Theor Appl Genet (2008) 116:903–913
DOI 10.1007/s00122-008-0721-7

ORIGINAL PAPER

Niger-wide assessment of in situ sorghum genetic diversity with microsatellite markers

M. Deu · F. Sagnard · J. Chantereau · C. Calatayud · D. Hérault · C. Mariac · J.-L. Pham · Y. Vigouroux · I. Kapran · P. S. Traore · A. Mamadou · B. Gerard · J. Ndjeunga · G. Bezançon

Received: 4 September 2007 / Accepted: 23 January 2008 / Published online: 14 February 2008

Abstract Understanding the geographical, environmental and social patterns of genetic diversity on different spatial scales is key to the sustainable in situ management of genetic resources. However, few surveys have been conducted on crop genetic diversity using exhaustive in situ germplasm collections on a country scale and such data are missing for sorghum in sub-Saharan Africa, its centre of origin. We report here a genetic analysis of 484 sorghum varieties collected in 79 villages evenly distributed across Niger, using 28 microsatellite markers. We found a high level of SSR diversity in Niger. Diversity varied between eastern and western Niger, and allelic richness was lower in the eastern part of the country. Genetic differentiation between botanical races was the first structuring factor (Fst = 0.19), but the geographical distribution and the ethnic group to which farmers belonged were also significantly associated with genetic diversity partitioning. Gene pools are poorly differentiated among climatic zones. The geographical situation of Niger, where typical western African (guinea), central African (caudatum) and eastern Sahelian African (durra) sorghum races converge, explained the high observed genetic diversity and was responsible for the interactions among the ethnic, geographical and botanical structure revealed in our study. After correcting for the structure of botanical races, spatial correlation of genetic diversity was still detected within 100 km, which may hint at limited seed exchanges between farmers. Sorghum domestication history, in relation to the spatial organisation of human societies, is therefore key information for sorghum in situ conservation programs in sub-Saharan Africa.

Communicated by F. Ordon.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-008-0721-7) contains supplementary material, which is available to authorized users.

M. Deu (✉) · F. Sagnard · C. Calatayud · D. Hérault
CIRAD, UMR DAP, Avenue Agropolis, TA A 96/03,
34398 Montpellier, France
e-mail: monique.deu@cirad.fr

F. Sagnard · P. S. Traore
International Crop Research Institute for the Semi-Arid Tropics (ICRISAT), Station de Samanko, BP320, Bamako, Mali

J. Chantereau
CIRAD, UPR Agrobiodiversité en savane,
34398 Montpellier, France

C. Mariac · J.-L.Pham
Institut de Recherche pour le Développement (IRD),
Montpellier, France

Y. Vigouroux · G. Bezançon
Institut de Recherche pour le Développement (IRD),
Niamey, Niger

I. Kapran · A. Mamadou
Institut National de la Recherche Agronomique du Niger (INRAN), Niamey, Niger

B. Gerard · J. Ndjeunga
International Crop Research Institute for the Semi-Arid Tropics (ICRISAT), Niamey, Niger

Introduction

Characterizing the patterns of in situ crop diversity and understanding the underlying evolutionary processes that have shaped the observed genetic structures are two prerequisites for both breeding and plant genetic resources programs. This is especially important for traditional cereals cultivated in their centre of diversity, where unique

genes or gene complexes can be found and often represent valuable genetic resources for breeders (Brush 1995; Maxted et al. 2002). In decentralised breeding programs, determining the genetic structure of local landraces, which are often preferred by farmers for their adaptation, taste or post-harvest processing traits, should help choose the best entries and delineate target release zones for improved varieties (Ceccarelli et al. 1997). The geographical distribution of landrace vernacular names, agromorphological types and genetic diversity on different spatial scales also provides valuable information to complement ex situ collections, and establish relevant criteria to initiate and monitor in situ conservation programs (Brush 2000).

Sorghum (*Sorghum bicolor* L. Moench) was domesticated in northeastern Africa. It is an annual, predominantly selfing cereal (Ollitrault et al. 1997; Djè et al. 2004). Cultivated sorghums (*Sorghum bicolor* ssp. bicolor) have been classified into five basic botanical races (bicolor, caudatum, durra, guinea and kafir) and ten intermediate ones, based on panicle and spikelet morphology (Harlan and de Wet 1972).

Most sorghum genetic surveys relied on gene bank accessions to assess large-scale geographical or taxonomic structures, either with allozyme or DNA markers (Aldrich et al. 1992; Deu et al. 1994, 1995, 2006; Cui et al. 1995; de Oliveira et al. 1996; Menkir et al. 1997; Djè et al. 2000; Grenier et al. 2000; Casa et al. 2005). Fewer studies have involved in situ collection of sorghum landraces and detailed information on locations, growing environments and farmers' practices. Surveys based on in situ collections generally reveal a lack of correlation between genetic diversity parameters and environmental factors (Ayana et al. 2000, 2001; Ghebru et al. 2002; Zongo et al. 2005) and weak genetic differentiation between regions (Ollitrault et al. 1997; Djè et al. 1999; Nkongolo and Nsapato 2003; Kayodé et al. 2006). As they are based on a relatively small number of villages that are not exhaustively sampled, limited inferences can be drawn as to the evolutionary factors responsible for the observed genetic structures.

In Niger, sorghum is the second most cultivated crop after pearl millet. It is grown under rainfed conditions in traditional farming systems. Water availability is the main limiting factor for agriculture. Rainfall varies from year to year and over short distances. A desiccation process has been observed especially in the dryer regions of the country, and Hulme (2001) evaluated that annual rainfall across the Sahel decreased by 20–30% between the 1930–1950 and the 1970–1990 periods. Over the past three decades, annual sorghum production has steadily increased, in response to the doubling of the population. This involved an extension of sorghum cultivation to marginal lands and an impoverishment of the soils due to shortened fallows (Wezel and Boecker 1998). Three major ethnic groups practice agriculture and small-scale plant breeding. The Zarma/Songhaï (22% of the population of Niger), the Hausa (56%) and the Kanuri (4.3%) are predominant in western, central and eastern regions, respectively (Fig. 1).

Adoption of improved varieties is limited in Niger, and the largest part of sorghum production comes from farmer-selected landraces. These landraces are locally adapted to their harsh, heterogeneous and unpredictable environments. Ethnic traditions, social organizations and food preferences also probably contribute to the extent and structure of crop diversity (Reenberg 2001). In situ conservation programs therefore require the identification and understanding of the drivers of genetic diversity dynamics.

The objectives of this study were to (1) characterize genetic diversity in a sorghum collection from 79 villages covering the complete geographical and agro-ecological range of sorghum growing areas in Niger, (2) assess the patterns of genetic diversity revealed by SSR markers in relation to botanical, climatic and ethnic factors, (3) measure the spatial structure of genetic diversity on different scales and (4) propose evolutionary scenarios of sorghum diversity in Niger as a basis for future in situ conservation programs.

Material and methods

Collection of sorghum varieties

From October 2003 to December 2003, an intensive collection of sorghum varieties was conducted in 79 villages covering the rainfall gradient and range of agro-ecological conditions of Niger's agricultural areas (Fig. 1). A multidisciplinary team composed of social scientists, agronomists and geneticists, assisted by Zarma and Hausa translators, collected seeds and interviewed farmers in two villages per day on average. Varietal inventory also included information on crop uses, seed origins, agricultural systems and social organizations in the villages. In each village, we sampled all local varieties listed by a representative group of farmers. Each variety was provided by one farmer in one village, either by grains (unselected panicles) or by seeds. The varieties were collected either in farmers' fields, in panicle-drying zones or in home granaries, depending on the advancement of harvest work. For each variety, a bulk of seeds was collected from 30 panicles. We assume that we sampled the majority of sorghum varieties grown in these 79 villages in 2003.

Racial characterisation of sorghum varieties

The whole collection was grown in 2004 at INRAN (Institut National de la Recherche Agronomique du Niger) experimental stations in Maradi and Bengou.

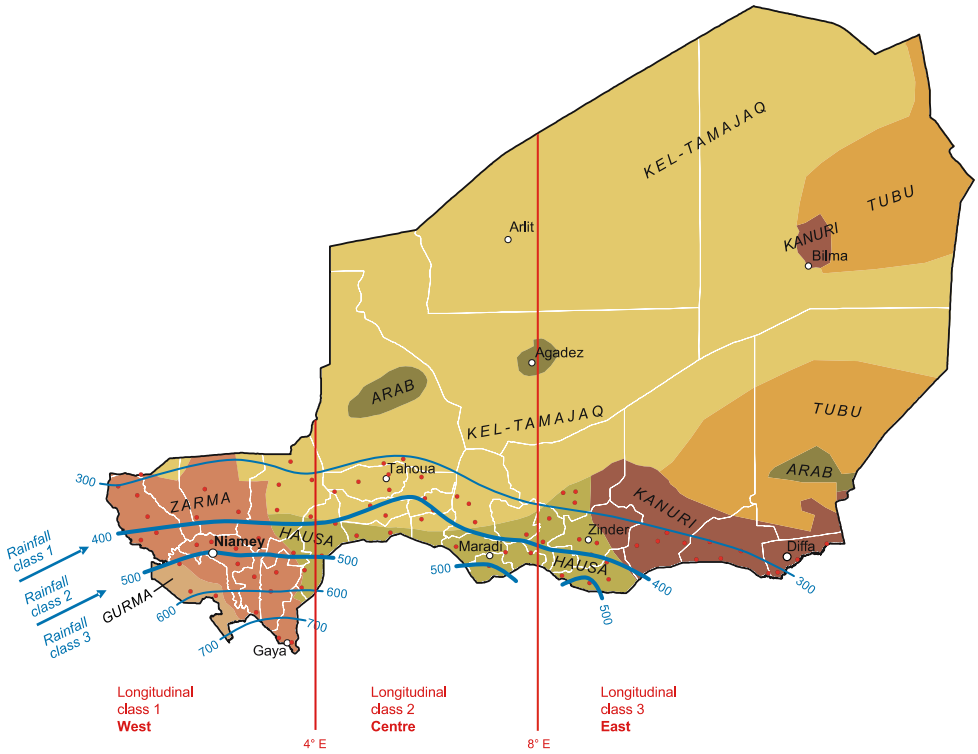

Fig. 1 Location of the 79 villages visited for the collection of 484 sorghum varieties in Niger. Isohyets were computed using ArcGIS geostatistical analyst (v. 9.0) based on 1971–2000 annual rainfall normals (linear kriging with smoothing). The backdrop ethnic map is a simplified distribution of dominant ethnic groups adapted from linguistic maps of Africa (http://www.muturzikin.com, 2007)

Racial characterisation based on panicle and spikelet morphology was carried out in accordance with Harlan and de Wet's (1972) classification. In addition, Snowden's taxon *margaritiferum* was distinguished, since those small-grain guinea sorghums appear to be genetically different from all other representatives of the guinea race (Folkertsma et al. 2005; Deu et al. 2006).

DNA extraction and SSR genotyping

Sorghum seeds were germinated in a greenhouse. For each variety, DNA was isolated from fresh leaves collected on one 2- to 3-week-old seedling following a CTAB protocol described by Deu et al. (1995). Hereafter, the genetic diversity of a variety refers to the multilocus genotype of one single plant representative of a maternal parent (i.e. landrace or improved variety). The genetic information provided by one single individual per variety, with no information on within-variety diversity, has proved to be sufficient to detect large-scale inter-varietal evolutionary trends even in outcrossing crops when the number of loci is sufficient (Matsuoka et al. 2002; Mariac et al. 2006).

Twenty-eight SSR markers were assayed (listed as electronic supplementary information S1). They formed a subset of 50 microsatellites selected for their reliability and scoring accuracy between laboratories, and their level of polymorphism and genome coverage for the Generation Challenge Program (http://sat.cirad.fr/sat/sorghum_SSR_kit). Most of these markers have been previously described and mapped: Sb (Brown et al. 1996; Taramino et al. 1997), Xcup (Schloss et al. 2002), Xtxp (Bhattramakki et al. 2000; Menz et al. 2002). Gpsb markers were developed in CIRAD and mapped in our RILs population (to be published elsewhere). PCR conditions and genotyping on Li-Cor automated sequencers were as described by Barnaud et al. (2007). Saga GT v. 2.2 (Li-Cor) was used to determine allele sizes. Genotyping was conducted at the Languedoc Roussillon Genotyping Platform hosted by CIRAD.

Investigated factors of genetic structure

The partitioning of microsatellite diversity was explored on different scales according to racial (botanical), eco-geographical, climatic, ethnic and seed origin criteria using information provided by the collection questionnaires or available in meteorological databases. Three longitudinal classes (4° interval) were created. They were roughly aligned on the administrative departments of Tillaberi and Dosso (Western), Tahoua and Maradi (Central) and Zinder and Diffa (Eastern). Three main annual rainfall classes were defined: "less than 400 mm", "from 400 to 500 mm" and "more than 500 mm". Varieties were also clustered according to the dominant ethnic groups in the villages (Hausa, Kanuri and Zarma/Songhaï) and their date of introduction (recent or ancient).

Genetic data analyses

Genetic diversity parameters were estimated for each defined group of varieties with GENETIX software 4.04 (Belkhir et al. 2002): total number of alleles (A^t), number of rare alleles (A^r, freq < 5%), observed heterozygosity (H_o) and expected heterozygosity or gene diversity (H_e) adjusted for low sampling size, according to Nei (1978), following formula

$$H = 2n\left[\left(1 - \sum_j \sum_i x_{ij}^2\right) \Big/ r\right] \Big/ (2n - 1)$$

where x_{ij} is the frequency of the ith allele of locus j, r is the number of genetic loci and n is the population size. We also calculated R_s (allelic richness) using the rarefaction method (Petit et al. 1998) implemented in FSTAT (Goudet 2001). This method permits to estimate the expected number of different alleles among equal-sized samples, based on the lowest sample size in groups of varieties that are defined for comparisons. Consequently, allelic richness was calculated for 16 varieties (32 genes) for botanical comparisons corresponding to the smallest size of the guinea margaritiferum sample but for 86 varieties (172 genes) for rainfall classes. The expectation of the number of alleles in a sample size of g ($g \leq N$) is

$$\hat{r}_{(g)} = \sum_i \left[1 - \left(\frac{C_{N-N_i}^g}{C_N^g}\right)\right]$$

where $a_1,\ldots,a_k$ are the different alleles at a single locus and $N_1,\ldots,N_k$ are the number of times they appeared in the N-size sample of each group.

The significance of differences in R_s and H_e between the defined groups was tested using a Wilcoxon signed-rank test across loci.

To investigate the genetic relationships between varieties, a genetic dissimilarity matrix was computed using the simple matching index. A dendrogram was then generated on the dissimilarity matrix with the Neighbour Joining (NJ) algorithm implemented in DARWIN V5 software (Perrier et al. 2003). Only one variety with too many missing data was not included in the cluster analysis.

We explored the population structure using Weir and Cockerham's (1984) θ as an estimate of Fst. Overall Fst and pairwise Fst were calculated and tested for their significance with GENETIX software 4.04. Fst was only interpreted as a descriptive differentiation parameter in our study, since the mixing of different varieties in the same group prevented us from drawing any evolutionary inferences from the estimated F-statistics.

However, analysis of the correlation between the genetic relatedness among varieties and the geographical distances provided indirect insight into several evolutionary factors, including limited gene flow responsible for isolation-by-distance processes (Sokal and Wartenberg 1983; Smouse and Peakall 1999; Rousset 2000). Under restricted gene flow through either pollen or seed movements, individuals or varieties growing close to each other are more genetically similar than those growing far apart. The mean pairwise genetic kinship among varieties within short distances is then higher than the average kinship between all sampled individuals. For spatial genetic structure analyses, distance classes were determined to ensure that a sufficient number of data were included in the computations for each distance interval. We used the individual kinship estimator proposed by Ritland (1996) averaged across loci as a measurement of genetic similarity between varieties. The statistical significance of kinship estimators was tested using 1,000 Monte Carlo simulations on the hypothesis of complete spatial randomness of genetic diversity. We first conducted the spatial analyses on the country-scale collection, and secondly, within three genetic sub-groups identified from the NJ dendrogram to assess fine scale patterns after correcting for the larger racial structure. All the spatial computations were performed using SPAGeDi (ver 1.2) software (Hardy et al. 2002).

Results

Racial sorghum diversity in Niger

The collection studied here comprises 484 sorghum varieties obtained in 79 villages (Fig. 1). The number of varieties collected per village ranges from zero to 16 (average: 6.13). Varietal richness varies across regions with 6.23, 8.25 and 3.88 varieties per village, in western, central and eastern Niger, respectively. No difference is observed between the two most important ethnic groups with 6.39 and 6.05 varieties per village in 48 Hausa and 18 Zarma/Songhaï villages, respectively.

All sorghum basic botanical races except kafir can be found in Niger (listed as electronic supplementary information S2). In this collection, durra (23.1%) and caudatum (21.3%) races are the most prevalent, followed by guinea (10.5%) and bicolor (8.1%). Among guinea, 43.1% could be identified as guinea margaritiferum varieties. Intermediate races (25.8%) between durra and bicolor (49 varieties), durra and caudatum (35 varieties), caudatum and bicolor (19 varieties) and caudatum and guinea (17 varieties) amount to a significant share of the collection. About 11% of the varieties could not be classified either because entire panicles were missing or because they were hybrids with a complex racial pedigree.

Racial distribution is not random between regions. Guinea varieties are mainly cultivated in western Niger, and principally in the southernmost areas where annual rainfall is more abundant. Caudatum and bicolor varieties are rare in eastern Niger. Durra varieties cover all the sorghum-growing areas, but are predominant in the eastern region. A high racial diversity is observed in central Niger, though caudatum and intermediate races prevail.

Overall genetic diversity

Twelve varieties were discarded from the SSR analysis due to their weak germination or insufficient yield during DNA extractions. Overall, we present the analysis of 472 varieties from 76 villages. The average missing data per accession was 2.6%.

The 28 microsatellite loci are polymorphic and reveal 292 alleles (listed as electronic supplementary information S1), from 2 to 26 alleles per locus with an average of 10.43. Sixty-four percent of alleles are rare at the 0.05 threshold. Observed heterozygosity ranges from 1.7% (gpsb067) to 7.8% (Xtxp15) with an average of 4.2%. The 28 SSR loci can discriminate 454 varieties out of 472 (96%).

Genetic diversity estimates

As expected, the allelic richness is highly sensitive to the rare alleles frequently revealed by SSR markers. This parameter increases with the number of sampled genes, which varies among groups that are compared. R_s is therefore only meaningful to assess the genetic diversity differences that occur within one categorical factor (i.e. between races for the racial structuration or between rainfall classes for the climatic groups).

Allelic richness is significantly higher in the caudatum ($R_s = 4.88$ alleles per locus) race than in durra (3.94; $P < 0.01$), guinea (4.11; $P < 0.05$) and guinea margaritiferum (3.21; $P < 0.01$) (Table 1). Within the guinea race, the margaritiferum group has significantly lower allelic richness ($P < 0.05$). The gene diversity estimates exhibit the same trends except that the durra race ($H_e = 0.393$) appears to be less genetically diverse than bicolor (0.571; $P < 0.001$) and guinea (0.496; $P < 0.05$). Overall Fst is high among the basic races (Fst = 0.278, $P < 0.01$) but remains substantial when intermediate races are also taken into account (Fst = 0.19, $P < 0.01$). Pairwise Fst are all significant at the 1‰ threshold and illustrate particularly high genetic differentiation between the margaritiferum

Table 1 Genetic diversity estimates between structuring factors

	N	A^t	A^p	R_s	H_e	H_o	Fst
Races	472(299[a])			5.50(32)			0.190(0.278[a])
Bicolor	36	144	9	4.43	0.571	0.038	
Caudatum	101	192	16	4.88	0.569	0.039	
Durra	111	174	17	3.94	0.393	0.029	
Guinea	29	130	8	4.11	0.496	0.071	
Guinea m.	22	92	7	3.21	0.427	0.041	
Intermediate	120	195	11	4.73	0.545	0.050	
Unclassified	53	160	4	4.65	0.582	0.044	
Regions	472			8.46(172)			0.070
West	186	239	38	7.92	0.646	0.060	
Centre	193	233	27	7.80	0.603	0.030	
East	93	172	12	6.12	0.422	0.033	
Rainfall classes	472			8.14(142)			0.031
<400 mm	246	253	32	7.70	0.580	0.040	
400–500 mm	145	235	17	7.87	0.624	0.035	
>500 mm	81	195	16	6.93	0.622	0.060	
Ethnic groups	459			6.21(46)			0.054
Hausa	307	257	51	5.96	0.590	0.032	
Zarma/Songhaï	124	226	27	6.22	0.650	0.068	
Kanuri	28	112	5	3.93	0.393	0.044	
Origins	428			8.94(248)			0.008
Recent introduction	136	221	18	7.88	0.586	0.034	
Ancient	292	269	65	8.75	0.618	0.046	
Total collection	472	292		10.43	0.613	0.042	

N number of landraces included in each group, A^t total number of alleles, A^p number of private alleles (present in a single group), R_s Allelic richness for each group based on a minimum sample of gene copies (the minimum sample is indicated in brackets), H_e gene diversity, H_o observed heterozygosity, *Fst* average genetic differentiation for each group

[a] For "racial" Fst, a second estimation of Fst has been calculated for basic races only on a population of 299 landraces after exclusion of intermediate and unclassified landraces

group and all other races (0.309 ≤ Fst ≤ 0.453), durra and guinea (Fst = 0.341) and durra and caudatum (Fst = 0.300) races.

Allelic richness appears significantly lower in eastern Niger ($R_s = 6.12$) than in central (7.80; $P < 0.001$) and western (7.92; $P < 0.001$) Niger. Between regions, the overall Fst of 0.07 is moderate while still largely significant ($P < 0.01$). Genetic differentiation is higher between western and eastern regions (Fst = 0.124, $P < 0.001$), which partially correspond to the geographical locations of the Zarma/Songhaï and Kanuri ethnic groups. The allelic richness of sorghums grown by Kanuri ($R_s = 3.93$) people is significantly lower than those grown by Zarma/Songhaï ($R_s = 6.22$; $P < 0.001$) and Hausa ($R_s = 5.96$; $P < 0.001$) villages. Genetic differentiation is weak between ancient and recently introduced varieties (Fst = 0.008) and between rainfall zones (overall Fst = 0.031).

Phenetic analysis

The NJ dendrogram (Fig. 2) reveals first-order differentiation between eastern and western Niger. The landraces collected in central Niger are intermixed in the two clusters. On a finer discrimination level, a racial pattern is found with seven clusters identified on the dendrogram. Cluster I is primarily composed of durra and intermediate durra (mainly DB and DC). Cluster II contains most of the caudatum varieties, while guinea margaritiferum and other guinea sorghums are grouped in clusters III and IV, respectively. Cluster V includes varieties identified as racial intermediate forms, and the two small remaining groups (clusters VI and VII) are composed of bicolor varieties.

Some varieties collected in different villages, but sharing name or exhibiting specific morphological traits, belong to the same subgroups. Among them, most of the caudatum varieties, called "Jan Jare" by Hausa farmers and characterised by their red grains and fusoid panicles, are clustered in subgroup II-a. Sub-cluster II-b is composed of sorghums identified as improved varieties. Sub-cluster II-c is made up of representatives from the variety called "Korbyu", a twin-seeded sorghum. Within durra and intermediate races too, some varieties are genetically discriminated: the "Matche Da Koumnya" sub-group contains late durra sorghums with a crossed peduncle, the "Makaho Da Wayo" sub-group is composed of varieties belonging to the Snowden's (1936) taxon membranaceum, intermediate between the durra and bicolor races.

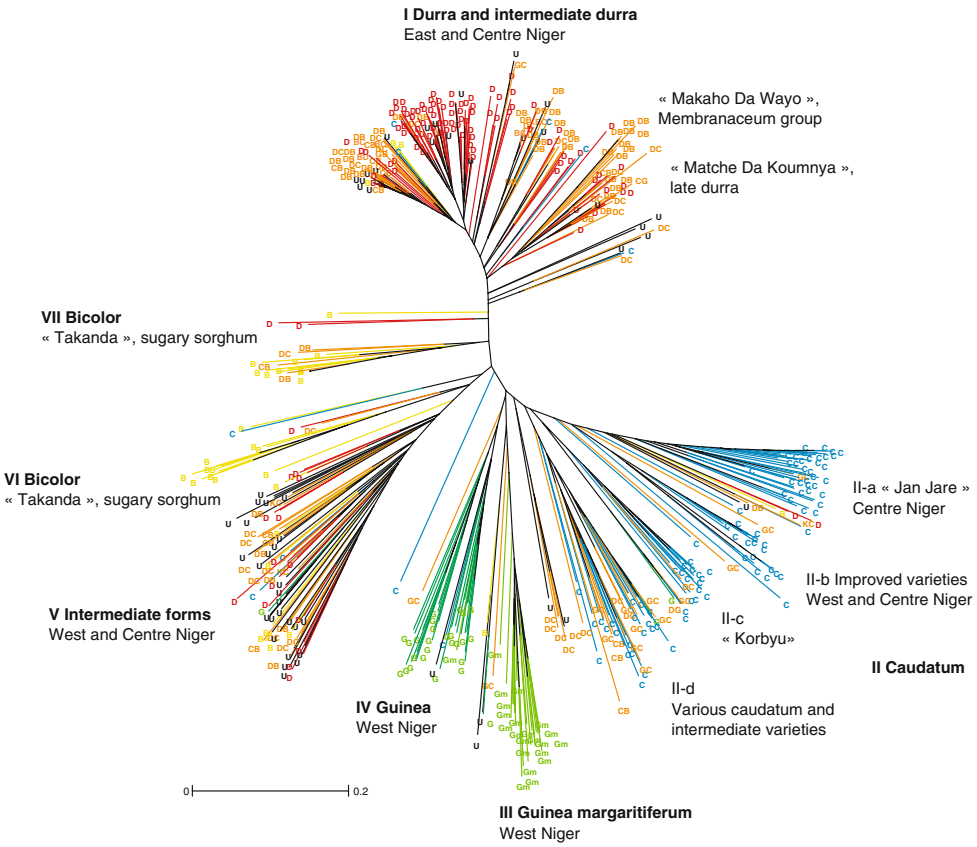

Fig. 2 Unrooted neighbour-joining tree based on allelic data from 28 SSR loci among 472 sorghum varieties using the simple matching index. Sorghum varieties are identified by their race, and each race has been represented by a specific color (B: bicolor in *yellow*, C: caudatum in *blue*, D: durra in *red*, G: guinea in *dark-green*, XY: intermediate races between race X and Y, all different intermediate races represented in *orange*, U: unclassified in *black*, and Gm: Guinea margaritiferum in *light green*)

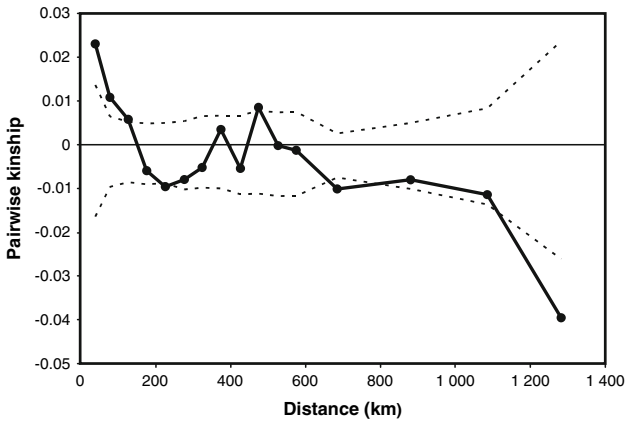

Fig. 3 Spatial regression of mean pairwise Ritland kinship on geographical distances between 472 sorghum varieties in Niger. *Closed circles* indicate mean kinship values for successive distance classes. The *dashed lines* represented upper and lower 95% confidences envelopes around the null hypothesis of complete spatial randomness

Spatial analysis

The analysis performed on all sorghum varieties collected in Niger reveals a strong spatial genetic structure within geographical distances under 200 km (Fig. 3). The observed spatial pattern, with significant positive relative kinships at short distances and significant negative relative kinships at longer distances, is typical of genetic diversity gradients rather than a patchy genetic structure.

The spatial analyses performed within genetically more homogeneous clusters (cluster I "durra", cluster II-a "caudatum-Jan Jare" and cluster III "guinea"), to remove the effect of the racial structure, still show a significant positive spatial autocorrelation of multilocus diversity at distances under 100 km (Fig. 4).

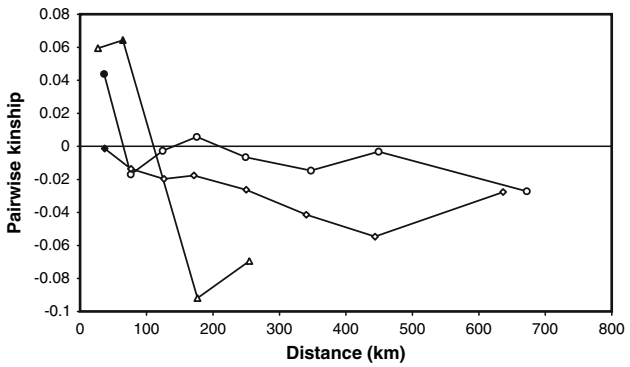

Fig. 4 Spatial regression of mean pairwise kinship on geographical distances performed within three genetically distinct groups of sorghum varieties. *Circles*, *diamond-shaped symbols* and *triangles* represent mean kinship values for durra, "Jan Jare" and guinea varieties, respectively. *Filled symbols* represent significant values

Discussion

Sorghum diversity in Niger

Results illustrate an outstanding racial diversity in Niger, in contrast with the racial distribution revealed in other Sahelian countries with partially similar agro-climatic conditions. In Burkina Faso, Zongo (1991) identified 93% of the 885 varieties collected across the country as guinea race sorghums. In Mali, Scheuring et al. (1980) also found a racial distribution skewed toward guinea (74% of 775 collected varieties) with only 19.5% of durra and 2% of caudatum sorghums. The more even racial distribution of sorghums in Niger and the significant presence of intermediate races are the source of the large genetic diversity measured by microsatellite markers. This can be explained by the geographical situation of Niger, at the crossroads between Arab-Berber and Sudan-Sahelian influences, with people of diverse origins and religions living together for thousands of years. Botanical races of sorghums and cultural or ethnolinguistic human groups have long been associated by several authors (reviewed in Doggett 1988, pp 43–56). Guinea varieties may have originally been cultivated in the humid zones of western Africa by animists. Durra varieties may have primarily been grown in arid Sahelian zones by African Muslims and Arabic people. The race caudatum may be historically associated with Chari-Nile speakers who are the dominant ethnolinguistic group in Chad and Sudan and moved towards western Africa more recently. In Niger, geographical and ethnic patterns of racial distribution also exist. The guinea sorghums are mainly cultivated by the Zarma/Songhaï group in western Niger, where rainfall is more abundant. Kanuri people who live in the eastern part of the country around Lake Chad cultivate almost exclusively durra sorghums. Botanical races can also be associated with food properties or specific uses as in the case of the "Takanda" group in Niger. These sweet sorghums that belong to the bicolor race were collected in 35% of the villages we visited. Their sweet stems are harvested before maturity and chewed by people.

In addition to food preferences and diversified uses, traditional crop diversity is generally maintained by farmers as a risk management strategy to increase the resilience in the face of strong environmental heterogeneity in both space and time, and numerous pest and diseases (Teshome et al. 1999). The high crop variety richness in a village might therefore be a characteristic of traditional agricultural societies as illustrated for instance by Teshome et al. (1999) with sorghum in Ethiopia, and Bellon (1997) with maize in Mexico. However, the average estimate of 6.13 sorghum varieties grown per village in Niger is much smaller than the diversity observed in other sorghum

farmer communities in sub-Saharan Africa: 13 indigenous varieties in a single community in Zimbabwe (Oosterhout 1997), 46 sorghum landraces identified by farmers in a village of North Cameroon (Barnaud et al. 2007), 76 varieties grown by farmers from three communities in northern Benin (Kayodé et al. 2006). This might be linked to farmers' preference for pearl millet in Niger, which is the main staple crop in the country (Mariac et al. 2006). It might also reflect scaling biases in our spatial collection protocol, where less time was devoted to fine-scale, village-level surveys, and more focus was put on a possibly more representative sampling on a coarser, country-level scale. The fact that the farmers collectively maintain and rely on a large number of landraces to meet their production needs indicates that they must play a key role in future food security and conservation programs.

Structure of genetic diversity

This study investigates the structure of sorghum genetic diversity using exhaustive in situ collections on a national scale. Gene diversity averaged over 28 microsatellite loci in Niger ($H_e = 0.61$) is similar to the published results with SSR in South Africa: $H_e = 0.60$ (Uptmoor et al. 2003), but slightly lower than the gene diversity estimated in Eritrea ($H_e = 0.78$) by Ghebru et al. (2002), and in Morocco ($H_e = 0.84$) by Djè et al. (1999). Comparisons of genetic diversity between these in situ studies are difficult, since estimates such as allelic richness or expected heterozygosity depend on the sampling schemes (single plant or DNA bulk), the number of surveyed SSR, the size of the SSR repeats and the location of the SSR on the genome (between coding or non-coding DNA regions). In this study, we use fewer dinucleotid microsatellites (less than 50% of the 28 SSR), which are generally more polymorphic than those containing longer repeats (Casa et al. 2005). Furthermore, eight of our SSR (Xcup), which are assumed to be located within or near genes, indeed reveal fewer alleles (4.63 alleles per locus). However, the presence in this study of 21 SSR, earlier genotyped in the CIRAD world core collection (unpublished data), highlights the high degree of microsatellite diversity in Niger. Allelic richness (10.0 alleles per locus) in Niger amounts to 86% of that detected in the core collection (11.7). That proportion is still 79% after correction for uneven sample sizes. Gene diversity in Niger computed on the same set of SSR markers is not statistically different compared to the core collection.

The botanical race proves to be the main structuring factor of sorghum genetic diversity in Niger, which confirms on a country scale the results obtained on a world scale using sorghum accessions from gene banks (Deu et al. 2006). The guinea margaritiferum group is genetically differentiated from all other races, including non-margaritiferum guinea sorghums. The same conclusion was reached by Deu et al. (1994, 1995, 2006) with RFLP markers and by Folkertsma et al. (2005) with SSR markers. The genetic history of guinea margaritiferum in western Africa is one of the last phylogenetic enigmas that still require further investigation within the species *Sorghum bicolor*.

The genetic differentiation appears moderate between geographical regions and ethnic groups, and much lower between climatic zones or varieties age. The association of individual ethnic groups with specific geographical regions causes strong geographical × ethno-cultural interactions in the structure of crop genetic diversity. However, the scientific statement that these two factors predominate over climatic conditions to explain the distribution of sorghum diversity in Niger deserves more investigation. It suggests that historical patterns of human migration and sorghum domestication exert a dominant influence on patterns of genetic diversity still observable today, more than recent environmental constraints. The NJ dendrogram confirmed that central Niger, at the crossroads between western and eastern influences, exhibited the highest genetic diversity.

Spatial patterns of genetic diversity and seed exchange systems

Because of our sampling design, in which villages at least 30 km apart were visited, and considering the preferentially selfed mating system of sorghum, estimated patterns of gene flow on a country scale are likely to have a stronger seed than pollen component. The overall spatial genetic structure is mainly due to the non-random spatial distribution of genetically distinct races. However, the residual spatial autocorrelation of genetic diversity within a race or within a more genetically distinct group could indicate gene flow limitations over short distances (under 100 km). Seed dispersal of a cultivated species occurred through both the traditional and commercial seed systems. In sub-Saharan Africa in general, traditional crops like sorghum are mainly exchanged between farmers, following collective social rules that involve relationships between relatives, neighbours and other social structures (Tescar 2004). Questionnaire results (not shown) support preferential seed exchanges between farmers, and reveal that among recent documented varieties, 65% originate in the traditional seed systems, against 35% from markets, extension services or seed emergency projects. The weak genetic differentiation found between ancient and recent varieties in Niger also suggests that the varieties introduced in a village are not of exotic origin. Indirect genetic marker

and agromorphological trait evidence about the social determinant of restricted seed-mediated gene flow has been provided by studies on pearl millet in India (vom Brocke et al. 2003), on barley in Syria (Parzies et al. 2004) and on maize in Mexico (Perales et al. 2005). Additional surveys following more refined sampling schemes based on anthropological data and within-variety genetic diversity should help precisely quantify the amount and extent of seed exchanges between farmers.

Implications for sorghum in situ conservation programs

Conservation genetics aim to identify and understand the evolutionary forces that have shaped the observed distribution of genetic diversity within a species on different scales, and identify populations or landraces that deserve priority conservation. According to the Convention for Biodiversity, in situ conservation policies are to be managed on a country scale. Results by Mariac et al. (2006) on pearl millet and by this study on sorghum diversity in Niger clearly indicate that genetic surveys within a country are indispensable to link genetic data with evolutionary processes. Much attention has been paid to the conservation of genetic resources in the centres of domestication (Brush 1995) that exhibit a high degree of diversity. Our results show that varieties grown in countries like Niger, at the crossroads between the primary sorghum domestication centre (Ethiopia-Sudan) and a putative secondary domestication centre in subhumid western Africa (Doggett 1988), cumulated and recombined a high number of alleles, as evidenced by the presence of original intermediate sorghum races in our collection.

Hypervariable markers are particularly useful for genetically identifying diverse regions or populations. Moreover, by simulations under certain conditions, diversity at neutral markers has shown to be correlated with allelic richness at loci undergoing selection (Bataillon et al. 1996). The large number of alleles revealed by microsatellite markers means that allelic richness can be used as a precise indicator of genetic diversity for conservation purposes (Petit et al. 1998). In our study, western and central Niger appear more genetically diverse than eastern Niger, but specific alleles are found in all regions, even in Kanuri villages that exhibit the lowest sorghum diversity. This implies that, to capture the whole range of existing diversity, all regions need to be sampled, with particular emphasis on western and central Niger.

A very small set of improved varieties collected in Niger (4.8% of collection) clustered into the genetically distinct sub-group II-b, among the caudatum races. This confirms the low adoption of improved materials in the farming systems of Niger, despite numerous NGO- and/or government-led seed emergency programs. The improved varieties targeting large-scale production areas fail to respond to the adaptive constraints in locally heterogeneous traditional agricultural systems or cannot satisfactorily fit the diversity of farmers' uses and preferences. These modern varieties found in our collection are not genetically diversified and have a pedigree that is probably poorly enriched by local germplasm. They only encompass a very small portion of the total sorghum genetic diversity in the country, and the widescale adoption of that homogenous material would certainly have triggered significant genetic erosion.

With rainfall being the most limiting factor for agriculture in the Sahel, most crop collections use agro-climatic gradients to stratify the sampling design despite little evidence of a correlation between genetic diversity and the environment (Ayana et al. 2001; Ghebru et al. 2002; Zongo et al. 2005). This study highlights that human factors, through the interactions between societies, geographical regions and sorghum botanical races, and the patterns of seed exchanges between farmers, explain a larger share of the observed genetic structure than do the climatic conditions. We thus suggest, like Brush (2000), that ethnic and social factors are key information to implement in situ sorghum conservation programs and conduct further ex situ collections of sorghum in sub-Saharan Africa.

Acknowledgments We thank the different people who took part in the 2003 sampling operation, and particularly Djibo Moussa, Moussa Tidjani and H. Yahaya Bissala. We are also grateful to Claire Billot and the Montpellier Languedoc-Roussillon Genopole platform for technical assistance. This work was supported by Institut Français de la Biodiversité (IFB). We thank Seydou B. Traoré (AGRHYMET) for providing access to 1971–2000 yearly rainfall normals.

References

Aldrich PR, Doebley J, Schertz KF, Stec A (1992) Patterns of allozyme variation in cultivated and wild *Sorghum bicolor*. Theor Appl Genet 85:451–460

Ayana A, Bryngelsson T, Bekele E (2000) Genetic variation of Ethiopian and Eritrean sorghum (*Sorghum bicolor* [L.] Moench) germplasm assessed by random amplified polymorphic DNA (RAPD). Genet Resour Crop Evol 47:471–482

Ayana A, Bryngelsson T, Bekele E (2001) Geographic and altitudinal allozyme variation in sorghum (*Sorghum bicolor* [L.] Moench) landraces from Ethiopia and Eritrea. Hereditas 135:1–12

Barnaud A, Deu M, Garine E, McKey D, Joly H (2007) Local genetic diversity of sorghum in a village in northern Cameroon: structure and dynamics of landraces. Theor Appl Genet 114:237–248

Bataillon TM, David JL, Schoen DJ (1996) Neutral genetic markers and conservation genetics: simulated germplasm collection. Genetics 144:409–417

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2002) Genetix 4.04, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier, France. http://www.univ-montp2.fr/~genetix/genetix/genetix.htm

Bellon MR (1997) On farm conservation as a process: an analysis of its components. In: Sperling L, Loevinsohn M (eds) Using diversity: enhancing and maintaining genetic resources on-farm. IDRC, Ottawa, Canada. http://www.idrc.ca/en/ev-85109-201-1-DO_TOPIC.html

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

Brown SM, Hopkins MS, Mitchell SE, Senior ML, Wang TY, Duncan RR, Gonzalez Candelas F, Kresovich S (1996) Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum (*Sorghum bicolor* [L.] Moench). Theor Appl Genet 93:190–198

Brush SB (1995) In situ conservation of landraces in centers of crop diversity: implications of germplasm conservation and utilization. Crop Sci 35:46–354

Brush SB (2000) The issue of in situ conservation of genetic resources. In: Brush SB (eds) Genes in the field: on-farm conservation of crop diversity. IPGRI/ IDRC/Lewis Publishers, Boca Raton, pp 3–26

Casa AM, Mitchell SE, Hamblin MT, Sun H, Bowers JE, Paterson AH, Aquadro CF Kresovich S (2005) Diversity and selection in sorghum: simultaneous analyses using simple sequence repeats. Theor Appl Genet 111:23–30

Ceccarelli S, Bailey E, Grando S, Tutwiler R (1997) Decentralized, participatory plant breeding: a link between formal plant breeding and small farmers. New frontiers in participatory research and gender analysis. In: Proceedings of the Seminar 'Participatory Research and Gender Analysis for Technology Development', Cali, Colombia, pp 65–74

Cui YX, Xu GW, Magill CW, Schertz KF, Hart GE (1995) RFLP-based assay of *Sorghum bicolor* (L) Moench genetic diversity. Theor Appl Genet 90:787–796

de Oliveira AC, Richter T, Bennetzen JL (1996) Regional and racial specificities in sorghum germplasm assessed with DNA markers. Genome 39:579–587

Deu M, Gonzalez de Leon D, Glaszmann JC, Degremont I, Chantereau J, Lanaud C, 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, Lanaud C (1995) Mitochondrial DNA diversity in wild and cultivated sorghum. Genome 38:635–645

Deu M, Rattunde F, Chantereau J (2006) A global view of genetic diversity in cultivated sorghums using a core collection. Genome 49:168–180

Djè Y, Forcioli D, Ater M, Lefèbvre C, Vekemans X (1999) Assessing population genetic structure of sorghum landraces from North-western Morocco using allozyme and microsatellite markers. Theor Appl Genet 99:157–163

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

Djè Y, Heuertz M, Ater M, Lefèbvre C, Vekemans X (2004) In situ estimation of outcrossing rate in sorghum landraces using microsatellite markers. Euphytica 138:205–212

Doggett H (1988) Sorghum, 2nd edn. Longman Scientific and Technical, London, UK

Folkertsma RT, Rattunde FH, Chandra S, Soma Raju W, Hash CT (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

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

Goudet J (2001) FSTAT, a program to estimate and test gene diversity and fixation indices (version 2.9.3). http://www.unil.ch/izea/softwares/fstat.html

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

Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol Ecol Notes 2:618–620

Harlan JR, de Wet JMJ (1972) Simplified classification of cultivated sorghum. Crop Sci 12:172–176

Hulme M (2001) Climatic perspectives on Sahelian desiccation: 1973–1998. Global Environ Change 11:19–29

Kayodé PAP, Linnemann AR, Nout RMJ, Hounhouigan JD, Stomph TJ, Smulders MJM (2006) Diversity and food quality properties of farmers' varieties of sorghum from Benin. J Sci Food Agric 86:1032–1039

Kim J-S, Klein PE, Klein RR, Price HJ, Mullet JE, Stelly DM (2005) Chromosome identification and nomenclature of *Sorghum bicolor*. Genetics 169:955–965

Mariac C, Luong V, Kapran I, Mamadou A, Sagnard F, Deu M, Chantereau J, Gerard B, Ndjeunga J, Bezançon G, Pham J-L, Vigouroux Y (2006) Diversity of wild and cultivated pearl millet accessions (*Pennisetum glaucum* [L.] R. Br.) in Niger assessed by microsatellite markers. Theor Appl Genet 114:49–58

Matsuoka Y, Vigouroux Y, Goodman MM, Sanchez GJ, Buckler E, Doebley J (2002) A single domestication for maize shown by multilocus microsatellite genotyping. Proc Natl Acad Sci USA 99:6080–6084

Maxted N, Guarino L, Myer L, Chiwona EA (2002) Towards a methodology for on-farm conservation of plant genetic resources. Genet Resour Crop Evol 49:31–46

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

Menz MA, Klein RR, Mullet JE, Obert JA, Unruh NC, Klein PE (2002) A high-density genetic map of *Sorghum bicolor* (L.) Moench based on 2926 AFLP (R), RFLP and SSR markers. Plant Mol Biol 48:483–499

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

Nkongolo KK, Nsapato L (2003) Genetic diversity in *Sorghum bicolor* (L.) Moench accessions from different ecogeographical regions in Malawi assessed with RAPDs. Genet Resour Crop Evol 50:149–156

Ollitrault O, Noyer JL, Chantereau J, Glaszmann JC (1997) Structure génétique et dynamique des variétés traditionnelles de sorgho au Burkina Faso. In: Begic A (ed) Gestion des ressources génétiques de plantes en Afrique des savanes. IER-BRG Solagral, Bamako, Mali, pp 231–240

Oosterhout SV (1997) What does in situ conservation mean in the life of a small-scale farmer? In: Sperling L, Loevinsohn M (eds) Using diversity: enhancing and maintaining genetic resources on-farm. IRDC, Ottawa, Canada. http://www.idrc.ca/en/ev-85112-201-1-DO_TOPIC.html

Parzies HK, Spoor W, Ennos RA (2004) Inferring seed exchange between farmers from population genetic structure of barley landrace Arabi Aswad from Northern Syria. Genet Resour Crop Evol 51:471–478

Perales HR, Benz BF, Brush SB (2005) Maize diversity and ethnolinguistic diversity in Chiapas, Mexico. Proc Natl Acad Sci USA 102:949–954

Perrier X, Flori A, Bonnot F (2003) Methods for data analysis. In: Hamon P, Seguin M, Perrier X, Glaszmann JC (eds) Genetic

diversity of cultivated tropical plants. Science Publishers, Inc. and CIRAD, Montpellier, pp 31–63

Petit RJ, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. Conserv Biol 12:844–855

Reenberg A (2001) Agricultural land use pattern dynamics in the Sudan-Sahel towards an event-driven framework. Land Use Policy 18:309–319

Ritland K (1996) Estimators for pairwise relatedness and individual inbreeding coefficients. Genet Res 67:175–185

Rousset F (2000) Genetic differentiation between individuals. J Evol Biol 13:58–62

Scheuring JF, Konate IM., Toure S (1980) The Malian sorghum collection. Sorghums Newsl 23:33635

Schloss SJ, Mitchell SE, White GM, Kukatla R, Bowers JE, Paterson AH, Kresovich S (2002) Characterization of RFLP probe sequences for gene discovery and SSR development in *Sorghum bicolor* (L.) Moench. Theor Appl Genet 105:912–920

Smouse PE, Peakall R (1999) Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. Heredity 82:561–573

Snowden JD (1936) The cultivated races of sorghum. Adlard, London, UK, pp 1–274

Sokal RR, Wartenberg DE (1983) A test of spatial autocorrelation analysis using an isolation-by-distance model. Genetics 105:219–237

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

Tescar RP (2004) Echanges Informels des Semences de sorgho (*Sorghum bicolor* L. Moench) au Burkina, études de cas: le Centre-Ouest, le Centre-Nord et la Boucle de Mouhoun, DEA Dissertation, INA P-G, Paris, 42 p

Teshome A, Fahrig L, Torrance JK, Lambert JD, Arnason TJ, Baum BR (1999) Maintenance of sorghum (*Sorghum bicolor*, Poaceae) landrace diversity by farmers' selection in Ethiopia. Econ Bot 53:79–88

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

vom Brocke K, Christinck A, Weltzien E, Presterl T, Geiger HH (2003) Farmers' seed systems and management practices determine pearl millet genetic diversity patterns in semiarid regions of India. Crop Sci 43:1680–1689

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370

Wezel A, Boecker R (1998) Fallow plant communities and site characteristics in semi-arid Niger, West Africa. J Arid Environ 40:269–280

Zongo JD (1991) Ressources génétiques des sorghos (*Sorghum bicolor* L. Moench) du Burkina Faso: Evaluation agromorphologique et génétique. PhD Thesis, University of Abidjan, Ivory Coast

Zongo JD, Gouyon PH, Sarr A, Sandmeier M (2005) Genetic diversity and phylogenic relations among Sahelian sorghums accessions. Genet Resour Crop Evol 52:869–878