

Received 2 March 2002
Accepted 3 June 2002

**A MICROSATELLITE MARKER BASED STUDY OF
CHROMOSOMAL REGIONS AND GENE EFFECTS ON YIELD AND
YIELD COMPONENTS IN MAIZE**

SAYYED ABOLGHASEM MOHAMMADI^{1*}, BODDUPALLI MARUTI
PRASANNA², CHARU SUDAN³ and NARSINGH NARAIN SINGH³

¹Dept. of Crop Production & Breeding, Faculty of Agriculture, University of
Tabriz, Tabriz 51664, Iran, ²Division of Genetics, Indian Agricultural Research
Institute, New Delhi 110012, India, ³Directorate of Maize Research, New Delhi,
India

Abstract: The genetic basis of different traits and the effect of chromosomal regions on the expression of such traits is the primary interest of this study. In the present investigation, the effect of chromosomal regions on yield and its two primary components was studied, and two genetic parameters, additive effect and degree of dominance, were estimated based on a single-locus model using markers with significant effects on the traits (informative markers). Eight inbred lines from diverse geographical regions of India were crossed in all possible pairs, and F₁s were evaluated for the yield and yield components in a replicated trial in two environments. Sequential path analysis was employed to find those yield components showing significant direct effects on yield with negligible multicollinearity. The parental lines were profiled using 56 polymorphic SSR (Simple Sequence Repeat) markers covering 10 chromosomes of maize. Stepwise multiple regression analysis was used to determine the informative markers on yield and its primary components detected through path analysis. Fourteen markers were found to have association with chromosomal regions showing significant effects on the total grain yield, 100-grain weight and total number of kernels per ear. Chromosome 1 with four informative markers revealed the highest genic effects on yield and its components. Markers bnlg594 and bnlg1360 on chromosome 10, and bnlg147 on chromosome 1 revealed the highest additive effects on the total grain yield, 100-grain weight and total number of kernels per ear, respectively. For the analyzed traits, overdominance occurred in all the loci and d/a values, varying from 8.60 for 100-grain weight to 1.40 for total grain yield.

* Corresponding author, E-mail: sa_mohammadi@yahoo.com

Key Words: Chromosomal Regions, Gene Effects, Informative Markers, Maize, Microsatellite Markers, Yield Components

INTRODUCTION

Estimation of the types of gene action involved in the expression of traits, the level of additive effects, and the degree of dominance are very important in developing a breeding method for the trait of interest [1]. Alleles with dominance or additive phenotypic effects influence heritability differently, depending on whether they are in homozygous or heterozygous conditions. Knowledge of the way genes act and interact will determine which breeding system can optimize gene action more efficiently, and will elucidate the role of breeding systems in the evolution of crop plants. Liu and Wu [2] strongly emphasized that neither genetic diversity nor heterozygosity is a good indicator for predicting heterosis. They suggested that, in the detection of chromosomal regions influencing yield and the prediction of hybrid performance, focus should be on 'gene combinations' rather than on single genes (or quantitative trait loci or QTLs). However, the inheritance of such combinations is closely related to the genetic background of the parental lines.

With the advent of polymerase chain reaction (PCR) technology, there has been an explosion of DNA markers that made it possible to solve various problems in plant genetics. Microsatellites or simple sequence repeats (SSRs) are one of the PCR-based molecular marker systems that have been extensively used in plant genetic studies in recent years [3-6]. Molecular techniques have been suggested to provide an efficient approach for the detection of chromosomal regions affecting different characteristics [4-7]. Bernardo [8] showed that criteria such as strong dominance effects and negative correlation between the allele frequencies of individual loci in parental inbreds are important in the effective prediction of hybrid performance using molecular heterozygosity. It was found that at least 30-50% of QTLs have to be linked to molecular markers, or that not more than 20-30% of molecular markers have to be randomly dispersed or unlinked to QTLs [8, 9]. Some researchers used a molecular marker based analysis of variance [4, 5] or multiple regression analysis [7] to identify the chromosomal regions associated with phenotypic variation in the traits of interest. This investigation used SSR markers on parental lines of a maize diallel cross to estimate the type of gene action on yield and its primary components and to determine the chromosomal region affecting the expression of these characteristics in the F₁ progeny.

MATERIALS AND METHODS

Genetic materials

Eight inbred lines from diverse geographical regions of India were used in this investigation: CM115, CM116, CM117, CM128, CM132, CM138, CM139, and NAI147. The lines were crossed in all possible pairs, and F₁s were evaluated for

yield and yield components in a replicated trial in Delhi during the monsoon season, and in Hyderabad during the winter season. The trials were carried out in a randomized block design with two replications per genotype, each plot consisting of two rows (5m length) with approximately 25 plants per row. Data for 17 morpho-agronomic characteristics were recorded on five representative plants for various characteristics for each genotype and for each replication.

DNA extraction and SSR assay

For each parental line, genomic DNA was extracted from two grams of young leaves, harvested in bulk from 10-15 four to five week old plants per genotype. The CTAB procedure [10] was used, with minor modifications. Agarose gel electrophoresis was carried out for DNA quantity and quality analyses, as per the standard procedures. Quantification of each sample was based on ethidium bromide fluorescence in comparison with uncut λ DNA samples of known quantity run on the same gel. This method also facilitated analysis of the DNA quality. PCR reactions were performed in a 10 μ l reaction mix consisting of 10ng template DNA, 0.1 mM dNTPs, 0.25 μ M SSR primers (forward + reverse), 0.75 mM MgCl₂, 0.025 U Taq DNA polymerase and 1X reaction buffer (10 mM Tris-HCl, 50 mM KCl, 1.5 mM MgCl₂). Controls run with each amplification included at least one sample of the reaction mix with no template DNA. PCR cycling consisted of initial denaturation at 94°C for 4 min, followed by 35 cycles of amplification, each having denaturation at 94°C for 1 min, annealing at 58-65°C (depending upon the primer) for 1 min, and extension at 72°C for 2 min. A final extension step at 72°C for 7 min was followed by termination of the cycle at 4°C. PCR amplified products were resolved using a 6% polyacrylamide sequencing gel followed by silver staining. Φ x-174/HinfI digest was used as a molecular weight standard in each run for allele scoring and designation. Inbreds revealing 'null' alleles or missing data, if any, for a specific SSR locus were included in the re-analysis for confirmation.

Statistical analysis

Field data were analyzed based on a linear model according to the experimental design employed. The sequential path model was used to determine those yield components with high significant effects on grain yield. These components had non-significant multicollinearity parameters. The discrimination power of each SSR marker was evaluated by the polymorphism information content (PIC) as $PIC = 1 - \sum p_{ij}^2$, where p_{ij} is the frequency of the i^{th} allele at the j^{th} SSR locus. For those genotypes showing heterozygosity for a specific SSR locus, PIC values were calculated by considering each allele from those genotypes as contributing one-half instead of one. The effect of chromosomal region on a characteristic was assessed based on informative markers. To identify the possible informative markers amongst the 56 polymorphic SSR markers dispersed across different chromosomes, a multiple stepwise regression analysis using the BACKWARD option in SPSS-10.0 was employed. Each quantitative characteristic in the

analysis was treated as a dependent variable and SSR alleles (scored as '1' for presence and '0' for absence) as independent variables. First, a full model using all the markers was assessed and the markers with significant regression coefficients were retained in the model. The model fitting was continued until all the markers with significant effects on traits were identified as informative markers. Two genetic parameters, additive effect and degree of dominance, were estimated using a single locus model for each informative marker. The additive (a) and dominance effects (d) for each locus were calculated using formula put forward by Saghai-Marooft *et al.* [4]:

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

where X_{ii} is the average character values of individuals homozygous for the i^{th} allele of the SSR locus, N_{ij} is the number of F_1 s heterozygous for the i^{th} and j^{th} alleles and k is the total number of alleles at the locus. The degree of dominance for a locus is given by d/a . Cluster analyses were performed to group parental lines using several cluster algorithms including UPGMA, complete linkage, and Ward's method. Three measures of similarity between parental lines were used, namely, Jaccard's coefficient, Nei and Li's coefficient, and the simple matching coefficient.

RESULTS

Significant differences in total grain yield (TGY) and its twelve component characteristics were observed between hybrids in each environment and in the results of the combined analysis. Sequential path analysis revealed that the 100-grain weight (100GW) and the total number of kernels per ear (TNK) had the highest direct effects, (surely these two factors are merely part of TGY, and do not have an effect on it per se – climatic conditions that may have an effect on TGY, TNK could be a significant portion of TGY) on TGY (0.74 and 0.78, respectively) and accounted for nearly 97% of the variation in TGY. The effects of these components on TGY were multiplicative. Among the 28 crosses under analysis, the highest TGY value was observed for those crosses that had CM139 as one of the parental lines, with either CM117 or CM138 as the other parent.

Out of the eighty primers screened, fifty-six polymorphic SSR markers with constant amplification products yielded 229 SSR alleles, with the lowest number (three markers) on chromosome 7, the maximum (eight markers) on chromosome 2, and an average of 5.8 markers per chromosome. The average number of SSR alleles per locus was 4.01, with a range from two alleles (in

bnlg1633, phi099, phi072, bnlgl155, umc1588, phi117) to eight alleles (in dupssr17). The presence of nine alleles in bnlgl121 was due to the heterozygosity of some of the parental lines at this locus. The PIC values of the various SSR loci ranged from 0.219 (bnlg155) to 0.867 (bnlg1614) with an average of 0.635 ± 0.152 . Seven SSR loci with PIC values greater than 0.800 were identified.

Analysis of all the resulting dendrograms from different algorithms revealed that the genetic relationships between the inbreds were maintained irrespective of the clustering algorithms and genetic distances. The analyses classified the inbred lines into three groups (Fig. 1). The high divergence of CM139 from the two inbreds, CM117 and CM138, was also reflected in performance of their cross combinations, which were the top two high-performing hybrids in the diallel analysis.

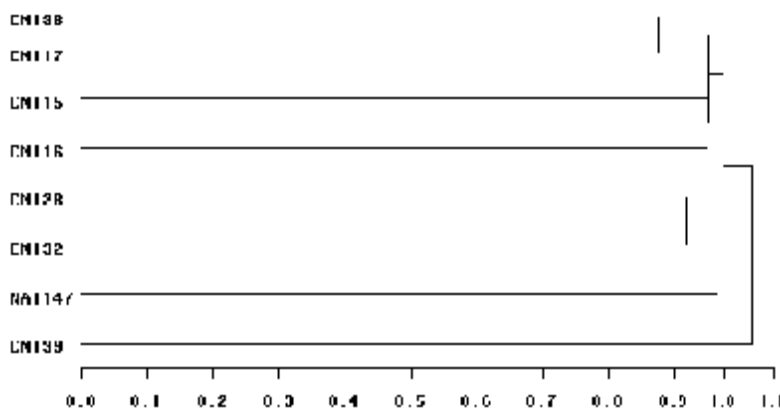


Fig. 1. Dendrogram depicting the genetic relationships between parental lines involved in diallel analysis, based on SSR data using UPGMA (Jaccard's coefficient).

A total of 14 informative SSR markers associated with TGY or its two component characteristics were detected distributed across all the chromosomes, except for chromosomes 7 and 8 (Table 1). Chromosome 1, with 4 markers, had the highest recorded R^2 values of the model (0.98). Bnlgl594 on chromosome 10, bnlgl360 on chromosome 10, and bnlgl147 on chromosome 1 had the highest additive effects on TGY, 100GW and TNK, respectively. Overdominance occurred in all the loci and d/a values varying from 8.60 for 100GW to 1.40 for TGY. The average d/a values for TGY, 100GW and TNK were 3.094, 3.56 and 2.58, respectively. The umc1689 revealed the highest overdominance effect on TGY and 100GW, whereas the highest overdominance effect for TNK was observed with bnlgl615. The significant effects detected by many of the informative markers were due to the outstanding performance of hybrids that had CM139 as one of their parents, for which the marker genotypes were also different from remaining parental lines.

Tab. 1. Chromosomal location, number of alleles, PIC^{*} values, additive effects and degrees of dominance estimated for informative markers (marker loci that had a high significant association with total grain yield and its two primary components) in a diallel set of eight Indian maize inbreds

Marker	Chromo- somal location	Num- ber of allele	PIC ¹	TGY ²		100GW ³		TNK ⁴	
				a ⁵	d/a ⁶	a	d/a	a	d/a
bnlg147	1	3	0.531	14.96	1.85			61.81	1.62
umc1689	1	3	0.664	7.13	5.37	0.47	8.60	34.90	3.87
bnlg615	1	5	0.774	7.64	4.51	1.20	3.92	26.01	4.76
umc331	1	3	0.603			1.30	3.13		
bnlg198	2	7	0.827	11.32	3.01			48.26	2.45
bnlg1018	3	4	0.597	11.37	1.92	1.32	1.53	50.02	1.42
bnlg1917	4	4	0.555	7.13	3.17			30.97	3.19
bnlg1836	5	4	0.694	7.69	3.48			41.33	2.54
bnlg1346	5	4	0.693	12.14	3.48	0.97	2.67	59.02	1.63
dupssr15	6	6	0.792					58.89	1.81
phi070	6	3	0.508			1.26	2.77		
dupssr9	9	4	0.602	10.31	2.93			10.01	2.33
bnlg594	10	7	0.857	21.71	1.40	1.29	2.89	42.20	2.63
bnlg1360	10	6	0.781	10.81	3.30	1.36	3.00	46.94	2.69

¹Polymorphic information content; ²Total grain yield per ear; ³100 grain weight; ⁴Total number of kernel per ear; ⁵Additive effect; ⁶Degree of dominance.

DISCUSSION

In maize, heterosis over mid-parents or the better parent is a common phenomenon. Therefore, the type of gene actions involved in hybrid performance and their relative contribution to the expression of heterosis is of particular importance in the establishment of an appropriate breeding methodology for the better utilization of hybrid vigor in commercial hybrid maize. The results indicated that the amount of heterosis for yield was higher than that for its components. This is understandable, because yield is a multiplicative function of its components. Several studies have also indicated the importance of multiplicative gene action versus arbitrary gene action in the expression of heterosis [5, 11]. Due to the interaction between marker loci because of the small number of parental lines involved in any given analysis, a possible problem in the estimation of single gene effect is the frequently observed sporadic correlations between marker loci which cause false positives in the detection of informative markers [4, 5]. To overcome this problem, besides the high degree of stringency that was adopted in the regression analysis,

a separate analysis was performed for each chromosome to eliminate the epistatic effects of other loci.

One of the hypotheses for the explanation of hybrid vigor is the overdominance hypothesis which assumes the existence of loci at which the heterozygote is superior to either homozygote. The analysis of gene action for informative markers in this study revealed the importance of overdominance gene action in maize for the expression of heterosis for yield and its components. The QTL studies in maize for yield and its components also indicated that the majority of yield and yield-related QTLs were associated with overdominance [12]. These results suggest that these QTL regions may play especially important roles in the phenomenon of heterosis. One possible explanation is that maize actually possesses a large number of genes for which alleles interact in a truly overdominant manner. In this study, the distinction between overdominance and pseudo-overdominance or the occurrence of dominant and recessive alleles in the coupling phase at closely-linked loci was not possible. It requires extensive recombinational separation of linked loci or fine mapping of QTLs displaying overdominance gene action.

The identification of QTLs affecting agronomically important traits is a key step in plant breeding programs using molecular markers to understand genetic phenomena such as heterosis. One of the important issues in QTL mapping is to choose an appropriate mapping population to efficiently identify QTLs with different gene action. Therefore, understanding the nature of the underlying gene action before QTL mapping experiments has great importance in planning an appropriate experimental design. Marker-based diallel analysis was used to detect the types of gene action involved in the expression of different characteristics [4, 5]. Informative markers identified using multiple regression analysis could correspond to those chromosomal regions containing genes of interest, although many of the tests in the analysis may not be independent of each other due to the small number of parents in the diallel set. It should be noted that the number of informative markers for a specific characteristic would differ from situation to situation depending on the germplasm analyzed and the relationships between the parents used in crosses. However, the results of such experiments could not only be useful in planning breeding programs for the improvement of the characteristics of interest, but could also provide preliminary information for QTL mapping experiments and efficient marker-assisted selection. However, it remains to be seen whether such informative markers can be found with reasonable experimental costs, and if the identified linkage relationships will exist across a wide range of germplasm within heterotic groups. In summary, the results of this investigation suggest that molecular marker-based estimation of gene effects controlling traits of interest can provide a useful and precise means for the establishment of appropriate breeding methods and preliminary information for accurate QTL mapping experiments.

REFERENCES

1. Hallauer, A.R. and Miranda, J.B. **Quantitative Genetics in Maize Breeding**. Iowa State Univ. Press, Ames, Iowa, (1988).
2. Liu, X.C. and Wu, J.L. SSR heterotic patterns of parents for making and predicting heterosis. **Mol. Breed.** 4 (1998) 263–268.
3. Senior, M.L. and Heun, M. Mapping maize microsatellites and polymerase chain reaction confirmation of the targeted repeats using a CT primer. **Genome** 36 (1993) 884–889.
4. Saghai Maroof, M.A., Zhang, Q. and Biyashev, R.M. Molecular marker analysis of powdery mildew in barley. **Theor. Appl. Genet.** 88 (1994) 733–740.
5. Zhang, Q., Gao, Y.J., Yang, S.H., Ragab, R., Saghai Maroof, M.A. and Li, Z.B. A diallel analysis of heterosis in elite hybrid rice based on RFLPs and microsatellites. **Theor. Appl. Genet.** 89 (1994) 185–192.
6. Vuylsteke, M., Kuiper, M. and Stam, P. Chromosomal regions involved in hybrid performance and heterosis: Their AFLP-based identification and practical use in prediction models. **Heredity** 85 (2000) 208–218.
7. Virk, P.S., Ford-Lloyd, B., Jacson, M.T., Pooni, H.S., Clemeno, T.P. and Newbury, H.J. Predicting quantitative variation within rice germplasm using molecular markers. **Heredity** 76 (1996) 296–304.
8. Bernardo, R. Relationship between single cross performance and molecular marker heterozygosity. **Theor. Appl. Genet.** 83 (1992) 628–634.
9. Lee, M. DNA markers for detecting genetic relationships among germplasm and for establishing heterotic groups. **Lecture During Maize Training Course**, CIMMYT, Mexico (1998).
10. Saghai Maroof, M.A., Soliman, K.M., Jorgensen, R.A. and Allard, R.W. Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location and population dynamics. **Proc. Natl. Acad. Sci. USA** 81 (1984) 8014–8018.
11. Schnell, F.W. and Cockerham, C.C. Multiplicative vs. arbitrary gene action in heterosis. **Genetics** 131 (1992) 461–469.
12. Stuber, C.W., Lincoln, S.E., Wolff, D.W., Helentjaris, T. and Lander, E.S. Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. **Genetics** 132 (1992) 823–839.