

THE USE OF MORPHOLOGICAL TRAITS AND GENETIC MARKERS TO ESTIMATE GENETIC RELATIONSHIPS IN MAIZE

E. Nagy and L. C. Marton

Agricultural Research Institute of the Hungarian Academy of Sciences, Martonvásár

Summary

The use of genetic markers allows the study of polymorphism and genetic distances between maize lines in greater depth than can be achieved on the basis of phenotype and DUS traits. The analysis of polymorphism between 46 maize inbred lines with known genetic background and the classification of these lines in related groups was carried out by means of morphological description, isoenzyme analysis, RAPD analysis, and identification using gene-linked microsatellite (SSR) markers. The genetic distance or degree of relationship between the lines was determined using cluster analysis. Only a very limited extent of allele polymorphism could be detected in isoenzyme analyses; the 46 lines formed only 18 gel electrophoresis groups. Nevertheless, on the basis of RAPD and SSR markers, all the lines could be distinguished from each other. This was reflected by the PIC (polymorphism index content) values, which ranged from 0.04 to 0.55 (mean 0.27) for the various enzyme loci, while far higher values were obtained for RAPD and SSR markers (0.20–0.91, mean 0.61, and 0.54–0.90, mean 0.73, respectively). Due to the large number of lines, two lines, derived from each other or from common parents, were chosen from each related group as the basis for grouping the lines according to genetic background. It was found that, while the individual marker systems only partially reflected the actual relationships between the lines, a joint processing of the genetic markers, supplemented with morphological data, revealed a close correlation between the groups formed on the dendrogram and the genetic background.

Index words: maize, polymorphism, isoenzyme, RAPD, SSR, cluster analysis

Introduction

The CPVO (Community Plant Variety Office) guidelines elaborated for the registration and Plant Variety Protection (PVP) of maize varieties are based on the criteria of distinctness, uniformity and stability (DUS). However, a description based on morphological traits is often not sufficient to satisfy these conditions. One of the main fields where the detection of differences between varieties is required is the establishment of intellectual property rights, so it is essential to use techniques, which allow the varieties to be distinguished accurately and reliably (Smith and Smith, 1989b). The rapid development of marker technologies has led to their widespread use in various fields of plant breeding, including maize breeding. Molecular markers have been used to facilitate the cataloguing genetic stocks, to measure genetic variability and genetic distances, and to provide variety descriptions for identification and patenting purposes (Smith and Senior, 1999). Identification techniques based on molecular markers form four main groups, depending on whether they are based on isoenzyme (Markert and Moller, 1959), RFLP (Botstein *et al.*, 1980), AFLP (Vos *et al.*, 1995) and PCR (Mullis and Faloona, 1987) analyses.

The CPVO regulations now make it compulsory to carry out starch gel electrophoresis to identify the isoenzyme patterns of maize varieties, based in most cases on the enzymes of the citrate and pentose phosphate cycles (Bourgoin-Greneche and Lallemand, 1993). As the genetic background of enzyme-coding genes, their location on the chromosomes and the possible alleles are all well known, a study of these genes can be used in most cases to check all three criteria included in the DUS tests (Stuber *et al.*, 1988; Goodman and Stuber, 1983; Smith and Smith, 1992). Although in some cases the isoenzymes possess several loci, with several alleles at each locus, they nevertheless represent only a fraction of the whole genome, so the information obtained in these analyses often falls short of that required from a full

genetic analysis. The practical application of isoenzyme polymorphism is also limited by the fact that certain alleles occur with considerably greater frequency than others, while the prediction of heterosis is impossible on the basis of allele polymorphism (Hunter and Kannenberg, 1971; Price *et al.*, 1986; Lamkey *et al.*, 1987).

Grouping of maize varieties into related groups is of great importance in breeding, since genetic distances between parental lines is a basic criterion for heterosis breeding. Molecular techniques which, directly or indirectly, map the genetic background of individual plants are indispensable for this work (Smith *et al.*, 1997; Pejic *et al.*, 1998).

In the course of the present studies, in addition to morphological description and isoenzyme analysis, 46 maize inbred lines with known genetic background were analysed using PCR-based techniques involving RAPD and gene-linked microsatellite (SSR) markers. The aim of the work was to determine how efficiently and in what combinations could these marker systems be used to demonstrate genetic polymorphism between maize inbred lines and to map relationships between the lines.

Materials and methods

Plant material: 46 maize inbred lines with known pedigrees, chosen from major germplasm groups were studied: Lancaster, Iodent, Iowa Stiff Stalk Synthetic (ISSS), Mindszentpuszta Yellow Dent (MYD) and OP Lacaune. Additional lines which cannot be classified in any of these above mentioned groups, whose origin is precisely known were also included: Argentinean flints, early Canadian lines, lines related to W 117 and derivatives of Co 125. Due to the large number of lines, two lines, derived either from each other or from common parents, were chosen from each related group for the analysis of degrees of relationship.

Morphological description: The lines were sown, one to a row, in small-plot experiments set up in a random block design with four replications at two locations in Hungary in two years. The scored traits were determined from the mean of 10 plants per plot, as described by the European Union CPVO-TP/2/2 (2001) and Smith and Smith (1989a).

Isoenzyme analysis: MDH (maleic acid dehydrogenase, 6 loci), IDH (isocitrate dehydrogenase, 2 loci), PGD (6-phosphogluconate dehydrogenase, 2 loci), PGI (phosphoglucose isomerase, 1 locus), PGM (phosphoglucumutase, 2 loci), ACP (acidic phosphatase, 1 locus) and ADH (alcohol dehydrogenase, 1 locus) was examined by starch gel electrophoresis (Goodman and Stuber, 1983; Stuber *et al.*, 1988).

Genetic analyses: Genomic DNA was isolated from seedlings of maize inbred lines according to Dweikat *et al.* (1994) and Gyulai *et al.* (2000).

The final DNA concentration of 25 ng/10 µl was adjusted in a dilution series on 0.8 % agarose gel electrophoresis. PCR reaction was carried out by Weining and Langridge, 1991.

The maize inbred lines were tested with 20 RAPD and 20 SSR primers. The random polymorphisms were analysed using a 10 bp OP/AB set (Operon Sci., USA), while the genetic polymorphisms occurring in the repetitive DNA sequences were analysed using 16–18 bp SSR and ISSR primers ([GACA]₄; AC[GACA]₄; [GACA]₄AC; CA[GACA]₄; [GACA]₄CA; [ACTG]₄

Statistical evaluation of the results: The data were evaluated using the SPSS (Norusis, 1993) cluster analysis programme on the basis of standardised morphological data, and a presence-absence data matrix for the alleles of PCR fragments and enzyme loci exhibiting polymorphism (Gyulai *et al.*, 2000).

The Dice similarity indexes were calculated by comparing pairs of lines according to the formula $GS = 2N_{ij}/(N_i + N_j)$, where GS represents the genetic similarity between inbred lines i and j , N_{ij} is the number of fragments common to i and j , N_i and N_j are the total number of fragments detected in lines i and j , respectively.

The PIC (polymorphism index content) values expressing the degree of polymorphism were calculated using the equation $1 - \sum f_i^2$, where f_i represents the frequency of the i^{th} allele at the given locus (Smith *et al.*, 1997).

Results and discussion

Polymorphism analysis on the basis of morphological descriptions

According to the results of analysis of variance, the environment had no significant effect on certain traits, such as the anthocyanine coloration of the spikelet husk, the shape of the ear, the kernel type, the colour of the kernel crown, the number of kernel rows, or the anthocyanine coloration of the cob. The anthocyanine coloration of the anther, the spikelet ring and the stigma, the degree of fertilisation of the ear and the width of the kernels exhibited only a slight dependence on the year (significant at the 10 % level of probability). The environment had a great influence on the kernel colour, where the data were significantly different at the 1 % level of probability. The remaining traits were significantly affected by the environment ($p = 0.1$ %).

On the basis of the CPVO-TP/2/2 (2001) application form, two lines, *Mo 17 Mv* and its isogenic derivative, *Mo 17 wx*, exhibited such strong similarity that they could not be distinguished on the basis of the fundamental traits listed on the form.

Polymorphism analysis on the basis of isoenzyme patterns

Thirteen of the 15 enzyme loci examined exhibited polymorphism on the basis of isoenzyme patterns. Of the 35 alleles possible at these loci in all, 29 were found in the tested lines, equivalent to a mean value of 2.2 alleles/locus.

When determining polymorphism on the basis of isoenzyme patterns, the 46 lines were found to form 29 different gel electrophoresis groups, indicating that some of the lines could not be distinguished from each other. In the majority of cases, this similarity was due to the relationship between the lines and could be attributed to their having the same genetic background. In a few cases, however, the similarity between the isoenzyme patterns could not be explained on the basis of pedigree.

On the basis of the isoenzyme pattern 18 lines exhibited distinct gel electrophoresis patterns, manifested not so much in the appearance of unique alleles, as in distinct combinations of alleles at the polymorphic loci.

A similar picture was obtained in the case of Dice similarity indexes. The highest value was obtained for lines exhibiting identical enzyme patterns (Dice index = 1), which for most lines could be explained by the genetic background. The lowest value was obtained between *Mv L2* and *F7* (0.38), which corresponds well with the origin of the lines, which are not related to each other at all.

The two Lancaster lines, *Mo 17 Mv* and its isogenic variant, *Mo 17 wx*, could not be distinguished from each other on the basis of isoenzyme pattern either.

The PIC (polymorphic index content) values expressing the degree of polymorphism ranged from 0.04 to 0.55 for the isoenzyme loci, with a mean value of 0.27. Among the polymorphic enzyme loci the lowest values were found for loci *Mdh3*, *Mdh5* and *Pgm1* (PIC = 0.04). The greatest amount of information with respect to the presence of polymorphism was provided by loci *Pgm2* and *Acp1*, but even these only had PIC values of 0.55. The extremely low mean value and the fact that even the highest value was not very high indicate that the efficiency of isoenzyme analysis in demonstrating polymorphism is extremely limited.

Polymorphism analysis using PCR-based markers (RAPD, SSR)

Analysis involving data from RAPD and microsatellites (SSR) was effective in detecting polymorphism among all the maize inbred lines tested.

Of the 20 RAPD primers tested the *OP/AB 2*, *OP/AB 3*, *OP/AB 6*, *OP/AB 9* and the *OP/AB 13* did not give clear results even after several replications.

Three primers (*OP/AB 9*, *OP/AB 15* and *OP/AB 19*) exhibited a 100 % monomorphic pattern. Other RAPD primers (*OP/AB/4*, *OP/AB7*, *OP/AB/10*, *OP/AB/14*, *OP/AB 18*, *OP/AB 20*) exhibited extremely selective polymorphism between the lines. Finally, 12 RAPD primers were chosen, with a total of 93 fragments, of which 54 (averaging 4.5 fragments/primer) gave a reliable polymorphic pattern in all the replications (Fig. 1).

The Dice similarity index was high (0.8–0.9) in several related groups (e.g. ISSS, Lancaster, the early Canadian lines), which can be explained by the origin of the lines. The lowest Dice indexes calculated during the RAPD analysis were obtained between lines which, in the light of their genetic background, were known to be unrelated to each other. The polymorphism between the two lines exhibiting the greatest similarity, *CM 105* and *CM 108* (Dice similarity coefficient: 0.91), was due to differences in two fragments on each of three primers (*OP/AB 4*, *OP/AB 10* and *OP/AB 14*).



Fig. 1. Analysis of polymorphism in maize lines (1-16) using the *OP/AB-20* RAPD primer. Mw: molecular weight marker

In the analysis of SSR markers, primer pairs linked to the genes *wx*, a *wx1*, a *cko1* and *adh2* did not give PCR patterns suitable for evaluation, as the pattern was 100% monomorphic for *zp1* and *p1* and almost 98% monomorphic for *mir2*.

Although two of the SSR loci linked to the genes *ohp2* and *mlg3* exhibited polymorphism, the patterns could not be identified because the differences between the fragments were extremely small, amounting to only a few base pairs. Several SSR loci linked to the *nacl1*, *mgs1*, a *pl1*, a *gln4* and *op2* genes gave extremely selective polymorphism between the lines (Fig 2).



Fig. 2. Analysis of polymorphism in maize lines (1-16) using the *o2* microsatellite primer pair (1: ATGGAGCACGTCTCATCTCAATGG; 2: AGCAGCAGCAACGTCTATGACACT). Mw: molecular weight marker

The data matrix was finally compiled on the basis of 71 polymorphic fragments from 11 primer pairs, representing 6.4 fragments/primer.

The selective polymorphism characteristic of the SSR markers was reflected in the Dice indexes, which had much lower values than in the case of isoenzyme patterns or RAPD analysis. High values were rarely found, and even these were lower than the highest values recorded for the other two methods. This indicates that the SSR markers provide an extremely selective analysis of polymorphism even with a low number of primers. The greatest similarity was found between lines *Mv L8* and *Mv L10* (Dice similarity coefficient: 0.85), the polymorphism of which was caused by two fragments from each of two microsatellite primer pairs linked to the genes *pl1* and *sus1*.

An analysis of genetic markers enabled polymorphism to be detected for all the lines, which is particularly noteworthy in the case of isogenic lines (a line and its *wx* variant, or a line and its fertility restoring, *rf*, variant).

The two Lancaster lines, *Mo 17 Mv* and *Mo 17 wx*, which did not exhibit polymorphism on the basis of either morphological description or isoenzyme pattern, could be distinguished by means of both SSR and RAPD marker analysis.

The PIC (polymorphic index content) values expressing the degree of polymorphism were in line with the above observations.

The PIC values obtained for RAPD and SSR markers were the highest, ranging from 0.20–0.91 for RAPD, with an average of 0.61, and from 0.54–0.90 for SSR, with an average of 0.73. This indicates that the RAPD and SSR markers gave a reliable picture of polymorphism even with a relatively low number of primers.

Analysis of degrees of relationship

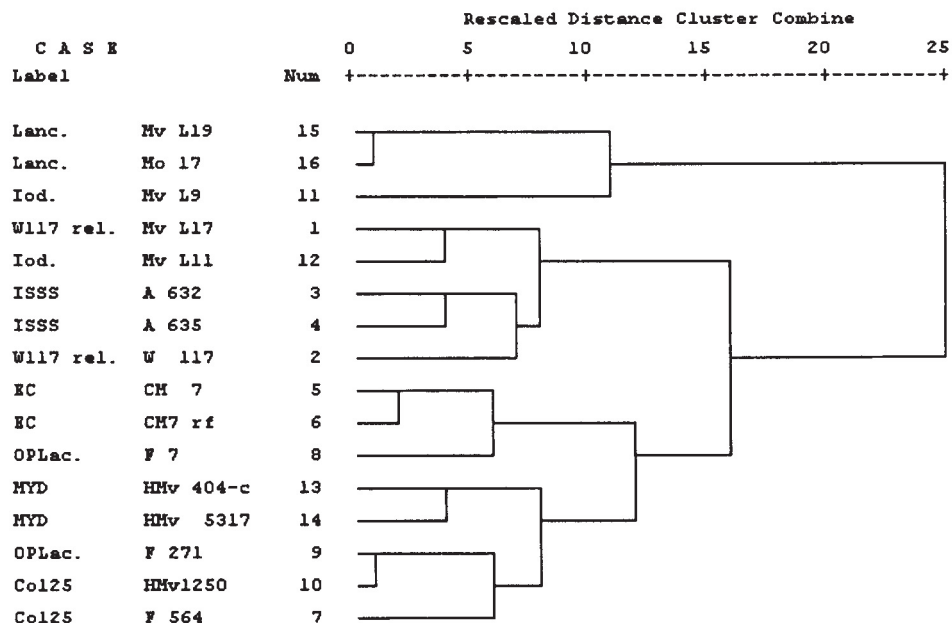


Fig. 3. Related groups on the basis of morphological traits

(ISSS: Iowa Stiff Stalk Synthetic; EC: Early Canadian; W117 rel.: lines related to W 117; Iod.: Iodent; Lanc.: Lancaster; OPLac.: OP Lacaune; MYD: Mindszentpuszta Yellow Dent; Co125: Co 125 derivatives)

To analyse degrees of relationship two lines, derived either from each other or from common parents, were chosen from each related group.

On the basis of morphological descriptions, the groups formed on the dendrogram did not reflect the real relationships for all the lines (Fig. 3).

Lines developed from the Lancaster (*Mo 17 Mv*, *Mv L19*), ISSS (*A 632*, *A 635*), Early Canadian (*CM 7*, *CM 7 rf*) and Mindszentpuszta Yellow Dent (*HMv 404-c*, *HMv5317*) varieties and derivatives of *Co 125* (*F 271*, *HMv 1250*) were found in the same cluster, as expected from the relationship between them, while lines related to Iodent (*Mv L9*, *Mv L11*), OP Lacaune (*F 564*, *F 7*) and *W 117* (*W 117*, *Mv L17*) were associated with members of other groups. A similar picture was obtained when genetic markers were analysed, among which the SSR patterns reflected real relationships to the least extent. Naturally, this was due to the low number of primer pairs, not to the inefficiency of the method. As the SSR markers are to be found in the hypervariable regions of the chromosomes, a far larger number of primer pairs are required than in the analysis of isoenzymes, which act as structural genes. When the genetic markers were subjected to combined analysis, all the lines with the exception of Iodent derivatives (*Mv L9*, *Mv L11*) were classified in the group expected from their origin (Fig. 4).

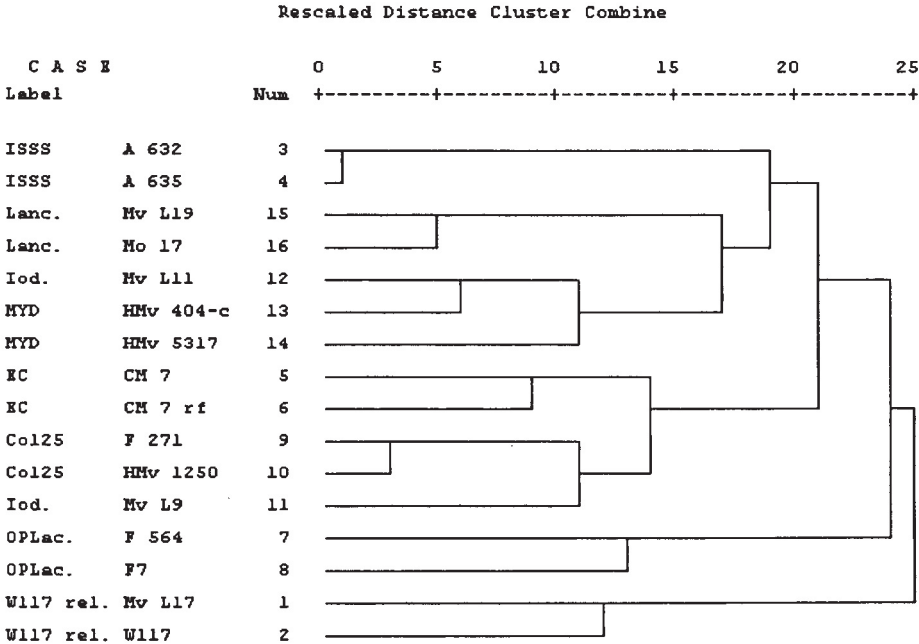


Fig. 4. Related groups on the basis of patterns given by genetic markers (isoenzyme, RAPD, SSR) (ISSS: Iowa Stiff Stalk Synthetic; EC: Early Canadian; W117 rel.: lines related to W 117; Iod.: Iodent; Lanc.: Lancaster; OPLac.: OP Lacaune; MYD: Mindszentpuszta Yellow Dent; Co125: Co 125 derivatives)

This means that when the number of lines exceeds the limits of analysis for the given marker, the best solution is not always to expand the number of primers or enzymes used for marker analysis; the correct combination of various genetic markers may give more exact results, suitable for evaluation. This may lead to a substantial reduction in costs, if for example the more expensive SSR analyses are combined with less costly isoenzyme analyses.

The genetic background was reflected most accurately by the dendrogram produced by the combined analysis of genetic markers, combined with morphological data, classified the related groups according to their genetic background (Fig. 5). The clusters indicating related groups could be clearly distinguished from each other, and only contained line pairs which were related to each other. It can thus be assumed that if the genetic background of genotypes of unknown origin is to be revealed, facilitating the prediction of heterosis, analysis at the DNA level will be required, combining both dominant (RAPD, AFLP, etc.) and codominant (SSR, RFLP, isoenzyme, etc.) markers, applied together with the heterosis test, morphological descriptions and pedigree analysis.

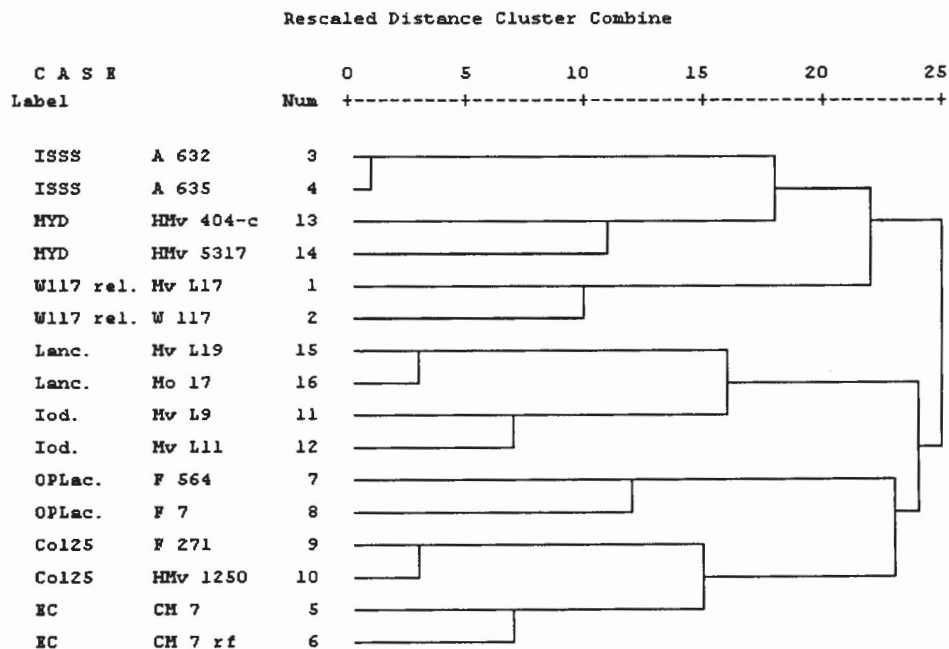


Fig. 5. Related groups on the basis of genetic markers and morphological traits (ISSS: Iowa Stiff Stalk Synthetic; EC: Early Canadian; W117 rel.: lines related to W 117; Iod.: Iodent; Lanc.: Lancaster; OPLac.: OP Lacaune; MYD: Mindszentpuszta Yellow Dent; Co125: Co 125 derivatives)

In summary molecular marker analyses were found to be an important part of maize breeding. They can be applied routinely to test for polymorphism, thus playing an indispensable role in checking the homogeneity of lines and hybrids and the genetic stability and purity of breeding stocks, as well as forming a basis for the protection of intellectual property rights.

The grouping of maize varieties according to their genetic background is of outstanding importance in breeding, since a sufficient genetic distance between the parental lines is a fundamental criterion for heterosis breeding. Maize lines can be put into groups which give an accurate and reliable reflection of their genetic background by combining various marker systems, supplemented with morphological data. In this way genetic markers can be of service to breeders in the planning of crossing programmes.

References

- Botstein D., R.L. Hite, M. Skonic, R.W. David, 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.*, **32**: 314-331.
- Bourgoin-Greneche, M., J. Lallemand, 1993. Electrophoresis and its application to the description of varieties. A presentation of the techniques used by GEVES, Paris. 1-63.
- Dweikat, I., H. Ohm, S. Mackenzie, F. Patterson, S. Cambron, R. Ratcliffe, 1994. Association of a DNA marker with Hessian fly resistance gene H9 in wheat. *Theor. Appl. Genet.*, **89**: 964-968.
- European Union Community Plant Variety Office 2001. Protocol for distinctness, uniformity and stability tests. *Zea mays* L. (Maize)
- Goodman, M. M., C.W. Stuber, 1983. Races of maize. VI. Isozyme variation among races of maize in Bolivia. *Maydica*, **28**: 169-187.
- Gyulai, G., J.A. Gémesné, Z. Sági, G. Venczel, P. Pintér, Z. Kristóf, O. Törjék, L. Heszky, S. Bottka, J. Kiss, L. Zatykó, 2000. Doubled haploid development and PCR-analysis of F₁ hybrid derived DH-R₂ paprika (*Capsicum annuum* L.) lines. *J. Plant Physiol.*, **156**: 168-174.
- Hunter, R. B., L.W. Kannenberg, 1971. Isozyme characterization of corn (*Zea mays* L.) inbreds and its relation to single cross hybrid performance. *Can. J. Cytol.*, **13**: 649-655.
- Lamkey. K.R., A.R. Hallauer, A.L. Kahler, 1987. Allelic differences at enzyme loci and hybrid performance in maize. *J. Herd.*, **78**: 231-234.
- Markert, C. L., F. Moller, 1959. Multimeric forms of enzymes: tissue, ontogenic, and species specific patterns. *Proc. Natl. Acad. Sci. USA*, **45**: 753-763.
- Mullis, K. B., F.A. Faloona, 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol.*, **155**: 335-350.
- Norusis, M. J., 1993. SPSS for Windows Professional Statistics Release 6.0, 385 pp.
- Pejic, I., P. Ajmone-Marsan, M. Morgante, V. Kozumplick, P. Castiglioni, G. Taramino, M. Motto, 1998. Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. *Maize Genet. Coop. Newsletter*, **72**: 18-19.
- Price, S. C., A.L. Kahler, A.R. Hallauer, P. Charmley, D.A. Giegel, 1986. Relationships between performance and multilocus heterozygosity at enzyme loci in single-cross hybrids of maize. *J. Herd.*, **77**: 341-344.
- Smith, J. S. C., E.C.L. Chin, H. Shu, O.S. Smith, S.J. Wall, M.L. Senior, S.E. Mitchell, S. Kresowitch, E.J. Ziegl, 1997. An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.): comparison with data from RFLPs and pedigree. *Theor. Appl. Genet.*, **95**: 163-173.
- Smith, J. S. C., M.L. Senior, 1999. The utility of simple sequence repeat (SSR) data to preferentially identify progeny lines of maize (*Zea mays* L.) that are bred from known inbred parents. *Maydica*, **45**: 205-213.
- Smith, J. S. C., O.S. Smith, 1989a. The description and assessment of distance between inbred lines of maize: I. The use of morphological traits as descriptors. *Maydica*, **34**: 141-150.
- Smith, J. S. C., O.S. Smith, 1989b. The description and assessment of distances between inbred lines of maize: II. The utility of morphological, biochemical, and genetic descriptors and a scheme for testing of distinctiveness between inbred lines. *Maydica*, **34**: 151-161.
- Smith, J. S. C., O.S. Smith, 1992. Measurement of genetic diversity among maize hybrids: A comparison of isozymic, RFLP, pedigree and heterosis data. *Maydica*, **37**: 53-60.
- Stuber, C. W., J.F. Wendel, M.M. Goodman, J.S.C. Smith, 1988. Techniques and scoring procedure for starch gel electrophoresis of enzymes from maize. *Tech. Bulletin*, **286**: 1-87.
- Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, M. Zabeau, 1995. *Nucleic Acids Res.*, **21**: 4407-4414.
- Weining, S., P. Langridge, 1991. Identification and mapping of polymorphisms in cereals based on the polymerase chain reaction. *Theor. Appl. Genet.*, **82**: 209-216.

Received 3 November, 2005, accepted 20 April, 2006