Theor Appl Genet (2008) 116:613–622
DOI 10.1007/s00122-007-0695-x

ORIGINAL PAPER

Fine mapping of a yield-enhancing QTL cluster associated with transgressive variation in an *Oryza sativa* × *O. rufipogon* cross

Xiaobo Xie · Fengxue Jin · Mi-Hee Song · Jung-Pil Suh · Hung-Goo Hwang · Yeon-Gyu Kim · Susan R. McCouch · Sang-Nag Ahn

Received: 30 July 2007 / Accepted: 4 December 2007 / Published online: 19 December 2007

Abstract A high-resolution physical map targeting a cluster of yield-related QTLs on the long arm of rice chromosome 9 has been constructed across a 37.4 kb region containing seven predicted genes. Using a series of BC_3F_4 nearly isogenic lines (NILs) derived from a cross between the Korean *japonica* cultivar Hwaseongbyeo and *Oryza rufipogon* (IRGC 105491), a total of seven QTLs for 1,000-grain weight, spikelets per panicle, grains per panicle, panicle length, spikelet density, heading date and plant height were identified in the cluster ($P \leq 0.0001$). All seven QTLs were additive, and alleles from the low-yielding *O. rufipogon* parent were beneficial in the Hwaseongbyeo background. Yield trials with BC_3F_4 NILs showed that lines containing a homozygous *O. rufipogon* introgression in the target region out-yielded sibling NILs containing Hwaseongbyeo DNA by 14.2–17.7%, and out-yielded the Hwaseongbyeo parent by 16.2–23.7%. While higher yielding plants containing the *O. rufipogon* introgression were also taller and later than controls, the fact that all seven of the QTLs were co-localized in the same 37.4 kb interval suggests the possibility that a single, pleiotropic gene acting as a major regulator of plant development may control this suite of agronomically important plant phenotypes.

Abbreviations
TGW 1,000-Grain weight

Communicated by Y. Xue.

Xiaobo Xie and Fengxue Jin have contributed equally to this study.

X. Xie · F. Jin · M.-H. Song · S.-N. Ahn (✉)
College of Agriculture and Life Sciences,
Chungnam National University, Daejeon 305-764, South Korea
e-mail: ahnsn@cnu.ac.kr

Present Address:
X. Xie
Institute of Horticulture,
Zhejiang Academy of Agricultural Sciences,
Hangzhou 310021, China

J.-P. Suh · H.-G. Hwang · Y.-G. Kim
National Institute of Crop Science,
Rural Development Administration,
Suwon 441-100, South Korea

S. R. McCouch
Department of Plant Breeding and Genetics,
Cornell University, Ithaca, NY 14850, USA

Introduction

Recent estimates suggest that a 40% increase in the production of rice is a must by 2030 to satisfy the needs of a growing population (Khush 2003). Wild relatives of rice are a rich source of desirable genes not only for yield but also for disease resistance, stress tolerance and other traits (Brar and Khush 1997; Xiao et al. 1998; Moncada et al. 2001; Thomson et al. 2003). Exploring wild and exotic rice germplasm for desireable genes and transferring them into cultivars through crossing and marker-assisted selection has been shown to be a feasible path for raising rice yields and augmenting stress resistance (Price et al. 2002; McCouch et al. 2007). At present, the large amount of germplasm preserved in gene banks (ex situ) and in situ throughout the world provides the groundwork for identifying new genes controlling yield and other valuable traits (Tanksley and McCouch 1997).

Recent progress in plant genome analysis has made it possible to examine the molecular basis of naturally occurring allelic variation underlying complex traits. Map-based

or positional cloning has been successful in isolating genes underlying QTLs in several plant species, including rice (Ashikari et al. 2005; Yano et al. 2000, Song et al. 2007), wheat (Yan et al. 2004), and tomato (Frary et al. 2000; Liu et al. 2002). A number of independent studies have reported the incidence of clustered QTLs for traits that are functionally related, including yield and yield components (Xiao et al. 1998; Moncada et al. 2001; Thomson et al. 2003), grain quality (Septiningsih et al. 2003; Li et al. 2004b), grain weight (Li et al. 2004a, Fan et al. 2006), as well as domestication-related traits and *indica/japonica* diagnostic traits (Thomson et al. 2003; Cai and Morishima 2002; Gu et al. 2005).

The present study is a part of a larger collaboration in which the same *O. rufipogon* donor was used to construct several BC_2F_2 populations using a diverse set of *O. sativa* cultivars as recurrent parents (RP) (McCouch et al. 2007). Each population was evaluated for a common set of phenotypes in the climatic and ecological zone where each RP was adapted. This experimental design made it possible to compare the phenotypic effects of specific *O. rufipogon* introgressions in multiple genetic backgrounds and expressed in different environments.

In this study, a nearly isogenic line (NIL) in the Hwaseongbyeo (temperate *japonica*) background contained a single *O. rufipogon*-derived introgression on chromosome 9 that was consistently associated with seven agronomically important traits. These included five yield components, plant height and heading date. This was consistent with previous reports from lower-resolution QTL studies in which an introgression from the same region of the *O. rufipogon* genome was associated with improved grain yield in a *tropical japonica* background (cv Jefferson) grown in the southern US (Thomson et al. 2003) and in an *indica* hybrid background grown in China (Xiao et al. 1998).

The development of a high-resolution map of the target region on chromsome 9 is the first step toward cloning the gene(s) underlying the grain yield QTLs and provides information that can be used to efficiently introgress the yield-enhancing *O. rufipogon* alleles into diverse rice cultivars to further evaluate their utility for plant improvement.

Materials and methods

Population development

The wild rice *O. rufipogon* (IRGC 105491) was used as a pollen parent in crosses with *O. sativa* spp. *japonica* cv. Hwaseongbyeo, an elite Korean cultivar, followed by two successive backcrosses made with Hwaseongbyeo as the recurrent parent (RP) as described previously (Cho et al. 2003). Among 820 BC_2F_1 plants, 172 individuals were selected based on phenotypic acceptability and selfed to generate BC_2F_2 families for QTL analysis. For this study, a single BC_2F_2 line, F132, was selected as the basis for fine-mapping (outlined in Fig. 1) because (a) it had an *O. rufipogon* introgression across the target region (identified by markers RM242 and RM215 on chromosome 9), (b) was associated with increased 1,000 grain weight (TGW), spikelets per panicle and panicle length and (c) had only eight non-target *O. rufipogon* segments elsewhere in the genome. F132 was backcrossed to Hwaseongbyeo and then allowed to self, generating an NIL-derived BC_3F_2 population that segregated for the region of interest on chromosome 9. One BC_3F_2 line, F10, was selected from this population using the same criteria as above, except that TGW was the only phenotype that was evaluated. Two BC_3F_3 derivatives, F10-67 and F10-70, were then used in progeny contrasts to fine-map the TGW QTL. Both of these lines were heterozygous across the target region on the long arm of chromosome 9 and contained an additional introgression on the short arm of chromosome 12 that was partly heterozygous and partly homozygous for *O. rufipogon* (Fig. 2). These two plants were selfed to form two BC_3F_3 sister populations, CR1416 and CR1417, consisting of 239 and 241 individuals, respectively (Fig. 1). The QTL for TGW, designated as *gw9.1*, was validated in both the populations. Subsequently, a series of BC_3F_4 sub-NILs (substitution lines) and controls were selected from CR1416 and CR1417 by marker assisted selection (Fig. 1). They were designated CR4340, CR4341, CR4344 (from CR1416), CR4350, CR4373, and CR4379 (from CR1417). Two of these lines, CR4340, containing *O. rufipogon* DNA across the entire region harboring *gw9.1* and CR4344, containing Hwaseongbyeo DNA were used as controls for yield trials.

Field trials

All the materials were grown in an experiment plot at Chungnam National University, Deajeon, Korea. They were sown on April 20 and transplanted on May 31. The growing density was 15 cm × 30 cm for a line × row. Yield trials were carried out using a completely randomized block design with three replications, and each block consisted of two rows with 30 plants per row. Before heading, the field was equipped with a bird-net to prevent bird damage.

Trait evaluation

Grains were allowed to dry naturally after harvesting and partially or un-filled seeds were removed with water. Fully-filled seeds were re-dried in an oven at 30°C for 24 h. 1,000-grain weight (TGW) was evaluated by measuring the

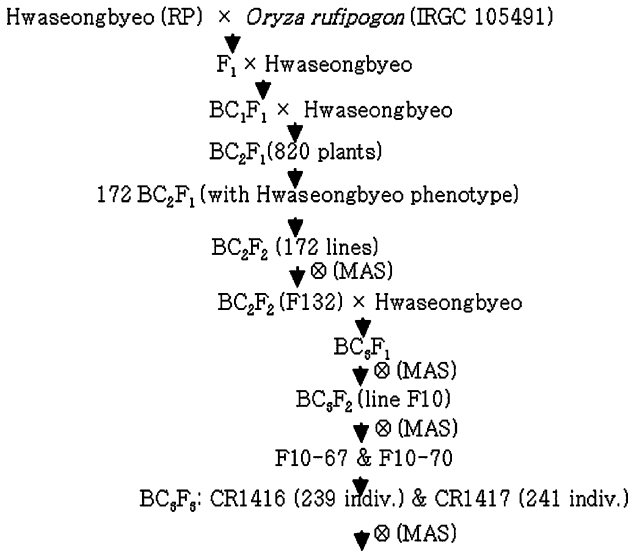

Fig. 1 Flow diagram showing how plant materials were developed

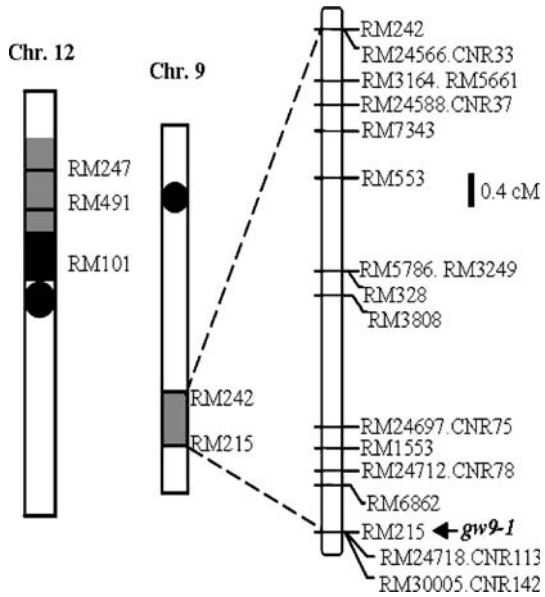

Fig. 2 Graphical genotype of BC3F3 pre-NIL CR1417 showing the location of *O. rufipogon* introgressions on chromosomes 9 (containing the QTL *gw9.1*) and 12 (spurious introgression, not associated with any QTL in this study). *Black circles* Centromeres, *white regions* Hwaseongbyeo homozygotes, *grey regions* heterozygotes, *black regions* homozygous *O. rufipogon* introgressions, *arrows* indicate position of the peak LOD score associated with *gw9.1* in this study based on an evaluation of 241 progeny derived from CR1417

weight of 100 randomly selected fully-filled grains averaged over three replications. Moisture content of the grain was measured using a Grain Moisture Meter (GMK-303) averaging over three replications. Days to heading, or heading date (HD) was evaluated as the number of days from seeding until heading, recorded as the emergence of the tip of the first panicle on a single plant. Plant height (PH) was measured from the ground surface to the tip of the tallest panicle (excluding the awn). Panicle length (PL) was measured from the panicle neck to the panicle tip (excluding the awn). Panicle number per plant (PN) was measured as the number of panicles per plant. Spikelets per panicle (SN), grain number per panicle (GN), and spikelet density on panicle (DEN) were measured by averaging three major panicles per plant. GN was measured as the number of filled seeds per panicle. DEN was measured as the number of spikelets per centimeter of panicle. Grain yield per plant (YD) was evaluated by averaging the grain yield (g) of 15 plants randomly selected from the center of each plot and dried to 10% moisture content. For 119 individual plants from one selected NIL, CR4379, mean grain length, grain width and grain thickness were measured on 100 randomly selected fully filled grains using the digimatic caliper (Mitutoyo Corp, Japan).

Marker development and analysis

New SSR markers were developed using the exon-anchored/intron-spanning approach described by Li et al. (2004a). Newly developed SSR markers were named according to the SSR nomenclature rules described in Temnykh et al. (2001). Briefly, rice microsatellite (RM) loci were identified based on the SSR locus annotation (International Rice Genome Sequencing Project 2005) and are available at http://sliver.plbr.cornell.edu/SSR and on the Gramene database (http://www.gramene.org) while the specific primer pairs used to detect the SSR loci (the *marker reagents*) which were designed as part of this study, are designated with the locus name and the suffix ".CNR" followed by a unique identifier (i.e., RM24718.CNR113). Primers were synthesized by Bioneer Co., Korea (http://www.bioneer.co.kr).

SSR markers (International Rice Genome Sequencing Project 2005; Temnykh et al. 2001; Panaud et al. 1996; Temnykh et al. 2000; McCouch et al. 2002) were used to survey the presence/absence of introgressions from *O. rufipogon* in the NILs and to evaluate the genotypes of segregating populations. DNA was extracted in micro-quantities using the method described in the study (McCouch et al. 1988) with slight modifications. Protocols for marker amplification using the polymerase chain reaction (PCR)and size separation using polyacrylamide gel electrophoresis (PAGE) and marker detection using the silver staining procedure were as described in Panaud et al. (1996). The silver staining kits were purchased from Bioneer Co., Korea (http://www.bioneer.co.kr).

Published SSR maps (Temnykh et al. 2001) were used as a reference to establish marker order. The recombinational distance between the SSR markers in the introgressed segment was determined based on segregation analysis in the

Table 1 Effect of *gw9.1* on GW in two sister lines at BC_3F_3

Line	QTL	*P*	Flanking markers	Mean ± SD (g)			Add effect[a]	R^2 (%)[b]
				HH	HR	RR		
CR1416	*gw9.1*	≤0.0001	RM24566.CNR33–RM215	25.0 ± 0.81	25.4 ± 0.71	26.5 ± 0.60	0.75	41.6
CR1417	*gw9.1*	≤0.0001	RM5661–RM215	25.7 ± 0.64	26.5 ± 0.79	27.7 ± 0.96	0.98	42.5

HH, RR Hwaseongbyeo and *O. rufipogon* homozygous genotypes, *HR* Hwaseongbyeo/*O. rufipogon* heterozygous genotype

[a] Additive effect = (RR − HH)/2

[b] Phenotype variation explained by QTL

BC_3F_3 generation using MAPMAKER Macintosh V2.0. Genetic distances were calculated in Kosambi centiMorgans (cM).

QTL fine mapping

QTLs were fine mapped by comparing the phenotypic means of genotypic classes of recombinants within the target region using the ANOVA feature in Data Desk 4.0. and based on the interval analysis ($P \leq 0.01$ and/or LOD ≥ 3.0) for markers within the target region using QGENE (Nelson 1997).

Results

NIL-based confirmation of *gw9-1*

To confirm results from a previous study (Cho et al. 2003) in which a QTL for 1,000 grain weight (TGW), *gw9.1*, was identified within a cluster of yield-related QTLs on the long arm of chromosome 9, two BC_3F_2 sister lines, F10-67 and F10-70, were developed from the original BC_2F_2 population as described in "Materials and methods" (Fig. 1). Each of these lines contained a 4.0 cM heterozygous introgression in the target region on chromosome 9, defined by markers RM242-RM215, as well as a second introgression on the short arm of chromosome 12 (Fig. 2). The introgression on chromosome 12 comprised a 7.6 cM region that was heterozygous and a smaller region that was fixed for *O. rufipogon* (Fig. 2). QTL analysis in BC_3F_3 progeny derived from F10-67 (the CR1416 population, comprised of 239 individuals) and F10-70 (the CR1417 population, comprised of 241 individuals) confirmed the association between TGW and markers in the target region on chromosome 9 ($P \leq 0.0001$ based on ANOVA). Interval analysis using the 18 markers shown in Fig. 2 indicated that *gw9.1* was tightly linked to RM215 (LOD = 29.67 in CR1417 and LOD = 31.9 in CR1416). There was no association between TGW and any of the markers in the introgression on chromosome 12 in either population.

Phenotypic characterization of *gw9-1*

The presence of *O. rufipogon* alleles at RM215 increased grain weight in the Hwaseongbyeo background. The phenotypic variation explained by the QTL *gw9.1* was 41.6–42.5% in the two sister lines, CR1416 and CR1417, respectively (Table 1). ANOVA analysis indicated that the additive effect of the *O. rufipogon* alleles was 0.75–0.98 g. When the CR1416 and CR1417 populations were divided into the three genotypic classes, the distributions overlapped, as illustrated in Fig. 3 for CR1417. The mean TGW for the Hwaseongbyeo homozygous class (HH), the heterozygotes (HR), and the *O. rufipogon* homozygous classs (RR) were 25.7 ± 0.64, 26.5 ± 0.79 and 27.7 ± 0.96, respectively, for CR1417 (Table 1). The means and distributions of TGW in CR1416 were almost identical to those observed in CR1417. While the three distributions over-

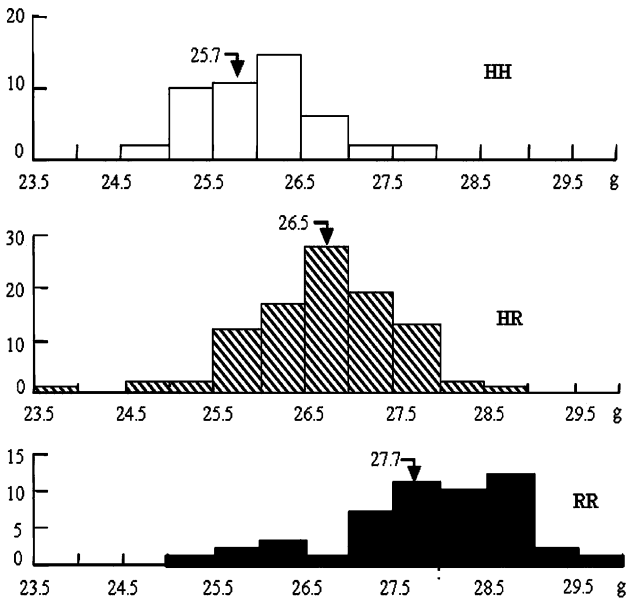

Fig. 3 Distribution of 1,000-grain weight (TGW) in BC_3F_3 progeny derived from CR1417 for the three genotypic classes defined by markers in the target region on chromosome 9, *white* corresponds to Hwaseongbyeo homozygotes (*HH*), *grey* corresponds to heterozygotes (*HR*) and *black* indicates *O. rufipogon* homozygotes (*RR*). *Arrows* show the mean TGW of each group

lapped with each other, the means were significantly different ($P < 0.01$, Duncan's Multiple Range Test).

Development of new SSR markers

Eight new SSR markers, confirmed to be polymorphic between Hwaseongbyeo and *O. rufipogon*, were developed in the introgressed region on chromosome 9 using sequence data from cv. Nipponbare (http://rgp.dna.affrc.go.jp/cgi-bin/statusdb/stattable.pl?chr=9). The names and primer sequences for these markers are provided in Table 2. Their orientations on chromosome 9 and their relationship to positions of previously published SSR markers in the interval RM242–RM215 are shown in Figs. 2 and 4. RM215, the marker most closely associated with *gw9.1*, was located near the interval RM24718.CNR113-RM30005.CNR142 on PAC clone P0229B10 (Table 2).

Substitution mapping of *gw9.1* using BC_3F_4 NILs

Six BC_3F_4 sub-NILs having informative recombination breakpoints within the introgressed region on chromosome

Table 2 Newly developed SSR markers at the interval RM242–RM215 on chromosome 9

Marker no.	Primer sequence	BAC/PAC clone	$(Motif)_n$[a]	Predicted size (bp)
RM24516.CNR31	5′ttgttgatcatggaggtgga3′; 5′gaccaccatgaggacgaagt3′	P0676H02	$(agc)_9$	156
RM24566.CNR33	5′gcttgaattcttgcccattg3′; 5′cgggaagaggtgtcattttc3′	OJ1011_C06	$(at)_{19}$	177
RM24583.CNR37	5′ctcccattcctcgctcattc3′; 5′ggccagtggtaggtgaaagt3′	OSJNBa005C18	$(gag)_8$	162
RM24697.CNR75	5′ttgtctctctctcccatgagc3′; 5′tggatcctagatatcgcatcgt3′	OJ1254_E07	$(tc)_{13}$	151
RM24712.CNR78	5′tccggtaacgtgatcttgaa3′; 5′aagatcgttggaggaacgag3′	P0569E11	$(cag)_{10}$	156
RM30004.CNR111	5′gctgtggccatttatgtatgc3′; 5′gatcatcgacgtgaatgaccta3′	P0229B10	$(tg)_5$	158
RM24718.CNR113	5′gaccaacgtgcatgtgactt3′; 5′gcttgcactagggctcctt3′	P0229B10	$(tcc)_7$	150
RM30005.CNR142	5′cgactcccactctcagatcc3′; 5′agattttcaccatcggttgc3′	P0229B10	$(tc)_5$	154

[a] The number to the right of the parenthesis indicates the repeat number

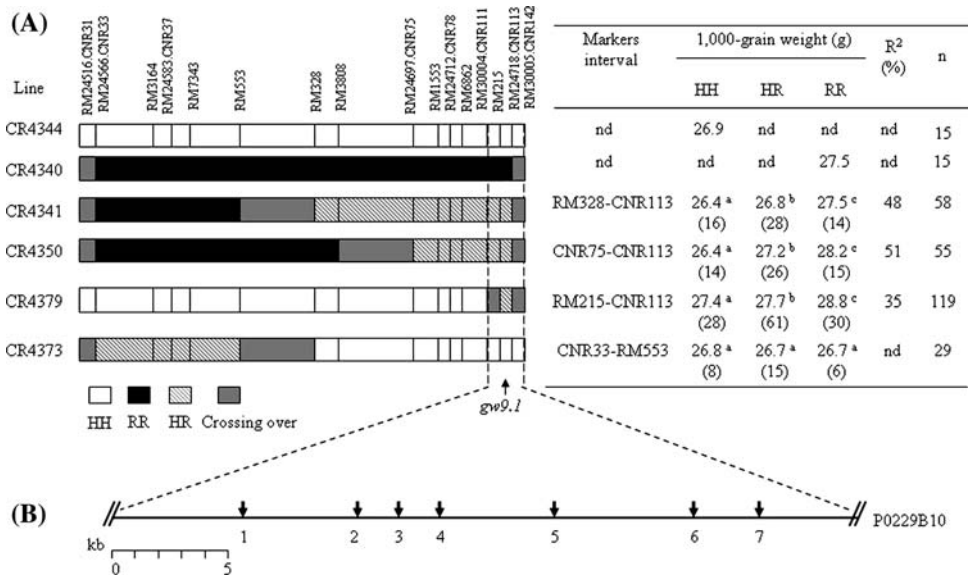

Fig. 4 Graphical representation of BC_3F_4 substitution lines and a fine-scale map of the target the region on chromosome 9. **a** *White*, *black* and *diagonal* slashes portions of the graph are homozygous Hwaseongbyeo, homozygous *O. rufipogon*, and heterozygous regions, respectively and *grey* portion regions are where crossing over occurred. *P* = probability value of ANOVA analysis. The table to the right of the graphical genotypes indicates mean TGW for each of the three genotypic classes of progeny (*HH* Hwaseongbyeo homozygotes, *HR* heterozygotes, and *RR O. rufipogon* homozygotes) derived from CR4341, CR4350, CR4379 and CR4373. *Broken vertical lines* define the interval containing the *gw9.1* locus. *N* Number of individuals evaluated in each line. Markers within heterozygous regions were tested and the highest R^2 scores are shown. [%]*nd* Not determined. Numbers followed by the same letter in each row are not significantly different at $P = 0.0001$. The difference of the mean values of TGW between CR4340 and CR4344 was significant at $P \leq 0.0001$. Numbers in () indicate the number of individual plants in each genotype class. **b** Detailed physical map showing the relative positions of the 37.4 kb candidate region containing 7 QTLs including *gw9.1* based on sequence annotation (http://ricegaas.dna.affrc.go.jp). Numbers indicate putative gene number

Table 3 QTLs for heading date and plant height in two BC_3F_4 lines

QTL	Line[a]	Pop. size	*P*	Linked marker	Mean ± SD			Add effect	R^2 (%)[b]
					HH	HR	RR		
hd9.1	CR4341	56	≤0.0001	RM215	100.3 ± 1.1	101.7 ± 1.5	104.0 ± 1.5	1.9	51.8
	CR4350	113	≤0.0001	RM215	99.4 ± 1.0	100.5 ± 1.1	103.2 ± 1.4	1.9	58.6
	CR4379	238	≤0.0001	RM215	84.2 ± 1.0	85.0 ± 1.4	88.4 ± 1.4	2.1	46.8
ph9.1	CR4350	113	≤0.0001	RM215	106.4 ± 4.3	109.0 ± 4.5	111.0 ± 3.4	2.3	15.8
	CR4379	238	≤0.0001	RM215	102.3 ± 3.0	103.1 ± 3.0	105.7 ± 3.1	1.7	13.2

[a] NIL CR4379 was sown on May 25 and transplanted on June 9, 2004. The other lines were sown on May 4 and transplanted on June 3, 2004

[b] Percent of the phenotypic variation explained by each QTL

9 were identified and used for fine mapping of *gw9.1*. The sub-NIL families were designated CR4340, CR4341, CR4344, CR4350, CR4373 and CR4379 (Fig. 4). All were fixed for recurrent parent (Hwaseongbyeo) DNA in the introgressed region on chromosome 12, except for sub-family CR4379 that remained heterozygous, and all retained the smaller segment of fixed *O. rufipogon* DNA near the centromere on chromosome 12.

As indicated in the table in Fig. 4, there were significant differences in TGW between NIL families CR4344 and CR4340 ($P \leq 0.0001$, ANOVA), which served as positive and negative controls, respectively, for the region as a whole. The TGW of CR4373 was not significantly different than that of CR4344 and the three genotypic classes for the CR4373 did not show significant differences, suggesting that CR4373 did not contain *O. rufipogon* alleles affecting grain weight. A comparison of TGW among NIL progeny showed significant differences among the three genotypic classes for the CR4341, CR4350 and CR4379 populations, while there were no significant differences in the CR4373 population ($P \leq 0.0001$, ANOVA). When interval analysis was conducted on each of the NIL-derived populations, *gw9.1* consistently mapped nearest to marker RM215, and this marker explained 35–51% of the phenotypic variation for TGW. This was consistent with the segregation data for CR4373 described above and allowed us to conclude that *gw9.1* was not located in the interval RM24566.CNR33–RM328. Segregation analysis of 119 progeny derived from NIL CR4379 further confirmed there was no significant association between the markers and phenotype in the introgressed region on chromosome 12.

Based on the size of the chromosome 9 introgression in CR4379, it was concluded that *gw9.1* was located in the interval RM24718.CNR111-RM30005.CNR142, a region of approximately 37.4 kb (Fig. 4) (http://rgp.dna.affrc.go.jp/cgi-bin/statusdb/stattable.pl?chr=9). This position, as determined by substitution mapping, was consistent with the results of interval analysis in the BC_3F_3 generation (Fig. 2). It should be noted that markers RM30004.CNR111 and RM30005.CNR142 represented the outside borders of the introgression in CR4379 (Fig. 4), and we therefore conclude that the *gw9.1* locus has actually been localized to a region, <37.4 kb in size.

To understand which aspect of grain shape trait was responsible for the increase in grain weight at *gw9.1*, ANOVA was carried out using three grain shape traits, grain length, grain width and grain thickness evaluated for progeny from CR4379. A comparison of three traits among NIL progeny showed significant differences among the three genotypic classes at the RM215 locus ($P \leq 0.0001$, $P = 0.0072$ and $P = 0.028$, respectively) (data not shown). Correlation analysis also indicated that 1,000-grain weight showed the highest significant correlation with grain length ($r = 0.625$***), followed by grain width ($r = 0.592$***) and grain thickness ($r = 0.239$*), and this result seems to imply that increase in grain length and width is primarily responsible for the increase in grain weight at *gw9.1*.

Mapping of QTLs for HD and PH

In the BC_3F_4 populations analyzed for TGW, it was noted that plants also appeared to be segregating for heading date and plant height. In a previous study (Cho et al. 2003), two QTLs, *hd9.1* and *ph9.1*, controlling heading date and plant height, respectively, were reported in the cluster of QTLs that mapped to the target region on chromosome 9. When all three phenotypes (TGW, HD and PH) were evaluated in the same NIL-derived populations, CR4341, CR4350 and CR4379, they all co-segregated with SSR marker RM215 ($P \leq 0.0001$) (Table 3). Based on the segregation observed in CR4379 (Fig. 4), both *hd9.1* and *ph9.1* were determined to co-localize with *gw9.1* within the interval RM24718.CNR111–RM30005.CNR142. In all cases, the *O. rufipogon* alleles showed an additive effect, with *O. rufipogon* delaying heading by about 4 days and increasing plant height by about 4 cm (Table 3). There was no association between heading date or plant height with any of the markers that segregated on chromosome 12 in NIL CR4379.

Table 4 QTLs for SN, GN, PL and DEN in NIL CR4379 (BC_3F_4)

QTLs	Pop size	*P*	Linked marker	Mean ± SD			Add effect	R^2 (%)
				HH[a]	HR	RR		
sn9.1	119	≤0.0001	RM215	115.8 ± 11.1a[b]	120.7 ± 12.8b	138.7 ± 11.7c	11.5	27.2
gn9.1	119	≤0.0001	RM215	109.6 ± 9.8a	114.2 ± 12.6b	130.4 ± 10.7c	10.4	25.7
pl9.1	119	≤0.0001	RM215	20.9 ± 1.1a	21.8 ± 0.9b	22.4 ± 0.8c	0.75	25.6
den9.1	119	≤0.0001	RM215	5.5 ± 0.4a	5.5 ± 0.5a	6.2 ± 0.5b	0.35	14.1
yd9.1	115	≤0.05	RM215	25.7 ± 3.1a	27.0 ± 4.3a	28.5 ± 3.7c	1.4	9.2

SN Spikelets per panicle, *GN* grains per panicle, *PL* panicle length, *DEN* panicle density, *YD* yield per plant

[a] HH, RR, and HR: Hwaseongbyeo, *O. rufipogon* homozygous and heterozygous genotypes, respectively

[b] Numbers followed by the same letter in each row are significantly different

Mapping of QTLs associated with other yield components

In the NIL-derived population CR4379, five additional phenotypes (spikelet number per panicle, SN; grain number per panicle, GN; spikelet density per panicle, DEN; panicle length, PL; and yield, YD) showed co-segregation with RM215 ($P \leq 0.0001$ for SN, GN, DEN and PL, $P \leq 0.05$ for YD) and coincided with the position of previously reported QTLs in this population (Cho et al. 2003) (Table 4). All *O. rufipogon* alleles associated with yield and yield components were transgressive and beneficial in the Hwaseongbyeo background (Table 4). Thus, *sn9.1, gn9.1, den9.1 pl9.1*, and *yd9.1* co-localized with *gw9*, *hd9.1* and *ph9.1* in the 37.4 kb interval flanked by markers RM24718.CNR113–RM30005.CNR142 on chromosome 9. Phenotype evaluation confirmed that all the QTLs showed additive effects, such that an *O. rufipogon* introgression in this region added 11.5 spikelets for SN, 10.4 grains for GN, 0.35 spikelets/cm for DEN, 0.75 cm for PH, and 1.4 g/plant for YD. There was no association between these phenotypes and any of the markers in the chromosome 12 introgression.

Impact of the QTL cluster on grain yield per plant

Two BC_3F_4 NILs, CR4340 and CR4344 containing a homozygous *O. rufipogon* and Hwaseongbyeo DNA in the target region, respectively, were used for yield trials together with Hwaseongbyeo parental controls in 2004 whereas the two NILs CR4340 and CR4379 (plants containing a homozygous *O. rufipogon* DNA in the target region were selected) were evaluated with the Hwaseongbyeo parent as a control in 2005. The yield trials were conducted using a completely randomized block design with three replications. Results showed that the average grain yield per plant of CR4340 was 17.7% and 23.7% higher than that of CR4344 and Hwaseongbyeo, respectively. And the average grain yield per plant of CR4379 was 14.2 and 16.2% higher than that of CR4344 and Hwaseongbyeo, respectively. The difference in grain yield per plant was significant ($P \leq 0.05$) between CR4340 and CR4344 in 2004, and between CR4344 and CR4379 in 2005, but there was no significant difference between CR4344 and Hwaseongbyeo (Table 5).

Candidate genes in the 37.4-kb region

On the basis of available sequence annotation (http://rice-gaas.dna.affrc.go.jp/), there are seven predicted genes (cnu1.AutoPredgene01–cnu1.AutoPredgene07) in the 37.4 kb target region. Of these genes, two have unknown functions and five have functional annotations; (1) cnu1.AutoPredgene01 (gene ID) is a gene with a transcript length of 756 bp containing two exons, classified as a putative anthocyanin biosynthesis regulatory protein P1; (2) cnu1.AutoPredgene03 is a gene with a transcript length of

Table 5 Comparison of grain yield and yield components among Hwaseongbyeo and two BC_3F_4 NILs

Line	HD[a] 2005	PH 2005	PL 2005	SN 2005	DEN 2005	YD 2004	YD 2005	YD Index 2004	YD Index 2005
Hwaseongbyeo	112a[b]	107a	19a	118a	5.2a	27.4a	27.7a	100	100
CR4344	112a	108a	19a	119a	5.4a	28.8a	28.2a	105.1	101.8
CR4379	115b	117b	21b	129b	5.8b	–	32.8b	–	118.4
CR4340	–	–	–	–	–	33.9b	–	123.7	–

[a] Refer to trait abbreviations in Tables 1, 3, and 4

[b] Numbers followed by the same letter in each column are not significantly different at $P < 0.05$

1548 bp with four exons and a putative kinesin light chain (isoform 2); (3) cnu1.AutoPredgene04 is a gene with a transcript of 795 bp with eight exons, classified as a putative crystal structure of recombinant ascorbate peroxidase; (4) cnu1.AutoPredgene06 is a gene with a transcript of 2,184 bp which was classified as a putative bZIP protein; and (5) cnu1.AutoPredgene07 is a gene with a transcript of 1,620 bp, classified as a putative CLB1 (Ca^{2+}-dependent lipid-binding domain) protein.

Discussion

The original target of this study was the QTL for TGW, *gw9.1*, originally mapped on the long arm of chromosome 9 (Cho et al. 2003). In the process of fine-mapping this trait, QTLs for four additional yield components, as well as heading date and plant height, were consistently detected in the same region. Substitution lines confirmed that the QTL for TGW resided in the 37.4 kb region and the six additional QTLs were co-localized in the same interval.

TGW is a trait that is difficult to measure in individual plants because grains can vary in size and weight along the same panicle or between individuals with the same genotype. The fact that we were able to fine-map TGW underscores the power of NIL development and substitution mapping in high-resolution genetic analysis (Li et al. 2004a; Wissuwa et al. 2002). As documented in this study, R^2 values steadily increased with advanced generations of backcrossing, from 12% in the BC_2F_2 generation to between 41.6 and 42.5% in the BC_3F_3 generation and up to 51% in the BC_3F_4 generation of NILs. As the number of spurious donor (i.e., *O. rufipogon*) introgressions in the genetic background decreased and the linkage between markers and the target gene(s) increased, the proportion of the phenotypic variation that could be explained by the markers was greatly enhanced. The ease of developing new markers was greatly facilitated by the availability of rice genome sequence (International Rice Genome Sequencing Project 2005; Goff et al. 2002; Yu et al. 2002) and the new markers were invaluable in defining the positions of recombinants and clarifying the borders of informative introgressions.

Two independent studies using the same *O. rufipogon* accession (IRGC 105491) as a donor in AB-QTL analysis, previously reported QTLs associated with the yield-components in a similar region on chromosome 9. Xiao et al. (1998) reported QTLs controlling TGW, panicle length and spikelet number per panicle linked to markers RZ422 and RG386, which define the same region as RM215. Thomson et al. (2003) reported one QTL each for yield, TGW, panicle length, spikelet number per panicle and grain number per panicle in association with RM242 and RM215. In that study all the *O. rufipogon* alleles for the yield-related traits were beneficial in the cv Jefferson (a *tropical japonica*) background, except that the *O. rufipogon* allele was associated with smaller grains (Thomson et al. 2003). This suggests an array of as-yet poorly understood interactions between the *O. rufipogon* introgressions and genes in the background of these varieties, given that the Hwaseong parent has small seeds typical of the temperate *japonica* subpopulation, while Jefferson is known for its long, slender grains (McClung et al. 1997). It appears that the *O. rufipogon* allele is associated with an increase in grain size in the temperate *japonica* but a decrease in the *tropical japonica* background. This offers an opportunity to further explore the molecular basis of the epistatic interactions governing grain size in rice. The detection of QTLs for the same yield-related traits in three independent studies, one in an *indica* background evaluated in southern China, one in a *tropical japonica* background evaluated in the southern US, and one in a temperate *japonica* background evaluated in Korea, provides strong support for the hypothesis that *O. rufipogon* alleles at this cluster of QTLs are likely to make an important contribution to enhancing yield in different elite cultivars of *O. sativa.*

Yield trials using fixed BC_3F_4 substitution lines confirmed that NILs containing *O. rufipogon* DNA in the target region significantly out-yielded NILs with Hwaseongbyeo DNA in the same region as well as parental (Hwaseongbyeo) controls. The grain yield per plant in the *O. rufipogon* NILs was 14.2–17.7% higher than that of the corresponding Hwaseongbyeo NILs, and 16.2–23.7% higher than that of the Hwaseongbyeo parent. This result confirmed that the yield-component QTLs identified in primary QTL analysis contributed directly to grain yield under field conditions. It also confirmed earlier reports about transgressive variation for yield and other traits coming from the *O. rufipogon* parent (Xiao et al. 1998; Moncada et al. 2001; Thomson et al. 2003; Septiningsih et al. 2003) because alleles from *O. rufipogon* enhanced the yield of the higher-yielding Hwaseongbyeo parent.

Two QTLs associated with grain shape traits were isolated in rice. Fan et al. (2006) elucidated the molecular mechanism of the gene underlying *GS3* for grain length and width. The gene encodes 232 amino acids with a putative PEBP-like domain, a transmembrane region, a putative TNFR/NGFR family cysteine-rich domain and a VWFC module and *GS3* might function as a negative regulator. A newly discovered quantitative trait locus, *GW2*, which encodes a new RING-type E3 ubiquitin ligase, has been found to regulate rice grain width and weight (Song et al. 2007). These two cloned genes appeared not to be functionally related to the five predicted genes with functional annotations in this study, raising the possibility that another unknown mechanism might be responsible for regulation of

grain weight at *gw9.1*. However, the sequence annotation using sequence data of cv. Nipponbare in this study would make it difficult to predict a candidate gene underlying *gw9.1*, which was derived from the *O. rufipogon*. This is mainly due to divergence in sequence level between the cultivated rice and its wild counterpart (Ammiraju et al. 2006). Ammiraju et al. (2006) revealed that the orthologous region of the *Adh1* gene in the *O. sativa* genome was 9.4–28% larger relative to four wild species including *O. rufipogon* and this size variation in the *Adh1* gene was mainly due to insertion of transposable elements as well as multiple genetic mechanisms. Thus, the candidate gene should be confirmed based on the complete sequence of *O. rufipogon* in the target region and we are sequencing the target region of *O. rufipogon* to determine the gene underlying *gw9.1* in this regard.

Exotic or wild-QTL alleles that are favorable for some traits may be associated with deleterious effects on other traits (Xiao et al. 1998; Ragot et al. 1995). In this research, *O. rufipogon* alleles in the target region on chromosome 9 had a favorable effect on yield, while at the same time, *O. rufipogon* alleles in this region delayed heading and increased plant height. Longer season rice varieties are undesirable if they limit the regions in which the variety can be grown or eliminate the possibility of planting a second or third crop (Xiao et al. 1998). Even in a single season, delayed maturity is often associated with a higher chance of encountering bad weather conditions. On the other hand, delayed maturity may contribute favorably to increased yields because of a longer grain-filling period. Increased plant height without a simultaneous increase in stem thickness makes rice plants more susceptible to lodging, leading indirectly to yield loss. In this study, the strong correlation between plant height and panicle length suggests that these structural features may be under common genetic control. Thus, plant breeders who wish to take advantage of these valuable yield-enhancing alleles from *O. rufipogon* will want to complement it with alleles that promote early heading and shorter plant stature at other loci in the rice genome, as demonstrated by Ashikari et al. (2005). Cloning the gene(s) underlying this cluster of QTLs on chromosome 9 is the first step toward understanding how the gene(s) functions and is likely to offer insights into how variation at this locus can be manipulated to provide the desired yield advantage without delaying heading or increasing plant height. Numerous genes associated with crop performance such as heading date (Yano et al. 2000; Kojima et al. 2002), tiller number (Li et al. 2003), erect leaves (Sakamoto et al. 2006), grain weight and length (Fan et al. 2006; Song et al. 2007) and spikelets per panicle (Ashikari et al. 2005) have already been cloned, and it will be of great interest to understand what gene(s) underlies this QTL cluster and how it/they interact with alleles at other loci to generate transgressive variation for this valuable suite of agronomically important phenotypes.

Acknowledgments This study was supported by grants to S.N.A. from the BioGreen 21 project (Code No. 20070301034034) of the RDA, from the Crop Functional Genomics Center of the 21st Century Frontier Research Program (Project no. CG3113), Republic of Korea.

References

Ammiraju JSS, Luo M, Goicoechea JL, Wang W, Kudma D et al. (2006) The *Oryza* bacterial artificial chromosome library resource: construction and analysis of 12 deep-coverage large-insert BAC libraries that represent the 10 genome types of the genus Oryza. Genome Res 16:140–147

Ashikari M, Sakakibara H, Lin S, Yamamoto T, Takashi T, Nishimura A, Angeles E, Qian Q, Kitano H, Matsuoka M (2005) Cytokinin oxidase regulates rice grain production. Science 309:741–745

Brar DS, Khush G (1997) Alien introgression in rice. Plant Mol Biol 35:35–47

Cai HW, Morishima H (2002) QTL clusters reflect character associations in wild and cultivated rie. Theor Appl Genet 104:1217–1228

Cho YC, Suh JP, Choi IS, Hong HC, Beak MK, Kang KH, Kim YG, Ahn SN, Choi HC, Hwang HG, Moon HP (2003) QTLs analysis of yield and its related traits in wild rice relative *Oryza rufipogon*. Treatises Crop Res Korea 4:19–29

Fan C, Xing YZ, Mao HL, Lu TT, Han B, Xu C, Li XH, Zhang Q (2006) *GS3*, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. Theor Appl Genet 112:1164–1171

Frary A, Nesbitt TC, Frary A, Grandillo S, Knaap EVD, Cong B, Liu JP, Meller J, Elber R, Alpert KB, Tanksley SD (2000) *fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. Science 289:85–88

Goff SA, Ricke D, Lan TH, Presting G, Wang R et al. (2002) A draft sequence of the rice genome (*Oryza sative* L. spp. *japonica*). Science 296:92–100

Gu XY, Kianian SF, Foley ME (2005) Dormancy genes from weedy rice respond divergently to seed development environments. Genetics 171:695–704

International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. Nature 436:793–800

Khush G (2003) What will it take to feed 5.0 billion rice consumers in 2030? Plant Mol Biol 59:1–6

Kojima S, Takahashi Y, Kobayashi Y, Monna L, Sasaki T, Araki T, Yano M (2002) *Hd3a*, a rice ortholog of the *Arabidopsis* FT gene, promotes transition to flowering downstream of *Hd1* under short-day conditions. Plant Cell Physiol 43:1096–1105

Li X, Qian Q, Fu Z, Wang Y, Xiong G, Zeng D, Wang X, Liu X et al (2003) Control of tillering in rice. Nature 422:618–621

Li J, Xiao J, Grandillo S, Jiang L, Wan Y, Deng Q, Yuan L, McCouch SR (2004a) QTL detection for rice grain quality traits using an interspecific backcross population derived from cultivated Asian (*O. sativa* L.) and African (*O. glaberrima* S.) rice. Genome 47:697–704

Li J, Thomson M, McCouch SR (2004b) Fine mapping of a grain-wieght quantitative trait locus in the pericentromeric region of rice chromosome 3. Genetics 168:2187–2195

Liu J, Van Eck J, Cong B, Tanksley SD (2002) A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. Proc Natl Acad Sci USA 99:13302–13306

McClung AM, Marchetti MA, Webb BD, Bollich CN (1997) Registration of "Jefferson" rice. Crop Sci 37:629–630

McCouch SR, Kochert G, Yu ZH, Wang ZY, Khush G, Coffman WR, Tanksley SD (1988) Molecular mapping of rice chromosomes. Theor Appl Genet 76:815–829

McCouch SR, Teytelman L, Xu Y, Lobos KB, Clare K, Walton M, Fu B, Maghirang R, Li Z, Xing Y et al. (2002) Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.). DNA Res 9:199–207

McCouch SR, Sweeney M, Li J, Jiang H, Thomson M, Septiningsih E, Edwards J, Moncada P, Xiao J, Garris A, Tai T, Martinez C, Tohme J, Sugiono M, McClung A, Yuan L, Ahn SN (2007) Through the genetic bottleneck: *O. rufipogon* as a source of trait-enhancing alleles for *O. sativa*. Euphytica 154:317–339

Moncada P, Martinez CP, Borrero J, Chatel M, Gauch H, Guimaraes E, Tohme J, McCouch SR (2001) Quantitative trai loci for yield and yield components in an *Oryza sativa* × *Oryza rufipogon* BC_2F_2 population evaluated in an upland environment. Theor Appl Genet 102:41–52

Nelson JC (1997) QGENE: software for marker-based genomic analysis and breeding. Mol Breed 3:239–245

Panaud O, Chen X, McCouch SR (1996) Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice (*Oryza sativa* L.). Mol Gen Genet 252:597–607

Price AH, Townend J, Jones MP, Audebert A, Courtois B (2002) Mapping QTLs associated with drought avoidance in upland rice grown in the Philippines and West Africa. Plant Mol Biol 48:683–695

Ragot M, Sisco PH, Hoisington DA, Stuber CW (1995) Molecular-marker-mediated characterization of favourable exotic alleles at quantitative trait loci in maize. Crop Sci 35:1306–1315

Sakamoto T, Morinaka Y, Ohnishi T, Sunohara H, Fujioka S, Ueguchi-Tanaka M, Mizutani M, Sakata K, Takatsuto S, Yoshida S, Tanaka H, Kitano H, Matsuoka M (2006) Erect leaves caused by brassinosteroid deficiency increase biomass production and grain yield in rice. Nat Biotechnol 24:105–109

Septiningsih EM., Trijatmiko S, 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

Song XJ, Huang W, Shi M, Zhu MZ, Lin HX (2007) A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. Nat Genet 39:623–630. doi:10.1038/ng2014

Tanksley SD, McCouch SR (1997) Seed banks and molecular maps; unlocking genetic potential from the wild. Science 277:1063–1066

Temnykh S, Park W, Ayres N, Cartinhour S, Hauck N, Lipovich L, Cho YG, Ishii T, McCouch SR (2000) Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). Theor Appl Genet 100:697–712

Temnykh S, DeClerk G, Lukashova A, Lipovich L, Cartinhour S, McCouch SR (2001) Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length, transposon associations, and genetic marker potential. Genome Res 11:1141–1452

Thomson M J, Tai TH, McClung AM, Lai XH, Hinga ME, Lobos KB, Xu Y, Martinez CP, McCouch SR (2003) Mapping quantitative trait loci for yield, yield components and morphological traits in an advanced backcross population between *Oryza rufipogon* and the *Oryza sativa* culivar Jefferson. Theor Appl Genet 107:479–493

Wissuwa M, Wegner J, Ae N, Yano M (2002) Substitution mapping of *Pup1*: a major QTL increasing phosphorus uptake of rice from a phosphorus-deficient soil. Theor Appl Genet 105:890–897

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

Yan L, Loukoianov A, Blechl A, Tranquilli G, Ramakrishna W, San-Miguel P, Bennetzen JL, Echenique V, Dubcovsky J (2004) The wheat VRN2 gene is a flowering repressor down-regulated by vernalization. Science 303:1640–1644

Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, Sasaki T (2000) *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene CONSTANS. Plant Cell 12:2473–2483

Yu J, Hu SN, Wang J, Wong GKK, Li S et al (2002) A draft sequence of the rice genome (*Oryza sativa* L. spp. *indica*). Science 296:79–92