Genetic Resources and Crop Evolution **50:** 139–148, 2003.

Development of a groundnut core collection using taxonomical, geographical and morphological descriptors

Hari D. Upadhyaya*, Rodomiro Ortiz, Paula J. Bramel and Sube Singh
*Genetic Resources and Enhancement Program, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), P.O. Patancheru 502 324, Andhra Pradesh, India; *Author for correspondence (e-mail: H.Upadhyaya@CGIAR.ORG)*

Received 26 July 2000; accepted in revised form 27 November 2000

Key words: Arachis hypogaea, Core collection, Diversity, Groundnut, Morphological descriptors

Abstract

Groundnut (*Arachis hypogaea* L.) is an important oilseed crop cultivated in 96 countries of world.World crop productivity (1.30 t ha^{-1}) is low. The available large variability contained in the germplasm accessions has not been adequately utilized in the crop improvement programs and most groundnut cultivars stand on a very narrow genetic base. This is due to lack of information on agronomic and other economic traits, which require extensive evaluation. The development of a core collection could facilitate easier access to groundnut genetic resources, enhance their use in crop improvement programs, and simplify the genebank management. This paper describes the development of a core collection from 14310 accessions of groundnut available from ICRISAT genebank. Germplasm accessions were stratified by country of origin within each of six botanical varieties. Data on 14 morphological descriptor traits were used for clustering by Ward's method. From each cluster $\approx$ 10 percent accessions were randomly selected to constitute a core collection consisting of 1704 accessions. Mean comparisons using 't' test and distribution using chi-square test and Wilcoxon's rank-sum non-parametric test on different descriptors indicated that the genetic variation available for these traits in the entire collection has been preserved in the core collection. The Shannon-Weaver diversity index for different traits was also similar in the entire collection and core collection. The important phenotypic correlations between different traits, which may be under the control of co-adapted gene complexes, were preserved in the core collection. This core collection provides an effective mechanism for the proper exploitation of groundnut germplasm resources for the genetic improvement of this crop.

Introduction

Groundnut is an important oilseed crop cultivated in 96 countries of world with an annual production of 30.97 million t on 23.80 million ha (Food and Agriculture Organization of the United Nations 1998). The world productivity of 1.30 t ha^{-1} is rather low. The average productivity in Asia(1.50 t ha^{-1}) and Africa(0.83 t ha^{-1}) is lower than in the USA (2.99 t ha^{-1}). To enhance groundnut productivity, breeding of groundnut cultivars with high yield potential and resistance to various biotic and abiotic constraints is the main objective in most groundnut improvement programs in the world.

The groundnut improvement programs have made significant progress in last two decades but exploitation of the available genetic variability is still limited. Jiang and Duan (1998) have reviewed the utilization of groundnut genetic resources in genetic improvement of crop in China and concluded that introduced foreign germplasm and wild relatives have seldom been utilized in developing cultivars. At ICRISAT also, a few genotypes out of 14310 accessions available in genebank, have been used by the breeding program. The improvement programs aim at rapid cultivar development and had relied mostly on use of established cultivars and elite breeding lines in the development of breeding material (Halward and

Wynne 1991). This approach has resulted in narrowing of the germplasm base of cultivated groundnut that needs to be broaden through enhancing the utilization of genetic resources in developing breeding populations.

The use of genetic resources in the breeding programs have been mainly as sources of resistance to pests and diseases (Knauft and Gorbet 1989). Even in case of resistance sources, breeders prefer their known sources and they are reluctant to use newly identified sources. This is due to lack of information on their reaction to other biotic and abiotic stresses, center of origin/adaptation, and whether they have same components of resistance governed by the same gene or alleles beside their poor agronomic potential (Singh and Simpson 1994). In fact there has been even fewer efforts in even identifying germplasm lines for increasing yield potential than for pest resistance and nutritional quality (Halward and Wynne 1991). Such a situation has arisen because traits like yield display a great deal of genotype by environment (G × E) interaction and to identify useful parents, multilocational evaluation of the large germplasm collection in the replicated field trials is required. This is a very costly exercise due to size of germplasm collection, and nearly impossible to carry out in view of dwindling resources. To overcome this problem, Frankel (1984) proposed sampling of the collection to a manageable sample or 'core collection'. A core collection contains a subset of accessions from entire collection that captures most of available genetic diversity of species (Brown 1989a). The core collection can be evaluated extensively and the information derived could be used to guide more efficient utilisation of the entire collection (Tohme et al. 1995; Brown 1989b).

Cultivated groundnut is an allotetraploid (2n = 40, x = 10) with two genomes, A and B. The evidence from molecular studies of genetic variation suggests that groundnut had a single origin involving a hybridization of *A. duranensis* Krapov et W. C. Gregory (A genome) and *A. ipaensis* Krapov et W. C. Gregory (B genome) (Kochert et al. 1996). Earlier reports (Kochert et al. 1991; Halward et al. 1991, 1992; Paik-Rao et al. 1992) suggested low levels of genetic variability, as assessed by molecular markers in the cultivated groundnut. In past this has limited the use of molecular markers in the genetic improvement of groundnut. However, in last few years, there is increasing evidence of detection of molecular polymorphism in the cultivated groundnut (He and Prakash 1997; Hopkins et al. 1999). Our studies at ICRISAT (Subramanian et al. 2000; Dwivedi et al. Unpublished) also provide evidence of substantial geneticvariation in the cultivated species suggesting that earlier reports of low genetic variation were due to limitation of techniques and materials used. For morphological and agronomic traits a large variation exists which can be used to develop core collection.

Frankel and Brown (1984), Brown (1989a, 1989b) described methods to select a core collection using information on the origin and characteristics of the accessions. In setting the core collection, the first issue is its size. Brown (1989a) using sampling theory of selectively neutral alleles argued that the entries in a core collection should be about 10% of total collection, with a ceiling of 3000 per species. This level of sampling is effective in retaining 70% alleles of the entire collection. The second issue is the degree of the genetic similarity or commonality among accessions and determining groups within the entire collection. The hierarchy of grouping begins with the groupings suggested by taxonomy (species, subspecies, and races) followed by assigning accessions to major geographic groups (country, state), climate or agroecological regions. The clustering within the broad geographic group could be done using strongly inherited traits. The number of accessions selected from each cluster will depend on the strategy used.

Since the original proposal of Frankel (1984), core collections have been established for many species including common bean (*Phaseolus vulgaris* L.) (Tohme et al. 1995), barley (*Hordeum vulgare* L.) (Knüpffer and van Hintum 1995), chickpea (*Cicer arietinum* L.) (Hannan et al. 1994; Upadhyaya et al. 2001), annual and perennial *Medicago* species (Diwan et al. 1994; Basigalup et al. 1995), perennial *Glycine* (Brown et al. 1987), cassava (*Manihot esculenta* Crantz) (Cordeiro et al. 1995), coffee (*Coffee* spp.) (Dussert et al. 1997), lentil (*Lens culinaris* Medic.) (Erskine and Muehlbauer 1991), okra (*Abelmoschus esculentus* L.) (Mahajan et al. 1996), peanut (*Arachis hypogaea* L.) (Holbrook et al. 1993), quinoa (*Chenopodium quinoa* Willd.) (Ortiz et al. 1998), and sweet potato (*Ipomoea batatas* L.) (Huaman et al. 1999). Holbrook et al. (1993) used data on 6 morphological variables, plant type, pod type, seed size, testa color, seed per pod, and seed weight in 7432 U. S. groundnut germplasm collection to develop a core collection comprising 831 accessions. ICRISAT holds the largest germplasm collection of 14310 groundnut accessions in its genebank in trust following the

agreement signed in 1994 with FAO. The objective of this study was to develop a core collection of groundnut from 14310 accessions contained in the ICRISAT genebank using data recorded on 14 morphological descriptors.

Materials and methods

A total of 14310 accessions from 92 countries were included to select the core collection. *Arachis hypogaea* L. subsp. *hypogaea* var. *hypogaea* had 6622 (46.3%) accessions followed by *A. hypogaea* L. subsp. *fastigiata* Wal. var. *vulgaris* Harz. with 5102 (35.7%) accessions, and *A . hypogaea* L. subsp. *fastigiata* var. *fastigiata* with 2302 (16.1%) accessions. Data were recorded in these 14 morphological descriptors: stem color, stem hair, branching pattern, leaf color, leaf shape, leaf hair, flower color, streak color, peg color, pod beak, pod constriction, pod reticulation, number of seeds per pod, and seed color pattern following IBPGR & ICRISAT (1992). The number of accessions for which data on morphological descriptors was recorded according to each botanical variety is given in Table 1.

The ICRISAT groundnut collection was stratified first by the botanical varieties within subspecies, i.e., subsp. *hypogaea* var. *hypogaea* and var. *hirsula* Kohl, and subsp. *fastigiata* var. *fastigiata*, var. *peruviana* Krapov et W. C. Gregory, var. *aequatoriana* Krapov et W. C. Gregory, and var. *vulgaris* (Krapovickas and Gregory 1994) followed by their country of origin. The accessions in a botanical variety from small and adjacent countries with similar agro-climate were grouped together following Brown (1989a). There were, therefore, 75 groups representing groundnut accessions held at ICRISAT genebank. The data on 14 morphological descriptors in each group was standardized using range of each variable to eliminate scale differences (Milligan and Cooper 1985). The standardized data was subjected to the hierarchical cluster algorithm of Ward (1963) at R^2 (squared multiple correlation) equal to 0.75, using SAS (SAS Institute 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 agglomerative procedure starts with n groups, i.e., one observation in one group (maximum between-group sum of squares) and proceeds by merging observations in groups so that the between-groups sum of squares decreases and within-groups sum of squares increases. In certain cases the within-groups sum of squares will remain the same but it will never decrease. From each cluster approximately 10% accessions were randomly selected for inclusion into the core collection. At least one accession was included even from those clusters that had less than 10 accessions.

The means of the entire collection and core collection were compared using t-tests for all morphological descriptors. The distribution homogeneity for each of 14 morphological descriptors between the entire collection and the core collection was also analysed by the Chi-square test. The Wilcoxon (1945) rank-sum non-parametric test was performed using the SAS NPAR1WAY procedure (SAS Institute 1989) to determine whether the core collection represents the entire germplasm collection for each of 14 traits. In this test the data in the entire and core collections are ranked. The smallest value is given a rank of 1 and the

Table 1. Number of accessions per botanical variety of groundnut on which morphological descriptors were recorded

Descriptor	*fastigiata*	*vulgaris*	*aequatoriana*	*peruviana*	*hypogaea*	*hirsuta*
Stem color	2295	5089	15	249	6592	20
Stem hairs	2295	5092	15	249	6592	20
Branching pattern	2301	5098	15	249	6602	20
Leaf color	2300	5095	15	249	6605	20
Leaf shape	2296	5096	15	249	6605	20
Leaf hairs	2300	5094	15	249	6601	20
Flower color	2293	5086	15	249	6596	20
Streak color	2293	5086	15	249	6596	20
Peg color	2292	5081	15	249	6592	20
Pod beak	2297	5091	15	249	6612	20
Pod constriction	2298	5091	15	249	6612	20
Pod reticulation	2298	5091	15	249	6612	20
Seeds per pod	2298	5090	15	249	6612	20
Seed color pattern	2249	5038	15	249	6540	20

Table 2. Means (± standard error) for morphological descriptors for the entire and core collections of groundnut

Descriptor	Entire collection	Core collection	Differences
Stem color	1.20 ± 0.003	1.22 ± 0.010	*
Stem hair	3.43 ± 0.008	3.49 ± 0.024	*
Branching pattern	1.53 ± 0.004	1.53 ± 0.012	NS
Leaf color	2.50 ± 0.004	2.50 ± 0.013	NS
Leaf shape	2.99 ± 0.001	2.99 ± 0.002	NS
Leaf hairs	1.08 ± 0.003	1.10 ± 0.012	*
Flower color	4.95 ± 0.002	4.95 ± 0.008	NS
Streak color	5.95 ± 0.002	5.98 ± 0.006	NS
Peg color	1.98 ± 0.001	1.97 ± 0.004	*
Pod beak	3.98 ± 0.011	4.00 ± 0.033	NS
Pod constriction	4.46 ± 0.010	4.45 ± 0.031	NS
Pod reticulation	4.42 ± 0.012	4.45 ± 0.038	NS
Seeds per pod	1.81 ± 0.034	1.89 ± 0.034	*
Seed color pattern	1.03 ± 0.001	1.03 ± 0.004	NS

NS and * indicate non-significant or significant differences at $p = 0.05$, respectively.

next small value will have a rank of 2, and so on. The ranks for each of entire collection and core collection are added together and sums of ranks are used to test the null hypothesis of no differences in distribution of entire collection and core collection. The diversity index (H′) of Shannon and Weaver (1949) was estimated and used as a measure of phenotypic diversity in the entire collection and core collection for each trait. The phenotypic correlations between descriptors in the entire collection and core collection were estimated independently, to know whether these associations, which may be under genetic control, were conserved in the core collection or not.

Results and discussion

A core collection consisting of 1704 accessions from the six groundnut botanical varieties was developed from the entire collection held by ICRISAT genebank. This was 11.9% of the accessions available in ICRISAT genebank. The core collection consists of 584 (34.3%) accessions belonging to var. *vulgaris*, 299 (17.5%) to var. *fastigiata*, 27 (1.6%) to var. *peruviana*, 6 (0.4%) to var. *aequatoriana*, 784 (46.0%) to var. *hypogaea*, and 4 (0.2%) to var. *hirsuta*. Except for *aequatoriana* and *hirsuta*, which have only 15 and 20 accessions, respectively in the entire collection, the representation of botanical varieties in the core collection corresponded with their contribution to the entire collection.

Differences between means of the core collection and entire collection were found significant for stem color, stem hairs, leaf hairs, peg color, and seeds per pod (Table 2). However, individual botanical variety wise, the differences were significant for stem color, stem hairs, branching pattern, and leaf color in *hypogaea*, for leaf color and pod reticulation in *fastigiata*, pod reticulation and seeds per pod in *peruviana*, and flower color, streak color and peg color in *vulgaris* . For one variety in both subspecies, variety *aequatoriana* in subsp. *fastigiata* and variety *hirsuta* in subsp. *hypogaea*, the differences were not significant for any of the 14 descriptors. Except branching pattern for which 50% variation in range was recovered, the core collection represented 100% variation of the entire collection. None of the seven accessions with alternate flowering on lateral branches with flowering on main stem or sequential flowering on lateral branches without flowering on main

Table 3. Comparison of frequency distribution for morphological descriptors in the entire and core collections of groundnut

Descriptor	Number of classes	χ^2	Probability
Stem color	2	6.144	0.014
Stem hairs	5	8.603	0.072
Branching pattern	4	5.673	0.129
Leaf color	5	0.581	0.965
Leaf shape	3	12.919	0.002
Leaf hairs	7	16.195	0.0128
Flower color	6	26.238	0.0001
Streak color	5	24.067	0.0001
Peg color	2	8.580	0.003
Pod beak	5	1.708	0.789
Pod constriction	5	9.449	0.051
Pod reticulation	5	10.487	0.033
Seeds per pod	8	12.663	0.081
Seed color pattern	2	1.996	0.158

Table 4. Means (± standard error) for the entire collection and for the core collection of groundnut for some of the traits not included in the development of core collection

Trait	Entire collection		Core Collection		Differences
	Number	Mean	Number	Mean	
Oil content (%)	7948	43.8 ± 0.03	991	43.7 ± 0.09	NS
Protein content (%)	5473	25.7 ± 0.03	682	25.8 ± 0.09	NS
Number of primary branches	9616	5.8 ± 0.02	1138	5.8 ± 0.06	NS
Primary seed color	14300	8.7 ± 0.03	1702	8.8 ± 0.08	*
Rust	8241	8.9 ± 0.01	1003	8.8 ± 0.03	NS
Late leaf spots	8236	8.9 ± 0.01	1000	8.9 ± 0.02	NS
Early leaf spots	10552	8.3 ± 0.01	1266	8.3 ± 0.02	NS
Rosette	11953	8.9 ± 0.01	1267	8.9 ± 0.02	NS
Plant height (cm)	9044	27.3 ± 0.12	1095	27.5 ± 0.35	NS
Growth habit	14285	4.6 ± 0.02	1700	4.6 ± 0.05	NS
Rainy season					
Days to emergence	12939	8.5 ± 0.01	1549	8.6 ± 0.04	NS
Days to 50% flowering	12952	25.1 ± 0.03	1549	25.0 ± 0.09	NS
Leaflet length (mm)	12992	52.7 ± 0.07	1571	52.8 ± 0.20	NS
Leaflet width (mm)	12994	23.5 ± 0.03	1571	23.6 ± 0.09	NS
Leaflet length/width ratio	12986	2.25 ± 0.002	1570	2.25 ± 0.005	NS
100-seed weight (g)	12943	43.7 ± 0.10	1563	44.2 ± 0.29	NS
Pod length (mm)	12726	28.6 ± 0.05	1534	29.0 ± 0.13	*
Pod width (mm)	12726	12.2 ± 0.01	1533	12.3 ± 0.04	*
Pod length/width ratio	12723	2.34 ± 0.003	1533	2.35 ± 0.008	NS
Seed length (mm)	12736	13.3 ± 0.02	1535	13.3 ± 0.05	NS
Seed width (mm)	12737	7.8 ± 0.01	1534	7.8 ± 0.02	NS
Seed length/width ratio	12734	1.70 ± 0.002	1534	1.71 ± 0.007	NS
Shelling (%)	12490	68.2 ± 0.07	1485	67.9 ± 0.19	*
Postrainy season					
Days to emergence	13492	12.0 ± 0.02	1627	12.0 ± 0.04	NS
Days to 50% flowering	13485	37.0 ± 0.05	1626	36.8 ± 0.13	NS
Leaflet length (mm)	13215	54.3 ± 0.08	1605	54.8 ± 0.23	*
Leaflet width (mm)	13215	25.3 ± 0.04	1605	25.5 ± 0.011	NS
Leaflet length/width ratio	13208	2.15 ± 0.002	1607	2.16 ± 0.005	NS
100-seed weight (g)	13240	51.4 ± 0.11	1605	51.9 ± 0.32	NS
Pod length (mm)	13143	30.4 ± 0.04	1592	30.8 ± 0.13	**
Pod width (mm)	13142	13.2 ± 0.01	1592	13.3 ± 0.04	**
Pod length/width ratio	13142	2.31 ± 0.002	1592	2.32 ± 0.007	NS
Seed length (mm)	13147	14.1 ± 0.02	1590	14.2 ± 0.06	*
Seed width (mm)	13148	8.7 ± 0.01	1592	8.7 ± 0.02	NS
Seed length/width ratio	13142	1.43 ± 0.002	1590	1.65 ± 0.006	**
Shelling (%)	12647	70.5 ± 0.07	1528	70.1 ± 0.19	NS

NS, *, and ** indicate non-significant or significant differences at $p = 0.05$ and $p = 0.01$, respectively.

stem in the entire collection were included in the core collection resulting in lowering the range for branching pattern. These results indicated that the chosen subset is representative of the entire collection and that the variation in the entire collection has been preserved in the core collection except for branching pattern.

Frequency distribution analysis indicated homogeneity of distribution, except for stem color ($p = 0.014$), leaf shape ($p = 0.002$), leaf hair ($p = 0.013$), flower color ($p = 0.0001$), streak color ($p = 0.0001$), and peg color ($p = 0.003$), between entire collection and core collection (Table 3). The frequency of accessions in core collection with pigmented stem, high hair density on stem surface, almost glabrous and woolly without bristles leaf sufrace, and pods with mainly three seeds was more than expected and resulted in to a high mean for these descriptors. The Wilcoxon rank-sum test indicated that all traits except stem color ($p = 0.02$), stem hairs ($p = 0.01$), and peg color ($p = 0.01$) have similar distribution in both core collection and entire collection. Variety wise also the Wilcoxon rank-sum test indicated that except for stem color, stem hairs, branching pattern, leaf color and

Table 5. Shannon-Weaver diversity index for 14 morphological descriptors in the entire and core collections of groundnut

Descriptor	Entire collection	Core collection
Stem color	0.217	0.231
Stem hair	0.246	0.264
Branching pattern	0.302	0.300
Leaf color	0.325	0.327
Leaf shape	0.009	0.022
Leaf hairs	0.111	0.128
Flower color	0.111	0.143
Streak color	0.061	0.090
Peg color	0.044	0.060
Pod beak	0.447	0.455
Pod constriction	0.396	0.419
Pod reticulation	0.468	0.493
Seeds per pod	0.489	0.512
Seed color pattern	0.054	0.063
Average ± s.e.	0.234 ± 0.0455	0.251 ± 0.0455

peg color in *hypogaea*, for leaf color and seed color pattern in *fastigiata*, seeds per pod in *peruviana*, and flower color, streak color and peg color in *vulgaris*, all the descriptors have similar distribution in the core collection and entire collection.

Similarly, for the 27 out of 36 traits, which were not considered for selecting the core collection due to availability of evaluation data on less number of accessions, the means of core collection and entire collection were not significantly different (Table 4). The core collection represented 100% range of entire collection for primary seed color, growth habit, rust (*Puccinia arachidis* Speig.), early leaf spot (*Cercospora arachidicola* Hori), and rosette virus disease. For ten other traits from 80% to 96.2% range of entire collection was recovered in the core collection. For the traits like resistance to rust and rosette, the percentage of accessions with a score in the core collection was similar to the entire collection. For example, in the entire collection only 126 (1.05% out of 11953 accessions evaluated for rosette had a rating of 1 on a 1 (resistant) to 9 (very susceptible) scale. Of these 126 accessions, 19 (1.33%) were represented in the core collection.

The Shannon-Weaver diversity index (H′) in the core collection was similar to that of the entire collection for all the descriptors (Table 5), which indicates that the diversity of the entire collection was represented in the core collection. The average H′ in the core collection was 0.171 in *vulgaris* (0.157 in the entire collection), 0.283 in *aequatoriana* (0.294), 0.257 in *fastigiata* (0.228), 0.223 in *hirsuta* (0.216), 0.188 in *hypogaea* (0.167), and 0.264 in *peruviana* (0.257) suggesting that the diversity in the each of botanical variety was adequately sampled in this core collection.

In developing a core collection it is important to have proper and adequate sampling to conserve the phenotypic associations arising out of co-adapted gene complexes (Ortiz et al. 1998). The phenotypic correlations observed in the entire collection have been preserved in the groundnut core collection (Table 6). This clearly demonstrates that the co-adapted gene complexes controlling these associations in the entire collection were properly sampled and that the selection of the core collection was adequate in this regard. Further the correlations among 21 evaluation traits were also examined. With such a large degree of freedom, 14000 or more, any correlation coefficient with an absolute value greater than 0.01 will be significant at $p = 0.0001$. However, because the proportion of variance in one trait that can be attributed to its linear regression on second trait is indicated by the square of the correlation coefficient (Snedecor and Cochran 1980). Therefore, 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, we found such meaningful relationship in the entire collection ($r = -0.89$) and core collection ($r = -0.78$) between leaf color and branching pattern (Table 6). Similarly for the evaluation traits also, we found strong meaningful relationships between several traits in the entire collection (Table 7). In core collection also, except between pod width in rainy season and pod width in postrainy season ($r = 0.67$) the magnitude of these correlations remained greater than 0.71. Further, these relationships suggested that it is not be necessary to measure all traits in future germplasm evaluations. For example, the strong correlation between reaction to rust and late leaf spots (*Phaeoisariopsis personata* Berk. & M. A. Curtis), which were $r = 0.83$ in the entire collection and $r = 0.86$ in the core collection, indicated that independent screenings against both of these important diseases are not required. However, the weak correlation between reaction to rust and early leaf spots ($r = 0.02$ in entire collection and 0.03 in core collection) and between late and early leaf spots ($r = 0.04$ in entire collection and 0.06 in core collection) suggested that independent screenings for rust or late and early leaf spots are required.

Some other relationships while not meeting the

Table 6. Correlation coefficients between morphological descriptors in the entire collection (above diagonal) and the core collection (below diagonal) of groundnut. A correlation > 0.01 will be significant at $p = 0.0001$

	Stem color	Stem hairs	Branching pattern	Leaf color	Leaf shape	Leaf hairs	Flower color	Streak color	Peg color	Pod beak	Pod constriction	Pod reticulation	Seeds per pod	Seed color pattern
Stem color		0.181	0.277	− 0.241	0.000	0.045	− 0.170	− 0.003	0.042	− 0.121	− 0.227	− 0.132	0.460	0.016
Stem hairs	0.168		0.139	− 0.082	0.060	0.064	− 0.012	− 0.125	− 0.211	0.034	− 0.069	0.175	0.250	0.097
Branching pattern	0.347	0.184		− 0.891	0.000	0.021	− 0.115	− 0.065	− 0.093	− 0.230	− 0.011	− 0.071	0.117	0.016
Leaf color	− 0.302	− 0.055	− 0.782		0.000	0.017	0.155	0.094	0.075	0.214	− 0.014	0.081	− 0.056	0.027
Leaf shape	0.000	0.150	0.000	0.000		0.000	0.000	0.000	0.000	0.028	0.030	0.024	− 0.013	0.000
Leaf hairs	0.000	0.115	0.106	0.126	0.000		0.065	0.029	− 0.037	0.025	− 0.050	0.058	0.052	− 0.005
Flower color	− 0.167	0.188	− 0.048	0.272	0.000	0.328		0.338	0.010	0.082	0.126	0.133	− 0.138	− 0.017
Streak color	− 0.062	− 0.224	− 0.067	0.184	0.000	0.193	0.558		0.346	− 0.002	0.059	− 0.101	− 0.111	− 0.139
Peg color	0.119	− 0.268	− 0.095	0.076	0.000	− 0.021	− 0.017	0.312		0.004	0.085	− 0.186	− 0.210	− 0.208
Pod beak	− 0.151	0.189	− 0.265	0.287	0.073	0.077	0.161	− 0.054	0.005		0.240	0.303	− 0.022	0.047
Pod constriction	− 0.215	− 0.012	− 0.054	0.049	0.077	− 0.089	0.154	0.016	0.127	0.276		0.168	− 0.450	− 0.102
Pod reticulation	− 0.108	0.359	0.052	0.074	0.054	0.166	0.223	− 0.236	− 0.352	0.287	0.168		− 0.042	0.212
Seeds per pod	0.335	0.337	0.271	− 0.056	− 0.030	0.206	− 0.026	− 0.212	− 0.278	0.004	− 0.443	0.124		0.142
Seed color pattern	− 0.013	0.139	0.066	0.024	0.000	0.072	0.012	− 0.295	− 0.407	− 0.003	− 0.213	0.387	0.292	

Table 7. Correlation coefficients with values more than 0.707 between traits in the entire and core collections of groundnut

Traits	Entire collection	Core collection
Rust-late leaf spots	0.834	0.861
Leaflet length rainy-leaflet width rainy	0.862	0.819
Pod length rainy-pod length width ratio rainy	0.725	0.781
Pod length rainy-pod length postrainy	0.864	0.867
Pod width rainy-pod width postrainy	0.724	0.674
Seed length rainy-seed length width ratio rainy	0.775	0.775
Seed length rainy-seed length postrainy	0.830	0.821
Seed length rainy-seed length width ratio postrainy	0.713	0.720
100-seed weight rainy-100-seed weight postrainy	0.738	0.705
Leaflet length postrainy-Leaflet width postrainy	0.838	0.778
Pod length postrainy-Pod length width ratio postrainy	0.715	0.759
Seed length postrainy-seed length width ratio postrainy	0.809	0.804

50% criterion, may be of interest to plant breeders. For example, branching pattern and number of primary branches per plant are two easily measurable traits. The former is negatively correlated ($r = -0.27$ in entire collection, $r = -0.29$ in core collection) and the later is positively correlated ($r = 0.29$ in entire collection, $r = 0.34$ in core collection) with protein content suggesting that either of these traits may be useful measure in choosing newer accessions for further evaluation for protein contents.

This core collection can be used very efficiently as a starting point for the research projects involving screening of germplasm collection for sources of desirable traits in groundnut. Aflatoxin contamination of the groundnut by the aflatoxin-producing fungi *Aspergillus flavus* and A. *parasiticus* is a serious health hazard and a major challenge to the groundnut industry. ICG 4749 (PI 337394 F), an accession from Argentina and resistant to seed infection and colonization by *A. flavus* (Mixon and Rogers 1973) was selected in this core collection. This accession belongs to cluster 5 from the Americas. There are 21 accessions in the cluster containing ICG 4749. These accessions may be similar to ICG 4749 and could also be valuable sources of resistance. The information on clusters to which particular accessions with traits of interest belong will assist thus in looking extensively for more accessions with similar traits. The core collection will also provide an efficient germplasm subset if it is not feasible to screen the entire germplasm collection. For example, in diseases like rosette, which is one of most destructive disease of groundnut in Africa, it is important to obtain the information on amount of variability present in the germplasm. In the 1995 rosette epidemic in the Eastern Province of Zambia, 43000 ha of groundnut crop were affected resulting in a loss of about US $ 4.9 million (Pala Subrahmanyam ICRISAT, personal communication). Using currently available screening technique it will take at least 7 to 8 years and significant resources to examine the entire germplasm collection. However, the core collection allows us to estimate the amount of genetic variability in the entire germplasm collection and possibly identify new sources of alleles for resistance in about one year with lesser resources.

The development of the groundnut core collection helps in tackling new constraints that may arise owing to new diseases or insect pests. Because, the core collection represents the entire germplasm collection and seeds of the core accessions are available, resistance sources to the new disease or an insect pest may be identified rapidly. For additional sources of resistance the accessions from the reserve collection can be examined selectively from the same cluster from which sources in the core collection have been identified.

The resources available for evaluation of germplasm are limited and dwindling steadily. Therefore, extensive evaluations of entire germplasm collection are not possible. This core collection provides a working collection of groundnut germplasm that can be extensively examined for all economically important traits. The data generated will provide the required information on genetic variability in groundnut and the possible relationships among the traits. This information will assist further in the decision making process to acquire new variability for a trait showing a very limited variation in the core collection. This groundnut core collection should be revised periodically as additional accessions are collected, particularly from botanical varieties *hirsuta* and *aequatoriana* as well as others from locations that are not represented or under represented in the ICRISAT genebank (Singh and Nigam 1997). These locations

include traditional groundnut areas in subsistence agriculture, areas of early introduction such as Laos, China, Angola, Malagasay Republic, Namibia, and South Africa, or areas of secondary center of diversity in Peru, Ecuador, Paraguay, and Uruguay.

The list of groundnut entries according to their botanical variety that are in the core collection includes name of country of origin, ICG number, and the cluster number are available from the corresponding author of this paper. The list is also available at http://grep.icrisat.cgiar.org/

References

Basigalup D.H., Barnes D.K. and Stucker R.E. 1995. Development of a core collection for perennial *Medicago* plant introductions. Crop Sci. 35: 1163–1168.

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. In: Brown A.H.D., Frankel O.H., Marshall D.R. and Williams J.T. (eds), The use of plant genetic resources. Cambridge Univ. Press, Cambridge, pp. 136–155.

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

Cordeiro C.M.T., Morales E.A.V., Ferreira P., Rocha D.M.S., Costa I.R.S., Valois A.C.C. et al. 1995. Towards a Brazilian core collection of cassava. In: Hodgkin T., Brown A.H.D., van Hintum Th.J.L. and Morales B.A.V. (eds), Core collections of plant genetic resources. International Plant Genetic Resources Institute (IPGRI), John Wiley & Sons, New York, pp. 155–168.

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

Dussert S., Chabrillange N., Anthony F., Engelmann F., Recalt C. and Hamon S. 1997. Variability in storage response within a coffee (*Coffea* spp.) core collection under slow growth conditions. Plant Cell Rep. 16: 344–348.

Dwivedi, S.L., Gurtu S., Chandra S., Upadhyaya H.D. and Nigam S.N. 2001. Assessment of genetic diversity among selected rosette virus resistant groundnut (Arachis hypogaea L.) germplasm. II. RAPD, AFLP, and phenotypic diversity. Unpublished.

Erskine W. and Muehlbauer F.J. 1991. Allozyme and morphological variability, outcrossing rate and core collection formation in lentil germplasm. Theor. Appl. Genet. 83: 119–125.

Frankel O.H. 1984. Genetic perspective of germplasm conservation. In: Arber W., Llimensee K., Peacock W.J. and Starlinger P. (eds), Genetic manipulations: impact on man and society. Cambridge University Press, Cambridge, pp. 161–170.

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. p. 1–13.

Food and Agriculture Organization of the United Nations 1998. Production Yearbook 52: 103–104.

Hannan R.M., Kaiser W.J. and Muehlbauer F.J. 1994. Development and utilization of the USDA chickpea germplasm core collection. In: Agronomy Abstracts 1994. ASA, Madison, WI p. 217.

Halward T.M., Stalker H.T., Larue E.A. and Kochert G. 1991. Genetic variation detectable with molecular markers among unadapted germ-plasm resources of cultivated peanut and related wild species. Genome 34: 1013–1020.

Halward T.M., Stalker H.T., Larue E.A. and Kochert G. 1992. Use of single-primer DNA amplification in genetic studies of peanut (*Arachis hypogaea* L.). Plant Mol. Biol. 18: 315–325.

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

He G. and Prakash C.S. 1997. Identification of polymorphic DNA markers in cultivated peanut (*Arachis hypogaea* L.). Euphytica 97: 143–149.

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

Hopkins M.S., Casa A.M., Wang T., Mtchell S.E., Dean R.E., Kochert G.D. et al. 1999. Discovery and characterization of polymorphic simple sequence repeats (SSRs) in peanut. Crop Sci. 39: 1243–1247.

Huaman Z., Aguilar C. and Ortiz R. 1999. Selecting a Peruvian sweet potato core collection on the basis of morphological, ecogeographical, and disease and pest reaction data. Theor. Appl. Genet. 98: 840–844.

IBPGR & ICRISAT (1992). Descriptors for groundnut. Int. Board of Plant Genetic Resources, Rome, Italy, and Int. Crops Res. Inst. For the Semi-Arid Tropics, Patancheru, A.P., India

Jiang H.F. and Duan N.X. 1998. Utilization of groundnut germplasm resources in breeding programs. Crop Genetic Resources 2: 24–25.

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

Knüpffer H. and van Hintum Th.J.L. 1995. The barley core collection: an international effort. In: Hodgkin T., Brown A.H.D., van Hintum Th.J.L. and Morales B.A.V. (eds), Core collections of plant genetic resources. International Plant Genetic Resources Institute (IPGRI), John Wiley & Sons, New York, pp. 171–178.

Kochert G., Halward T., Branch W.D. and Simpson C.E. 1991. RFLP variability in peanut (*Arachis hypogaea* L.) cultivars and wild species. Theor. Appl. Genet. 81: 565–570.

Kochert G., Stalker H.T., Gimenes M., Galgaro L., Romero Lopes C. and Moore K. 1996. RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). Am. J. Bot. 83: 1282–1291.

Krapovickas A. and Gregory W.C. 1994. Taxonomía del genéro *Arachis* (Leguminosae). Bonplandia VIII: 1–187.

Mahajan R.K., Bisht I.S., Agrawal R.C. and Rana R.S. 1996. Studies on south Asian okra collection: a methodology for establishing a representative core set using characterization data. Genet. Resour. Crop Evol. 43: 244–255.

Milligan G.W. and Cooper M. 1985. An examination of procedures for determining the number of clusters in a data set. Psychometrika. 50: 159–179.

Mixon A.C. and Rogers K.M. 1973. Peanut accessions resistant to seed infection by *Aspergillus flavus*. Agron. J. 65: 560–562.

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

Paik-Rao O.G., Smith R.L. and Knauft D.A. 1992. Restriction

fragment length polymorphism evaluation of 6 peanut species within the *Arachis* section. Theor. Appl. Genet. 84: 201–208.
SAS Institute 1989. SAS/STAT User's Guide Version 6, 4th edn. SAS Institute, Inc., Cary, NC.
Shannon C.E. and Weaver W. 1949. The mathematical theory of communication. Univ. Illinois Press, Urbana.
Singh A.K. and Simpson C.E. 1994. Biosystematics and genetic resources. In: Smartt J. (ed.), The groundnut crop: a scientific basis for improvement. Chapman & Hall, London, pp. 96–137.
Singh A.K. and Nigam S.N. 1997. Groundnut. In: Fuccillo D., Sears L. and Stapleton P. (eds), Biodiversity in trust – conservation and use of plant genetic resources in CGIAR centers. Cambridge Univ. Press, Cambridge, pp. 114–128.
Skinner D.Z., Bauchan G.R., Auricht G. and Hughes S. 1999. A method for the efficient management and utilization of large germplasm collections. Crop Sci. 39: 1237–1242.
Snedecor G.W. and Cochran W.G. 1980. Statistical methods. 7th edn. Iowa State Univ. Press, Ames.
Subramanian V., Gurtu S., Nageswara Rao R.C. and Nigam S.N. 2000. Identification of DNA polymorphism in cultivated groundnut using random amplified polymorphic DNA (RAPD) assay. Genome 43: 656–660.
Tohme J., Jones P., Beebe S. and Iwanaga M. 1995. The combined use of agroecological and characterisation data to establish the CIAT *Phaseolus vulgaris* core collection. In: Hodgkin T., Brown A.H.D., van Hintum Th.J.L. and Morales B.A.V. (eds), Core collections of plant genetic resources. International Plant Genetic Resources Institute (IPGRI), John Wiley & Sons, New York, pp. 95–108.
Upadhyaya H.D., Bramel P.J. and Singh S. 2001. Development of a chick pea core subset using geographic distribution and quantitative traits. Crop Sci. 41: 206–210.
Ward J. 1963. Hierarchical grouping to optimize an objective function. J. Am. Stat. Assoc. 38: 236–244.
Wilcoxon F. 1945. Individual comparisons by ranking methods. Biometrics Bull. 1: 80–83.