Jun *et al. Appl Biol Chem* (2019) 62:31
https://doi.org/10.1186/s13765-019-0439-z

KSABC
The Korean Society for Applied Biological Chemistry

ARTICLE Open Access

Development of *Tos17* insertion mutants from Korean cultivars 'Ilmibyeo' and 'Baegjinju1ho' (*Oryza sativa* L.)

Kyong Mi Jun[1†], Joung Sug Kim[2†], Songhwa Chae[2], Yoon-Mok Pahk[1], Gang-Seob Lee[3], Joon-Hui Chung[4], Yeon-Ki Kim[2*] and Baek Hie Nahm[1,2*]

Abstract

Rice is one of the most important crops globally and a model plant for genomic studies of monocots. With the release of complete genome sequences, the next challenge is to develop various resources based on functional analyses of genes. In this study, we generated mutants via the insertion of *Tos17*, a mobile endogenous retrotransposon active during tissue culture. Two rice cultivars, *Oryza sativa* L. japonica **'**Ilmibyeo**'** (IM) and **'**Baegjinju1ho**'** (BJJ1), which represent white and brown rice in the Korean domestic market, respectively, were selected for this study. We analyzed 7608 flanking sequences of newly transposed *Tos17* insertions by the flanking adaptor-ligation polymerase chain reaction method and identified 1672 and 843 mutants (M_2 generation) in IM and BJJ1, respectively. An analysis of these *Tos17* insertions showed the preferential insertion of *Tos17* into rice chromosome genic regions (approximately 70%). We found new insertional mutants in 830 genes among the 1533 genes representing 2515 IM and BJJ1 mutants that did not overlap with the 3280 genes affected in the 'Nipponbare' (NP) mutants from the National Institute of Agrobiological Sciences database. Of the 1000 lines of *Tos17* insertion mutants, we observed semi-dwarf and various leaf-type mutants, including those with narrow, pale-green, and striped leaves at the vegetative stage. At the reproductive stage, 10 lines showed a 17–56% increase in 100-grain weight compared with the wild type. This study demonstrates the potential utility of *Tos17* mutants via an efficient tissue culture method in various rice cultivars for improving agronomic traits, including seed weight.

Keywords: Baegjinju1ho, Ilmibyeo, Phenotype, Retrotransposon, *Tos17* mutant

Introduction

Rice is one of the world's most important staple crops. Because rice has the smallest genome size (420 Mb) among major cereal crops, it is the most popular model plant for agronomic, genetic, and physiological studies. Therefore, rice is important for deciphering the molecular control of agronomical traits and improving seed production and quality.

Since the 1970s, rice production has more than doubled (130%) due to the green revolution in cultivation technology and breeding. Breeding programs are large in scale and tend to require several years; as a consequence, data collection can be time-consuming and tedious. Although breeding has played a crucial role in the past, future challenges, such as climate change, natural resource depletion, and increasing population, will create an even higher demand for improved rice varieties.

Recently, various genetic resources for the functional analysis of the rice genome have been rapidly established, including T-DNA and transposon-tagged rice mutant populations. Among the most important outcomes is that large numbers of genes have been functionally characterized, many of which are directly related to rice traits. Most of these advances have been achieved through T-DNA insertional mutagenesis [1, 2]. In total,

*Correspondence: kim750a11@gmail.com; bhnahm@gmail.com
†Kyong Mi Jun and Joung Sug Kim have contributed equally to this work
[2] Department of Biosciences and Bioinformatics, Myongji University, Yongin-si, Gyeonggi-do 17058, Republic of Korea
Full list of author information is available at the end of the article

154,391 insertional flanking sequence tags (FSTs) have been identified in the rice genome using T-DNA [3, 4]. POSTECH RISD has produced approximately 100,000 T-DNA-transformed lines and 20,889 genes disrupted by T-DNA insertion [3, 5]. In addition, RMD (Rice Mutant Database) and CIRAD-INRA have applied the Agrobacterium-mediated T-DNA insertion method to generate 31,892 and 27,870 FSTs harboring 6641 and 3602 genes, respectively [2]. Although T-DNA insertion has been used in the functional analysis of rice genes, it has several limitations. One disadvantage of this method is the potential deletion or chromosomal rearrangement of surrounding genomic DNA [6]. Additionally, the use of T-DNA causes social concerns because of the potential adverse effects of GM foods on human health and environmental safety [7].

Transposon insertion is another major strategy to obtain a large number of insertional mutants. [8, 9]. Transposable elements are divided into two groups: DNA-type (class II elements) elements that catalyze excision and reinsertion, i.e., the cut-and-paste mode, and retrotransposons (class I elements) that function in copy-and-paste mode via an RNA intermediate. In the rice genome, 32 families of retrotransposons have been identified [10], and *Tos17* is the most active among active elements. Tissue culture-induced activation of *Tos17* has been a useful tool for insertional mutagenesis and the functional analysis of genes.

Only one to five copies of *Tos17* are present in various cultivars of rice; moreover, it is inherited stably, with heavy DNA methylation occurring under normal growth conditions and transposition only occurring during prolonged tissue culture [11]. 'Nipponbare' (NP) has two almost identical copies of *Tos17* on chromosomes 7 (chr.07) and 10 (chr.10), and the only differences are a 90-bp insertion and 6 point mutations on chr.10 (Fig. 1a). Only *Tos17* on chr.07 is transpositionally active, whereas the other copy is inactive [12]. *Tos17* tends to insert into genic regions [13], although little is known about the occurrence and diversity of *Tos17* within Korean rice cultivars.

In addition to facilitating functional analyses of rice, the development of endogenous mutants can not only overcome social concerns but also contribute to the development of useful agronomic traits that suit the cultural and geological environments of selected cultivar lines. In a previous study, Miyao et al. [8] analyzed approximately 42,292 flanking sequences from 4316 mutant lines produced from 5-month tissue cultures of NP. Of these 42,292 flanking sequences, 3536 genes were aligned by BLASTX searches. A total of 53 types of abnormal phenotypes were observed from the seedling to harvest stages of these mutants [14]. However, this number of mutants is not sufficient to provide reasonable coverage of the total predicted 33,239 genes to develop rice varieties. Until now, the production of *Tos17* mutants has been limited to the study of NP.

In this study, we generated *Tos17* mutants using the Korean domestic rice cultivars *O. sativa* L. japonica 'Ilmibyeo' (IM) and 'Baegjinju1ho' (BJJ1) through a 1-month tissue culture. IM, which is representative of white rice in the Korean market, was developed from the three-way cross of Milyang `96//Milyang `95/Seomjinbyeo. BJJ1, which represents brown rice, was developed from a cross of Ilpum (MNU)-10-2-GH1-3 and 'Seoanbyeo'. We analyzed 7608 flanking sequences of the newly transposed *Tos17* mutants and produced 1672 and 843 mutants (M_2 generation) from IM and BJJ1, respectively. In addition, we collected phenotypic data for the discovery of agronomically important varieties and genes, including mutants with a higher seed yield.

Materials and methods

Generation of *Tos17* mutants

Peeled rice (IM and BJJ1) seeds were sterilized in 70% alcohol for 2 min, shaken in 50% Clorox for 15 min and washed with distilled H_2O four times. Calli derived from embryos were grown in 2N6 media containing 4 g of Chu media (Duchefa Biochemie B.V.), 2 g of casamino acids, 0.5 g of proline, 0.5 g of glutamine, 30 g of sucrose, 2 mg of 2,4-D, and 2.5 g of Gelrite per liter, pH 5.8, at 28 °C for 1 month in the dark. Each round, hard, light-yellow callus with a diameter of 1–3 mm was chosen and transferred into fresh 2N6 medium. After growing at 28 °C for 4 days in the dark, the calli were transferred to fresh MSR media containing 4.4 g of MS salt (Duchefa Biochemie B.V.), 0.5 g of MSE, 5 mg of kinetin, 1 mg of NAA, 30 g of sucrose, and 4 g of Gelrite per liter, pH 5.8, and grown at 28 °C for 1 month under 12 h light/12 h dark conditions until greening. Then, the greening calli were grown into whole plants on MS0 media containing 4.4 g of MS salt (Duchefa Biochemie B.V.), 30 g of sucrose, 2.5 g of Gelrite per liter, pH 5.8, at 28 °C for 1 month under continuous light. In total, approximately 15,000 M_0 plants were regenerated on MS0 media and sampled for genomic DNA extraction.

Amplification of flanking regions of *Tos17* by adaptor-ligation PCR

Plant genomic DNA was extracted and purified from young leaves using NucleoSpin Plant II (Macherey–Nagel GmbH & Co. KG) according to the manufacturer's protocol. Genomic DNA (500 ng) was digested with 2 U of restriction enzyme at 37 °C for 1 h and ligated with adaptors (50 pmol) by 5 U T4 DNA ligase (Takara, Japan) at 16 °C for 1 h in a 20-μl reaction volume. The first PCR

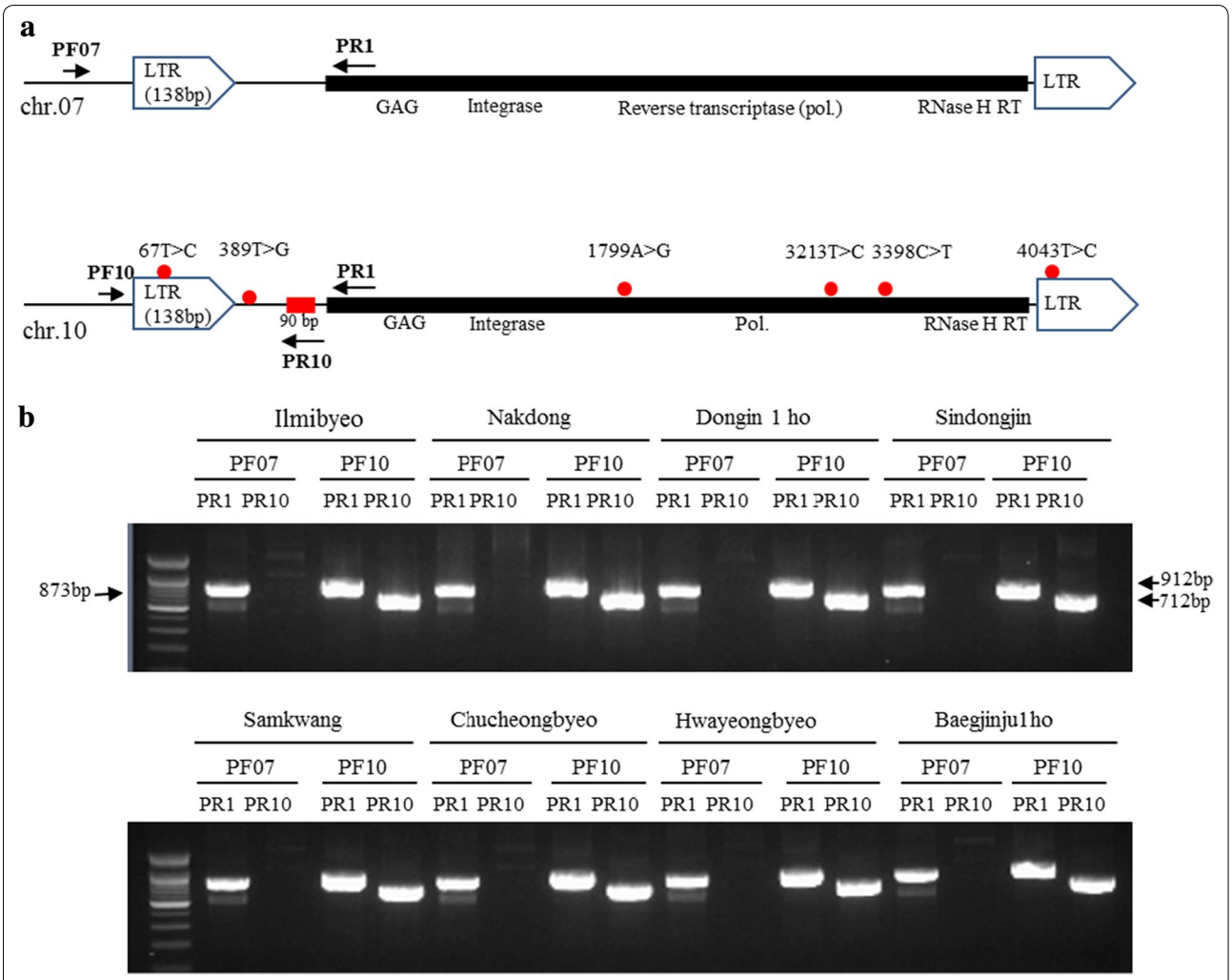

Fig. 1 Comparison of *Tos17* in various Korean domestic rice cultivars to *Tos17* in Nipponbare, **a** structure of *Tos17* on chromosomes 7 (chr.07) and 10 (chr.10). Long terminal repeats (LTRs) are symbolized by open arrows, and open reading frames (ORFs) are indicated by black bars. Six red dots and one box indicate the point mutations and 90-bp insertion in chr.10. **b** Amplified *Tos17* from various Korean domestic rice cultivars

was conducted with 5 µl of digested and ligated mixture, 0.5 pmol A1 (5′-GCGTAATACGACTCACTATAGCAA TTAACC-3′) and T1 (5′-TGCTCTCCACTATGTGCC CTCCGAGCTA-3′) primers, and PCR premixture (Solgent, Korea) in a 20-µl reaction mixture with the following protocol: initial denaturation step at 95 °C for 5 min, 20 cycles at 94 °C for 30 s and at 72 °C for 1 min, and a final elongation step at 72 °C for 10 min. Then, the second PCR was conducted with 5 µl of the first PCR product using the A2 (5′-GACTCACTATAGCAATTAAC-3′) and T2 (5′-ACAAGTCGCTGATTTCTTCAC-3′) primers with the following conditions: an initial denaturation step at 94 °C for 5 min; 40 cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for one minute; and a final elongation step at 72 °C for 10 min. Amplified products were loaded on a 1% agarose gel, and the PCR products were purified using a HiYield Gel/PCR DNA Extraction Kit (RBC, Taiwan) and sequenced using an ABI3730XL system and the T2 primer.

Analysis of the FSTs of *Tos17*

The FSTs of *Tos17* mutants were analyzed with the FST-VAL web tool [13]. We considered only alignments with an E-value of $1e^{-5}$ using the IRGSP-1.0 Rice Genome Annotation Database (http://rapdb.dna.affrc.go.jp/).

M_2 generation in a paddy field

M_1 seeds of the *Tos17* mutants were obtained in the greenhouse from 2010 to 2015, and twelve lines of M_1

seeds were planted in a paddy field located in Jeonju (35°N, 127°E), Korea, from 2016 to 2017. The M_2 seeds were harvested and weighed.

Results and discussion

Determination of endogenous *Tos17* position

To confirm the position of endogenous *Tos17* on chr.07 and chr.10 in various Korean domestic rice cultivars, we used a Tos17-flanking primer in either chr.07 (PF07) or chr.10 (PF10) and a Tos17-common primer (PR1) and a Tos17-specific primer (PR10) in chr.10 based on the NP genome sequence (Fig. 1a). We isolated each genomic DNA from leaves of IM, BJJ1, 'Nakdong', 'Dongin 1 ho', 'Sindongjin', 'Samkwang', 'Chucheongbyeo' and 'Hwayeongbyeo'.

Subsequently, 712-bp and 912-bp fragments were isolated using the PF10 primer and either the PR10 or PR1 primer, respectively, from the genomic DNA of the Korean domestic rice cultivars examined. In addition, an 873-bp fragment of the genomic DNA was amplified with the PF07 and PR1 primers but not with the PF07 and PR10 primers (Fig. 1b). The results of this sequencing showed that many Korean cultivars, including IM and BJJ1, harbor at least 2 copies of *Tos17* on chr.07 and chr.10.

Generation of *Tos17* mutant lines

To produce *Tos17* mutant lines from calli, we chose IM and BJJ1 as representatives of white and brown rice in the Korean market, respectively. Each callus was induced from seed for 1 month on 2N6 media and regenerated on MS0 media (Additional file 1: Fig. S1). To determine the fraction of these putative newly transposed *Tos17* copies, we extracted genomic DNA from the leaves of the M_0 generation. In addition, we amplified the *Tos17* flanking region from genomic DNA digested with the restriction enzyme *MspI* using the 3′ primer of *Tos17* (T1 and T2 primers) and an adaptor-specific primer (A1 and A2 primers) (Additional file 1: Fig. S1). All of the PCR products exhibited two common fragments at 513 and 588 bp, which were derived from chr.07 and chr.10, respectively (Additional file 1: Fig. S2). The PCR products amplified from newly inserted *Tos17* sequences were isolated from the gels and sequenced with the T2 primer (Additional file 1: Fig. S2). In this study, *Tos17* was newly transposed as a low-copy-number region in the genome of IM and BJJ1 (0 ~ 3 copies).

Furthermore, we isolated 7608 FSTs. When we sequenced the newly transposed FSTs, the long terminal repeat (LTR) region of *Tos17* was found to be identical to that of chr.07. As shown in a previous report, we confirmed that *Tos17* in chr.07 is active only in tissue culture [12].

By analyzing the flanking sequences of *Tos17*, we found that all 4.2-kb fragments of *Tos17* with the LTR region at both ends were cleanly inserted into the rice genome. In contrast, T-DNA tagging, another method used for insertional mutagenesis, involves incomplete insertion. The border sequence of the T-DNA is often broken before insertion, and in some cases, the T-DNA backbone is inserted into the genome without cleavage [15].

Analysis of *Tos17* flanking sequences

To analyze the 7608 FSTs, we used FSTVAL, which was previously developed as a web tool to manage bulk flanking sequence tags [13]. Among the analyzed FSTs, 6959 were aligned to the rice genome of IRGSP1.0, and 279 FSTs (4%) exclusively matched the repeat region (Table 1). The frequencies of *Tos17* insertions in genic and intergenic regions were 70% and 26%, respectively (Table 1), which indicates that the insertion of *Tos17* was more frequent in genic than in intergenic regions.

The distribution of the *Tos17* insertions on the 12 rice chromosomes is shown in Additional file 1: Fig. S3 and Additional file 2: Table S1. Overall, three chromosomes (chr.01, chr.02, and chr.03) had high densities of

Table 1 Insertion frequencies of *Tos17* in genic and intergenic regions

Type	All *Tos17* flanking sequence		Ilmibyeo		Baegjinju1ho	
	Number of insertions	Ratio (%)	Number of insertions	Ratio (%)	Number of insertions	Ratio (%)
Intergenic	1901	26	467	23	251	25
Exon	2700	37	814	40	384	39
Intron	1517	21	494	24	213	22
5′ Upstream-1000	652	9	194	9	112	11
3′ Downstream-300	189	3	60	3	21	2
Repeat	279	4	25	1	7	1
Total	7238	100	2054	100	988	100

Tos17 insertions, and chr.10 and chr.12 had relatively low densities of *Tos17* insertions (Additional file 2: Table S1). *Tos17* preferentially inserted into the distal ends of chromosomes rather than the centers, whereas *Tos17* rarely inserted into centromeric regions (Additional file 1: Fig. S3).

To regenerate *Tos17* mutants from the calli of IM and BJJ1, we independently harvested M_1 seeds in a greenhouse and cultivated 1672 IM and 843 BJJ1 mutants (M_2 generation) (Additional file 2: Table S2). The frequencies of *Tos17* insertions in genic regions were 76% (1562 FSTs) and 74% (730 FSTs) in IM and BJJ1, respectively (Table 1). Among 1495 genes with FSTs in their genic regions, 1147 (76.7%) and 199 (13.3%) genes contained one site and two sites of *Tos17* insertion, respectively (Additional file 2: Table S3).

Burr et al. [16] have reported that gene densities are generally highest at the distal ends of the chromosome arms and that fewer ESTs are found in centromeric regions. Rice has an average gene density of one gene per 9.9 kb, especially on chr.01, chr.02, and chr.03, which have high gene densities of one gene per 8.9, 9.1, and 8.7 kb, respectively [17]. These results support our observation that *Tos17* predominantly inserted into genic regions (Table 1), which is one of the advantages of using *Tos17* for insertional mutagenesis.

We found that 1533 genes were tagged by *Tos17* insertion from 2515 mutants of the IM and BJJ1 cultivars. In a previous study, Miyao et al. [8] analyzed 3536 genes representing 42,292 NP mutants. A comparison of the differences among the mutant groups revealed that 830 genes among the *Tos17* mutants (54%) were found only in the IM and BJJ1 mutants and not in the NP mutants.

Until now, a total of 65 *Tos17* mutants from NP have been studied in 42 published papers, among which 29 domestic mutants with 13 mutated genes were obtained as IM and BJJ1 mutants. We produced 6 independent lines in which the *Phosphate2* (*PHO2*, *Os05g0557700*) gene was knocked out. Among these lines, Y17MJ353 mutants showed the lesion mimic phenotype, which was similar to the phenotype of *pho2* mutants reported by Lorieux et al. [18]. We also produced a BJJ002B07 mutant in which *Tos17* is inserted into a SULTR-like phosphorus distribution transporter (*SPDT*, *Os06g0143700*). A previous study [19] reported that the concentration of phytate in brown rice was 25–32% lower in *spdt* mutants than in wild-type rice. In addition, by observing the phenotypes of the *Tos17* mutants under biotic or abiotic stress conditions, such as pathogen, drought, cold, or salt stress, it will be possible to develop new rice lines with tolerance to various stresses.

Functional classification of *Tos17* mutants via MapMan analysis

To functionally characterize genes with genic-region FSTs, we categorized 1533 genes from the IM and BJJ1 mutants with MapMan [20], excluding repeated genes. We also obtained 3280 genes from NP mutants and analyzed them with MapMan [3]. After mapping, 953 and 2269 genes were assigned to different MapMan terms (bin) in the IM/BJJ1 and the NP mutants, respectively (Additional file 2: Table S4). We subsequently compared the patterns of the categorized genes in the IM/BJJ1 and NP mutants (Additional file 1: Fig. S4B and C). Overall, 159 genes in the IM and BJJ1 mutants were mapped to the metabolism overview, with 16 bins (Additional file 1: Fig. S4B, Additional file 2: Table S4). As shown in Additional file 1: Fig. S4, the genes presented similar patterns in the overview of both metabolism and cell functions. In addition, we showed that *Tos17* insertions were distributed throughout the genes of the rice genome without bias based on the assigned gene ontology (GO) terms by AgriGO analysis (Additional file 1: Fig. S5). To identify *Tos17* mutants that could be useful for new variety development, it is important to examine whether *Tos17* insertions were maintained over successive generations.

To confirm the accuracy and stability of the next generation of *Tos17* insertion lines, 22 lines were selected from the various groups categorized by MapMan and then evaluated by tissue-direct PCR conducted on leaves from each line. We selected three genes each from the signaling and transport categories. Four genes from each of the stress, RNA, DNA and protein groups were also selected and examined (Additional file 1: Fig. S6). PCR fragments were amplified with gene-specific primers within a 1-kb region from the *Tos17* insertion site and a *Tos17* primer targeted to 239 bp from the 3′ end of *Tos17*. The PCR product size from IM (wild type) was approximately 1 kb, which was approximately 240 bp larger than that from the *Tos17* insertion lines (+/+) (Additional file 1: Fig. S6). The insertion position of *Tos17* confirmed by adaptor-ligation PCR in the M_0 generation was identical to the genotyping PCR results from the M_2 generation. We found that the *Tos17* insertions were maintained through the M_2 generations and did not appear in new positions without tissue culture (Additional file 1: Fig. S6).

Classification of phenotypes in 1000 lines of *Tos17* insertion mutants at the vegetative stage

Plant heights are classified into two types: "semi-dwarf" and "long culm". The "semi-dwarf" condition is characterized by plant heights that are 70 ~ 80% of the wild-type heights [14]. Y17MJ030 and Y17MJ818 showed phenotypes with heights of 75 and 50 cm, respectively (Fig. 2a).

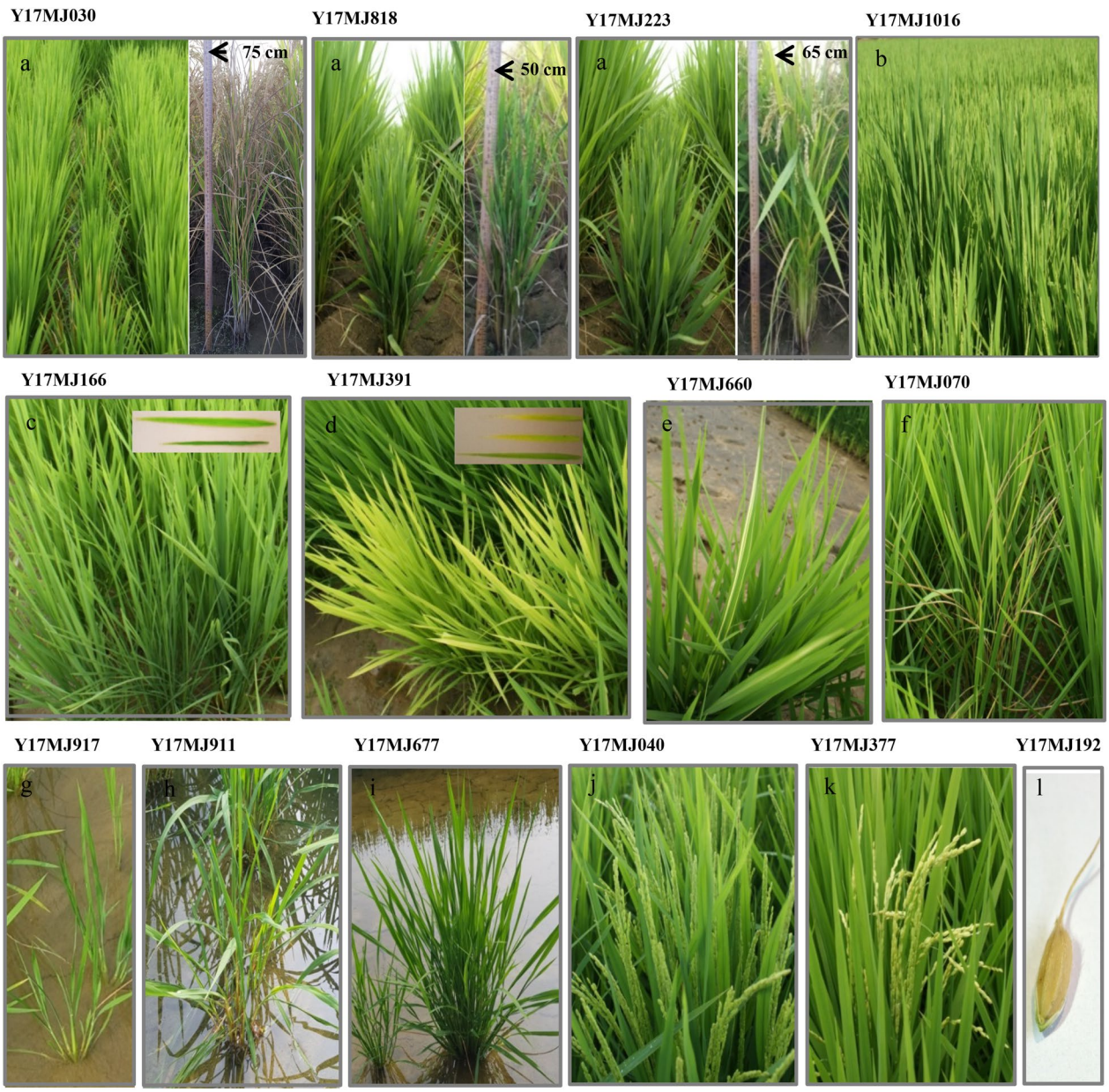

Fig. 2 Representative phenotypes of *Tos17* insertional mutants, **a** semi-dwarf, **b** long culms, **c** narrow leaves, **d** pale green leaves, **e** striped leaves, **f** lesion mimic, **g** weak growth, **h** low tillering, **i** high tillering, **j** dense panicles, **k** early heading, **l** long awns

In the Y17MJ030 and Y17MJ818 lines, *Tos17* was inserted into the introns of *Os02g0131800* (*Trivalent AI influx transporter*) and *Os06g0725100* (Lipase), respectively, as confirmed by the adaptor-ligation PCR method (Additional file 1: Fig. S7). In addition, Y17MJ223, which was 65 cm tall, also showed the semi-dwarf phenotype and had *Tos17* inserted into *Os02g0823000*, *peptidase A22B* (Fig. 2a and Additional file 1: Fig. S7). The Y17MJ1016 line showed the long culm phenotype (Fig. 2b) and a *Tos17* insertion that was not found in the IRGSP1.0 database but matched the *Oryza sativa* IM scaffold 750_cov161 (unpublished).

We observed a narrow leaf shape (Y17MJ166), in which *Tos17* was inserted 14.326 kb upstream from *Os04g0441600* (Fig. 2c), which encodes a protein similar to *Androgen-induced 1*; a pale green leaf (Y17MJ391), in which *Tos17* was inserted in an exon of *Os02g0828100* (Fig. 2d), which encodes a *BRO1* domain containing

protein; a striped leaf (Y17MJ660), in which *Tos17* was inserted 12-bp downstream of *Os06g0111300*, which encodes a hypothetical conserved gene (Fig. 2e and Additional file 1: Fig. S7); and a lesion mimic phenotype (Y17MJ070), in which *Tos17* was inserted 3.525 kb upstream of *Os11g0187150* (hypothetical protein) (Fig. 2f and Additional file 1: Fig. S7).

A weak growth phenotype with slim seedlings presenting retarded growth (Y17MJ917), in which *Tos17* was inserted into an exon of *Os06g0705350* (pentatricopeptide repeat-containing protein) (Fig. 2g). We also observed low (Y17MJ911) and high (Y17MJ677) tillering mutants that contained new *Tos17* insertions in an exon of *Os11g0617700* (DUF594 domain-containing protein) and 264 bp downstream of *Os06g0115300* (Acyl-CoA-binding protein), respectively (Fig. 2h and i).

Classification of phenotypes at the reproductive stage

Dense panicles (Y17MJ040) were observed in mutants in which *Tos17* was inserted into an intron of *Os02g0673500*, which encodes a *bHLH* domain protein (Fig. 2j). Early heading phenotype and flowering occurred approximately 20 days earlier than the wild type (Y17MJ377) (Fig. 2k). In the long seed awn phenotype (Y17MJ192) (Fig. 2l), *Tos17* was inserted into an exon of *Os05g0380300*, which encodes an NBS-LRR protein (Additional file 1: Fig. S7).

Among the *Tos17* mutants, 1000 mutants from IM were grown and harvested, and the 100-grain weight from each of these mutants was measured for 5 events in each line (Additional file 1: Fig. S8A). The average 100-grain weight across the 1000 mutants was 2.46 g, which was similar to the average 100-grain weight of the IM wild type. Among the mutants, 10 lines were selected, and these samples showed a weight increase of 17–56% (2.9–3.9 g) compared with that of the IM wild type (Table 2, Additional file 1: Fig. S8B). In the Y17MJ368-1, Y17MJ380-1, and Y17MJ192-1 mutant lines, *Tos17* was inserted into the exons of *Os01g0110100* (*phosphate transporter*), *Os02g0796600* (esterase/lipase/thioesterase domain containing protein), and *Os05g0380300* (similar to NBS-LRR protein) (Fig. 3). Y17MJ940-4, Y17MJ600-2, Y17MJ760-1, and Y17MJ459-4 contained the newly transposed *Tos17* in the introns of *Os02g0118800* (similar to NBS-LRR disease resistance protein), *Os03g0101100* (similar to palmitoyl-protein thioesterase-like), *Os01g0178700* (similar to protein binding protein), and *Os06g0702700*, encoding butirosin biosynthesis (Fig. 3), respectively. In addition, the *Tos17* insertions of Y17MJ644-1 and Y17MJ112-4 were positioned 5′ upstream of *Os05g0137200* (similar to *MDR-like ABC transporter*) and *Os02g0730000* (similar to mitochondrial aldehyde dehydrogenase), respectively. The Y17MJ550-1 line included two *Tos17* insertions in the exons of the *Os04g0632400* and *Os02g0131850* genes, both encoding hypothetical proteins (Fig. 3). In the ten *Tos17* mutants, the 100-grain weight and area of dehulled grain were greater than those in IM (Table 2). However, the area density of the mutants was similar to that of IM, with values of 97–106% (Table 2). Additionally, the Y17MJ192-1 mutant showed a high area density (117% of the IM area density). The area density is the degree of compactness of the deposition of nutrients, which determines the grain weight and the area of grain space [21]. A high area density may not always be related to the yield, rate of seedling growth, or earliness of plant

Table 2 Ten *Tos17* mutant lines showing an increase in weight of 16% (2.9 g) compared to that of Ilmibyeo (wild type)

Sample name	Weight (100 grains)		Length (1 grain)		Area (1 dehulled grain)		Area density* (1 grain)	
	g	%	cm	%	cm^2	%	g/cm^2	%
Ilmibyeo-1	2.5	100	0.68	100	0.123	100	0.204	100
Y17MJ940-4	2.9	116	0.752	111	0.141	115	0.206	101
Y17MJ368-1	2.9	116	0.779	115	0.141	115	0.206	101
Y17MJ380-1	3	120	0.781	115	0.148	121	0.203	100
Y17MJ600-2	3.1	124	0.852	125	0.149	121	0.209	102
Y17MJ644-1	3.2	128	0.796	117	0.161	131	0.199	98
Y17MJ112-4	3.2	128	0.788	116	0.161	132	0.198	97
Y17MJ760-1	3.3	132	0.8	118	0.16	131	0.206	101
Y17MJ550-1	3.4	136	0.821	121	0.158	129	0.215	105
Y17MJ459-4	3.5	140	0.889	131	0.163	133	0.215	106
Y17MJ192-1	3.9	156	0.879	129	0.164	134	0.238	117
Ilmibyeo-2	2.6	104	0.687	101	0.129	106	0.201	98

*Area density = one grain weight/one dehulled grain area

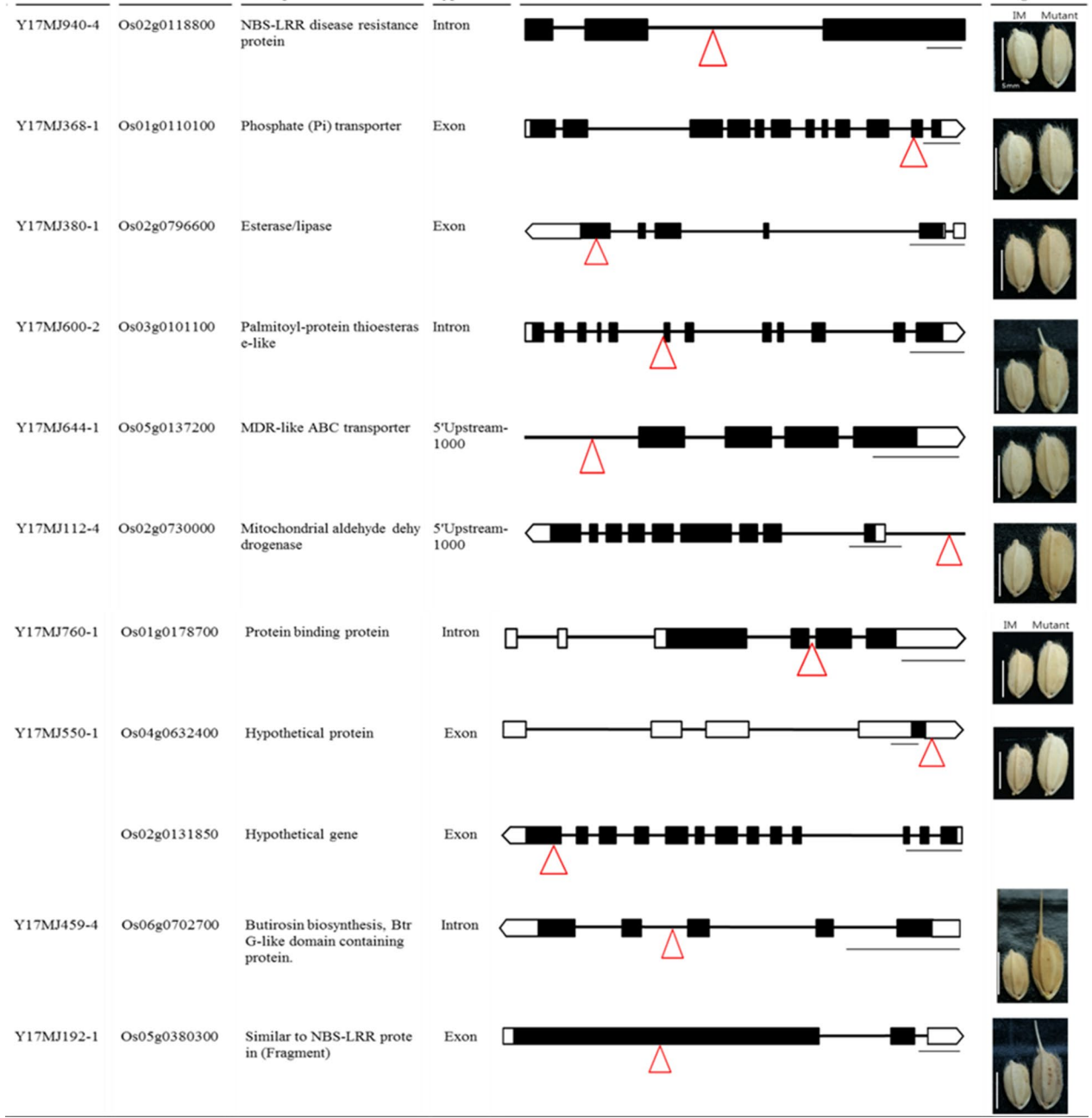

Fig. 3 Insertion positions of *Tos17* and grain phenotypes of *Tos17* insertional mutants, Open and filled boxes indicate non-coding and coding regions, respectively. The insertion sites are shown by red triangles. Scale bar indicates 5 mm

development; however, this parameter is useful for producing new rice varieties.

After analyzing 7608 flanking sequences, 1672 and 843 mutants were produced from IM and BJJ1, respectively (M_2 generation). As an example, we identified *Tos17* insertion mutants with high seed yields from phenotypic data. The production of a large number of *Tos17* insertion mutants with insertion site information is a powerful method that can be performed in a short period of time compared to breeding programs using random mutations, such as chemical and radiation mutagenesis. Until now, the *Tos17* insertion mutants of rice have been

developed using NP as resources for the functional analysis of genes. In this study, we demonstrated the potential use of these mutants not only for functional analysis of genes related to agricultural trails but also for developing new rice varieties using commercial cultivars of brown and white rice in Korea. Furthermore, considering the natural mutation in *Tos17* insertion, mutants with elite traits can be directly used in breeding. Such mutants can be used as a parental resource for breeding to produce secondary varieties with improved traits.

Additional files

Additional file 1: Additional figures. Fig. S1. Scheme for high-throughput *Tos17* insertion mutagenesis by tissue culture in rice. **Fig. S2.** Results of the amplification of flanking regions of *Tos17* by adaptor-ligation PCR in regenerated plants. **Fig. S3.** Distribution maps of the *Tos17* insertions in the rice genome. **Fig. S4.** MapMan classification. **Fig. S5.** Categorization of 1533 genes from *Tos17* mutants through AgriGO analysis. **Fig. S6.** Identification of *Tos17* insertion mutants. Genomic DNA was isolated from the leaves of *Tos17* insertion lines for PCR analysis. **Fig. S7.** Insertion position of *Tos17* according to phenotypes. **Fig. S8.** Comparison of 100-grain weight in IM and *Tos17* mutants.

Additional file 2: Additional tables. Table S1. Number of *Tos17* mutants generated. **Table S2.** *Tos17* mutants with different numbers of insertion sites in a single gene. **Table S3.** MapMan classification of genes with FSTs in genic regions of the IM and BJJ1 mutants and NP mutants.

Acknowledgements
This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (JSK, Grant no. NRF-2018R1D1A1B07049288 and NRF-2017R1A6A3A01076563; YKK, Grant no. NRF-2018R1D1A1B07049348) and the Next-Generation BioGreen 21 Program (BHN, Grant No. PJ01105703), RDA, the Republic of Korea.

Authors' contributions
KMJ and JSK generated the data and wrote the paper. SC and YMP performed the flanking DNA sequencing analysis. GSL and JHC observed the field phenotypes of the rice lines. BHN and YKK inspired the overall work and revised the final manuscript. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Author details
[1] Plant Molecular Genetics Institute, GreenGene Biotech Inc., Yongin-si, Gyeonggi-do 17058, Republic of Korea. [2] Department of Biosciences and Bioinformatics, Myongji University, Yongin-si, Gyeonggi-do 17058, Republic of Korea. [3] Biosafety Division, National Institute of Agricultural Science, Jeonju-si, Jeollabuk-do, 54874, Republic of Korea. [4] The Agricultural Genome Center, Rural Development Administration, Jeonju-si, Jeollabuk-do, 54874, Republic of Korea.

Received: 17 February 2019 Accepted: 29 May 2019
Published online: 08 June 2019

References

1. Jeong DH, An S, Park S, Kang HG, Park GG, Kim SR, Sim J, Kim YO, Kim MK, Kim SR, Kim J, Shin M, Jung M, An G (2006) Generation of a flanking sequence-tag database for activation-tagging lines in Japonica rice. Plant J 45:123–132
2. Zhang J, Li C, Wu C, Xiong L, Chen G, Zhang Q, Wang S (2006) RMD: a rice mutant database for functional analysis of the rice genome. Nucleic Acids Res 34:D745–D748
3. Chandran AKN, Lee GS, Yoo YH, Yoon UH, Ahn BO, Yun DW, Kim JH, Choi HK, An G, Kim TH, Jung KH (2016) Functional classification of rice flanking sequence tagged genes using MapMan terms and global understanding on metabolic and regulatory pathways affected by *dxr* mutant having defects in light response. Rice 9:17
4. Wang N, Long T, Yao W, Xiong L, Zhang Q, Wu C (2013) Mutant resources for the functional analysis of the rice genome. Mol Plant 6:596–604
5. Jeon JS, Lee S, Jung KH, Jun SH, Jeong DH, Lee J, Kim C, Jang S, Yang K, Nam J, An K, Han MJ, Sung RJ, Choi HS, Yu JH, Choi JH, Cho SY, Cha SS, Kim SI, An G (2000) T-DNA insertional mutagenesis for functional genomics in rice. Plant J 22:561–570
6. Nacry P, Camilleri C, Courtial B, Caboche M, Bouchez D (1998) Major chromosomal rearrangements induced by T-DNA transformation in Arabidopsis. Genetics 149:641–650
7. Bawa AS, Anilakumar KR (2013) Genetically modified foods: safety, risks and public concerns-a review. J Food Sci Technol 50:1035–1046
8. Miyao A, Tanaka K, Murata K, Sawaki H, Takeda S, Abe K, Shinozuka Y, Onosato K, Hirochika H (2003) Target site specificity of the *Tos17* retrotransposon shows a preference for insertion within genes and against insertion in retrotransposon-rich regions of the genome. Plant Cell 15:1771–1780
9. Piffanelli P, Droc G, Mieulet D, Lanau N, Bes M, Bourgeois E, Rouviere C, Gavory F, Cruaud C, Ghesquiere A, Guiderdoni E (2007) Large-scale characterization of *Tos17* insertion sites in a rice T-DNA mutant library. Plant Mol Biol 65:587–601
10. Hirochika H (1997) Retrotransposons of rice: their regulation and use for genome analysis. Plant Mol Biol 35:231–240
11. Ding Y, Wang X, Su L, Zhai J, Cao S, Zhang D, Liu C, Bi Y, Qian Q, Cheng Z, Chu C, Cao X (2007) SDG714, a histone H3K9 methyltransferase, is involved in *Tos17* DNA methylation and transposition in rice. Plant Cell 19:9–22
12. Hirochika H (2001) Contribution of the *Tos17* retrotransposon to rice functional genomics. Curr Opin Plant Biol 4:118–122
13. Kim JS, Kim J, Lee TH, Jun KM, Kim TH, Kim YH, Park HM, Jeon JS, An G, Yoon UH, Nahm BH, Kim YK (2012) FSTVAL: a new web tool to validate bulk flanking sequence tags. Plant Methods 8:19
14. Miyao A, Iwasaki Y, Kitano H, Itoh J, Maekawa M, Murata K, Yatou O, Nagato Y, Hirochika H (2007) A large-scale collection of phenotypic data describing an insertional mutant population to facilitate functional analysis of rice genes. Plant Mol Biol 63:625–635
15. Brunaud V, Balzergue S, Dubreucq B, Aubourg S, Samson F, Chauvin S, Bechtold N, Cruaud C, DeRose R, Pelletier G, Lepiniec L, Caboche M, Lecharny A (2002) T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites. EMBO Rep 3:1152–1157
16. Burr B (2002) Mapping and sequencing the rice genome. Plant Cell 14:521–523
17. Sasaki T (2005) The map-based sequence of the rice genome. Nature 436:793–800
18. Lorieux M, Blein M, Lozano J, Bouniol M, Droc G, Dievart A, Perin C, Mieulet D, Lanau N, Bes M, Rouviere C, Gay C, Piffanelli P, Larmande P, Michel C, Barnola I, Biderre-Petit C, Sallaud C, Perez P, Bourgis F, Ghesquiere A, Gantet P, Tohme J, Morel JB, Guiderdoni E (2012) In-depth molecular and phenotypic characterization in a rice insertion line library facilitates gene identification through reverse and forward genetics approaches. Plant Biotechnol J 10:555–568
19. Yamaji N, Takemoto Y, Miyaji T, Mitani-Ueno N, Yoshida KT, Ma JF (2017) Reducing phosphorus accumulation in rice grains with an impaired transporter in the node. Nature 541:92–95
20. Thimm O, Blasing O, Gibon Y, Nagel A, Meyer S, Kruger P, Selbig J, Muller LA, Rhee SY, Stitt M (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J 37:914–939
21. Arain A, Vergara B, Visperas R (1990) Seed density in relation to seedling quality and crop establishment in *Oryza sativa* L. Philipp J Crop Sci 15:137–145

Publisher's Note