Theor Appl Genet (1994) 89:185–192

Qifa Zhang · Y. J. Gao · S. H. Yang · R. A. Ragab
M. A. Saghai Maroof · Z. B. Li

A diallel analysis of heterosis in elite hybrid rice based on RFLPs and microsatellites

Received: 11 September 1993 / Accepted: 12 October 1993

Abstract Hybrid rice has contributed significantly to the dramatic increase of rice production in the world. Despite this, little attention has been given to studying the genetic basis of heterosis in rice. In this paper, we report a diallel analysis of heterosis using two classes of molecular markers: restriction fragment length polymorphisms, (RFLPs) and microsatellites. Eight lines, which represent a significant portion of hybrid rice germ plasm, were crossed in all possible pairs, and the F_1s were evaluated for yield and yield component traits in a replicated field trial. The parental lines were surveyed for polymorphisms with 117 RFLP probes and ten microsatellites, resulting in a total of 76 polymorphic markers well-spaced in the rice RFLP map. The results indicated that high level heterosis is common among these crosses: more than 100% midparent and 40% better-parent heterosis were observed in many F_1s, including some crosses between maintainer lines. Heterosis was found to be much higher for yield than for yield component traits, which fits a multiplicative model almost perfectly. Between 16 and 30 marker loci (positive markers) detected highly significant effects on yield or its component traits. Heterozygosity was significantly correlated with several attributes of performance and heterosis. Correlations based on positive markers (specific heterozygosity) were large for midparent heterosis of yield and seeds/panicle and also for F_1 kernel weight. These large correlations may have practical utility for predicting heterosis.

Key words *Oryza sativa* · Hybrid vigor · Molecular marker · Yield

Communicated by P. M. A. Tigerstedt

Q. Zhang (✉) · Y. J. Gao
Biotechnology Center, Huazhong Agricultural University, Wuhan 430070, China

S. H. Yang · Z. B. Li
Department of Agronomy, Huazhong Agricultural University, Wuhan 430070, China

R. A. Ragab[1] · M. A. Saghai Maroof
Department of Crop and Soil Environmental Sciences, Virgina Polytechnic Institute and State University, Blacksburg, VA 24061, USA

Permanent address:
[1] Department of Genetics, Faculty of Agriculture, Minia University, Minia, Egypt

Introduction

The development of molecular genetic markers that detect variation at the DNA sequence level has made it possible to obtain solutions to problems that were previously inaccessible to genetic manipulation. Extensive genome mapping based on DNA restriction fragment length polymorphism (RFLP) has been accomplished in many crop species (O' Brien 1992). The availability of RFLP-based linkage maps has led to the widespread application of molecular techniques in the genetic studies of crop plants, ranging from mapping genes of economical importance (e.g. McCouch et al. 1990; Liu et al. 1992), to population diversity and systematic analysis (Wang et al. 1992; Zhang et al. 1993), to germ plasm identification and evaluation (e.g. Dudley et al. 1992; Zhang et al. 1993), and to the detection and genetic analysis of quantitatively inheritated agronomic characters (e.g. Keim et al. 1990; Paterson et al. 1990; Stuber et al. 1992).

Short tandem repeats in DNA sequences termed microsatellites or simple sequence repeats (SSRs) have been extensively explored for use as genetic markers in mammals (Tautz 1989; Weber and May 1989). Such markers have demonstrated a number of appealing features relative to RFLPs. Genetic linkage maps based solely on SSR polymorphisms have been constructed for several mammalian species including man (Weissenbach et al. 1992) and mouse (Dietrich et al. 1992). SSR markers are also being rapidly developed in several plant species including soybean (Akkaya et al. 1992; Morgante and Olivieri 1993; Yu et al. 1993) and rice (Wu and Tanksley 1993; Zhao and Kochert 1993). Close

linkage between an SSR marker and a gene conferring resistance to soybean mosaic virus was recently reported by Yu et al. (1993).

The genetic basis of heterosis in plants has been of interest to several generations of scientists. A large body of theories, mostly based on mathematical modeling addressing various aspects of hybrid vigor, has been developed during the last several decades (see Schnell and Cockerham 1992 for a recent review). However, experimental data have largely been unavailable until recently with the application of RFLP techniques in corn (e.g. Smith et al. 1990; Dudley et al. 1991; Stuber et al. 1992). A common finding from these studies was that heterosis was significantly correlated with the heterozygosity of marker loci but that the level of correlation varied from one data set to another. While variable perspectives were obtained regarding the predictability of heterosis and performance of the hybrids, based on the molecular marker genotypes, the relative importance of the various types of gene actions (e.g., additive, dominance, and epistasis) also seemed to differ greatly from one study to another. Such variable observations suggested the complexity of the genetic basis underlying heterosis.

Hybrid rice has contributed significantly to the dramatic increase in rice production of the world (Yuan 1992). An enormous amount of breeding effort has been invested in the development of elite rice hybrids, and this has resulted in the release of many hybrid combinations in China and other Asian countries during the last 20 years. Consequently, a very large area is now being planted with hybrid rice. It is estimated that in China alone hybrid rice occupies an annual acreage of about 17 million ha, which is approximately 55% of the total rice area under cultivation in this country. While tremendous success has been achieved in hybrid rice breeding, experimental data pertinent to the genetic basis of heterosis in rice have remained scarce. The few reported studies are mainly concerned with estimating the amount of heterosis in varietal crosses and demonstrating performance of hybrids (e.g., Rice research group 1978; Virmani et al. 1982; Young and Virmani 1990).

In this paper we describe a diallel analysis of heterosis in some of the best performing rice hybrids using two classes of molecular markers: RFLPs and SSRs. The objectives of this study were: (1) to assess the amount of heterosis and yielding potential among crosses between selected lines that represent a significant portion of the most superior germ plasm in hybrid rice; (2) to identify chromosomal regions with significant effects on yield and its component traits, and (3) to determine the level of correlations of heterozygosity with performance and heterosis.

Materials and methods

Rice lines and crosses

Eight rice lines were used in this study: 'Ce 64-7' (abbreviated as CE hereafter), 'Guang B' (GB), 'Ma Xie' (MX), 'Ming Hui 63' (MH), 'Qing Si Ai' (QS), 'Te Qing' (TQ), 'Xian Gai' (XG) and 'Zhen Shan 97' (ZS). These lines include the parents of several of the best performing hybrids grown in China as well as the parents of some newly released hybrids. Three of these eight lines have been used as restorers, which carry the genes for restoring the fertility of male-sterile lines with several types of cytoplasm; the other five are maintainers for their respective male-sterile lines.

All of the eight lines were selfed for one generation, and seeds of the bagged heads from a single plant per line were used to produce the parents for making the crosses. The eight lines were crossed in all possible combinations to form a diallel set of 28 crosses.

Field experiment

All 28 F_1s and eight parents (36 entries in total) were grown in the 1992 rice growing season at the Agricultural Experimental Station of Huazhong Agricultural University, Wuhan, China. Twenty seedlings per entry were transplanted into a two-row plot following a randomized complete block design (Steel and Torrie 1980) with three replications. All field management was essentially the same as that found under normal agricultural conditions except that plants were spaced 33 × 33 cm apart.

True hybrid plants were identified during the growing season. They varied in number from 5 to 20 per plot, with the majority of plots having 8–9. A maximum of 10 plants per plot were scored for the following four characters: (1) seed-setting tillers per plant, measured as the number of tillers bearing 5 or more filled seeds; (2) number of filled seeds per panicle, which is the total number of seeds per plant divided by the number of seed-setting tillers; (3) weight (g) of 1000 kernels; and (4) grain yield (g) of the whole plant.

Markers and laboratory assays

Two classes of markers were used to survey DNA polymorphism among the parental lines: RFLPs and SSRs. In the first class, about 60 well-spaced probes from the rice RFLP linkage map (Tanksley et al. 1992) were selected for initial screening. A marker from a nearby region was added to the survey whenever a probe failed to detect polymorphism among the parents. In all, 117 probes were assayed. The second class consisted of ten SSR markers originally developed by Wu and Tanksley (1993).

Leaf tissue was harvested in bulk from seedlings of selfed seeds for each parent. DNA extraction followed essentially the CTAB method (Saghai Maroof et al. 1984) with three modifications: (1) Fresh leaf tissue was ground to fine powder under liquid nitrogen; (2) DNA extraction buffer, its concentration increased to 1.67 × and preheated to 95 °C, was added at 2.0 ml per gram fresh tissue; (3) the resulting DNA was purified with RNase digestion followed by a phenol/chloroform-chloroform extraction.

For surveying with the first 60 probes, DNA samples were digested singly with three restriction enzymes, *Bam*HI, *Eco*RI, and *Hin*dIII. Six restriction enzymes (*Bam*HI, *Dra*I, *Eco*RI, *Eco*RV, *Hin*dIII, and *Xba*I) were used for the remaining 57 probes. The only exception was the survey for ribosomal DNA (rDNA) spacer-length variants in which DNA was digested with only one enzyme (*Bam*HI). Thus, in all, 523 probe/enzyme combinations (P/E) were surveyed. Digestion, electrophoresis, blotting and hybridization followed previously described procedures (Saghai Maroof et al. 1984; Zhang et al. 1992).

The primers for the SSR markers were synthesized by Operon Technologies, Calif., USA, according to the published sequences by Wu and Tanksley (1993). SSR procedures were basically as described by Wu and Tanksley (1993). Briefly, the polymerase chain reaction (PCR) was conducted in a volume of 20 µl containing 50 ng of template DNA, 0.2 µ*M* each primer, 160 µ*M* each dATP, dGTP, dTTP and 2 µ*M* dCTP, 1 µCi [α-^{32}P]dCTP, 50 m*M* KCl, 10 m*M* Tris pH 8.3, 0.01% gelatin, 3 m*M* $MgCl_2$, and 1.0 unit of Taq polymerase. Samples were covered with mineral oil and the reaction was processed at 94 °C for 1 min, 55 °C for 2 min, and 72 °C for 1.5 min for 30 cycles followed by an extension step of 72 °C for 5 min. After the PCR reaction, 10 µl stop solution containing 95% formamide, 20 m*M* EDTA was added to the amplification product, and 3 µl per sample

was loaded on a 5.1% polyacrylamide denaturing gel containing 8 *M* urea.

Statistical analysis

The amount of hybrid vigor (H) was evaluated using two measurements: (1) midparent heterosis:

$$H = (F_1 - \bar{P})/\bar{P}$$

where F_1 is the mean of the F_1 hybrid, and

$$\bar{P} = (P_1 + P_2)/2$$

in which P_1 and P_2 are the means of the corresponding parents; and (2) better-parent heterosis:

$$H = (F_1 - P_b)/P_b$$

in which P_b is the value of the better-parent.

Heterozygosity of an F_1 hybrid was measured as the percentage difference of marker genotypes between the two parents. This was also used as a measure of distance between parents for grouping the parental lines using several clustering algorithms including single linkage, complete linkage (Sokal and Sneath 1963), and Ward's method (Ward 1963).

The effect of a chromosomal region on a trait as marked by a molecular marker was assessed with one-way analysis of variance using marker genotypes as the group effect and entries within genotypes as the error term.

Two genetic parameters, additive effect and degree of dominance, were estimated using a single-locus model for markers that detected significant effects on the trait (positive markers). The additive effect (a), measuring the half difference between the two parents, is calculated as

$$a = \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} (N_{ij}|X_{ii} - X_{jj}|/2)/\sum N_{ij},$$

where X_{ii} is the average of the individuals homozygous for the i^{th} allele of the RFLP locus, N_{ij} is the number of F_1s heterozygous for the i^{th} and j^{th} alleles, and k is the total number of alleles at the locus. The dominance effect for each locus is calculated by

$$d = \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} [N_{ij}|X_{ij} - (X_{ii} + X_{jj})/2|]/\Sigma N_{ij}.$$

It is necessary to take the absolute values for both estimates since the effects of different genotypes may cancel out each other for loci with multiple alleles. The degree of dominance for a locus is given by d/a. A zero d/a indicates no dominance, d/a values between 0 and 1 measure partial dominance and d/a values greater than 1 signify overdominance.

Results

Heterosis of yield and yield components

Highly significant differences in yield and yield components were observed among the 28 crosses and the eight parents (Table 1). The highest yielding F_1s had 'Min Hui 63' (MH) and 'Guang B' (GB) as one or both of the parents. The crosses corresponding to the best performing commercial hybrids were among the highest yielding F_1s. It is interesting to note that the two highest yielding F_1s were from crosses between maintainer lines: XG × GB and MX × GB.

Table 1 Measurements of yield and its components in the diallel set involving eight parental rice lines

Genotype	Tillers/plant	Seeds/panicle	Kernel weight	Yield/plant
MH × QS	43.8	136	24.3	145.6 *CD*[c]
MH × TQ	36.8	185	24.2	166.6 *AB*
MH × CE	44.9	135	24.2	148.1 *CD*
MH × XG[a]	33.9	206	25.0	172.4 *AB*
MH × ZS[a]	38.7	162	26.8	171.1 *AB*
MH × MX[a]	43.4	142	28.7	176.0 *A*
MH × GB	39.3	177	25.0	173.6 *AB*
QS × TQ	40.8	162	22.2	145.4 *CD*
QS × CE	45.1	108	21.3	103.0 *GH*
QS × XG	35.3	168	22.4	130.7 *DE*
QS × ZS	37.5	146	24.2	132.6 *DE*
QS × MX	41.9	137	24.9	143.3 *CD*
QS × GB	39.0	157	21.9	135.2 *DE*
TQ × CE	42.2	119	22.1	110.5 *FG*
TQ × XG	30.6	230	22.4	157.6 *BC*
TQ × ZS	36.5	178	25.1	161.2 *AB*
TQ × MX	39.9	164	25.1	163.3 *AB*
TQ × GB	36.3	187	21.7	146.9 *CD*
CE × XG	40.5	128	23.1	118.1 *FG*
CE × ZS[a,b]	47.4	121	23.6	137.1 *DE*
CE × MX	43.3	115	23.9	121.9 *EF*
CE × GB	38.8	137	21.6	114.1 *FG*
XG × ZS	26.8	165	23.3	104.3 *G*
XG × MX	31.6	136	24.2	102.7 *GH*
XG × GB	33.5	227	23.2	176.8 *A*
ZS × MX	29.3	58	24.1	40.9 *J*
ZS × GB	37.6	162	23.9	144.0 *CD*
MX × GB	44.3	153	25.6	178.0 *A*
MH	40.5	114	27.0	123.7 *EF*
QS	43.9	121	20.5	107.5 *FG*
TQ	35.9	182	24.0	157.1 *BC*
CE	48.6	83	21.0	84.2 *HI*
XG	19.9	166	22.5	74.1 *I*
ZS	25.7	57	23.6	35.0 *J*
MX	32.9	47	23.9	36.4 *J*
GB	37.0	165	21.5	131.3 *DE*
Average	37.9	145	23.7	129.7

[a] Crosses corresponding to commercial hybrids
[b] An early maturing hybrid
[c] Ranked using Duncan's multiple range test (Duncan 1955) at the 0.01 probability level

The amount of heterosis differed greatly from one character to another, and also varied widely among the 28 crosses (Table 2). The highest heterosis was observed for yield, with greater than 100% midparent heterosis and more than 40% better-parent heterosis in a number of crosses. In general, commercial hybrids were among those that demonstrated the highest heterosis. A cross between two maintainer lines, MX × GB, also showed a very high level of heterosis.

Among the three yield components, seeds/panicle displayed the highest level of heterosis, followed by tillers/plant; very little heterosis was detected for kernel weight.

Polymorphism of marker loci

A total of 203 probe/enzyme combinations (P/E), involving 68 probes and all six restriction enzymes de-

Table 2 Mid- and better- parent heterosis for yield and its component traits in 28 crosses of the diallel set (*MP* Midparent heterosis, *BP* better-parent heterosis)

	Tillers/plant		Seeds/panicle		Kernel weight		Yield/plant	
	MP	BP	MP	BP	MP	BP	MP	BP
MH × QS	3.79	−0.23	15.69	12.45	2.10	−10.00	25.95	17.70
MH × TQ	−3.66	−9.14	24.81	1.59	−5.10	−10.37	18.66	6.05
MH × CE	0.90	−7.61	36.50	18.01	0.83	−10.37	42.54	19.73
MH × XG	12.25	−16.30	46.65	23.74	0.81	−7.41	74.32	39.37
MH × ZS	16.92	−4.44	88.68	41.35	5.93	−0.74	115.76	38.32
MH × MX	18.26	7.16	74.01	22.90	12.55	6.30	119.73	42.28
MH × GB	1.29	−2.96	26.39	7.03	2.88	−7.41	36.16	32.22
QS × TQ	2.26	−7.06	6.79	−11.03	−0.45	−7.50	9.90	−7.45
QS × CE	−2.59	−7.20	5.66	−10.80	2.40	1.43	7.52	−4.19
QS × XG	10.66	−19.59	16.47	0.72	4.19	−0.44	43.94	21.58
QS × ZS	7.76	−14.58	63.90	20.53	10.00	2.54	85.97	23.35
QS × MX	9.11	−4.56	62.75	13.11	12.16	4.18	99.30	33.30
QS × GB	−3.70	−11.16	9.92	−4.66	4.29	1.86	13.23	2.97
TQ × CE	−0.24	−13.17	−10.24	−34.58	−1.78	−7.92	−8.45	−29.66
TQ × XG	9.68	−14.76	32.13	26.40	−3.86	−6.67	36.33	0.32
TQ × ZS	18.51	1.67	48.83	−2.31	5.46	4.58	67.74	2.61
TQ × MX	15.99	11.14	43.33	−9.77	4.58	4.58	68.70	3.95
TQ × GB	−0.55	−1.89	7.78	2.69	−4.82	−9.58	1.87	−6.49
CE × XG	18.08	−16.67	2.56	−23.02	5.96	2.67	49.30	40.26
CE × ZS	27.42	−2.47	72.51	45.20	5.83	0.00	130.03	62.83
CE × MX	6.13	−10.91	76.26	38.01	6.22	0.00	102.16	44.77
CE × GB	−9.35	−20.16	9.98	−17.20	1.41	0.47	5.84	−13.10
XG × ZS	17.54	4.28	47.81	−0.78	1.30	−1.27	91.38	40.76
XG × MX	19.70	−3.95	27.72	−18.03	4.31	1.26	85.71	38.60
XG × GB	17.54	−9.46	36.73	36.24	5.45	3.11	72.15	34.65
ZS × MX	0.00	−10.94	10.94	1.40	1.26	0.84	14.57	12.36
ZS × GB	19.75	1.62	45.72	−1.94	6.22	1.27	73.08	9.67
MX × GB	26.57	19.73	44.54	−7.03	12.78	7.11	112.16	35.57
Average	9.29	−5.84	34.82	6.08	3.68	−1.34	56.98	19.37

tected polymorphism among the eight parents. Out of the 68 probes, 2 (RG 146 and rDNA) were known to detect two distinct loci each. Among the remaining 66 probes, 59 displayed banding patterns that are in agreement with typical single locus variation, whereas multiple variable bands were detected by the other 7 probes with at least one enzyme. Data from the P/E combinations displaying multiple variable bands were scored and analyzed on the basis of individual bands.

Banding patterns resolved by different enzymes within probes were often perfectly correlated with each other, thus generating redundant information. In such cases, data from one of those enzymes were used in the analysis. The same data processing scheme also applies to those probes whose RFLPs were scored as individual bands. Consequently, the amount of data used from different probes was not necessarily equal. Scorable bands were resolved with eight of the ten microsatellite markers; variants detected by each marker appeared to be alleles of a single locus. Altogether, the 76 variable markers, 68 RFLPs and eight SSRs, produced 157 pieces of non-redundant information in detecting polymorphism among the eight parents.

The level of polymorphism defined as percentage differences between each pair of parents is illustrated in Table 3. This can also be viewed as the distance between parents as well as the level of heterozygosity of the resulting F_1s.

Cluster analysis

A cluster analysis was conducted using the percentage differences of non-redundant data between lines at marker loci as the distance measure. Identical hierarchical trees were obtained using complete linkage and Ward's method, which placed these eight parents into three groups with QS, TQ and GB in the first group, XG, ZS and MX in the second group, and MH and CE in the third group.

Table 3 Distances between paired parents as measured by percent disagreement of genotypes at marker loci

	MH	QS	TQ	CE	XG	ZS	MX
QS	0.527						
TQ	0.515	0.248					
CE	0.442	0.418	0.455				
XG	0.552	0.455	0.479	0.491			
ZS	0.576	0.509	0.461	0.545	0.424		
MX	0.479	0.430	0.455	0.497	0.448	0.351	
GB	0.515	0.230	0.255	0.412	0.497	0.497	0.497

Both MH and CE were derivatives of IRRI lines, and all three lines in the first group, QS, TQ and GB, were released from the Guang Dong Academy of Agricultural Sciences. However, the available pedigree information (Lin and Ming 1991) did not allow for a detailed assessment concerning the goodness of match between similarity of marker genotypes and the parentage among the eight lines.

Marker loci that detected significant effects on yield and yield components

Marker loci that detected significant effects (referred to as positive markers hereafter) with the one-way AOV at the 0.01 probability level for each trait are listed in Table 4. The number of positive markers are 35, 17, 27, and 16 for tillers/plant, seeds/panicle, kernel weight, and yield, respectively. It is clear from Table 4 that there is a large proportion (11/16) of positive markers in common between yield and seeds/panicle but that this is much less so between yield and tillers/plant or kernel weight. Many positive markers for tillers/plant and kernel weight did not seem to contribute directly to yield.

The numerical values of the two genetic parameters (a and d/a) varied greatly among the chromosomal segments underlying these positive marker loci (Table 4). Various degrees of partial dominance were observed for a majority of the loci for the three component traits (tillers/plant, seeds/panicle, and kernel weight), and overdominance occurred only in a small proportion of the loci. Interestingly, all of the positive markers for yield demonstrated overdominance, with d/a values varying from 1.13 to 3.66.

Correlation of heterozygosity with performance and heterosis

Two measurements of heterozygosity were used for calculating the correlations of heterozygosity with performance and heterosis: heterozygosity at all marker loci (general heterozygosity) and heterozygosity calculated on the basis of positive markers for each trait (specific heterozygosity).

The correlation between general heterozygosity and performance was generally small and it was only significant with kernel weight (Table 5). On the other hand, heterosis appeared to be more closely correlated with general heterozygosity. Significant correlations were detected between general heterozygosity and midparent heterosis for three traits, tillers/plant, seeds/panicle, and yield, and between general heterozygosity and better-parent heterosis for two traits, seeds/panicle and yield.

When specific heterozygosity for each trait was used in the calculation, both increases and decreases in correlations were observed among the various attributes compared to those obtained using general heterozygosity. The most noteworthy instances of increased correlations were F_1 kernel weight, and midparent heterosis of seeds/panicle and yield (Table 5).

Discussion

Exploiting hybrid vigor has become a major goal in many rice breeding programs. Our results clearly established that strong heterosis is common in crosses of this diallel set from parents comprising elite germ plasm for hybrid rice breeding.

One striking observation is the substantially larger amount of heterosis for yield than for yield component traits. This is understandable because yield is a generally multiplicative function of the three component traits, tillers/plant (T), seeds/panicle (S), and kernel weight (W); i.e., $Y = T \times S \times W$, a relationship which holds perfectly in our data set. Thus, heterosis for yield can be expressed as:

$$Y_{het} = Y_{F_1}/Y_{\bar{P}} = T_{F_1} \cdot S_{F_1} \cdot W_{F_1}/[(Y_{P_1} + Y_{P_2})/2]$$

$$= T_{F_1} \cdot S_{F_1} \cdot W_{F_1}/[(T_{P_1} \cdot S_{P_1} \cdot W_{P_1} + T_{P_2} \cdot S_{P_2} \cdot W_{P_2})/2]$$

It is easy to show from this equation that the slight heterosis of component traits would multiplicatively amplify each other to produce a much larger heterosis in the ultimate trait. Such amplification effects would be very pronounced if heterosis in the component traits were sizable. For example, about 26% heterosis for each of the three component traits would result in 100% heterosis in the composite trait. The above relationship is well exemplified by data from the various crosses included in the present study (Table 2), indicating that the amount of heterosis for yield was always in close agreement with predicted values based on the above equation. This multiplicative relationship of component heterosis has been the subject of a number of theoretical studies (e.g., Grafius 1959; Schnell and Cockerham 1992). A similar phenomenon has been noted in a number of previous studies, including one on cross breeding in guinea pig by Wright (1922), who observed that the relative improvements were much larger for total performance than for its component traits, and in a study of barley by Immer (1941), who found a close agreement between the observed heterosis of yield and the product of the heterosis of the three component traits.

This same reasoning also applies to the types of gene action (e.g., dominance vs. overdominance) involved in heterosis. In theory, a locus has an effect on the ultimate trait (yield) only when it exerts its effect on one or more of the component traits (tillers/plant, seeds/panicle, and kernel weight). Based on the single-locus model, a d/a (degree of dominance) value larger than 0.26 on each component trait would result in a d/a larger than 1.0 (apparent overdominance) in the ultimate trait if this locus has an effect on all three component traits. Similarly, for a locus having an effect on two of the three component traits, a d/a value of 0.41 is necessary in

Table 4 Additive effects and degrees of dominance estimated for marker loci that had a highly significant ($P \leq 0.01$) effect (positive markers) on yield and yield components in this diallel set

Marker	Chromosomal location	Tillers/plant	Seeds/panicle	Kernel weight	Yield
RG381	1		1.50/0.22[a]		
RG101	1		51.46/0.84	0.48/3.25	48.49/1.21
RG462	1			1.68/0.12	
RG532	1	6.58/0.71	51.46/0.84	1.57/0.46	33.64/1.58
RG472	1	6.58/0.71	19.85/1.90		33.64/1.58
RG520	2	5.37/1.13		1.17/0.54	
CDO1091	2			1.95/0.21	
RG151	2		51.46/0.84	1.80/0.30	48.49/1.21
RG139	2	9.90/0.34			
RG171	2	4.81/0.26			
RZ599	2		25.61/0.85		25.87/1.13
RG152	2	7.49/0.20			
RG191	3	5.64/0.19			
RG722	3	9.90/0.34			
RG393	3			1.95/0.21	
RG620	4		27.92/0.39		
RZ819	4	9.90/0.34			
RG214	4		34.19/0.59	1.45/1.01	27.20/1.16
RG449	4	3.61/0.81			
RZ262	4	4.81/0.26			
RM122	5			1.42/1.18	
RG207	5	4.86/0.44			
RG360	5	4.81/0.26			
RG403	5	4.81/0.26			
RM164	5	5.61/0.78		1.24/1.28	24.76/2.23
RG470	5			1.30/0.86	
RZ516	6	5.37/1.13		1.17/0.54	
RZ398	6		29.38/0.36		
RZ588	6	3.78/0.96			22.43/1.66
RG213	6	5.87/0.42			
RG648	6	9.90/0.34			
RG424	6			1.34/0.21	
RZ828	6	6.58/0.71	19.85/1.90		33.64/1.58
RG634	7	4.72/0.34		1.57/0.30	
RG128	7	4.25/0.72	19.85/1.90	1.57/0.46	9.93/3.66
RG1034	8	6.22/0.13			
RZ66	8		51.46/0.84	0.48/3.25	48.49/1.21
RZ649	8	8.38/0.42	29.51/0.22	1.57/0.46	9.93/3.66
RZ698	9	6.33/0.74			37.13/1.51
RG358	9	5.67/0.30	33.77/0.29		
RG667	9	6.58/0.71	19.85/1.90		33.64/1.58
RG570	9	6.27/0.31			
RZ404	9	5.37/1.13		1.17/0.54	
RZ892	10			1.95/0.21	
RZ337	10			1.96/0.10	
RG134	10	9.90/0.34			
RG1109	11	8.78/0.55			
RG2	11			1.41/0.56	
RG98	11		36.65/0.34	1.95/0.21	
RG396	12	6.22/0.13			
RZ76	12	4.29/0.65		1.68/0.38	
RG9	12	9.90/0.34		1.12/1.11	29.74/2.18
rDNA(1)				1.18/0.04	
rDNA(2)				1.95/0.21	
rDNA(3)					15.08/1.89
RM163			32.06/1.42		
RM123		9.90/0.34			

[a] Additive effect/degree of dominance.

order for overdominance to be observed in the ultimate trait.

In the present study, 3 of the 16 positive marker loci for yield demonstrated highly significant effects on all three component traits (Table 4), and another 9 had highly significant effects on two component traits. The majority of the d/a values estimated for the component traits at these positive marker loci exceeded their respective thresholds (0.26 and 0.41), as described above. It should be noted that the data presentation in Table 4

Table 5 Correlations of heterozygosity with performance, midparent (MP) heterosis and better-parent (BP) heterosis in the F_1s of this diallel set ($r_{0.05} = 0.374$, $r_{0.01} = 0.478$)

	Performance	MP heterosis	BP heterosis
Tillers/plant	0.132/0.097[a]	0.487/0.349	0.075/−0.376
Seeds/panicle	0.126/0.179	0.539/0.714	0.435/0.227
Kernel weight	0.527/0.702	0.301/0.250	0.076/−0.153
Yield	0.355/0.479	0.561/0.773	0.529/0.392

[a] Correlations calculated using: general heterozygosity/specific heterozygosity. See text for the definition of general and specific heterozygosity.

was truncated at the 0.01 probability level, and probe/trait combinations that demonstrated smaller effects are not listed. Also, each of the three component traits can be further divided into subtraits, and in turn, the expression of each subtrait may be ascribed to a set of genes. Clearly, overdominance for yield does not need to be imposed as a genetic explanation for the large observed d/a values, at least for the loci that did not show overdominance in the component traits. Crow (1952) pointed out another possibility for heterosis, which he termed pseudo-overdominance (linked loci with advantageous alleles in repulsion phase). Further, Minvielle (1987) suggested that dominance may not be necessary for heterosis with the presence of epistasis. However, the present data did not permit evaluation of such propositions and, indeed, such possibilities cannot be distinguished without extensive recombination analysis.

Molecular marker-based diallel analysis has also been used to detect the effects of specific genes. In a diallel study of powdery mildew resistance in barley, Saghai Maroof et al. (1993) identified a number of markers that detected significant effects on resistance to one or more of the mildew isolates among the F_1s and their parents. Interestingly, a majority of these markers are located in the chromosomal regions where genes for powdery mildew resistance had been reported in previous studies based on segregation analysis. Such a remarkable agreement between the locations of resistance genes and positive markers strongly indicated that marker-based diallel analysis may be very useful for detecting the presence of a gene as well as determining its approximate map location. Such usefulness may even be enhanced by prior knowledge of the presence and locations of the genes of interest.

A possible problem associated with the detection of a single gene effect in a diallel system is the frequent sporadic correlations among marker loci even when located on different chromosomes due to the small number of parental lines used in the analysis. Such sporadic correlations would often cause false positives in the detection. However, because of the high degree of stringency we adopted for selecting the positive markers from the AOV, it is reasonable to assume that a marker is underlining a locus for yield if: (1) it is simultaneously showing highly significant effects on yield and one or more component traits, and (2) the marker genotype is not perfectly correlated with other positive markers for yield. Thus, 5 of the positive markers, RZ599 on chromosome 2, RG214 on chromosome 4, RZ588 on chromosome 6, RZ698 on chromosome 9, and RG9 on chromosome 12, are likely to be true positive markers, each of them marking a locus for yield. There are, however, at least two possibilities for those unlinked positive markers that are perfectly correlated with each other: (1) each positive marker detected a distinct locus for yield, and the estimated numerical value of the genetic parameters represented the joint effect of those loic; (2) some markers detected loci for yield, but other markers were simply hitchhikers. These possibilities will be differentiated by segregating population analysis in subsequent studies, which will also lead to a detailed characterization of the types of gene action and the amounts of genetic effects at each locus, as well as in the entire genome.

One of the major objectives of this study was to assess the relationship of marker locus heterozygosity, performance, and heterosis among the progenies from parents of best-performing rice hybrids. The results showed that correlation is, in general, low between heterozygosity and performance, is higher between heterozygosity and better-parent heterosis, and is highest between heterozygosity and midparent heterosis. Moreover, when only data from the positive markers (specific heterozygosity) for each trait were used, large correlation coefficients were obtained in several instances, indicating that heterozygosity has contributed an important component to heterosis in hybrid rice. This finding is in agreement with results from previous studies in corn (e.g., Stuber et al. 1992). These large correlations between heterosis and specific heterozygosity as defined in the present study may be useful for prediction purposes. In reality, it may be much more practical to predict heterosis on the basis of a small number of informative markers than to use a large number of markers covering the entire genome, although many studies are certainly necessary before we can identify such a small number of markers that will give accurate predictions.

Acknowledgement The authors thank Dr. S.D. Tanksley for kindly providing the probes and K.S. Wu for technical assistance. This research was supported in part by a grant from the Rockefeller Foundation and a grant from the Education Commission of PRC. R.A. Ragab was a Rockefeller Foundation Biotechnology Fellowship recipient.

References

Akkaya MS, Bhagwat AA, Cregan PB (1992) Length polymorphisms of simple sequence repeat DNA in soybean. Genetics 132: 1131–1139

Crow JF (1952) Dominance and overdominance. In Gowen JW (ed) Heterosis. Iowa State College Press, Ames, pp 282–297

Dietrich W, Katz H, Lincoln SE, Shin HS, Friedman J, Dracopoli NC, Lander ES (1992) A genetic map of the mouse suitable for typing interspecific crosses. Genetics 131:423–477

Dudley JW, Saghai Maroof MA, Rufener GK (1991) Molecular markers and grouping of parents in maize breeding programs. Crop Sci 31:718–723

Dudley JW, Saghai Maroof MA, Rufener GK (1992) Molecular marker information and selection of parents in corn breeding programs. Crop Sci 32:301–304

Duncan DB (1955) Multiple range and multiple F tests. Biometrics 11:1–42

Grafius JE (1961) The complex trait as a geometric construct. Heredity 16:225–228

Immer FR (1941) Relation between yielding ability and homozygosis in barley crosses. J Am Soc Agron 33:200–206

Keim P, Diers BW, Olson TC, Shoemaker RC (1990) RFLP mapping in soybean: association between marker loci and variation in quantitative traits. Genetics 126:735–742

Lin S, Ming S (1991) Rice varieties and their genealogy in China. Shanghai Science and Technology Press, Shanghai (in Chinese)

Liu A, Zhang Q, Li H (1992) Location of a gene for wide compatibility in the RFLP linkage map. Rice Genet Newsl 9:134–136

McCouch SR, Kush, GS, Tanksley SD (1990) Tagging genes for disease and insect resistance via linkage to RFLP markers. In: IRRI (ed) Rice genetics II. (Proc 2nd Int Rice Genet Symp. Manila, Philippines, pp 443–449

Minvielle F (1987) Dominance is not necessary for heterosis: a two-locus model. Genet Res 49:245–247

Morgante M, Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. Plant J 3:175–182

O'Brien SJ (1992) Genetic maps. Book 6: plants. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Paterson AH, De Verna JW, Lanini B, Tanksley SD (1991) Fine mapping of quantitative trait loci using selected overlapping recombinant chromosomes in an interspecies cross of tomato. Genetics 124:735–742

Rice research group of Mian Yang Agricultural Institute of Sichuan Province (1978) An analysis on heterosis of F_1 hybrids in rice. Acta Sin Genet 5:158–164

Saghai Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci USA 81:8014–8018

Saghai Maroof MA, Zhang Q, Biyashev RM (1994) Molecular marker analysis of powdery mildew resistance using a barley diallel cross. Theor Appl Genet (in press)

Schnell FW, Cockerham CC (1992) Multiplicative vs. arbitrary gene action in heterosis. Genetics 131:461–469

Smith OS, Smith JSC, Bowen SL, Tenborg RA, Wall SJ (1990) Similarities among a group of elite maize inbreds as measured by pedigree, F_1 grain yield, grain yield heterosis and RFLPs. Theor Appl Genet 80:833–840

Sokal R, Sneath PHA (1963) Principles of numerical taxonomy. Freeman WH and Company, San Francisco and London

Song KM, Osborn TC, William PH (1988) Brassica taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs) Theor Appl Genet 79:497–506

Steel RGD, Torrie JH (1980) Principles and procedures of statistics. McGraw-Hill Book Co

Stuber CW, Lincoln SE, Wolff DW, Helentjaris T, Lander ES (1992) Identification of genetic factors contributing to heterosis in a hybrid from elite maize inbred lines using molecular markers. Genetics 132:823–839

Tanksley S, Causse M, Fulton T, Ahn N, Wang Z, Wu K, Xiao J, Yu Z, Second G, McCouch S (1992) A high density molecular map of the rice genome. Rice Genet Newsl 9:111–115

Tautz D (1989) Hypervariability of simple sequences as a general source of polymorphic DNA markers. Nucleic Acids Res 17: 6463–6471

Virmani SS, Aquino RC, Khush GS (1982) Heterosis breeding in rice. Theor Appl Genet 63:373–380

Wang ZY, Second G, Tanksley SD (1992) Polymorphism and phylogenetic relationship among species in the genus *Oryza* as determined by analysis of nuclear RFLPs. Theor Appl Genet 83:565–581

Ward JH (1963) Hierarchical grouping to optimize an objective function. J Am Stat Assoc 58:236–244

Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet 44:388–396

Weissenbach J, Gyapay G, Dip C, Virgnal A, Morissette J, Millasseau P, Vaysseix G, Lathrop M (1992) A second-generation linkage map of the human genome. Nature 359:794–801

Wright S (1992) The effects of inbreeding and cross breeding on guinea pigs. III. Crosses between highly inbred families. Tech Bull US Dept Agric 1121

Wu KS, Tanksley SD (1993) Abundance, polymorphism and genetic mapping of microsatellites in rice. Mol Gen Genet 241:225–235

Young JB, Virmani SS (1990) Heterosis in rice over environments. Euphytica 51:87–93

Yu YG, Saghai Maroof MA, Buss GR, Maughan PJ, Tolin SA (1993) RFLP and microsatellite mapping of a gene for soybean mosaic virus resistance. Phytopathology 84:60–64

Yuan LP (1992) Development and prospects of hybrid rice breeding. In: You C, Chen ZL (eds) Agricultural biotechnology. (Proc Asia-Pacific Conf Agric Biotechnol). China Science and Technol Press, Beijing, pp 97–105

Zhang Q, Saghai Maroof MA, Lu TY, Shen BZ (1992) Genetic diversity and differentiation of *indica* and *japonica* rice detected by RFLP analysis. Theor Appl Genet 83:495–499

Zhang Q, Saghai Maroof MA, Kleinhofs A (1993) Comparative diversity analysis of RFLPs and isozymes within and among populations of *Hordeum vulgare* ssp. *spontaneum*. Genetics 134:909–916

Zhao X, Kochert G (1993) Phylogenetic distribution and genetic mapping of a (GGC)n microsatellite from rice (*Oryza sativa* L.). Plant Mol Biol 21:607–614