Molecular Breeding **12:** 97–106, 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 with SSR data

Martin Heckenberger[1], Jeroen Rouppe van der Voort[2,4], Albrecht E. Melchinger[1,4], Johan Peleman[2] and Martin Bohn[1,3,*]

*[1]Institute of Plant Breeding, Seed Science, and Population Genetics, University of Hohenheim, 70593 Stuttgart, Germany; [2]Keygene, P.O. Box 216, Wageningen, 6700, AE, The Netherlands; [3]Current address: Crop Science Department, University of Illinois, S-110 Turner Hall, 1102 South Goodwin Avenue, Urbana, IL 61801, USA; [4]These authors contributed equally to this work; *Author for correspondence (e-mail: mbohn@uiuc.edu; fax: ++217 333 9817)*

Received 14 May 2002; accepted in revised form 12 February 2003

Key words: AFLPs, Essentially derived varieties, Genetic distances, Intra-varietal variation, Reproducibility, SSRs

Abstract

Accuracy and reproducibility of genetic distances (GDs) based on molecular markers are crucial issues for identification of essentially derived varieties (EDVs). Our objectives were to investigate (1) the amount of variation for amplified fragment length polymorphism (AFLP) markers found among different accessions within maize inbreds and doubled haploid (DH) lines, (2) the proportion attributable to genetic and technical components and marker system specific sources, (3) its effect on GDs between maize lines and implications for identification of EDVs, and (4) the comparison to published SSR data from the same plant materials. Two to five accessions from nine inbred lines and five DH lines were taken from different sources of maintenance breeding or drawn as independent samples from the same seed lot. Each of the 41 accessions was genotyped with 20 AFLP primer combinations revealing 988 AFLP markers. Map positions were available for 605 AFLPs covering all maize chromosomes. On average, six (0.6%) AFLP bands were polymorphic between different accessions of the same line. GDs between two accessions of the same line averaged 0.013 for inbreds and 0.006 for DH lines. The correlation of GDs based on AFLPs and SSRs was tight (r = 0.97**) across all 946 pairs of accessions but decreased (r = 0.55**) for 43 pairs of accessions originating from the same line. On the basis of our results, we recommend specific EDV thresholds for marker systems with different degree of polymorphism. In addition, precautions should be taken to warrant a high level of homogeneity for DNA markers within maize lines before applying for plant variety protection.

Introduction

In all major crops, genetic distances (GDs) based on reliable molecular marker data have been found to reflect accurately the degree of pedigree relationships between genotypes (Melchinger 1999). In maize, several studies reported highly significant correlations between GDs based on molecular markers and the coefficient of coancestry (for review, see Lübberstedt et al. 2000). Consequently, molecular markers, especially amplified fragment length polymorphisms (AFLPs) and simple sequence repeats (SSRs), were recommended as an appropriate tool to test for essential derivation in plant varieties (Smith et al. 1991; Knaak et al. 1996; ASSINSEL 2000).

Scientifically reliable criteria must be developed to differentiate between EDVs and independent varieties because of the severe legal consequences for the

breeders. Indispensable prerequisites are accuracy and reproducibility of GD estimates. Reproducibility problems were investigated by Jones et al. (1997), who reported scoring differences of up to 2 base pairs (bp) among the same SSR fragments detected by different labs. In sugar beet (*Beta vulgaris* L.) and wild *Beta* species, a reproducibility of AFLP bands of 97.6% was determined by performing all necessary analytical steps twice (Hansen et al. 1999). In contrast to these results, Jones et al. (1997) and Bagley et al. (2001) reported an extremely high reproducibility of AFLP bands close to 100%. In addition, Heckenberger et al. (2002) revealed variation in GD estimates based on SSRs of up to 0.12 on a 0 to 1 scale between different generations of maintenance breeding of the same inbred line or the same inbred line maintained by different breeders. However, critical information on the reproducibility of AFLP bands and their stability during maintenance is still lacking.

The overall goal of our study was to determine the variation of AFLP markers among different accessions of maize inbreds and doubled haploid (DH) lines. In detail, our objectives were to investigate (1) the amount of variation for amplified fragment length polymorphism (AFLP) markers found among different accessions within maize inbreds and doubled haploid (DH) lines, (2) the proportion attributable to genetic and technical components and marker system specific sources, (3) its effect on the GD between maize lines and implications for identification of EDVs and (4) the comparison to published SSR data from the same plant materials (Heckenberger et al. 2002).

Materials and methods

Plant materials

For nine maize inbred lines, six from the flint and three from the dent pool, and five DH lines from the dent pool, two to five accessions per line were fingerprinted. Accessions were obtained from different generations of maintenance breeding conducted by the University of Hohenheim (UHOH) and three commercial breeding companies (B1–B3). Two accessions per DH line were obtained by drawing two independent samples of 20 kernels out of the same seed lot. All DH lines were derived from the cross 941118 of inbred lines RG2302 and 69117. A detailed description of the plant materials analyzed and the applied method of maintenance breeding is given in our companion paper (Heckenberger et al. 2002).

AFLP and SSR analyses

AFLP fingerprints were generated by Keygene N.V. from leaf punches from a bulk of 10 individual plants per accession as described by Vos et al. (1995), visualized by use of a Fuji BAS/2000 phosphorimager, and scored dominantly on the set of maize lines with proprietary software developed by Keygene N.V. AFLP markers were referred to a proprietary integrated map of maize. This integrated map combines linkage information of 5650 molecular markers and is based on 23 separate mapping populations (Peleman et al. 2000). Comparison of the 988 AFLP markers scored in this project with the integrated maize map revealed that for 605 of these AFLP markers map information was available.

SSR analyses were carried out by Celera using 100 publicly available SSR primer pairs, equally distributed across the maize genome using an ABI Prism™ 377 DNA Sequencer with 5% polyacrylamide gels. Internal fragment size standards were used in each lane to increase accuracy of DNA fragment size determination. Fragment sizes were determined automatically by using the GeneScan® and Genotyper® software packages. For a detailed description of the procedures, see our companion paper (Heckenberger et al. 2002). AFLP and SSR analyses of each accession were carried out using seeds from the same seed lot.

Statistical analyses

The polymorphic information content (PIC) was calculated for each primer combination using the formula $PIC = 2p_i\,(1 - p_i)$, where p_i is the frequency of the i^{th} AFLP band (Roldan-Ruiz et al. 2000). The marker index was calculated for each AFLP primer combination as $MI = \overline{PIC} \times n\beta$, where $\overline{PIC}$ is the mean PIC value, n is the number of bands, and β is the proportion of polymorphic bands (Powell et al. 1996).

Genetic distances (GDs) were calculated as 1 – genetic similarity (GS). For AFLPs, GS was calculated using the GS coefficient of Jaccard (1908) and for SSRs the GS coefficient of Dice (1945) was employed. In the case of missing values, *i.e.*, if one or several primer combinations did not yield amplification products in one of the two accessions compared,

the corresponding alleles of the other accession were not used for GD calculation. GD estimates were based on the whole AFLP data set or on AFLP data for single chromosomes. The cluster analysis was performed with GD estimates using the 'unweighted pair group method using arithmetic averages' (UPGMA) (Nei et al. 1983). The reliability of the cluster was assessed by applying a bootstrap procedure (Efron 1979). The cophenetic correlation (r_{cpe}) was calculated to test for the goodness-of-fit between GD values obtained from the cluster and the original GD estimates. The significance of r_{cpe} was determined by the Mantel test (Mantel 1967) based on 10000 permutations. Standard deviation (SD) of GD estimates was calculated using the formula of Bar-Hen and Charcosset (1994),

$$SD = \sqrt{\frac{1}{N} GD \times (1 - GD)},$$

where N is the number of polymorphic bands and GD is the genetic distance between two genotypes or the mean GD between two groups of genotypes.

The null hypothesis (H_0) that the markers were randomly distributed across the genome was tested against the alternative hypothesis (H_1) that the markers were not randomly distributed across the genome by a dispersion analysis (Johnson et al. 1992). The test statistic was

$$X^2 = \frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{\bar{y}}.$$

Here, n denotes the number of chromosomal intervals between two fixed marker loci (BINs), y_i the number of markers located in a particular BIN, and $\bar{y}$ the mean number of markers per BIN. Under H_0, X^2 follows a χ^2-distribution with n-1 degrees of freedom. H_0 was rejected, if X^2 was smaller than $\chi^2_{1-\frac{\alpha}{2};n-1}$ or larger than $\chi^2_{\frac{\alpha}{2};n-1}$.

GDs between accessions of the same line were defined as ″GD within″ (GD_w). Lines were screened for outliers based on their mean GD_w values, by applying the test of Anscombe and Tukey (1963). For all analyses where information on map positions was necessary, only the mapped AFLP markers were used. The genomic locations of variation within maize lines detected with AFLPs and SSRs were compared based on their BIN positions.

Calculation of GDs, cluster analysis, cophenetic correlation and Mantel test were performed with software NTSYS-PC (Rohlf 1989). The bootstrap procedure was carried out with the Winboot computer program (Yap et al. 1996). All other statistical calculations were performed with SAS (SAS Institute 1988).

Results

Description of markers

A total of 988 AFLP bands was identified. The number of polymorphic bands per PC varied from 40 to 83 with an average of 49. No significant differences between *Pst*I/*Mse*I and *Eco*RI/*Mse*I primers were found for the number of polymorphic bands, mean *PIC*, *MI*, and mean GD_w values (Table 1). In addition, the proportion of missing values had no influence on these quality parameters. Comparison of the 988 AFLP markers scored in this project with the Keygene integrated maize map (Peleman et al. 2000) revealed that for 605 of these AFLP markers map information was available. The average marker interval of the mapped markers was 2.5 cM, the total map length amounted to 1512 cM.

By comparison, the average marker interval for SSRs employed in the companion study was 12.4 cM and the total map length was 1210 cM (Heckenberger et al. 2002). Evaluation of the number of markers per BIN revealed a higher standard deviation for AFLPs than for SSRs. The null hypothesis of a random marker distribution across the genome was rejected for both AFLPs and SSRs. For AFLPs, X^2 was significantly ($P < 0.05$) larger than $\chi^2_{\frac{\alpha}{2};n-1}$, indicating that AFLPs were heterogeneously distributed over the genome, whereas X^2 was smaller than $\chi^2_{1-\frac{\alpha}{2};n-1}$, indicating that SSRs were uniformly distributed across the genome.

Genetic distances of accessions within maize lines

Between different accessions of the same inbred or DH line, an average of six AFLP fragments (0.6%) was polymorphic. The mean GD_w for inbred lines calculated separately for each line varied from 0.000 to 0.022 (Table 2).

Table 1. Statistics characterizing the degree of polymorphism and quality of AFLP data generated with 20 primer combinations

	Enzymes		Polymorphic bands					
PC[†]	*Eco*RI/*Mse*I	*Pst*I/*Mse*I	No.	Proportion (%)	$\overline{PIC}$ [‡]	MI[§]	$\overline{GD_w}$ [#]	Prop. Of missing values (%)[¶]
A	x		45	66	0.36	16.5	0.010	2
B		x	46	73	0.36	16.6	0.010	6
C	x		40	47	0.38	15.2	0.005	2
D	x		58	69	0.34	19.8	0.023	6
E		x	40	58	0.35	13.9	0.012	6
F		x	56	68	0.26	14.7	0.024	28
G		x	47	62	0.30	14.0	0.019	9
H		x	41	59	0.37	15.2	0.019	5
I	x		59	68	0.36	21.2	0.009	8
J		x	46	69	0.35	16.1	0.007	3
K	x		43	58	0.35	15.3	0.030	6
L	x		83	71	0.34	27.8	0.006	4
M		x	45	62	0.38	17.0	0.012	4
N	x		51	68	0.36	18.4	0.012	3
O	x		48	67	0.34	16.3	0.016	9
P		x	46	70	0.34	15.5	0.023	5
Q	x		48	62	0.36	17.4	0.016	2
R		x	50	68	0.35	17.6	0.018	6
S	x		47	64	0.27	12.5	0.006	38
T		x	49	68	0.35	17.2	0.013	7
Mean			**49**	**65**	**0.34**	**16.5**	**0.015**	**8**

[†] PC = Primer combination.
[‡] $\overline{PIC}$ = Mean PIC value observed for AFLPs of the particular PC.
[§] MI = Marker index.
[#] $\overline{GD_w}$ = Mean GD value between accessions of the same inbred or DH line (GD_w), obtained only with markers from the particular PC.
[¶] Proportion of missing values based on all datapoints.

The dendrogram obtained from UPGMA cluster analysis resulted in a clear separation of flint and dent inbred lines (Figure 1). Using genotypic data from all chromosomes, accessions derived from the same line were always clustered together. However, this pattern was not consistently obtained comparing the 10 dendrograms based on marker data of individual chromosomes (data not shown).

GD estimates based on AFLP and SSR data were significantly ($P < 0.01$) correlated with each other (r = 0.97**) for all pairwise 946 combinations of all 41 entries (Table 3). The correlation dropped to r = 0.07 when considering only the 46 GD_w values between accessions of the same line and increased to r = 0.55** after omitting the GD_w values of outliers ZS 467 (AFLP) and D146 (SSR) (Figure 2).

Locations of variation within maize lines

A total of 58 mapped AFLP markers, localized within 31 BINs, showed variation between accessions of the same line. Only six of these BIN locations displayed also variation with SSRs for the same accessions. These were located on chromosomes 1 (DK105), 2 (D149, D171, UH200), 7 (D149), and 8 (UH002). For all other BINs with varying AFLP markers adjacent SSR markers were not polymorphic within lines.

Genetic distance between lines

The effect of the variation of GD_w on the GD between lines was investigated by calculating the span for GD values for each pairwise combination of accessions between the same two lines. These span values varied from 0.001 to 0.114 with an average of 0.017. If outliers D146 and ZS467 were omitted from the anal-

Table 2. Mean, maximum, and minimum of genetic distances (GD_w) between accessions of the same maize inbred or DH line

		AFLP			SSR		
Line	No. of accessions	Mean	Min.	Max.	Mean	Min.	Max.
Inbred lines							
D146	3	0.004	0.000	0.006	***0.090*** †	0.061	0.116
D149	4	0.020	0.008	0.033	0.031	0.020	0.042
D171	4	0.004	0.000	0.008	0.016	0.005	0.028
D503	2	0.000	–‡	–‡	0.005	–‡	–‡
DK105	4	0.022	0.008	0.032	0.036	0.017	0.056
UH002	3	0.009	0.005	0.016	0.026	0.021	0.031
D06	5	0.011	0.003	0.020	0.011	0.000	0.021
UH200	3	0.013	0.010	0.015	0.034	0.025	0.047
UH300	3	0.020	0.019	0.022	0.033	0.016	0.044
Mean §		**0.011**			**0.024**		
DH lines							
ZS264	2	0.008	–‡	–‡	0.005	–‡	–‡
ZS265	2	0.008	–	–	0.000	–	–
ZS337	2	0.003	–	–	0.000	–	–
ZS467	2	***0.124*** †	–	–	0.000	–	–
ZS595	2	0.008	–	–	0.000	–	–
Mean §		**0.007**			**0.001**		
Grand mean §		**0.010**			**0.015**		

† Outliers based on the test of Anscombe and Tukey (1963).
‡ For lines with only two accessions, only a single GD_w value was available, therefore, no maximum and minimum was calculated.
§ Means were calculated leaving out outliers.

ysis, span-values for AFLPs and SSRs were of similar magnitude.

Discussion

AFLPs proved to be an appropriate tool to distinguish between flint and dent lines. This is in agreement with results published by Lübberstedt et al. (1999), who genotyped a set of 51 European flint and dent lines with eight AFLP primer combinations and found a clear separation of flint and dent germplasm. *PIC* values and marker indices were almost identical to those published by Lübberstedt et al. (2000). In addition, DH lines and their parental lines could be clearly separated from other dent germplasm developed by the University of Hohenheim. However, using only data from individual chromosomes, accessions of the same line did not cluster together in some cases. This lack of association can be explained by sampling effects due to the smaller number of markers for individual chromosomes. In addition, we used the Jaccard coefficient for AFLPs and the Dice coefficient for SSRs following the proposal by Link et al. (1995) for dominant and codominant marker systems, respectively. This may also have slightly decreased the correlation between AFLP- and SSR-derived GD values.

GD within inbred and DH lines

Inbred lines

In general, the dominant AFLP markers yielded lower GD_w values than the codominant SSR markers. As heterogeneity due to residual heterozygosity was the major cause of SSR variation within inbred lines, the lower GD_w values of AFLPs can be explained by the fact that heterogeneity cannot be detected using a dominant marker system like AFLPs. In addition, the

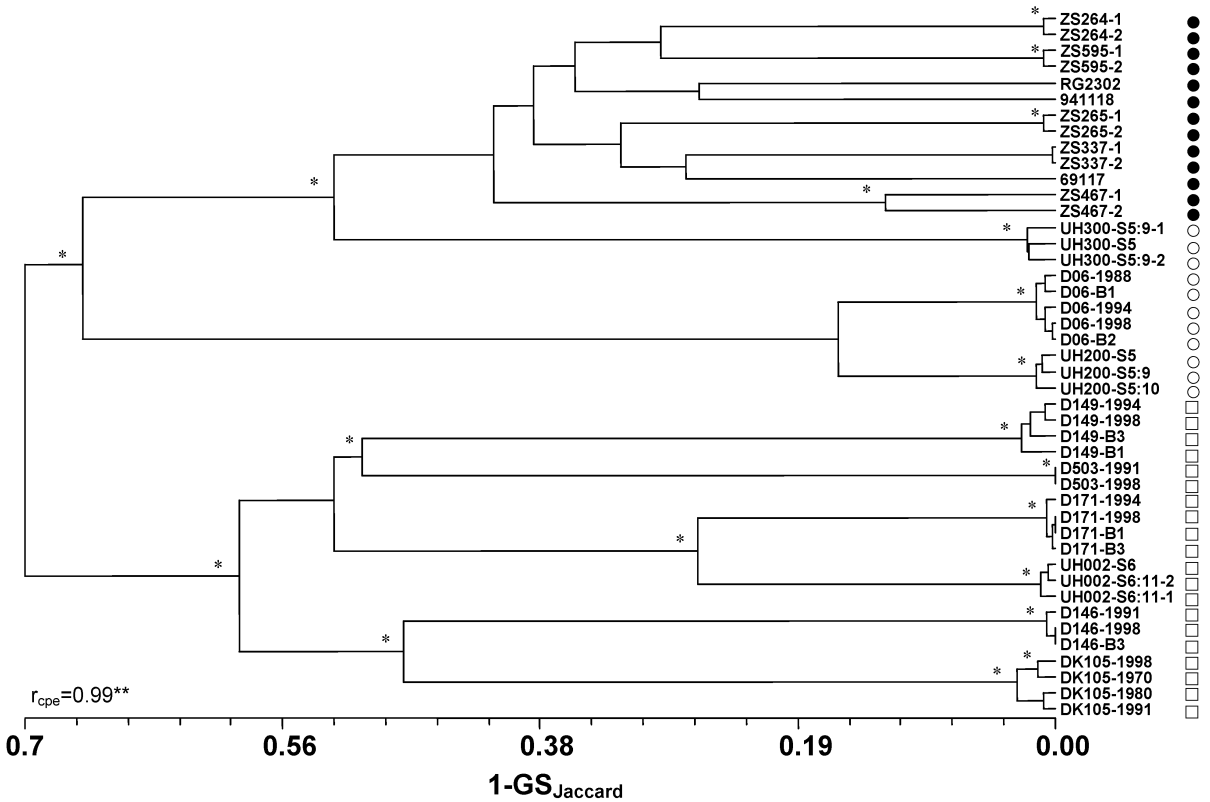

Figure 1. Associations among accessions of maize inbred lines revealed by UPGMA cluster analysis based on genetic distances calculated from AFLP data. Asterisks (*) at the forks indicate that the group to the right of the fork was found in at least 95% of 10000 bootstrap runs. DH lines and their parents are marked by filled circles (●). Flint and dent lines are marked with squares (□) and circles (●;○), respectively. DH lines were derived from F_1-hybrid 941118 generated by crossing lines s69117 and RG2302.

Table 3. Correlations between GD_{SSR} and GD_{AFLP} based on GDs of all 946 pairs of accessions or only GDs between 46 pairs of accessions from the same line (GD_w) for single chromosomes

	$r(GD_{SSR}, GD_{AFLP})$	
Chromosome	All GDs	GD_w
1	0.82**	0.01
2	0.91**	0.32*
3	0.84**	–0.02
4	0.87**	–0.19
5	0.84**	–0.16
6	0.81**	–0.03
7	0.89**	0.18
8	0.76**	0.02
9	0.83**	–0.08
10	0.60**	–0.10
All chromosomes	0.97**	0.07

generally lower degree of polymorphism of AFLPs compared to SSRs (Powell et al. 1996) may also contribute to the observed low GD_w values for AFLPs.

A variety of genetic and technical causes can contribute to the observed variation of GD_w estimates between accessions of the same line. First, the occurrence of point mutations is a possible cause. A mutation can result in a loss of an AFLP band, if it renders or disrupts the recognition site of the restriction enzyme or the selective bases of the primer. The natural mutation rate for higher eukaryotes was estimated to range from 10^{-9} to 10^{-7} per bp and generation (Drake et al. 1998). No data on mutation rates of AFLPs is yet available. However, if it is not significantly higher than the above mentioned values, mutations can explain only a minor fraction of the observed variation for GD_w values.

Second, lines in higher selfing generations of maintainance breeding are fixed for different alleles because the parental S_5 or S_6 lines were still heterozygous for a minor proportion of the genome.

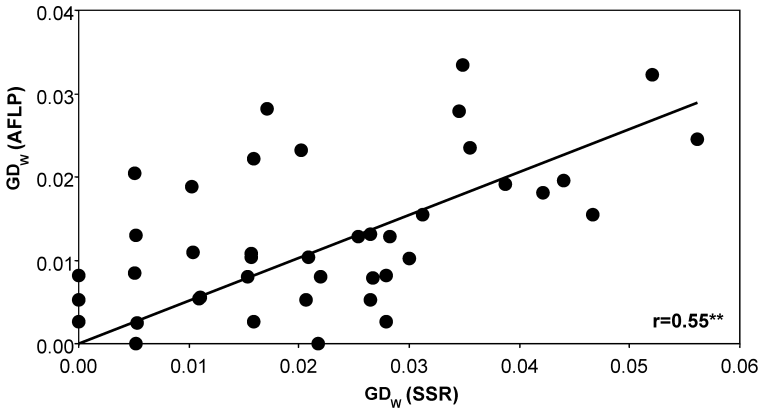

Figure 2. Plot of GD_w values revealed by 100 SSR markers and 20 AFLP primer combinations without outliers.

Third, contamination with foreign pollen could contribute to the observed variation within lines. However, this can be ruled out, because it should affect a large number of loci and, consequently, would be detectable as outliers. In addition, also our SSR results do not support the hypothesis of contamination with foreign pollen.

Fourth, heteroduplex bands can be a source for heterogeneity in SSR fingerprints (Hatcher et al. 1993). Heteroduplex bands emerge by annealing of two DNA fragments of unequal sequence or length and, therefore, tend to retard the migration of the particular band during electrophoresis (Hatcher et al. 1993; Nataraj et al. 1999). In AFLPs, no information on heteroduplex bands is available. Even though the occurrence of heteroduplexes in AFLPs cannot be ruled out as a source of variation in GD estimates, heteroduplex AFLP fragments were never encountered when the sequences derived from both strands of single AFLP markers were compared (J. Rouppe van der Voort and J. Peleman, pers. comm.).

In contrast, genotyping samples from bulked individuals may slightly decrease the variation within lines.

DH lines

DH lines were included in our study to distinguish between genetic and technical reasons for variation of GD within accessions of the same line. Since DHs are homogeneous and genetically uniform, differences among samples are only attributable to technical causes. While SSRs yielded identical fingerprints, replications of DH lines were not scored identical with AFLPs. As the seed samples of the two accessions fingerprinted in our study were drawn out of the same seed lot, they should be scored as identical. Therefore, segregation and bulking effects can be ruled out as reasons for the observed variation of AFLPs within DH lines.

Technical variation due to poor DNA quality, incomplete digestion of DNA, inconsistent amplification, or scoring problems of the applied software are the most probable reasons for the observed variation between identical samples of DH lines. In addition, heterogeneity within DH lines was observed by Murigneux et al. (1993). However, this is very unlikely, as our SSR results do not support this hypothesis.

Further investigations revealed that the DNA of accession ZS467-2 was incompletely digested, which explains the extraordinarily high variation between the two accessions of this DH line. This indicates that even for a highly reproducible marker technique such as AFLPs, routine analyses could lead to incorrect results in the case of a suspected EDV. Therefore, we recommend replication of the lab assays to minimize the experimental error. Given the high value of litigation involved in EDV claims, additional costs for replicated lab assays are well justified.

Locations of variation

Matches in the locations of variation detected with AFLPs and SSRs could be caused by the fact that certain genomic regions were still segregating by the time the accessions were separated. In addition, it cannot be ruled out that these matches are attributable to chance. However, when the BIN positions of markers contributing to the observed variation within lines were compared, only a low coincidence between the

locations of variation for AFLPs and SSRs was detected. For example, chromosome 9 revealed the highest number of SSRs displaying variation between accessions of the same line but the lowest for AFLPs. This observed discrepancy is presumably attributable to the different methods by which dominant AFLPs and codominant SSRs generate polymorphisms. In addition, the number of markers displaying variation within maize lines was too small in order to assign the observed variation unambiguously to certain genomic regions. Therefore, inferences about genomic regions showing variation in GD_w estimates must be considered with caution.

Use of AFLPs and SSRs for identification of EDVs

The rationale underlying the use of molecular markers in testing for EDVs is to take the GD estimated from marker data as an indicator for the true GD between two genotypes across the entire genome. If the estimated GD exceeds a certain threshold, the two genotypes are considered as independently derived, whereas otherwise this is taken as evidence for a putative EDV. Obviously, any variation between different accessions of the same inbred or DH line attributable to genetic and/or technical reasons is not only reflected by the GD_w but will also affect the GD between accessions of different lines. High levels of variation within lines have as consequence a decreased resolution to distinguish related lines. In general, this type of variation will inflate the Type I error in testing for essential derivation (H_0: One of the two lines is an EDV of the other) and reduce the Type II error. If the range of GD values between accessions of the same line is large, it might be possible that a true EDV could be judged as independently derived or a truly independent variety as essentially derived just by genotyping different accessions of each variety (Heckenberger et al. 2002). Subsequently, we compare RFLP, RAPD, AFLP, and SSR markers under this aspect and other criteria relevant in investigations of EDVs.

First, the reproducibility of molecular marker data using different sources of DNA extracted from the same accession or across laboratories must be high. This criteria excludes RAPDs from the list of possible marker systems because they were found to be less reproducible than AFLPs or SSRs (Bagley et al. 2001; Jones et al. 1997).

Second, the applied marker system should reveal a high degree of polymorphism. In the ideal case, if unrelated genotypes have no marker bands in common (GD = 1), identical bands in related individuals are exclusively attributable to 'identity by descent' and, thus, directly reflect the degree of relatedness. However, with GD < 1 for unrelated genotypes, some marker bands are 'identical in state' and this must be taken into account when defining EDV thresholds. In our study, the mean GD for unrelated dent lines was significantly different between AFLPs and SSRs suggesting that the breeders must agree on different EDV thresholds for each marker system.

Third, the applied marker system should warrant at low costs a uniform and dense coverage of the entire genome to obtain unbiased GD estimates with small standard errors. In principle, this requirement can be met by AFLPs, RFLPs, and SSRs, but AFLPs offer the advantage that a large number of markers can be produced per primer combination. In comparison with AFLPs, RFLPs and SSRs have a lower MI, but their information content is higher due to the codominant inheritance and the higher degree of polymorphism. In addition, detailed marker information, including primer sequences and map position, are publicly available for RFLPs and SSRs in maize (see MaizeDB http://nucleus.agron.missouri.edu/index.html, confirmed 16 April 03). Owing to limitations in the automation and standardization of RFLPs, they represent no longer a competitive alternative to SSRs and AFLPs for EDV analyses.

AFLP markers can be produced at lower costs per data point than SSRs, because one primer combination produces a large number of bands. Despite this economic advantage of AFLPs, the French association of maize breeders (SEPROMA) recently recommended a set of uniformly distributed SSR markers for EDV identification. SSRs provide a higher degree of transparency for legal issues than AFLPs due to their codominant inheritance, their known map positions, and their public availability. However, since the set of SSRs is known to all breeders, it is, in principle, possible to use this information to select for genetic diversity at some SSR markers to avoid an EDV, while maintaining a high degree of relatedness in other genomic regions. Complementary use of AFLP markers would prevent this situation, because selection for variants at specific marker loci would be more difficult for AFLPs than for a fixed set of SSRs. In addition, the redundancy in the genotypic information caused by clustered AFLPs could be omitted by applying map based genetic distances (Dillmann et al. 1997).

Taking into account all criteria, both AFLPs and SSRs are suitable marker systems for EDV identification. In order to counterbalance advantages and disadvantages of each marker system, AFLPs and SSRs could be used in a complementary way to unambiguously distinguish EDVs and independent varieties. We further conclude that the stability of marker data across different generations of multiplication or accessions maintained by different breeders is primarily a function of marker reproducibility and residual heterozygosity or heterogeneity. Therefore, with regard to the use of DNA markers for resolving EDV issues, it is important to reduce residual heterozygosity before applying for plant variety protection. This can be achieved by further selfing and/or pre-screening of lines with molecular markers for homogeneity or by production of DH lines.

Acknowledgements

The present study was supported by a grant from the European Union, Grant No. QLK-CT-1999-01499 (MMEDV) (for further information, see http://www.niab.com/bbp/mmedv/, confirmed 16 April 03). In addition, the authors wish to thank Semundo (Hadmersleben, Germany) for supplying the DH lines.

References

Anscombe F.J. and Tukey J.W. 1963. The examination and analysis of residuals. Technometrics 5: 141–160.

ASSINSEL 2000. DUS testing: Phenotype vs. Genotype. ASSINSEL, Position Paper adopted at the Rome Congress in May 2000.

Bagley M.J., Anderson S.L. and May B. 2001. Choice of methodology for assessing genetic impacts of environmental stressors: polymorphism and reproducibility of RAPD and AFLP fingerprints. Ecotoxicology 10: 239–244.

Bar-Hen A. and Charcosset A. 1994. Relationship between molecular and morphological distances in a maize inbred lines collection. Application for breeder's rights protection. Biometrics in Plant Breeding: Applications of Molecular Markers. In: Proceedings of the Ninth Meeting of the EUCARPIA section Biometrics in Plant Breeding. pp. 57–66.

Dice L.R. 1945. Measures of the amount of ecologic association between species. Ecology 26: 297–302.

Dillmann C., Bar-Hen A., Guerin D., Charcosset A. and Murigneux A. 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.

Drake J.W., Charlesworth B., Charlesworth D. and Crow J.F. 1998. Rates of spontaneous mutation. Genetics 148: 1667–1686.

Efron B. 1979. Bootstrap methods: another look at the jackknife. Ann. Stat. 7: 1–26.

Hansen M., Kraft T., Christiansson M. and Nilsson N.O. 1999. Evaluation of AFLP in *Beta*. Theor. Appl. Genet. 98: 845–852.

Hatcher S.L., Lambert Q.T., Raymond L.T. and Carlson J.R. 1993. Heteroduplex formation: a potential source of errors from PCR products. Prenatal Diagnosis 13: 171–177.

Heckenberger M., Bohn M., Ziegle J.S., Joe L.K., Hauser J.D., Hutton M. et al. 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.

Jaccard P. 1908. Nouvelles recherches sur la distribution florale. Bull. Soc. Vaud. Sci. Nat. 44: 223–270.

Johnson N.L., Kotz S. and Kemp A.W. 1992. Univariate Discrete Distributions. J. Wiley and Sons, Inc., New York.

Jones C.J., Edwards K.J., Castaglione S., Winfield M.O., Sala F., Vandewiel C. et al. 1997. Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. Mol. Breed. 3: 381–390.

Knaak C., Förster J. and Jäger-Gussen M. 1996. 'Abgeleitete' Sorten aus praktischer Sicht. Bericht über die 47. Arbeitstagung 1996 der Saatzuchtleiter im Rahmen der 'Vereinigung österreichischer Pflanzenzüchter' BAL Gumpenstein. pp. 167–172.

Link W., Dixkens C., Singh M., Schwall M. and Melchinger A.E. 1995. Genetic diversity in European and mediterranean faba bean germ plasm revealed by RAPD markers. Theor. Appl. Genet. 90: 27–32.

Lübberstedt T., Melchinger A.E., Dussle C., Vuylsteke M. and Kuiper M. 2000. Relationships among early European maize inbreds: IV. Genetic diversity revealed with AFLP markers and comparison with RFLP, RAPD, and pedigree data. Crop. Sci. 40: 783–791.

Lübberstedt T., Vuylsteke M., Dussle C., Kuiper M. and Melchinger A.E. 1999. Untersuchungen zur genetischen Diversität zwischen europäischen Maisinzuchtlinien mit AFLPs, RFLPs, RAPDs, und SSRs – ein Vergleich verschiedener Markersysteme. Bericht über die 49. Arbeitstagung 1998 der Vereinigung österreichischer Pflanzenzüchter, BAL Gumpenstein. pp. 45–49.

Mantel N.A. 1967. The detection of disease clustering and a generalized regression approach. Cancer Res. 27: 209–220.

Melchinger A.E. 1999. Genetic Diversity and Heterosis. In: Coors J.G. and Pandery S. (eds), The Genetics and Exploitation of Heterosis in Crops. Crop Science Society of America, Inc., Madison, USA, pp. 99–118.

Murigneux A., Barloy D., Leroy P. and Beckert M. 1993. Molecular and morphological evaluation of doubled haploid lines in maize. 1. Homogeneity within DH lines. Theor. Appl. Genet. 86: 837–842.

Nataraj A.J., Olivos-Glander I., Kusukawa N. and Highsmith W.E. 1999. Single-strand conformation polymorphism and heteroduplex analysis for gel-based mutation detection. Electrophoresis 20: 1177–1185.

Nei M., Tajima F. and Tateno Y. 1983. Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. J. Mol. Evol. 19: 153–170.

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

Powell W., Morgante M., Andre C., Hanafey M., Vogel J., Tingey S. et al. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol. Breed. 2: 225–238.

Rohlf F.J. 1989. NTSYS-pc Numerical Taxonomy and Multivariate Analysis System. Exeter Publishing Co, Ltd., Setauket, NY, USA.

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

SAS Institute 1988. SAS/STAT User's guide, Release 6.03 edn. SAS, Cary.

Smith J.S.C., Smith O.S., Bowen S.L., Tenborg R.A. and Wall S.J. 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.

Vos P., Hogers R., Bleeker M., Reijans M., Lee T.v.de., Hornes M. et al. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23: 4407–4414.

Yap I.V., Nelson R.J. and WinBoot 1996. A program for performing bootstrap analysis of binary data to determine confidence limits of UPGMA-based dendrograms. International Rice Research Institute, Manila, Philippines.