A global view of genetic diversity in cultivated sorghums using a core collection

M. Deu, F. Rattunde, and J. Chantereau

Abstract: We report here an analysis of the structure of genetic diversity in cultivated sorghums. A core collection of 210 landraces representative of race, latitude of origin, response to day length, and production system was analysed with 74 RFLP probes dispersed throughout the genome. Multivariate analyses showed the specificity of the subrace guinea margaritiferum, as well as the geographical and racial pattern of genetic diversity. Neighbour-joining analysis revealed a clear differentiation between northern and southern equatorial African accessions. The presence of Asian accessions in these 2 major geographical poles for sorghum evolution indicated two introductions of sorghum into Asia. Morphological race also influenced the pattern of sorghum genetic diversity. A single predominant race was identified in 8 of 10 clusters of accessions, i.e., 1 kafir, 1 durra, 4 guinea, and 2 caudatum clusters. Guinea sorghums, with the exception of accessions in the margaritiferum subrace, clustered in 3 geographical groups, i.e., western African, southern African, and Asian guinea clusters; the latter two appeared more closely related. Caudatum were mainly distributed in 2 clusters, the African Great Lakes caudatum cluster and those African caudatum originating from other African regions. This last differentiation appears related to contrasting photoperiod responses. These results aid in the optimization of sampling accessions for introgression in breeding programs.

Key words: sorghum, core collection, genetic diversity, RFLP.

Résumé : Une analyse approfondie de la structuration de la diversité génétique des sorghos cultivés est présentée. Une core collection de 210 variétés représentatives de la diversité des races, latitudes, réponses à la durée du jour et systèmes de culture a été analysée au moyen de 74 sondes PLFR réparties sur le génome. Les analyses multivariées ont montré la spécificité des guinea margaritiferum ainsi qu'une structuration géographique et raciale de la diversité génétique. L'analyse « neighbour joining » a montré une différenciation entre variétés africaines du nord et du sud de l'équateur. La présence de variétés asiatiques dans chacun de ces grands pôles géographiques d'évolution des sorghos indique 2 introductions en Asie. La race est aussi un des facteurs impliqués dans la structuration de la diversité génétique des sorghos. Parmi les 10 groupes de variétés identifiés, 8 groupes sont constitués majoritairement d'accessions appartenant à la même race, soit 1 groupe de kafir, 1 de durra, 4 groupes de guinea et 2 groupes de caudatum. Les sorghos guinea, à l'exception des guinea margaritiferum, se répartissent dans trois groupes géographiques, guinea d'Afrique de l'ouest, d'Afrique Australe et d'Asie, les 2 derniers étant les plus proches génétiquement. Les caudatum sont principalement répartis dans 2 groupes, le groupe des variétés originaires de la région des grands lacs d'Afrique et celui constitué par les autres caudatum africains. Cette dernière différenciation apparaît liée à des différences de comportements vis-à-vis de la photopériode. Ces résultats permettent d'optimiser le choix des variétés à utiliser dans un programme de sélection.

Mots clés : Sorghum, core collection, diversité génétique, PLFR.

Introduction

Sorghum is a major staple food and fodder crop in tropical and semi-tropical Africa and Asia (Doggett 1988). Cultivated sorghums (*Sorghum bicolor* subsp. *bicolor*) exhibit great phenotypic variability. Five major races (bicolor, caudatum, durra, guinea, and kafir) and 10 intermediate races, corresponding to the pairwise combination of major races, are identified according to panicle and spikelet morphology (Harlan and de Wet 1972).

Many studies have examined patterns of genetic diversity among sorghum accessions from ex situ germplasm collections using RFLP or RAPD markers (Aldrich and Doebley 1992; Tao et al. 1993; Deu et al. 1994; Cui et al. 1995; de

Received 21 March 2005. Accepted 13 September 2005. Published on the NRC Research Press Web site at http://genome.nrc.ca on 15 February 2006.

Corresponding Editor: J.P. Gustafson.

M. Deu.[1] Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), UMR 1096-PIA, TA 40/03, Avenue Agropolis, 34398 Montpellier CEDEX 5, France.
F. Rattunde. International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), B.P. 320, Bamako, Mali.
J. Chantereau. Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), TA 70/03, Avenue Agropolis, 34398 Montpellier CEDEX 5, France.

[1]Corresponding author (e-mail: monique.deu@cirad.fr).

 doi:10.1139/G05-092

Oliveira et al. 1996; Menkir et al. 1997) and more recently using SSR markers (Djé et al. 2000; Folkertsma et al. 2005). Most studies have indicated that geographic origin and (or) racial classification are associated with the organization of genetic diversity. However, RAPD markers alone failed to separate accessions into discrete racial or geographic groups in a large collection of sorghums (Menkir et al. 1997). This was probably due to different factors, such as the nature of evolutionary mechanisms underlying the variation measured by the different markers, the repetitive sequence revealed by some RAPD markers, the lack of allelism of RAPD bands of similar sizes, and their unequal distribution in the genome (N'Goran et al. 1994; Powell et al. 1996; Noli et al. 1997). Similarly, differentiation among races or geographic origins of sorghums was not revealed by Djé et al. (2000) using a limited number of SSR markers.

Thus, to date, there is no satisfactory assessment of the global structure of diversity of sorghums worldwide and based on a large number of markers. Previous studies were mostly based on relatively small numbers of accessions. Furthermore, sampling was not conducted on large collections appropriately stratified to assure representation of the full genetic diversity of cultivated sorghums worldwide. A useful approach to appropriately sample the diversity of world sorghum collections was recently developed by Grenier et al. (2001*a*). This approach relies on stratification of the world collection based on both a 2-way matrix of racial classification and latitude of origin and clustering on empirical assessments for photoperiod sensitivity, one of the principal determinants of adaptation. Knowledge of the major sorghum crop systems and the landraces prevalent in each could be used to further enhance the choice of accessions to best represent the full genetic diversity. For example, very contrasting sorghums, with contrasting adaptation are grown in the rainy season and post-rainy season within the same geographic region. In addition, knowledge of patterns of genetic diversity gained from previous studies could be considered to assure more complete representation of global diversity.

To enable a more comprehensive assessment of the structure of genetic diversity in tropical and sub-tropical landrace sorghums, this study established a large core collection based on the stratification technique developed by Grenier et al. (2001*a*) and complemented with knowledge of major crop systems and previously reported patterns of genetic diversity. A thorough analysis of the structure of global genetic diversity of cultivated sorghums was conducted using this worldwide core collection. A large number of RFLP markers have been applied to finely analyse the 210 accessions entering the core collection to (*i*) characterize the levels and patterns of diversity within and among clusters of accessions and (*ii*) examine the different factors (i.e., racial and geographical) involved in sorghum genetic differentiation. The manageable size of this core collection better permits the exploration of the range of variation of numerous and complex traits of agricultural interest, such as photoperiod sensitivity in trials conducted in different locations and (or) years owing to its reduced management and evaluation costs. Lastly, the characterization with molecular markers and the assessment of the extent and distribution of genetic diversity in such a collection permit a more relevant choice of the accessions suitable either to optimize a breeding program or to perform association studies between markers and traits based on linkage disequilibrium.

Materials and methods

Sampling and characterization of accessions

The 210 sorghum accessions assembled for this study were chosen to provide representation of cultivated landrace sorghums from the whole world, with sampling based on race classification, latitude of origin, response to day length, and form of cultivation. An extensive sampling exercise to develop core collections from the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) sorghum collection was conducted by Grenier et al. (2001*a*, 2001*b*). Three core collections were obtained after stratification of the ICRISAT landraces collection into 4 clusters based on photoperiod sensitivity. The majority of sorghum accessions examined in this study were derived from the logarithmic sampling "L sorghum core" of 2247 accessions obtained by Grenier et al. (2001*b*). The L sorghum core was used because it best represented the full range of genetic diversity among the core collections. Subsequent random sampling (10%) within the L sorghum core resulted in a subset of 225 accessions (Grenier et al. 2000). We retained 128 of the 225 accessions, excluding the majority of the kafir race and intermediate kafir forms, which were over represented in the L subset and known for their low level of genetic and phenotypic diversities (Morden et al. 1989; Menkir et al. 1997; Djé et al. 2000). Supplementary accessions were added to this collection to provide more complete representation of diverse sorghum landraces, taking advantage of prior information on specific sets of germplasm and assuring appropriate representation of the variability of the guinea race, bicolor race, and transplanted sorghums of western Africa. Thus, 49 landrace accessions from a collection representing racial and geographical diversity (Deu et al. 1994), 16 from the ICRISAT world sorghum collection, and 17 from the CIRAD collection were added. Our final core collection, upon which this study is based, comprised 210 sorghum landrace accessions (with 179 and 31 accessions, respectively, from the ICRISAT and CIRAD collections).

Racial characterization based on panicle and spikelet morphology was performed in accordance to the Harlan and de Wet classification (1972). Additionally, the taxonomic characterization proposed by Snowden (1936) and recognised by de Wet and Harlan (1972) was used to distinguish the margaritiferum subrace from other guinea landraces. The accessions of the core collection were grown and independently observed for racial characterization in 3 different sites; ICRISAT Patancheru (1996–1998, India) for all except the 31 CIRAD accessions, ICRISAT Samanko, Mali, 2000, and CIRAD Montpellier, France, 2001.

RFLP analysis

A set of 90 probes was selected, according to their location on the CIRAD sorghum reference genetic map (Dufour et al. 1997; Boivin et al. 1999; Ventelon et al. 2001), to provide good coverage of the genome. These probes were rice (RZ prefix) and oat (CDO prefix) cDNA probes, maize genomic and cDNA probes (BNL, CSU, UMC prefixes),

sorghum cDNA probes (SbRPG prefix), and sugarcane genomic probes (SSCIR prefix).

The 210 accessions were maintained by self-pollination. Selfed seed from a variable number of plants was bulked for each accession, and the number of generations of multiplication varied by accession. DNA was extracted from a mixture of leaves collected on 5 plants for each accession so as to enable assessment of any heterogeneity within accessions. The RFLP procedures were carried out as previously described (Deu et al. 1994, 1995). Four restriction endonucleases, *Dra*I, *Eco*RI, *Eco*RV, and *Hin*dIII, were used to digest 5 µg of DNA from each sample. Each probe was used in combination with a unique restriction enzyme.

Data analyses

Genetic diversity was estimated within each race (or cluster defined according to the NJ tree, see below) using 5 statistics: the percentage of polymorphic loci (using the 0.99 criterion), the total number of alleles, the number of rare alleles (frequency ≤1% in the total collection), the mean number of alleles per polymorphic loci, and the average genetic diversity index. The genetic diversity index was computed for each locus according to Nei's unbiased estimator (1978) as

$$H = 2n(1 - \sum P_i^2) / (2n - 1)$$

where P_i is the frequency of the ith allele and n is the population size. These computations were performed with Genetix software, v. 4.04 (Belkhir et al. 2002).

A factor analysis of correspondences (FAC) method was performed with Genetix software to depict the organisation of molecular variation and to identify the most discriminative markers.

The presence or absence of a band was scored as 1 or 0, respectively, for each probe enzyme combination detecting polymorphism. The proportion of shared fragments, proposed by Nei and Li (1979)

$$F = 2 \times \left(\frac{\text{number of shared fragments}}{\text{number of total fragments}} \right)$$

was calculated for all pairwise comparisons. The index

$$D = 1 - F$$

was calculated to obtain the dissimilarity matrix between all accessions. A neighbour-joining (NJ) analysis was performed on the dissimilarity matrix to determine the aggregation of the accessions into clusters. Support for clustering was determined by a bootstrap procedure applied on 219 RFLP bands (1000 replications). Furthermore, to test the goodness of fit for the clustering to represent the similarities between accessions, the cophenetic value matrix was calculated for the resulting tree matrix and compared with the original dissimilarity matrix. This comparison produces a cophenetic correlation coefficient, generally varying between 0.6 and 0.95 (Sneath and Sokal 1973). These analyses were conducted with Darwin software (Perrier et al. 2003).

Results

Racial characterization

The racial classification of accessions was consistent for 168 accessions, but showed some discrepancies for the remaining 42 across the assessments in France, Mali, and India (Table 1). Most discrepancies were minor, for example, an accession characterized either as a basic race or a related intermediate form as in the case of IS 3771. When discrepancies did occur, accessions were classified according to the most consensual racial characterization. Major discrepancies occurred only for IS 7287, which did not conform to the ICRISAT grain description and was classified as X, and IS 9331 and IS 16186, for which interpretations of morphological characters by ICRISAT and CIRAD gene bank curators differed. The racial distribution of the accessions in this study, based on the consensual classification, was 24 bicolors; 62 guineas, of which 14 are margaritiferum; 44 caudatums; 29 durras; 19 kafirs; and 32 intermediate forms.

Marker polymorphism

Ninety probes were tested on the 210 accessions. A total of 16 probes were discarded because they provided either a poor level of polymorphism or a complex pattern of hybridization, suggestive of duplicated sequences. The 74 polymorphic probes yielded complete data for 205 accessions, revealing 219 RFLP bands for statistical analysis. Each band could be associated with one allele, since all but two of the selected probes revealed unique loci and unique bands in most of the accessions. The maximum number of RFLP alleles detected by a probe was 5 and the average number detected by polymorphic probes was 3.00, which is similar to values in previous sorghum studies (Deu et al. 1994; de Oliveira et al. 1996).

All races other than kafir exhibited similar levels of polymorphism, as indicated by the percentage of polymorphic loci, the total number of observed alleles, and Nei's unbiased estimator of genetic diversity (Table 2*a*). The kafir accessions exhibited only 45% polymorphic loci, approximately half the frequency shown by other races, and only 110 alleles, in contrast to 169 to 176 in other races. The kafir accessions also present a very restricted genetic diversity with $H = 0.14$, as compared with values ranging from 0.33 to 0.37 found in other races.

A large number of alleles were rare (frequency ≤1%), totalling 31 out of 219 total alleles (Table 2*a*). Most of the rare alleles, 24 in total, occurred in a single accession. Rare alleles were encountered most often in bicolor and guinea accessions (11 in each race). The rare alleles in bicolor accessions were mostly found in 3 Ethiopian accessions (3 rare alleles) and in the Algerian and Turkish accessions (2 rare alleles each). The majority of rare alleles in the guinea accessions were detected in the margaritiferum subrace (4 alleles) and in the Asian (3 alleles) and Tanzanian (2 alleles) accessions. An intermediate number of rare alleles were observed in caudatum (4 alleles, 2 of which were from China) and durra (6 alleles, with 2 each from Cameroonian and Indian accessions) races. The kafir race, however, exhibited only one rare allele.

Table 1. List of the accessions of the core collection with comparative racial classification.

Accession No.	Origin	Racial classification	
		ICRISAT	CIRAD
IS 13	USA	B	B
IS 303	CHN	KB	B
IS 929	SDN	D	D
IS 1398	SDN	D	D
IS 2156	NGA	B	B
IS 2262	SDN	KC	C
IS 2263	SDN	D	D
IS 2398	ZAF	K	K
IS 2416	ZAF	K	K
IS 2430	NGA	G	Gm
IS 2787	KEN	C	C
IS 2807	ZWE	C	C
IS 2814	ZWE	C	C
IS 2848	ZAF	C	C
IS 3073	SDN	C	C
IS 3421	IND	G	G
IS 3511	SDN	KC	C
IS 3771	CHN	C	CB
IS 3780	CHN	GB	B
IS 3905	MLI	G	G
IS 3957	NPL	G	G
IS 3958	NPL	G	G
IS 3959	NPL	G	G
IS 3967	IND	G	G
IS 3971	IND	D	D
IS 4027	IND	B	B
IS 4285	IND	D	D
IS 4821	IND	D	D
IS 4963	IND	G	G
IS 5430	IND	B	B
IS 5867	IND	DC	DC
IS 5972	IND	D	D
IS 6118	IND	DC	DC
IS 6193	IND	D	D
IS 6351	IND	D	D
IS 6745	BFA	G	G
IS 6828	BFA	G	G
X (IS 7287)	NGA	C	G
IS 7861	NGA	G	G
IS 7889	NGA	G	G
IS 8685	SDN	C	C
IS 8882	UGA	C	C
IS 9303	ZAF	K	K
IS 9331	ZAF	K	C
IS 9468	ZAF	K	K
IS 9527	ZAF	K	K
IS 9597	NER	G	Gm
IS 10194	BFA	GB	B
IS 10234	CAF	GC	C
IS 10801	TCD	GC	GC
IS 10844	TCD	G	GC
IS 10876	NGA	GC	C
IS 10882	NGA	GC	C
IS 11026	ETH	D	D
IS 11119	ETH	C	C

Table 1 (*continued*).

Accession No.	Origin	Racial classification	
		ICRISAT	CIRAD
IS 11827	ETH	D	D
IS 12169	ETH	B	B
IS 12179	ETH	B	B
IS 12447	SDN	DC	GC
IS 12531	ETH	B	B
IS 12542	ETH	D	D
IS 12804	TUR	B	B
IS 12931	CHN	G	G
IS 13113	IND	B	B
IS 13452	ZWE	GC	C
IS 13791	ZAF	K	KC
IS 13845	ZAF	K	K
IS 13848	ZAF	C	C
IS 13926	ZAF	K	KC
IS 13989	ZAF	K	K
IS 14317	SWZ	G	G
IS 14331	ZAF	G	G
IS 14351	MWI	G	G
IS 14414	MWI	G	G
IS 14417	MWI	G	G
IS 15148	CMR	C	C
IS 15443	CMR	C	C
IS 15752	CMR	C	C
IS 16044	CMR	C	C
IS 16101	CMR	D	D
IS 16125	CMR	C	C
IS 16173	CMR	C	C
IS 16186	CMR	C	D
IS 16396	CMR	G	G
IS 16545	CMR	C	C
IS 17658	GHA	G	G
IS 19026	SDN	C	DC
IS 19132	SDN	GC	CB
IS 19453	BWA	D	D
IS 19455	BWA	G	Gm
IS 19466	LKA	G	G
IS 19685	IND	G	G
IS 19847	IND	DC	DC
IS 19953	SEN	G	Gm
IS 20016	SEN	G	G
IS 20064	SEN	G	Gm
IS 20097	SEN	G	G
IS 20351	NER	D	D
IS 20689	USA	B	B
IS 20706	USA	B	B
IS 20727	USA	B	B
IS 20864	CHN	C	C
IS 21124	KEN	C	C
IS 21519	MWI	G	G
IS 21622	MWI	G	G
IS 21849	SLE	G	Gm
IS 21891	USA	C	C
IS 22239	BWA	K	K
IS 22282	BWA	B	B
IS 22287	BWA	K	K
IS 22294	BWA	K	K

Table 1 (*continued*).

Accession No.	Origin	Racial classification	
		ICRISAT	CIRAD
IS 22330	BWA	D	D
IS 22332	BWA	K	KC
IS 22334	BWA	K	K
IS 22893	SDN	GC	GC
IS 23100	TZA	G	G
IS 23142	TZA	D	D
IS 23178	ZMB	B	B
IS 23254	ZMB	B	B
IS 23644	GMB	G	Gm
IS 23645	GMB	G	Gm
IS 23666	GMB	G	Gm
IS 23669	GMB	G	Gm
IS 23777	MWI	G	G
IS 24009	USA	C	C
IS 24072	TZA	G	G
IS 24139	TZA	G	G
IS 24481	ZAF	K	K
IS 24887	NGA	G	G
IS 25077	GHA	G	G
IS 25499	BDI	C	C
IS 25596	RWA	C	C
IS 25702	MLI	GC	GC
IS 25733	MLI	G	G
IS 26041	MLI	G	G
IS 26110	MLI	G	G
IS 26457	BEN	G	G
IS 26554	BEN	G	G
IS 26731	ZAF	B	B
IS 26833	SDN	C	C
IS 27146	ZWE	D	D
IS 27164	ZWE	K	K
IS 27390	BFA	G	G
IS 27490	BFA	G	G
IS 27891	ZAF	GC	GC
IS 28409	YEM	D	D
IS 28645	YEM	DC	DC
IS 29226	SWZ	GC	KB
IS 29233	SWZ	K	K
IS 29310	SWZ	GC	KC
IS 29375	LSO	GC	GC
IS 29407	LSO	K	KC
IS 29409	LSO	K	K
IS 29496	LSO	K	K
IS 29569	LSO	K	K
IS 29606	ZAF	K	K
IS 29691	ZWE	G	G
IS 29872	ZWE	C	KC
IS 29876	ZWE	GC	C
IS 29911	ZWE	C	KC
IS 30030	ZWE	K	K
IS 30175	ZWE	G	G
IS 30317	CHN	C	C
IS 30352	CHN	DC	CB
IS 30385	CHN	C	C
IS 30400	CHN	C	KC
IS 30405	CHN	DC	BC

Table 1 (*concluded*).

Accession No.	Origin	Racial classification	
		ICRISAT	CIRAD
IS 30417	CHN	DC	BC
IS 30436	CHN	C	C
IS 30441	CHN	C	C
IS 30443	CHN	C	CB
IS 30451	CHN	C	C
IS 30538	KOR	B	B
IS 31559	BDI	C	C
IS 31681	DZA	B	B
IS 32569	SOM	D	D
IS 33116	CMR	C	C
IS 33261	CMR	C	C
IS 33353	KEN	C	C
SSM 12	CMR		D
SSM 19	CMR		D
SSM 29	CMR		D
SSM 205	BFA		G
SSM 215	ETH		C
SSM 232	BFA		G
SSM 249	BFA		G
SSM 261	BFA		G
SSM 275	BFA		Gm
SSM 276	BFA		Gm
SSM 379	MLI		G
SSM 501	NER		DC
SSM 505	NER		Gm
SSM 546	NER		D
SSM 547	NER		C
SSM 552	NER		C
SSM 557	NER		G
SSM 625	NGA		DC
SSM 964	SEN		D
SSM 973	SEN		D
SSM 1046	SEN		G
SSM 1049	SEN		B
SSM 1057	SEN		Gm
SSM 1102	TCD		DC
SSM 1103	TCD		DC
SSM 1123	NER		C
SSM 1267	CMR		DC
SSM 1284	ZAR		B
SSM 1370	ZAF		B
SSM 1592	TCD		D
SSM 1611	TCD		D

Note: Bold font indicates missing RFLP data; italics indicate accession for which only RFLP data are available. Racial classification performed in accordance with Harlan and de Wet (1972); B, bicolor; C, caudatum; D, durra; G, guinea; Gm, Guinea margaritiferum; K, kafir; XY, intermediate between X and Y. Country codes are as follows: BEN, Benin; BDI, Burundi; BFA, Burkina Faso; BWA, Botswana; CAF, Central African Republic; CHN, China, CMR, Cameroon; DZA, Algeria; ETH, Ethiopia; GHA, Ghana; GMB, Gambia; IND, India; KEN, Kenya; KOR, Korea; LKA, Sri Lanka; LSO, Lesotho; MLI, Mali; MWI, Malawi; NER, Niger; NGA, Nigeria; NPL, Nepal; RWA, Rwanda; SDN, Sudan; SEN, Senegal; SLE, Sierra Leone; SOM, Somalia; SWZ, Swaziland; TCD, Chad; TUR, Turkey; TZA, Tanzania; UGA, Uganda; USA, United States of America; YEM, Yemen; ZAF, Republic of South Africa; ZAR, Democratic Republic of Congo (Zaïre), ZMB; Zambia; ZWE, Zimbabwe.

Table 2. Genetic variation within race and cluster defined according to the NJ analysis.

(*a*) Variation within races.

Race (no. of accessions)	% polymorphic loci[a]	Total no. of alleles	No. of rare alleles[b]	Mean no. of alleles/ polymorphic loci[c]	H_{exp}[d]
Bicolor (22)	0.92	175	11	2.38	0.37
Caudatum (43)	0.93	172	4	2.34	0.34
Durra (29)	0.91	169	6	2.30	0.33
Guinea (62)	0.86	176	11	2.38	0.35
Kafir (18)	0.45	110	1	1.49	0.14
Total (205)	0.99	219	31	3.00	0.40

(*b*) Variation within clusters defined by NJ analysis.

Cluster (no. of accessions)	% polymorphic loci[a]	Total no. of alleles	No. of rare alleles[b]	Mean no. of alleles/ polymorphic loci[c]	H_{exp}[d]
Cluster I (26), guinea from western Africa	0.61	129	0	1.75	0.19
Cluster II (14), guinea margaritiferum	0.39	106	4	1.44	0.10
Cluster III (25), durra	0.76	151	6	2.05	0.25
Cluster IV (12), Chinese sorghums	0.31	99	2	1.34	0.10
Cluster V (25), caudatum	0.76	140	1	1.90	0.25
Cluster VI (11), transplanted sorghums	0.66	126	1	1.71	0.24
Cluster VII (30) kafir and intermediatekafir	0.46	112	0	1.52	0.14
Cluster VIII (13), guinea from southern Africa	0.41	110	0	1.49	0.18
Cluster IX (8), Asian guinea	0.32	103	3	1.40	0.11
Cluster X (7), caudatum	0.61	124	0	1.68	0.27

[a]Percentage of polymorphic loci at the 0.99 criterion.
[b]Frequency ≤1% in the total collection.
[c]At the 0.99 criterion in the total collection.
[d]Average Nei's unbiased estimator of genetic diversity.

Multivariate analysis

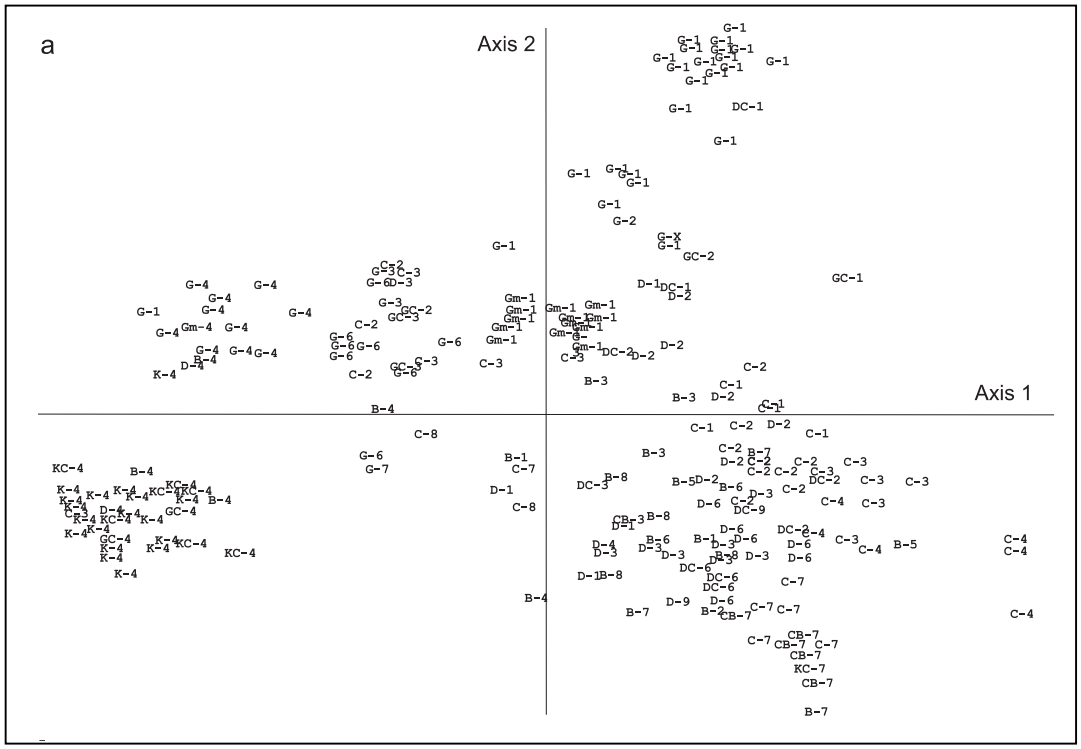
Considering all loci simultaneously, 194 genotypes were identified among the 205 accessions. The first FAC on 219 RFLP bands identified one distinct group composed almost entirely of guinea margaritiferum accessions (along with one guinea accession from Sri Lanka and one GC accession from Chad). A second FAC was conducted on 180 active RFLP bands (i.e., bands that were taken into account for the analysis), with the rare bands (≤1% frequency) and the specific bands of the previously identified, distinct, margaritiferum group of accessions set as inactive variables. However, this second FAC still resulted in a guinea margaritiferum group distinctly separated from all other cultivated sorghums (data not shown). A third FAC was performed on 175 active bands, with the guinea margaritiferum accessions as inactive, to examine the pattern of molecular variation in the rest of the cultivated sorghums.

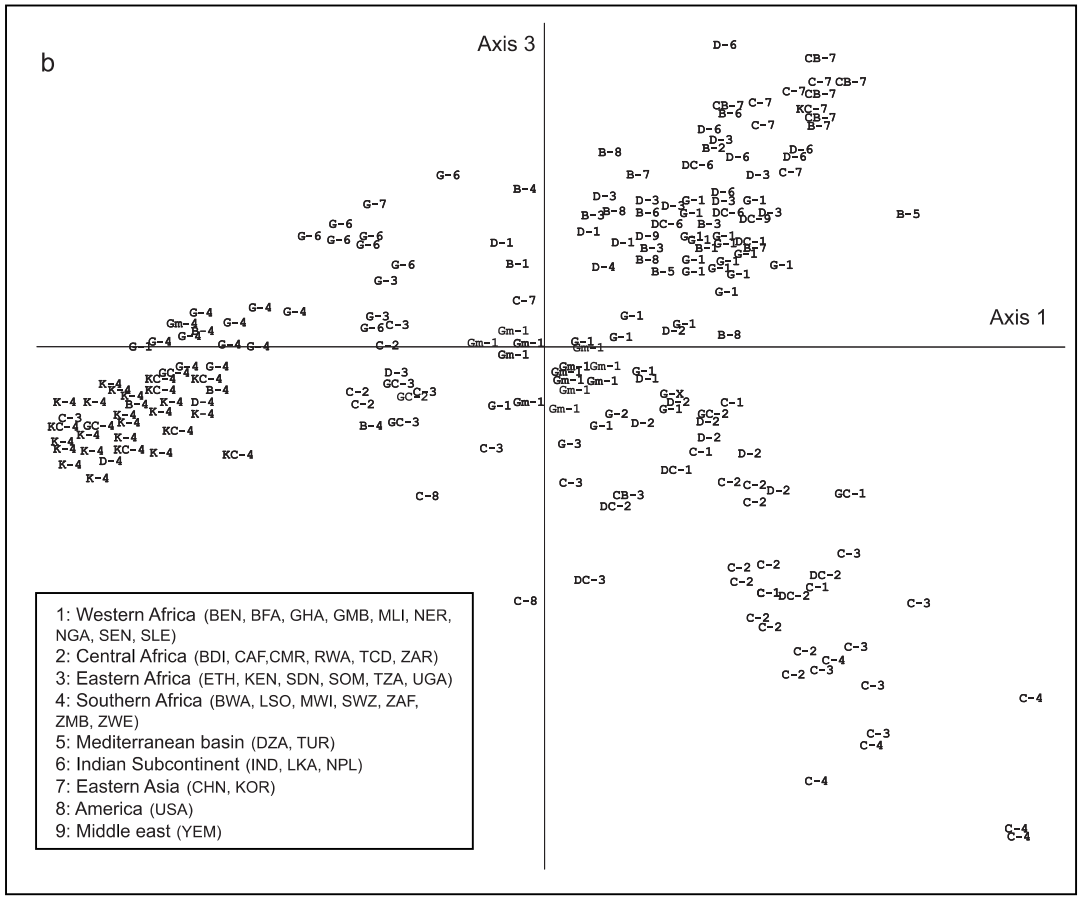
The 3 main axes of the third FAC accounted for 12.9%, 8.7%, and 7.7% of the variation. The distribution of the accessions on the first 2 axes is shown in Fig. 1*a*. Differentiation based on geographic origin and racial classification is observed on this plane. The first axis separated the kafir accessions and southern African and Asian guinea accessions (negative coordinates) from other accessions. Axis 2 separated southern African accessions into 2 groups; a kafir group (lower) and a guinea race group (upper) with closely associated Asian guinea accessions. This axis also separated western African guinea accessions (upper right) from durra and caudatum accessions. The third axis (Fig. 1*b*) separated African caudatum accessions (lower right) from a mixed group composed primarily of durra and Asian caudatum accessions.

Discriminative markers and specific alleles

We examined the loading values of the 219 RFLP bands on the main axes for the first and third FACs. This could permit the identification of alleles exhibiting the highest power of discrimination between groups. The first FAC permitted identification of alleles (and consequently probes) that discriminate margaritiferum accessions from the rest of the cultivated sorghums, whereas the discriminative alleles within non-margariteferum accessions were identified by the third FAC. Forty-eight discriminative probes were identified. They appeared scattered over the whole sorghum genome (Fig. 2). For example, alleles with strong negative contributions to axis 1 (mostly southern African guinea and kafir and Asian guinea accessions) were identified by probes located on linkage groups A, B, F, G, H, I, and J on our reference map (Boivin et al. 1999). Similarly, alleles with strong positive contributions to axis 2 (primarily guinea accessions

Fig. 1. Distribution of 205 sorghum accessions on the first planes of the FACs performed with 175 RFLP bands as active variables. The accessions are represented on plane (1, 2) in *a* and on plane (1, 3) in *b*, according to their morphological race and geographic origin. Each letter represents a race code designation (described in Table 1), and each number represents a geographic origin: 1, western Africa; 2, central Africa; 3, eastern Africa; 4, southern Africa; 5, Mediterranean basin; 6, Indian subcontinent; 7, eastern Asia; 8, America; 9, Middle East.





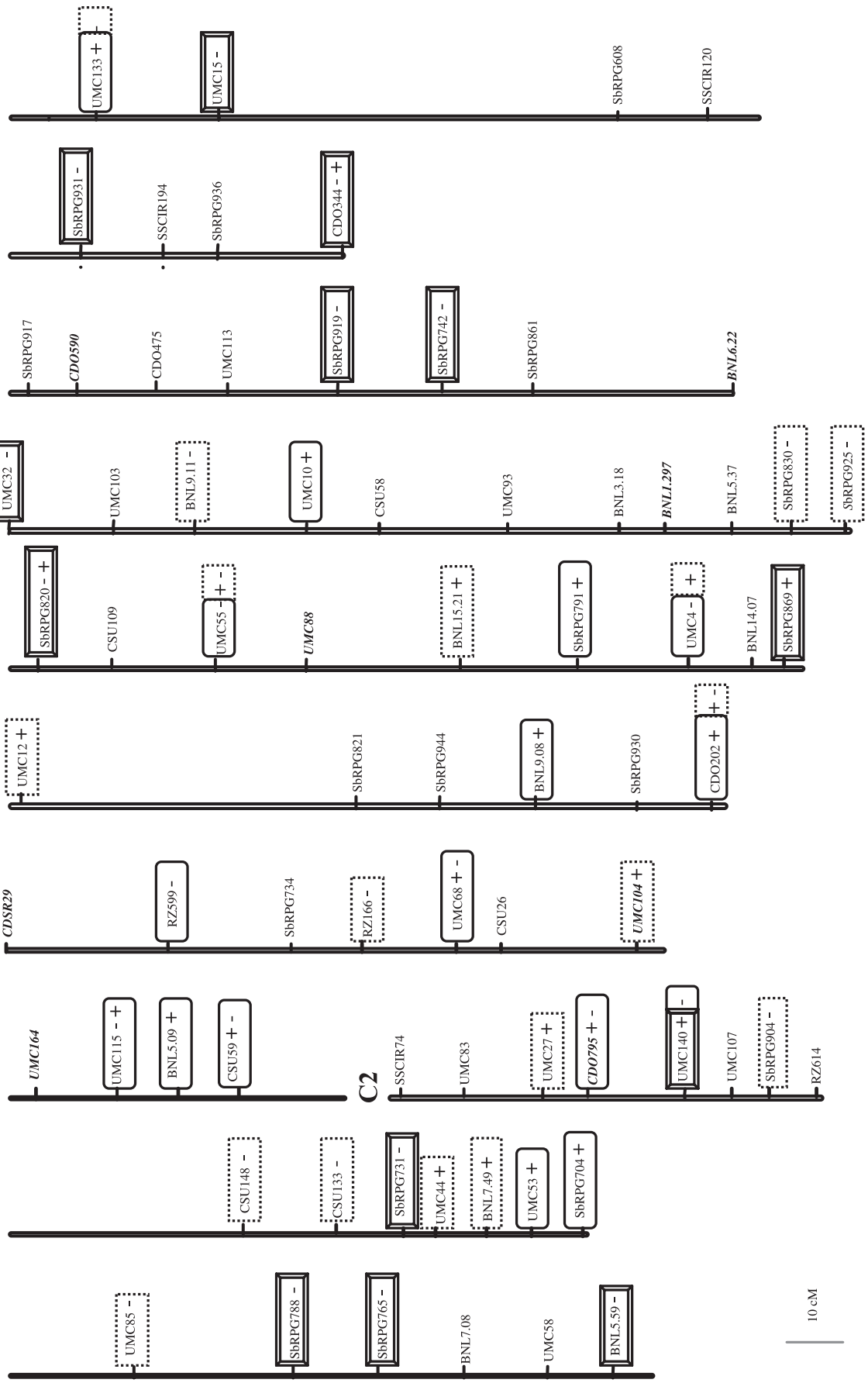


Fig. 2. Location on the sorghum genetic map of the 74 probes used for sorghum core collection analysis. The genetic map was described in Boivin et al. (1999) and Ventelon et al. (2001). Linkage groups are named according to these authors. Discriminative probes detecting specific guinea margaritiferum alleles in the first FAC are noted in bold and italic. Discriminative probes identified in the third FAC are outlined by boxes; those probes for which alleles have a large contribution to axis 1, 2, and 3 are outlined by double lines, single lines, and dashed lines, respectively. + and – in the box indicates whether the contribution of alleles to each axis is positive or negative.

from western Africa) were revealed by probes on linkage groups B, C, D, E, F, G, and J. Alleles with positive contribution to axis 3 (primarily durra and Asian caudatum accessions) were identified by probes on linkage groups B, C, D, E, and F. Nevertheless, some alleles at linked loci (separated by fewer than 10 cM) contributed greatly to the discrimination between groups. One such example is the alleles detected by probes UMC53 and SbRPG704 (linkage group B), for which the former differentiated western African guinea accessions from the other cultivated sorghums, while the latter differentiated all guinea accessions from the rest of the collection.

No allele appeared to be unique to a specific geographical or racial cluster of accessions. Nevertheless, some alleles were found to occur predominantly in specific clusters. The margaritiferums in this study exhibited 8 specific alleles, 5 of which were present in most of the margaritiferum accessions. The probes CDSR29, CDO590, CDO795, UMC104, and UMC164 revealed these 5 specific alleles. Alleles quasi-specific to most of the Chinese caudatum and bicolor accessions were revealed by probes UMC55, BNL 7.49, and SSCIR194. No allele specific to the transplanted sorghum accessions was observed.

Neighbour-joining analysis

The dissimilarity values calculated between all pairwise accessions with the 219 RFLP bands ranged from 0 to 0.718 with an average of 0.196. The neighbour-joining (NJ) analysis (Fig. 3) permitted identification of the following 10 major clusters of accessions: cluster I, guinea accessions (not including the margaritiferum subrace) from western Africa; cluster II, guinea margaritiferum accessions; cluster III, durra accessions from Asia and Africa and durra–caudatum from Asia; cluster IV, Chinese accessions of mainly caudatum, bicolor, and caudatum–bicolor intermediate races; cluster V, caudatum accessions from diverse African countries; cluster VI, durra and caudatum accessions of transplanted sorghums from Chad and Cameroon; cluster VII, kafir and intermediate kafir varieties from southern Africa; cluster VIII, guinea accessions from southern Africa; cluster IX, guinea accessions from Asia; cluster X, caudatum accessions predominantly from the Great Lakes region of Africa.

These clusters formed 2 main geographical groups, with clusters I–VI primarily including accessions from north equatorial Africa and clusters VII–X containing accessions from south equatorial Africa. The Asian accessions were distributed among these two main geographical groups. A single predominant race could be identified in 8 of these clusters, with guinea accessions predominant in 4 clusters, caudatum accessions in 2 clusters, and durra and kafir accessions each in 1 cluster. Although the two remaining clusters appeared quite mixed at the racial level, they could be characterized by a specific geographic origin (cluster IV, composed of Chinese sorghums) or by a particular crop system carried out in a narrow geographic area (cluster VI, composed of transplanted sorghums). These transplanted sorghums are both durra and caudatum accessions cultivated in the post-rainy season on residual moisture in inundated clays from Chad and Cameroon after flood subsidence. Bicolor accessions were distributed widely across different clusters, although a portion of them was present in cluster IV.

Bootstrap values could confirm the consistency of some clusters, i.e., margaritiferum accessions (bootstrap value of 100% for cluster II) and Chinese accessions (bootstrap value of 98% for cluster IV). Relatively low bootstrap values observed for other clusters, i.e., durra (cluster III) or caudatum (cluster V), could be explained by the moderately strong structure observed in cultivated sorghums, the hybrid or intermediate origin of some accessions that could not be classified in a specific cluster, or the larger genetic variation in these clusters (as confirmed by the long branches found in these clusters compared with branches found in guinea margaritiferum and Chinese sorghums). Nevertheless, the high value found for the cophenetic correlation ($r_{coph} = 0.92$) indicated a good agreement between the tree and the original dissimilarity matrix, thus corroborating the consistency of the tree.

Discussion

This study provides one of the most detailed and comprehensive examinations of genetic diversity in tropical and sub-tropical landrace sorghums through analysis of a core collection, representative of geographic origin, race, photoperiod sensitivity, and production system, using a large number of markers (74 RFLP probes) dispersed over the entire genome. To evaluate the use of molecular markers for large-scale germplasm diversity analysis, Dillmann et al. (1997) have proposed several criteria that have been applied in this study: a large number of markers, a large repartition of these markers on the genetic map of the species, the selection of monolocus probes, and the use of a single enzyme per probe. These authors have reported that the precision of the estimation of the genetic distance calculated between all pairwise maize lines increases with the number of markers, as each marker can be considered as a sample of the genome.

The RFLP markers permitted the identification of 194 genotypes among the 205 accessions and 8 sets of accessions with identical RFLP patterns were found. They included 18 accessions. We could observe that each set was composed of 2 or 3 accessions originating from the same country or belonging to the same race and presenting a phenotypic similarity as shown by the analysis of morpho-agronomic traits in the field trials conducted in Mali and France. These identities could highlight the duplications of entries in the ex situ collections. These identities could be confirmed or weakened, since a large number of highly polymorphic SSR markers developed and mapped during these last years are now available for the sorghum research community (Bhattramakki et al. 2000; Kong et al. 2000; Schloss et al. 2002). The characterization of this sorghum collection with a set of carefully selected SSR markers is now in progress.

Loci involved in the differentiation of the clusters were scattered over the whole sorghum genome. Furthermore, alleles at linked loci are also involved in discrimination between clusters. Interestingly, some of the discriminative loci mapped in genomic regions associated with morphology-related QTLs (plant height, panicle compactness, and length) identified by Rami et al. (1998) on linkage groups A, C1, and F. At least, the pattern of diversity revealed in this sorghum core collection permits investigation of genome-wide

LD owing to crop history (founder effect, genetic drift, and selection) and local LD owing to genetic linkage. We are engaged in an exploratory study on the extent and intensity of linkage disequilibrium useable for association studies in this species (Deu and Glaszmann 2004).

Geographic and racial pattern of clustering

The apparent first order differentiation between northern and southern equatorial African accessions indicated by the NJ analysis suggests 2 major geographic poles for sorghum evolution and differentiation. This observation is compatible with the postulated origin of this species in the northeast quadrant of Africa (Harlan and Stemler 1976; Doggett 1988; Wendorf et al. 1992). The absence of rare alleles (Table 2*b*) in southern equatorial accessions (kafir, guinea, and caudatum found in cluster VII, VIII, and X, respectively) fits also with the expectation that southern equatorial African sorghums evolved later from other African sorghums (Doggett 1988). The suggested bipolar evolution of sorghums also agrees with indications of ethnic divisions between northern (Nilotic and Sudanian) and southern equatorial Africa (Bantu) (de Wet and Huckabay 1967; Gourou 1970; Doggett 1988) that could have contributed to isolation of gene pools and divergent evolution. The presence of Asian accessions in both the north equatorial group (cluster IV, Fig. 3) and south equatorial group (cluster IX) suggests that sorghums from both pools could have been introduced to Asia as suggested by Harlan and Stemler (1976).

The NJ analysis indicated that morphological race also has substantial influence on the pattern of genetic diversity, with 8 of 10 clusters based on a single predominant race (Fig. 3). Previous studies have also shown associations between racial characterization and the pattern of genetic diversity as measured by RFLP (Deu et al. 1994; Cui et al. 1995) and RAPD (Tao et al. 1993) and by combining RFLP, RAPD, and ISSR molecular markers (de Oliveira et al. 1996). However, the racial differentiation observed in most of these studies was less clear than in this study. The stronger association of racial classification with pattern of diversity observed in this study could be due to many factors, including stratification of world germplasm and a greater number of accessions sampled for effectively sampling diversity, inclusion only of landrace accessions, and rigorous verification of racial classification.

Genetic diversity within race

The high diversity of the bicolor race, shown by its estimator of gene diversity, its mean number of alleles (Table 2*a*), and its presence in multiple clusters (Fig. 3), corresponds with expectations for this race, which is considered the most ancient with such wide geographic distribution and diversity of uses (forage, broom-corn, and sweet stems) (Doggett 1988).

The guinea race also exhibited high diversity, with 3 main and very distinct groups (western African (non-margaritiferum), margaritiferum, and southern African) (Fig. 3). These groups correspond to those identified previously with isozymes, RFLP, and SSR markers (Ollitrault et al. 1989; Deu et al. 1994, Cui et al. 1995; Folkertsma et al. 2005). The genetic distinctness of guinea margaritiferum sorghums from other guinea forms was previously discussed by Deu et al. (1994, 1995), de Oliveira et al. (1996), and Folkertsma et al. (in press). Our study showed that margaritiferums differed from other guineas not only by possessing rare and specific alleles, but also by having distinct genotypes at other loci, as indicated by the second FAC performed without the alleles specific to margaritiferums. The genetic distinctness of margaritiferums from other guinea sorghums from western Africa is remarkable, since both are interfertile and cultivated in sympatry in the same season by the same farmers. The only margaritiferum from southern Africa (IS 19455) included in this study did not cluster with the western African margaritiferums, but rather clustered with other southern African guinea accessions (Fig. 3), even though it was found to present a mitotype closely related to western African guinea margaritiferums mitotypes (Deu et al. 1995). These results suggest that this southern African margaritiferum had a common ancestor with western African margaritiferums and that human selection and geographic isolation resulted in marked changes of its nuclear genetic background.

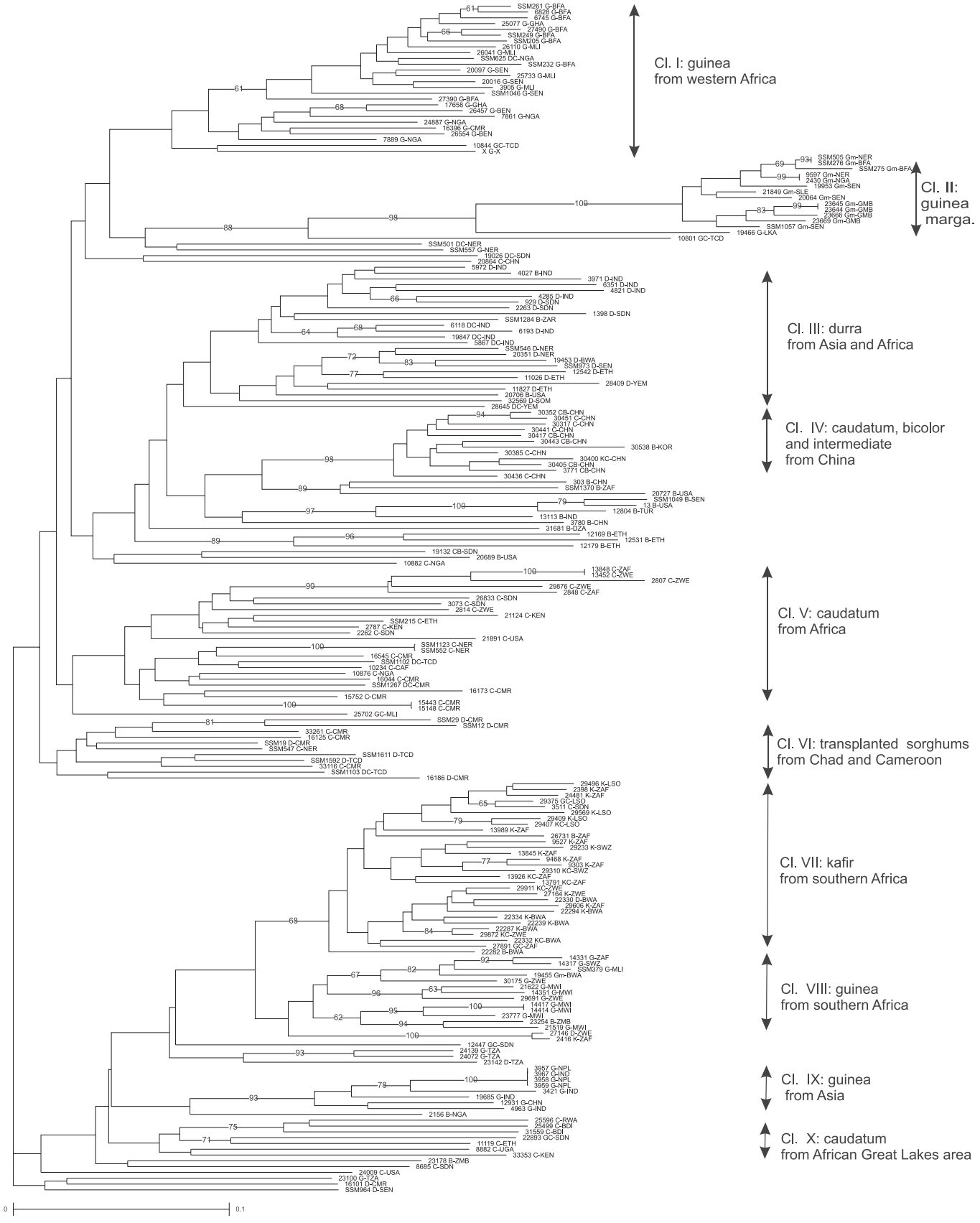
Our study has also revealed an additional small cluster of guinea accessions originating from Asia, not shown in previous studies. The close relationship between these Asian accessions and southern African guinea accessions suggests recent introduction of Asian forms from southern Africa. However, the presence of 3 alleles (frequency <10%) in both the Asian guinea and Asian durra accessions (data not shown) suggests weak gene flow between these races or limited introgression from wild local sorghums.

The differentiation of caudatum accessions into 3 groups (Fig. 3), although not previously reported as such, does correspond with previous studies. The 2 main groups, African Great Lakes (cluster X) and remaining African countries (cluster V), were clearly differentiated both in this study and in that of Deu et al. (2003) based on observations of 21 morphological traits. These 2 groups also showed considerable difference for photoperiod sensitivity when sown in Mali at 1-month intervals, with the Great Lakes accessions showing much higher sensitivity than that of the others (data not shown).

The third and smallest group of caudatums and caudatum–bicolor intermediates (cluster IV) was composed of Chinese accessions. This group showed very restricted genetic diversity (Table 2*b*), as was previously reported (de Oliveira et al. 1996). This group was not highly distinct from other cultivated sorghums in contrast to the finding of de Oliveira et al. (1996), although the small number of accessions involved in our study ($n = 12$) means that results could be influenced by sampling.

The observation of durra sorghums comprising 2 separate groups (Fig. 3) has not been previously reported. The main group (cluster III) consists of accessions of the most widely cultivated durra sorghums. These sorghums are known to have superior adaptation to droughty rain-fed conditions, and are considered to have originated in northeast Africa from where they migrated throughout Africa and on to Asia (Doggett 1988). The other durras, belonging to the mixed group composed of durra and caudatum accessions (cluster VI), consist of particular transplanted sorghums from Chad and Cameroon, which are cultivated in the post-rainy season by transplantation in receding moisture systems. The cultivation of these 2 groups in different seasons would likely limit

Fig. 3. Neighbour-joining analysis based on RFLP data among 205 cultivated accessions using the Nei and Li similarity index. The numbers on the branches indicate bootstrap values (expressed in percentages) and are shown for all clusters with > 60% bootstrap support. Cl, cluster.



genetic exchange between them and other sorghums that did not compete with them in dry conditions.

A low genetic diversity was observed in the kafir race as previously reported in studies using isozymes, RFLP, RAPD, or SSR markers (Ollitrault et al. 1989; Deu et al. 1994; Cui et al. 1995; Menkir et al. 1997; Djé et al. 2000). All but one of the kafir accessions constituted a specific cluster. This race exhibited the lowest number of alleles and the unique rare allele was encountered in the accession not included in the kafir cluster. These results are in agreement with the recent origin and restricted geographic distribution of this race (Doggett 1988).

Genetic resources management and crop improvement

The patterns of genetic relationships observed in this study should provide more detailed insights for genetic resource conservation and use of sorghum. The patterns of diversity within the guinea and caudatum races are particularly relevant to genetic base broadening efforts within those races. The association of level of photoperiod sensitivity with genetic differentiation observed within the caudatum race would have significant consequences for accessing the full range of genetic diversity in applied breeding programs, since they primarily use germplasm of similar photoperiod sensitivity. This is particularly critical in the guinea race, which is the most photoperiod sensitive (Grenier et al. 2001*a*), and any relationship between photoperiod sensitivity and genetic differentiation would have significant consequences for efforts to use genetic diversity within this race to develop guinea-race hybrids for western Africa.

Finally, the pattern of genetic diversity, as revealed in this study with RFLP markers, may offer new opportunities to relate that diversity to the structure of diversity for important agronomic traits, since this collection is being phenotyped for components of photoperiod sensitivity, putative components of drought tolerance, and grain-quality characteristics.

Acknowledgements

This work was financially supported by a grant from the Bureau des Ressources Génétiques (Paris, France), which is gratefully acknowledged. We thank Fanny Vanel for technical assistance and J.L. Noyer (CIRAD) and Dr. S. Chandra (ICRISAT) for their constructive comments on a previous version of the manuscript.

References

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

Bhattramakki, D., Dong, J., Chhabra, A.K., and Hart, G.E. 2000. An integrated SSR and RFLP linkage map of *Sorghum bicolor* (L.) Moench. Genome, **43**: 988–1002.

Belkhir, K., Borsa P., Chikhi, L., Raufaste, N., and Bonhomme, F. 2002. Genetix 4.04, logiciel sous Windows™ pour la génétique des populations. Laboratoire génome, populations, interaction, CNRS UMR 5000, Université de Montpellier II, Montpellier, France.

Boivin, K., Deu, M., Rami, J.F., Trouche, G., and Hamon, P. 1999. Towards a saturated sorghum map using RFLP and AFLP markers. Theor. Appl. Genet. **98**: 320–328.

Cui, Y.X., Xu, G.W., Magill, C.W., Schertz, K.F., and Hart, G.E. 1995. RFLP-based assay of *Sorghum bicolor* (L.) Moench genetic diversity. Theor. Appl. Genet. **90**: 787–796.

de Oliveira, A.C., Richter, T., and Bennetzen, J.L. 1996. Regional and racial specificities in sorghum germplasm assessed with DNA markers. Genome, **39**: 579–587.

Deu, M., Gonzalez-de-Leon, D., Glazsmann, J.C., Degremont, I., Chantereau, J., Lanaud, C., and Hamon, P. 1994. RFLP diversity in cultivated sorghum in relation to racial differentiation. Theor. Appl. Genet. **88**: 838–844.

Deu, M., Hamon, P., Chantereau, J., Dufour, P., D'Hont, A., and Lanaud, C. 1995. Mitochondrial DNA diversity in wild and cultivated sorghum. Genome, **38**: 635–645.

Deu, M., Hamon, P., Bonnot, F., and Chantereau, J. 2003. Sorghum. *In* Genetic diversity of cultivated tropical plants. *Edited by* P. Hamon, M. Seguin, X. Perrier, and J.C. Glaszmann. Science Publishers, Inc., and CIRAD, Montpellier, France. pp. 307–336.

Deu, M., and Glaszmann, J.C. 2004. Linkage disequilibrium in sorghum. *In* Proceedings of the Plant and Animal Genome XII Conference, San Diego, Calif., 10–14 January 2004. Applied Biosystems, Foster City, Calif. p. 7. Available from http://www.intl-pag.org/12/abstracts/W02_PAG12_10.html.

de Wet, J.M.J., and Huckabay, J.P. 1967. The origin of *Sorghum bicolor* II distribution and domestication. Evolution, **21**: 787–802.

de Wet, J.M.J., and Harlan, J.R. 1972. Origin and evolution of guinea sorghums. E. Afr. Agric. For. J. **38**: 114–119.

Dillmann, C., Bar-Hen, A., Guerin, D., Charcosset, A., and Murigneux, A. 1997. Comparison of RFLP and morphological distances between maize *Zea may*s L. inbred lines. Consequence for germplasm protection purposes. Theor. Appl. Genet. **95**: 92–102.

Djé, Y., Heuertz, M., Lefèbre, C., and Vekemans, X. 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, New York, N.Y.

Dufour, P., Deu, M., Grivet, A., D'Hont, A., Paulet, F., Bouet, A., Lanaud, C., Glaszmann, J.C., and Hamon, P. 1997. Construction of a composite sorghum genome map and comparison with sugarcane, a related complex polyploid. Theor. Appl. Genet. **92**: 1034–1030.

Folkertsma, R.T., Rattunde, F.H., Chandra, S., Soma Raju, G., and Hash, C.T. 2005. The pattern of genetic diversity of Guinea-race *Sorghum bicolor* (L.) Moench landraces as revealed with SSR markers. Theor. Appl. Genet. **111**: 399–409.

Gourou, P. 1970. L'Afrique. Hachette, Paris, France.

Grenier, C., Bramel-Cox, P.J., Noirot, M., Prasadao Rao, K.E., and Hamon P. 2000. Assessment of genetic diversity in three subsets constituted from the ICRISAT sorghum collection using random vs. non-random sampling procedures. A. Using morpho-agronomical and passport data. Theor. Appl. Genet. **101**: 190–196.

Grenier, C., Hamon, P., and Bramel-Cox, P.J. 2001*a*. Core collection of sorghum: I. Stratification based on eco-geographical data. Crop Sci. **41**: 234–240.

Grenier, C., Hamon, P., and Bramel-Cox, P.J. 2001*b*. Core collection of sorghum: II. Comparison of three random sampling strategies. Crop Sci. **41**: 241–246.

Harlan, J.R., and de Wet, J.M.J. 1972. A simplified classification of cultivated sorghum. Crop Sci. **12**: 127–176.

Harlan, J.R., and Stemler, A. 1976. The races of Sorghum in Africa. *In* Origin of African plant domestication. *Edited by* J.R. Harlan, J.M.J. de Wet, and A. Stemler. Mouton Press, The Hague, Netherlands. pp. 465–478.

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

Menkir, A., Golsbrough, P., and Ejeta, G. 1997. RAPD based assessment of genetic diversity in cultivated races of sorghum. Crop Sci. **37**: 564–569.

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

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

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

Noli, E., Salvi, S., and Tuberosa, R. 1997. Comparative analysis of genetic relationships in barley based on RFLP and RAPD markers. Genome, **40**: 607–616.

N'Goran, J.A.K., Laurent, V., Risterucci, A.M., and Lanaud, C. 1994. Comparative genetic diversity studies of *Theobroma cacao* L. using RFLP and RAPD markers. Heredity, **73**: 589–597.

Ollitrault, P., Arnaud, P., and Chantereau, J. 1989. Polymorphisme enzymatique des sorghos. II Organisation génétique et évolutive des sorghos cultivés. Agron. Trop. **44**: 211–222.

Perrier, X., Flori, A., and Bonnot, F. 2003. Methods of data analysis. *In* Genetic diversity of cultivated tropical plants. *Edited by* P. Hamon, M. Seguin, X. Perrier, and J.C. Glaszmann. Science Publishers, Inc., and CIRAD, Montpellier, France. pp. 31–63.

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

Rami, J.F., Dufour, P., Trouche, G., Fliedel, G., Mestres, C., Davrieux, F., Blanchard, P., and Hamon, P. 1998. Quantitative trait loci for grain quality, productivity, morphological and agronomical traits in sorghum (*Sorghum bicolor* L. Moench). Theor. Appl. Genet. **97**: 605–616.

Schloss, S.J., Mitchell, S.E., White, G.M., Kukatla, R., Bowers, J.E., Paterson, A.H., and Kresovich, S. 2002. Characterization of RFLP probe sequence for gene discovery and SSR development in *Sorghum bicolor* (L.) Moench. Theor. Appl. Genet. **105**: 912–920.

Sneath, P.H.A., and Sokal, R.R. 1973. Numerical taxonomy. WH Freeman and Company, San Francisco, Calif.

Snowden, J.D. 1936. The cultivated races of sorghum. Adlard, London, UK.

Tao, Y., Manners, J.M., Ludlow, M.M., and Henzell, R.G. 1993. DNA polymorphisms in grain sorghum (*Sorghum bicolor* (L.) Moench). Theor. Appl. Genet. **86**: 679–688.

Ventelon, M., Deu, M., Garsmeur, O., Doligez, A., Ghesquiere, A., Lorieux, M., Rami, J.F., Glaszmann, J.C., and Grivet, L. 2001. A direct comparison between the genetic maps of sorghum and rice. Theor. Appl. Genet. **102**: 379–386.

Wendorf, F., Close, A.E., Schild, R., Wasylikowa, K., Housley, R.A., Harlan, J.R., and Krolik, H. 1992. Saharan exploitation of plants 8000 years BP. Nature (London), **359**: 721–724.