Molecular Breeding **1**: 133–142, 1995.

Molecular divergence and hybrid performance in rice

Qifa Zhang [1], Y.J. Gao [1], M.A. Saghai Maroof [2], S.H. Yang [3] and J.X. Li [1]
[1] *Biotechnology Center and* [3] *Department of Agronomy, Huazhong Agricultural University, Wuhan 430070, China;* [2] *Department of Crop and Soil Environmental Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA*

Received 18 April 1994; accepted in revised form 1 September 1994

Key words: genetic distance, heterosis, *Oryza sativa*, restriction fragment length polymorphism (RFLP), simple sequence repeat (SSR)

Abstract

This study was undertaken to determine the relationship between genetic distance of the parents based on molecular markers and F_1 performance in a set of diallel crosses involving eight commonly used parental lines in hybrid rice production. The F_1s and their parents were measured for five traits including heading date, plant height, straw weight, grain yield and biomass. The parental lines were assayed for DNA polymorphisms using two classes of markers: 140 probes for restriction fragment length polymorphisms (RFLPs) and 12 simple sequence repeats (SSRs), resulting in a total of 105 polymorphic markers well spaced along the 12 rice chromosomes. SSRs detected more polymorphism than RFLPs among the eight lines. A cluster analysis based on marker genotypes separated these eight lines into three groups which agree essentially with the available pedigree information. Correlations were mostly low between general heterozygosity based on all the markers and F_1 performance and heterosis. In contrast, very high correlations were detected between midparent heterosis and specific heterozygosity based on the markers that detected significant effects for all the five traits; these correlations may have practical utility in predicting heterosis. The analyses also suggest the existence of two likely heterotic groups in the rice germplasm represented by these eight lines.

Introduction

Prediction of hybrid performance has been of primary interest to essentially all hybrid crop breeding programs which have attracted enormous amounts of efforts over the years [6, 11, 20]. Recently, genetic linkage maps based on molecular markers have been constructed for many crop species [12]. It is hoped that the availability of molecular marker-based linkage maps will provide an effective means for predicting heterosis and thus expedite the time- and labor-consuming field screening.

Large numbers of studies have been conducted in corn to investigate the relationship between marker genotype divergence of the parents and performance of hybrids, which have produced variable results. Lee *et al.* [7] and Smith *et al.* [17] detected strong correlations between genetic distance based on DNA restriction fragment length polymorphisms (RFLPs) and hybrid performance or heterosis in crosses between elite inbreds from the US Corn Belt. In contrast, Godshalk *et al.* [5] and Dudley *et al.* [4] observed low correlations between marker distance and yield. Whereas the results of Melchinger *et al.* [9] and

Boppenmaier *et al.* [2] indicated that the correlations between marker distance and F_1 performance are dependent on the origin of lines studied; the correlations are high for crosses between lines from the same heterotic groups, low between lines from different heterotic groups, and intermediate for mixtures consisting of lines both within and between groups. Despite the inconsistency in the correlations, results from all these studies indicate that genetic distances based on marker genotypes are in close agreement with pedigree information and can unambiguously resolve lines into their respective heterotic groups.

Rice is the main staple for a large segment of the world population. Hybrid rice has contributed significantly to the dramatic increase of rice production in the world [27], which parallels in many ways the role of hybrid corn in the corn industry. While tremendous success has been achieved in hybrid rice breeding, experimental data pertinent to the genetic characterization of hybrid rice germplasm 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. 14, 22, 26]. There has been only one reported study addressing the relationship between isozyme diversity and hybrid performance [13], and no study has been conducted to assess such relationship using DNA markers. Thus, it can be expected that knowledge concerning the genetic characteristics of the breeding lines and relationship of genotypic heterozygosity with hybrid performance would greatly improve the efficiency of rice breeding programs.

We recently reported a molecular marker-based analysis of yield and its three component traits in a set of diallel crosses involving eight lines that are commonly used as parents in hybrid rice production [29]. This study identified several chromosomal regions which may have significant effects on yield and yield component traits. A number of agronomic characters also play major roles in the performance of hybrid rice and they are known to manifest high levels of heterosis, which warrants detailed studies. The objectives of this paper were to determine the relationship between marker locus heterozygosity and F_1 performance and heterosis for several important agronomic characters, and to assess possible heterotic relationships among the lines using both data from the marker assay and field tests.

Materials and methods

Rice lines, crosses and field experiments

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. These lines and many of their F_1 hybrids have been repeatedly tested over the years in many locations of China. Three of these eight lines, CE, MH and TQ, have been used as restorers which carry the fertility restoring genes for 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, and the F_1s and the eight parents were tested in a replicated field trial. Details of the design and cultural practices of the field experiments were described previously [29]. Briefly, 20 plants per F_1 (or parent) were transplanted into a two-row plot following a randomized complete block design [19] with each plot replicated three times. Data were collected for five traits on each plant including: (1) heading date, date of emergence of the first panicle scored as number of days after July 1; (2) plant height (cm), length of the longest tiller; (3) straw weight (g), total mass of the dried straw; (4) grain yield (g) of the whole plant; (5) biomass (g), the sum of straw weight and grain yield.

Markers and laboratory assays

Two classes of markers were used to survey DNA polymorphism among the parental lines: RFLPs and simple sequence repeats (or SSRs) which are also referred to as microsatellites. The chromosomal distribution of these markers is given in Table 1.

The procedures of DNA extraction and RFLP assay were described previously [15, 28]. DNA samples were digested singly with one to six restriction enzymes, and surveyed with each of 140 probes selected from a published rice RFLP linkage map [21]. A total of 673 probe/enzyme combinations (PE) were surveyed for RFLPs among the eight lines.

A set of 12 SSR markers – 10 developed by Wu and Tanksley [24] and 2 by Yang *et al.* [25] – were used in this study. The primers for SSR markers were synthesized by Operon Technologies, CA, according to the published sequences [24, 25]. The assay procedures were essentially as described previously [29].

Table 1. Markers used for surveying DNA polymorphisms among the eight parental lines of hybrid rice.

Chromosome	Molecular marker					
1	RZ566	RG536	RG236	RM200	RG381	RG109
	RG101	RG957	RG462	RG146A[a]	*RG811*	RG532
	RZ288	RG472	*RG400*[b]			
2	RG520	CDO1091	RG151	*RG256*	*RG89*	*RZ906*
	RG95	*RG744*	*RG324*	RG139	*RG25*	RG171
	RZ643	RZ599	RG152			
3	*RG104*	RG348	RG191	*RG450*	RG722	*RG117*
	RG335	RG69	RG393	*RG913*	RZ403	RM203
	CDO795	RM168	*RZ745*	CDO87	RZ142	*RG910*
	RM148	*RZ393*	RG163[b]			
4	RG620	RZ819	RG214	*RG122*	RG449	RZ262
	CDO456					
5	RM122	RG207	RG360	*RZ244*	RG403	RG573
	RM164	RG13	*CDO89*	RG470	RZ470	RG346
6	*CDO475*	RZ516	*RG408*	RZ398	RZ588	RZ667
	RG213	*RG138*	RG64	RG264	RG123	RG648
	RG424	*RG778*	RZ242	RZ828		
7	*RG351*	CDO38	*CDO405*	RG634	RZ395	RZ264
	RG146B[a]	*RG650*	*RG711*	*RG678*	CDO533	RG30
	RG477	*RG511*	RZ272	RG128		
8	*RZ143*	*RG597*	*RG20*	RG333	RG1034	*RZ617*
	RZ66	*RZ28*	*RG598*	RZ649	*RZ997*	
9	R45s	RZ698	RZ206	RG358	RZ228	RG667
	RG570	RG451	RZ404			
10	RZ892	*RZ561*	RZ625	RZ337	RG752	RG134
11	*RG303*	RG1109	RG2	RG167	RG1094	*RG131*
	RM120	RM167	*RG118*	*RG1022*	RZ638	RG98
	RG553[b]					
12	*RG181*	RG396	CDO344	RZ76	*RG869*	RG9
	RZ397	RZ816				
Unmapped	RM1	RM2	RM123	RM163		

Note. The markers are listed essentially in the order of appearance on each chromosome [21, 24]. Those prefixed RM are SSRs and all others are RFLP probes. The ones set in italics are monomorphic among the eight lines.

[a] RG146 is mapped to two chromosomal locations.

[b] Exact location can not be found in the map.

Statistical analysis

The marker genotypes of the F_1s were inferred from the parental genotypes. Heterozygosity of an F_1 hybrid was measured as the percent difference of marker genotypes between the two parents. The effect of a chromosomal region on a trait as marked by a molecular markers was assessed with an one-way analysis of variance using marker types (genotypes or bands) as groups and entries within marker types as the error term. Markers that detected significant effect on a trait at the 0.01 probability level were referred to as positive markers of that trait. It should be noted that this probability level does not hold experimentwise, as many of the tests are not independent of each other because of the small number of parents included in the study. The amount of hybrid vigor of a cross for each trait was evaluated using mid-parent heterosis. Cluster analyses were performed to group the eight lines using several clustering algorithms including single linkage, complete linkage [18], and Ward's method [23]. Two measurements of dissimilarities between the parents were adopted in the grouping: marker heterozygosity and mid-parent heterosis.

Results

Polymorphisms of marker loci

Polymorphisms were detected with a total of 105 markers (Table 1) including 12 SSRs and 92 RFLP clones (one clone is mapped to two chromosomal locations), with an average distance of 11.6 cM between adjacent markers. A single band per line was resolved by 11 of the 12 SSRs which was in accord to single locus variation. Multiple variable bands were detected by the twelfth SSR marker presumably due to the presence of multiple copies of the same sequence in different chromosomal locations. Among the RFLP markers, polymorphisms were detected with a total of 268 PEs. Banding patterns detected by the majority (78/92) of the probes with all enzymes were in agreement with single locus variation. Four probes detected multiple variable bands with all enzymes. Within each of the remaining 11 probes, single-band variation was detected with some enzymes, but multiple variable bands were observed with others.

PEs and SSRs whose banding patterns agreed with single-locus variation were scored as genotypes, and those showing multiple variable bands were scored by presence or absence of individual bands. Banding patterns resolved by different enzymes within a given probe were often perfectly correlated with each other. In such cases, data from one of those enzymes were used in the analysis to avoid redundant information. The same data processing scheme also applies to those PEs and SSRs whose variation was scored by bands. Consequently, the amount of data from different probes was not equal. In all, 217 pieces of non-redundant information were obtained, including 137 entries scored as genotypes and 80 scored by individual bands.

Different levels of RFLP and SSR polymorphisms were detected among the eight lines. To compare these two classes of markers, we refer to each of the above 137 non-redundant RFLP and SSR entries as a 'locus', and the variants at each locus as alleles. Thus, a majority (96/126) of the RFLP loci had 2 alleles each, with an average of 2.3 alleles per locus, whereas a much smaller proportion (5/11) of the SSR loci were diallelic with an average of 3.3 alleles per locus.

Grouping of the eight lines by RFLPs and SSRs

Data from a total of 217 non-redundant molecular marker entries were used to classify the eight lines. Cluster analyses with both complete linkage and Ward's method based on a distance measure of marker heterozygosity yielded exactly the same grouping. This analysis resolved the eight parents into three groups (Fig. 1), with MH and CE in the first group, XG, ZS and MX in the second, and GB, QS and TQ in the third. The two lines in the first group, MH and CE, are among the most widely used restorer lines, and both have large portions of their genome derived from IRRI

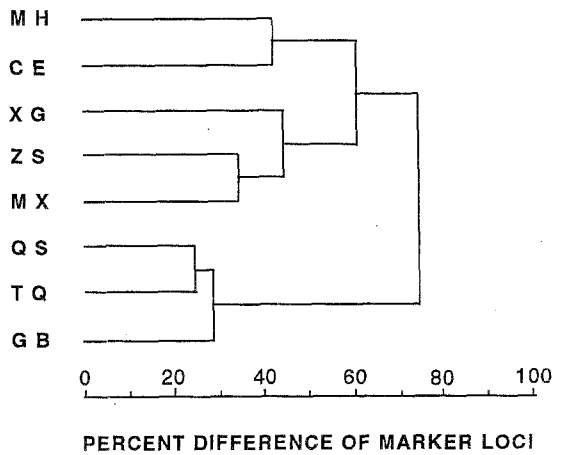

Fig. 1. Cluster dendrogram of eight commonly used parental lines of hybrid rice resolved by Ward's method using marker genotype heterozygosity as the distance measure. Abbreviations for the line names: MH, Ming Hui 63; CE, Ce 64-7; XG, Xian Gai; ZS, Zhen Shan 97; MX, Ma Xie; QS, Qing Si Ai; TQ, Te Qing; GB, Guang B.

(International Rice Research Institute) lines [8]. The three lines in the second group, XG, MX and ZS, are among the most widely used male sterile lines of South China origin, and the three lines in the third group, TQ, QS and GB, were developed by the Guangdong Academy of Agricultural Sciences. Published information [8] indicated that QS and TQ have an ancestor in common in their pedigrees, and ZS was one of the ancestors of MX. Since results from many previous studies in maize indicate that grouping based on RFLPs agrees well with the pedigree information [2, 4, 7, 10], it can be inferred that GB is closely related to QS and TQ, and as is XG to ZS and MX.

The non-redundant entries of 11 SSRs and 126 PEs of RFLPs, scored as genotypes, were used to compare the grouping efficiency of these two classes of markers. The 126 PEs resolved the same clustering structure as the total data set, while a different dendrogram was obtained using the 11 SSR loci (data not shown). The main difference is an alteration in the position of CE and TQ, such that TQ was placed with MH to form a separate group and CE was placed with GB and QS in a different group. To substantiate whether such difference in grouping is due to sampling error of small number of markers, or there is real difference between SSRs and RFLPs in grouping the rice lines, a computer simulation was performed in which the eight parents were clustered using 20 runs each consisting of 11 marker loci drawn at random from the 126 PEs of the RFLP data set. Only in 4 of the 20 cases, was the same grouping as that of the total data set obtained, indicating that the difference in grouping between the 11 SSR markers and the large RFLP data is most likely to be a result of sampling error.

F_1 performance and heterosis

There are large and highly significant differences in performance of all the five traits among the eight parents and their F_1 hybrids (Table 2). The relative ranking of these F_1s and parents agrees well with breeders' perception. It is also clear from Table 2 that these traits are highly correlated; high yielding F_1s and parental lines are usually late heading, tall and big stature, and conversely; early and short lines and hybrids are often low yielding. However, there are noteworthy exceptions. For example, the two highest yielding F_1s, derived from crosses between maintainer lines: XG × GB and MX × GB, appear to be shorter statured and headed earlier than even commercial hybrids.

Large amounts of heterosis were observed for all five traits. As can be deduced from Table 2, grain yield showed the highest heterosis followed by biomass, straw weight, plant height, and heading date displayed the lowest heterosis. Also, heterosis for grain yield, biomass and plant height was positive in almost all F_1s.

Grouping of the eight lines by mid-parent heterosis

A cluster analysis was performed to group the eight parental lines using mid-parent heterosis as the measure of dissimilarity. The overall cluster structures were exactly the same with all three algorithms (single linkage, complete linkage and Ward's method) using four characters that are

Table 2. Measurements of five traits in a set of diallel crosses involving eight parental lines of hybrid rice. See text for the abbreviations of the line names.

Cross or parent	Heading date[a]	Plant height (cm)	Straw weight (g)[b]	Grain yield (g)[b]	Biomass (g)[b]
MH × QS	35.8	107.5	115.9	145.6	261.5
MH × TQ	34.8	120.1	124.8	166.6	291.4
MH × CE	28.2	105.6	100.6	148.1	248.7
MH × XG	38.1	127.2	152.6	172.4	325.0
MH × ZS	39.6	127.6	141.7	171.1	312.8
MH × MX	41.4	124.4	155.3	176.0	331.3
MH × GB	34.7	121.5	132.1	173.6	305.7
QS × TQ	36.4	104.2	105.7	145.4	251.1
QS × CE	24.6	88.6	91.5	103.0	194.5
QS × XG	32.6	99.1	118.7	130.7	249.4
QS × ZS	39.0	108.4	113.7	132.6	246.3
QS × MX	40.3	106.8	123.3	143.3	266.6
QS × GB	32.4	104.0	105.4	135.2	240.6
TQ × CE	25.0	102.6	118.3	110.5	228.8
TQ × XG	30.6	114.8	115.4	157.6	273.0
TQ × ZS	38.1	120.9	123.4	161.2	284.6
TQ × MX	39.1	119.0	130.6	163.3	293.9
TQ × GB	33.9	114.1	105.9	146.9	252.8
CE × XG	24.7	100.2	99.6	118.1	217.7
CE × ZS	31.1	109.6	112.5	137.1	249.6
CE × MX	26.6	103.3	106.8	121.9	228.7
CE × GB	28.9	102.8	98.4	114.1	212.5
XG × ZS	24.3	102.1	81.6	104.3	185.9
XG × MX	19.9	93.0	77.2	102.7	179.9
XG × GB	31.2	114.4	122.7	176.8	299.5
ZS × MX	8.7	68.2	40.5	40.9	81.4
ZS × GB	34.5	119.0	115.3	144.0	259.3
MX × GB	37.3	116.9	133.9	178.0	311.9
MH	42.0	111.0	134.2	123.7	257.9
QS	40.9	82.2	95.0	107.5	202.5
TQ	38.0	117.9	127.8	157.1	284.9
CE	27.6	93.8	102.3	84.2	186.5
XG	23.3	96.8	73.1	74.1	147.2
ZS	8.9	67.9	38.0	35.0	73.0
MX	6.5	68.1	44.7	36.4	81.1
GB	31.0	104.9	94.0	131.3	225.3
LSD[c]	2.8	5.1	17.5	27.8	36.9

[a] Number of days after July 1.
[b] Measured on a per plant basis.
[c] Least significant difference at $p<0.01$ [19].

related to yield and plant size (plant height, straw weight, grain yield and biomass), either individually or in combinations. An example of grouping based on grain yield is illustrated in Fig. 2. The grouping using the fifth trait (heading date) differed slightly from those of the other four traits in that XG was clustered with MH, CE, QS, TQ and GB (data not shown).

Similar to the marker-based analysis, the grouping based on the four yield and size related traits also placed ZS and MX in one group, and GB, QS and TQ in another group (Fig. 2). However, there were remarkable differences between the two groupings: (1) CE, which was in a group with MH by the marker-based analysis, was closely clustered with QS, TQ and GB, and MH also joined this cluster at a higher dissimilarity level; (2) XG, which was placed in the group with MX and ZS by the marker-based analysis, became dissociated from that group, and formed a group by itself.

As illustrated in Fig. 2, this analysis revealed an unambiguous and interesting heterotic relationship among the eight parents: crosses of ZS and MX with the five lines in the other group usually produced high heterosis, and those between XG and lines in the other two groups resulted in intermediate heterosis, whereas crosses between lines within each group had a low level of heterosis (also see Table 2).

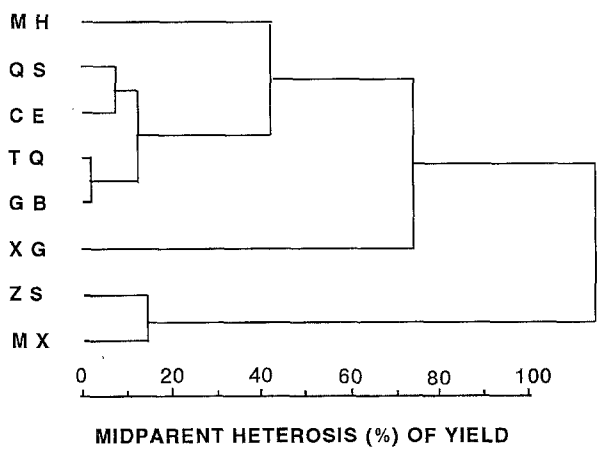

Fig. 2. Cluster dendrogram of the eight parental lines grouped by complete linkage using mid-parent heterosis of grain yield as the measure of dissimilarity. See legend of Fig. 1 for abbreviations of the line names.

Relationship of marker heterozygosity with F_1 performance and heterosis

The relationship of marker genotype heterozygosity with various attributes of performance and heterosis was evaluated using general and specific heterozygosity measurements. As described previously [29], general heterozygosity of an F_1 is the percent difference between the parents with all markers included in the study, and specific heterozygosity for a trait of an F_1 refers to the percent difference between the parents with only the positive markers of that trait as determined by the one-way analysis of variance.

The product-moment correlations between general heterozygosity and the performance of the five traits were low in general, although some of them were significantly different from zero (Table 3). The correlations of mid-parent heterosis with general heterozygosity were higher than those of performance for all five traits.

The correlations were greatly improved when specific heterozygosity for each trait was used in the calculation, especially for those between specific heterozygosity and mid-parent heterosis (Table 3). As a result, high correlations were obtained between mid-parent heterosis and specific heterozygosity for all five traits. Grain yield and biomass were used as examples to illustrate such correlations (Fig. 3), which showed that the amount of mid-parent heterosis increased linearly with the specific heterozygosity. Since F_1 measurements can be completely specified by the parental means and mid-parent heterosis, the hybrid performance can well be described with specific heterozygosity using such linear relationship along with the parental means.

Table 3. Product-moment correlations of heterozygosity with F_1 performance and mid-parent heterosis in the diallel set.

	Performance	Heterosis
Heading date	0.232/0.458[a]	0.338/0.857
Plant height	0.381/0.506	0.454/0.864
Straw weight	0.439/0.661	0.467/0.778
Grain yield	0.309/0.482	0.530/0.789
Biomass	0.374/0.464	0.516/0.895

Note. Critical values for significant correlations are 0.374 and 0.478 (26 degrees of freedom) at $p<0.05$ and $p<0.01$ respectively.

[a] Correlations based on: general heterozygosity/specific heterozygosity.

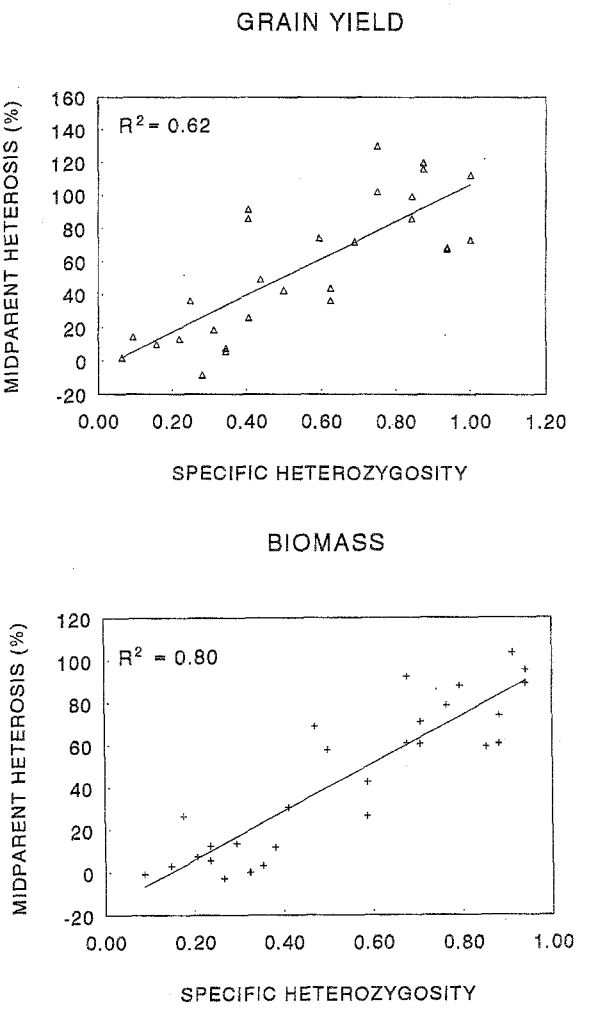

Fig. 3. Relationship between mid-parent heterosis and specific heterozygosity. Specific heterozygosity is the percentage of difference between parents based on positive markers of each trait.

Discussion

We have surveyed genetic divergence using RFLP and SSR markers among eight commonly used parental lines of hybrid rice and assessed the

relationship between molecular marker polymorphisms and hybrid performance in the diallel set involving these eight parents. We observed that about one third of the previously mapped RFLP probes failed to detect polymorphism in these eight parents, whereas all the 12 SSRs detected polymorphism. Moreover, more alleles were detected by SSRs than RFLPs on the basis of per variable locus. Thus, as a class of markers, SSRs detected both higher proportion of polymorphic loci and larger number of alleles per locus than did RFLPs. This is also likely to be the case across a wide range of rice germplasm, as similar results have been observed previously [24, 25].

The analyses have demonstrated two features concerning the correlations of marker genotype heterozygosity with F_1 performance and heterosis. First, correlations of heterozygosity with mid-parent heterosis of all five traits were higher than those with F_1 performance, indicating that mid-parent heterosis may be better predicted than performance by marker genotypes. And second, correlations calculated using specific heterozygosity based on the small number of positive markers are much larger than those using general heterozygosity which is based on the total molecular markers in the study. This indicates that a subset of informative markers may much better predict hybrid performance than a large number of random markers, corroborating the suggestion from theoretical calculations [1, 3]. It should be noted, in particular, that the correlations between the specific heterozygosity and mid-parent heterosis are high for all the five traits, and, statistically, such large correlations have reached the level of being practically useful for predictive purposes.

Saghai Maroof *et al.* [16] used a marker-based diallel analysis to detect the existence of genes for powdery mildew resistance in barley, and found a remarkable agreement in the chromosomal locations between positive markers and resistance genes that had been reported in previous studies. This in turn suggests that marker-based analysis may be useful for detecting the existence of genes controlling the traits of interest. It is therefore highly likely that a large portion of the positive markers identified by the analyses of variance in the present study correspond to the chromosomal regions containing the genes of interest, although many of the tests may not be independent of each other because of the small number of parents included in the diallel cross. It should be noted, however, some of the parents are closely related to each other which differs from the situations in screening for heterotic crosses in hybrid breeding. Thus, it remains to be determined how many of these markers will still be informative in other data sets, and whether such correlations will still be valid across different types of germplasm.

Melchinger *et al.* [9] and Boppenmaier *et al.* [2] expressed doubt on the usefulness of increased genome coverage with additional markers for increasing the correlations between marker distance and hybrid performance to improve the efficiency of the prediction. They alternatively suggested identification of marker loci and genotypes significantly associated with traits of interest. In our study, the 157 pieces of non-redundant information included in the previous analysis [29] and the 217 non-redundant entries in the present analysis provided an assessment of the effect of increased genome coverage with additional markers. Taking grain yield as an example, adding such a large number of markers did not increase (actually decreased) the correlations of general heterozygosity with F_1 performance and heterosis. However, this increased number of markers has led to an increase in the correlations of specific heterozygosity with F_1 performance and heterosis, presumably due to the inclusion of additional positive markers. This suggested that although increasing the number of markers did not directly contribute to the correlations between general heterozygosity and hybrid performance, it did add to the totality of positive markers which provide useful information for prediction purposes.

The analysis of molecular marker data resolved these eight lines into three well separated groups, which agree essentially with the available pedigree information. Inspection of the heterotic patterns among the eight parents indicated that the two lines, ZS and MX (perhaps the third one, XG, as

well) in the second group, can produce high level of heterosis when crossed to lines in the other two groups. However, intercrossing lines of the other two groups did not show very much heterosis. Furthermore, both TQ and QS have certain amounts of germplasm from landraces of Southeast Asia origin according to the available pedigree information. Although the marker-based analysis indicates that these two lines are not related to MH and CE (both of which had germplasm from IRRI lines), they display similar heterotic pattern to that of MH and CE when crossed to other lines. Thus, these results seem to suggest the existence of two likely heterotic groups: rice strains of Southeast China origin in one group and ecotypes in certain parts of Southeast Asia in the other group. This likely heterotic relationship agrees well with breeders' perception. In fact, this is indeed the parental configuration of many commercial hybrid combinations which often have a short statured male sterile line of South China origin, and a medium height restorer line containing certain proportion of germplasm derived from Southeast Asian rice strains. However, detailed studies are necessary to determine the extent and the genetic characteristics of such heterotic relations.

In summary, results from this study suggest that molecular marker based analysis provides a useful means for germplasm characterization. Correlations between marker heterozygosity on the basis of a subset of informative markers and mid-parent heterosis may have a practical utility in predicting hybrid performance in rice. Molecular markers may also help identify heterotic groups when combined with the heterotic patterns observed in field tests.

Acknowledgements

The authors thank Dr S.D. Tanksley's group for kindly providing the rice RFLP clones. This research was supported in part by a grant from the Rockefeller Foundation and a grant from the Education Commission of PRC.

References

1. Bernardo R: Relationship between single-cross performance and molecular marker heterozygosity. Theor Appl Genet 83: 628–634 (1992).
2. Boppenmaier J, Melchinger AE, Seitz G, Geiger HH, Herrmann RG: Genetic diversity for RFLPs in European maize inbreds. III. Performance of crosses within versus between heterotic groups for grain traits. Plant Breed 111: 217–226 (1993).
3. Charcosset A, Lefort-Buson M, Gallais A: Relationship between heterosis and heterozygosity at marker loci: a theoretical computation. Theor Appl Genet 81: 571–575 (1991).
4. Dudley JW, Saghai Maroof MA, Rufener GK: Molecular markers and grouping of parents in maize breeding programs. Crop Sci 31: 718–723 (1991).
5. Godshalk EB, Lee M, Lamkey KR: Relationship of restriction fragment length polymorphisms to single-cross hybrid performance of maize. Theor Appl Genet 80: 273–280 (1990).
6. Hallauer AR, Miranda JB: Quantitative Genetics in Maize Breeding. Iowa State University Press, Ames (1988).
7. Lee M, Godshalk EB, Lamkey KR, Woodman WW: Association of restriction fragment length polymorphisms among maize inbreds with agronomic performance of their crosses. Crop Sci 29: 1067–1071 (1989).
8. Lin S, Ming S: Rice Varieties and Their Genealogy in China. (In Chinese.) Shanghai Science and Technology Press, Shanghai (1991).
9. Melchinger AE, Lee M, Lamkey KR, Woodman WL: Genetic diversity for restriction fragment length polymorphisms: relation to estimated genetic effects in maize inbreds. Crop Sci 30: 1033–1040 (1990).
10. Messmer MM, Melchinger AE, Herrmann RG, Boppenmaier J: Relationships among early European maize inbreds. II. Comparison of pedigree and RFLP data. Crop Sci 33: 944–950 (1993).
11. Moll RH, Lonnquist JH, Vélez Fortuno J, Johnson EC: The relationship of heterosis and genetic divergence in maize. Genetics 52: 139–144 (1965).
12. O'Brien SJ: Genetic Maps, Book 6, Plants. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1992).
13. Peng JY, Glaszmann JC, Virmani SS: Heterosis and isozyme divergence in indica rice. Crop Sci 28: 561–563 (1988).
14. Rice Research Group of Mian Yang Agricultural Institute of Sichuan Province: an analysis on heterosis of F_1 hybrids in rice. Acta Sin Genet 5: 158–164 (1978).
15. Saghai Maroof MA, Soliman KM, Jorgensen RA, Allard RW: Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomallocation, and population dynamics. Proc Natl Acad Sci USA 81: 8014–8018 (1984).

16. Saghai Maroof MA, Zhang Q, Biyashev RM: Molecular marker analyses of powdery mildew resistance in barley. Theor Appl Genet 88: 733–740 (1994).
17. Smith OS, Smith JSC, Bowen SL, Tenborg RA, Wall SJ: 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 (1990).
18. Sokal R, Sneath PHA: Principles of Numerical Taxonomy. Freeman, San Francisco/London (1963).
19. Steel RGD, Torrie JH: Principles and Procedures of Statistics. McGraw-Hill, New York (1980).
20. Stuber CW: Molecular markers in the manipulation of quantitative characters. In: Brown AHD, Clegg MT, Kahler AL, Weir BS (eds) Plant Population Genetics, Breeding, and Genetic Resources, pp. 334–350. Sinauer Associates, Sunderland, MA (1989).
21. Tanksley S, Causse M, Fulton T, Ahn N, Wang Z, Wu K, Xiao J, Yu Z, Second G, McCouch S: A high density molecular map of the rice genome. Rice Genet Newsl 9: 111–115 (1992).
22. Virmani SS, Aquino RC, Khush GS: Heterosis breeding in rice. Theor Appl Genet 63: 373–380 (1982).
23. Ward JH: Hierarchical grouping to optimize an objective function. J Am Stat Ass 58: 236–244 (1963).
24. Wu KS, Tanksley SD: Abundance, polymorphism and genetic mapping of microsatellites in rice. Mol Gen Genet 241: 225–235 (1993).
25. Yang GP, Saghai Maroof MA, Xu CG, Zhang Q, Biyashev RM: Comparative analysis of microsatellite DNA polymorphism in landraces and cultivars of rice. Mol Gen Genet, in press (1995).
26. Young JB, Virmani SS: Heterosis in rice over environments. Euphytica 51: 87–93 (1990).
27. Yuan LP: Development and prospects of hybrid rice breeding. In: You C, Chen ZL (eds) Agricultural Biotechnology. Proceedings Asia-Pacific Conference on Agricultural Biotechnology, Beijing, August 20–24, 1992, pp. 97–105 (1992).
28. Zhang Q, Saghai Maroof MA, Lu TY, Shen BZ: Genetic diversity and differentiation of *indica* and *japonica* rice detected by RFLP analysis. Theor Appl Genet 83: 495–499 (1992).
29. Zhang Q, Gao YJ, Yang SH, Ragab R, Saghai Maroof MA, Li ZB: A diallel analysis of heterosis in elite hybrid rice based on RFLPs and microsatellites. Theor Appl Genet, in press (1995).