5. Identification of essentially derived varieties (EDVs) derived from biparental crosses of homozygous lines. II. Morphological distances and heterosis in comparison with SSR and AFLP data in maize

M. Heckenberger, M. Bohn, D. Klein, and A. E. Melchinger

Abstract

Morphological traits and heterosis have been proposed apart from genetic distances (GDs) based on molecular markers as possible tools to assess the genetic conformity between putative essentially derived varieties (EDVs) and their initial varieties (IVs). However, for maize and other crops no consensus has been reached regarding methods and thresholds for identification of EDVs, because reliable benchmark data are lacking. The objectives of this study were to (1) evaluate the power of morphological traits and heterosis to discriminate between homozygous progenies derived from F_2, BC_1, and BC_2 populations, (2) compare the findings to published data based on SSRs and AFLPs, and (3) draw conclusions about the usefulness of the various distance measures for identification of EDVs. Morphological distances (MDs) based on 25 traits and mid-parent heterosis for 12 traits were observed for a total of 58 European maize inbred lines comprising 38 triplets. A triplet consisted of one homozygous line derived from a F_2, BC_1 or BC_2 population and both parental inbreds. In addition, all inbreds were genotyped with 100 uniformly distributed SSRs and 20 AFLP primer combinations in companion studies for calculation of genetic distances (GDs). Correlations between the coancestry coefficient, GDs, MDs, and midparent heterosis were significant and high for most traits. However, thresholds for EDVs to discriminate between F_2- and BC_1-derived or BC_1- and BC_2-derived progenies using only morphological distances or heterosis yielded considerably higher values for Type I (α) and Type II ($1-\beta$) errors than observed with GDs based on SSRs and AFLPs. Consequently, we suggest a multi-stage procedure with the initial use of morphological data and a consecutive fingerprinting with molecular markers for identification of EDVs.

Introduction

Plant variety protection (PVP) systems and their laws and regulations should balance commercial interests and warrant sustainable development of new cultivars. On the one hand, registered plant varieties need to be protected against plagiarism and misuse. On the other hand, protected germplasm should be accessible to secure future breeding progress. Therefore, the concept of "breeder's exemption" was introduced into the UPOV convention to solve the obvious conflict between the different stakeholders within the PVP system (UPOV, 1978). Accordingly, plant breeders have access to protected germplasm for the development of new varieties.

New methods such as doubled haploids, marker-assisted backcrossing, and genetic engineering have provided the technical basis to undermine the breeder's exemption in the original sense of the UPOV convention. These tools allow to add a small number of genes to a protected variety and apply for PVP for this "new" variety. In addition, it is possible to select on purpose for lines that are similar to their parents. Therefore, the efforts invested in breeding the original variety can be exploited by the breeder of the plagiarized variety without indemnification. For this reason, the concept of essentially derived varieties (EDVs) was implemented into the revised UPOV convention (UPOV, 1991) and several national plant variety protection acts.

Accordingly, a variety is deemed to be essentially derived from an initial variety (IV), if it is clearly distinguishable but genetically conform to the IV. If the extent of conformity exceeds a certain threshold, the concept of essentially derived varieties (EDVs) indicates that the breeder of the EDV has to arrive at an agreement with the breeder of the IV. However, no consensus has currently been reached on the methods for determining the genetic conformity to distinguish between EDVs and independently derived varieties (IDVs). In addition, accepted or non-accepted breeding procedures have not yet been defined.

Molecular markers, especially simple sequence repeats (SSRs) and amplified fragment length polymorphisms (AFLPs) have been recommended as appropriate tools for determining EDVs in various crops including maize (Dillmann et al., 1997; Bernardo and

Kahler, 2001; Roldan Ruiz et al., 2000a). By contrast, the use of morphological traits or heterosis is still under debate (ASSINSEL, 1999). Hitherto, accurate morphological and agronomic descriptions of cultivars and varieties are the basis of tests for distinctness, uniformity and stability (DUS) within world-wide PVP systems and assure farmers and breeders that they are using clearly identifiable varieties to high standards of purity and quality (Smith and Smith, 1989a). In addition, numerous studies showed significant correlations between midparent heterosis and the coefficient of parentage (*f*) (Melchinger, 1999; Smith et al., 1991). For these reasons, proponents of the use of morphological traits or heterosis for identifications of EDVs argue that phenotypic information provides the basis for PVP and should also be used for identification of EDVs. First studies on the ability of morphological traits to estimate the genetic conformity between related ryegrass (*Lolium perenne* L.) varieties were performed by Gilliland et al. (2000) and Roldan Ruiz et al. (2000b) but revealed only a limited power to distinguish between IDVs and EDVs.

In maize, a triangular instead of a linear relationship was observed between morphological distances and genetic distances or the coancestry coefficient (*f*) (Dillmann and Guérin, 1998). In addition, genetic relationships among maize inbred lines on the basis of morphology were essentially random compared to any relation derived from heterosis or pedigree data (Smith and Smith, 1989b). However, data on the usefulness of heterosis or morphological traits that reflect the degree of relatedness of maize inbred lines in terms of essential derivation is scanty.

The main goal of this study was to investigate the relationship of homozygous progeny lines in maize derived from F_2, BC_1, or BC_2 populations to their parental inbreds based on heterosis and morphological distances (MDs) in comparison with SSR- and AFLP-based genetic distances (GDs). In detail, our objectives were to (1) evaluate the power of heterosis and MDs to discriminate between progenies derived from F_2, BC_1, and BC_2 populations, (2) compare the findings to published data based on SSRs and AFLPs, and (3) draw conclusions about the usefulness of the various distance measures for identification of EDVs.

Materials and Methods

Plant Materials

A total of 58 elite maize inbred lines were analyzed comprising 24 European flint and 34 European dent lines. These lines originated from the maize breeding programs at the University of Hohenheim (Stuttgart, Germany) and two commercial breeding companies in Germany. The 58 lines comprised 38 triplets. A triplet consisted of one progeny line O and both parental lines P1 and P2. The materials consisted of 26 intra-pool triplets of European dent and 12 intra-pool triplets of European flint lines. Progenies were either derived from F_2, BC_1 or BC_2 populations.

For each combination of lines within a triplet (P1xP2, P1xO, and P2xO), seeds from the corresponding F_1 hybrid were generated. In addition, if more than one progeny line (O_1, O_2, …., O_j) was derived from a cross of the same two parental lines, each possible F_1 hybrid (O_1xO_2, …, $O_{j-1}xO_j$) was generated. In total 114 intra-pool F_1 hybrids were tested in this study. Detailed information on all 38 triplets, the 58 maize inbreds and the hybrids included in this study is available as supplemental data in Tables C and D in the appendix of this thesis.

Molecular Analyses

All lines were genotyped with a set of 100 SSR markers uniformly covering the entire maize genome as described in detail by Heckenberger et al. (2002). The 100 SSRs were selected on the basis of reliable single-locus amplification, absence of null alleles, high degree of polymorphism, and high reproducibility of the bands. SSR analyses were performed on a commercial basis by Celera (1756 Picasso Avenue, Davis CA 95616, USA). In addition, all lines were genotyped for AFLPs by Keygene N.V. (P.O. Box 216, 6700 AE Wageningen, The Netherlands). A total of 20 AFLP primer combinations (PCs) was used as described in detail by Heckenberger et al. (2003). AFLP markers were referred to a proprietary integrated map of maize as described by Peleman et al. (2000).

Experimental Design

Field experiments were conducted in 2000 and 2001 at three locations in South Germany with two replications per location. All sites (Bad Krozingen, Eckartsweier, and Scherzheim) are located in the Upper Rhine Valley, a major area of grain-maize production in Germany. All inbred lines and hybrids of a triplet were grown together in one block. Within each triplet block, F_1 hybrids were grown side-by-side with their parental lines to guarantee heterosis estimates with high accuracy. All trials received standard cultural practices of fertilization as well as control of insects and weeds.

The experimental unit was a three-row plot with a row spacing of 0.75 m and a plot length of 4.0 m. Trials were overplanted and later thinned manually to 26 plants per row with a final plant density of 8.7 plants/m^2. Each row was harvested separately. To reduce neighbor effects between adjacent plots with different vigor (inbreds *vs.* hybrids), only data of the middle row of each plot were used for further analyses. The experiment was performed using a randomized block design. Parameter values were observed for 23 morphological traits according to the UPOV guidelines (UPOV, 1978) and 6 additional agronomic traits (Tab. 1) by measuring a minimum of 5 individual plants of a particular plot or by visual observation of the whole plot.

Statistical Analyses

Grain yield for each single row was adjusted to 84.5% dry matter content (DMC). Heterosis was determined as midparent-heterosis $MPH = (F_1 - MP)/MP$, where F_1 is the F_1 hybrid performance and $MP = (P_1 + P_2)/2$ the mid-parent value in which P_1 and P_2 are the performance of the inbred parents, respectively. Analyses of variance (ANOVA) were performed for morphological traits and midparent heterosis using a statistical model considering genotypes as fixed effects and environments as random effects. Heritabilities (h^2) were estimated on an entry-mean basis for all traits. Likewise, heritabilities (h^2_{MPH}) was calculated on a triplet-mean basis for heterotic traits (Hallauer and Miranda, 1981). For calculation of morphological distances (MDs), observations for each trait were standardized by dividing with the phenotypic standard deviation of the particular trait. Euclidean (MD_{EUC}) and Mahalanobis (1936) (MD_{MAH}) distances were calculated based on standardized observations for each pairwise comparison of inbred lines. Malécot's (1948)

coancestry coefficient (*f*) was calculated between all pairwise line combinations. Genetic distances (GDs) between lines based on SSR (GD_{SSR}) or AFLP (GD_{AFLP}) data were estimated using Rogers' distance (1972). The linear relationship between 1-*f*, GDs, MDs, and heterosis estimates was evaluated with a lack-of-fit test (Snedecor and Cochran, 1980). Empirical and approximated frequency distributions of MD values were compared with a Kolmogorov-Smirnov test (Lehmann, 1986) to check for significant deviations. Simple correlations (*r*) were calculated between 1-*f*, GDs, MDs, and heterosis estimates. Homogeneity of variance components of data from flint and dent germplasm was evaluated with Levene's test (1960). Variance components and correlations were not significantly different between flint and dent lines. Consequently, only results from pooled data were reported.

In order to evaluate potential EDV thresholds, the cumulative frequency distributions for genetic distances were approximated with beta distributions (Johnson et al., 1995) as described in detail by Heckenberger et al. (2004a) Frequency distributions for morphological distances and midparent heterosis for F_2-, BC_1-, or BC_2-derived progeny lines were approximated by normal distributions with parameters chosen such that the mean and variance of the original distribution were conserved. Based on these distributions, we calculated Type I (α) and Type II (1-β) errors for various EDV thresholds and various types of populations as suggested in a companion paper for molecular marker data (Heckenberger et al., 2004a). Here, α corresponds to the probability that a true IDV will be wrongly judged as EDV and 1-β corresponds to the probability that a true EDV will not be recognized as such and judged as IDV. We first investigated the situation that an F_2-derived progeny will be considered as IDV, but a BC_1-derived progeny as EDV. Alternatively, we regarded a BC_1-derived progeny as IDV, but a BC_2-derived progeny as EDV.

Statistical analyses of marker data and *f* values were performed as described by Heckenberger et al. (2004a) using the PLABSIM software package (Frisch et al., 2000). ANOVA for field experiments were calculated with the PLABSTAT software (Utz, 2001). All other statistical calculations were carried out with the R software package (Ihaka and Gentleman, 1996).

Results

Morphological Traits and Heterosis Data

Estimates of genotypic variances ($\hat{\sigma}_g^2$) pooled across flint and dent inbred lines were significant (P<0.01) for all traits (Table 1). In addition, significant (P<0.01) genotype x environment interactions were observed for most traits due to cool and wet weather conditions in 2000 and hot and dry weather conditions in 2001. In most cases $\hat{\sigma}_{ge}^2$ was considerably smaller than $\hat{\sigma}_g^2$.

Significant (P<0.01) estimates of σ_g^2 among triplets for MPH were observed for most traits (Table 2). However, considerable differences were found between traits depending on the relative amount of MPH with highest values for grain yield (GYD), grain yield of hand harvested ears (GYE), number of kernels per ear (NKE), and plant length (PLG). Heritabilities for MPH (h^2_{MPH}) of heterotic traits ranged from 0.66 to 0.97 and were slightly smaller than for line *per se* performance (h^2).

SSR and AFLP Marker Data

A total of 580 SSR alleles and 1053 polymorphic AFLP bands was identified for the set of 58 maize lines. The number of alleles per SSR ranged from 3 to 12 with a mean of 5.9. PIC values for SSRs varied between 0.08 and 0.86 and averaged 0.64. The number of polymorphic bands per AFLP primer combination varied from 40 to 70 with an average of 54. PIC values for individual AFLP bands ranged from 0.03 to 0.50 with a mean of 0.33. A detailed description of the genetic diversity revealed by SSRs and AFLPs in this set of lines is given elsewhere.

Relationships Among Distance Measures, Heterosis, and Coancestry

Correlations (r) between 1-f and genetic distances based on SSRs (GD_{SSR}) and AFLPs (GD_{AFLP}) were highly significant (P<0.01) and exceeded 0.85 in both flint and dent lines with a single exception (Table 3). By comparison, r values between GDs and MDs

were medium ($0.40 \leq r \leq 0.68$). Likewise, *r* values between MD_{EUC} and MD_{MAH} were only of moderate size. Correlations for flint lines were consistently higher than for dent lines. Coancestry was moderately correlated with MD_{EUC}, but poorly correlated with MD_{MAH}; both relationships showed a triangular form (Fig. 1).

In contrast, the relationships of GD_{SSR}, GD_{AFLP} and 1-*f* with MPH were linear for most heterotic traits (Fig. 2). Corresponding *r* values were highly significant ($P<0.01$) and moderate to high depending on the trait (Table 2). In general, these correlations were considerably higher than those of MD_{EUC} or MD_{MAH} with GD_{SSR}, GD_{AFLP}, or 1-*f*.

Threshold Scenarios for Identification of EDVs

Observed frequency distributions of MD_{EUC} and MD_{MAH} for F_2-, BC_1-, and BC_2-derived progenies fitted well the approximated normal distributions in the joint analysis of flint and dent lines. Considerable overlaps between the frequency distributions of MD_{EUC} and MD_{MAH} were observed for F_2- *vs.* BC_1- as well as for BC_1- *vs.* BC_2-derived progenies (Fig. 3).

For thresholds based on MDs, the power β to classify a BC_1-derived progeny line as EDV amounted to 18% for MD_{EUC} and 3% for MD_{MAH}, when choosing $\alpha=0.05$ for F_2-derived lines (Table 4). Assuming $\alpha=0.05$ for BC_1-derived lines, corresponding values of β for BC_2-derived lines were considerably higher for MD_{EUC} and MD_{MAH}. The power β for thresholds determined by $\alpha=1-\beta$ to classify BC_1- or BC_2-derived progenies as EDVs increased considerably compared to the values for $\alpha=0.05$. This increase in β was associated with higher values for α. When potential thresholds were based on MPH, the power β to classify a BC_1-derived progeny line as EDV ranged from 2% to 30% assuming $\alpha=0.05$ between F_2-derived lines (Table 4). Choosing $\alpha=0.05$ for BC_1-derived lines, the values of β increased for BC_2-derived lines. For $\alpha=1-\beta$, the power to classify BC_1- or BC_2-derived progenies as EDVs increased substantially, however, this was again associated with higher values for α. In general, values for α and 1-β were of similar magnitude for MDs and MPH.

A detailed description of threshold scenarios based on SSR- and AFLP-based GDs is given in companion papers (Heckenberger et al., 2004a; 2004b).

Discussion

The maize inbred lines examined in our study represent a cross-section of modern elite flint and dent inbred lines from commercial and public maize breeding programs in Germany. Morphological traits were chosen according to the UPOV guidelines of distinctness, uniformity and stability (DUS). In addition, heterosis and morphological traits were determined in extensive field trials over two years and three locations, which exceeds by far the number of environments employed for DUS testing within PVP systems. Furthermore, SSRs were selected as a suitable marker system due to their known map positions and high degree of polymorphism. AFLPs were chosen due to the greater number of markers per assay unit and their high reproducibility (Heckenberger et al., 2003). Thus, the present study is the first larger investigation after a series of pioneering papers based on isozymes and RFLPs (Smith and Smith, 1989a; 1989b; Smith et al., 1991) to provide critical data on the ability of morphological distances and heterosis for identification of EDVs in maize in direct comparison with SSR and AFLP data. For this reason, our results provide a well-founded comparison of different distance measures for identification of EDVs in maize and may serve as an example for other crops.

Data Quality and Relatedness of Different Measures for Genetic Conformity

Despite the contrasting climatic conditions during the vegetation seasons in 2000 and 2001, high heritabilities were observed for morphological traits and midparent heterosis, the former being considerably higher than those reported by Rebourg et al. (2001). In addition, UPGMA cluster analysis based on MD_{EUC} showed a clear grouping of flint and dent lines, which further corroborates the high quality of morphological data (available as supplemental data in Figure A in the appendix of this thesis). The dendrogram based on MD_{MAH} showed as well a grouping of flint and dent lines, but contained several inbreds that were clustered together with lines of the opposite pool. In addition, only moderate correlations between MC_{EUC} and MD_{MAH} were observed. This can be explained by the different statistical properties of MC_{EUC} and MD_{MAH}, because MD_{MAH} adjusts for the correlations of traits.

The graphs between MDs and GDs (Fig. 1) confirmed the triangular relationship of morphological and genetic distances reported in previous studies (Dillmann et al., 1997; Rebourg et al., 2001). This indicates that low GDs correspond necessarily with low MDs, whereas the reverse does not necessarily hold true because high GDs can correspond with both high and low MDs (Van Eeuwijk and Baril, 2001). In addition, the triangular shape has several biological explanations (Nuel et al., 2001) and is also expected, if only molecular markers tightly linked with the genes controlling the phenotypic trait(s) were used (Burstin and Charcosset, 1997).

In the present investigation, flint and dent lines and triplets showed similar estimates of $\hat{\sigma}_g^2$ and correlations among the various criteria. This is in harmony with a previous study of Bar-Hen et al. (1995), who examined 974 maize inbred lines with morphological traits and RFLPs. Correlations of 1-*f*, GD_{SSR}, or GD_{AFLP}, with MPH were higher than reported by Ajmone Marsan et al. (1998) for AFLPs but similar to correlations of MPH with 1-*f* and GDs based on RFLPs in intra-pool crosses (Boppenmaier et al., 1993; Smith et al., 1990). In addition, our study confirms the findings of Smith and Smith (1989b) that correlations of molecular markers or 1-*f* were considerably higher with MPH than with MD_{EUC} or MD_{MAH}.

Distinctness *vs.* Conformity

To confirm an essential derivation in the intention of the UPOV convention (UPOV, 1991), three separate criteria must be fulfilled. An EDV must be (i) distinct from the IV, (ii) 'predominantly derived' from the IV, and (iii) conform to it in the expression of it's 'essential characteristics'. Distinctness can be determined based on morphological traits by established procedures for DUS testing. Establishing a predominant derivation will either require a directly documented evidence, *e.g.*, by breeding books, or could be determined with molecular evidence similar to forensic approaches in the human sector (Gill et al., 1995). However, the question of whether conformity in the expression of essential characteristics should be assessed by phenotypic rather than molecular data is still unsolved. While differences in the expression of one single trait are sufficient to prove distinctness between two varieties, assessment of conformity should be based on a large number of

morphological traits and still could not give a definite answer due to the triangular relationship mentioned above.

Proponents of phenotypic data argue that the term 'conform in the expression of it's essential characteristics that result from the genotype' (UPOV, 1991) implies the use of phenotypic data rather than molecular data. In contrast, opponents state that even highly heritable phenotypic traits can only offer an indirect measure of the relatedness of two cultivars. In contrast, molecular data provide a direct estimate of the true relatedness of two genotypes because it is unbiased from environmental effects and reflects the percentage of the genome in common between the IV and a putative EDV. Based on our results, genetic distances based on molecular markers have clear advantages for identification of EDVs.

Power of Morphological Distances and Heterosis for Identification of EDVs

For MD_{EUC} as well as for MD_{MAH}, extensive overlaps of the frequency distributions of F_2-, BC_1-, and BC_2-derived progenies were found in spite of the significant correlations with 1-*f*. Thus, Type I (α) and Type II (1-β) errors observed for MDs were considerably higher than observed for GDs based on molecular markers (Table 3). Consequently, MDs provide only a rough estimate of the true relatedness of two lines and can only poorly discriminate F_2-, BC_1-, and BC_2-derived progenies. These results confirm data from ryegrass (Gilliland et al., 2000) and maize (Smith and Smith, 1989b) showing that morphological conformity could give an initial indication of the relatedness of two cultivars, particularly for highly conforming pairs of inbreds. However, a small MD between two varieties cannot be taken as a definitive proof that they are in fact closely related because of the triangular relationship between 1-*f* and MDs.

In contrast to the triangular relationship between GDs and MDs, a linear relationship of MPH with GDs or 1-*f* was observed as expected by quantitative genetic theory (Melchinger, 1999). However, in spite of the higher correlation of MPH with 1-*f* or GDs, MPH was not markedly superior to MDs regarding the power to discriminate between F_2-, BC_1- or BC_2-derived progeny. This is attributable to the larger experimental error and GxE interactions of MPH in comparison with line *per se* performance (data not shown) as reflected by the comparison of h^2_{MPH} vs. h^2.

For nearly all scenarios examined, GDs based on SSRs or AFLPs were superior over MDs or MPH for any trait or combination of traits in their power β to discriminate among F_2-, BC_1, and BC_2-derived progenies for given values of α. However, different from MDs and MPH, the use of SSR or AFLP markers would require thresholds specific for a given germplasm pool. This is necessary because flint and dent lines differed significantly in their mean GD between unrelated lines due to the different levels of polymorphism within each germplasm pool (Heckenberger et al., 2004a)

Conclusions

Based on our results, morphological distances and midparent heterosis can provide only an initial indication for putative EDVs. However, a reliable identification of EDVs by MPH or MDs alone is not possible due to the large overlaps in the frequency distributions of MDs and MPH of F_2-, BC_1-, and BC_2-derived progenies. In addition, MDs and MPH have several disadvantages compared to molecular markers. First, assessment of morphological traits and MPH requires extensive field trials over several years and locations to minimize environmental effects. Therefore, these measurements are more expensive and time consuming than molecular marker analyses. Second, heterosis estimates requires production and testing of hybrids. In addition, reciprocal crosses should be evaluated to minimize the risk of maternal effects (Melchinger et al., 1986). Third, the scoring of morphological traits is to some extent subjective. Therefore, a number of check inbreds must be included in the study to warrant a high quality of morphological traits across different years and scoring persons.

In conclusion, we recommend a multi-stage procedure for identification of EDVs with the initial use of morphological data from DUS testing and a consecutive fingerprinting with a first set of at least 100 SSR markers or 20 AFLP PCs for putative EDVs. If doubts still prevail, whether a variety has been derived independently from another variety or not, the corresponding genotypes should be fingerprinted with a second set of SSRs or AFLPs. Use of MPH for identification of EDVs is problematic, because the rationale for using MPH is merely its linear relationship with 1-*f* and the biological mechanisms underlying heterosis are not fully understood.

Acknowledgements

We are indebted to the 'Gesellschaft zur Förderung der privaten Pflanzenzüchtung' (GFP), Germany, for a grant to support Dr. Bohn and the molecular analyses of the present study. Financial support for M. Heckenberger and the field trials was provided by a grant from the European Union, Grant No. QLK-CT-1999-01499 (MMEDV).

References

Ajmone Marsan, P., P. Castiglioni, F. Fusari, M. Kuiper, and M. Motto. 1998. Genetic diversity and its relationship to hybrid performance in maize as revealed by RFLP and AFLP markers. Theor. Appl. Genet. 96:219-227.

ASSINSEL. 1999. Essential Derivation and Dependence. Practical Information.

Bar-Hen, A., A. Charcosset, M. Bougoin, and Guiard J. 1995. Relationship between genetic markers and morphological traits in a maize inbred lines collection. Euphytica. 84:145-154.

Bernardo, R. and A. L. Kahler. 2001. North American study on essential derivation in maize: inbreds developed without and with selection from F_2 populations. Theor. Appl. Genet. 102:986-992.

Boppenmaier, J., A. E. Melchinger, G. Seitz, H. H. Geiger, and R. G. Herrmann. 1993. Genetic diversity for RFLPs in European maize inbreds. III. Performance of crosses within versus between heterotic groups for grain traits. Plant Breed. 111:217-226.

Burstin, J. and A. Charcosset. 1997. Relationship between phenotypic and marker distances: theoretical and experimental investigations. Heredity. 79:477-483.

Dillmann, C., A. Bar-Hen, D. Guerin, A. Charcosset, and A. Murigneux. 1997. Comparison of RFLP and morphological distances between maize (*Zea mays* L.) inbred lines - consequences for germplasm protection purposes. Theor. Appl. Genet. 95:92-102.

Dillmann, D. and D. Guérin. 1998. Comparison between maize inbred lines: genetic distances in the expert's eye. Agronomie. 18:659-667.

Frisch, M., M. Bohn, and A. E. Melchinger. 2000. Plabsim: Software for simulation of marker-assisted backcrossing. J. Hered. 91:86-87.

Gill, P., C. P. Kimpton, A. Urquhart, N. Oldroyd, E. S. Millican, S. K. Watson, and T. J. Downes. 1995. Automated short tandem repeat (STR) analysis in forensic casework: a strategy for the future. Electrophoresis 16:1543-1552.

Gilliland, T. J., R. Coll, E. Calsyn, M. de. Loose, M. J. T. v. Eijk, and I. Roldan Ruiz. 2000. Estimating genetic conformity between related ryegrass (*Lolium*) varieties. 1. Morphology and biochemical characterisation. Mol. Breed. 6:569-580.

Hallauer, A. R. and J.B. Miranda. 1981. Quantitative Genetics in Maize Breeding. The Iowa State University Press, Ames, Iowa.

Heckenberger, M., M. Bohn, and A. E. Melchinger. 2004a. A statistical framework for identification of essentially derived varieties derived from biparental crosses of homozygous lines. II. SSR data from maize. Crop Sci. submitted.

Heckenberger, M., M. Bohn, J. Rouppe van der Voort, and A. E. Melchinger. 2004b. A statistical framework for identification of essentially derived varieties derived from biparental crosses of homozygous lines. IV. AFLP data from maize and comparison with SSR data. In preparation.

Heckenberger, M., M. Bohn, J. S. Ziegle, L. K. Joe, J. D. Hauser, M. Hutton, and A. E. Melchinger. 2002. Variation of DNA fingerprints among accessions within maize inbred lines and implications for identification of essentially derived varieties. I. Genetic and technical sources of variation in SSR data. Mol. Breed. 10:181-191.

Heckenberger, M., J. Rouppe van der Voort, A. E. Melchinger, J. Peleman, and M. Bohn. 2003. Variation of DNA fingerprints among accessions within maize inbred lines and implications for identification of essentially derived varieties. II. Genetic and technical sources of variation in AFLP data and comparison to SSR data. Mol. Breed. 12:97-106.

Ihaka R. and R. Gentleman. 1996. R: A Language for data analysis and graphics. J. Comp. Graph. Stat. 5:299-314.

Johnson, N. L., S. Kotz, and N. Balakrishnan. 1995. Continuous Univariate Distributions Vol.2. p. 210-275. J. Wiley and Sons, Inc., New York.

Lehmann, E. L. 1986. Testing Statistical Hypotheses. John Wiley and Sons, New York.

Levene, H. 1960. Robust tests for equality of variances. p. 278-292. *In* Contributions to probability and statistics: Essays in honour of Harold Hotelling. Stanford University Press, Stanford, CA.

Mahalanobis P.C. 1936. On the generalized distance in statistics. Proc. Natl. Acad. Sci. India. 2:49-45.

Malécot, G. 1948. Les mathematiques de l'heredité. p. Masson & Cies, Paris.

Melchinger, A.E. 1999. Genetic diversity and heterosis. Int. Symp. on Genetics and Exploitation of Heterosis in Crop Plants. p. 99-118. *In* J.G. Coors and S. Pandey (ed.) The genetics and exploitation of heterosis in crops. ASA, CSSA, and SSSA, Madison, WI.

Melchinger, A. E., H. H. Geiger, and F. W. Schnell. 1985. Reciprocal differences in single-cross hybrids and their F_2 and backcross progenies in maize. Maydica. 30:395-405.

Nuel, G., S. Robin, and C.P. Baril. 2001. Predicting distanced using a linear model: the case of varietal distinctness. J. Appl. Stat. 28:607-621.

Peleman, J., R. van Wijk, J. van Oeveren and R. van Schaik. 2000. Linkage map integration:an integrated genetic map of *Zea mays* L. Poster P472. Plant & Animal Genome Conference VIII. San Diego.

Rebourg, C., B. Gouesnard, and A. Charcosset. 2001. Large scale molecular analysis of traditional European maize populations. Relationships with morphological variation. Heredity. 86:574-587.

Rogers, J. S. 1972. Measures of genetic similarity and genetic distance. Studies in Genetics VII. Univ. Texas Publ. 7213:145-153.

Roldan Ruiz, I., E. Calsyn, T. J. Gilliland, R. Coll, M. J. T. van Eijk, and De LooseM. 2000a. Estimating genetic conformity between related ryegrass (*Lolium*) varieties. 2. AFLP characterization. Mol. Breed. 6:593-602.

Roldan Ruiz, I., F. A. v. Eeuwijk, T. J. Gilliland, P. Dubreuil, C. Dillmann, J. Lallemand, M. de. Loose, and C. P. Baril. 2000b. A comparative study of molecular and morphological methods of describing relationships between perennial ryegrass (*Lolium perenne* L.) varieties. Theor. Appl. Genet. 103:1138-1150.

Smith J.S.C. and Smith O.S. 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. and Smith O.S. 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, S. L. Bowen, R. A. Tenborg, and S. J. Wall. 1991. The description and assessment of distances between inbred lines of maize. III. A revised scheme for the testing of distinctiveness between inbred lines utilizing DNA RFLPs. Maydica. 36:213-226.

Smith, O. S., J. S. C. Smith, S. L. Bowen, R. A. Tenborg, and S. J. Wall. 1990. Similarities among a group of elite maize inbreds as measured by pedigree, F1 grain yield, grain yield, heterosis, and RFLPs. Theor. Appl. Genet. 80:833-840.

Snedecor, G. and W. Cochran. 1980. Statistical methods. Iowa State University Press, Ames, IA.

UPOV. 1978. International convention for the protection of new varieties of plants.

UPOV. 1991. International convention for the protection of new varieties of plants.

Utz, H. F. 2001. Plabstat, ein Computerprogramm zur statistischen Analyse von pflanzenzüchterischen Experimenten, Version 2P vom 14. Juli 2001 (in German). University of Hohenheim, Germany.

Van Eeuwijk, F. A. and C. P. Baril. 2001. Conceptual and statistical issues related to the use of molecular markers for distinctness and essential derivation. Acta Hort. 546:35-53.