions/cultivars
1	− 1197	C/T[b]	*O. rufipogon* (IRGC89224)
2	− 1192	C/T[b]	*O. rufipogon* (IRGC104425); IR64, N22, Pusa44, PR114, PAU201, Minghui63, Feng-Ai-Zhan
3	− 1189	A/G[b]	*O. glaberrima* (IRGC100854, IRGC101800, IRGC102489, IRGC102512, IR102925); *O. barthii* (IRGC100117, IRGC106294); *O. nivara* (CR100400, CR100429, IRGC92713, IRGC105880); *O. rufipogon* (IRGC80610, CR100472A, CR100013, IRGC81976, IRGC83823, IRGC89224, IRGC103308, IRGC104308, IRGC104425, IRGC104867, IRGC105569, IRGC105902, IRGC106162, IRGC106336, IRGC113652); *O. longistaminata* (IRGC105206); Kitake, IR64, N22, Pusa44, PR114, PAU201, Minghui63, Feng-Ai-Zhan
4	− 1139	C/T[b]	*O. nivara* (CR100429); *O. rufipogon* (CR100472A, IRGC89224, IRGC113652)
5	− 1111	A/G[b]	*O. glaberrima* (IRGC100854, IRGC101800, IRGC102489, IRGC102512, IRGC102925); *O. barthii* (IRGC100117, IRGC106294); *O. nivara* (CR100429); *O. rufipogon* (CR100013, CR100472A, IRGC104308, IR105902, IRGC113652, IRGC106336, IRGC105569, IRGC83823, IRGC81976, IRGC89224); *O. longistaminata* (IRGC105206)
6	− 1108	A/G[b]	*O. nivara* (IRGC105880)
7	− 1103	C/A[c]	*O. rufipogon* (IRGC80610, IRGC104425, IRGC105569); *O. nivara* (IRGC92713, IRGC105880, CR100400); IR64, N22, Pusa44, PR114, PAU201, Minghui63, Feng-Ai-Zhan
8	− 1103	C/T[b]	*O. rufipogon* (IRGC113652)
9	− 1093	T/C[b]	*O. glaberrima* (IRGC100854, IRGC101800, IRGC102489, IRGC102512, IRGC102925); *O. barthii* (IRGC100117, IRGC106294); *O. rufipogon* (IRGC81976, IRGC89224, IRGC104308, IRGC103308, IRGC105569, IRGC105902, IRGC106162, IRGC106336, IRGC113652, CR100472A); *O. nivara* (CR100429); *O. longistaminata* (IRGC105206);
10	− 1072	C/T[b]	*O. glaberrima* (IRGC100854, IRGC101800, IRGC102489, IRGC102512, IRGC102925)
11	− 1071	C/A[c]	*O. glaberrima* (IRGC100854, IRGC101800, IRGC102489, IRGC102512, IRGC102925); *O. barthii* (IRGC100117)
12	− 1064	C/A[c]	*O. rufipogon* (IRGC80610, IRGC104425); *O. nivara* (IRGC105880, IRGC92713, CR100400); IR64, N22, Pusa44, PR114, PAU201, Minghui63, Feng-Ai-Zhan
13	− 1041	A/G[b]	*O. rufipogon* (IRGC106336)
14	− 1038	C/A[c]	*O. rufipogon* (CR100013, IRGC83823, IRGC106162); *O. longistaminata* (IRGC105206)
15	− 1024	C/G[c]	*O. barthii* (IRGC106294)
16	− 1021	G/C[c]	*O. barthii* (IRGC106294)
17	− 927	A/G[b]	*O. rufipogon* (IRGC81976)
18	− 902	C/T[b]	*O. rufipogon* (IRGC103308, IRGC105902)
19	− 895	C/A[c]	*O. longistaminata* (IRGC105206)
20	− 858	G/T[c]	*O. nivara* (IRGC105880, IRGC92713, CR100400); *O. rufipogon* (IRGC80610); IR64, N22, Pusa44, PR114, PAU201, Minghui63, Feng-Ai-Zhan
21	− 828	G/A[b]	*O. rufipogon* (IRGC103308)
22	− 644	C/T[b]	*O. rufipogon* (IRGC81976); *O. rufipogon* (IRGC106336, IRGC113652)
23	− 595	C/G[c]	*O. rufipogon* (IRGC81976, IRGC106336, IRGC113652)
24	− 588	G/A[b]	*O. glaberrima* (IR100854, IR101800, IR102489, IR102512, IR102925); *O. barthii* (IR100117)
25	− 455	A/G[b]	*O. rufipogon* (CR100013, CR100472A, IRGC81976, IRGC83823, IRGC105569); *O. nivara* (CR100429, IRGC92713); *O. longistaminata* (IRGC105206); IR64, N22, Pusa44, PR114, PAU201, Minghui63, Feng-Ai-Zhan
26	− 443	A/G[b]	*O. nivara* (IRGC105880)
27	− 437	C/G[c]	*O. rufipogon* (IRGC103308, IRGC104308, IRGC106162, IRGC106336, IRGC113652)
28	− 396	G/A[b]	*O. rufipogon* (IRGC103308)
29	− 247	C/A[c]	*O. nivara* (IRGC105880); N22
30	− 227	C/A[c]	*O. glaberrima* (IRGC100854, IRGC101800, IRGC102489, IRGC102512, IRGC102925)
31	− 226	C/A[c]	*O. glaberrima* (IRGC100854, IRGC101800, IRGC102489, IRGC102512, IRGC102925)
32	− 154	C/A[c]	*O. rufipogon* (IRGC83823)
33	+ 37	C/A[c]	IR64
34	+ 39	A/C[c]	IR64

Table 2 (continued)

S. No.	Position[a]	Alleles	Wild species accessions/cultivars
35	+48	A/G[b]	*O. glaberrima* (IRGC100854, IRGC101800, IRGC102489, IRGC102512, IRGC102925); *O. barthii* (IRGC100117, IRGC106294); *O. nivara* (CR100400, CR100429, IRGC92713, IRGC105880); *O. rufipogon* (IRGC80610, CR100472A, CR100013, IRGC81976, IRGC83823, IRGC89224, IRGC103308, IRGC104308, IRGC104425, IRGC104867, IRGC105569, IRGC105902, IRGC106162, IRGC106336, IRGC113652); *O. longistaminata* (IRGC105206); Kitake, IR64, N22, Pusa44, PR114, PAU201, Minghui63, Feng-Ai-Zhan
36	+93	A/C[c]	IR64
37	+114	A/G[b]	*O. glaberrima* (IRGC100854, IRGC101800, IRGC102489, IRGC102512, IRGC102925); *O. barthii* (IRGC100117, IRGC106294); *O. nivara* (CR100400, CR100429, IRGC92713, IRGC105880); *O. rufipogon* (IRGC80610, CR100472A, CR100013, IRGC81976, IRGC83823, IRGC89224, IRGC103308, IRGC104308, IRGC104425, IRGC104867, IRGC105569, IRGC106162, IRGC106336, IRGC113652); IR64, N22, Pusa44, PR114, PAU201, Minghui63, Feng-Ai-Zhan
38	+173	T/C[b]	*O. nivara* (CR100400, CR100429, IRGC92713); *O. rufipogon* (IRGC89224, CR 100472A)
39	+352	T/C[b]	*O. nivara* (CR100400, CR100429); *O. rufipogon* (IRGC89224)

[a]SNPs were calculated from translation start site of *OsPLDα1* gene

Transitions[b] and transversions[c] observed as nucleotide substitutions

promoter, based on their alignment with the Nipponbare promoter. Moreover, SNPs C/T_{-1072}, A/C_{-227}, and A/C_{-226} were found specific to the *O. glaberrima* species. Out of the 9 SNPs identified in *O. barthii* accessions, C/G_{-1024} and G/C_{-1021} were specific to the IRGC106294 accession; C/A_{-1071} and G/A_{-588} were detected in the IRGC100117 accession and in *O. glaberrima* accessions as well; and the remaining five SNPs were detected in all the accessions of *O. barthii* and *O. glaberrima*. Among the 15 SNPs detected in five *O. nivara* accessions, only two SNPs including A/G_{-1108} and A/G_{-443}, detected in the IRGC105880 accession, were found specific to this species while rest of the SNPs were present in the accessions of other species as well.

Of all the wild species and cultivars under study, the maximum number of SNPs (24) has been identified in the *O. rufipogon* species from which 11 were found specific to the *O. rufipogon* accessions (Table 2). In addition, two small insertions including CTC$_{-253}$ and GCTT$_{-128}$ have also been detected in the *O. rufipogon* accessions. The CTC insertion was found in IRGC89224, IRGC104425, and IRGC105902 accessions while the GCTT insertion was found in IRGC104308, IRGC103308, IRGC106162, and IRGC81976 accessions. Of the 11 SNPs identified in the eight cultivars under study, only 3 including C/A_{+37}, A/C_{+39}, and A/C_{+93} were found specific to the cultivar IR64 while rest of the SNPs were found in the wild species accessions as well. Kitake, in comparison to other cultivars, showed minimal variations as the SNPs C/T_{-1192}, C/A_{-1064}, G/T_{-858}, and A/G_{-455} were detected in all the cultivars except Kitake.

Phylogenetic relationships among *Oryza* accessions based on *OsPLDα1* sequence variation

Phylogenetic analysis of the *OsPLDα1* promoter region revealed the divergence of 26 wild *Oryza* species accessions and 8 *Oryza* cultivars from the reference sequence of *japonica* cultivar Nipponbare (Fig. 1). Phylogenetic tree showed that all accessions of the wild *Oryza* species having African centre of origin, *O. glaberrima, O. barthii,* and *O. longistamina*, were more closely related to each other in comparison to the accessions of *O. rufipogon* and *O. nivara* species, which originated in South-Asian nations. Moreover, three *O. rufipogon* accessions (IRGC89224, IRGC104425, and IRGC105902) were found to be at the maximum distance from the reference, thus highlighting the presence of useful sequence variations which may further lead to the presence/absence of *cis*-elements involved in regulation of *OsPLDα1* expression. Furthermore, differences observed between the *indica* and *japonica* cultivars were greater than within *indica* and *japonica* cultivars as *japonica* cvs. Nipponbare and Kitake were more similar to each other than to *indica* cvs. IR64, N22, Pusa44, PR114, PAU201, Minghui63, and Feng-Ai-Zhan.

Regulatory sequences in *OsPLDα1* promoter

The analysis revealed the presence of 92 classes of sequence motifs representing known and novel regulatory *cis*-elements (341 including copy numbers) in the *OsPLDα1* promoters of the wild species accessions and cultivars of rice. Further, comparison of the identified *cis*-elements with Nipponbare showed the copy number variations of 48 *cis*-elements (Table 3). In addition, 10 *cis*-elements including ARFAT, IBOXCORENT, MYB1LEPR, NTBBF1ARROLB, PYRIMIDINEBOXOSRAMY1A, ROOTMOTIFTAPOX1, SEBFCONSSTPR10A, SORLIP2AT, TBOXATGAPB, and TGTCACACMCUCUMISIN, which were detected in at least one of wild species accession, were found missing in Nipponbare. On the contrary, ACGTOSGLUB1 was detected in Nipponbare, but found missing in four wild

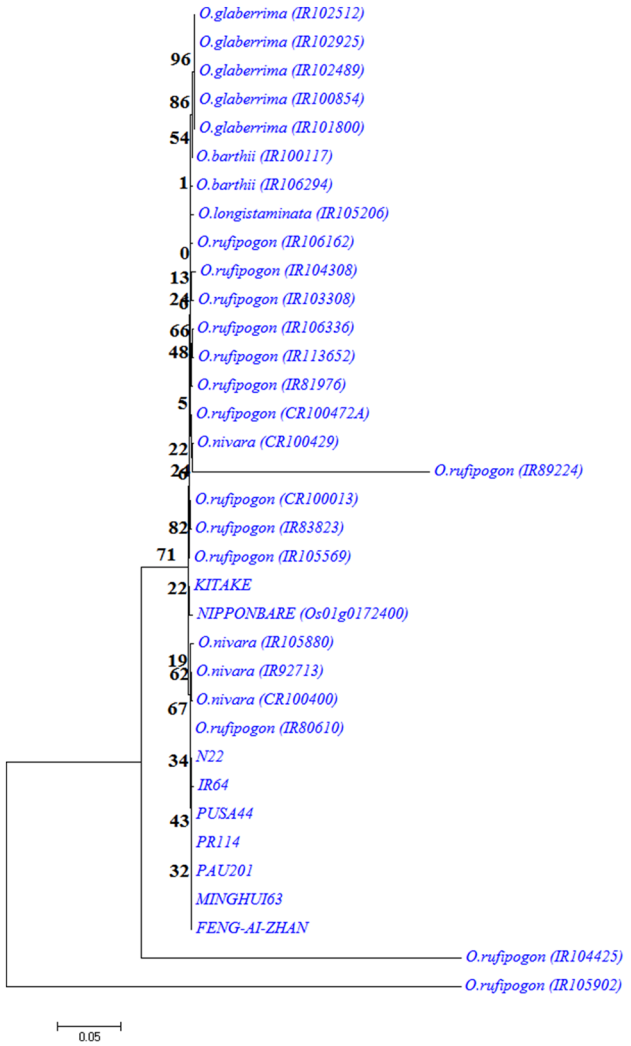

Fig. 1 Phylogenetic relationship across Nipponbare (Locus ID *Os01g0172400*), wild species accessions, and cultivars of rice based on *OsPLDα1* promoter sequence variation. The neighbor-joining tree was drawn based on the nucleotide sequences of the *OsPLDα1* promoters from 26 wild *Oryza* species accessions, 8 *Oryza* cultivars, and using Nipponbare *OsPLDα1* promoter as the reference sequence. In all, 1,000 bootstrap replicas were applied. The tree indicates the three *Oryza rufipogon* accessions (IRGC105902, IRGC104425 and IRGC89224) at maximum distance from Nipponbare. *Japonica* cultivar Kitake shows close relatedness to the Nipponbare in comparison to the *indica* cultivars (IR64, N22, Pusa44, PR114, PAU201, Minghui63, and Feng-Ai-Zhan)

species accession (CR100429, IRGC92713, IRGC105880, and IRGC80610). The role played by *cis*-elements, in regulating the gene expression, and their location on the promoter sequence was identified (Supplementary Table S3) and most of the regulatory elements (268 including copy number) were found to be located in the upstream region of the promoter. Furthermore, mapping and alignment of the *cis*-elements identified in the *OsPLDalpha1* promoter of Nipponbare to the wild species accessions showed the relative position of all the detected *cis*-elements, variations in copy numbers, and presence/absence of *cis*-elements between as well as within the accessions of different species (Fig. 2).

Differential expression of *OsPLDα1*

The *OsPLDα1* gene expression was compared across the accessions selected on the basis of presence/absence and copy number variations of *cis*-elements in *OsPLDα1* promoter. Expression profiling, generated using four exon-specific qRT-PCR primer pairs (Supplementary Table S2), indicated significant expression differences in the selected accessions (Fig. 3). The four qRT-PCR primer pairs were designed in a way to determine expression differences as well as abundance of different *OsPLDα1* transcript forms (Fig. 5). Expression analysis using qRT-primer PLDE1, designed from the first exon (Fig. 3a), revealed the presence as well as variations in abundance of 5′ truncated *OsPLDα1* mRNA in the accessions. Minimum expression for first exon was detected in the IRGC104425 accession of *O. rufipogon* while IRGC105279 accession of *O. longistaminata* had the highest expression levels. Further, expression levels for qRT-PCR primers designed from 5′ (PLDE2.1) and 3′ (PLDE2.2) ends of second exon, revealed differences in the expression levels within the same accession (Fig. 3b, c). It suggests the presence of truncations, within the second exon of *OsPLDα1* gene, during the process of transcription. For the second exon, IRGC89224 had the lowest level for 5′end transcripts while for the 3′end transcript levels were lowest and comparable in IRGC89224 and IRGC105902 accessions. In addition, there were differences between all the selected accessions for the abundance of transcripts containing 3′end and 5′end of the second exon.

Lowest level of *OsPLDα1* mRNA transcripts having third exon (Fig. 3d) were detected in the IRGC89224 accession of *O. rufipogon* while the highest levels were observed in *O. barthii* accession (IRGC106294). Among all the wild species accessions, IRGC105569 accession of *O. rufipogon* was found closely related to the Nipponbare cultivar, and thus it was used as control to compare the transcript levels in the other wild species accessions. The levels of cumulative relative expression for all the four exon specific primers were 0.20-, 0.26-, 0.41-, 0.44-, 0.61-, 0.62-, 0.75-, 1.37-, 1.91-, and 2.49-fold in IRGC89224 (*O. rufipogon*), IRGC105902 (*O. rufipogon*), IRGC104425 (*O. rufipogon*), IRGC106162 (*O. rufipogon*), IRGC104308 (*O. rufipogon*), CR100013 (*O. rufipogon*), IRGC103308 (*O. rufipogon*), CR100400 (*O. nivara*), IRGC106294 (*O. barthii*), IRGC102489 (*O. glaberrima*) accessions in comparison to the control accession (IRGC105569). Thus, the expression profiling showed lowest expression of *OsPLDα1* gene in IRGC89224 followed by IRGC105902

Table 3 Presence/absence and copy number variations of *cis*-elements in the *OsPLDα1* promoter region of wild species accessions and cultivars of rice in comparison to the reference promoter of Nipponbare

Cis-element[a]	Wild *Oryza* species accessions/*Oryza* cultivars[b]	
	+	–
2SSEEDPROTBANAPA (1x), ACGTCBOX (2x), ACGTTBOX (2x), AMMORESIVDCRNIA1(1x), ANAERO2CONSENSUS (3x) BOXIINTPATPB (1x), CANBNNAPA (2x), CPBCSPOR (1x), EBOXBNNAPA (18x), ELRECOREPCRP1 (2x), GCCCORE (1x), GT1CORE (1x), INTRONLOWER (1x), LTRE1HVBLT49 (1x), MYB2AT (1x), MYB2CONSENSUSAT (2x), MYBCORE (5x), MYBCOREATCYCB1 (1x), MYCATERD1 (1x), MYCATRD22 (1x), MYCCONSENSUSAT (18x), NODCON1GM (1x), OSE1ROOTNODULE (1x), P1BS (2x), POLASIG1 (1x), POLASIG3 (1x), RAV1AAT (1x), RYREPEATBNNAPA (1x), SORLIP5AT (1x), TATABOX5 (1x), TATABOXOSPAL (1x), TATCCAOSAMY (1x), WBBOXPCWRKY1 (2x), WBOXNTCHN48 (1x)	1x (W1 to W26; C1 to C8)	None
ABRELATERD1 (2x)	2x (W1, W2, W3, W4, W5, W6, W7, W9, W13, W14, W15, W16, W17, W18, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8); 1x (W8, W10, W11, W12)	None
ABRERATCAL (2x)	2x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W13, W14, W15, W16, W17, W20, W21, W22, W26; C1 to C8); 1x (W18, W19, W23, W24, W25)	None
ACGTATERD1 (12x)	12x (W1, W2, W3, W4, W5, W6, W7, W9, W13, W14, W15, W16, W17, W19, W20, W21, W22, W23, W24, W25, W26); 10x (W8, W10, W11, W12, W18; C1 to C8)	None
ACGTOSGLUB1 (1x)	1x (W1, W2, W3, W4, W5, W6, W7, W9, W13, W14, W15, W16, W17, W18, W19, W20, W21, W22, W23, W24, W25, W26)	W8, W10, W11, W12; C1 to C8
ARFAT (–)	1x (W17)	W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W13, W14, W15, W16, W18, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8
ARR1AT (12x)	13x (W13, W15, W16, W24, W25); 12x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W14, W17, W18, W19, W20, W21, W22, W23, W26; C1 to C8); 11x (W12)	None
BIHD1OS (4x)	4x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W12, W13, W14, W15, W16, W17, W18, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8); 3x (W11)	None
CAATBOX1 (6x)	7x (W11); 6x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W12, W13, W14, W15, W16, W17, W18, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8)	None

Table 3 (continued)

Cis-element[a]	Wild *Oryza* species accessions/*Oryza* cultivars[b]	
	+	–
CACTFTPPCA1 (22x)	23x (W1, W2, W3, W4, W5, W6, W10, W11, W12; C1 to C8); 22x (W7, W8, W13, W14, W15, W16, W18, W19, W20, W21, W22, W23, W24, W25, W26); 21x (W9, W17)	None
CBFHV (3x)	3x (W1 to W25; C1 to C8); 2x (W26)	None
CCAATBOX1 (3x)	3x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W12, W13, W14, W15, W16, W17, W18, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8); 4x (W11)	None
CGACGOSAMY3 (16x)	16x (W7, W8, W9, W10, W11, W12, W13, W15, W16, W17, W18, W19, W20, W21, W22, W26; C1 to C8); 15x (W1, W2, W3, W4, W5, W6, W14, W23, W24, W25)	None
CGCGBOXAT (18x)	18x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W13, W14, W15, W16, W17, W20, W21, W26; C1 to C8); 16x (W19, W22, W23, W24, W25); 14x (W18)	None
CIACADIANLELHC (3x)	3x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W14, W15, W17, W18, W19, W20, W21, W22, W23, W24, W25, W26, C1, C3, C4, C5, C6, C7, C8); 2x (W13, W16, C2)	None
CURECORECR (6x)	6x (W1, W2, W3, W4, W5, W6, W7, W13, W14, W15, W16, W18, W19, W20, W21, W22, W23, W24, W25, W26); 4x (W9, W10, W11, W12, W17; C1 to C8); 2x (W8)	None
DOFCOREZM (11x)	12x (W13, W18, W19); 11x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W14, W15, W16, W17, W20, W21, W22, W23, W24, W25, W26; C1 to C8)	None
DPBFCOREDCDC3 (5x)	6x (W8, W9, W10, W13, W17); 5x (W1, W2, W3, W4, W5, W6, W7, W11, W12, W14, W15, W16, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8); 4x (W18)	None
DRE2COREZMRAB17 (2x)	2x (W1 to W25; C1 to C8); 1x (W26)	None
DRECRTCOREAT (3x)	3x (W1 to W25; C1 to C8); 2x (W26)	None
EECCRCAH1 (2x)	3x (W18); 2x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W13, W14, W15, W16, W17, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8)	None
GATABOX (8x)	8x (W8, W10, W11, W12, W16, W19, W20, W23; C1 to C8); 7x (W1, W2, W3, W4, W5, W6, W7, W9, W13, W14, W15, W17, W18, W21, W22, W24, W25, W26)	None
GT1CONSENSUS (11x)	12x (W18); 11x (W1, W2, W3, W4, W5, W6, W13, W14, W15, W16, W17, W19, W21, W22, W23, W24, W25, W26); 10x (W7, W8, W9, W10, W11, W12, W20; C1 to C8)	None
GT1GMSCAM4 (4x)	5x (W18); 4x (W1, W2, W3, W4, W5, W6, W7, W8, W9 W10, W11, W12, W13, W14, W15, W16, W17, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8)	None

Table 3 (continued)

Cis-element[a]	Wild *Oryza* species accessions/*Oryza* cultivars[b]	
	+	–
GTGANTG10 (7x)	8x (W1, W2, W3, W4, W5, W6, W7, W9, W13, W14, W15, W16, W17, W18, W19, W21, W22, W24, W25, W26); 7x (W8, W10, W11, W12, W20, W23; C1 to C8)	None
HEXAMERATH4 (2x)	3x (W18, W19, W23, W24, W25); 2x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W13, W14, W15, W16, W17, W20, W21, W22, W26; C1 to C8)	None
IBOX (1x)	2x (W16); 1x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W13, W14, W15, W17, W18, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8)	None
IBOXCORE (2x)	3x (W16); 2x (W1, W2, W3, W4, W5, W6, W7, W9, W13, W14, W15, W17, W18, W19, W22, W23, W24, W25, W26); 1x (W8, W10, W11, W12, W20, W21; C1 to C8)	None
IBOXCORENT (–)	1x (W16)	W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W13, W14, W15, W17, W18, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8
LTRECOREATCOR15 (4x)	5x (W15); 4x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W13, W14, W16, W17, W18, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8)	None
MYB1LEPR (–)	1x (W18)	W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W13, W14, W15, W16, W17, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8
MYBPZM (3x)	4x (C2); 3x (W1 to W26; C1, C3, C4, C5, C6, C7, C8)	None
MYBST1 (2x)	3x (W16); 2x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W13, W14, W15, W17, W18, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8)	None
NODCON2GM (7x)	8x (W9, W14, W17); 7x (W1, W2, W3, W4, W5, W6, W7, W14, W15, W16, W18, W19, W22, W23, W24, W25, W26, C2); 6x (W8, W10, W11, W12, W13, W20, W26, C1, C3, C4, C5, C7, C8)	None
NTBBF1ARROLB (–)	1x (W1, W2, W3, W4, W5, W6)	W7, W8, W9, W10, W11, W12, W13, W14, W15, W16, W17, W18, W19, W20, W21, W22, W23, W24, W25, W26, C1 to C8
OSE2ROOTNODULE (7x)	8x (W9, W13, W14, W17); 7x (W1, W2, W3, W4, W5, W6, W7, W15, W16, W18, W19, W21, W22, W23, W24, W25, W26, C2); 6x (W8, W10, W11, W12, W20, C1, C3, C4, C5, C7, C8)	None
POLLEN1LELAT52 (6x)	6x (W10, W11, W12, W13, W14, W15, W16, W20; C1 to C8); 5x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W17, W18, W19, W21, W22, W23, W24, W25, W26)	None

Table 3 (continued)

Cis-element[a]	Wild *Oryza* species accessions/*Oryza* cultivars[b]	
	+	–
PRECONSCRHSP70A (3x)	3x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W14, W15, W16, W17, W18, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8); 2x (W13)	None
PYRIMIDINEBOXOSRAMY1A (–)	2x (W15); 1x (W6, W7, W9, W13, W17, W18, W19, W21, W22, W23, W24, W25, W26)	W1, W2, W3, W4, W5, W8, W10, W11, W12, W14, W16, W20; C1 to C8
RAV1BAT (1x)	1x (W1, W2, W3, W4, W5, W6, W8, W9, W10, W11, W12, W13, W14, W15, W16, W17, W18, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8)	None
RBCSCONSENSUS (1x)	1x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W13, W14, W15, W16, W17, W18, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8)	W12
REBETALGLHCB21 (1x)	2x (W16); 1x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W13, W14, W15, W17, W18, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8)	None
RHERPATEXPA7 (2x)	2x (W1, W2, W3, W4, W5, W6, W7, W9, W13, W14, W15, W16, W17, W18, W19, W20, W21, W22, W23, W24, W25, W26); 1x (W8, W10, W11, W12; C1 to C8)	None
ROOTMOTIFTAPOX1 (–)	1x (W14, W16, W23, W26)	W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W13, W15, W17, W18, W19, W20, W21, W22, W24, W25; C1 to C8
SEBFCONSSTPR10A (–)	1x (W17)	W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W13, W14, W15, W16, W18, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8
SEF4MOTIFGM7S (2x)	2x (W1, W2, W3, W4, W5, W6, W7, W9, W13, W14, W15, W16, W17, W19, W22, W23, W24, W25, W26); 1x (W8, W10, W11, W12, W18, W20, W21; C1 to C8)	None
SITEIIATCYTC (1x)	2x (W19, W24); 1x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W13, W14, W15, W16, W17, W18, W20, W21, W22, W23, W25, W26; C1 to C8)	None
SORLIP1AT (6x)	6x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W12, W13, W14, W15, W16, W17, W19, W20, W21, W23, W24, W25, W26, C1, C3, C4, C5, C6, C7, C8); 5x (W11, W18, W22, C2)	None
SORLIP2AT (–)	2x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W13, W14, W15, W16, W17, W18, W19, W20, W21, W23, W24, W25; C1 to C8)	W22, W26
SREATMSD (1x)	2x (W16); 1x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W13, W14, W15, W17, W18, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8)	None

Table 3 (continued)

Cis-element[a]	Wild *Oryza* species accessions/*Oryza* cultivars[b]	
	+	–
SURECOREATSULTR11 (4x)	3x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W13, W14, W15, W16, W17, W18, W19, W21, W22, W23, W24, W25); 2x (W20, W26; C1 to C8)	None
TAAAGSTKST1 (2x)	2x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W13, W14, W16, W17, W18, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8); 1x (W15)	None
TBOXATGAPB (–)	1x (W18)	W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W13, W14, W15, W16, W17, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8
TGTCACACMCUCUMISIN (–)	1x (W1, W2, W3, W4, W5, W6, W7, W9, W13, W14, W16, W18, W19, W21, W22, W24, W25, W26)	W8, W10, W11, W12, W15, W17, W20, W23; C1 to C8
WBOXATNPR1 (5x)	5x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W13, W14, W15, W16, W17, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8); 4x (W18)	None
WBOXHVISO1 (3x)	3x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W13, W14, W15, W16, W17, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8); 2x (W18)	None
WBOXNTERF3 (7x)	7x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W11, W12, W13, W14, W15, W16, W17, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8); 6x (W18)	None
WRKY71OS (11x)	11x (W1, W2, W3, W4, W5, W6, W7, W8, W9, W10, W12, W13, W14, W15, W16, W17, W19, W20, W21, W22, W23, W24, W25, W26; C1 to C8); 10x (W11, W18)	None

W1 *O. glaberrima* (*IR100854*)	W10 *O. nivara* (*IR92713*)	W19 *O. rufipogon* (*IR104308*)	C1 KITAKE
W2 *O. glaberrima* (*IR101800*)	W11 *O. nivara* (*IR105880*)	W20 *O. rufipogon* (*IR104425*)	C2 IR64
W3 *O. glaberrima* (*IR102489*)	W12 *O. rufipogon* (*IR80610*)	W21 *O. rufipogon* (*IR105569*)	C3 N22
W4 *O. glaberrima* (*IR102512*)	W13 *O. rufipogon* (*CR 100472A*)	W22 *O. rufipogon* (*IR105902*)	C4 PUS44
W5 *O. glaberrima* (*IR102925*)	W14 *O. rufipogon* (*CR100013*)	W23 *O. rufipogon* (*IR106162*)	C5 PR114
W6 *O. barthii* (*IR100117*)	W15 *O. rufipogon* (*IR81976*)	W24 *O. rufipogon* (*IR106336*)	C6 PAU201
W7 *O. barthii* (*IR106294*)	W16 *O. rufipogon* (*IR83823*)	W25 *O. rufipogon* (*IR113652*)	C7 MINGHUI63
W8 *O. nivara* (*CR100400*)	W17 *O. rufipogon* (*IR89224*)	W26 *O. longistaminalis* (*IR105206*)	C8 FENG-AI-ZHAN
W9 *O. nivara* (*CR100429*)	W18 *O. rufipogon* (*IR103308*)		

[a]Numbers in parenthesis indicate the copy number (x) of *cis*-elements in Nipponbare; (-) indicates the absence of respective *cis*-element in the Nipponbare *OsPLDα1* promoter

[b]Presence (+)/absence (-) and copy numbers (x) of *cis*-elements in the wild species accessions and cultivars of rice. W and C numbers denotes wild species accessions and cultivars, respectively

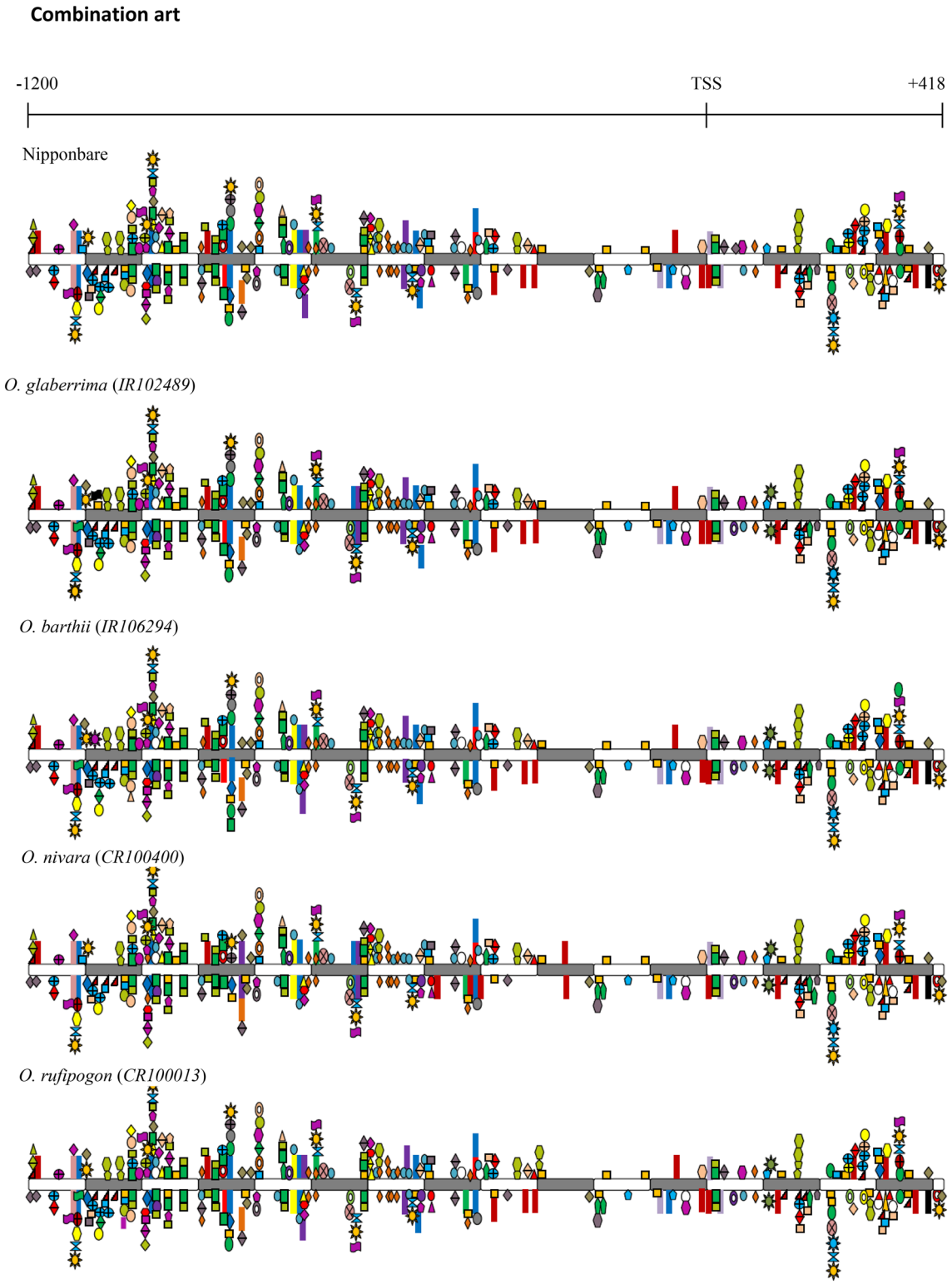

Fig. 2 *Cis*-regulatory elements maps in the *OsPLDα1* promoter (1.2 kb upstream and 0.4 kb downstream of the Translation Start Codon) in the Nipponbare and wild *Oryza* species accessions. The colored shapes on the promoter sequence indicate the position of the respective *cis*-elements. This mapping and alignment of *cis*-elements indicates the presence/absence as well as copy number variations between as well as within the accessions of different species

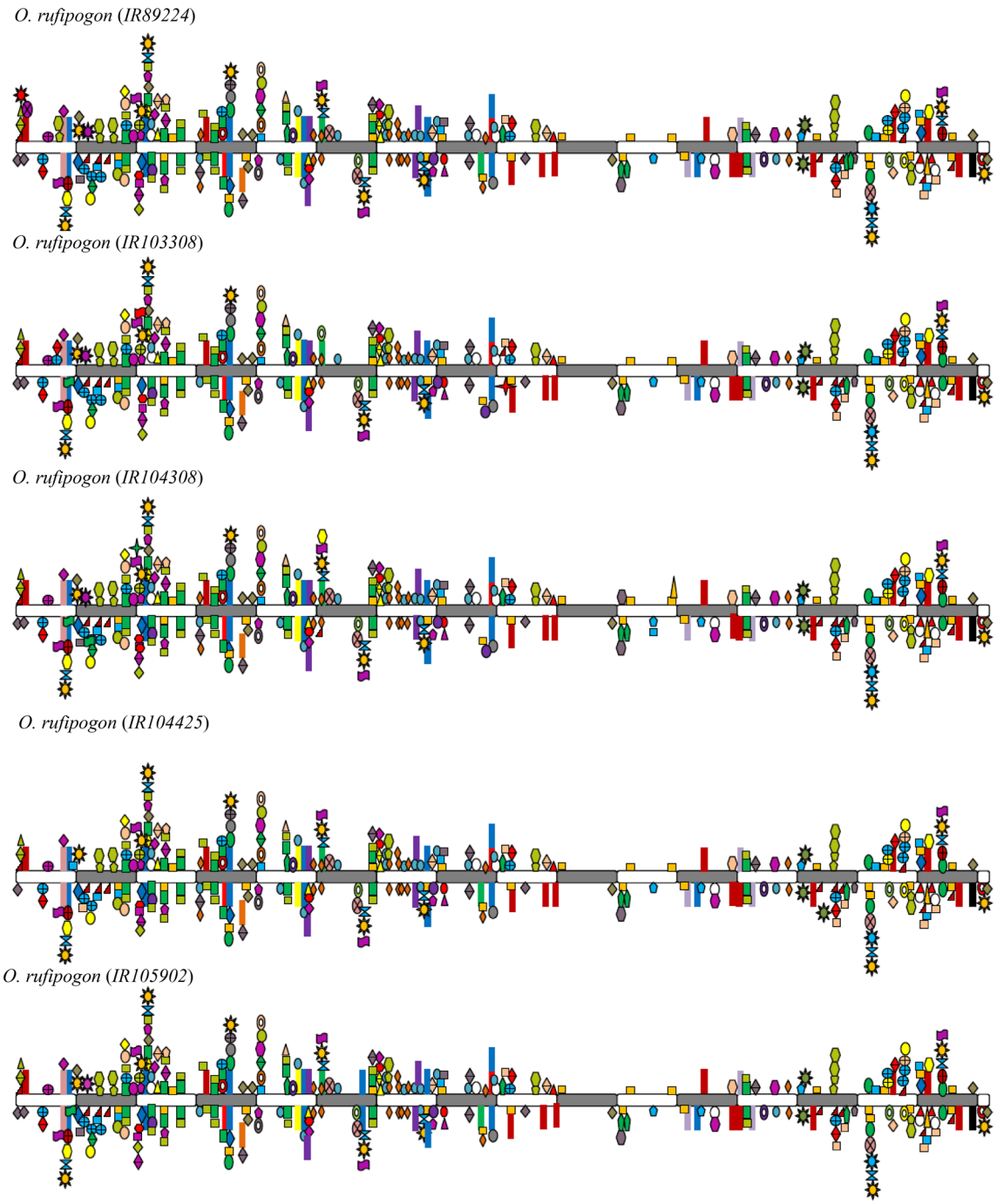

Fig. 2 (continued)

and IRGC104425 accessions of *O. rufipogon*, while O. *nivara, O. barthii,* and *O. glaberrima* accessions had *OsPLDα1* gene expressions higher than the control. Furthermore, hierarchical clustering dendrogram represents that expression values for all the primers in IRGC104425 and IRGC105902 accessions of *O. rufipogon* were falling in the green category (low expression values); however, the black color obtained in case of IRGC104425 was more skewed towards the lower side i.e., the green side (Fig. 4). For all the other accessions, at least one of the exon specific expression values was found falling in the red category and/or was more skewed towards the higher value (depicted in the red color). Thus, in addition to IRGC89224 and IRGC105902 accessions, IRGC104425 accession of *O. rufipogon* was concluded to have low level of *OsPLDα1* gene expression.

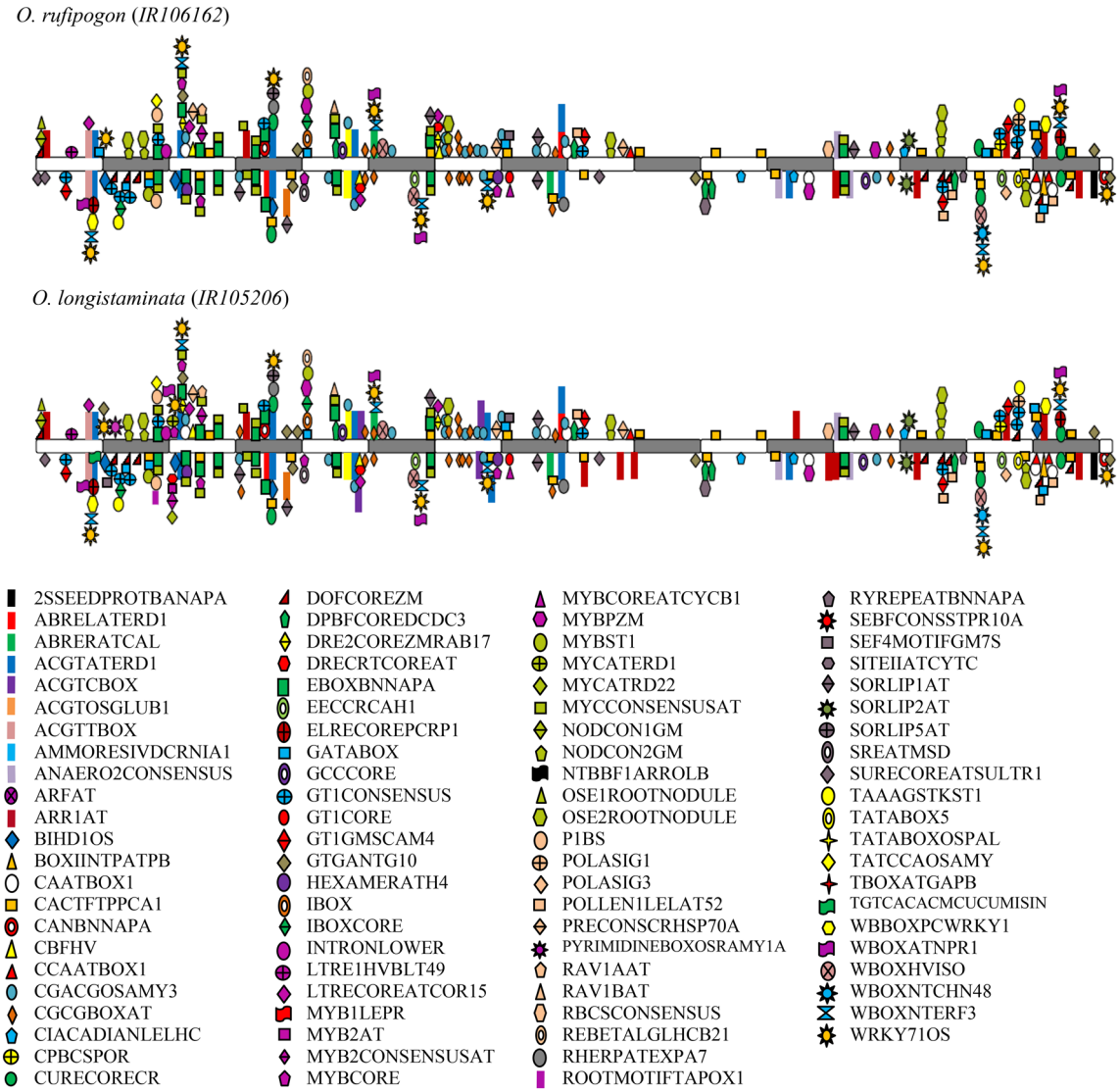

Fig. 2 (continued)

Discussion

The PLD enzyme has been known responsible for the rapid degradation of rice bran oil, and consequently severely limits the use of rice bran for human consumption. Triacylglycerols, the main lipids of rice bran, occur in oil bodies with phospholipid membranes that are disintegrated by PLD [31]. Ueki et al. (1995) purified a PLD protein (*OsPLDα1*) from the rice bran fraction which plays an important role in the rice bran oil rancidity. However, there is no report in literature on detection of superior *OsPLDα1* promoters in the wild germplasm of rice to improve the rice bran quality. Moreover, the molecular mechanisms of *OsPLDα1* gene expression are still unclear. In this study, the *OsPLDα1* promoter regions in wild *Oryza* species accessions and *Oryza* cultivars were screened to identify the novel *cis* elements and to describe the regulation of gene expression. We identified SNPs and InDels at *OsPLDα1* promoters in the wild germplasm of rice in comparison to the reference sequence of Nipponbare. Interestingly, three accessions of *O. rufipogon* including IRGC89224, IRGC104425 and IRGC105902 were found to harbor useful variability which contributed towards lower expression of *OsPLDα1* gene. The wild species of *Oryza* serve as an important reservoir of genetic variability for agronomic traits such as biotic and abiotic stresses and for improved yield potential [32–35]. Many useful genes viz. genes for resistance to bacterial blight (BB), blast, tungro virus, grassy stunt virus; tolerance to soil toxicity (acid sulphate, iron toxicity); and cytoplasmic male sterility have been transferred from wild species of rice [36, 37]. Among the most successful examples of utilizing wild species is the use of *Oryza nivara* genes to provide long-lasting resistance

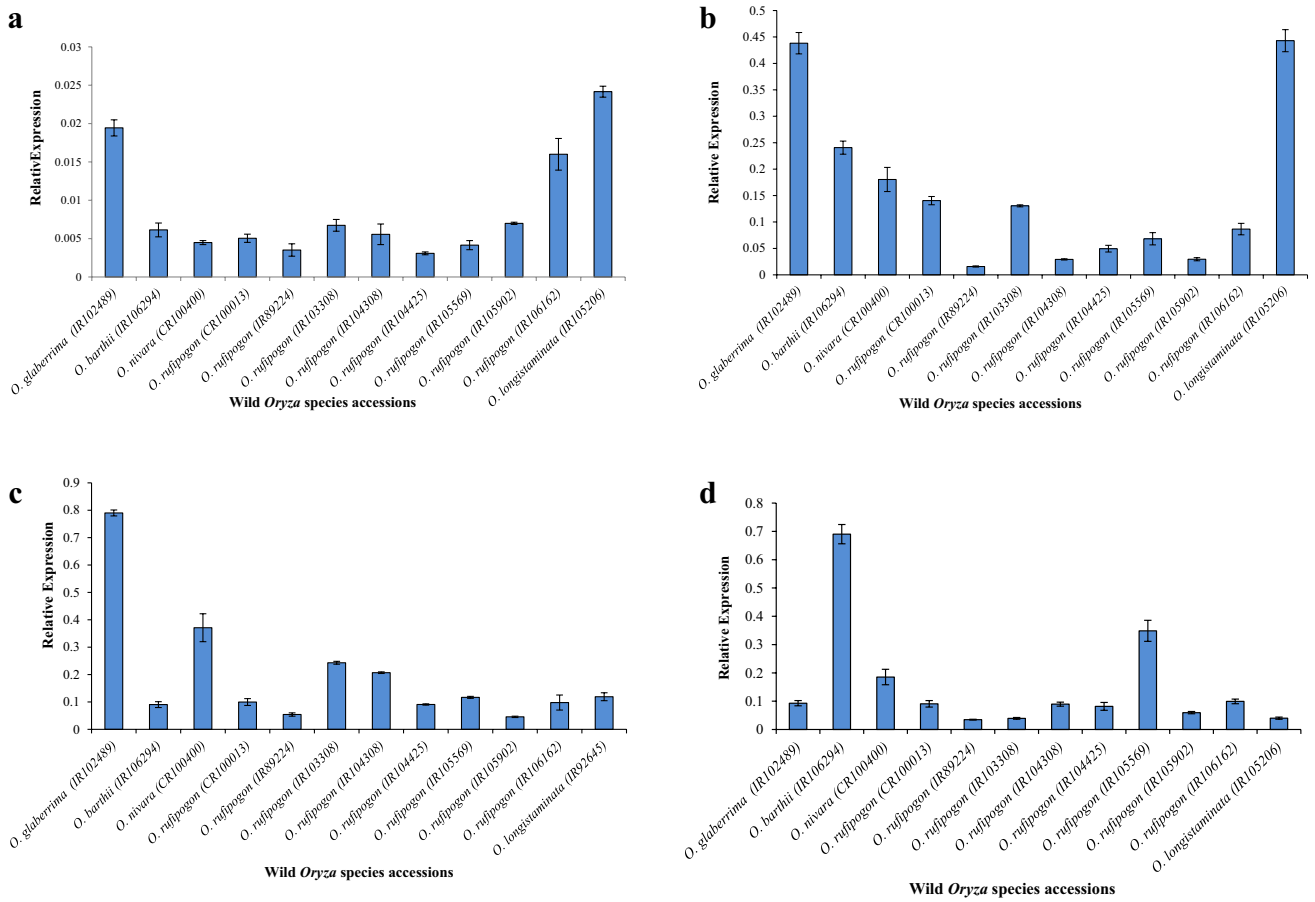

Fig. 3 The expression patterns of *OsPLDα1* gene in the wild *Oryza* species accessions. **a** The expression patterns for qRT-PCR primer (PLDE1) designed from first exon of the gene. **b** The expression patterns for qRT-PCR primer (PLDE2.1) designed from 5′end of second exon. **c** The expression patterns for qRT-PCR primer (PLDE2.2) designed from 3′end of second exon. **d** The expression patterns for qRT-PCR primer (PLDE3) designed from third exon of the gene. The actin gene (Locus ID *Os10g0510000*) was used as an internal control to normalize gene expression levels

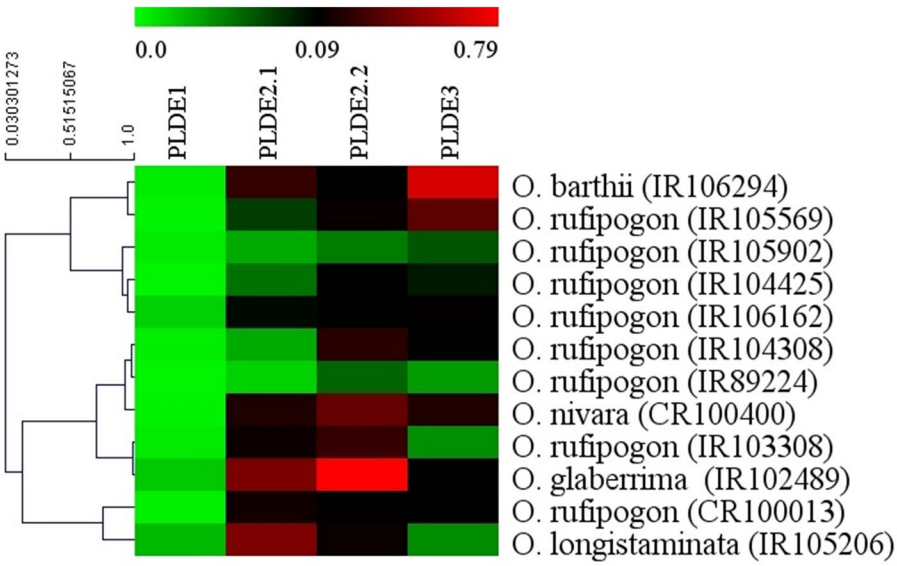

Fig. 4 Heatmap showing differential expression of *OsPLDα1* transcripts between as well as within the accessions of wild *Oryza* species. PLDE1, PLDE2.1, PLDE2.2 and PLDE3 denotes the exon-specific qRT-PCR primers designed from different exons of *OsPLDα1*. Wild species accessions (horizontal) were hierarchially clustered (Pearson correlation, average linkage). Color patterns from green to red indicate low to high transcript levels, thus IRGC89224 had the lowest expression for all the four exon specific qRT primers

to grassy stunt virus [38]. Similarly, *O. rufipogon* has been the source of beneficial alleles for diverse traits including grain size, grain weight [39, 40], grain yield [41–43], grain quality [44], cold tolerance [45], aluminium tolerance [46], and flowering time [47]. In the similar fashion, *O. rufipogon* accessions showing variations in the *OsPLDα1* promoter region could serve as the useful source to improve the rice bran quality in elite rice lines.

Further, our results on molecular diversity at *OsPLDα1* promoter, revealed the presence of lower diversity between Nipponabre and *japonica* cultivar (Kitake) in comparison to the *indica* cultivars (N22, IR64, Pusa44, PR114, PAU201, Minghui63, and Feng-Ai-Zhan). The study conducted on the assessment of genetic diversity using both nuclear and cytoplasmic molecular markers also showed that *indica* and *japonica* are associated with different subgroups of *O. rufipogon* [48–50]. The accessions of African rice (*O. glaberrima*) and its wild progenitors viz. *O. barthii* and *O. longistaminata* were found closely related to each other. Similar results were obtained in the genetic diversity study conducted by using multiple gene sequences by Li et al. [51]. Similarly, the *O. nivara* accessions were found closely related to most of the *O. rufipogon* accessions. In a recent study, Kim et al. (2016) have grouped *O. nivara* and *O. rufipogon* together in a single *O. rufipogon* Griff. species complex (ORSC) on the basis of ease of cross-hybridization between the two species, supporting our results [52]. Our results also suggest the existence of lesser diversity among the *O. rufipogon* accessions belonging to Southeast Asian nations. Similar results were obtained during the several studies conducted for evaluating the genetic diversity of Asian wild rice using RFLP, microsatellite markers, SINEs, sequence based polymorphism, ISSRs, chloroplast, and low-copy nuclear markers [53–55].

Furthermore, to evaluate the expression patterns of *OsPLDα1* gene in selected wild species accessions, the appropriate stage of plant development was selected by performing expression profiling of *OsPLDα1* and its 16 other paralogs using the expression data from *RiceXProv3.0* database. Expression profiling at 49 different time points and using different plant tissues including eight inflorescence (I1–I8), four anther (A1–A4), three pistil (P1–P3), three lemma (L1–L3), three palea (PA1–PA3), ten ovary (O1–O10), twelve embryo (EM0–EM42), and six endosperm (EN07–EN42) stages revealed that *OsPLDα1* keeps on expressing under all the stages of plant development in each selected tissue (Supplementary Fig. S1). Further, an experiment conducted by Suzuki (2011) demonstrated that the PLD content in embryos of Nipponbare increased during seed development up to 3 wk after flowering, becoming constant thereafter. Moreover, PLD-null rice mutant (03-s108), having < 0.01% PLD enzymatic activity in rice bran fraction when compared to Nipponbare, showed no PLD protein band in the seeds 1 week after the flowering. It concluded that the reproductive tissue at early stage of grain development could be used for RNA isolation for *OsPLDα1* expression studies. Therefore, for the present study, immature seeds (1 week after flowering) were selected for RNA isolation to carry out the expression studies.

For expression profiling, four exon-spcific qRT-primer pairs (PLDE1, PLDE2.1, PLDE2.2, and PLDE3) were designed from the coding sequence of *OsPLDα1* (Supplementary Table S2). The exon-specific primers were designed in a way to evaluate the selected accessions for differences in the gene expression as well as to reveal if the gene is alternatively spliced (Fig. 5). qRT-PCR generated expression profiles revealed the presence of interspecific and intraspecific differences in the *OsPLDα1* transcript abundance (Fig. 3). In addition, it also indicated the existence of significant expression variations between the different accessions for the same splice form as well as within the same accession for different splice forms. The expression differences observed in this study could arise from *cis*-regulatory changes that affect transcription initiation, transcription rate and/or transcript stability in an allele-specific manner or from differences in *trans*-regulatory changes that modify the expression of factors that interact with *cis*-regulatory sequences [56, 57]. However, Wittkopp et al. (2004) reported that the expression differences are not caused by *trans*-regulatory changes with widespread effects, but rather by many *cis-acting* changes spread throughout the genome [58]. In this study, significant expression differences were observed for the transcripts obtained with primer designed from the 5′ (PLDE2.1) and 3′ (PLDE2.2) ends of the second exon, which reveals the presence of truncations, within the second exon of *OsPLDα1* gene, during transcription. In addition, the IRGC106294 accession of *O. barthi* and IRGC105569 accession of O. *rufipogon* had high abundance of third exon containing transcripts, in comparison to the transcript abundance for the other two exons (Fig. 3d). The probable reason for this could be the presence of a new transcript form which consists of only third exon, in addition to the five already existing alternate splice forms of *OsPLDα1*. Alternate splicing is a widespread process used in higher eukaryotes to regulate the gene expression and functional diversification of proteins. It has been reported that alternative splicing of the Ca^{2+}-independent phospholipase A_2 ($iPLA_2$) pre-mRNA in humans can result in the production of regulatory subunits that can modify $iPLA_2$ in vivo activity [59].

Also, it has been observed that primer PLDE1 shows low levels of expression of the gene for all the tested accessions (Fig. 4). The reason for this could be the presence of first exon only in three of the five *OsPLDα1* transcript forms (Fig. 5), indicating that it might not play an important role towards determining the *OsPLDα1* expression. Moreover, first exon is 108 base pairs in size which

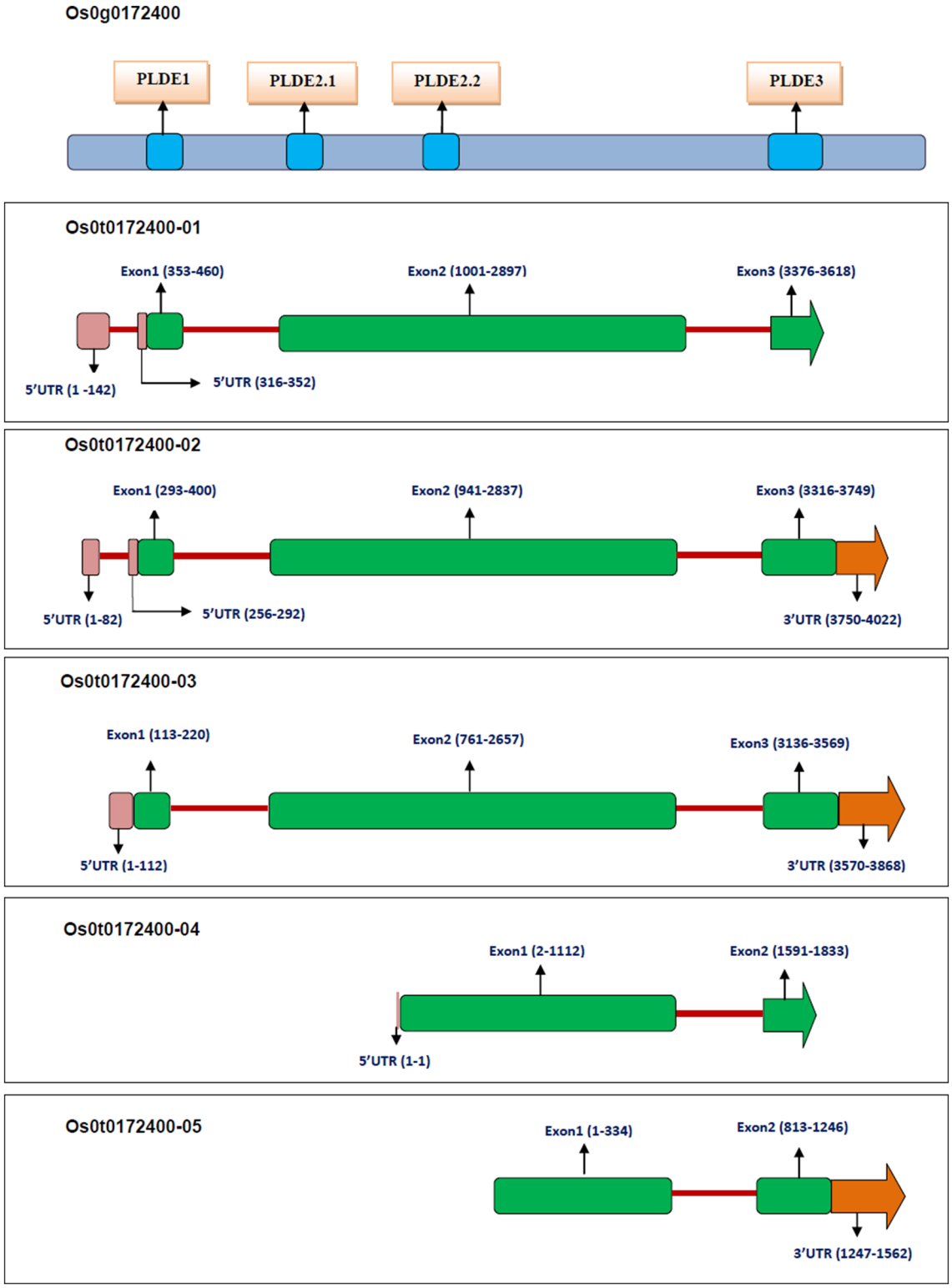

Fig. 5 Architecture of five transcipt variants of Os*PLDα1* gene. It shows the location of four qRT-PCR primers (PLDE1, PLDE2.1, PLDE2.2, and PLDE3) which were designed from different exons, in a way to evaluate variations in relative abundance of different splice forms. The transcript form *Os0t0172400-01* has the similar architecture as of the Os*PLDα1* gene, and thus shows the position of exons/introns of the gene under study

is the shortest one in comparison to the size of second (1897 base pairs) and third (243 base pairs) exons. PLDE1 denotes the exon specific primer designed on the first exon which covers the position from 7 to 101 nucleotide bases, and thus generates the amplicon of only 95 base pairs size. In contrast, the amplicon sizes obtained using the primers designed from 5′ end of second exon (PLDE2.1), 3′ end of second exon (PLDE2.2), and from third exon (PLDE3) were 116, 122, and 223 base pairs, respectively (Table S2). Due to the shorter amplicon length obtained with PLDE1, in comparison to the other three primers, it showed relatively lower expression for all the accessions under study.

The knowledge of the regulatory mechanisms for *OsPLDα1* gene would help in the understanding of how the gene expression is controlled in plants. Unfortunately, however, there is no literature regarding promoter elements, transcription factors, or molecular mechanisms involved in the regulation of *OsPLDα1* gene. To begin to explore these questions, we performed in silico analysis of the *OsPLDα1* promoters, from wild species accessions and cultivars of rice. The results revealed copy number variations and presence/absence of *cis*-elements in comparison to the Nipponbare, and it is well known that the promoter activity of a gene is largely affected by the copy number, inter motif distance, position, deletion or mutation in the core sequences of the *cis-regulatory* regions [60]. In IRGC104425 accession, one copy of GT1CONSENSUS, IBOXCORE, NODCON2GM, OSE2ROOTNODULE, and SEF4MOTIFGM7S; two copies of SURECOREATSULTR11 elements were found missing. However, two copies of CGCGBOXAT and one copy of SORLIP1AT elements were found missing in IRGC105902 accession. The alteration in number of *cis*-elements is a consequence of segmental duplication over the course of evolution. Also, some conservative motifs lost their previous function and some specific elements appeared among genomes during evolution [61]. Owing to the changes in *cis* elements, importance of certain genes increased or decreased over the time, thereby, requiring an increase or decrease in its expression [62]. Moreover, the accumulated mutations in the promoter region of a gene have been reported to be the underlying cause for restricted expression pattern in cultivars as compared to the broad expression pattern of its wild type allele in the wild species of *Oryza* [63, 64]. Different *cis-regulatory* elements function differently when their copy number, distance between multiple copies, and distance from the TATA box are altered [65]. A study has already confirmed that GT elements play different functions depending on their copy number [66]. Therefore, variations in the copy number of GT element and other *cis* elements as detected in the current study, could be responsible for low *OsPLDα1* gene expression in IRGC104425 and IRGC105902 accessions.

Furthermore, we report in this study presence of SEBF (TTGTCTC; Silencing Element Binding Factor) and ARFAT (TGTCTC; auxin response factor) elements within the *OsPLDα1* promoter of *O. rufipogon* accession IRGC89224, which were absent in Nipponbare, rice cultivars, and other wild *Oryza* species. The identified elements shared the motifs and location in the upstream region of *OsPLDα1* gene (Table S3). Auxin is recognized as the universal plant growth hormone because of its principal role in the regulation of cell expansion, division and differentiation [67]. Aux/IAA regulates its own transcription, and functions as transcriptional regulator of other auxin responsive genes in cooperation with Auxin Response Factors (ARF) [68]. In the absence of auxin, Aux/IAA inhibits ARF function and represses other auxin early responsive genes. Moreover, it has been estimated that between 2 and 5% of the total genes in *Arabidopsis* and rice are either upregulated or downregulated by auxin [69]. Many studies have revealed that the *cis-acting* elements function as activators or repressors in regulating tissue-specific gene expression which is governed by the combined action of *cis-acting* elements in its promoter region, as well as different nuclear proteins interacting with the elements [70]. There are several defense genes which have been known to be negatively regulated be auxin [71]. Also, PLD gene is reported to be involved in the plant's response to pathogenic elicitation, seed germination, seedling growth, and leaf senescence, in addition to signaling by various plant hormones and stress responses [72–76]. However, identification, cloning, and expression of novel types of plant PLDs established that they are a family of heterogeneous enzymes that differ in catalytic and regulatory properties [77–80]. Through genome-wide analysis in rice, 17 PLD members have been found on different chromosomes including PLD α(8), β(2), γ(3), κ, ζ(2) and φ [17]. The function of plant PLDs has been studied using knockout, knockdown, and over expression of corresponding genes, resulting in the identification of the physiological functions of several PLDs. Many studies conclude that there are different isoforms of PLD which are responsible for similar kind of plant functions and defense responses [81–83]. Moreover, Suzuki (2011) identified *OsPLDα1* null mutant plants, having no rice bran deterioration, which grew normally in a paddy field indicating that seed PLD (*OsPLDα1*) is not necessary for seed maturation and germination [20]. These studies indicate that down regulation of the *OsPDLα1* activity, by the action of ARFAT and SEBF elements, will not compromise the other functions of the plant. Hence, ARFAT and SEBF elements which appeared to act as repressors in regulating the *OsPLDα1*

gene expression have been considered as 'novel' elements for maintaining the rice bran quality.

Conclusion

In summary, our efforts for identifying novel *cis*-elements in the *OsPLDα1* promoter of wild *Oryza* species accessions revealed the presence of ARFAT and SEBF *cis* elements in the *Oryza rufipogon* accession (IRGC89224) which had the lowest expression of *OsPLDα1* enzyme as well. Though, the further confirmation of low enzyme activity should be done by estimating the *OsPLDα1* activity in the rice bran by biochemical means. For this purpose, maintenance and seed multiplication of the identified low expressing wild species accessions has been undertaken at the Punjab Agricultural University to produce the amount of rice bran (~ 10 g) needed to estimate the *OsPLDα1* activity. In addition, the *O. rufipogon* accessions (IRGC89224, IRGC105902, and IRGC104425), with low *OsPLDα1* expressions, have been deployed as the potential donors in the back-crossing programs to transfer the desired trait into elite cultivars of rice.

Acknowledgements The authors are thankful to the International Rice Research Institute (IRRI), Philippines, Manila and National Rice Research Institute (NRRI), Cuttack, India for providing wild species germplasm of rice. This work was supported by the Monsanto Beachell Borlaug International Scholarship Program, USA.

Compliance with ethical standards

Conflict of interest The authors declare that no conflict of interest exists.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Salunkhe DK, Chavan JK, Adsule RN, Kadam SS (1992) Rice. World oilseeds: chemistry, technology, and utilization. Van Nostrand Reinhold, New York, pp 424–428
2. Krishna AGG, Khatoon S, Babylatha R (2005) Frying performance of processed rice bran oils. J Food Lipids 12:1–11
3. Yoshino G, Kazumi T, Amano M, Takeiwa M, Yamasaki T, Takashima S, Iwai M, Hatanaka H, Baba S (1989) Effects of γ-oryzanol on hyperlipidemic subjects. Curr Ther Res 45:543–552
4. Nicolosi RJ, Ausman LM, Hegsted DM (1991) Rice bran oil lowers serum total and low density lipoprotein cholesterol and apo B levels in nonhuman primates. Atherosclerosis 88:133–142
5. Takano K, Kamoi I, Obara T (1986) Changes in lipid components and lipolytic enzyme activities of rice bran during storage (Studies on the mechanisms of lipids-hydrolysing in rice bran part I). J Jpn Soc Food Sci Technol 33:310–315
6. Barnes P, Galliard T (1991) Rancidity in cereal products. Lipid Technol 3:23–28
7. Orthoefer FT, Eastman J (2004) Rice bran and oil. In: Champagne ET (ed) Rice: chemistry and technology. American Association of Cereal Chemists, Inc., St Paul, pp 569–593
8. List GR, Mounts TL, Lanser AC (1992) Factors promoting the formation of nonhydratable soybean phosphatides. J Am Oil Chem Soc 69:403–410
9. Nakayama Y, Saio K, Kito M (1981) Decomposition of phospholipids in soybean during storage. Cereal Chem 58:260–264
10. Thompson JE, Froesea CD, Madeya E, Smitha MD, Yuwen H (1998) Lipid metabolism during plant senescence. Prog Lipid Res 37:119–141
11. Samama AM, Pearce RS (1993) Ageing of cucumber and onion seeds: phospholipase D, lipoxygenase activity and changes in phospholipid content. J Exp Bot 44:1253–1265
12. Takano K (1993) Advances in cereal chemistry and technology in Japan. Cereal Foods World 38:695–698
13. Takano K, Kamoi I, Obara T (1987) Purification and properties of rice bran phospholipase D. J Jpn Soc Food Sci Technol 34:8–13
14. Takano K, Kamoi I, Obara T (1989) Properties and degradation of rice bran spherosome. J Jpn Soc Food Sci Technol 36:468–474
15. Feussner I, Wasternack C (2002) The lipoxygenase pathway. Annu Rev Plant Biol 53:275–297
16. Suzuki Y, Ise K, Li C, Honda I, Iwai Y, Matsukura U (1999) Volatile components in stored rice [*Oryza sativa* (L.)] of varieties with and without lipoxygenase-3 in seeds. J Agric Food Chem 47:1119–1124
17. Zhou Z, Robards K, Helliwell S, Blanchard C (2002) Ageing of stored rice: changes in chemical and physical attributes. J Cereal Sci 35:65–78
18. Gang L, Fang L, Hong-Wei X (2007) Genome-wide analysis of the phospholipase D family in *Oryza sativa* and functional characterization of *PLDβ1* in seed germination. Cell Res 17:881–894
19. Ueki J, Morioka S, Komari T, Kumashiro T (1995) Purification and characterization of phospholipase D (PLD) from rice (*Oryza sativa* L.) and cloning of cDNA for PLD from rice and maize (*Zea mays* L.). Plant Cell Physiol 36:903–914
20. Suzuki Y (2011) Isolation and characterization of a rice (*Oryza sativa* L.) mutant deficient in seed phospholipase D, an enzyme involved in the degradation of oil-body membranes. Crop Sci 51:567–573
21. Gangadhara Kumar PR, Prakash V (2010) The structure functional catalytic activity of rice bran lipase in the presence of selenium and lithium. Eur Food Res Technol 230:551–557
22. Malekian F, Rao MR, Prinyawiwatkul W, Marshall WE, Windhauser M, Ahmedna M (2000) Lipase and lipoxygenase activity, functionality, and nutrient losses in rice bran during storage, Bull 870. Louisiana Agricultural Experiment Station, LSU AgCenter, Baton Rouge, pp 1–56
23. Ramezanzadeh FM, Rao RM, Windhauser M, Prinyawiwatkul W, Marshall WE (1999) Prevention of hydrolytic rancidity in rice bran during storage. J Agric Food Chem 47:2997–3000
24. Khush GS, Ling KC (1974) Inheritance of resistance to grassy stunt virus and its vector in rice. J Hered 65:135–136
25. Amante-Bordeos A, Sitch LA, Nelson R, Damacio RD, Oliva NP, Aswidinnoor H (1992) Transfer of bacterial blight and blast resistance from the tetraploid wild rice *Oryza minuta* to cultivated rice *Oryza sativa*. Theor Appl Genet 84:345–354
26. Sidana K (2012) Molecular mapping of QTL for yield and yield components transferred from *Oryza longistaminata* (A. Chev. et. Roehr.) to *Oryza sativa* (L.). M.Sc. thesis, Punjab Agricultural University, Ludhiana, India
27. Venter M, Botha FC (2004) Promoter analysis and transcription profiling: integration of genetic data enhances understanding of gene expression. Physiol Plant 120:74–83
28. Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer length polymorphism in barley:

mendelian inheritance, chromosomal location and population dynamics. Proc Nat Acad Sci USA 81:8014–8019
29. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729
30. Schmittgen TD, Livak KJ (2008) Analyzing real-time PCR data by the comparative C_T method. Nat Protoc 3:1101–1108
31. Godber JS, Juliano BO (2004) Rice lipids. In: Champagne ET (ed) Rice: chemistry and technology. American Association of Cereal Chemists, Inc., St Paul, pp 163–190
32. Brar DS, Khush GS (1997) Alien introgression in rice. Plant Mol Biol 35:35–47
33. Xiao J, Li J, Grandillo S, Ahn SN, Yuan L, Tanksley SD, McCouch SR (1998) Identification of trait-improving quantitative trait loci alleles from a wild rice relative, *Oryza rufipogon*. Genetics 150:899–909
34. Moncada P, Martinez CP, Borrero J, Chatel M, Gauch H Jr, Guimaraes E, Tohme J, McCouch SR (2001) Quantitative trait loci for yield and yield components in an *Oryza sativa* x *Oryza rufipogon* BC_2 F_2 population evaluated in an upland environment. Theor Appl Genet 102:41–52
35. Yoon DB, Kang KH, Kim HJ, Ju HG, Kwon SJ, Suh JP, Jeong OY, Ahn SN (2006) Mapping quantitative trait loci for yield loci for yield components and morphological traits in an advanced backcross population between *Oryza grandiglumis* and the *Oryza sativa japonica* cultivar Hwaseongbyeo. Theor Appl Genet 112:1052–1062
36. Brar DS, Khush GS (2002) Transferring genes from wild species into rice. In: Kang MS (ed) Quantitative genetics, genomics, and plant breeding. CABI, Wallingford, pp 197–217
37. Brar DS, Khush GS (2006) Cytogenetic manipulation and germplasm enhancement of rice (*Oryza sativa* L.). In: Singh RJ, Jauhar PP (eds) Genetic resources, chromosome engineering and crop improvement. CRC, Boca Raton, pp 115–158
38. Khush GS, Ling KC, Aguiero VM (1977) Breeding for resistance to grassy stunt in rice. Proc 3rd Int Congress, SABRAO, Canberra, Australia, pp 3–9
39. Li J, Thomson MJ, McCouch SR (2004) Fine mapping of a grain-weight quantitative trait locus in the pericentromeric region of rice chromosome 3. Genetics 168:2187–2195
40. Xie X, Song MH, Jin F, Suh JP, McCouch SR, Ahn SN (2006) Fine mapping of a grain weight quantitative trait locus on rice chromosome 8 using near isogenic lines derived from a cross between *Oryza sativa* and *O. rufipogon*. Genetics 113:885–894
41. Xiao J, Grandillo S, Ahn SN, Yua L, Tanksley SD, Li J, Yuan L (1996) Genes from wild rice improve yield. Nature 384:223–224
42. Li D, Sun C, Fu Y, Li C, Zhu Z, Chen L, Cai H, Wang X (2002) Identification and mapping of genes for improving yield from Chinese common wild rice (*O. rufipogon* Griff.) using advanced backcross QTL analysis. Chin Sci Bull 18:1533–1537
43. Tian F, Zhu Z, Zhang B, Tan L, Fu Y, Wang X, Sun C (2006) Fine mapping of a quantitative trait locus for grain number per panicle from wild rice (*Oryza rufipogon* Griff.). Theor Appl Genet 113:619–629
44. Septiningsih EM, Trijatmiko KR, Moeljopawiro S, McCouch SR (2003) Identification of quantitative trait loci for grain quality in an advanced backcross population derived from the *Oryza sativa* variety IR64 and the wild relative *O. rufipogon*. Theor Appl Genet 107:1433–1441
45. Liu F, Sun C, Tan L, Fu Y, Li D, Wang X (2003) Identification and mapping of quantitative trait loci controlling cold-tolerance of Chinese common wild rice (*O. rufipogon* Griff.) at booting to flowering stages. Chin Sci Bull 48:2068–2071
46. Nguyen BD, Brar DS, Bui BC, Nguyen TV, Pham LN, Nguyen HT (2003) Identification and mapping of the QTL for aluminum tolerance introgressed from the new source, *Oryza rufipogon* Griff. into *indica* rice (*Oryza sativa* L.). Theor Appl Genet 106:583–593
47. Thomson MJ, Edwards JD, Septiningsih EM, Harrington S, McCouch SR (2006) Substitution mapping of dth1.1, a flowering time QTL associated with transgressive variation in rice, reveals a cluster of QTLs. Genetics 172:2501–2514
48. Cheng Z, Motoshahi R, Tsuchimoto S, Fukuta Y, Ohtsubo H, Ohtsubo E (2003) Polyphyletic origin of cultivated rice: based on the interspersion pattern of SINEs. Mol Biol Evol 2:67–75
49. Zhu Q, Ge S (2005) Phylogenetic relationships among A-genome species of the genus *Oryza* revealed by intron sequences of four nuclear genes. New Phytol 167:249–265
50. Civan P, Craig H, Cox CJ, Brown TA (2015) The geographically separate domestication of Asian Rice. Nat Plants. https://doi.org/10.1038/nplants.2015.164
51. Li Z, Zheng XM, Ge S (2011) Genetic diversity and domestication history of African rice (*Oryza glaberrima*) as inferred from multiple gene sequences. Theor Appl Genet 123:21–31
52. Kim HJ, Jung J, Singh N, Greenberg A, Doyle JD, Tyagi W, Chung JW, Kimball J, Hamilton RS, McCouch SR (2016) Population dyanamics among six major groups of the *Oryza rufipogon* species complex, wild relative of cultivated Asian rice. Rice 9:56. https://doi.org/10.1186/s12284-016-0119-0
53. Joshi SP, Gupta VS, Aggarwal RK, Ranjekar PK, Brar DS (2000) Genetic diversity and phylogenetic relationship as revealed by inter simple sequence repeat (ISSR) polymorphism in the genus *Oryza*. Theor Appl Genet 100:1311–1320
54. Rakshit S, Rakshit A, Matsumura H, Takahashi Y, Hasegawa Y, Ito A et al (2007) Large-scale DNA polymorphism study of *Oryza sativa* and *O. rufipogon* reveals the origin and divergence of Asian rice. Theor Appl Genet 114:731–743
55. Huang P, Molina J, Flowers JM, Rubinstein S, Jackson SA, Purugganan MD et al (2012) Phylogeography of Asian wild rice, *Oryza rufipogon*: a genome wide view. Mol Ecol 21:4593–4604
56. Davidson EH (2001) Genomic regulatory systems: development and evolution. Academic, San Diego
57. Carroll SB, Grenier JK, Weatherbee SD (2001) From DNA to diversity: molecular genetics and the evolution of animal design. Blackwell, Oxford
58. Wittkopp PJ, Haerum BK, Clark AG (2004) Evolutionary changes in cis and trans gene regulation. Nature 430:85–88
59. Larsson PKA, Claesson HE, Kennedy BP (1998) Multiple splice variants of the human calcium-independent *Phospholipase* A_2 and their effect on enzyme activity. J Biol Chem 273:207–214
60. Mehrotra R, Sangwan C, Khan ZH, Mehrotra S (2015) *Cis* regulatory elements in regulation of plant gene expression: an overview. Gene Technol 4:129
61. Shamloo-Dashtpagerdi R, Razi H, Aliakbari M, Lindlöf A, Ebrahimi M, Ebrahimie E (2015) A novel pair wise comparison method for in silico discovery of statistically significant *cis*-regulatory elements in eukaryotic promoter regions: application to *Arabidopsis*. J Theor Biol 364:364–376
62. Bhattacharjee S, Ranganaath K, Mehrotra R, Mehrotra S (2013) Combinatorial control of gene expression. Biomed Res Int. https://doi.org/10.1155/2013/407263
63. Rosas U, Mei Y, Xie Q, Banta JA et al (2014) Variation in *Arabidopsis* flowering time associated with *cis-regulatory* variation in CONSTANS. Nat Commun 5:3651. https://doi.org/10.1038/ncomms4651
64. Wang E, Wang J, Zhu X, Hao W et al (2008) Control of rice grain-filling and yield by a gene with a potential signature of domestication. Nat Genet 40:1370–1374
65. Mehrotra R, Kiran K, Chaturvedi CR, Ansari SA, Lodhi N, Sawant S, Tuli R (2005) Effect of copy number and spacing of the ACGT

and GT *cis* elements on transient expression of minimal promoter in plants. J Genet 84:183–187
66. Mehrotra R, Panwar J (2009) Dimerization of GT element interferes negatively with gene activation. J Genet 88:257–260
67. Kieffer M, Neve J, Kepinski S (2010) Defining auxin response contexts in plant development. Curr Opin Plant Biol 13:12020
68. Chapman EJ, Estelle M (2009) Mechanism of auxin-regulated gene expression in plants. Annu Rev Genet 43:265–285
69. Nemhauser JL, Mockler TC, Chory J (2004) Interdependency of brassinosteroid and auxin signaling in *Arabidopsis*. PLoS Biol 2:e258
70. Cai M, Wei J, Li X, Xu C, Wang SA (2007) Rice promoter containing both novel positive and negative cis-elements for regulation of green tissue-specific gene expression in transgenic plants. Plant Biotechnol J 5:664–674
71. Jouanneau JP, Lapous D, Guern J (1991) In plant protoplasts, the spontaneous expression of defense reactions and the responsiveness to exogenous elicitors are under auxin control. Plant Physiol 96:459–466
72. Bastiaan OB, Teun M (2006) The role of phospholipase D in plant stress responses. Curr Opin Plant Biol 9:515–522
73. Yamaguchi T, Minami E, Ueki J, Shibuya N (2005) Elicitor-induced activation of phospholipases plays an important role for the induction of defense responses in suspension-cultured rice cells. Plant Cell Physiol 46:579–587
74. Testerink C, Munnik T (2005) Phosphatidic acid: a multifunctional stress signaling lipid in plants. Trends Plant Sci 10:368–375
75. Wang X (2005) Regulatory functions of phospholipase D and phosphatidic acid in plant growth, development, and stress responses. Plant Physiol 139:566–573
76. Li M, Qin C, Welti R, Wang X (2006) Double knockouts of *phospholipases Dζ1* and *Dζ2* in *Arabidopsis* affect root elongation during phosphate-limited growth but do not affect root hair patterning. Plant Physiol 140:761–770
77. Li G, Xue HW (2007) *Arabidopsis PLDζ2* regulates vesicle trafficking and is required for auxin response. Plant Cell 19:281–295
78. Pappan K, Qin W, Dyer JH, Zheng L, Wang X (1997) Molecular cloning and functional analysis of polyphosphoinositide dependent phospholipase D, *PLDβ*, from *Arabidopsis*. J Biol Chem 272:7055–7061
79. Pappan K, Zheng S, Wang X (1997) Identification and characterization of a novel phospholipase D that requires polyphosphoinositide and submicromolar calcium for activity in *Arabidopsis*. J Biol Chem 272:7048–7054
80. Pappan K, Austin-Brown S, Chapman KD, Wang X (1998) Substrate selectivities and lipid modulation of phospholipase D α, β and γ from plants. Arch Biochem Biophys 353:131–140
81. Qin W, Pappan K, Wang X (1997) Molecular heterogeneity of phospholipase D (PLD): cloning of PLD δ and regulation of plant PLD α, β and δ by polyphosphoinositides and calcium. J Biol Chem 272:28267–28273
82. Li J, Wang X (2019) Phospholipase-D and phosphatidic acid in plant immunity. Plant Sci 279:45–50
83. McGee JD, Roe JL, Sweat TA, Wang X, Guikema JA, Leach JE (2003) Rice phospholipase D isoforms show differential cellular location and gene induction. Plant Cell Physiol 44:1013–1026