Identification of Trait-Improving Quantitative Trait Loci Alleles From a Wild Rice Relative, *Oryza rufipogon*

Jinhua Xiao,*,[1] Jiming Li,*,† Silvana Grandillo,* Sang Nag Ahn,*,[2] Longping Yuan,† Steven D. Tanksley* and Susan R. McCouch*

**Department of Plant Breeding, Cornell University, Ithaca, New York 14853 and †China National Hybrid Rice Research and Development Center, Changsha, Hunan 410125, People's Republic of China*

Manuscript received November 7, 1997
Accepted for publication July 13, 1998

ABSTRACT

Wild species are valued as a unique source of genetic variation, but they have rarely been used for the genetic improvement of quantitative traits. To identify trait-improving quantitative trait loci (QTL) alleles from exotic species, an accession of *Oryza rufipogon*, a relative of cultivated rice, was chosen on the basis of a genetic diversity study. An interspecific BC_2 testcross population (V20A/*O. rufipogon*//V20B///V20B////Ce64) consisting of 300 families was evaluated for 12 agronomically important quantitative traits. The *O. rufipogon* accession was phenotypically inferior for all 12 traits. However, transgressive segregants that outperformed the original elite hybrid variety, V20A/Ce64, were observed for all traits examined. A set of 122 RFLP and microsatellite markers was used to identify QTL. A total of 68 significant QTL were identified, and of these, 35 (51%) had beneficial alleles derived from the phenotypically inferior *O. rufipogon* parent. Nineteen (54%) of these beneficial QTL alleles were free of deleterious effects on other characters. *O. rufipogon* alleles at two QTL on chromosomes *1* and *2* were associated with an 18 and 17% increase in grain yield per plant, respectively, without delaying maturity or increasing plant height. This discovery suggests that the innovative use of molecular maps and markers can alter the way geneticists utilize wild and exotic germplasm.

WILD relatives of crop species have been given considerable attention in germplasm collections, because they are known to contain a large proportion of the existing genetic variation for these species. In rice, the majority of genetic variation in the genus Oryza still lies untapped in wild relatives (Wang *et al.* 1992). This is presumably due to the genetic bottlenecks that accompanied the domestication process. Intensive modern breeding efforts have contributed to a narrowing of the gene pool by further concentrating favorable alleles already present in early domesticates (Simmonds 1976; Ladizinsky 1985; Debouck 1991).

Though wild and unadapted germplasm is phenotypically less desirable than modern varieties in its overall appearance and performance, breeders have long recognized the intrinsic value of wild species for the improvement of simply inherited traits, including disease and insect resistance or cytoplasmic male sterility. Among the most successful examples of utilizing wild germplasm in the history of rice breeding include the use of *Oryza nivara* genes to provide long-lasting resistance to grassy stunt virus (Khush *et al.* 1977; Plucknett *et al.* 1987) and the use of *O. spontanea* as the source of wild abortive cytoplasmic male sterility, which has provided the cornerstone for today's hybrid rice (Li and Zhu 1988).

Despite these successes, it has been virtually impossible to utilize wild germplasm for the improvement of quantitatively inherited traits, such as yield, because the superior trait of interest cannot be identified phenotypically in the wild accessions. For most quantitative traits, a phenotype is conditioned by several genes having either trait-enhancing ("positive") or trait-depressing ("negative") alleles. In elite cultivars, the agriculturally "positive" alleles are represented in high frequency, while the agriculturally "negative" alleles, though still present in the gene pool, are relatively rare. For undomesticated germplasm, agriculturally desirable alleles are present in low frequency and are often masked by the effects of deleterious alleles. Because the overall phenotype of most wild species is agronomically undesirable, it is frequently concluded that this germplasm has low breeding value; *i.e.*, there are no trait-enhancing alleles present in the genotype.

The advent of molecular markers and maps makes it possible to identify individual quantitative trait loci (QTL) associated with yield and its components, environmental stress tolerance, disease and insect resistance, and quality traits in a variety of crop plants (for reviews see Tanksley 1993; McCouch and Doerge 1995; Stuber 1995; Plant Genome Database: http://

Corresponding author: Susan R. McCouch, Department of Plant Breeding, 252 Emerson Hall, Cornell University, Ithaca, NY 14853-1902. E-mail: srm4@cornell.edu

[1] *Present address:* Monsanto Corporation, St. Louis, MO 63198.

[2] *Present address:* Crops Experiment Station, Rural Development Administration, 109 Seodum Dong, Suweon, 441-100, Korea.

probe.nalusda.gov:8300/). The maps and markers represent a powerful tool not only for identifying positive QTL alleles but also for facilitating the selection of recombinant genotypes.

An early study by Frey *et al.* (1983) in cereals (oats, barley, sorghum, and pearl millet) has shown that, despite their overall inferior agronomic performance, wild and weedy species are likely to contain genetic factors that can increase the yield of modern varieties. Recent evidence from a study in tomato demonstrated that molecular genetic maps can be used to efficiently exploit the genetic potential of wild species for the improvement of yield and quality in elite processing tomatoes (Tanksley *et al.* 1996). Using a strategy referred to as advanced backcross QTL analysis (Tanksley and Nelson 1996), valuable QTL were simultaneously discovered and transferred from wild and unadapted germplasm into elite breeding lines. Here we report a molecular marker-facilitated study in rice to determine the potential for using the wild relative, *O. rufipogon*, to improve key quantitative traits of agronomic importance in an elite Chinese hybrid variety.

MATERIALS AND METHODS

Selection of wild species and cultivated parents: The genus Oryza includes about 20 wild species, as well as 2 cultivated species, *Oryza sativa* L., of Asian origin, and *O. glaberrima*, of African origin (Chang 1984). Six of the wild species share the AA genome with cultivated rice and can be hybridized through sexual crossing.

In 1991, we obtained 42 accessions of rice germplasm containing the AA genome from six relatives of *O. sativa* (*O. glaberrima*, *O. barthii*, *O. glumaepatula*, *O. nivara*, *O. spontanea*, and *O. rufipogon*) from the International Rice Germplasm Collection (IRGC) at the International Rice Research Institute (IRRI). Thirty-four of these accessions, along with 15 accessions from the cultivated species, were probed with 25 RFLP markers distributed on the 12 chromosomes of rice to determine the degree of genetic distance between the wild and the cultivated gene pools. Molecular data were subjected to a principal component analysis (Figure 1). One accession of *O. rufipogon* (IRGC 105491), indicated by a filled circle in Figure 1, was chosen as the wild donor for this study because *O. rufipogon* was genetically very close to cultivated rice, and accession was selected because it gave fertile F_1 offspring when crossed with both *Indica* and *Japonica* cultivars. V20B, the maintainer line of V20A [a widely used cytoplasmic male sterile (CMS) line in hybrid seed production] having the same nuclear genome as V20A, was used as a recurrent female parent. Ce64, a widely used restoration line in hybrid seed production in China, was used as a tester. The F_1 hybrid between V20A and Ce64 shows very strong heterosis and is one of the top-performing hybrid varieties in China.

Population development: *O. rufipogon* (IRGC 105491) was crossed as the male parent to V20A. The F_1 plants showed strong vegetative heterosis over V20B and were backcrossed twice with V20B. Fifty-two BC_1 plants were generated, which were field-grown in China during the summer of 1993. The best 10 BC_1 plants, selected for desirable plant type, maturity, and fertility, were backcrossed a second time to V20B to generate >3000 BC_2 plants. From these a subset of 300 BC_2 plants were selected based on the same criteria as the BC_1 population and crossed with Ce64 to generate 300 BC_2 testcross families.

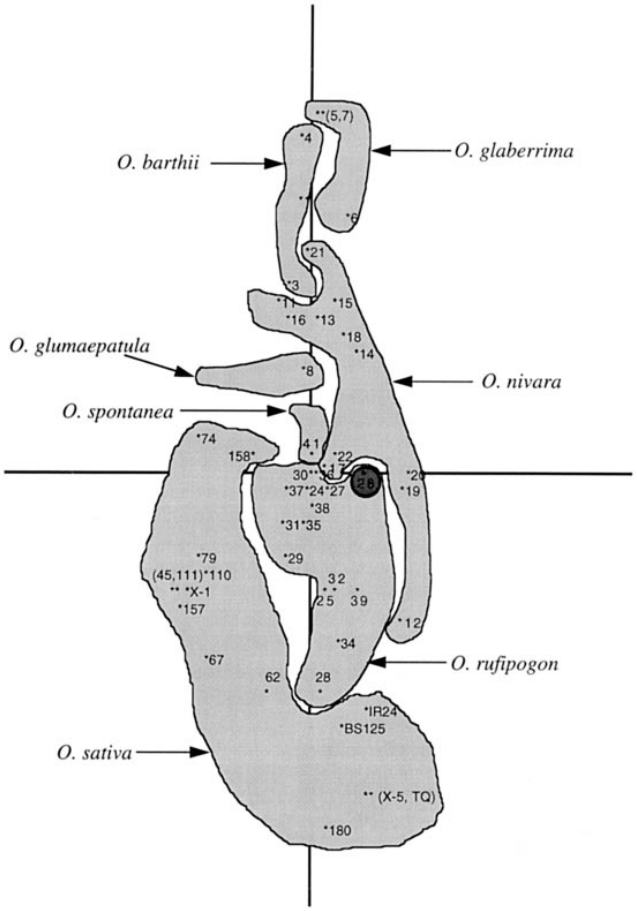


Figure 1.—Principal component analysis (PCA) of 34 wild accessions and 15 accessions from cultivated species (*O. sativa*). Gray dot, *O. rufipogon* accession (IRGC 105491). Cultivated species (*O. sativa*) accessions: 45, Shinano Mochi; 62, Aiyeh Lu; 67, Uz Ros 275; 74, Wu Tao Yeh Tao; 79, Yu Tao; 110, Chacarero; 111, Promisk 6; 157, Chang Pai; 158, Chipda; 180, Hunan Early Dwarf 7; TQ, Teqing.

Field trial and trait evaluation: The 300 BC_2 testcross families along with V20B, the *O. rufipogon* accession, and the commercial F_1 hybrid (V20A/Ce64) were grown in a field during the summer of 1994 at the China National Hybrid Rice Research and Development Center, where the majority of hybrid rice varieties are developed and released. Three-row plots with 11 plants per row were planted in a randomized complete block design with two replications. Grain yield per plot was evaluated based on grain harvest of all plants in each plot. The middle 10 plants in the central row of each plot were evaluated for 12 additional traits as follows. *Days to heading* was evaluated as the number of days from sowing in the field until 10% of the panicles on the 10 plants had headed. *Days to maturity* was evaluated as the number of days from sowing until an average of 80% of the grains on the 10 plants had reached a golden yellow. *Plant height* was calculated as the average number of centimeters from the ground to the tip of the tallest panicle (excluding the awn). *Panicle length* was measured as the average number of centimeters from the panicle neck to the panicle tip (excluding awn) based on an evaluation of all panicles from the 10 plants. *Panicles per plant* was the average number of panicles on the 10 plants (panicles having less than five

seeds were not counted). *Spikelets per panicle* was calculated by counting the total number of spikelets from the 10 plants divided by the number of panicles from all 10 plants. *Grains per panicle* was the average number of filled spikelets from the 10 plants divided by the number of panicles from all 10 plants. *Seed set rate* was calculated as a percentage: the number of filled spikelets per panicle divided by the number of spikelets per panicle. *Spikelets per plant* was calculated as the average number of spikelets on each of the 10 plants analyzed. *Grains per plant* was the average number of filled spikelets on each of the 10 plants analyzed. *1000-grain weight* was measured in grams as the average weight of three different samples of 1000 fully filled grains from each plot. *Yield per plant* was measured in grams and calculated as the average weight per plant of bulked grain harvested from the 10 plants per plot.

Marker genotype determination: DNA from the parents (*O. rufipogon*, V20B, and Ce64) was surveyed for polymorphism using two kinds of markers: RFLP (restriction fragment length polymorphism) and microsatellites or SSLP (simple sequence length polymorphism). Twenty seeds from each of the 300 BC_2 testcross families were bulked for DNA extraction. RFLP genotypes were determined as previously described in McCouch *et al.* (1988). The 102 RFLP probes were used for marker analysis in the BC_2 testcross population. These probes were a subset of those previously mapped in two different rice mapping populations (Causse *et al.* 1994; Kurata *et al.* 1994) that showed polymorphism between V20A and the *O. rufipogon* accession in genomic DNA digests with at least one of four restriction enzymes (*Eco*RI, *Eco*RV, *Dra*I, and *Hin*dIII).

In addition to the 102 RFLP markers, 20 SSLP markers, showing polymorphisms between V20A, Ce64, and the wild accession, were used to amplify microsatellites of DNA from the 300 BC_2 testcross families. The procedures used for the microsatellite assay were as described in Panaud *et al.* (1996).

Data analysis: Statistical analyses were performed using qGene (Nelson 1997) and Data Desk 4.0 (Data Description, Inc., 1992). Segregation ratios of individual markers were statistically determined at each marker locus for deviations from the expected Mendelian segregation ratio (3:1) by χ^2 tests. Genome composition was estimated in terms of the proportion of alleles transmitted from the cultivated or the wild germplasm. When an interval was bordered by two markers having the same genotype (originating from the same parent), the interval was treated as being composed entirely of DNA from the specified parental genome. When an interval was bordered by two consecutive markers having a different genotype, the interval was assumed to be composed of half of each parental genome. QTL mapping was conducted on BC_2 testcross data by regression of field performance on marker genotype using standard analysis of variance (ANOVA) procedures and $P < 0.01$ and assuming regular segregation of wild and cultivated alleles within testcross families.

To determine whether there were any significant interactions between trait-improving "wild QTL" alleles and loci elsewhere in the genome, the difference between the phenotypic mean of the genotypic class composed of heterozygotes, with one allele from *O. rufipogon* and one from Ce64 (the testcross parent), and the genotypic class corresponding to the original elite hybrid, with one allele from V20A and the other from Ce64, were statistically analyzed using Data Desk 4.0 (Data Description, Inc., 1992).

RESULTS

Polymorphism of markers: A total of 380 RFLP probes from the 12 chromosomes were used in the polymorphism survey. Of these, 106 (28%) were polymorphic between *O. rufipogon* and V20B with at least one of the four restriction enzymes. This is much less than the polymorphism previously reported for intersubspecific (indica/japonica or indica/javanica) crosses (McCouch *et al.* 1988; Wang *et al.* 1994; Xiao *et al.* 1996). Ninety-nine SSLP markers were also surveyed on V20B and *O. rufipogon*, and 59 (59.6%) detected polymorphism. The frequency of polymorphism detected by microsatellites was twice that detected by RFLPs, demonstrating the greater power of resolution of microsatellite markers.

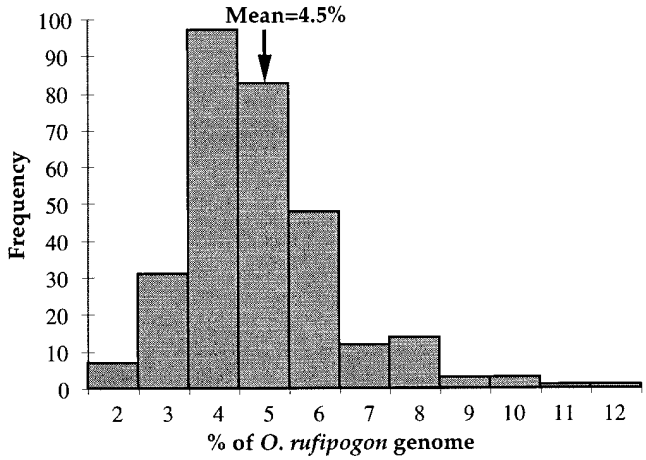


Figure 2.—Frequency distribution of the percentage of *O. rufipogon* genome in the 300 BC_2 testcross families.

Some regions of the genome were particularly monomorphic with all markers surveyed. For example, among the 22 RFLP markers and 6 SSLP markers on chromosome *10*, only 1 RFLP marker (RG561) and 1 microsatellite marker (RM222) showed polymorphism between *O. rufipogon* and V20B. These results suggest that some regions of the cultivated and wild genomes may be common by descent or that the *O. rufipogon* accession used in this study may be a derivative of a hybrid between wild *O. rufipogon* and cultivated *O. sativa*, resulting from the proximity of wild relatives to farmers' fields throughout Asia.

Genome composition of BC_2 testcross families: The genetic constitution of the BC_2 testcross families was visualized using qGene (Nelson 1997). Each of these families was expected to contain a Ce64 (tester) allele at every locus in combination with either a V20-derived allele or the *O. rufipogon* allele. Because the population was a selected BC_2 testcross, we expected to see the allelic constitution of the original V/64 hybrid variety, with the exception of small substitutions of the V20B genome, with chromosome segments from *O. rufipogon*. The frequency distribution of the proportion of *O. rufipogon* genome introgressed into the 300 families is shown in Figure 2. The percentage of *O. rufipogon* genome, which was calculated on the basis of the size (based on centimorgan distance) and the number of the wild introgressions contained in the 300 BC_2 testcross

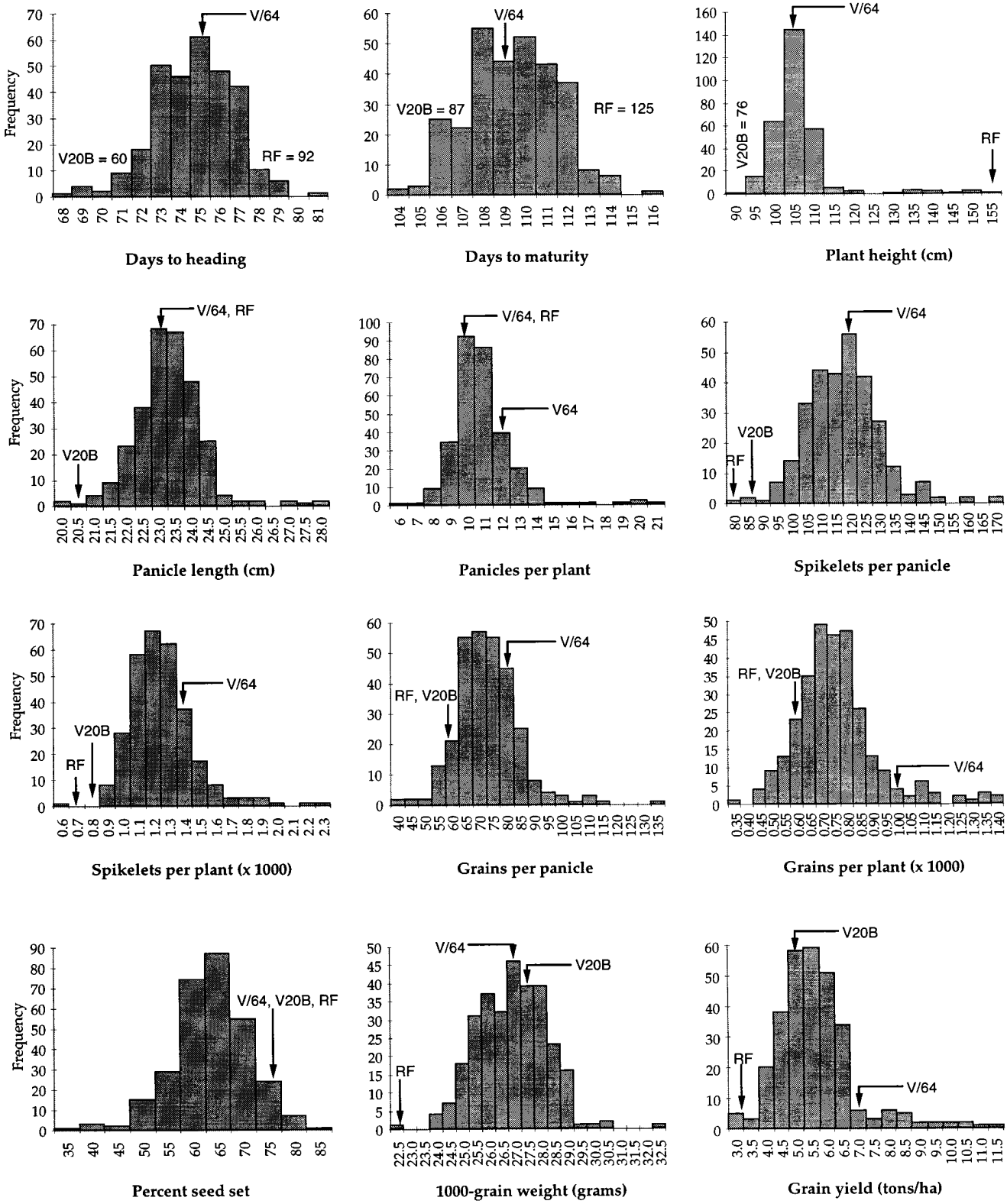


Figure 3.—Frequency distribution of phenotypes for each trait in the 300 BC_2 testcross families. Phenotypes of *O. rufipogon*, V20B and V/64 (V20A × Ce64) are shown by arrows.

families, ranged from 1.5 to 11.5% with a mean of 4.5%, which was not different from the expected for an unselected BC_2 testcross population (6.25% at $P \leq 0.05$).

Trait performance of BC_2 testcross families: The frequency distributions of phenotypes for each trait in the 300 testcross families are shown in Figure 3. All traits

showed approximately normal distributions. As shown in Figure 3, the *O. rufipogon* accession is phenotypically inferior for all of the traits examined here. However, the transgressive segregants (having phenotypic values less than or greater than both *O. rufipogon* and the V/64 hybrid) were observed for all traits studied. For example, 15 and 14% of the BC_2 testcross families outperformed V/64 with respect to grain yield and grains per plant, respectively. Fifty-six percent of the families had a higher 1000-grain weight, although excessively heavy grains are not considered a favorable trait in rice. The *O. rufipogon* accession was very late flowering and maturing; however, 43 and 35% of BC_2 testcross lines headed and matured earlier than V/64, respectively. These results suggest that genes introgressed from *O. rufipogon* into an elite genetic background can improve key agronomic traits of an elite rice variety, even though *O. rufipogon* itself is phenotypically inferior to the cultivated variety.

Trait correlations: The correlation between traits was estimated by regressing phenotypic values of one trait on those of another trait. The significant correlation coefficients among the 12 traits are presented in Table 1. For the majority of correlations, the degree and direction (positive or negative) of the correlation was consistent with that observed in a recombinant inbred population derived from an intersubspecific (indica/japonica) cross (Xiao *et al.* 1996). However, the negative correlations between 1000-grain weight and grains per plant or grains per panicle were greatly reduced in the present study.

QTL controlling heading and maturity dates: Seven QTL were significantly associated with days to heading (Figure 4 and Table 2). The phenotypic variance (R^2) explained by these individual QTL ranged from 3.00 to 15.21%. Compared to the performance of the V/64 hybrid, these individual QTL decreased days to heading by up to 3.48 or increased it by as much as 4.86. For QTL on chromosomes *6* and *12* (dth6.1 and dth12.1), the *O. rufipogon* alleles resulted in earlier heading. For the other QTL, the *O. rufipogon* alleles caused later heading.

Eight putative QTL were found for days to maturity. All the QTL except *dtm3.1* mapped to approximately the same locations as the QTL controlling days to heading. The *O. rufipogon* alleles at the two QTL on chromosomes *6* and *12* that were associated with earlier heading were also associated with earlier maturity. For the other six QTL, the *O. rufipogon* alleles delayed both heading and maturity.

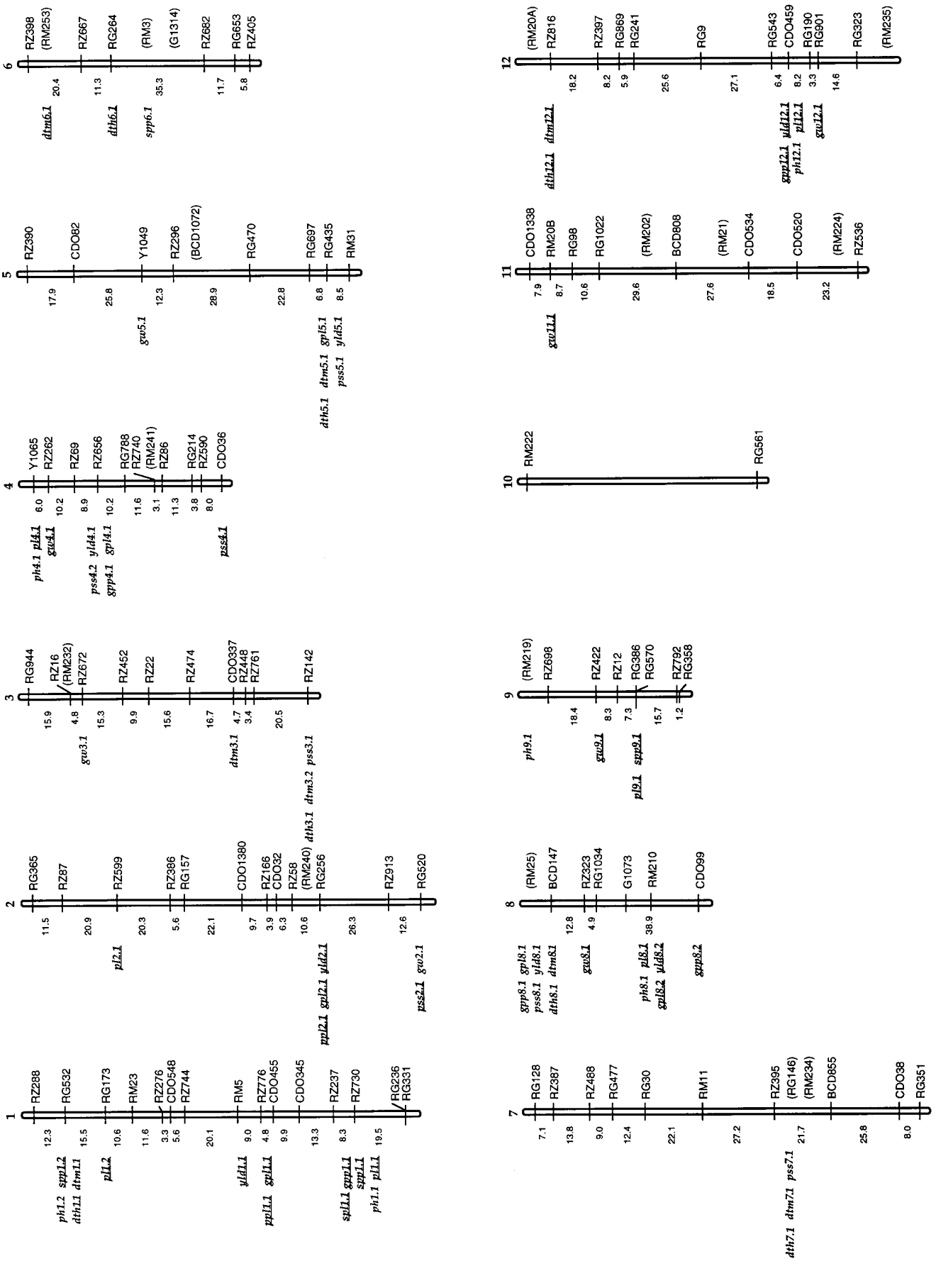
QTL influencing plant height: Six QTL on chromosomes *1*, *4*, *8*, *9*, and *12* were significantly associated with plant height (Figure 4 and Table 2). In all of these cases, the *O. rufipogon* alleles increased plant height. The phenotypic effect of each QTL ranged from 12.7 to 42.5 cm, which corresponds to a 12.57–41.96% increase in plant height over V/64. The phenotypic vari-

TABLE 1

Significant correlation coefficients (*r*) among traits in an *O. rufipogon*-derived BC_2 testcross population

	dth	dtm	ph	pl	ppl	spp	spl	gpp	gpl	pss	gw
Days to heading (dth)											
Days to maturity (dtm)	0.868										
Plant height (ph)	0.139	0.088									
Panicle length (pl)	0.234	0.168	0.683								
Panicles per plant (ppl)	−0.274	−0.195	−0.050	−0.085							
Spikelets per panicle (spp)	0.388	0.360	0.528	0.601	−0.254						
Spikelets per plant (spl)	−0.009	0.046	0.289	0.305	0.777	0.402					
Grains per panicle (gpp)	0.058	0.078	0.615	0.469	−0.088	0.623	0.329				
Grains per plant (gpl)	−0.153	−0.082	0.399	0.281	0.681	0.276	0.835	0.660			
Percent seed set (pss)	−0.257	−0.203	0.270	0.028	0.110	−0.084	0.062	0.704	0.581		
1000-grain weight (gw)	0.251	0.245	0.284	0.290	−0.112	0.099	−0.056	−0.115	−0.175	−0.244	
Grain yield (yld)	−0.100	−0.028	0.471	0.348	0.659	0.303	0.830	0.640	0.968	0.541	0.043

$r = 0.138$ at $P = 0.05$; $r = 0.181$ at $P = 0.01$.


1
RZ288 12.3
RG532 15.5
RG173 10.6
RM23 11.6
RZ276 3.3
CDO548 5.6
RZ744 20.1
RM5 9.0
RZ776 4.8
CDO455 9.9
CDO345 13.3
RZ237 8.3
RZ730 19.5
RG236
RG331
ph1.2 spp1.2 dth1.1 dtm1.1
pl1.2
yld1.1
gpl1.1
ppl1.1
spl1.1 gpp1.1 spp1.1 ph1.1 pl1.1
2
RG365 11.5
RZ287 20.9
RZ599 20.3
RZ386 5.6
RG157 22.1
CDO1380 9.7
RZ166 3.9
CDO32 6.3
RZ58
(RM240) 10.6
RG256
RZ913 26.3
RG620 12.6
pl2.1
ppl2.1 gpl2.1 yld2.1
pss2.1 gw2.1
3
RG944 15.9
RZ16 (RM232) 4.8
RZ672 15.3
RZ452 9.9
RZ22 15.6
RZ474 16.7
CDO337 4.7
RZ448 3.4
RZ761 20.5
RZ142
gw3.1
dtm3.1
dth3.1 dtm3.2 pss3.1
4
Y1065 6.0
RZ262 10.2
RZ69 8.9
RZ656 10.2
RG788 11.6
RZ740 (RM241) 3.1
RZ86 11.3
RG214 3.8
RZ590 8.0
CDO36
ph4.1 pl4.1 gw4.1
pss4.2 yld4.1 gpp4.1 gpl4.1
pss4.1
5
RZ390 17.9
CDO82 25.8
Y1049 12.3
RZ296 (BCD1072) 28.9
RG470 22.8
RG697 6.8
RG435 8.5
RM31
gw5.1
dth5.1 dtm5.1 gp15.1 pss5.1 yld5.1
6
RZ398 (RM253) 20.4
RZ667 11.3
RG264 35.3
(RM3) (G1314) 11.7
RZ682 5.8
RG653
RZ405
dtm6.1
dth6.1
spp6.1
7
RG128 7.1
RZ387 13.8
RZ488 9.0
RG477 12.4
RG30 22.1
RM11 27.2
RZ395 (RG146) (RM234) 21.7
BCD855 25.8
CDO38 8.0
RG351
dth7.1 dtm7.1 pss7.1
8
(RM25) 12.8
BCD147 4.9
RZ323
RG1034 7.3
G1073 38.9
RM210 15.7
CDO99 1.2
RG358
gpp8.1 gp18.1 pss8.1 yld8.1 dth8.1 dtm8.1
gw8.1
ph8.1 pl8.1 gp18.2 yld8.2
gpp8.2
9
(RM219) 18.4
RZ698 8.3
RZ422 7.3
RZ12
RG386
RG570
RZ792
ph9.1
gw9.1
pl9.1 spp9.1
10
RM222
RG561
11
CDO1338 7.9
RM20B 8.7
RG98 10.6
RG1022 29.6
(RM202) 27.6
BCD808 18.5
(RM21) 23.2
CDO534
CDO520
(RM224)
RZ536
gw11.1
12
(RM20A)
RZ816 18.2
RZ397 8.2
RG869 5.9
RG241 25.6
RG9 27.1
RG543 6.4
CDO459 8.2
RG190 3.3
RG901 14.6
RG323
(RM235)
dth12.1 dtm12.1
gpp12.1 yld12.1 ph12.1 pl12.1 gw12.1

ance explained by each QTL ranged from 8.49 to 44.77%.

QTL affecting yield traits: Yield and its related traits were also subjected to QTL analysis. The map locations of significant QTL are indicated in Figure 4, and the characteristics of these QTL are listed in Table 2.

Panicle length: Seven significant genomic regions were associated with this trait. For all the QTL, the *O. rufipogon* alleles increased panicle length. These individual QTL explained from 4.48 to 14.20% of observed phenotypic variance and increased panicle length from 1.05 to 3.22 cm, which corresponds to a 4.61 to 14.25% increase over V/64.

Panicles per plant: Two genomic regions were identified for the number of panicles per plant. The *O. rufipogon* alleles increased the average number of panicles per plant by 1.94 and 1.76, respectively, compared to the original V/64 hybrid.

Spikelets per panicle: Four QTL significantly influenced the number of spikelets per panicle, and, for three of these, the *O. rufipogon* alleles correlated with an increase of 28.36, 20.48, and 12.20 spikelets, which corresponds to a 23.89, 17.25, and 10.28% increase over V/64, respectively.

Spikelets per plant: Only one significant QTL was associated with spikelets per plant. The *O. rufipogon* allele at *spl1.1* on chromosome *1* increased the number of spikelets per plant by 341.32, or 25%, over V/64. However, this QTL explained only 4.23% of the phenotypic variance associated with this trait.

Grains per panicle: Five QTL were significantly associated with grains per panicle. For three of these QTL, *O. rufipogon* alleles caused an increase in grains per panicle with a phenotypic effect of 32.92, 16.06, and 22.16 grains, respectively, compared to V/64. The QTL with the largest effect explained 13.36% of the phenotypic variance.

Grains per plant: Six QTL significantly influenced grains per plant. The *O. rufipogon* alleles increased the number of grains per plant at *gpl1.1*, *gpl2.1*, and *gpl8.2*, corresponding to a 13.98, 12.95, and 22.27% increase over V/64, respectively.

Percentage seed set: Seven significant QTL were associated with spikelet fertility. The *O. rufipogon* alleles increased seed set at *pss2.1* and *pss4.1* and decreased seed set at the other four QTL. These individual QTL explained 3.00 to 14.61% of the total phenotypic variation, and had a positive phenotypic effect of increasing seed set by 6.24 and 7.58% and a negative effect of decreasing seed set by up to 15.54% compared to the V/64. Markers linked to these *O. rufipogon* alleles that reduce seed set can be used for negative selection in interspecific crosses with wild relatives.

1000-grain weight: Eight QTL had significant effects on grain weight. For five of these cases, the *O. rufipogon* alleles increased grain weight. The magnitude of the phenotypic effect of these individual QTL ranged from 1.30 to 2.22 g on a 1000-grain weight basis, corresponding to a 4.91 to 8.38% increase over V/64.

Grain yield: Seven QTL showed significant association with grain yield. The *O. rufipogon* alleles were associated with yield increases at four of these loci ranging from 0.98 to 1.22 tons/ha, which corresponds 14.61 to 18.26% of V/64. The two yield-enhancing QTL on chromosomes *1* and *2* (*yld1.1* and *yld 2.1*) were correlated with positive *O. rufipogon*-derived QTL for the yield components, panicles per plant and grains per plant, which mapped to approximately the same locations. The *O. rufipogon* alleles at the other QTL decreased grain yield from 1.37 to 1.76 tons/ha, or 20.57 to 26.32% of V/64.

Digenic interactions of yld1.1 and yld2.1 with other markers: Four chromosomal regions, RG331-CDO345 on chromosome *1*, RZ69 on chromosome *4*, RZ422-RG570 on chromosome *9*, and RG561 on chromosome *10*, showed significant interactions with *yld1.1* at $P < 0.01$. The wild alleles in these four regions, together with the wild allele at *yld1.1*, increased grain yield. Significant interactions with *yld2.1* ($P < 0.01$) were detected for RM240 on chromosome *2* and RZ422-RG570 on chromosome *9*. For the region RZ422-RG570, the wild allele interacted with the wild allele at *yld2.1* to enhance grain yield. However, at RM240, the cultivated allele interacted with the wild allele at *yld2.1* to increase grain yield. Because RM240 and *yld2.1* are linked on chromosome *2* at a distance of about 25 cM, these results suggest that a deleterious linkage was broken in the BC_1 and/or BC_2 plants that were selected for high yield in this population.

It is noteworthy that the region RZ422-RG570 on chromosome *9* showed significant interactions with both *yld1.1* and *yld2.1*, with the wild-wild combination outyielding the wild-cultivated combination in both cases.

DISCUSSION

Trait-improving QTL alleles of wild origin: For each QTL, the direction (negative or positive) of the *O. rufipogon* allele's effect on the target trait was determined.

Figure 4.—Linkage map of markers used for BC_2 testcross QTL analysis. The marker order and relative distances (in Kosambi mapping units) are based on the rice molecular genetic map (Causse *et al.* 1994). Markers in parentheses are mapped in populations other than the SL population in Causse *et al.* (1994). Left, map positions of significant ($P < 0.01$) QTL. Underlined QTL, *O. rufipogon* allele at that QTL is considered favorable for the trait. Abbreviations for traits: *dth*, days to heading; *dtm*, days to maturity; *ph*, plant height; *pl*, panicle length; *ppl*, panicles per plant; *spp*, spikelets per panicle; *spl*, spikelets per plant; *gpp*, grains per panicle; *gpl*, grains per plant; *pss*, percentage seed set; *gw*, 1000-grain weight; *yld*, yield.

TABLE 2

Characteristics of *O. rufipogon*-derived QTL-affecting traits in a BC_2 testcross population

Trait	QTL	Chromosome	Marker	PV (%)[a]	*P*[b]	VC class[c]	Mixed class[d]	RC class[c]	Allele effect[e]	V/64 (%)[f]
Days to heading	*dth1.1*	*1*	RG532	4.32	0.0003	74.63	76.23	77.83	3.20	4.27
	dth3.1	*3*	RG944	4.04	0.0005	74.59	75.78	76.97	2.38	3.17
	dth5.1	*5*	RG435	3.00	0.0028	74.61	75.67	76.73	2.12	2.83
	dth6.1	*6*	RG264	15.21	<0.0001	75.31	73.57	71.83	−3.48	−4.64
	dth7.1	*7*	RG146	6.60	<0.0001	74.64	77.07	79.50	4.86	6.48
	dth8.1	*8*	BCD147	6.91	<0.0001	74.53	76.00	77.47	2.94	3.92
	dth12.1	*12*	RZ816	4.01	0.0005	75.07	74.25	73.43	−1.64	−2.19
Days to maturity	*dtm1.1*	*1*	RG532	5.81	<0.0001	109.30	111.23	113.16	3.86	3.54
	dtm3.1	*3*	CDO337	3.24	0.0019	109.29	110.42	111.55	2.26	2.07
	dtm3.2	*3*	RG944	3.68	0.0009	109.29	110.46	111.63	2.34	2.15
	dtm5.1	*5*	RG435	2.86	0.0035	109.30	110.38	111.46	2.16	1.98
	dtm6.1	*6*	RM253	10.57	<0.0001	109.85	108.39	106.93	−2.92	−2.68
	dtm7.1	*7*	RG146	6.97	<0.0001	109.32	111.93	114.54	5.22	4.79
	dtm8.1	*8*	BCD147	6.09	<0.0001	109.21	110.66	112.11	2.90	2.66
	dtm12.1	*12*	RZ816	3.40	0.0014	109.75	108.96	108.17	−1.58	−1.45
Plant height (cm)	*ph1.1*	*1*	RZ730	44.77	<0.0001	101.81	123.04	144.27	42.46	41.96
	ph1.2	*1*	RG532	11.28	<0.0001	102.48	112.68	122.88	20.40	20.16
	ph4.1	*4*	Y1065	8.49	<0.0001	102.20	108.56	114.92	12.72	12.57
	ph8.1	*8*	RM210	14.47	<0.0001	102.46	114.50	126.54	24.08	23.79
	ph9.1	*9*	RM219	9.87	<0.0001	102.53	111.84	121.15	18.62	18.40
	ph12.1	*12*	CDO459	17.82	<0.0001	102.44	117.30	132.16	29.72	29.37
Panicle length (cm)	*pl1.1*	*1*	RZ730	14.20	<0.0001	22.98	24.59	26.20	3.22	14.25
	pl1.2	*1*	RG173	5.12	0.0001	23.03	24.07	25.11	2.08	9.20
	pl2.1	*2*	RZ599	5.15	0.0001	22.94	23.48	24.01	1.05	4.61
	pl4.1	*4*	Y1065	4.87	0.0002	22.98	23.62	24.26	1.28	5.66
	pl8.1	*8*	RM210	6.90	<0.0001	23.03	24.13	25.23	2.20	9.73
	pl9.1	*9*	RG386	5.33	0.0001	22.95	23.57	24.19	1.24	5.49
	pl12.1	*12*	CD0459	4.48	0.0002	23.04	24.04	25.04	2.00	8.85
Panicles per plant	*ppl1.1*	*1*	RZ776	3.81	0.0007	10.27	11.24	12.21	1.94	16.87
	ppl2.1	*2*	RG256	3.44	0.0014	10.27	11.15	12.03	1.76	15.30
Spikelets per panicle	*spp1.1*	*1*	RZ730	7.81	<0.0001	114.71	128.89	143.07	28.36	23.89
	spp1.2	*1*	RG532	4.45	0.0003	114.91	125.15	135.39	20.48	17.25
	spp6.1	*6*	RM3	3.76	0.0010	117.96	113.01	108.06	−9.90	−8.34
	spp9.1	*9*	RG386	3.50	0.0018	114.47	120.57	126.67	12.20	10.28
Spikelets per plant	*spl1.1*	*1*	RZ730	4.23	0.0004	1187.54	1358.20	1528.86	341.32	25.00
Grains per panicle	*gpp1.1*	*1*	RZ730	13.36	<0.0001	69.41	85.87	102.33	32.92	36.98
	gpp4.1	*4*	RZ656	5.18	0.0001	71.34	62.19	53.04	−18.30	−20.56
	gpp8.1	*8*	RM25	4.43	0.0003	71.62	65.17	58.72	−12.90	−14.49
	gpp8.2	*8*	CD099	5.18	0.0001	69.40	77.43	85.46	16.06	18.04
	gpp12.1	*12*	CDO459	4.92	0.0001	69.93	81.01	92.09	22.16	24.89

(*Continued*)

TABLE 2

(Continued)

Trait	QTL	Chromosome	Marker	PV (%)[a]	P^b	VC class[c]	Mixed class[d]	RC class[e]	Allele effect[e]	V/64 (%)[f]
Grains per plant	*gpl1.1*	*1*	RZ776	2.47	0.0067	722.38	793.96	865.54	143.16	13.98
	gpl2.1	*2*	RG256	2.31	0.0088	722.06	788.36	854.66	132.60	12.95
	gpl4.1	*4*	RZ656	4.23	0.0006	744.50	624.05	505.60	−240.90	−23.53
	gpl5.1	*5*	RG435	3.96	0.0006	746.21	647.25	548.29	−197.92	−19.33
	gpl8.1	*8*	RM25	6.57	<0.0001	751.31	638.24	525.17	−226.14	−22.09
	gpl8.2	*8*	RM210	3.08	0.0027	726.91	840.90	954.89	227.98	22.27
Percentage seed set	*pss2.1*	*2*	RG520	3.98	0.0007	60.60	64.39	68.18	7.58	10.11
	pss3.1	*3*	RG944	4.14	0.0004	61.81	57.45	53.09	−8.72	−11.63
	pss4.1	*4*	CDO36	3.00	0.0028	60.57	63.69	66.81	6.24	8.32
	pss4.2	*4*	RZ656	7.00	<0.0001	62.03	54.97	47.91	−14.12	−18.83
	pss5.1	*5*	RG435	3.92	0.0006	61.85	57.51	53.17	−8.68	−11.57
	pss7.1	*7*	RG146	3.99	0.0005	61.55	54.65	47.75	−13.80	−18.40
	pss8.1	*8*	RM25	14.61	<0.0001	62.47	54.70	46.93	−15.54	−20.72
1000-grain weight (g)	*gw2.1*	*2*	RG520	2.78	0.0046	26.78	26.19	25.60	−1.18	−4.45
	gw3.1	*3*	RZ672	10.22	<0.0001	27.04	26.18	25.32	−1.72	−6.49
	gw4.1	*4*	RZ262	5.17	0.0001	26.54	27.43	28.32	1.78	6.72
	gw5.1	*5*	Y1049	9.56	<0.0001	26.93	26.01	25.09	−1.84	−6.94
	gw8.1	*8*	RZ323	3.14	0.0022	26.58	27.23	27.88	1.30	4.91
	gw9.1	*9*	RZ422	4.81	0.0001	26.59	27.70	28.81	2.22	8.38
	gw11.1	*11*	RM20B	6.81	<0.0001	26.55	27.57	28.59	2.04	7.70
	gw12.1	*12*	RG190	4.49	0.0003	26.57	27.33	28.09	1.52	5.74
Grain yield (tons/ha)	*yld1.1*	*1*	RM5	2.81	0.0040	5.77	6.38	6.99	1.22	18.26
	yld2.1	*2*	RG256	2.56	0.0058	5.76	6.33	6.90	1.14	17.07
	yld4.1	*4*	RZ656	3.52	0.0017	5.94	5.06	4.19	−1.76	−26.32
	yld5.1	*5*	RG435	3.00	0.0028	5.96	5.27	4.58	−1.37	−20.57
	yld8.1	*8*	RM25	5.16	0.0001	5.99	5.20	4.40	−1.59	−23.80
	yld8.2	*8*	RM210	4.16	0.0005	5.81	6.34	6.87	1.07	15.94
	yld12.1	*12*	CDO459	2.88	0.0034	5.81	6.30	6.79	0.98	14.61

[a] Percentage of the phenotypic variance explained by the QTL.

[b] Probability that the marker genotype had no effect on the trait.

[c] Genotypic classes defined by single markers: VC is the heterozygote (V20A/Ce64) and RC is the interspecific heterozygote (*O. rufipogon*/Ce64) in the BC_2 testcross population.

[d] A BC_2 testcross family that had a heterozygous marker genotype, with one allele from the maternal parent (V20A or V20B) and the other allele from *O. rufipogon* in the BC_2 plant from which the BC_2 testcross family was produced. Theoretically, half of the individuals have a VC genotype and half have a VR genotype in such a BC_2 testcross family.

[e] The allele effect of a QTL is the difference of the phenotypic means between RC and VC classes.

[f] Percentage increase or decrease of the RC class in comparison to the elite V/64 hybrid.

A trait-improving QTL allele is defined as favorable for a given trait if, within the BC_2 testcross population where the QTL were detected, the mean performance of individuals having a heterozygous genotype with one allele from *O. rufipogon* and the other from Ce64 was better than that of individuals having the standard elite hybrid combination of one allele from V20 and the other from Ce64. Because the population was a testcross population, only those wild-QTL alleles having a dominant (or partially dominant) gene action or overdominant or additive action over the Ce64 counterpart could be detected. Recessive wild-QTL alleles with either a positive or negative effect would not have been detected. The BC_2 testcross is also inherently inefficient at detecting epistasis, because of the imbalance in the size of the group carrying a wild *vs.* a cultivated allele at any specific locus. The population structure may partly explain why few significant digenic interactions were observed in this study.

Of 68 QTL identified in this BC_2 testcross population, 35 (51%) had trait-improving alleles derived from the *O. rufipogon* accession. Trait-improving QTL alleles from the wild relative were detected for all traits (except for plant height, where any increase or decrease is considered deleterious to modern rice varieties). Yet, as evidenced by Ragot *et al.* (1995) and Tanksley *et al.* (1996), exotic or wild-QTL alleles that are favorable for some traits are often associated with deleterious effects on other traits. This phenomenon was observed in 16 of the 35 (46%) cases in the present study. For example, *O. rufipogon* alleles at *yld8.2* and *yld12.1* contributed positively to grain yield, but increased plant height, and *O. rufipogon* alleles in the genomic region around RZ730 on chromosome *1* increased panicle length, spikelets per panicle, spikelets per plant, and grains per panicle, but also increased plant height (Figure 4). Increases in plant height make rice plants more susceptible to lodging, leading indirectly to yield loss. Either pleiotropy, where a single gene affects multiple characters, or tight linkage of multiple QTL, where each linked gene affects a separate character, can be the genetic cause. In cases where linkage was the reason behind the association, marker-assisted disruption of the deleterious linkage would potentially make it possible to utilize these putatively positive wild-QTL alleles for genetic improvement. Fine mapping and further genetic dissection of the target regions containing these QTL would be needed to distinguish between pleiotropy and linkage of multiple genes.

Of the 35 trait-improving "wild-QTL" alleles 19 had no detectable negative effects on any measured trait (Figure 4). Although further evaluation is required to determine whether new, previously undetected, secondary effects would be observed in near isogenic lines containing single-QTL introgressions, it is possible that these 19 wild-QTL alleles would be immediately useful in improving traits of agronomic importance. For example, the wild alleles at *yld1.1* and *yld2.1* (Figure 4 and Table 2) increased grain yield by 1.2 and 1.1 tons/ha, respectively, corresponding to an 18 and 17% increase over V/64, with no deleterious effects on plant height or maturity. Transfer of these yield-enhancing alleles into elite (hybrid and inbred) varieties is underway to test the possibility that they may be of value to breeders seeking to substantially raise rice yield potential.

QTL alleles of wild origin increase grain yield without delaying maturity: Breeding for high yield potential in a target environment is an important objective in almost every rice breeding program. Yet, gains in yield potential often come at the cost of prolonging growth duration. Longer-season rice varieties are generally not preferred, as this feature limits the regions in which the variety can be produced and often eliminates the possibility of planting a second or third crop. More desirable are crop varieties that generate higher yields without delaying maturity. The present study found that the two QTL, *yld1.1* and *yld2.1*, increased grain yield with no detectable effect on maturity.

QTL alleles of wild origin shorten growth duration without decreasing grain yield: Although the *O. rufipogon* accession used in this study is late maturing, we identified wild alleles at two loci (*dtm6.1* on chromosome *6* and *dtm12.1* on chromosome *12*) that reduced growth duration in comparison with V/64 (an early maturing variety), and earliness was expressed without sacrificing grain yield. This finding represents a disruption of the commonly detected association between early maturity and reduced grain yield in cultivated rice. Obtaining varieties that combine earliness with high yield potential is an important breeding objective because, as mentioned above, early varieties can be cultivated in climatic regions where longer-growth-duration varieties cannot be grown, and they offer flexibility in planting dates and allow double or triple cropping.

Trait-improving QTL alleles of wild origin exhibit decreased pleiotropic effects on negatively correlated traits: Trait correlations may be attributable to either pleiotropic effects of single genes or to tight linkage of several genes that individually influence specific traits. In a previous study by Xiao *et al.* (1996), pleiotropy was suggested for three chromosomal regions that were simultaneously associated with 1000-grain weight and grains per plant or 1000-grain weight and grains per panicle. These yield components showed a highly negative correlation ($r = -0.703$ and $r = -0.608$, respectively), such that a plant produced either heavier grains (1000-grain weight) or a larger number of small grains per plant or per panicle. In that study, three significant QTL associated with 1000-grain weight mapped to the same positions as three QTL-affecting grains per plant and grains per panicle.

Negative correlations between 1000-grain weight and grains per plant or grains per panicle were also observed in the current study, but the degree of the correlations was greatly reduced ($r = -0.175$ and -0.115, respectively) (Table 1). The low correlation coefficients in the

present study were due to the fact that, of the eight *O. rufipogon*-derived QTL-controlling 1000-grain weight, (mapped on chromosomes *2, 3, 4, 5, 8, 9, 11*, and *12*) (Figure 3 and Table 2), five increased grain weight but had no detectable effect on grains per plant or grains per panicle. The *O. rufipogon* alleles at three of the six QTL (mapped on chromosomes *1, 2, 4, 5*, and *8*) (Figure 3 and Table 2) increased grains per plant and had no significant effect on 1000-grain weight. Three of the six significant *O. rufipogon*-derived QTL increased grains per panicle (Figure 3 and Table 2) and had no effect on grain weight.

When positions of QTL associated with the same trait were compared in different studies, it was observed that two of the three QTL for grain weight identified by Xiao *et al.* (1996) were in similar locations on chromosomes *4* and *5* as grain weight QTL reported in the present study. In the 1996 study, which was based on an *indica* × *japonica* (intersubspecific) cross, pleiotropic effects were observed in both cases, such that the QTL for increased grain weight was simultaneously associated with a decrease in grains per plant and grains per panicle. In contrast, QTL for grain weight in the same regions in the present study, *gw4.1* and *gw5.1*, were not associated with any significant effect on either grains per plant or grains per panicle.

The discovery of QTL from *O. rufipogon* that are free of linkage drag and the negative pleiotropic effects observed in studies using cultivated rice species suggests that an additional value of introducing new alleles from wild species may be the disruption of linkage relationships observed as negative correlations among traits important to agriculture. Our QTL mapping results suggest that these trait-improving QTL alleles, acting singly or epistatically with other loci introgressed from the wild relative, *O. rufipogon*, have reduced pleiotropic effects on several traits that are known to be highly negatively correlated in populations derived from cultivated varieties.

The results from the present study indicate that one of the closest wild relatives of cultivated rice, *O. rufipogon*, despite its overall inferior appearance, contains QTL alleles that are likely to substantially improve agronomically important traits, including yield. This discovery implies that the world's reservoir of wild and unadapted germplasm may hold the key to future productivity increases in rice and possibly other crop species.

We thank C. Nelson for providing the qGene software used extensively in this analysis and for comments on the manuscript, S. Hur for assistance in DNA extraction, and Y. Xu and D. Lawson for critical reviews of this article. The work was funded by The Rockefeller Foundation and the United States Department of Agriculture Plant Genome Research Program (National Research Initiative grant number 96-35300-3645).

LITERATURE CITED

Causse, M. A., T. M. Fulton, Y. G. Cho, S. N. Ahn, J. Chunwongse *et al.*, 1994 Saturated molecular map of the rice genome based on an interspecific backcross population. Genetics **138:** 1251–1274.

Chang, T. T., 1984 Conservation of rice genetic resources: luxury or necessity? Science **224:** 251–256.

Data Description, Inc., 1992 *Data Desk: Statistics Guide.* Data Description, Inc., Ithaca, New York.

Debouck, D. G., 1991 Genetic variation in crop species and their wild relatives: a viewpoint for their conservation, pp. 41–51 in *Genetic Diversity, and Crop Strategies for Roots and Tubers.* Bonn, Germany, Arbeitsgemeinschaft Tropische und Subtropische Agrarforschung e.V. and International Board for Plant Genetic Resources.

Frey, K. J., T. S. Cox, D. M. Rodgers and P. Bramel-Cox, 1983 Increasing cereal yields with genes from wild and weedy species, pp. 51–68 in *Proceedings of the 15th International Genetics Congress, New Delhi, India, December 12–21, 1983.* Oxford and IBH Publishing Co., New Delhi, India.

Khush, G. S., K. C. Ling, R. C. Aquino and V. M. Aguiero, 1977 Breeding for resistance to grassy stunt in rice. Proceedings of the 3rd International Congress of the Society for the Advancement of Breeding Researchers in Asia and Oceania (SABRAO). Plant Breeding Paper **1**(4b)**:** 3–9.

Kurata, N., Y. Nagamura, K. Yamamoto, Y. Harushima, N. Sue *et al.*, 1994 A 300 kilobase interval genetic map of rice including 883 expressed sequences. Nature Genetics **8:** 365–372.

Ladizinsky, G., 1985 Founder effect in crop-plant evolution. Econ. Bot. **39:** 191–199.

Li, Z., and Y. Zhu, 1988 Rice male sterile cytoplasm and fertility restoration, pp. 85–102 in *Hybrid Rice.* International Rice Research Institute. Manila, Philippines.

McCouch, S. R., and R. W. Doerge, 1995 QTL mapping in rice. Trends Genet. **11:** 482–487.

McCouch, S. R., G. Kochert, Z. H. Yu, Z. Y. Wang, G. S. Khush *et al.*, 1988 Molecular mapping of rice chromosomes. Theor. Appl. Genet. **76:** 815–829.

Nelson, J. C., 1997 QGene manual. Electronic address: greengenes.cit.cornell.edu port 70; directory "Software for genetics / qGene."

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

Plucknett, D. L., N. J. H. Smith, J. T. Williams and N. M. Anishetty, 1987 A case study in rice germplasm: IR36, pp. 171–185 in *Gene Banks and The World's Food*, edited by D. L. Plucknett, N. J. H. Smith, J. T. Williams and N. M. Anishetty. Princeton University Press, Princeton, NJ.

Ragot, M., P. H. Sisco, D. A. Hoisington and C. W. Stuber, 1995 Molecular-marker-mediated characterization of favourable exotic alleles at quantitative trait loci in maize. Crop Science **35:** 1306–1315.

Simmonds, N. W., 1976 *Evolution of Crop Plants.* Longman, London, New York.

Stuber, C. W., 1995 Mapping and manipulating quantitative traits in maize. Trends Genet. **11:** 477–481.

Tanksley, S. D., 1993 Mapping polygenes. Annu. Rev. Genet. **27:** 205–233.

Tanksley, S. D., and J. C. Nelson, 1996 Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theor. Appl. Genet. **92:** 191–203.

Tanksley, S. D., S. Grandillo, T. M. Fulton, D. Zamir, Y. Eshed *et al.*, 1996 Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium.* Theor. Appl. Genet. **92:** 213–224.

Wang, Z. W., G. Second and S. D. Tanksley, 1992 Polymorphism and phylogenetic relationships among species in the genus *Oryza* as determined by analysis of nuclear RFLPs. Theor. Appl. Genet. **83:** 565–581.

Wang, G., D. J. Mackill, J. M. Bonman, S. R. McCouch, M. C. Champoux *et al.*, 1994 RFLP mapping of genes conferring complete and partial resistance to blast in a durably resistant rice cultivar. Genetics **136:** 1421–1434.

Xiao, J., J. Li, L. Yuan and S. D. Tanksley, 1996 Identification of QTLs affecting traits of agronomic importance in a recombinant inbred population derived from a subspecific rice cross. Theor. Appl. Genet. **92:** 230–244.

Communicating editor: J. A. Birchler