Sorghum

Monique Deu, Perla Hamon, François Bonnot
and Jacques Chantereau

Sorghum, *Sorghum bicolor* (L.) Moench, is cultivated in tropical as well as temperate countries. In 1995, it was cultivated on over 43 million ha and the global production was more than 54 million t (FAO, 1995). It is the fifth most important cereal in the world, after wheat, rice, maize, and barley. Despite the growing importance of rice and maize, sorghum remains an essential element of human nutrition for several countries in Africa (Sudan, Botswana, Burkina Faso, Rwanda, Chad, and Cameroon) and Asia (India and China). In 1995, the production of sorghum grain was more than 16 million t in Africa and 15 million t in Asia. The grains are made into flour and consumed as porridge or pancakes. Certain better-adapted varietal types are used in making beer, sweets, or popcorn.

Sorghum has other uses also. In Argentina, Australia, South Africa, Mexico, and especially the United States, where the production in 1995 exceeded 17 million t, sorghum is mainly used as animal feed. Finally, in certain regions of Africa and Asia, the panicle may be used to make brooms while the stems are used as forage, fuel, or construction materials or in leather dyeing or paper making. The pith may yield sugar, syrup, glue, and alcohol.

Sorghum is a hardy plant, adapted to difficult environments and tolerant of poor soils, drought, high temperatures, and even flooding. It can be cultivated where more valuable crops cannot be.

TAXONOMY AND GENETIC RESOURCES

Taxonomy and Geographic Distribution of the Genus *Sorghum*

Sorghum (genus *Sorghum*, family Poaceae) is a cereal closely related to maize and sugarcane; all three belong to the tribe Andropogoneae. The great morphological diversity of the genus *Sorghum* led botanists to distinguish a large number of taxa—712 were described by Snowden (1936). A simplified classification, taking genetic exchanges into account, is widely used today (de Wet, 1978). Recommended by the IPGRI (International Plant Genetic

Resources Institute), it divides the genus into five sections. The section Sorghum includes the following: all the cultivated sorghums with grains (*S. bicolor* ssp. *bicolor*, diploid, 2x = 20); wild sorghums, diploid, annual, originating from Africa (*S. bicolor* ssp. *verticilliflorum*); wild sorghums, diploid, with rhizomes, perennial, present in India, Sri Lanka, and Southeast Asia (*S. propinquum*); and wild sorghums, tetraploid, also with rhizomes and perennial, found in Southeast Asia, India, Middle East, and around the Mediterranean Sea (*S. halepense*). According to Mann et al. (1983), sorghum was domesticated around 3000 years BCE in northeastern Africa. However, later archaeological data (Wendorf et al., 1992) suggest that it was first used more than 6000 years BCE. According to Harlan and Stemler (1976) and Doggett (1988), the cultivated sorghums derive from wild African sorghums *S. bicolor* ssp. *verticilliflorum*.

The cultivated sorghums present a wide phenotypic diversity. From the form of the inflorescence and above all the structure of the panicle, it is classified into five basic races (bicolor, caudatum, durra, guinea, and kafir) and ten intermediate races obtained by the combination of any two of those races (guinea-bicolor, durra-caudatum; Harlan and de Wet, 1972). However, according to Doggett (1988), the bicolor sorghums, selected on traits not linked to the structure of the panicle, such as sweet stem and forage, do not constitute a race but rather form a fairly heterogeneous group. These sorghums would be close to the primitive sorghums from which they were domesticated.

The various races of sorghum presently occupy contiguous areas of distribution, even though one race may be predominant in a region. The sorghums cultivated in India are principally the durra, which are also found in East Africa (Ethiopia) and the Middle East (Turkey, Syria). The bicolor type is widespread in Africa and Asia, while the kafir are found essentially in southern Africa, from Tanzania to South Africa. The caudatum race, predominant in central Africa, has become an important source of genetic material in breeding programmes for temperate regions. The guinea, sorghums typical of West Africa, are also present in East Africa and southern Africa.

The dispersion and present geographic distribution of the cultivated sorghums (Figs. 1 and 2) are closely linked to past human migrations, cultural traditions, and dietary practices of the ethnic groups who cultivate them. In Africa, for example, the durra race is associated with Islam, as opposed to the kafir sorghums cultivated by 'infidels'. The wide geographic distribution of the cultivated sorghums also shows that the races adapted to different ecological zones. The guinea race—large sorghums with a very loose panicle that permits quick drying and corneous grains that are resistant to moulds—is often cultivated in the humid zones, while the durra—short sorghums with a generally compact panicle and large grains—are well adapted to drought.

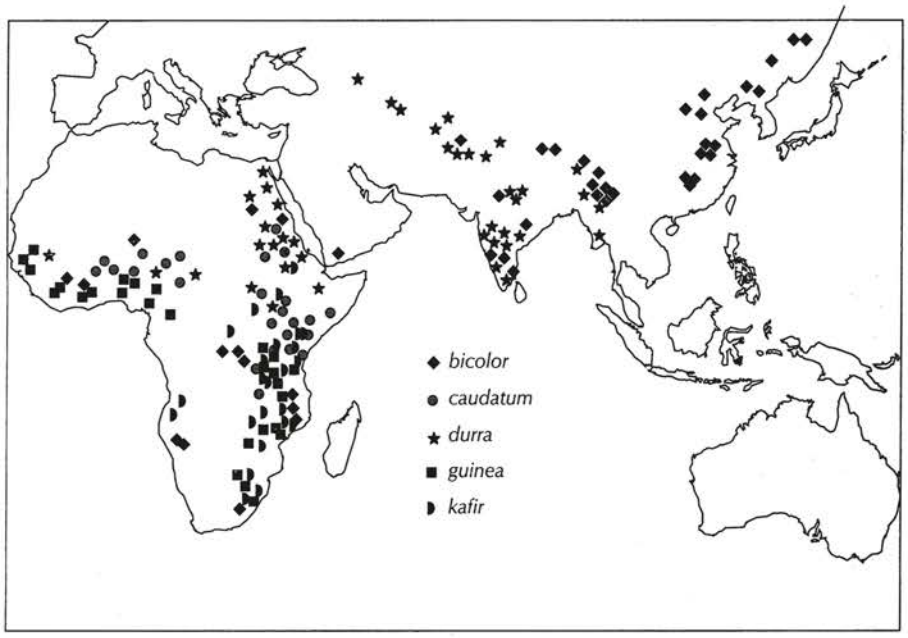

Fig. 1. Distribution of cultivated sorghum races.

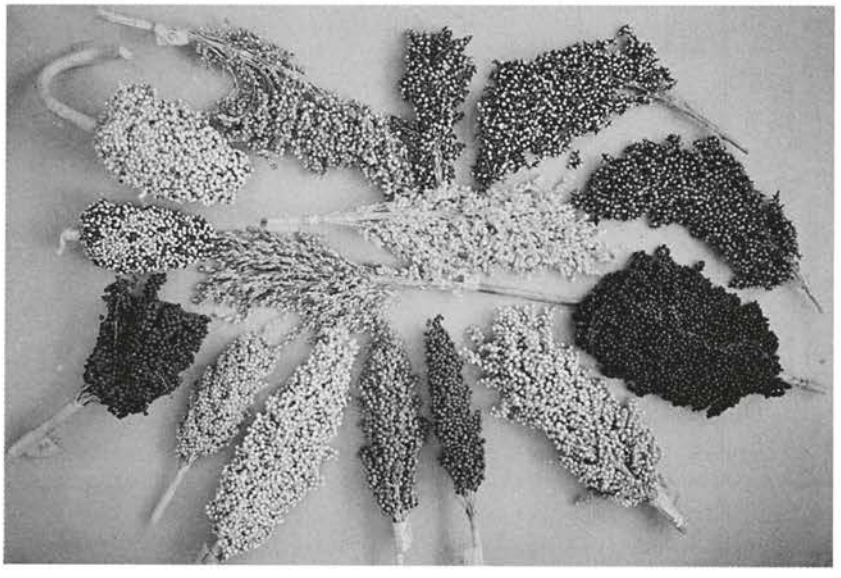

Fig. 2. Races of cultivated sorghum and their hybrids: (left) durra, (top) guinea, (below) caudatum, and (centre) bicolor (photo by J.L. Noyer—CIRAD).

The cultivated sorghums are monoecious and preferentially autogamous even though relatively high rates of allogamy (10 to 30%) were indicated in guinea populations (Ollitrault, 1987; Chantereau and Kondombo, 1994). All the races of cultivated sorghums can be crossed with each other and hybridize with the related wild diploid sorghums (*S. bicolor* ssp. *verticilliflorum*). Similarly, gene flows are possible between *S. bicolor* and *S. propinquum*, wild diploid (de Wet et al., 1976), and between *S. bicolor* and *S. halepense*, wild tetraploid (Arriola, 1995).

Genetic Resources

In the 1970s, the international scientific community acknowledged the disturbing extinction of traditional varieties of several cultivated plants as well as that of numerous wild populations. This extinction was linked to the abandonment of traditional varieties in favour of new, improved varieties and the rapid destruction of natural habitats, often caused by uncontrolled urbanization. On the global scale, ICRISAT (International Crops Research Institute for the Semi-Arid Tropics) was mandated to ensure the collection and conservation of genetic resources of sorghum. In 1972, the genetic resources unit of ICRISAT, in India, received around 10,000 accessions collected earlier by the Rockefeller Foundation. By the early 1990s it contained around 33,100 accessions collected from 86 countries. This collection was expected to reach 45,000 accessions by the year 2000 (Prasada Rao and Ramanatha Rao, 1995). The accessions were conserved and multiplied under the responsibility of ICRISAT at Patancheru, India. They were for the most part duplicated in the United States, where over 42,000 accessions were enumerated by the National Plant Germplasm System. They are conserved for the long term in three major sites: Fort Collins, Colorado, by the National Seed Storage Laboratory; Griffin, Georgia; and Mayguez, Puerto Rico, by the United States Department of Agriculture.

In parallel, smaller collections were built up in many countries. In China, 10,386 accessions collected from the 1950s onward were conserved by the national gene bank at Beijing. The South African collection of sorghum numbers around 4000 accessions, and the Australian collection around 3800. In France, 5850 accessions are conserved, 3850 of them by the IRD (Institut de Recherche pour le Developpement, earlier Orstom) and 2000 by CIRAD. All the collections have two weaknesses in common: the low representation of wild sorghums—only 1.2% of the United States collection and 1.1% of the ICRISAT collection—and an unequal representation of races and geographic origins. The ICRISAT collection, for example, has 17.4% guinea, 20.2% durra, 21.8% caudatum, against 2.3% kafir and 3.2% bicolor, and comes essentially from five countries: India (17.3%), Cameroon (19.2%), Yemen (8.5%), Sudan (8.1%), and Ethiopia (8.1%).

At the Indian centre of ICRISAT, the accessions were evaluated using a set of 20 morphological descriptors, of which 19 belong to the list recommended by IBPGR (1993): The 'passport' data and evaluation are complete for around 28,000 accessions On the other hand, in the United States, only 13% of the accessions of the global collection have been completely evaluated.

The management and evaluation of such collections is costly. Nevertheless, ICRISAT pursues a more complete evaluation. According to Dahlberg and Spinks (1995), efforts must also be made by the United States, where the field evaluation (40 descriptors) of all the accessions in the collection must be undertaken.

An essential step is the constitution of a data bank that is easily accessible to allow the effective use of these genetic resources in national research programmes. In the United States, the data are regularly made available via the database of GRIN (Germplasm Resources Information Network), accessible on the internet (http://www.ars-grin.gov; Dahlberg and Spinks, 1995). At ICRISAT, the database can also be consulted online (http://noc1.cgiar.org/seartype.htm).

Genetic material is distributed naturally through seeds. Between 1985 and 1995, 118,381 samples were also distributed throughout the world from the Untied States (Dahlberg and Spinks, 1995) and 237,265 by ICRISAT (Mengesha and Appa Rao, 1994). Other countries that maintain smaller collections also participate in exchanges of plant material.

In this chapter, we review studies on the genetic diversity of cultivated sorghums. Then we compare the structurations obtained by analysis of a sample using three types of markers (morphological, enzymatic, and molecular markers) and several methodologies. The results are considered in terms of application to the constitution of core collections. In conclusion, we examine the strategy of genetic resource conservation and the utility of molecular markers in understanding the organization of genetic diversity within cultivated sorghums.

ORGANIZATION OF GENETIC DIVERSITY

Genetic Diversity Revealed by Morphological Descriptors

There are few studies on the organization of genetic diversity as revealed by morphological traits. The first study, by Chantereau et al. (1989), was carried out on 157 ecotypes of widely diverse race and geographic origin. These authors showed that, on the basis of the 25 agromorphological traits studied, of which 14 figure in the IBPGR list, sorghums can be classified into three groups: the guinea and bicolor group, the caudatum and kafir group, and the durra group. These groups are distinguished primarily by their behaviour

in culture. The results reveal the adaptation of races to cultivation practices, specific uses, and biotic and abiotic constraints. For example, the guinea, like the bicolor, are hardy sorghums of wet zones, adapted to extensive cultivation. The durra are hardy sorghums of dry zones and occasionally of degrading environments, while the caudatum and kafir behave like the most modern sorghums, best adapted to semi-intensive or even intensive cultivation.

Subsequently, Appa Rao et al. (1996) studied close to 4000 accessions of sorghums originating from different Indian states and present in the ICRISAT collection using 14 morphological and agronomic descriptors. In their work, the data were not treated by multivariate analysis. Rather, the univariate descriptive analyses indicate a great morphophysiological diversity with a greater diversity among states than within states. All the races are present in India, but the durra and their intermediates are widely predominant. No information is given as to the distribution of races as a function of various traits.

In Ethiopia, Teshome et al. (1997) focused on a particular region comprising northern Shewa and southern Welo. In this restricted geographic area, the analysis of 14 morphological traits, of which 7 figure in the IBPGR list, reveals a great phenotypic diversity. Four pure races—bicolor, caudatum, durra, and guinea—and an intermediate race, durra-bicolor, were maintained by human selection. The dendrograms obtained did not reveal a clear taxonomic pattern. On the other hand, three groups could be distinguished from the multivariate analyses, mainly according to two criteria: presence or absence of sugar in the stem and plumpness of grains. This classification, different from that of Chantereau et al. (1989), relies in fact on the use of a different set of descriptors.

In our study, the analysis was conducted on a sample of 230 accessions comprising a larger proportion of guinea and caudatum but fewer durra, kafir, and intermediates than that of Chantereau et al. (1989). Twenty-one of the 25 morphophysiological traits earlier studied by Chantereau et al. (1989) were taken into account. Seventy modalities were defined from 21 variables and then a correspondence analysis (CA) was conducted on a binary table (all the active individuals and modalities). The projection of variables on the first planes of the CA show that axis 1 (12.7%) separates the small sorghums, with thick stem and glumes shorter than the grain, from the tall sorghums with long and loose panicle. Axis 2 (7.2%) isolates the sorghums with crossed peduncle and compact panicle. Axis 3 (5.4%) distinguishes the early-flowering sorghums, presenting few internodes, from the later-flowering and tall varieties. Representation of the accessions according to their racial classification in plane 1-2 shows results similar to those obtained by Chantereau et al. (1989), with a differentiation of three groups: (1) guinea and bicolor, (2) caudatum and kafir, and (3) durra.

Hierarchical clustering (HC; Fig. 3) gives a more global picture of the organization of genetic diversity and provides indications on the relationships

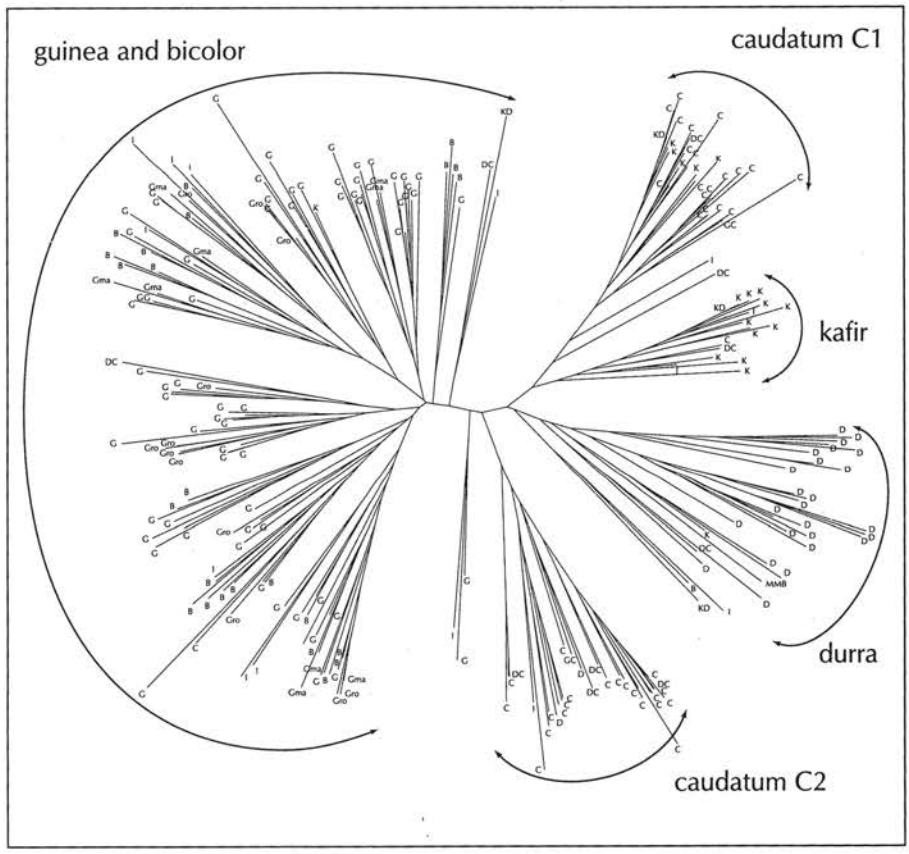

Fig. 3. Genetic diversity of cultivated sorghums revealed by morphological markers. The tree is constructed from the Sokal and Michener index.

between the groups in terms of proximity and distance. The guinea (G) and bicolor (B) form a highly variable group; the kafir (K) constitute a relatively homogeneous group; the caudatum (C), on the other hand, are separated into two subgroups, C1 and C2. The subgroup C1 is characterized by varieties presenting a short cycle and few internodes (less than 10), while subgroup C2 covers sorghums of medium cycle and a larger number of internodes. The two subgroups of caudatum appear quite divergent from each other, with distances between the subgroups that appear to be of the same magnitude as distances detected between the races; subgroup C1 is closer to kafir group than subgroup C2. The separation into two subgroups could reflect a difference in behaviour with respect to the photoperiod. Subgroup C1, close to kafir group, must thus be less sensitive or insensitive to photoperiod, while subgroup C2, close to the durra (D), must be sensitive to photoperiod (Chantereau et al., 1997). This sub-structuration of caudatum passed unnoticed in the study of Chantereau et al. (1989) and in the correspondence

analysis (CA) presented above, which can be explained by the fact that the two subgroups, located on either side of axis 1 in the 1-3 plane of our CA (Fig. 4), form an apparently continuous group.

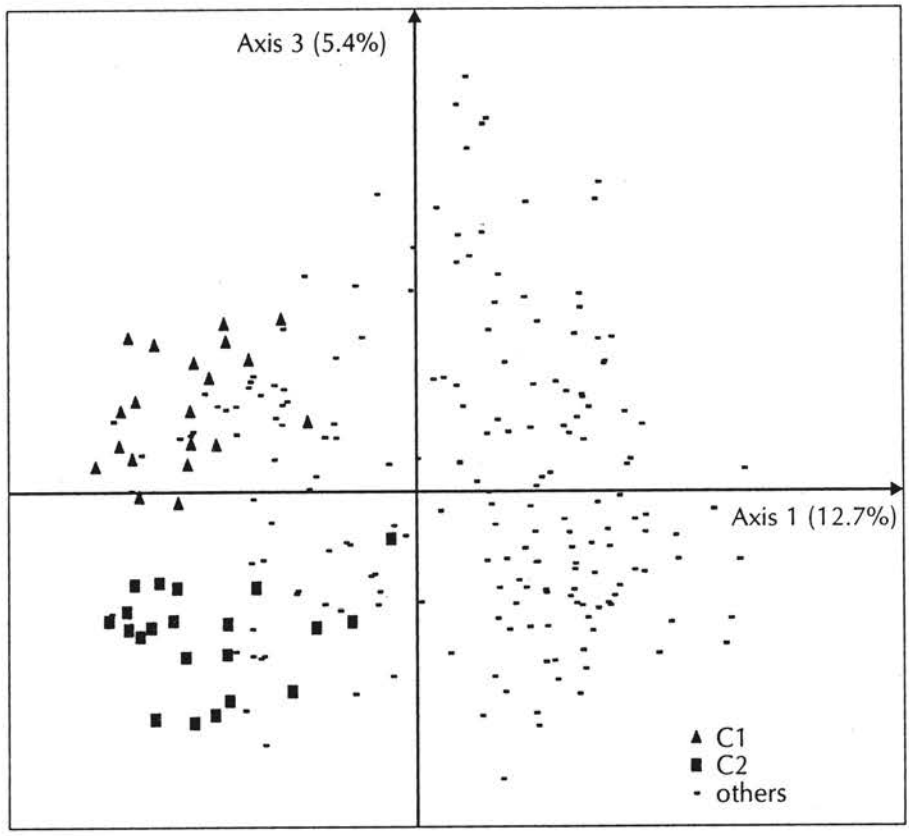

Fig. 4. Correspondence analysis on morphological traits.

No group is totally homogeneous with respect to the racial classification. As a general rule, individuals belonging to intermediate races are distributed in different groups depending on their degree of likeness to sorghums of the purest race. This observation could be reflected in a dynamic in the evolution of sorghums linked to several genetic exchanges favoured by traditional cultivation practices (association of several varieties, even several races, in a single field) and to various natural and human selection pressures. It would be interesting to observe the relative position of these intermediates in relation to other criteria of classification—such as those used by Teshome et al. (1997), or criteria of technological quality of grain, resistance to biotic and abiotic stresses—in order to better understand their role in the evolution of sorghums. Do they contribute to the widening of the genetic variability while evolving towards a parental type? Are they maintained or constantly produced and

eliminated? Their competitiveness with pure races is very high in Chad when judged by the nature of the races that are traditionally cultivated in this region (Yagoua, 1995). On the other hand, the situation seems different in Ethiopia, where Teshome et al. (1997) indicate the presence of five cultivated races, of which four are 'pure' races.

Genetic Diversity Revealed by Enzymatic Markers

The analysis of enzymatic polymorphism of cultivated sorghums was the focus of studies conducted in the United States and in France for over ten years (Morden et al., 1989; Ollitrault et al., 1989b; Aldrich et al., 1992; Degremont, 1992).

The study of Morden et al. (1989), complemented by that of Aldrich et al. (1992), shows that most of the total genetic diversity is due to differences of geographic origin rather than race. They also indicate that the western and eastern African regions have the highest level of heterozygosity and that southern Africa has the narrowest genetic diversity.

For Ollitrault et al. (1989b), this geographic differentiation seemed more marked. These authors distinguished three geographically distinct groups: a western African group, an eastern and central African group, and a southern African group. Moreover, structuration according to the geographic zones was observed for two races, bicolor and guinea. The differentiation of three distinct groups within the guinea sorghums—guinea of West Africa, guinea of southern Africa, and guinea of the margaritiferum type—was subsequently largely demonstrated by the in-depth analyses of Degremont (1992).

Our study was conducted on 230 accessions representative of the geographic and racial diversity. A good genetic diversity was revealed in the sample, with 11 polymorphic loci (at a threshold of 99%) and a mean number of alleles per polymorphic locus of 2.8. The pattern of diversity indicated by the AHC (Fig. 5) is in agreement with the western-southern African geographic differentiation noted by Ollitrault et al. (1989b). Moreover, the study leads to the following conclusions. The accessions of East Africa and Central Africa present a wide range of variability and do not form a well-differentiated group. No group is totally homogeneous, racially or geographically. The guinea sorghums of southern Africa present a narrower genetic diversity than those of West Africa. The guinea sorghums of southern Africa are closer to kafir than to any other group. The guinea sorghums of Central and East Africa are found preferentially in the West African group. The guinea margaritiferum (Gma) form a highly homogeneous group that is clearly distinct from all the other guinea. Very curiously, they are relatively closely related to a small group of 12 accessions, of which two thirds are the caudatum sorghums belonging in equal parts to subgroups C1 and C2 described earlier.

While structuration using morphological descriptors indicates groups with common cropping performance, geographical structuration obtained

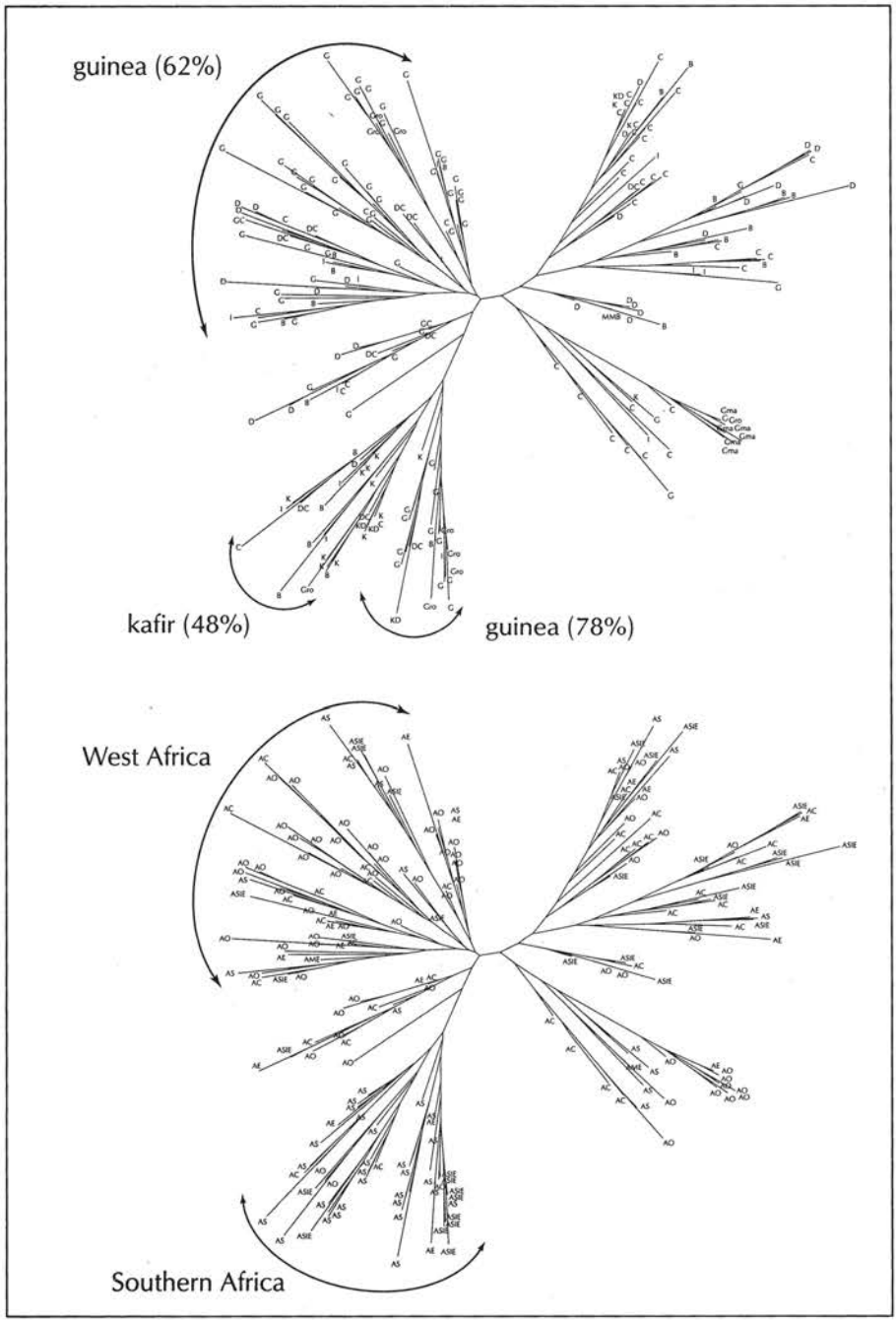

Fig. 5. Genetic diversity of cultivated sorghums revealed by enzymatic markers. The racial affiliation is indicated on the tree above and the geographic origin on the tree below. The tree is constructed from the Dice index.

using enzymatic markers indicates in its major lines the areas of distribution preferred by different races: guinea in West Africa, guinea and kafir in southern Africa, and durra and caudatum in East and Central Africa. However, the fact that durra of West Africa, for example, are found in the group of guinea sorghums of West Africa could indicate that these durra sorghums (barring labelling or identification errors) are of ancient introduction in West Africa. They could thus have followed a different evolutionary path, under the constraints of West Africa, from that of the other durra and integrated themselves totally in the West Africa group because of gene flow. On the contrary, the caudatum sorghums recently introduced in West Africa still retain the genotypic structures linked to their Central and East African origin and are not found in the West Africa group.

Genetic Diversity Revealed by Molecular Markers

Genetic diversity in sorghum was first studied by Aldrich and Doeley (1992) using RFLP markers (38 probes) and by Tao et al. (1993) using RFLP markers (16 probes) and RAPD markers (29 primers). These studies on restricted samples, of fewer than 50 accessions, do not enable us to observe a marked racial or geographic pattern of genetic organization. On the other hand, the authors conclude that RFLP and RAPD markers indicate a greater allelic diversity than enzymatic markers. However, the capacity of markers to reveal a structuration depends on whether there is a structuration, on the representativity of the sample analysed, and on the number and type of markers used. Aldrich and Doebley (1992) studied 31 accessions representing the five major races equally but sampled from 10 countries distributed unequally between the four major regions of Asia and Africa. Similarly, the 36 accessions studied by Tao et al. (1993) are not representative of the racial diversity or of large areas of distribution of sorghum.

Deu et al. (1994) used 33 RFLP probes to study a sample of 94 accessions, taking into account the racial and geographic diversity of sorghums. The study indicated a southern African group, the genetic variability of which is more restricted than that of West Africa or East and Central Africa. This study also revealed a racial differentiation. Except the bicolor accessions, which do not constitute a homogeneous group, the caudatum, durra, and kafir races form three distinct groups, while the guinea are divided into three subgroups: the guinea of West Africa, those of southern Africa, and guinea margaritiferum. This study confirms the internal structuration of guinea observed with the isozymes and allows a racial differentiation not revealed by morphological traits (caudatum and kafir) or enzymatic markers (caudatum and durra).

A study conducted by Cui et al. (1995) with 61 RFLP probes on 41 accessions shows a less clear racial differentiation since the kafir and guinea of southern Africa and Asia form a single group. The study also indicates the

originality of a guinea margaritiferum, which is found to be more closely related to the wild sorghums than to the cultivated sorghums. For the authors, this grouping is consistent with the rather 'wild' phenotype of this accession.

The singularity of guinea margaritiferum is also mentioned in a study by de Oliveira et al. (1996) using 20 RFLP probes, 13 RAPD primers, and 4 ISSR: the three accessions of this type analysed differed from other cultivated sorghums mainly by the absence of common alleles found in the rest of the collection rather than specific alleles unique to the three accessions.

An important study was conducted by Menkir et al. (1997) on 190 accessions representative of five principal races, with widely varying geographic origins—13 countries for the kafir and 28 to 32 countries for the other races. The analyses performed using 82 RAPD primers show that genetic diversity is much greater within the bicolor and the guinea than in the other races. The kafir sorghums are the least diversified. Moreover, 86% of the total variability occurred within a race and 14% among races. Only 13% of the total variability is linked to geographic differentiation. In these conditions, the authors do not observe racial or geographical pattern of genetic diversity. These results, which do not agree with those of Deu et al. (1994), could be explained either by the mode of sampling or by the nature of markers used.

First, the sampling of geographical origins for each race would be too artificial and not reflect the traditional distribution of races—the countries of origin are not specified in the study. Human migrations and exchanges of material favour contacts and gene flows between races, which help widen the genetic base of each race. Nevertheless, even though there are numerous intermediates, at various degrees, the major racial types persist in cultivation. In northern Shewa and southern Welo, in Ethiopia, four of the five major races coexist—bicolor, caudatum, durra, and guinea (Teshome et al., 1997). Although genetically close, the kafir and guinea of southern Africa retain different morphological characteristics. This could indicate the existence of very strong selection pressures on the morphological traits and thus the fact that the major races correspond to types with high ecogeographical specialization. The recent migrations, on their part, help to artificially enlarge the variability within regions and reduce the interregional differentiation.

Second, the RAPD markers used in this study are mostly multilocus, with an average of 4.2 bands revealed per primer, while close to 75% of the probes tested in RFLP by Deu et al. (1994) correspond to single copy loci. In these conditions, the RAPD could correspond to non-coding regions or to repeat sequences in the genome, while the RFLP would correspond rather to single copy, coding sequences. The rates of evolution of these regions are probably different and, because of this, these two types of markers do not result from the same evolutionary processes. Moreover, phenomena of homoplasy, more frequent with RAPD than with RFLP, could account for observed differences in structure (Powell et al., 1996).

Comparison of Genetic Diversity Organization Revealed by the Three Types of Markers

Since RFLP data are available for 92 accessions already characterized by morphological and enzymatic markers, we first compared the organization of genetic diversity obtained on this subgroup with those revealed by study of the initial sample of 230 accessions.

For morphological diversity, all the modalities found in the initial collection, including those of low frequency, are presented in the subgroup. The radial representations obtained in the two cases (230 and 92) are very similar. The group showing the greatest diversity is made up of guinea and bicolor. Four other groups are observed: one group of durra, one of kafir, and two of caudatum. Eighteen accessions occupy positions intermediate between these groups.

The 31 alleles detected from the 11 enzymatic loci in the 230 accessions were conserved in the subgroup. The structuration observed shows three major identifiable groups: the guinea margaritiferum, a group from southern Africa including the kafir and guinea, and a group comprising the caudatum and durra. Between these three groups can be seen the presence of numerous accessions occupying intermediate positions. The West Africa group that can be observed in the initial sample does not appear clearly in the subgroup.

Genetic structuration revealed by the three types of markers has been compared on this subsample of 92 accessions.

The average numbers of alleles per polymorphic locus appear equivalent whatever the nature of marker considered: 2.8 for isozymes against 2.9 for RFLP (if the restriction enzyme is *Hind*III) and 3.0 (with *Xba*I). In this particular case, RFLP does not allow access to a greater allelic polymorphism than the isozymes.

We were able to identify 11 polymorphic loci for the isozymes and 33 for the RFLP. Even more than the percentage of polymorphic loci, what seems interesting is the number of polymorphic loci that are relatively easily accessible. In the study of Ollitrault et al. (1989b) on 348 accessions, 18 loci out of the 25 tested appeared polymorphic. It thus seems difficult to achieve more than 20 enzymatic polymorphic loci, while Cui et al. (1995) used 61 RFLP polymorphic loci.

In our study, the 11 enzymatic loci enabled identification of 83.6% of genotypes (or 77 out of 92), while the 33 RFLP loci led to 94.5% genotypic identification. Nevertheless, in some cases, the isozymes discriminate between accessions not differentiated by RFLP markers. Thus, the enzymatic loci *EST-D* allow differentiation of accessions of each of two pairs of accessions that appear identical according to RFLP. Two other accessions identical with the RFLP have different alleles at three enzymatic loci (*END*, *HEX* and *LAP*). The combination of isozyme markers and RFLP allows us to discriminate all the accessions with the exception of two pairs formed by two caudatum on

the one hand and two guinea roxburghii (Gro) on the other. In these conditions, it would be interesting to identify the minimal combination of isozyme and RFLP markers that allows identification of all the genotypes, and then to test its efficiency on other groups of accessions.

The comparison of structures revealed by the different markers shows that the great morphological diversity of guinea sorghums is accompanied by great genetic diversity, with three groups of differentiation clearly marked by the RFLP. In addition, it indicates that the caudatum and durra sorghums, well differentiated in morphological terms, constitute two groups whose genetic proximity is indicated by isozyme markers and RFLP. Moreover, the kafir sorghums, relatively homogeneous from the morphological point of view, are genetically closer to guinea of southern Africa than to other sorghums.

In fact, although the organization of genetic diversity obtained with the three types of markers cannot be perfectly superimposed, they are consistent with the information acquired on sorghum.

- Sorghums are preferentially autogamous, but existing natural allogamy, even though sometimes low, is favoured by traditional cultivation practices. Because of this, gene exchange enhances the morphological diversity (within the limit of selection practised) and above all genetic diversity for the isozymic and molecular characters that are inherently neutral with respect to selection.
- The guinea margaritiferum are the most differentiated among the cultivated sorghums. Deu et al. (1995) showed that they present a mitochondrial polymorphism that distinguish them from all the other sorghums, cultivated or wild, belonging to the species *S. bicolor*. Because of their genetic and agromorphological characteristics, these guinea sorghums could sustain the interest of breeders.
- The guinea of southern Africa are genetically closer to kafir than to other guinea. Because these two groups of sorghum share the same area of distribution, the gene flows occur naturally. The intragroup diversity increases at the expense of intergroup diversity.

There are other methods for comparing the information given by various markers. The use of the Mantel test and the construction of consensus trees or common minimal trees are examples. In the first case, correlation coefficients, r, between the matrixes of similarity indexes produced from different markers are calculated. Thus, the value of distances between individuals is emphasized. The calculated significance of r is totally subjective. According to Rohlf (1990), the correlation could be considered very high if r is greater than or equal to 0.9, high if r is between 0.8 and 0.9, low if r is between 0.7 and 0.8, and very low if r is lower than 0.7. Several studies refer to the use of this test (Messmer et al., 1991; Engqvist and Becker, 1994;

Thormann et al., 1994; Powell et al., 1996). However, few studies take into account the enzymatic markers. Engqvist and Becker (1994) calculated the correlation coefficients obtained in *Brassica napus* from RFLP, RAPD, and enzymatic markers. The lowest coefficient was obtained with RFLP-isozyme pairs (0.53 against 0.67 for the RAPD-isozyme pair and 0.76 for the RFLP-RAPD pair). Similarly, in maize, Messmer et al. (1991) obtained a low correlation ($r = 0.23$) for the RFLP-isozyme pairs.

In our case, in accordance with the patterns of diversity observed, the correlations are very low: r equals 0.19 for the isozyme-morphology pair, 0.2 for the RFLP-morphology pair, and 0.3 for the RFLP-isozyme pair.

Comparison by construction of trees considers the structures obtained without taking into account the edge lengths.

In our case, the consensus trees obtained are of the star or 'open umbrella' type, which indicates the absence of strictly identical groups (common structures) between the trees produced from different markers. This constraint seems too great in our case and not appropriate to the biological reality. If two groupings differ only by the presence or absence of an individual (poorly classified for whatever reason), the edge is not conserved in the consensus tree.

In contrast, the construction of common trees takes into account the fact that all the individuals could not be correctly represented in all the trees (different sources of possible error). This method allows us to find the largest subgroup of individuals forming the same structure in the trees compared.

According to Perrier et al. (1999), at a threshold of 5% with a population of around 100 individuals, a maximum of 19 points common between two trees can be obtained at random. In our case, the three trees thus constructed have 19 to 24 common points (Figs. 6 and 7). This result is in agreement with the low correlation coefficients calculated between the distance matrixes. Nevertheless, in considering the structures common to trees constructed from morphological markers and RFLP (Fig. 6), we note that the edges in caudatum (C2) and durra are present in the two trees; similar observations can be made for those that appear in part of the guinea. It may seem surprising not to find the kafir group in the common structures. In fact, this group is more homogeneous in genetic terms than in morphological terms. The individuals closest morphologically are not genetically the most similar.

Similar observations can be made during comparison of trees constructed from morphological and enzymatic markers (Fig. 7). However, if some large structures are conserved (opposition between guinea and caudatum, durra), the same individuals in these two common trees do not constitute the framework of the common structures. This observation takes into account the fact that it is not possible to construct a consensus tree.

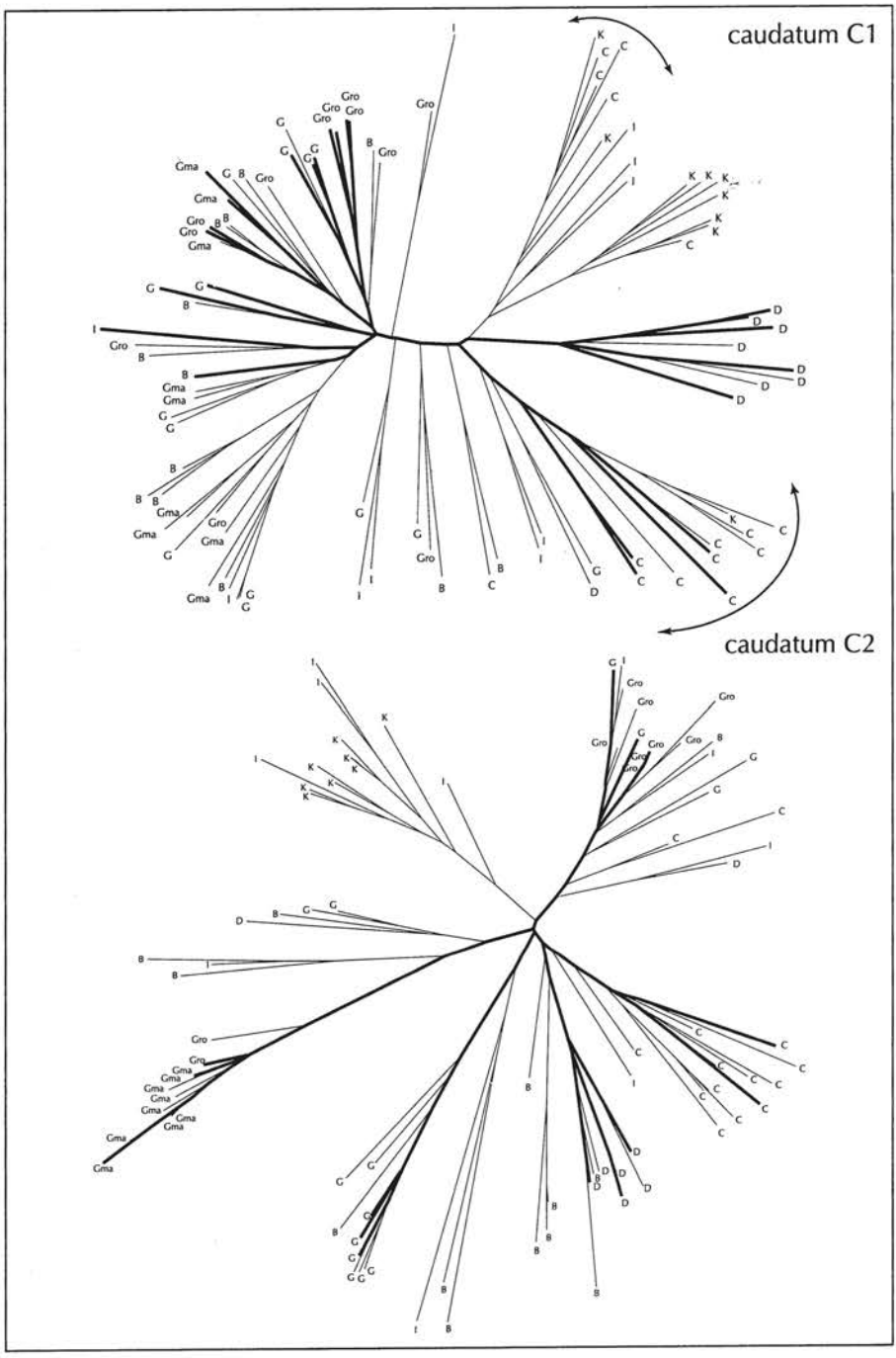

Fig. 6. Representation of common edges (bold) in trees constructed from morphological data (above) and molecular data (below).

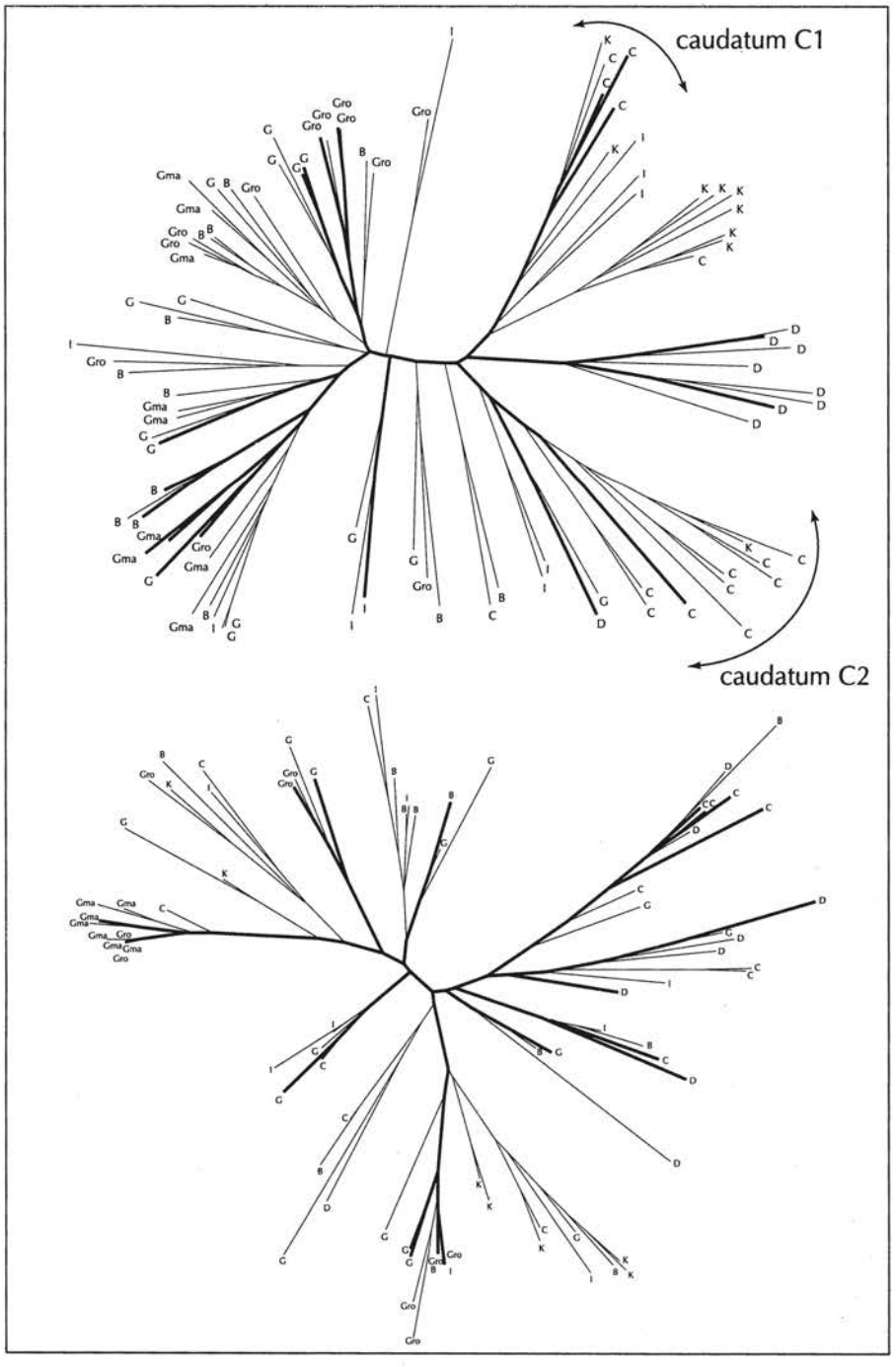

Fig. 7. Representation of common edges (bold) in trees constructed from morphological data (above) and enzymatic data (below).

Application to Constitution of Core Collections

In order to facilitate the use of large collections, the concept of core collection was developed by Frankel and Brown (1984) and subsequently by Brown (1989). Various sampling strategies were proposed (Hamon et al., 1995, for review). They were applied to several Poaceae, such as barley (van Hintum et al., 1990), ryegrass (Charmet et al., 1993), and durum wheat (Spagnoletti-Zeulli and Qualset, 1993).

For sorghum, collections of limited size were developed from the early 1990s onward. A basic collection was constituted in India by ICRISAT. It comprised 1400 accessions selected according to criteria of taxonomy, geographic distribution, and ecological adaptation, as defined by Harlan (1972). Very rapidly, this collection posed practical problems due to photoperiodism, a character that was not taken into account during the constitution of the collection. At Patancheru, many tropical ecotypes, notably from Cameroon, Ethiopia, Nigeria, and Sudan, did not flower at all or flowered too late. This collection was thus declared to be specific to one location and poorly representative of the phenotypic diversity of the world collection. It was then decided that other, smaller collections would be created.

A core collection comprising only cultivated sorghums was then established at ICRISAT by Prasada Rao and Ramanatha Rao (1995) with two objectives: to facilitate access to the world collection and to better represent the genetic diversity of cultivated sorghums.

For this, the world collection was stratified in subgroups defined according to geographic, taxonomic, and agronomic criteria. Then, other clusters within subgroups were defined through a principal components analysis on seven characters—cycle from seed to spike formation, plant height, inflorescence exsertion, inflorescence length and width, grain covering, and 100 grain weight. Taking into account the two modes of structuration of the initial collection, a sample proportionate to the number of accessions present in each subgroup was taken. This core collection comprised 3475 accessions, or around 10% of the ICRISAT collection.

The process of stratification before the selection aimed to take into account a possible pattern (geographic or taxonomic, for example) of the genetic diversity. Prior knowledge of the extent and pattern of genetic diversity in the gene pool considered is indispensable to maximize the genetic diversity to be preserved in a core collection. In this context, are the seven morphological characters taken into account for the constitution of the core collection sufficient and relevant to obtain a good representation of the overall diversity of sorghums? Various studies on the genetic diversity of sorghums, including ours, indicate that phenotypic and genetic divergence are not totally in agreement.

Using the PCS (principal component score) strategy developed by Noirot et al. (1996) and Hamon et al. (1998), we have constituted, from the initial

sample of 230 accessions, two cores by using either morphological or enzymatic data. These two core collections were compared with the initial sample in order to understand the modifications induced by the process of selection.

The curves representing the cumulative number of individuals maximizing the diversity as a function of cumulative percentage of the total diversity show that 50% of the total enzymatic diversity is represented by 50 accessions, while 77 accessions are necessary to obtain 50% of the phenotypic diversity. These two sets of accessions constituted respectively CoreI and CoreM.

With respect to the distribution of botanical races (Table 1), the χ^2 test indicates significant differences between the initial sample and the two core collections. In particular, it is noted that the kafir race is not represented in CoreI. This observation is in agreement with the low genetic diversity of the kafir race noted by various authors. On the other hand, the large geographic regions are all represented in the two core collections without significant modifications in relation to the initial sample (Table 2).

For the allelic frequencies, as expected, highly significant differences (χ^2 test, $P < 0.01$) were observed for CoreI, while the differences were not significant for CoreM. However, in CoreI, the global divergence is due to two loci—endopeptidase and acid C phosphatase—affected by the selection. In CoreM, selection based on morphological characters did not enable conservation of all the rare alleles: three alleles out of five of frequency less than 5% were absent (Table 3). Nevertheless, it can be asked what will be the future of these rare alleles subjected to natural selection.

For phenotypic diversity, the situation is logically the inverse: the differences are not significant with CoreI but are highly significant with CoreM. In this case, the global divergence is due to significant differences (at least at a threshold of 5%) for 12 characters out of the 21 studied.

These comparisons indicate that the selection on neutral, enzymatic markers (CoreI) is interesting because it allows conservation of allelic diversity while preserving the initial phenotypic diversity. However, the large collections are subjected to more or less systematic morphoagronomic evaluations. It is thus also important to consider these descriptors for constituting core collections. For this purpose, one could question whether all the information available in the initial collection is useful to take into account. In other words, is there redundant information? Is it possible to define associations between markers? Some answers can be provided by the use of the Cramer coefficient of association V, calculated for each pair of variables (Bishop et al., 1975). The coefficients calculated from data of the initial sample (230 accessions) have values between -0.3 and + 0.65 (this coefficient varies theoretically between -1 and +1). When the value 0.5 is taken for the lower significant limit of association, associations are detected for five groups of variables: the brown layer (Cbr) and vitreousness (Vit);

leaf width (Laf) and stem diameter (Dtp); days to flowering (Nje) and number of internodes (Nen); compactness of panicle (Cpa) and length of panicle (Lpa) and of peduncle (Lpe); and grain shape (Fgr) and brown layer (Cbr) as well as length and form of the peduncle (Lpe, Fpe).

Table 1. Racial distribution in the initial sample and in the two core collections constituted by the PCS method (number and percentage of different races)

	bicolor	caudatum	durra	guinea	kafir	others
Initial sample	25 (10.9)	41 (17.8)	27 (11.7)	88 (38.3)	18 (7.8)	31 (13.5)
CoreI	4 (8)	5 (10)	10 (20)	25 (50)	0	6 (12)
CoreM	13 (16.9)	6 (7.8)	16 (20.7)	32 (41.6)	2 (2.6)	8 (10.4)

$\chi^2_{CoreM} = 12.4$ (ddl = 4), significant at 5%.
$\chi^2_{CoreI} = 9.53$ (ddl = 3), significant at 5%.

Table 2. Geographic distribution of accessions in the initial sample and in the two core collections (number and percentage of accessions)

	West Africa	East Africa	Central Africa	Southern Africa	Asia	America
Initial sample	74 (32.2)	19 (8.3)	39 (16.9)	55 (23.9)	39 (16.9)	4 (0.02)
CoreI	21 (42)	5 (10)	7 (14)	9 (18)	7 (14)	1 (2)
CoreM	26 (33.8)	7 (9)	10 (13)	14 (18.2)	18 (23.4)	2 (2.6)

$\chi^2_{CoreM} = 3.17$ (ddl = 4), ns.
$\chi^2_{CoreI} = 9.5$ (ddl = 3), ns.

Table 3. Total number of alleles of different frequency categories in the initial sample and in the two core collections

	Total number of alleles of different frequency categories in		
Frequency of alleles (x)	initial sample	CoreI	CoreM
$x < 0.05$	5	5	2
$0.05 \leq x < 0.1$	2	2	2
$0.1 \leq x < 0.2$	3	3	3
$0.2 \leq x < 0.4$	9	9	9
$x > 0.4$	12	12	12

Given that this initial sample does not correspond to a random sampling from the world collection, we could ask whether the associations indicated are fortuitous (sampling bias), or produced by linkage disequilibrium (foundation effect, genetic drift), or linked to the genetic organization of the genome, or represent coadapted gene complexes.

The genome mapping of different traits could provide elements of reflection even if the analysis of progeny is not comparable to that of a population.

Some of these traits have been mapped (Rami et al., 1998). It is interesting to note that the gene *B2/b2* of the brown layer and a major QTL for vitreousness (Vit, $r^2 = 54\%$) were co-located on the linkage group F. Similarly, two QTL for compactness and length of panicle are co-located on the linkage group A (Cpa, $r^2 = 22\%$, Lpa, $r^2 = 35\%$) and on the linkage group F (Cpa, $r^2 = 13.5\%$, Lpa, $r^2 = 20\%$). Note that these QTLs mapped on the linkage group F are genetically independent of the *B2/b2*-Vit pair.

In agreement with the results obtained on our two core collections, no association (coefficient $V \geq 0.5$) was noted between morphological and enzymatic markers. Similarly, no association was revealed between the enzymatic markers at the fixed threshold. Despite the absence of major groups of linkages between the enzymatic loci (Ollitrault et al., 1989a), two situations are interesting to consider: the loci *LAP-2* and *PA-B*, 12 cM apart, which present a Cramer coefficient V of 0.24; and the locus pairs *HEX* and *LAP*, *DIA* and *LAP*, *AMY* and *EST-C*, with coefficients of association 0.45, 0.36, and 0.34 respectively and genetically independent. In considering CoreI and CoreM, it can be observed that, for the pairs *AMY-EST* and *HEX-LAP*, the associations are maintained and are not random (exact Fisher test significant for CoreI, highly significant for CoreM, in both cases). For the pairs *DIA-LAP* and *PAB-LAP*, on the contrary, the associations are ruptured in the two core collections (exact Fisher test not significant).

The mapping of all the morphological and enzymatic markers as well as the search for association including molecular markers is still to be probed. They will certainly provide useful information for fine-tuning the effect of markers to be taken into account in the constitution of core collections and for evaluating the relative contributions of coadapted structures to be included in the core collection and linkage disequilibrium.

CONCLUSION

Studies on genetic diversity show first of all that there is great variability among cultivated sorghums. Eleven enzymatic loci are polymorphic in our study, 13 in that of Morden et al. (1989), and 18 in that of Ollitrault (1989b). Close to 75% of heterologous maize probes, in combination with two restriction enzymes, hybridize with sorghum DNA and reveal polymorphism

(Deu et al., 1994). In this study, each locus (enzymatic or revealed by RFLP) is represented on average by 2.8 to 3 alleles.

Molecular RFLP markers allow us to discriminate four to five major botanical races described by Harlan and de Wet (1972). These races have a rather large and structured genetic variability. The kafir race, for example, seems highly homogeneous, while strong differentiation is observed within the guinea, partly due to secondary centres of diversification.

The patterns of genetic organization observed using enzymatic and morphological markers do not perfectly agree with the racial classification. Some races constitute groups that have common behaviours in cultivation (bicolor and guinea on the one hand, caudatum and kafir on the other, with morphological descriptors) or a common area of origin (caudatum and durra, with enzymatic markers).

Generally, molecular and enzymatic markers are considered neutral with respect to selection. Do the differences in the pattern of genetic diversity revealed by the two types of markers lie in the number of loci considered, in their genomic location, or in their own specificity? The study of genetic diversity revealed by analysis of a large number of RFLP loci very regularly distributed on the sorghum genome will certainly help answer these questions.

It has not been possible to construct a consensus tree after taking into account different morphological and enzymatic markers, even though some common elements of structure could be observed. This result indicates the absence of a strong structuration within the cultivated sorghums. Neither racial affinity, nor area of origin, nor cropping performance in cultivation led to the establishment of strong mating barriers. The presence of individuals in intermediate position in the tree representations indicates the existence of gene flows. Thus, in natural conditions, gene exchanges occur and, even though they occur at a low rate, they contribute to local broadening of the genetic diversity. At the same time, strong selection pressures are exerted, allowing the major races to be maintained where they are of particular interest. Such dynamic traditional management favours the establishment of sites reserved for *in situ* and participatory conservation of sorghum genetic resources. This aspect is important to consider since during propagation of the plant material there is the problem of genetic drift linked to self-fertilization of populations that are not always fixed, even though they are preferentially autogamous. This absence of strong structuration also allows us to consider the cultivated sorghums as a single gene pool.

The PCS strategy has thus been applied directly on the initial sample (without prior stratification) to constitute core collections, considering the morphological traits and then the enzymatic markers. Examination of two core collections constituted in this manner shows that 22% of the accessions for CoreI and 33.4% for CoreM are needed to represent 50% of the total initial diversity. The morphological diversity thus seems more fragmented than the enzymatic diversity. However, this comparison remains very limited to

the extent that the number of loci considered is different in each case. It would be particularly interesting to have access to RFLP data in order to have another measure of the distribution of genetic diversity in the initial sample.

Besides, examination of two core collections indicates that there is no association between the two types of characters. In our study, selection on the enzymatic markers has not affected the initial morphological variability. Selection on morphological traits leads to the loss of some rare alleles at certain enzymatic loci. It would thus be interesting to consider simultaneously the two types of characters to constitute core collections. In these conditions, one could consider a morphological selection on the enzymatic groups previously obtained or, on the contrary, an enzymatic selection on the morphological groups. On the other hand, some associations within morphological traits and within enzymatic markers have been detected, which could indicate some redundancy of information. Investigations must be pursued to better understand the origin of these associations and to define sets of informative characters. Genome mapping provides reflective trails and will certainly allow us to optimize the tools for evaluation of diversity. Molecular markers are of great interest in sorghum because they are susceptible to mapping (the enzymatic loci are almost all genetically independent). Besides, they allow us to study the determinism of different characters no matter what their degree of complexity. Molecular markers and genome mapping are particularly promising tools for genetic resource management.

In the case of sorghum, a good number of agromorphological traits have been mapped or are being mapped (Lin et al., 1995; Pereira et al., 1995; Pereira and Lee, 1995; Rami et al., 1998). The access to information on genomic order will necessarily enhance the 'legibility' of genetic resources.

APPENDIX

Plant Material

The initial sample of cultivated sorghum (*S. bicolor* ssp. *bicolor*) comprised 230 accessions, of which 136 were studied by Chantereau et al. (1989) and Ollitrault et al. (1989b) and 63, mainly of the guinea race, by Degremont (1992). These sorghums are traditional varieties from the ICRISAT or CIRAD collections. Morphological and enzymatic data were available for the 230 accessions.

RFLP analysis was done on 92 accessions, chosen from the 230 according to two criteria: geographic origin and racial classification. Seventy-four were in the sample analysed by Deu et al. (1994); most of the 18 others belonged to the guinea race.

Enzymatic Study

Analysis was done on 8 enzymatic systems indicating 11 polymorphic loci: alcohol dehydrogenase (ADH), amylase (AMY), diaphorase (DIA), endopeptidase (END), esterase (EST), hexokinase (HEX), leucine aminopeptidase (LAP), and acid phosphatase (PA). The experimental protocols as well as the genetic interpretations of zymograms are described by Ollitrault et al. (1989a) and Degremont (1992).

Morphological Study

The experimental set-up is described by Chantereau et al. (1989) and Degremont (1992). The analysis was done on 21 morphological characters common to the two studies. The descriptors marked by an asterisk belong to the list recommended by the IBPGR. Ten qualitative variables were retained: anthocyan of leaves (Ant)*, aristation (Ari)*, brown layer of grain (Cbr)*, grain colour (Cgr)*, panicle compactness (Cpa)*, form of grain (Fgr)*, form of peduncle (Fpe), length of glumes (Log)*, opening of glumes (Oug), and grain vitreousness (Vit)*. Eleven quantitative variables were measured on the principal stem: stem diameter (Dtp), stem height (Htp)*, panicle length (Lpa)*, width and length of the third subpanicle leaf (Laf and Lof), length of peduncle (Lpe), number of internodes (Nen), number of days between seed and 50% earing (Nje)*, number of fertile tillers (Ntu)*, weight of grains per panicle (Pgp), and 500-grain weight (P5g)*.

RFLP Studies

Thirty-one genomic probes of maize, which correspond to 50 probe-enzyme combinations revealing polymorphism, were used. These probes, distributed

throughout the genome, represent single copy or very slightly repeated sequences. The probe-enzyme combinations were identical to those described by Deu et al. (1994), with just two exceptions: the combination Umc 38-*Hind*III was eliminated, and the pair Bnl 7.49-*Xba*I was added.

Treatment of Data

For enzymatic data and RFLP, the CA (Benzecri, 1973) was performed on a binary table on which each allele or band is coded by two variables: presence and absence. For the morphological traits, in order to treat the qualitative and quantitative traits jointly, the latter were transformed into qualitative variables by coding into 2 to 4 classes, according to the distribution of each trait. These multivariate analyses were performed using Addad software.

The distances between individuals were calculated using the Dice index of similarity (1945) for the enzymatic data and RFLP, and the Sokal and Michener index (or simple matching) for the qualitative morphological data. The similarity matrixes were calculated and compared according to the Mantel test using Ntsys software PC version 1.80. Hierarchical clustering was done for the three types of data using as a criterion of aggregation the neighbour joining method developed by Saitou and Nei (1987) and implemented in the Darwin software created by CIRAD. The consensus trees and common minimal trees (Perrier et al., 1999) were constructed using Darwin software.

The Cramer coefficients of association V (Bishop et al., 1975) were calculated using Sas software. The PCS method (Noirot et al., 1996) was applied to the collection of 230 accessions using enzymatic and morphological data separately. Selection on enzymatic data enabled definition of CoreI, and that on morphological data enabled definition of CoreM.

REFERENCES

Aldrich, P.R. and Doebley, J. 1992. Restriction fragment variation in the nuclear and chloroplast genomes of cultivated and wild *Sorghum bicolor*. *Theoretical and Applied Genetics*, 85: 293-302.

Aldrich, P.R., Doebley, J., Schertz, K.F., and Stec, A. 1992. Patterns of allozyme variation in cultivated and wild *Sorghum bicolor*. *Theoretical and Applied Genetics*, 85: 451-460.

Appa Rao, S., Prasada Rao, K.E., Mengesha, M.H., and Gopal-Reddy, V. 1996. Morphological diversity in sorghum germplasm from India. *Genetic Resources and Crop Evolution*, 43: 559-567.

Arriola, P.E. 1995. Crop to weed gene flow in *Sorghum*: implications for transgenic release in Africa. *African Crop Science Journal*, 3(2): 153-160.

Benzecri, J.P. 1973. *L'Analyse des Données, Vol. II. L'Analyse des Correspondances*. Paris, Dunod, 616 p.

Bishop, Y.M.M., Fienberg, S.E., and Holland, P.W. 1975. *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, Mass., Massachusetts Institute of Technology Press.

Brown, A.H.D. 1989. Core collections: a practical approach to genetic resources management. *Genome*, 31: 817-824.

Chantereau, J., Arnaud, M., Ollitrault, P., Nabayaogo, P., and Noyer, J.L. 1959. Etude de la diversité morphophysiologique et classification des sorghos cultivés. *L'Agronomie Tropicale*, 44(3): 223-232.

Chantereau, J. and Kondombo, C. 1994. Estimation des taux d'allogamie chez les sorghos de la race guinea. In: *Progress in Food Grain Research and Production in Semi-Arid Africa*. J.M. Menyonga et al. eds., Ouagadougou, Burkina, SAFGRAD, pp. 309-314.

Chantereau, J., Vaksmann, M., Bahmani, I., Ag-Hamada, M., Chartier, M., and Bonhomme, R. 1997. Characterization of different temperature and photoperiod responses in African sorghum cultivars. In: Amélioration du sorgho et de sa culture en Afrique de l'Ouest et du Centre, A. Ratnadass et al. eds., Montpellier, France, CIRAD, pp. 29-35.

Charmet, G., Balfourier, F., Ravel, C., and Denis, J.B. 1993. Genotype × environment interaction in a core collection of French perennial ryegrass populations. *Theoretical and Applied Genetics*, 86: 731-736.

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. *Theoretical and Applied Genetics*, 90: 787-796.

Dahlberg, J.A. and Spinks, M.S. 1995. Current status of the US sorghum germplasm collection. *International Sorghum and Millets Newsletter*, 36: 4-12.

de Wet J.M.J. 1978. Systematics and evolution of *Sorghum* sect. *Sorghum* (Graminae). *American Journal of Botany*, 65(4): 477-484.

de Wet J.M.J., Harlan, J.R., and Price, E.G. 1976. Variability in *Sorghum bicolor.* In: *Origins of African Plant Domestication.* J.R. Harlan et al. eds., The Hague, Mouton, pp. 453-463.

de Oliveira, A.C., Richter, T., and Bennetzen, J.L. 1996. Regional and racial specificities in sorghum germplasm assessed with DNA markers. *Genome*, 39(3): 579-587.

Degremont, I. 1992. Evaluation de la diversité génétique et du comportement en croisement des sorghos (*Sorghum bicolor* L. Moench) de race guinea au moyen de marqueurs enzymatiques et morphophysiologiques. Doct. thesis, Université Paris XI, Orsay, 191 p.

Deu, M., Gonzalez de Leon, D., Glaszmann, J.C., Degremont, I., Chantereau, J., Lanaud, C., and Hamon, P. 1994. RFLP diversity in cultivated sorghum in relation to racial differentiation. *Theoretical and Applied Genetics*, 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.

Dice, L.R. 1945. Measures of the amount of ecologic association between species. *Ecology*, 26: 297-302.

Doggett, H. 1988. *Sorghum*, 2nd ed. London, Longman, 512 p.

Engqvist, G.M. and Becker, H.C., 1994. Genetic diversity for allozymes, RFLPs and RAPDs in resynthetized rape. In: *Biometrics in Plant Breeding: Applications of Molecular Markers: IXth Meeting of the EUCARPIA Section Biometrics in Plant Breeding.* J. Jansen, eds., Wageningen, Pays-Bas.

FAO, 1995. *Production Yearbook: 1994* (vol. 47). Rome, FAO, 243 p.

Frankel, O.H. and Brown, A.H.D. 1984. Current plant genetic resources: a critical appraisal. In: *Genetics, New Frontiers* (vol. IV). New Delhi, Oxford and IBH, pp. 1-11.

Hamon, S., Dussert, J., Deu, M., Hamon, P., Seguin, M., Glaszmann, J.C., Grivet, L., Chantereau, J., Chevallier, M.H., Flori, A., Lashermes, P., Legnate, H. and Noirot, M. 1998. Effects of quantitative and qualitative principal component score strategies on the structure of coffee, rice, rubber tree and sorghum core collections. *Genetics, Selection, Evolution, 30 (suppl. 1)*: 237-258.

Hamon, S., Hodgkin, T., Dussert, S., and Noirot, M. 1995. Core collection: accomplishments and challenges. *Plant Breeding Abstracts*, 65(8): 1125-1133.

Harlan J.R. 1972. Genetic resources in sorghum. In: *Sorghum in the Seventies.* N.G.P. Rao and L.R. House eds., New Delhi, Oxford and IBH.

Harlan, J.R. and Stemler, A. 1976. The races of sorghum in Africa. In: *Origins of African Plant Domestication.* J.R. Harlan et al., eds., The Hague, Mouton, pp. 465-478.

Harlan, J.R., and de Wet, J.M.J. 1972. A simplified classification of cultivated sorghum. *Crop Science,* 12: 172-176.

van Hintum, T., von Bothmer, R., Fischbeck, G., and Knüpffer, H. 1990. The establishment of the barley core collection. *Barley Newsletter,* 34: 41-42.

IBPGR. 1993. *Descriptors for Sorghum,* Sorghum bicolor (L.) Moench. Rome, IBPGR-ICRISAT, 38 p.

Lin, Y.R., Schertz, K.F., and Paterson, A.H. 1995. Comparative analysis of QTL affecting plant height and maturity across the Poaceae in reference to an interspecific sorghum population. *Genetics,* 141: 391-411.

Mann, J.A., Kimber, C.T., and Miller, F.R. 1983. *The Origin and Early Cultivation of Sorghums in Africa.* College Station, Texas A and M University Press, Texas Agricultural Experiment Station Bulletin no. 1454.

Mengesha, M.H., and Appa Rao, S. 1994. Management of plant genetic resources at ICRISAT. In: *Evaluating ICRISAT Research Impact (IAC).* Patancheru, India, ICRISAT, pp. 11-14.

Menkir, A., Goldsbrough, P., and Ejeta, G. 1997. RAPD-based assessment of genetic diversity in cultivated races of sorghum. *Crop Science,* 37(2): 564-569.

Messmer, M.M., Melchinger, A.E., Lee M., Woodman, W.L., Lee E,A., and Lamkey, K.R. 1991. Genetic diversity among progenitors and elite lines from the lowa Stiff Stalk Synthetic (ISSS) maize population: comparison of allozyme and RFLP data. *Theoretical and Applied Genetics,* 83: 97-107.

Morden, C.W., Doebley, J., and Schertz, K.F. 1989. Allozyme variation in Old World races of *Sorghum bicolor* (Poaceae). *American Journal of Botany,* 76(2): 247-255.

Noirot, M., Hamon, S., and Anthony, F. 1996. The principal component scoring: a new method of constituting a core collection using quantitative data. *Genetic Resources and Crop Evolution,* 43: 1-6.

Ollitrault, P. 1987. Evaluation génétique des sorghos cultivés, *Sorghum bicolor* (L.) Moench, par l'analyse conjointe des diversités enzymatique et morphophysiologique: relations avec les sorghos sauvages. Doct. thesis, Université Paris XI, Orsay, 187 p.

Ollitrault, P., Arnaud, M., and Chantereau, J. 1989b. Polymorphisme enzymatique des sorghos. 2.. Organisation génétique et évolutive des sorghos cultivés. *L'Agronomie Tropicale,* 44(3): 211-222.

Ollitrault, P., Escoute, J., and Noyer, J.L. 1989a. Polymorphisme enzymatique des sorghos. l. Description de 11 systèmes enzymatiques, déterminisme et liaisons génétiques. *L'Agronomie Tropicale,* 44(3): 203-210.

Pereira, M.G. Ahnert, D., Lee M. and Klier K. 1995. Genetic mapping of quantitative trait loci for panicle characteristics and seed weight in sorghum. *Revista Brasileira de Genética*, 18(2) : 249-257.

Pereira, M.G. and Lee M. 1995. Identification of genomic regions affecting plant height in sorghum and maize. *Theoretical and Applied Genetics*, 90: 380-388.

Perrier, X., Flori, A., and Bonnot, F. 1999. Les méthodes d'analyse des données. In: *Diversité Génétique des Plantes Tropicales Cultivées*. P. Hamon et al. eds., Montpellier, France, CIRAD, collection Repèeres, pp. 43-46.

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. *Molecular Breeding*, 2: 225-238.

Prasada Rao, K.E. and Ramanatha Rao, V. 1995. The use of characterisation data in developing a core collection of sorghum. In: *Core Collections of Plant Genetic Resources*. T. Hodgkin et al. eds., Chichester, Wiley, pp. 109-116.

Rami, J.F., Dufour, P., 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. *Theoretical and Applied Genetics*, 97(4) : 605-616.

Rohlf, F.J. 1990. Fitting curves to outlines. In: Proceedings of the Michigan Morphometrics Workshop. F.J. Rohlf and F.L. Bookstein eds., Ann Arbor, University of Michigan, Museum of Zoology Special Publication no. 2, pp. 167-177.

Saitou, N. and Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4: 406-425.

Snowden, J.D. 1936. *The Cultivated Races of Sorghum*. London, Adlard, 274 p.

Spagnoletti-Zeulli, P. and Qualset, C.O. 1993. Evaluation of five strategies for obtaining a core collection of durum wheat. *Theoretical and Applied Genetics*, 78: 295-304,

Tao, Y., Manners, J.M., Ludlow, M.M., and Henzell, R.G. 1993. DNA polymorphisms in grain sorghum, *Sorghum bicolor* (L.) Moench. *Theoretical and Applied Genetics*, 86(6): 679-688.

Teshome, A., Baum, B.R., Fahrig, L., Torrance, J.K., Arnason, T.J. and Lambert, J.D. 1997. Sorghum, *Sorghum bicolor* (L.) Moench, landrace variation and classification in North Shewa and South Welo, Ethiopia. *Euphytica*, 97: 255-263.

Thormann, C.E., Ferreira, M.E., Camargo, L.E.A., Tivang, J.G., and Osborn, T.C. 1994. Comparison of RFLP and RAPD markers to estimating genetic

relationships within and among cruciferous species. *Theoretical and Applied Genetics*, 88: 973-980.

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*, 359: 721-724.

Yagoua, N.D. 1995. Caractérisation du sorgho pluvial, *Sorghum bicolor* (L.) Moench, de la zone soudanienne du Tchad. Montpellier, France, CIRAD (internal document).