Epistasis for Three Grain Yield Components in Rice (*Oryza sativa L.*)

Zhikang Li,*,† Shannon R. M. Pinson,‡ William D. Park,§ Andrew H. Paterson* and James W. Stansel†

**Department of Soil and Crop Sciences and §Department of Biochemistry and Biophysics, Texas A&M University, College Station, Texas 77843, †Texas A&M University Agricultural System Research and Extension Center, Beaumont, Texas 77713 and ‡United States Department of Agriculture, Agricultural Research Service, Beaumont, Texas 77713*

Manuscript received June 10, 1996
Accepted for publication September 27, 1996

ABSTRACT

The genetic basis for three grain yield components of rice, 1000 kernel weight (KW), grain number per panicle (GN), and grain weight per panicle (GWP), was investigated using restriction fragment length polymorphism markers and F_4 progeny testing from a cross between rice subspecies *japonica* (cultivar Lemont from USA) and *indica* (cv. Teqing from China). Following identification of 19 QTL affecting these traits, we investigated the role of epistasis in genetic control of these phenotypes. Among 63 markers distributed throughout the genome that appeared to be involved in 79 highly significant ($P < 0.001$) interactions, most (46 or 73%) did not appear to have "main" effects on the relevant traits, but influenced the trait(s) predominantly through interactions. These results indicate that epistasis is an important genetic basis for complex traits such as yield components, especially traits of low heritability such as GN and GWP. The identification of epistatic loci is an important step toward resolution of discrepancies between quantitative trait loci mapping and classical genetic dogma, contributes to better understanding of the persistence of quantitative genetic variation in populations, and impels reconsideration of optimal mapping methodology and marker-assisted breeding strategies for improvement of complex traits.

EPISTASIS, or interactions between nonallelic genes, is an important factor that affects phenotypic expression of genes and genetic variation in populations. Gene interaction is the core of WRIGHT's theory of the genetic basis for evolution (WRIGHT 1932, 1951) and plays a key role in founder effect models of speciation (TEMPLETON 1979, 1980). While molecular evidence of interacting genes influencing development has been accumulated at a rapid rate, our knowledge of how epistatic genes influence quantitative phenotypes remains incomplete because of the complexities of studying quantitative traits. Nevertheless, a considerable body of classical evidence has strongly suggested prevalence of epistasis affecting quantitative traits in populations (SPICKETT and THODAY 1966; FALCONER 1981; MATHER and JINKS 1982; POONI *et al.* 1987; ALLARD 1988).

Quantitative trait loci (QTL) mapping using DNA markers has improved our understanding of the genetic basis of quantitative traits. Many studies have identified QTL in plants and animals (EDWARDS *et al.* 1987, 1992; STUBER *et al.* 1987, 1992; PATERSON *et al.* 1988, 1990, 1991, 1995a,b; FATOKUN *et al.* 1992; DEVICENTE and TANKSLEY 1993; KOWALSKI *et al.* 1994; ANTHONY *et al.* 1995; GEORGES *et al.* 1995; HORVAT and MEDRANO 1995; LI *et al.* 1995a,b). However, QTL mapping studies to date have also raised several puzzling questions. First, only a very limited number of chromosomal regions or QTL, each having relatively large phenotypic effects, have been identified for most quantitative traits regardless of the complexity of traits and amount of genetic variation in mapping populations. Second, identified QTL often explained only a portion of the total variation despite nearly complete genome coverage by genetic markers. These results deviate from what would be expected from classical quantitative genetic theory and numerous selection studies that suggest that a large number of genes each having a small effect are involved in quantitative genetic variation (*cf.* MATHER and JINKS 1982). Third, while different QTL for the same trait generally do not interact with each other, the total genotypic variance explained by all QTL for the same trait is generally much smaller than the sum of genotypic variances explained by different QTL. This contradicts the statistical prediction of additivity under independence. Factors such as environmental influences, physiological pleiotropy, QTL with effects too small to detect, covariances between closely linked QTL, and epistasis may partially explain the observed dilemmas (PATERSON *et al.* 1991; ZENG 1993; COCKERHAM and ZENG 1996).

Although most QTL mapping studies reveal little evidence for the presence of epistasis between QTL (EDWARDS et al. 1987, 1992; PATERSON *et al.* 1988; 1991; FATOKUN *et al.* 1992; STUBER *et al.* 1992; DEVICENTE and TANKSLEY 1993; LI *et al.* 1995a,b), "genetic background effects" on quantitative traits have been well documented in Drosophila (SPASSKY *et al.* 1965), tomato

Corresponding author: Zhikang Li, Department of Soil and Crop Science, Texas A&M University, College Station, TX 77843-2474.
E-mail: zli@tam2000.tamu.edu

(TANKSLEY and HEWITT 1988), rice (KINOSHITA *et al.* 1982; SATO and SAKAMOTO 1983), and most recently in soybean (LARK *et al.* 1995) and maize (DOEBLEY *et al.* 1995; COCKERHAM and ZENG 1996). These observations suggest that there exist epistatic loci that may affect the expression of genes or QTL, causing background effects.

Epistasis is generally thought to be responsible for reduced fitness in hybrid progenies between related species, since genes causing problems in hybrids are apparently conferring normal functions in parental species. In crop improvement, an important fact that is observed by plant breeders is that the yield of most progenies from a cross between any two elite parents of the same species or subspecies is lower than the parents, suggesting epistasis may also be an important genetic basis for yield and its components in crop plants.

We here report detection of epistasis affecting three grain yield components of rice using replicated progeny testing and a complete restriction fragment length polymorphism (RFLP) linkage map. Our primary objectives were to reassess the prevalence of digenic interactions, identify specific genomic regions involved in digenic interactions, and study patterns by which alleles at two loci interact, to better reconcile classical evidence of epistasis with measured gene actions of QTL. Our studies differed from prior studies in the level of replication associated with calculation of breeding value of each F_2 plant, and also introduce new approaches to separating interactions from sampling variation.

MATERIALS AND METHODS

Plant materials: *Oryza sativa* L. cvs., "Lemont" and "Teqing" were used as parents in the study. The female parent, Lemont, belongs to *japonica* varietal group. The male parent, Teqing, is a typical semidwarf *indica* variety from China. From each of 255 F_2 plants from the Lemont/Teqing cross, seven to 11 (10 in most of the cases) F_3 plants from each of the F_2 plants were randomly selected and used to produce a total of 2418 F_2-derived F_4 lines for the phenotyping experiment, as described previously (LI *et al.* 1995a).

RFLP marker genotyping and the field experiment: Reconstruction of genotypes of the original 255 F_2 plants for 113 RFLP and two morphological markers (*C* and *gl-1*) using their derived F_3 plants were described previously (LI *et al.* 1995a). The resulting 115 loci were spaced at an average 19.1 cM across the 12 rice chromosomes. Phenotyping of the 2418 F_4 lines was conducted in 1990 at the Texas A&M University System Agricultural Research and Extension Center in Beaumont, as described (LI *et al.* 1995a). For assay of grain yield components, 10–13 panicles (each from a single plant) were collected in each of the 2418 F_4 lines and dried at 50° for 72 hr. Panicles from each of the F_4 lines were hand threshed and measured for the number of grains per panicle (GN), 1000 kernel weight (KW, two samples of 200 grains were weighed and converted to 1000 grain weight), and grain weight per panicle (GWP). The "breeding values" of each of the 255 F_2 plants for these traits, each calculated as the mean value of 70–130 observations from 10 F_4 lines derived from each of the 255 F_2 plants, were used in the data analyses.

Statistical model for detecting digenic epistasis: For the simplest case of digenic epistasis affecting a quantitative trait in an unreplicated F_2 mapping population, the most commonly used linear model in a two-way ANOVA (EDWARDS *et al.* 1987; STUBER *et al.* 1987, 1993; PATERSON *et al.* 1988; DEVICENTE and TANKSLEY 1993; XIAO *et al.* 1995) can be shown as follows:

$$y_{ijm} = \mu + \alpha_i + \alpha_j + \tau_{ij} + \epsilon_{ijm}, \quad \text{for } m = 1,2, \ldots, n_{ij}, \quad (1)$$

where y_{ijm} is the phenotype of the mth F_2 plant with the digenic genotype at loci i and j, α_i and α_j are the main effects (if any, which include the additive and the dominance effects) associated with the loci i and j, respectively; τ_{ij} are the effects arising from interactions between alleles at the loci i and j, and ϵ_{ijm} is the residual effect including the genetic effect unexplained by the two loci in the model plus measurement errors and other factors, which is assumed to be an independent random variable having a normal distribution with zero mean and a variance of σ^2. The expected mean squares from different sources in the model are given in Table 1. In the full model, three hypotheses are tested (the main effects associated with two loci and the interactions between them) where i and j are markers near QTL. To detect genomic regions that have epistatic effects on quantitative traits, this model is extended to cover random genomic loci. To assay the whole genome for digenic epistatic effects on quantitative traits, one has to evaluate $n(n-1)/2$ possible interactions for a map of n markers.

Data analyses: Three steps in data analyses were taken to identify epistasis. First, we conducted conventional QTL mapping to identify QTL affecting the three traits using interval mapping (LANDER and BOTSTEIN 1989). A LOD $\geq$ 2.4 was used as a threshold for claiming the presence of putative QTL. Then, all putative QTL were further confirmed using multi-QTL models of interval mapping (MapMaker/QTL) and multiple regression (ZENG 1993, 1994) until all QTL in the model were highly significant ($P < 0.001$ in a multiple regression model, and each QTL contributed to the multi-QTL model by an additional LOD $\geq$ 2.0 in the interval mapping).

Second, we selected a subset of 95 representative codominant DNA markers from the original linkage map (LI *et al.* 1995) and conducted all possible two-way ANOVA between these markers using the model (1) and SAS PROC GLM (SAS Institute 1987) (dominant markers and those with more than 35 missing data were not used). The threshold to claim a statistically significant interaction was $P \leq 0.001$ (F tests with a degree of freedom of 4), and $R^2_{\text{interaction}} \geq 5\%$. For cases when two or more significant interactions were found to be due to linkage [by examining the F-statistic profiles along the two genomic regions and the corresponding digenic effects τ_{ij} (see the following section)], the one with the highest F and R^2 values was retained.

Third, to remove false positive interactions due to the background genetic effects arising from segregating QTL, multiple regression with all QTL (identified in the first step) fixed in the model was utilized to reanalyze the highly significant interactions detected in two-way ANOVA, based on the following model:

$$y_{ijm} = b_0 + \sum_k b_k x_{mk} + b_i x_{mi} + b_j x_{mj} + b_{mij} x_{ij} + \epsilon_{ijm}$$

$$\text{for k} = 1, 2, \ldots k, \; m = 1,2, \ldots, n_{ij}, \quad (2)$$

where y_{ijm} is the trait value of the individuals with the same digenic genotype at marker loci i and j ($i, j = 1, 2, 3$), b_0 is the mean of the model, b_k is the partial regression coefficient of the phenotype on the kth QTL (or the main effects of the kth QTL), b_i, b_j, and b_{ij} [equivalent to α_i, α_j, and τ_{ij} in the model (1)] are partial coefficients (the main effects, if any,

TABLE 1

Expected mean squares for digenic interactions between unlinked marker loci

Sources of variation	d.f.[a]	Expected mean square
Between marker i genotypes	$n_i - 1$	$\sigma^2 + 27 \sum_{i=1}^{n_i} [\alpha_i^2/(n_i - 1)]$
Between marker j genotypes	$n_j - 1$	$\sigma^2 + 27 \sum_{j=1}^{n_j} \{\alpha_j^2/(n_j - 1)\}$
Interaction between markers i and j	$(n_i - 1)\ (n_j - 1)$	$\sigma^2 + 9 \sum_{i=1,j=1}^{n_{ij}} \{\tau_{ij}^2/(n_i - 1)\ (n_j - 1)\}$
Residual	$N - n_{ij} - 1$	σ^2

[a] $N = \sum_{i=1}^{3} n_i = \sum_{j=1}^{9} n_j = \sum_{m=1}^{9} n_{ij}$ is the number of F_2 plants in the experiment.

and the interaction effects) of phenotype y on the ith and jth markers conditional on all k QTL, and ϵ_{ijm} is the residual, which is assumed to be an identically and independently distributed variable with zero mean and a variance of σ^2. Since the interaction effects b_{ij} are tested conditional to those detected QTL, not only the background genetic effects from nonrandom sampling of the QTL will be effectively controlled, but the power to detect epistasis can also be increased (ZENG 1993, 1994), provided that the QTL are indenpendent from one another. Interaction effects highly significant in both two-way ANOVA and multiple regression analyses were considered to be due to epistasis and used for further analyses.

The maximum likelihood estimates of interaction effects τ_{ij} involved in each interaction highly significant in above analyses (with and without QTL included) were obtained using the mean of the nine digenic genotypes (GRAYBILL 1976) in which $\hat{\tau}_{ij} = \hat{\mu}_{ij} - \hat{\mu}_{i.} - \hat{\mu}_{.j} + \hat{\mu}$. A t test was performed to test the Ho: $\tau_{ij} = 0$ using $\sqrt{\frac{\hat{\sigma}^2}{n_{ij}}}$ as standard errors, where $\hat{\sigma}^2$, the residual variance, was approximated by averaging the 4465 observed s_{ij}^2 (the observed interaction variance component), which is an empirical estimate of $\hat{\sigma}^2$, of the experiment, and n_{ij} was the observed sample size of individual digenic genotypes.

RESULTS

Trait means, variation, heritability, and correlation: The differences between the parents for the three traits were small (significant at $P = 0.05$ for KW, but not significant for GN and GWP based on t tests). Lemont had KW of 23.3 ± 0.20 g, GN of 137.5 ± 23.1, and GWP of 31.9 ± 5.2 g. Teqing had KW of 24.0 ± 0.16 g, GN of 148.0 ± 23.6, and GWP of 35.5 ± 5.6 g. The mean of the F_2 breeding values showed only normally lower KW (22.9 g) but significantly ($P < 0.01$) lower GN and GWP (112.2, and 25.4 g) than the parents.

The F_4 lines showed tremendous variation for the three traits. The phenotypic values of the 2418 F_4 lines and the breeding values of the 255 F_2 plants were approximately normally distributed for KW, GN and GWP (Figure 1), typical of polygenic inheritance. KW showed transgressive segregation in both directions, ranging from 14.0 to 35.7 g in the F_4 lines and from 17.2 to 29.5 g in the F_2 breeding values, respectively. GN and GWP also showed transgressive segregation but primarily toward reduced GN and GWP. Such a reduction in yield traits is typical of progenies from crosses between distantly related rice varieties such as *indica/japonica*

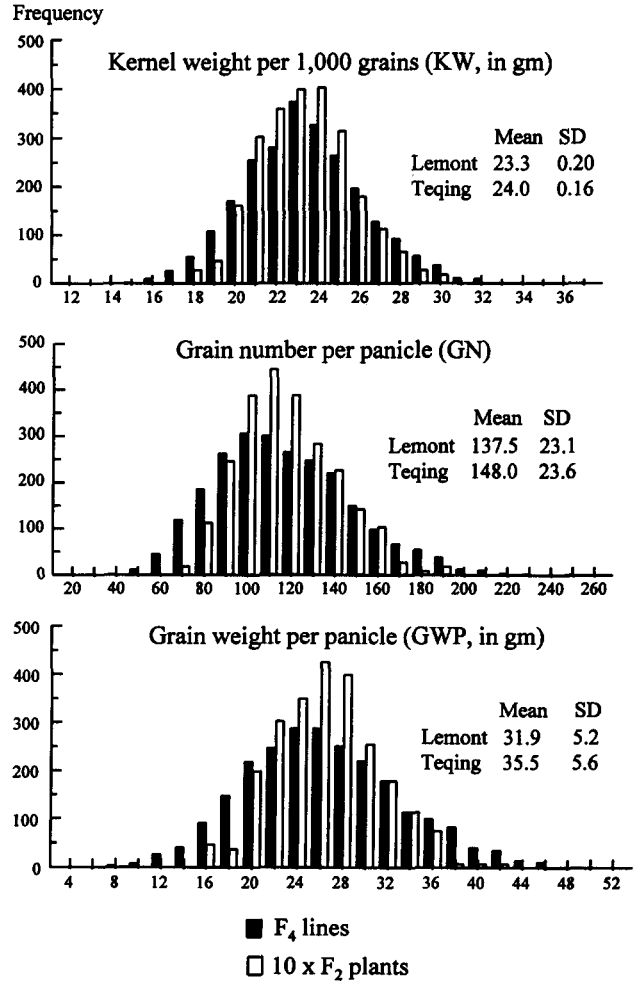

FIGURE 1.—Frequency distribution of the phenotypic values of 2418 F_4 rice lines and the breeding values of 255 F_2 plants (10 times) of the Lemont/Teqing cross for 1000 KW, GN, and GWP.

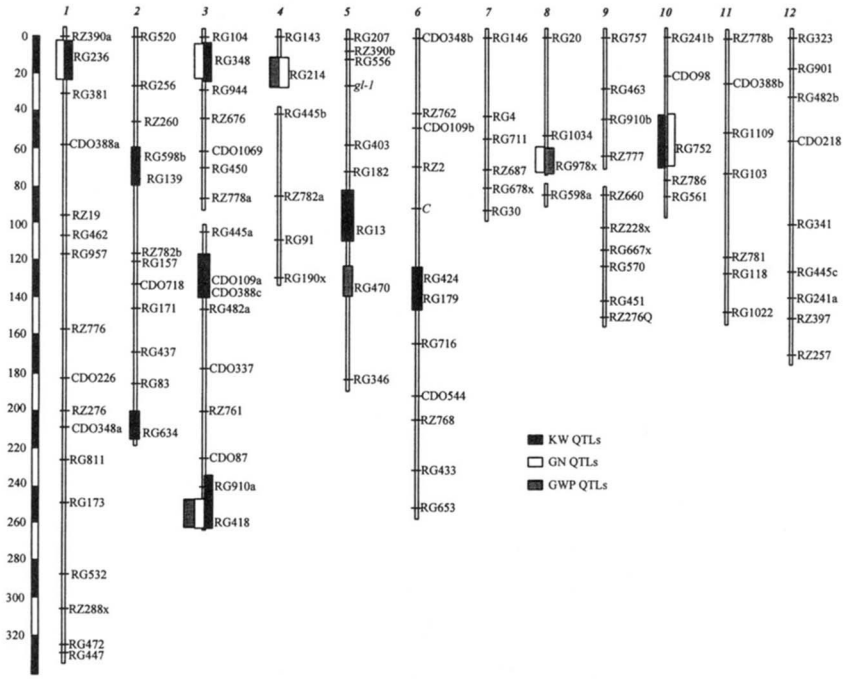

FIGURE 2.—The RFLP map with 115 markers constructed from 255 F_2 plants of the Lemont/Teqing rice cross. The boxes indicate the genomic locations (1 LOD confidence intervals) for 19 QTL identified for KW, GN, and GWP.

crosses. KW had a heritability of 0.877, while GN and GWP had lower heritabilities of 0.535 and 0.413, respectively. GN and GWP were positively correlated (r_G = 0.865, $P < 0.0001$). KW was negatively correlated with GN ($r_G = -0.461$, $P < 0.0001$), but uncorrelated with GWP ($r_G = 0.014$).

Identification of QTL affecting KW, GN and GWP: Table 2 and Figure 2 show 19 QTL (LOD > 2.4) influencing the three traits, identified by interval mapping and confirmed by multiple regression analyses.

QTL affecting KW: Eight QTL affecting KW were identified in this population using interval mapping based on LOD > 2.4, all confirmed by multiple regression analyses. In addition, four genomic regions showed effects on these traits that fell slightly below the threshold (2.4 > LOD ≥ 2.0). Of these, three became insignificant by multiple regression analyses, one on chromosome *2* (*QKw2b*) became highly significant in the multiple regression models. The nine QTL, located on chromosomes *1, 2, 3, 5, 6,* and *10,* collectively accounted for ~84.6% of genotypic variation (the variance among the breeding values of the F_2 plants). These QTL had additive effects ranging from 0.56 to 1.51 g, and explained 4.1–16.3% of the genotypic variation in the F_2 plants (Table 2). The alleles at five QTL (*QKw1a, QKw2a, QKw3b, QKw3c,* and *QKw6*) for increased KW were from Lemont, and the remaining four (*QKw2b, QKw3a, QKw5,* and *QKw10*) from Teqing. Large dominance effects at *QKw1a, QKw2a* and *QKw5* were associated with reduced KW, but with increased KW at *QKw2b* and *QKw10.*

GN: Six QTL were mapped to chromosomes *1, 3, 4, 8,* and *10,* collectively accounted for 37.8% of the genotypic variance. These QTL had additive effects ranging from 0.8 to 11.9 grains per panicle and explained 4.8–11.8% of the genotypic variance. Four alleles (*QGn3a, QGn4,* and *QGn10*) for increased GN were from Lemont, and two (*QGn1* and *QGn3b*) from Teqing. *QGn8* did not have an appreciated additive effect, but showed overdominance for increased GN. Large dominance for increased GN was also observed at *QGn4,* but for reduced GN at *QGn3a* and *QGn10.*

GWP: Two definite QTL (*QGwp3* and *QGwp4*) influencing GWP were mapped to chromosomes *3* and *4.* When these QTL were fixed, we were able to detect two additional QTL (*QGwp5* and *QGwp8*) on chromosomes *5* and *8* using both interval mapping and multiple regression. Collectively, the four QTL accounted for 22.7% of the genotypic variance. The alleles for increased GWP at *QGwp3, QGwp5* and *QGwp8* were from Teqing, and that at *QGwp4* was from Lemont, with additive effects ranging from 0.58 to 1.79 g per panicle. Dominance effects at all QTL were associated with increased GWP, with overdominance detected at *QGwp4* and *QGwp8.*

Pleiotropic effects of QTL: The three GWP QTL (*QGwp3, QGwp4,* and *QGwp8*) were related to the GN QTL, *QGn3b, QGn4* and *QGwp8.* In all cases, alleles for increased GN were associated with increased GWP, in agreement with the high positive correlation between GN and GWP. Four GN QTL *QGn1, QGn3a, QGn3b,* and *QGn10* were mapped to approximately the same positions as KW QTL *QKw1a, QKw3a, QKw3c,* and *QKw10.* In all these cases, the additive effects of GN

TABLE 2

QTL affecting 1000 KW, GN, and GWP in the Lemont × Teqing cross

QTL[a]	Flanking markers[b]	a[c]	d	R^2_G %	LOD
QKw1	RG236–RG381	−0.86	−2.32	8.1	3.46
QKw2a	RG598b–RG139	−0.84	−1.94	6.6	3.33
QKw3a	RG104–RG348	0.80	0.16	5.1	2.88
QKw3b	RG445a–CDO109a	−1.07	−0.68	11.7	6.82
QKw3c	RG910a–RG418	−0.95	0.10	7.2	3.57
QKw5	RG182–RG13	1.51	−1.80	16.3	5.10
QKw6	RG424–RG179	−0.87	−0.10	6.0	3.09
QKw10	CDO98–RG752	1.07	1.44	10.3	3.77
QKw2b	RG83–RG634	0.56	2.48	4.1	2.59
R^2_G % from multi-QTL model[d]				84.6	
QGn1	RG236–RG381	7.2	6.4	4.8	2.64
QGn3a	RG104–RG348	−7.3	−15.0	5.7	3.20
QGn3b	RG910a–RG418	11.9	1.2	11.8	6.70
QGn4	RG143–RG214	−10.2	20.5	8.8	4.93
QGn8	RG1034–RG978	−0.8	40.8	6.0	2.93
QGn10	CDO98–RG752	−9.9	−18.4	9.7	2.46
R^2_G % from multi-QTL model[d]				37.8	
QGwp3	RG910a–RG418	1.71	0.66	5.9	3.25
QGwp4	RG143–RG214	−1.79	4.59	6.8	3.89
QGwp5	RG13–RG470	1.60	1.08	5.7	2.46
QGwp8	RG1034–RG978	0.58	7.65	4.3	2.64
R^2_G % from multi-QTL model[d]				22.7	

[a] Individual QTL are designated with Q indicating a QTL, abbreviation of the trait name and the chromosome number. When more than one QTL affecting a trait was identified on the same chromosome, they are distinguished by different letters.

[b] The underlined markers are those closer to the QTL.

[c] a is the additive effect due to substitution of a Lemont allele by the corresponding Teqing allele. The d is the dominance effect associated with the heterozygote. Units for KW and GWP are grams.

[d] R^2 was obtained by running the model with all QTL (LOD > 2.4) included.

QTL were negatively associated with those of the corresponding KW QTL, which also agreed well with the negative correlation between GN and KW.

Detection of digenic epistasis: Of 4465 possible two-way ANOVAs between the 95 selected markers, we detected 81, 79, and 79 highly significant ($P < 0.001$, $R^2_{\text{interaction}} > 5\%$) interactions for KW, GN and GWP, respectively. A portion (34.5, 26.6, and 35.4% for KW, GN and GWP, respectively) of the interactions significant at $P \leq 0.001$ involved groups of linked markers. After removing interactions that were due to linkage effects, the remaining number of interactions significant at $P \leq 0.001$ was 57 for KW, 58 for GN, and 55 for GWP, respectively.

Interactions due to main effects of the segregating QTL: In the present study, because the sample size for each of the digenic genotypes was relatively small and there were a number of QTL segregating for each of the traits in the mapping population, statistically significant interactions may have arisen from the nonrandom sampling of segregating QTL. When these highly significant interactions were reanalyzed by the multiple regression analyses in which the QTL (each represented by its most closely linked marker) were fixed in models, 41 of the 57 interactions affecting KW, 30 of the 58 interactions affecting GN, and 20 of the 55 interactions influencing GWP became insignificant, while all QTL remained highly significant in the model. This indicated that a high probability of false positive interactions detected by two-way ANOVA arose from the main effects of the segregating QTL. The significance level of 16 interactions affecting KW, 28 interactions affecting GN, and 35 interactions influencing GWP remained largely unchanged or even became more significant, and all occurred between unlinked markers.

Decomposition of digenic interactions: In our experimental design, the variance component resulting from interaction effects $\hat{\tau}_{ij} = (\hat{\mu}_{ij} - \hat{\mu}_{i.} - \hat{\mu}_{.j} + \hat{\mu})$ between alleles at two loci can be partitioned into three subcomponents: one degree of freedom due to interaction effects between homozygotes ($\hat{\tau}^2_{11} + \hat{\tau}^2_{12} + \hat{\tau}^2_{21} + \hat{\tau}^2_{22}$), a second and third degrees of freedom between a homozygote at one locus and the heterozygote at the other ($\hat{\tau}^2_{13} + \hat{\tau}^2_{23} + \hat{\tau}^2_{31} + \hat{\tau}^2_{32}$), and a fourth degree of freedom between heterozygotes ($\hat{\tau}^2_{33}$). Based on the quantitative genetic model (MATHER and JINKS 1982), each of the interaction effects ($\hat{\tau}_{11}$, $\hat{\tau}_{12}$, $\hat{\tau}_{21}$, and $\hat{\tau}_{22}$) between the homozygotes in our experiment design, consisted primarily of the additive component of epistasis [for instance, the genetic expectation for $\hat{\tau}_{11}$ and $\hat{\tau}_{22}$ is $i_{ab} - [{}^1/_8(j_{ab} + j_{ba}) + {}^1/_{64}l_{ab}]$, while the interaction effects ($\hat{\tau}_{13}$, $\hat{\tau}_{23}$, $\hat{\tau}_{31}$, and $\hat{\tau}_{32}$) between heterozygotes and homozy-

TABLE 3

Decomposition of interaction variances affecting 1000 KW, GN, and GWP into between homozygotes (*aa*), homozygote × heterozygote (*ad*), and heterozygote × heterozygote (*dd*) variance components

Trait	R^2	V_{aa}	N_{aa}	V_{ad}	N_{ad}	V_{dd}	N_{dd}
KW							
Mean	6.50	66.2	18.8	30.9	47.8	3.0	33.4
SD	0.89	13.2	2.4	11.9	4.0	3.4	4.1
GN							
Mean	7.02	64.4	19.4	31.8	47.0	3.8	33.6
SD	1.63	12.8	4.0	11.5	4.5	3.0	6.1
GWP							
Mean	6.46	65.0	18.6	31.6	47.9	3.4	33.8
SD	1.21	12.9	4.1	12.2	3.8	2.9	5.7

[a] V and N are percentages of variance components and frequencies of their contributions to F_2 plants. R^2 is also in percentage.

gotes estimated only one-eighth of the nonadditive components of epistasis (j_{ab} or j_{ba}). The genetic expectation for $\hat{\tau}_{33}$ was only $^1/_{64}l_{ab}$.

In agreement with the genetic expectations, decomposition of the interaction variances of the significant interactions affecting KW, GN and GWP showed that these interactions were largely due to the additive component of epistasis (Table 3). For example, the double homozygous genotypes (F_2) accounted for ~19% of the population, but explained ~65% of the total interaction variances for the three traits. The heterozygotes at one or both loci had frequencies of ~47 and 34%, but explained ~31.5 and 3.5% of the total interaction variances, respectively. Consistent results were obtained from t tests in which there was at least one significant interaction effect (different from zero) between homozygotes in each of these interactions, but none of interaction effects associated with heterozygotes were significantly different from zero. These results suggested that the observed digenic interactions were due primarily to the additive epistatic gene action, as would be expected from the F_4 progeny testing in our experimental design.

The magnitudes of the additive digenic epistatic effects: The mean R^2 of the 16, 28, and 35 interactions affecting KW, GN and GWP were slightly larger than those of the identified QTL described previously. Of the possible 64, 120, and 140 interaction effects between homozygotes (there are four possible interaction effects in each of the interactions), 31, 70, and 85 were significantly different from zero based on t tests. There were three, five, and 20 interaction effects for KW, GN, and GWP, that were larger than the largest doubled additive effect of the QTL.

For KW, the mean R^2 explained by 16 interactions was 6.50%, ranging from 5.28 to 8.09%, which was slightly higher than the mean R^2 (5.73%, obtained by one-way ANOVA) explained by the nine KW QTL. The average magnitudes of the 33 interaction effects were 1.66 ± 0.54 g, similar to the doubled mean additive effect (1.56 ± 0.12 g) of the QTL. For GN, the mean R^2 explained by 28 interactions was 7.02%, ranging from 4.94 to 12.46%, which was slightly larger than the mean R^2 (6.77%) explained by the six GN QTL. The mean magnitude of the 70 interaction effects was 15.0 ± 6.5, similar to the doubled mean additive effects of the QTL (15.1 ± 5.1). For GWP, the mean R^2 explained by 35 interactions was 6.46%, ranging from 5.07 to 8.99%, which was larger than the mean R^2 (5.13%) explained by the four GWP QTLs. The mean magnitude of the 86 interaction effects was 3.02 ± 1.3 g, slightly larger than the doubled mean additive effects of the four QTL (2.53 ± 1.40 g).

Evidence for coadapted epistatic gene complexes: When the four interaction effects in an interaction were classified into two parental types (1L/2L or τ_{jj}, and 1T/2T or τ_{ii}) and two recombinant types (1L/2T or τ_{ji}, and 1T/2L or τ_{ij}), all three traits showed interesting characteristics in terms of the additive digenic epistatic effects (Table 4). Interactions between alleles from the same parents (the parental type) tended to result in increased productivity, while interactions between alleles from the different parents (the recombinants) tended to result in decreased productivity (Table 4). For instance, for GN, only 24.4% of the parental epistatic effects (τ_{ii} or τ_{jj}) were associated with reduced GN, but 69.0% of the recombinants were associated with reduced GN. As a result, the sum of the parental type interaction effects resulted in 217.8 more grains per panicle, but the cumulative effects of the recombinant interaction effects resulted in 120.9 fewer grains per panicle.

Evidence of possible high order interactions: Tables 5 and 6 list markers involved in the significant interactions affecting the three traits. Although it is expected that there should be a maximum of $2n$ loci for n digenic interactions, the actual number of the loci (markers) involved was much smaller. In fact, some loci appeared to be involved in interactions with more than one other locus. Here, we initially define a term, multiepistativity (ME), to describe a locus that interacts simultaneously with more than one other locus. When each marker involved in the significant interactions was counted as an independent locus, the number of loci involved in the digenic interactions was 24, 39 and 44 for KW, GN and GWP, respectively. The mean ME of the loci involved in digenic interactions for KW, GN, and GWP was 1.19, 1.44, and 1.59, respectively. In particular, a larger number of loci with ME ≥ 2 were influencing the complex traits GN and GWP (14 and 18) than the simpler trait KW (seven), and there were six loci affecting GWP that had a ME ≥ 3 (but only one for KW and two for GN). These results suggested possible presence of high order interactions affecting the traits, and that the more complicated traits, GN and GWP, appeared

TABLE 4

Summary of interactions ($\hat{\tau}_{ij}$) between homozygous genotypes for KW, GN and GWP

	Parental types						Recombinants (1L/2T or 1T/2L)	
	Lemont (1L/2L)		Teqing (1T/2T)		Total			
Trait	+[a]	−	+	−	+	−	+	−
KW								
No.	5	2	4	4	9	6	7	9
$\sum\hat{\tau}_{ij}$ (g)	5.35		0.89		6.24		−0.21	
GN								
No.	14	5	17	5	31	10	9	20
$\sum\hat{\tau}_{ij}$	137.2		80.6		217.8		−120.9	
GWP								
No.	16	9	17	7	33	16	15	21
$\sum\hat{\tau}_{ij}$ (g)	19.0		21.15		40.15		−12.50	

[a] + and − represent the interaction effects of positive and negative trait values, respectively.

to be determined by more complex higher order interactions than the highly heritable trait KW.

Relationships between epistatic loci and QTL: When the loci (markers) involved in the digenic interactions (Tables 5 and 6) were classified into three categories [(1) markers flanking the detected QTL affecting the same trait, (2) markers flanking the detected QTL affecting related trait(s), and (3) random genomic markers], we found that the majority of the interactions (85.5%) occurred between markers not linked to QTL affecting the same traits, and 47.4% of the identified QTLs (four of nine for KW, three of six for GN, and two of four for GWP) were not independent from the background loci.

Of the 24 loci affecting KW, 19 (79.2%) loci were random genomic markers (type 3) that were involved in 12 (75%) of the 17 interactions. Five were type 1 markers flanking *QKw1, QKw2b, QKw3a,* and *QKw3c,* and were involved in four of the 17 interactions. No interactions were detected between the KW QTL themselves. Of the 39 loci influencing GN, 32 (82.1%) were random markers (type 1) involved in 20 (71.4%) of the 28 interactions. Four were type 1 markers flanking three of the six GN QTL (*QGn1, QGn4b,* and *QGn10*), which were involved in five interactions. Three were type 2 markers, flanking *QKw2b, QKw3b* and *QGwp5,* which were involved in three interactions. Of the 44 loci affecting GWP, 36 (81.8%) were random markers that were involved in 24 (68.6%) of the interactions. Three were type 1 flanking *QGwp4a* and *QGwp5.* Five were type 2 markers flanking QTL for KW and/or GN that were involved in eight additional interactions.

Overall, a total of 63 different loci (markers) were involved in the digenic interactions affecting the three traits. Of these, 46 (73%) were random markers not associated with any detected QTL. Eight loci were involved in interactions affecting all three traits, 28 involved in interactions affecting two traits, and the remaining 27 affecting only one trait.

Pleiotropic effects of the additive digenic interactions on KW, GN, and GWP: Table 6 shows three interactions affecting both KW and GN, 16 interactions affecting both GN and GWP, and one interaction affecting both KW and GWP, respectively, in which epistatic effects for the correlated traits matched almost perfectly in both directions and magnitudes.

In the three interactions affecting both KW and GN, six of the seven significant *aa* effects for KW were negatively associated with the significant effects for GN. The coefficient of determination (R^2) of the six paired effects for KW and GN was 0.93 ($P < 0.0001$). In the 16 interactions affecting both GN and GWP, 35 of the 36 significant *aa* effects for GWP were positively associated with the corresponding significant effects for GN. The R^2 of the 35 paired *aa* effects were 0.86 ($P < 0.0001$). There was only one interaction influencing both KW and GWP, in which one (positive) of three significant *aa* effects for GWP were matched by a negative *aa* effect for KW.

DISCUSSION

These results indicate that a substantial portion of the genetic variances for complex traits that are *inexplicable* solely by QTL with relatively large phenotypic effects may be due to epistasis. Moreover, "main effects" of individual QTL may be somewhat modified as a result of epistatic relationships. The present study had two advantages in detecting epistasis. First, we selected three related grain yield component traits in the progeny of an intersubspecific rice cross, since the productivity and its components of progeny from such a cross are expected to be affected by epistasis. Second, the use of the breeding values of individual F_2 plants each ~100 observations of the F_4 progeny testing reduced the experimental errors in the measurements of the quantitative traits. It is also realized that our experimental design was unable to detect the nonadditive component

TABLE 5

Interaction effects between homozygotes at unlinked markers which affect KW, GN, and GWP

Trait	Chromosome	Marker 1	Marker 2	Chromosome	R^2 (%)	Digenic genotypes 1L/2L[b]	1L/2T	1T/2L	1T/2T
KW	*1*	RG236^{G}	RZ782b	*2*	5.57	−0.10	−0.42	1.61***	1.32**
	1	RG173	RG171	*2*	7.14	−1.39**	1.77****	2.45****	−0.64
	2	RG634	RZ761	*3*	5.28	−0.64	−1.71****	−0.77	0.58
	2	CDO718	CDO348b	*6*	8.09	−0.03	−3.39****	−0.94*	1.52***
	3	RG450	RG418^{G,P}	*3*	5.68	0.99*	−0.69	−2.30****	0.71
	3	RG348^{G}	RG470	*5*	5.44	0.13	0.54	−2.01****	−0.34
	3	RG944	RZ2	*6*	6.18	2.21****	−0.13	−0.71	−0.35
	3	CDO1069	RG4	*7*	7.06	−1.63***	0.07	1.30**	1.90****
	3	CDO1069	RG1022	*11*	7.75	0.42	−0.01	−2.15****	−0.19
	5	RG556	RG561	*10*	5.29	1.88****	−0.17	0.81	−1.19*
	6	CDO348b	RG716	*6*	6.02	2.06****	−0.81	−1.16*	0.75
	6	RZ2	RG910b	*9*	6.47	1.52***	−0.98*	−2.17****	1.11*
	8	RG598a	RZ397	*12*	7.52	−0.88	0.45	1.44**	−0.93*
	11	RG1022	RG901	*12*	6.44	2.15****	−0.84	−0.19	−1.05*
GN	*1*	RZ776	RG437	*2*	4.96	5.5	−7.1*	−0.8	15.4****
	1	RG811	RG256	*2*	6.02	0.5	−7.5*	−19.8****	8.4*
	1	RZ776	RG214^{P}	*4*	5.97	−3.4	−3.4	15.2****	7.1
	1	RG447	RG190	*4*	8.18	−2.2	−3.1	−9.3*	20.1****
	2	RG83	RG20	*8*	6.62	9.3*	−2.0	−15.9***	9.7*
	2	RG83	RG752^{K}	*10*	6.10	−14.4***	−4.1	6.1	−7.6*
	2	RG171	CDO98^{K}	*10*	7.71	13.9**	−1.4	−27.4****	8.9*
	3	RZ778a	RG103	*11*	6.35	4.2	3.3	27.4****	−7.8*
	6	*C*	RG20	*8*	6.90	17.0****	−1.8	13.4**	−11.7**
	6	RG716	RG752^{K}	*10*	6.57	−18.2****	9.6*	6.5	3.8
GWP	*1*	RZ390a	RG957	*1*	5.22	−3.69****	−1.29	−0.71	1.18
	1	RG381	CDO718	*2*	6.00	4.11****	−0.27	−1.66*	3.48****
	1	RZ776	RG634	*2*	5.31	3.22****	0.02	−2.26*	2.16*
	1	RZ19	RZ761	*3*	6.52	2.62**	−0.98	−0.94	5.12****
	1	RG957	RG13	*5*	5.79	0.26	1.78*	2.28*	−3.37***
	1	RG236^{K,G}	CDO98^{K,G}	*10*	6.86	3.32***	0.23	1.09	−1.55*
	2	RG256	RG634	*2*	6.08	−1.44	2.31**	−2.47**	1.92*
	2	RG520	RG182	*5*	8.10	0.12	−1.09	−1.65*	4.60****
	2	RZ782b	RG4	*7*	5.51	1.04	−3.48****	1.15	−1.31
	2	RG634	RG598a	*8*	5.07	−0.12	−2.03*	−0.30	1.82*
	2	RG83	CDO98^{K,G}	*10*	6.83	4.49****	−0.12	−4.06****	1.50
	2	RG256	RG118	11	5.30	−2.76**	1.21	0.95	−2.69**
	3	R761	RG214^{G}	*4*	6.45	−0.45	−2.18*	4.60****	3.15***
	3	RZ676	RG4	*7*	8.79	−3.04***	2.69**	3.53****	−4.61****
	6	RG716	RG463	*9*	6.73	−3.36***	−1.31	−1.00	1.58*
	6	RG716	RG561	*10*	5.51	−5.46****	2.00*	0.52	0.53
	6	*C*	RG241a	*12*	5.20	−1.75*	2.33**	−2.20*	1.23
	7	RG678b	RG118	*11*	5.91	3.70***	−0.20	0.97	2.75**
	10	RG752^{K,G}	RG1109	*11*	5.28	−0.92	2.63**	0.62	−3.03***

[a] Underlined are markers flanking the identified QTL. The superscripts K, G, and P represent markers flanking QTL for KW, GN and GWP, respectively.

[b] *, **, ***, and **** represent the significance levels of $P = 0.05$, 0.01, 0.001, and 0.0001, respectively.

of epistasis. Since our results were obtained strictly on the analyses of digenic interactions affecting yield components and based on a relatively small population size, extrapolation of our results on epistasis to highly heritable quantitative traits or to quantitative traits in crosses involving closely related parents (*cf.* the same subspecies) should be made cautiously. Nevertheless, several points concerning the importance of epistasis affecting quantitative trait variation in populations and the detection of epistatic loci in QTL mapping studies merit discussion.

Detection of epistasis: In the present study, we were able to detect large numbers of statistically significant ($P < 0.001$) interactions between random markers. We

TABLE 6

Homozygous interlocus interactions affecting multiple traits

Trait	Chromosome	Marker 1	Marker 2	Chromosome	R^2 (%)	Digenic genotypes[b] 1L/2L	1L/2T	1T/2L	1T/2T
GWP	*2*	RZ782b	CDO98	*10*	7.39	2.84**	1.03	−3.93****	1.77*
KW					6.80	0.16	1.32**	0.28	−2.33***
GN					6.72	11.5*	−1.0	−17.7****	19.8****
KW	*3*	RZ761	RG463	*9*	7.16	−1.23*	1.13*	1.41*	−0.86
GN					8.69	21.5****	−12.1*	−21.3****	6.5
KW	*5*	RG556	RG20	*8*	5.92	−0.30	−0.61	1.63***	−1.21*
GN					8.59	−2.8	5.2	−17.3****	16.8****
GWP	*1*	RG381	RG811	*1*	6.08	0.43	−0.10	−0.79	3.89***
GN					5.37	2.7	1.8	−3.5	16.7***
GWP	*1*	RG236[K,G]	RZ782a	*4*	5.93	3.44****	−2.23*	−0.44	1.56
GN					5.10	14.1***	−11.0**	−1.8	8.1
GWP	*1*	RZ19	RG207	*5*	5.19	2.31**	−3.99****	0.10	3.05***
GN					4.94	10.3*	−12.3**	4.4	15.1***
GWP	*1*	RZ19	RG1022	*11*	5.66	−0.03	−1.83	−2.92**	0.62
GN					5.96	−0.5	−6.9	−15.0***	−2.8
GWP	*1*	RG381	RW257	*12*	5.56	0.73	2.62**	−1.52*	3.33****
GN					7.07	7.6*	8.9*	−11.4**	18.3****
GWP	*2*	RG634[K]	CDO544	*6*	6.07	3.24***	−3.42***	−0.40	2.51**
GN					6.62	22.5****	−15.1***	−1.6	11.5**
GWP	*3*	RZ761	RG207	*5*	8.80	3.81****	−0.62	−0.63	4.09****
GN					9.77	20.0****	−0.0	−7.8*	19.5****
GWP	*3*	CDO109a[K]	CDO98	*10*	7.95	−4.67****	−1.46*	1.98*	0.90
GN					7.77	−21.0****	−10.6*	4.2	5.6
GWP	*4*	RG445b	CDO109b	*6*	7.62	−3.76****	−0.02	0.74	0.69
GN					5.62	−15.6****	0.6	1.1	1.8
GWP	*4*	RZ782a	RG1034	*8*	5.77	2.48**	2.68***	−1.87*	0.84
GN					6.97	11.5*	17.7****	−8.4*	5.2
GWP	*5*	RG346	RZ660	*9*	5.93	1.00	3.26***	1.31	−0.77
GN					5.54	4.9	12.8**	8.9*	−8.2*
GWP	*5*	RG470[P]	CDO98	*10*	8.99	−2.06*	−1.54*	6.23****	−2.71**
GN					8.00	−8.7*	−7.5*	30.7****	−11.5*
GWP	*6*	CDO348b	RZ660	*9*	6.49	1.95*	−0.49	0.25	−11.28****
GN					7.63	12.1*	−3.1	5.9	−51.6****
GWP	*7*	RG4	RG463	*9*	8.85	2.12*	0.96	−0.07	−3.16**
GN					12.46	17.6***	3.0	0.6	−16.7**
GWP	*8*	RG1034	RZ777	*9*	7.46	2.86***	−3.76****	1.08	0.18
GN					8.17	26.2****	−17.2****	1.0	−2.6

[a] The superscripts K, G, and P represent markers flanking the identified QTL for KW, GN and GWP, respectively.

[b] L and T represent homozygous Lemont and Teqing alleles at the interacting markers, 1 and 2 represent markers 1 and 2. *, **, ***, and *** represent the significance levels of $P = 0.05$, 0.01, 0.001, and 0.0001, respectively, based on t tests.

found that use of multiple regression (ZENG 1993, 1994) had significantly improved the power in detecting QTL, as we identified two additional QTL and removed two false positive QTL in the present study. This method also proved to be a very effective way to control the background genetic effects caused by segregating QTL in the detection of epistasis. Given the common population size in most mapping studies, such a control of background genetic effects of QTL is critical in identifying real loci involved in epistasis since highly significant false positive interactions may well arise in the two-way ANOVA as a result of nonrandom sampling of multiple segregating QTL. Our sample size of the 255 F_2 families (or seven to 17 F_2 plants for each of the homozygous digenic genotypes at two loci) was fairly small, and we indeed found a high probability of false positive interactions arising from the background genetic effects of segregating QTL in the population. For the same reason, we acknowledge that a significant proportion of digenic interactions may have gone unde-

TABLE 7

Comparison of QTL mapping and analyses of digenic interactions

	KW	GN	GWP
Heritability (h^2)	0.877	0.535	0.413
QTL mapping			
No. of identified QTL	9	6	4
Mean R^2 (%) explained by individual QTL from ANOVA	6.0	6.8	5.1
R^2 (%) collectively explained by all QTL	84.6	37.8	22.7
Standardized mean doubled additive effect of QTL	2.3	0.7	0.3
Analyses of digenic interactions			
No. of digenic interactions detected	16	28	35
Mean R^2 (%) explained by individual interaction variances	6.5	7.0	6.5
No. of QTL (markers) interacting with random loci	4	3	2
No. of interactions between QTL	0	0	0
% of interactions between QTL and random markers	25.0	14.3	8.6
% of interactions between random markers	76.5	88.5	90.9
Minimum number of epistatic loci involved	24	39	44
Mean multiepistativity (ME[a]) of individual loci	1.19	1.44	1.59
No. of loci with ME $\geq$ 2	3	14	18
Standardized mean additive interation effect	2.7	0.7	0.4

[a] ME is the number of other loci a specific locus (marker) interacts with.

tected, as several highly significant interactions were detected only in multiple regression models.

Relationships between QTL and epistatic loci: Two major points can be made concerning the relationships between the QTL and the epistatic loci identified in the present study (Table 7). First, the majority (87.1%) of loci involved in the digenic interactions did not appear to have significant main (additive and/or dominance) effects on the three traits when assayed alone. It is conceivable that an epistatic locus can be detected as a QTL when alleles at the other locus it interacts with become fixed. Thus, QTL mapping using one-way ANOVA or interval mapping based on the classic quantitative genetics model is expected to be able to detect genes with relatively large effects, and/or loci that have intermediate effects but act largely independently of other genes in the genetic backgrounds and the environment studied. The methodology and the experimental designs used in most QTL mapping studies *preferentially* identify genes that either have large effects and/or act independently, as suggested by PATERSON *et al.* (1991). This argument is supported by the results that the same QTL with large effects are mappable in very different crosses and environments (PATERSON *et al.* 1991; SCHON *et al.* 1994; LI *et al.* 1995a; XIAO *et al.* 1995). Thus, like the results from most previous mapping studies (EDWARDS *et al.* 1987; STUBER *et al.* 1987, 1993; PATERSON *et al.* 1988, 1990, 1991; KOWALSKI *et al.* 1994), paucity of interactions between QTL in the present study is expected.

Second, a significant proportion (45.0%) of the identified QTL (linked markers) were involved in digenic interactions with background loci. Thus, the usual estimates of main effect of a QTL can be confounded by interactions, which may change according to genetic backgrounds, environments, and other factors, as reported in tomatoes and maize (TANKSLEY and HEWITT 1988; DEVICENTE and TANKSLEY 1993; STROMBERG *et al.* 1994; DOEBLEY *et al.* 1995; LARK *et al.* 1995). In other words, QTL and the epistatic loci are interchangeable depending on the genetic backgrounds and probably environments where they are identified. Furthermore, many of the QTL detectable by the interval mapping or ANOVA may actually represent groups of tightly linked epistatic genes (COCKERHAM and ZENG 1996). Thus, lack of interactions between QTL should not be considered as evidence supporting the absence of epistasis.

The number of loci involved in epistasis, higher order interactions and the magnitude of epistatic effects: In agreement with the classical expectations, our results (Table 7) indicated that more complex traits like GN and GWP were indeed influenced by a greater number of epistatic markers and more complicated forms of epistasis. A dilemma arising from this result was that the cumulative contribution of all these interactions was well beyond the total variation of individual traits if these digenic interactions were independent from one another. Statistically, this indicated that many of the tests and estimates were highly correlated. Genetically, our results that interactions between QTL and background loci and high ME values of many of the markers involved in digenic interactions suggest possible presence of higher order gene interactions.

It was noted that markers affecting the same trait and having high ME did not necessarily interact with other markers in the same way. If they did so, they would have been identified as additional QTL, as discussed above. Genetically, there is no reason to assume that a

locus involved in high-order interactions should interact similarly with other loci since different loci may represent different genes, regulatory genes, etc. Also, one-to-one correspondence between a marker and a gene is not a realistic assumption because of the relatively low resolution (19 cM) of our genetic map and the maximum linkage disequilibrium in an F_2 population. This argument may hold true even for situations where a marker interacted similarly with several other markers. Nevertheless, the observation that some markers interact differently with several other loci at the digenic level would imply the presence of high order interactions.

Although our estimates of individual interaction effects were biased because of the presence of segregating QTL and higher order interactions, the mean R^2 and the interaction effects obtained from a large number of interactions suggest that epistatic gene effects are equivalent in magnitude to QTL, at least for the three traits studied.

The evidence of coadapted epistatic gene complexes: Although the apparent presence of complementary (or compatible) genes that affect the fitness and its components of progenies in interspecific hybrid backgrounds has been widely observed (*cf.* STEBBINS 1954), the hard evidence came only recently in Drosophila when interactions between conspecific genes affecting male sterility were experimentally demonstrated (CLARK 1987; CABOT *et al.* 1994; DAVIS *et al.* 1994; PALOPOLI and WU 1994; LOREN *et al* 1996). Our result that interactions between alleles from the same parents tend to result in increased productivity and the recombinant type interactions tend to result in reduced fitness strongly suggests that epistatic loci affecting the three grain yield components in the Lemont/Teqing cross act in a predominantly complementary manner. This has important implications for quantitative genetic theory underlying genetic variation and evolution of quantitative traits in populations, as well as gene mapping experiments.

First, the predominantly complementary gene actions of large numbers of epistatic loci affecting fitness traits are consistent with the observed nonhierarchical multilocus structure of isozyme polymorphism in *O. sativa,* which was predicted to be generated and maintained by strong selection on adaptedness determined by coadapted epistatic gene complexes (LI and RUTGER 1997). The presence of such coadapted epistatic gene complexes has long been realized (WRIGHT 1931; ALLARD 1988, 1996). The cross between Lemont and Teqing should be considered a wide cross (*japonica/indica*). Then, the observation that the majority of the interactions between alleles from different parents (the recombinant types) resulted in reduced productivity strongly suggests the presence of such coadapted epistatic gene complexes and provides an appropriate explanation for the reduced GN and GWP in the F_4 progeny of the Lemont/Teqing cross, which could not be explained by the identified QTL exhibiting large dominance effects for increased productivity. In other words, epistasis has played an important role in maintaining the integrity of both *indica* and *japonica* genomes (genes) in rice.

The third implication is that the complementary loci may play an important role in the maintenance of genetic variation for quantitative traits in populations. The observation that all of the interactions occurred between unlinked markers is not surprising in a selfing plant species like rice since such associations between unlinked loci can be easily fixed by inbreeding. It is also realized that the low resolution of genetic maps and the common experimental designs plus the statistical methods used in most mapping experiments do not allow detection of epistasis between closely linked markers (COCKERHAM and ZENG 1996). While the theories of mutation-selection balance (LANDE 1975, 1982) and mutation-drift balance (KIMURA 1971), alone or together, do not appear to adequately explain the high level of genetic variation for quantitative traits in self-pollinated plant populations, pronounced epistasis together with mutation and different kinds of selection provide a viable explanation (GILLESPIE 1978; CLARK 1991; SZATHMARY 1993; GAVRILETS and JONG 1993; ORR 1995).

Implications for evolution and crop breeding: The parents, Lemont and Teqing, are representatives of two highly differentiated gene pools [the *indica* gene pool and the *japonica* gene pool, as indicated by the high level of polymorphisms between them in both isozymes and RFLP (LI and RUTGER 1997)], although both adapt well to similar environments. This indicates that phenotypic similarity between parents may provide little information regarding the loci contributing to genetic variation of quantitative traits in their progenies. When large numbers of complementary loci are involved, considerable variation for quantitative traits could be maintained in populations that have similar overall phenotypes.

The same phenotype of a quantitative trait may be generated by very different allelic combinations of QTL and/or epistatic loci. This also provides an explanation for the observation that many natural populations of self-pollinated plant species (or traditional landraces) are mixtures of many different genotypes (CHANG 1967; HARLAN 1969; OKA 1988). It is conceivable that together with very limited outcrossing rate and low mutation rates such mixtures of different genotypes in self-pollinated plant species might have provided themselves with considerable "buffer capacity" to adapt to changing environments. With such a population structure and pronounced epistasis, population bottlenecks and subdivision expose hidden additive genetic variation to selection and lead to rapid differentiation of subpopulations (CARSON and TEMPLETON 1984; WADE 1992). In

evolution, this implies that the same peak(s) in the fitness landscape of a self-pollinated plant species, as defined by WRIGHT (1951), may be occupied by individuals with very different genetic compositions. Shifts from one peak to another (a higher peak) may be initiated by genetic drift, but more likely by fluctuating selection pressure if different gene combinations in the peak(s) show relatively large but varied G × E interaction (or more pronounced but varied phenotypic plasticity), which appears to be the case for self-pollinated plant species (JAIN and MARSHALL 1967).

In plant breeding, one could imagine that considerable variation for quantitative traits is hidden in the F_2 generation when linkage disequilibrium is maximal, and is released in later generations by recombination. As we noted, although interactions between alleles from different parents (the recombinants) tended to result in reduced fitness or yield, a significant proportion of the recombinants had resulted in increased GN and GWP, which would provide opportunities for crop improvement.

Finally, prevalence of complementary loci affecting complex quantitative traits in rice implies that classification of alleles as "favorable" or "unfavorable" may be misleading since the effect of an allele may be positive, neutral or negative depending on interactions with other loci and on environments. In other words, what plant and animal breeders have been looking for are probably not the "best" genes but the best gene combination(s). Thus, selection for increased trait values should be more efficient when it is practiced on individual allelic combinations at two or more loci. If so, gene/QTL mapping needs to place more emphasis on identifying the best multilocus gene combination(s).

Genetic basis of relationships between complex traits and component traits: It is normally assumed that genotypic correlation between complicated traits such as yield and its component traits arises presumably as a result of pleiotropy. Pleiotropic effects of the detected QTL and the epistatic loci appeared to be responsible for the negative correlation between the two sink-size traits KW and GN, presumably as a result of competition for limited resources (carbohydrates). However, a significant proportion of the identified QTL and the digenic interactions affecting GN, but none affecting KW, contribute to GWP. The negative association resulting primarily from pleiotropic effects of loci affecting GN and KW is partially responsible. Lack of association between loci influencing GWP and those affecting KW but not GN was perhaps because effects of these loci on GWP were too small to be detected (canceled out by large environmental effects associated with GWP) even though they indeed contribute to GWP. Our result that six (18.2%) of the 33 interactions influencing GWP involved markers affecting KW suggests that epistasis may also be a factor. In these cases, it is possible that some of the QTL for component traits may have indeed acted as modifying factors and contribute to the genotypic correlation. Thus, detailed information about the loci involved in the different yield components and their genetic relationships will certainly be helpful to improve yield potential by deliberately manipulating these loci through marker-assisted selection.

We are grateful to Dr. Z.-B. ZENG and two anonymous reviewers for many critical comments and suggestions in the manuscript. We thank Drs. S. D. TANKSLEY and S. MCCOUCH for providing DNA probes, Dr. S.-C. LIU for valuable discussions, and Dr. Y.-W. WANG for technical help. This research was supported by USDA-Agricultural Research Service, Southern Plains Area; The Texas A&M University System Agricultural Research and Extension Center; The Texas Rice Research Foundation; and grants from the Texas Advanced Technology Program to W.D.P. and to Z.L.

LITERATURE CITED

ALLARD, R. W., 1988 Future directions in plant population genetics, evolution and breeding, pp. 1–19 in *Plant Population Genetics and Germplasm Resources*, edited by A. H. D. BROWN, M. T. CLEGG, A. L. KAHLER and B. S. WEIR. Sinauer Associates Inc., Sunderland, MA.

ANTHONY, D. L., S. L. MULLANEY, L. A. REID, J. D. FRY, C. H. LANGLEY *et al.*, 1995 High resolution mapping of genetic factors affecting abdominal bristle number in *Drosophila melanogaster*. Genetics **139:** 1273–1291.

CABOT, E. L., A. W. DAVIS, N. A. JOHNSON and C. I. WU, 1994 Genetics of reproductive isolation in the *Drosophila simulans* clade: complex epistasis underlying hybrid male sterility. Genetics **137:** 175–189.

CARSON, H. L., and A. R. TEMPLETON, 1984 Genetic revolutions in relation to speciation phenomena: the founding of new populations. Annu. Rev. Ecol. Syst. **15:** 97–131.

CHANG, T. T., 1975 Exploration and survey in rice, pp. 159–165 in *Crop Genetic Resources for Today and Tomorrow*, edited by O. H. FRANKEL and J. G. HAWKES. Cambridge University Press, Cambridge.

CLARK, A. G., 1987 A test of multilocus interaction in *Drosophila melanogaster*. Am. Nat. **130:** 283–299.

CLARK, A. G., 1991 Mutation-selection balance and metabolic control theory. Genetics **129:** 909–923.

COCKERHAM, C. C., and Z. B. ZENG, 1996 Design III with marker loci. Genetics **143:** 1437–1456.

DAVIS, A. W., E. G. NOONBERG and C. I. WU, 1994 Evidence for complex gene interactions between conspecific chromosomes underlying hybrid female sterility in the *Drosophila simulans* clade. Genetics **137:** 191–199.

DEVINCENTE, M. C., and S. D. TANKSLEY, 1993 QTL analysis of transgressive segregation in an interspecific tomato cross. Genetics **134:** 585–596.

DOEBLEY, J., A. STEC and C. GUSTUS, 1995 *teosinte branchedl* and the origin of maize: evidence for epistasis and the evolution of dominance. Genetics **141:** 333–346.

EDWARDS, M. D., C. W. STUBER and J. F. WENDEL, 1987 Molecular-marker-facilitated investigations of quantitative-trait loci in maize. I. Numbers, genomic distribution and types of gene action. Genetics **116:** 113–125.

EDWARDS, M. D., T. HELENTJARIS, S. WRIGHT and C. W. STUBER, 1992 Molecular-marker-facilitated investigations of quantitative trait loci in maize. 4. Analysis based on genome saturation with isozyme and restriction fragment length polymorphism markers. Theor. Appl. Genet. **83:** 765–774.

FALCONER, D. S., 1981 *Introduction to Quantitative Genetics*, Ed. 2. Longman Press, New York.

FATOKUN, C. A., D. I. MENANCIO-HAUTEA,D. DANESH and N. D. YOUNG, 1992 Evidence for orthologous seed weight genes in cowpea and mung bean based on RFLP mapping. Genetics **132:** 841–846.

FREUND, E. J., 1971 *Mathematical Statistics*, Ed. 2. Prentice-Hall, Inc., Englewood Cliffs, NJ.

GAVRILETS, S., and G. DE JONG, 1993 Pleiotropic models of polygenic

variation, stabilization selection, and epistasis. Genetics **134:** 609–625.

GEORGES, M., D. NIELSON, M. MACKINNON, A. MISHRA, R. OKIMOTO *et al.*, 1995 Mapping quantitative trait loci contributing milk production in dairy cattle by exploiting progeny testing. Genetics **139:** 907–920.

GRAYBILL, F. A., 1976 *Theory and Application of the Linear Model.* Wadsworth and Brooks/Cole, Advanced Books and Software, Pacific Grove, CA.

GILLESPIE, J., 1978 A general model for enzyme variation in natural populations. V. The SAS-CFF model. Theor. Pop. Biol. **14:** 1–45.

HARLAN, J. R., 1969 Evolutionary dynamics of plant domestication. Jap. J. Gene. **44(Suppl.1):** 337–343.

HEDRICK, P. W., 1983 *Genetics of Populations.* Jones and Bartlett Publishers, Inc. Boston.

HORVAT, S., and J. F. MEDRANO, 1995 Interval mapping of *high growth* (*hg*), a major locus that increases weight gain in mice. Genetics **139:** 1737–1748.

JAIN, S. K., and D. R. MARSHALL, 1967 Population studies in predominantly self-pollinated species. X. Variation in natural populations of *Avena fatua* and *A. barbata.* Am. Nat. **101:** 19–33.

KIMURA, M., 1971 *Theoretical Aspects of Population Genetics.* Princeton University Press, Princeton, NJ.

KINOSHITA, T., and N. SHINBASHI, 1982 Identification of dwarf genes and their character expression in the isogenic background. Jpn. J. Breed. **32:** 219–231.

KOWALSKI, S. P., T. H. LAN, A. FELDMANN and A. H. PATERSON, 1994 QTL mapping of naturally-occurring variation in flowering time of *Arabidopsis thaliana.* Mol. Gen. Genet **245:** 548–555.

LANDE, R., 1975 The maintenance of genetic variability by mutation in a polygenic character with linked loci. Genet. Res. **26:** 221–235.

LANDE, R., 1982 A quantitative genetic theory of life history evolution. Ecology **63:** 607–615.

LANDE, R., and R. THOMPSON, 1990 Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics **124:** 743–756.

LANDER, E. S., and D. BOTSTEIN, 1989 Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics **121:** 185–199.

LARK, K. G., K. CHASE, F. ADLER, L. M. MANSUR and J. H. ORF, 1995 Interactions between quantitative trait loci in soybean in which trait variation at one locus is conditional upon a specific allele at another. Proc. Natl. Acad. Sci. USA **92:** 4656–4660.

LI, Z. K., and J. N. RUTGER, 1997 Geographic distribution and multilocus structure of isozyme variation of rice (*Oryza sativa* L.). Theor. Appl. Genet. (in press).

LI, Z. K., S. R. M. PINSON, J. W. STANSEL and W. D. PARK, 1995a Identification of QTL for heading date and plant height in rice using RFLP markers. Theor. Appl.Genet. **91:** 374–381.

LI, Z. K., S. R. M. PINSON, M. A. MARCHETTI, J. W. STANSEL and W. D. PARK, 1995b Characterization of quantitative trait loci contributing to field resistance to sheath blight (*Rhizonctonia solani*) in rice. Theor. Appl. Genet. **91:** 382–388.

LOREN, H. R., B. SINERVO, C. R. LINDER, M. C. UNGERER and D. M. ARIAS, 1996 Role of gene interactions in hybrid speciation: evidence from ancient and experimental hybrids. Science **272:** 741–745.

MATHER, K., and J. L. JINKS, 1982 *Biometrical Genetics,* Ed. 3. Chapman and Hall, London.

OKA, H. I., 1988 Evolution of cultivated rice (*Oryza sativa* L.). Japan Scientific Societies Press, Tokyo and Elsevier Science Publishers, Amsterdam, Oxford, New York, and Tokyo.

ORR, H. A., 1995 The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics **139:** 1805–1813.

PALOPOLI, M. F., and C. I. WU, 1994 Genetics of hybrid male sterility between *Drosophila* sibling species: a complex web of epistasis is related in interspecific studies. Genetics **138:** 329–341.

PATERSON, A. H., S. E. LANDER, J. D. HEWITT, S. PETERSON, H. D. LINCOLN *et al.*, 1988 Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature **335:** 721–726.

PATERSON, A. H., D. S. HEWITT, J. D. ZAMIR, D. RABINOWITCH, H. D. LINCOLN *et al.*, 1991a Mendelian factors underlying quantitative traits in tomato: comparison across species, generations and environments. Genetics **127:** 181–197.

PATERSON, A. H., J. W. DEVERNA, B. LANINI and S. D. TANKSLEY, 1991b Fine mapping of quantitative trait loci using selected overlapping recombinant chromosomes, in a interspecies cross of tomato. Genetics **124:** 735–742.

POONI, H. S., D. J. COOMBS and P. S. JINKS, 1987 Detection of epistasis and linkage of interacting genes in the presence of reciprocal differences. Heredity **58:** 257–266.

SATO, S., and I. SAKAMOTO, 1983 Inheritance of heading time in isogenic line rice cultivar, Taichung 65 carrying earliness genes from a reciprocal translocation homozygote, T3-7. Jpn. J. Breed. **33 (Suppl. 1):** 118–119.

SCHON, C. S., A. E. MELCHINGER, J. BOPPENMAIER, E. B. JUNG, R. G. HERRMANN *et al.*, 1994 RFLP mapping in maize: quantitative trait loci affecting testcross performance of elite European flint lines. Crop Sci. **34:** 378–389.

SPASSKY, B., T. DOBZHANSKY and W. W. ANDERSON, 1965 Genetics of natural populations. XXXVI. Epistatic interactions of the components of the genetic load in *Drosophila pseudoobscura.* Genetics **52:** 653–664.

SPICKETT, S. G., and J. M. THODAY, 1966 Regular response to selection 3. Interaction between located polygenes. Genet. Res. **7:** 96–121.

STUBER, C. W., M. D. EDWARDS and J. F. WENDEL, 1987 Molecular marker-facilitated investigations of quantitative trait loci in maize. II. Factors influencing yield and its components traits. Crop Sci. **27:** 239–248.

STUBER, C. W., S. E. LINCOLN, D. W. WOLFF, T. HELENTJARIS and E. S. LANDER, 1992 Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. Genetics **132:** 823–839.

STROMBERG, L. D., J. W. DUDLEY and G. K. RUFENER, 1994 Comparing conventional early generation selection with molecular marker assisted selection in maize. Crop Sci. **34:** 1221–1225.

SZATHMARY, E., 1993 Do deletion mutations act synergistically? Metabolic control therapy provides a partial answer. Genetics **133:** 127–132.

TANKSLEY, S. D., and J. D. HEWITT, 1988 Use of molecular markers in breeding for soluble solids in tomato—a re-examination. Theor. Appl. Genet. **75:** 811–823.

TEMPLETON, A. R., 1979 The unit of selection in *Drosophila mercatorum.* II. Genetic evolution and the origin of coadapted genomes in parthenogenetic strains. Genetics **92:** 1265–1282.

TEMPLETON, A. R., 1980 The theory of speciation via the founder principle. Genetics **94:** 1011–1038.

VAN DER VEEN, J. H., 1959 Tests of non-allelic interaction and linkage for quantitative characters in generations derived from two diploid pure lines. Genetic XXX **(3):** 201–232.

WADE, M. J., 1992 Sewall Wright: gene interaction and the shifting balance theory. Oxf. Surv. Evol. Biol. **8:** 33–62.

WRIGHT, S., 1931 Evolution in Mendelian populations. Genetics **16:** 97–159.

WRIGHT, S. 1932 The roles of mutation, inbreeding, crossbreeding and selection in evolution. Proc. VI. Intern. Congr. Genet. **1:** 356–366.

WRIGHT, S., 1951 The genetic structure of populations. Ann. Eugenics **15:** 323–354.

ZENG, Z. B., 1993 Theoretical basis of separation of multiple linked gene effects on mapping quantitative trait loci. Proc. Natl. Acad. Sci. USA **90:** 10972–10976.

ZENG, Z. B., 1994 The precision mapping of quantitative trait loci. Genetics **136:** 1457–1468.

Communicating editor: Z-H. ZENG