PLANT GENETIC RESOURCES

Developing a Mini Core of Peanut for Utilization of Genetic Resources

Hari D. Upadhyaya,* Paula J. Bramel, Rodomiro Ortiz, and Sube Singh

ABSTRACT

Peanut (*Arachis hypogaea* L.) breeding programs with a goal of rapid cultivar development have used mainly elite breeding lines and cultivars, which has resulted in the development of breeding materials with a narrow genetic base. Utilization of exotic germplasm resources in breeding programs is needed to enhance the diversity of cultivars. Scientific plant breeding and its need for large variability, concern about potential loss of variability, and nonavailability of low cost tools to identify similarities or differences among accessions led genebanks to hold large germplasm collections. Core collections, generally contain about 10% of total accessions, represent the genetic variability of entire germplasm collection, and have been suggested as a way to enhance use of genetic resources in crop improvement. The objective of this study was to develop a peanut mini core subset. The peanut core subset was evaluated for morphological, agronomic, and quality traits in the rainy and postrainy seasons. Ward's method of clustering was used to separate core collection accessions into groups of similar accessions. A mini core subset consisting of 184 accessions was selected. Newman Keuls' test for means, Levene's test for variances, and chi-square test for frequency distribution analysis for different traits indicated that the variation available in the core collection has been preserved in the mini core subset. This mini core subset will enhance exploitation of peanut genetic resources.

Peanut improvement has made significant progress in the last two decades, resulting in enhanced productivity worldwide. Productivity increased at the rate 14.7 kg ha^{-1} yr^{-1} in the USA (Mozingo et al., 1987). Genetic gains were estimated between 1.3 to 3.2% yr^{-1} under rainfed cultivation in India (Nigam et al., 1994). However, little of the large genetic variability contained in the germplasm accessions has been utilized in crop improvement programs. The International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) holds in trust in excess of 14 000 accessions of cultivated peanut but very few have been used in the improvement program (Upadhyaya et al., 2002), suggesting that most peanut cultivars have a very narrow genetic base. Jiang and Duan (1998) reviewing the utilization of peanut genetic resources in genetic improvement in China, concluded that introduced foreign germplasm and wild relatives have seldom been utilized in cultivar development. In the USA, the cultivar Dixie Giant was a germplasm source in all pedigrees of runner market-type peanuts and Small White Spanish-1 cultivar in 90% or more pedigrees. The two lines contributed nearly 50% of the germplasm of runner cultivars (Knauft and Gorbet, 1989).

There are numerous examples where plant breeders have effectively exploited the exotic germplasm by introgressing gene(s) for disease resistance or single genes controlling other traits (Stalker, 1980). The use of exotic germplasm in improvement of quantitative traits is conspicuously rare, although large proportions of breeding efforts in different breeding programs are directed towards improving such traits. There are many reasons for the low use of diverse germplasm for improvement of the quantitative traits in the adapted germplasm pool. Foremost among these, is the supposition that these germplasm lines have little to offer the improvement of elite cultivars, or that it would require such an extended effort to exploit that the investment of time and resources are not justified (Goodman, 1985; Hallauer, 1978). Improvement programs aimed at short-term rapid cultivar development rely mostly on established cultivars and elite breeding lines in developing breeding materials, rather than long-term germplasm development using exotic germplasm (Halward and Wynne, 1991). The large variability in the germplasm in genebanks rather than prompting greater utilization, creates the problem of not knowing what germplasm to use to begin the genetic enhancement of the crop breeding pool(s). This situation has arisen because of incomplete knowledge of germplasm accessions, the relationships among them, unavailability of descriptive characters, and uncertainty about the best evaluation methods for tapping germplasm resources. Core collections are becoming important tools to overcome this situation and enhance utilization of genetic resources in crop improvement programs.

A core collection is a subset of accessions from the entire collection that capture most of available genetic diversity of the species (Brown, 1989a). A core subset can be extensively evaluated and the information derived be used to plan a more efficient utilization of the entire collection (Tohme et al., 1995; Brown, 1989b). Upadhyaya et al. (2002) developed a peanut core subset of 1704 accessions, on the basis of taxonomic, geographic, and morphologic descriptors of 14 310 peanut accessions from the ICRISAT genebank. This core subset includes 11.9% of accessions and represents the genetic variation available in the entire collection. This core also preserves the coadapted gene complexes represented in the entire collection (Upadhyaya et al., 2002). Brown (1989a) using sampling theory of selectively neu-

Genetic Resources and Enhancement Program, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), P.O. Patancheru 502 324, Andhra Pradesh, India. Received 27 Aug. 2001. *Corresponding author (H.Upadhyaya@CGIAR.ORG).

Published in Crop Sci. 42:2150–2156 (2002).

tral alleles argued that the size of a core subset should be about 10% of total accessions, with a ceiling of 3000 per species. This sample size seems to be effective in retaining 70% of alleles of the entire collection. Holbrook et al. (1993) developed a peanut core collection of 831 accessions on the basis of six morphological variables (plant type, pod type, seed size, testa color, seed per pod, and seed weight) of 7432 U.S. peanut accessions. The core subset of peanut developed at ICRISAT contains 1704 accessions, which is too large for assessing quantitative traits of economic importance. These traits display genotype × environment interaction, and multilocation testing should be considered to identify useful parents. The problem is how to reduce the size of a core subset further without losing the spectrum of species diversity. Upadhyaya and Ortiz (2001) suggested a strategy for sampling the entire and core collections for developing a mini core subset which contains about 1% of total accessions in the entire collection but captures most of the useful variation of the crop. The objective of this research was to develop a mini core subset of peanut.

MATERIAL AND METHODS

The ICRISAT genebank holds in trust a peanut collection of 14 310 accessions from 92 countries. The botanical variety *hypogaea* is represented by 6622 (46.3%) accessions followed by *vulgaris* with 5106 (35.7%) accessions, and *fastigiata* with 2298 (16.1%) accessions. The experimental materials for this study consisted of 584 accessions belonging to variety *vulgaris*, 299 to *fastigiata*, 27 to *peruviana*, six *aequatoriana*, 784 *hypogaea*, and four *hirsuta* (Upadhyaya et al., 2002). These 1704 entries were planted in the alfisol-Patancheru Soil Series (Udic Rhodustolf) fields in the 1999 rainy and 1999-2000 postrainy seasons at Patancheru, Andhra Pradesh, India. Each plot consisted of a single 4-m row on a ridge. The distance between rows was 600 mm and between plants 100 mm. Care was taken to ensure uniform planting depth of 30 mm. Seeds were treated with ethrel (2-chloroethylphosphonic acid) before sowing to overcome the possible effects of postharvest seed dormancy of genotypes belonging to *hypogaea* and *hirsuta* botanical varieties.

The experiments received 60 kg P_2O_5, 400 kg gypsum ha^{-1}, full irrigation (12 irrigations in the postrainy and six in the rainy season, each irrigation 50 mm of water) and protection against diseases, insects, and weeds. In each accession, five competitive plants were selected randomly to record observations on number of primary branches, plant height (mm), leaflet length and width (mm), number of pods per plant, length and width of 10 mature pods (mm), number of seeds per pod, length and width of 10 mature seeds (mm), yield per plant (g), shelling percentage, and 100-seed weight (g). Data on morphological descriptors, growth habit, branching pattern, stem color, stem hairs, leaf color, leaf hairs, flower (standard petal) color, streak (markings on standard petal) color, and peg color were recorded per plot, and on pod beak, pod reticulation, pod constriction, and primary seed color on the five selected plants (IBPGR and ICRISAT, 1992). Days to emergence (days from sowing to the stage when 50% of seedlings have emerged), days to 50% flowering (days from sowing to the stage when 50% of plants have begun flowering), and plot yield were also recorded by plot. The yield of the five selected plants was added to the plot yield to determine the total plot yield.

An equal weight of sound mature seeds from each of the selected plants was bulked to determine the oil and protein contents in the 1999-2000 postrainy season. Oil content was measured with a nuclear magnetic resonance spectrometer following the procedure described by Jambunathan et al. (1985). All readings were taken on oven dried (110°C, 16 h) samples and data were corrected to a uniform 50 g kg^{-1} seed moisture content. Protein content was estimated with a Technicon Autoanalyser (Pulse Instrumentation Ltd., Saskatoon, SK) (Singh and Jambunathan, 1980).

A phenotypic distance matrix was created by calculating differences between each pair of accessions for each of the 47 traits. The diversity index was calculated by averaging all the differences in the phenotypic values for each trait divided by the respective range (Johns et al., 1997). The distance matrix was subjected to the hierarchical cluster algorithm of Ward (1963) at an R^2 (squared multiple correlation value) of 0.75 by means of SAS (1989). This method optimizes an objective function because it minimizes the sum of squares within groups and maximizes the sum of squares between groups. The proportional sampling strategy was used, and from each cluster approximately 10% of the accessions were randomly selected for the mini core subset. At least one accession was included from each cluster even if they had 10 accessions or less.

The means of the core subset and the mini core subset were compared by Newman-Keuls procedure (Newman, 1939; Keuls, 1952). The homogeneity of variances between the core and mini core subsets was tested by Levene's test (Levene, 1960). The percentage of significant differences between the core collection and mini core subset was calculated for the mean difference percentage (MD%) or the variance difference percentage (VD%) (Hu et al., 2000). The coincidence rate (CR%) and the variable rate (VR%) were calculated to evaluate properties of the mini core subset (Hu et al., 2000). The distribution homogeneity for each of the 13 morphological descriptor traits and composition of mini core collection in terms of accessions in each botanical variety and region were analyzed using the chi-square test. The medians of the 40 traits of the core collection and mini core subset were compared using SAS NPAR1WAY procedure (SAS, 1989). The phenotypic correlations among different traits in the core and mini core were estimated independently, to determine whether these associations, which may be under the same genetic control, were conserved in the mini core subset. The diversity index (H′) of Shannon and Weaver (1949) was used as a measure of phenotypic diversity of each trait. The index was calculated independently in both the core collection and the

Table 1. Number and percentage (within brackets) of accessions in the entire collection, core subset, and mini core subset and chi square values, and probability for number of accessions according to botanical varieties in the mini core subset of peanut at International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

Botanical varieties	Entire collection	Core subset	Mini core subset
fastigiata	2298 (16.1%)	299 (17.6%)	37 (20.1%)
vulgaris	5106 (35.7%)	584 (34.3%)	58 (31.5%)
aequatoriana	15 (0.1%)	6 (0.4%)	1 (0.5%)
peruviana	249 (1.7%)	27 (1.6%)	2 (1.1%)
hypogaea	6622 (46.3%)	784 (46.0%)	85 (46.2%)
hirsuta	20 (0.1%)	4 (0.2%)	1 (0.5%)
χ^2	3.643	1.199	
Probability	0.602	0.945	

Table 2. Means and variances for six morphological descriptors recorded in the core and mini core subsets of peanut at International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) in the 1999 rainy season.

Morphological descriptor	Means†			Variances†			
	Core subset	Mini core subset	Significance‡	Core subset	Mini core subset	*F* value	*P*
Stem hairs	3.50	3.50	NS	0.96	0.97	0.15	0.704
Leaf color	2.48	2.49	NS	0.27	0.27	0.003	0.954
Leaf hairs	1.02	1.05	NS	0.12	0.28	0.91	0.339
Pod beak	4.07	4.04	NS	2.18	2.40	0.72	0.397
Pod constriction	3.74	3.75	NS	1.28	1.21	0.24	0.627
Pod reticulation	5.05	5.08	NS	1.44	1.59	0.79	0.375

† Differences between means of core and mini core subsets were tested by Newman-Keuls' test and variance homogeneity by Levene's test.
‡ NS indicates non-significant differences.

mini core subset to determine whether the diversity for each trait was retained in the mini core subset.

RESULTS AND DISCUSSION

The 1704 accessions included in the core subset were grouped into 77 clusters. The number of core accessions in individual clusters ranged from 1 (0.06%) to 76 (4.46%). The number of clusters with 1 to 10 accessions was 20, with 11 to 20 accessions 25, 21 to 30 accessions 15, 31 to 40 accessions 8, 41 to 50 accessions 1, 51 to 60 accessions 5, 61 to 70 accessions 1, and 71 to 76 accessions 2. Following the procedure used to develop the mini core subset, 184 accessions were selected from the core subset.

All six botanical varieties were represented in the mini core subset. The number of entries included in the mini core was 37 *fastigiata* (20.1%), 58 *vulgaris* (31.5%), 85 *hypogaea* (46.2%), two *peruviana* (1.1%), and one each of *aequatoriana*, and *hirsuta* (0.5%). This corresponded very well with the number of *fastigiata* (2298 accessions or 16.1%), *vulgaris* (5106 or 35.7%), *hypogaea* (6622 or 46.3%), and *peruviana* (249 or 1.7%) in the entire

Table 3. Means and variances for agronomic and quality traits recorded in the rainy and postrainy seasons in the core and mini core subsets of peanut at International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

Agronomic and quality traits	Means†			Variances†			
	Core subset	Mini core subset	Significance‡	Core subset	Mini core subset	*F* value	*P*
Rainy season							
Days to emergence	8.3	8.2	NS	1.13	1.09	0.015	0.903
Days to 50% flowering	23.3	23.4	NS	15.62	16.47	0.259	0.611
Number of primary branches	4.8	4.8	NS	0.73	0.98	7.626	0.006
Plant height (mm)	310.2	307.8	NS	5 131.6	5 596.3	0.055	0.459
Leaflet length (mm)	50.3	49.7	NS	41.78	46.36	0.746	0.388
Leaflet width (mm)	21.9	21.9	NS	7.79	7.84	0.018	0.893
Number of seeds per pod	1.6	1.7	NS	0.91	0.98	0.922	0.337
Pod length (mm)	28.1	28.3	NS	31.00	31.81	0.179	0.672
Pod width (mm)	11.9	12.0	NS	1.84	2.40	6.803	0.009
Number of pods per plant	8.8	8.9	NS	13.53	15.75	1.912	0.167
Seed length (mm)	12.1	12.2	NS	4.30	4.54	0.443	0.506
Seed width (mm)	7.0	7.1	*	0.40	0.53	7.127	0.008
100-seed weight (g)	30.1	30.5	NS	36.29	36.28	0.080	0.778
Yield per plant (g)	5.4	5.5	NS	6.59	7.92	2.192	0.139
Yield per plot (kg ha^{-1})	605.2	597.8	NS	73 856.65	92 510.79	3.806	0.051
Shelling percentage	63.4	63.5	NS	35.20	34.97	0.001	0.981
Postrainy season							
Days to emergence	10.5	10.5	NS	0.98	0.94	0.042	0.837
Days to 50% flowering	41.4	41.4	NS	38.36	40.79	1.143	0.281
Number of primary branches	4.5	4.5	NS	0.78	0.94	2.689	0.101
Plant height (mm)	130.2	129.4	NS	1 383.5	1 297.9	0.220	0.637
Leaflet length (mm)	43.4	42.8	NS	57.23	54.22	0.217	0.641
Leaflet width (mm)	19.5	19.6	NS	9.38	17.17	4.032	0.045
Number of seeds per pod	1.6	1.7	NS	0.90	0.96	0.586	0.444
Pod length (mm)	29.7	29.7	NS	27.65	30.25	0.784	0.376
Pod width (mm)	12.7	12.7	NS	2.08	2.54	2.669	0.103
Number of pods per plant	12.8	13.3	NS	23.90	26.34	0.416	0.519
Seed length (mm)	13.9	13.8	NS	6.32	6.27	0.0001	0.994
Seed width (mm)	8.7	8.6	NS	0.99	1.02	0.048	0.826
100-seed weight (g)	49.0	48.5	NS	89.67	88.51	0.0002	0.990
Yield per plant (g)	12.1	12.1	NS	26.04	27.17	0.070	0.792
Yield per plot (kg ha^{-1})	1 176.9	1 171.2	NS	154 737.92	183 820.23	3.282	0.070
Shelling percentage	70.9	70.8	NS	25.34	30.68	1.294	0.255
Oil content (%)	49.4	49.3	NS	4.87	4.91	0.035	0.851
Protein content (%)	22.8	22.9	NS	5.36	5.71	0.496	0.481

† Differences between means of core and mini core subsets were tested by Newman-Keuls' test and variance homogeneity by Levene's test.
‡ NS and * indicate nonsignificant and significant differences at $P = 0.05$, respectively.

collection (Table 1). The *aequatoriana* and *hirsuta*, which have only 15 and 20 accessions, respectively in the entire collection, were over represented in the mini core subset because they were also in excess in the core subset itself (Upadhyaya et al., 2002). The chi square test indicated that the mini core collection represented the distribution of botanical varieties in the core collection ($\chi^2 = 1.199$, $P = 0.945$) and the entire collection ($\chi^2 = 3.643$, $P = 0.602$)

Composition of the mini core subset reflected the predominance of lines from Asia, Africa, and America in the core subset and entire collection. In the mini core subset, the number of accessions included were 60 (32.6%) from Asia, 61 (33.2%) from America, 43 (23.4%) from Africa, and 3 (1.6%) from Europe. Chi square tests indicated that the mini core collection represented the distribution of regions in the core collection ($\chi^2 = 5.021$, $P = 0.413$) and the entire collection ($\chi^2 = 8.589$, $P = 0.127$). South America, which is where the primary center of diversity (Gregory and Gregory, 1976) and seven secondary centers of diversity are located, contributed 29 (15.8%) accessions to the mini core subset, which compared favorably with the number of accessions from the continent in the entire collection (2142 or 15.0%). India, and China which are considered important centers of diversity because of their long history of peanut cultivation were adequately sampled in the mini core subset with 41 accessions (22.3%) from India and 7 accessions (3.8%) from China.

Differences between means of the core and mini core subsets were found to be nonsignificant for the morphological traits (Table 2) and for 15 agronomic traits in the 1999 rainy season, and 18 agronomic and quality traits in the 1999-2000 postrainy season (Table 3). The 2.12% value for MD% indicated that the mini core subset represented adequately the core collection (Hu et al., 2000). The variances of the core and mini core subsets were homogeneous for the morphological traits (Table 2), 13 agronomic traits in the rainy season, and 17 agronomic and quality traits in the postrainy season

Table 4. Median, range, and coefficient of variation for 40 traits in the core and mini core subsets of peanut at International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

	Median			Range		Coefficient of variation (%)	
	Core subset	Mini core subset	*P*	Core subset	Mini core subset	Core subset	Mini core subset
Morphological descriptor							
Stem hairs	3.0	3.0	0.976	1.0–9.0	1.0–9.0	27.25	28.18
Leaf color	2.0	2.0	0.797	2.0–4.0	2.0–4.0	20.99	21.00
Leaf hairs	1.0	1.0	0.180	1.0–8.0	1.0–8.0	32.21	50.29
Pod beak	5.0	5.0	0.739	1.0–9.0	1.0–9.0	36.30	38.33
Pod constriction	3.0	3.0	0.902	1.0–7.0	1.0–7.0	30.03	29.27
Pod reticulation	5.0	5.0	0.756	1.0–9.0	1.0–9.0	23.23	24.84
Agronomic and quality traits							
Rainy season							
Days to emergence	8.0	8.0	0.100	6.0–12.0	6.0–11.0	12.38	12.68
Days to 50% flowering	22.0	22.0	0.999	15.0–31.0	15.0–30.0	17.04	17.33
Number of primary branches	5.0	5.0	0.986	3.0–8.0	3.0–8.0	18.03	20.45
Plant height (mm)	310.0	320.0	0.756	66.0–634.0	80.0–520.0	23.45	24.30
Leaflet length (mm)	50.2	49.0	0.030	24.8–75.4	33.0–73.0	12.85	13.71
Leaflet width (mm)	21.6	21.0	0.876	7.0–30.8	15.0–30.8	12.66	12.81
Number of seeds per pod	1.0	1.0	0.680	1.0–5.0	1.0–5.0	57.36	58.82
Pod length (mm)	28.0	28.0	0.352	12.6–55.7	13.0–48.0	19.51	19.91
Pod width (mm)	11.9	12.0	0.013	5.7–17.5	6.0–16.0	11.31	12.90
Number of pods per plant	8.0	8.0	0.788	1.0–28.0	1.0–24.0	41.30	44.63
Seed length (mm)	12.0	12.0	0.375	7.2–19.5	8.0–18.0	17.02	17.54
Seed width (mm)	7.0	7.0	0.156	4.8–9.5	5.0–9.5	8.97	10.25
100-seed weight (g)	30.0	30.0	0.683	16.0–68.0	18.0–52.0	19.62	19.76
Yield per plant (g)	5.2	5.0	0.530	0.4–18.0	0.4–14.0	47.06	51.13
Yield per plot (kg ha^{-1})	587.5	556.0	0.641	29.2–1875.0	33.3–1875.0	44.47	50.88
Shelling percentage	63.9	64.0	0.494	40.0–78.7	40.0–75.0	9.29	9.31
Postrainy season							
Days to emergence	11.0	11.0	0.776	9.0–13.0	9.0–13.0	9.34	9.26
Days to 50% flowering	38.0	38.0	0.513	29.0–56.0	29.0–54.0	14.83	15.42
Number of primary branches	4.2	4.0	0.0002	2.2–8.4	3.0–8.0	19.34	21.38
Plant height (mm)	128.0	130.0	0.583	36.0–299.0	70.0–252.0	28.36	27.84
Leaflet length (mm)	43.0	43.0	0.569	23.4–75.0	25.0–65.0	17.34	17.22
Leaflet width (mm)	19.0	19.0	0.533	8.6–58.8	12.0–58.8	15.48	21.19
Number of seeds per pod	1.0	1.0	0.765	1.0–5.0	1.0–5.0	57.30	58.57
Pod length (mm)	29.0	29.0	0.956	14.0–58.0	14.0–51.0	17.45	18.50
Pod width (mm)	13.0	12.0	0.574	7.0–19.0	7.0–17.0	11.35	12.59
Number of pods per plant	12.4	13.0	0.876	2.0–34.8	2.0–32.0	38.26	38.56
Seed length (mm)	14.0	14.0	0.880	9.0–22.0	9.0–20.0	18.00	18.20
Seed width (mm)	9.0	9.0	0.428	6.0–13.0	7.0–12.0	11.46	11.69
100-seed weight (g)	47.8	48.0	0.756	25.0–80.0	28.0–78.0	19.18	19.39
Yield per plant (g)	11.3	11.0	0.877	0.38–37.88	1.0–31.0	42.21	42.93
Yield per plot (kg ha^{-1})	1163.6	1146.5	0.756	65.0–2697.9	153.0–2697.9	33.13	36.61
Shelling percentage	72.1	72.0	0.212	37.6–83.5	42.0–82.0	7.00	7.82
Oil content (%)	49.7	50.0	0.436	41.0–55.3	43.0–55.3	4.43	4.49
Protein content (%)	22.9	23.0	0.043	15.5–29.1	17.0–29.1	10.09	10.42

Table 5. Comparison of frequency distribution for morphological descriptors in the core and mini core subsets of peanut at International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

Morphological descriptor	Number of classes	χ^2	Probability
Growth habit	5	0.616	0.533
Branching pattern	3	0.025	0.302
Stem color	2	0.234	0.629
Stem hairs	5	1.949	0.745
Leaf color	3	0.073	0.964
Leaf hairs	4	3.007	0.391
Flower color	3	4.045	0.061
Streak color	3	3.332	0.189
Peg color	2	0.485	0.486
Pod beak	5	2.821	0.588
Pod constriction	4	0.353	0.950
Pod reticulation	5	2.088	0.720
Primary seed color	15	15.671	0.334

Table 7. Shannon-Weaver diversity index for 13 morphological descriptor traits in the core and mini core subsets of peanut at International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

Morphological descriptor	Core subset	Mini core subset
Growth habit	0.533	0.518
Branching pattern	0.302	0.313
Stem color	0.195	0.204
Stem hairs	0.273	0.275
Leaf color	0.322	0.323
Leaf hairs	0.028	0.051
Flower color	0.061	0.078
Streak color	0.120	0.155
Peg color	0.050	0.062
Pod beak	0.483	0.501
Pod constriction	0.360	0.349
Pod reticulation	0.362	0.396
Primary seed color	0.452	0.498
Average ± s.e.	0.273 ± 0.045	0.286 ± 0.045

(Table 3). The variances for number of primary branches, pod width, and seed width in the rainy season and leaflet width in the postrainy season were significantly greater in the mini core subset than in the core subset (Table 3). The variances and the coefficients of variation in the selected subset should be higher than in the initial collection (Hu et al., 2000). The 8.51% value for VD% observed in the mini core subset, suggested the adequacy of the mini core subset. The coefficients of variation for most traits were higher in the mini core subset than the core collection (Table 4) resulting in a higher VR% (107.35%), which indicates that the sample size of the mini core subset was adequate.

No significant differences were found between median values of core collection and mini core subset for any of the traits except leaflet length and pod width in the rainy season and number of primary branches and protein content in the postrainy season (Table 4). The 100% variation range of the core subset was included in the mini core subset for morphological descriptors (Table 4). The variation included in the mini core was 80 to 100% for 11 agronomic traits in the rainy season and 15 traits in the postrainy season. In the remaining five traits in the rainy season the variation included ranged from 65.4 to 79.6%, and in the three traits in the postrainy season the variation included ranged from 69.2 to 77.5% (Table 4). The high CR% (89.3%) retained in the mini core subset indicated that it was representative of the core collection.

The analysis of frequency distribution confirmed homogeneity of distribution between the core collection and the mini core subset (Table 5). These results suggested that the mini core subset chosen is representative of the core collection, which in turn was representative of the entire collection (Upadhyaya et al., 2002).

An adequate and proper sampling, essential in developing a representative core collection, should consider the conservation of phenotypic associations arising out of co-adapted gene complexes (Ortiz et al., 1998). Phenotypic correlations were calculated between the 47 traits in the core collection and mini core subset independently. With more than 1700 degrees of freedom a large number of correlation coefficients which had an absolute value greater than 0.10 were significant at $P = 0.0001$. However, the proportion of variance in one trait that can be attributed to its linear relationship with a second trait is indicated by the coefficient of determination (Snedecor and Cochran, 1980). Considering this criterion, the correlation coefficients with an absolute value greater than 0.71 have been suggested to be as meaningful (Skinner et al., 1999), so that more than 50% of the variation in one trait is predicted by the other. In our study, 19 such meaningful relationship occurred in the core collection. Eight meaningful relationships between the agronomic traits were found (Table 6). In the mini core subset the magnitude of these relationships was also greater than 0.71, except for leaflet width and length in the postrainy season ($r = 0.66$).

This mini core subset preserves the phenotypic correlations observed in the core subset (Table 6). These relationships suggested that it is not necessary to measure all traits in future germplasm evaluations and only easily measurable traits should be given priority. For example, branching pattern and number of primary branches per plant are correlated in the rainy ($r = -0.73$ in the core and $r = -0.68$ in mini core subset) and

Table 6. Correlation coefficients with values more than 0.707 between agronomic traits in the core and mini core subsets of peanut at International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

Traits	Core subset	Mini core subset
Leaflet length postrainy season - leaflet length rainy season	0.74	0.74
Leaflet width rainy season - leaflet length rainy season	0.82	0.79
Leaflet width postrainy season - leaflet length postrainy season	0.82	0.66
Number of seeds per pod postrainy season - number of seeds per pod rainy season	0.99	0.97
Pod length postrainy season - pod length rainy season	0.76	0.72
Pod width rainy season - pod length rainy season	0.70	0.71
Seed length postrainy season - seed length rainy season	0.76	0.75
Days to 50% flowering postrainy season - days to 50% flowering rainy season	0.75	0.77

Table 8. Shannon-Weaver diversity index for agronomic and quality traits in the core and mini core subsets of peanut in the rainy and postrainy seasons at International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

	Core subset			Mini core subset		
Agronomic and quality traits	Rainy	Postrainy	Average ± s.e.	Rainy	Postrainy	Average ± s.e.
Days to emergence	0.599	0.543	0.571 ± 0.028	0.590	0.560	0.575 ± 0.015
Days to 50% flowering	0.564	0.547	0.555 ± 0.009	0.579	0.543	0.561 ± 0.018
Number of primary branches	0.525	0.529	0.527 ± 0.002	0.579	0.539	0.559 ± 0.020
Plant height (mm)	0.623	0.598	0.611 ± 0.013	0.616	0.577	0.597 ± 0.019
Leaflet length (mm)	0.625	0.609	0.617 ± 0.008	0.616	0.620	0.618 ± 0.002
Leaflet width (mm)	0.635	0.600	0.617 ± 0.017	0.590	0.512	0.551 ± 0.039
Number of seeds per pod	0.447	0.446	0.447 ± 0.001	0.465	0.459	0.462 ± 0.003
Pod length (mm)	0.605	0.634	0.620 ± 0.015	0.603	0.607	0.605 ± 0.002
Pod width (mm)	0.594	0.604	0.599 ± 0.005	0.614	0.599	0.606 ± 0.008
Number of pods per plant	0.619	0.612	0.615 ± 0.004	0.605	0.612	0.609 ± 0.003
Seed length (mm)	0.614	0.629	0.621 ± 0.008	0.620	0.584	0.602 ± 0.018
Seed width (mm)	0.455	0.606	0.531 ± 0.076	0.741	0.496	0.618 ± 0.123
100-seed weight (g)	0.626	0.618	0.622 ± 0.004	0.597	0.624	0.611 ± 0.014
Yield per plant (g)	0.601	0.614	0.608 ± 0.006	0.572	0.614	0.593 ± 0.021
Yield per plot (kg ha^{-1})	0.615	0.633	0.624 ± 0.009	0.599	0.169	0.609 ± 0.010
Shelling (%)	0.622	0.582	0.602 ± 0.019	0.612	0.533	0.583 ± 0.030
Oil content (%)	–†	0.627	0.627	–	0.641	0.641
Protein content (%)	–	0.627	0.627	–	0.603	0.603
Average ± s.e.	0.586 ± 0.015	0.592 ± 0.011	0.591	0.600 ± 0.013	0.576 ± 0.012	0.589

† Data not recorded.

postrainy seasons ($r = -0.72$ in the core and $r = -0.64$ in mini core subset). Likewise, the former trait is an easily measurable trait and is negatively correlated with protein content ($r = -0.27$ in entire collection, $r = -0.29$ in core collection), while the later is positively correlated with protein content ($r = 0.29$ in entire collection, $r = 0.34$ in core collection) (Upadhyaya et al., 2002). These results suggest that either of these traits may be a useful measure in choosing newer accessions for further evaluation for protein content.

The Shannon-Weaver diversity index (H′) was calculated to compare phenotypic characters in the core and mini core subset. The index is used in genetic studies as a convenient measure of both allelic richness and allelic evenness. A low H′ indicates an extremely unbalanced frequency of classes for an individual trait and a lack of genetic diversity. The average H′ for the 13 morphological descriptors and agronomic traits in the mini core subset was similar to the core subset (Tables 7 and 8) indicating that the diversity of the core was represented in the mini core subset.

Scientific plant breeding and its need for using variability within the species as well as the concern about loss of landraces, wild types, and cultivars led to large genebank collections. Genebank curators adopted the philosophy of keeping everything in absence of low cost technology to identify unique accessions. This resulted in rapid growth of these germplasm collections. Paradoxically, as these collections grew their utilization did not (Duvick, 1984). Extensive evaluations of entire germplasm collections or even core collections of a thousand or more accessions are very expensive and difficult. This mini core subset of 184 accessions, representing 10.8% of the core collection but including the six botanical varieties of peanut preserves the variation present in the core collection.

The core collection preserved the variation in the entire collection of peanut (Upadhyaya et al., 2002), and the mini core accessions, which only includes 1.29% of the entire collection, represents the total diversity contained in the core collection. This mini core subset drastically reduces the number of entries to be evaluated and provides a working collection of peanut germplasm that can be extensively examined for all economically important traits. The multilocational evaluation of this mini core subset will help in identifying useful parents for improvement programs, which will result in enhanced use of genetic resources for the improvement of quantitative traits. The peanut core collections (Holbrook et al.1993; Upadhyaya et al., 2002) have been evaluated for various traits to identify useful parents (Anderson et al., 1996; Hammond et al., 1997; Franke et al., 1999; Holbrook and Anderson, 1995; Holbrook et al., 1997; Holbrook et al., 1998; Holbrook et al., 2000; Isleib et al., 1995; Upadhyaya et al., 2001).

When selecting the exotic germplasm lines for inclusion in breeding programs, it is important to consider the genetic background of the line as it will be useful in predicting its behavior in hybrid combinations with adapted genotypes. The less divergent the germplasm line and adapted lines are, the more likely it will be that the additive gene effects will play a primary role in inheritance of quantitative traits (Isleib and Wynne, 1983). As the diversity between parents increases, dominance effects and epistatic variations have significant roles in the inheritance of quantitative traits (Halward and Wynne, 1991). In a self-pollinated crop like peanut, this would have implications in choosing an appropriate selection strategy. This mini core subset can be used for molecular characterization to select distinct parents for maximizing diversity in peanut breeding populations. The list of peanut cultivars included in the mini core subset with the ICG number and country of origin are available on diskette, free of charge from the corresponding author. This list is also available on the web site at http://www.icrisat.org/text/research/grep/homepage/project1/gnmncore.htm; verified 3 May 2002.

REFERENCES

Anderson, W.F., C.C. Holbrook, and A.K. Culbreath. 1996. Screening the peanut core collection for resistance to tomato spotted wilt virus. Peanut Sci. 23:57–61.

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

Brown, A.H.D. 1989b. The case for core collections. p. 136–155. *In* A.H.D. Brown et al. (ed.) The use of plant genetic resources. Cambridge Univ. Press, Cambridge, England.

Duvick, D. 1984. Genetic diversity in major farm crops on the farm and in reserve. Econ. Bot. 38:161–178.

Franke, M.D., T.B. Brennemen, and C.C. Holbrook. 1999. Identification of resistance to Rhizoctonia limb rot in a core collection of peanut germplasm. Plant Dis. 83:944–948.

Goodman, M.M. 1985. Exotic maize germplasm: Status, prospects and remedies. Iowa State J. Res. 59:497–527.

Gregory, W.C., and M.P. Gregory. 1976. Groundnut. p. 151–154. *In* N.W. Simmonds (ed.) Evolution of crop plants. Longman Group Ltd., London.

Hallauer, A.R. 1978. Potential of exotic germplasm for maize improvement. p. 229–247. *In* D.B. Waldon (ed.) Maize breeding and genetics. John Willey & Sons, New York.

Halward, T.M., and J.C. Wynne. 1991. Generation means analysis for productivity in two diverse peanut crosses. Theor. Appl. Genet. 82:784–792.

Hammond, E.G., D. Duvick, T. Wang, H. Dodo, and R.N Pittman. 1997. Survey of fatty acid composition of peanut (*Arachis hypogaea*) germplasm and characterization of their epoxy and eicosenoic acids. J. Am. Oil Chem. Soc. 74:1235–1239.

Holbrook, C.C., and W.F. Anderson. 1995. Evaluation of a core collection to identify resistance to late leaf spot in peanut. Crop Sci. 35:1700–1702.

Holbrook, C.C., W.F. Anderson, and R.N Pittman. 1993. Selection of a core collection from the U.S. germplasm collection of peanut. Crop Sci. 33:859–861.

Holbrook, C.C., J. Bruniard, K.M. Moore, and D.A. Knauft. 1998. Evaluation of the peanut core collection for oil content. p. 159. *In* Agronomy Abstracts. ASA, Madison, WI.

Holbrook, C.C., M.G. Stephenson, and A.W. Johnson. 2000. Level and geographical distribution of resistance to *Meloidogyne arenaria* in the U.S. peanut germplasm collection. Crop Sci. 40:1168–1171.

Holbrook, C.C., D.W. Wilson, and M.E. Matheron. 1997. Results from screening the peanut core collection for resistance to preharvest aflatoxin contamination. *In* J. Robens and J. Dorner (ed.) Proc. Aflatoxin Elimination Workshop, Memphis, TN. 27–28 Oct. 1997. USDA-ARS, Beltsville, MD.

Hu, J., J. Zhu, and H.M. Xu. 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.

IBPGR and ICRISAT. 1992. Descriptors for groundnut. Int. Board of Plant Genetic Resources, Rome, Italy, and Int. Crops Res. Inst. for the Semi-Arid Tropics, Patancheru, AP, India.

Isleib, T.G., M.K. Beute, P.W. Rice, and J.E. Hollowell. 1995. Screening the core collection for resistance to *Cylindrocladium* black rot and early leaf spot. Proc. Am. Peanut Res. Educ. Soc. 27:25.

Isleib, T.G., and J.C. Wynne. 1983. Heterosis in testcrosses of 27 exotic peanut cultivars. Crop Sci. 23:832–841.

Jambunathan, R., S.M. Raju, and P. Barde. 1985. Analysis of oil contents of groundnut by nuclear magnetic resonance spectrometry. J. Sci. Food Agric. 36:162–166.

Jiang, H.F., and N.X. Duan. 1998. Utilization of groundnut germplasm resources in breeding programme. Crop Genet. Resour. 2:24–25.

Johns, M.A., P.W. Skroch, J. Nienhuis, P. Hinrichson, G. Bascur, and C. Munoz-Schick. 1997. Gene pool classification of common bean landraces from Chile based on RAPD and morphological data. Crop Sci. 37:605–613.

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

Knauft, D.A., and D.W. Gorbet. 1989. Genetic diversity among peanut cultivars. Crop Sci. 29:1417–1422.

Levene, H. 1960. Robust tests for equality of variances. p. 278–292. *In* I. Olkin (ed.) Contributions to probability and statistics: Essays in honour of Harold Hotelling. Stanford University Press, Stanford, CA.

Mozingo, R.W., T.A. Coffelt, and J.C. Wynne. 1987. Genetic improvement in large-seeded virginia-type peanut cultivars since 1944. Crop Sci. 27:228–231.

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.

Nigam, S.N., L.J. Reddy, H.D. Upadhyaya, and S.L. Dwivedi. 1994. Genetic enhancement in groundnut. p 28–37. *In* Sustainability in oilseeds. Indian Society of Oilseeds Research, Hyderabad, India.

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

SAS Institute. 1989. SAS/STAT user's guide. Version 6, 4th ed. SAS Institute, Inc., Cary, NC

Shannon, C.E., and W. Weaver. 1949. The mathematical theory of communication. Univ. Illinois Press, Urbana.

Singh, U., and R. Jambunathan. 1980. Evaluation of rapid method for the estimation of protein in chickpea (*Cicer arietinum* L.). J. Sci. Food Agric. 31:247–254.

Skinner, D.Z., G.R. Bauchan, G. Auricht, and S. Hughes. 1999. A method for the efficient management and utilization of large germplasm collections. Crop Sci. 39:1237–1242

Snedecor, G.W., and W.G. Cochran 1980. Statistical methods. 7th ed. Iowa State Univ. Press, Ames.

Stalker, H.T. 1980. Utilization of wild species for crop improvement. Adv. Agron. 33:111–147.

Tohme, J., P. Jones, S. Beebe, and M. Iwanaga. 1995. The combined use of agroecological and characterisation data to establish the CIAT *Phaseolus vulgaris* core collection. p. 95–108. *In* T. Hodgkin et al. (ed.) Core collections of plant genetic resources. International Plant Genetic Resources Institute (IPGRI). John Wiley & Sons, New York.

Upadhyaya, H.D., S.N. Nigam, and S. Singh. 2001. Evaluation of groundnut core collections to identify sources of tolerance to low temperature at germination. Indian J. Plant Genet. Resour. 14:165–167.

Upadhyaya, H.D., and R. Ortiz. 2001. A mini core subset for capturing diversity and promoting utilization of chickpea genetic resources in crop improvement. Theor. Appl. Genet. 102:1292–1298.

Upadhyaya, H.D., R. Ortiz, P.J. Bramel, and Sube Singh. 2002. Development of a groundnut core collection using taxonomical, geographical and morphological descriptors. Genet. Resour. Crop Evol. (in press).

Ward, J. 1963. Hierarchical grouping to optimize an objective function. J. Am. Stat. Assoc. 38:236–244.