Population Structure and Diversity in Finger Millet (Eleusine coracana) Germplasm

Article *in* Tropical Plant Biology · January 2008

DOI: 10.1007/s12042-008-9012-3

Tropical Plant Biol. (2008) 1:131–141
DOI 10.1007/s12042-008-9012-3

Population Structure and Diversity in Finger Millet (*Eleusine coracana*) Germplasm

Mathews M. Dida · Nelson Wanyera · Melanie L. Harrison Dunn · Jeffrey L. Bennetzen · Katrien M. Devos

Received: 30 March 2008 / Accepted: 29 April 2008 / Published online: 18 June 2008

Abstract A genotypic analysis of 79 finger millet accessions (*E. coracana* subsp. *coracana*) from 11 African and five Asian countries, plus 14 wild *E. coracana* subsp. *africana* lines collected in Uganda and Kenya was conducted with 45 SSR markers distributed across the finger millet genome. Phylogenetic and population structure analyses showed that the *E. coracana* germplasm formed three largely distinct subpopulations, representing subsp. *africana*, subsp. *coracana* originating from Africa and subsp. *coracana* originating from Asia. A few lines showed admixture between the African and Asian cultivated germplasm pools and were the result of either targeted or accidental intercrossing. Evidence of gene flow was also seen between the African wild and cultivated subpopulations, indicating that hybridizations among subspecies occur naturally where both species are sympatric. The genotyping, combined with phylogenetic and population structure analyses proved to be very powerful in predicting the origin of breeding materials. The genotypic study was complemented by a phenotypic evaluation. The wild and cultivated accessions differed by a range of domestication-related characters, such as tiller number, plant height, peduncle length, seed color and grain yield. Significant differences in plant architecture and yield were also identified between the Asian and African subpopulations. The observed population structure within cultivated finger millet is consistent with the theory that, after the introduction of finger millet from Africa into India via the trade routes some 3000 years ago, the two germplasm pools remained largely isolated until recent times. The significantly lower diversity present within the Asian subpopulation also suggests that it arose from a relatively small number of founder plants.

Keywords Biodiversity · *Eleusine coracana* · Finger millet · Phylogenetic analysis · Population structure

Communicated by Ray Ming

Electronic supplementary material The online version of this article (doi:10.1007/s12042-008-9012-3) contains supplementary material, which is available to authorized users.

M. M. Dida
Department of Botany and Horticulture, Maseno University, P.O. Box 333, Maseno, Kenya

N. Wanyera
National Agricultural Research Organization, SAARI, P.O., Soroti, Uganda

M. L. Harrison Dunn
USDA-ARS, 1109 Experiment Station, Griffin, GA 20223, USA

J. L. Bennetzen
Department of Genetics, University of Georgia, Athens, GA 30602, USA

M. M. Dida · K. M. Devos (✉)
Department of Crop and Soil Sciences, and Department of Plant Biology, 3111 Miller Plant Sciences, University of Georgia, Athens, GA 30602, USA
e-mail: kdevos@uga.edu

Introduction

Finger millet, *Eleusine coracana* L. Gaertn., is a tetraploid crop ($2n=4x=36$; genome constitution AABB) belonging to the grass family *Poaceae*, subfamily Chloridoideae. It is cultivated mainly in East Africa and Southern India. Precise figures on the amount of finger millet grown worldwide are not available, as finger millet production data are generally grouped with those of the Panicoids pearl, foxtail and proso millet. However, it is estimated that some 10% of the

world's 30 million ton of millet produced is finger millet. The yield potential for finger millet is in the range of 4–5 tons/ha [2, 15, 19], but yields vary greatly depending on the country and region. In India, yields might reach 1 ton/ha on dryland sites and average at least 2 tons/ha under irrigated conditions [17]. Millet yields in Kenya, Uganda, Ethiopia, Tanzania, Zambia, Zimbabwe and Malawi vary from 0.3 tons/ha in Zimbabwe to 1.6 tons/ha in Uganda [6]. In these East-African countries, finger millet is the main millet cultivated, hence the 'millet' figures can be used to represent trends in finger millet. Yields in Uganda have steadily increased from about 0.9 ton/ha in 1960s to 1.6 tons/ha in 2006. Yield increases of some 0.5 ton/ha have also been achieved in Ethiopia over the past 30 years. In Kenya, on the other hand, yields have sharply declined, in particular between 1978 (1.6 tons/ha) and 1981 (0.7 ton/ha). Some of the factors that might have played a role in this decline are an increased focus on maize as a food crop which has resulted in the displacement of finger millet to areas with more marginal soils and growing conditions, and the lack of a continuous finger millet breeding program. Even today, Kenyan farmers rely largely on landraces for finger millet production. The improved varieties that are grown in Kenya have mostly been bred in Uganda, where a breeding program has been in existence since 1965. Finger millet breeding in Africa has been largely limited to pure line selection of local landraces. Hybridization-based breeding has been conducted in Uganda and India since the late 1960s–early 1970s and included the crossing of Indian and African cultivars in an attempt to widen the germplasm base. The first Indaf variety was released in India in 1976. The Ugandan variety Engeny, supposedly the result of an intercross between African and Indian germplasm, was released in 1969.

The oldest archeological *Eleusine coracana* sample currently available was discovered in Axum, Ethiopia and is thought to date back to 3000 B.C. [11]. Cultivated finger millet, *E. coracana* subsp. *coracana*, is considered to have been domesticated some 5,000 years ago from the wild *E. coracana* subsp. *africana* ($2n=4x=36$) in the highland that stretches from Ethiopia to Uganda. Subsp. *africana* is the result of a spontaneous hybridization event between the diploid *E. indica* (AA genome) and an unknown B-genome donor [18]. Domesticated finger millet was subsequently moved to the lowlands of Africa and around 1000 B.C. introduced into India through the sea trade that existed between India and Africa [10]. India became a secondary centre of diversity for finger millet.

Genetic variation within *Eleusine coracana* subsp. *coracana* is limited. All isozyme and DNA marker studies carried out to date have shown an overall low level of variation between cultivated finger millet lines, including lines with adaptation to different geographical regions [4, 12, 16, 22, 26]. Morphological variation, on the other hand, is large. A 1997 study of some 2,000 finger millet lines from the ICRISAT, Asia Center, collection has shown a considerable range in flowering time (54–120 days), plant height (45–165 days), number of basal tillers (1–70), peduncle length (2–28 cm), inflorescence length (1–32 cm) and other morphological traits [20].

The aim of the current study was to evaluate the genotypic variation that is present in *Eleusine* germplasm sampled across a wide geographical range, using robust simple sequence repeat (SSR) markers. Correlation of the SSR genotypes with the phenotypic diversity measured during field trials in Kenya and Uganda provides breeders with information on those lines that have potential as crossing parents for improving locally adapted varieties.

Results

Diversity Statistics

The average number of alleles per taxon or population analyzed, frequency of the major allele and gene diversity averaged over each taxon are given in Table 1. The corresponding values for each SSR can be found in Supplementary Table 1. The wild *E. coracana* subsp. *africana* germplasm has the highest level of divergence (0.39) followed by the African cultivated material (0.33). The Asian finger millet germplasm that was analyzed shows the least variation (0.22).

Phylogeny

The bootstrap consensus phylogenetic tree is presented in Fig. 1. All but two wild subsp. *africana* accessions formed a clade on the tree with moderate support (70%). Wild accessions EK6 and IE 2595 were placed towards the base of a well-supported (81%) clade including all cultivated subsp. *coracana* accessions. Within the subsp. *africana* clade, the accessions are separated by geographic region, although this separation is only weakly supported by bootstrapping (57%). *E. indica* appears as a sister clade to subsp. *africana*. Bootstrap values for all branches within the subsp. *coracana* clade are generally low.

Population Structure

The results of the population structure analysis indicate that the *E. coracana* germplasm analyzed belongs to three subpopulations (Fig. 2). An accession was considered to belong to a single subpopulation if ≥90% of its alleles were allocated to this subpopulation. This was the case for 90% of the *E. coracana* subsp. *coracana* accessions and 86% of

Table 1 Diversity statistics for 45 finger millet SSRs

Taxon	No. of accessions	Average no. of alleles per SSR	Major allele frequency	Gene diversity
E. coracana subsp. *coracana*—Asian subpop. *ss* (and geogr)[a]	19 (26)	2.20 (2.80)	0.84 (0.81)	0.22 (0.27)
E. coracana subsp. *coracana*—African subpop. *ss* (and geogr)[a]	40 (47)	3.36 (3.55)	0.75 (0.75)	0.33 (0.33)
E. coracana subsp. *coracana*	79	4.04	0.74	0.34
E. coracana subsp. *africana*[b]	12	2.47	0.68	0.39
Subsp. *coracana* + *africana*	93	5.71	0.67	0.45
All *Eleusine* species	96	6.42	0.66	0.47

[a] *sensu strictu* (*ss*): includes only Asian/African varieties that grouped within the Asian/African subpopulation based on the program STRUCTURE. Geographic (geogr) includes all varieties that were listed as originating from Asia/Africa. The diversity statistics for the geographic groupings are given in parenthesis

[b] Does not include IE 2595 and EK6, which were identified as hybrids between subsp. *coracana* and subsp. *africana*

the subsp. *africana* lines. The remaining lines are likely to be hybrids. The first subpopulation, hereafter referred to as the African *coracana* subpopulation, consists largely of *E. coracana* subsp. *coracana* of African origin. A total of 95% of the non-hybrid African subsp. *coracana* accessions, 17% of the Asian subsp. *coracana* accessions and 6 lines of unknown origin belong to this subpopulation with 90% or more of their alleles. The Asian *coracana* subpopulation contains 83% and 5% of the non-hybrid cultivated Asian and African accessions, respectively. Subpopulation *africana* consists entirely of subsp. *africana* accessions. Under our definition of a hybrid accession (<90% of the alleles belonging to a single subpopulation), 10% of the cultivated accessions exhibited evidence of admixture between the Asian and African *coracana* subpopulations. Two accessions (14%) that were characterized morphologically as belonging to subsp. *africana* showed evidence of gene flow between finger millet and its wild progenitor, *E. coracana* subsp. *africana*.

Trait Variation Between Subpopulations

A number of traits were significantly different (95% confidence limit) between accessions belonging to the Asian and African *coracana* subpopulations (Fig. 3; Supplementary Table 2). Average plant height, depending on the trial, was between 27% and 42% higher across the African compared to the Asian *coracana* subpopulation. Similar increases were also seen for the length (26%) and width (32%) of the flag leaf. The number of basal and axillary productive tillers, on the other hand, was 30% and 60% lower, respectively, across the African compared to the Asian *coracana* subpopulation (Fig. 3A,B). The lower tiller number in the African *coracana* subpopulation is compensated for by the presence of some 20% more fingers per head and some 20% more grains per spikelet (Fig. 3C). Mean grain yield as determined by the dry weight of seed heads per plot in the African *coracana* subpopulation was more than twice that of the Asian *coracana* subpopulation in the tested African locations (Fig. 3D).

When comparing traits across the cultivated African and wild *africana* subpopulations, wild accessions flowered and matured significantly earlier than the cultivated lines. Significant differences in plant height between the wild and cultivated accessions were seen in the Ugandan, but not in the Kenyan, trial. Subspecies *africana* accessions tillered more profusely, but tillers had fewer axillary branches compared to African *coracana* finger millet accessions. Other significant morphological differences included the presence of longer peduncles, longer fingers and darker seeds in the *africana* compared to the African *coracana* accessions. Head blast infection levels were higher in the wild accessions compared to the cultivated African accessions in the Kenyan field trial, but lower in the Ugandan trial.

Discussion

Phylogenetic Analysis

The analysis was conducted on a sample of 79 *E. coracana* subsp. *coracana* accessions, 26 from Asia (mainly India), 47 from 11 African countries, and six of unknown origin using 45 SSR markers. In addition, a total of 14 *E. coracana* subsp. *africana*, two *E. indica* and one *E. kigeziensis* accession were analyzed. This is the largest study to date, both in terms of the number of markers and the number of accessions analyzed, that aims to characterize the diversity present in finger millet germplasm across its range of cultivation. Of the 45 SSRs used, 24 have been mapped and are distributed over 13 of the 18 finger millet chromosomes [4]. The location of the remaining SSRs is unknown, but the assumption is that the SSRs are spread across the genome.

The phylogenetic tree based on those 45 SSRs and rooted on the wild species *E. kigeziensis* shows two major

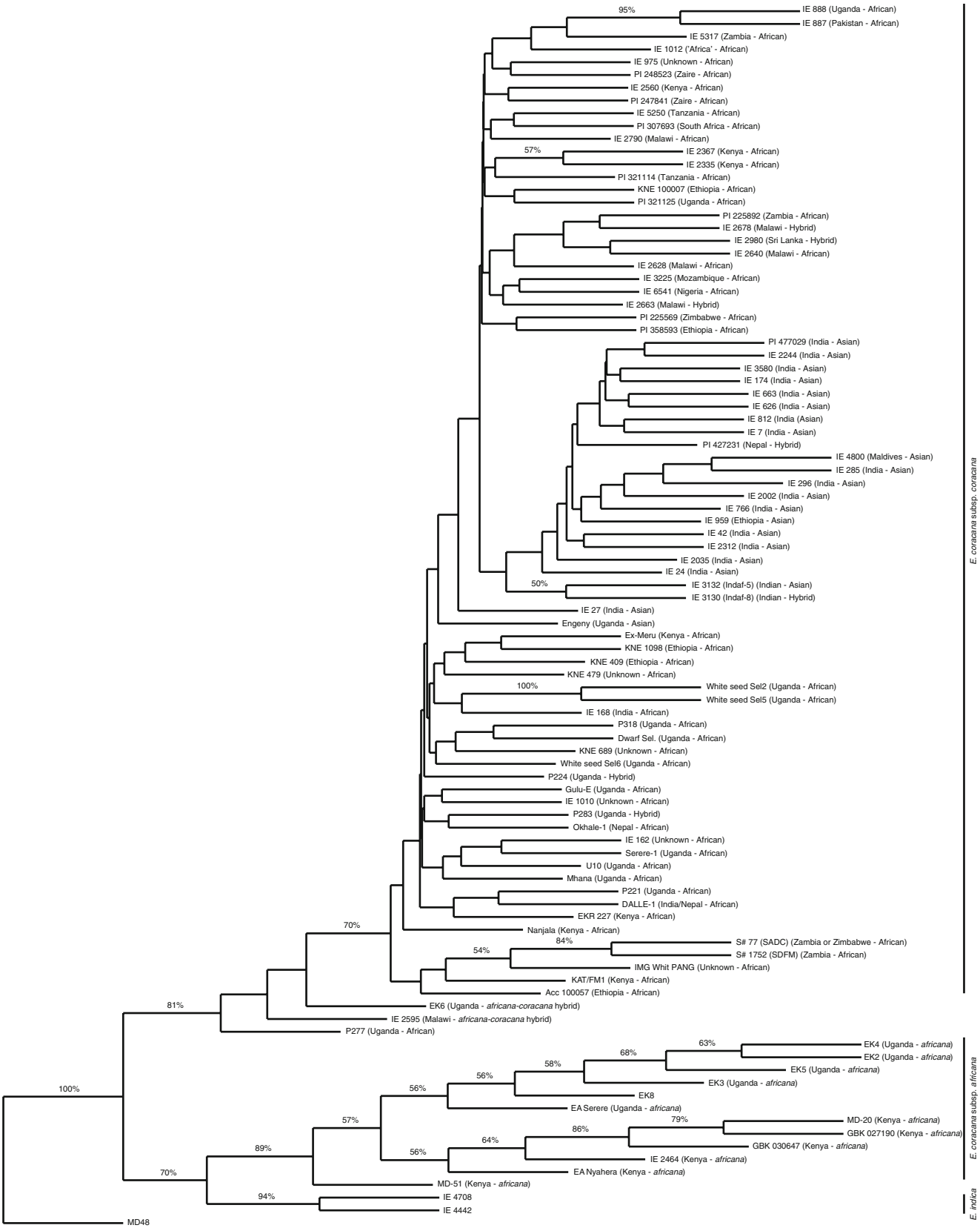

Fig. 1 Bootstrap consensus tree of 79 *E. coracana* subsp. *coracana* accessions, 14 *E. coracana* subsp. *africana* accessions, two *E. indica* accessions and one *E. kigeziensis* (MD-48) accession. The latter was used as the outgroup to root the tree. Bootstrap values ≥50% are shown. For each accession, the geographic origin of the line and the subpopulation to which the line belongs based on the STRUCTURE analysis, are given in *parenthesis*

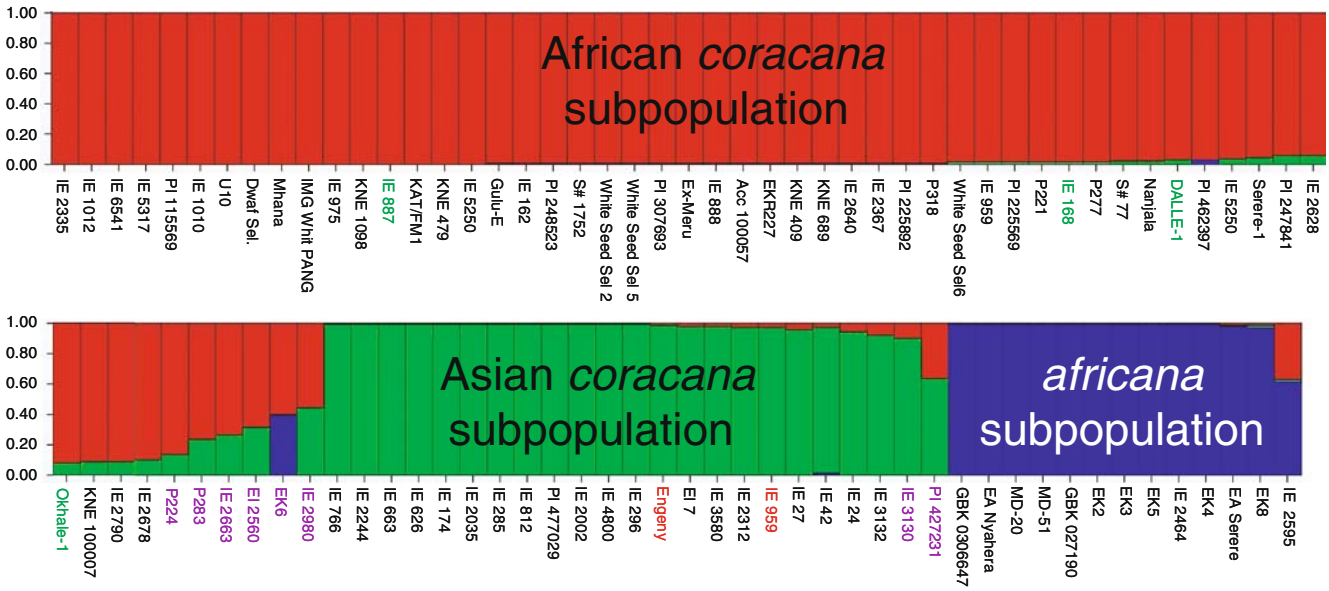

Fig. 2 Classification of the *E. coracana* germplasm analyzed into three subpopulations referred to as African *coracana*, Asian *coracana* and *africana* subpopulations. Accessions written in *purple* exhibit more than 10% admixture. Accessions that belong to the African *coracana* subpopulation and are written in *green* are of Asian origin. Accessions that belong to the Indian *coracana* subpopulation and are written in *red* are of African origin

clades, one consisting of mainly *E. coracana* subsp. *coracana* and a clade consisting of subsp. *africana* and the two *E. indica* accessions analyzed. Since the wild subsp. *africana* is hypothesized to have originated from a hybridization event between an *E. indica* accession and an unknown B-genome donor, one would expect the diploid *E. indica* to be sister to *E. coracana* subsp. *africana* and this is what we observed in the phylogenetic analysis (Fig. 1). One of the *E. indica* lines was collected from what was described as a 'farm or cultivated habitat' and its closer phylogenetic relationship to the other *indica* accession and the wild subsp. *africana* rather than the cultivated subsp. *coracana* suggests that gene flow from the cultivated crop to *E. indica* is not ubiquitous.

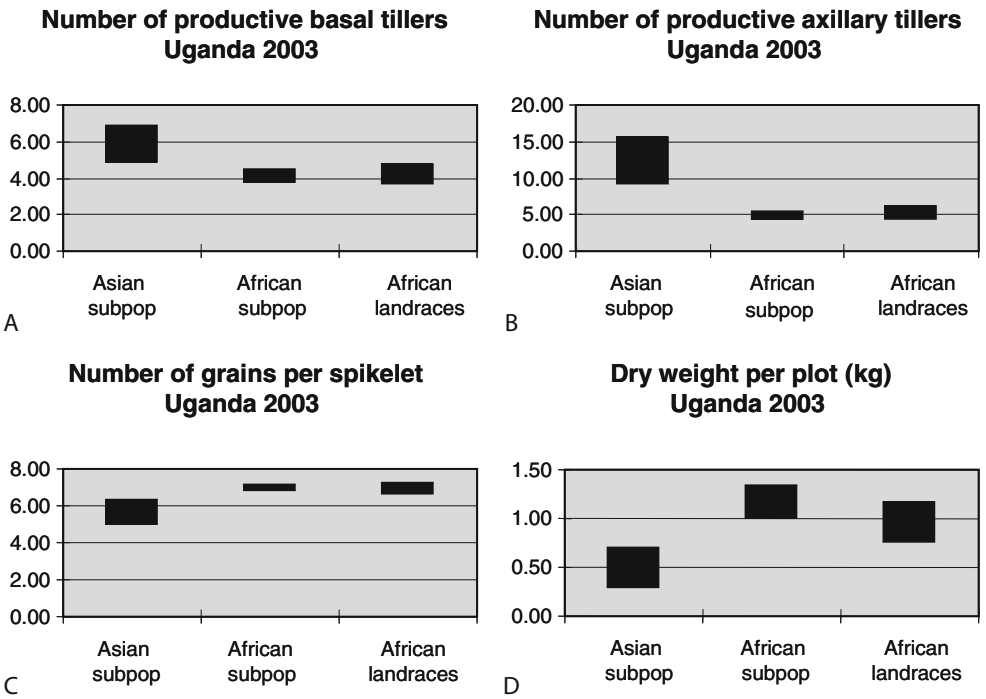

Fig. 3 Examples of traits that are significantly different among the Asian *coracana* and African *coracana* subpopulations. Because the African *corcacana* subpopulation includes both landraces and elite lines, the comparison was extended to a subset of the African *coracana* subpopulation that comprises only landraces (African landraces). The blocks delineate the 95% confidence intervals of the trait means in the subpopulations. Dry weight/plot (3D) represents the weight in kilogram of the sun-dried heads of all plants in a plot

The subsp. *africana* accessions largely group by country (bootstrap value of 57%). The only exception is MD-51, an accession collected in Kitale, Kenya, which is sister to the remainder of the Kenyan and Ugandan *africana* accessions. This suggests that MD-51 carries alleles of both Ugandan and Kenyan subsp. *africana* germplasm. Two further accessions that were morphologically characterized as subsp. *africana*, EK6 and IE 2595, grouped with the cultivated subsp. *coracana*. The placement of EK6 and IE 2595 with the cultivated clade suggests that both accessions are of hybrid origin with genomic regions drawn from wild and cultivated germplasms. This inference is supported by the STRUCTURE analysis (see below).

Within the cultivated clade, very few branches are supported by high bootstrap values, but some geographic patterning can nevertheless be observed. The majority of the accessions of Asian origin appear to group together (Fig. 1) and are sister to a group of accessions that originate mostly from African lowland countries such as South Africa, Mozambique, Zimbabwe, Malawi, Tanzania, Zambia and Zaire. The Indian and African lowland accessions are more distantly related to the wild subsp. *africana* than the majority of Kenyan, Ugandan and Nepalese varieties, which are more adapted to the highlands. This fits with the hypothesis that finger millet was first domesticated in the African highlands, then spread south where the crop became adapted to the lowland, and from there was introduced into India. Some close phylogenetic relationships, supported by bootstrap values ≥50%, can be observed. A high similarity can be expected for lines derived from the same breeding program such as the two Indaf lines analyzed (bootstrap value of 50%). The same may also be true for the two Kenyan accessions IE 2335 and IE 2367 (57%) and for the Ugandan White Seed Sel2 and 5 (100%). There is some doubt on the origin of accession S# 77 (either Zambia or Zimbabwe), but its high similarity with the Zambian accession S# 1752 (bootstrap value of 84%) suggests that both lines might have their origins in Zambia. More difficult to explain are the tight clustering of the Pakistani accession IE 887 with the Ugandan accession IE 888 (95%), and of the Zambian accessions S# 1752 and S# 77 with the Kenyan accession Kat/FM1 (54%). Kat/FM1 is a selection made between 1979 and 1985 from a local cultivar from the Machakos district in Kenya [14]. Two possible hypotheses can be put forward. The simplest one is that the origin of the lines has been incorrectly recorded or that sample mix-up occurred during collection, storage or seed multiplication. Alternatively, the similarity between lines from different geographic origins might be indicative of germplasm exchange between different breeding programs. It was interesting to note that, although White Sel2 and 5 are highly similar and presumably the result of selection within the same population, a third line, White Sel6, is much more diverse (Fig. 1) and, despite the similar naming, has likely been selected from a different white-grained population. This hypothesis is supported by the phenotypic differences between White Sel2 and 5, which have purple stems, inflorescences and leaves, and White Sel6, which is green.

Population Structure

Since some geographical grouping, albeit with little bootstrap support, was observed in the phylogenetic analysis, we wanted to evaluate whether alternative clustering methods such as model-based clustering would resolve the geographic patterning. From the phylogenetic analysis, we knew that there were at least two population subgroups, largely corresponding to the subsp. *coracana* and *africana*. The phylogenetic analysis had also provided some evidence for gene flow between the two subpopulations. Using the software program STRUCTURE and allowing for admixture between the subpopulations, three subpopulations, a wild subsp. *africana* subpopulation, an African subsp. *coracana* and an Asian subsp. *coracana* population, were identified within the *E. coracana* germplasm analyzed. There was good correspondence between the geographic patterning observed in the phylogenetic tree and the population structure identified using STRUCTURE. Although the population subgroups corresponded largely to geographic regions, there were some notable exceptions. Excluding any lines that showed evidence of admixture (<90% of alleles in one subpopulation), 95% of the cultivated African accessions belonged to the African subpopulation. The two African lines that grouped with the Asian subpopulation are Engeny, a selection from the progeny of a cross between Indian and African germplasm released in 1969 by the East African Agriculture and Forestry Research Organization (EAAFRO) in Uganda, and IE 959, an accession originating from Ethiopia. If Engeny truly was the result of hybridization breeding between adapted Ugandan and Indian accessions, a degree of admixture would have been seen. The data suggest that Engeny was a selection from progeny from either a selfing event of an Indian accession or from a cross between two Indian accessions. Considering that hybridization breeding was definitely not routine at that time, selfing of a female Indian parent rather than the intentional cross-pollination with African germplasm seems a logical explanation. We have no information on the origin of IE 959.

Seventeen percent of the Asian accessions fell within the African *coracana* subpopulation. A closer look at the origin of these lines indicates that one originates from Northern India (IE 168), one from Nepal (Okhale-1) and one from Pakistan (IE 887). DALLE-1, although originally bred in India, was released in Nepal in 1980 (http://www.narc.org.np/publicaton/pdf/varieties_released/VarietiesEng.pdf) and

is adapted to the Nepalese mid-hills. Four other lines (IE 162, IE 975, IE 1010 and IE 1012), listed in the Germplasm Resources Information Network (GRIN) database as originating from India, also grouped with the African *coracana* subpopulation. However, IE 1012 was referred to by Gowda et al. [8] as an African cultivar exploited in India as a source of blast resistance. Accessions IE 975 and IE 1010 are registered as accessions of unknown origin by ICRISAT, which is the seed source of the IE-lines present in the USDA-PGRCU collection. No information could be obtained on accession IE 162. If we exclude IE 162, IE 975 and IE 1010, which might be listed incorrectly, from the analysis, then all four Asian accessions that group with the African *coracana* subpopulation have been bred for the Northern regions (Northern India, Pakistan and Nepal) so they may be selections from recent introductions from the African highlands. Such a scenario is supported by the close phylogenetic relationship between IE 887 from Pakistan and IE 888 from Uganda (Fig. 1; 95% bootstrap support).

Evidence of admixture between Indian and African germplasm was found in the Indian accessions Indaf-8 and PI 427231. Indaf varieties are the result of intercrossing between Indian and African germplasm, a breeding strategy that was adopted in India in the 1970s. The STRUCTURE program classified some 10% of the alleles present in Indaf-8 as being of African origin. A second Indaf variety analyzed, Indaf-5, carried only about 8% African alleles. A more detailed marker study would be needed to determine how the African alleles are distributed across the Indaf genomes. The Nepalese line PI 427231, on the other hand, carried 36% of African alleles and was clearly of hybrid origin. Considering our earlier observation that at least some Nepalese germplasm appears to be derived from African introductions, this high level of admixture may not be too surprising. A high level of admixture with 56% African and 43% Asian alleles was also seen in IE 2980, a line from Sri Lanka. Finger millet in Sri Lanka is grown mainly in upland locations, particularly in dry areas. Since only one line from Sri Lanka was included in our analysis, it is unclear whether this high level of admixture is representative for the finger millet lines grown in Sri Lanka. Evidence of admixture with Asian accessions was also found in African germplasm. IE 2560 (Kenya), IE 2663 (Malawi), IE 2678 (Malawi), P224 (Uganda) and P283 (Uganda) have Indian alleles at between 10 and 32% of their loci tested. P224 and P283, the lines developed by the Ugandan breeding program, are known to be selections from crosses between Indian and African germplasm. No information is available on the other accessions. Consistent with the results of the phylogenetic analysis, two of the 14 lines with a subsp. *africana* phenotype showed admixture with the African cultivated subsp. *coracana* subpopulation. EK6 was collected as a wild accession, but the majority of the alleles present in this line are of cultivated origin. Since EK6 was collected near a farm, we can conclude that this accession resulted from a cultivated–wild intercross. Although the majority of the alleles are cultivated, the wild phenotype presumably provides a higher fitness in the natural environment. IE 2595, on the other hand, is a supposedly cultivated line that was reclassified within this project as belonging to subsp. *africana* based on its phenotype. IE 2595 has a genome that is 62% wild. EK6 and IE 2595 provide the first genotypic evidence that unambiguously demonstrates that gene flow between subsp. *coracana* and *africana* occurs naturally where both species are sympatric [3]. Neither of these accessions are first generation hybrids, indicating that some hybrid derivatives are competitive and can persist in the wild.

Diversity Analysis

In the first instance, we assessed the variation present both within cultivated subsp. *coracana* and within wild subsp. *africana* germplasm. Although the wild accessions were obtained from only two countries, Kenya and Uganda, the variation present within subsp. *africana* (diversity of 0.39) was higher than in the cultivated germplasm (0.34) whose collection range spanned two continents (Table 1). The decreased diversity of the cultivated compared to the wild germplasm is indicative of a domestication bottleneck. Next, we investigated whether the introduction of finger millet from Africa into Asia some 3000 years ago represented a second bottleneck. *E. coracana* subsp. *coracana* accessions were classified as Asian or African based on (1) the results of the population structure analysis and (2) their reported region of origin, and the mean diversity within each subpopulation across all 45 SSR loci was established. The diversity within the Asian subgroup is considerably lower (diversity of 0.22) compared to the African subgroup (0.33; Table 1). When only the geographic information is used to classify the lines into subgroups, the diversity of the Asian and African subgroups is 0.27 and 0.33, respectively. The diversity data suggest that the Indian germplasm pool was created from a limited number of founder populations and contradicts an observation from an earlier study comprising seven African and ten Asian accessions that the diversity is fairly evenly represented within both Africa and Asia [22]. Our data suggest that Indian finger millet germplasm has undergone two bottlenecks, a first one resulting from the domestication of the crop and a second one upon its introduction from Africa into India. Hybridization-based improvement of finger millet within the Indian germplasm pool will therefore have limited scope. Finger millet breeders in India are addressing the lack of novel alleles within Indian

germplasm by intercrossing adapted Indian and African accessions. Several of the popular varieties released in India since the 1970s, including the Indaf, MR and GPU lines, are Indian–African hybrids [23].

Trait Variation Between Subpopulations

Considering the existence of a clear population structure in the *E. coracana* germplasm analyzed, it is not surprising that distinct differences can be found in morphological characters between the subpopulations. The differences between the wild *africana* subpopulation and cultivated African germplasm reflect mainly the change in morphology from wild to domesticated phenotype. Typical characteristics that are selected during domestication include non-shattering of the grain, larger grain, paler grain color, increased panicle size, reduced tillering, and adaptations in flowering time and plant height. Seed size and shattering were not determined in our study, but significant differences were found in the number of basal tillers, peduncle length, flowering time, seed color and grain yield (Supplementary Table 2). Significant differences in plant height were also observed between the wild and cultivated African germplasm in the Ugandan trial but not in the Kenyan field trial. It should be noted that, while there was a significant correlation between trait measurements such as plant height and flowering time made in the Ugandan and Kenyan field trials for cultivated accessions, this was not the case for the wild germplasm. Wild accessions tend to show considerable phenotypic plasticity depending on the prevailing environmental conditions, a characteristic that appears to have been largely lost in the domestication process.

When comparing the Asian and African *coracana* subpopulations, the overall trend observed was that Asian accessions are smaller in stature, have smaller flag leaves, have around one to two more basal tillers and have three times more axillary tillers than African accessions. They also have fewer fingers per panicle, they have fewer seed per spikelet and they exhibited much lower yield than the African accessions (Fig. 3; Supplementary Table 2). Two possible reasons for the poorer performance of the Asian compared to the African *coracana* accessions in Kenyan and Ugandan field trials spring to mind. Firstly, the Indian accessions may not be as well adapted to the Kenyan and Ugandan environments as the lines belonging to the African subpopulation. Although the geographic range of the African subpopulation stretches from South Africa to Ethiopia and includes Nepal, more than half of the accessions in this subpopulation originate from Kenya and Uganda and are thus locally adapted. The effect of the environment on the performance of the lines could have been assessed by growing the germplasm selection in Southern India, an experiment we did not conduct. Secondly, some 30% of the studied African subpopulation are elite varieties, while the Indian accessions analyzed are mainly landraces. When the elite lines were removed from the analysis and trait values of only the Asian and African landraces were compared, all differences remained significant (Fig. 3; Supplementary Table 2). The average yield of the African *coracana* subpopulation, expressed as dry weight per plot, decreased by some 20% when the elite lines were removed (Fig. 3D), but the difference in yield between the Asian and African *coracana* subpopulations remained significant. It was interesting to note that the elite varieties were also significantly taller than the landraces but, as a group, flowered earlier than the landraces. It thus appears that the trend of selecting for increased plant height, either directly or indirectly by selecting for higher yields, which started with the domestication of finger millet is still continuing, at least in Africa.

The overall level of resistance to the blast fungus, *Pyricularia oryzae*, was comparable between the Asian and African subpopulations. Early flowering accessions from both continents were very susceptible to blast. Blast infection levels were negatively correlated with flowering time, and it has been proposed that late flowering lines tend to escape head and neck infections, which are the most devastating in terms of yield losses [5]. Some later flowering accessions such as IE 2312, Ex-Meru, IE 2560 and IE 2663 had lower blast incidence ratings but scored higher in disease severity, which is indicative of disease escape. The wild accessions, on the other hand, had lower overall infection levels than the cultivated lines in Uganda, but higher infection levels in Kenya. It is not clear whether this is due to differences in the pathogen population or environmental conditions. Nevertheless, lines were identified such as Mhana and the wild/cultivated hybrid EK6 that were intermediate and early, respectively, in flowering time but displayed good levels of resistance in both the Ugandan and Kenyan field trials.

Conclusions

Genotyping combined with phylogenetic and population structure analyses is a powerful method for characterizing germplasm. Based on the genotypic data, we were able to predict origins of breeding materials that were unknown at the time the experiments were performed. In several cases, we were able to provide support for the predicted origin through extensive literature searches and information provided by finger millet breeders. The existence of two population groups within *E. coracana* subsp. *coracana*, corresponding largely to the African and Asian continents and the lower diversity within the latter group is consistent with the theory that finger millet was domesticated in Africa and then introduced into India with the two groups

remaining largely isolated until recent times [3]. The two subpopulations basically form two distinct germplasm pools that could be employed to enhance finger millet germplasm in both India and Africa. Hybridization-based breeding between Indian and African germplasm has been ongoing since the 1970s, and has resulted in the release of several high-yielding varieties. Although the lower performance of the Indian lines in Africa would seem to indicate that African germplasm will not benefit from the introduction of alleles found in the Indian germplasm pool, one has to remember that plants with apparently unsuitable phenotypes may carry favorable alleles, as has been unambiguously demonstrated by Tanksley and colleagues [24, 25]. An interesting observation in this respect is that P224 and P283, two Ugandan elite lines resulting from intercrossing between African and Indian germplasm that carried 14 and 23% of Indian alleles, respectively, are high-performing lines. In fact, P224 is known as one of the most stable varieties in terms of yield. It thus seems likely that the Indian alleles contributed to the varietal enhancement. Favorable alleles for finger millet improvement for traits such as resistance to blast and high protein content might also be obtained from the wild subsp. *africana* [1]. However, the fact that subsp. *africana* is rarely represented in germplasm collections indicates that wild germplasm is not yet receiving extensive use as a potential gene donor.

Methods

Plant Material

Seventy-nine *E. coracana* subsp. *coracana* accessions from 11 African and five Asian countries were obtained from the Kenyan Agricultural Research Institute (KARI), Kakamega, Kenya; Serere Agricultural and Animal Production Research Institute (SAARI), Soroti, Uganda; the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Nairobi, Kenya and ICRISAT, Patancheru, India; and the Plant Genetic Resources Conservation Unit (PGRCU) of the USDA-ARS, Georgia, USA. Fourteen *E. coracana* subsp. *africana* accessions, originating from Kenya, Uganda and Malawi were obtained from ICRISAT, SAARI and the Gene Bank of Kenya, Muguga, Kenya, or were collected by two of the authors (N. Wanyera and M.M. Dida). The two *E. indica* lines used in the analysis were obtained from ICRISAT. The *E. kigeziensis* accession that was used as an outgroup was collected in Uganda by M.M. Dida. A list of all germplasm used, their country and, where known, region of origin and the source of the seed is given in Supplementary Table 3.

Trait Analyses

Because not all 96 lines could be assessed in a single field trial due to space and labor limitations, a subset of 62 finger millet accessions were selected for phenotyping based on geographic distribution and seed availability. The finger millet accessions were grown in 2003 during the first rainy season at SAARI, Uganda and during the 2003 and 2004 long rainy seasons at Maseno in Kenya. The accessions were planted in a randomized complete block design with three replications. Plots were 3 m long single rows with spacing of 45×10 cm in the Kenyan trial and consisted of two 5 m rows with spacing of 30×10 cm in the Ugandan trial. For the 2003 trial in Maseno, diammonium phosphate (DAP 18:46) was applied at a rate of 70 kg/ha at the time of planting. No fertilizer was applied in the other two trials. Hand weeding was done twice. Measurements, including plant height, the number of productive basal tillers, number of productive axillary tillers, days to 50% flowering, days to 50% maturity, ear shape, finger number, finger length, finger branching, number of grains per spikelet, flower color (1=green; 2=purple), grain color, blade width and length of the flag leaf, peduncle length and severity of blast infections caused by *Pyricularia oryzae* on foliage, neck and finger were taken on five (Kenya) or ten (Uganda) randomly selected plants for each accession as described in the list of descriptors for finger millet [9]. Not all traits were measured in each trial. The traits measured in the 2003 Ugandan trial and the 2003 and 2004 Kenyan trials are listed in Supplementary Table 2. Blast severity at the Maseno site was recorded on a scale of 1 to 5 with 1 representing the least and 5 the highest infection. Blast incidence was calculated as the percentage of plants in the plot with head or finger blast. Yield parameters measured include seed yield/head (mean weight in grams of ten randomly selected heads from different plants per replicate; Kenyan trial) and dry weight/plot (weight in kilogram of the sun-dried heads of all plants in a plot; Ugandan trial).

SSR Markers

Forty five SSR markers, described in Dida et al. [4], were used for genotyping. SSR amplification was carried out using an M13-tailed SSR-specific forward primer, a SSR-specific reverse primer and a fluorescently labeled M13 primer as previously described [4]. Four SSRs, labeled with different fluorochromes, were pooled and run on an ABI 3730xl. Patterns were analyzed using GeneMapper v. 3.5 (Applied Biosystems). Allele scores were verified manually. Finger millet is an inbreeding species and heterozygous loci were therefore considered to be derived from outcrossing or seed mixtures and entered as missing data.

Diversity Analyses

Summary statistics, including the number of alleles, the frequency of the major allele and the gene diversity (expected heterozygosity) for all *Eleusine* lines, for *E. coracana* subsp. *coracana*, for subsp. *africana* and for the African and Asian subpopulations within subsp. *coracana* were calculated using PowerMarker [13].

Phylogenetic Analysis

Genetic distances (Dps) were calculated based on the proportion of shared alleles (ps) with Dps=1−ps using the program MICROSAT v. 1.5 (Eric Minch, Stanford University, USA; http://hpgl.stanford.edu/projects/microsat/). Five hundred bootstrapped distance matrices were used to construct a consensus phylogenetic tree as implemented in the FITCH program in Phylip v. 3.66 ([7]; http://evolution.genetics.washington.edu/phylip.html). The input of samples was randomized (J option in FITCH). The resulting tree was rooted using *E. kigeziensis* acc. MD-48 as the outgroup.

Population Structure Analysis

Analysis of the population structure and of gene flow between *E. coracana* subsp. *coracana* and subsp. *africana* was carried out using a model-based clustering method as implemented in the software program STRUCTURE v. 2.1 [21]. In this method, it is assumed that a number of subpopulations exists in the sample analyzed. Each accession can have membership in different subgroups (admixture model; ALPHAPROPSD=0.20). The number of subgroups (K) in the population was determined by running the program at different K values with K varying from 2 to 6. Three independent runs were assessed for each K value. We used a burn-in period of 100,000 and 1,000,000 replications. No geographic information was used to assist with the clustering.

Trait Variation Between Subpopulations

In order to identify which traits varied significantly between the different subpopulations, all accessions, irrespective of their country of origin, were placed in the subpopulation with which they shared ≥90% of their alleles as determined by the population structure analysis. Lines that displayed evidence of admixture (<90% of the alleles belonging to a single subpopulation) were not included in the analysis. The means for each trait were calculated for each of the subpopulations, and the data were resampled 1,000 times with replacement to determine the 95% confidence intervals of the means. We judged traits as varying significantly among subpopulations if the confidence intervals of the means were non-overlapping.

Acknowledgments We thank KARI, SAARI, ICRISAT and PGRCU-USDA-ARS for providing seed of the finger millet accessions. We also thank Jim Leebens–Mack for assisting with the phylogenetic analysis and critical reading of this manuscript. This research was achieved with funding from USAID, the McKnight Foundation and the Kirkhouse Trust.

References

1. Barbeau WE, Hilu KW (1993) Protein, calcium, iron and amino acid content of selected wild and domesticated cultivars of finger millet. Plant Foods Hum Nutr 43:97–104
2. Bondale KV (1993) Present status of small millets production in India. In: Riley KW, Gupta SC, Seetharam A, Mushonga JN (eds) Advances in small millets. Oxford & IBH Publishing Co. Pvt. Ltd., New Delhi, pp 117–121
3. de Wet JMJ, Prasada Rao KE, Brink DE, Mengesha MH (1984) Systematics and evolution of *Eleusine coracana* (Gramineae). Am J Bot 71:550–557
4. Dida MM, Srinivasachary, Ramakrishnan S, Bennetzen JL, Gale MD, Devos KM (2007) The genetic map of finger millet, *Eleusine coracana*. Theor Appl Genet 114:321–332
5. Esele JP (1989) Cropping systems, production technology, pests and diseases of finger millet in Uganda. In: Segal A, Riley KW, Harinarayana G (eds) Small millets in global agriculture. Oxford & IBH Publishing Co. Pvt. Ltd., New Delhi, pp 293–300
6. FAOSTAT data (2006) FAO Statistical Databases (http://faostat.fao.org)
7. Felsenstein J (1993) PHYLIP (Phylogeny Inference Package) version 3.5c.
8. Gowda BTS, Seetharam A, Viswanath S, Sannegowda S (1986) Incorporation of blast resistance to Indian elite finger millet cultivars from African cultivar IE 1012. SABRAO J 18:119–120
9. Harinarayana G, Seetharam A (1985) Descriptors for finger millet. IBPGR Secretariat, Rome
10. Hilu KW, de Wet JMJ (1976) Racial evolution in *Eleusine coracana* ssp. *coracana* (finger millet). Am J Bot 63:1311–1318
11. Hilu KW, de Wet JMJ, Harlan JR (1979) Archaeobotanical studies of *Eleusine coracana* ssp. *coracana*. Am J Bot 66:330–333
12. Hilu KW, Johnson JL (1992) Ribosomal DNA variation in finger millet and wild species of *Eleusine* (Poaceae). Theor Appl Genet 83:895–902
13. Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128–2129
14. M'Ragwa LRF, Watson CE Jr (1994) Registration of 'KAT/FM-1' finger millet. Crop Sci 34:1690–1691
15. Mushonga JN, Muza FR, Dhiwayo HH (1993) Development, current and future research strategies on finger millet in Zimbabwe. In: Riley KW, Gupta SC, Seetharam A, Mushonga JN (eds) Advances in small millets. Oxford & IBH Publishing Co. Pvt. Ltd., New Delhi, pp 11–19
16. Muza FR, Lee DJ, Andrews DJ, Gupta SC (1995) Mitochondrial DNA variation in finger millet (*Eleusine coracana*). Euphytica 81:199–205
17. National Research Council (1996) Lost crops of Africa, volume I. Grains. National Academy Press, Washington, DC
18. Neves SS, Swire-Clark G, Hilu KW, Baird WV (2005) Phylogeny of *Eleusine* (Poaceae: Chloridoideae) based on nuclear ITS and plastid trnT-trnF sequences. Mol Phylogenet Evol 35:395–419

19. Odelle SE (1993) Improvement of finger millet in Uganda. In: Riley KW, Gupta SC, Seetharam A, Mushonga JN (eds) Advances in small millets. Oxford & IBH Publishing Co. Pvt. Ltd., New Delhi, pp 75–83
20. Prasada Rao KE, de Wet JMJ (1997) Small millets. In: Fuccillo D, Sears L, Stapleton P (eds) Biodiversity in trust. University Press, Cambridge, UK, pp 259–272
21. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959
22. Salimath SS, de Oliveira AC, Godwin ID, Bennetzen JL (1995) Assessment of genome origins and genetic diversity in the genus *Eleusine* with DNA markers. Genome 38:757–763
23. Seetharam A, Patel DP, Halaswamy BH (2006) Small millets. In: Dhillon BS, Saxena S, Agrawal A, Tyagi RK (eds) Plant genetic resources: foodgrain crops. Narosa Publishing House, New Delhi, pp 204–222
24. Tanksley SD, McCouch SR (1997) Seed banks and molecular maps: unlocking genetic potential from the wild. Science 277:1063–1066
25. Tanksley SD, Nelson JC (1996) Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theor Appl Genet 92:191–203
26. Werth CR, Hilu KW, Langner CA (1994) Isozymes of *Eleusine* (Gramineae) and the origin of finger millet. Am J Bot 81:1186–1197