Published online June 24, 2005

Genetic Variation in Sorghum Germplasm from Sudan, ICRISAT, and USA Assessed by Simple Sequence Repeats (SSRs)

A. H. Abu Assar, R. Uptmoor, A. A. Abdelmula, M. Salih, F. Ordon, and W. Friedt*

ABSTRACT

Assessment of genetic variability in crops has a strong impact on plant breeding and conservation of genetic resources. It is particularly useful in the characterization of individuals, accessions, and cultivars in determining duplications in germplasm collections and for selecting parents. The objective of this study was to estimate genetic diversity and to obtain information on the genetic relationship among 96 sorghum [*Sorghum bicolor* (L.) Moench] accessions from Sudan, ICRISAT, and Nebraska, USA, using 16 simple sequence repeats (SSRs). In total, 117 polymorphic bands were detected with a mean of 7.3 alleles per SSR locus. By this approach each accession is uniquely fingerprinted. Genetic similarity estimates ranged from 0 to 0.91, with a mean of 0.30. The polymorphic information content (PIC) for SSRs ranged from 0.46 (SB4-72) to 0.87 (SBAGF06). Diversity index (DI) for all accessions was 0.71. Within subgroups, DI was 0.63 for Sudanese landraces and improved cultivars, 0.49 for PI accessions, 0.42 for Nebraska derivatives, 0.39 for the ICRISAT advanced breeding lines (ABLs), 0.65 for the Feterita group, 0.71 for the Milo group, 0.63 for a Synthetic group (new breeding materials), 0.68 for the Hegiri group, and 0.47 for the Mugud group. Mantel statistics revealed a good fit of the unweighted pair-grouped method with arithmetic average (UPGMA) cluster to the original genetic similarity (GS) data ($r = 0.867$). UPGMA clustering produced two main clusters comprising mainly nonimproved germplasm (gene bank accessions and Nebraska population derivatives), and improved genotypes (cultivars, Gadarif collections, and ICRISAT advanced lines). Grouping of accessions by UPGMA cluster analysis matched with the geographical origin and/or pedigree information (Sudan, USA, ICRISAT), the adaptation zone (Gadarif area, Sudan), and morphological characters (Feterita, Mugud, and Milo types), indicating the strong differentiation among the sorghum materials.

Assessment of the genetic variability within cultivated crops has a strong impact on plant breeding strategies and conservation of genetic resources (Dean et al., 1999; Simioniuc et al., 2002). It is particularly useful in the characterization of individuals, accessions, and cultivars in determining duplications in germplasm collections and for the choice of parental genotypes in breeding programs (Davila et al., 1998; Ribaut and Hoisington, 1998). In the past, indirect estimates of similarity based on morphological information have been widely used in many species including sorghum (Ayana and Bekele, 1999). However, morphological variation does not reliably reflect the real genetic variation because of genotype-environment interactions and the largely unknown genetic control of polygenically inherited morphological and agronomic traits (Smith and Smith, 1992).

Sorghum is fifth in acreage among the world's cereals (Doggett, 1988). It consists of cultivated and wild species. *Sorghum bicolor* subsp. *bicolor* ($2n = 20$) is the taxon that includes agronomically important grain races, that is, bicolor, caudatum, durra, guinea, and kafir, several hybrid races and working groups (for a review see Doggett, 1988).

Sorghum is an important staple food throughout semiarid Asian and African regions (Ahmed et al., 2000). Studying the genetic variation of sorghum germplasm collections from Sudan attracts special interest for several reasons. Beyond the economic importance of the crop, Sudan is within the geographical range where sorghum is believed to be domesticated for the first time (Mann et al., 1983) and where the largest genetic variation for both cultivated and wild sorghum is found (Doggett, 1988). However, phenotypic variation does not reliably reflect genetic variation because of the role of environmental interaction in determining the phenotype (Smith and Smith, 1989; Smith et al., 1991). In recent years, the number of molecular assays available for application in this area has increased dramatically, with each method differing in principles, applications, type and amount of polymorphism detected, as well as cost and time requirements (Karp et al., 1998). The molecular assays include restriction fragment length polymorphism (RFLP) (Botstein et al., 1980), random amplified polymorphic DNA (RAPD) (Williams et al., 1990), SSR polymorphism (Tautz, 1989), and amplified fragment length polymorphism (AFLP) (Zabeau and Vos, 1993).

Microsatellites (i.e., SSRs) are becoming the markers of choice for fingerprinting and genetic diversity studies in a wide range of living organisms (Gupta and Varshney, 2000). Simple sequence repeats represent an ideal marker system due to their codominant inheritance, locus specificity, and multi-allelic character. Therefore, they have been established as useful genetic markers in many plant species (Cregan et al., 1999; Goulão et al., 2001). Here we report on the use of SSRs for molecular characterization of sorghum accessions derived from and collected in Sudan with the main objectives of estimating genetic diversity and determining the genetic relationship among these accessions.

A.H. Abu Assar and M. Salih, Agricultural Research Corporation (ARC), P.O. Box 126, Wad Medani-Sudan; R. Uptmoor and W. Friedt, Institute of Crop Science and Plant Breeding I, Heinrich-Buff-Ring 26-32, D-35392 Giessen, Germany; A.A. Abdelmula, Department of Crop Production, Faculty of Agriculture, Khartoum North, Shambat 13314, P.O. Box 32; F. Ordon, Institute of Epidemiology and Resistance, Federal Centre for Breeding Research on Cultivated Plants, Theodor-Roemer-Weg 4, D-06449 Aschersleben, Germany. Received 1 Aug. 2003. *Corresponding author (wolfgang.friedt@agrar.uni-giessen.de).

Published in Crop Sci. 45:1636–1644 (2005).
Plant Genetic Resources
doi:10.2135/cropsci2003.0383

677 S. Segoe Rd., Madison, WI 53711 USA

Abbreviations: ABL, advanced breeding line; bp, base pairs; ddH_2O, double distilled water; DI, diversity index; GS, genetic similarity; ICRISAT, International Crops Research Institute for the Semi-Arid Tropics; PIC, polymorphic information content; SSR, simple sequence repeat; UPGMA, unweighted pair-grouped method with arithmetic average.

MATERIALS AND METHODS

Plant Material

A total of 96 sorghum genotypes was analyzed: 2 released cultivars, 35 landraces collected from sorghum-growing areas in Sudan, 12 ABL introduced from ICRISAT, 38 gene bank accessions (Sudanese landraces collection), and 9 advanced lines derived from a population developed from local landraces and genotypes introduced from Nebraska (Table 1). According to nomenclature of the farmers, germplasm can be classified into four groups based on head shape, seed color, and/or presence or absence of the pericarp. All Feterita genotypes have pericarp and white seed, whereas Milo genotypes have creamy seed color and no pericarp. The Hegiri group has dark brown seed and the Mugud group has spherical compact heads, white or red seeds, and no pericarp. Besides this, a synthetic group comprises the ABLs and population derivatives.

DNA Extraction and SSR Procedure

Leaf tips of 2-week-old seedlings were collected from five plants of each of the 96 genotypes. Seedlings were grown in quick-pot standard plates (33 by 51.5 cm) with one plant per pot in the greenhouse. A leaf sample of 100 to 150 mg was ground and DNA was extracted according to Doyle and Doyle (1990). The DNA content was measured fluorometerically (Hoefer Scientific Instrument, Model TKO 100, San Francisco, CA). The 16 SSR markers published by Brown et al. (1996) and Taramino et al. (1997) were used for the estimation of GS (Table 2). DNA amplifications were performed according to Brown et al. (1996) and performed in a Geneamp 9700 thermal cycler (Applied Biosystems, Foster City, CA) with 20-μL reaction volume. The SSR reaction contained 2 μL (25 ng) genomic DNA, 0.75 μL forward primer (2pmol μL^{-1}) labeled with IRD 700 or IRD 800, respectively, 0.75 μL reverse primer (2pmol μL^{-1}), 0.8 μL $MgCl_2$ (25 m*M*), 2 μL reaction buffer (10×), 0.4 μL dNTPs (100 m*M*), 0.1 μL *Taq* DNA polymerase (10 U μL^{-1}, Eppendorf, Hamburg, Germany) and 13.45 μL ddH_2O. The PCR reaction conditions consisted of 5 min at 94°C for initial denaturation, followed by 18 cycles of polymerization reaction, each consisting of a denaturation step of 30 s at 94°C, an annealing step of 30 s at 65°C (annealing temperature was reduced by 0.5°C in each of the 18 cycles), and a polymerization step of 1 min at 72°C. The next 20 cycles of polymerization reaction consisted of 15 s at 94°C, an annealing step of 30 s at 55°C, and a polymerization step of 1 min at 72°C, followed by a final polymerization step of 7 min at 72°C. A total of 16 primer pairs were used for PCR amplifications. An equal volume of formamide loading buffer was added and the samples were denatured at 94°C for 3 min; 1.0 μL of each sample was loaded on to a 25-cm, 8% denaturing polyacrylamide gel (Long Ranger, FMC Biozym, Hessisch Oldendorf, Germany) that had been preheated for 30 min. Electrophoresis was conducted in 1.0 Long Ranger TBE buffer at 1500 V, 50 W, 35 mA, and 48°C using a Li-Cor DNA Analyzer Gene Readir 4200 (Licor Biosciences, Bad Homburg, Germany). The fragment sizes were compared to a 50- to 350-bp standard (MWG Biotech AG, Ebersberg, Germany).

Data Analysis

The presence (1) or absence (0) of bands was scored using the software RFLP-Scan 2.1 (Scanalytic, Fairfax, VA). Each column of the resulting binary (0/1) matrix represented one allele of the corresponding SSR locus. Pairwise GS was estimated using SIMQUAL of the software package NTSYS-pc (Rohlf, 2000) according to Nei and Li (1979):

$$GS = \frac{2a}{2a + b + c}$$

where *a* refers to alleles shared between two accessions, and *b* and *c* refer to alleles present in either one of the accessions of a pairwise genotypic comparison. The similarity matrices were used to construct the dendrogram for all 96 accessions using SAHN of NTSYS-pc (Rohlf, 2000) based on UPGMA. The fit of the UPGMA cluster to the original similarity indices was computed according to the Mantel test procedure (Mantel, 1967) by using MxComp of the software package NTSYS-pc (Rohlf, 2000).

The PIC for each SSR was estimated by determining the frequency of alleles per locus:

$$PIC = 1 - \sum x_i^2$$

where x_i is the relative frequency of the *i*th allele of the SSR loci. The mean genetic DI across all loci was calculated according to Nei (1973):

$$DI = n_a(1/n_l \sum_j (1 - \sum_i x_{ij}^2))/(n_a - 1)$$

where x_{ij} is the frequency of the ith allele of locus, j, n_l is the number of genetic loci, and n_a is the number of accessions.

RESULTS

Genetic Relationships among Sorghum Accessions

The 16 SSR primer pairs used in this study were able to uniquely fingerprint each of the 96 sorghum accessions. Genetic similarity ranged from zero (Red Mugud versus PI 569798, Feterita Eriana versus PI 570413, and PI 569629 versus Dwarf White Milo) to 0.91 for SAR 16 versus SAR 35 with a mean of 0.30. The dendrogram generated from the UPGMA cluster analysis based on Nei and Li (1979) similarity indices grouped all 96 genotypes into two main clusters and 18 significant subclusters related to geographical origin, morphological characters, and/or adaptation zone (Fig. 1). The Mantel test (Mantel, 1967) showed a good fit of the cophenetic values to the original data set ($r = 0.867$).

The first cluster includes cultivars, landraces, and ABLs which are further subdivided into 12 groups, whereas the second cluster covers gene bank accessions and Nebraska population derivatives, which are further subdivided into six subgroups. The cluster from Red Mugud to White Mugud covers the Mugud group. This cluster is followed by two clusters of the Feterita landraces that were collected from El Gadarif state, covering the landraces from Feterita Eriana to Gadamballia, and from Feterita Rass Girid (Kassab) to 'Wad Ahmed'. Dabar Habashi clusters within Feterita landraces but belongs to the Milo group. Next to these clusters are the ICRISAT ABLs that are grouped into two clusters, from SRN 39 to ICSR 93004 and from ICSR 91030 to ICSR 93002. The clusters from Wad Akar to Dabar Baladi, from Abu Teman to Teteron, and from Abu Shy to Dabar Nigiri consist of a mixture of landraces belonging to the Feterita, Milo, and Hegiri group that were collected in El Gadarif, Sudan. Furthermore, the cluster from Gadamel Hamam to Ingaz, and that from IS 9830 to LRB 6 consists of Milo and Feterita types and is followed by two clusters of the Milo group (from Abu

Table 1. Names, collection sites, categories, pericarp status, and morphological groups of sorghum accessions used in this study.

Serial no.	Entry name†	Collection site‡	Category§	Pericarp¶	Grain color	Morphological group#
1	Red Mugud	southern Gadarif	landrace	A	light red	Mugud
2	Feterita Eriana	southern Gadarif	landrace	P	white	Feterita
3	Abu Teman	southern Gadarif	landrace	A	white	Milo
4	Wad Akar	southern Gadarif	landrace	P	brown	Hegiri
5	Arfa Gadamak	northern Gadarif	landrace	P	spotted-white	Feterita
6	Gadamel Hamam	northern Gadarif	landrace	P	white	Feterita
7	Dwarf White Milo	southern Gadarif	landrace	A	creamy	Milo
8	Abu Shy	southern Gadarif	landrace	A	creamy	Milo
9	Dabar Zera Zera	southern Gadarif	landrace	A	creamy	Milo
10	Teteron	southern Gadarif	landrace	A	white	Feterita
11	Feterita Rass Girid (Kassab)	southern Gadarif	landrace	P	white	Feterita
12	El Safra	southern Gadarif	landrace	A	yellow	Milo
13	Wad El-Mubarak	southern Gadarif	landrace	P	white	Feterita
14	Dabar Baladi	southern Gadarif	landrace	A	creamy	Milo
15	Wad Ahmed	Medani/northern Gadarif	cultivar	P	white	Feterita
16	Ajab Sedo	northern Gadarif	landrace	P	white	Feterita
17	White Mugud	southern Gadarif	landrace	A	white	Mugud
18	Feterita Arafa	southern Gadarif	landrace	P	white	Feterita
19	Abu Nafain	Eastern Gadarif	landrace	A	creamy	Milo
20	Feterita Rass Girid (Umblail)	southern Gadarif	landrace	P	white	Feterita
21	El-Najada	southern Gadarif	landrace	A	brown	Hegiri
22	Fakai Mustahi	southern Gadarif	landrace	A	creamy	Milo
23	Red Mugud (G)	southern Gadarif	landrace	A	light red	Mugud
24	Dabar Nigiri	southern Gadarif	landrace	A	creamy	Milo
25	Sham Sham	southern Gadarif	landrace	A	brown	Hegiri
26	Ahaimir	southern Gadarif	landrace	A	dark brown	Hegiri
27	White Milo	southern Gadarif	landrace	A	white	Milo
28	Koracola	northern Gadarif	landrace	P	white	Feterita
29	Gadamballia	northern Gadarif	landrace	P	white	Feterita
30	Dabar Habashi	southern Gadarif	landrace	A	creamy	Milo
31	Ingaz	Medani	landrace	A	creamy	Milo
32	Tabat	Medani	landrace	A	white	Milo
33	IS 9830	Medani	landrace	P	white	Feterita
34	Serena	Medani	landrace	A	brown	Hegiri
35	LRB 6	Medani	landrace	A	creamy	Milo
36	Tuzee	Gadarif	landrace	P	white	Feterita
37	SRN 39	Gadarif	cultivar	A	creamy	Milo
38	SAR 1	ICRISAT	ABL	A	creamy	Synthetic
39	SAR 16	ICRISAT	ABL	A	creamy	Synthetic
40	SAR 34	ICRISAT	ABL	A	creamy	Synthetic
41	SAR 35	ICRISAT	ABL	A	creamy	Synthetic
42	SAR 41	ICRISAT	ABL	A	creamy	Synthetic
43	SAR 42	ICRISAT	ABL	A	creamy	Synthetic
44	ICSR 91030	ICRISAT	ABL	A	creamy	Synthetic
45	ICSR 92001	ICRISAT	ABL	A	creamy	Synthetic
46	ICSR 92003	ICRISAT	ABL	A	creamy	Synthetic
47	ICSR 93002	ICRISAT	ABL	A	creamy	Synthetic
48	ICSR 93003	ICRISAT	ABL	A	creamy	Synthetic
49	ICSR 93004	ICRISAT	ABL	A	creamy	Synthetic
50	N765-1-1	Nebraska	PD	A	creamy	Synthetic
51	N770-1-1	Nebraska	PD	P	white	Synthetic
52	N77-1-1	Nebraska	PD	A	creamy	Synthetic
53	N789-1-1	Nebraska	PD	A	creamy	Synthetic
54	N799-1-1	Nebraska	PD	A	creamy	Synthetic
55	AB-5-4-1	Nebraska	PD	A	creamy	Synthetic
56	AB-7-1-1	Nebraska	PD	A	creamy	Synthetic
57	AB-19-3-1	Nebraska	PD	P	white	Synthetic
58	OB-9-3-1	Nebraska	PD	P	white	Synthetic
59	PI 569579	SGB	accession	P	white	Feterita
60	PI 569593	SGB	accession	A	creamy	Milo
61	PI 569582	SGB	accession	A	creamy	Milo
62	PI 569537	SGB	accession	P	white	Feterita
63	PI 569620	SGB	accession	P	white	Feterita
64	PI 569628	SGB	accession	A	creamy	Milo
65	PI 569629	SGB	accession	A	creamy	Milo
66	PI 569630	SGB	accession	A	brown	Hegiri
67	PI 569634	SGB	accession	A	white	Feterita
68	PI 569695	SGB	accession	P	white	Feterita
69	PI 569704	SGB	accession	A	creamy	Milo
70	PI 569706	SGB	accession	A	creamy	Milo
71	PI 569798	SGB	accession	A	creamy	Milo
72	PI 569799	SGB	accession	A	creamy	Milo
73	PI 569802	SGB	accession	A	creamy	Milo
74	PI 569805	SGB	accession	P	white	Feterita
75	PI 569850	SGB	accession	P	white	Feterita
76	PI 569851	SGB	accession	P	white	Feterita
77	PI 569853	SGB	accession	P	white	Feterita
78	PI 569926	SGB	accession	P	white	Feterita

Continued on next page.

Table 1. Continued.

Serial no.	Entry name†	Collection site‡	Category§	Pericarp¶	Grain color	Morphological group#
79	PI 569945	SGB	accession	P	white	Feterita
80	PI 569949	SGB	accession	P	white	Feterita
81	PI 569950	SGB	accession	P	white	Feterita
82	PI 569951	SGB	accession	P	white	Feterita
83	PI 569953	SGB	accession	A	creamy	Milo
84	PI 569976	SGB	accession	A	creamy	Milo
85	PI 570413	SGB	accession	A	creamy	Milo
86	PI 570446	SGB	accession	P	white	Feterita
87	PI 570543	SGB	accession	P	white	Feterita
88	PI 570552	SGB	accession	P	white	Feterita
89	PI 570553	SGB	accession	P	white	Feterita
90	PI 570554	SGB	accession	P	white	Feterita
91	PI 570601	SGB	accession	A	brown	Hegiri
92	PI 570688	SGB	accession	A	brown	Hegiri
93	PI 570698	SGB	accession	A	brown	Hegiri
94	PI 570710	SGB	accession	A	creamy	Milo
95	PI 570767	SGB	accession	P	white	Feterita
96	PI 570784	SGB	accession	P	white	Feterita

† PI, plant introduction.
‡ SGB, Sudan Gene Bank.
§ ABL, advanced breeding line; PD, population derivative.
¶ A, absent; P, present.
# Grouping was based on the grain color and/or presence of pericarp.

Nafain to PI 569704, and from Dwarf White Milo to White Milo). The second main cluster shows significant subclusters that include the following. The cluster from N765-1-1 to N789-1-1 covers the Nebraska derivatives (Nebraska drought population derivatives) and reflects pedigree relationships as well as geographical origin (Nebraska). The cluster from N799-1-1 to PI 569597 covers both Nebraska and gene bank accessions. The clusters from PI 569628 to PI 570553, from PI 569537 to PI 570688, from PI 579853 to PI 570413, and from PI 569706 to PI 569799 cover only gene bank accessions (Sudan collection) with the exception of the landrace Tuzee. The accessions PI 569537 and PI 569776 cluster separately.

Genetic Diversity

In total 117 alleles were detected in 16 SSR loci, with an average of 7.3 alleles per locus. The number of amplification products per primer pair varied from 3 to 18, and the size of the amplified fragments ranged from 95 to 325 bp. The total number of putative alleles at each locus and the observed size range of these alleles are given in Table 2. The PIC values ranged from 0.46 for SB4-72 to 0.87 for SBAGF06. In some cases (e.g., SB-36, SB5-236, SB-72, and SBAGF06) the observed number of alleles was much higher than reported in other publications, which may be due to the larger number and wider geographic origin of accessions used here. For some loci the size range of PCR products obtained in this study is substantially wider than that reported earlier.

The 96 genotypes were analyzed in two ways; first, based on genetic improvement status (landraces and improved cultivars from Gadarif and Medani, Sudan, and ICRISAT ABLs versus nonimproved gene bank accessions and Nebraska derivatives), which are further differentiated into subgroups. In this respect, the DI was found to be 0.58 for the 49 landraces, improved cultivars, and ABLs; 0.63 for the 37 landraces and improved cultivars from Gadarif and Medani; and 0.59 for the 31 Gadarif landraces. Diversity index for the ICRSIAT ABLs was 0.39. Diversity index for the 47 gene bank accessions (PI accessions) and Nebraska derivatives was 0.52. Gene bank accessions and Nebraska derivatives could further be partitioned into gene bank accessions with a DI estimated at 0.49 and Nebraska derivatives that revealed a DI of 0.42 (Table 3).

The second analysis is based on the morphological groups, and DI estimates were as follows: 0.65 within Feterita, 0.71 within Milo, 0.63 within the Synthetics, 0.68 within Hegiri, and 0.47 within the Mugud group (Table 4).

DISCUSSION

Genetic Relationships Revealed by UPGMA-Clustering

The 96 sorghum accessions analyzed had unique fingerprints. All SSR markers were polymorphic, confirming their usefulness for genetic analysis. The range of

Table 2. Microsatellite primer sets, linkage group, and repeat motif used for the study: their PCR product range, number of alleles per locus, and polymorphic information content (PIC) on 96 sorghum genotypes.

SSR ID	LG†	Repeat type	Observed PCR product range	Number of alleles	PIC
SB6-34	I	$[(AC)/(CG)]_{15}$	218–243	5	0.56
SB5-236	G	$(AG)_{20}$	173–193	7	0.77
SB4-121	D	$(AC)_{14}$	209–216	3	0.46
SB6-57	C	$(AG)_{18}$	292–325	6	0.67
SB1-10	D	$(AG)_{27}$	230–311	9	0.81
SB4-15	E	$(AG)_{16}$	106–132	5	0.65
SB4-32	E	$(AG)_{15}$	172–223	9	0.76
SB6-342	A	$(AC)_{25}$	281–300	4	0.71
SB1-1	H	$(AG)_{16}$	251–270	7	0.71
SB5-206	E	$(AC)_{13}/(AG)_{20}$	100–145	9	0.74
SB4-72	F	$(AG)_{16}$	181–208	4	0.57
SB6-84	B	$(AG)_{14}$	162–224	12	0.78
SBAGB02	A	$(AG)_{35}$	95–145	7	0.62
SBAGF06	A	$(AG)_{35}$	100–179	18	0.87
SBAGH04	F	$(AG)_{39}$	127–155	6	0.75
SBKAFGKI	J	$(ACA)_{9}$	140–165	6	0.70

† LG, linkage group according to Taramino et al. (1997).

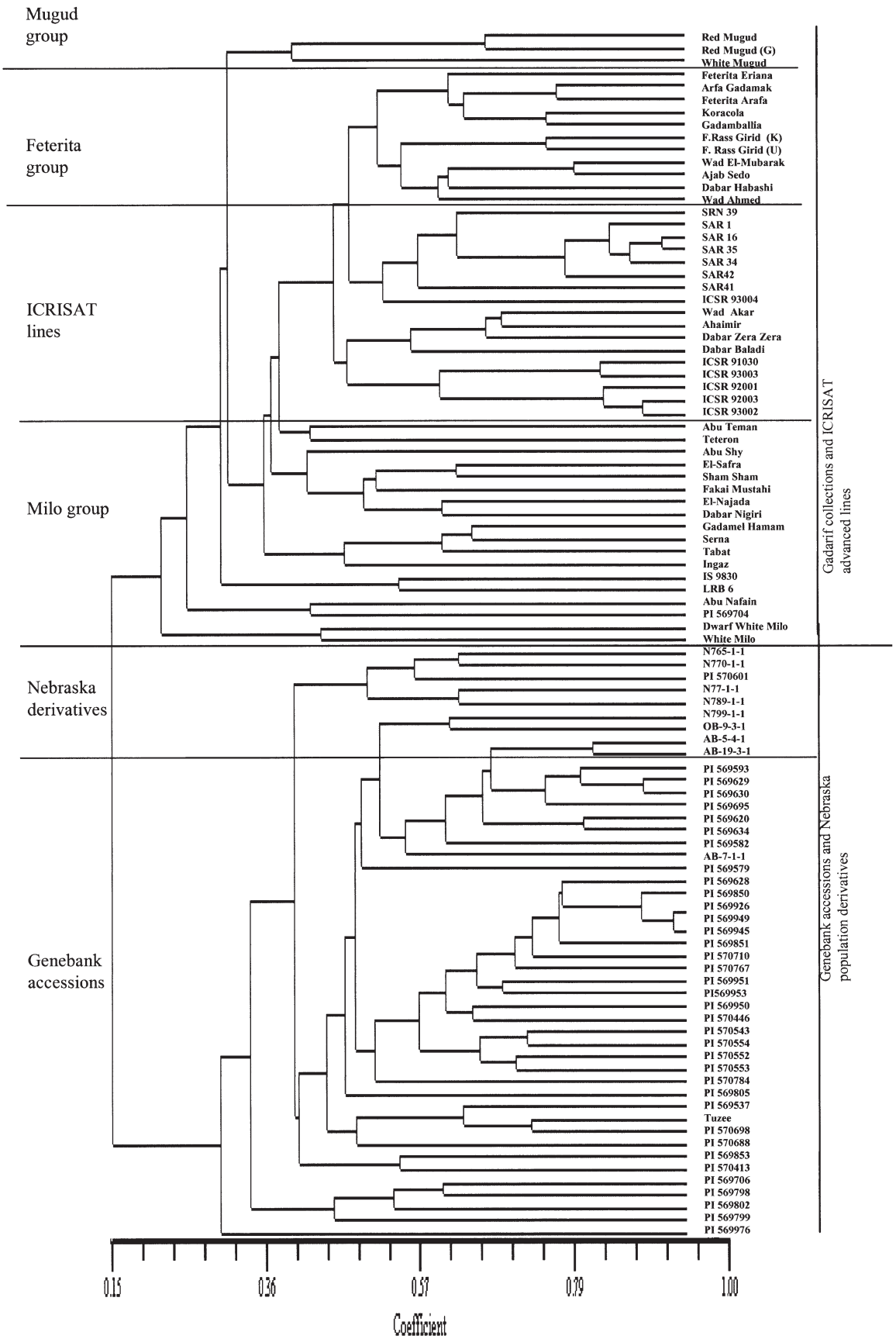

Fig. 1. Dendrogram generated by UPGMA cluster analysis showing the relationships among 96 sorghum accessions based on Nei and Li similarity estimates (Nei and Li, 1979). (Note: The two accessions PI569926 and PI569949 belong to the same subcluster.)

Table 3. Genetic diversity within groups of sorghum accessions from Sudan, ICRISAT, and the USA.

Category	No. of accessions	Diversity index (DI)	Alleles/locus
1. Landraces, improved cultivars, and advanced breeding lines	49	0.58	5.9
i. Landraces (35) and improved cultivars (2) from Gadarif and Medani	37	0.63	5.4
– Gadarif landraces (30) and cultivars (1)	31	0.59	5.3
– Medani landraces (5) and cultivars (1)	6	0.62	2.9
ii. ICRISAT ABL	12	0.39	2.7
2. Gene bank accessions and Nebraska derivatives	47	0.52	5.5
– Gene bank accessions (PI accessions)	38	0.49	5.1
– Nebraska derivatives	9	0.42	3.5
All accessions	96	0.71	7.3

Table 4. Genetic diversity within morphological groups of sorghum germplasm from Sudan, ICRISAT, and the USA.

Category	No. of accessions	Diversity index (DI)	Alleles/locus
Feterita	35	0.65	5.7
Milo	29	0.71	6.1
Synthetic	20	0.63	4.4
Hegiri	9	0.68	3.7
Mugud	3	0.47	1.9
All	96	0.71	

GS was very wide (0.0–0.91) and resulted in low mean GS (0.30). Similar results have been reported by Uptmoor et al. (2003). The use of SSRs with many alleles per locus enables uniquely fingerprinting a large number of accessions by relatively few loci (cf., McCouch et al., 1997). The construction of stable dendrograms, objectively reflecting genetic relationships, mainly depends on the number of alleles analyzed (Zhang et al., 2002). Doyle et al. (1998) suggested that inferring relationships from microsatellites could be problematic particularly among species but also within species because of the hypervariability of microsatellites. However, Matsuoka et al. (2002) found that dendrograms based on SSRs in maize (*Zea mays* L.) were in good agreement with expected genetic relationships. The results on genetic relatedness and genetic diversity within sorghum accessions from Sudan, ICRISAT, and the USA depict a clear separation between improved cultivars and gene bank accessions (Fig. 1). The clustering of the three Mugud landraces in one cluster reflects morphological relationships. These three cultivars are from the same morphological group (Mugud) and are morphologically difficult to differentiate. In part it also holds true for the Feterita group where 83% of the genotypes clustered together. However, the other 17% of the accessions were scattered in between other morphological groups within the first main cluster of Fig. 1. The two landraces Feterita Rass Girid (Kassab) and Feterita Rass Girid (Umbelail) fall in the same subcluster, which indicates that they are closely related and might have the same genetic background. The clustering of ICRISAT SAR and ICSR series in the same cluster together with SRN 39 reflect pedigree relationships as well as common geographical origin (ICRISAT). The presence of the cultivar SRN 39 in the ICRISAT group was expected, because this cultivar is originally from ICRISAT. SRN 39 was developed for resistance to *Striga hermonthica* (Del.) Benth. and later introduced to Sudan and released as striga-resistant cultivar (Sorghum National Program Sudan, A.G.T. Babiker, personal communication, 2001). Furthermore, the cluster of the Nebraska drought population derivatives reflects pedigree relationships and geographic origin (Nebraska). The clustering of the landrace Tuzee with PI 579853 can be explained by the fact that both landraces belong to the Feterita group; their clustering might reflect the close relationship between Tuzee and this accession. The cluster of the gene bank accessions (Sudan collections) is evidence for geographic origin. These results are in agreement with comparable results on barley genotypes published by Ordon et al. (1997). The grouping of all Feterita, Mugud, Synthetic, and Milo types separately into different subclusters confirms to morphological characters. Accordingly, these results suggest that the dendrogram based on the estimated GS reflects pedigree and morphological relationships as reported by Ahnert et al. (1996) for sorghum inbred lines, as well as geographic and adaptation zones as reported by Hormaza (2002) on apricot (*Prunus armeniaca* L.) accessions. Simioniuc et al. (2002) concluded that dendrograms based on GS estimated for pea (*Pisum sativum* L.) cultivars using AFLPs and RAPDs only reflect pedigree data to some extent. However, Ayana et al. (2000) reported a weak differentiation of Ethiopian and Eritrean sorghum accessions according to both agro-ecological adaptation zones and regions of origin. Similar results were found by Uptmoor et al. (2003). According to Graner et al. (1994), a better knowledge and measurement of GS of accessions could help to maintain genetic diversity. Therefore, information about GS among germplasm would be helpful for plant breeders to choose diverse parents for crossing which may lead to transgressive segregates for quantitative traits, promoting further breeding progress. However, the process of parent selection may be enhanced in the future by high throughput marker system facilitating efficient haplotyping and procedures of association genetics (Powell and Russell, 2000).

Genetic Diversity within Cultivated Sorghum

A critical premise for using markers to assess genetic diversity is the number of loci studied and their adequacy in representing the whole genome (Akkaya et al., 1995; Pejic et al., 1998). This study tried to meet the latter requirement in selecting the SSR probes that cover all of the 10 linkage groups, A through J (Table 2). The 16 primer pairs used in this study generated multiple alleles across the complete range of genotypes. The 96 accessions that were included in this study encompass a relatively broad array of germplasm diversity. For example, the set of germplasm includes accessions from different geographic areas (Sudan, India, and the USA). Germplasm groups that are represented include Feterita, Milo, Hegiri, and Synthetic. Within these groups,

there is a wide variation with respect to kernel color, plant height, and maturity.

The overall DI in this study was 0.71. High genetic diversity was found within a group of 37 Sudanese landraces and improved cultivars (DI = 0.63). This high variability could be due to the different morphological groups covered. But it also has to be taken into account that Sudan is considered to be part of the origin of diversity for sorghum (Doggett, 1988). The results of this study are in agreement with Djè et al. (2000) who also found high genetic diversities in accessions belonging to the race bicolor and/or originating from Eastern Africa. When accessions were analyzed based on morphological groups, DI ranged from 0.47 within the Mugud group to 0.71 within the Milo group. The low DI value within the Mugud type group may be due to the small sample size (three landraces) and the close relationship that is revealed by the UPGMA analysis. It is interesting to note that when ICRISAT ABL and Nebraska population derivatives were analyzed separately, DI was estimated at 0.39 and 0.42, respectively, and when analyzed together as a Synthetic group the DI increased to 0.63. The same holds true for the Feterita group from El Gadarif state (DI = 0.37) and that from the gene bank (DI = 0.50). When the two groups are combined the DI rose to 0.65. When all genotypes analyzed are taken together the overall DI increased to 0.71, which is indicative of the large genetic diversity present within these sorghum accessions. Many studies suggested that accessions from Eastern Africa are highly variable (Aldrich and Doebley, 1992; Deu et al., 1994). Morden et al. (1989) concluded that the genetic variation is more closely associated with geographic origin than racial classification. Taramino et al. (1997) used 13 SSRs to reveal moderate to high levels of diversity among a group of nine sorghum lines of different racial classification and from different geographic origin. The low genetic diversity within ICRISAT lines (DI = 0.39) and their close relationship revealed by UPGMA-cluster analysis suggest that these lines share a common genetic background. The same holds true for the Nebraska population derivatives (DI = 0.42). The gene bank accessions revealed relatively low genetic diversity (DI = 0.49) and their clustering can be explained by the fact that these accessions were selected out of 600 original accessions screened under Sudan drought conditions, based on morphological characters such as earliness and plant height. Therefore, genotyping could result in a decreased DI within these accessions. However, the diversity for the 31 landraces and improved cultivars, which were collected in Gadarif, was only intermediate (DI = 0.59), which could be ascribed to a small adaptation zone. Uptmoor et al. (2003) found low DI for landraces derived from the Northern Province in South Africa. However, Djè et al. (2000) estimated high genetic diversity for 25 sorghum landraces derived from a restricted area of northwestern Morocco using three SSRs. The high genetic diversity estimated on all accessions (0.71) is comparable to the results of Uptmoor et al. (2003), Djè et al. (2000), and Grenier et al. (2000), who found values of 0.665, 0.897, and 0.80, respectively. In our study sorghum accessions derived from different parts of the world were included and the SSRs used generated 7.3 alleles per locus on average for the sample analyzed. Uptmoor et al. (2003) detected 8.68 alleles per locus on average when analyzing 46 sorghum accessions with 25 SSRs. In contrast, Djè et al. (2000) detected as many as 19.2 alleles per locus on average for 25 accessions, and Ghebru et al. (2002) detected 13.9 alleles per locus for 28 accessions analyzed with 15 SSRs.

The variability in the number of alleles per locus (3–18) may result from different locus specific mutation rates (Estoup et al., 2002) and reflects strong differences in allelic diversity between SSR loci, which affects estimating genetic diversity since the DI, according to Nei (1973), depends on both the number of alleles per locus and the respective allele frequency (McCouch et al., 1997). Polymorphic information content values of Table 2 represent the variation in locus specific genetic diversity for the sorghum accessions used in this study. Besides locus specific mutation rates, the number of alleles per locus and gene diversity can be affected (i.e., reduced) by size homoplasy which occurs when different copies of a locus are identical in state, although they are not identical by descent (Estoup et al., 2002). However, microsatellites are typically multi-allelic markers (Matsuoka et al., 2002) with heterozygosity values much higher than those of RFLPs (McCouch et al., 1997). Accordingly, different authors have shown that microsatellites with three or more alleles per locus are more common than those with less than three alleles per locus in sorghum (Taramino et al., 1997; Kong et al., 2000) and in maize (Matsuoka et al., 2002). As genetic diversity is calculated as arithmetic mean of locus specific diversities, the set of primers used for analyzing genetic diversity should represent the variation among loci as good as possible.

This study provides a first detailed analysis and quantification of genetic diversity in Sudanese sorghum germplasm. The data also support the findings that microsatellites can be effectively used for studying genetic diversity in sorghum. The SSR data proved to be useful in identifying genetic relationships among a diverse collection of accessions, with the majority of the accessions clustering in concordance with pedigree relationships and/or morphological information, adaptation zones and/or geographic origin.

Molecular markers have an important role to play in many aspects of genetic resource conservation (Karp et al., 1997). The choice of diverse parents for crossing based on molecular information would be helpful for plant breeders. A breeding strategy may involve choosing high-yielding parents that possess many random genetic differences in the hope of finding an increased number of transgressive recombinants (Tinker et al., 1993; Graner et al., 1994). This information in connection with results from field experiments under drought stress conditions (data not shown) could be very useful for sorghum breeding programs in Sudan. Finally, combining the molecular information and morphological traits is expected to enhance the process of incorporation of many desirable genes into well-adapted cultivars and landraces.

ACKNOWLEDGMENTS

We thank Mr. Roland Kürschner for taking care of the plants in the greenhouse. Many thanks to Deutscher Akademischer Austauschdienst (DAAD) for its contribution to the successful completion of this research work by offering a scholarship grant for one of us (AHA).

REFERENCES

Ahmed, M.M., J.H. Sanders, and W.T. Nell. 2000. New sorghum and millet cultivar introduction in Sub-Saharan Africa: Impacts and research agenda. Agric. Syst. 64:55–65.

Ahnert, D., M. Lee, D.F. Austin, C. Livini, S.J. Openshaw, J.S.C. Smith, K. Proter, and G. Dalton. 1996. Genetic diversity among elite sorghum inbred lines assessed with DNA markers and pedigree information. Crop Sci. 36:1385–1392.

Akkaya, M.S., R.C. Shoemaker, J.E. Sprecht, A.A. Bhagwat, and P.B. Cergan. 1995. Integration of simple sequence repeat DNA markers into a soybean linkage map. Crop Sci. 35:1439–1445.

Aldrich, P.R., and J. Doebley. 1992. Restriction fragment variation in the nuclear and chloroplast genomes of cultivated and wild *Sorghum bicolor*. Theor. Appl. Genet. 85:293–302.

Ayana, A., and E. Bekele. 1999. Multivariate analysis of morphological variation in sorghum (*Sorghum bicolor* (L.) Moench) germplasm from Ethiopia and Eritrea. Genet. Res. Crop Evol. 46: 378–384.

Ayana, A., T. Bryngelsson, and E. Bekele. 2000. Genetic variation of Ethiopian and Eritrean sorghum (*Sorghum bicolor* (L.) Moench) germplasm assessed by random amplified polymorphic DNA (RAPD). Genet. Res. Crop Evol. 47:471–481.

Botstein, D., R.L. White, M. Skolnick, and R.W. Davis. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. 32:314–331.

Brown, S.M., M.S. Hopkins, S.E. Mitchell, M.L. Senior, T.Y. Wang, R.R. Duncan, F. Gonzalez-Candelas, and S. Kresovich. 1996. Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum (*Sorghum bicolor* (L.) Monech). Theor. Appl. Genet. 93:190–198.

Cregan, P.B., I. Jarvik, A.L. Bush, R.C. Shoemaker, K.G. Lark, A.L. Khler, N. Kaya, T.T. Vantoai, D.G. Lohnes, J. Chung, and J.E. Specht. 1999. An integrated genetic linkage map of the soybean genome. Crop Sci. 39:1464–1490.

Davila, J.A., M.O. Sanchez de la Hoz, Y. Loarce, and E. Ferrer. 1998. DNA and coefficients of parentage to determine genetic relationships in barley. Genome 41:477–486.

Dean, R.E., J.A. Dahlberg, M.S. Hopkins, C.V. Mitchell, and S. Kresovich. 1999. Genetic redundancy and diversity among 'orange' accessions in the U.S. national sorghum collection as assessed with simple sequence repeat (SSR) markers. Crop Sci. 39:1215–1221.

Deu, M., D. Gonzalenz-de-leon, J.C. Glasmann, I. Degremont, J. Chantereau, C. Lanaud, and P. Hamon. 1994. RFLP diversity in cultivated sorghum in relation to racial differential. Theor. Appl. Genet. 88:838–844.

Djè, Y., M. Heurtz, C. Lefèbre, and X. Vekemans. 2000. Assessment of genetic diversity within and among germplasm accessions in cultivated sorghum using microsatellite markers. Theor. Appl. Genet. 100:918–925.

Doggett, H. 1988. Sorghum. 2nd ed. Longman Scientific and Technical, London.

Doyle, J.D., and J.L. Doyle. 1990. Isolation of plant DNA from fresh tissue. BRL Focus 12:13–15.

Doyle, J.J., M. Morgante, S.V. Tingey, and W. Powell. 1998. Size homoplasy in chloroplast microsatellites of wild relatives of soybean (*Glycine* subgenus *Glycine*). Mol. Biol. Evol. 15:215–218.

Estoup, A., P. Jarne, and J.M. Cornuet. 2002. Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. Mol. Ecol. 11:1591–1604.

Ghebru, B., R.J. Schmidt, and J.L. Bennetzen. 2002. Genetic diversity of Eritrean sorghum landraces assessed with simple sequence repeat (SSR) markers. Theor. Appl. Genet. 105:229–236.

Goulão, L., L. Cabrita, C.M. Oliverita, and J.M. Leitao. 2001. Comparing RAPD and AFLP™ analysis in discrimination and estimation of genetic similarities among apple (*Malus domestica* Borkh.) cultivars. Euphytica 119:259–270.

Graner, A., W.F. Ludwig, and A.E. Melchinger. 1994. Relationships among European barley germplasm. Crop Sci. 34:1199–1205.

Grenier, C., M. Deu, S. Kresovich, P.J. Bramel-Cox, and P. Hamon. 2000. Assessment of genetic diversity in three subsets constituted from ICRISAT sorghum collection using random vs non-random sampling procedures B. Using molecular markers. Theor. Appl. Genet. 101:197–202.

Gupta, P.K., and R.K. Varshney. 2000. The development and use of microsatellite markers for genetic analysis and plant breeding with the emphasis on bread wheat. Euphytica 113:163–185.

Hormaza, J.I. 2002. Molecular characterisation and similarity relationships among apricot (*Prunus armeniaca* L.) genotypes using simple sequence repeats. Theor. Appl. Genet. 104:321–328.

Karp, A., P.G. Isaac, and D.S. Ingram. 1998. Molecular tools for screening biodiversity. Chapman and Hall, London.

Karp, A., S. Kresovich, K.V. Bhat, W.G. Ayad, and T. Hodgkin. 1997. Molecular tools in plant genetic resources conservation: A guide to the technologies. Tech. Bull. No. 2. IPGRI, Rome.

Kong, I., J. Dong, and G.E. Hart. 2000. Characteristics, linkage map positions, and allelic differentiation of *Sorghum bicolour* (L.) Moench DNA simple-sequence repeats (SSRs). Theor. Appl. Genet. 101: 438–448.

Mann, J.A., C.T. Kimber, and F.R. Miller. 1983. The origin and early cultivation of sorghums in Africa. Bull. No. 1454. Texas Agric. Exp. Stn., College Station, TX.

Mantel, M. 1967. The detection of disease clustering and generalised regression approach. Cancer Res. 27:209–220.

Matsuoka, Y., S.E. Mitchell, S. Kresovich, M. Goodmann, and J. Doebley. 2002. Microsatellites in *Zea*—variability, patterns of mutations, and use for evolutionary studies. Theor. Appl. Genet. 104:436–450.

McCouch, S.R., X. Chen, O. Panaud, S. Temnykh, Y. Xu, Y.G. Cho, N. Huang, T. Ishii, and M. Blair. 1997. Microsatellite marker development, mapping and application in rice genetics and breeding. Plant Mol. Biol. 35:89–99.

Morden, C., J.F. Doebley, and K.F. Schertz. 1989. Allozyme variation in old world races of *Sorghum bicolor* (Poaceae). Am. J. Bot. 76: 247–255.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. USA 70:3321–3323.

Nei, M., and W.H. Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA 76:5269–5273.

Ordon, F., A. Schiemann, and W. Friedt. 1997. Assessment of the genetic relatedness of barley accessions (*Hordeum vulgare* s.l.) resistant to soil-borne Mosaic inducing viruses (Ba MMV, BaYMV-2) using RAPDs. Theor. Appl. Genet. 94:325–330.

Pejic, I., P. Ajmone-Marsan, M. Morgante, V. Kozumplik, P.Castiglioni, G.Taramino, and M. Motto. 1998. Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs. Theor. Appl. Genet. 97:1248–1255.

Powell, W., and J.R. Russell. 2000. Molecular analysis of barley diversity. p. 29–31. *In* S. Logue (ed.) Proc. 8th Int. Barley Genet. Symp., Adelaide, Australia. 22–27 Oct. 2000. Dep. Plant Sci., Adelaide Univ., Glen Osmond, Australia.

Ribaut, J.M., and D. Hoisington. 1998. Marker-assisted selection: New tools and strategies. Trends Plant Sci. 3:236–239.

Rohlf, F.J. 2000. NTSYS-pc numerical taxonomy and multivariate analysis system, version 2.1. Exeter Publ., New York.

Simioniuc, D., R. Uptmoor, W. Friedt, and F. Ordon. 2002. Genetic diversity and relationships among pea cultivars (*Pisum sativum* L.) revealed by RAPDs and AFLPs. Plant Breed. 121:429–435.

Smith, J.S.C., and O.S. Smith. 1989. The description and assessment of distance between lines of maize. II. The utility of morphological, biochemical, and genetic descriptors and a scheme for the testing of distinctiveness between inbred lines. Maydica 34:151–161.

Smith, J.S.C., and O.S. Smith. 1992. Fingerprinting crop varieties. Adv. Agron. 47:85–140.

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

Taramino, G., R. Tarchini, S. Ferrario, M. Lee, and M.E. Pe. 1997. Characterisation and mapping of simple sequence repeats (SSRs) in *Sorghum bicolor*. Theor. Appl. Genet. 95:66–72.

Tautz, D. 1989. Hyper-variability of simple sequences as a general source of polymorphic DNA markers. Nucleic Acids Res. 17:6463–6471.

Tinker, N.A., M.G. Fortin, and D.E. Mather. 1993. Random amplified polymorphic DNA and pedigree relationships in spring barley. Theor. Appl. Genet. 95:66–72.

Uptmoor, R., W. Wenzel, W. Friedt, G. Donaldson, K. Ayisi, and F. Ordon. 2003. Comparative analysis on the genetic relatedness of sorghum bicolor accessions from South Africa by RAPDs, AFLPs and SSRs. Theor. Appl. Genet. 106:1316–1325.

Williams, J.G.K., A.R. Kubelik, K.J. Livak, J.A. Rafalski, and S.V. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18:6531–6535.

Zabeau, M., and P. Vos. 1993. Selective restriction fragment amplification: A general method for DNA fingerprinting. European patent application no. 92402639.7. Publication no. 0 534 858 A1.

Zhang, X.Y., C.W. Li, L.F.H.M. Wang, G.X. You, and Y.S. Dong. 2002. An estimation of minimum number of SSR alleles needed to reveal genetic relationships in wheat varieties. I. Information from large-scale planted varieties and cornerstone breeding parents in Chinese wheat improvement and production. Theor. Appl. Genet. 106:112–117.