Theor Appl Genet (2000) 101:197–202

C. Grenier · M. Deu · S. Kresovich · P.J. Bramel-Cox
P. Hamon

Assessment of Genetic Diversity in Three Subsets Constituted from the ICRISAT Sorghum Collection Using Random vs Non-Random Sampling Procedures

B. Using molecular markers

Received: 28 October 1999 / Accepted: 17 November 1999

Abstract The large size of the sorghum [*Sorghum bi-color* (L.) Moench] landrace collection maintained by ICRISAT lead to the establishment of a core collection. Thus, three subsets of around 200 accessions were established from: (1) a random sampling after stratification of the entire landrace collection (L), (2) a selective sampling based on quantitative characters (PCS), and (3) a selection based on the geographical origin of landraces and the traits under farmers' selection (T). An assessment was done of the genetic diversity retained by each sampling strategy using the polymorphisms at 15 microsatellite loci. The landraces of each subset were genotyped with three multiplex polymerase chain reactions (PCRs) of five fluorescent primer-pairs each with semi-automated allele sizing. The average allelic richness for each subset was equivalent (16.1, 16.3 and 15.4 alleles per locus for the subsets PCS, L, and T, respectively). The average genetic diversity was also comparable for the three subsets (0.81, 0.77 and 0.80 for the subsets PCS, L, and T, respectively). Allelic frequency distribution for each subset was compared with a chi-square test but few significant differences were observed. A high percentage of rare alleles (71 to 76% of 206 total rare alleles) was maintained in the three subsets. The global molecular diversity retained in each subset was not affected by a sampling procedure based upon phenotypic characters.

Communicated by P.M.A. Tigerstedt

C. Grenier (✉) · M. Deu
CIRAD, TA 70/03, Avenue Agropolis,
F-34398 Montpellier Cedex 5, France
Fax: +33 (0) 4 67 61 55 09
e-mail: grenier@cirad.fr

P.J. Bramel-Cox
ICRISAT-GREP-GRD Patancheru, 502 324 A.P., India

S. Kresovich
Insitute for Genomic Diversity, Cornell University, Ithaca,
NY 14853–2703, USA

P. Hamon
Université Montpellier III/IRD, route de Mende,
34199 Montpellier Cedex 5, France

Key words Core collection · Sorghum · SSRs · Genetic diversity

Introduction

To assist in the management of a large collection, core collections have been defined to target maintenance and evaluation on a subset that retains a large part of the diversity encompassed in the entire collection (Brown 1989a, b). Most core collections have been established from random procedures applied either directly on the initial set or on a stratified initial set. In this latter case, the stratification has often been based upon biological or eco-geographical considerations. The degree of diversity in the core collection was essentially based on phenotypic values (Diwan et al. 1994; Basigalup et al. 1995; Casler 1995; Charmet and Balfourier 1995; Cordeiro et al 1995; Hodgkin 1995; Rincon et al. 1997; Balfourier et al. 1998; Ortiz et al. 1998; Huaman et al. 1999). Recently, some of these core collections went through specific evaluation procedures. For example, the core collection of US perennial alfalfa established by Basigalup et al. (1995) was assessed for acid soil tolerance (Bouton 1996). The results demonstrated that: (1) the diversity of the current core collection was sufficient to initially identify acid soil-tolerant germplasm, and (2) that germplasm with exceptional acid soil tolerance will be difficult to isolate from the alfalfa collection. The same alfalfa core collection was assessed for forage quality variation and it provided a wider range of variation for forage quality traits than present in contemporary cultivars (Jung et al. 1997). A core subset established from a geographically stratified Peruvian quinoa collection and with a proportional sampling was screened for salt tolerance by Ruiz-Tapia et al. in 1997, for details see Ortiz et al. (1999). The most-tolerant accessions were found within a particular geographic origin and showed distinct morphological characteristics. From this evaluation for salt tolerance, the authors found that the core collection could serve as a starting point for further exploitation of

the genetic resources available in the Peruvian quinoa genebank.

For sorghum [*Sorghum bicolor* (L.) Moench], core collections of 10% size have been established with random procedures and assessed for their phenotypic diversity (for details see Grenier et al. 2000). Efforts were then concentrated on three subsets of 1% size in order to assess the genetic diversity retained when random vs non-random sampling strategies are applied.

Comparisons between the entire landrace collection and the three subsets (Grenier et al. 2000) found that these represented the overall phenotypic diversity of the entire landrace collection. Furthermore, they captured a high level of phenotypic diversity as measured by Shannon-Weaver Diversity indices.

However, only one selective sampling did not affect the morpho-agronomic characters that were not previously considered in the sampling. Two hypotheses were formulated in order to explain such results and it was suggested that molecular markers could help in resolving such a situation. Thus, the aims of the present study were: (1) to compare the genetic diversity of the three subsets using microsatellites, (2) to assess the impact of the sampling strategy at the DNA diversity level, and (3) to determine the rate of rare alleles retained by each procedure.

Materials and Methods

The logarithmic subset (L) was sampled from the ICRISAT sorghum landrace collection that had been stratified into four clusters of photoperiod sensitivity. This sampling gives a subset of 225 accessions. The PCS subset is obtained using the principal component scoring strategy (Noirot et al. 1996) applied to a set of 2000 randomly identified accessions. This subset includes 229 accessions. The taxonomic subset (T) has 246 accessions selected on their particular characters of interest and their specific geographic origin (Grenier et al. 2000).

Genomic DNA of each accession was extracted from on average 30 4-day old seedlings using the DNA extraction protocol described in Hoisington (1994). PCR-amplification was carried out at the USDA-ARS Plant Genetic Resources Conservation Unit's laboratory, Griffin, Ga., USA. Multiplex PCR of 15 SSRs loci (see Table 1) was performed with fluorescently labelled primer pairs, as described in Dean et al. (1999). To check for repeatability between gels, each gel run included the amplification products of a single-plant DNA template from the breeding line BTx623. Automated sizing and labelled alleles were determined relative to an internal lane standard using ABI Genescan 672 software version 2.1. Limits of detection for fragments were arbitrarily chosen so that one longer repeat band with a peak height of less than 10% of the highest peak was not accepted. Thus, it eliminated bands that were due to the slippage of the *Taq* polymerase. Additionally, bands with one extra base that resulted from the terminal deoxynucleotidyl transferase activity of *Taq* polymerase were not considered.

The gene diversity index per locus was calculated as $1-\Sigma p_i^2$ where p_i is the frequency of the i^{th} SSR allele (Nei 1973). The effective number of alleles per locus was obtained from the number of alleles whose frequency was higher than 5%. On these frequent alleles, chi-square tests were performed between the three subsets and for the global set.

Results

The rate of homogeneity of the global set (i.e. all three subsets) is variable between the 15 loci (Table 1). For example, 57% of the accessions of the global set are homogeneous at the locus Sb6–42, while 98% of the accessions are homogeneous at the locus Sb5–256. The rates observed are very similar between the three subsets; loci that revealed the highest or the least heterogeneity are the same in the three subsets.

For sorghum, heterogeneity within an accession could reflect heterozygosity at a locus and/or seed mixture. In this study it would be difficult to distinguish between these cases; thus hereafter we will only consider homogeneous accessions. With this condition, a total number

Table 1 Multiplex primer sets used for the amplification of SSR loci in *S. bicolor* and the total number of accessions studied. SSR allele size ranges are those obtained in this study. The percentage of homogeneous accessions is given in parentheses for the global set (GS=the three subsets together) and each subset (principal component score strategy: PCS, logarithmic strategy: L, and taxonomic strategy: T)

Locus	Repeat	Allele size (bp)	GS	PCS	L	T
Multiplex set 1						
Sb4–15	$(AG)_{16}$	118–134	562 (86)	208 (85)	196 (87)	158 (87)
Sb4–121	$(AC)_{14}$	213–229	609 (85)	221 (80)	220 (86)	168 (91)
Sb4–32	$(AG)_{15}$	167–235	575 (82)	201 (80)	219 (79)	155 (87)
Sb5–236	$(AG)_{20}$	158–208	621 (76)	227 (70)	219 (81)	175 (75)
Sb6–342	$(AC)_{25}$	268–294	555 (86)	211 (87)	201 (85)	143 (85)
Multiplex set 2						
Sb6–36	$(AG)_{19}$	162–200	624 (58)	226 (61)	224 (55)	174 (59)
Sb1–1	$(AG)_{16}$	242–300	616 (74)	223 (70)	219 (80)	174 (70)
Sb1–10	$(AG)_{27}$	252–496	594 (75)	220 (69)	209 (81)	165 (77)
Sb5–256	$(AG)_8$	166–182	537 (98)	203 (100)	198 (97)	136 (97)
Sb6–84	$(AG)_{14}$	178–222	611 (80)	219 (74)	223 (87)	169 (77)
Multiplex set 3						
Sb6–42	$(AG)_{26}$	166–212	614 (57)	224 (52)	219 (71)	171 (46)
Sb6–57	$(AG)_{18}$	284–320	607 (72)	223 (68)	217 (84)	167 (61)
Sb6–34	$[(AC)/(CG)]_{15}$	186–208	619 (86)	223 (83)	222 (89)	174 (86)
Sb5–206	$(AC)_{13}/(AG)_{20}$	104–156	605 (76)	219 (78)	212 (85)	174 (63)
Sb4–72	$(AG)_{16}$	182–212	618 (83)	225 (83)	218 (87)	175 (78)

Table 2 Molecular diversity assessed with the allelic richness (No.=total number of alleles) and with the genetic diversity [DI=1–$\Sigma(p_i^2)$, where p_i is the frequency of the i^{th} allele (Weir 1989)] for the global set (GS=the three subsets together) and each subset (principal component score strategy: PCS, logarithmic strategy: L, and taxonomic strategy: T) considering only homogeneous accessions

Locus	GS		PCS		L		T	
	No.	(DI)	No.	(DI)	No.	(DI)	No.	(DI)
Sb4–15	9	(0.80)	8	(0.81)	9	(0.76)	9	(0.77)
Sb4–121	9	(0.84)	9	(0.84)	9	(0.82)	8	(0.84)
Sb4–32	33	(0.91)	21	(0.91)	26	(0.89)	27	(0.91)
Sb5–236	24	(0.86)	19	(0.87)	19	(0.81)	19	(0.88)
Sb6–342	14	(0.79)	14	(0.80)	14	(0.77)	13	(0.79)
Sb6–36	20	(0.85)	17	(0.84)	19	(0.83)	18	(0.86)
Sb1–1	28	(0.86)	22	(0.83)	20	(0.88)	21	(0.85)
Sb1–10	31	(0.82)	26	(0.89)	20	(0.69)	23	(0.82)
Sb5–256	7	(0.27)	5	(0.30)	7	(0.23)	5	(0.28)
Sb6–84	23	(0.86)	21	(0.84)	20	(0.88)	16	(0.82)
Sb6–42	23	(0.93)	22	(0.93)	21	(0.92)	19	(0.90)
Sb6–57	16	(0.81)	12	(0.81)	15	(0.81)	11	(0.80)
Sb6–34	11	(0.74)	11	(0.76)	9	(0.71)	8	(0.74)
Sb5–206	26	(0.93)	25	(0.93)	24	(0.92)	23	(0.93)
Sb4–72	15	(0.71)	9	(0.72)	12	(0.66)	11	(0.75)
Total	289		241		244		231	
Mean	19.3	(0.80)	16.1	(0.81)	16.3	(0.77)	15.4	(0.80)

Table 3 Number of frequent alleles (A) and χ^2 test used to compare the distribution of frequent alleles between the three subsets (principal component score strategy: PCS, logarithmic strategy: L, and taxonomic strategy: T) and to the global set (GS=the three subsets together) considering homogeneous accessions only. One letter (a, b or c) shared by two sets indicates non-significant differences while different letters mean significant differences in allelic distribution between the subset and the global set at $P<0.05$. The number of rare alleles in GS (B) and the number of alleles that were rare in the GS and present in each subset (C) are given

Locus	GS			PCS			L			T		
	(A)	χ^2	(B)	(A)	χ^2	(C)	(A)	χ^2	(C)	(A)	χ^2	(C)
Sb4–15	6	a	3	6	a	2	5	b	3	6	a	3
Sb4–121	7	a	2	7	a	2	7	a	2	8	a	1
Sb4–32	8	a	25	8	a	11	8	b	18	8	a	19
Sb5–236	5	a, b	19	5	a	13	5	b	14	6	a, b	14
Sb6–342	6	a	8	7	a	8	5	a	8	6	a	7
Sb6–36	5	a	15	5	a	12	4	a	14	4	a	12
Sb1–1	5	a, b	23	4	a	17	6	a, b	15	4	b	16
Sb1–10	3	a	28	3	b	23	3	c	17	3	a, b, c	20
Sb5–256	2	a	5	2	a	3	2	a	5	2	a	3
Sb6–84	6	a, b	17	6	a	15	7	b	12	6	a	10
Sb6–42	9	a, b	14	9	a	13	9	a, b	12	7	b	10
Sb6–57	4	a, b	12	6	a, b	8	4	b	11	4	a	7
Sb6–34	4	a	7	4	a	7	4	a	5	4	a	4
Sb5–206	7	a	19	7	a	18	8	a	17	7	a	15
Sb4–72	6	a, b	9	6	a	3	4	a, b	5	5	b	5
Total	83		206	85		157	81		158	80		146
Mean	5.5			5.7			5.4			5.3		

of 289 alleles was recorded for the 15 loci (Table 2). In the global set, the mean allelic richness is 19.3 alleles per locus, and ranges from seven alleles at the locus Sb5–256 to 33 alleles at the locus Sb4–32. The gene diversity ranges between 0.71 and 0.93, except for the locus Sb5–256 that has a gene-diversity value of 0.27.

Comparisons between the three subsets reveal similar tendencies. The mean allelic richness is of the same magnitude (16.3, 16.1 and 15.4 alleles per locus for the L, the PCS and the T subsets, respectively). The genetic diversity is high in all three subsets. It ranges for 14 of the 15 loci from 0.72 to 0.93 (mean of 0.81) in the PCS subset, 0.66 to 0.92 (mean of 0.77) in the L subset, and 0.74 to 0.93 (mean of 0.80) in the T subset. For one locus, Sb5–256, low and comparable genetic diversity values are found (0.23, 0.28 and 0.3 for the L, the T and the PCS subsets, respectively).

When only the frequent alleles are considered, mean allelic richness for the global set is 5.5 alleles per locus, with a range of two (locus Sb5–256) to nine alleles (locus Sb6–42) (Table 3). Once again, for each subset, the mean allelic richness is of same magnitude (5.7, 5.4 and 5.3 for PCS, L and T subsets respectively). The chi-square tests between the three subsets and relative to the global set were performed only on the frequent alleles. When compared to the global set, the T subset does not have a significantly different allelic distribution at any of the 15 loci. The PCS subset differs only at one locus (Sb1–10). The L subset has three loci (Sb4–15, Sb4–32 and Sb1–10) with a significantly different allelic distri-

bution at $P<0.05$; one locus (Sb6–84) has a limit of significance of 5%. Comparisons between subsets show significant differences between the PCS and L subsets for allelic distribution at five loci (Sb4–15, Sb4–32, Sb5–236, Sb1–10 and Sb6–84). Also, the PCS and T subsets have three loci (Sb1–1, Sb6–42 and Sb4–72) with significantly different allelic distributions. The allelic distribution in the L and T subsets is significantly different for four loci (Sb4–15, Sb4–32, Sb6–84, and Sb6–57).

In the global set, 206 alleles have a frequency lower than 5%. From these 206 rare alleles, 157 are present in the PCS subset from which 16 are found specifically in this subset. An equivalent number of rare alleles (158) is maintained in the L subset and 13 of them are specific. In the T subset, 146 of the rare alleles are maintained, and 17 of them are specifically present in this subset. Some of the rare alleles are also specifically lost in the subset: 31 in the taxonomic subset, 19 and 15 in the PCS and logarithmic subsets, respectively.

Discussion

The efficiency of ex situ conservation and its use could be improved with greater genetic information on the germplasm. The study of 15 SSRs markers on three subsets of the sorghum landrace collection held at ICRISAT found that the homogeneity rate varied between 46 and 100% depending on the subset and the locus considered. The origin of heterogeneity in landraces of self-pollinated species needs to be considered for their secure management. Although sorghum is a preferentially autogamous crop, the out-crossing rate can range between 5 and 30% (Ollitrault 1987; Pedersen et al. 1998). As a result, some landraces could still represent traditional populations, although the possibility of seed mixtures either at the harvest or in the seed stocks cannot be excluded. Accession heterogeneity is a key point and needs to be considered in a meaningful way. For genebank management, much care is necessary during the regeneration and evaluation process.

Molecular markers have an important role to play in many aspects of conservation (Karp et al. 1997). For instance they can be used for assessing genetic diversity in a large collection. However, their uses need to be considered in terms of both their cost and the status of the collection. For a well-characterized collection, with as extensive and valuable data set, morpho-agronomic characters can be considered as efficient tools to assess genetic diversity. In this case too, molecular markers can be important tools for fine-resolution questions on plant genetic resource conservation (Kresovich et al. 1993). However, when collections are not fully characterized, or if genetic information provided by the characters is confounded by association or bias, then the molecular markers can be helpful for genetic diversity assessment. In any event, molecular studies should be considered as complementary to phenotypic characterization.

In comparison with other markers, SSRs can be very appropriate tools to assess the genetic diversity in sorghum. Between two to four alleles per polymorphic locus were found with isozymes, RFLP, or RAPD markers (Morden et al. 1989; Ollitrault and Noyer 1989; Deu et al. 1994; de Oliveira et al. 1996; Deu et al. 1999). The 15 SSRs markers developed by Brown et al. (1996) revealed from two to six fragment sizes per polymorphic locus and a diversity index from 0.21 to 0.73 on 17 temperately and tropically adapted lines of sorghum (Brown et al. 1996). Moreover, the 15 microsatellite loci are widely spread on the sorghum genome and 14 of them have been mapped for nine of the ten sorghum linkage groups (Dean et al. 1999). With the same SSR set, Dean et al. (1999) assessed diversity among 95 Orange accessions and found from three to 11 alleles per locus and a genetic diversity ranging from 0.16 to 0.77 (Dean et al. 1999). In our study, using this SSR set, from seven to 33 alleles per locus and a genetic diversity range of 0.71 to 0.93 (one locus at 0.27) were found. These differences between molecular markers in the number of alleles per locus are in concordance with what is known about the mechanisms that generate variability for these different categories of markers, e.g. punctual mutation, insertion/deletion for all the markers, plus polymerase slippage and/or unequal crossing-over for microsatellite markers.

The high values of allelic richness and genetic diversity found in the global set quantify the degree of genetic diversity in the entire landrace collection. These results suggest that the high phenotypic diversity of the sorghum landrace collection is associated with a high level of diversity at the DNA level. It is also possible to use such markers for further genotyping and the elimination of redundancies as reported by Hokanson et al. (1998) on apple, by Chavarriaga-Aguirre et al. (1999) on cassava, and Dean et al. (1999) on the Orange sorghum accessions. The high values for genetic diversity reported here support the hypothesis that in sorghum none these alleles have a selective advantage. For frequent alleles the allelic frequencies are well-distributed in the global set.

Comparisons between the three subsets highlighted the similar genetic diversity retained by each sampling. From the previous comparison of the three subsets with morpho-agronomic data, we found the same global phenotypic diversity, as measured by the Shannon-Wearer Diversity index. However, two hypotheses were considered to explain the discrepancy in morpho-agronomic distribution between the three sampling strategies, i.e. bias introduced during the sampling and bias due to the specific constitution of the subsets (Grenier et al. 2000). Our present results of equal global molecular genetic diversity between the three subsets indicate a similar efficiency of the three strategies to retain genetic diversity. However, allelic frequencies are found significantly different for some loci between the three subsets. Interestingly, the L subset differs from the PCS and T subsets, for the same three loci. These results would support the hypothesis that the phenotypic characters that are influ-

enced by the sampling are not affected because of a bias during the sampling. Indeed, in such a case, a different allelic distribution would have been expected for the three subsets at any of the loci; for two pair-wise comparisons the probability that subsets share three random modified loci out of the 15 is less than 5%. By contrast, our results suggest that phenotypic characters, like molecular markers, could be affected as a result of their associations with the characters used for the constitution of the specific core collections. For example, if the stratification of the entire landrace collection was based on racial or geographic considerations, it would have been expected that the sampling does not affect the same characters as in Grenier's et al. (2000) study, nor the same loci as in this study. Thus, could molecular markers be used to constitute a core collection that would retain the best sample of the agro-morphological diversity? This case could be considered for collections that have not yet been evaluated and for which the cost of field analyses is much higher than for molecular analyses. In the case of large collections, such as the ICRISAT sorghum collection, it is not feasible to assess molecular markers for the entire collection because of the cost. To assess the potential in a limited manner, we considered the global set evaluated in this study as the target set and the five microsatellite loci for which data were the most complete (Sb4–121, Sb1–1, Sb6–84, Sb6–34 and Sb4–72) as the sampling criteria. Then a selective sampling strategy on an initial set of 521 accessions with maximization of the molecular diversity (PCS strategy 'corequal' obtained from htpp://www.mpl.ird.fr/genetrop/rg/pcss/index.htm) was applied in order to establish a subset that represents 10% of the size of the target set. This subset includes 52% of total variability (data not shown). Comparisons of the phenotypic diversity between the subset and the target set based on the same characters as those used by Grenier et al. (2000) found that molecular PCS sampling does not induce changes in the distribution of the morpho-agronomic characters, except for flowering days in the rainy season. In these conditions, it would be interesting to consider an initial set that represents the entire landrace collection as well as others sets of microsatellite loci. Furthermore, any association should be assessed with an investigation of the relationship between molecular markers, the specific morpho-agronomic characteristics and passport data.

Acknowledgements The authors would like to thank Sharon Mitchell, Rob Dean and Mark Hopkins for their valuable assistance provided at the USDA-ARS Plant Genetic Resources Conservation Unit's laboratory, Griffin GA

References

Balfourier F, Charmet G, Prosperi J-M, Goulard M, Monestiez P (1998) Comparison of different spatial strategies for sampling a core collection of natural populations of fodder crops. Genet Sci Evol 30 (Suppl1):S215–S235

Basigalup DH, Barnes DK, Stuker RE (1995) Development of a core collection for perennial *Medicago* plant introduction. Crop Sci 35:1163–1168

Bouton JH (1996) Screening the Alfalfa core collection for acid soil tolerance. Crop Sci 36:198–200

Brown AHD (1989a) The case for core collections. In: Brown AHD, Frankel OH, Marshall DR, Williams JT (eds) The use of plant genetic resources. Cambridge University Press, Cambridge, UK, pp 135–156

Brown AHD (1989b) Core collections: a practical approach to genetic resources management. Genome 31:818–824

Brown SM, Hopkins MS, Mitchell SE, Senior ML, Wang TY, Duncan RR, Gonzales-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

Casler MD (1995) Patterns of variation in a collection of perennial ryegrass accessions. Crop Sci 35: 1169–1177

Charmet G, Balfourier F (1995) The use of geostatistics for sampling a core collection of perennial ryegrass populations. Genet Res Crop Evol 42:303–309

Chavarriaga-Aguirre P, Maya MM, Tohme J, Duque MC, Iglesias C, Bonierbale MW, Kresovich S, Kochert G (1999) Using microsatellites, isozymes and AFLPs to evaluate genetic diversity and redundancy in the cassava core collection and to assess the usefulness of DNA-based markers to maintain germplasm collection. Molecular Breeding 5:263–273

Cordeiro CMT, Morales EAV, Ferreira P, Rocha DMS, Costa IRS, Valois ACC, Silva S (1995) Towards a Brazilian core collection of cassava. In: Hodgkin T, Brown AHD, Hintum ThJL van, Morales EAV (eds) Core collections of plant genetic resources. John Wiley and Sons, Chichester, UK, pp 155–168

Dean RE, Dahlberg JA, Hopkins MS, Mitchell SE, Kresovich S (1999) Genetic redundancy and diversity among 'Orange' accessions in the U.S. national sorghum collection as assessed with simple sequence repeat (SSR) markers. Crop Sci 39:1215–1221

Deu M, Gonzalez-de-Leon D, Glaszman JC, Dégremont I, Chantereau J, Lanaud C, Hamon P (1994) RFLP diversity in cultivated sorghum in relation to racial differenciation. Theor Appl Genet 88:838–844

Deu M, Hamon P, Bonnot F, Chantereau J (1999) Le Sorgho. In: Hamon P (ed) Diversité génétique des plantes tropicales cultivées, CIRAD, Montpellier, France, pp 351–382

Diwan N, Bauchan GR, McIntosh MS (1994) A core collection for the United States annual *Medicago* germplasm collection. Crop Sci 34:279–285

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

Hodgkin T, Brown AHD, Hintum ThJL van, Morales EAV (1995) Future directions. In: Hodgkin T, Brown AHD, Hintum ThJL van, Morales EAV (eds) Core collections of plant genetic resources, John Wiley and Sons, Chichester, UK, pp 253–260

Hoisington D, Khairaliah M, Gonzales-de-Leon D (1994) Laboratory protocols, 2nd edn. CIMMYT Applied Molecular Genetics Laboratory

Hokanson SC, Szewc-McFadden AK, Lamboy WF, McFerson JR (1998) Microsatellites (SSR) markers reveal genetic identities, genetic diversity and relationships in a *Malus*×*domestica* borkh. core subset collection. Theor Appl Genet 97:671–683

Huaman Z, Aguilar C, Ortiz R (1999) Selecting a Peruvian sweetpotato core collection on the basis of morphological, eco-geographical, and disease and pest reaction data. Theor Appl Genet 98:840–844

Jung HG, Sheaffer CC, Barnes DK, Halgerson JL (1997) Forage quality variation in the U.S. alfalfa core collection. Crop Sci 37:1361–1366

Karp A, Kresovich S, Bhat KV, Ayad WG, Hodgkin T (1997) Molecular tools in plant genetic resources conservation: a guide to the technologies. IPGRI Technical Bulletin No. 2

Kresovich S, Lamboy WF, Szewc-McFadden AF, McFerson J, Forsline PL (1993) Molecular diagnostics and plant genetic resources conservation. Ag Biotech News and Information Vol 5 No.7:225N–258N

Morden CW, Doebley, JF, Schertz, KF (1989) Allozyme variation in old world races of *Sorghum bicolor* (Poaceae). Am J Bot 76: 247–255

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl. Acad. Sci. U.S.A. 70:3321–3323

Noirot M, Hamon S, Anthony F (1996) The principal component scoring: a new method of constituting a core collection using quantitative data. Genet Res Crop Evol 43:1–6

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

Ollitrault P (1987) Evaluation génétique des sorghos cultivés (*Sorghum bicolor* L. Moench) par l'analyse conjointe des diversités enzymatique et morphophysiologique; relation avec les sorghos sauvages. Doctorat es sciences, Paris-XII, Orsay

Ollitrault P, Noyer J-L (1989) Polymorphisme enzymatique des sorghos. III. Identification et classification de variétés améliorées de l'IRAT. L'agronomie Tropicale 45(1):59–66

Ortiz R, Ruiz-Tapia,EN, Mujica-Sanchez A (1998) Sampling strategy for a core collection of Peruvian quinoa germplasm. Theor Appl Genet 96:475–483

Ortiz R, Madsen S, Ruiz-Tapia EN, Jacobsen S-E, Mujica-Sanchez A, Christiansen JL, Stolen O (1999) Validating a core collection of Peruvian quinoa germplasm. Genet Res Crop Evol 46:285–290

Pedersen JF, Toy JJ, Johnson B (1998) Natural outcrossing of sorghum and Sudangrass in the central great plains. Crop Sci 38:937–939

Rincon F, Johnson B, Crossa J, Taba S (1997) Identifying subsets of maize accessions by three-mode principal component analysis. Crop Sci 37:1936–1943

Weir BS (1989) Sampling properties of gene diversity. In: Brown AHD, Clegg MT, Kahler AL, Weir BS (eds) Plant population genetics, breeding, and genetic resources. Sinauer Associates Inc. Sunderland, Massachusetts, pp 23–42

THEORETICAL AND APPLIED GENETICS

International Journal of Plant Breeding Research

Volume 101 Number 1–2 July 2000

(Continued on page A 3)