Genetic Resources and Crop Evolution (2006) 53: 515–521
DOI 10.1007/s10722-004-2032-4

Sampling a core collection of Island cotton (*Gossypium barbadense* L.) based on the genotypic values of fiber traits

Haiming Xu[1,☆], Yongjun Mei[2,☆], Jin Hu[1], Jun Zhu[1,*] and Ping Gong[2]
[1]*Department of Agronomy, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou 310029, China;* [2]*College of Plant Science and Technology, Tarimu University of Agricultural Reclamation, Alar, Xinjiang 843300, China; *Author for correspondence (e-mail: jzhu@zju.edu.cn; phone: +86-571-86971731; fax: +86-571-8971498)*

Received 8 March 2004; accepted in revised form 6 August 2004

Key words: Core collection, Fiber traits, Genotypic value, Germplasm, *Gossypium barbadense* L.

Abstract

A genetic model, including effects of environments, genotypes, and genotype by environment interaction, was employed to analyze five fiber traits of Island cotton (*Gossypium barbadense* L.). Genotypic values of 304 accessions were predicted by the adjusted unbiased prediction (AUP). Genetic similarities between different accessions were measured by Mahalanobis distances based on genotypic values. Appropriate sampling strategies, linkage rules in stepwise clustering, and sampling proportion were evaluated. To form a core collection of Island cotton, 60 accessions were sampled by the deviation sampling strategy combined with single linkage rule of hierarchical clustering. The genetic variation and structure captured by the core collection were examined in means, variances, ranges and coefficients of variation, correlation coefficients of quantitative traits, and the accessions distribution plotted by first two principal components between two collections. It was showed that the initial collection was well represented by the core collection for exploiting the Island cotton germplasm.

Introduction

Since the concept of germplasm core collection was first proposed (Frankel and Brown 1984), many core collections have been constructed for different genetic resources, using different sampling strategy and sampling proportion (Brown et al. 1987; Diwan et al. 1994; Bisht et al. 1998; Igartua et al. 1998; Huaman et al. 1999; Zhang et al. 2000; Li et al. 2002; Rodiño et al. 2003; Zewdie et al. 2004). A core collection needs to contain as much diversity as possible, reducing the amount of diversity but increasing the utility of the core (van Hintum 1999). The development of a core collection will alleviate the burden in management of germplasm collection on curators; it will also be simplified for plant breeder to access to the collection for screening exotic materials, when reducing size of the core.

Many kinds of data, such as morphological, agronomic, ecogeographical traits, molecular and biochemical markers, etc., have been used in sampling a core of a germplasm collection. Different kinds of data might have their own properties and efficiencies in measuring genetic diversity in collection. Molecular markers can measure the genetic similarities in DNA sequence among accessions without any influences from environments (Ghislain et al. 1999); however, it is

☆These authors contributed equally to this work.

impractical to genotyping all entries of whole collection because of its expensive cost and cumbersome work involved (van Treuren et al. 2004). Most traits of crop varieties are quantitative traits, controlled not only by genotypes, but also by environments as well as genotype by environment interaction. Phenotypic similarity of accessions cannot correctly reflect their similarities in genetic variation (van Raamsdonk and Wijnker 2000; Upadhyaya et al. 2002), especially, when there are large differences in weather, cultivation and field management in different years. Therefore, biases will be introduced when classifying populations and developing core collections. The genotypic values of traits can be predicted by an adjusted unbiased prediction (AUP) method (Zhu 1993; Zhu and Weir 1996). On the basis of the genotypic values, core collection can be constructed using appropriate cluster methods combining sampling strategy. The core collection based on genotypic values should be more representative in genetic diversity of the initial collection than that of phenotypic data (Hu et al. 2000).

Island cotton (*Gossypium barbadense* L.) is an important slap-up textile material, which has many merits, such as long fiber, excellent fineness, high strength, and other advantages like high photosynthesis efficiency. In the present study, 304 accessions, which cover about 90% of germplasm resources of Island cotton in China, were used to construct core collection. Twelve potential core subsets, based on genotypic values of five fiber traits, were sampled by the stepwise clustering procedure (Hu et al. 2000), resulting in one core collection of the Island cotton tested.

Materials and methods

Predicting genotypic values

A germplasm collection of 304 varieties of Island cotton was planted in the experimental farm of Tarimu University of Agricultural Reclamation, Alar, Xinjiang Province of China in three years (1990–1992). A randomized complete block design with two replications in each year was carried out. Five fiber quality traits (2.5% Span-Length (mm), Uniform (%), Fiber Strength (cn/tex), Elongation (%), and Micronaire) were analyzed. The observed value of the jth accession in the kth block within the ith year could be expressed as:

$$y_{ijk} = \mu + E_i + G_j + GE_{ij} + B_{k(i)} + e_{ijk}$$

where μ is the population mean; E_i is the random effect of the ith environment (year), $E_i \sim (0, \sigma^2_{\mathrm{E}})$, $i = 1, 2, 3$; G_j is the random effect of the jth accession, $G_j \sim (0, \sigma^2_{\mathrm{G}})$, $j = 1, 2, 3, \ldots, 304$; GE_{ij} is the random effect of, $G_j \times E_i$, $GE_{ij} \sim (0, \sigma^2_{\mathrm{GE}})$; $B_{k(i)}$ is the random effect of the kth block within the ith environment, $B_{k(i)} \sim (0, \sigma^2_{\mathrm{B}})$, $k = 1, 2$; e_{ijk} is the residual effect, $e_{ijk} \sim (0, \sigma^2_{\mathrm{e}})$.

Mixed model approach was used for estimating variance components by MINQU(1) method and for predicting genotypic values of each accession by AUP method (Zhu 1993; Zhu and Weir 1996).

Constructing and evaluating core collection

The procedure of stepwise clustering proposed by Hu et al. (2000) was employed in developing a core collection of the Island cotton. Mahalanobis distances (Mahalanobis 1936), based on predicted genotypic values, were used for stepwise clustering. A core collection was determined by following two steps. In the first step, twelve potential core subsets were sampled by 12 diverse combining schemes of three sampling strategies (random sampling, preferred sampling, and deviation sampling) and four linkage rules (single linkage, complete linkage, unweighted pair-group average, and the Ward's methods). An optimal combination was chosen based on the magnitudes of genetic diversity captured by these potential subsets. In the second step, three subsets were sampled by the screened optimal sampling strategy and linkage rule, respectively, at 10, 15 and 20% proportion; an appropriate sampling proportion was specified. On the basis of these results, the core collection consisted of 60 accessions was easily constructed.

For each potential core subset and the initial collection, the mean, variance, range of variation, and coefficient of variation of quantitative trait were calculated; the homogeneity of variances between subset and the entire collection was tested by Levene's test (Levene 1960), and the significant difference in mean was tested by the Newman–Keuls procedure (Newman 1939; Keuls 1952). Other four indices, the mean difference percentage (MD%), the variance difference percentage

(VD%), the coincidence rate (CR%) and the variable rate (VR%) were compared for screening optimal ampling strategy and linkage rule (Hu et al. 2000).

The principal components analysis (PCA) was applied in validation of the core collection. Distribution of the reserve accessions and the core collections was plotted by the first two principal components. The relationship among traits retained by the core was checked by correlation coefficients.

Results

Phenotypic variation was controlled not only by genotypes, but also by environments and GE interaction (Table 1). Significant genotypic variation was detected for fiber length, uniformity and Micronaire, especially for fiber length with genotypic variance being over 50% to the total phenotypic variation. Except of fiber length, other four fiber traits had significant variation in GE interaction. Therefore, genotypic values were rationally selected to measure the genetic similarity among accessions and develop core collection of Island cotton.

In procedure of developing core collection, there are 12 different combinations between four linkage rules (C1 = single linkage, C2 = complete linkage, C3 = UPGMA, and C4 = Ward's method) and three sampling strategies (S1 = random sampling, S2 = preferred sampling, and S3 = deviation sampling). By sampling at 20%, 12 potential subsets were constructed, and denoted as C1S1, C2S1, C3S1, C4S1, C1S2, C2S2, C3S2, C4S2, C1S3, C2S3, C3S3, and C4S3, respectively (Table 2). There was no significant difference ($\alpha = 0.05$) in means between the potential core subsets and the initial collection. The CR% of each subset was larger than 80%, indicating that the range of variation of traits was kept well. Seven subsets had 100% VD% and over 120% VR%, which could be explained by the fact that the genetic variation was significantly increased after eliminating redundant accessions. Under the same sampling method, there was a little difference in four indices among potential cores for different linkage rules. It was noted C1S1 had relative larger VD%, CR% and VR%, as compared with C2S1, C3S1 and C4S1. Similar inclination in magnitudes could also be found in the other subsets sampled by the preferred sampling method or the deviation sampling method. From these results, it could be inferred that single linkage rule was a better choice in stepwise clustering procedure for constructing core collection.

From the deviation sampling (S3), four potential cores (C1S3, C2S3, C3S3, C4S3) all had the largest value in VD%, and also maintained the same original range of variation (> 99% in CR%), and relative larger VR% to the other subsets from

Table 2. Comparison of 12 potential core subsets.

Core collections	MD%	VD%	CR%	VR%
C1S1	0.0	100.0	98.7	132.4
C2S1	0.0	40.0	92.3	119.3
C3S1	0.0	20.0	82.4	114.3
C4S1	0.0	20.0	87.6	118.6
C1S2	0.0	100.0	100.0	136.3
C2S2	0.0	60.0	100.0	121.6
C3S2	0.0	40.0	100.0	119.3
C4S2	0.0	100.0	100.0	125.9
C1S3	0.0	100.0	100.0	144.3
C2S3	0.0	100.0	99.1	133.6
C3S3	0.0	100.0	99.1	129.5
C4S3	0.0	100.0	99.1	134.2

Note. MD% (VD%) is the percentage of significant difference ($p \leq 0.05$) in means (variances) of traits between core subsets and initial collection, CR% is coincidence rate, and VR% is variable rate.

Table 1. Variance analysis of five fiber traits.

Variance	Length	Uniformity	Strength	Elongation	Micronaire
$\hat{\sigma}^2_E/\hat{\sigma}^2_P$	0.029	0.027	0.020	0.432*	0.239
$\hat{\sigma}^2_G/\hat{\sigma}^2_P$	0.522*	0.132*	0.100	0.045	0.160*
$\hat{\sigma}^2_{GE}/\hat{\sigma}^2_P$	0.068	0.161*	0.163*	0.184*	0.270*
$\hat{s}^2_B/\hat{s}^2_P$	0.003	0.000	0.004	0.003	0.004
$\hat{\sigma}^2_E/\hat{\sigma}^2_P$	0.378*	0.681*	0.713*	0.337*	0.327*

*Indicate significance at $p \leq 0.01$; $\hat{\sigma}^2_E, \hat{\sigma}^2_G, \hat{\sigma}^2_{GE}, \hat{\sigma}^2_B, \hat{\sigma}^2_e, \hat{\sigma}^2_P$ are estimated variance components of environment, genotype, block, genotype by environment, residual, and phenotype, respectively.

the same linkage rule. For the preferred sampling (S2), four sampled potential cores (C1S2, C2S2, C3S2, C4S2) had the same CR% as that of the entire collection, due to the accessions with extreme values of traits were preferred in sampling as core accessions. However, C2S2 and C3S2 had lower VD% (60% and 40%) and VR% (121.6% and 119.3%), indicating preferred sampling combined with UPGMA or Ward's method could not capture enough genetic variation from the initial collection. In the case of random sampling (S1), only relative higher VD% was found for C1S1, but lowest CR% and VR% for C2S1, C3S1 and C4S1. Therefore, it was not a good strategy for sampling a core collection. As a result of the preceding analysis, deviation sampling combined with single linkage (C1S3) in stepwise clustering could be the best choice for constructing a core collection of the Island cotton out of 12 combining schemes.

In the present analysis, three sampling proportion was employed and their corresponding subsets were compared. It was showed by the results (Table 3) that the variance and C.V. increased when sampling proportion decreased. C1S3-20 completely captured the range of variation of five traits in entire collection; no significant change in means was detected, indicating the pattern of genetic variation was kept unchanged in the course of sampling at 20% proportion. When the sampling proportion was reduced from 20 to 15% or 10%, the mean of fiber length was significantly changed. Therefore, sampling proportion of 20% was employed in this research.

According to the preceding results, the potential core subset C1S3-20 (Table 3), consisted of 60 core accessions, was regarded as a core collection of the Island cotton; other 244 accessions constituted a reserve collection. In order to validate this core collection, another core subset, sampled directly by complete random strategy at 20% proportion without clustering, was compared in pattern of genetic variation with the C1S3-20 by the principal component analysis (PCA). The distribution of accessions could be approximately presented by the first two principal components, which could account for 73% of the total genetic variation in the initial collection. Two plots (Figure 1) illustrated clearly in visual that many overlapped accessions in central area were very similar in genetic diversity to each other. With regard to core accessions marked by circle, less genetic redundancy could be found in plot A or B. It was also clearly showed that more extreme accessions were included in the core C1S3-20 relative to the core subset from complete random sampling. As a result, much better coverage or pattern of distribution could be found in plot A than in plot B. Therefore, it could be inferred that the genetic structure and variation of the initial Island cotton collection was well represented by the core collection (C1S3-20).

Table 3. Comparison of genetic variation of five fiber traits between the entire collection and subsets at three sampling proportions.

Traits	Collection	Mean	Variance	Range	CV
Length	Initial	34.077	2.638	10.820	0.048
	C1S3-10	32.896*	6.433**	10.820	0.077
	C1S3-15	33.120*	5.462**	10.820	0.070
	C1S3-20	33.607	5.180**	10.820	0.068
Uniformity	Initial	48.277	1.357	6.559	0.024
	C1S3-10	48.478	2.817**	6.551	0.035
	C1S3-15	48.419	2.861**	6.559	0.035
	C1S3-20	48.429	2.465**	6.559	0.032
Strength	Initial	22.578	0.602	6.666	0.034
	C1S3-10	22.234	1.775**	6.666	0.060
	C1S3-15	22.301	1.600**	6.666	0.057
	C1S3-20	22.458	1.530**	6.666	0.055
Elongation	Initial	6.542	0.021	0.951	0.022
	C1S3-10	6.600	0.056**	0.908	0.036
	C1S3-15	6.594	0.051**	0.908	0.034
	C1S3-20	6.577	0.046**	0.951	0.033
Micronaire	Initial	3.688	0.061	1.555	0.067
	C1S3-10	3.811	0.167**	1.555	0.107
	C1S3-15	3.786	0.136**	1.555	0.097
	C1S3-20	3.739	0.121**	1.555	0.093

C1S3-10, C1S3-15, C1S3-20 are subsets sampled by deviation sampling strategy combined with single linkage method at proportion of 10, 15 and 20%, respectively.
* (**) Indicates significant difference in mean (variance) at $p \leq 0.05$ (0.01) between the subsets and the entire collection.

An adequate core collection should maintain genetic associations arising out of co-adapted gene complexes (Ortiz et al. 1998) in entire collection. Using the predicted genotypic values, correlation coefficients among traits were calculated and laid out in Table 4 for the entire collection, the core collection (C1S3-20), and the core subset from complete random strategy without clustering. In general, the core collection (C1S3-20) preserved the genetic correlation observed in the entire collection. Between any pair of traits, very close magnitude of correlation coefficients could be found between the entire collection and C1S3-20. However, in the respect of significance, the

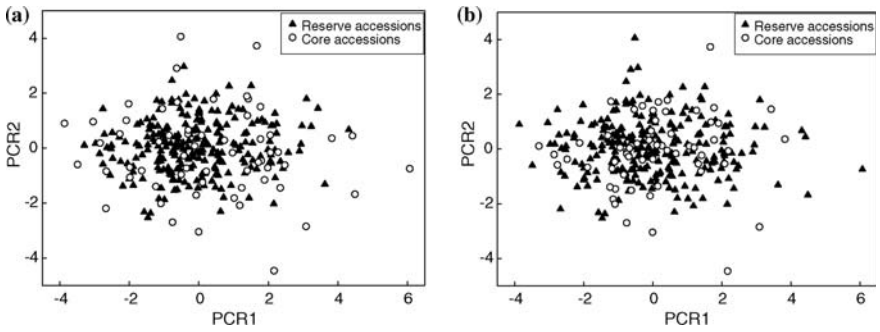

Figure 1. Principal component plots of the reserve and core collections at 20% sampling proportion; (a). Plot for core collection (C1S3-20) constructed by deviation sampling combined with single linkage in stepwise clustering; (b). Plot for core subset sampled by complete random strategy without clustering.

Table 4. Correlation coefficients among traits in the entire collection and the core collection.

	Uniformity	Strength	Elongation	Micronaire
Length	−0.287[a,**]	0.279[**]	−0.335[**]	−0.459[**]
	−0.246[b,*]	0.369[**]	−0.333[**]	−0.489[**]
	−0.159[c]	0.160	−0.189	−0.382[**]
Uniformity		0.162[**]	0.479[**]	0.547[**]
		0.100	0.418[**]	0.552[**]
		0.285[*]	0.540[**]	0.613[**]
Strength			0.125[*]	−0.120[*]
			0.120	−0.225
			0.216	−0.013
Elongation				0.611[**]
				0.585[**]
				0.516[**]

[a]Correlation coefficients for the entire collection.
[b]Correlation coefficients for the core collection (C1S3-20) sampled by deviation strategy combined with single linkage.
[c]Correlation coefficients for the core subset sampled by complete random strategy without clustering.
* (**) Indicate significant correlation at $p \leq 0.05$ (0.01).

significance of association between Uniformity and Strength, Strength and Elongation, Strength and Micronaire, observed in the initial collection, were not detected in the C1S3-20. With regards to the core subset from complete random sampling at 20% proportion, the correlation coefficients among traits are also close to that of the entire collection, but much difference could be seen when compared to that of the C1S3-20 and the entire collection.

Discussion

As proposed by Frankel and Brown (1984), the sampling strategy to obtain a core subset should maximize the diversity while attempting to reduce the redundancy of identical genotypes. Therefore, selection of an appropriate sampling strategy is an important prerequisite to establish core collections of appropriate size. Different sampling strategies have been evaluated and used in constructing practical core collection (Spagnoletti and Qualset 1993; Yonezawa et al. 1995; Ortiz et al. 1998; Jane et al. 2000; Marcos and Abadie 2001; Chandra et al. 2002; Upadhyaya et al. 2003), such as random sampling without replacement, random-systematic by chronology, random-stratified by geographic origin and frequency, proportional, logarithmic, genetic diversity dependent, and so on. In most cases, phenotypic data are used with clusters resulting from a cluster analysis. Hu et al. (2000) proposed a stepwise clustering procedure to construct core collection based on genotypic values, and the properties of three sampling strategies (random, preferred and deviation sampling) combined with three linkage rules (complete linkage, UPGMA, Ward's method) were evaluated. Genetic diversity dependent strategy, regarded as a better strategy for constructing core collection, was proposed by Yonezawa et al. (1995). In essential, the deviation sampling employed in this research is the same as this strategy. The UPGMA was adopted very often in cluster analysis. However, in the present study, the deviation sampling combined with single linkage rule was found to be a good strategy to construct core collection, which could maximize the genetic variation, while, the genetic structure of population was retained by the core collection.

Specifying an appropriate sampling proportion (or core collection size) was another issue in

setting a core collection. On the basis of the sampling theory of selectively neutral alleles, Brown (1989) proposed that a core collection should contain about 10% of the whole collection. Diwan et al. (1995) concluded that a core collection of 10% of the size of the total collection of annual *Medicago* species is too small. Their recommendation is to use a size of 17%. van Raamsdonk and Wijnker (2000) compared representation of three core collections of tulip sampled at 10, 15 and 20% proportion, respectively; 20% was found to be the best proportion in their research. Balakrishnan et al. (2000) proposed a method to specify the size of core collection via fitting a logistic regression model. Therefore, there is no invariable proportion in a core development. Most researchers believe that 10–30% of accessions may be enough to represent 70–90% of the genetic diversity of the whole collection. In the present study, 20% turned out to be an appropriate sampling proportion, and final core collection of the Island cotton was determined at this level.

Acknowledgements

This work was supported in part by the 'The 973 project of National Natural Science Foundation of China', State key Basic Research and Development Plan of China and National Natural Science Foundation of China.

References

Balakrishnan R., Nair N.V. and Sreenivasan T.V. 2000. A method for establishing a core collection of *Saccharum officinarum* L. germplasm based on quantitative-morphological data. Genet. Resour. Crop Evol. 47: 1–9.

Bisht I.S., Mahajan R.K. and Patel D.P. 1998. The use of characterisation data to establish the Indian mungbean core collection and assessment of genetic diversity. Genet. Resour. Crop Evol. 45: 127–133.

Brown A.H.D. 1989. Core collection: a practical approach to genetic resources management. Genome 31: 818–824.

Brown A.H.D., Grace J.P. and Speer S.S. 1987. Designation of a core collection of perennial *Glycine*. Soybean Genet. Newslett. 14: 59–70.

Chandra S., Huaman Z., Hari Krishna S. and Ortiz R. 2002. Optimal sampling strategy and core collection size of Andean tetraploid potato based on isozyme data – a simulation study. Theor. Appl. Genet. 104: 1325–1334.

Diwan N., Bauchan G.R. and McIntosh M.S.A. 1994. Core collection for the United States annual *Medicago* germplasm collection. Crop Sci. 34: 279–285.

Diwan N., McIntosh M.S. and Bauchan G.R. 1995. Methods of developing a core collection of annual Medicago species. Theor. Appl. Genet. 90: 755–761.

Frankel O.H. and Brown A.H.D. 1984. Current plant genetic resources – a critical appraisal. In: Chopra V.L., Joshi B.C., Sharma R.P. and Bansal H.C. (eds), Genetics: New Frontiers, Vol. IV. Oxford & IBH Publ. Co., New Delhi, pp. 1–13.

Ghislain M., Zhang D., Fajardo D., Huamán Z. and Hijmans R.J. 1999. Marker-assisted sampling of the cultivated Andean potato Solanum phureja collection using RAPD markers. Genet. Resour. Crop Evol. 46: 547–555.

Hu J., Zhu J. and Xu H.M. 2000. Methods of constructing core collections by stepwise clustering with three sampling strategies based on the genotypic values of crops. Theor. Appl. Genet. 101: 264–268.

Huaman Z., Aguilar C. and 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.

Igartua E., Gracia M.P., Lasa J.M., Medina B., Molina-Cano J.L., Montoya J.L. and Romagosa I. 1998. The Spanish barley core collection. Genet. Resour. Crop Evol. 45: 475–481.

Jane M.M., Rodriguez J.M. and Nienhuis J. 2000. Development of an algorithm identifying maximally diverse core collections. Genet. Resour. Crop Evol. 47: 515–526.

Keuls M. 1952. The use of the "Studentized range" in connection with an analysis of variance. Euphytica 1: 112–122.

Levene H. 1960. Robust tests for equality of variances. In: Olkin I. (ed.), Contributions to Probability and Statistics: Essay in Honour of Harold Hotelling. Stanford University Press, Stanford, pp. 278–292.

Li Z., Zhang H., Zeng Y., Yang Z., Shen S., Sun C. and Wang X. 2002. Studies on sampling schemes for the establishment of core collection of rice landraces in Yunnan China. Genet. Resour. Crop Evol. 49: 67–74.

Mahalanobis P.C. 1936. On the generalized distance in statistics. Proc. Natl. Inst. Sci. India 2: 49–55.

Marcos M. and Abadie T. 2001. Sampling strategy to develop a core collection of Uruguayan maize landraces based on morphological traits. Genet. Resour. Crop Evol. 48: 381–390.

Newman D. 1939. The distribution of range in samples from a normal population expressed in terms of an independent estimate of standard deviation. Biometrika 31: 20–30.

Ortiz H., Ruiz-Tapia E.N. and Mujica-Sanchez A. 1998. Sampling strategy for a core collection of Peruvian qulnoa germplasm. Theor. Appl. Genet. 96: 475–483.

Rodiño A.P., Santalla M., De Ron A.M. and Singh S.P. 2003. A core collection of common bean from the Iberian peninsula. Euphytica 131: 165–175.

Spagnoletti Zeuli P.L. and Qualset C.Q. 1993. Evaluation of five strategies for obtaining a core subset from a large genetic resource collection of durum wheat. Theor. Appl. Genet. 87: 295–304.

Upadhyaya H.D., Ortiz R., Bramel P.J. and Singh S. 2002. Phenotypic diversity for morphological and agronomic characteristics in chickpea core collection. Euphytica 123: 333–342.

Upadhyaya H.D., Ortiz R., Bramel P.J. and Singh S. 2003. Development of a groundnut core collection using taxonomical, geographical and morphological descriptors. Genet. Resour. Crop Evol. 50: 139–148.

van Hintum T.J.L. 1999. The general methodology for creating a core collection. In: Johnson R.C. and Hodgkin T. (eds), Core Collections for Today and Tomorrow. International Plant Genetic Resources Institute, Rome, Italy, pp. 10–17.

van Raamsdonk L.W.D. and Wijnker J. 2000. The development of a new approach for establishing a core collection using multivariate analyses with tulip as case. Genet. Resour. Crop Evol. 47: 403–416.

van Treuren R., Magda A., Hoekstra R. and van Hintum Th.J.L. 2004. Genetic and economic aspects of marker-assisted reduction of redundancy from a wild potato germplasm collection. Genet. Resour. Crop Evol. 51: 277–290.

Yonezawa K., Nomura T. and Morishima H. 1995. Sampling strategies for use in stratified germplasm collections. In: Hodgkin T., Brown A.H.D., van Hintum Th.J.L. and Morales E.A.V. (eds), Core Collections of Plant Genetic Resources. John Wiley and Sons, Chichester, UK, pp. 35–53.

Zewdie Y., Tong N. and Bosland P. 2004. Establishing a core collection of Capsicum using a cluster analysis with enlightened selection of accessions. Genet. Resour. Crop Evol. 51: 147–151.

Zhang X.R., Zhao Y.Z., Cheng Y., Feng X.Y., Guo Q.Y., Zhou M.D. and Toby H. 2000. Genet. Resour. Crop Evol. 47: 273–279.

Zhu J. 1993. Methods of prediction genotype value and heterosis for offspring of hybrids. J. Biomath. 8: 32–44.

Zhu J. and Weir B.S. 1996. Diallel analysis for sex-linked and maternal effects. Theor. Appl. Genet. 92(1): 1–9.