Theor Appl Genet (1992) 83:903–911

THEORETICAL AND APPLIED GENETICS

Use of RFLP markers to search for alleles in a maize population for improvement of an elite hybrid

B. E. Zehr [1,*,], J. W. Dudley [1], J. Chojecki [2], M. A. Saghai Maroof [3,***], and R. P. Mowers [3]**

[1] Department of Agronomy, University of Illinois, Urbana, IL 61801, USA
[2] ICI Seeds, Jealott's Hill Research Station, Bracknell, Berkshire, RG12 6EY, UK
[3] Garst Seed Co., Research Department, Slater, IA 50244, USA

Received December 4, 1990; Accepted September 10, 1991
Communicated by A. L. Kahler

Summary. Molecular markers can be used to detect alleles in donor genetic material for improvement of existing cultivars or hybrids. DNA restriction fragment length polymorphisms (RFLPs) were used as markers to search for favorable alleles at quantitative trait loci in the maize (*Zea mays* L.) population BS11(FR)C7 which were not in the hybrid 'FRB73 × FRMo17.' Thirty-four RFLP markers were used to determine RFLP 'fingerprints' for 220 [BS11(FR)C7 × FRMo17] F_2 individuals; multiple morphs (bands) were observed at most markers. Statistical associations between RFLPs and trait expression in F_2 × FRB73 progeny were found for grain yield, stalk and root lodging, plant and ear height, maturity, and seven grain yield component traits. Associations were found using linear contrasts among RFLP marker classes to estimate trait effects. Estimated effects for grain yield ranged from 213 to 538 kg ha^{-1}, 3.0–7.5% of the experimental mean, respectively. RFLP markers with greatest probability of association with grain yield were NPI234 (short arm of chromosome 1) and UMC16 (long arm of chromosome 3). Digenic epistasis appeared to be important in grain yield expression, as indicated by a 12% increase in the proportion of genotypic variation accounted for when significant di-marker interactions were added to a linear model, including all markers individually associated with grain yield. The majority of interactions associated with grain yield involved markers NPI234 and UMC21 (long arm of chromosome 6). Many RFLP markers were associated with multiple traits. At some markers, the same bands were associated with favorable effects for stalk lodging, grain yield, and yield components. RFLP bands unique to BS11(FR)C7 showed associations favorable over those from FRMo17 for at least one marker in all but one trait. The results of this study will be useful in future RFLP marker-assisted selection programs aimed at developing lines for improved performance in combination with FRB73.

Key words: *Zea mays* L. – RFLPs – Quantitative traits – Favorable alleles – Epistasis

* Present address: Department of Agronomy, Purdue University, West Lafayette, IN 47907-1150, USA
** To whom correspondence should be addressed
*** Present address: Department of Crop and Soil Environmental Sciences, Virginia Polytechnic Institute & State University, Blacksburg, VA 24061-0404, USA

Introduction

Molecular markers are being studied for their potential to enhance selection efficiency in plant breeding. With the development of molecular marker linkage maps, isozyme and restriction fragment length polymorphism (RFLP) markers have been used to locate and manipulate loci affecting expression of simply and quantitatively inherited traits (Tanksley et al. 1989; Stuber 1992). Associations between molecular markers and quantiative trait loci (QTL) have been reported in a number of plant species (Stuber 1992). QTL detection is dependent upon disequilibrium in the genetic linkage between marker forms and alleles at a gene of interest. In order to maximize such disequilibrium, F_1 backcross or F_2 populations from a cross of two inbred parents have commonly been used. Marker and agronomic data are collected on F_2 individuals or recombinant inbred lines derived from them. Genetic linkage is then inferred from statistical correlations between trait expression and segregation of marker forms among individuals or lines.

Recurrent mating among random or selected individuals in a population over generations provides opportunity for genetic recombination. In such a population, the

ability for marker detection of QTL and estimation of their effects will be diminished as linkage equilibrium is approached. However, crossing a population with an inbred line will create a degree of linkage disequilibrium in the F_2, even if the population is in linkage equilibrium (J. W. Dudley, unpublished results). Any disequilibrium in the population resulting from recurrent selection should enhance disequilibrium in the F_2 from a population × line cross, and thus detection of marker-QTL associations. Also, disequilibrium in the population would have been maintained where close marker-QTL linkages are present.

In maize (*Zea mays* L.), marker polymorphism is common within populations and among inbred lines (Stuber and Goodman 1983; Helentjaris et al. 1985; Smith et al. 1985; Kahler et al. 1986). Differences in magnitude, action, and genomic distribution of QTL effects measured by marker associations have been found in a number of F_2 populations (Kahler and Wehrhan 1986; Stuber et al. 1987; Edwards et al. 1987; Abler et al. 1991). Although F_2 individuals are useful for detecting marker-QTL associations, the goal in a commercial maize breeding program is to maximize performance of the cross between two inbred lines. Crosses of F_2 individuals, or recombinant inbred lines, with an unrelated tester can be used to evaluate effects related to the substitution of marker forms in an inbred tester background (Cowen 1988). Previous studies have shown genomic locations of detected QTL for testcross performance to be different from those in F_2 individuals or lines per se (Guffy et al. 1989; Stuber 1992).

Dudley (1988 a) developed quantitative genetic theory and methodology for identification of donor populations containing favorable alleles not already present in elite single-cross hybrids. Dudley (1988 b) tested 19 maize populations and identified BS11(FR)C7 as having the greatest relative number of favorable alleles for grain yield not already present in the hybrid FRB73 × FRMo17. BS11(FR)C7 also was found to contain relatively large numbers of favorable alleles for early maturity and root lodging, while having a heterotic pattern more similar to FRMo17 than FRB73.

This study uses RFLP markers to: (1) search for QTL segregating among testcrosses to FRB73 of F_2 plants from the cross of population BS11(FR)C7 with inbred line FRMo17; (2) identify useful alleles from BS11(FR)C7 that could improve the agronomic performance of FRMo17 in combination with FRB73; (3) search for digenic epistasis among marker pairs individually determined to have possible QTL associations.

Materials and methods

Genetic material and field evaluation

Population BS11(FR)C7 was used as a donor of alleles for improvement of hybrid FRB73 × FRMo17. BS11(FR)C7 resulted from seven cycles of full-sib reciprocal selection for grain yield in BS11, a synthetic of corn belt lines and southern prolific material developed at Pioneer Hi-Bred International, Inc. (Hallauer et al. 1974). FRB73 and FRMo17 are Illinois Foundation Seeds, Inc. backcross versions of inbred lines B73 and Mo17 containing the *Ht-1* gene (Henderson 1976).

To develop a segregating F_2 population for QTL mapping, FRMo17 individuals were crossed to 150 representative plants from BS11(FR)C7. One hundred fourty four F_1 progeny were then randomly mated (RM) by making plant-to-plant crosses, never using the same plant as both a male and female, and bulking all resulting seed. From this seed bulk, 220 F_2 individuals were selected at random; each was self-pollinated and testcrossed to FRB73 for further evaluation.

In the summer of 1988, the 220 F_2 testcross progeny plus the hybrid FRB73 × FRMo17 and three commercial hybrid checks were grown in generalized lattice designs with 32 blocks, seven entries each, in three replications at two Illinois locations (both at Urbana) and two Iowa locations (Slater and Dunlap). At the Illinois locations, single-row plots 5.33 m long with 0.76 m between rows were over-planted and thinned to stands of 61,728 plants ha^{-1}. Iowa locations were planted in two-row plots with 5.33 m long rows, 0.76 m apart, and thinned to similar stands. Trait measurements were recorded on a per-plot basis and plots were machine harvested. Shelled grain yield, grain moisture at harvest, final stand, and number of stalk lodged plants (broken below the ear) were measured at all locations. Number of root-lodged plants (leaning more than 45° from vertical), plant and ear height (visual estimate of the plot average) were measured at Illinois locations only. Maturity (days from planting to 50% pollen shed) was measured at one Illinois location.

Trial entries also were grown separately in two replications at one Illinois (Urbana) and one Iowa (Slater) location for measurement of grain yield components. All plots consisted of single rows similar to those previously described. Plots were hand-harvested and the top ears from five competitive plants in each plot were sampled. Plot averages were calculated for the following traits: ear length, ear diameter, kernel depth, number of kernels and kernel rows per ear, total grain weight per ear, and 100-kernel weight. Ears were placed end-to-end and side-by-side to measure length and diameter. Kernel depth was calculated as one-half the difference between average ear diameter and average cob diameter. Seed bulks from each plot were used to obtain weights for 100 kernel samples.

Laboratory analysis

Procedures for DNA extraction, restriction enzyme digestion, gel electrophoresis, Southern blotting, and hybridization were as described by Saghai Maroof et al. (1984); oligo-labelling of probes (Feinberg and Vogelstein 1984) was used in place of nick translation. In this paper, each DNA probe/restriction enzyme combination will be referred to as an RFLP marker, individual morphs at a marker as bands, and pair-wise banding combinations in F_2 individuals as marker classes.

Banding patterns were determined for FRB73, FRMo17, and 18 random individuals from BS11(FR)C7, using 34 DNA probes and restriction digests of genomic DNA with four endonucleases (*Eco*RI, *Eco*RV, *Hin*dIII, *Sst*I). Not all probe/restriction digest combinations were tried; enzymes most often used in defining markers were *Eco*RI and *Hin*dIII. A total of 34 RFLP markers was selected, each based on maximum number of bands per probe among the inbred lines and population. Probes were obtained from Brookhaven National Laboratory (BNL), Native Plants, Inc. (NPI) and the University of Missouri at Columbia (UMC). Relative genomic locations are known (E. H. Coe, unpublished results, Maize Genet Coop Newsletter 65:129–164) and represent 18 of 20 chromosome arms (Table 1).

Table 1. DNA probes, representation of chromosome arms, and number of bands (morphs) used in statistical analyses

DNA probe[a]	Chromosome arm[b]	No. of bands used in analyses[c]
NPI234	1-S	2
BNL5.62	1-S	4
NPI238	1-L	4
UMC44-B	2-S	4
NPI402	2-S	3
NPI239	2-S	6
NPI297	2-L	4
UMC92	3-S	3
UMC16	3-L	3
UMC96	3-L	5
UMC47	4-S	2
UMC66	4-C	3
BNL15.07	4-L	3
NPI451	4-L	5
UMC27	5-S	3
UMC01	5-C	5
BNL5.71	5-L	2
UMC104	5-L	4
NPI373	6-L	3
UMC21	6-L	3
NPI223	6-L	2
UMC38	6-L	2
NPI391	7-S	2
NPI455	7-L	2
BNL16.06	7-L	3
UMC80	7-L	3
NPI114	8-S	3
UMC12	8-L	2
UMC30	8-L	3
BZ-1	9-S	3
NPI211	9-S	3
UMC81	9-C	3
NPI285	10-S	4
UMC44-A	10-L	4

[a] DNA probes were obtained from Brookhaven National Laboratory (BNL), Native Plants, Inc. (NPI), and the University of Missouri at Columbia (UMC); BZ-1 hybridizes at the *bronze-1* locus

[b] S = short arm, L = long arm, C = centromeric region

[c] For most probes, all observed bands were used in statistical analyses; however, bands with frequency <0.02 in the F_2 population were treated as missing data

Marker classes were determined for the 220 [FRMo17 × BS11(FR)C7]RM F_2 individuals used in testcrosses with FRB73. At each marker, individual bands were given letter designations. The number of bands across F_2 individuals ranged from two to six per marker (Table 1). Multiple bands were determined to be of a single marker by their pair-wise segregation among F_2 individuals, and based on knowledge of probe hybridization from previous studies.

Statistical analysis

The INCANOV program (Carmer and Kratzke 1988) for analysis of variance in a generalized lattice was used to obtain adjusted means for trait measurements of F_2 testcross and F_1 hybrid entries at each location. Adjusted entry means (except those of the three commercial hybrid checks) averaged over locations and the SAS procedure GLM (SAS Institute 1987) were then used in all subsequent analyses.

The following approach was used to determine statistical associations between segregation among F_2 individuals at RFLP markers and quantitative trait expression in F_2 × FRB73 progeny. Multiple linear contrasts among marker classes were calculated at each marker, each contrast indicating the presence or absence of a specific band (see Appendix). For each band, phenotypic observations in marker classes were given an indicator value of 1.0 if homozygous for the band, 0.5 if heterozygous, and 0.0 if the band was not present. For markers with multiple bands, this essentially pools bands into two categories: e.g., A and K, where A is the band of interest and K is the pooling of all other bands at the marker. In an analysis of variance (ANOVA), variance among the resulting three marker classes was partitioned into effects due to linear regression and residual. A separate ANOVA was calculated for each band per marker. Marker classes involving bands with extremely low frequencies were treated as missing data because any related effects would be estimated with low precision. A cut-off point of frequency <0.02 was suggested by the data; resulting contrasts had at least five observations per marker class after pooling of bands. For most probes, all observed bands were used in analyses.

For a marker with n bands, significance tests for n linear contrasts were obtained. A marker was considered for association with a QTL if at least one of its linear contrasts was significant. Linear regression coefficients were used to estimate relative QTL effects. At each marker showing multiple bands and significant trait association, t-tests for differences between pairs of linear regression coefficients were calculated to determine which bands were associated with QTL alleles having similar effect. Individual linear regressions at a single marker are not independent (see Appendix). The formula given by Steel and Torrie (1980) for testing homogeneity of two independent regressions was expanded to account for covariance due to non-independence. The t-statistic was calculated as follows:

$$t=(b_1-b_2)/[S_p^2(1/\Sigma X_1^2+1/\Sigma X_2^2-2\Sigma X_1X_2/(\Sigma X_1^2\Sigma X_2^2))]^{0.5},$$

where b_1-b_2 is the difference between regression coefficients, S_p^2 is pooled error, and the sum of squares and cross products for independent variables, X_1 and X_2, are corrected for their means.

For those markers that showed significant linear effects, all possible pairs were evaluated for the presence of digenic epistasis. The linear contrast with lowest probability of a greater F-value at each marker was chosen to represent the marker. These contrasts were then used in two-way ANOVAs to test for di-marker interactions. Interactions involving only linear effects were partitioned from those involving residuals. Only significant linear-by-linear interactions were accepted as indicating possible digenic epistasis.

Results and discussion

Detection of RFLP-QTL associations and estimation of effects

Markers with trait associations at $\alpha \leq 0.05$ levels of significance and estimated effects are listed in Tables 2 and 3. All measured traits showed marker association at these levels. The linear contrast with lowest probability value (showing greatest degree of trait association) at each marker was chosen to estimate the associated effect. For

Table 2. Estimated linear effects for substitution of bands at RFLP markers associated with grain yield at $\alpha \leq 0.05$

Chromosome arm[a]	RFLP		Estimated effect[b] (kg ha^{-1})	SE	Prob. $>F$[c]
	Marker	Band			
1-S	NPI234	B[de]	332	91	0.0003
		A	−332		
1-L	NPI238	B[d]	373	240	
		D[d]	237	231	
		A[de]	117	98	
		C	−231	99	0.0204
2-S	UMC44-B	C[d]	118	119	
		B[d]	100	150	
		A[de]	68	95	
		D	−305	121	0.0127
2-L	NPI297	A[de]	78	102	
		B[d]	72	159	
		D[d]	34	114	
		C	−426	179	0.0182
3-S	UMC92	B[d]	360	110	0.0013
		C[e]	−236	106	
		A	−342	211	
3-L	UMC16	B[d]	483	119	0.0001
		A[e]	−205	105	
		C	−319	151	
3-L	UMC96	C[d]	360	130	0.0063
		A[e]	26	101	
		B	− 60	120	
		D	−490	279	
		E	−548	205	
4-L	BNL15.07	C[d]	394	171	0.0225
		A	− 44	113	
		B[e]	−114	108	
5-S	UMC27	B[de]	102	110	
		A[d]	76	120	
		C	−461	181	0.0116
5-L	UMC104	A[de]	213	89	0.0181
		B	−104	95	
		C	−239	212	
		D	−274	170	
6-L	UMC21	A[d]	538	197	0.0068
		C[e]	49	108	
		B	−252	118	
6-L	UMC38	A[de]	221	80	0.0066
		B	−221		
7-S	NPI391	B[de]	265	104	0.0117
		A	−265		
8-S	NPI114	A[d]	117	105	
		B[de]	80	99	
		C	−288	122	0.0196
9-C	UMC81	C[de]	172	106	
		A[d]	13	123	
		B	−364	148	0.0147

grain yield, 15 markers were associated at $\alpha \leq 0.05$. Those with greatest degree of association were NPI234, on the short arm of chromosome 1 (1–S), and UMC16, on the long arm of chromosome 3 (3–L) (Table 2). Estimates of associated effects for grain yield ranged from 213 kg ha^{-1} (UMC104) to 538 kg ha^{-1} (UMC21), approximately 3.0–7.5% of the experimental mean of 7220 kg ha^{-1}.

The percentage of genotypic variation ($R^2 \times 100$) for grain yield explained by the linear contrast with greatest degree of trait association at each marker ranged from 2.4–7.1% (Table 4). Percentages were generally lower for markers associated with maturity and yield component traits, the exception being NPI391, which accounted for 10.8% of genotypic variation measured for ear diameter. Multiple regression models fitting the linear contrast representing each marker associated at $\alpha \leq 0.05$ explained 30% of genotypic variation for grain yield and stalk lodging, and 40% for grain moisture at harvest. R^2 percentages for multiple regression models were lower for other traits.

Much discussion has taken place in the literature (Lander and Botstein 1989) and elsewhere as to acceptable Type I error rates (α) for declaring marker-trait associations. The prevalent opinion has been that Type I errors (probability of falsely declaring an association) should be on a per genome (experiment-wise, α_e) rather than per marker or marker interval rate. This is intended to minimize mistaken associations in analyses over the entire genome. For the data in this paper, two to six linear contrasts were made per marker, across 34 markers, 110 contrasts per trait (Table 1). Hypothesis tests were conducted using a per contrast Type I error rate of $\alpha_c \leq 0.05$. A rough adjustment of $\alpha_e = \alpha_c/110 = 0.0004$ could be made to account for contrasts both within and among markers. Even using this very conservative α_e level, linear contrasts representing two of 15 markers at $\alpha_c \leq 0.05$ for grain yield are interpreted as having QTL association (Table 2). However, this Type I error adjustment highly inflates the Type II error rate (probability of not detecting a true effect). The appropriate balance between Type I and Type II error rates in corn breeding

[a] S = short arm; L = long arm; C = centromeric region

[b] Vertical bars represent no significant difference between estimated effects by t-test ($\alpha < 0.10$). The experimental mean for grain yield was 7,220 kg ha^{-1}

[c] Only the probability level for the linear contrast showing greatest grain yield association is presented at each marker. This degree of association and its corresponding estimated effect were chosen to represent the marker

[d] RFLP band(s) to be favored in marker-assisted selection for increased grain yield

[e] RFLP band in common with the FRMo17 parent; all other bands originated from the donor population BS11(FR)C7

Table 3. RFLP marker associations at $\alpha \leq 0.05$ for grain moisture, lodging, height, maturity, and grain yield components

Trait	Markers associated at ranges of Prob. $> F$[a]			Range for estimates of associated effects[b]	Experimental mean
	0.0500–0.0100	0.0090–0.0010	0.0009–0.0001		
Grain moisture (g 100 g^{-1})	BZ-1, B15.07, N297, N373, N455, U44-A, U44-B, U80	U16, N114	N238, U92	0.6–1.7	19.9
Stalk lodging (% standing plants)	B15.07, N373, N391, N451	N238, N455, U44-B	N297, U44-A	2.1–4.2	90.0
Root lodging (% standing plants)	N238, N391, N402	U44-A, N223, U16, U92	N239	3.8–6.6	86.0
Plant height (cm)	BZ-1, N239, U01, U27, U44-A, U47, U92	U16	N238	2.9–5.8	170.7
Ear height (cm)	U01, U47	U16, U21	N238, U92	2.3–4.1	82.8
Maturity (days to pollen)	N238, N239, U01, U12, U92	N373, U16, U80	–	0.6–1.3	75.3
Ear length (mm)	BZ-1, B5.62, N297, N373, N451, U44-A, U81	N239, U01, U104	–	3.5–7.6	15.7 cm
Ear diameter (mm)	N114, N223, N239, N373, U27, U92	N34	N391	0.9–2.3	4.2 cm
Kernel depth (mm)	N234, U80, U92	N223, N391	–	1.0–1.9	0.2 cm
Kernel rows (number per ear)	N297, N391	U66	–	0.5–0.6	15.8
Kernel number (per ear)	N239, N297, U01, U27	–	–	35.5–49.4	532.7
Grain weight (g per ear)	N114, N223, N234, N239, N297, N373, N391, U27, U92	U01	–	5.7–14.2	112.1
100-kernel weight (g)	N402, U44-A, U96	U01	–	0.8–2.5	21.4

[a] Marker abbreviations are 'B' for BNL, 'N' for NPI, and 'U' for UMC; BZ-1 hybridizes at the *bronze-1* locus. Horizontal dashes indicate no marker associations at the level of significance

[b] Estimates are for effects associated with a single marker

programs where the goal is to use marker-QTL information for selection is not clear. The stringent controls on Type I error necessary when the objective is to locate and clone a gene (such as in human genetics) may be unwarranted. The intent of this study is to form a basis for marker-assisted selection of a complex trait such as grain yield, and the emphasis is on detection and relative estimation of true effects, which may be small. Consequently, hypothesis tests at the α_c level are reported.

Reduced linkage disequilibrium in the F_2 of BS11(FR)C7 × FRMo17, caused by the failure of BS11(FR)C7 to be homozygous, increases the difficulty of detecting strong marker-QTL associations. If, for example, FRMo17 has the genotype $M_1M_1t_1t_1$ where M_1 is a band at marker 1 and t_1 is an unfavorable allele at QTL locus 1, then maximum disequilibrium in the F_2 would occur when BS11(FR)C7 is $m_1m_1T_1T_1$ and recombination between M and T is low. If BS11(FR)C7 is segregating, then the degree of F_2 disequilibrium is a function of the recombination value and product of the frequencies of m_1 and T_1 (J.W. Dudley, unpublished results). Therefore, detection of strong marker-trait associations implies close marker-QTL linkages and a high frequency of both marker bands and QTL alleles in BS11(FR)C7. Weaker linkages and lower frequencies would result in reduction of the levels of significance of association and estimates of true effects, or an inability to detect any statistical associations. The fact that a relatively high percentage of markers showed significant associations, and as much as 47% of the variation was

Table 4. Percentage of genotypic variation ($R^2 \times 100$) explained by RFLP markers and di-marker interactions with trait associations at $\alpha \leq 0.05$

Trait	Number of markers[a]	R^2 (×100) for linear effect models		
		Range for individual markers[b]	All markers[c]	All markers with di-marker interactions[d]
Grain yield	15	2.4– 7.1	30.0	42.3
Grain moisture	13	2.3– 7.4	40.3	47.4
Stalk lodging	9	1.9– 7.9	30.4	30.5
Root lodging	8	1.9– 7.3	18.5	–
Plant height	9	1.9– 7.0	25.3	–
Ear height	6	2.1– 7.1	20.5	–
Maturity	8	2.0– 4.3	17.4	–
Ear length	10	1.9– 4.6	22.3	30.9
Ear diameter	8	2.0–10.8	18.7	20.7
Kernel depth	5	1.9– 6.4	13.9	–
Kernel rows	3	1.9– 4.1	8.4	9.2
Kernel number	4	1.9– 2.7	8.8	–
Grain weight	10	2.2– 4.2	17.9	18.0
100 kernel weight	4	2.1– 3.8	11.0	–

[a] Number of RFLP markers associated at $\alpha \leq 0.05$
[b] The linear contrast representing each marker
[c] Multiple linear regression model including the contrast for each marker
[d] Multiple linear regression model including the contrast for each marker plus di-marker interactions ($\alpha \leq 0.05$) among them. Numbers of interactions included in the model were eight for grain yield, three for grain moisture, two for stalk lodging, five for ear length, and one each for ear diameter, kernel rows, and grain weight. Horizontal dashes indicate that no interactions were detected at $\alpha \leq 0.05$

accounted for by markers with significant effects (Table 4), suggests a relatively high degree of disequilibrium in the F_2.

The presence of multiple bands at a marker presents new problems in statistical analyses aimed at identifying RFLP-QTL associations. Multiple bands do not necessarily mark multiple QTL alleles. If an RFLP marker contains n bands and n-1 are linked to the same QTL allele, then n (n-1)/2 marker classes are expected to be associated with similar phenotypic expression. In a single factor ANOVA, this would reduce both variability among marker classes and the power of an F-test for significance. Analyses using linear contrasts among marker classes address this problem by testing effects associated with substitution of each marker band and pooling those that are similar.

Analysis of linear effects associated with changes in marker band representation is appropriate from a genetic standpoint. Although RFLP data were collected from F_2 individuals, measurements of agronomic traits were made in $F_2 \times$ FRB73 testcross progeny. Therefore, at a single marker, changes in phenotypic expression from one homozygous F_2 marker class to another are measures of the effect associated with substitution of bands in a common genetic background (FRB73). This is twice the average effect of a gene, as defined by Falconer (1981). For example, an RFLP marker with two bands would have three marker classes among F_2 individuals: AA, AB, BB. However, actual marker classes in testcross progeny are AH, ½ AH + ½ BH, and BH, where H = the FRB73 band. Any change in phenotypic expression from AH to BH should be additive only, regardless of the dominance relationship of H with A or B. Because A and B are expected to have equal representation in testcross progeny derived from F_2 individididuals of marker class AB, phenotypic information from this class should also be used when estimating linear effects. The same principles apply for a marker with multiple bands, except that sequential linear contrasts can be made testing the substitution of each band for all others, pooled as a group. Because a computer simulation showed that pooling marker classes of unequal size in a completely linear model can result in significant quadratic effects (data not shown), significant non-linear effects in ANOVAs were disregarded as artifacts. Linear contrasts were weighted by number of observations to account for unequal size among pooled marker classes.

Favorable bands at RFLP markers and multiple trait associations

Estimation of linear effects at a marker with multiple bands allows for grouping of bands associated with similar QTL effects. Because each band is contrasted against the grouping of all others at a marker, those associated with similar trait effects will have statistically similar linear regression coefficients (see Appendix). Table 2 shows linear substitution effects of all bands at each marker associated with grain yield at $\alpha \leq 0.05$. For markers with multiple bands, vertical bars indicate regression coefficients for specific bands not significantly different by t-test at $\alpha \leq 0.10$. Marker bands in the same t-group were interpreted as being associated with QTL alleles having similar effects. No t-groupings are necessary for markers with two bands.

Grouping of similar linear effects also has implications for marker-assisted selection. When bands associated with the most favorable effects have statistically similar regression coefficients, then selection should be against bands not in the t-group. For example, selection for increased grain yield at marker NPI238 should be against band C and for any of the other three bands (Table 2). When the most favorable band is not in a group, then selection should be for that band only. Use of a less stringent Type I error rate ($\alpha \leq 0.10$) for acceptance of pair-wise differences among regression coefficients is conservative from a selection standpoint. Asso-

ciated effects of bands are not as likely to be declared statistically similar and grouped together as favorable. Therefore, selection based on marker bands is less likely to be affected by mistaken groupings of unequal effects.

Markers associated with expression of multiple traits can be seen among Tables 2 and 3. At $\alpha \leq 0.05$, marker NPI297 was associated with grain yield, four of seven grain yield component traits, and stalk lodging. Favorable band designations were: A, B, and D for increased grain yield; B and/or D for increased yield component values; A for reduced stalk lodging. These data suggest that QTL alleles for component traits and standability combine to influence the expression of machine-harvestable grain yield. Other markers associated with both grain yield and several component traits were NPI114, NPI234, NPI391, UMC27, UMC81, UMC92, and UMC96. In most cases, the same bands detected favorable effects for all associated traits. Common favorable marker bands also were seen across other traits. At NPI238, the same bands were favorable for decrease in lodging (stalk and root), height (plant and ear), grain moisture at harvest, and early maturity. This type of information will be useful in marker-directed selection for improvement of multiple traits.

Bands unique to BS11(FR)C7 and associated with favorable effects were found at markers for all traits except ear diameter (Table 5). For grain yield, favorable bands at UMC92 (3-S), UMC16 (3-L), UMC96 (3-L), BNL15.07 (4-L), and UMC21 (6-L) were from BS11(FR)C7 only; UMC16 had the highest degree of grain yield association and one of the largest estimated effects (Table 2). For the other ten markers associated with grain yield at $\alpha \leq 0.05$, at least one favorable band was representative of the FRMo17 parent. Markers identifying favorable QTL alleles from BS11(FR)C7 may be useful in selecting F_2-derived lines or backcross progeny with increased grain yield over FRMo17 in crosses to FRB73. More data will be needed to determine whether or not associated effects are consistent for other inbred testers.

Digenic epistasis

Significant ($\alpha \leq 0.05$) interactions between linear contrasts representing pairs of markers were detected for grain yield, grain moisture, stalk lodging, and four grain yield components. For most traits, the number of interactions compared to the total number possible was less than expected by chance. However, epistasis associated with marker pairs appeared to be important in the expression of grain yield and ear length. The total number of interactions for both traits was slightly more than expected by chance (8 of 105 for grain yield, 5 of 45 for ear length). For grain yield, genotypic variation explained by the genetic model increased 12% when di-

Table 5. RFLP markers with trait-associated bands ($\alpha \leq 0.05$) unique to BS11(FR)C7 and favorable over those from FRMo17 in combination with FRB73[a]

Trait	Markers[b]
Grain yield	BNL15.07, UMC16, UMC21, UMC92, UMC96
Grain moisture	NPI114, NPI238, NPI297, UMC44-A, UMC44-B
Stalk lodging	BNL15.07, NPI238, NPI373, NPI451, NPI455
Root lodging	NPI223, NPI238, NPI239, NPI391
Plant height	NPI238, NPI239, UMC01, UMC27
Ear height	NPI238, UMC01, UMC21
Maturity	NPI238, NPI239, UMC01, UMC80
Ear length	NPI297, NPI451, UMC44-A, UMC81
Ear diameter	–
Kernel depth	UMC80
Kernel rows	NPI297, UMC66
Kernel number	NPI297
Grain weight	NPI297
100-kernel weight	NPI402, UMC44-A

[a] BS11(FR)C7-derived bands showed associations favorable for increase in grain yield and value of component traits; decrease in grain moisture at harvest, lodging and height; and early maturity

[b] No markers associated with ear diameter showed favorable bands unique to BS11(FR)C7

marker interactions were added to a multiple linear model containing all markers individually associated at $\alpha \leq 0.05$ (Table 4); the increase was more than 8% for ear length. Representation of markers in the eight significant interactions for grain yield appeared to be nonrandom; NPI234 (1-S) and UMC21 (6-L) were each involved in four interactions, once with each other. Only two markers (UMC81, UMC104) showed significant interactions for both grain yield and ear length.

Acknowledgements. This research was supported by a grant from ICI Seeds, UK, and Garst Seed Co., USA. All laboratory procedures were preformed by ICI-Garst. The assistance of Drs. I. Bridges, K. Rufener II, K. Russell, and D. Foster at Garst Seed Co., Dr. S. Barnes and Jacqueline Brannigan at ICI Seeds, and Dr. S. Carmer at the University of Illinois is gratefully acknowledged.

Appendix

The association of RFLP marker UMC16 with grain yield can be used as an example for calculating linear contrasts among marker classes, to estimate effects and test for bands associated with similar effects. For each of the three bands (A, B, C) at UMC16, indicator values were given to testcross observations in F_2 marker classes according to the presence or absence of the band.

F_2 marker classes	Marker bands			N	Average grain yield of F_2 testcross progeny (kg ha^{-1})
	A	B	C		
AA	1.0	0.0	0.0	116	7187
AB	0.5	0.5	0.0	50	7374
AC	0.5	0.0	0.5	36	6953
BB	0.0	1.0	0.0	9	7589
BC	0.0	0.5	0.5	6	7521
CC	0.0	0.0	1.0	3	7358

Regressions of grain yield on their corresponding indicator values were calculated, measured in testcross progeny from F_2 individuals. Regressions were weighted by the number of observations (N) per indicator value (pooled marker class). The ANOVA for each linear contrast using Mg ha^{-1} of grain yield in sum of squares calculations was:

ANOVA for UMC16 – band A					
Source of variation	Degrees of freedom	Sum of squares	Mean square	F-value	Prob. $>F$
Marker classes (pooled)	2	1.87			
linear	1	0.95	0.95	3.85	NS
residual	1	0.92	0.92	3.74	NS
Error	217	53.44	0.25		
ANOVA for UMC16 – band B					
Marker classes (pooled)	2	3.92			
linear	1	3.91	3.91	16.51	0.0001
residual	1	0.01	0.01	0.07	NS
Error	217	51.38	0.24		
ANOVA for UMC16 – band C					
Marker classes (pooled)	2	1.81			
linear	1	1.11	1.11	4.49	0.0353
residual	1	0.70	0.70	2.84	NS
Error	217	53.50	0.25		

Hypothesis test were conducted at $\alpha \leq 0.05$ in each ANOVA. The linear contrast for band B was chosen to represent UMC16 because it had the lowest probability value. A linear regression coefficient was calculated in each contrast to estimate the grain yield effect associated with substitution of the corresponding band. Significant differences between pairs of coefficients were determined by t-test ($\alpha \leq 0.10$), taking into account the variance of coefficient estimates as well as covariances among them. Estimated effects associated with substitution at bands A and C were not significantly different (NS).

UMC16 band	Estimate of associated effect (kg ha^{-1})		SE	Prob. $>F$
B	+483		119	0.0001
A	−205	NS	105	NS
C	−319		151	0.0353

The estimated effect associated with substituting band B for the others was favorable for increased grain yield (+483 kg ha^{-1}), and B was unique to the donor population BS11(FR)C7.

One possible interpretation of the behavior of grain yield expression associated with substitution of bands at marker UMC16 is as follows: bands A and C may be linked to QTL allele(s) with similar unfavorable effects, while band B is linked to different QTL allele(s) having a relatively favorable effect. When B is contrasted against A and C, a strong linear effect associated with substituting B for the other bands is detected. When either band A is contrasted against C and B, or band C is contrasted against A and B, the effect of substition is reduced because bands associated with similar QTL effects are being contrasted against each other. The linear coefficients calculated for bands A and C are statistically similar because effects of alleles associated with A and C are common in their relationship to effect(s) of allele(s) associated with B.

References

Abler BSB, Edwards MD, Stuber CW (1991) Isoenzymatic identification of quantitative trait loci in crosses of elite maize inbreds. Crop Sci 31:267–274

Carmer SG, Kratzke DT (1988) A user's guide to PCAGRON: personal computer FORTRAN programs for the statistical design and analysis of agronomic research experiments. University of Illinois, Dept. of Agronomy, Stat. Lab, Tech. Rep. No. 29

Cowen NM (1988) The use of replicated progenies in marker-based mapping of QTL's. Theor Appl Genet 75:857–862

Dudley JW (1988a) Theory for identification of lines or populations useful for improvement of elite single crosses. In: Proc 2nd Int Conf Quantitative Genet, Raleigh/NC, May–June, 1987. Sinauer Associates, Sunderland/MA, pp 451–461

Dudley JW (1988b) Evaluation of maize populations as sources of favorable alleles. Crop Sci 28:486–491

Edwards MD, Stuber CW, Wendel JF (1987) Molecular marker-facilitated investigations of quantitative trait loci in maize. I. Numbers, genomic distribution and types of gene action. Genetics 116:113–125

Falconer DS (1981) Introduction to quantitative genetics, 2nd edn. Longman Scientific & Technical, London

Feinberg AP, Vogelstein B (1984) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity (Addendum). Anal Biochem 137:266–267

Guffy RD, Stuber CW, Edwards MD (1989) Dissecting and enhancing heterosis in corn using molecular markers. In: Proc 25th Annu Illinois Corn Breeders School, Champaign/IL, March, 1989. University of Illinois, Urbana-Champaign/IL, pp 99–119

Hallauer AR, Eberhart SA, Russell WA (1974) Registration of maize germplasm. Crop Sci 14:341–342

Helentjaris T, King G, Slocum M, Siedenstrang C, Wegman S (1985) Restriction fragment polymorphisms as probes for plant diversity and their development as tools for applied plant breeding. Plant Mol Biol 5:109–118

Henderson C (1976) Maize research and breeders manual. No. VIII. Illinois Foundation Seeds, Champaign/IL

Kahler AL, Wehrhahn CF (1986) Associations between quantitative traits and enzyme loci in the F_2 population of a maize hybrid. Theor Appl Genet 72:15–26

Kahler AL, Hallauer AR, Gardner CO (1986) Allozyme polymorphisms within and among open-pollinated and adapted exotic populations of maize. Theor Appl Genet 72:592–601

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199

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

SAS Institute (1987) SAS/STAT Guide for personal computers. Version 6 ed. SAS Institute, Cary/NC

Smith JSC, Goodman MM, Stuber CW (1985) Genetic variability within U.S. maize germplasm II. Widely-used inbred lines 1970 to 1979. Crop Sci 25:681–685

Steel RDG, Torrie JH (1980) Principles and procedures of statistics: a biometrical approach, 2nd edn. McGraw-Hill, New York

Stuber CW (1992) Biochemical/molecular markers in plant breeding. Plant breeding reviews, Vol. 9. Wiley Press (in press)

Stuber CW, Goodman MM (1983) Allozyme genotypes for popular and historically important inbred lines of corn, *Zea mays* L. USDA-ARS, ARR-S-16. US Govt Printing Office, Washington, DC

Stuber CW, Edwards MD, Wendel JF (1987) Molecular marker-facilitated investigations of quantiative trait loci in maize. II. Factors influencing yield and its component traits. Crop Sci 27:639–648

Tanksley SD, Young ND, Paterson AH, Bonierbale MW (1989) RFLP mapping in plant breeding: new tools for an old science. Bio/Technol 7:257–264