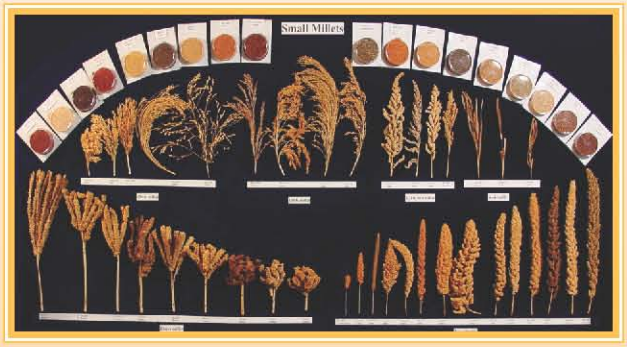
Technical Manual no. 10

Managing and Enhancing the Use of Germplasm – Strategies and Methodologies



ICRISAT
Science with a human face

International Crops Research Institute for the Semi-Arid Tropics

Technical Manual no. 10

Managing and Enhancing the Use of Germplasm – Strategies and Methodologies

Hari D Upadhyaya and CL Laxmipathi Gowda

ICRISAT
Science with a human face

International Crops Research Institute for the Semi-Arid Tropics

Patancheru 502 324, Andhra Pradesh, India
2009

Germplasm Utilization

Although there is an increase in the number of germplasm accessions in genebanks, there is no corresponding increase in their use by the crop improvement scientists, indicating that the collections were not being used to their full potential. Thus, a very large gap exists between availability and actual utilization of the materials. For example, very few of the 15,445 groundnut and 20,267 chickpea accessions conserved in the genebank have been utilized in cultivar development of these two crops at ICRISAT. Similarly, in the national programs, the number of germplasm lines used in breeding programs is very limited. Not only is the limited use of germplasm a worrisome issue, but also the large-scale deployment of a very few genotypes complicates the whole situation even more. Extensive use of fewer and closely related parents in crop improvement is contrary to the purpose of collecting a large number of germplasm accessions, and could result in vulnerability of cultivars to pests and diseases. The fears of epidemics similar to the southern corn leaf blight in the USA during 1970 (resulting in huge economic loss) and late blight of potato (that resulted in the famine in Europe during 1840s) due to a narrow genetic base of crop cultivars looms large even today.

Reasons for low use of germplasm are:

- Difficulty of crop improvement in handling large collections,
- Lack of information on a large number of accessions, particularly for traits of economic importance, which display large genotype × environment interactions and require multilocation and replicated evaluation,
- Limited capacity of breeding programs to absorb new material,
- Restricted access to the germplasm collections due to insufficient seed quantities, and
- Inadequate linkage between genebanks and the users of germplasm.

7A. Enhancing germplasm utilization

Utilization of germplasm can be enhanced by:

- Developing representative core and mini core collections to overcome the size related problems of collections,
- Identifying trait-specific germplasm for use in crop improvement programs,
- Multilocation evaluation of germplasm, including core and mini core collections, and organizing field days facilitating the selection of germplasm by crop improvement scientists (see section 10C),
- Developing trait-specific genepools in case of cross pollinating crops to provide partially converted populations to the breeders, and
- Ease of accessibility to all accessions of the collection by the users of germplasm.

7B. Core collections

The establishment of *ex situ* germplasm collections has been the result of several decades of efforts to conserve plant biodiversity. As collections rapidly grew beyond easily-manageable sizes, the task of quantifying diversity became daunting. Also, with the increase in size of collections, the realization that they are little used by breeders also grew.

7B.1. Concept of core collection

Frankel (1984) proposed sampling of the collections to a manageable sample or 'core collection'. A core collection consists of a limited set of accessions derived from a germplasm collection, which would 'represent, with a minimum of repetitiveness, the genetic diversity of a crop species and its relatives'. The accessions excluded from the core collection are retained as the reserve collection. Due to its reduced size, the core collection can be studied extensively and the information derived can be used to guide more efficient utilization of the much larger reserve collection.

Basic elements of core collection concept are:

- the original collection is large in size in view of management or use, but has taxonomic integrity;
- the core collection from this large original collection has a small size;
- the core collection is a representative sample of the collection; and
- the core collection retains sufficient diversity.

7B.2. Steps in setting up the core collection*:*

- Defining the collection to be represented in the core collection, and assembling all relevant data on these accessions
- Deciding the size of the core collection
- Grouping of accessions into groups that reflect the major genetic and ecological categories within the entire collection
- Selecting the core entries – how many from each group and which ones
- Representativeness of the core collection (in terms of diversity and inclusiveness)
- Managing the core collection for supply to potential users.

Defining the core collection and assembling data on accessions

It is important to ascertain what collection is to be used for developing a core collection. The core collection should serve as many users and uses as possible, and it should be comprehensive. The passport data on taxonomy, geographical origin, and ecological adaptation for each accession in the collection should be assembled. Characterization data on morphological traits, genetic markers and evaluation data (if available) should be assembled. All the available information should be used to develop the most representative core collection.

Size of core collection

The first decision to make in setting up a core collection is regarding its size. The size of the core (number of core entries) may be set by the resources available. In general, 10% of the parent collection is accepted as core collection. This proportion, in theory, should retain more than 70% of the alleles in the parent collection. For a very large collection, a core of 10% might still amount to a very large number of accessions to deal with, and hence suggest 3,000 as an upper limit for the size of a core collection. The reduced collection size will also help in reducing expenses for the genebank management.

Grouping of accessions

The grouping of accessions into categories of genetic similarity or commonality among accessions and determining groups in the entire collection is one of the most crucial steps. The hierarchy of grouping begins with the groupings suggested by taxonomy (species, subspecies, races), followed by assigning accessions to major geographic groups, climatic or agroecological regions. The accessions from larger countries can be divided into ecological regions; and those from small and adjacent countries can be grouped together. Clustering could be done within the broad geographical group to sort accessions into clusters using hierarchical clustering methods. At ICRISAT, we have used Ward (1963) method for clustering. This method optimizes an objective function because it minimizes the sums of squares within groups and maximizes the sums of squares between the groups.

Selecting the core entries

The number of accessions in different groups is likely to vary greatly. The accessions allocated to a cluster will share genetic affinity. Once the decision on size of the core is taken, the decision on the number of accessions from each cluster will depend on the strategy to be adopted. The following three strategies have been suggested to decide on the number of accessions from each cluster:

- *Constant strategy (C)* – equal number of accessions are sampled from each cluster into the core irrespective of the total number of accessions in different groups.
- *Proportional strategy (P)* – a fixed proportion of each group is selected for inclusion into the core collection, so that the group is represented in proportion to its frequency in the entire collection.
- *Logarithmic strategy (L)* – The number of accessions included into core are in proportion to the logarithm of the number of accessions in that cluster.

Representativeness of core collection

It is important that the core collection is representative of the entire collection, and its diversity needs to be assessed while setting up the core collection. Various parametric and non-parametric statistical methods can be used to compare the adequacy of core as a representative sample of the entire collection in terms of means, variances, frequency distributions, etc, between core and entire collections.

Managing the core collection

Managing the accessions included in a core collection is important so that it truly becomes a point of entry to the proper exploitation of genetic resources for crop improvement. The core accessions may be multiplied, conserved, evaluated and kept ready for dispatch to the researchers on short notice.

7B.3. Types of core collections

It has been generalized that a core collection is about 10% of the parent collection. However, the spectrum of germplasm diversity and relevance of the proposed core collection could vary depending on its relevance. Accordingly, these could be of three types: Global core collection, Regional core collection and Demand driven core collection.

Global Core Collection: The international agricultural research centers (IARCs) working under the CGIAR have the wider responsibility to assist the researchers globally, and therefore conserve germplasm from all geographical regions and countries. A core collection developed from these global germplasm collections could be termed as *Global core collection.* Sets of seven core collections of ICRISAT mandate crops (chickpea, groundnut, pigeonpea, sorghum and pearl millet) and two small millets (finger millet and foxtail millet) have already been developed. These core collections represent >80% diversity of the respective entire collections.

Regional Core Collection: Ecological environment differs from location to location, and so will differ adaptation of the germplasm accessions. For example, groundnut is a crop of worldwide importance and cultivated in over 90 countries. Considering its importance in Asia, it was decided to develop a core collection based on accessions of Asian origin, presuming that all accessions will carry higher level of adaptation and provide better options to scientists in the Asia region.

Demand Driven Core Collection: Some research programs may have special focus on developing a collection of particular biological/market type cultivars such as large seeded kabuli chickpea, Valencia type groundnut, etc, since researchers will find a specific collection to be more useful. Scientists at New Mexico State University (NMSU), USA and ICRISAT have developed a core collection of Valencia market type groundnut having 77 accessions representing germplasm conserved at NMSU and facilitating better focus in research on this type of groundnut in New Mexico and West Texas regions. Similarly, a *Guinea* core collection of 293 accessions was selected from 3,907 accessions using data on nine quantitative traits at ICRISAT.

7B.4. Global core collections developed at ICRISAT (Table 7A.1)

Chickpea: From the collection of 16,991 accessions (as on June 2000), a core collection was developed using data on country of origin and 13 quantitative traits through clustering by Ward's method and 10% of the accessions or a minimum of one accession were randomly selected from each cluster to constitute a core collection of 1,956 accessions. The validity of core accessions was tested using standard statistical parameters and the core collection was found to be a good representative of the entire collection.

Groundnut: A core collection was developed from the collection of 14,310 accessions using data on taxonomic affiliation, geographical origin and 14 morphological traits. Similar to the case of chickpea, clusters were formed and 10% of the accessions were picked from each cluster to constitute a core collection of 1,704 accessions which was 11.9% of the parent collection. On testing the validity of core accessions, it was found to be adequate.

Pigeonpea: The pigeonpea core collection was constituted based on the data on geographical origin and 14 qualitative traits and comprised 1,290 accessions. The core accessions were derived from the entire collection of 12,153 pigeonpea accessions in the ICRISAT genebank representing 56 countries. The statistical estimates on core as well as the entire collection indicated that the core developed is a good representative of the parent collection.

Sorghum: To develop a sorghum core collection from 22,473 landraces in the ICRISAT genebank, accessions were stratified in four clusters based on four classes of photoperiod sensitivity: photoperiod insensitive, mildly sensitive, sensitive and highly sensitive. The landrace accessions were classified into 60 groups as results of the combinations of 15 basic and intermediate races and four photoperiod groups. From each group following a logarithm strategy of sampling, a core collection of 2,247 accessions was constituted.

Pearl millet: The pearl millet core collection was developed based on the data on geographical origin and 11 quantitative traits. This was derived from the entire collection of 16,063 representing 25 countries. The core collection contained 1,600 accessions, which is 10% of the entire collection. The estimates of various statistical parameters revealed that the core collection was a fairly good representative of the entire collection.

Finger millet: A core collection of finger millet was developed from 5,940 accessions held in the ICRISAT genebank. These accessions represented 23 countries. The core collection was constituted using the data on geographic origin and 14 quantitative traits and contained 622 accessions, which was 10.47% of the entire collection.

Foxtail millet: A core collection of foxtail millet was developed from 1,474 accessions held in ICRISAT genebank from 23 countries. The core collection was constituted using data on geographic origin and 12 qualitative traits and contained 155 accessions, which accounted for 10.52% of the entire collection.

7B.5. Augmenting the core collection

Core collections are dynamic and need updating/augmenting when new accessions or information becomes available. For example, a core collection of pearl millet (1,600 accessions) was developed in 1998 using available data for 11 agronomic traits on 16,063 accessions. Meanwhile, by the year 2007, a total of 4,717 germplasm accessions were assembled additionally and characterized. This necessitated augmenting the core collection. For this, a phenotypic distance matrix was created for 4,717 accessions by calculating differences between each pair of accessions for each of the 22 (10 morphological and 12 quantitative) traits. The diversity index was calculated by averaging all the differences in the phenotypic values for each trait divided by the respective range. This distance matrix

was subjected to hierarchical cluster algorithm of Ward (1963) at an R^2 (squared multiple correlation value) of 0.75. 10% of the accessions or a minimum of one from each cluster was randomly selected to form a representative sample of 4,717 accessions. Various statistical analyses indicated that the selected sample not only represented the 4,717 accessions, but when added to the core collection, the augmented core collection (2,094 accessions) represented the entire collection (20,844 accessions).

7C. Mini core collections

Most often, the germplasm collections at IARC genebanks are very large. Thus, the number of core accessions will be too high for meaningful replicated evaluations and evaluation at different locations. To overcome this, Upadhyaya and Ortiz (2001) postulated the concept of mini core collection. A mini core collection consists of 10% accessions of the core collection, and hence only 1% of the entire collection. This mini core collection still represents the diversity of the entire collection. In fact, constituting a mini core collection is a two-stage strategy. The first stage involves developing a representative core collection (about 10%) from the entire collection using available information on origin, characterization and evaluation data of the accessions. The second stage involves evaluation of the core collection for various morphological, agronomic, and grain quality traits; and then selecting a further core of about 10% accessions. At both the stages, standard clustering procedures are used to cluster groups of similar accessions. Following this strategy, mini core collections of chickpea, groundnut, pigeonpea and sorghum have been developed; and development of mini core of pearl millet and finger millet is in progress.

Global mini core collections developed at ICRISAT (Table 7A.1)

Chickpea: The chickpea core collection of 1,956 accessions was evaluated in the 1999-2000 postrainy season at ICRISAT, Patancheru, India. The data was recorded on 16 quantitative and qualitative traits. A phenotypic distance matrix was created by calculating differences between each pair of accessions for each of 22 traits following Johns et al. (1997). The distance matrix was subjected to the hierarchical cluster algorithm of Ward (1963) and 28 clusters were formed. From each of the 28 clusters, 10% accessions were randomly selected, which resulted in the formation of global mini core collection of chickpea germplasm of 211 accessions, which was 10.8% of the core collection and 1.24% of the entire collection.

Groundnut: To constitute a mini core collection of groundnut, the core collection comprising 1,704 accessions was evaluated during the 1999 rainy season for 13 qualitative and 16 quantitative traits and in the 1999-2000 postrainy season for 18 quantitative traits at ICRISAT, Patancheru, India. A phenotypic distance matrix was created by calculating differences between each pair of accessions for each trait following Johns et al. (1997). The distance matrix was subjected to the hierarchical cluster algorithm of Ward (1963) and 77 clusters were formed. From each cluster, 10% accessions were randomly selected, which resulted in the formation of global mini core collection of groundnut germplasm of 184 accessions, which was 10.8% of the core collection and 1.29% of the entire collection.

Pigeonpea: The core collection of 1,290 accessions was evaluated at ICRISAT, Patancheru, India for 18 qualitative and 16 quantitative traits. A phenotypic distance matrix was created by calculating differences between each pair of accessions for each of 34 traits following Johns et al. (1997). The distance matrix was subjected to the hierarchical cluster algorithm of Ward (1963) and 79 clusters were formed. From each cluster, 10% accessions were randomly selected, which resulted in the formation of global mini core collection of pigeonpea containing 146 accessions, which was 11.3% of the core collection or 1.20% of the entire collection.

Sorghum: To develop a mini core collection of sorghum, the core collection of 2,246 accessions was evaluated in 2004-05 postrainy season (October-April) at ICRISAT, Patancheru, India, for 11 qualitative and 10 quantitative traits. As was the practice for the other crops, a phenotypic distance matrix of 2,246 accessions for each of the 21 traits was created. The distance matrix was subjected to the hierarchical cluster algorithm of Ward (1963) and 21 clusters were formed. From each cluster, 10% accessions or a minimum of one were randomly selected, which resulted in the formation of global mini core collection of sorghum, which had 242 accessions. The mini core strength was 11.3% of the core collection or 1.20% of the entire collection.

Table 7A.1. Core and mini core collections of ICRISAT mandate crops.

Crop	Number of accessions used	Number of traits involved	Number of accessions in subset	Reference
Core collections				
Sorghum	33,100	7	3,475	Prasada Rao and Ramanatha Rao 1995
	22,473	2[1]	2,246[2]	Grenier et al. 2001a, b
Pearl millet	16,063	11	1,600	Bhattacharjee et al. 2007
	20,766	12	2,094	Upadhyaya et al. 2009a
Chickpea	16,991	13	1,956	Upadhyaya et al. 2001a
Pigeonpea	12,153	14	1,290	Reddy et al. 2005
Groundnut	14,310	14	1,704	Upadhyaya et al. 2003
Groundnut (Asia)	4,738	15	504	Upadhyaya et al. 2001c
Finger millet	5,940	14	622	Upadhyaya et al. 2006c
Foxtail millet	1,474	12	155	Upadhyaya et al. 2008
Mini core collections				
Sorghum	2,246[2]	21	242	Upadhyaya et al. 2009b
Pearl millet	2,094	18	238	ICRISAT unpublished data
Chickpea	1,956	22	211	Upadhyaya and Ortiz 2001b
Pigeonpea	1,290	34	146	Upadhyaya et al. 2006d
Groundnut	1,704	31	184	Upadhyaya et al. 2002

1 Photoperiod response as measured by flowering and plant height was used to group the accessions and 19 traits used to validate core.

2 One accession has been denotified, therefore, the core is of 2,246 accessions.

7D. Identification of new sources

The core and mini core collections of various crops were evaluated to identify trait-specific diverse parents. Due to the reduced size, the core and the mini core sets have been evaluated and characterized precisely and useful trait-specific accessions have been identified for use in breeding programs to develop cultivars with a broad genetic base.

- *Drought tolerance*: 18 new sources of drought tolerance have been identified in groundnut and chickpea. These new sources are similar or better than the known sources for drought resistance, and are superior or similar for the agronomic traits.
- *Salinity tolerance*: 12 tolerant sources in chickpea and 16 in pigeonpea were identified.
- *Diseases resistance in chickpea*: 67 accessions resistant to wilt; 6 resistant to dry root rot; 3 tolerant to Ascochyta blight; 55 tolerant to Botrytis gray mold disease and 18 accessions for multiple resistance.
- *Early-maturity in groundnut and chickpea*: 21 diverse landraces of groundnut, which are similar to the earliest maturing Chico, but have high yield and better pod and seed traits in groundnut. Similarly, 28 new diverse sources of early maturity in chickpea, which mature as early as earliest maturing germplasm Harigantars (85-90 days) but produce up to 23% more yield.
- *Productivity traits in groundnut, chickpea and pigeonpea*: A number of high-yielding sources from the Asia region core collection in different botanical varieties (20 Spanish, 15 Valencia and 25 Virginia). Similarly, high-yielding and diverse sources have been identified in chickpea and pigeonpea.
- *Large-seeded kabuli chickpea*: Sixteen diverse germplasm lines, which have 100-seed weight up to 55 g compared to 20 g of the popular Indian cultivar L 550 have been identified. Scientists at Indian Institute of Pulses Research, Kanpur, India have identified 12 accessions for large scale evaluation and five accessions for breeding large seeded kabuli cultivars.
- *Agronomic traits in Thailand and China*: Thai scientists have identified five accessions each for high pod yield, shelling percentage and seed size. Similarly, scientists in China have identified five accessions with large seed size, 14 accessions with resistance to bacterial wilt, and four accessions with high oleic and low linoleic acid content.

7E. Composite collections

The revolution in molecular biology, bioinformatics and information technology has provided the scientific community with tremendous opportunities to address some of the world's most serious agricultural and food security issues. ICRISAT in collaboration with the Generation Challenge Program (GCP) on "Unlocking Genetic Diversity in Crops for the Resource-Poor (www.generationcp.org)" has constituted composite collections of chickpea, sorghum, groundnut, pigeonpea, finger millet and foxtail millet that encompass the crop diversity. Phenotypic and genotypic characterization of these sets will provide opportunities and scope of identifying useful and unique germplasm resources for utilization in crop improvement. The composite collections have been genotyped using SSR markers to study genetic diversity, population structure and select a reference set of 200-400 most diverse accessions for research use (see section 10B).