Genet Resour Crop Evol (2007) 54:1437–1446
DOI 10.1007/s10722-006-9128-y

RESEARCH ARTICLE

AFLP fingerprinting of sesame (*Sesamum indicum* L.) cultivars: identification, genetic relationship and comparison of AFLP informativeness parameters

Hernán Laurentin · Petr Karlovsky

Received: 3 April 2006 / Accepted: 28 September 2006 / Published online: 16 February 2007

Abstract Amplified fragments length polymorphism (AFLP) was used to distinguish 20 cultivars of sesame (*Sesamum indicum* L.) and to elucidate the genetic relationship among these genotypes. The data were also used to estimate the usefulness of parameters currently used to assess the informativeness of molecular markers. A total of 339 markers were obtained using 8 primer combinations. Of the bands, 91% were polymorphic. Five primer combinations were able to distinguish all 20 cultivars used. None of the remaining three primer combinations could distinguish all accessions if used alone, but using all three combinations reduced the probability of a random match to 5×10^{-5}. Polymorphic information content (PIC), resolving power (*Rp*) and marker index (MI) of each primer combination failed to correlate significantly with the number of genotypes resolved. Jaccard's similarity coefficients ranged from 0.31 to 0.78. Fifteen cultivars were grouped by four UPGMA-clusters supported by bootstrapping values larger than 0.70. The grouping pattern was similar to the grouping generated by principal coordinate analysis. The results demonstrated that AFLP-based fingerprints can be used to identify unequivocally sesame genotypes, which is needed for cultivar identification and for the assessment of the genetic variability of breeding stocks. We recommend to use the number of cultivars identified by a primer combination instead of PIC, *Rp* and MI; and to calculate the maximal, instead of average probability of identical match by chance in the assessment of the informativeness of a marker for cultivar identification.

Keywords AFLP · DNA fingerprinting · Genotype identification · Genotyping · *Sesamum indicum*

H. Laurentin (✉)
Department of Biologic Sciences, Agronomy Faculty, Universidad Centroccidental Lisandro Alvarado, vía Agua Viva, Cabudare, estado Lara, Venezuela
e-mail: hlaurentin@ucla.edu.ve

P. Karlovsky
Department of Crop Sciences, Georg-August University, Grisebachstrasse 6, 37077 Goettingen, Germany

Introduction

Sesame (*Sesamum indicum* L.) is an important crop in tropical and subtropical areas (Ashri 1998). Over six millions hectares were harvested worldwide in 2004, producing over three million tons of seeds (FAO 2005). India, Sudan, Myanmar and China are the most important sesame producers, with 68% of world production. The production in America is 170,000 tons per year; Mexico, Guatemala and Venezuela contribute 60% to the production on the continent with a

little contribution to world production (only 6%), but very important exportation trade (22%). Venezuela is the 16th biggest world producer (30,000 tons per year), and the seventh biggest exporter of sesame seed (24,000 tons per year). Its sesame is considered to be of high quality.

Sesame production in Venezuela is important in the Western Llanos, specifically around Turen town, in Portuguesa state. In the early years of Venezuelan sesame production, since 1940 until 1990, sesame was used as oil source for the national market, and some as oil for export. Presently it is used for export as processed grain. Because of its importance for export, sesame breeding attained a high priority in Venezuela leading to the development of over 30 cultivars during the last 60 years. Reliable identification of these cultivars is a requirement. DNA fingerprinting has been used for checking the identity and purity of cultivars in different crops and for assessing the genetic variability of breeding stocks (Fernandez et al. 2002; Archak et al. 2003; Rajora and Rahman 2003; de Moretzsohn et al. 2004; Dangi et al. 2004; Buhariwalla et al. 2005). It has been particularly useful for the selection of germplasm in crossing schemes. Amplified fragment length polymorphism (AFLP) is a reliable genotyping method with a high degree of reproducibility and discriminatory power (Savelkoul et al. 1999). AFLP has proved to be a robust marker technique to distinguish plant genotypes (Milbourne et al. 1997; Zhang et al. 1999; Muminovic et al. 2004). A recently developed database format for AFLP data allows for storage and comparison of profiles of cultivars and accessions (Hong and Chuah 2003). The ability of markers to discriminate between genotypes is usually estimated by means of probability of identical match by chance (Pi) (Ramakrishna et al. 1994), marker index (MI) (Powell et al. 1996), resolving power (*Rp*) (Prevost and Wilkinson 1999), polymorphic information content (PIC) (Roldán-Ruiz et al. 2000) and recording both the number of fingerprints or haplotypes observed, and the number of genotypes with unique fingerprints (Rajora and Rahman 2003).

The aims of the present study were to evaluate the ability of AFLP markers for distinguishing 20 sesame cultivars, to determine the genetic relationship among these genotypes and to estimate the usefulness of parameters currently used to assess the informativeness of molecular markers for genotyping.

Material and methods

Plant materials

Twenty cultivars, coming from different sesame breeding programs and representative of the commercial cultivars used in Venezuela, were used in the present study. They are listed in Table 1 with information regarding their origin.

DNA extraction

Three grams of apical young leaves from 6 plants per accession were collected and used for DNA extraction. Leaves were ground in liquid nitrogen and the tissue powder was dispersed in CTAB buffer (2.3 g sorbitol, 1 g N-laurylsarcosine, 0.8 g CTAB, 4.7 g sodium chloride, and 1 g polyvinylpyrodidone in a total volume of 100 ml of 20 mM EDTA, 10 mM Tris, pH set to 8.0) containing 0.4 mg proteinase K and 20 µl mercaptoethanol. The homogenates were incubated for 10 min at 42°C and 10 min at 65°C, cooled to room temperature and extracted with 8 ml of chloroform/isoamylalcohol (24:1). Phases were separated by centrifugation for 10 min at 12000 RCF (relative centrifugal force or g value). Polyethyleneglycol (PEG 6000, SERVA Electrophoresis, Germany) stock solution (30%) was added to the aqueous phase to a final concentration of 6%, mixed, and after 30 min of incubation at room temperature the precipitated DNA was sedimented by centrifugation for 20 min at 12000 RCF. The pellets were washed twice with 70% ethanol and dissolved in 200 µl TE buffer (10 mM Tris HCl pH 8.0, 0.1 mM EDTA). 500 µl of 5 M ammonium acetate solution were added and samples were kept at 0°C for 30 min, centrifuged for 30 min at 4°C and 18000 RCF. 500 µl of isopropanol were added to the supernatant and DNA was precipitated (10 min at room temperature). Samples were centrifuged at 18000 RCF at room temperature for 10 min; pellets were

Table 1 Commercial cultivars used in the present study and their respective origin

Cultivar	Origin
Venezuela 51	Originated by individual selection from the offspring of a Chinese accession (Langham and Rodriguez 1946).
Acarigua	A high performance F2 plant obtained by the cross between a cultivar from Nicaragua and a cultivar from China, was crossed with the cultivar Venezuela 51, its offspring was selected for three seasons, resulting in ''Acarigua'' (Mazzani 1952).
Inamar	Individual selection from the offspring from the same Acariguás parents (Mazzani 1953).
Maporal	Selected from cultivar Arapatol, from Ethiopia (Mazzani et al. 1973).
Caripucha	Unknown
Felicidad	Introduced from Mexico. Unknown origin
Chino Amarillo	Introduced from Mexico. Unknown origin
UCV-1	Elite line selected from first cycle of recurrent selection toward high yield. The original population was obtained by cross, one to one, among 50 exotic accessions (Laurentin et al. 2000).
43 × 32,19 × 10	Selected lines from second cycle of recurrent selection toward high yield, under heavy whitefly infestation. The original population was obtained by cross, one to one, among 50 exotic accessions (Laurentin et al. 2000).
UCV-3	Individual selection from Arawaca (unpublished data).
Fonucla	Selection from cultivar Arawaca (Montilla and Cedeño 1991). Arawaca was obtained by selection of the mixture of 496 F1 plants obtained from crosses among 32 cultivars without reciprocal. The origin of these cultivars is unknown.
UCLA1	Individual selection from a USA accession (Montilla and Teran 1996). Unknown origin
UCLA37-1, UCLA65, UCLA83, UCLA90, UCLA249, UCLA295	Elite lines from Universidad Centrooccidental Lisandro Alvarado Sesame Breeding Program. Unknown origin
Glauca	Unknown origin.

washed twice with 70% ethanol, dried and dissolved in 200 µl of TE buffer. DNA concentration was determined by electrophoresis in a 0.8% agarose gel with lambda DNA standard.

AFLP analysis

AFLP analysis was performed as originally proposed (Voss et al. 1995) with minor modifications (Reineke and Karlovsky 2000; Laurentin and Karlovsky 2006). In general, AFLP were carried out in the following way: 250 ng of DNA were used for each reaction, which was replicated twice for each cultivar. DNA was digested with 10 U EcoRI and 3 U of Tru1I (both entzymes from MBI Fermentas, Germany). 10 µl of a solution with final concentration of 5 pmol of EcoRI adapter, 50 pmol of Tru1I adapter, 1× T4 DNA ligase buffer and 1U T4 DNA ligase (MBI Fermentas, Germany) were added to the digested DNA. The solution was incubated at 20°C for 2 h, and diluted 10-fold with TE buffer. Following ligation, a first amplification was carried out with primers containing one selective nucleotide (cytocine and adenine for MseI and EcoRI primers, respectively) (Table 2), in a total volume of 10 µl. PCR was performed for 20 cycles, which consisted of 30 s at 94°C, 1 min at 56°C and 1 min at 72°C in the thermocycler Tpersonal (Biometra, Göttingen, Germany). The PCR products were diluted 10-fold with TE buffer. The second amplification was carried out with eight primer combinations using labeled EcorRI-primer (Cy5)E_ACA combined with one of the eight MseI primers listed in Table 2. The thermocycler program consisted of two segments. The first segment comprised 12 cycles with the annealing temperature decreased from 65°C by 0.7°C in each cycle: 30 s at 94°C, 30 s at 65°C to 57.3°C and 1 min at 72°C. The second segment consisted of 23 cycles of 30 s at 94°C, 1 min at 56°C and 1 min at 72 °C. PCR products were mixed with 10 µl of loading buffer (98% formamide, 10 mM EDTA and 0.025% brom-

Table 2 Primer sequences used in the first and second amplification

Primer name	Sequence 5′–3′
AFLP_E_A	GACTGCGTACCAATTCA
AFLP_E_ACA	(Cy5)GACTGCGTACCAATTCACA
AFLP_M_C	GATGAGTCCTGAGTAAC
AFLP_M_CAA	GATGAGTCCTGAGTAACAA
AFLP_M_CAT	GATGAGTCCTGAGTAACAT
AFLP_M_CAG	GATGAGTCCTGAGTAACAG
AFLP_M_CAC	GATGAGTCCTGAGTAACAC
AFLP_M_CCA	GATGAGTCCTGAGTAACCA
AFLP_M_CCC	GATGAGTCCTGAGTAACCC
AFLP_M_CTCA	GATGAGTCCTGAGTAACTCA
AFLP_M_CGAA	GATGAGTCCTGAGTAACGAA

ophenolblue), denatured for 4 min at 90°C and 5 µl of each reaction (twice per cultivar) were loaded onto a 7% polyacrylamide gel (ReproGel™ LongRead,Amersham Pharmacia Biotech, Uppsala, Sweden) and run in the ALFexpress II DNA analyser (Amersham Pharmacia Biotech, Uppsala, Sweden). Three microliters of Genemark 500 Fluorescent DNA ladder, labeled with Cy5 (Northernbiothech, Weston, USA), were loaded on each gel and the electrophoresis was performed for 700 min at 1500 V, 25 W, 60 mA and 55°C. The electropherogram recorded by software ALFwin™ Sequence Analyser 2.00 (AmershamPharmacia Biotech, Uppsala, Sweden) was transformed to a pseudogel image in TIFF-format, visualized in Adobe® ImageReady™ version 3.0 (Adobe Systems Inc., USA) and analyzed using GelCompar II (Applied Maths, Belgium).

Statistical analysis

Bands were automatically recognized by GelCompar II using threshold values of 5% of profiling (relative to the maximum value within each lane). Band matching was performed and repeatable fragments between the two AFLP reactions for each cultivar were identified (in all the cases between 94 and 100%). Repeatable fragments were exported as a binary matrix, which was used for all the analysis. Discriminatory power of AFLP markers was evaluated by means of three parameters. The Polymorphic Information Content (PIC) for each AFLP was calculated as previously proposed (Roldan-Ruiz et al. 2000): $PIC_i = 2f_i(1-f_i)$, where PIC_i is the polymorphic information content of marker i, f_i the frequency of the marker bands which were present and $1-f_i$ the frequency of marker bands which were absent. Dominant markers as AFLP have a maximum PIC of 0.5 when half of the accessions have the band and the other half does not have the band (De Riek et al. 2001). *PIC* was averaged over the bands for each primer. Marker index (MI) was calculated as proposed by Powell et al. (1996) and used by Milbourne et al. (1997): MI is the product between diversity index (equivalent to PIC) and effective multiplex ratio (EMR), where EMR is defined as the product of the fraction of polymorphic loci and the number of polymorphic loci. This parameter was calculated for each primer. Resolving power (Rp) of each primer was calculated according Prevost and Wilkinson (1999): $Rp = \sum Ib$ where Ib (band informativeness) takes the values of: $1-[2x|0.5-p|]$, where p is the proportion of the genotypes containing the band. In addition to these parameters, number of different fingerprints per primer and number of elite lines with unique fingerprints per primer were recorded. Pearson correlation coefficients were calculated between the three parameters and both number of different fingerprints per primer and number of cultivars with unique fingerprint per primer. To get the level of confidence in identifying the 20 cultivars, the probability of identical match by chance (Pi) was calculated as proposed by Wetton (1987) and Ramakrishna et al. (1994): $Pi = X^n$, where X is a similarity index between 2 genotypes and n is the average number of bands in the two genotypes compared. Pi expresses the probability that a band present in one genotype is also present in the other. X was calculated as $2N_{AB}/(N_A + N_B)$, where N_{AB} is the number of bands present in both genotypes, N_A the total number of bands in genotype A, and N_B the total number of bands in genotype B. This index represents the probability that the bands present in one cultivar are also present in the other. This probability was calculated for each possible comparison between pairs of cultivars and for each primer and the highest probability was recorded. We believe that for genotype identification purposes it is desirable to know the highest

probability of identical match by chance rather than the average value for each primer combination, as proposed by Ramakrishna et al. (1994).

Jaccards similarity coefficient and unweighted pair group method with arithmetic mean (UPGMA) were used to perform the clustering analysis, which was tested with two statistical significance tests. They were the bootstraping analysis for the assessment of the robustness of the dendrogram topology, and the cophenetic correlation as an estimation of the faithfulness of the cluster analysis. Firstly, the bootstraping analysis was carried out using WinBoot software. Dendrogram-derived similarities were compared with experimental similarities to get cophenetic correlation. Principal coordinates analysis (PCA) was also carried out to display the location of the 20 cultivars in three-dimensions. All numerical taxonomic analyses were conducted using NTSYS-PC software, version 2.11T (Exeter Software, New York).

Results

Table 3 displays the number of bands (NB), number of polymorphic bands (NPB), number of exclusive bands (NEB), number of haplotypes (NH), number of genotypes with exclusive haplotypes (NGEH), polimorphic information content (PIC), resolving power (*Rp*) and marker index (MI) obtained per AFLP primer combination. The total number of bands was 339; ninety one percent of them being polymorphic. Number of bands per primer combination ranged from 22 (E_ACA + M_CGAA) to 70 (E_ACA + M_CAT), and polymorphism ranged from 71% (E_ACA + M_CCA) to 100% (E_ACA + M_CGAA and E_ACA + M_CAG). Fifty unique bands were obtained for 13 genotypes, where Maporal, UCLA83 and UCLA37-1 had most with 18, 7 and 6 respectively. Five primer combinations were able to discriminate the 20 cultivars. Combination E_ACA + M_CAG recorded the highest values for PIC, *Rp*, and MI. With PIC, *Rp* or MI, no significant correlation was found between either the number of fingerprints or elite lines with exclusive fingerprints. Number of bands per genotype ranged between 106 for 19 × 10 and 197 for UCV-1, with an average of 160 bands per genotype.

Table 4 displays minimum, maximum and average probability of identical match by chance per primer. Using the 8 primer combinations, the maximum probability of identical match by chance was 1:20000 between cultivars Fonucla and UCLA65. For three primer combinations, some genotypes generated identical AFLP patterns, leading to Pi value of 100%. Even then, the average probability of identical match by chance in patterns generated by these three primers were either low, medium or high as compared with the other primers. This shows that the average probability is not suitable as a measure for the assessment of the capability of primer pairs to distinguish among genotypes.

Jaccard's similarity coefficients ranged from 0.31 between Chino Amarillo and Maporal to 0.78 between Fonucla and UCLA65, with an average of 0.52 (Table 5). The UPGMA-based

Table 3 Number of bands (NB), number of polymorphic bands (NPB), number of exclusive bands (NEB), number of haplotypes (NH), number of genotypes with exclusive haplotype (NGEH), polymorphic information content (PIC), resolving power (*Rp*) and marker index (MI) obtained per AFLP primer combination

Primer E_ACA combined with:	NB	NPB	NEB	NH	NGEH	PIC	*Rp*	MI
M_CTCA	27	26	3	20	20	0.27 ± 0.16	10.30	6.88
M_CAA	41	36	3	18	17	0.29 ± 0.18	18.20	9.10
M_CCA	42	30	6	20	20	0.22 ± 0.19	13.10	4.62
M_CGAA	22	22	6	20	20	0.30 ± 0.17	10.10	6.65
M_CAT	70	68	15	19	18	0.29 ± 0.16	29.30	18.89
M_CAG	50	50	3	20	20	0.39 ± 0.13	30.50	19.39
M_CCC	33	30	4	19	18	0.22 ± 0.14	19.40	6.12
M_CAC	54	48	10	20	20	0.26 ± 0.18	20.60	11.09
Average	42.40	38.90	6.25	19.38	18.88	0.28 ± 0.16	18.94 ± 7.34	10.34 ± 5.40

Table 4 Minimum, maximum and average of probability of identical match by chance for each primer combination

Primer E_ACA combined with:	Probability of identical match by chance		
	Minimum	Average	Maximum between:
M_CTCA	2.39×10^{-17}	0.102	0.3568, UCLA249 and UCLA83
M_CAA	4.11×10^{-16}	0.079	1.000, UCLA295, UCLA37–1 and 19×10
M_CCA	4.36×10^{-8}	0.126	0.8521, UCLA83 and UCLA65
M_CGAA	0.000	0.039	0.7500, Acarigua and Inamar
M_CAT	2.42×10^{-35}	0.031	1.000, Inamar and 19×10
M_CAG	0.000	0.014	0.5398, Caripucha and Glauca
M_CCC	6.11×10^{-11}	0.134	1.000, UCLA295 and UCLA37-1
M_CAC	8.77×10^{-19}	0.040	0.7023, UCLA249 and UCLA295
Total	3.11×10^{-78}	2.70×10^{-7}	5.19×10^{-5}, Fonucla and UCLA65

phenogram (Fig. 1) and biplot from principal coordinates analysis (Fig. 2) showed a similar pattern: cultivars Maporal, Chino Amarillo, 19×10, Felicidad, Inamar, and UCLA37-1 were different, and they appeared separated from the others for the two analyses. The other cultivars were grouped in two clusters by dendrogram. Both analyses failed to group together all the cultivars that are related by pedigree: Acarigua and Inamar coming from Venezuela 51; UCV-1, 19×10, 43×32 coming from the same basic population; Fonucla and UCV-3 selected from the same cultivar. The cophenetic correlation coefficient was 90%. Bootstrapping values were > 70% in cluster grouping 15 cultivars at 0.58 similarity value. The principal coordinate analysis (PCO) showed that the first three axes accounted for 95% of total variation.

Table 5 Minimum, maximum and mean of Jaccard́s similarity coefficients of 20 sesame cultivars based on 339 AFLP markers

Cultivar	Similarity coefficient		
	Minimum	Maximum	Mean ± SD
Chino Amarillo	0.309	0.476	0.385 ± 0.042
Felicidad	0.376	0.587	0.463 ± 0.059
Venezuela 51	0.389	0.747	0.551 ± 0.109
Acarigua	0.394	0.744	0.560 ± 0.098
UCV-1	0.359	0.747	0.555 ± 0.107
Maporal	0.309	0.490	0.405 ± 0.040
Caripucha	0.376	0.705	0.571 ± 0.097
Inamar	0.370	0.646	0.491 ± 0.063
Glauca	0.434	0.659	0.555 ± 0.056
43×32	0.338	0.681	0.560 ± 0.101
19×10	0.337	0.500	0.434 ± 0.042
UCLA249	0.379	0.685	0.563 ± 0.093
UCLA83	0.361	0.684	0.540 ± 0.100
UCLA1	0.397	0.674	0.542 ± 0.076
UCLA90	0.356	0.674	0.560 ± 0.085
UCLA295	0.384	0.672	0.525 ± 0.083
UCLA37-1	0.398	0.646	0.500 ± 0.073
Fonucla	0.377	0.781	0.558 ± 0.101
UCLA65	0.352	0.781	0.582 ± 0.100
UCV-3	0.341	0.613	0.509 ± 0.068

Discussion

AFLPs from eight primer combinations have been a successful tool for identifying commercial cultivars with a low probability of getting identical match by chance. E_ACA + M_CAG is considered a valuable primer combination, because it is the most informative for all the indexes calculated: polymorphic information content, marker index, resolving power, number of haplotypes, number of cultivars resolved, average probability of identical match by chance, and it presented the second lowest value for the maximum probability of identical match by chance when all the possible comparisons were carried out. Polymorphic information content (PIC) and marker index (MI) have been used to measure informativeness of AFLP primer combinations in other self-pollinated crops such as soybean (*Glycine max* (L.) Merr.) (PIC = 0.32 and MI = 6.14, Powell et al. 1996), wheat (*Triticum aestivum* L.) (PIC = 0.32 and MI = 3.41, Bohn et al. 1999; PIC = 0.31, Stodart et al. 2005), cornsalad (*Valerianella locusta* (L.) Lat.) (PIC = 0.25 and MI = 4.47, Muminovic et al. 2004) and triticale (×*Triticosecale* Wittm.) (PIC = 0.25 and MI = 8.60, Tams et al. 2005). Why these parameters related with the informa-

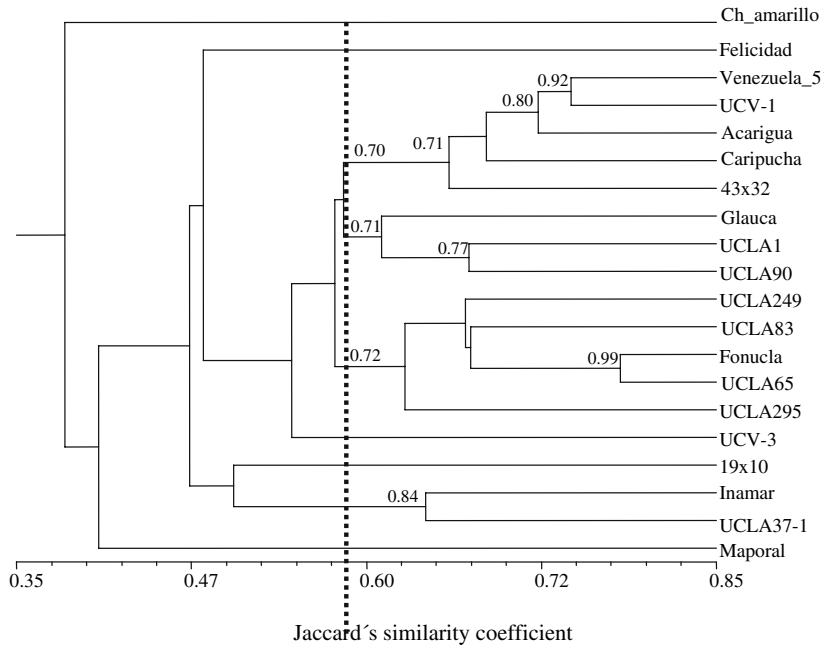

Fig. 1 Phenogram of 20 sesame cultivars based on Jaccards similarity coefficients using 339 AFLP markers. Bootstrapping values are indicated for clusters with values equal or higher than 0.70

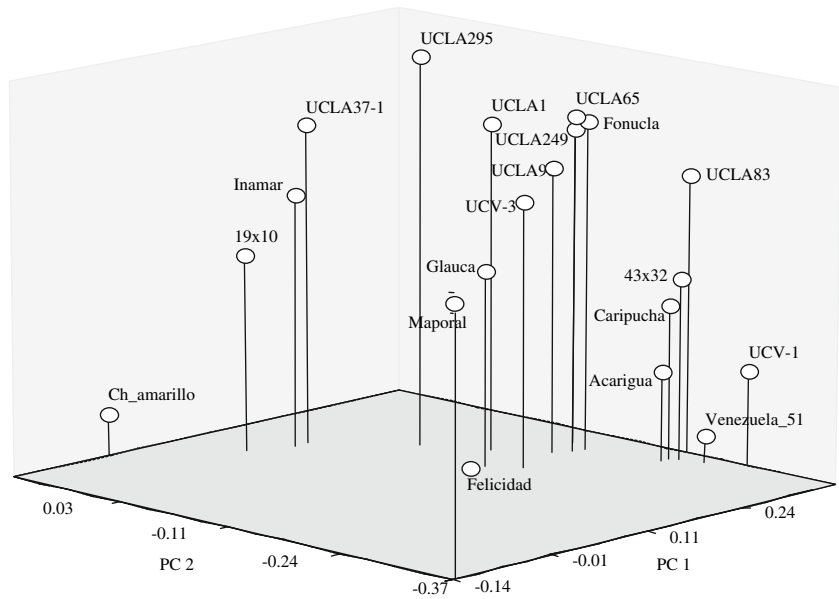

Fig. 2 Three-dimensional graph from the principal coordinate analysis of 339 AFLP markers of 20 sesame cultivars

tiveness of molecular markers for identifying genotypes is not totally clear. Most of the studies have used these indexes for comparing AFLPs with other molecular markers. Discrimination of as many cultivars as possible would be the most important feature of one primer combination, when the purpose of the evaluation is to identify unequivocally a specific genotype. Prevost and Wilkinson (1999) and Fernandez et al. (2002) found a strong and linear relationship between the ability of a primer to distinguish genotypes and resolving power (*Rp*), but not with marker index (MI). The data reported by Rajora and Rahman (2003) indicate significant correlation ($P < 0.05$) between PIC and number of genotypes observed, but not with number of cultivars with

unique genotype. Lack of correlation between PIC, MI, *Rp* and both number of haplotypes and number of cultivars resolved in our study, or lack of consistency in the correlation in other studies, make it clear that these indexes do not ever evaluate the informativeness of a primer combination. It would be more suitable to record how many genotypes are discriminated by primer, than to calculate parameters such as PIC, *Rp* and MI, when fingerprints are carried out for identifying genotypes.

Probability of identical match by chance is an important calculation when the purpose of fingerprinting is to identify genotypes for checking identity and purity of a cultivar. Our results suggest that averages of this probability for each primer do not seem suitable for this purpose. Primer combinations with low (E_ACA + M_CAT) and intermediate (E_ACA + M_CAA) average probability of identical match by chance, resulted with probability of 100% of identical match by chance between at least one pair of comparisons. This result indicates the importance for calculating of maximal probability of identical match by chance, and also suggests the use of several primer combinations for identifying genotypes. Maximal probability of 100% for individual primer combinations became maximal probability of 0.000052 when 8 primer combinations were used.

AFLP fingerprint showed an average of 6 unique bands per primer combination; this may be very useful for genotyping cultivars, because these unique bands can be converted into STS (sequence tagged site) markers. This in turn may be useful for detecting mixes between cultivars (Fernadez et al. 2002).

The UPGMA-based phenogram and principal coordinate analysis displayed similar pattern. Cultivars UCV-1, 19×10 and 43×32, derived from the same basic population, did not group together. This basic population resulted from crosses among 50 accessions to produce a highly variable population. Three best lines of the recurrent selection program mentioned above, which cover a broad genetic diversity as compared with the remaining 17 cultivars, originate from this highly variable population. Cultivars Fonucla and UCV-3, selected from the same cultivar (Arawaca), differed considerably from each another. Arawaca was obtained by a bulk population method, which is characterized by selecting a mixture of genotypes with similar phenotypic traits; therefore there is a theoretical explanation for this result. Inamar and Acarigua come from the same single cross between Venezuela 51 and an F_2 plant. This F_2 plant must have had a high level of heterozygocity to obtain such dissimilar cultivars. Genetically dissimilar cultivars coming from the same single cross have been reported in other crops such as barley (*Hordeum vulgare* L.) (Fernandez et al. 2002) and cashew (*Anacardium occidentale* L.) (Archak et al. 2003). This observation is so interesting in sesame breeding, because it is revealing that only two parents may be enough for generating a base population with broad genetic variability.

Nine cultivars used in this study (Fonucla, UCLA1, UCLA249, UCLA295, UCLA37-1, UCLA65, UCLA83, UCLA90 and UCV-3) were characterized using morphological traits (Laurentin et al. 2004) and RAPD (Salazar et al. 2006). When only these cultivars were subjected to principal coordinate analysis with our AFLP data, the three studies were similar only in grouping closely UCLA90 and UCLA1. But when we compare AFLP and RAPD studies with each other, even though Mantel test showed non-significant correlation between similarity matrices ($P < 0.05$), two clusters grouped the same cultivars: UCLA1, UCLA90 in one cluster, and UCLA65, UCLA295, Fonucla in other one; furthermore both analysis failed to assign UCV-3 and UCLA37-1 to some cluster. Why this differs from the morphological characterization can be explained because molecular characterization covers the entire genome variability (Ovesná et al. 2002) excluding the environmental influence (Rao 2004), whereas morphological characterization, mainly of quantitative traits in multi- environment experiments as those studied by Laurentin et al. (2004), are subjected to strong environmental influence (Karp et al. 1997; Rao 2004). According to our study, and in agreement with the previous study using RAPD (Salazar et al. 2006), UCLA37-1 appear to be the most suitable parents of an eventual new ''white seed'' population when crossed with some of the other white seed cultivars studied (elite lines from

Universidad Centrooccidental Lisandro Alvarado sesame breeding program, Table 1), under the assumption that the more is the genetic distance between parents the more is the possibility to identify potential and suitable new cultivars from a segregant population.

Five cultivars used in the present study (19×10, 43×32, Fonucla, UCV-3 and UCLA37-1) were characterized according presence or absence of secondary metabolites in roots, stems, leaves, fruits and seeds (Laurentin et al. 2003). The three-dimensional graph from the principal coordinate analysis, even though Mantel test showed non-significant correlation between similarity matrix (AFLP) and correlation matrix (secondary metabolites) ($P < 0.05$), displayed a similar grouping to that of these cultivars using AFLP. This close relationship between AFLP and secondary metabolites could be useful in future breeding programs, even more so when these secondary metabolites were related with resistance against whitefly.

The results of the present study have demonstrated that AFLP-based fingerprints are a useful tool to identify sesame genotypes unequivocally. This information could be used for cultivar identification and protection of breeder's rights. Also, AFLP-fingerprints have been used successfully in our study for assessing genetic variability of breeding stocks and for the determination of the genetic relationship among cultivars.

Acknowledgements This work was supported by the Programme Alban, European Union Programme of High Level Scholarships for Latin America, Identification Number E03D13301VE, International PhD program for Agricultural Sciences in Göttingen University (IPAG) and Universidad Centroccidental Lisandro Alvarado.

References

Archak S, Gaikwad B, Gautam D, Rao E, Swamy K, Karihaloo J (2003) DNA fingerprinting of Indian cashew (*Anacardium occidentale* L.) varieties using RAPD and ISSR techniques. Euphytica 230:397–404

Ashri A (1998) Sesame breeding. Plant Breed Rev 16:179–228

Bohn M, Utz H, Melchinger A (1999) Genetic similarities among winter wheat cultivars determined on the basis of RFLPs, AFLPs and SSRs and their use for predicting progeny variance. Crop Sci 39:228–237

Buhariwalla HK, Jayashree B, Eshwar K, Crouch JH (2005) Development of ESTs from chickpea roots and their use in diversity analysis of the *Cicer* genus. BMC Plant Biol 5:16

Dangi RS, Lagu MD, Choudhary LB, Ranjekar PK, Gupta VS (2004) Assessment of genetic diversity in *Trigonella foenum-graecum* and *Trigonella caerulea* using ISSR and RAPD markers. BMC Plant Biol 4:13

De Riek J, Calsyn E, Everaert I, Van Bockstaele E, De Loose M (2001) AFLP based alternatives for the assessment of distinctness, uniformity and stability of sugar beet varieties. Theor Appl Genet 103:1254–1265

FAO (2005) FAOstat Databases. http://www.faostat.fao.org/

Fernandez M, Figueiras A, Benito C (2002) The use of ISSR and RAPD markers for detecting DNA polymorphism, genotype identification and genetic diversity among barley cultivars with known origin. Theor Appl Genet 104:845–851

Hong Y, Chuah A (2003) A format for databasing and comparison of AFLP fingerprint profiles. BMC Bioinformatics 4:7

Karp A, Kresovich S, Bhat K, Ayad W, Hodgkin T (1997) Molecular tools in plant genetic resources conservation: a guide to the technology. IPGRI Technical Bulletin No. 2. International Plant Genetic Resources Institute, Rome, Italy

Langham D, Rodriguez M (1946) Dos nuevas variedades de ajonjolí Venezuela 51 y Venezuela 52. Circular No. 15. Departamento de Genetica. Dirección de Agricultura. Maracay. Venezuela

Laurentin H, Layrisse A, Quijada P (2000) Evaluación de dos ciclos de selección recurrente para altos rendimientos de semilla en una población de ajonjolí. Agronomía Tropical (Maracay) 50:521–535

Laurentin H, Pereira C, Sanabria M (2003) Phytochemical characterization of six sesame (*Sesamum indicum* L.) genotypes and their relationships with resistance against whitefly (*Bemisia tabaci* Gennadius). Agronomy J 95(6):1577–1582

Laurentin H, Montilla D, Garcia V (2004) Relación entre el rendimiento de ocho genotipos de ajonjolí (*Sesamum indicum* L.) y sus componentes. Comparación de metodologías. Bioagro 16:153–162

Laurentin H, Karlovsky P (2006) Genetic relationship and diversity in a sesame (*Sesamum indicum* L.) germplasm collection using amplified fragment length polymorphism. BMC Genet 7:10

Mazzani B (1952) Una nueva variedad de ajonjolí producica en el Instituto Nacional de Agricultura. Agronomía Tropical (Maracay) 1:269–277

Mazzani B (1953) Inamar: nueva variedad de ajonjolí producidad en el Instituto Nacional de Agricultura. Agronomía Tropical (Maracay) 3:211–213

Mazzani B, Nava C, Martinez A, Layrisse A (1973) Maporal, una nueva variedad de ajonjoli para los Llanos Occidentales. Agronomia Tropical 23:501–508

Milbourne D, Meyer R, Bradshaw J, Baird E, Bonar N, Provan J, Powell W, Waught R (1997) Comparison of PCR-based marker systems for the analysis of genetic

relationships in cultivated potato. Mol Breed 3:127–136
de Moretzsohn MC, Hopkins MS, Mitchell SE, Kresovich S, Valls JF, Ferreira ME (2004) Genetic diversity of peanut (*Arachis hypogaea* L.) and its wild relatives based on the analysis of hypervariable regions of the genome. BMC Plant Biol 4:11
Montilla D, Cedeño T (1991) Fonucla: una nueva variedad de ajonjolí (*Sesamum indicum* L.). Bioagro 3:52–54
Montilla D, Teran H (1996) UCLA1, una nueva variedad de ajonjolí (*Sesamum indicum* L.). Bioagro 8:26–29
Muminovic J, Melchinger A, Lübberstedt T (2004) Genetic diversity in cornsalad (*Valerianella locusta*) and related species as determined by AFLP markers. Plant Breed 123:460–466
Ovesná J, Poláková K, Leisová L (2002) DNA analyses and their applications in plant breeding. Czech J Genet Plant Breed 38:29–40
Powell W, Margenta M, Andre C, Hanfrey M, Vogel J, Tingey S, Rafalsky A (1996) The utility of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol Breed 2:225–238
Prevost A, Wilkinson M (1999) A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. Theor Appl Genet 98:107–112
Rajora O, Rahman M (2003) Microsatellite DNA and RAPD fingerprinting, identification and genetic relationships of hybrid poplar (*Populus* × *canadensis*) cultivars. Theor Appl Genet 106:470–477
Ramakrishna W, Lagu M, Gupta V, Ranjekar P (1994) DNA fingerprinting in rice using oligonucleotide probes specific for simple repetitive DNA sequences. Theor Appl Genet 88:402–406
Rao N (2004) Plant genetic resources: advancing conservation and use through biotechnology. Afr J Biotechnol 3:136–145
Reineke A, Karlovsky P (2000) Simplified AFLP protocol: replacement of primer labeling by the incoporation of α-labeled nucleotides during PCR. BioTechniques 28:622–623
Roldán-Ruiz I, Dendauw J, VanBockstaele E, Depicker A, De Loose M (2000) AFLP markers reveal high polymorphic rates in ryegrasses (*Lolium* spp.). Mol Breed 6:125–134
Salazar B, Laurentin H, Dávila M, Castillo M (2006) Reliability of the RAPD technique for germplasm analysis of sesame (*Sesamum indicum* L.) from Venezuela. Interciencia 31:456–460
Savelkoul P, Aarts H, DeHaas J, Dijkshoorn L, Duim B, Otsen M, Rademaker J, Schouls L, Lenstra J (1999) Amplified-fragment length polymorphism analysis: the state of an art. J Clin Microbiol 37:3083–3091
Stodart B, Mackay M, Raman H (2005) AFLP and SSR analysis of genetic diversity among landraces of bread wheat (*Triticum aestivum* L. em. Thell.) from different geographic regions. Aust J Agric Res 56:691–697
Tams S, Melchinger A, Bauer E (2005) Genetic similarity among European winter triticale elite germplasm assessed with AFLP and comparisons with SSR and pedigree data. Plant Breed 124:154–160
Voss P, Hogers R, Bleeter M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414
Wetton J, Carter R, Parkin D, Walters D (1987) Demographic study of a wild house sparrow population by DNA fingerprinting. Nature 327:147–149
Zhang LH, Ozias-Akins P, Kochert G, Kresovich S, Dean R, Hanna W (1999) Differentiation of bermudagrass (*Cynodon* spp.) genotypes by AFLP analysis. Theor Appl Genet 98:895–902