Theor Appl Genet (2001) 102:1292–1298

H.D. Upadhyaya · R. Ortiz

A mini core subset for capturing diversity and promoting utilization of chickpea genetic resources in crop improvement

Received: 20 August 2000 / Accepted: 25 September 2000

Abstract A core collection is a chosen subset of large germplasm collection that generally contains about 10% of the total accessions and represents the genetic variability of entire germplasm collection. The purpose of a core collection is to improve the use of genetic resources in crop improvement programs. In many crops the number of accessions contained in the genebank are several thousands, and a core subset consisting of 10% of total accessions would be an unwieldy proposition. In this article we have suggested a two-stage strategy to select a chickpea mini core subset consisting of only about 1% of the entire collection held in trust at ICRISAT's genebank (16,991 accessions). This mini core subset still represents the diversity of the entire core collection. The first stage involves developing a representative core subset (about 10%) from the entire collection using all the available information on origin, geographical distribution, and characterization and evaluation data of accessions. The second stage involves evaluation of the core subset for various morphological, agronomic, and quality traits, and selecting a further subset of about 10% accessions from the core subset. At both stages standard clustering procedure was used to separate groups of similar accessions. A mini core subset consisting 211 accessions from 1,956 core subset accessions, using data on 22 morphological and agronomic traits, was selected. Newman-Keuls' test for means, Levene's test for variances, the chi-square test and Wilcoxon's rank-sum non-parametric 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. The most important phenotypic correlations which may be under the control of coadapted gene complexes, were also preserved in the mini core. This mini core subset, due to its drastically reduced size, will prove to be a point of entry to proper exploitation of chickpea genetic resources.

Keywords Agronomic traits · Chickpea · Diversity · Mini core subset · Morphological traits

Communicated by P.M.A. Tigerstedt

H.D. Upadhyaya · R. Ortiz (✉)
Genetic Resources and Enhancement Program,
International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), P.O. Patancheru 502 324, Andhra Pradesh, India
Fax: +91-40-3241239
e-mail: R.Ortiz@CGIAR.ORG

Introduction

Plant breeders have successfully improved the yield potential of most crops, which has resulted in higher production in last four decades. However, in several crops yields have reached a plateau, and further progress is not very significant. One of the main reasons for such a situation is that plant breeders tend to confine themselves to their working collection that consists largely of highly adapted material. It is through the use of this kind of resource that breeders have been able to maintain and, in some cases, steadily increase the yield potential over the decades (Evans 1983).

There are numerous examples where plant breeders have effectively exploited exotic germplasm by introgressing gene(s) for diseases resistance or single genes controlling other traits (Stalker 1980). However, the use of exotic germplasm in the improvement of quantitative traits is conspicuously rare even though the majority of breeding efforts are directed towards improving them. There are many reasons for the low use of diverse germplasm for improving quantitative traits in the adapted germplasm pool. Foremost among these is the supposition that such germplasm has little to offer the elite cultivars, or that it would require such an extended efforts to exploit it that the time and resource investment is not justified (Goodman 1985). Several efforts were begun, only to be abandoned when it became evident that the forthcoming gains would be difficult and slow (Hallauer 1978). The priority of breeders has been for achieving the demonstrable short-term gains rather than long-term germplasm development using exotic germplasm. Thus,

the large pool of variability instead of prompting more use has created a situation of not knowing where to begin. This has arisen due to an incomplete knowledge of germplasm accessions and the relationships among them, the unavailability of descriptive characters, and an uncertainty about the best evaluation methods for tapping the germplasm resources.

The development of core collections has been suggested as means to enhance the use of genetic resources in the improvement programs. A core collection is a subset of accessions from the entire collection that captures most of available genetic diversity of the species (Brown 1989a). The core subset can be evaluated extensively, and the information derived could be used to guide more efficiently the utilization of the entire collection (Tohme *et al*. 1995; Brown 1989b). Upadhyaya *et al*. (2001) developed a core subset of 1,956 accessions using the geographic distribution accessions and data on 13 quantitative traits: days to 50% flowering, plant height, plant width, days to maturity, basal primary branches, apical primary branches, basal secondary branches, apical secondary branches, tertiary branches, pods per plant, seeds per pod, seed yield, and 100-seed weight. This core represented 11.5% of the accessions and preserved the genetic variation available in the entire collection; it also preserved the coadapted gene complexes represented in the entire collection. In setting a core subset, the first issue is its size. Brown (1989a), using the sampling theory of selectively neutral alleles, argued that the entries in a core subset should be about 10% of the total collection, with a ceiling of 3,000 per species. This sampling is effective in retaining 70% of the alleles of the entire collection. In several crop species where the entire collection contains more than 80,000 accessions, for example, the rice germplasm collection at IRRI, a 10% core subset should contain more than 8,000 entries or 3,000 per species. This itself is a very large size to assess traits of economic importance, which often display genotype × environment interaction, and to identify useful parents through multilocation trials. Consequently, the main issue is how to reduce the size of a core collection further without losing species diversity. We discuss here a strategy for sampling the entire and core collections leading to a mini core subset in chickpea (*Cicer arietinum* L.) with the number of entries that although small captures most of the useful variation of the crop.

Materials and methods

The International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India, holds in trust the world's largest chickpea collection of 16,991 accessions from 44 countries. The experimental materials for this study comprised 1,956 chickpea core collection entries, consisting of 1,465 Desi, 433 Kabuli, and 58 Intermediate types. All 1,956 entries were planted in the Vertisols (Kasireddipally series-Isohyperthermic Typic Pellustert) field in the 1999/2000 postrainy season at the ICRISAT Center at Patancheru, India. Each treatment consisted of a 4-m row on ridge. The distance between rows was 60 cm and between plants 10 cm. Care was taken to ensure a uniform depth of planting. The experiment received 46 kg P_2O_5, 18 kg N ha^{-1} and 3 irrigations (7 cm water per irrigation). In each entry five competitive plants were selected randomly on which to record observations on plant height (cm), plant width (cm), number of apical primary branches, apical secondary branches, basal primary branches, basal secondary branches, and tertiary branches, number of pods per plant and seeds per pod, 100-seed weight (g), and plant yield (g). The observations on morphological descriptors, such as flower color, plant color, growth habit, seed color, seed shape, dots on seed testa, and seed testa texture, were recorded (IBPGR, ICRISAT, and ICARDA, 1993) on a whole-plot basis. Days to 50% flowering (days from sowing to the stage when 50% plants have begun flowering), days to maturity (from sowing to the stage when 90% pods have matured and turned yellow), flowering duration (days between 50% flowering and end of flowering in 50% plants), and plot yield were also recorded on a plot basis. The yield of five plants was added to determine plot yield.

A phenotypic distance matrix was created by calculating differences between each pair of accessions for each of the 22 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 was subjected to the hierarchical cluster algorithm of Ward (1963) at an R^2 (squared multiple correlation value) of 0.75 using 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 aglomerative procedure starts with 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. For sampling, the proportional strategy was used, and from each cluster approximately 10% of the accessions were randomly selected for constituting the ''mini core'' subset. At lease one accession was included even from those clusters which had less than 10 accessions.

The means of the core subset and the mini core subset were compared using the Newman-Keuls procedure (Newman 1939; Keuls 1952) for all the 22 traits. The homogeneity of variances of the core and mini core subsets was tested using 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 22 traits was analyzed using the chi-square test. The Wilcoxon (1945) rank-sum non-parametric test was performed using the SAS NPAR1WAY procedure (SAS 1989) to determine whether the mini core subset represents the core subset for each of the 22 traits. The medians of these traits of the core collection and mini core subset were compared using the SAS NPAR1WAY procedure (SAS 1989). The phenotypic correlations between different traits in the core and mini core were estimated independently, to know whether these associations, which may be under 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 the phenotypic diversity of each trait. The index was calculated independently in both the core collection and mini core subset to determine whether the diversity for each trait was retained in the mini core subset.

Results and discussion

The clustering procedure we used resulted in grouping the 1,956 core subset entries into 28 clusters. The number of core entries in the clusters ranged from 18 (0.9%) in cluster number 28 to 152 (7.8%) in cluster number 4. The procedure we used to develop the mini core subset

Table 1 Number and percentage (within brackets) of Desi, Kabuli, and Intermediate types of accessions in entire collection, core subset, and mini core subset of chickpea

Types	Entire collection	Core subset	Mini core subset
Desi	12,779 (75.5%)	1,465 (74.9%)	159 (75.4%)
Kabuli	3,528 (20.8%)	433 (22.1%)	44 (20.9%)
Intermediate	621 (3.7%)	58 (3.0%)	8 (3.8%)

Table 2 Means and variances for 22 traits recorded in the core subset and mini core subset of chickpea

Trait	Means			Variances			
	Core subset	Mini core subset	Significance	Core subset	Mini core subset	*F*-value	*P*
Days to 50% flower	62.9	62.2	NS[a]	81.9	78.3	0.222	0.638
Flower color	2.09	2.11	NS	3.02	3.09	0.041	0.840
Plant color	1.69	1.70	NS	0.264	0.24	1.548	0.214
Flowering duration (days)	41.0	41.2	NS	33.14	33.10	0.002	0.969
Growth habit	2.34	2.39	NS	0.609	0.716	1.959	0.162
Plant height (cm)	47.1	46.7	NS	72.05	70.81	0.027	0.869
Plant width (cm)	46.5	46.4	NS	39.5	32.7	2.165	0.141
Apical primary branches (no.)	1.61	1.79	NS	2.08	2.20	0.246	0.620
Apical secondary branches (no.)	6.52	6.55	NS	7.21	6.20	0.987	0.321
Basal primary branches (no.)	2.87	2.89	NS	0.54	0.56	0.165	0.685
Basal secondary branches (no.)	4.06	4.07	NS	2.71	2.63	0.107	0.744
Tertiary branches (no.)	2.30	2.28	NS	3.14	3.38	0.531	0.466
Days to maturity	116.1	116.0	NS	33.5	34.8	0.118	0.731
Pods per plant (no.)	81.7	83.3	NS	1,462.7	1,425.8	0.049	0.826
Seeds per pod (no.)	1.30	1.28	NS	0.097	0.094	0.082	0.774
Seed color	9.65	9.82	NS	28.6	31.7	1.752	0.186
Dots on seed testa	1.67	1.67	NS	0.22	0.22	0.000002	0.999
Seed shape	1.28	1.28	NS	0.26	0.28	0.291	0.590
Seed testa texture	1.29	1.28	NS	0.25	0.25	0.012	0.912
100-seed weight (g)	17.30	17.21	NS	65.3	58.8	0.291	0.590
Plant yield (g)	15.0	15.0	NS	35.8	29.8	1.508	0.220
Plot yield (kg ha^{-1})	1,632.0	1,663.8	NS	278,499	250,145	1.271	0.260

[a] NS, Non-significant differences at $P = 0.05$

Differences between means of core and mini core subsets were tested by Newman-Keuls test and variance homogeneity by Levene's test

resulted in the selection of 211 entries from the core subset. The mini core subset composition reflected the predominance of Asian entries in the core subset and entire collection. In the mini core subset the number of entries included were 172 (81.5%) from Asia, 26 (12.3%) from Africa, 7 (3.3%) from America, and 5 (2.4%) from Europe. This compared favorably with the number of accessions in the entire collection and core subset from Asia (14,393, 84.75% in entire collection; 1,579, 80.7% in core subset), Africa (1,436, 8.5%; 200, 10.2%), America (619, 3.6%; 87, 4.5%), and Europe (371, 2.2%; 60, 3.1%) (Upadhyaya *et al*. 2001). Southwest Asia and the Mediterranean region, which are two primary centers of diversity, accounted for 60 (34.9%) and 6 (3.5%) entries in the mini core subset, respectively. This compared favorably with the number of accessions in the entire collection (southwest Asia 5,540, 32.6%; Mediterranean 402, 2.4%) and core subset (588, 30.1%; 53, 2.7%). Ethiopia, which is secondary center of diversity of chickpea, was represented by 14 entries (6.6%) in the mini core, 120 entries (6.1%) in the core subset, and 928 (5.5%) in the entire collection, suggesting that this country was slightly underrepresented in the ICRISAT collection (Upadhyaya *et al*. 2001).

Three types of chickpea based on seed shape – desi, kabuli, and intermediate – are recognized. Desi types are angular-shaped, small-seeded, and dark-colored, whereas kabuli types are owl-shaped, large-seeded, and cream-colored; the intermediate types have pea-shaped seeds. The number of entries included in the mini core was 159 Desi (75.4%), 44 kabuli (20.9%), and 8 intermediate (3.8%) types. This corresponded very well with the number of desi (12,779, 75.5%), kabuli (3,528, 20.8%) and intermediate (621, 3.7%) types in the entire collection (Table 1).

Differences between the means of the core and mini core subsets were found to be non-significant for all traits (Table 2). The recorded MD% (0) indicated that the mini core subset represented the core collection well (Hu *et al*. 2000). The variances of the core and mini core subsets were homogeneous for all the traits (Table 2). There were no significant differences between the medians of the core and mini core subsets for any of the 22 traits (Table 3). Between 90.7% and 100% of the variation exhibited by the core collection was included in the mini core subset for flower color, plant color, flowering duration, growth habit, days to maturity, seed color, dots on seed testa, seed testa texture, seed shape, yield per

Table 3 Median, range, and coefficient of variation for 22 traitsin the core subset and mini core subset of chickpea

Trait	Median			Range		Coefficient of variation (%)	
	Core subset	Mini core subset	*P*	Core subset	Mini core subset	Core subset	Mini core subset
Days to 50% flower	62.0	62.0	0.308	27–90	33–82	14.38	14.21
Flower color	1.00	1.00	0.849	1–8	1–8	83.05	83.32
Plant color	2.00	2.00	0.635	1–3	1–3	30.38	28.73
Flowering duration (days)	41.0	41.0	0.926	21–71	24–71	14.05	13.96
Growth habit	2.00	2.00	0.453	1–5	1–5	33.31	35.35
Plant height (cm)	46.6	46.2	0.833	12.8–91.6	12.8–78.6	18.03	18.01
Plant width (cm)	46.2	45.8	0.312	26.8–72.2	31.2–66.6	13.52	12.33
Apical primary branches (no.)	1.33	1.67	0.043	0.00–7.67	0.00–6.67	89.61	82.96
Apical secondary branches (no.)	6.00	6.33	0.406	0.0–20.7	0.0–15.7	41.21	38.02
Basal primary branches (no.)	2.67	2.67	0.879	1.0–6.0	1.3–5.3	25.47	26.01
Basal secondary branches (no.)	4.00	4.00	0.254	0.0–12.0	0.3–8.7	40.52	39.82
Tertiary branches (no.)	2.00	2.00	0.501	0.0–9.7	0.0–7.0	77.04	80.69
Days to maturity	117.0	117.0	0.893	100–129	100–127	4.99	5.08
Pods per plant (no.)	76.7	78.3	0.724	9.3–292.7	13.3–247.7	46.79	45.34
Seeds per pod (no.)	1.20	1.20	0.391	1.0–2.6	1.0–2.2	23.86	23.97
Seed color	11.00	11.00	0.706	1.0–24.0	1.0–24.0	55.46	57.28
Dots on seed testa	2.00	2.00	0.999	1.0–2.0	1.0–2.0	28.23	28.29
Seed shape	1.00	1.00	0.884	1.0–3.0	1.0–3.0	39.92	41.25
Seed testa texture	1.00	1.00	0.768	1.0–3.0	1.0–3.0	38.99	39.07
100-seed weight (g)	14.5	14.6	0.803	7.4–64.2	8.3–57.2	46.72	44.54
Plant yield (g)	14.0	14.0	0.796	3.3–46.0	5.3–46.0	39.80	36.45
Plot yield (kg ha^{-1})	1625	1638	0.556	158–3129	388—3,083	32.34	30.06

plant, and yield per plot. In 9 out of remaining 11 traits the variation included ranged from 72.4% to 87.0% (Table 3). For both the apical and basal secondary branches the range variation included in the mini core subset was 69.4%. The high CR% retained in the mini core subset (87.9%) indicated that it was representative of the core collection. The coefficients of variation for all 22 traits were similar in the core and the mini core subsets (VR% = 98.5%).

The analysis of frequency distribution, except apical primary branches (P = 0.041) confirmed the homogeneity of the distribution (data not given). The Wilcoxon rank-sum test indicated that for all 22 traits both the core and mini core subsets have similar distributions (data not given). These results suggest that the mini core subset chosen is representative of the core collection, which in turn was representative of the entire collection (Upadhyaya *et al*. 2001). Thus, the variation contained in the entire collection of 16,991 accessions has been preserved in the mini core subset of 211 entries.

Phenotypic correlations were conducted between all 22 traits in the core subset and mini core subset independently. With more than 1,950 degrees of freedom a large number of correlation coefficients which have an absolute value greater than 0.10 were significant at P = 0.0001 in the core subset (Table 4). However, the proportion of variance in one trait that can be attributed to its linear relationship with a second trait is indicated by the square of the correlation coefficient (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, we found this type of meaningful relationship in the core collection between days to 50% flowering and flowering duration (r = –0.753), flower color and plant color (r = –0.761), flower color and seed shape (r = 0.782), flower color and dots on seed testa (r = –0.707), and pod number and plant yield (r = 0.745). In the mini core subset also, except between flower color and dots on seed testa (r = 0. 703), the magnitude of these correlations remained greater than 0.71.

An adequate and proper sampling, essential in developing a representative core collection, should consider the conservation of phenotypic associations arising out of coadapted gene complexes (Ortiz *et al*. 1998). This mini core subset preserves the phenotypic correlations observed in the core subset (Table 4). This clearly indicated that the selection of the mini core was adequate in this regard and that the coadapted gene complexes controlling these associations were sampled properly and adequately. Further, these relationships suggested that it is not necessary to measure all traits in future germplasm evaluations; only easily measurable traits – days to 50% flowering, flower color, and plant yield – need be used. Other relationships which did not meet the 50% criterion may be of interest to breeders. For example, an easily measurable trait like 100-seed weight is significantly associated with both pod number per plant (r = –0.458) and plot yield (r = 0. 416) in the mini core subset, suggesting that 100-seed weight could serve a useful purpose in choosing the accessions for further evaluations.

Table 4 Correlation coefficients between 22 characters in the core collection (above diagonal) and mini core (below diagonal) entries in the 1999/2000 post-rainy season at the ICRISAT Center

	DF	FLCL	PLCL	FD	GH	PLHT	PLWD	APB	ASB	BPB	BSB	TB	DM	PN	SDPD	SDCL	SDD	SDSH	SDT	SDWT	YPP	YKGH
DF[a]		0.241	–0.21	–0.753	–0.264	0.47	0.381	–0.231	0.028	0.201	0.218	0.06	0.626	–0.27	–0.068	0.142	–0.204	0.196	0.075	0.139	–0.223	–0.428
FLCL	0.304		–0.761	0.01	–0.167	0.291	0.262	–0.019	–0.045	0.043	–0.092	–0.062	0.383	–0.354	–0.239	0.202	–0.707	0.782	0.684	0.58	–0.09	–0.376
PLCL	–0.272	–0.793		–0.044	0.148	–0.308	–0.256	0.014	0.029	–0.041	0.106	0.076	–0.367	0.333	0.253	–0.264	0.508	–0.673	–0.547	–0.527	0.08	0.309
FD	–0.741	–0.018	–0.025		0.072	–0.207	–0.149	0.228	0.095	–0.155	–0.075	0.115	–0.145	0.071	–0.019	–0.094	–0.046	0.018	0.062	0.076	0.127	0.205
GH	–0.256	–0.135	0.124	0.092		–0.354	–0.172	–0.007	–0.211	–0.067	–0.196	–0.204	–0.366	0.179	0.028	0.023	0.206	–0.141	–0.089	–0.119	0.137	0.208
PLHT	0.485	0.214	–0.268	–0.205	–0.401		0.621	–0.002	0.039	0.162	0.071	0.05	0.468	–0.287	–0.16	0.09	–0.301	0.271	0.18	0.258	–0.159	–0.202
PLWD	0.362	0.254	–0.287	–0.066	–0.173	0.630		–0.026	–0.038	0.17	0.016	0.066	0.419	–0.201	–0.132	0.087	–0.267	0.234	0.183	0.256	–0.079	–0.082
APB	–0.256	–0.086	0.069	0.221	0.071	–0.050	0.025		0.168	–0.046	–0.054	0.122	–0.044	0.151	–0.017	–0.089	–0.028	–0.025	–0.01	0.039	0.192	0.106
ASB	–0.04	–0.049	0.032	0.129	–0.234	0.112	0.101	0.072		0.047	0.5	0.509	0.215	0.263	0.032	–0.065	–0.034	–0.021	–0.045	–0.075	0.244	0.08
BPB	0.277	0.093	–0.011	–0.185	–0.174	0.265	0.190	–0.150	–0.007		0.145	0.049	0.097	0.079	–0.001	0.067	–0.023	0.014	–0.009	–0.005	0.068	–0.024
BSB	0.143	–0.048	0.060	–0.004	–0.181	0.090	0.052	–0.169	0.589	0.128		0.528	0.287	0.275	0.083	–0.018	0.045	–0.062	–0.113	–0.11	0.245	0.109
TB	0.059	–0.022	0.038	0.123	–0.199	0.088	0.185	0.094	0.579	0.011	0.550		0.295	0.225	0.038	–0.08	–0.049	–0.029	–0.064	–0.07	0.216	0.069
DM	0.558	0.385	–0.381	–0.033	–0.320	0.411	0.408	–0.107	0.236	0.109	0.380	0.377		–0.299	–0.159	0.101	–0.36	0.324	0.189	0.306	–0.173	–0.389
PN	–0.267	–0.392	0.392	0.064	0.120	–0.249	–0.203	0.230	0.273	–0.020	0.311	0.184	–0.276		0.146	–0.005	0.326	–0.303	–0.244	–0.39	0.745	0.401
SDPD	–0.016	–0.256	0.296	–0.026	0.035	–0.041	–0.024	–0.017	0.107	0.046	0.091	0.077	–0.051	0.130		–0.089	0.125	–0.231	–0.233	–0.384	0.039	0.148
SDCL	0.128	0.238	–0.300	–0.054	0.046	0.010	–0.086	–0.105	–0.063	0.016	–0.038	–0.174	0.027	0.073	–0.096		0.173	0.139	0.174	0.117	0.012	–0.029
SDD	–0.303	–0.703	0.559	0.031	0.209	–0.320	–0.393	–0.007	–0.041	–0.098	0.001	–0.140	–0.409	0.416	0.122	0.193		–0.703	–0.604	–0.417	0.146	0.43
SDSH	0.204	0.713	–0.682	0.034	–0.144	0.237	0.297	–0.089	0.006	0.038	–0.027	0.055	0.300	–0.343	–0.234	0.105	–0.687		0.698	0.526	–0.059	–0.321
SDT	0.133	0.637	–0.553	0.009	–0.070	0.159	0.175	–0.062	–0.059	–0.008	–0.110	0.020	0.145	–0.254	–0.245	0.156	–0.594	0.651		0.491	–0.041	–0.224
SDWT	0.172	0.618	–0.590	0.044	–0.114	0.198	0.214	0.008	–0.112	0.020	–0.114	–0.020	0.284	–0.458	–0.383	0.057	–0.489	0.570	0.510		0.057	–0.304
YPP	–0.202	–0.077	0.095	0.128	0.079	–0.121	–0.034	0.292	0.238	–0.032	0.247	0.249	–0.176	0.732	0.053	0.076	0.191	–0.076	–0.017	–0.021		0.355
YKGH	–0.393	–0.357	0.361	0.129	0.087	–0.113	–0.029	0.109	0.141	–0.071	0.144	0.108	–0.389	0.416	0.064	–0.058	0.432	–0.258	–0.168	–0.274	0.384	

[a] DF Days to flower; FLCL, flower color; PLCL, plant color; FD, flowering duration (days); GH, growth habit; PLHT, plant height (cm); PLWD, plant width (cm); APB, apical primary branches (no.); ASB, apical secondary branches (no.); BPB, basal primary branches (no.); BSB, basal secondary branches (no.); TB, tertiary branches (no.); DM, days to maturity; PN, pods per plant (no.); SDPD, seeds per pod (no.); SDCL, seed color; SDD, dots on seed testa; SDSH, seed shape; SDT, seed testa texture; SDWT, 100-seed weight (g); YPP, plant yield (g); YKGH, plot yield (kg ha^{-1})

Table 5 Shannon-Weaver diversity index for 22 morphological and agronomic traits in the core and mini core subsets of chickpea

Trait	Core subset	Mini core subset
Days to 50% flower	0.611	0.629
Flower color	0.372	0.377
Plant color	0.324	0.300
Flowering duration (days)	0.619	0.587
Growth habit	0.249	0.288
Plant height (cm)	0.620	0.606
Plant width (cm)	0.625	0.635
Apical primary branches (no.)	0.623	0.512
Apical secondary branches (no.)	0.601	0.607
Basal primary branches (no.)	0.650	0.618
Basal secondary branches (no.)	0.495	0.622
Tertiary branches (no.)	0.536	0.511
Days to maturity	0.602	0.603
Pods per plant (no.)	0.612	0.613
Seeds per pod (no.)	0.443	0.417
Seed color	0.856	0.871
Dots on seed testa	0.276	0.276
Seed shape	0.284	0.289
Seed testa texture	0.286	0.281
100-seed weight (g)	0.505	0.493
Plant yield (g)	0.606	0.591
Plot yield (kg ha^{-1})	0.603	0.604
Average ± SE	0.519±0.034	0.515±0.034

The Shannon-Weaver diversity index (H`) was calculated to compare phenotypic diversity among characters in the core and mini core subsets. The index is used in genetic studies as a convenient measure of both allelic richness and allelic evenness, however, because of log transformation it is not readily interpretable in the genetic terms (Brown and Weir 1983). A low H` indicates an extremely unbalanced frequency classes for an individual trait and a lack of genetic diversity. The average H` in the mini core subset was similar to that of the core subset (Table 5), indicating that the diversity of the core was represented in the mini core subset.

The resources available for any evaluation of germplasm are limited and dwindling steadily. Therefore, extensive evaluations of an entire germplasm collection, or even a large core collection, would run into the thousands and be very expensive and difficult. This mini core subset (211 entries) which represents 10.8% of the 1,956 core subset entries but preserves the variation present in the core subset provides an easy approach to access the genetic resources. The chickpea core collection (1,956 entries) preserved the variation of the entire collection (16,991 accessions); the mini core subset (1.24% of entire collection) should therefore represent the total diversity contained in this entire collection. This mini core subset drastically reduces the number of entries to be evaluated and provides a working collection of chickpea 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 that will result in the enhanced use of genetic resources for improving quantitative traits. This mini core subset can be used for molecular characterization research, and the extent of diversity can be inferred for the entire collection.

References

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

Brown AHD (1989b) The case for core collections. In: Brown, AHD, Frankel, OH, Marshal, RD, Williams, JT (eds). The use of plant genetic resources. Cambridge University Press, Cambridge, UK, pp 136–155

Brown AHD, Weir BS (1983) Measuring genetic variability in plant populations. In: Tanksley A, Orton TJ (eds) Isozymes in plant genetic and breeding. Part A. Elsevier, Amesterdam, pp 219–238

Evans LT (1983) Raising the yield potential: by selection or design? In: Kosuge T, Meredith CP, Hollaender A, (eds) Genetic engineering of plants. Plenum Press, New York

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

Hallauer AR (1978) Potential of exotic germplasm for maize improvement In: Waldon DB (ed) Maize breeding and genetics. John Willey & Sons, New York, pp 229–247

Hu J, Zhu J, Xu HM (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, ICRISAT, and ICARDA (1993) Descriptors for Chickpea (*Cicer arietinum* L.). International Board for Plant Genetic Resources, Rome, Italy; International Crops Research Institute for the Semi-Arid Tropics, Patancheru, India; International Center for Agriculture Research in the Dry Areas. Aleppo, Syria

Johns MA, Skroch PW, Nienhuis J, Hinrichson P, Bascur G, Munoz-Schick C (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

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

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 R, Ruiz-Tapia EN, Mujica-Sanchez A (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 edn. SAS Institute, Cary, N.C.

Shannon CE, Weaver W (1949) The mathematical theory of communication. University of Illinois Press, Urbana

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

Snedecor GW, Cochran WG (1980) Statistical methods, 7th edn. Iowa State University Press, Ames

Stalker HT (1980) Utilization of wild species for crop improvement. Adv Agron 33:111–147

Tohme J, Jones P, Beebe S, Iwanaga M (1995) The combined use of agroecological and characterisation data to establish the CIAT *Phaseolus vulgaris* core collection. In: Hodgkin T, Brown AHD, van Hintum ThJL, Morales EAV (eds) Core collections of plant genetic resources. International Plant Genetic Resources Institute (IPGRI). John Wiley & Sons, New York, pp 95–108

Upadhyaya HD, Bramel PJ, Singh Sube (2001) Development of a chickpea core subset using geographical distribution and quantitative traits. Crop Sci (in press)

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