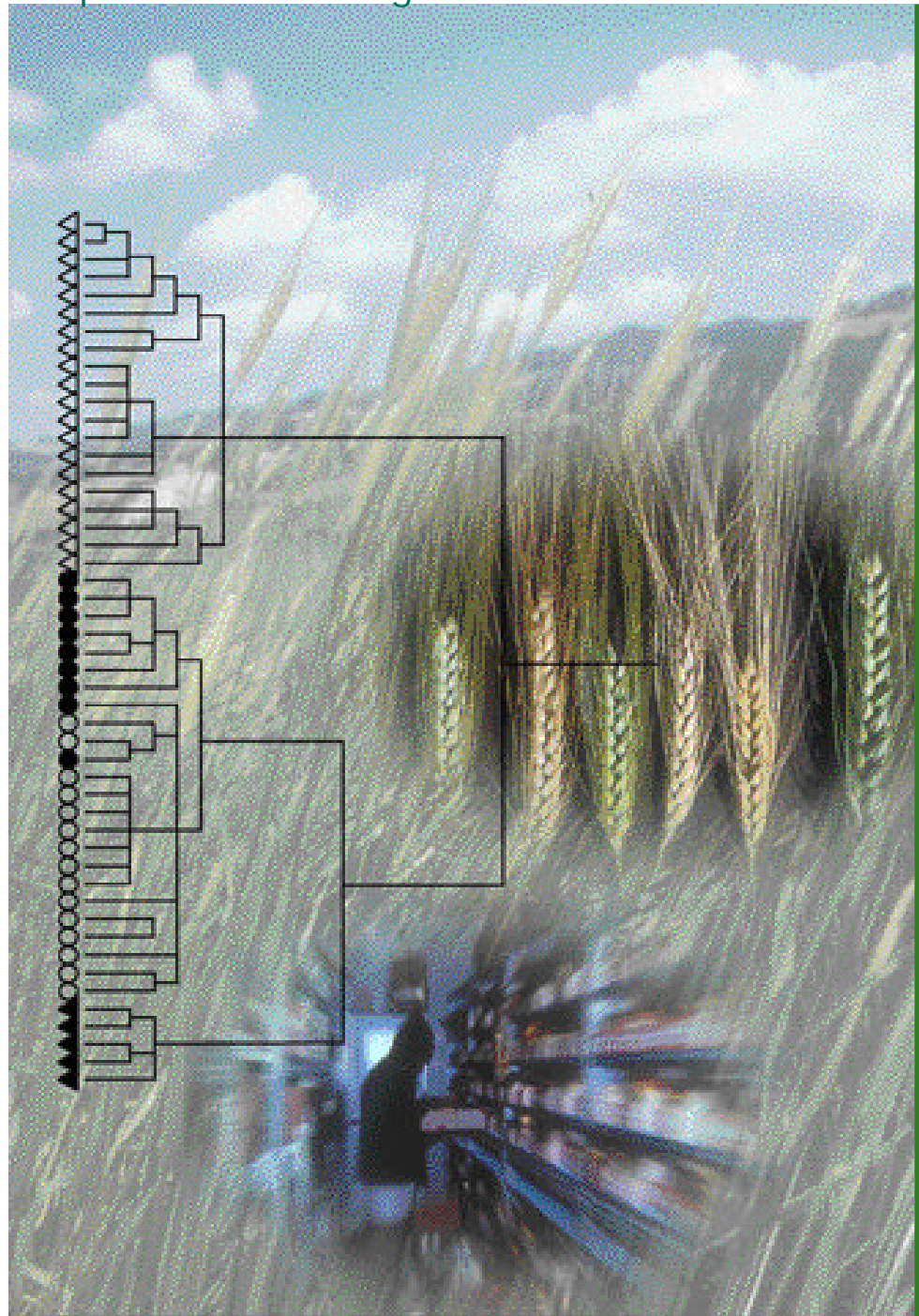





Core collections of plant genetic resources

Th.J.L. van Hintum, A.H.D. Brown,
C. Spillane and T. Hodgkin






IPGRI Technical Bulletins are published by the International Plant Genetic Resources Institute with the intention of putting forward definitive recommendations for techniques in genetic resources. They are specifically aimed at National Programme and genebank personnel.



Core collections of plant genetic resources

Th.J.L. van Hintum, A.H.D. Brown,
C. Spillane and T. Hodgkin



The **International Plant Genetic Resources Institute** (IPGRI) is an autonomous international scientific organization, supported by the Consultative Group on International Agricultural Research (CGIAR).

IPGRI's **mandate** is to advance the conservation and use of genetic diversity for the well-being of present and future generations. IPGRI's headquarters is based in Rome, Italy, with offices in another 19 countries worldwide. It operates through three programmes: (1) the Plant Genetic Resources Programme, (2) the CGIAR Genetic Resources Support Programme, and (3) the International Network for the Improvement of Banana and Plantain (INIBAP).

The **international status** of IPGRI is conferred under an Establishment Agreement which, by January 2000, had been signed and ratified by the Governments of Algeria, Australia, Belgium, Benin, Bolivia, Brazil, Burkina Faso, Cameroon, Chile, China, Congo, Costa Rica, Côte d'Ivoire, Cyprus, Czech Republic, Denmark, Ecuador, Egypt, Greece, Guinea, Hungary, India, Indonesia, Iran, Israel, Italy, Jordan, Kenya, Malaysia, Mauritania, Morocco, Norway, Pakistan, Panama, Peru, Poland, Portugal, Romania, Russia, Senegal, Slovakia, Sudan, Switzerland, Syria, Tunisia, Turkey, Uganda and Ukraine.

The geographical designations employed and the presentation of material in this publication do not imply the expression of any opinion whatsoever on the part of IPGRI or the CGIAR concerning the legal status of any country, territory, city or area or its authorities, or concerning the delimitation of its frontiers or boundaries. Similarly, the views expressed are those of the authors and do not necessarily reflect the views of these participating organizations.

Citation: van Hintum, Th.J.L., A.H.D. Brown, C. Spillane and T. Hodgkin. 2000. Core collections of plant genetic resources. IPGRI Technical Bulletin No. 3. International Plant Genetic Resources Institute, Rome, Italy.

Cover: illustration by Paul Neate, IPGRI.

ISBN 92-9043-454-6

IPGRI
Via delle Sette Chiese, 142
00145 Rome
Italy

© International Plant Genetic Resources Institute, 2000

Introduction to the Series

The Technical Bulletin series is targeted at scientists and technicians managing genetic resources collections. Each title will aim to provide guidance on choices while implementing conservation techniques and procedures and in the experimentation required to adapt these to local operating conditions and target species. Techniques are discussed and, where relevant, options presented and suggestions made for experiments. The Technical Bulletins are authored by scientists working in the genetic resources area. IPGRI welcomes suggestions of topics for future volumes. In addition, IPGRI would encourage, and is prepared to support, the exchange of research findings obtained at the various genebanks and laboratories.

Contents

Acknowledgements	5
1 Introduction	6
1.1 Why have a core collection?	6
1.2 What is a core collection?	7
2 Establishing a core collection	9
2.1 Identifying the material to be represented	10
2.2 The size of the core collection	11
2.3 Division into genetically distinct groups	13
2.4 The number of entries per group	16
2.4.1 <i>Procedures based on group size</i>	17
2.4.2 <i>Procedures based on marker diversity</i>	19
2.4.3 <i>Procedures based on informal knowledge</i>	20
2.5 Choosing the entries	21
2.5.1 <i>Random and systematic procedures</i>	21
2.5.2 <i>Analytical procedures</i>	22
2.5.3 <i>Pragmatic procedures</i>	22
3 Managing the core collection	24
3.1 Maintaining the core collection	24
3.2 Distributing the core collection	26
3.3 Information management	27
3.4 Validating the core collection	28
3.5 Altering the core collection	30
4 Using the core collection	32
4.1 Improving genebank operations	32
4.2 Improving the use of plant genetic resources	34
4.2.1 <i>Applied research: Screening for useful traits and characteristics</i>	36
4.2.2 <i>Basic research and education</i>	37
4.3 Increasing the use of core collections	38
5 Future developments	40
References	43

Acknowledgements

The authors of this volume wish to acknowledge the valuable assistance given by a number of genebank managers and IPGRI staff in the preparation of this volume through their encouragement and helpful comments. The support of CSIRO Plant Industry, Canberra, Australia, and the Centre for Genetic Resources (CGN), Plant Research International, Wageningen, the Netherlands is also gratefully acknowledged. Particular thanks are also due to Imke Thormann, Silvana Bonas and Carole Salas for their help in supporting the work of preparing this Bulletin.

1 Introduction

Genebanks around the world hold collections of the genetic resources of crop plants for long-term conservation and for ease of access by plant breeders, researchers and other users. The last 25 years have seen remarkable progress in assembling and conserving these resources. Indeed, many plant germplasm collections now face major problems of size and organization. Some collections have grown so large as to hinder the very purposes for which they exist, namely, the conservation and the use of the genetic diversity they hold.

Perceiving that the large size of some collections could deter use, Frankel (1984) proposed that a limited or “core collection” could be established from an existing collection. With minimum similarity between its entries¹ the core collection is of limited size and chosen to represent the genetic diversity of a large collection, a crop, a wild species or group of species. It does not replace the existing collection or material from which it is obtained.

Since Frankel put forward his proposal, a body of literature on the theory and practice of core collections has appeared including many examples of the approach. Core collections have become accepted as efficient tools for improving conservation and use of collections. The Global Plan of Action for the Conservation and Sustainable Utilization of Plant Genetic Resources for Food and Agriculture (FAO 1996) recommends core collection development as one of the activities needed to improve use of plant genetic resources. This technical bulletin sets out the procedures that can be used to establish, manage and use a core collection, drawing on the accumulated experience so far.

1.1 Why have a core collection?

At the outset, it is helpful to ask what the main issues and problems in the management of a genebank are – in the work it does on conservation, regeneration, duplication, documentation, evaluation and its use. Specific questions that often face the manager include:

- Are some activities of the genebank made difficult because of the large size of the collection?

¹ For clarity in this bulletin, we will use the term accession to refer to a sample maintained in the whole collection, and the term entry to refer to any accession or subsample chosen to be in the core collection.

- Are genebank activities and the use of the collection limited by lack of knowledge of the way in which genetic diversity is distributed in the collection?
- Is it difficult to decide what the priorities are, whether gaps exist or when new material should be added to the collection?
- Does the size of the collection mean that users are unaware of variation in the collection; variation that might perform well in their environments, or benefit their breeding programmes, or enrich their research projects?

If the answer to any of these questions is “yes”, then an attractive option for meeting the challenges is to form a core collection. A core collection provides a structured sample from the collection; one that is a more manageable size than the whole collection. Its structure is such as to represent the diversity of the collection. It forms a reference set, and, when choices have to be made, an automatic priority for attention.

1.2 What is a core collection?

Since its inception, the core collection has been variously interpreted and implemented. Box 1 gives three definitions. For simplicity this bulletin uses the one that attaches to a specific crop collection in a genebank. This definition readily extends to a collection that includes a group of related species, or to one that is the aggregate of several collections of the same taxa held in a network of cooperating genebanks.

The word “core” suggests the central or innermost part, the heart and the most important part. This idea makes most sense when the core stands as a reference point to an identified set of material, most commonly a collection. In this way the link between the core and the whole collection remains firmly in place so that the core provides efficient access to the resources of the whole collection. The core collection idea also carries with it the aim that the set of entries it contains has a stand-alone function. Ideally, the core will also provide a focus for evaluation where information on a growing set of variables can be obtained and assessed on a structured and limited set of accessions. In this way studies on the core collection provide an overall view of the properties to be found in the whole collection.

A core collection will always be substantially smaller than the collection from which it is formed. Brown (1989b) suggested that it should be no more than 10% of the whole collection and always less than 2000 entries. In practice, most core collections are between 5% and 20% of the collections from which they were established and the largest to date is about 2000 accessions.

Box 1. Definitions of the core collection

The original definition:

- A core collection is a limited set of accessions representing, with a minimum of repetitiveness, the genetic diversity of a crop species and its wild relatives (Frankel 1984).

From this definition, two operational definitions have followed:

- For an individual genebank, a core collection consists of a limited number of the accessions in an existing collection, chosen to represent the genetic spectrum in the whole collection. It should include as much as possible of its genetic diversity (Brown 1995).

Such a set of a set of accessions has also been called a "core subset" of the whole collection.

- For a whole crop species, a core collection consists of a limited number of entries chosen to represent the genetic diversity of the whole crop species and its wild relatives. It is a synthetic and comprehensive core collection, assembled cooperatively by national and international genebanks and supplemented with fresh samples of wild or crop populations where needed to fill gaps. The best example of such a core is the international Barley Core Collection.

The overall objective is to get a collection conserved better and used more effectively. It would be counterproductive to have a core so large that it suffers the same problems as the whole collection. On the other hand, a core that failed to contain a significant fraction of the whole collection's diversity would not serve its purpose.

Frankel's proposal of core collections some 15 years ago led to some debate over what a core collection is, what its advantages and disadvantages are and whether it can be modified and used in other ways (Brown 1995). This bulletin takes the simplest and most straightforward approach to forming a core collection and leaves open to genebank managers how they might modify, elaborate or repeat the steps. Section 2 of this bulletin describes the methods employed to form a structured sample that is a core collection and Section 3 considers the questions that arise in managing a core collection. Section 4 concentrates on uses of the core collection, both in managing the whole collection and how clients of the genebank might use the core to improve their programmes. Finally, in Section 5 we outline some likely developments in the establishment and use of core collections.

2 Establishing a core collection

Creating a core collection can be very simple. It can be done for any germplasm collection. It does not require complete documentation or fully reliable data. There is no need to have data on genetic markers or to have specialist knowledge of mathematics. All that is needed is a germplasm collection, someone with some basic knowledge about the collection and the species involved, and some time for selecting the core. Box 2 gives some simple approaches to sampling a collection to illustrate how simple the process can be. However simple or complex a procedure is followed, it is always worth ensuring that there is consultation between genebank managers, plant breeders and other research workers interested in the crop and the use of its genetic diversity.

A general procedure for the selection of a core collection can be divided into five steps, which are described in the following sections (see also Fig. 1).

1. Identify the material (collection) that will be represented.²
2. Decide on the size of the core collection.
3. Divide the set of material used into distinct groups.
4. Decide on the number of entries per group.
5. Choose the entries from each group that will be included in the core.

Each of these steps can be more or less complex depending on the information available and the procedures used. Sometimes an iterative process is adopted in which some or all of the steps are repeated with a decreasing core size. Thus, a first cycle might create a preliminary core of, say, twice the expected final size of the core collection. The material in this selection is further characterized and, with the help of the additional data, the preliminary core is reduced to its final size.

Box 2. Simple ways of sampling collections

Even a random selection from a germplasm collection can be used as a core collection. The simplest and almost the least efficient sampling process for a core would still probably give a better representation of the genetic diversity than a sequential set such as the accessions with numbers 8201 to 8485. A simple improvement to this would be regular or systematic sampling by accession number, for example for a core collection of 10% which was obtained by including all accessions with accession numbers ending with a zero. However, none of these is satisfactory since they make no use of any information on the accessions. Using such information will always improve the representativeness of the core collection.

² Van Hintum (1998, 1999) describes the collection or set of material used for forming a core as the “domain”.

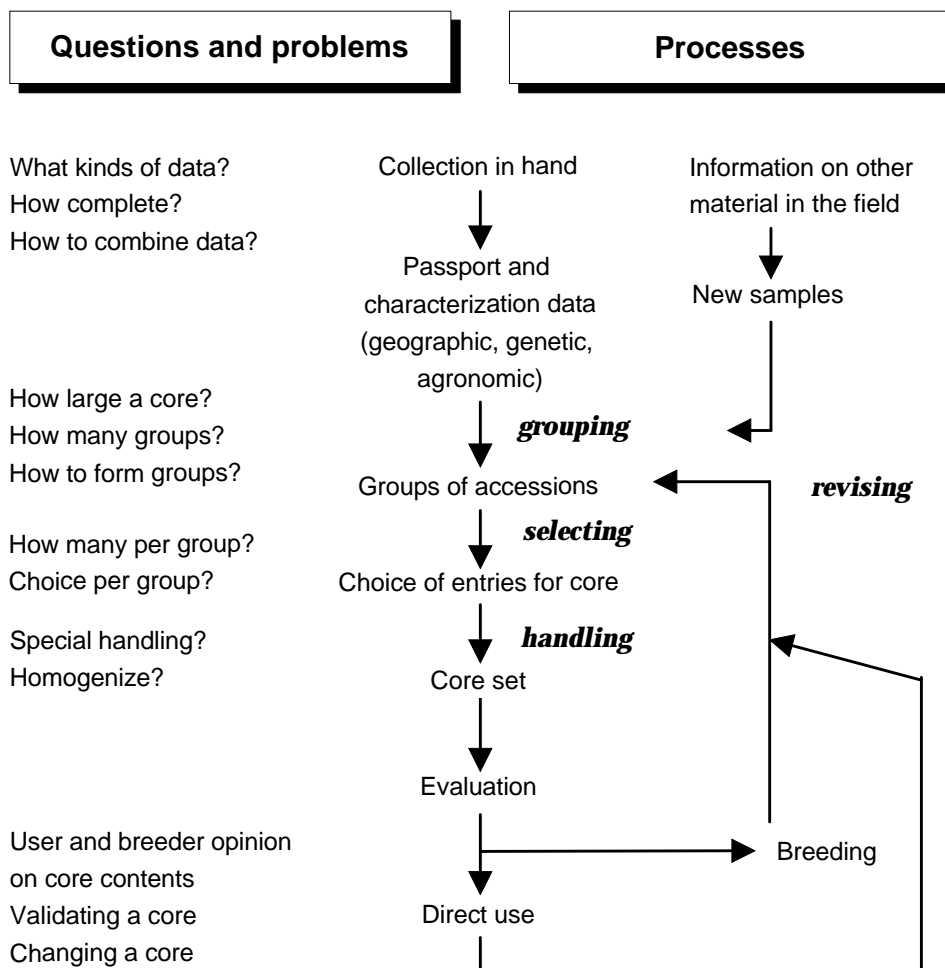


Fig. 1. Flow chart to illustrate steps in developing a core collection.

2.1 Identifying the material to be represented

The material or collection that is to be represented by the core will obviously differ from one case to another. It is dependent on the material that is available, on what constitutes a sensible set of material for core development, and on the objectives behind the establishment of a core collection.

Most often a core collection will aim to represent all material of a certain crop in a genebank collection. Examples of this include the cores of the CSIRO (Commonwealth Scientific and Industrial Research Organisation, Australia) collection of perennial *Glycine* (Brown *et al.* 1987), the complete US germplasm collection of peanut (Holbrook *et al.* 1993) and that of the Peruvian quinoa collection (Ortiz *et al.* 1998). Other core collections have been established to include the material of a

certain crop in several collections, such as the core of cultivated *Brassica oleracea* in European collections (Boukema *et al.* 1997). The objective may extend to all accessions of a crop and its wild relatives in any collection such as the international Barley Core Collection (Knüpfner and van Hintum 1995).

In some cases, cores have been established to represent only a part of a collection, such as the core of sesame landraces from India (Bisht *et al.* 1998) or of lentil accessions from Chile, Greece and Turkey (Erskine and Muehlbauer 1991). In another case, Diwan *et al.* (1994) developed a core from a 40% cross-section of the USA annual *Medicago* collection.³

This bulletin deals largely with the formation of core collections from individual genebank collections. However, where representation of diversity is the objective, the methods described can be readily extended to any set of germplasm accessions.

2.2 The size of the core collection

After defining what material the core collection is to represent, the next step is to decide on its size; that is, how many entries it will contain. Given the objectives of the core collection, its size will be much less than the source collection. Most core collections described so far are an order of magnitude smaller than the collection from which they came. Thus, most of the core collections surveyed by Spillane *et al.* (unpublished) were 5–20% of the size of the collection from which they were established (Fig. 2).

When the collection from which the core is developed is very large, the percentage size may be much smaller than 5%. The international Barley Core Collection (1600 accessions) is less than 0.3% of the world barley holding and the ICRISAT (International Crops Research Institute for the Semi-Arid Tropics, Patancheru, India) sorghum core collection of approximately 600 accessions was formed from a collection of 40 000 accessions (1.5%). In terms of absolute numbers, the largest core collection is about 2500 accessions. This is a multi-

³ Some “core” collections or sets have been described which are restricted to material with specific traits, such as local maize populations with good combining ability (Radovic and Jelovac 1994) or *Pisum sativum* germplasm with disease resistance (Matthews and Ambrose 1994). These may be better described as character core sets since the representation of diversity is restricted to specific characters. In these cases, the sets were obtained from whole genebank collections.

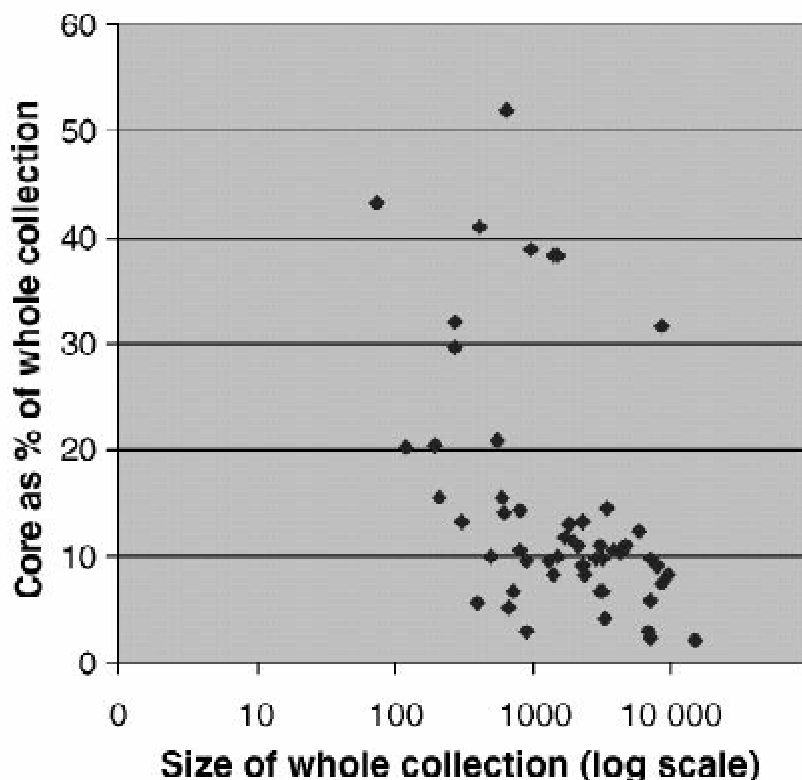


Fig. 2. Percentage core versus overall collection size for 37 core collections of seed-propagated crops (Spillane *et al.*, unpubl.).

species core of wild and cultivated *Solanum* species and the core is about 31% of the collection from which it was developed (Spillane *et al.*, unpubl.).

Clearly, no single size or fixed proportion for the core will be appropriate for all cases. A number of different studies provide some general guidelines that can be useful. Thus, Brown (1989b) described the consequences of simple random sampling from either theoretical gene frequency profiles or from empirical examples. These analyses indicate that a core of about 10% of the total collection was likely to contain at least 70% of the variation in the whole collection. In practice, the proportion of variation captured in a core of 10% is likely to be higher because of the division of the collection into genetically meaningful groups (see section 2.3 below). For example, the sorghum core collection developed at ICRISAT contains less than 3% of the collection but is reported to contain over 90% of the variation in the whole collection (Paula Bramel, pers. comm.).

Yonezawa *et al.* (1995) concluded that the optimum sample fraction depends largely upon the degree of genetic redundancy

among accessions, the resources available for maintenance of core entries and the frequency of regeneration of the entries. A single optimum could not be identified but a proportion of 20–30% was considered best under certain circumstances. Charmet and Balfourier (1995) and Bisht *et al.* (1998) have analyzed size and grouping strategies (see sections 2.3, 2.4) and found that sizes of 5–10% were optimal, capturing 75–90% of the diversity. In contrast, Noirot *et al.* (1996) have suggested that higher percentages (20–30%) are needed, particularly where the objective is to capture the genetic diversity of quantitatively inherited characters.

In addition to the genetic aspects described above, a number of other considerations will guide the curator when deciding the size of the core collection. Where there are very many small groups, as in the case of multi-species core collections, a higher percentage of entries might be taken from each group. This is often the case where it is necessary to include every group and many groups have only one or two accessions in them.

The resources available for maintaining the collection will be important as will the method used for maintaining it. For example, the area of the field available for a clonal genebank might place a strict absolute limit on the size. Alternatively, a core collection maintained under *in vitro* slow growth conditions will need suitable tissue culture and incubation facilities and this may affect the size of core chosen. The needs of users also will be important. Core collections that are too large for the intended users are unlikely to meet the objectives of establishing them. In these cases it might be preferable to have a smaller, better-used core than a larger one which is seldom requested.

2.3 Division into genetically distinct groups

An effective representation of the genetic diversity of the whole collection in a core depends on first separating the accessions into meaningful groups. This is often called stratification. The groups should be constructed so as to maximize variation between groups and minimize variation within groups. The way in which this is done is a key part of developing a good core collection. There are a number of different approaches which have been tested either singly or in combination. The optimum choice will depend not only on the information available but also on the way in which genetic diversity varies within crop gene pools and the collections being used.

A stepwise, hierarchical procedure is usually followed to define the groups: first making the major divisions and subsequently splitting these subgroups into smaller ones (van

Hintum 1994). Often, the first divisions are based on taxonomy, separating the wild from the cultivated species and, within these groups, separating the species and then subspecies in a stepwise manner. Using taxonomy and knowledge about domestication, distribution, breeding history, cropping pattern and utilization, a structured hierarchy can be developed which forms a diversity tree (see Fig. 3).

At each stage in the stratification process one needs to identify meaningful subgroups that are expected to be genetically distinct from each other. It is also important to ensure that all the material in a group finds a place in a subgroup; adding groups such as “other” and “unclassified” might be useful for this reason.

The process of stratification often can often be relatively straightforward, with taxonomic groups being divided into crop types or into ecogeographic groups which are well established and have a long history. Examples of relatively simple classifications of this type include spring and winter-sown barley (Knüpfper and van Hintum 1995) and sesame from different Chinese production regions (Zhang *et al.* 2000). More complex stratification processes have been developed, as detailed in Box 3.

This process of dividing groups ends at the point where it is no longer sensible or possible to divide the subgroups any further. This might be because there is no further reliable

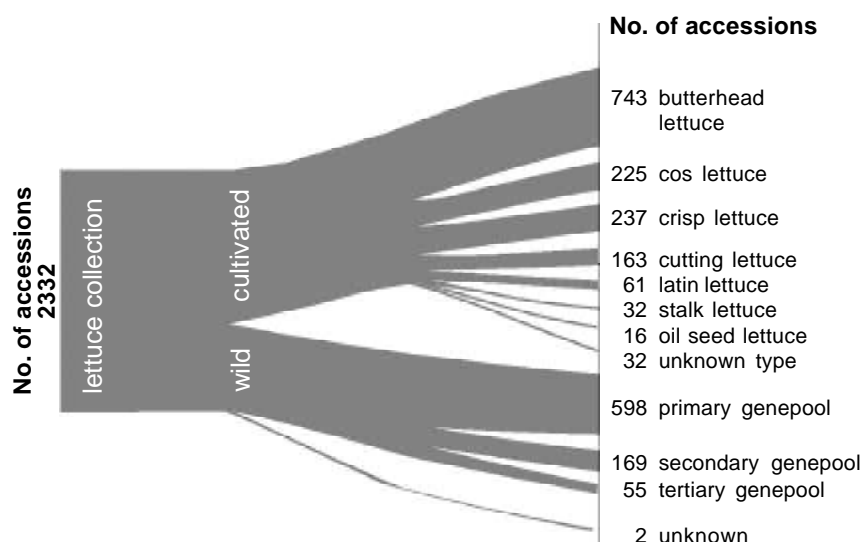


Fig. 3. The first part of the diversity tree of the CGN *Lactuca* collection. The widths of the branches correspond with the number of accessions in the corresponding groups.

information available that could be used for separation of genetically different groups, or because the groups are genetically homogeneous.

Stratification might end, as in the previous example, with groups such as “beans from short-day, dry, lowland areas in Mexico” (Tohme *et al.* 1995). In other, less refined cases, stratification produced groups such as “white cabbages from France” (Boukema and van Hintum 1994), or the wild species “*Medicago arabica*” (Diwan *et al.* 1994, 1995).

An alternative approach to creating groups of similar accessions is using multivariate analysis (for a discussion of the techniques see Crossa *et al.* 1995). If data are available on genetic markers, agromorphological characteristics or other characters, it is possible to construct a dendrogram and to group accessions, using a range of different cluster, discriminant or principal components analysis methods (e.g. Fig. 4, Box 4). This approach can be used together with the classification procedures described above, as in the case of a Chinese sesame core collection (Zhang *et al.* 2000). In this case 14 different groups were established using geography, type of variety (modern or landrace) and agroecological production zone. Cluster analysis of 14 agromorphological characters using Ward’s procedure was then used to identify the final sets of groups for core formation.

In some cases, it is not possible to include all accessions in the cluster analysis. This may result from incomplete data sets or from limits in the capacity of the statistical software. In these cases the next step will be to allocate the remaining accessions to the clusters formed. Bisht *et al.* (1998) developed a procedure to carry out this process when software capacity was limiting and used it to allocate 4000 sesame accessions to clusters developed from a clustering analysis of a test set of 100 accessions. More work is required to test this type of procedure and determine its adequacy. In particular, care should be

Box 3. The hierarchical classification used for the CIAT *Phaseolus* core collection

Tohme *et al.* (1995) used an elaborate classification in developing their *Phaseolus* core collection. They divided the CIAT (Centro Internacional de Agricultura Tropical, Colombia) common bean (*Phaseolus vulgare*) collection into the groups “cultivated in primary centres of diversity”, “cultivated in non-primary centres of diversity” and “wild”. Within the first group a division according to country of origin was made, resulting in 11 subgroups. Within each subgroup a division according to similar adaptation was made, separating the areas with a long history of bean production from those where beans had been introduced recently. These adaptation zones were further divided into agroecological classes according to the specific combination of soil type, altitude, water stress and daylength that was present. This gave a set of 54 possible environments and bean accessions were classified according to the environment from which they had been collected.

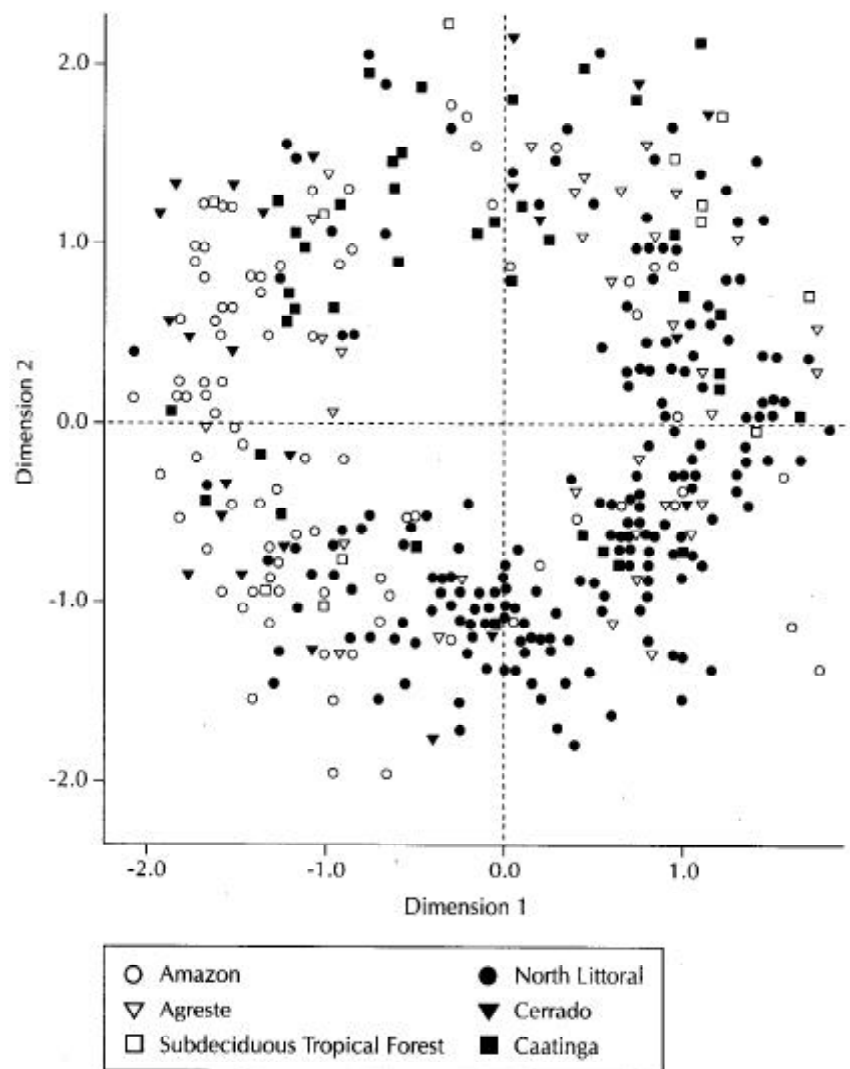


Fig. 4. First two dimensions of a multidimensional scaling of 389 Brazilian cassava accessions showing the close association between agromorphological characteristics and ecological classification (Cordeiro *et al.* 1995).

taken that the data used produce genetically consistent groups; the data (and characters used) must have genetic significance. In the case of data on quantitative traits, this might be a problem given genotype \times environment interactions and experimental error variation.

2.4 The number of entries per group

Once the groups have been defined, the number of entries that are to be included within each group has to be decided. This

can be done in a stepwise manner as the process of stratification is carried out. After the first stratification step the number of entries allocated to each major group is determined and this process is repeated with each subsequent division. Alternatively, it may be carried out as a separate activity at the end of the grouping process.

There are several ways of deciding on the number of entries that should be in each group and these are described in some detail in the paragraphs below. Broadly speaking, three approaches have been used. One involves allocating entries according to the number of accessions that occur within a group. Alternatively, if sufficient material in the groups has been characterized with genetic markers, it is possible to compare marker diversity within the groups and base the allocation on this comparison. Finally, a subjective approach: the allocation can also take account of user needs or other information available to the curator.

The decision as to which of these approaches to follow is largely up to the curator and a mixture of different strategies may often be most appropriate. For example, information on marker diversity can be used to determine allocation numbers for some groups, while for other groups a different allocation strategy could be used. Similarly, a decision on the number of entries of wild or cultivated materials that should be included in a core can reflect user capacity or needs. However, within each of these groups a different allocation strategy based on diversity or numbers can be followed. Another common decision made by curators is to ensure that a least one accession is included from every group, even when group sizes are so small that a proportional allocation would result in their omission from the core.

2.4.1 Procedures based on group size

Procedures for determining the number of entries per group are available that are based on the number of accessions in each group. The number of accessions in the various groups in the collection is often a good overall indicator of the utility of

Box 4. The clustering procedure used for developing a French core collection of ryegrass

An interesting example of using multivariate analysis is given by Charmet *et al.* (1993; Balfourier *et al.* 1999). They used a clustering algorithm in which wild ryegrass populations collected in France were grouped on the basis of agronomic traits. The clustering procedure had the constraint that two populations could only be clustered together if their collection sites were less than 60 km apart. This made it impossible that populations growing in different parts of the country would occur in the same group simply because they resembled each other agronomically.

the groups or the diversity within them (or both). They will reflect the results of collecting trips, research efforts, historical needs and users' interest. Of course they will not be infallible indicators and often some subjective adjustment may be desirable, as discussed in section 2.4.3 below.

Brown (1989a) proposed three procedures based on group size which are usually known as the constant (C), proportional (P) and logarithmic (L) strategies. Brown recommended that, whatever strategy was followed, at least one accession from each group should be included in the core, no matter how small the group.

- The constant strategy (C) simply allocates an equal number of entries to each group, independent of the number of accessions in the group.
- The proportional strategy (P) allocates the entries to a group in proportion to the number of accessions in each group. Also here, the number of entries in the subgroups need to be integers, and an appropriate rounding function should be used.
- The logarithmic strategy (L) allocates entries to each group in proportion to the logarithm of the number of accessions in the group. The number of entries in each group is rounded to the nearest whole number, ensuring that the total number of entries allocated equals the number available.

The three strategies will give different number of entries in different groups (Table 1) and their efficiency has been compared by a number of workers (e.g. Brown 1989a; van Hintum *et al.* 1995; Mahajan and Bisht 1999). While differences were usually slight, the constant strategy generally performed less well than the other two. The evidence therefore favours the use of a proportional or logarithmic strategy and these have been most widely used. As a general rule, the logarithmic

Table 1. Allocation of entries to different groups in a core collection according to the constant (C), proportional (P) and logarithmic (L) strategies, if 30 entries are to be selected from 205 accessions

Group	No. of accessions in collection	No. of entries using different strategies		
		C	L	P
1	120	8	10	18
2	50	8	8	7
3	25	7	7	4
4	10	7	5	1
Total	205	30	30	30

algorithm might be the first one to consider, since it gives intermediate results.

2.4.2 *Procedures based on marker diversity*

The next group of procedures involves the use of data that provide estimates of the amount of diversity within groups or the divergence between them (Brown and Schoen 1994; Schoen and Brown 1995).

- In the H strategy, marker genes such as allozymes or DNA markers provide estimates of allele frequencies at polymorphic loci from which to calculate a gene diversity index (h). This index is equivalent to the expected heterozygosity if a population is fully outbreeding. Theory based on a strict maximization of the allelic richness of the total core (the total number of alleles in a combined sample) suggests that the relative contribution of each group to the core should be in direct proportion to its diversity as measured by the function $\{h / (1-h)\}$.

This approach can be extended to use quantitative data. Estimates of additive genetic variance can be used instead of estimates of $\{h/(1-h)\}$ as weighting factors. If environmental variance can be presumed constant, the same theory justifies using phenotypic variances as weighting factors for group contributions to the core. Other options include the use of the Shannon-Weaver diversity index based on qualitative morphological trait data (see Mahajan and Bisht 1999 for an example of this).

- The M strategy is an alternative approach to using marker data that focuses on the distinct allele types present for a range of marker loci in each accession. It makes a deterministic use of the marker data, rather than a statistical use. Using a linear program it searches for the combination of accessions that will maximize the number of observed alleles at the marker loci in the core while keeping the total number of samples to a specified limit and ensuring a minimum number from every group. Thus the M strategy guarantees including the maximum allelic richness for the marker loci used and takes direct account of the amount of variation and the pattern divergence at those loci. It is worth noting that the M strategy not only determines the number of accessions that should come from different groups, it also identifies the accessions that should be included.

Both strategies assume that all the elements of diversity information used (the alleles in strategy M) are equivalent in

their value as indicators of total diversity within groups. More complex value systems are conceivable that would attach particular weighting factors to variation at specific loci, if variation at these loci was felt to be more important than variation at other loci.

It is worth emphasizing that the effectiveness of these different strategies depends on a successful classification that gives genetically meaningful groups. The grouping procedure ensures that the major components of adaptation across all habitats or separate areas are properly represented so that a genetically rich area will not entirely displace samples from areas with less polymorphism but different adaptive characteristics.

2.4.3 Procedures based on informal knowledge

Subjective procedures, based on informal knowledge, can be used to determine the number of entries in a group, and may also be used to adjust the numbers obtained through some other procedure. As noted above, while group size may often be a reasonable indicator of group diversity or utility, this is not necessarily so. The accessions in collections often reflect past priorities which are no longer relevant, historical accidents (such as gifts of material from other centres) or specific research projects. These can all increase the number of accessions in certain groups. Whatever strategy is used, the curator may want to adjust the size of certain groups either upwards or downwards (see Table 2).

The most common adjustments that curators might make are:

- to increase the size of a group that is expected to be particularly important to the user community
- to decrease the size of a group which is of little interest to the

Table 2. Theoretical example of allocation of entries to groups following adjustment of the L strategy to take account of expected diversity level in the groups

Group	No. of accessions in collection	Allocation through L strategy	Expected diversity in group	Adjusted no. allocated
1	120	10	Moderate	10
2	50	8	Low	6
3	25	7	Moderate	7
4	10	5	High	7
Total	205	30	30	30

user community despite its high diversity (common for wild crop relatives in collections)

- to decrease the size of a group that is present in large numbers but is not believed to contain much diversity
- to increase the size of a group believed to contain high levels of diversity that is poorly represented in the whole collection, even when a formal estimate of diversity is not available (this may be the case for wild crop relatives).

2.5 Choosing the entries

The final step in establishing a core collection is choosing the actual entries. At this point in the procedure the collection has been divided into many small groups of similar accessions and the number of entries that have to be chosen in group is decided. The question that remains is which entries should be chosen in each group.

The selected entries should be those that best represent the group and best serve the function and purposes of the core. This is not only with respect to the genetic diversity in the group but also with respect to other considerations such as the quality of the documentation of the entries, availability of seeds or the role of some entries as standards or as important parents in breeding programmes. As described below, several approaches are possible. The choice can be more or less random, it can be based on some formal analytical procedure or on pragmatic considerations. The method used can differ between groups since the available information may well differ. When the choice is between well-known cultivars in a group, an informed decision often can be made because information on them is available. In contrast, when the choice is between accessions in a group of wild material from the same region, there may be no additional data and only random procedures can be used.

2.5.1 *Random and systematic procedures*

The quickest and easiest way to choose the entries is to make a random sample of accessions in the group. Systematic selection is an alternative to simple random sampling but care should be taken to avoid selecting a set of sequential accession numbers since this might be biased. Often, blocks of accessions of similar origin will enter a collection together, and will be given sequential accession numbers. For this reason it may be better to choose spaced accession numbers. For example, if accession numbers 23, 24, 25, 53, 65, 66, 67, 68 and 69 form a group and three entries have to be chosen, it is better to take 24, 53 and 67 rather than 23, 24 and 25.

2.5.2 Analytical procedures

There are several ways in which additional data from the accessions in a group can be used to select the entries. For example:

- If some marker, characterization/evaluation or additional passport data are available, one can ensure that all possible variants are represented in the core. Thus, if two entries are to be selected from a group, and data on flower colour are available, one might include accessions with different flower colour.
- If extensive marker, characterization or evaluation data are available, it is possible to carry out a multivariate analysis of the accessions (Crossa *et al.* 1995). This could be a principal component analysis or a clustering analysis. It should ideally result in a number of clusters equal to the number of entries allocated to the group. Each cluster can then be represented by a random accession, by the most representative accession of the cluster or by an accession meeting pragmatic criteria listed above.
- A formal procedure for selecting the entries has been developed by Noirot *et al.* (1996). This method (Principal Component Scoring) maximizes the sample diversity of a group of selected accessions using quantitative data and principal component analysis and can be used for selecting accessions within any identified group.
- If the pedigrees of the accessions in a certain group are known, this information can be used to maximize the diversity of the selected entries by avoiding accessions with common parentage (van Hintum and Haalman 1994).

2.5.3 Pragmatic procedures

Practical considerations will always play a part in the choice of entries from a group. These can include:

- *Reliability of classification.* As a general rule, the entries in the core collection should have reliable data; accessions for which the classification into groups is uncertain should, if possible, not be selected.
 - *Amount of additional information.* The amount of information available on an entry increases its potential usefulness in a core collection and entries with additional information, such as evaluation or pedigree data, are to be preferred whenever possible.
 - *Reputation.* Accessions with a high reputation, for example that played an important role in breeding history, or that are being used as a standard in research, should be favoured for inclusion in the core collection.
 - *Availability of material.* It will probably be important to have relatively large quantities of material available for entries in the
-

core collection and those accessions for which substantial quantities are available will be preferred.

- *Policy.* There may be some restrictions on distributing some accessions and these should be omitted from the core collection. One important aspect, for example, may be the date of collecting. Material collected after the entry into force of the Convention on Biological Diversity in 1992 is governed by different procedures with regard to its distribution than material collected before this date.

At the end of this stage a core collection will have been established. In the next sections its management and use are discussed.

3 Managing the core collection

Once the entries for the core collection have been identified, genebank managers need to make a series of decisions about the management of the core collection. These include deciding on how to store the entries, how to regenerate and multiply them and how to manage their documentation. Once the core collection is established, the manager will also want to evaluate the extent to which it has been successful in including diversity present in the whole collection without unnecessary redundancy. This process has been called validation. Finally, procedures for altering the core entries will need to be established that can take account of new knowledge, or the addition of new accessions to the whole collection (some of which it might be desirable to include in the core collection).

3.1 Maintaining the core collection

A first question to be asked is whether sets of the core collection entries need to be maintained separately from the rest of the genebank accessions or whether, even if maintained as part of the whole collection, they need any special treatment. In genebanks with good facilities and ample storage or maintenance capacity, it may seem unnecessary to maintain core entries separately from the rest of the collection. However, even in these situations, additional quantities of seed or propagating material of the core entries will probably be required and space and facilities allocated accordingly.

If the collection from which the core has been developed is maintained as a base collection⁴ a separate set of the selected entries will certainly be required. Where an active collection already exists and is used to provide the core collection entries, then the genebank manager can decide if physical separation of a core set is useful. There are a number of reasons why it may be desirable:

- Minimizing the possibility of errors by ensuring that a single standard source is used which is maintained in one place as the core collection
- Simplifying distribution and use by retaining all the core collection in the same physical location

⁴ A base collection is a collection of genetic resource samples which is kept for long-term secure conservation and is not used as a routine distribution source. Materials are only removed from a base collection for infrequent regeneration (IBPGR 1991).

- Simplifying the additional monitoring that is desirable for a core collection of, for example, the amounts and quality of the entries
- Providing added security by maintaining the core collection in a more secure site (e.g. under close supervision near appropriate facilities in the case of field-maintained material)
- Meeting the expected needs for increased quantities of the core collection entries.

In some cases, the reduced number of accessions will allow the core collection to be maintained using alternative (improved) conservation methods. Cold-temperature facilities (a deep freeze cabinet) may be available for the entries in the core collections even when it is not available for the whole collection. The use of *in vitro* storage methods can become feasible for core collection entries, as is the case for the core collection of coffee developed by IRD⁵ (Institut de recherche pour le développement), where accessions are maintained under slow growth conditions as shoot cultures (Dussert *et al.* 1997).

The entries selected for core collections are not normally modified in any way when they become part of the core. More seed or other propagating material may be required but, in most cases, genebank managers will want to ensure that the properties and characteristics of the core entries remain unchanged. This has the advantage that the core entry remains identifiably the same as the accession in the base (or active) collection that was originally chosen. However, exceptions are possible. In the international Barley Core Collection (Box 5) each entry is being converted to a pure line

Box 5. The international Barley Core Collection

The international Barley Core Collection (BCC) is an example of a core collection which is not part of any other collection or combination of collections (Knüpffer and van Hintum 1995). It is 'a selected and limited set of accessions, optimally representing the genetic diversity of cultivated barley (*H. vulgare* s.lat.) and wild species of *Hordeum*, and providing well-known genetic standards.' The objectives are 'to increase the efficiency of evaluation and thus of utilization of existing collections, to provide a manageable and representative selection of the available barley germplasm for use in research and breeding, and to provide adequate material for the needs of standardization in scientific work with barley.' The BCC was compiled by an international committee on the basis of expertise on barley diversity in specific regions or particular wild taxa of *Hordeum*. The size has been kept restricted, and will not exceed 2000 entries. It is maintained and distributed by a number of genebanks throughout the world, some of which are responsible for specific parts of the core collection. Wherever possible the accessions in the BCC are homozygous lines, permitting identical multiplication of the material over generations and locations and avoiding genetic change due to natural selection or random drift.

⁵ Formerly ORSTOM (Institut français de recherche scientifique pour le développement en coopération).

by single seed descent from an individual plant typical of the entry (Knüpfper and van Hintum 1995). The arguments in favour of this were that the core entry would always remain identical and unchanging and could be used anywhere in the world for any purpose in the sure knowledge that the material used was genetically identical. On the other hand, the advantages of retaining the entries as heterogeneous as they originally were means that intrapopulation variation can be studied, and used.

Another suggestion is to constitute the entries of a core collection by combining or bulking the accessions within a group. Once a final group of accessions has been identified, samples would be taken from each accession and mixed together to form the core entry. This would be used with seed-propagated crops and might retain the allelic richness of the whole collection in the core. However, it breaks the connection of equivalence between the entry in the core and that in the original collection. Some of the data on the original accessions such as passport data would no longer be relevant. This procedure has been suggested for a perennial ryegrass core collection being developed in France (Balfourier *et al.* 1994) and is further discussed in Section 5.

3.2 Distributing the core collection

The demand for core collection entries is likely to be substantially greater than for other accessions in the whole collection. Genebank managers will want to have larger quantities of seed or other planting material available of the core collection entries. It is therefore desirable to use the best available procedures for regenerating and multiplying the core entries. Thus, seed-propagated outbreeding species should be multiplied using sufficiently large plant populations, with optimum isolation practices, so as to minimize genetic drift and avoid geneflow between accessions (Breese 1989). Appropriate precautions also will be needed for vegetatively propagated species. Where the species is maintained *in vitro*, more samples should be maintained but where this is not possible the entries will often have to be multiplied "on demand".

An aspect of genebank management that is increasingly important is the distribution of material free of pathogens and pests. Many countries impose severe restrictions on germplasm introduction for plant health reasons. Genebank managers will want to ensure that core collection entries meet appropriate plant health standards, both because these are the accessions that are most likely to be distributed, and because it is desirable

that the highest possible maintenance standards are adopted for this set of material (see FAO/IBPGR Technical Guidelines for the Safe Movement of Germplasm). At the same time, with a relatively small number of entries, it often will be possible to carry out testing, quarantine and pathogen-elimination activities for a core collection that are impractical for the whole collection.

3.3 Information management

As the core collection is developed and used, more and more information will become available on the entries. No special procedures may be required for this and the genebank's normal documentation systems will be appropriate. Normally, the core collection will be fully characterized and, where possible, it should be the first material used when new characters (e.g. biochemical or molecular markers) are assessed. The information available on the core entries will therefore soon become more extensive than that available for other accessions. This is likely to be particularly the case for information on the agronomically important characteristics that constitute evaluation data.

The process of core collection establishment and entry selection should itself be documented. Information on the procedures that were followed and the data used needs to be available so that it can be used in validation, substitution, amendment or in exploring properties of accessions from the same core group. Publication of the procedures used to establish the core collection is one important part of this process.

Information on the core collection entries can provide valuable additional information on the whole collection. The stratification procedures used to establish the core can be used in reverse to provide information about other accessions in the group from which a specific entry was obtained. Of course, there are limitations to this process depending on the grouping procedure used and the characteristics involved.

Some information management implications will be apparent where core collections are developed through collaboration between different genebanks (e.g. in the case of the international Barley Core Collection or the proposed *Beta* core collection). Arrangements will need to be made to ensure that all partners have access to, and can provide the same information on, the core accessions. As core collections are distributed and used, more information should become available from a variety of different sources. The ways in which this information is incorporated and made available may require some special arrangements by database managers.

3.4 Validating the core collection

Once a core collection has been established, an important question for genebank managers is the extent to which it meets its original objectives in terms of the representation of diversity and lack of repetition. This process of validating the core collection usually involves comparing it in some appropriate way with the original collection from which it was developed. Validation should not interfere with the use of the core collection but should, if possible, be integrated with it. A comparison between the core collection and the whole collection can be carried out using characteristics that were involved in establishment of the core or using characteristics that were not used in its development. In general, both types of comparisons are desirable for validation. A critical analysis of whether entries meet the objectives of the core collections is also useful (Ortiz *et al.* 1999)

A preliminary evaluation can be made by comparing means, ranges, frequencies and variances of specific characters in the different groups of the core collection with those of the groups from which they were derived. Ranges are expected to remain similar while means will move toward the median and variances may increase in the core compared with the whole collection (Figs. 5, 6). However, certain core collection selection methods may not give this result. The principal component score strategy developed by Noirot *et al.* (1996) tends to select entries with extreme expressions of character states used. With this method, entries with median expressions may be under-represented.

Biochemical and molecular markers have been suggested as suitable characters to assess the success of the core in meeting the objectives of representing genetic diversity. For example, core collections of about 10% should possess about 70% of the alleles found in the whole collection. Of course, it is impractical (and unnecessary) to test a core sample against the complete collection from which it is derived. However, samples of the core and the whole collection can be compared to determine whether they have broadly similar biochemical or molecular marker alleles. Thus, Skroch *et al.* (1998) used RAPDs to test Mexican *Phaseolus* core entries and compare them to a random set of the collection from which they were derived. Interestingly, in this case no difference was found between the two samples. This suggested, on the one hand, that the core was no more successful than a random set of accessions in capturing diversity of RAPD polymorphism and, on the other, that all polymorphisms found in the random set were also present in the core.

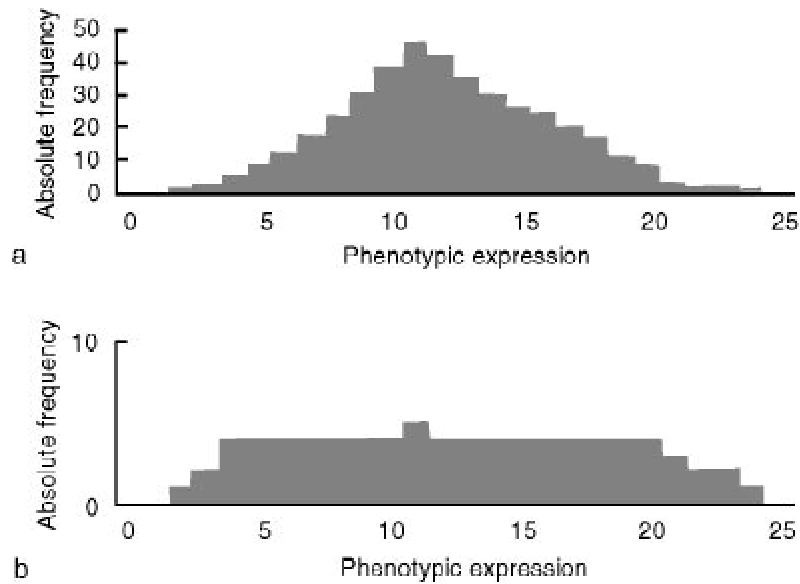


Fig. 5. Distribution of a quantitative trait in the whole collection and a core collection. (a) The whole collection. Number of accessions=400, average=11.3, standard deviation=3.9, minimum=1, maximum=24. (b) A core collection. Number of entries=80, average=12.4, standard deviation=5.9, minimum=1, maximum=24.

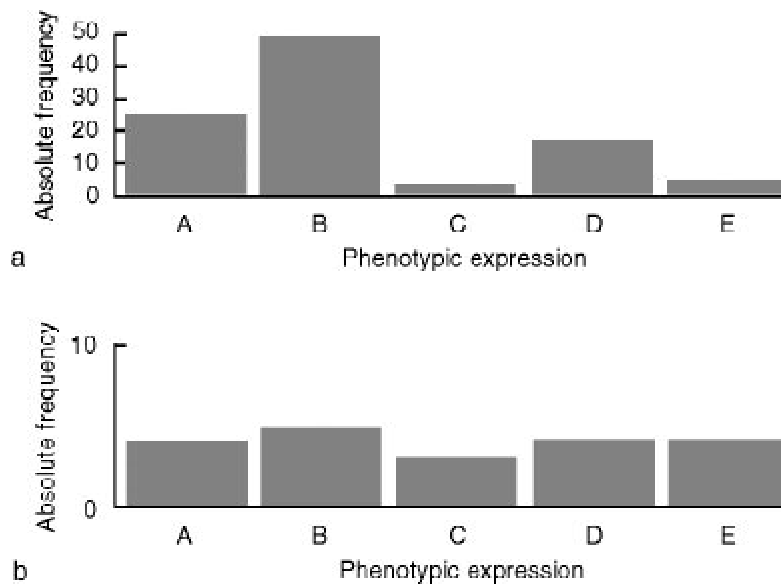


Fig. 6. Distribution of a qualitative trait in the whole collection and in a core collection. (a) The whole collection (number of accessions=100, number of classes=5). (b) The core collection (number of entries=20, number of classes=5).

Testing for redundancy is more difficult and will depend to a great extent on the experience of the genebank manager. Where costs of maintenance are high it may be worth using molecular and biochemical markers for such tests (van Hintum *et al.* 1996; Phippen *et al.* 1997). However, it will always be important to look at all characteristics for which information is available and not to rely on a few characters to detect presumed duplication or redundancy.

3.5 Altering the core collection

Once established, a core collection should be fairly stable. Its management will become more difficult if its constitution is continually changing and its use will be inhibited if the entries are constantly subject to change. There is a need for a certain level of constancy of the core over time so that evaluations of the core at different times can be compared and the bulk of the entries should remain constant. At the same time, genebank managers will want to change the number and identity of the entries in a core collection from time to time. The changes may be slight (substituting one entry for another in a specific group), but they may be more substantial (changing the balance between groups). Some of the factors that genebank managers should consider when deciding to change the constitution of a core collection include:

- Identifying accessions with specific desirable traits that should be included in the core collection such as new forms of resistance to specific diseases or to particular stresses.
- Meeting requests from users to include specific accessions on the grounds that they have important characteristics for the user community such as standard cultivars for specific tests or anchor genotypes used in molecular genetic mapping or genome sequencing.
- Acquiring additional accessions with new diversity that needs to be represented in the core collection.
- Increased knowledge on the extent and distribution of genetic diversity within the material from which the core is drawn. Finding that some groups have substantially more diversity than other groups may lead to considering the possibility of altering the number of entries taken from the groups in question.
- Taking account of new breeding methods or knowledge that change the value of a group to users. Thus, the detection of yield-related QTLs (quantitative trait loci) in wild relatives might increase the value of some wild species to users.

Whatever the reasons for change, the process should be planned and organized in such a way as to avoid compromising

the integrity of the core collection or causing problems to users. Ideally, users should be informed of any significant changes to the core and the reasons for such changes. The quality of the material that is introduced into the core and the data on it should be of the same standard as the rest of the core, or rapidly brought to that standard. Since the emphasis of core collections is on size limitation, the objective of the manager should be to replace rather than to add new material.

4 Using the core collection

Core collections are established to improve the conservation and use of genetic resources. They can help in genebank management, in the decisions that need to be taken on what should be conserved and in the improved use of material held in genebanks (Table 3). In this section, the ways in which core collections can be used are described and ways to increase their use are outlined.

From Table 3 it can be seen that core collections can be used in a number of different ways:

- As a *limited* set, in tasks where a small number of accessions is needed (e.g. 12, 15 in Table 3).
- As a list of entries that will have *priority* for attention (e.g. 3, 4, 6 in Table 3)
- As a *reference* set, when an indicator is needed for sections of the whole collection (e.g. 1, 5, 10 in Table 3)
- As an optimum set of *experimental* material, for assessment of the collection or the development of methods (e.g. 2, 7, 9, 11, 16 in Table 3).

Often the core will function in more than one way with respect to a particular task. Thus, all or part of a core can function as both an experimental set and a reference set in developing the optimum method for testing for a complex trait, evaluating the core response to the trait, and using some (or all) of the core in genetic studies to determine trait inheritance. In discussing the use of the core it is helpful to distinguish between the ways in which core collections can help genebank managers conserve material better, and the ways in which cores can help improve use of collections.

4.1 Improving genebank operations

The work of establishing a core collection usually leads to a substantial increase in knowledge of the extent and distribution of genetic diversity within a collection. Once it has been established, it becomes the framework for further studies. Patterns of diversity can be expected to reflect among-group and within-group structures of the core collection. Specific hypotheses concerning the distribution of diversity in certain groups or sets of groups can be tested (Tohme *et al.* 1999), e.g. are samples from centres of origin more diverse than from outlying areas of a species distribution? Data failing to confirm this hypothesis might indicate a serious gap in the collection.

Table 3. Examples of the ways in which core collections can be used

Task	Function of core
Collection characteristics	
1 Dealing with new accessions	Provide a reference set for determining the group of accessions with which new material should be compared
2 Detection of gaps or uneven collecting	Allow identification of discontinuities in variation indicating missing material or sets where large amounts of variation are associated with a few accessions
Collection management	
3 Developing regeneration schedules	Core entries have priority, especially when upgrading collections
4 Prioritizing handling	Provide set for priority handling when needed
5 Monitoring viability	Provide appropriate set of accessions for monitoring whole
6 Duplication	Act as a priority group for safety-duplication, for further distribution to regional or international genebanks or for maintenance in different conditions (e.g. as DNA libraries, in field banks or <i>in vitro</i>)
7 Development and application of new conservation methods	Provide test material of choice for possible improved maintenance procedures (e.g. ultra-dry seeds, <i>in vitro</i> and cryopreservation)
Information management	
8 Database organization	Provide benchmark standard for documentation and allow stratification of whole collection to be recorded
Study and use of collections	
9 Developing descriptor lists	Entries appropriate to test sufficiency of descriptors to discriminate accessions
10 Testing expensive and complex traits (e.g. photoperiod response)	Allow development of efficient two-step sampling procedure – first between, then within groups
11 Method development	Provide set of material which is likely to cover full range of characteristic expression
12 Relationships between different characters	Provide restricted set likely to cover full range of different character expressions to maximize efficiency of correlation studies
13 Genetic studies	Allow selection of optimal material for studies of trait inheritance and estimation of general combining ability
14 Prebreeding	Provide dissimilar groups likely to assist in identifying heterosis or bringing together new gene combinations
Distribution from collections	
15 Having adequate seed supplies on hand	Larger amounts of seed can be produced of the limited set of core entries
16 Distribution of diverse and representative samples	Provide maximum diversity in limited sets of accessions for assessment by users

More specifically, the core provides evidence of gaps in collections (groups with very small number of accessions or substantial discontinuities between groups) which may indicate that further collecting is required. The core accessions also provide a suitable reference set for new material entering the genebank which can be assigned to a specific group and tested with a set of accessions expected to show similar properties.

Genebanks are usually interested in applying and improving a variety of different conservation activities (e.g. *in vitro* methods, methods for assaying purity, regeneration protocols, development and testing of descriptor lists). Such methods must be applicable and reliable across the spectrum of genetic variation that they handle and the core provides the most appropriate group for such studies (Dussert *et al.* 1997). The core furnishes ideal material for this purpose. It also provides good material for the various routine monitoring activities that genebanks have to carry out, such as seed viability testing.

The core is never intended to replace the whole collection but there may be a number of situations where it can function as a priority set for security purposes. It can be given priority if resources for regeneration are limited or if safety-duplication of a collection has to be carried out in stages. It can also provide the optimum set of material for emergency situations, such as civil strife or environmental disasters, where only a portion of the collection can be secured.

A continuing area of development once a core collection has been established will be the linkage between the core collection and the whole collection. An important question for any collection is: how reliable are inferences based on core entries in predicting the performance of other accessions in the same group? An ideal strategy for curators would be to accumulate case histories and experience relevant to this question as a way of assessing and improving their core collections.

4.2 Improving the use of plant genetic resources

The wider the range of different users for the core collection the wider the range of uses the core will be put to. While plant breeders and other researchers may be the mainstay of the user community, it is worth trying to identify other potential users who could use a core collection in their work could be identified. Possible users for all or part of a core collection could include educationalists, agronomists, extension agents, NGOs (non-governmental organizations), farmers' organizations, and growers themselves.

Table 4. Examples of the use of core collections for detecting germplasm with specific characteristics

General characteristic	Specific trait	Core collection	Reference
Biotic stress resistance	Novel white mold resistance	Common bean	Miklas <i>et al.</i> 1997
	Resistance to late leafspot; tomato spotted wilt virus; Rhizoctonia limb rot	Peanut	Holbrook and Anderson 1995; Anderson <i>et al.</i> 1996; Franke <i>et al.</i> 1999; Holbrook 1999
	Anthraxnose, Fusarium wilt	Lentil	Kaiser <i>et al.</i> 1998; Baya <i>et al.</i> 1997
	Cabbage aphid, <i>Brevicoryne brassicae</i> resistance	<i>Brassica oleracea</i>	Ellis <i>et al.</i> 1998
Abiotic stress tolerance	<i>Peronospora parasitica</i> resistance	<i>Brassica oleracea</i>	Coelho <i>et al.</i> 1998
	Eyespot (<i>Pseudocercospora herpotrichoides</i>) resistance	<i>Triticum monococcum</i>	Cadle <i>et al.</i> 1997
Chemical content	Acid soil tolerance	Alfalfa	Bouton 1996
	Salinity	Quinoa	Ruiz-Tapia <i>et al.</i> 1997
Physiological traits	Oil and meal quality factors	Safflower	Bergman <i>et al.</i> 1997
	Fatty acid composition – epoxy and eicosenoic acids	Peanut	Hammond <i>et al.</i> 1997
	Leaf and stem forage quality	<i>Medicago sativa</i>	Jung <i>et al.</i> 1997
	Digestibility	<i>Lolium perenne</i>	Boller <i>et al.</i> 1998
Other breeding-related traits	Storage response under slow growth <i>in vitro</i> conditions	Coffee	Dussert <i>et al.</i> 1997
	Turf quality, seed production	<i>Poa pratensis</i>	Johnson <i>et al.</i> 1999
	Cyanogenesis and climate adaptability	White clover	Pederson <i>et al.</i> 1996
	Nitrogen uptake and biomass accumulation	Potato	Errebhi <i>et al.</i> 1998
Other breeding-related traits	Combining ability/heterotic patterns	Wheat	Spagnoletti Zeuli and Qualset 1995
	Genotype x environment interactions	Perennial ryegrass	Charmet <i>et al.</i> 1993; Balfourier <i>et al.</i> 1997

All users of plant genetic resources should be encouraged to use the core collection wherever relevant. In this way they will be able to use the genetic diversity in a crop more effectively and comprehensively to detect the characteristics or properties in which they are interested. This will improve the quality of their work and will provide opportunities to ensure that information on the core collection continually improves. The advantages that the core provides to users in terms of both increasing the quality and efficiency of their work should be continually stressed.

Genebank managers will want to ensure that they maintain flexibility in using a core to address the specific needs of different users. These may not always require all of the core collection. Subsets of the core, which are maximized for genetic variation of certain characters, or in certain types of material, can be used to address specific requests. For instance, subsets might be provided from a core subset which included only the wild relatives of a crop as part of a search for new yield-increasing genes (Tanksley and McCouch 1997).

A major interest of many users concerned with either applied or basic research will be to locate specific characters or variation in specific characteristics in a crop genepool, as part of an efficient search strategy. The core collection provides an efficient entry point to the whole collection for a genepool-wide search for characteristics of interest. For instance, the US peanut core collection has been used as an entry point to the overall collection by identifying which groups in the core subset contain accessions with resistance to tomato spotted wilt virus (TSWV). It is then planned to screen all accessions in the whole collection from groups possessing resistance (Anderson *et al.* 1996; Holbrook 1999).

4.2.1 Applied research: Screening for useful traits and characteristics

A common use of the core collection will be to screen for agronomically useful traits or characteristics. Core collections can be screened for a wide range of qualitative and/or quantitative traits. The presence or absence of the traits or characteristics in the core might be used to make cost-benefit analyses of how comprehensive search strategies should be and what procedures will optimize benefits. In fact, core collections already have been used to detect material with specific useful traits in a range of crops and the evidence suggests that they have been an effective way of identifying useful new germplasm (Table 4). Core collections also have been used in more general multi-trait evaluation experiments (e.g. Willner *et al.* 1998;

Hodgkin *et al.* 1999).

In some crops, users may wish to identify useful genotypes rather than genes (e.g. many forages or clonal crops). In such instances, the core will allow a rational sampling of the broadest representation of genetic diversity in the selection and improvement activities. A wider range of users, such as agronomists, extension workers or farmers, may be interested in making direct selections of genotypes from core collections and evaluating them for their own needs.

So far, there have been few reports of the use of core collections for identifying lines with good combining ability (Frankel and Brown 1984; Spagnoletti Zeuli and Qualset 1995). However, core collections provide a very good basis for choosing lines for such studies. Knowing the relationship between commonly used tester lines and relevant core entries may provide added incentives for use of the core.

Core collections will be powerful tools where researchers are interested in background effects on gene expression (epistasis, epigenetics, etc.) and need a range of highly divergent background genotypes for the planned studies. Similarly, end users with an interest in assessing or exploiting the extent of either genotype by environment ($G \times E$) interactions or gene or QTL by environment interactions can use core entries to assess the extent of such interactions in a gene pool (Charmet *et al.* 1993; Balfourier *et al.* 1997).

4.2.2 Basic research and education

Core collections will be useful wherever a wide range of diversity is needed for either research or illustrative purposes.

Box 6. Some options and links for promoting use of the core collection

Options

- Direct mail to all existing users
- Posters or presentations at conferences or meetings
- Scientific and other publications (e.g. *Crop Science*, *Plant Genetic Resources Newsletter*, *Genetic Resources and Crop Evolution*, *Euphytica*)
- Crop newsletters (e.g. *InfoMusa*, *Barley Newsletter*)
- E-mail listservs or newsgroups (e.g. Plant Tissue Culture Listserv, GrainGenes)
- Crop-specific networks (e.g. Cassava Biotechnology Network, INBAR, Nitrogen Fixing Tree Association, International Network for Genetic Evaluation of Rice)
- Regional genetic resources or agricultural research networks (e.g. ASARECA, ECP/GR, AgREN)
- Crop or germplasm committees
- International information systems such as WIEWS.

Links

- Relevant Internet web sites (e.g. germplasm databases: SINGER, pcGRIN)
- Genome mapping/sequencing databases (e.g. GrainGenes, SolGenes, Rice Sequencing project, Grasses Genome Initiative etc).
- Crop evaluation networks (e.g. INGER, LAMP, UPWARD).
- Educational initiatives (e.g. modules of university courses, practical work).
- Agronomists and extension networks involved in agricultural development (e.g. AgREN).
- NGOs involved in agricultural development (e.g. SRISTI).

Core collections have been used to examine the extent and distribution of genetic diversity within crop gene pools (Tohme *et al.* 1996, 1999) and relationships between gene pools. Cores also may be used to establish correlation between traits and environmental parameters and this approach was taken to establish the relationship between cyanogenesis and climate in US white clover germplasm (Pederson *et al.* 1996).

4.3 Increasing the use of core collections

Once a core collection has been established, it is important to let users know of its existence and of how it might be useful to them in their work. It should not be assumed that users of germplasm are automatically aware of the existence and utility of the core collection. Reaching as wide a range of users as possible with relevant information about the core collection will increase the utility of the collection to users and hence strengthen the link between the core and its users. The more clients that use the core **and** provide feedback on its utility, the more the value and relevance of the core to end user needs can be enhanced over time.

Many different ways can be used to reach potential users with relevant information about the core. Being proactive in seeking new users will be desirable and many different ways of “advertising” the core to a wide range of potential users, both at the national and international levels, should be considered. A list of some possible options is given in Box 6.

Continuous feedback from users will help to evaluate and enhance the utility of the core collection. Additional characterization and evaluation data from users should always be sought and will enhance the information quality of accessions in the core collection. In this respect, it is important to establish clear procedures and requirements for users to provide whatever relevant data.

It is important to provide clear guidance on how the core collection should be cited or referenced by end users. If the results of different users are to be cross-comparable, it is essential that all users use the same nomenclature for accessions in references to the core collection. All users could be encouraged to publish or make known their work with the core collection. Constructing and continuously updating a bibliography of documented information and usages of the core collection will provide an invaluable resource over time to new users of the core and assist in core collection management.

Needs, preferences and procedures of germplasm users change over time and it may be useful to survey core collection

needs of users periodically (e.g. every 2–5 years). Such surveys could be used as an input to any efforts to modify the core, and may help to tailor the core toward the needs of different groups (McFerson *et al.* 1996).

Opportunities will arise when multiplying core accessions to demonstrate the range of variability available to the end users. Users could be invited to open days to inspect what phenotypic variability is available in the core or in particular parts of it (e.g. groups within the core). If a wide range of different users is invited to participate in such exercises, their feedback could be particularly valuable to the core collection curator.

Crop improvement approaches and methodologies are continually changing. Linkages or “bridges” between the core collection and other crop improvement initiatives (see Box 6) and having the core collection as either a reference or a relational point for crop improvement activities will be extremely useful.

Future developments

The core collection concept has been around since 1984 and, since then, many core collections have been established and used, and the number is continuing to grow. There will undoubtedly be improvements in the procedures used to establish core collections, as results from different studies become available. The relative effectiveness of taxonomic, geographic, ecological and agromorphological methods of grouping will become clearer. The extent to which different methods ensure that genetic diversity is captured in core collections will be established. New ways of choosing entries for the core from specific groups will be tested and improved methods of using core collections devised.

Information from the increased use of biochemical and molecular markers will certainly help to improve core collections as will the increasing use of geographic information systems in the analysis of patterns of distribution of genetic diversity (Guarino *et al.* 2001). Confirmation and revision of core collections as such data sets become available means that the core collection will function as a “summary” of existing knowledge about the broad variation patterns found in crop gene pools. At the same time, better access to information about collections via the Internet can increase both the popularity and usefulness of core collections.

Core collections will have a major influence at yet another level. The procedure fosters an approach which validates the knowledge that exists about a collection and which provides a powerful incentive to acquire new knowledge. In this way a core collection adds value to the whole collection. Furthermore, it fosters a methodology of sampling and analysis that allows new and better ways of accessing and utilizing genetic diversity to be developed.

An essential step in the methodology of forming a core collection is the selection of a genetically representative set from any larger set of accessions. This means that it is possible not only to represent the entire collection, but also smaller parts of it, or different parts in different densities. It provides a way in which genebank managers can meet requests such as “50 accessions representing the West European cultivars of butterhead lettuce”, “50 accessions representing all other butterhead lettuce in the world”, or, “25 accessions representing the other types of cultivated *Lactuca*”.

Van Hintum (1999) describes a procedure that can be used to identify a representative sample from a germplasm collection

on-line, where there is direct access to computerized data on a genebank collection. This procedure can be used to meet specific demands, such as those listed above and, more generally, to select a sample of any specified size according to established and defined principles. He calls these samples “core selections”, and notes that the procedure can be used for selection of a representative sample of any size based on any set of material. Such systems, in combination with improved Internet interfaces to germplasm databases, can be expected to greatly improve the access to germplasm collections.

The sampling methodology used to develop core collections can be used for other conservation work where defined sets of samples of limited size are required. It could be used to help identify those populations of a wild species, such as a crop relative or forestry species, that should be maintained *in situ* in order to maximize the genetic diversity maintained. It has been suggested that it could be used in a similar way to support on-farm conservation of traditional cultivars (Bhuwon Sthapit, pers. comm.). Here again, the procedures would allow national programmes to work with communities and farmers in identifying local cultivars whose continued use would maximize the genetic diversity maintained.

Another possible methodological development that has been proposed is the “bulk core collection” (van Hintum 1998). This was briefly referred to in Section 3.1. In most core collections, groups of accessions are formed from which one or more entries are selected. In creating a “bulk core collection” all or some of the accessions within a group are bulked to form a single entry. These entries will thus, initially, contain all genetic diversity available within the corresponding groups. The material within a bulk entry can be recombined by crossing or, in the case of cross-pollinators, by growing them out in a single isolation plot.

Bulk core collections might provide a way of increasing the amount of genetic diversity within a core collection since one does not need to select a single accession from groups that are still heterogeneous. However, their utility for use in future work depends on the ability of curators to identify these genes in bulk populations, a problem that is similar in some ways to that of identifying useful genes in large collections. Bulk populations might be used in situations where variable populations are desirable. For example:

- They could function as a starting point for creating adapted populations with broadly similar characteristics. By exposing the bulk entries to natural and mild mass selection, and “feeding”
-

them with germplasm with desired new characters, the germplasm would slowly introgress into the adapted background of the bulks. Experiences with composite crosses of barley (Suneson 1963) and alternative bulk approaches (e.g. Kannenberg and Falk 1995; Veteläinen *et al.* 1996) show the potential of population improvement and suggest that this may be a useful additional approach.

- Bulk core collections, or the products after several rounds of mild selection and adaptation, could be used in dynamic conservation programmes. In these programmes genebank material is exposed to the changing environment and can thus adapt to new stresses. Population conservation programmes are not new (Simmonds 1962) but may be an increasingly relevant way of maintaining adaptive capacity to cope with climate change or increasing UV radiation. In the final analysis such bulk core collections are breeding tools and not core collections in the original sense.

References

- Anderson, W.F., C.C. Holbrook and A.K. Culbreath. 1996. Screening the peanut core collection for resistance to tomato spotted wilt virus. *Peanut Sci.* 23:57-61.
- Balfourier, F., G. Charmet and C. Grand Ravel. 1994. Conservation of allelic multiplicity and genotypic frequency by pooling wild populations of perennial ryegrass. *Heredity* 73: 386-396.
- Balfourier, F., J.A. Oliveira, G. Charmet and E. Arbones. 1997. Factorial regression analysis of genotype by environment interaction in ryegrass populations, using both isozyme and climatic data as covariates. *Euphytica* 98:37-47.
- Balfourier, F., J.M. Prosperi, G. Charmet, M. Goulard and P. Monastiez. 1999. Using spatial patterns of diversity to develop core collections. Pp. 37-48 in *Core Collections for Today and Tomorrow* (R.C. Johnson and T. Hodgkin, eds.). IPGRI, Rome, Italy.
- Baya, B., W. Erskine and M. Singh. 1997. Screening lentil for resistance to Fusarium wilt: methodology and sources of resistance. *Euphytica* 98:1-2, 69-74.
- Bergman, J.W., C.R. Flynn and R.C. Johnson. 1997. Evaluation of safflower accessions for oil and meal quality factors. Pp. 232-243 in *Safflower: A multipurpose species with unexploited potential and world adaptability*. Proc IVth International Safflower Conference, Bari, Italy, 2-7 June 1997 (A. Corleto and H.H. Mundel, eds.). Adriatica Editrice, Bari, Italy.
- Bisht, I.S., R.K. Mahajan, T.R. Loknathan and R.C. Agrawal. 1998 Diversity in Indian sesame collection and stratification of germplasm accessions in different diversity groups. *Genet. Resour. and Crop Evol.* 45:325-335.
- Boller, B., D. Schmid, F.X. Schubiger, B. Boller and F.J. Stadelmann. 1998. Variation in digestibility among European accessions of *Lolium perenne*. Pp. 227-229 in *Breeding for a Multifunctional Agriculture*, Proceedings of the 21st Meeting of the Fodder Crops and Amenity Grasses Section of EUCARPIA, Kartause Ittingen, Switzerland, 9-12 September 1997. Swiss Federal Research Station for Agroecology and Agriculture, Zurich, Switzerland.
- Boukema, I.W. and Th.J.L. van Hintum. 1994. *Brassica oleracea*, a case of an integrated approach to genetic resources conservation. Pp. 123-129 in *Evaluation and Exploitation of Genetic Resources, Pre-Breeding. Evaluation and exploitation of genetic resources: pre-breeding. Proceedings of the Genetic Resources Section Meeting of Eucarpia*, 15-18 March 1994, Clermont-Ferrand, France (F. Balfourier and M.R. Perretant, eds.). Institut National de la Recherche Agronomique, Versailles, France.
- Boukema, T.W., Th.J.L. van Hintum and D. Astley. 1997. The creation and composition of the *Brassica oleracea* core collection. *Plant Genet. Resour. Newsl.* 111:29-32.
- Bouton J.H. 1996. Screening the alfalfa core collection for acid soil tolerance. *Crop Sci.* 36:1, 198-200.
-

- Breese, E.L. 1989. Regeneration and multiplication of germplasm resources in seed genebanks: the scientific background. IBPGR, Rome.
- 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. Pp. 136-156 in *The Use of Plant Genetic Resources* (A.H.D. Brown, O.H. Frankel, D.R. Marshall and J.T. Williams, eds.). Cambridge University Press, Cambridge, UK.
- Brown, A.H.D. 1995. The core collection at the crossroads. Pp. 3-19 in *Core Collections of Plant Genetic Resources* (T. Hodgkin, A.H.D. Brown, Th.J.L. van Hintum and E.A.V. Morales, eds.). John Wiley and Sons, UK.
- Brown, A.H.D. and D.J. Schoen. 1994. Optimal sampling strategies for core collections of plant genetic resources. Pp. 357-369 in *Conservation Genetics* (V. Loeschke, J. Tomiuk and S.K. Jain, eds.). Birkhäuser Verlag Basel, Switzerland.
- Brown, A.H.D., J.P. Grace and S.S. Speer. 1987. Designation of a "core" collection of perennial *Glycine*. *Soybean Genet. Newsl.* 14:59-70.
- Cadle, M.M., T.D. Murray and S.S. Jones. 1997. Identification of resistance to *Pseudocercospora herpotrichoides* in *Triticum monococcum*. *Plant Disease* 81(10):1181-1186.
- Charmet, G. and F. Balfourier. 1995. The use of geostatistics for sampling a core collection of perennial ryegrass populations. *Genet. Resour. and Crop Evol.* 42:4, 303-309.
- Charmet, G., F. Balfourier, C. Ravel and J.B. Denis. 1993. Genotype x environment interactions in a collection of French perennial ryegrass populations. *Theor. Appl. Genet.* 86:731-736.
- Coelho, P., K. Bahcevandziev, L. Valerio, A. Monteiro, D. Leckie, D. Astley, I.R. Crute, I. Boukema, G. Thomas and A.A. Monteiro. 1998. The relationship between cotyledon and adult plant resistance to downy mildew (*Peronospora parasitica*) in *Brassica oleracea*. *Acta Hort.* 459:335-342.
- Cordeiro, C.M.T., E.A.V. Morales, P. Ferreira, D.M.S. Rocha, I.R.S. Costa, A.C.C. Valois and S. Silva. 1995. Towards a Brazilian core collection of cassava. Pp. 155-168 in *Core Collections of Plant Genetic Resources* (T. Hodgkin, A.H.D. Brown, Th.J.L. van Hintum and E.A.V. Morales, eds.). John Wiley and Sons, UK.
- Crossa, J., I.H. DeLacy and S. Taba. 1995. The use of multivariate methods in developing a core collection. Pp. 77-92 in *Core Collections of Plant Genetic Resources* (T. Hodgkin, A.H.D. Brown, Th.J.L. van Hintum and E.A.V. Morales, eds.). John Wiley and Sons, UK.
- Diwan, N., G.R. Bauchan and M.S. McIntosh. 1994. A core collection for the United States annual *Medicago* germplasm collection. *Crop Sci.* 34:279-285.
- Diwan, N., M.S. McIntosh and G.R. Bauchan. 1995. Methods of developing a core collection of annual *Medicago* species. *Theor. Appl. Genet.* 90:755-761.
- Dussert, S., N. Chabrillange, F. Anthony, F. Engelmann, C. Recalt and S. Hamon. 1997. Variability in storage response within a coffee (*Coffea* spp.) core collection under slow growth conditions. *Plant Cell Reports* 16:344-348.
-

- Ellis, P.R., D.A.C. Pink, K. Phelps, P.L. Jukes, S.E. Breeds and A.E. Pinnegar. 1998. Evaluation of a core collection of *Brassica oleracea* accessions for resistance to *Brevicoryne brassicae*, the cabbage aphid. *Euphytica* 103(2):149-160.
- Errebhi, M., C.J. Rosen, F.I. Lauer, M.W. Martin, J.B. Bamberg and D.E. Birong. 1998. Screening of exotic potato germplasm for nitrogen uptake and biomass production. *Am. J. Potato Res.* 75(2):93-100.
- Erskine, W. and F.J. Muehlbauer. 1991. Allozyme and morphological variability, outcrossing rate and core collection formation in lentil germplasm. *Theor. Appl. Genet.* 83:119-125.
- FAO. 1996. Global Plan of Action for the Conservation and Sustainable Utilization of Plant Genetic Resources for Food and Agriculture. FAO, Rome, Italy.
- Franke, M.D., T.B. Brenneman and C.C. Holbrook. 1999. Identification of resistance to *Rhizoctonia limb rot* in a core collection of peanut germplasm. *Plant Disease* 83(10):944-948.
- Frankel, O.H. 1984. Genetic perspectives of germplasm conservation. Pp. 161-170 in *Genetic Manipulation: Impact on Man and Society* (W. Arber, K. Llimensee, W.J. Peacock and P. Starlinger, eds.). Cambridge University Press, Cambridge.
- Frankel, O.H. and A.H.D. Brown. 1984. Plant genetic resources today: A critical appraisal. In *Crop Genetic Resources: Conservation and Evaluation* (J.H.W. Holden and J.T. Williams, eds.). Allen and Unwin, Winchester, Massachusetts, USA.
- Guarino, L., A. Jarvis, R.J. Hijmans and N. Maxted. 2001. Geographic information systems (GIS) and the conservation and use of plant genetic resources. In *Managing Plant Genetic Diversity* (J.M.M. Engels, V. Ramanatha Rao, A.H.D. Brown and M.T. Jackson, eds.). CABI, Oxon, UK. (in press)
- Hammond, E.G., D. Duvick, T. Wang, H. Dodo and R.N. Pittman. 1997. Survey of the fatty acid composition of peanut (*Arachis hypogaea*) germplasm and characterization of their epoxy and eicosenoic acids. *JACOS* 74(10):1235-1239.
- Hodgkin, T., Guo Qingyuan, Zhang Xiurong, Zhao Yingzhong, Feng Xiangyun, P.L. Gautam, R.K. Mahajan, I.S. Bisht, T.R. Loknathan, P.N. Mathur and Zhou Mingde. 1999. Developing sesame core collections in China and India. Pp. 74-81 in *Core Collections for Today and Tomorrow* (R.C. Johnson and T. Hodgkin, eds.). IPGRI, Rome, Italy.
- Holbrook, C.C. 1999. Testing and utilization of a core collection for the US germplasm collection of peanut. Pp. 68-73 in *Core Collections for Today and Tomorrow* (R.C. Johnson and T. Hodgkin, eds.). IPGRI, Rome, Italy.
- Holbrook, C.C. and W.F. Anderson. 1995. Evaluation of a core collection to identify resistance to late spot in peanut. *Crop Sci.* 35:1700-1702.
- Holbrook, C.C., W.F. Anderson and R.N. Pittman. 1993. Selection of a core collection from the U.S. germplasm collection of peanut. *Crop Sci.* 33:859-861.
-

- IBPGR. 1991. Elsevier's Dictionary of Plant Genetic Resources. Elsevier Science Publishers, Amsterdam.
- Johnson, R.C., W.J. Johnston, M.C. Nelson, C.J. Simon and C.T. Golob. 1999. Core utilization and development – an example with *Poa pratensis*. Pp. 49-60 in Core Collections for Today and Tomorrow (R.C. Johnson and T. Hodgkin, eds.). IPGRI, Rome, Italy.
- Jung, H.G., C.C. Sheaffer, D.K. Barnes and J.L. Halgerson. 1997. Forage quality variation in the U.S. alfalfa core collection. *Crop Sci.* 37(4):1361-1366.
- Kaiser, W.J., M. Mihov, F.J. Muelbauer and R.M. Hannan. 1998. First report of anthracnose of lentil incited by *Colletotrichum truncatum* in Bulgaria. *Plant Disease* 82(1):128.
- Kannenberg, L.W. and D.E. Falk. 1995. Models for activation of plant genetic resources for crop breeding programs. *Can. J. Plant Sci.* 75:45-53.
- Knüpffer, H. and Th.J.L. van Hintum. 1995. The Barley Core Collection - an international effort. Pp. 171-178 in Core Collections of Plant Genetic Resources (T. Hodgkin, A.H.D. Brown, Th.J.L. van Hintum and E.A.V. Morales, eds.). John Wiley and Sons, UK.
- Mahajan, R.K. and Bisht, I.S. 1999. Sampling strategies for developing an Indian sesame core collection. *Plant Foods for Human Nutrition* (in press).
- Matthews, P. and M.J. Ambrose. 1994. Development and use of a 'core' collection for the John Innes *Pisum* collection. Pp. 99-107 in Evaluation and Exploitation of Genetic Resources, Pre-Breeding. Evaluation and exploitation of genetic resources: pre-breeding. Proceedings of the Genetic Resources Section Meeting of Eucarpia, 15-18 March 1994, Clermont-Ferrand, France (F. Balfourier and M.R. Perretant, eds.). Institut National de la Recherche Agronomique, Versailles, France.
- McFerson, J.R., W.F. Lamboy and S. Kresovich. 1996. Assessing user perceptions of genetic resource collections in crucifer crops. *Crop Sci.* 36:831-838.
- Miklas, P., R. Delorme, R. Hannan and M. Dickson. 1997. Using the core concept to identify new white mold resistance sources in common bean. *Bean Improvement Coop. Abstr.* 18.
- Noirot, M., S. Hamon and F. Anthony. 1996. The principal component scoring: a new method of constituting a core collection using quantitative data. *Genet. Resour. and Crop Evol.* 43:1-6.
- Ortiz, R., E.N. Ruiz-Tapia and A. Mujica-Sanchez. 1998. Sampling strategy for a core collection of Peruvian quinoa germplasm. *Theor. Appl. Genet.* 96:475-483.
- Ortiz, R., S. Madsen, E.N. Ruiz-Tapia, S.-E. Jacobsen, A. Mujica-Sanchez, J.L. Christiansen and O. Stolen. 1999. Validating a core collection of Peruvian quinoa germplasm. *Genet. Resour. and Crop Evol.* 46:285-290.
- Pederson, G.A., T.E. Fairbrother and S.L. Greene. 1996. Cyanogenesis and climatic relationships in U.S. white clover germplasm collection and core subset. *Crop Sci.* 36:427-433.
- Phippen, W.B., S. Kresovich, F.G. Candelas and J.R. McFerson. 1997. Molecular characterization can quantify and partition variation among
-

- genebank holdings: a case study with phenotypically similar accessions of *Brassica oleracea* var *capitata* L. (cabbage) 'Golden Acre'. *Theor. Appl. Genet.* 94:227-234.
- Radovic, G. and D. Jelovac. 1994. The possible approach in maize core collection development. Pp. 109-115 *in* Evaluation and Exploitation of Genetic Resources, Pre-Breeding. Evaluation and exploitation of genetic resources: pre-breeding. Proceedings of the Genetic Resources Section Meeting of Eucarpia, 15-18 March 1994, Clermont-Ferrand, France (F. Balfourier and M.R. Perretant, eds.). Institut National de la Recherche Agronomique, Versailles, France.
- Ruiz-Tapia, E.N., O. Stolen, J.L. Christiansen, R. Ortiz and L. Stolen. 1997. Assessment of quinoa (*Chenopodium quinoa* Willdenow) accessions for tolerance to salinity. *Crop Development for the cool and wet regions of Europe. Spelt and quinoa*, pp. 69-78. Working group meeting, Wageningen, Netherlands, 24-25 October 1997. European Commission, DG XII, Science, Research and Development, COST, Luxembourg.
- Schoen, D.J. and A.H.D. Brown. 1995. Maximising genetic diversity in core collections of wild relatives of crop species. Pp. 55-76 *in* Core Collections of Plant Genetic Resources (T. Hodgkin, A.H.D. Brown, Th.J.L. van Hintum and E.A.V. Morales, eds.). John Wiley and Sons, UK.
- Simmonds, N.W. 1962. Variability in crop plants, its use and conservation. *Biol. Rev.* 37:442-465.
- Skroch, P.W., J. Nienhuis, S. Beebe, J. Tohme and F. Pedraza. 1998. Comparison of Mexican common bean (*Phaseolus vulgaris* L.) core and reserve germplasm collections. *Crop Sci.* 38(2):488-496.
- Spagnoletti Zeuli, P.L.S. and C.O. Qualset. 1995. The durum wheat core collection and the plant breeder. Pp. 213-228 *in* Core Collections of Plant Genetic Resources (T. Hodgkin, A.H.D. Brown, Th.J.L. van Hintum and E.A.V. Morales, eds.). John Wiley and Sons, UK.
- Suneson, C.A. 1963. Breeding techniques - composite crosses and hybrid barley. Pp. 303-309 *in* Barley Genetics I. Wageningen.
- Tanksley, S.D. and S.R. McCouch. 1997. Seed banks and molecular maps: Unlocking genetic potential from the wild. *Science* 277:1063-1066.
- Tohme, J., O.D. Gonzalez, S. Beebe and M.C. Duque. 1996. AFLP analysis of gene pools of a wild bean core collection. *Crop Sci.* 36:1375-1384.
- Tohme, J., P. Jones, S. Beebe and M. Iwanaga. 1995. The combined use of agroecological and characterisation data to establish the CIAT *Phaseolus vulgaris* core collection. Pp. 95-107 *in* Core Collections of Plant Genetic Resources (T. Hodgkin, A.H.D. Brown, Th.J.L. van Hintum and E.A.V. Morales, eds.). John Wiley and Sons, UK.
- Tohme, J., S. Beebe and C. Iglesias. 1999. Molecular characterization of the CIAT bean and cassava core collections. Pp. 28-36 *in* Core Collections for Today and Tomorrow (R.C. Johnson and T. Hodgkin, eds.). IPGRI, Rome, Italy.
- van Hintum, Th.J.L. 1994. Hierarchical approaches to the analysis of genetic diversity in crop plants. Pp. 23-34 *in* Core Collections of Plant Genetic Resources (T. Hodgkin, A.H.D. Brown, Th.J.L. van Hintum and E.A.V. Morales, eds.). John Wiley and Sons, UK.
-

- van Hintum, Th.J.L. 1998. Concept and application of core collections in genebanks. Proceedings of a symposium "Züchterische Nutzung pflanzengenetischer Ressourcen – Ergebnisse und Forschungsbedarf", 29 Sep – 1 Oct. Gatersleben, Germany, ZDI, Bonn (in press).
- van Hintum, Th.J.L. 1999. The Core Selector, a system to generate representative selections of germplasm accessions. *Plant Genet. Resour. Newsl.* 118:64-67.
- van Hintum, Th.J.L. and D. Haalman. 1994. Pedigree analysis for composing a core collection of modern cultivars, with examples from barley (*Hordeum vulgare* s. lat.). *Theor. Appl. Genet.* 88:70-74.
- van Hintum, Th.J.L., I.W. Boukema and D.L. Visser. 1996. Reduction of duplication in a *Brassica oleracea* germplasm collection. *Genet. Resour. and Crop Evol.* 43:343-349.
- van Hintum, Th.J.L., R. von Bothmer and Dirk L. Visser. 1995. Sampling strategies for composing a core collection of cultivated barley (*Hordeum vulgare* s. lat.) collected in China. *Hereditas* 122:7-15.
- Veteläinen, M., E. Nissilä, P.M.A. Tigerstedt and R. von Bothmer. 1996. Utilization of exotic germplasm in Nordic barley breeding and its consequences for adaptation. *Euphytica* 92:267-273.
- Willner, E., U. Kastirr, F. Rabenstein, B. Boller and F.J. Stadelmann. 1998. Evaluation of European ryegrass core collection and possible use in breeding. Pp. 164-167 in *Breeding for a multifunctional agriculture. Proceedings of the 21st Meeting of the Fodder Crops and Amenity Grasses Section of EUCARPIA, Kartause Ittingen, Switzerland, 9-12 September 1997.* Swiss Federal Research Station for Agroecology and Agriculture, Zurich, Switzerland.
- Yonezawa, K., T. Nomura and H. Morishima. 1995. Sampling strategies for use in stratified germplasm collections. Pp. 35-54 in *Core Collections of Plant Genetic Resources* (T. Hodgkin, A.H.D. Brown, Th.J.L. van Hintum and E.A.V. Morales, eds.). John Wiley and Sons, UK.
- Zhang, X., Y. Zhao, Y. Cheng, X. Feng, Q. Guo, M. Zhou and T. Hodgkin. 2000. Establishment of sesame germplasm core collection in China. *Genet. Resour. and Crop Evol.* 47:273-279.
-



IPGRI is a
**FUTURE
HARVEST**
Centre

ISBN 92-9043-454-6