Theor Appl Genet (2001) 103:1254–1265

J. De Riek · E. Calsyn · I. Everaert · E. Van Bockstaele
M. De Loose

AFLP based alternatives for the assessment of Distinctness, Uniformity and Stability of sugar beet varieties

Received: 26 June 2000 / Accepted: 13 March 2001

Abstract Three approaches for addressing criteria for Distinctness, Uniformity and Stability (DUS) assessment by means of AFLP data are presented. AFLP data were obtained for three consecutive seed deliveries of 15 sugar beet varieties that were under investigation for the official Belgian list ('93, '94 and '95). In total, 696 AFLP markers were scored on 1350 plants. As a first approach, a cluster analysis based on Nei's standard genetic distances between varieties and/or seed deliveries was made. Three major groups put together varieties belonging to corresponding breeding programmes. Statistical procedures, involving bootstrapping and random sampling of subsets of markers, were applied to test the reproducibility of the ordinations and the redundancy present in the data set. In a second approach, the genetic structure inferred by varieties and seed deliveries was submitted to an Analysis of Molecular Variance (AMOVA). Major genetic variation was attributed to individual plant differences within seed deliveries. Differences among seed deliveries seemed to be as important as differences among varieties or breeding programmes. Individual plant data were used for assignment tests. The computation of the assignment was based on the ranking of individual genotypes to one other (based on Jaccard similarity coefficients). The distribution over the accessions for each variety or seed delivery was used to check what group of plants each individual is genetically most similar to. Varieties were classified according to the degree to which the distribution over the different accessions was mainly allocated to their appropriate seed deliveries (from the same variety) or cross-allocated to other varieties. Criteria for DUS-evaluation could be set by each of the approaches; it is discussed in what way the result obtained differs and agrees.

Communicated by F. Salamini

J. De Riek (✉) · E. Calsyn · I. Everaert · E. Van Bockstaele
M. De Loose
Department for Plant Genetics and Breeding – CLO Gent,
Caritasstraat 21, 9090 Melle, Belgium
e-mail: j.deriek@clo.fgov.be
Tel: +31 9 272 28 81, Fax: +31 9 272 29 01

Keywords *Beta vulgaris* L. · Plant variety testing · Genetic variation · Molecular markers · Cultivar identification

Introduction

Sugar beet provides about 40$ of the world's sugar production. Present world production is about 295 million tons of beets from about 9.5 million ha. In Europe sugar beet is grown on 1980000 hectares producing 17 million tons of sugar per year. This yields an average turnover of 5006 million Euro per year for European farmers (Sugar-online 2000; WABCG 2000). Productivity (root yield and sugar yield), quality (sucrose content, juice purity) and disease resistance have an important economic impact on the efficiency of sugar processing and are therefore important breeding characteristics. Besides, sugar beet seed-companies focus on different aspects of crop husbandry involving the reduction of tare by improvement of the shape of the club root, the quality of the seeds (monogerminity, earliness, seed size, pelleting and priming) and weed control. As a result of all these efforts, and because of better crop management, annual production increased to 70 kg of sugar per ha during the last 50 years. At the same time the number of new varieties to be evaluated and registered expanded every year. In Europe, until now, variety testing for sugar beets is organised at the national level and is often accomplished by specialised sugar-beet institutes. The applicants, the farmers, the sugar industry and the governments provide the necessary funding. Because of the lack of UPOV-guidelines (UPOV 1991) for the determination of Distinctness, Uniformity and Stability (DUS) of sugar beet varieties, testing for inscription on national variety lists often mingles DUS-evaluation with the assessment of the Value for Culture and Use (VCU). Variety testing is expensive due to the large trials at different locations needed for reliable VCU evaluation. Although based on regional diversity in agricultural conditions, the establishment of the different evaluation standards, at least for

DUS-assessment, is profitable. This can be partly achieved by adopting general guidelines for evaluation. However, morphology often yields poor descriptors for the discrimination of sugar beets. Genetic characterisation by means of molecular markers can provide a new approach for DUS assessment. Molecular markers can provide a fast and reproducible identification tool, not biased by changing environments and applicable over all stages of seed production, trading, agricultural production and processing. The above properties converge with the demand by the seed companies for better protection of hybrids and inbred lines. So far, different types of molecular markers have been developed and mapped in sugar beet for diverse purposes (Jung et al. 1993; Schondelmaier et al. 1996; Schumacher et al. 1997). Marker-assisted selection procedures mainly pinpoint valuable traits by markers linked to genes and quantitative trait loci (Francis et al. 1998). For the exploration of new genetic resources, genetic distances and variation by means of marker profiles are currently estimated (Kraft et al. 1997; Hansen et al. 1999, McGrath et al. 1999). Both offer new tools for variety description: (1) dissection of the variety traits into a specific combination of trait-linked markers, and (2) overall techniques to compare genotypes and to assess variability within and between groups. In the present study, we aimed to examine different strategies for the evaluation of distinctness, uniformity and stability based on AFLP fingerprints of sugar beet varieties that can be classified under the second group of techniques. Firstly, the marker frequency data per variety and/or seed delivery were used for classification. Secondly, the genetic structure inferred by varieties and seed deliveries was tested by the analysis of molecular variance (AMOVA). Finally, individual plant data were used for assignment tests, searching for the most-related genotype or group of genotypes.

Materials and methods

Plant material

Fifteen sugar beet varieties were included. Seeds of three consecutive seed deliveries (1993, 1994 and 1995) were obtained from the KBIVB-Tienen (Belgium). The same seed lot was used as in the official variety trials (Table 1). Thirty individual plants per variety per seed delivery were analysed (in total 1350 plants). These plants were processed in blocks of 150 (ten individual plants per variety per seed delivery) through sowing, DNA preparation and AFLP analysis. Blocks were randomised over seed deliveries. Plants were always grown for 1 month under a 16-h day (22°C, 80% relative humidity) and a 8-h night (16°C, 80% relative humidity) conditions.

DNA isolation

At harvest, approximately 1 g fresh weight of leaf material was immediately immersed in liquid nitrogen and subsequently lyophilised during 48 h. The dry material was vacuum-packed for storage at –20°C until DNA extraction. Stored material was ground using a Culatti mechanical mill. The DNA isolation protocol was based on the CTAB method of Doyle and Doyle (1987). To 25 mg of lyophilised ground tissue, 1 ml of CTAB extraction buffer (100 mM Tris-HCl pH 8, containing 2% CTAB, 20 mM EDTA, 1.4 M NaCl, 0.5 mM $Na_2S_2O_5$, 0.4% β-mercaptoethanol and 1% PVP MW 40000) and RNase (10 U) was added. Samples were incubated for 40 min at 65°C. Afterwards, samples were homogenised with 1 ml of chloroform/isoamylalcohol (24/1) and centrifuged for 15 min at 10000 g. The supernatant was transferred to a fresh tube and the DNA precipitated with 1 ml of ice-cold (–20°C) isopropanol. After centrifugation (5000 g; 15 min.), the pellet was washed with EtOH (76%) – 0.2 M NaOAc, dried and dissolved in water. DNA concentration and quality was constantly checked compared to a standard series of lambda-DNA on a 1.5% TAE buffered agarose gel after electrophoresis.

AFLP reactions and PAGE

AFLP (Vos et al. 1995) was performed using the commercially available kit from Perkin-Elmer Biosystems for fluorescent fragment detection (Perkin-Elmer 1995). *Eco*RI and *Mse*I were used for DNA digestion. Adapter ligation, pre-selective and selective amplification was performed as in the specified protocols. Selec-

Table 1 List of the varieties tested

Variety	Ploidy	Seed company	Testing in '93, '94, '95[a]	Listed on the Belgian variety list since
Victoria	3n	Kleinwanzlebener Saatzucht (Germany)[b]	C4, C5, C6	1990
Claudia	3n	Kleinwanzlebener Saatzucht (Germany)[b]	C2, C3, C4	1992
Pascal	2n	Hilleshög (Sweden)	C2, C3, C4	1992
Winner	3n	Kuhn (The Netherlands)[c]	R2, C1, C2	1994
Stratos	2n	Ulrich Dieckmann (Germany)[c]	R1, R2, C1	1995
Opus	3n	Ulrich Dieckmann (Germany)[c]	P, R1, R2	1995
Nevada	3n	Agrosem (France)	P, R1, R2	1995
Jackpot	3n	Delitzsch Pflanzenzucht (Germany)[b]	P, R1, R2	1995
Avalon	3n	Delitzsch Pflanzenzucht (Germany)[b]	P, R1, R2	1995
Orion	3n	Van der Have (The Netherlands)[c]	P, R1, R2	1995
Robusta	3n	Van der Have (The Netherlands)[c]	P, R1, R2	Rejected for listing in 1995
Sakara	2n	Hilleshög (Sweden)	P, R1, R2	1995
Gerda	2n	Kleinwanzlebener Saatzucht (Germany)[b]	P, R1, R2	Rejected for listing in 1995
Olivia	3n	Kleinwanzlebener Saatzucht (Germany)[b]	P, R1, R2	1995
Sfinx	3n	SES-Europe (Belgium)	P, R1, R2	1995

[a] P=preliminary testing year; R1=first year of registration trials; R2=second year of registration trials; Cn=nth year of trial for a listed variety
[b, c] Indicating varieties belonging to related breeding programs

tive amplification was done using fluorescent-labelled *Eco*RI-*Mse*I primer combinations with six selective bases. The primer combinations used were *Eco*RI-ACA/*Mse*I-CTG (PC1), *Eco*RI-ACT/*Mse*I-CAT (PC2) and *Eco*RI-AGG/*Mse*I-CTT (PC3). PCR amplifications were performed using a Perkin-Elmer 9600. AFLP fragments were separated by PAGE on an ABI Prism 377 DNA Sequencer on 36-cm gels using 4.25% denaturing polyacrylamide (4.25% acrylamide/bisacrylamide 19/1, 6 M urea in 1× TBE). A GS-500 ROX-labelled size standard (Perkin Elmer) was loaded in each lane in order to facilitate the automatic analysis of the gel and the sizing of the fragments.

Band scoring

During a run on the ABI 377, the fluorescent signal in each lane was recorded continuously. Genescan 2.1 was used to estimate detection time, signal peak height and surface for each fragment. Sizing of the fragments was performed by the Genescan software module by interpolation to the internal lane standard according to the Local Southern algorithm (as recommended by the manufacturer). Only the fragments between 70 bp and 450 bp, and having a signal peak height above the apparatus standard base line setting of 50 (recommended by the manufacturer), were used for scoring. Resolution of the gel system (i.e. the capacity to separate two subsequent bands in one lane) is 1 bp. However, due to lane-to-lane variation and differences in interpolation of the standard, the scoring of the same band position between different lanes varied within 1 bp. After transfer of the Genescan data to Microsoft Access, variations in fragment size (within 1 bp) were assigned to the corresponding categories (i.e. markers) and a scoring table (1/0) was generated. Using Access queries, marker selection thresholds were set towards average signal peak height of a category, and frequency of appearance, in the data set: average signal peak height >120 and frequency >0.15. The first threshold excluded categories that collect fragments that are in general close to the apparatus detection limit in order to avoid non-detection of fragments in AFLP reactions that are somewhat lower in average intensity. The second threshold (f >0.15) was in the first place set to remove artefact bands, both real gel artefacts but also artefacts from imprecise band definition by the Genescan software. For the comparison of individual genotypes De Riek et al. (1999) proved that the positive effect of removing artefacts overrules the information loss of also discarding rare, but real, fragments. Moreover, the best discrimination capacity between single genotypes was obtained when both rare and abundant markers were excluded. In this study, more precaution was taken as groups of plants (varieties, seed deliveries) are compared: (1) all abundant markers were kept in the analysis because the absence of an abundant marker was considered to be more informative than the presence of a rare marker taking into account the inaccuracy level of the automated scoring; and (2) as an exception to the general setting of frequency >0.15 over the total data set, in order not to discard markers that were prevalently present in just one variety or seed delivery; all markers with a frequency >0.5 within one single-seed delivery of whatever variety were kept in the analysis.

Statistical analyses

Polymorphic Information Content (PIC) for dominant markers was calculated as:

$$PIC = 1 - [f^2 + (1 - f)^2],$$

where ‘f’ is the frequency of the marker in the data set. PIC for dominant markers is a maximum of 0.5 for ‘f’=0.5.

Euclidean distances and Jaccard similarities were calculated using the R-package (Legendre and Vaudor 1991). The Jaccard similarity coefficient between two genotypes was calculated by:

$$S_{Jacc} = a/(a + b + c),$$

where ‘a’ is the number of fragments present in both genotypes, ‘b’ the number of fragments only present in genotype 1, and ‘c’ the number of fragments only present in genotype 2. The Euclidean distance between two populations was calculated by:

$$D_{Eucl} = [\Sigma_m[f_{1m} - f_{2m}]^2]^{1/2},$$

where ‘m’ is summed over loci and where ‘f_{1m}’ is the frequency of the present (1) allele at the m-th locus in population 1. The same formula was applied on binary marker data from individual genotypes (allele frequencies being 1 or 0).

Calculation of Nei genetic distance and standard errors, together with the construction of dendrograms (UPGMA) and bootstrapping, were performed by Dispan (Nei 1972, 1978; Ota 1993) and Phylip (Felsenstein 1993). The Nei genetic distance between two populations was calculated by:

$$D_{Nei} = -\ln \frac{\Sigma_m \Sigma_i \, f_{1mi} f_{2mi}}{[\Sigma_m \Sigma_i \, f_{1mi}^2]^{1/2} [\Sigma_m \Sigma_i \, f_{2mi}^2]^{1/2}}$$

where ‘m’ is summed over loci, ‘i’ over alleles at the m-th locus, and where ‘f_{1mi}’ is the frequency of the i-th allele at the m-th locus in population 1.

Two kinds of random re-sampling tests were applied on ordinations: bootstrapping, and random permutation and selection of subsets. Bootstrapping evaluates the reproducibility of the ordination in a statistical way. It involves creating a new data set by sampling characters randomly with replacement, so that the resulting data set has the same size as the original, but some characters have been left out and others are duplicated. In Phylip the amount of replacement is fixed to 35% per bootstrap. Multiple data sets were generated by 100 or 500 re-sampling cycles. The random variation of the results from analysing these bootstrapped data sets can be shown statistically to be typical of the variation from collecting new data sets (Felsenstein 1993). Random permutation and selection of different subsets of markers without replacement was used to test the redundancy in data sets. Multiple data sets were generated in a gradually decreasing series (subsets ranging from 75% to 15% of the initial number of markers) by 100 selection cycles per subset level. Phylip was used to generate permutated data sets; the different subsets were selected from these.

Analysis of molecular variance (AMOVA) was applied (Schneider et al. 2000) on the Euclidean distance matrix between individual genotypes to attribute the distribution of genetic variation to different sources (breeding programs, varieties and seed deliveries). AMOVA is a statistical tool originally developed for population genetics (Excoffier et al. 1992) and based on the scheme of analysis of variance. AMOVA does not allow that genetic structures with more than two levels are tested at once. Therefore, in a double approach, first the allocation of the genetic variation to breeding programmes, and within breeding programmes, to varieties; and second the allocation of the variation to varieties, and within varieties, to seed deliveries was tested. In addition, AMOVA generates a population pair-wise F-statistic (comparable e.g. to the Nei genetic distance calculated from marker frequency data) for which the probability is tested by permutation analysis. Although this approach is computationally more complicated than the use of marker frequency data, it also takes the variation within the group of genotypes into account.

For the assignment tests first a ranking of the 30 most-resembling partners was made per individual plant. For this, a (1350× 1350)-resemblance matrix, using the Jaccard coefficient, was constructed. Pairs of plants with a Jaccard similarity below a bottom threshold of 0.65 were excluded from this ranking. Secondly, it was evaluated in what way, for each plant, the 30 most-resembling partners were distributed over the total set of different seed deliveries and varieties.

Results

Generation of AFLP data

The three AFLP primer combinations used generated approximately 70 to 90 AFLP fragments per reaction. In-

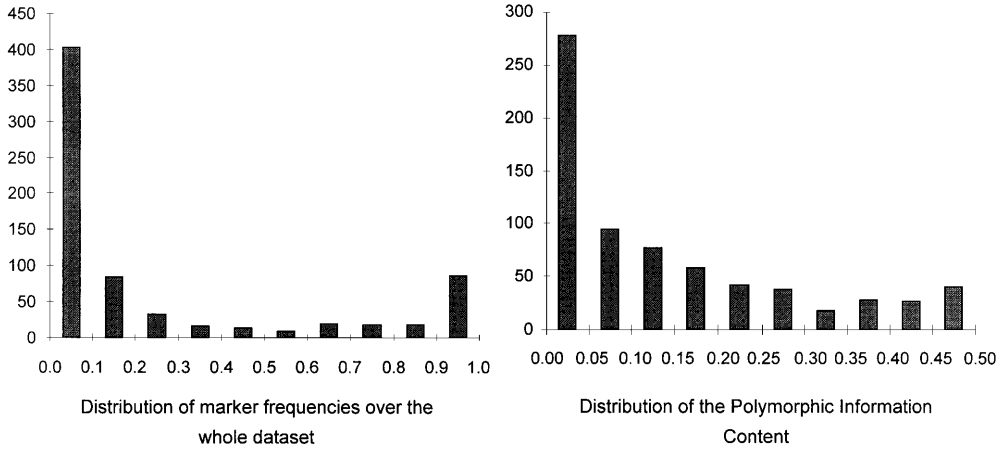

Fig. 1 Distribution of the marker frequency and the Polymorphic Information Content of all scored markers before selection thresholds towards frequency were set

tense bands, indicative for repetitive fragments, were not observed. The number of fragments generated corresponded to what is reported for primer combinations having an A/T content of the selective nucleotides of 3 and 4 for sugar beet (Hansen et al., 1999). After scoring all the AFLP profiles and applying the marker selection thresholds as defined, in total 696 markers were retained (244 for PC1, 268 for PC2 and 184 for PC3). No markers could be identified that were exclusively present or absent in a single variety. The distribution of the frequency of appearance of the markers in the data set and the corresponding Polymorphic Information Content (PIC) are shown in Fig. 1. Although having a low PIC-value indicative for a low discriminatory capacity, the group of nearly monomorphic markers (63 markers with f >0.95) was not excluded from the analysis in order not to overrule the specific absence of a marker in a certain variety.

Classification based on marker-frequency data

Starting from marker-frequency data, taking all plants of a variety or of a certain seed delivery of a variety as a group, Nei standard genetic distances were calculated. Figure 2 shows the relationship between varieties; the distance matrix is given in Table 2. Three major clusters could be distinguished that are supported by high bootstrap values (>450 on 500 multiple data sets): (1) 'Avalon', 'Claudia', 'Victoria' and 'Olivia'; (2) 'Opus', 'Robusta', 'Orion' and 'Winner'; and (3) 'Pascal' and 'Sakara.' The varieties 'Jackpot', 'Nevada', 'Sfinx', 'Stratos' and 'Gerda' did not cluster to a specific group. A good separation was obtained between diploid and triploid varieties (clusters 1 and 2 only hold triploid varieties). The minimal distance that was observed between varieties belonging to different clusters was 0.0065 ('Claudia' versus 'Orion'). In most cases, distances between varieties of different clusters were larger than 0.01. In the first cluster, 'Claudia' and 'Victoria' appeared to be the most-similar varieties (D=0.0026) whereas in the second cluster, 'Opus' and 'Orion' were

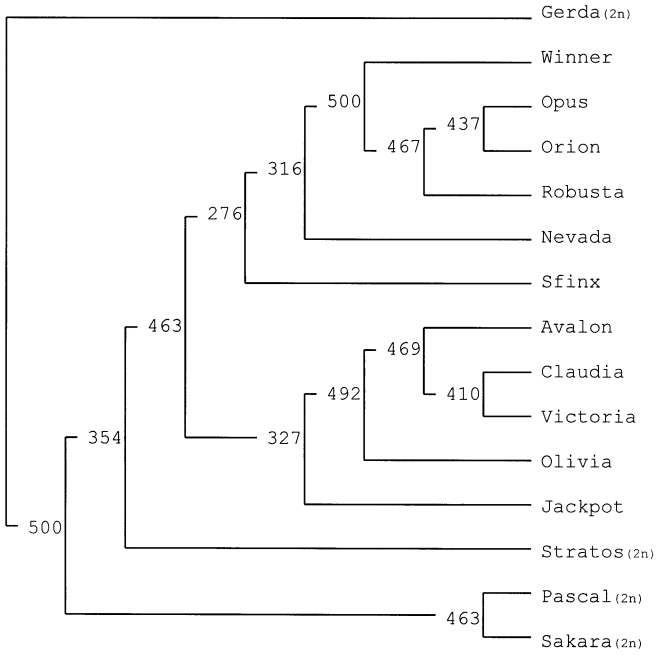

Fig. 2 Ordination of the different varieties (standard Nei genetic distance; UPGMA clustering) with indication of boostrap values of 500 datasets (the branch length does not reflect the actual genetic distance)

the closest ones (D=0.0020). All observed distances were at least 5-times their standard error.

In a second ordination, the different seed deliveries for each variety were grouped (Fig. 3). The matrix of pair-wise distances between seed deliveries is not shown. Distances were at least 3-times their standard error. Reproducibility of the grouping below each node of the dendrogram and the redundancy in the data supporting this grouping were verified by analysing 100 multiple datasets from bootstrapping or random permutation and selection of reduced subsets of markers. Both techniques revealed approximately the same data structure. Nodes with high bootstrap values were already stable using 100 or 200 markers (data not shown).

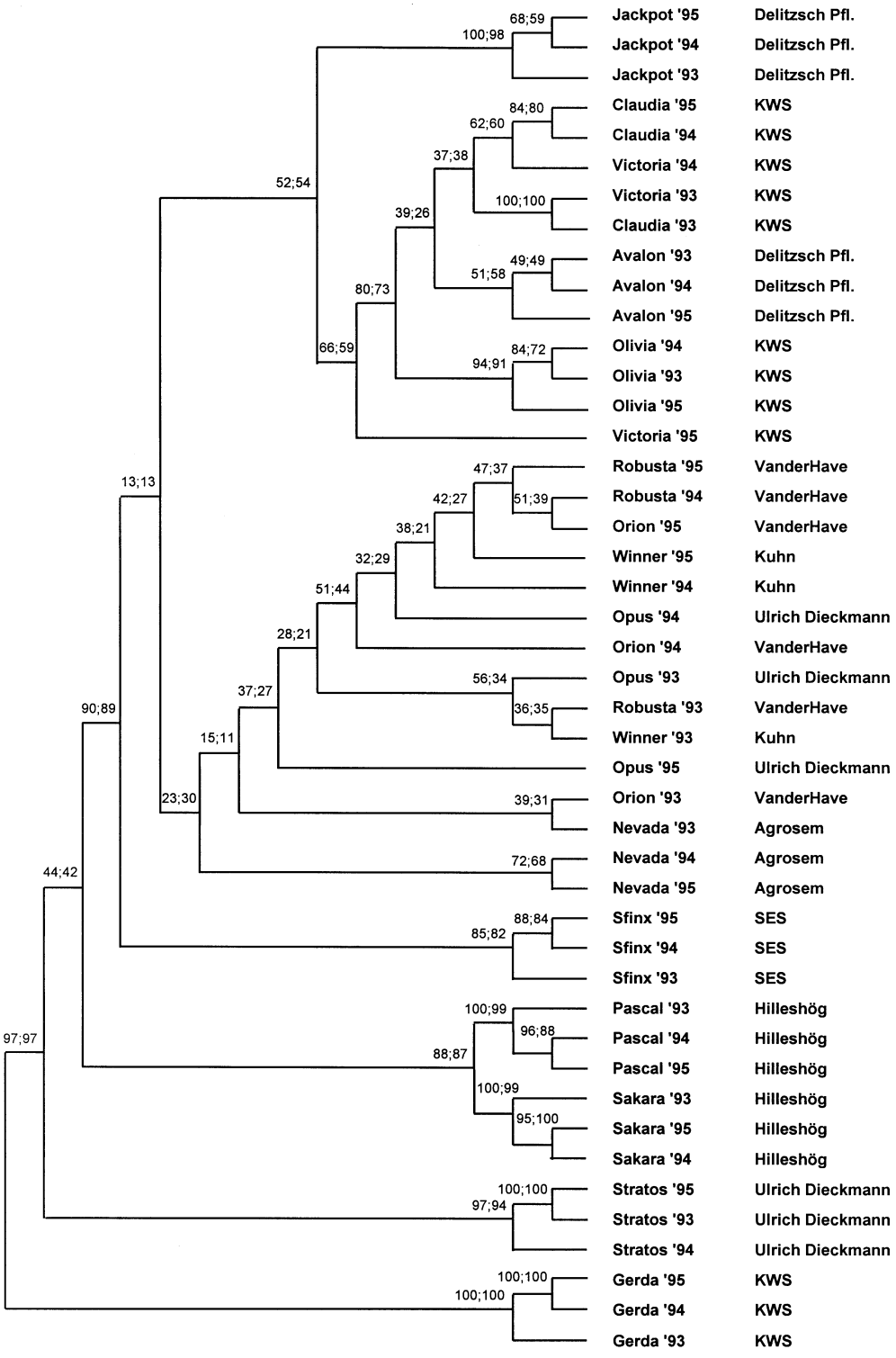

Fig. 3 Ordination of the different seed deliveries (standard Nei genetic distance: UPGMA clustering; the branch length does not reflect the actual genetic distance); the first number on the nodes refers to the boostrap values; the second to random permutation and the selection of 300 markers out of 606 (100 re-sampled data sets)

Within Cluster 1 the three consecutive seed deliveries were grouped together for ‘Jackpot’, ‘Olivia’ and ‘Avalon’; for ‘Jackpot’ (node 15) and for ‘Olivia’ (node 18) quite high bootstrap values (>80 on 100 multiple data sets) were obtained. Within Cluster 1 ‘Claudia’ and ‘Victoria’ were the most-similar varieties. Especially the deliveries of ’93 and ’94 of ‘Claudia’ and ‘Victoria’ were very close (D=0.0031 and 0.0033 respectively). This is in general below the distances obtained between delivery years for each variety (D ranges from 0.0029 to 0.0089). As a result, the resolution obtained with this set of markers was not sufficient to assign the delivery years of each variety to the same cluster. ‘Claudia-’94’ and ‘Claudia-’95’ were reliably grouped (bootstrap value=84 for node 2); ‘Victoria’ was completely dispersed.

Table 2 Standard Nei genetic distance (10^{-2}) between pairs of varieties (lower triangle) – standard errors on Nei distances (upper triangle)

Variety	Avalon	Claudia	Gerda	Jackpot	Nevada	Olivia	Opus	Orion	Pascal	Robusta	Sakara	Sfinx	Stratos	Victoria	Winner
Avalon		0.06	0.26	0.18	0.13	0.13	0.19	0.14	0.21	0.19	0.29	0.20	0.28	0.08	0.17
Claudia	0.45		0.27	0.12	0.12	0.08	0.12	0.09	0.20	0.14	0.27	0.18	0.20	0.04	0.12
Gerda	1.87	1.84		0.34	0.31	0.30	0.35	0.34	0.30	0.36	0.38	0.42	0.30	0.24	0.34
Jackpot	1.09	0.86	2.56		0.13	0.18	0.18	0.17	0.19	0.23	0.27	0.18	0.26	0.13	0.19
Nevada	0.98	0.90	2.12	1.03		0.17	0.10	0.09	0.20	0.13	0.25	0.11	0.21	0.12	0.12
Olivia	0.73	0.58	2.24	1.30	1.27		0.18	0.15	0.25	0.20	0.36	0.22	0.25	0.07	0.20
Opus	1.17	0.79	2.45	1.15	0.72	1.20		0.03	0.19	0.05	0.25	0.18	0.19	0.13	0.05
Orion	0.95	0.65	2.28	1.09	0.68	1.14	0.20		0.21	0.06	0.26	0.17	0.18	0.10	0.05
Pascal	1.47	1.28	2.23	1.26	1.36	1.85	1.46	1.57		0.23	0.18	0.25	0.20	0.17	0.23
Robusta	1.34	0.98	2.36	1.51	1.04	1.41	0.33	0.36	1.76		0.26	0.20	0.20	0.13	0.05
Sakara	1.98	1.80	2.96	1.79	1.89	2.62	1.91	1.94	1.11	1.97		0.29	0.30	0.26	0.25
Sfinx	1.30	1.23	2.88	1.15	0.85	1.31	1.05	1.01	1.61	1.33	1.93		0.28	0.17	0.20
Stratos	1.91	1.39	2.33	1.86	1.56	1.82	1.13	1.23	1.49	1.33	2.07	1.78		0.18	0.20
Victoria	0.54	0.26	1.80	0.88	0.88	0.62	0.83	0.69	1.20	0.97	1.71	1.12	1.10		0.13
Winner	1.22	0.88	2.32	1.16	0.73	1.36	0.30	0.31	1.63	0.32	1.97	1.14	1.29	0.83	

Table 3 Stability over seed deliveries expressed as the average standard Nei distance between seed deliveries

Variety	Av. st. Nei distance between seed deliveries
Jackpot	0.00370
Avalon	0.00483
Olivia	0.00497
Claudia	0.00527
Sakara (2n)	0.00600
Pascal (2n)	0.00637
Robusta	0.00660
Sfinx	0.00683
Winner	0.00737
Nevada	0.00823
Victoria	0.00830
Gerda (2n)	0.00860
Orion	0.00960
Opus	0.00963
Stratos (2n)	0.01443

Different seed deliveries appeared to be much less structured in the second cluster. In fact, low bootstrap values (around 50 or lower) were obtained here indicating a low reliability of structure. 'Nevada', which was grouped apart from the second cluster using the overall marker frequency data for the three seed deliveries, was tailed now to the second cluster (Fig. 3). The analysis of a set of 20 extra samples for the varieties belonging to this cluster, bringing the total number of analysed plants per seed delivery to 50 (data not shown), enabled the separation for 'Nevada' from the remainder. It also increased bootstrap values. However, the tailing of the seed deliveries of 'Opus', 'Orion', 'Robusta' and 'Winner' could not be resolved by increasing the numbers of plants analysed.

Varieties belonging to the third cluster and the remainder showed a well-structured grouping between seed deliveries. This was characterised by high bootstrap values; stable groupings could often be based on a random selection of just 100 markers.

The average standard Nei distance between seed deliveries from the same variety is shown in Table 3. Varieties were put in ascending order of average standard Nei distance, showing on top of the list the varieties with more-similar seed deliveries. Average distances between seed deliveries from the same variety ranged from 0.0037 to 0.0144 and, in some cases, pair-wise distances between varieties grouped to the same cluster (Table 2) fell within this range. This confirms what was observed in Fig. 3 for clusters 1 and 2: no clear delineation of the delivery years within varieties is evident.

Analysis of molecular variance

The AMOVA procedure provides a general framework for the analysis of population genetic structure based on any distance matrix (Euclidean distances were used) between individual plant genotypes. Two genetic structure designs were applied (Table 4): (1) allocation of the variation to breeding programmes, and within breeding programmes, to varieties, and (2) allocation of the variation to varieties, and within varieties, to seed deliveries. The population pair-wise *F*-statistics matrix (data not shown) revealed an identical data structure as the use of the standard Nei distance; all Fst had a probability that was significant at 0.05 (100 permutations). When testing both genetic structure designs (Table 4), care must be taken in the interpretation of the results because both designs cannot be nested. Different breeding programmes did account for 2.6% of the total variation whereas differences among varieties within breeding programmes accounted for 5.5% (first design). Major variation remained attributed to the variation within cultivars. Under the second design, where varieties represented the highest level, only 4.5% of the total variation could be attributed to varieties; seed deliveries within varieties accounted for 9.5% of the total variation. For this design, the variation within seed deliveries was detailed by their sum of squares (Table 5). These values can be used as an estimate for the Uniformity of the seed delivery. As a general conclusion of both AMOVA designs, it can be stated that

Table 4 AMOVA designs and results

(1) Among breeding programmes, among varieties within breeding programmes and among individuals within varieties

Source of variation	Degrees of freedom	Sum of squares	Variance components	Percentage of variation
Among breeding programmes	4	3315.007	1.81060**	2.63
Among varieties within breeding programmes	10	4033.184	3.81031**	5.54
Within varieties	1334	83640.191	63.17235**	91.83
Total	1348	90988.382	68.34451	

(2) Among varieties, among seed deliveries within varieties and among individuals within seed deliveries

Source of variation	Degrees of freedom	Sum of squares	Variance components	Percentage of variation
Among varieties	14	7348.191	3.06447**	4.48
Among seed deliveries within varieties	30	7537.654	6.46819**	9.46
Within seed deliveries	1304	76102.537	58.81185**	86.05
Total	1348	90988.382	68.34451	

** Significant at 0.01 levels, evaluated by 1000 permutations

Table 5 Uniformity expressed as the sum of squares (AMOVA) per seed delivery

Variety	Sum of squares				Ranking			
	Total	Year 1	Year 2	Year 3	**Total**	Year 1	Year 2	Year 3
Olivia	**4344**	1173	1379	1792	**1**	1	1	7
Claudia	**4537**	1527	1625	1385	**2**	8	7	1
Jackpot	**4835**	1305	1737	1793	**3**	3	9	8
Stratos (2n)	**4845**	1235	2222	1388	**4**	2	15	2
Robusta	**4875**	1514	1438	1923	**5**	6	2	13
Orion	**4951**	1825	1521	1605	**6**	13	4	3
Victoria	**5053**	1526	1793	1734	**7**	7	10	4
Opus	**5073**	1491	1604	1978	**8**	5	6	15
Winner	**5135**	1769	1568	1798	**9**	10	5	9
Gerda (2n)	**5140**	1852	1514	1774	**10**	14	3	6
Sakara (2n)	**5248**	1377	1938	1933	**11**	4	13	14
Avalon	**5260**	1643	1861	1756	**12**	9	11	5
Sfinx	**5426**	1786	1723	1917	**13**	11	8	12
Pascal (2n)	**5527**	1802	1873	1852	**14**	12	12	11
Nevada	**5856**	2011	2003	1842	**15**	15	14	10

Table 6 Uniformity expressed as the average Jaccard similarity between plants of the same seed delivery

Variety	Average Jaccard similarity				Ranking			
	Total	Year 1	Year 2	Year 3	**Total**	Year 1	Year 2	Year 3
Olivia	**0.6568**	0.6838	0.6722	0.6143	**1**	1	1	6
Jackpot	**0.6406**	0.6730	0.6282	0.6206	**2**	2	7	5
Claudia	**0.6379**	0.6223	0.6281	0.6632	**3**	8	8	1
Robusta	**0.6269**	0.6273	0.6550	0.5984	**4**	5	2	10
Winner	**0.6183**	0.5976	0.6517	0.6057	**5**	11	3	8
Opus	**0.6171**	0.6253	0.6299	0.5960	**6**	6	6	12
Avalon	**0.6163**	0.6153	0.6064	0.6272	**7**	9	10	4
Orion	**0.6152**	0.5737	0.6430	0.6288	**8**	12	4	3
Victoria	**0.6132**	0.6224	0.6031	0.6140	**9**	7	11	7
Stratos	**0.6126**	0.6575	0.5349	0.6455	**10**	3	15	2
Sfinx	**0.6079**	0.5991	0.6270	0.5975	**11**	10	9	11
Sakara	**0.6039**	0.6425	0.5782	0.5909	**12**	4	14	15
Gerda	**0.5949**	0.5520	0.6379	0.5948	**13**	15	5	13
Nevada	**0.5866**	0.5616	0.5945	0.6037	**14**	14	12	9
Pascal	**0.5841**	0.5730	0.5871	0.5922	**15**	13	13	14

Table 7 Stability expressed in function of the average, the minimum, the maximum and the range of the Jaccard similarities between plants of different seed deliveries

Variety	Jaccard similarity between different seed deliveries				Ranking			
	Average	Minimum	Maximum	Range	Average	Minimum	Maximum	Range
Olivia	**0.6435**	0.6312	0.6652	0.0339	**1**	1	1	12
Jackpot	**0.6305**	0.6186	0.6411	0.0225	**2**	2	2	6
Claudia	**0.6242**	0.6092	0.6374	0.0282	**3**	3	3	7
Robusta	**0.6102**	0.5935	0.6223	0.0287	**4**	5	5	8
Aavalon	**0.6056**	0.6037	0.6088	0.0051	**5**	4	9	1
Winner	**0.6005**	0.5837	0.6128	0.0291	**6**	7	7	10
Sfinx	**0.5951**	0.5825	0.6030	0.0205	**7**	8	11	5
Opus	**0.5944**	0.5821	0.6113	0.0292	**8**	9	8	11
Victoria	**0.5926**	0.5891	0.5953	0.0063	**9**	6	12	2
Orion	**0.5924**	0.5781	0.6206	0.0425	**10**	11	6	13
Sakara	**0.5881**	0.5799	0.5944	0.0146	**11**	10	13	3
Gerda	**0.5744**	0.5522	0.6052	0.0530	**12**	14	10	14
Stratos	**0.5726**	0.5464	0.6234	0.0770	**13**	15	4	15
Pascal	**0.5706**	0.5645	0.5797	0.0153	**14**	12	15	4
Nevada	**0.5697**	0.5573	0.5863	0.0289	**15**	13	14	9

most of the variation is attributed to variation between individual plants within seed deliveries of cultivars; differences among seed deliveries seem to be as important as differences among varieties or breeding programmes.

Tests based on pair wise comparison of individual plants

Estimates for the internal genetic variation for varieties or seed deliveries can also directly be assessed from the pair-wise resemblance data for individual plants (Jaccard similarity coefficients were used). The average over all pair-wise Jaccard coefficients of plants belonging to identical seed delivery lots of a certain variety (delivery 1–1, 2–2 and 3–3 relationships) was calculated. A ranking was made per year and as an average over the 3 testing years (Table 6). The average similarity over all pair wise Jaccard coefficients of plants belonging to different seed delivery lots of a certain variety (delivery 1–2, 1–3 and 2–3 relationships were grouped) is shown in Table 7.

The distribution for the assignment of individual genotypes to a certain variety is given in Table 8. In the bottom panel the global assignment of the 30 most-similar partners (this number equals the number of individuals tested per seed delivery) is given. Table 8 must be read horizontally: e.g. for all individual plants analysed from ‘Opus’, the 30 most-similar partners were 649-times tracked back to ‘Opus’ itself, 404-times to ‘Robusta’, 289-times to ‘Orion’, 256-times to ‘Winner’ and so on. The upper panels of Table 8 show that the distribution of the assignment over the varieties depends on the size of the sample of the resembling partners taken into account for the test. Compared to the 30 most-similar plants, the number of top 3 or top 10 allocations differed. In general, when reducing the assignment to only the high-ranking plants, varieties tend to be less dispersed. Here, a remark must be formulated concerning the way the genotyping over the different seed deliveries was performed. As analyses were always grouped in blocks of ten plants belonging to the same seed delivery, a reduction of the allocation to only the top three might cause deviations, as plants are in some cases preferably assigned to individuals of the same block.

The assignment tests were further detailed up to the level of seed deliveries (data not shown but presented in the text). From these tests, ‘Sfinx’ appeared to be an example of a variety that is well distinguishable from the others, although it seems to refer to a common genetic pool as e.g. ‘Avalon’, ‘Claudia’, ‘Jackpot’ or ‘Olivia.’ There was a preferential allocation to the same seed delivery in the detailed assignment tests, but plants were also equally assigned to the other deliveries of the same variety. The assignment for the deliveries of ‘Sfinx’ to the other varieties in the data set was generally low and dispersed among deliveries from the same target variety. On the other hand, ‘Opus’ can be taken as an example of a variety that cross-attributed to several varieties. Although a good year-to-year assignment for the seed deliveries is retained for ‘Opus’, a structured assignment to ‘Robusta’ was also observed (a consistent allocation among all deliveries of ‘Robusta’). A similar allocation, but to a lesser extent, was observed to ‘Orion’ and ‘Winner.’ In the same way, based on Table 8 and their appropriate seed delivery allocations (data not shown), the other varieties were also classified according to their degree of cross-allocation. The assignment patterns for ‘Gerda’, ‘Jackpot’, ‘Pascal’, ‘Sakara’, ‘Sfinx’ and ‘Stratos’ appeared to be very typical; they were rather characteristic for ‘Olivia’ and ‘Robusta’; for ‘Avalon’, ‘Claudia’, ‘Nevada’, ‘Winner’, ‘Opus’, ‘Orion’ and ‘Victoria’ patterns were most cross-attributing. ‘Victoria’, which was the oldest variety in trial, was the most-dispersed variety.

Table 8 Assignment tests per variety (upper panel: three most-similar plants; middle: ten most-similar plants; lower panel: 30 most-similar plants)

	Variety	Avalon	Claudia	Gerda	Jackpot	Nevada	Olivia	Opus	Orion	Pascal	Robusta	Sakara	Sfinx	Stratos	Victoria	Winner
3 most-similar plants	Avalon	**139**	28		18	2	53	2	1		1	1	2	1	2	1
	Claudia	39	**126**	1	19	2	52	3			4		1	4	11	1
	Gerda			**231**			3			1		1			1	
	Jackpot	6	7		**214**	2	9	3	3		2	1	4		3	5
	Nevada	11	11	2	16	**107**	12	3	13		11		29	1	5	9
	Olivia	22	13		5		**215**	1	1		2		1		4	1
	Opus	2	10		9	4	17	**111**	23		37		10	2	4	20
	Orion	12	20		12	9	14	19	**88**		42	2	5	1	1	32
	Pascal	5	7		16		4	2	1	**153**	1	32	6	1	2	
	Robusta	6	4		4	3	14	22	22		**121**		5	2	3	35
	Sakara		1	1		1				26		**182**	2			2
	Sfinx	6	7		11	8	12	1			2		**196**		3	1
	Stratos	3	5		3	1	5	7	1	3	4			**185**	4	4
	Victoria	23	49		29	4	59	6	4	1	4		5	1	**81**	6
	Winner	3	11		16	5	9	18	20	1	33		7		4	**127**
10 most-similar plants	Avalon	**357**	116	1	58	6	192	2	8		11	2	16	3	16	9
	Claudia	137	**309**	1	76	14	229	11	22	2	13	1	17	5	35	12
	Gerda	12	5	**658**		4	19		2	1	1	1			3	
	Jackpot	25	52		**636**	11	54	6	9	2	3	3	24		18	12
	Nevada	33	59	2	64	**230**	48	37	34	1	46	3	91	1	17	43
	Olivia	60	71	1	21	6	**662**	8	6		6		12		18	3
	Opus	9	35		47	25	56	**290**	79		142		48	3	16	74
	Orion	43	55	1	50	38	55	77	**213**		149	2	33	2	18	103
	Pascal	20	31		78	8	32	15	4	**400**	5	94	24	8	12	12
	Robusta	14	28		13	26	48	82	86		**379**		24	5	17	118
	Sakara	1	10	1	13	3	1		6	71	5	**514**	10	2		3
	Sfinx	23	40		39	26	61	28	8	2	13	1	**532**		8	12
	Stratos	18	29	1	15	6	33	32	12	7	14	1	16	**490**	18	24
	Victoria	96	176		114	13	184	24	15	4	23	1	25	7	**202**	27
	Winner	20	56		47	21	38	87	80	5	150	1	32	5	16	**312**
30 most-similar plants	Avalon	**788**	348	14	199	33	567	35	59	1	47	7	60	5	73	38
	Claudia	404	**776**	2	339	53	678	73	86	10	81	8	70	15	143	77
	Gerda	132	47	**1327**	28	6	133	6	8	2	27	3	1	7	47	9
	Jackpot	113	264	2	**1548**	48	247	58	57	18	29	17	103	3	73	59
	Nevada	121	209	3	203	**472**	202	152	149	11	150	9	252	6	49	130
	Olivia	256	319	5	163	23	**1644**	40	34	2	39		82	7	88	33
	Opus	69	178		175	102	221	**649**	289	5	404	4	164	16	59	256
	Orion	158	221	2	187	111	257	292	**509**	4	399	8	120	13	74	328
	Pascal	72	130	3	282	44	120	51	30	**781**	42	269	81	19	37	54
	Robusta	78	150	4	92	91	198	364	299	3	**869**	7	109	30	50	412
	Sakara	29	52	2	99	25	29	23	22	164	30	**1099**	79	17	18	15
	Sfinx	114	186		184	105	309	127	65	10	89	11	**1058**	2	36	68
	Stratos	77	120	4	67	39	198	135	53	15	101	18	82	**899**	84	82
	Victoria	291	471	9	364	54	564	104	79	14	100	9	106	31	**400**	120
	Winner	86	208	1	206	103	182	304	277	10	461	9	119	21	73	**668**

Discussion

Value for DUS-assessment

Standard procedures for DUS-assessment, as laid down in UPOV-guidelines (UPOV 1991) are based on a single trait by trait evaluation. For measured phenotypic characteristics, relative tolerance limits such as standard deviation or variance are commonly used as criteria for the comparison with known varieties, and validated statistical testing procedures are available. The statistical procedure of choice is a *t*-test where the average difference between a candidate and an established variety is compared with the standard error for the difference. Therefore, DUS-testing is typically a univariate problem. However, procedures to combine information of a set of marker data as provided by phylogenetic and population genetic research are typically multivariate. There is quite a consensus that uncharacterised molecular-marker data are not to be handled in a single trait by trait evaluation. There are two main reasons for this. Firstly, the information content provided by a single uncharacterised marker should be considered as unknown and probably low. A link to crop performance is most often sought-after. For that purpose, markers must have been mapped and linkage been demonstrated between a set of markers and particular traits. Secondly, due to the current UPOV-convention, the same characteristics used for Distinctness evaluation should be used for Uniformity and Stability testing. However, it can be questioned where the benefits for the breeder and for the end-user are, except to solve distinction problems, to impose a continuous selection and control on constant marker profiles. This might be achievable in self-pollinating crops or hybrids, based on pure lines, but is hard to fulfil in crossbreeding crops when the crops show a considerable within-variety genetic heterogeneity. Without wishing to enter the discussion on what is feasible on a crop by crop basis, integration of molecular marker data in DUS-assessment will have to take different concepts than the current UPOV-methodology. Nevertheless, for the ease of interpretation, most proposals to integrate marker data in DUS-testing opt to transform marker data to a univariate general measure for genetic distance or similarity (UPOV-BMT 1997, 1998). However, for genetic distances, no statistical testing procedures are unequivocally accepted as standard yet: e.g. their statistical properties are complicated, one reason only being the non-independence of the markers within a set (Dillman et al. 1997; Foulley and Hill 1999).

In the study presented we also opted to make use of genetic distance estimators. They were applied at different levels: from an overall analysis (classifications based on marker frequency data) to a plant by plant comparison (assignment tests), but all within the option not to broaden the trait by trait evaluation to single marker data.

Distinctness evaluation

Distinctness testing could be performed by the three procedures described. The calculation of genetic distances (standard Nei distance from marker frequency data or Fst based on Euclidean distances between single plants) is very powerful to reveal differences between accessions. All pair-wise standard Nei distances calculated were at least 3-times larger than the standard error on their value both on the level of varieties and of seed deliveries. However, in some cases, differences between varieties were sometimes smaller than differences between seed deliveries. Therefore, the significance testing as provided by the standard errors on Nei distance is probably too sensitive.

It was supposed that the estimation of Fsts and permutation testing, taking into account the inside variation (AMOVA), should better deal with the major variation within varieties or seed deliveries that is not accounted for when using marker-frequency data. Surprisingly, the permutation testing in AMOVA also only yielded highly significant Fst-values. Permutation testing in AMOVA concerns the randomisation of individuals in order to test if both sub-populations are drawn from an identical group. The input for AMOVA is a matrix with pair-wise distances between individual genotypes. The method does not account for any uncertainty on the pair-wise distances itself, e.g. due to sampling of a limited number of plants per accession or the choice of a specific subset of markers the distance was calculated on.

Bootstrapping, random permutation and the selection of different subsets of markers appeared to be more appropriate techniques to handle this kind of ambiguity. Both techniques gave very similar results on the stability of the obtained groupings. Bootstrapping is computationally much more feasible and actually reflected well the results from the more extended random permutation and selection procedures. However, both techniques are biased by the composition of the set of varieties in the analysis. For example, seed deliveries of a large set of more closely related varieties are apt to mix up more easily when creating multiple datasets by re-sampling strategies, resulting in a lower stability of nodes at a smaller genetic distance. An ideal testing procedure should therefore include the randomisation of individuals between groups and account for uncertainty on the pair-wise distances.

When starting from marker frequencies or in AMOVA, the ordering of genotypes in a certain genetic structure (breeder, variety, and seed delivery) is taken for granted. Such approaches can only give an indirect indication, e.g., for mixing of seed lots of related varieties, or use of different pollinator or mother lines by a small shift in resemblance between varieties or an increased variation within the variety. Because of the high variation observed within varieties and deliveries, one might consider the calculation of a global distance between accessions as not satisfactory. To overcome this, assignment tests were performed (Law et al. 1999). For Distinctness

evaluation assignment tests appeared to reflect well the results obtained by the previous methods. The assignment tests, however, did much better reveal the underlying data structure in order to look for the most-similar variety. For example, 'Opus' and 'Orion' always appeared to be the closest couple in the ordination based on marker frequencies, this is not directly confirmed by the assignment tests. 'Robusta' was the best "target" variety after 'Opus' itself. For 'Orion', 'Robusta' is the best second target far before 'Opus.' The computation of the assignment, based on the ranking of individual genotypes to one other, is straightforward, although demanding when large numbers of plants are to be analysed. While thresholds for distinction can be defined in a clear way, e.g. 70% of the most-similar genotypes must belong to the same variety, the interpretation of data can also be biased. If Distinction criteria are directly to be set from such assignment tests, one has to carefully consider the set of varieties to be used as a reference framework. New varieties originating from related breeding pools are likely to show a more-dispersed assignment than a product from a totally new genetic background. Nevertheless, the intrinsic simplicity of the concept of assignment tests that has much in common with, e.g., the screening for off-types in self-pollinated crops which probably fits best to the current testing procedures.

Uniformity and stability testing

Uniformity and stability testing based on AFLP data appeared to be much more troublesome. A first indication for Uniformity and Stability of the tested varieties can be obtained from the bootstrapping and random re-sampling experiment (Fig. 3). Here, the distinction between seed deliveries is evaluated. Some varieties tend to have very typical seed deliveries, other are more diffused. The pattern observed for the better-clustered seed deliveries in Fig. 3 corresponds well to the order obtained by checking the degree of cross-allocation in the assignment test for the seed deliveries. However, Uniformity and Stability testing, if accessed directly from the assignment by looking for deviations in the allocation of seed deliveries from the same variety, appeared to be little discriminative. Although based on the same observations as for Distinctness, the assessment of Uniformity and Stability is essentially concentrated on variability within a variety and reproducibility of this variability. Their assessment should be directly derived from data on the seed deliveries of each variety, independent from distinctness estimation. This can most straightforwardly be accomplished by averaging all pair-wise similarities between single plants within or between specific seed deliveries (Tables 6 and 7).

In general, the average Jaccard similarity within and between seed deliveries provided a high correlation between the rankings for Uniformity and Stability. This can be intrinsic to well-established varieties. However, the correlation observed between both parameters might also indicate that an element, that is inherent to a specific variety, is interfering with the genetic origin of the material or the composition of the components that made the hybrid. Moreover, comparison with alternative approaches often yielded conflicting conclusions.

After AMOVA, uniformity has been expressed by the means of the specific Sum of Squares per seed delivery (Table 5). Compared to this, in Table 6 some local rearrangements are observed but the general ranking over the 3 delivery years is more or less maintained. For quite some varieties considerable shifts in the ranking between testing years are observed. Both in Tables 5 and 6 a general appreciation of uniformity over the 3 testing years is given by calculating the average score and the average ranking over the 3 years. However, by doing so, all agreement between calculations is lost. Without having an appropriate reference framework, it is tentative to base general conclusions on the present rankings. Both diploid and triploid varieties have been combined in the analysis. The use of a dominant marker system not capable of discerning between 1, 2 or 3 copies of the same allele will force triploid varieties to be more equal. This can already be observed here. Moreover, variety concepts might differ a lot (two-way versus three-way hybrids, number of components in the hybrids) which all can yield agronomically valuable varieties.

Procedures for stability testing were, even more, not in agreement. The average standard Nei distance between seed deliveries (Table 3) has to be compared to the average of all pair-wise similarities between single plants between specific seed deliveries (Table 7). Although 'Olivia', 'Jackpot' 'Claudia' and 'Avalon' are both among the five highest-ranking varieties, the rearrangements in the middle and bottom parts of the tables make the total ranking unmatched.

Although the parameters for Stability and Uniformity were both directly derived from the same data input as for Distinctness testing, and their definition is clear (better than or below the level of the reference set), non-consistent results were observed when comparing the different approaches used. Opposite to the granting of Distinctness, that rewards the creation of a new breed in comparison to all the existing ones, Uniformity and Stability are much more labels of guarantee for agricultural application. Therefore, they are much closer to VCU evaluation. Moreover, field-testing for Uniformity and Stability is also statistically well defined: pro of that there are no differences within or between subsequent seed deliveries. Re-seeding spare seed from previous years in adjacent plots to the current ones most easily does this. The most-direct benefit of molecular markers in DUS-testing currently seems to be for Distinction purposes. A first application that can be accomplished at relatively low cost is its use for the pre-screening of new applicant varieties and the grouping of similar varieties in appropriate field trials.

Acknowledgements The authors thank the KBIVB-IRBAB for their excellent co-operation and Veerle Buysens, Laurence Desmet and Veerle Cools for their efforts in DNA preparations and AFLP reactions. The staff of the biotech laboratory is very much appreciated for their useful support and skilful assistance. We also want to acknowledge all participants to the UPOV-BMT meetings for their contribution in the discussions for the subject. The Belgian Ministry of Small Enterprises, Traders and Agriculture –DG4 and DG6 funded this research. The experiments comply with current Belgian laws.

References

De Riek J, Dendauw J, Mertens M, Deloose M, Heursel J, Van Bockstaele E (1999) Validation of criteria for the selection of AFLP markers to assess the genetic variation of a breeders' collection of evergreen azaleas. Theor Appl Genet 99:1155–1165

Dillman C, Charcosset A, Goffinet B, Smith JSC, Dattée Y (1997) Best linear unbiased estimator of the molecular genetic distance between inbred lines. Advances in biometrical genetics. In: Posnan, Krajewski P, Kaczmarek (eds) Proc 10th meeting of EUCARPIA Section Biometrics in Plant Breeding, pp 105–110

Doyle J, Doyle JL (1987) Genomic plant DNA preparation from fresh tissue – the CTAB method. Phytochem Bull 19:11

Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Felsenstein J (1993) PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author, Department of Genetics, University of Washington, Seattle, Washington

Foulley JL, Hill WG (1999) On the precision of estimation of genetic distance. Genet Selec Evol 31:457–464

Francis SA, Redfern M, Chwarszczynska DM, Asher MJC, Lewellen RT (1998) Use of molecular markers in breeding for disease resistance in sugar beet (*Beta vulgaris* L.). Aspects Appl Biol 52:279–285

Hansen M, Kraft T, Christiansson M, Nilsson NO (1999) Evaluation of AFLP in *Beta*. Theor Appl Genet 98:845–852

Jung C, Pillen K, Frese L, Fähr S, Melchinger AE (1993) Phylogenetic relationships between cultivated and wild species of the genus *Beta* revealed by DNA "fingerprinting." Theor Appl Genet 86:449–457

Kraft T, Fridlund B, Hjerdin A, Säll T, Tuvesson S, Halldén C (1997) Estimating genetic variation in sugar beets and wild beets using pools of individuals. Genome 40:527–533

Law JR, Cooke RJ, Reeves JC, Donini P, Smith JSC (1999) Most similar variety comparisons as a grouping tool. Plant Var Seeds (in press)

Legendre P, Vaudor A (1991) The R package: multidimensional analysis, spatial analysis. Département de sciences biologiques, Université de Montréal

McGrath JM, Derrico CA, Yu Y (1999) Genetic diversity in selected, historical US sugar beet germplasm and *Beta vulgaris ssp. maritima*. Theor Appl Genet 98:968–976

Nei M (1972) Genetic distances between populations. Am Nat 106:283–292

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590.

Ota T (1993) DISPAN: genetic distance and phylogenetic analysis. Institute of Molecular Evolution and Genetics, Pennsylvania State University, University Park, Pennsylvania

Perkin-Elmer (1995) AFLP plant mapping kit: protocol

Schneider S, Roessli D, Excoffier L (2000) Arlequin ver. 2.000: A software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland

Schondelmaier J, Steinrücken G, Jung C (1996) Integration of AFLP markers into a linkage map of sugar beet (*Beta vulgaris* L.). Plant Breed 115:230–237

Schumacher K, Schondelmaier J, Barzen E, Steinrücken G, Borchardt D, Weber WE, Jung C, Salamini F (1997) Combining different linkage maps in sugar beet (*Beta vulgaris* L.) to make one map. Plant Breed 116:23–38

Sugaronline (2000) Statistics, analysis and reports: production/consumption, prices, trends ..., http://www.sugaronline.com/

UPOV (1991) International convention for the protection of new varieties of plants, Geneva, Publication no 221(E)

UPOV-BMT (1997) BMT/4/21 Report of the 4th session of the Working group on biochemical and molecular techniques and DNA-profiling in particular, Cambridge, UK

UPOV-BMT (1998) BMT/5/17 Report of the 5th session of the Working group on biochemical and molecular techniques and DNA-profiling in particular, Beltsville, USA

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Fijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 21:4407–4414

WABCG (2000) World association of beet and cane growers, http://www.ifap.org/wabcg/