Theor Appl Genet (1995) 91:1086–1091

P. Skroch · J. Nienhuis

Impact of scoring error and reproducibility of RAPD data on RAPD based estimates of genetic distance

Received: 17 November 1994 / Accepted: 26 May 1995

Abstract RAPD band reproducibility and scoring error were evaluated for RAPDs generated by 50 RAPD primers among ten snap bean (*Phaseolus vulgaris* L.) genotypes. Genetic distances based on different sets of RAPD bands were compared to evaluate the impact of scoring error, reproducibility, and differences in relative amplification strength on the reproducibility of RAPD based genetic distance estimates. The measured RAPD data scoring error was 2%. Reproducibility, expressed as the percentage of RAPD bands scored that are also scored in replicate data, was 76%. The results indicate that the probability of a scored RAPD band being scored in replicate data is strongly dependent on the uniformity of amplification conditions between experiments, as well as the relative amplification strength of the RAPD band. Significant improvement in the reproducibility of scored bands and some reduction in scoring error was achieved by reducing differences in reaction conditions between replicates. Observed primer variability for the reproducibility of scored RAPDs may also facilitate the selection of primers, resulting in dramatic improvements in the reproducibility of RAPD data used in germplasm studies. Variance of genetic distances across replicates due to sampling error was found to be more than six times greater than that due to scoring error for a set of 192 RAPD bands. Genetic distance matrices computed from the RAPD bands scored in replicated data and RAPD bands that failed to be scored in replicated data were not significantly different. Differences in the ethidium bromide staining intensity of RAPD bands were not associated with significant differences in resulting genetic distance matrices. The assumption of sampling error as the only source of error was sufficient to account for the observed variation in genetic distance estimates across independent sets of RAPD bands.

Communicated by G. E. Hart

P. Skroch (✉) · J. Nienhuis
Department of Horticulture, University of Wisconsin, Madison, Wisconsin 53706, USA

Key words *Phaseolus vulgaris* · RAPD variation Reproducibility · Scoring error · Genetic distance

Introduction

An important application of molecular marker technology has been in the measurement of genetic diversity and genetic relationships among individuals and populations. RFLPs and isozymes have been extensively applied in this area. More recently, the use of RAPD (Welsh and McClelland 1990; Williams et al. 1990) molecular markers has become widespread. RAPD derived estimates of genetic relationships have been shown to be in general agreement with relationships based on pedigree, RFLP, and isozyme data (pedigree: Vierling and Nguyen 1992, Yang and Quiros 1993; RFLP: Dos Santos et al. 1994; isozyme: Yang and Quiros 1993; Heun et al. 1994). However, problems with the reproducibility of RAPD amplification and with scoring error for RAPD data have also been reported.

Using a genetic test of reliability, Weeden et al. (1992) found that error due to mistakes in the scoring of RAPD markers as present or absent ranged from 0 to 8% depending on the relative amplification strength as measured by the intensity of ethidium bromide staining. Possible sources of scoring error have been identified (Hunt and Page 1992; Heun and Helentjaris 1993; Huang and Jeang 1994). In addition, theoretical studies (Lamboy 1994a, b) have shown how marker scoring errors can bias genetic similarity estimates and have suggested replication as a means of obtaining estimates of the frequencies of scoring errors. Nevertheless, the magnitudes and effects of these biases have not been established experimentally.

The problem of RAPD data reproducibility has been addressed differently by different researchers. Dos Santos et al. (1994), Beebe et al. (1995) and Nienhuis et al. (1995) reported results based on unreplicated data. Vierling and Nguyen (1992) and Stiles et al. (1993) merely

referred to their RAPD banding patterns as reproducible. Aruna et al. (1993) only used intensely staining, and clearly resolvable bands and reported that replicated DNA samples showed no variation between them. Ellsworth et al. (1993) recommended standardization and internal control to mitigate the problem of artifactual variation. Others have used partial or complete replication and eliminated bands that were not consistently amplified (Kresovich et al. 1992; Yang and Quiros 1993; Novy et al. 1994).

The above studies, and others, either do not report specific levels of reproducibility or do not demonstrate a relationship between data replication and an improvement in the accuracy or precision of the resulting genetic distance or similarity estimates. Failure of RAPDs to be scored consistently across replicates (PCR runs) may reflect the sensitivity of RAPD amplification to subtle variation in the reaction environment and not the validity of the data.

The specific objectives of the research reported in this paper are: (1) estimation of RAPD data reproducibility and evaluation of the factors that may influence reproducibility, (2) estimation of RAPD scoring error and evaluation of the factors which may influence RAPD scoring error for RAPD bands that can be scored across replicates, and (3) evaluation of the impact of reproducibility, scoring error, and RAPD amplification strength on estimates of genetic distance among ten bean genotypes.

Materials and methods

Plant material, DNA extraction, RAPD reactions, and data scoring

Bean genotypes, DNA extractions, and RAPD reactions were as described by Skroch and Nienhuis (1995). Data was scored and RAPDs classified into faint, medium and bold categories as described by Skroch and Nienhuis (1995). The term "RAPD band" is used here, as it was in Skroch and Nienhuis (1995), to describe a set of comigrating RAPD fragments amplified by the same RAPD primer.

Comparison of replicated RAPD data sets

Three experiments were performed, in each of which two replications of data were generated and compared. Gels were scored independently and subsequently compared to identify RAPDs scored in only one or both replicates. For RAPD bands scored in both replicates, scoring error was measured as the percentage of RAPD scores which differed across replicates for comparisons of the same genotype. RAPD data reproducibility was defined as the average percentage of RAPD bands scored in one replicate that were also scored in the other replicate. RAPD primer reproducibility was defined as RAPD data reproducibility computed on a per-primer basis.

These three experiments represent a series in which conditions between replications became more and more similar. In experiment 1, original data which was collected in 1991–1992 was compared with replicated data collected a year later in 1993. In experiment 2, both replications were completed on the same day using identical stock solutions for mixing reactions. In experiment 3, a further reduction in variability was achieved by mixing reactions for both replications from the same stock solutions and in the same tube, and then running reaction products for both replications side by side on the same gel.

Experiment 1: scoring error and reproducibility

The original data, as described by Skroch and Nienhuis (1995), was used to rank primers by the number of polymorphisms scored per primer, and the 50 primers with the largest number of polymorphic bands per primer were chosen for replication (in a few cases substitutions were made because primers were not available). To avoid possible confounding effects due to within line variability, original DNA extractions were used for replication. Reactions were performed for these 50 primers and ten genotypes as already described (Skroch and Nienhuis 1995). To avoid bias, results were scored without reference to the original data. This replicated data was then compared to the original data (Skroch and Nienhuis 1995) for the selected 50 primers. In the following discussion, the original and replicate data for the 50 primers used in this experiment correspond to replicate 1 and replicate 2 of experiment 1, respectively.

Experiment 1: comparison of distance matrices

Two estimators of genetic distance commonly used in the analysis of RAPD marker data were employed: the complement to the simple matching coefficient ($1 - S$) and the complement of Jaccard's similarity coefficient ($1 - J$). The difference between two individuals estimated as $1 - J$ or $1 - S$ for RAPD marker data is the proportion of differences found for a specified set of RAPD bands (Gower 1972). Using the formula for sample variance (Steel and Torrie 1980) and recognizing that $1 - S$ and $1 - J$ are means over loci, the sample variance of a genetic distance, d, can be written as $\text{var}(1 - J) = J(1 - J)/(n_j - 1)$ and $\text{var}(1 - S) = S(1 - S)/(n_s - 1)$, where n_j and n_s are the number of observations in each estimate.

RAPD data generated in experiment 1 was partitioned by the classification of RAPD bands into faint, medium, or bold categories in replicate 1 and by whether RAPDs were scored in only one or both replicates. Distance matrices derived from these data subsets were compared by the analysis of the expected and observed rates of type 1 errors for t tests ($\alpha = 0.05$) computed for the pairwise comparison of the component distances. Comparisons were restricted to distance matrices derived from disjoint subsets of RAPD bands. Since the same distances were being computed in both matrices, any significant differences were considered as type 1 errors. Greater then expected type 1 errors would lead to rejection of the null hypothesis that the two specified independent band subsets were estimating the same genetic distances. Since all comparisons of distance matrices involved the comparison of independent data sets, statistical tests for the pairwise comparison of individual distances were legitimate. However, since component distances within a distance matrix were not independent, rates of type 1 error for the set of comparisons between matrices may not have followed the theoretical distribution. Thus, to evaluate the significance of the observed results, randomization tests were used (Edgington 1980). P values for observed rates of type 1 errors were computed for each comparison of distance matrices based on the analysis of 1000 random permutations of bands between groups.

Sampling error variance was compared to scoring error variance for distances estimated as the complement to the simple matching coefficient. The mean scoring error variance of a genetic distance estimate was computed from replicate 1 and replicate 2 distances, based on RAPD bands scored in both replicates. The average sampling error variance was computed based on genetic distance estimates averaged across replicates using the formula for sample variance given above.

Experiment 2

The results from the first experiment were used to rank primers by the number of bands per primer which were not scored in both replications. Ten of the twelve primers which had the largest number of

bands not scored in both replications were selected for the second and third experiments. In each of these experiments two sets of RAPD data were generated using these ten primers and the ten original genotypes. In experiment 2, reactions for both replications were mixed from the same stock solutions of enzyme, reaction buffer, dNTPs, primer, and template, on the same day and run in the same machine. However, reaction mixes for each replication were mixed independently and the amplification products were electrophoresed on separate gels.

Experiment 3

In experiment 3, reaction mixes for each of the ten primers, including *Taq*, buffer, primer, and dNTPs, for both replications were mixed from the same stock solutions and in the same micro-centrifuge tube so there could be no differences between replications due to pipetting or mixing polymerase, buffer, primer, or dNTPs. The reaction products were electrophoresed with replications for each primer side by side on the same gel.

Results

Experiment 1: scoring error and reproducibility

A high level of variation for RAPD bands scored was observed between replicates for 50 RAPD primers (Table 1). A total of 318 RAPDs were scored, including 48 in replicate 1 only, 78 in replication 2 only, and 192 in both replicates. The 126 RAPD bands not scored in both replicates were distributed as 94 faint, 27 medium and 5 bold bands. The difference between the number of bands scored only in replication 1 and the number of bands scored only in replication 2 was due to more bands being scored in the faint category in replication 2. RAPD band reproducibility was 60, 83, and 96% for faint, medium and bold bands, respectively. Averaged over all band categories, reproducibility was 76% indicating that, on average, 76% of RAPD bands scored in one replicate were also scored in the other replicate.

Scoring error for RAPD bands scored in both replicates varied from a low of 1.3% for bands classified as bold, to 2.2% for RAPDs classified as faint, to a maximum of 2.4% for bands classified as medium. The average error rate based on 1835 total observations over all band categories was 2.0%. The 37 scoring errors detected were distributed into 18 cases where a RAPD was scored for a genotype in replicate 1 but was not scored for the same genotype in replicate 2, and 19 cases where a RAPD was not scored for a genotype in replicate 1 but was scored for the same genotype in replicate 2. The observed distribution of errors among genotypes did not differ from a random distribution as tested by chi-square ($P > 0.05$). However, the distribution of errors among RAPD bands was non random. Scoring errors were distributed among only 20 RAPD bands scored for 17 RAPD primers. Furthermore, four RAPD bands out of 192 accounted for 18 of the 37 scoring errors.

Table 1 Distribution and reproducibility of RAPD bands classified as faint, medium, and bold in original data (rep 1) and replicated data (rep 2) for 50 RAPD primers and ten bean genotypes

Band category	Number of bands scored				RAPD band reproducibility
	Rep1[a]	Rep2[b]	Rep1 and rep2[c]	Total	
Bold	3	2	57	62	96
Medium	13	14	66	93	83
Faint	32	62	69	163	60
All bands	48	78	192	318	76

[a] Scored in replication 1 only, classified as faint, medium or bold in replication 1
[b] Scored in replication 2 only, classified as faint, medium or bold in replication 2
[c] Scored in both replicates, classified as faint, medium or bold in replication 1

Experiment 1: variation among primers for number and reproducibility of scored RAPD bands

Primers were ranked by the reproducibility of scored RAPD bands. The highest ranking 25 primers, in order from highest to lowest ranking, were A14, H16, H12, G3, A2, B10, J4, O13, H15, A4, P1, A17, W6, I16, F13, I10, M12, O19, U19, A9, R4, O16, and G5. For these primers, the average primer reproducibility was 88% compared to 78% for all 50 primers. Averaged across replicates, 5.11 polymorphic RAPD bands were scored per primer while the average for all 50 primers was 5.24.

Primers were also ranked by the number of polymorphic bands scored per primer, averaged across replicates. The highest ranking 25 primers were U1, W6, O5, G5, O16, R4, A9, G8, U19, O19, M12, I16, T15, J5, A7, G16, F20, I3, B5, I7, X11, F6, J4, B10, and A2, again in order from highest to lowest ranking. For these primers, average primer reproducibility was 76% compared to 78% for all 50 primers and 88% for the best 25 primers ranked by primer reproducibility. Averaged across replicates, 5.94 polymorphic RAPD bands were scored per primer.

Experiment 1: comparison of genetic distance matrices

Genetic distances were compared for different subsets of data as defined by RAPD band classification into faint, medium and bold categories in replication 1 and by whether RAPDs were scored in only one or both replicates (Table 2). Pairwise comparisons among distance matrices were made to evaluate the consistency of genetic distance estimates across different categories of data. The numbers of statistically different distance estimates varied from zero to five. For both genetic distance estimators, computed P values based on randomization tests indicated that the observed distribution of significant t tests were within the range expected

Table 2 Total numbers of significant t-tests found for 45 t-tests performed in the comparison of different genetic-distance matrices generated using subsets of data from experiment 1

Data set 1[a]	Data set 2[a]	1-SM		1-J	
		T-tests[b]	P[c]	T-tests[d]	P
Bands scored only in replication 1	Bands only scored in replication 2	1	0.57	2	0.44
Bands scored in both replications	Bands scored in only one replicate	2	0.39	1	0.57
Bands scored and classified in replication 1					
Faint bands	Bold bands	1	0.63	0	0.83
Faint bands	Medium bands	2	0.39	0	0.81
Medium bands	Bold bands	3	0.25	3	0.25
Bands scored in both replications and classified in replication 1					
Faint bands	Bold bands	1	0.59	3	0.28
Faint bands	Medium bands	5	0.09	3	0.32
Medium bands	Bold bands	4	0.15	2	0.43

[a] Genetic distances, for each genotype pair, based on RAPD bands specified under the column heading 'Data set 1', were compared to the corresponding genetic-distance estimates calculated from RAPD bands specified under the column heading 'Data set 2'
[b] Number of significant t-tests found in comparison of the 45 genetic-distance estimates calculated for 'Data set 1' with those calculated from 'Data set 2' using the compliment to the simple matching coefficient to calculate distance
[c] Probability of obtaining a larger number of significant comparisons by chance, computed using a randomization test employing 1000 random permutations of bands between specified groups.
[d] Number of significant t-tests found in comparison of the 45 genetic-distance estimates calculated for 'Data set 1' with those calculated from 'Data set 2' using the complement of Jaccard's genetic-similarity coefficient

Table 3 Summary of results from three experiments comparing paired replicates of RAPD data for ten bean genotypes

Experiment[a]	RAPD bands				Data points	
	Faint: med: bold[b]	No. scored in both reps	Total no. scored	Reproducibility (%)[c]	Number of observations	Scoring error (%)[d]
Exp. 1a	1.0:0.8:0.5	192	318	76	1835	2.0
Exp. 1b	1.0:0.6:0.4	32	78	59	315	1.9
Exp. 2	1.0:0.5:0.3	62	84	85	620	1.6
Exp. 3	1.0:0.5:0.4	56	64	94	558	0.7

[a] Exp. 1a designates results from experiment 1, exp. 1b designates the subset of experiment 1 which included only the ten primers selected for experiment 2 (exp. 2) and experiment 3 (exp. 3)
[b] Relative numbers of RAPD bands classified as faint, medium and bold in each replicate, averaged over two replications for each experiment
[c] The average percentage of RAPD bands scored in one replicate that were also scored in the other replicate
[d] The percentage of RAPD scores which differed across replicates for comparisons of the same genotype

due to chance. The distribution of significant t tests was random. In almost all comparisons the significant differences found were different for the two genetic distance estimators.

Bands that were scored exclusively in replication 1 or exclusively in replication 2 gave genetic distance values which were statistically consistent with those calculated from RAPD bands scored in both replicates (Table 2, row 2) indicating no detectable difference between replicated and unreplicated data. Furthermore, statistical differences were not found for distance matrices computed from bands of different amplification strength.

Experiment 1: comparison of sampling and scoring error variance

For genetic distances estimated as the complement to the simple matching coefficient, sampling error variance (11.1×10^{-4}) was 6.65 times greater, on average, than scoring error variance (1.67×10^{-4}) for the 192 RAPD bands scored in both replicates. Doubling the number of RAPD bands scored from 192 to 384 would result in a reduction in sampling variance (5.55×10^{-4}) equal to 1/2 the sampling variance for 192 RAPD bands. This reduction in sampling variance is greater than 3 times the estimated variance due to scoring error. Thus, even if it were possible to completely eliminate scoring error variance by data replication, doubling the number of bands scored would result in a 3-fold greater reduction in the total variance (sampling variance + marker scoring error variance) of genetic distance estimates.

When combining data generated from independent PCR runs, eliminating bands that are not scored in all sets of genotypes increases the sampling error for comparisons of genotypes within each set. For each replicate, the average sampling variance was computed for genetic distances estimated from the 192 RAPD bands

scored in both replicates and for genetic distances estimated from the total number of bands scored in each replicate separately. Averaged over replicates, sampling variance would be increased by 24.5% if RAPD bands not scored in both replicates were eliminated. The magnitude of this increase (2.74×10^{-4}) is 1.64 times the estimated variance of genetic distances due to scoring error.

Experiments 2 and 3

Based on the data from experiment 1, ten primers were chosen for experiments 2 and 3. Expressed as a percentage, RAPD data reproducibility for bands scored for these ten primers was only 59% in experiment 1. Reproducibility increased to 85% in experiment 2 (Table 3). This increase in reproducibility was correlated with the removal of factors contributing to differences in reaction conditions. RAPD data reproducibility increased to 94% in experiment 3 after additional factors contributing to variability among replicates had been eliminated. The relative abundance of faint, medium, and bold bands was relatively stable across experiments. Scoring error for RAPDs scored in both replications was low in all experiments, decreasing from 1.9% in experiment 1 to 0.7% in experiment 3 (Table 3).

Discussion

The series of experiments in which reproducibility increased with the removal of sources of difference in reaction conditions between replications, establishes the problem of reproducibility as one of replicating an adequate amplification environment for a particular RAPD band. Furthermore, as shown by the classification of RAPD bands into faint, medium, and bold categories, the reproducibility of RAPD fragment amplification was correlated with band quality. The failure of all RAPD bands to be scored in replicated data sets is most likely due to the sensitivity of the RAPD technique and the particular RAPD band, as well as the difficulty of exact duplication of reaction conditions.

One possible strategy for improving the reproducibility of RAPD data would be to select primers based on the reproducibility of the data they generate. Results of experiment 1 indicated variability among primers, suggesting this possibility. Furthermore, the results indicate that this could be achieved without a reduction in the average number of bands scored per primer. Alternatively, primers may be selected to maximize the number of RAPD bands scored. The average reproducibility for the highest ranking 25 primers, ranked on the basis of the number of bands scored per primer, was comparable to results for all 50 primers. Thus, selection of primers generating larger numbers of polymorphisms can be accomplished without a reduction in the reproducibility of the resulting data.

Scoring error was estimated to be only 2% in this study. Thus, we did not find rates of scoring error as high as those reported by Weeden et al. (1992). Furthermore, in contrast to their results, the results reported here indicated no relationship between the rate of scoring error and the relative amplification strength of RAPD bands. One possible explanation for this inconsistency is that our method results in a lower frequency of scoring errors. Alternatively the discrepancy could be due to methodological differences in the detection of error. Certain types of errors, such as those occurring because of competitive inhibition of weakly amplifying RAPDs, could be reproducible and not detectable by data replication. However, the frequencies of such errors could be estimated by genetic analysis such as that reported by Weeden et al. (1992) or by analysis of fragment frequencies as reported by Skroch and Nienhuis (1995). These studies may indicate a similar relationship between scoring error and the classification of RAPD bands into faint, medium and bold categories.

The results indicated that for 192 RAPD bands, doubling the number of RAPD bands scored would have a 3-fold greater impact on the variance of genetic distance estimates than the elimination of scoring error. The effect of sample size is important to the interpretation of our results because sampling variance is inversely proportional to the number of bands scored while scoring error variance is independent of band number. For example, if genetic distances were based on 96 RAPD bands, instead of 192, the relative importance of sampling error would be twice as great. RAPD data sets used in germplasm studies are typically much smaller than 192 bands. Thus, with respect to the relative efficiency of replication, the conclusions based on our results are probably conservative.

There are other sources of error which cannot be directly quantified. For example, violation of the assumptions underlying the estimation of genetic distance based on RAPD data. Specifically, for genetic distance estimates based on the complement to the simple matching coefficient, it is assumed that co-migrating RAPDs are amplified from the same genomic locus and that two genotypes sharing the absence of a RAPD do so because of similarity at that RAPD locus. In addition, it is assumed that the level of sequence similarity between genomes indicated by monomorphic comparisons, and the amount of sequence difference indicated by presence versus absence of RAPD amplification, are the same for all comparisons. Violation of these assumptions would lead to an increase in the variance of genetic distance estimates. However, the assumption of sampling error as the only source of error was sufficient to account for variation in genetic distance estimates across independent data sets. Thus, the consistency of genetic distance estimates suggests the absence of significant, but unidentified, sources of error and also supports the assumptions on which estimated genetic distances were based.

The demonstration that disjoint RAPD marker data sets estimate genetic distances that differ by no more than would be expected due to sampling error alone is important; however, the comparisons in our study were also based on statistically important classification criteria. The results from the comparison of genetic distances based on unreplicated data with genetic distances based on replicated data were consistent with the hypothesis that these two sets of RAPD bands actually represent random samples from the same population of RAPDs relative to the detection of genomic similarity or difference. This conclusion counters the suggestion by other authors that RAPD bands that cannot be consistently scored across PCR runs should not be used in studies of genetic relationship (for example: Kresovich et al. 1992; Yang and Quiros 1993; Novy et al. 1994). Furthermore, these results strengthen conclusions reached in other studies based on unreplicated RAPD data, especially those which have used the same version of the RAPD technique (for example: Skroch et al. 1992a,b; Dos Santos et al. 1994; Beebe et al. 1995; Nienhuis et al. 1995; Skroch and Nienhuis 1995).

Differences in reproducibility (this study), scoring error (Weeden et al. 1992), and frequency of occurrence (Skroch and Nienhuis 1995) have been observed for RAPDs of different amplification strength. Weak amplification of some RAPD bands compared to others may reflect differences in the specificity of amplification and underlying differences in the level of sequence polymorphism that each type of RAPD represents. Distance matrices were compared to evaluate the impact that differences in statistical and physical properties of RAPDs of varying amplification strength may have on RAPD based estimates of genetic distance. Although statistical and physical differences may exist, they were not significant enough to result in differences in genetic distance estimates greater than expected by chance.

References

Aruna M, Ozias-Akins P, Austin ME, Kochert G (1993) Genetic relatedness among rabbiteye blueberry (*Vaccinium ashei*) cultivars determined by DNA amplification using single primers of arbitrary sequence. Genome 36:971–977

Beebe SE, Ochoa I, Skroch PW, Nienhuis J, Tivang J (1995) Genetic diversity among common-bean breeding lines developed for Central America. Crop Sci 35:1173–1178

dos Santos JB, Nienhuis J, Skroch P, Tivang J, Slocum MK (1994) Comparison of RAPD and RFLP genetic markers in determining genetic similarity among *Brassica oleracea* L. genotypes. Theor Appl Genet 87:909–915

Edgington ES (1980) Randomization tests. Marcel Dekker, Inc., New York

Ellsworth DL, Rittenhouse KD, Honeycutt RL (1993) Artifactual variation in randomly amplified polymorphic DNA banding patterns. Biotechniques 14:214–217

Gower JC (1972) Measures of taxonomic distances and their analysis. In: Weiner JS, Huizinga J (eds) The assessment of population affinities in man. Clarendon Press, Oxford, pp 1–24

Heun M, Helentjaris T (1993) Inheritance of RAPDs in F_1 hybrids of corn. Theor Appl Genet 85:961–968

Heun M, MurphyJP, Phillips TD (1994) A comparison of RAPD and isozyme analyses for determining the genetic relationships among *Avena sterilis* L. accessions. Theor Appl Genet 87: 689–696

Huang LM, Jeang KT (1994) Long-range jumping of incompletely extended polymerase chain reaction fragments generates unexpected products. Biotechniques 16:242–246

Hunt GJ, Page RE (1992) Patterns of inheritance with RAPD molecular markers reveal novel types of polymorphism in the honey bee. Theor Appl Genet 85:15–20

Kresovich S, Williams JGK, McFerson JR, Routman EJ, Schaal BA (1992) Characterization of genetic identities and relationships of *Brassica oleracea* L. via a random amplified polymorphic DNA assay. Theor Appl Genet 85:190–196

Lamboy (1994a) Computing genetic similarity coefficients from RAPD data: the effects of PCR artifacts. PCR Methods Applic 4:31–37

Lamboy (1994b) Computing genetic similarity coefficients from RAPD data: correcting for the effects of PCR artifacts caused by variation in experimental conditions. PCR Methods Applic 4:38–43

Nienhuis J, Tivang J, Skroch PW, dos Santos JB (1995) Genetic relationships among cultivars and landraces of lima bean (*Phaseolus lunatus* L.) as measured by RAPD markers. J Am Soc Hort Sci 120:300–306

Novy RC, Kobak C, Goffreda J, Vorsa N (1994) RAPDs identify varietal misclassifcation and regional divergence in cranberry [*Vaccinium macrocarpon* (Ait.) Pursh]. Theor Appl Genet 88:1004–1010

Skroch PW, Nienhuis J (1995) Qualitative and quantitative characterization of RAPD variation among snap-bean (*Phaseolus vulgaris*) genotypes. Theor Appl Genet 91:1078–1085

Skroch PW, dos Santos JB, Nienhuis J (1992a) Genetic relationships among *Phaseolus vulgaris* genotypes based on RAPD marker data. Annu Rep Bean Improv Coop 35:23–24

Skroch PW, Tivang J, Nienhuis J (1992b) Analysis of genetic relationships using RAPD marker data. In: Applications of RAPD technology to plant breeding. Crop Science Society of America, Madison, Wisconsin

Steel RGD, Torrie JH (1980) Principles and procedures of statistics: a biometrical approach. McGraw-Hill Publishing Company,

Stiles JI, Lemme C, Sondur S, Morshidi MB, Manshardt R (1993) Using randomly amplified polymorphic DNA for evaluating genetic relationships among papaya cultivars. Theor Appl Genet 85:697–701

Vierling RA, Nguyen HT (1992) Use of RAPD markers to determine the genetic diversity of diploid, wheat genotypes. Theor Appl Genet 84:835–838

Weeden NA, Timmerman GM, Hemmat M, Kneen BE, Lofhi MA (1992) Inheritance and reliability of RAPD markers. In: Applications of RAPD markers to plant breeding. Crop Science Society of America

Welsh J, McClelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res: 18:7213–7218

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535

Yang X, Quiros C (1993) Identification and classification of celery cultivars with RAPD markers. Theor Appl Genet 86:205–212