Breeding Science 66: 599–605 (2016)
doi:10.1270/jsbbs.16058

Research Paper

Variations in *DENSE AND ERECT PANICLE 1* (*DEP1*) contribute to the diversity of the panicle trait in high-yielding *japonica* rice varieties in northern China

Mingzhu Zhao[1)], Jian Sun[1)], Zhiqiu Xiao[2)], Fei Cheng[1)], Hai Xu[1)], Liang Tang[1)], Wenfu Chen[1)], Zhengjin Xu[1)] and Quan Xu*[1,3)]

[1)] *Rice Research Institute of Shenyang Agricultural University/Key Laboratory of Northern Japonica Rice Genetics and Breeding, Ministry of Education and Liaoning Province/Key Laboratory of Northeast Rice Biology and Genetics and Breeding, Ministry of Agriculture*, Shenyang 110866, China

[2)] *Laboratory of Agricultural Resource and Environment, College of Land and Environmental Science of Shenyang Agricultural University*, Shenyang, 110866, China

[3)] *College of Agronomy of Shenyang Agricultural University*, Shenyang, 110866, China

Rice is one of mankind's major food staples, and the erect panicle architecture in rice is an important morphological improvement. The *dense and erect panicle 1* (*DEP1*) locus corresponds with the formation of erect panicles and has been widely used in rice breeding. However, the genetic diversity of *DEP1* remains narrow. In order to improve the genetic diversity of *DEP1*, we used a rice germplasm collection of 72 high yielding *japonica* rice varieties to analyze the contribution of *DEP1* to the panicle traits. We found 45 SNPs and 26 insertions and deletions (indels) within the DNA fragment of *DEP1*. We further detected 7 haplotypes and found that the replacement of 637 bp by a 12 bp fragment could explain the erect panicle architecture in all 72 germplasms. An SNP (G/C) at the -1253 bp of the promoter region caused a core sequence shift (TGGGCC) of a site II transcriptional regulatory element. The association analysis showed that the SNP(G/C) largely affects the number of primary and secondary branches, and grain number per panicle. Our results provide novel insights into the function and genetic diversity of *DEP1*. The SNP (G/C) at the promoter region will contribute to the flexible application of *DEP1* in rice breeding.

Key Words: *DEP1*, genetic diversity, panicle architecture, high-yield breeding.

Introduction

In rice breeding, two morphological improvements have dramatically increased rice production in recent years. The first was the green revolution of the 1960s in which the development of semi-dwarf lines greatly enhanced rice yield (Suh and Heu 1978). The second was the ideal plant architecture that allowed for the continued increase in rice yield to significantly higher levels than before (Khush 1995, Peng *et al.* 2008). Despite these two breakthrough improvements in rice yields, the world population faces an imminent food crisis in the next 50 years. Thus, breeders and scientists are keen to develop new morphological improvements that can create elite high-yielding cultivars.

Since the 1980s, a number of high-yielding *japonica* rice varieties, which are characterized by dense and erect panicles, have been released in northern China. As the dense and erect panicle varieties began to occupy a dominant place among *japonica* rice in northern China, scientists and breeders began to pay close attention to the dense and erect panicle architecture. Xu *et al.* (1995) found that the erect panicle trait was controlled by a single dominant gene using the posterity of the cross between 'Liaojing 5' and 'Toyonishiki' (a famous *japonica* variety) (Xu *et al.* 1995). A major QTL that conferred the erect panicle was located between the DNA markers RM5652 and H90 on Chromosome 9 in 2007 (Yan *et al.* 2007). In 2009, the target gene of the erect panicle was identified to *Os09g0441900* by three independent research teams and was named *DEP1/EP/qPE9-1* (Huang *et al.* 2009, Wang *et al.* 2009, Zhou *et al.* 2009). In the *DEP1* locus, the replacement of a 637 bp stretch from the middle of the 5th exon with a 12 bp sequence results in erect panicle architecture. The *DEP1* locus is pleiotropic for the erect panicle, the number of grains per panicle, and nitrogen uptake and metabolism (Huang *et al.* 2009, Sun *et al.* 2014). *DEP1* has a conventional plant-specific Gγ subunit protein domain in its N-terminus followed by a cysteine-rich domain

Communicated by Q. Qian
Received March 29, 2016. Accepted May 25, 2016.
First Published Online in J-STAGE on July 27, 2016.
*Corresponding author (e-mail: kobexu34@live.cn)

at the C-terminus. *DEP1* mutants eliminate the TNFR cysteine-rich domain, but not the Gγ domain, resulting in two contrasting effects on plant architecture. First, it represses longitudinal cell division and plant height during the vegetative growth period. Second, it promotes cell proliferation and panicle branching during the reproductive stage, which increases meristematic activity, resulting in reduced inflorescence internode lengths (Huang *et al.* 2009, Sun *et al.* 2014). These converse functions of the *DEP1* mutant result in an erect panicle architecture, well-developed vascular bundles, an increase in the number of grains per panicle and a consequent increase in the grain yield. A recent study demonstrated that *DEP1* is involved in rice nutrient-use efficiency (Sun *et al.* 2014), which makes *DEP1* unique in rice breeding.

Much progress has been made in our understanding of the *DEP1* molecular function in rice; however, the genetic diversity of *DEP1* remains largely elusive. In recent years, a wide range of sequencing has contributed to a better understanding of the function and application of major genes. An association study of major heading time genes using a core collection of 64 rice cultivars from around the world revealed that variations in the rice flowering time gene *Hd1*, *Heading date 3a* (*Hd3a*)—a rice orthologue of FLOWERING LOCUS (FT), and *Early heading date 1* (*Ehd1*) contribute to diversity in the flowering time (Takahashi *et al.* 2009). Nucleotide diversity in *Ghd7*, a gene pleiotropic for grain number, plant height, and heading date, causes phenotypic changes in multiple traits, and some SNPs affect grain number regardless of photoperiod (Lu *et al.* 2012). Thus, large-scale sequencing of major genes that are related to important agronomic traits may provide us with more conclusive information regarding the function of major genes and the flexible application of these genes in rice programs.

In this study, we sequenced a germplasm collection of 72 high-yielding cultivated *japonica* rice varieties from northern China to identify diverse alleles, haplotypes, and key SNPs (in *DEP1*) that affect yield components. Our results showed that the 72 germplasms could be classified into seven haplotypes based on the SNPs and indels that were detected in the full-length genome of *DEP1*. The replacement of 637 bp by a 12 bp fragment explained the erect panicle architecture among 72 germplasms. An SNP (G/C) at the -1253 bp of the promoter region caused a core sequence shift (TGGGCC) of a site II transcriptional regulatory element, which largely affects the number of primary and secondary branches and grain number per panicle among haplotypes. Our results provide novel insights into the function and genetic diversity of *DEP1*. Moreover, the new allele detected in this study will contribute to the flexible application of *DEP1* in rice breeding.

Materials and Methods

Plant materials

A total of 72 accessions of high-yield *Oryza sativa japonica* varieties that have been widely released in northern China were used in this study. Among the 72 varieties, 19 varieties were collected from Heilongjiang Province, 16 from Jilin Province, 23 from Liaoning Province, and 14 were original Japanese varieties. The basic information for each germplasm is shown in **Supplemental Table 1**. All of the germplasms were grown in a paddy field at the experimental farm of Shenyang Agricultural University, Shenyang, China (41.8°N, 123.4°E) during the summer of 2014. Seeds were sown in a seedling nursery on April 24, 2014 with one seedling transplanted per hill on May 23. The seedlings were transplanted at 30 cm × 15 cm spacing. The germplasms were arranged in a randomized block design with three replications, and each replication included at least 40 plants. Fertilizer was applied as a basal dressing at an application rate of 75 kg ha^{-1} N, 150 kg ha^{-1} P, and 75 kg ha^{-1} K.

At the maturation stage (35 days after the full heading), the above-ground portions of 9 plants for each germplasm were harvested from each plot. After counting panicle number and measuring plant height, panicles were hand-threshed and placed in water. Filled grains, which sank in water, were separated from the unfilled grains. To determine dry weight, the filled and unfilled grains were then oven-dried at 80°C for two days. The number of grains per panicle and grain-filling percentage were calculated using the above data. Three average-sized panicles were taken from each plot to observe the number of primary branches, the number of secondary branches, and the number of spikelets on each branch.

DNA extraction, PCR, and sequencing

Three weeks after transplanting, we sampled and extracted DNA from the leaves of 8 plants as a bulk for each germplasm. Genomic DNA was isolated from fresh-frozen leaves using the CTAB method (Doyle 1991). The total 7158-bp length of *DEP1* including 1837-bp promoter regions, 344 bp 5′ UTR, 1272 bp coding region, 2877 bp intro, 228 bp 3′ UTR and 600-bp downstream of *DEP1* were amplified from genomic DNA using KOD plus Neo FX (Toyobo, Japan). The primer used to sequence the *DEP1* locus is listed in **Supplemental Table 2**. The sequencing of the PCR production was performed at BGI Corporation (China). In order to understand the relationship of yield related genetic basis among the 72 accessions, we detected the population structure based on polymorphism of molecular markers in functional yield-related gene regions. Moreover, we expected to assess the contribution of *DEP1* on yield components in diverse yield-related genetic backgrounds by considering the subgroups of 72 accessions according to their genetic clusters. Fifty-eight DNA markers related to the yield traits were selected for genotyping the 72 germplasms according to previous studies conducted by Xing *et al.* (2010) and Huang *et al.* (2009). The primers are listed in **Supplemental Table 3**. The PCR for SSR markers and Caps markers was performed in a reaction mixture (5 mL total volume) containing 0.5 mL 10_Ex-Taq buffer, 0.5 mL

dNTPs (2 mM each), 0.25 mL DMSO, 0.02 mL Ex Taq, l00 ng genomic DNA and five pmol of each primer. The PCR conditions were as follows: 94°C for 10 min; 35 cycles of 94°C for 30 s, 50°C for 1 min and 72°C for 30 s; and finally 72°C for 7 min. Reaction products were loaded on 12% polyacrylamide gels, run at 500 v for 2 h, stained with ethidium bromide, and then the gels were scanned with a molecular imager FX (Nippon Bio-Rad Laboratories, Tokyo). The PCR for indel markers was performed as described above and the amplicons of markers were resolved in 1.0% agarose gels, stained with ethidium bromide, and visualized under UV light.

Data analysis

The genomic sequences and protein sequences were aligned by ClustalX 2.0.0, and the alignments were used as an input format into TASSEL V2.1. Nucleotide diversity and Tajima's D statistics were calculated using the DnaSP 5.0 program. To eliminate the effect of population structure (Q) and genetic relatedness (K) on the traits for marker-trait associations, Q and Q+K models were used. The Q model was performed using the general linear model (GLM) and the Q+K model was performed using the mixed linear model (MLM) in TASSEL V2.1 (Yu *et al.* 2006, Zhang *et al.* 2009). Fifty-eight DNA markers were used to calculate the kinship matrix (K) using the SPAGeDi software package (Hardy and Vekemans 2002), and all negative values between individuals were set to 0 (Yu *et al.* 2006). The Duncan multiple range test and critical test were conducted if the analyses were significant ($P < 0.05$). Statistical analysis was performed by the STATISTICA software (StatSoft 1995). The evolutionary relationship among the 7 haplotypes was inferred using the UPGMA method, and phylogenetic analyses were conducted using the MEGA5 software. The physical and chemical properties of *DEP1* were predicted using Predictprotein (www.predictprotein.org). The secondary structure of *DEP1* protein was predicted using the tools available at swissmodel.expasy.org. The transmembrane structure of *DEP1* was predicted using the tools available at bioinf.cs.ucl.ac.uk. The transcriptional regulatory elements of the promoter region of *DEP1* were analyzed using the tools available at www.dna.affrc.go.jp/PLACE/signalup.html. STRUCTURE 2.1 software (Pritchard *et al.* 2000) was used to examine the population structure. Using a burn-in length of 1,000 steps followed by a run length of 1,000 Monte Carlo Markov Chain replicates, the number of subgroups from $k = 1$ to $k = 10$ were tested with a model that assumed admixture and correlated allele frequencies for simulation. The most probable structure number of k was calculated based on Evanno *et al.* (2005). The effects of population structure on all traits were tested using PROC GLM. The model statement included one of the two components of the k = 2 Q matrix from the STRUCTURE analysis (Yang *et al.* 2011).

Results

Population structure

Genotyping studies showed that all of the 58 DNA markers were polymorphic among the 72 germplasms. We identified a significant population structure in the collection that could be classified into two subpopulations because the highest log-likelihood scores of the population structure were observed when the number of populations was set at 2 ($k = 2$). The genetic constitution of the germplasms from Heilongjiang and Jilin were distinct from those that of the germplasm from Liaoning and Japan (**Fig. 1**). According to the neighbor-joining tree, the germplasms from Heilongjiang and Jilin were classified into subgroup1 and germplasms from Liaoning and Japan were classified into subgroup2 (**Supplemental Fig. 1**). Relative kinship coefficients based on the DNA marker data showed that 53.72% of the pairwise kinship estimates were equal to 0, and the remaining estimates were between 0.025 and 0.5, suggesting that there was either no relatedness or weak relatedness between these pairs of varieties (**Supplemental Fig. 2**).

Nucleotide diversity

The analysis of the nucleotide diversity of the whole genomic DNA sequence of *DEP1* (7158 bp) from 72 germplasms showed 45 single nucleotide polymorphics (SNPs) and 26 insertions and deletions (indels) were detected. The detailed information is listed in **Supplemental Table 4**. Varied DNA polymorphisms were observed in different regions of the *DEP1* genome and the pairwise nucleotide diversity parameter (π) in the 5th exon was higher than that in other regions in the whole germplasm population. The pairwise nucleotide diversity parameter (π) in subgroup1 was lower than that in subgroup2. The Tajima's D values reached a significant positive level in the 5th exon region (5636–6330 bp) in subgroup2, whereas Tajima's D values were negative in other regions in both subgroup1 and subgroup2 (**Fig. 2**).

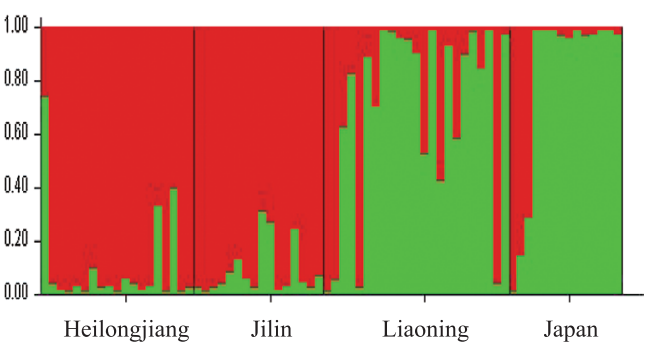
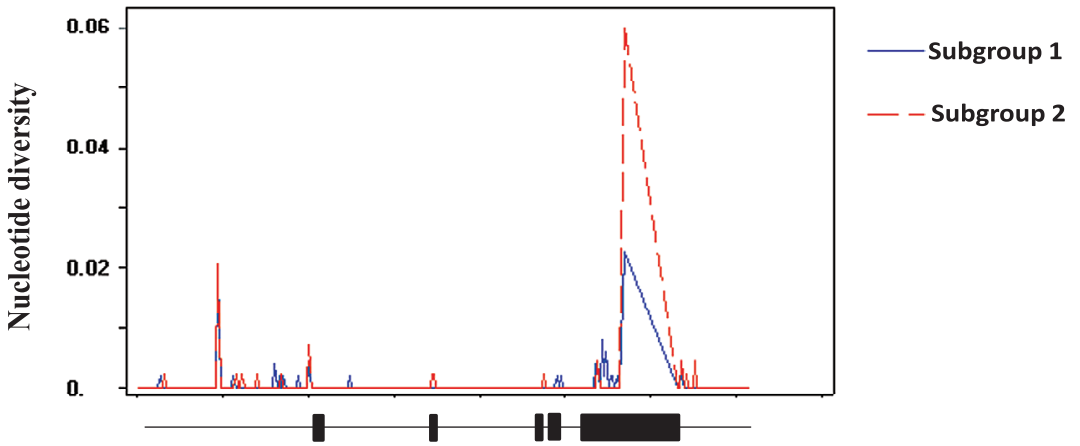
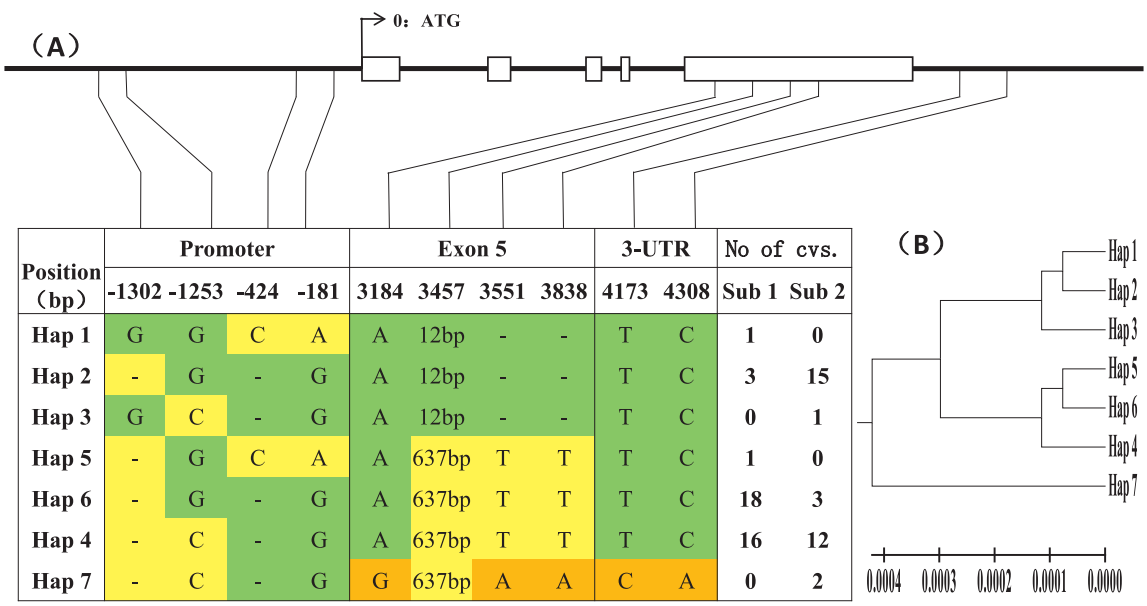


Fig. 1. Population structure for 72 accessions obtained by analysis with the STRUCTURE program. Two clusters ($k = 2$, indicated by green and red shading) were obtained from the simulation using all 72 accessions based on 58 DNA markers. Most of the Heilongjiang and Jilin cultivars were assigned to the red group while most of the Liaoning and Japanese cultivars were assigned to the green group.



Tajima D	Entire region	Promoter	Exon 5	
	(1-7158bp)	(1-2181bp)	(5328-5635bp)	(5636-6330bp)
Subgroup 1	-2.2412*	-1.61896	-2.4351*	-0.4414
Subgroup 2	-1.15648	-1.6716	-0.79933	2.36622*

Fig. 2. Nucleotide diversity analysis and test for neutral selection in subgroup1 and subgroup2. *, significant at $P < 0.05$.



Position (bp)	Promoter				Exon 5				3-UTR		No of cvs.	
	-1302	-1253	-424	-181	3184	3457	3551	3838	4173	4308	Sub 1	Sub 2
Hap 1	G	G	C	A	A	12bp	-	-	T	C	1	0
Hap 2	-	G	-	G	A	12bp	-	-	T	C	3	15
Hap 3	G	C	-	G	A	12bp	-	-	T	C	0	1
Hap 5	-	G	C	A	A	637bp	T	T	T	C	1	0
Hap 6	-	G	-	G	A	637bp	T	T	T	C	18	3
Hap 4	-	C	-	G	A	637bp	T	T	T	C	16	12
Hap 7	-	C	-	G	G	637bp	A	A	C	A	0	2

Fig. 3. Haplotype analysis of the *DEP1* gene region in the 72 cultivars. (A) The promoter region and entire length of the *DEP1* genome is shown in the graphic on the top. The positions of every SNP (SNP frequency > 2.5%) and indels are shown in the first row. Seven haplotypes (Hap1–Hap7) were detected in the 72 cultivars. The number of cultivars in each subgroup is shown in the right columns. (B) Phylogenetic tree of the 7 haplotypes (Hap1–Hap7).

Comparison sequences and haplotype analysis

Among the 45 SNPs and 26 indels, we selected 7 SNPs and 7 indels with a bi-allele frequency > 2.5% for haplotype analysis. We constructed 7 haplotypes (H1–H7) from the 72 germplasms (**Fig. 3**). Among the 7 SNPs and 3 indels, the G/C SNP at the -1253 bp of the promoter region and the 637 bp/12 bp substitution showed higher frequency (**Supplemental Table 5**). The G/C SNP at 1253 bp caused an amino acid change in the promoter region while the 637 bp/12 bp substitution caused a frame shift mutation in the protein in *DEP1*. The latter mutation results in the erect panicle architecture. The H1, H2, and H3 haplotypes had the 12 bp substitution, which exhibited the erect panicle architecture. The other 4 haplotypes, H4, H5, H6, and H7, carried a 637 bp substitution resulting in a curved panicle architecture. Among subgroup1, H6 (G/637 bp) and H4 (C/637 bp) accounted for a large proportion; in subgroup2, most germplasms had H2 (G/12 bp) and H4 (C/637 bp) haplotypes.

Bioinformatics analysis of DEP1

The analysis of transcriptional regulatory elements in the promoter region in *DEP1* showed that G/C SNPs at -1253 bp caused a core sequence shift (TGGGCC) of site II transcriptional regulatory elements. The coding region analysis showed that 3 SNPs and 1 indel could classify the DEP1 protein from 72 germplasms into 3 main types: types I, II, and III. Type I includes the H7 haplotype, type II includes H4, H5, and H6. Type II includes H1 and H2 while type III has H3. Although there was a 3 amino acid change between type I and type II, the physical and chemical properties, secondary structure, and transmembrane structure of type I and type II were similar between the types. Compared with type I and type II, type III lost 231 amino acids, which caused a change in physical and chemical properties, and secondary and transmembrane structures (**Supplemental Table 6**, **Supplemental Fig. 3**).

Association between haplotype and panicle traits

Six traits were collected for varieties in this association, and the measured traits showed relatively high broad-sense heritability with the exception of the1000-grain weight (**Supplemental Table 7**). Only a small part of the phenotypic variation for the trait was due to the presence of subgroups, with an average of 8% across all traits with the exception of heading date (**Supplemental Table 7**). The GLM (Q) and MLM (Q+K) models (**Table 1**) showed that the 12 bp/637 bp substitution of *DEP1* was significantly associated with heading date, panicle length, number of primary branches, number of secondary branches and grain number per panicle. The G/C SNP (-1253) at the promoter region of *DEP1* only showed a significant association with panicle length.

The investigation of yield components among the 7 haplotypes confirmed that the G/C SNP and 12 bp/637 bp substitution strongly affected yield components. In subgroup1, Hap1 and Hap2 had the 12 bp substitution that corresponds to the erect panicle architecture, and they showed a significant increase in the number of primary and secondary branches, and number of grains per panicle compared with Hap4, Hap5, and Hap6, which had the 637 bp substitution (**Table 2**). The G/C SNP at -1253 bp in the promoter region caused a different phenotype in the haplotypes compared with a 637 bp substitution (Hap4, Hap5, and Hap6). The G allele in Hap5 and Hap6 is related to a decrease in the number of primary and secondary branches and number of grains per panicle compared with Hap4 which has a C allele. Similarly, in subgroup2, the G allele was related to a

Table 1. Association analysis results of traits by GLM (Q) and MLM (Q+K) models

Region	Position (bp)	Allele	Model	P-Marker					
				HD	PL	NPB	NSB	GPP	TGW
Promoter	-1302	1 bp/0 bp	Q	0.443	0.429	0.245	0.487	0.964	0.054
			Q+K	0.529	0.468	0.262	0.345	0.806	0.051
	-1253	G/C	Q	0.629	0.002	0.304	0.713	0.731	0.686
			Q+K	0.716	0.003	0.767	0.762	0.675	0.335
	-424	1 bp/0 bp	Q	0.137	0.563	0.631	0.159	0.192	0.664
			Q+K	0.283	0.694	0.792	0.435	0.563	0.795
	-181	A/G	Q	0.561	0.880	0.830	0.924	0.998	0.670
			Q+K	0.494	0.856	0.654	0.828	0.950	0.640
Exon 5	3184	A/G	Q	0.891	0.444	0.861	0.807	0.907	0.608
			Q+K	0.850	0.408	0.759	0.417	0.456	0.584
	3457	12 bp/637 bp	Q	0.044	0.001	0.000	0.001	0.000	0.247
			Q+K	0.015	0.002	0.000	0.005	0.000	0.247

P-marker indicates the significance between marker and the phenotype, and a level of $P < 0.05$ was taken as significant.
HD: Heading date, PL: panicle length, NPB: number of primary branches, NSB: number of secondary branches, GPP: number of grains per panicle, TGW: 1,000 grain weight, GL: grain length, GW: grain width.

Table 2. The yield components of each haplotype between subgroup1 and subgroup2

Group	Haplotype	HD (d)	PL (cm)	NPB	NSB	GPP	TGW (g)	GL (mm)	GW (mm)
Subgroup 1	Hap 1 (G/*dep1*)	109.5 ± 0.7 a	16.5 ± 0.9 a	11.0 ± 1.4 ab	29.0 ± 7.9 a	139.8 ± 24.9 a	24.4 ± 0.0 ab	6.9 ± 0.0 a	3.2 ± 0.0 ab
	Hap 2 (G/*dep1*)	109.3 ± 0.8 a	16.9 ± 1.6 a	12.7 ± 0.5 a	25.3 ± 8.1 ab	143.1 ± 20.6 a	26.5 ± 1.9 a	7.6 ± 0.7 a	3.2 ± 0.2 ab
	Hap 4 (C/*DEP1*)	101.4 ± 8.0 ab	18.9 ± 1.9 a	11.5 ± 1.1 ab	24.3 ± 5.8 ab	129.0 ± 23.4 ab	23.9 ± 2.2 b	7.4 ± 0.5 a	3.2 ± 0.2 b
	Hap 5 (G/*DEP1*)	102.5 ± 0.7 ab	18.6 ± 1.3 a	10.0 ± 0.0 bc	19.6 ± 3.4 bc	108.8 ± 11.9 b	26.3 ± 1.2 a	7.8 ± 0.3 a	3.3 ± 0.1 ab
	Hap 6 (G/*DEP1*)	98.6 ± 5.2 b	18.1 ± 4.2 a	8.8 ± 1.4 c	17.1 ± 2.9 c	96.6 ± 11.7 c	26.4 ± 1.4 a	7.3 ± 0.3 a	3.3 ± 0.2 a
Subgroup 2	Hap 2 (G/*dep1*)	113.4 ± 1.1 b	17.4 ± 1.0 b	13.3 ± 1.2 ab	27.5 ± 5.0 a	153.5 ± 19.8 a	24.0 ± 1.9 b	7.4 ± 0.2 a	3.2 ± 0.1 b
	Hap 3 (C/*dep1*)	108.5 ± 0.7 b	17.0 ± 0.4 b	14.0 ± 1.4 a	19.1 ± 5.0 bc	133.8 ± 16.6 b	29.2 ± 2.7 a	7.7 ± 0.1 a	3.4 ± 0.1 a
	Hap 4 (C/*DEP1*)	112.0 ± 3.8 a	20.2 ± 1.8 a	10.0 ± 1.1 c	21.3 ± 7.8 ab	117.9 ± 28.2 b	26.2 ± 2.4 a	7.6 ± 0.3 a	3.3 ± 0.1 a
	Hap 6 (G/*DEP1*)	99.3 ± 9.1 b	17.7 ± 2.1 b	8.3 ± 1.9 d	12.6 ± 2.8 c	82.7 ± 22.0 c	24.0 ± 2.1 b	7.3 ± 0.4 a	3.3 ± 0.2 a
	Hap 7 (C/*DEP1*)	110.5 ± 1.0 a	20.2 ± 0.8 a	11.8 ± 0.5 b	22.5 ± 2.4 ab	129.5 ± 12.2 b	25.8 ± 1.7 ab	7.5 ± 0.2 a	3.3 ± 0.2 a

HD: Heading date, PL: panicle length, NPB: number of primary branches, NSB: number of secondary branches, GPP: number of grains per panicle, TGW: 1,000 grain weight, GL: grain length, GW: grain width. Within a column, data followed by different lowercase letters indicate significant difference at 5% probability levels.

decrease in the number of primary and secondary branches and number of grains per panicle compared with the C allele when the haplotypes had 637 bp substitution (Hap4, Hap6, and Hap7). Interestingly, the G allele (Hap2) was related to an increase in the number of primary and secondary branches and the number of grains per panicle compared with the C allele (Hap3) when the haplotypes had a 12 bp substitution (**Table 2**).

Discussion

Erect panicle architecture in rice is an important morphological improvement, and significantly enhances the yield of rice production. Besides the *DEP1* locus, *DEP2*, allelic to *EP2* (*Os07g0616000*), was identified from the mutations of 'Zhonghua 11' and 'Nipponbare' (Abe *et al.* 2010, Li *et al.* 2010). *DEP3* (*Os06g0677000*) was identified from a mutation in an elite Korean *japonica* cultivar (Qiao *et al.* 2011), and both the *dep3* and *dep2/ep2* alleles were recessive alleles that caused a dense and erect panicle architecture. In this study, we found the 637 bp/12 bp substitution at the *DEP1* locus can explain all the erect panicle architectures found among 72 high-yielding *japonica* rice varieties. The *DEP1* thought to be a transmembrane protein (Botella 2012), and the replacement of a 637-bp stretch by a 12-bp sequence in the 5th exon eliminates the TNFR cysteine-rich domain (Sun *et al.* 2014). A recent study showed experimental evidence that AGG3, a homologue gene of *DEP1* in Arabidopsis, contains a functional transmembrane domain, and the cysteine-rich domain at C-terminus is extracellular (Wolfenstetter *et al.* 2015). Our bioinformatics analysis confirmed the replacement of a 637 bp stretch by a 12 bp sequence in type III protein, which may cause a change of transmembrane structure compared to type I protein and type II protein, and subsequently change the function of *DEP1*. For example, the grain number per panicle of haplotypes (H1, H2 and H3) that contain the type III protein was significantly greater than that of the other two types (**Supplemental Table 6**). The significant positive Tajima's D parameters in the 5th exon (5636 bp–6330 bp) region of *DEP1* in subgroup2 suggested that the positive selection may occur in this region during rice improvement and breeding. Moreover, a comparison of the pairwise nucleotide diversity parameter (π) with the genome-wide average level (0.0013) of *japonica* rice in 517 landraces in China (Huang *et al.* 2010) showed that the π value of *DEP1* (0.00045) was lower than the average level. Thus, *DEP1* provides a target for selection for high yields during rice improvement in northern China.

The dominant allele at the *DEP1* locus in the *japonica* variety Shao313 caused a 40.9% increase of grain yield per plant, resulting from an increase in the number of grains per panicle (Huang *et al.* 2009). However, some studies did not report any significant change in grain number per panicle in erect panicle plant (Yi *et al.* 2011, Zhou *et al.* 2009). These results and observations indicate that the effects of *DEP1* may differ under different genetic backgrounds. In this study, the comparison of the *DEP1* allele using 72 high-yielding varieties gave us a conclusive function of *DEP1*. The 12 bp allele in 3 (Hap1, Hap2, and Hap3) of 7 haplotypes caused the formation of erect panicles with an increase in the number of primary and secondary branches, and a number of grains per panicle as compared with a 637 bp allele in the 4 other haplotypes (Hap4, Hap5, Hap6 and Hap7). Thus, *DEP1* increases the grain number per panicle in a wide genetic background.

The G/C SNPs at -1253 bp caused a core sequence shift (TGGGCC) of a site II transcriptional regulatory element. Site II plays an important role in the meristematic tissue-specific expression of rice proliferating cell nuclear antigen gene, presumably by mediating putative enhancer activities dependent on the far upstream region (Kosugi *et al.* 1995). As previous studies and this study showed, the haplotypes that 12 and 637 bp substitutions showed contrasting panicle traits in primary and secondary branches and grains per panicle (Huang *et al.* 2009, Wang *et al.* 2009, Zhou *et al.* 2009). The G allele (composing site II) seems to amplify this contrast effect. Among the haplotypes with the 637 bp allele (Hap4, 5, 6, 7), the G allele resulted in a decrease in the number of primary and secondary branches, and the number of grains per panicle compared with the C allele haplotypes. The G allele in the haplotypes with 12 bp substitution causes an increase in the number of primary and secondary branches and the number of grains per panicle compared with the C allele haplotype (Hap3). We found that among the 52 varieties that had a 637 bp substitution, 22 varieties carried the G allele, and 30 varieties had the C allele, whereas all but one of the 16 varieties with the 12 bp substitution carried the G allele. These results indicate that the G allele is more important than the C allele for the *DEP1* allele, which has the 12 bp substitution to exert its functions.

Our previous study showed that the plant showed superior agronomic traits when the *DEP1* locus was heterozygous (Xu *et al.* 2014). Moreover, the 12 bp substitution allele that caused the erect panicle trait exhibited better values for yield components in *indica* genetic background than in *japonica* (Xu *et al.* 2015). In this study, we found that the diversity in the promoter region of *DEP1* led to a variation in panicle traits. Thus, these results will help to ensure that *DEP1* be used more flexibly and comprehensively and may lead to new strategies in the use of *DEP1* for rice breeding.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (No. 31501284, No. 31371587 and No. 31430062), China Postdoctoral Science Foundation Grant (No. 2014M560221 and 2015T80270). The authors declare no conflicts of interest associated with the research. We thank LetPub (www.letpub.com) for its linguistic assistance during the preparation of this manuscript.

Literature Cited

Abe, Y., K. Mieda, T. Ando, I. Kono, M. Yano, H. Kitano and Y. Iwasaki (2010) The *SMALL AND ROUND SEED1* (*SRS1/DEP2*) gene is involved in the regulation of seed size in rice. Genes Genet. Syst. 85: 327–339.

Botella, J.R. (2012) Can heterotrimeric G proteins help to feed the world? Trends Plant Sci. 17: 563–568.

Doyle, J. (1991) DNA protocols for plants. Molecular techniques in taxonomy, vol. 57. NATO ASI Series. Springer, Berlin, pp. 283–293.

Evanno, G., S. Regnaut and J. Goudet (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14: 2611–2620.

Hardy, O.J. and X. Vekemans (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol. Ecol. Notes 2: 618–620.

Huang, X., Q. Qian, Z. Liu, H. Sun, S. He, D. Luo, G. Xia, C. Chu, J. Li and X. Fu (2009) Natural variation at the *DEP1* locus enhances grain yield in rice. Nat. Genet. 41: 494–497.

Huang, X., X. Wei, T. Sang, Q. Zhao, Q. Feng, Y. Zhao, C. Li, C. Zhu, T. Lu, Z. Zhang *et al.* (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. Nat. Genet. 42: 961–967.

Khush, G.S. (1995) Breaking the yield frontier of rice. GeoJournal 35: 329–332.

Kosugi, S., I. Suzuka and Y. Ohashi (1995) Two of three promoter elements identified in a rice gene for proliferating cell nuclear antigen are essential for meristematic tissue-specific expression. Plant J. 7: 877–886.

Li, F., W. Liu, J. Tang, J. Chen, H. Tong, B. Hu, C. Li, J. Fang, M. Chen and C. Chu (2010) Rice DENSE AND ERECT PANICLE 2 is essential for determining panicle outgrowth and elongation. Cell Res. 20: 838–849.

Lu, L., W. Yan, W. Xue, D. Shao and Y. Xing (2012) Evolution and association analysis of *Ghd7* in rice. PLoS ONE 7: e34021.

Peng, S., G.S. Khush, P. Virk, Q. Tang and Y. Zou (2008) Progress in ideotype breeding to increase rice yield potential. Field Crops Res. 108: 32–38.

Pritchard, J.K., M. Stephens and P. Donnelly (2000) Inference of population structure using multilocus genotype data. Genetics 155: 945–959.

Qiao, Y., R. Piao, J. Shi, S.-I. Lee, W. Jiang, B.-K. Kim, J. Lee, L. Han, W. Ma and H.-J. Koh (2011) Fine mapping and candidate gene analysis of *dense and erect panicle 3*, *DEP3*, which confers high grain yield in rice (*Oryza sativa* L.). Theor. Appl. Genet. 122: 1439–1449.

Suh, H. and M. Heu (1978) The segregation mode of plant height in the cross of rice varieties. XI. Linkage analysis of the semi-dwarfness of the rice variety "Tongil". Korean J. Breed. 10: 1–6.

Sun, H., Q. Qian, K. Wu, J. Luo, S. Wang, C. Zhang, Y. Ma, Q. Liu, X. Huang, Q. Yuan *et al.* (2014) Heterotrimeric G proteins regulate nitrogen-use efficiency in rice. Nat. Genet. 46: 652–656.

Takahashi, Y., K.M. Teshima, S. Yokoi, H. Innan and K. Shimamoto (2009) Variations in Hd1 proteins, *Hd3a* promoters, and *Ehd1* expression levels contribute to diversity of flowering time in cultivated rice. Proc. Natl. Acad. Sci. USA 106: 4555–4560.

Wang, J., T. Nakazaki, S. Chen, W. Chen, H. Saito, T. Tsukiyama, Y. Okumoto, Z. Xu and T. Tanisaka (2009) Identification and characterization of the erect-pose panicle gene *EP* conferring high grain yield in rice (*Oryza sativa* L.). Theor. Appl. Genet. 119: 85–91.

Wolfenstetter, S., D. Chakravorty, R. Kula, D. Urano, Y. Trusov, M.B. Sheahan, D.W. McCurdy, S.M. Assmann, A.M. Jones and J.R. Botella (2015) Evidence for an unusual transmembrane configuration of AGG3, a class C Gγ subunit of Arabidopsis. Plant J. 81: 388–398.

Xu, Q., N. Xu, H. Xu, L. Tang, J. Liu, J. Sun and J. Wang (2014) Breeding value estimation of the application of *IPA1* and *DEP1* to improvement of *Oryza sativa* L. ssp. *japonica* in early hybrid generations. Mol. Breed. 34: 1933–1942.

Xu, Q., T. Liu, W. Bi, Y. Wang, H. Xu, L. Tang, J. Sun and Z. Xu (2015) Different effects of *DEP1* on vascular bundle- and panicle-related traits under *indica* and *japonica* genetic backgrounds. Mol. Breed. 35: 173.

Xu, Z., W. Chen, L. Zhang and C. Zhang (1995) The heredity of the erect panicle character and relations with other characters in rice. Journal of Shenyang Agricultural University 26: 1–7.

Yan, C.J., J.H. Zhou, S. Yan, F. Chen, M. Yeboah, S.Z. Tang, G.H. Liang and M.H. Gu (2007) Identification and characterization of a major QTL responsible for erect panicle trait in *japonica* rice (*Oryza sativa* L.). Theor. Appl. Genet. 115: 1093–1100.

Yang, X., S. Gao, S. Xu, Z. Zhang, B.M. Prasanna, L. Li, J. Li and J. Yan (2011) Characterization of a global germplasm collection and its potential utilization for analysis of complex quantitative traits in maize. Mol. Breed. 28: 511–526.

Yi, X., Z. Zhang, S. Zeng, C. Tian, J. Peng, M. Li, Y. Lu, Q. Meng, M. Gu and C. Yan (2011) Introgression of *qPE9-1* allele, conferring the panicle erectness, leads to the decrease of grain yield per plant in japonica rice (*Oryza sativa* L.). J. Genet. Genomics 38: 217–223.

Yu, J., G. Pressoir, W.H. Briggs, I.V. Bi, M. Yamasaki, J.F. Doebley, M.D. McMullen, B.S. Gaut, D.M. Nielsen, J.B. Holland *et al.* (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet. 38: 203–208.

Zhang, Z., E.S. Buckler, T.M. Casstevens and P.J. Bradbury (2009) Software engineering the mixed model for genome-wide association studies on large samples. Brief. Bioinformatics 10: 664–675.

Zhou, Y., J. Zhu, Z. Li, C. Yi, J. Liu, H. Zhang, S. Tang, M. Gu and G. Liang (2009) Deletion in a quantitative trait gene *qPE9-1* associated with panicle erectness improves plant architecture during rice domestication. Genetics 183: 315–324.