REGULAR ARTICLE

Development and evaluation of advanced backcross families of rice for agronomically important traits

Kalluvettankuzhi K. Sabu[1]*, Mohammed Z. Abdullah[2], Li Sze Lim[3], Ratnam Wickneswari[3]

[1] Tropical Botanic Garden and Research Institute, Thiruvananthapuram 695562, India.
[2] Gene Bank, MARDI Seberang Perai, PO Box 203, 13200 Kepala Batas, Pulau Pinang, Malaysia.
[3] School of Environmental and Natural Resource Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Malaysia.
* Corresponding author: Kalluvettankuzhi K. Sabu, E-mail: krishnan.sabu@gmail.com

ABSTRACT

Wild relatives of crop species may contain trait-enhancing genes. Our studies showed that grain yield in cultivated rice can be increased significantly using the genes found in common wild rice. Advanced backcross progenies (*Oryza rufipogon* IRGC105491 x *O. sativa* cv. MR219) exhibited higher values for most of the yield and yield-related traits analyzed in comparison to the cultivated parent, MR219. About 10% of the BC_2F_2 families exhibited 33.5% higher grain yield per plant in comparison to MR219. Even though there was an increase in empty grains per panicle, number of tillers per plant also increased considerably (41.5%). This resulted in higher grain yield per plant, which ranged from 52.4 to 77.5g for the selected families in contrast to 42.4g for MR219. Two families, viz., S163 and S165, were identified through GGE biplot analysis as most promising genotypes for cultivar development. The backcross families offer an efficient source of genetic diversity for future breeding programmes.

Key Words: *Oryza rufipogon; Oryza sativa; wild rice; grain yield; advanced backcross; trait enhancing genes*.

INTRODUCTION

Increasing world rice (*Oryza sativa* L.) production is vital to meet dietary demands of the growing global population. Rice cultivars with higher yield potential must be developed to reduce the gap between yield potential and average farm yields (Peng et al., 1999). Irrigated land in Asia is most suitable for rice production, as rainfed regions are constrained by

drought, flooding and poor soil quality (Cassman, 1999). In Malaysia, rice is the most important food crop and is cultivated on about 0.6 million ha of arable land. Conventional rice breeding carried out during the last 37 years has resulted in the release of 34 modern varieties with an average yield of about 5 tons ha-1. In spite of these advances, rice production is inadequate and has been on a plateau in recent years in the country (IRRI, 2001). The introduction of genes from exotic sources is one approach towards further improvement of yield and yield-related traits, which has been successful in many crop species (Percival, 1987; Wagoner, 1990; Sacks et al., 2003b; McCarty et al., 2004a, 2004b). However, genetic improvement from wild germplasm can result in the transfer of undesirable traits, as reported in some QTL mapping studies in rice (Xiao et al., 1998; Moncada et al., 2001; Septiningsih et al., 2003a). Hence it is imperative that breeders have a better knowledge of heritability, and genetic and environmental effects of yield and yield-related traits.

Oryza rufipogon Griff. (common wild rice) is considered to be the progenitor of Asian cultivated rice, *O. sativa* L. (Second, 1982; Cheng et al., 2003). In general, *O. sativa* is predominantly self-pollinated whereas *O. rufipogon* shows partial outcrossing (Oka and Morishima, 1997). Conventional and molecular breeding have demonstrated that introgression of *O. rufipogon* genes could add new traits or improve the existing agronomic traits of *O. sativa*, such as perenniality (Sacks et al., 2003a), yield performance in upland (Moncada et al., 2001) or lowland conditions (Martinez et al., 1998; Xiao et al., 1998; Septiningsih et al., 2003a), grain quality (Septiningsih et al., 2003b), and floral morphology (Cai and Morishima, 2002; Uga et al., 2003).

Oryza rufipogon accession no. IRGC105491 is an annual form that was originally collected by Japanese researchers from Malaysia (D.S. Brar, 2005, pers. commun.). It grows to a height of about 143 cm and produces on average 39 tillers in Malaysia. *Oryza sativa* cv. MR219 (hereafter referred to as MR219) is an elite cultivar bred from about 15 parents by the Malaysian Agricultural Research and Development Institute (MARDI). It can grow to be 110 cm in height and have about 20 tillers. The objectives of this research were (i) to develop advanced backcross families, (ii) evaluate their field performance under standard farming conditions to select possible superior families for cultivar development, and (iii) to estimate heritability of yield and yield-related traits, a measure of possible genetic advancement under selection (Allard, 1966). The same backcross families are also currently being genotyped for QTL mapping in another study. Due to difficulties in obtaining large number of pureline (selfed) seeds, the current experiment was conducted for only one season in two replications.

MATERIALS AND METHODS

POPULATION DEVELOPMENT

Crosses were made during 2002-2004 and subsequent evaluations were conducted from April to July 2004 at MARDI field experimental station in Seberang Perai, Penang, Malaysia (5°32´N, 100°27´E). *O. rufipogon* was used as the female parent and crossed to MR219. The F_1 progenies were backcrossed twice to MR219 - the recurrent female parent. Twenty-seven BC_2F_1 lines were self-pollinated and seeds were harvested from 10 plants per line (except for one seed line where only 6 plants produced sufficient seeds) which resulted in a final count of 266 BC_2F_2 families.

EXPERIMENTAL DESIGN AND MEASUREMENT OF AGRONOMIC TRAITS

The field performance of the BC_2F_2 families was conducted at MARDI and the experimental site was a homogenous natural rice field. The 266 BC_2F_2 families and two parental lines were regarded as treatments and assigned to plots in a randomized complete block design with two replications. Thus a single replication consisted of 268 plots. The field was ploughed before planting. Pre-germinated seeds were sown in large troughs (2.5 m long, 1 m wide and 45 cm deep) containing wet soil procured from a nearby paddy field. Healthy seedlings were transplanted to the rice field after 24 days of growth. Fifty-four plants were

planted in each plot (4.5 long × 1.2 m wide). Each plot consisted of 3 rows and there were 18 plants per row. The row-to-row spacing was 30 cm and plants were spaced 25 cm apart within rows. The space between adjacent plots was 60 cm. The rice plants were fertilised at recommended rates (170:80:150 N:P:K, kg ha^{-1}). To avoid border effect, 10 plants in the centre of the 2nd row were used for recording various yield traits.

Plant protection agents were used as required, following the standard practices of the MARDI experimental station. Furadan (insecticide) was applied at 17-25 kg ha^{-1} every 40 days after transplanting. Basmi 1% Jumbo (weedicide) was applied at 2.1 kg ha^{-1} and 2.5 kg ha^{-1} after 7 and 14 days of transplanting, respectively. Rumputox (weedicide) was used at 14 kg ha^{-1} before panicle initiation. Bayluscide 70wp (against snails) was applied as and when required at 315 g ha^{-1}.

Important vegetative traits were recorded from 10 plants in each plot at near-maturity. Culm length was measured from the soil level to the base of the panicle. This was done by holding all the tillers together and then selecting the longest tiller and measuring against a meter-stick. Panicle length was measured (from the same tiller used to record culm length) from the base to the tip of the panicle (excluding awn). Number of tillers per plant was recorded as the total number of grain-bearing and non grain-bearing tillers, whereas panicles per plant were recorded as the total number of panicles in the grain-bearing tillers on a plant. Mean values of the vegetative traits in a family were computed from the 10 plants. Thirty panicles from each plot (3 panicles from each of the 10 plants used) were used to record filled and empty grains per panicle. The 1000-grain weight was estimated in grams as the mean weight of 3 different samples of 1000 fully filled grains. The moisture content (MC) was determined by drying the grains at 130°C for 2 h in a hot-air oven, followed by storage in a desiccator for 20 min before recording the weight. The grain weight was then adjusted to 14% MC by multiplying the dry grain weight with an adjustment coefficient (Gomez and Gomez, 1984). Yield per plant, expressed in g, was calculated from the weight of bulked grain harvested from 10 plants in each plot.

The BC$_2$F$_2$ families and the parents were subjected to field evaluation in both replications for morphological characteristics, such as awn and grain type. Samples with long awns were classified as *O. rufipogon* and those without awns were classified as MR219 types. If small awns were found, they were regarded as intermediate type. For grain colour, there were two classes corresponding to *O. rufipogon* (black grains) and MR219 (golden yellow grains). Families possessing different types of awns or grains were designated as mixed type.

DATA ANALYSIS

Analysis of variance was performed on yield and yield-related traits using the General Linear Model (GLM) procedure assuming a random effects model and type III sum of squares (SAS Institute Inc., 2006; SAS version 9.1 was used for the data analysis using individual plant data). Various descriptive statistics were determined using PROC MEANS for each type (parental or progeny) separately. Duncan's new multiple range test ($\alpha = 0.05$) was used for mean separation. Normal distributions and histograms of various traits were determined using PROC CAPABILITY. Regression analysis was conducted using PROC REG. Ranking of families was based on grain yield per plant, culm length, tillers per panicle, filled grains per panicle and 1000-grain weight via PROC RANK. The qualitative (awn and grain types) data were subjected to χ^2 analyses using the CHISQ option in PROC FREQ.

Variance components for the statistical design were calculated using the TYPE=REML option in PROC MIXED of SAS. In this TYPE, the PROC MIXED uses a ridge-stabilized Newton-Raphson algorithm to optimise a restricted or residual maximum likelihood (REML) function. Broad-sense heritability was estimated as the ratio of the estimate of the genetic variance to the phenotypic variance of a family mean. Phenotypic correlations were calculated using PROC CORR. The GGE biplot analysis of the trait data was performed using GGEbiplot (Yan and Kang, 2003).

RESULTS

ASSESSMENT OF YIELD AND YIELD RELATED TRAITS

Differences between replications and among families for yield and yield-related traits were highly significant (Table 1). Histograms showing frequency distribution of various agronomic traits in the BC_2F_2 families are shown in Figure 1.

Table 1. Analysis of variance for yield and yield related traits for BC_2F_2 families.

Trait	Source	df	Mean Square
Culm length, cm	Replication	1	35756.29**
	Family	265	4407.00**
	Plants (Family)[a]	2394	92.76
	Error	2650	101.29
Panicle length, cm	Replication	1	671.66**
	Family	265	23.67**
	Plants (Family)	2394	3.03
	Error	2650	3.34
Tillers per plant	Replication	1	3211.36**
	Family	265	179.70**
	Plants (Family)	2394	36.89
	Error	2656	36.98
Panicles per plant	Replication	1	2959.95**
	Family	265	180.05**
	Plants (Family)	2394	37.11
	Error	2656	37.12
Filled grains per panicle	Replication	1	135918.65**
	Family	265	10627.53**
	Error	15690	788.96
Empty grains per panicle	Replication	1	279487.98**
	Family	265	5868.86**
	Error	15691	359.86
Grain yield per plant, g	Replication	1	1055.64**
	Family	265	78.97**
	Error	265	42.25
1000-grain weight, g	Replication	1	63.74**
	Family	265	6.73**
	Plants (Family)	509	0.85
	Error	264	1.55

** Significant at the 0.01 probability level.

[a] Plants (Family) analysis was not performed for filled grains per panicle and empty grains per panicle as these traits were analysed by collecting 30 panicles from a family. Plants (Family) analysis was also not performed for grain yield per plant as the trait was calculated from the weight of bulked grain harvested from 10 plants in each family.

Culm length did not show a continuous distribution. About 82% of the BC_2F_2 families were classified as the MR219 types, with a mean culm length of 87 cm, whereas the remaining families, designated as *Oryza rufipogon* types, had a mean culm length of 124 cm. All other traits demonstrated almost normal distribution (skewness ranged from –0.11 to 0.74) where the BC_2F_2 families did not fall into the two distinguishable parental types.
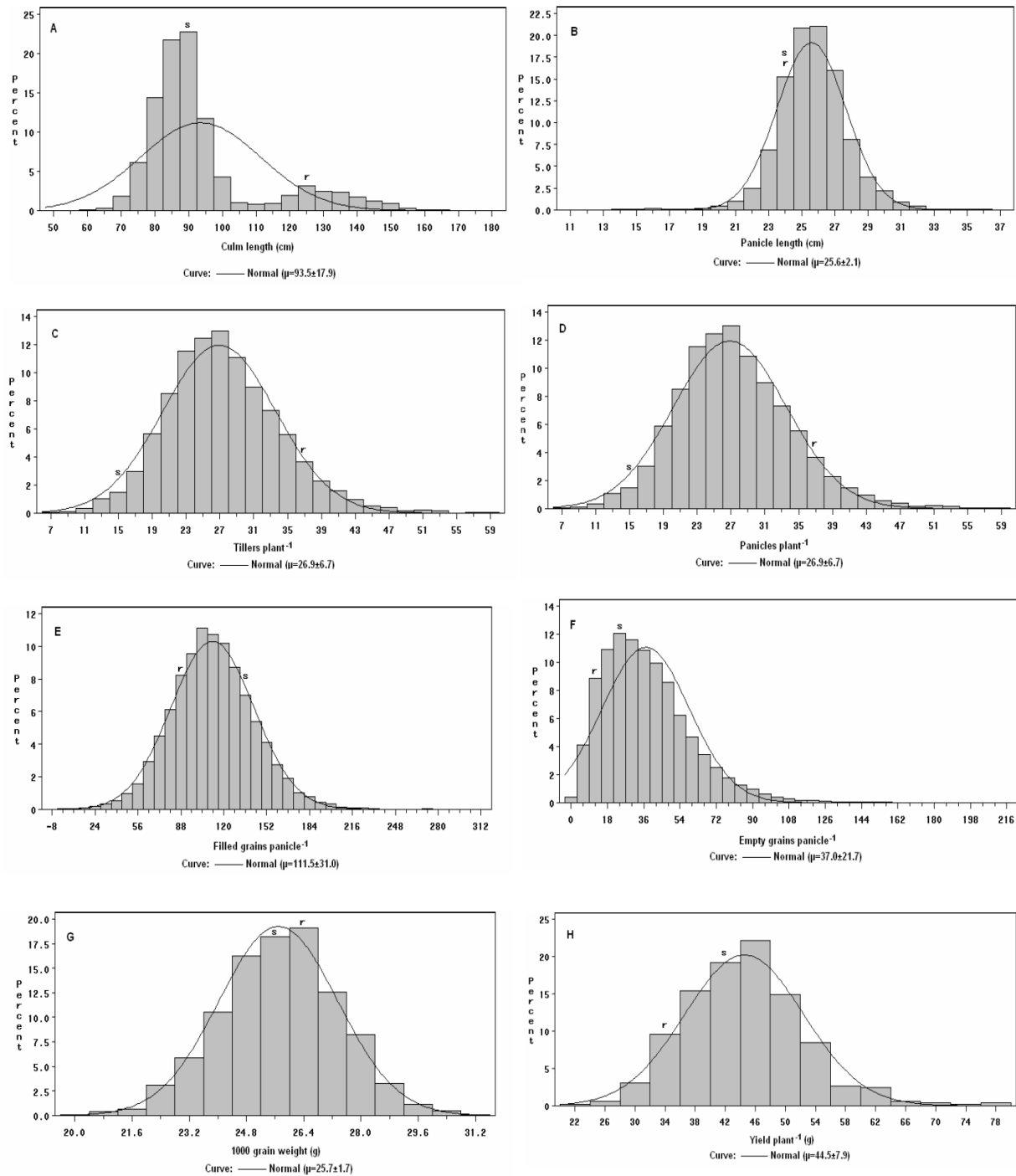
Figure 1. Histograms (A – H) showing frequency distribution of agronomic traits in BC_2F_2 families (*Oryza rufipogon* acc. no. IRGC105491/*O. sativa* cv. MR219).

REGRESSION OF GRAIN YIELD ON OTHER TRAITS

Regression of grain yield on culm length, panicle length, tillers per plant, panicles per plant, filled grains per panicle, empty grains per panicle and 1000-grain weight (Case I) resulted in a good model fit (F = 11.95, P < 0.001). In a subsequent analysis, culm length, tillers per plant, filled grains per panicle and 1000-grain weight (Case II) were used as independent variables. This resulted in an F value of 20.83 (P < 0.001).

Table 2. Mean values of agronomic traits for parents and BC_2F_2 families.

TYPE	CL	PL	TPL	PPL	FGP	EGP	GW	PY
O. rufipogon IRGC105491	118.7*a[a]	24.3b	39.7a	39.5a	92.4c	10.2c	26.1a	36.3a
O. sativa cv. MR219	86.2**c	24.5b	20.7c	20.7c	137.3a	23.8b	25.0b	42.4a
BC_2F_2 families	93.5**b	25.6**a	26.9**b	26.9**b	111.5**b	34.0**a	25.7**ab	44.5**a
SD	17.9	2.1	6.7	6.7	31.0	21.7	1.7	7.9
Min	56.0	11.6	6.0	6.0	0.0	0.0	20.2	23.3
Max	180.0	37.0	59.0	59.0	313.0	217.0	31.2	77.5

Abbreviations used: Parental or progeny types (TYPE), culm length, cm (CL), panicle length, cm (PL), tillers per plant (TPL), panicles per plant (PPL), filled grains per panicle (FGP), empty grains per panicle (EGP), 1000-grain weight, g (GW) and grain yield per plant (PY).
[a] Mean values followed by the same letter are not significantly different at 0.05 probability level of Duncan's Multiple Range Test.
Means followed by * and ** indicate significant differences within each type at 0.05 and 0.01 probability level respectively. Significance within parental or progeny types are based on TYPE III Sum of Squires (PROC GLM).

SELECTION OF SUPERIOR BC_2F_2 FAMILIES

Grain yield per plant, culm length, tillers per plant, filled grains per panicle and 1000-grain weight were used to rank BC_2F_2 families. After ranking, 26 superior families (10.2%) were selected for detailed analysis (Table 3). The short-listed families showed grain yield of 52.4 to 77.5 g and had 25.1 to 35.7 tillers per plant.

BC_2F_2 FAMILY ANALYSIS

Means, coefficient of variation, mean maximum and minimum trait values, and broad-sense heritability for BC_2F_2 population are presented in Table 4. Results from the ANOVA showed highly significant ($P < 0.001$) differences for replications and families for culm length, panicle length, tillers per plant, panicles per plant, filled grains per panicle, empty grains per panicle and 1000-grain weight. Empty grains per panicle varied greatly (CV=26.8) in the 266 families. Variation in panicle length and 1000-grain weight was the lowest (CV=4.3 and 5.2 respectively).

Estimates of broad-sense heritability were the highest for culm length (0.78) followed by 1000-grain weight (0.55). For rest of the traits, heritability ranged from 0.19 to 0.33.

CORRELATION BETWEEN VARIOUS YIELD AND YIELD-RELATED TRAITS

Correlations among traits (Table 5) revealed highest correlation ($r = 0.99$, $P< 0.001$) for tillers per plant vs. panicles per plant. Among other traits, grain yield per plant was correlated with tillers per plant ($r = 0.27$), panicles per plant ($r = 0.27$) and filled grains per panicle ($r = 0.24$).

GGE BIPLOT ANALYSIS

Biplot analysis was carried out based on data presented in Table 3 (variables such as culm length and empty grains per panicle were excluded as these were not relevant to the analysis). Biplot is essentially a scatter plot that graphically displays both the entries (e.g., families) and the testers (e.g., traits) of a two-way data. Figure 2 showed which genotypes were the best in what traits. Only two traits stood out: filled grains per panicle and yield per plant. Family S132 was the highest performing for filled grains per panicle and family S163 was the highest for yield per plant. The graph in Figure 3 showed the correlations among traits. Filled grains per panicle and yield per plant lacked obtuse angle, and were thus positively correlated. The graph in Figure 4 shows comparisons of all genotypes with the "ideal" genotype. Families S163 and S165 were the closest to the ideal genotype, followed by S162,

S132, and S257; other families were below average (when all traits were considered). The GGE biplots explained 73.4% (PC1) and 22.1% (PC2) variability, i.e., 95.5% of the total variation.

Table 3. Mean values of agronomic traits of 26 outperforming BC_2F_2 families.

Family #	Type[a]	CL	PL	TPL	PPL	FGP	EGP	GW	PY
S163	A	118.4	27.5	31.9	31.9	112.3	28.4	27.6	77.5
S165	D	124.5	26.9	31.8	31.8	138.6	17.1	23.4	67.6
S164	A	95.3	25.3	32.3	32.3	99.7	34.2	25.2	60.0
S132	E	87.9	26.1	27.0	27.0	143.9	39.3	26.0	59.9
S257	D	99.0	25.6	33.1	33.1	118.8	47.3	27.8	59.4
S202	A	131.1	27.4	25.8	25.6	109.1	48.8	26.2	59.2
S162	A	88.1	24.0	31.5	31.5	121.7	37.9	28.9	57.2
S211	C	89.2	26.1	31.2	31.2	107.3	38.6	27.3	56.3
S127	C	80.1	24.0	29.0	29.0	122.6	43.6	25.4	56.2
S37	B	86.6	24.4	27.9	27.9	107.0	56.0	27.1	56.0
S62	A	75.2	25.9	29.5	29.5	121.0	46.9	23.6	55.7
S207	C	132.3	26.8	29.4	29.3	103.5	23.1	25.4	55.7
S201	E	124.5	28.6	27.0	27.0	98.8	56.4	23.9	55.1
S161	A	119.2	26.6	25.8	25.8	103.3	35.9	27.1	55.0
S39	C	82.8	26.0	27.9	27.9	110.0	38.3	27.4	54.3
S76	E	91.1	25.4	28.3	28.3	117.2	32.3	24.7	54.1
S227	A	122.6	27.1	30.7	30.6	86.8	63.8	24.2	54.1
S98	E	91.2	23.8	32.0	31.9	107.7	22.2	28.1	53.6
S100	A	87.6	24.6	25.1	25.0	131.8	32.3	23.8	53.6
S70	E	83.4	26.4	29.6	29.6	100.8	59.0	27.9	53.6
S97	A	82.4	24.4	30.9	30.9	117.2	26.7	24.9	53.3
S203	A	123.7	26.3	27.9	27.9	100.9	34.7	26.3	52.8
S125	A	89.3	24.6	28.2	28.2	125.0	35.3	25.7	52.7
S154	C	88.9	25.7	26.7	26.7	125.7	36.0	26.6	52.6
S210	A	121.7	27.1	25.8	25.8	117.0	33.6	26.0	52.5
S109	C	78.9	27.7	35.7	35.7	115.6	40.8	25.2	52.4
Mean									
Selected BC_2F_2		99.8a[b]	25.9a	29.3a	29.3a	114.0a	38.8a	26.0a	56.6a
MR219		86.2a	24.5a	20.7b	20.7b	137.3ab	23.8a	25.0a	42.4b
O. rufipogon		118.7a	24.3a	39.7c	39.5c	92.4b	10.2a	26.1a	36.3b

Abbreviations used: CL—culm length, cm; PL—panicle length, cm; TPL—tillers per plant; PPL—panicles per plant; FGP—filled grains per panicle; EGP—empty grains per panicle; GW—1000-grain weight, g; PY—grain yield per plant.
[a] A—both awn and grain MR219 type; B—awn *O. rufipogon* type but grain MR219 type; C—awn intermediate type but grain MR219 type; D—awn and grain mix types; E—awn mix type but grain MR219 type.
[b] Mean values followed by the same letter are not significantly different at 0.05 probability level of Duncan's Multiple Range Test.

Table 4. Grand means (*X*), coefficient of variation (*CV*) and mean maximum and minimum trait values, and broad-sense heritability (h_b^2) for the agronomic traits in BC$_2$F$_2$ families.

Traits	X	CV	XH	XL	h^2_b
CL	93.5	15.9	136.4	74.2	0.78
PL	25.6	4.3	29.6	22.3	0.33
TPL	26.9	11.1	36.1	18.8	0.23
PPL	26.9	11.1	36.1	18.8	0.23
FGP	111.5	11.9	149.8	72.9	0.19
EGP	37.0	26.8	65.8	12.3	0.20
GW	25.7	5.2	29.9	22.2	0.55

Abbreviations used: CL—culm length, cm; PL—panicle length, cm; TPL—tillers per plant; PPL—panicles per plant; FGP—filled grains per panicle; EGP—empty grains per panicle; GW—1000-grain weight, g.

Table 5. Pearson correlation coefficients among various traits in BC$_2$F$_2$ population (*O. rufipogon*/MR219).
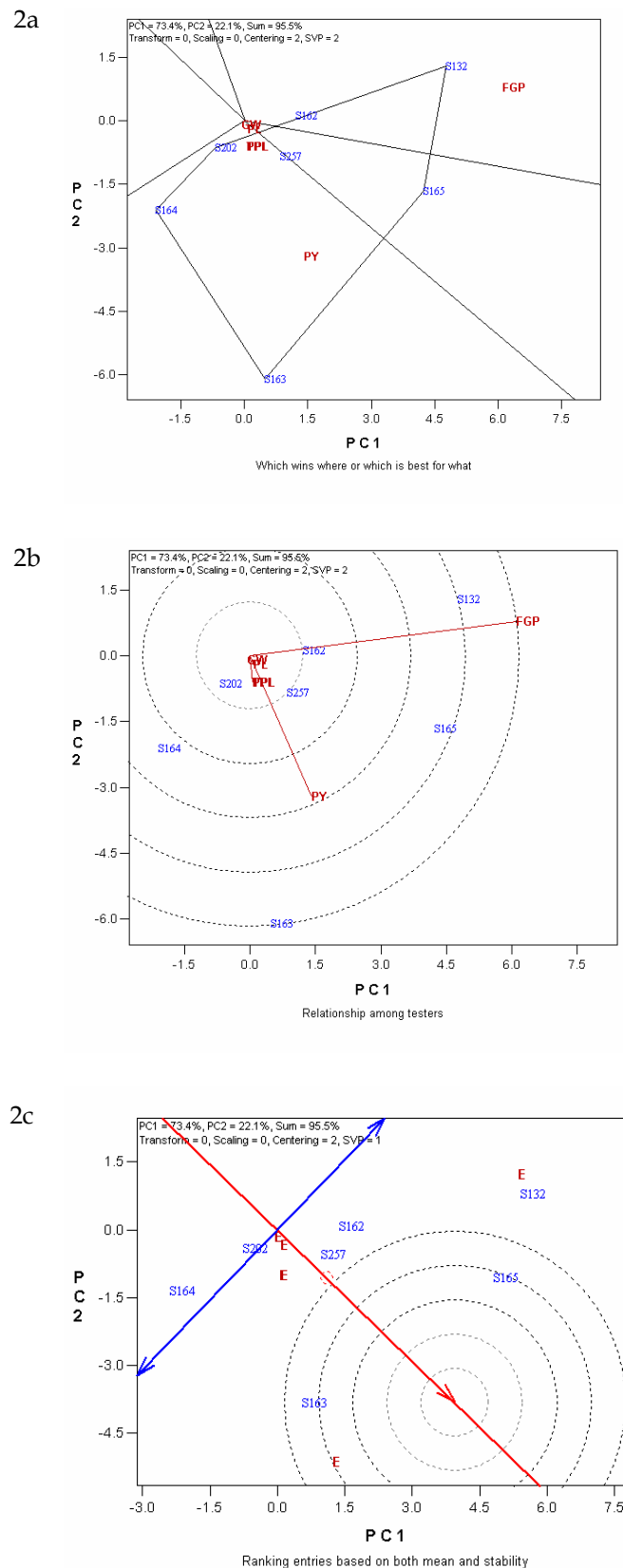
Culm length (CL)	CL	PL	TPL	PPL	FGP	EGP	GW
Panicle length (PL)	0.531***						
Tillers per plant (TPL)	–0.225***	–0.106					
Panicles per plant (PPL)	–0.224***	–0.109	0.999***				
Filled grains per panicle (FGP)	–0.027	0.136*	–0.322***	–0.322***			
Empty grains per panicle (EGP)	0.101	0.176**	–0.030	–0.028	–0.325***		
1000–grain weight (GW)	0.047	–0.105	–0.044	–0.049	–0.173**	0.038	
Grain yield per plant (PY)	0.072	0.100	0.271***	0.271***	0.236***	–0.078	0.071

Significance at *0.05, **0.01, ***0.001 probability level.

DISCUSSION

Rice breeding through incorporation of yield-enhancing genes from wild germplasm is not a new approach. In a previous study, Xiao et al. (1998) showed that quantitative trait loci (QTL) alleles from *Oryza rufipogon* acc. no. IRGC105491 could increase yield potential of cultivated rice by as much as 20%. Despite the apparent yield improvement, traits deleteriously affecting an otherwise good plant type may show linkage with yield- enhancing traits and this can lead to unpredictable results. For an ideal plant type, the optimal culm length of rice should be ~80-90 cm, as in the case of the cultivar parent MR219 used in the current study, and the tillers should be of closed type with flag leaves pointing up. This kind of plant structure is greatly preferred to avoid lodging, to space the plants equally in the field and for maximal capture of sunlight, which increases the plant's yield potential. About 82% of the BC$_2$F$_2$ families were the MR219 types, with a mean culm length of 87 cm, which is similar to that of MR219 (86 cm). *O. rufipogon* types had a mean culm length of 124 cm and 47 families exhibited this kind of phenotype. Consequently, the skewness was much higher for culm length (1.47). All other traits demonstrated almost normal distribution (skewness ranged from –0.11 to 0.74) where the BC$_2$F$_2$ families did not fall into the two distinguishable parental types.

Bimodal frequency distribution of agronomic traits was observed in previous studies; it was detected for plant height in backcross progenies by Xiao et al. (1998) and for kernel elongation ratio in F$_2$ progenies by Faruq et al. (2004).

2a

Which wins where or which is best for what

2b

Relationship among testers

2c

Ranking entries based on both mean and stability

Abbreviations: PL—Panicle length; TPL—Tillers per plant; PPL—Panicles per plant; FGP—Filled grains per panicle; and PY—Grain yield per plant.

Figure 2. GGE biplot analysis of selected quantitative traits in BC_2F_2 families. (a) Which wins where or which is best for what; (b) Relationship among testers; (c) Ranking entries based on both mean and stability.

This kind of distribution of culm length might be due to fewer genes with significant effects (e.g., QTL for plant height had the highest percent phenotypic variance or significance as reported by Xiao et al. (1998); Moncada et al. (2001) and Septiningsih et al. (2003a). This contrasts with the general nature of quantitative traits where there are several distinct genes with small effects as reported in many QTL mapping studies in rice (Lin and Ritland, 1997; Xiao et al., 1998; Moncada et al., 2001; Thomson et al., 2003; Uga et al., 2003). Faruq et al. (2004) suggested that bimodal distribution with one of the peaks was smaller than the other (as for culm length in the current study), suggesting the action of one major gene along with a few modifier genes.

In the present study (based on one season), the grain yield per plant did not show any significant differences among parental or progeny types. This might be due to the small sample size (n = 1) in comparison to the vegetative traits (n = 10). In a similar study, Martinez et al. (1998) also found no significant difference in grain yield between the parental cultivar and its progenies. However, analysis of variance within parental or progeny types showed differences for culm length among individual observations in *O. rufipogon* (P < 0.001) and MR219 (P < 0.001). Similarly, highly significant differences (P < 0.001) were obtained among 266 BC_2F_2 families for all traits measured.

Xiao et al. (1998) reported that 15, 14 and 56% of the BC_2 testcross families outperformed the commercial hybrid (V29A/Ce64) with respect to grain yield, grains per plant and 1000-grain weight, respectively. For the same traits, 62.8, 3.4 and 74.8% of the BC_2F_2 families outperformed parent MR219. Thus, even though *O. rufipogon* is inferior in most of the traits analysed, it positively contributed to yield enhancement in the BC2F2 families. Grain yield in *O. rufipogon* was 36.3 g per plant while MR219 produced on average 42.4 g filled grains (based on one season data). But mean grain yield per plant for 266 BC_2F_2 families was 44.5 g with a highest value of 77.5 g. This increase might be transgressive segregation (genotype × genotype interaction) where *O. rufipogon* alleles augment effect of MR219 alleles as reported by Moncada et al. (2001).

Regression analysis helped to determine the most important variables that affect grain yield in rice. In Case I where grain yield was regressed as a dependent variable with all other traits, an *F*-value of 11.95 was obtained. The *F*-statistics indicated that 22% of the variance in grain yield was attributed to the independent variables (adjusted R^2 = 0.22; Neter et al., 1996). However, not all independent variables were significant or their regression coefficient was too low. Hence re-analysis with culm length, tillers per plant, filled grains per panicle and 1000-grain weight resulted in an *F* value of 20.83 (adjusted R^2 = 0.23; P < 0.001), which indicated that the Case II is more suitable than Case I. These traits were used to rank and select candidate families for subsequent field evaluation and cultivar development.

Analysis of variance (ANOVA) indicated that the variance components in the linear model are required to adequately describe the variation in the agronomic traits (P < 0.001). The BC_2F_2 families were highly variable for filled grains per panicle (MS error, 788.96; coefficient of variation, 26.8).

Estimates of heritabilities for use in selection among BC_2F_2 families were low to moderate and ranged from 0.19 to 0.78 (Table 4). Culm length was found to be the most heritable trait in the BC_2F_2 families, with a heritability estimate of 78%, followed by grain weight (55%). This indicates that selection for these traits in these populations would be most effective for the expression of these traits in the succeeding generations. However, none of these traits showed statistically significant and strong correlation with grain yield per plant, the economic trait. On the other hand, traits having low heritabilities such as tillers per plant (23%), panicles per plant (23%) and filled grains per panicle (19%) showed highly significant (P < 0.001) but low correlation (*r* ranged from 0.24 to 0.27) with grain yield per plant.

Correlation analysis revealed the highest correlation values (*r* = 0.99, P < 0.001) for tillers per plant vs. panicles per plant, which is a clear indication of a lower number of infertile tillers among BC_2F_2 progenies (Table 5). Some traits showed negative correlation. For example,

1000-grain weight was negatively correlated with filled grains per plant ($r = -0.17$). A similar negative correlation between the same traits was reported in previous studies ($r = -0.18$, Xiao et al. (1988); $r = -0.21$, Septiningsih et al. (2003a)). In another report (Moncada et al., 2001) involving BC_2F_2 progenies derived from the same *O. rufipogon* accession as used in the current study but evaluated under upland conditions, the trait correlations were largely different from the present study. For example, 1000-grain weight was positively correlated with grains per plant ($r = 0.23$). These variations in correlation coefficients between traits in different studies can be explained by the genetic makeup of the populations and climatic or edaphic conditions in which they were evaluated.

Twenty-six top-performing families were short-listed (Table 3) which accounted for ~10% of the BC_2F_2 families. Even though BC_2F_2 families in general showed an increase of ~5% grain yield, the short-listed 26 families showed ~27% higher yield per plant in comparison to the parent, MR219. The selected families had 25.1 to 35.7 tillers per plant in contrast to 20.7 and 39.5 tillers for MR219 and *O. rufipogon* respectively. The increase in grain yield is due to higher numbers of tillers per plant, which the BC_2F_2 families most likely inherited from *O. rufipogon*.

The biplot analyses infer interesting and useful implications for breeding. Of much interest was the fact that only filled grains per panicle and yield per plant were useful for identification of outperforming families (Figure 2a). Most of the other traits were clustered together in the center of the biplot indicating little or no impact. There was not a single family that performed well for these two traits. Instead, family S132 recorded highest value for filled grains per panicle and S163 was highest for yield per plant. These two traits were positively correlated (Figure 2b). That was the only important correlation. Pearson correlation analysis also resulted in positive values for these two traits ($r = 0.236$; $P = 0.001$). Biplot analysis could help identify 'ideal' genotypes (highest performing in all traits and most stable across all traits), which must be helpful for breeders. Two families were most promising, viz., S163 and S165 (Figure 2c) and could be used as a starting point for cultivar development programmes.

Not all of the 26 top-performing families are suitable for cultivar development. This is mainly because the tall plants are prone to lodging when they approach maturity and in the face of strong wind or rain. Another attribute requiring consideration is the long awn that is inherited from *O. rufipogon* and is found in family S37. Whereas most of the families inherited grain colour and awn type from MR219 (220 and 194 families, respectively), the inheritance was very low (5 and 55 families, respectively) from *O. rufipogon*. Among the families inheriting grain color and awn type from MR219, 89 families gained both traits from MR219 simultaneously. The lower chi-square statistic for this observation (1.65) is an indication that these attributes are associated with each other.

Although the yield potential of the BC_2F_2 families is high, there is a much higher increase (43%, comparing to MR219) in empty grains per panicle which adversely affected the overall grain yield. For the short-listed families, mean increase in empty grains was 63% (ranged from –28 to 168%). Future breeding programmes based on these BC_2F_2 families should consider this problem to achieve maximum yield potential. To develop improved cultivars, the selected families should be subjected to cultivar performance trials (together with check cultivars) in a sample composed of trials from multiple locations across multiple seasons within an ecogeographic region (Feyerherm et al., 2004), following guidelines as suggested by Fisher (1951) and Snedecor and Cochran (1980).

Even though wild or weedy forms are inferior in most of the agronomic traits, they have been demonstrated to contribute positively to yield enhancement in many crop species (Frey et al., 1983; Xiao et al., 1998; Moncada et al., 2001; Thomson et al., 2003; McCarty et al., 2004a). The present study also confirmed this phenomenon and concludes that at least some of the genotypes excluded during the domestication process still contain valuable trait enhancing genes. These genes could be used to break the yield ceiling in rice to achieve much higher grain yield and to ensure food security for the future.

REFERENCES

Allard, R.W. (1966). *Principles of Plant breeding*, New York. John Wiley & Sons, Inc.

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

Cassman, K.G. (1999). Ecological intensification of cereal production systems: Yield potential, soil quality, and precision agriculture. *Proc. National Acad. Sci.* (USA) 96, 5952–5959.

Cheng, C.Y., Motohashi, R. Tsuchimoto, S. Fukuta, Y., Ohtsubo, H. (2003). Polyphyletic origin of cultivated rice: Based on the interspersion pattern of SINEs. *Mol. Biol. Evol.* 20, 67–75.

Faruq, G., Mohamad, O., Hadjim, K. Meisner, C. (2004). Inheritance of kernel elongation in rice. *Int. J. Agri. Biol.* 6, 813–815.

Feyerherm, A.M., Paulsen, G.M., Fritz, A.K. (2004). A statistical method for summarizing results of cultivar performance trials. *Agron. J.* 96, 992–996.

Fisher, R. A. (1951). *The design of experiments. 6th ed*. Oliver and Boyd Ltd., Edinburgh, UK.

Gomez, K.A., Gomez, A.A. (1984). *Statistical procedures for agricultural research. 2nd Ed.* John Wiley & Sons, Inc. Singapore.

IRRI (2001). *World Rice Statistics 2001.* International Rice Research Institute (IRRI), Manila, Philippines.

Lin, J.Z., Ritland, K. (1997). Quantitative trait loci differentiating the outbreeding Mimulus guttatus from the inbreeding *M. platycalyx*. *Genetics* 146, 1115–1121.

Martínez, C.P., Tohme J., López J., Borrero, J., McCouch S.R., Almeida, A. 1998. Identification and utilization of genes from wild rice germplasm for the improvement of yield and stress resistance. *Proceedings of 27th Rice Technical Working Group (1–4 March 1998, Reno, Nevada, USA).* Texas Agricultural Experiment Station, The Texas A&M University System, College Station, 57.

McCarty, C.J., Jenkins, J.N., Wu, J. (2004a). Primitive accession derived germplasm by cultivar crosses as sources for cotton improvement: I. Phenotypic values and variance components. *Crop Sci.* 44, 1226–1230.

McCarty, C.J., Jenkins, J.N, Wu, J. (2004b). Primitive accession derived germplasm by cultivar crosses as sources for cotton improvement: II. Genetic effects and genotypic values. *Crop Sci.* 44, 1231–1235.

Moncada, P., Martínez, C.P., Borrero, J. Chatel, M., Gauch Jr, H., Guimaraes, E., Tohme, J. McCouch, S.R. (2001). Quantitative trait loci for yield and yield components in an *Oryza sativa* × *O. rufipogon* BC$_2$F$_2$ population evaluated in an upland environment. *Theor. Appl. Genet.* 102, 41–52.

Neter, J., Kutner, M.H., Nachtsheim, C.J., Wasserman, W. (1996). *Applied Linear Regression Models, 3rd edition*. Irwin Publishing, Chicago, Illinois

Oka, H., Morishima, H. (1997). Wild and cultivated rice, in: Matuo, T., Hoshikawa, K. (Eds.), *Science of the rice plant, 3. Genetics*, pp 88–111. Food and Agriculture Policy Research Center, Tokyo.

Peng, S., Cassman, K.G. Virmani, S.S., Sheehy, J., Khush, G.S. (1999). Yield potential trends of tropical rice since the release of IR8 and the challenge of increasing rice yield potential. *Crop Sci.* 39, 1552–1559.

Percival, A.E. (1987). The national collection of *Gossypium* germplasm. *So. Coop. Series Bull.* 321.

Sacks, E.J., Roxas, J.P., Sta. Cruz, M.T. (2003a). Developing perennial upland rice I: Field performance of *Oryza sativa/O. rufipogon* F₁, F₄, and BC₁F₄ progeny. *Crop Sci.* 43, 120–128.

Sacks, E.J., Roxas, J.P., Sta. Cruz., M.T. (2003b). Developing perennial upland rice II: Field performance of S1 families from an intermated *Oryza sativa/O. longistaminata* population. *Crop Sci.* 43, 129–134.

SAS Institute, Inc. (2006). *SAS procedures guide. Release 9.13 edition*. SAS Institute, Inc., Cary, NC.

Second, G. (1982). Origin of the genic diversity of cultivated rice (*Oryza* spp.): Study of the polymorphism scored at 40 isozyme loci. *Jap. J. Genet.* 57, 25–57.

Septiningsih, E.M., Prasetiyono, J., Lubis, E., Tai, T.H., Tjubaryat, T., Moeljopawiro, S., McCouch, S.R. (2003a). Identification of quantitative trait loci for yield and yield components in an advanced backcross population derived from the *Oryza sativa* variety IR64 and the wild relative *O. rufipogon*. *Theor. Appl. Genet.* 107, 1419–1432.

Septiningsih, E.M., Trijatmiko, K.R., Moeljopawiro, S., McCouch, S.R. (2003b). 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.

Snedecor, G.W., Cochran, W.G. (1980). *Statistical methods. 7th ed*. The Iowa State Univ. Press, Ames.

Thomson, M.J., Tai, T.H., McClung, A.M., Lai, X.H., Hinga, M.E., Lobos, K.B., Xu, Y., Martinez, C.P., McCouch, S.R. (2003). Mapping quantitative trait loci for yield, yield components and morphological traits in an advanced backcross population between *Oryza rufipogon* and the *Oryza sativa* cultivar Jefferson. *Theor. Appl. Genet.* 107, 479–493.

Uga, Y., Fukuta, Y., Cai, H.W., Iwata, H., Ohsawa, R., Morishima, H., Fujimura, T. (2003). Mapping QTL influencing rice floral morphology using recombinant inbred lines derived from a cross between *Oryza sativa* L. and *Oryza rufipogon* Griff. *Theor. Appl. Genet.* 107, 218–226.

Wagoner, P. (1990). Perennial grain development: past efforts and potential for the future. *Crit. Rev. Plant Sci.* 9, 381–409.

Xiao J, Li, J., Grandillo, S., Ahn, S.N, Yuan, L., Tanksley, S.D, McCouch, S.R. (1998). Identification of trait-improving quantitative trait loci alleles from a wild rice relative, *Oryza rufipogon*. *Genetics* 150, 899–909.

Yan, W., Kang, M. S. (2003). *GGE Biplot Analysis: A Graphical Tool for Breeders, Geneticists, and Agronomists*. CRC Press. Boca Raton, FL.