Accession Management

Sackville Hamilton, N. Ruaraidh; Engels, J. M. M.; van Hintum, Th J. L.; Koo, B.; Smale, M.

Published in:
Accession Management Trials of Genetic Resources Collections

Publication date:
2002

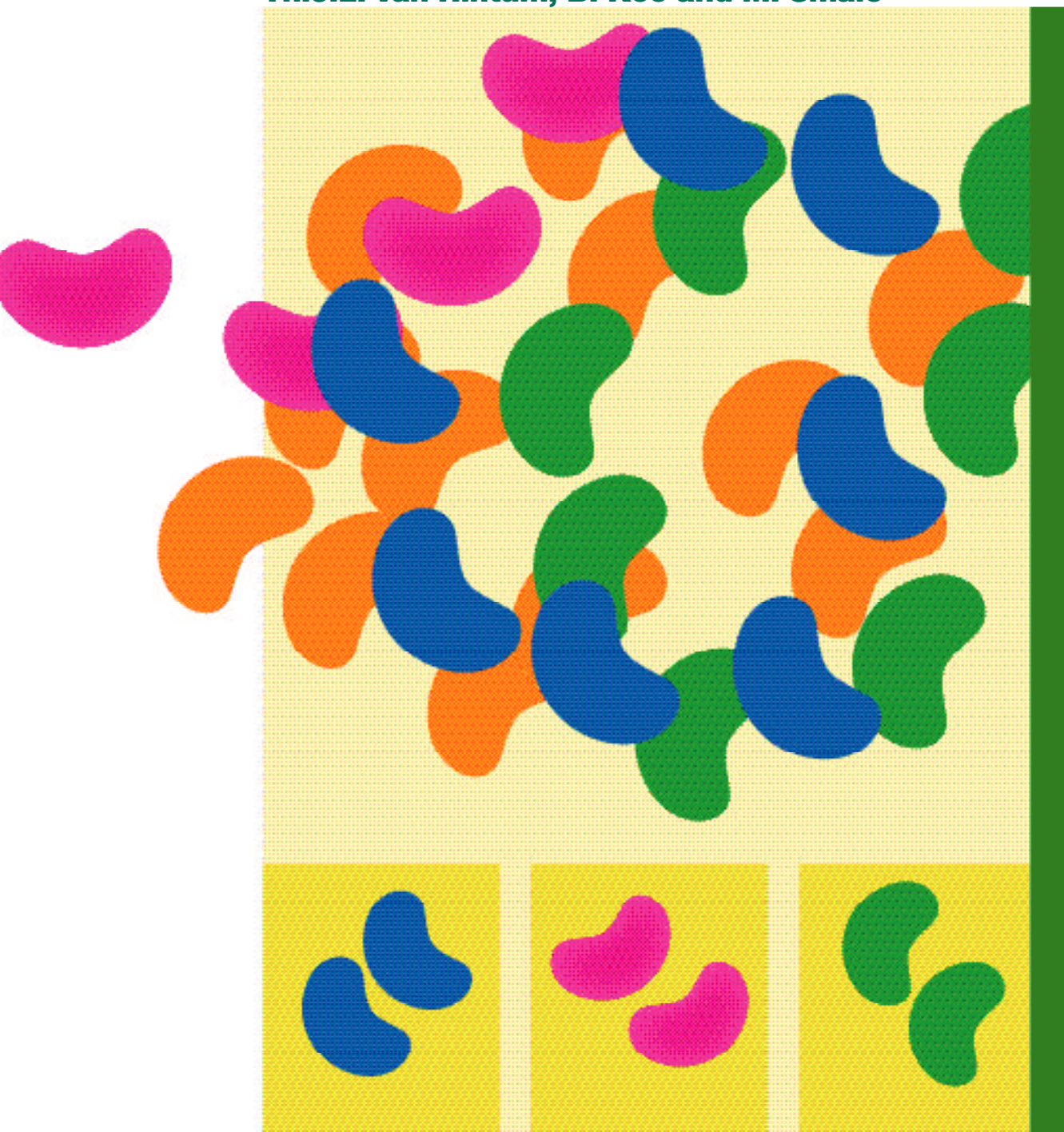
Citation for published version (APA):
Sackville Hamilton, N. R., Engels, J. M. M., van Hintum, T. J. L., Koo, B., & Smale, M. (2002). Accession Management: Combining or splitting accessions as a tool to improve germplasm management efficiency. In *Accession Management Trials of Genetic Resources Collections* Bioversity International.
http://hdl.handle.net/2160/3822

IPGRI TECHNICAL BULLETIN NO. 5

IPGRI

Accession management

N.R. Sackville Hamilton, J.M.M. Engels,
Th.J.L. van Hintum, B. Koo and M. Smale



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.

Previous titles in this series:

A protocol to determine seed storage behaviour
T.D. Hong and R.H. Ellis
IPGRI Technical Bulletin No. 1, 1996.

Molecular tools in plant genetic resources conservation: a guide to the technologies
A. Karp, S. Kresovich, K.V. Bhat, W.G. Ayad and T. Hodgkin
IPGRI Technical Bulletin No. 2, 1997.

Core collections of plant genetic resources
Th.J.L. van Hintum, A.H.D. Brown, C. Spillane and T. Hodgkin
IPGRI Technical Bulletin No. 3, 2000.

Design and analysis of evaluation trials of genetic resources collections
Statistical Services Centre and University of Reading
IPGRI Technical Bulletin No. 4, 2001.

Copies can be obtained in PDF format from IPGRI's Website (www.ipgri.cgiar.org) or in printed format by sending a request to ipgri-publications@cgiar.org.

Accession management

combining or splitting accessions as a tool to improve germplasm management efficiency

N.R. Sackville Hamilton, J.M.M. Engels, Th.J.L. van Hintum, B. Koo and M. Smale

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.

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 Maccarese, near Rome, Italy, with offices in another 19 countries worldwide. The Institute 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 2001, 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.

In 2000 financial support for the Research Agenda of IPGRI was provided by the Governments of Armenia, Australia, Austria, Belgium, Brazil, Bulgaria, Canada, China, Croatia, Cyprus, Czech Republic, Denmark, Estonia, F.R. Yugoslavia (Serbia and Montenegro), Finland, France, Germany, Greece, Hungary, Iceland, India, Ireland, Israel, Italy, Japan, Republic of Korea, Latvia, Lithuania, Luxembourg, Macedonia (F.Y.R.), Malta, Mexico, the Netherlands, Norway, Peru, the Philippines, Poland, Portugal, Romania, Slovakia, Slovenia, South Africa, Spain, Sweden, Switzerland, Thailand, Turkey, Uganda, the UK and the USA and by the African Development Bank (AfDB), Asian Development Bank (ADB), Center for Development Research (ZEF), Center for Forestry Research (CIFOR), Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), Centro Agronómico Tropical de Investigación y Enseñanza, Costa Rica (CATIE), Common Fund for Commodities (CFC), Technical Centre for Agricultural and Rural Cooperation (CTA), European Environmental Agency, European Union, Food and Agriculture Organization of the United Nations (FAO), Food and Fertilizer Technology Center for the Asia and Pacific Region (FFTC), Future Harvest, Global Forum on Agricultural Research (GFAR), Instituto Colombiano para el Desarollo de la Cienca y la Technología (COLCIENCIAS), Inter-American Drug Abuse Control Commission (CICAD), International Association for the Promotion of Cooperation with Scientists from the New Independent States of the former Soviet Union (INTAS), International Development Research Centre (IDRC), International Foundation for Science (IFS), International Fund for Agricultural Development (IFAD), International Service for National Agricultural Research (ISNAR), Japan International Research Centre for Agricultural Sciences (JIRCAS), National Geographic Society, Natural

Resources Institute (NRI), Programme on Participatory Research and Gender Analysis for Technology Development and Institutional Innovation (PGRA), Regional Fund for Agricultural Technology (FONTAGRO), Rockefeller Foundation, Taiwan Banana Research Institute (TBRI), Technova, United Nations Development Programme (UNDP), UNDP Global Environment Facility (UNDP-GEF), United Nations Environment Programme (UNEP), UNEP Global Environment Facility (UNEP-GEF), United States Department of Agriculture (USDA), Vlaamse Vereiniging voor Ontwikkelingssasamenwerking en Technische Bijstand (VVOB) and the World Bank.

Citation: Sackville Hamilton, N.R., J.M.M. Engels, Th.J.L. van Hintum, B. Koo and M. Smale. 2002. Accession management. Combining or splitting accessions as a tool to improve germplasm management efficiency. IPGRI Technical Bulletin No. 5. International Plant Genetic Resources Institute, Rome, Italy.

Cover: Design by Patrizia Tazza.

ISBN 92-9043-516-X

IPGRI
Via dei Tre Denari 472/a
00057 Maccarese
Rome, Italy

Acknowledgements

The authors acknowledge the contributions made to this publication by numerous genebank staff members around the world, through discussions with the authors as well as through the discussions after the presentation of the ideas during the International Conference on Science and Technology for Managing Plant Genetic Diversity in the 21st Century in Kuala Lumpur in June 2001. The keen interest of the System-wide Genetic Resources Programme of the Consultative Group on International Agricultural Research has been reflected in a substantial financial contribution to this work as well as through the contributions made by several of its members. Both contributions are duly acknowledged. The authors also acknowledge the written comments and corrections received on the draft manuscript from the following: Charles Nkhoma, Daniel Debouck, Florent Engelmann, James Were, KarLing Tao, Luigi Guarino, Ramanatha Rao, Rudolf Schachl, Steve Jarvis, Suketoshi Taba and Tony Brown.

Contents

Preface

Production of this publication has been triggered by the following observations:

(a) It is essential to document good accession management practices that have developed in practical genebank situations, as part of the process of establishing guidelines for optimal genebank management.
(b) To date there has been no systematic treatment of issues related to the genetic composition of genebank accessions.
(c) Genebank operations are becoming increasingly expensive, not only because of ever increasing collection sizes but also due to permanently rising labour and recurrent costs and the introduction of high-cost molecular and information technologies. This requires new approaches to optimize genebank operations, for example to benefit from the gains in productivity and performance now achievable through the new technologies.
(d) There is a need to ensure that genebank management practices are indeed optimal, based on appropriate scientific and economic principles.

Rising costs, falling budgets, developing molecular and information technologies, changing expectations are rapidly changing conditions, which present new challenges and opportunities that genebanks must face. To maximize the genetic and economic efficiency of conserving and utilizing *ex situ* collections of plant genetic resources, genebanks must develop innovative approaches more suited to modern conditions and make use of these opportunities.

This Technical Bulletin deals with one aspect of the broader objective of helping to improve genebank management in response to changing conditions. Specifically, it aims to encourage discussion and consideration of the optimum genetic composition of the unit of management that we call the accession: if it is not optimal, should we split or combine accessions?

The publication is aimed at curators and other genebank staff. It is meant to be a discussion guide, to provide ideas and suggestions on how management procedures can be improved, and to point to possible implications of a given management procedure. It does not aim to provide a definitive theory on composing genebank collections or on the management of accessions, nor does it attempt to provide a complete overview of possible approaches and procedures. The reader is invited to give feedback to IPGRI (comments, additions, alternative approaches, etc.) on

any aspect of the contents, and thus contribute to the discussion started with this publication.

The publication has had a long evolution. Following recognition of the problem, IPGRI developed an outline plan, which was revised and developed during a small brainstorming workshop in 2000 to create its final basic structure. This led to a first draft that was presented as a paper and discussed at the **International Conference on Science and Technology for Managing Plant Genetic Diversity in the 21st Century** (van Hintum *et al*. 2001). The first draft was finalized among co-authors, following which wide consultation was sought with many other members of the PGR community.

It should be noted that IPGRI is currently producing a publication on germplasm collection management strategies with very complementary ideas and approaches to this Technical Bulletin. It addresses conservation strategies in the national and institutional context, which are broader than the more technical procedures presented in this Bulletin.

Coosje Hoogendoorn
Deputy Director General Programmes

1 Introduction

1.1 The need for more effective and efficient germplasm management

In the past, the strategy for creating collections of plant genetic resources for *ex situ* conservation has often been straightforward: choose a crop, and gather some or all of the available samples. This approach has resulted in many huge collections. World-wide, more than six million samples are stored and conserved as seed and more than half a million are held in field genebanks, in more than 1300 genebanks (FAO 1996).

After a first period of collecting, documenting and evaluating accessions, the genebank community has gradually shifted towards a phase of developing strategies for improving the composition and management of collections. With increasing numbers of accessions conserved, increasing requirements for regeneration as seed lose viability and seed stocks become exhausted, decreasing budgets, and the changing role of governments as privatization increases, there is now a need to streamline conservation activities and increase efficiency. The increasing global emphasis on short-term solutions has further increased the need to justify and streamline long-term conservation, and consequently increased the need to ensure that decisions are optimal in the long-term.

Optimizing genebank efficiency involves choosing between many different conflicting demands for limited resources. For example, allocation of resources to running or improving management of a genebank must be assessed against allocating those resources to networking and sharing efforts with other genebanks. The potential benefits of investment in collecting new diversity must be judged against investing in better conservation or utilization of existing collections. Such broader issues are beyond the scope of this publication, which focuses on the issue of optimizing the genetic information contained in a given set of accessions (i.e. a collection).

Moreover, genebank management involves many issues that fall beyond the scope of this publication, such as methods for viability monitoring, germplasm health issues, and exploiting new molecular and information technologies. Here we are concerned only with the issue of optimizing the composition of individual accessions.

1.2 Accessions: optimizing the unit of genebank management

Decisions on what should be included in a collection and how the collection is organized are fundamental to efficient operation

of a genebank. In an *ex situ* collection, each clone or genotype may be managed as a distinct entity. More commonly, a population of genotypes is grouped together and managed as a single unit. For seed samples of obligate outbreeders, we cannot manage each genotype as a separate entity, and the management unit has to be a population of inter-breeding genotypes. Whatever its composition, this unit of management is termed the accession.

Grouping different clones or genotypes into a single management unit raises the question, 'What is the optimum grouping?' This question is not new. Decisions have had to be made concerning the composition of accessions ever since the first genebanks started conserving germplasm. The decision-making process starts during the collecting phase (Guarino *et al.* 1995), and continues during preparation of the samples for inclusion in the genebank collection. A good example is the common practice to separate the different genera or species in a mixed sample. In many genebanks it stops at that point. Other genebanks go further, by separating accessions of selfing crops into different morphotypes (Lehmann and Mansfeld 1957; Hammer *et al.* 1995).

It is often supposed that the unit of management should correspond to the sample of seed or plants received for inclusion in the collection. For example, material received as a single clone is usually managed as a single clone, and material received as a genetically heterogeneous population (e.g. a wild population or mixed landrace) is often managed as a single accession. In some cases this may indeed be appropriate. For example, to conserve a mixed landrace we should conserve not just all its component alleles but all its component genotypes (the combinations of alleles across all loci) at their original frequencies. Depending on its breeding system and the traditional system for maintaining the landrace, and on the conditions or understanding under which the material was obtained, managing the landrace as a single accession may be the only acceptable option.

However, a decision to manage each original sample as an accession is not necessarily optimal for efficient conservation or for efficient utilization. If the composition of existing accessions is not optimal, a genebank manager may consider restructuring accessions. Two options are available: accessions may be split or combined. The objective of this publication is to help genebank managers decide what is the optimal composition of accessions, and whether they should consider splitting or combining.

Increasing emphasis is being placed on reducing costs by eliminating or combining duplicates. It is obvious that if two accessions in a collection are identical in all aspects, one of them is

redundant and should be removed. However, except for pure homozygous lines, accessions will very rarely be identical duplicates. Therefore, the question "Are the accessions identical?" should be reformulated as "Are the accessions sufficiently different to justify managing them as distinct entities?"

Combining accessions is irreversible. Depending on breeding system and genotypic composition and the curator's knowledge of the original genotypic composition, splitting may also be at least partly irreversible. There is therefore a need to consider the consequences of making 'wrong' decisions, and to consider economically feasible ways of correcting wrong decisions by keeping 'backup' copies of the original.

It is also necessary to have due regard for the principles of the Convention on Biological Diversity (CBD 1992) and the International Treaty on Plant Genetic Resources for Food and Agriculture (CGRFA 2001). In particular, the sovereign rights of nations over their biodiversity must be respected. For example, if a curator would combine accessions from different countries of origin, this could lead to problems if one of those countries requested repatriation of their native accessions. It is possible that the only solution would be that the genebank also keeps the original samples, possibly in an archive collection. However, since the exact implications of these international agreements for accession management are not resolved yet, they will not be considered any further in this Technical Bulletin.

1.3 Management to achieve genebank objectives

A recurring and fundamental theme throughout this document is that management decisions depend critically on the objectives of the genebank, in line with its institutional mandate and mission. Genebanks differ in their objectives for conservation and utilization, in the economic constraints under which they operate and in the completeness of their collections.

For some genebanks, the need to reduce the cost of maintaining existing collections is becoming of over-riding importance. For others, genetic efficiency and the cost of filling gaps in the collection remain the primary objectives.

For some genebanks, emphasis is on efficient facilitation and promotion of the immediate utilization of the active collection as a working collection for breeding, pre-breeding and research. In this case, decisions on accession management should be aimed at maximizing the efficiency of identifying and deploying valuable genes to the user community. For other genebanks, efficient conservation is the primary objective. Many genebanks aim at both objectives, maintaining an active collection for utilization

and a base collection under optimal conditions for conservation. Such genebanks have the opportunity to make different decisions for their active and base collections, optimizing them independently for utilization and conservation.

Objectives for conservation also differ between genebanks. The focus may be on conserving certain combinations of genes or genotypes, such as landraces or wild populations evolved in a particular habitat. In this case, decisions on accession management must be aimed at maintaining the genetic composition of each accession as close as possible to its original state. Opportunities for splitting or combining accessions are then relatively few. Alternatively the focus may be on maximizing the total genetic diversity conserved in the collection, such as the total number of alleles or total genetic variance or genetic variance among accessions. In this case, maintaining the precise genetic identity of an accession is not as important as ensuring that its genes remain represented somewhere in the collection. Choices to split or combine accessions can then be a more important part of the genebank management process. A single genebank may span both these objectives. For example, it may hold some accessions that must be conserved in their original genetic state, such as heritage landraces, or accessions with detailed genetic data that are valuable only so long as they accurately describe the accessions. At the same time it may hold other accessions whose value lies not in their particular genetic composition *per se* but rather in their contribution to the total diversity conserved in the collection. Curators of such genebanks may make different management decisions for the different types of accession.

1.4 Economic pressures and genebank objectives

Obviously there is a positive correlation between the number of accessions being maintained and the costs of the genebank operation; more accessions means each operation must be undertaken more times. The relationship is not perfect since the cost per accession is not constant, but nevertheless, controlling the number of accessions is a key consideration in managing the genebank's budget. Confronted with increasingly severe financial limitations, genebanks must look critically at collection management procedures. The economic pressure is to limit or reduce the size of a collection, in order to meet the universal genebank objective of minimizing cost. The task facing the genebank manager is to reconcile this increasing economic pressure with other objectives.

The size of a collection may be limited by choosing not to add new accessions, and reduced by eliminating existing accessions.

Taking such decisions constitutes abrogation of responsibility for conserving those accessions, and as such should be undertaken only after due discussion with, and agreement of, the genebank's funding body, parent organization, user community and related genebanks. Ideally, it should be undertaken only as part of rationalization of collections across genebanks, with inter-genebank agreement on responsibilities for conservation to avoid unnecessary duplication. Such issues are beyond the scope of this publication.

The only other option for reducing the size of a collection is to combine accessions. The genebank manager needs to determine whether this option does in fact reduce costs and whether it facilitates or conflicts with other genebank objectives. For example, combining accessions requires additional documentation and could even increase costs per accession for other operations so much as to increase total costs. It can have major genetic consequences, such as increasing the loss of alleles or genotypes during regeneration. The genetic and operational aspects of combining accessions are dealt with in this publication.

Conversely, splitting accessions could be economically undesirable because it might increase costs. The general expectation is an increase in costs because of the increase in numbers of accessions to be managed, although, as discussed in this document, in certain situations it may reduce costs if it generates accessions that are easier to manage. It may help improve quality aspects, such as maintenance of genetic integrity, quality of documentation and ease of use; but these potential improvements come with a real economic price to pay. The genebank manager needs to determine whether splitting brings any quality improvements and, if it does, whether these benefits are worth the price to be paid. Genetic and operational aspects of splitting accessions are also dealt with in this publication.

1.5 The decision-making process

A decision concerning the genetic composition of an accession is always made, even if only implicitly, at the moment of including the material as an accession in the collection. Decisions can also be made at any point in *ex situ* conservation from the point of collecting onwards. At the collecting phase, the future desired composition of the accession could influence the sampling strategy, as will be elaborated in later sections. Decisions can be made when the collection is being reorganized or restructured or when the material in the collection is optimized for use. Decisions can be changed as more information is acquired on

accessions. If sub-samples for conservation are kept separately from sub-samples for use, different decisions may be made for different sub-samples.

Genebank management operations are interdependent; as a consequence of making a decision for one operation, we may then have to modify other aspects of genebank management. The decision for one operation must therefore not be based on analysis of that operation in isolation, but rather on optimizing the entire genebank performance. For example, harvesting plants individually in a regeneration plot can help reduce genetic drift and shift. However, the higher costs of doing so may make it economically unacceptable unless we use only a small number of plants per plot. So if we choose to harvest plants separately to reduce drift and shift, the consequence is we may have to reduce the number of parents, which has a genetic cost of increasing drift. We should choose balanced bulks (see Section 2.3) only if the genetic benefit for drift and shift exceeds the genetic cost for drift and the economic cost.

1.6 Structure of document

A conceptual framework for the management of germplasm accessions is presented in Sections 2–4. Genetic issues of the conceptual framework are described in Section 2 (i.e. distribution of genetic variation, correlation between characters, and reproductive characters), and the operational aspects in Section 3 (i.e. regeneration, characterization, evaluation and documentation, storage, viability monitoring, as well as facilitating use). Both reasons for and consequences of deliberate changes of the composition of accessions will be examined in both Sections 2 and 3. Section 4 considers the economics of optimization, formalizing the central objective, which is to maximize value at minimum cost. Based on these concepts, Section 5 presents the implications for the management of different categories of accessions, i.e. clones, true-breeding lines and outbreeders. These implications are illustrated with examples of actual cases where accessions were combined or split to improve efficiency and cost-effectiveness. Based on these theoretical considerations and the actual case studies, Section 6 summarizes some provisional recommendations.

2 Conceptual framework: genetic issues

In this section we outline the genetic issues that are relevant to a decision to split or combine, i.e. the genetic impact of splitting or combining relative to retaining the original designations. These genetic issues must be considered in the total context of the genebank storing the accessions, which includes the biology of the species concerned, the objectives of the genebank, the economic impacts of splitting or combining accessions, and the impacts of and consequences for optimal management of accessions.

The evolutionary processes of mutation, recombination, drift, selection and gene flow form an intrinsic, natural and necessary part of *in situ* conservation. The same processes also occur during *ex situ* conservation, except that evolutionary changes occurring in response to the *ex situ* environment involve qualitatively and quantitatively different patterns of drift, selection and gene flow. Their continuation *ex situ* thus represents a loss of genetic integrity of the conserved germplasm, which we aim to halt as much as is economically possible. Remembering that the rate of evolutionary change tends to be greatest when the environment and consequent selection pressures change, we expect high rates of change when a population sample is collected and transferred *ex situ*; the challenge to halt the change is therefore difficult.

The principles of *ex situ* evolutionary change during regeneration have been reviewed by Breese (1989). Alleles and genotypes may be lost from an accession by drift or by selection. Losses may occur by chance sampling effects when sub-sampling for regeneration or storage, by differential mortality during storage or during any stage of the regeneration cycle, and by differential seed set during regeneration. Losses by drift are stochastic[1] and are highest for rare alleles and genotypes. They therefore tend to be highest in genetically variable accessions, since such accessions tend to have most rare alleles and genotypes. Losses by selection are less stochastic and tend also to be highest in genetically variable accessions, although for different reasons and in different ways.[2] In general, therefore, the more variable an accession is, the more difficult is the task of maintaining its genetic composition intact.

So, if an accession is genetically variable, we may wish to consider splitting it to create two or more new accessions. To justify splitting an accession, the following conditions must be met.

---

[1] That is they occur by chance. This means that at best we can only determine the likelihood that an allele will be lost: we cannot determine exactly when it will be lost.

1. Splitting must significantly reduce genetic variance and/or increase the frequency of rare alleles (the subject of this section).
2. The lower variability of the resulting accessions must significantly improve their conservation and utilization (the subject of the next section).
3. The benefits for conservation and utilization must be worth the extra cost of managing more accessions (the subject of Section 4).

Conversely, if two or more accessions are genetically similar, we may wish to consider combining them to create a single new accession that is cheaper to conserve or utilize. The key question to resolve is whether combining two accessions increases variability to an extent that may interfere with efficient conservation or utilization. This is a two-part question:

1. What level of intra-accession variability is acceptable or desirable for efficient conservation and utilization? (The subject of the next two sections.)
2. How similar must two accessions be to achieve that level of variability by combining them? (The subject of this section.)

Three major genetic issues determine the effect of splitting and combining on the genotypic composition of the resulting accession(s). The first is the distribution of the genetic variation for a character, the second the possible correlation between characters, and the third is the role of the character in reproduction. We shall consider each of these in turn.

2.1 Distribution of genetic variation for a character

The possibility of combining accessions on the basis of a character, or separating one accession into groups with different genotypes for a character depends on the ease of scoring the character, the distribution of its expression, and its genetic background. If it is

---

[2] The theory of response to selection is complex (see e.g. Mather 1973; Mather and Jinks 1971). The most basic rule is Fisher's (1930) fundamental theorem of natural selection, "the rate of increase in fitness of any organism is equal to its genetic variance in fitness at that time". Exactly the same principle applies to artificial selection. Crudely, selection operates through differences in fitness, i.e. differences in reproduction and mortality. The rate of response to selection for a character is proportional to the heritability of that character and the strength of selection, and depends also on the mode of inheritance of the character. Selection for one character will also cause changes in other characters that are pleiotropic expressions of the first or that are genetically linked to the first. Consequences in terms of probabilities and rates of loss of the alleles selected against depend further on the mode of inheritance of the character: for example, recessive alleles are less likely than dominant alleles to be eliminated by selection.

difficult or impossible to observe the character state, it obviously is not a good criterion for combining or separation. If the expression of the character has low heritability, it is also not a good criterion for combining or separating.

If the expression of a character within an accession follows a discrete distribution function i.e. with a number of distinct states, and each state corresponds to a genotype, and it is easy, cheap and quick to identify the distinct genotypes, then splitting may be a workable option. It is still more likely to be a workable option if the species is inbreeding and the accession comprises a set of homozygous, true breeding lines. Then splitting may produce accessions that are pure lines at least for the character concerned. A small number of states is also better, because a large number of states would require creation of a large number of new accessions, which would have a high cost.

If the expression follows a continuous distribution, due to a high environmental effect or a large number of genes involved in the expression, splitting the accession to form two or more new accessions is usually less beneficial. Splitting may produce accessions with lower genetic variance within accessions than the original, but the extent of reduction is smaller, especially if heritability is not high. The resulting small reduction in within-accession genetic variance will in most cases have relatively small impact on drift and selection, and the increased costs of maintenance are unlikely to be justifiable. Fitness characters are an exception to this rule, as discussed below.

The efficacy of splitting is still lower for outbreeding populations, since they contain heterozygotes with genetic variation 'hidden' as recessive alleles in heterozygotes. In this case, even if identification of parental genotypic states is perfect, the progeny will include genotypes that do not conform to the parental state. In outbreeding populations, it is also necessary to avoid inbreeding depression, which may be caused by splitting using too few plants for a new accession in an attempt to reduce variance.

Conversely, forming one new accession by combining two or more original accessions may be a workable option if the accessions have similar distributions of genotypes and it is easy, cheap and quick to identify the distribution of genotypes. This will result in little increase in drift and selection, or even no increase if they are identical. It is still necessary for the character to have high heritability, since otherwise we cannot confidently measure their similarity. In contrast, the presence or absence of heterozygotes and outbreeding will not strongly influence a decision to combine. Accessions should not be combined if they are homozygous for different discrete states.

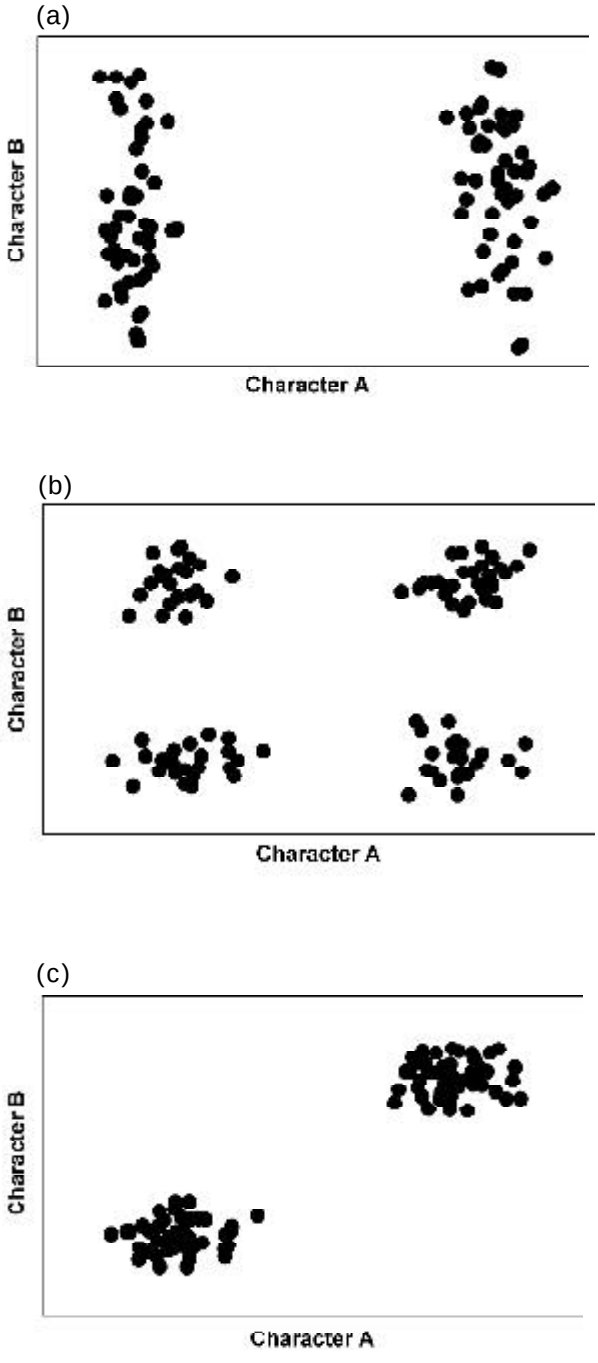
Thus, to be potentially useful in decisions on splitting or combining accessions, a character must be highly heritable and easy, cheap and quick to measure. To be useful for splitting accessions, a character must be present within accessions as a small number of discrete homozygous states. To be potentially useful for combining accessions, the character may be continuous or discrete, homozygous or heterozygous; we just need to be able to quantify genetic variation within and between accessions.

At this point we will have decided either that a character is definitely not useful for splitting or combining, or that it might be useful. If it is an agronomically important character but difficult or expensive to measure, it may be appropriate to split solely on the basis of that character, as discussed in the next section. If a character state is rare within a given accession so that many plants of that accession must be screened to find another example, it may be appropriate to split opportunistically when the rare character state is detected. More generally, if we decide a character might be useful, whether it is indeed useful then depends on further genetic considerations discussed in the next two sections.

2.2 Correlations between characters

If genetic variation within an accession for one character is not correlated with genetic variation in other characters, then splitting an accession on the basis of that character is likely to be relatively ineffective. If the character is present as a small number of discrete, easily identifiable homozygous states, splitting will be effective for the character concerned, but without correlations it would not reduce genetic variation within accessions for any other character. This is illustrated in Fig. 1 (a) and (b), in which splitting on the basis of character A would not help reduce variation in character B. The only acceptable reason for splitting in this situation would be that character A is expensive to measure or of high importance for current utilization, such that it is considered desirable to fix the genotype of that character regardless of the rest of its genome.

Conversely, if genetic variation within an accession for one character is correlated with genetic variation in other characters, the character is more likely to be useful as a basis for splitting accessions. Splitting an accession in this way first requires that the character be present as a small number of discrete, easily identifiable homozygous states, but its efficacy depends on the magnitude and cause of correlation. Under appropriate conditions, splitting will reduce variance not only for the



character concerned but also for the other correlated characters, as illustrated in Fig. 1(c).

Similarly, knowledge of correlations between characters is also required for decisions on combining similar or duplicate accessions. If two homozygous accessions are historically duplicate (i.e. derived by random sub-sampling from the same original seed lot), they would need to be scored for only a few distinguishing characters to confirm whether they are biologically duplicate. In the case of outbreeders, and of inbreeders without prior expectation of being duplicates, many more characters are needed.

Different characters may be correlated by three mechanisms. First, they may be pleiotropic expressions of a single gene locus. For example, genes controlling 'size' may simultaneously influence the size of leaves, petioles, flowers etc. Expression of pleiotropic characters will inevitably be correlated. This form of correlation should be disregarded since it involves the same loci.

Second, characters may be correlated through genetic linkage. This occurs when the gene loci controlling the characters are so close together on a chromosome that they tend to be inherited together. Information on linkage between genes is becoming widely available in the form of genetic linkage maps for many species (e.g. Lombard and Delourme 2001; Ortiz *et al.* 2001): defining length in terms of linkage, chromosomes tend to

Fig. 1. Some hypothetical distributions of character expressions within an accession, illustrating the importance of correlations between characters. In all cases, character A shows a discontinuous distribution with two distinct states, and is therefore a potential candidate for splitting accessions. Splitting into two accessions on the basis of the character state for character A will also reduce variance within-accessions for character B only in example (c) where the two characters are correlated.

be of the order of 100 centiMorgans (cM) in length.[3] A few species have mechanisms that unite the entire genome into a single 'supergene' (such as the translocation ring joining all 14 chromosomes of most species of *Oenothera*: see e.g. Chapman and Mulcahy 1997). In this case, splitting by one character will effectively reduce intra-accession diversity of all characters. However, this is an extreme case. In most species, only a few characters will be correlated by genetic linkage, and splitting will help reduce variance only for these characters.

Thirdly, characters may be correlated by descent from a common ancestor: all the descendants of one plant share the set of characters inherited from the common ancestor. For mixtures of inbred lines or populations of obligate inbreeders or apomicts, the correlation by descent is complete and the complete correlation spans the entire genome. In these cases, splitting an accession on the basis of one appropriate character is fully effective in reducing intra-accession diversity for all loci. For wild populations of inbreeding species, the correlation will also span the entire genome but may be less complete because selfing rates are rarely 100%. Correlation by descent can be increased by forced inbreeding, e.g. by preventing cross-pollination during regeneration, and this can be a useful pre-cursor to splitting.

The genetic neighbourhood area of wild plant populations (Crawford 1984) is usually remarkably small. It is usually much smaller than the area sampled as a single accession by collectors (Sackville Hamilton and Chorlton 1995; Hayward and Sackville Hamilton 1997). The area sampled may then comprise a set of genetically distinct, but often overlapping, sub-populations. In such cases the original sample of collected seed, if treated as a single population, shows a relatively high level of apparent inbreeding even if the species is an obligate outbreeder, and plants from the same sub-population will have high correlation by descent. One cycle of seed multiplication is enough to remove this correlation completely if the sample is grown as a single panmictic unit. However, depending on collecting and storage methodology, in some cases it may still be possible to classify seed according to their original sub-population, and this can provide a good basis for subdividing samples of variable

[3] One cM is defined as the length of DNA separating two genes between which there is a 1% probability of recombination during meiosis. On average, 1 cM is roughly equivalent to 10^6 nucleotide base pairs, although this varies widely between 'hot' and 'cold' spots of recombination.

wild populations, regardless of breeding system, into more homogeneous units.

2.3 Reproductive characters

Variation in reproductive characters within an accession has a direct impact on drift and selection, i.e. on random and non-random changes in genetic composition of the accession.

Micro-environmental variation within regeneration plots results in drift. For example, if one plant in a regeneration plot happens to be in a relatively nutrient-rich patch of soil, that plant may have a higher yield and so contribute more pollen and seed than the other plants to the next generation. This represents a chance change in genetic composition of the accession in favour of the genotype of that plant. Since this is purely a chance factor, with no systematic bias favouring any plant of that genotype, by definition it constitutes random drift and not systematic selection.

Conversely, genetic variation for reproductive characters results in selection. That is, any genotype that increases male or female contribution to the progeny generation by definition results in a systematic increase in frequency of that genotype in the progeny seed. The direction and strength of selection may vary with year, location (including location of the regeneration site and location of the plot within the regeneration site) or management.

The existence of any kind of heritable variation within an accession in fitness characters may provide justification for splitting it. Conversely, the existence of any kind of heritable variation between two accessions in reproductive characters may rule out the possibility of combining them.

For example, the number of flowers produced by a plant is a major determinant of its male and female contribution to the progeny generation. Heritable variation in male and female contributions constitutes selection pressure favouring the most productive genotypes. The selection also favours any character that contributes to high pollen and seed production in the regeneration environment (such as plant size, growth rate, developmental triggers controlling the onset of flowering, thermotolerance when regenerating in a hot climate, etc). Splitting an accession into high- and low-yielding genotypes may help reduce the selection even if the distribution is continuous. The potential benefit of splitting by fecundity must be compared with alternative approaches. Alternatives include pruning to equalize the number of flowers or inflorescences, and taking a balanced bulk (i.e. harvesting seed separately from each mother plant, taking

an equal number of seed from each and bulking them to form the accession). A balanced bulk eliminates both drift and selection through female contributions but does not control male contribution, whereas splitting reduces both female and male fitness differentials but does not control drift.

Another important character is maturity date. If intra-specific genetic diversity for maturity date is high, there may be no single satisfactory harvest date. An early harvest will select for earliness by eliminating the late plants. A late harvest may select for lateness if ripe seed drop from the early plants before harvest. An intermediate harvest may lose variation by eliminating both early and late plants. Splitting an accession by maturity date may help reduce such selection. As with other reproductive characters, it will also reduce selection for all characters that contribute to earliness or lateness. An alternative approach is to take multiple non-destructive harvests, ideally all of the same size.

In summary, splitting accessions can improve conservation only if it is highly effective in reducing variation within accessions. This condition applies primarily to inbreeding species, but in the case of reproductive characters it can be used for a wider range of species. Combining accessions can be economically desirable and genetically acceptable for all types of species, but only if the economic benefits of combining exceed the costs of identifying which accessions to combine. This means it must be easy to quantify genetic variation within and between accessions, and there must be some prior expectation of duplicity (i.e. identical passport data), so that observing genetic duplication for a few alleles can be extrapolated to assume genetic duplication at all alleles.

3 Conceptual framework: operational issues

3.1 Introduction

Splitting increases and lumping decreases the number of accessions in a collection. This has direct implications for all aspects of genebank management. In general, increasing the number of accessions will increase the number of times each operation must be repeated, so that splitting accessions will tend to increase the cost of all genebank operations. If the genebank's budget is fixed, the curator cannot increase the cost of genebank operations; then increasing the number of accessions can be accommodated only by reducing the resources that are available for conserving and utilizing each accession. However, this universal effect has to be balanced with other effects, which can be slightly more complicated. For example, if management protocols can be simplified for genetically uniform accessions, a large number of uniform accessions may be cheaper to conserve than a small number of variable accessions. Thus there is not a simple direct relationship between the size of a collection and the cost of maintenance, and total cost of conserving many uniform accessions is not necessarily higher than that of conserving fewer more variable accessions. These effects will be discussed below separately for routine genebank activities.

3.2 Regeneration

The main issue here is maintenance of the genetic integrity of the accession during regeneration, and the price to pay. The greater the number of accessions, the less input for regeneration per accession is available for a given genebank capacity. But reducing the number of accessions by lumping usually increases diversity within the accessions; and the more diversity within the accession, the higher the magnitude of shift and selection if not properly managed. Choosing to split or combine accessions may therefore require changes to the regeneration protocol (Sackville Hamilton and Chorlton 1997), and consideration of the costs and implications of these changes must be included in the decision-making process.

3.2.1 Maintaining genetic integrity

The major threats to loss of genetic integrity during regeneration are drift as a result of random effects, selection as a result of a higher fitness of some genotypes, and contamination with alien pollen or seed (Sackville Hamilton and Chorlton 1997). In general, the more diversity there is in an accession, the higher

the threat of drift and selection, and the more difficult it is to recognize contamination. At one extreme, for an accession consisting of one homozygous genotype of an inbreeding species (e.g. a modern variety of barley) there is no risk of loss of genetic integrity by drift or selection. The only risk is of contamination by mechanical mixing with seed other accessions, and the regeneration protocol must be chosen to minimize that risk and to identify contaminants.

For a mixture of genotypes (irrespective of whether homo–zygous or heterozygous, inbreeding or outbreeding), care must be taken to ensure that the composition of the accession stays the same. The regeneration protocol should seek to avoid selection and drift, since both processes will result in a change of the frequencies of genotypes, or even the loss of genotypes. Furthermore contamination is more difficult to detect since the accession already contains different types. Therefore, maintaining genetic integrity of a heterogeneous accession is more demanding than for homogeneous homozygous accession, and can therefore be more expensive.

Splitting one heterogeneous accession to create several more homogeneous accessions might prove beneficial for achieving these aims with the given resources, particularly with inbreeding species. If the process of splitting is taken too far with outbreeding species, there is a risk that the resulting relatively uniform lines will suffer inbreeding depression.

However, even for an 'easy' crop such as barley, maintaining the genetic integrity of genebank accessions can prove very difficult. A recent study revealed that the effective population size in barley regenerations using an estimated 600 plants, was only 4.7 (Parzies *et al.* 2000). van Hintum and Visser (1995) showed that duplicate barley accessions had developed into quite different mixtures in different genebanks. Both studies looked at the consequences of historical protocols, and therefore may not accurately reflect the consequences of modern regeneration protocols.

3.2.2 Prioritizing prevention of contamination, selection and drift

Different measures are required to prevent contamination, selection and drift. This means that, with limited resources, the curator may have to choose which component is most important to prevent. For example, avoidance of drift involves the use of a large number of parent plants, but increasing the number of plants per accession reduces the resources that can be invested in each plant, and this may prevent effective application of measures to control contamination and selection.

This need for prioritizing the factors causing loss of genetic integrity complicates the decision on whether regeneration protocols should be altered following splitting or lumping. Part of the decision-making process should include a consideration of whether the curator should also change the relative priorities attached to contamination, selection and drift.

In some cases, a valid strategy is to use a small number of plants per accession despite the resulting greater loss of diversity due to drift. The reason for this is the easier management of the regenerations: less space is needed per accession and more care can be taken per accession. For example in one case curators chose to take only 30 plants per accession of the outbreeding ryegrass, so that using balanced bulks and isolation chambers became economically feasible. Possible drift is accepted, since eliminating pollen contamination and reducing selection were considered higher priorities.

In another case, a curator chose only eight parent plants per accession of the inbreeding crop lettuce. The intra-accession diversity was considered small and relatively unimportant, and the cost per plant was very high as a result of the complicated and expensive treatments of the individual plants during regeneration such as cutting the heads, GA_3 treatment, staking, etc. The acceptance of drift allowed inclusion of a considerably higher number of accessions in the collection.

3.3 Characterization, evaluation and documentation

Intra-accession variation complicates efforts to describe and record character states. Although some systems for recording this diversity have been devised, it remains a problem for both qualitative and quantitative diversity. As a result, it is often neglected although the deviating minority genotypes might determine the value of an accession. Furthermore the experiments needed for efficient description of intra-accession diversity require more plants and are therefore more expensive than those where only the majority or the average is described.

3.3.1 Qualitative characters

To be able to score intra-accession diversity of qualitative characters, the experiment should allow the expression of differences between the character states. Depending on the genetic background and the minimum frequency that should be detected, a considerable number of plants may be needed to reliably detect an allele. For example, if a recessive allele occurs at a frequency of 5% in a panmictic population, 1197 plants must be observed to be 95% sure that the allele is detected.[4]

Moreover, once the intra-accession diversity is observed, the question arises how it should be recorded and documented. Several systems have been devised to solve this problem, including a scoring protocol that denotes the presence of diversity without specifying what, the recording of estimates of the frequencies of each score, or intermediate approaches (e.g. Rana *et al.* 1991; van Hintum 1993). Further complications arise at the point that the scores need to be retrieved, e.g. how to query the database, how to extract data in a form suitable for statistical analysis, how to combine data scored or stored following different systems.

3.3.2 Quantitative characters

In the case of quantitative characters there is another problem connected to intra-accession diversity, namely that environmental variance is entangled with genetic diversity. Large scale experiments are needed to quantify the genetic component. This is very rarely done, and consequently within-accession variation for such characters will often simply be neglected altogether.

Splitting accessions into more homogeneous accessions may help reduce these problems, especially for inbreeding species. The extra costs of evaluating heterogeneous accessions need to be assessed against the cost and genetic efficiency of separating accessions into uniform components. In the case of lumping accessions, care should be taken that these problems are not created or aggravated.

3.3.3 Documenting management decisions

A choice to split or combine of accessions itself needs to be documented. Data on the original accession(s) should be retained on the curator's database regardless of whether seed of the original accession(s) are retained in the seed store. Pedigree information must be recorded relating the new to the original accession(s). Some of the required data elements should already be available in the database, such as who did it and when. The database manager may need to establish new elements of pedigree information, such as the reason and methodology for splitting or combining.

---

[4] The chance of finding a phenotype occurring with the frequency f_p in a sample of n plants is $1-(1-f_p)^n$. The frequency of a phenotype (f_p) corresponding with a homozygous recessive allele occurring with frequency f_g in a panmictic population will be f_g^2.

3.4 Storage

The most obvious effect of splitting or lumping accessions on the storage of genetic material will be, as for other operations, that of the numbers of accessions. It can also have implications for other aspects of storage.

3.4.1 Number of seeds per accession

The minimum number of seeds that must be stored and distributed depends in part on the requirement to maintain the genetic diversity within the original sample, and to distribute that same range of diversity to users. If there is no diversity in the accession, or the curator accepts the loss of diversity, the minimum numbers can be smaller. This might be relevant especially in the case where the size of the seeds determines the capacity of the collection or where the size of container limits the number of seeds that can be stored for each accession. If the crop is inbreeding, it might be preferable to split the accessions up into homogeneous sub-samples. This would allow the curator to regenerate on the basis of only the number of plants needed for sufficient seed production, and to send users only very few seeds.

3.4.2 Active and base collections

Many genebanks maintain separate active and base collections. The base collection is held in optimal conditions for long-term storage, and exists for optimal conservation. The active collection exists to facilitate utilization, and is held in conditions where access is easy.

The existence of more than one seed sample of each accession raises the possibility that different management decisions can be applied to different samples. For example, if combining accessions facilitates utilization but adversely affects the quality of conservation, the curator may consider combining them in the active collection but retaining the original accessions in the base collection.

This concept can be taken further as an insurance against wrong decisions. For example, if two accessions have been identified as duplicates and therefore combined, the curator may consider 'archiving' small samples of the original accessions in optimal storage conditions along with the base collection. Then, if future evidence suggests the decision to combine was wrong, the curator can revert to the original accessions. Such a decision is feasible where the marginal costs of archiving is low, for example where the long-term storage facility has unused space available. The process of identifying duplicates will usually involve a statistical analysis that enables a degree of confidence

to be attached to the statement that two given accessions are duplicates. To reduce costs of archiving, a curator may choose to archive the original accessions only for the decisions where the level of confidence is relatively low.

3.4.3 Multiple containers

A completely different implication of accession management on seed storage relates to the conservation of wild populations that are subdivided into genetically distinct sub-populations. Storage of the population sample in a single container loses this genetic population structure, permits random crossing between genotypes that would not normally cross and therefore creates new recombinants, and a panmictic population that is more variable than the original subpopulations and that is therefore subject to greater drift and selection.

An option to help overcome this problem is to use a different container to store the seed of each subpopulation in the base collection while managing the entire population as a single accession. That is, seed of all subpopulations would be mixed to form a single seed sample in the active collection, and documentation and utilization would be based on the entire population. Repeat regeneration for utilization would involve taking an equal number of seed from each subpopulation to form a mixed parental generation. Benefits of such a system include improved control of drift and shift during regeneration, and the potential to study genetic subdivision of populations. It may be justified if the additional costs are low.

An alternative option is to split into sub-populations completely and maintain each as a different accession. The additional costs of doing so are potentially high, since the number of accessions to be regenerated, documented and characterized is multiplied by the number of subpopulations. It may be justified if the population comprises a small number of highly distinct sub-populations.

A decision on use of multiple containers must be made at the time of collection, since once seeds are mixed, it is not possible to return to the original subdivided population structure.

3.5 Monitoring viability

The main issue here is that splitting or combining directly affects the numbers of seed lots whose viability must be monitored. One additional point, however, needs to be raised. In the case of variable accessions, there might be a difference of viability between the different components. Consider the case of a mixture of two genotypes, of which one occurring in 90% 'stores

well', and the other with frequency 10% 'stores less well'. At a certain moment in time the small fraction will have completely died whereas the viability test does not indicate the necessity to regenerate.

In theory this danger could be reduced by reducing the intra-accession diversity, although this is likely to be difficult. Detecting intra-accession diversity for seed longevity is exceptionally expensive and slow. Splitting on the basis of other characteristics will not reduce intra-accession diversity for seed longevity unless they are correlated, such as in the case of mixtures of homozygous lines. Moreover, non-genetic variation in seed storage characteristics, e.g. associated with harvesting at different stages of seed maturity, can mask any genetic variation in some species. Therefore, it is generally not cost effective to attempt to reduce this danger.

An equivalent situation can arise when combining accessions if the seed lots being combined have different seed quality. This situation should be avoided by always rejuvenating seed before combining to ensure the same seed quality.

3.6 Facilitating use

Facilitating the use of genebank material is a key element in genebank operation. For the user the quality of a genebank depends to a large extent on the ease of use of the conserved germplasm. Combining or splitting accessions can have a direct effect on this.

3.6.1 Single plant selections in bulked accessions

If the user is looking for character states that can be observed from single plants in a population, there might be a preference for highly variable accessions, allowing the screening of many plants from a limited number of accessions. If a collection consists of 1000 accessions that can be classified in 50 groups, it might be an option to include in the active collection 50 bulked samples each containing all accessions within the each group. A user might prefer to evaluate the 50 bulked accessions rather than the 1000 individual accessions, if the character states of interest can be more easily found in the bulked samples. This is likely to apply only for highly visible character states such as disease resistance in a heavily infected plot, and only where it is possible to easily to recognize individual plants within a plot, and only if the scoring system is tailored to recording individual plants. The system may also have benefits for the curator, since it reduces the number of accessions that need to be distributed. The curator will need to consider whether all users prefer to

evaluate bulked accessions, or whether it is necessary to retain the original as well as the bulked accessions in the active collection.

Maintenance of genetic integrity of the bulks is likely to be more difficult. The curator should therefore consider bulking in the active collection only, and retaining the unbulked accessions in the base collection.

3.6.2 Evaluating whole populations

Many characters are scored on whole populations rather than single plants, by choice or necessity. It is necessary in grass swards because it is difficult to distinguish single plants within a dense population. It also applies for characters where the contrast between desirable and undesirable genotypes is not visually striking, so that it is difficult to detect elite plants within a population. It also applies where the user's methodology is based on scoring whole populations, and observing individual plants is inconvenient. In all three cases, alleles present at low frequency will not be detected, and identifying elite populations from among a large number of relatively uniform accessions will be more efficient than trying to identify elite individuals within a few populations.

In these cases, improving the quality of genebank accessions from the user perspective can involve selections within accessions or splitting accessions into more homogeneous groups.

In addition, modern varieties must be genetically uniform. When using landraces or wild populations, breeders have to select for uniformity. In this case, splitting variable accessions constitutes pre-breeding that helps achieve the level of intra-accession diversity that the breeder requires.

4 Conceptual framework: economic issues

4.1 Introduction

Analysis of the costs and benefits of genebank operations is fundamental to efficient genebank management. In the context of this document, we need to know quantitatively the economic implications of splitting or combining accessions, relative to leaving them 'as is'. However, no simplified methodology is on hand that genebank managers can use to conduct a comprehensive cost-benefit calculation. This is because of conceptual and measurement problems associated with the benefits calculation. There is sufficient methodology, however, to enable genebank managers to conduct a thorough analysis of costs. With this methodology, managers can assess the relative cost of their operations. Moreover, the methodology is an appropriate one when managers have the objective of minimizing costs subject to maintaining some level of viable, genetic diversity. For this section we will briefly summarize what is known, highlight what is missing from our economic understanding and then outline the issues that should be addressed.

4.2 A basic economics of genebank operation

The most comprehensive information available to date details the costs of conservation as estimated by compiling data from records kept by genebank managers (Burstin *et al.* 1997; Epperson *et al.* 1997; Pardey *et al.* 2001). Pardey *et al.* (2001) give comprehensive details of the production economics theory and calculations they have used to construct cost estimates for the maize and wheat genebanks held at CIMMYT. The essential notion of production economics is that outputs are produced with some combination of inputs. The institutions and technological environment that prevail at a point in time predetermine the combination of inputs, though these factors change over time. Applied to the case of a genebank, the inputs of labour, equipment, and acquired seeds are processed to produce outputs in the form of stored, viable seeds and accompanying information. Properly stored seeds and relevant information can be disseminated immediately for current use, or placed in the storage facility as options that can be exercised (repeatedly if necessary) in future years.

4.2.1 Costs categories

Costs of inputs to the genebank operation are broadly classified as variable, capital, and quasi-fixed. Variable inputs are

those that are sensitive to the size of the operation, capital inputs as those that are not, and quasi-fixed inputs as a group of inputs that are neither fixed nor variable, but 'lumpy.' A quasi-fixed input is 'lumpy' in the sense that it is a discrete, indivisible unit that cannot be adjusted easily with fluctuation in the extent of genebank operations; it is variable in that it is more easily adjusted than a capital item such as the building itself. In the framework outlined by Pardey *et al.* (200), skilled labour with scientific expertise, such as the genebank manager and laboratory scientists, are classified as quasi-fixed inputs. Technicians and temporary workers, or those paid on an hourly basis, are treated as variable labour inputs. Over a sufficiently small size range—i.e. the size range over which we would not alter the complement of skilled labour—quasi-fixed inputs can be regarded as fixed, so that costs can be classified simply into fixed and variable.

Consider the effect of varying the size of the genebank operation (Fig. 2). By definition the fixed costs do not vary with size, so are represented by a horizontal line on a plot of total cost vs. size. Some variable costs would vary in direct proportion to the size of the operation; for example the number and therefore total cost of labels used for regeneration plots will increase in direct proportion to the number of accessions regenerated each year. However, overall the 'law of diminishing marginal returns' is typically assumed for production factors as the size of the operation becomes large. This law reflects the classical proposition, still widely recognized as empirically valid, that the physical productivity of an individual factor, such as labour, declines as more is added while all other inputs are held constant.

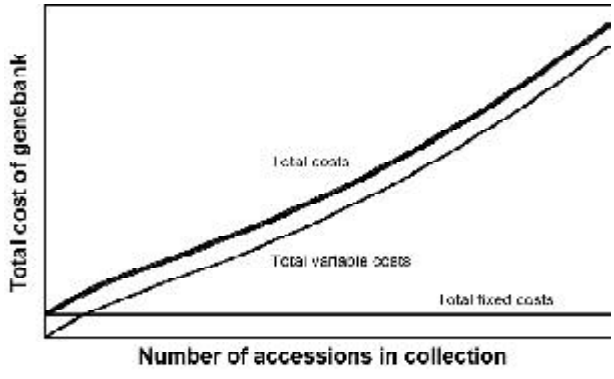


Fig. 2. Hypothetical breakdown of the total costs of running a genebank into fixed and variable costs, as a function of the number of accessions held in the collection.

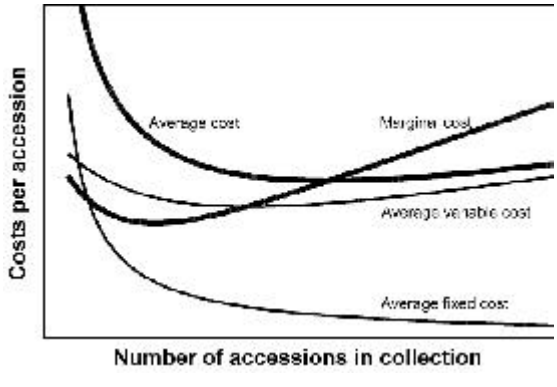


Fig. 3. Hypothetical average and marginal costs per accession. The values correspond to the case depicted in Fig. 2.

Diminishing marginal returns is reflected in Fig. 2 in as an upwards curvature in the line for total variable cost towards the right of the graph.

4.2.2 Costs per accession

Costs per accession can be expressed in two ways (Fig. 3), as *average* and *marginal* costs.

The *marginal* cost per accession is the increase in total cost of the genebank caused by adding one more accession to the collection. By definition this increase in cost involves only variable cost elements. The law of diminishing marginal returns is reflected in the increase in marginal cost as the collection becomes large.

The *average* cost per accession is the total cost of running the genebank divided by the number of accessions in the collection. It has two components, *average fixed cost* and *average variable cost*. Average fixed cost is the total fixed cost divided by the number of accessions, which, because total fixed costs are constant, monotonically decreases as the number of accessions increases. By contrast, average variable cost is in general U-shaped. As the number of accessions increases from a small size, the operation becomes more efficient and average variable cost decreases. After a certain minimum level of cost, it increases with the number of accessions due to excessive use of variable resources given fixed factors.

Because of the fixed costs, average cost per accession typically decreases as the size of the collection increases. It reaches a minimum at the point where average cost = marginal cost. Above this point, average cost increases with the number of accessions, but this reflects excessive and inefficient use of variable resources for a given level of fixed costs, and genebanks should not normally operate in this region.

4.3 Limitations of economic theory

We identify two particular difficulties that restrict our ability to apply a simple formula in order to optimize the composition of accessions: the quantification of cost per accession and the quantification of benefits.

4.3.1 Cost per accession

Published estimates of costs based on records kept by genebank managers (Burstin *et al.* 1997; Epperson *et al.* 1997; Pardey *et al.* 2001) are for total costs of running the genebank. From this information, the average cost per accession can be easily and accurately calculated.

On the other hand, marginal costs are difficult to estimate. They cannot be calculated from genebank managers' financial records even with a long time series of historical data. The genebank manager does not keep records of what the genebank would have cost if one more or one less accession had been kept during year—yet that is the information needed to calculate marginal costs.

Yet for the purpose of guiding management decisions, what we need to know is the marginal cost, not the average cost per accession. If we need to choose whether to change from action A to action B for an accession, the average costs of actions A and B are irrelevant. Instead, we need to know what would be the cost saving achieved by not doing action A, and the marginal extra cost of doing action B instead.

Economists typically apply theoretical principles to make assumptions that allow estimation of marginal cost. One of three assumptions can be made as follows.

(i) Over the relevant size range, marginal costs are constant.
(ii) Curators are operating at the most efficient point possible, where average costs have reached a minimum level. At this point, marginal cost equals average cost.
(iii) Fixed cost items (capital or quasi-fixed inputs) are utilized at a less-than-full capacity. In this case, marginal costs are always less than average cost.

For practical purposes, the third case is generally assumed and the average costs are interpreted as upper bounds of the corresponding marginal costs. However, clearly this is not entirely satisfactory.

4.3.1 Benefits

Much of this document is concerned with genetic impacts, on the assumption that improving conservation of genetic diversity has a value and a benefit. To choose an optimal conservation strategy by maximizing benefits relative to costs, we must be able to quantify benefits. How can this be achieved?

Conceptual advances in estimating benefits have been hindered by the fact that crop genetic resources generate values with multiple dimensions. Progress in empirical analysis has also been hampered by measurement difficulties, since only some dimensions of the value of crop genetic resources are revealed in market prices. The value derived from crop genetic resources is broadly categorized as *use value* and *non-use value*. Sometimes referred to as *existence values*, non-use values reflect the satisfaction individuals or societies may derive simply from knowing that something exists, independently of whether it is

used (Krutilla 1967). It is difficult to imagine, however, that many people, other than a few specialists, derive pleasure only from being assured that crop genetic resources are housed somewhere in a genebank. Instead, crop species are conserved precisely because they are thought to embody alleles of potential use to human society. Most value associated with the accessions in a genebank collection is derived from their use rather than their mere existence. Use value includes *current use value* and *expected future use value*, as well as the value of retaining the flexibility to respond to some unknown, future event—called *option value*. Overviews and surveys discussing the sources of economic value in crop genetic resources are numerous, including Pearce and Moran (1994) and Swanson (1996).

Both current and future use values can be estimated through market prices when a product or good, such as grain or seed, is traded. We can use forms of 'hedonic analysis' to ascertain the current value for productivity enhancement of crop genetic resources embodied in crop varieties (Evenson *et al.* 1998). A genebank collection, in contrast to a breeder's working collection, exists to a large extent in order to respond to future, unforeseen challenges, and therefore the expected future use value of a genebank collection is an important component of its total value. We can, with some methodological difficulty and a number of caveats, calculate a *present value of expected future benefits from direct use of germplasm in crop improvement.* We do so by combining the probability of finding useful material with its predicted productivity benefit once it is found and incorporated into new varieties. The time required to search for and incorporate useful genes into well-adapted germplasm affects the magnitude of expected benefits in a major way because of the time value of money.

Option value is similar to expected future use value conceptually, but distinct from it in practice. For example, we might use the past incidence of changes in rust disease pathogens or other major pest outbreaks to predict the expected future value of certain types of accessions as sources for new sources of resistance for a known pest. However, there are some pests and other environmental events for which we have no prior knowledge at all. Accessions, and collections of accessions, can have option value related to this uncertainty—but determining its magnitude is difficult.

Crop genetic resources are public goods and market prices generally fail to capture the full value of public goods. While recent changes in intellectual property rights may alter the public good nature of crop genetic resources, the problem of relying

on market prices to assign value to streams of direct use benefits from utilization of accessions in crop improvement is likely to persist. Finally, there are many current and future uses of genebank accessions other than their direct use in breeding new crop varieties—and many of these are contributions to other types of public goods, such as knowledge.

Alongside conceptual overviews of the sources of value, several theoretical economic models have analysed the value of genetic resources (For example, Brown and Goldstein 1984; Weitzman 1993; Polasky and Solow 1995; Simpson *et al*. 1996; Evenson and Lemarié 1998). By contrast, there are few published examples that use empirical data to estimate the value of genebank collections. Evenson and Gollin (1997) traced the flow of rice germplasm from the International Rice Research Institute into improved varieties grown in the developing world, and estimated that adding 1000 accessions to the collection was associated with annual income of US$325 million in present value terms. Gollin *et al*. (2000) studied several cases of the search for resistance among germplasm stored in a wheat collection at CIMMYT genebank, drawing inferences about the optimal size of collections and the conditions under which marginal accessions may or may not have high value. Zohrabian (2000) estimated the lower-bound value of an additional accession in the U.S. soybean collection, concluding that while the absolute value may not be great in absolute terms, it more than justified its cost.

Unfortunately, none of the above treatments of value addresses the question of how much the use value might be increased by improving the genetic efficiency of conservation. Part of the problem lies in the need to estimate the expected benefit B in terms that can be compared with the cost C. That is, we must be able to assign a value in monetary equivalents (a price in some currency) to all of the multiple dimensions of benefits mentioned above. In fact, most genebank managers face fixed budgets in the short-term. Their objective in such circumstances is to maximize expected net benefits given their budget constraint, or to be as 'cost-effective' as possible in their management. In that case, regardless of how they measure benefits, if expenditures are fixed, they can pursue their objective by choosing the strategy that maximizes expected benefits for the outlay. Stated differently, they can maximize the ratio of expected benefits to costs. It is important, however, that costs be calculated correctly, as outlined in the next section. Benefits could be measured by whatever criterion we judge appropriate, which in the case of genebanks means some measure of genetic

efficiency of conservation and utilization, rather than a figure expressed in monetary terms. Then management decisions should be based on comparing the expected benefits associated with a set of options, all of which have the same total cost.

However, even so, benefits of genebanks, and more especially the options available to us for measuring genetic efficiency, are multi-dimensional. For example, maximizing genetic efficiency embodies, *inter alia*, the following components: minimizing drift, minimizing selection, maximizing the number of distinct alleles conserved, maximizing between-accession genetic variance, minimizing the genetic distance between the current and original seed samples of each accession, and maximizing the ease of identifying and locating an allele in the collection. If we could quantitatively assess the relative impacts of each of these on current and expected future use values, we would be able to derive a single objective function to maximize. However, this is not currently possible. We do not know which component of genetic efficiency is most important, yet without this knowledge we cannot identify the management that will maximize expected benefits.

In practice this brings us back to the need to establish clear genebank objectives in the context of the genebank's institutional mandate and mission. We first define what we are seeking to do in terms of conserving and utilizing genetic diversity, in the form of a quantifiable criterion. Maximizing an agreed quantifiable criterion should be set as one of the genebank objectives. If we accept that criterion as our best predictor of current or expected future use value, and fix our budget so that we have only to maximize our selected criterion—then we can begin to solve the problem of choosing the optimal management strategy.

4.4 Determining consequences for lumping and splitting

With the above caveats, we can outline in principle the steps that are required assess the economic costs of either lumping or splitting accessions. We need to begin by identifying all the cost elements of running the genebank (Table 1). We then need to identify the basic operations like those presented by Pardey *et al.* (2001)—e.g. acquisition, medium-term storage, long-term storage, germination testing, dissemination, safety duplication, regeneration, information management, general management. Next, we need to consider the impact of lumping or splitting on each of these activities presented, in accordance with the relevant considerations described in Section 3.

Table 1. Examples of cost elements in genebank operation

Operations	Non-capital			Capital
	Quasi-fixed	Labour	Non-labour	
Information management (including data analysis)	• Information manager • Data analyst	• For data entry • For equipment maintenance	• Computer supplies • Publication related expenses • Software licenses	• Servers • Computer equipment
General management	• Genebank head or genebank manager	• Secretaries • Unallocatable labour	• Office expenses • Electricity • Unallocatable expenses	• Buildings • Unallocatable equipment
Storage (medium term and long term)	• Genebank curator	• For maintaining and operating refrigeration equipment and facility	• Electricity for storage rooms	• Cold storage room • Refrigeration equipment • Storage shelves and seed container
Viability testing	• Genebank curator	• Lab technician • worker	• Chemicals and supplies	• Lab equipment and supply
Acquisition	• Genebank curator • Scientist for seed health testing	• Lab technician • Temporary worker	• Chemicals and supplies • Seed envelope	• Lab equipment and facility
Safety duplication	• Genebank curator	• Temporary worker	• Packing supplies • Shipping cost	
Dissemination	• Genebank curator	• Lab technician • Temporary worker	• Chemicals and supplies • Packing supplies • Shipping cost	• Equipment and facility
Regeneration	• Genebank curator • Field manager	• Field worker • Equipment technician • Temporary worker	• Chemicals and supplies for fields • Fuel for vehicle • Electricity for drying machine	• Farming land • Screenhouse • Seed dryer • Seed cleaning equipment
Characterization	• Field manager • Lab scientist	• Field worker for agronomic characterization • Lab technician for molecular characterization	• Lab chemicals and supplies	• Lab equipment and facility
Evaluation	• Field manager • Lab scientist	• Lab technician • Field worker	• Lab chemicals and supplies	• Lab equipment and facility
Pre-breeding	• Field manager • Lab scientist	• Lab technician • Field worker	• Lab chemicals and supplies	• Lab equipment and facility
Other researchers	• Genebank curator • Lab scientist	• Lab technician	• Lab chemicals and supplies	• Lab equipment and facility

The nature of the economic assessment depends on the magnitude and conditions of re-structuring that is being considered:

- If the splitting and/or lumping is to be undertaken with no change in capital infrastructure, then we can simply ignore the capital costs.
- If it is to be undertaken with no change in the complement of senior scientists, then the total quasi-fixed costs must be held constant. This means that any increase in senior scientist inputs for one activity must be accompanied by an equivalent reduction for other activities.
- If the genebank is assigned a non-negotiable fixed total budget, then the same applies to the total for variable (labour plus non-labour) costs.

Then, for each cost component of each activity, two key estimates must be generated. First, we must determine the impact on the number of accessions to be processed each year. For storage this is simple: splitting an accession into two and discarding the original increases the number stored by one; and combining two accessions into one and discarding the originals reduces the number stored by one. For germination testing there will be, at least in the long-term, a *pro rata* increase/decrease associated with splitting/lumping. For regeneration there might be complex dependencies on usage. For dissemination and chatacterization the genebank manager may have greater flexibility in choosing which and how many accessions are disseminated and characterized each year, independently of the number stored.

Second, we must determine the impact on the efforts required to maintain the quality of the accessions processed each year - and hence estimate the effect on average costs per accession. This is probably one of the most difficult parts of the process. As outlined above, existing analyses of genebank costs are retrospective, first calculating total economic cost and then estimating average cost by dividing the total cost by the number of accessions. For planning to lump or split, managers need to estimate future average costs per accessions that will result from a change in management procedures. For example, after careful consideration of the elements of the handling procedure, we might estimate that the average costs per accession will rise by 10%.

Then, the manager's best estimate of the change in total costs would be given by multiplying the expected change in costs per accession multiplied by the change in total numbers of accessions processed annually. By doing this we shall have completed the achievable half of the economic analysis—impacts on economic

costs. The remainder—impacts on value—is beyond the scope of this document and beyond the achievements of any genebank analysis undertaken to date.

5 Accession management in specific situations

In this section we consider some specific situations where splitting or combining accessions may be appropriate. We classify accessions according to the breeding system of the species and the magnitude of genetic variation present within them. We also present a selection of examples for different collections where different genetic, operational and economic constraints have resulted in the application of contrasting management decisions.

Different factors are relevant for clonal species, true-breeding genotypes (inbreeders and apomicts) and outbreeders. For seed collections (true-breeding and outbreeders), the relative importance of different factors varies with the magnitude of intra-accession diversity. The scale of intra-accession diversity is illustrated in three categories: varieties and breeders' lines are the most uniform, landraces are generally more variable, and wild populations are usually the most variable. Each of these situations is exemplified here.

The difference between cultivated and wild populations lies not only in the amount but also in the distribution of intra-population diversity. It is usually distributed at random among the genotypes of a landrace but non-randomly in wild populations. Wild populations of many species of plant have a remarkably small 'genetic neighbourhood area' (Crawford 1984; Cahalan and Gliddon 1985; Beattie and Culver 1979; Kerster and Levin 1968; Levin and Kerster 1968; Richards and Ibrahim 1978; Schmitt 1980).[5] The distribution of pollen and seed dispersal is highly skewed: a small percentage of pollen and seed can often be dispersed very large distances, but the vast majority falls close to their parents. The precise distribution of dispersal distances depends on the dispersal mechanism (e.g. the longest documented dispersal distances are recorded for dispersal by water and birds). However, the general principle that dispersal is skewed, with most progeny falling close to the parent, applies

---

[5] The area within which plants can be regarded as crossing at random. For obligate outcrossers, this equals the variance of pollen dispersal distance plus the variance of seed dispersal distance. The genetic neighbourhood size is the number of plants of the species present within that area. The genetic neighbourhood area defines the minimum scale for population subdivision and should not be confused with larger scale effects of gene flow. For example, although the genetic neighbourhood area of *Lolium perenne* is estimated to be 8 m^2, effects of gene flow on population structure can be detected at distances of over 100 km (Monestiez *et al*. 1994).

to most species regardless of dispersal mechanism. Thus, a small genetic neighbourhood area is a feature of many plant species. Some estimates are 2.5 m^2 for the self-incompatible insect-pollinated herbaceous legume *Trifolium repens* L., and 8 m^2 for the self-incompatible wind-pollinated grass *Lolium perenne* L. (Hayward and Sackville Hamilton 1997). This can be much smaller than what appears as one physically continuous population. Where the area covered by such a physical population exceeds the genetic neighbourhood area, the population will be subdivided into a large number of overlapping but genetically distinct sub-populations, even if the site is homogeneous. The implications of distribution of intra-population diversity will also be considered below.

5.1 Clones

In the case of plants conserved as clones, each clone will normally be managed as a separate accession, and the concepts of splitting and combining are then meaningless. The size of the collection may, however, be reduced by eliminating accessions that are found to be the same clone as another accession in the collection (i.e. perfect duplicates). In such species, if there is a need to reduce costs and duplicates cannot be found, alternative efficiency measures must be considered, e.g. core collections. When eliminating duplicate clones, just as when combining or splitting accessions, it is important not to eliminate data on the original accessions. For example, passport data on eliminated clones remains valuable for studies of the ecogeographic distribution of genetic diversity.

For species that are normally propagated and used as clones, this will apply to all cultivated lines. Opportunities for combining are limited to accessions collected and stored as seed, which may include wild relatives, landraces, and pre-commercial breeders' lines. Case Study 1 shows an example where the high cost of conserving clones has made it economically effective to identify and eliminate duplicates.

Clonal species that may be collected as clones but managed as seed (for example temperate forages) would normally be conserved as seed rather than as clones. Therefore, for genebank management purposes, such species would be classified according to their sexual reproduction, as described below. However, the clonal population structure also has implications for collecting methods (Sackville Hamilton and Chorlton 1995) including the choice to split or combine at the point of collecting, collecting each sub-population as a separate accession, or combining all sub-populations to form one accession).

Case Study 1. Potatoes at CIP
(Centro Internacional de la Papa, Lima, Peru)

Eliminating duplicate clones and splitting seed accessions in the Potato Collection

Since conservation of potato as clones in a field genebank is expensive, we decided it would be cost effective to undertake a comprehensive assessment of duplication, and have achieved major savings by reducing the size of the collection from 15 000 to 3500. Accessions stored as seed are genetically variable so we decided we should assess whether they should be split, but found it was unnecessary.

Cultivars are maintained as clones in a field genebank, re-grown every year. The field genebank is backed up by storage of tubers in cold stores, by *in vitro* culture of diverse accessions at two locations, and by conventional seed conservation of the non-sterile accessions. Botanical seed are stored dried in medium- and long-term storage in accordance with international standards for seed conservation. This multiple system is necessary for secure conservation, but is expensive.

We therefore decided to look for and eliminate duplicate clones. The search involved a sequential process of initial morphological characterization for preliminary identification of potential duplicates, followed by full morphological characterization of those potential duplicates, followed by electrophoretic analysis of tuber proteins and esterase isozymes in morphologically identical accessions (Huamán 1994, 1998). By this means we have identified 3500 genetically distinct cultivars in the original untyped collection of 15 000 clonal accessions, and have eliminated the duplicates.

For accessions stored as seed, we had the opposite concern that maintenance costs are not high, but high genetic variation within accessions means we risk losing diversity by drift. RAPD markers have been used to test for genetic drift in the seed collection *ex situ*, and repeat collections have been made to compare with drift *in situ*. Large changes have been detected *in situ* (Rio *et al*. 1997a), but no significant drift has been detected during regeneration ex situ (Rio *et al*. 1997b). It is concluded that there is no need to split accessions to reduce drift *ex situ*.

Head of the Genetic Resources Unit, CIP, Lima, Peru
For more information see http://www.cipotato.org/projects/germplasm.htm

5.2 True-breeding lines

These include apomicts, obligate inbreeders and artificially inbred lines. As described below, inbreeders may present good opportunities for splitting. There may also be opportunities for combining, as illustrated in Case Study 2 for accessions with little associated passport data.

5.2.1 Varieties and breeders' lines

Commercial varieties and late-generation breeders' lines will normally be genetically completely or almost completely uniform, comprising a single genotype. Splitting in this situation is not an option.

For synthetic varieties and simple mixtures, all characters will show full correlation by descent, so that splitting will be fully effective in minimizing genetic variation within mixed accessions. The normal recommended procedure would be to split mixtures both in the base collection for conservation and in the active collection for utilization. The mixture can and should be reconstituted for specific evaluation of the mixture. The

Case Study 2. Flax at CGN
(Centre for Genetic Resources Netherlands, Plant Research International, Wageningen, Netherlands)

Making the most of data-poor accessions

The flax collection at the Centre for Genetic Resources, Netherlands (CGN) contains many accessions with almost no passport data. We decided it was wasteful to fill a collection with such poorly documented and possibly similar accessions, so we chose to investigate the feasibility of identifying and combining genetically similar ones. We were able to reduce the number of accessions in the group studied by over 50% with minimal loss of variation between accessions.

For about 30% of the flax accessions at the CGN, the only passport data available was a coded accession name consisting of a few letters and a number, such as 'M 25-341' or '324-Rm' for example. These names allowed grouping of the material in series, such as the M 25 or the Rm series. To investigate the genetic relationships of the accessions within and between the different series, an AFLP study was carried out on 29 accessions belonging to three such series. Subsequently, an analysis of molecular variance (Excoffier *et al.* 1992) was used to compare the genetic variation observed within and among accessions. Substantial differences in intra-accession variation were observed and accessions that were not significantly different were bulked into groups. As a result, more or less homogeneous accessions remained separate entries while more heterogeneous accessions could often be lumped. Combining to reduce 29 accessions to 14, reduced the among-population component of variance by only 2.6% while at the same time maintaining similar levels of variation among accessions (Treuren *et al.* 2001).

Loek van Soest
Curator CGN Flax collection
For more information see http://www.plant.wageningen-ur.nl/cgn/

genebank may also choose to retain a sample of the original mixture, if doing so presents benefits that justify the additional cost.

Combining identical accessions in this category is also likely to be cost-effective, because of a relatively low cost of identifying identical duplicates (i.e. accessions that have identical alleles at all loci) and relatively high benefit of combining them. The high benefit results from reduced cost and improved efficiency of conservation and utilization, with zero loss of genetic integrity. The low cost arises because two accessions with identical registered names can be considered to have a high probability of being biologically identical, so that confirmation that they are indeed identical would require evaluation of only a few loci. Therefore, combining identical duplicates is generally recommended for such accessions. However, care needs to be taken in identifying duplicates, because of possible identification errors and possible contamination with alien pollen, genes or genotypes. If two accessions with the same registered name were genetically different, it would be necessary to determine which

is the true variety—but that is a question beyond the scope of this document.

Much of the relative simplicity of research and breeding varieties of such species lies in being able to produce and study genetically uniform lines. Combining similar but not identical accessions may remove some of this benefit. The curator will therefore need carefully to evaluate the advantages and disadvantages of combining such accessions before deciding whether to do so.

5.2.2 Landraces and other mixtures

Landraces of true-breeding species can be 'simple' mechanical mixtures from a genetic perspective. If they are genuinely 100% true-breeding, they can be treated as synthetic varieties as above, i.e. split for conservation in the base collection, and split and/or maintained as the original for utilization in the active collection.

If, as is often the case, they are not entirely true-breeding, then more care will need to be taken in how they are maintained. Conservation should be based on the normal practice for maintaining the landrace *in situ* under the original cultural practices in the original environment. We note an important corollary: efficient *ex situ* conservation of a landrace requires knowledge of traditional cultural practices and environments of that landrace—knowledge that is often missing. Collectors should be trained to document this information when collecting, and any collecting forms that do not already have relevant data fields should be revised. Moreover, *ex situ* it is often difficult to mimic the original *in situ* environment and traditional cultural practices. Therefore, consideration should be given to true *in situ* conservation, and indeed to complementary *in situ* and *ex situ* conservation.

In some cases, such as Sorghum in Yemen (Sackville Hamilton and Al Khawlani 1981) and *Phaseolus vulgaris* L. in some African countries, a landrace can be deliberately maintained as a mixture by the farmer reconstructing the mixture each year. At each harvest the farmer selects seed of each component, and re-mixes them to form the seed mixture to sow the crop for the following year. A curator conserving such a landrace *ex situ* should likewise split it into its components for conservation.

In other cases, the landrace is maintained by farmers as a population, being harvested and sown as a self-maintaining mixture. In such cases, splitting the landrace into its components may be irreversible as it may not be possible to reconstitute the original landrace by re-mixing the components. If possible and if consistent with genebank objectives, splitting such landraces should

generally be avoided. On the other hand, it is often not possible for a curator to mimic the environment and traditional cultural practices used for the landrace *in situ*, in which case, since the landrace is genetically heterogeneous, it is expected to lose genetic integrity rapidly, which is also undesirable.

The curator then faces a difficult choice—which is worse, progressively losing genetic integrity by conserving the landrace intact but in the wrong conditions, or losing genetic integrity in one step by splitting with minimal subsequent further degradation? One option could be to do both: maintain the original intact in long-term storage and use it only for critical landrace restoration or assessment; and split a sample into components that can be used, regenerated and re-mixed frequently with minimal progressive loss of genetic integrity.

If a curator decides to split a landrace, but considers it not justifiable to maintain the original as well, the curator should first at least examine the composition of the original sample.

Case Study 3 illustrates a situation where accessions are routinely split during regeneration if they are found to be a mixture of easily distinguishable genotypes.

As illustrated in Case Study 4, for some purposes it can also be acceptable to separate and use only one genotype from each mixed accession while retaining the original.

As illustrated in Case Study 5, for some purposes it can also be acceptable to increase uniformity within accessions by reducing population size, without trying to split and retain the original within-accession diversity.

5.2.3 Wild populations

In the extreme case of zero outcrossing, pollen dispersal contributes nothing to dispersal of genes, and the variance of seed dispersal distances is sufficient as an estimate of the area of random maternity. The extent of genetic subdivision of wild populations of inbreeders is therefore generally greater than for outbreeders. Being true-breeding means that the characteristics of progeny derived from each mother plant will show perfect correlation by descent, or at least high correlation if they are not quite 100% inbreeding or apomictic.

This situation largely fulfils the criteria for splitting on the basis of the identity of the mother plant. Collecting and conserving the seed of each sampled mother plant as a separate accession will be highly effective in reducing genetic variance within accessions, compared with bulking all seed collected to create and conserve one accession to represent the entire physical population. However, it is not ideal for splitting, since the

Case Study 3. Landrace populations of self-pollinating crops

(Institut für Pflanzengenetik und Kulturpflanzenforschung, Gatersleben, Germany)

Splitting morphologically variable accessions to prevent loss of rare alleles

Landraces and other locally adapted germplasm can show a high degree of morphological diversity in a collected sample. To counteract the possible loss of rare alleles while regenerating accessions of such landraces of self-pollinating crops, the genebank of IPK for more than fifty years has been practising the splitting of variable accessions into morphologically distinct lines.

The main aim of splitting is to retain as many as possible of the rare components of a population, thus maintaining as much of the within-accession diversity as possible. Another reason is related to genebank management practices: it is easier to detect results of unconscious selection, admixtures or mistaken identity of accessions if they are homogeneous. Finally, the morphological infraspecific classification of cultivated plant species, as practised at IPK's genebank since its foundation in 1943, is only possible if accessions are homogeneous and can be assigned, using botanical keys, to certain taxa which, in turn, are denominations for combinations of stable morphological characters.

The first splitting usually takes place during collecting, or when processing the collected seeds in the genebank immediately after collecting. Mixtures of different crops, or crop-weed mixtures, are first divided into components comprising single species. Variable landraces may also be dissected using characters that can be seen on the seeds or plants (e.g. two-rowed vs. six-rowed barley, seed colour in garden beans). The collector appends lower-case letters to the collection number to designate the selections. For example, from a bean sample with the collection number 3456 within a particular expedition, the sub-samples 3456a, 3456b, 3456c, etc. may be derived.

Upon arrival in the genebank, a preliminary accession number is assigned to each (sub)-sample. At the first multiplication in Gatersleben (for which Lehmann and Mansfeld 1957, recommend a larger plot size than the standard size of 2.5 m^2 for cereals), the within-accession variation is carefully observed, and morphologically distinct lines are harvested separately. From each of these lines, a herbarium sample, or in case of cereals, a spike bundle, is taken to document the morphological characteristics of the accession, together with a seed sample, for later comparison purposes. The original population is also maintained as a separate accession and designated as such. This process is continued at up to three multiplications until a stable situation is reached (i.e. no further segregation occurs), or until it becomes obvious that the offspring of the 'pure' lines will be segregating again, in which case the accession will be maintained as a variable population only. Segregations occurring at later generations are eliminated, because the probability that they are the result of admixtures, spontaneous mutations or cross-pollination, increases with each regeneration cycle. The offspring of the single lines is being compared with the herbarium or spike samples to ensure stability of character expression. Upper-case letters are appended to the preliminary, or later the final, accession number to designate derived accessions. For example, a legume accession with the preliminary Gatersleben accession number V 147 may have been split up into V 147 A, V 147 B, V 147 C, etc.

For example, several oat samples collected in Czechoslovakia have been split up into 20 and more lines, the maximum being 34 for collection number 574 from Hvozdnica, western Slovakia of the IPK expedition in 1981; the derived lines were classified into 4 different botanical varieties. Collected samples from Iran, Ethiopia, Libya and the Balkan area have been subdivided into 20 and more lines. H. Kuckuck (expeditions in Iran in 1952-1954) seems to have assigned one collection number per site. Collection number 9 from Behbahan gave rise to a total of 44 genebank accessions of *Hordeum vulgare* (2 lines, 1 botanical variety), *Linum usitatissimum* (19 lines, 2 varieties), *Triticum aestivum* (13 lines, 7 varieties), and *T. durum* (10 lines, 7 varieties).

This practice was first described in 1957 (Lehmann and Mansfeld, 1957), and with minor adjustments, it is still practised today.

This approach leads to an increase of the number of accessions that to be maintained in the genebank. Among ca. 90 000 accessions maintained at IPK's genebank headquarter in Gatersleben, ca. 38 000 originated from collecting expeditions. Approximately 14 500 of these accessions resulted from splitting-up collected samples of landraces.

Helmut Knüpffer
Head IPK Genebank Documentation
For more information see Lehmann and Mansfeld (1957)

Case Study 4. The International Barley Core Collection

Even landraces are forced to be homogeneous in the Barley Core Collection

The International Barley Core Collection provides standard material for barley research. Therefore, it was decided to include as far as possible homogeneous homozygous lines, to ensure that different scientists use the same genotype when they use the same accession from the Barley Core Collection.

The idea behind the International Barley Core Collection (BCC) was to create a stable set of between 2000 and 3000 barley genotypes that represent the genetic diversity in barley, including landraces and wild species. We wanted to guarantee full comparability of results obtained by different research groups around the globe and across years by ensuring that all studies are based on an identical set of genotypes. For example, we wanted the AFLP fingerprints generated at one place to be produced by the same genotypes that expressed a level of drought tolerance determined somewhere else. This combination of results creates a tremendous added value.

To achieve this we chose to select only one homozygous line by SSD (single seed descent) to represent each genebank accession to be included in the BCC, even if it was initially heterogeneous, such as in the case of landraces and the wild progenitor of barley (*Hordeum vulgare* ssp. *spontaneum*). From each heterogeneous genebank accession we selected only the single most abundant genotype and split it from the remainder of the genebank accession to form the BCC accession. We had three important reasons for this:

1. we avoid the risk of one scientist choosing one genotype and another scientist choosing another from the same accession, so that results can be truly comparable across laboratories;
2. we avoid changes in genetic composition of accessions during regeneration, so that results can be truly comparable across generations,
3. we allow the documentation to be simple, one score per trait per accession per trial.

Owing to the difficulty of multiplication, accessions of other wild inbreeding species in the BCC have been established from few original seeds (not SSD) and may thus be heterogeneous. The two outbreeding, self-incompatible *Hordeum* species are maintained as populations.

Because of the benefits of homogeneity, we accepted the disadvantage that it would not be possible to use the BCC to study diversity within landraces and wild barley populations. The link with the original accession is documented, so that for such studies it is always possible to go back from the BCC accession to the original heterogeneous accession maintained somewhere in a genebank.

Roland von Bothmer
Chairman Barley Core Collection committee
For more information see Knüpffer and van Hintum (1995)

population would have to be split into a large number of accessions—one for each mother plant sampled. Therefore, doing so risks making the genebank collection very much larger and more expensive—for example, if seed are collected from 100 plants per field, then the choice to split or not is a choice to make 100 or 1 accession from one field. Therefore, the curator needs to be sure that the advantages of splitting are large enough to justify its high costs.

The decision to split or combine on the basis of the identity of the mother plant must be undertaken at the time of collecting. Once seed from different plants are mixed, it is not possible to separate them again later, and therefore a decision to bulk seed and to conserve them as one accession is irreversible. It should also be noted that many population genetics studies require seed to be harvested separately from each mother plant. For example, estimates of inbreeding coefficients in the field, the magnitude of population subdivision, the extent of gene flow and the strength of selection pressures are impossible from a

Case Study 5. Lettuce at CGN
(Centre for Genetic Resources Netherlands, Plant Research International, Wageningen, Netherlands)

Accepting drift to increase number of accessions

Since regenerating lettuce is rather difficult and expensive, the Centre for Genetic Resources, Netherlands (CGN), has chosen to concentrate on the diversity between accessions rather than the diversity within accessions.

Regenerating lettuce is very labour intensive. Seeds are sown, young plants are potted in a greenhouse, flowering plants staked to prevent lodging, and seeds are harvested manually. The plants have to be protected from a number of pests and diseases. Furthermore, depending on the species, a treatment of seeds to break dormancy, vernalization of germinating seeds or a gibberellic acid treatment of young plants is required to induce flowering. A few species are cross-pollinating and need isolation and insect pollination. The inflorescences of wild species need to be wrapped up in perforated polythene bags to prevent the seeds from floating all through the green house. Apart from labour costs there are also high costs involved in the taking and testing samples of each plant for *Lactuca* Mosaic Virus.

Since using the 'normal' number of plants (100 according to 'Genebank Standards': FAO 1994) for regeneration was so expensive, we decided that only 8 plants would be used for cultivars, which can be assumed to be homogeneous, and 16 plants for heterogeneous landraces and wild species. This implies that the chance of losing alleles as a result of drift is considerable (see table), but, on the other hand, it became feasible to manage a much larger number of accessions. This approach increases diversity between accessions, by accepting a reduction of within-accession variation, and thus allows the genebank to conserve more accessions without increasing the overall cost for conservation.

Chance of losing a selfing genotype in two regenerations with different initial frequencies and population sizes.

Number	Initial frequency of genotype			
of plants	0.05	0.10	0.20	0.50
8	0.77	0.58	0.32	0.03
16	0.59	0.34	0.10	0.00
50	0.20	0.04	0.00	0.00
100	0.04	0.00	0.00	0.00

Ietje Boukema
Curator CGN lettuce collection
For more information see www.cpro.dlo.nl/cgn

bulk sample. Therefore, a decision to bulk is not only irreversible but also rules out a large area of research.

Often the collecting site is clearly heterogeneous. For example,

- the species being collected may occur in more than one distinct type of vegetation on the same site. Spatial heterogeneity of vegetation is a feature of almost all plant ecosystems, whether ruderal or permanent, whether herbaceous, shrubby or forested. That is, an ecosystem does not comprise a uniform mixture of all its species; rather there are more or less discrete patches of different kinds dominated by different species. Each species in the ecosystem may occur in more than one type of patch.

- grasses may be found on and off a path;
- in an undulating site, plants may be found on the dry tops as well as in the wet hollows;
- in a forest site, understorey species may be found in the shade of a tree or in an unshaded gap in the forest.

In these cases, separate samples should normally be collected from each distinct microenvironment. The decision to split or combine must be taken at the time of collecting. A decision to combine at the point of collecting is irreversible, and makes the accession valueless for research on evolutionary adaptation to the micro-environmental heterogeneity.

A decision can be determined by the objectives of the genebank; Case Study 6 illustrates a situation where a decision to split by parent or not when collecting depends on whether there is a need to study population genetics.

Case Study 6. Lettuce at CGN
(Centre for Genetic Resources Netherlands, Plant Research International, Wageningen, Netherlands)

A wild inbreeding population: one or many accessions?

A scientist might be interested in a population, a breeder in a pure line. A genebank has, apart from these users demands, also to consider biological and financial constraints. The result is a pragmatic approach, as will be illustrated with the CGN lettuce collection.

The common sampling strategy for wild populations is to collect seeds randomly, from as many plants as possible at one site, trying to collect all alleles present in the population. Hawkes (1980) advises to collect 100 plants from highly variable populations and 50 plants from uniform populations. In practical situations this is not always achievable.

For genebank material it is often not known how many plants were collected. The most common wild relative of lettuce is *L. serriola*, a species with a high selfing rate. The material of this species in the CGN lettuce collection, was, as far as we know, collected from between 1 and 60 plants. This high variation in plant number was caused by differences in size of the collected population, or availability of seeds on the plants in the population at the time of collecting. Another important factor is the collector's strategy; some collectors collect single plants, and call each plant an accession.

Users of the lettuce collection sometimes complained because they expected homogeneous accessions, but obtained variable results when they screened accessions for resistances or other characters. This causes problems in discriminating between resistant plants and plants that were not properly inoculated. In such cases progeny testing is required before parent plants can be selected for further breeding.

The preference of users for homogeneous material from genebanks implies that lines of a population should be stored instead of the population itself. Splitting a population into lines can be done during collecting, by sampling single plants instead of populations, or later, for example during the first regeneration. CGN stores each complete wild lettuce population as an accession. Only if very different morphotypes appear during the first regeneration, these will become separate accessions. Since, due to limited regeneration capacity, only 16 plants per population are regenerated, the conserved lettuce populations will suffer from drift narrowing down the diversity within the population.

Collection of populations from as many different origins as possible should be attempted. Especially in the case of wild species, the diversity between accessions can be considered more important than the diversity within accessions.

Ietje Boukema
Curator CGN lettuce collection
for more information: Hawkes (1980)

5.3 Outbreeders

All kinds of populations of outbreeding species generally show relatively high levels of genetic variation within populations, even when selected for uniformity as is the case with varieties and breeders' lines. Many agronomically important characteristics show continuous distributions with low correlations between characters. It is therefore usually not appropriate to split, but in some situations, it may be appropriate to combine.

5.3.1 Varieties and breeders' lines

It can be appropriate to combine duplicate varieties, for similar reasons outlined for the inbreeding species described above. However, because they are genetically variable, the costs of identifying duplicates can be higher and the benefits can be less clear. Two accessions that share the same registered name are likely to be similar but not identical. Combining them will there-

Case Study 7. Managing temperate forages at IGER
(Institute of Grassland and Environmental Research, Aberystwyth, UK)

Temporary archiving: an efficient alternative to combining

The Genetic Resources Unit (GRU) at IGER has determined that it would not be economically viable to attempt to rationalize by combining duplicate accessions. We chose instead to reduce costs by identifying a subset of the collection most relevant to current breeding and research objectives, keeping only that subset active, and temporarily 'archiving' the rest of the collection.

In our situation, combining accessions would bring little economic benefit because conservation costs*per se* are low (full cost ~€0.2 year^{-1} accession^{-1}), would be detrimental to utilization because it is difficult to evaluate variable accessions and identify rare alleles, and would be detrimental to conservation because of the resulting increase in within-accession diversity. The costs of identifying duplicate accessions to be combined are high and the probable frequency of duplicates low, because accessions are highly variable and historically duplicate accessions are likely not to be biologically duplicate. Even historically duplicate copies of commercial cultivars with identical registered names are known to be biologically distinct. As a consequence, identification of duplicates would require an exceptionally detailed molecular characterization of large numbers of plants from each accession. Therefore attempting to reduce the size of the collection by combining accessions would be economically and genetically detrimental, and we choose not to do so.

Utilization costs are orders of magnitude higher (full cost to the GRU ~€80 year^{-1} donation^{-1} for donating accessions to external users, and ~€1000 year^{-1} accession^{-1} for utilization by the GRU for routine characterization and evaluation). By definition, most of any good PGR collection has low value for current utilization, and is being conserved for its potential value for future breeding and research objectives. Our users do not want immediate access to most of the collection. Therefore, maintaining the entire collection active for current utilization is unnecessary and, given the high cost of utilization, economically wasteful. Therefore, given the low cost of conservation, we choose to conserve the entire collection intact but to keep only a targeted subset in the active collection available for current use, temporarily 'archiving' the remainder for future use. Success of this approach depends on close collaboration with users, on being highly responsive to changing users needs, on being able to define a subset optimized for current use, and on being able easily to bring an archived accession back into use when required.

N R Sackville Hamilton
IGER GRU
For more information see http://www.igergru.bbsrc.ac.uk/

fore generally increase diversity within the accession, and the curator needs to determine whether this is acceptable. For example, the genebank at IGER contains 15 accessions of the *Lolium perenne* cultivar 'S23'. This is an old variety that remained popular for many years and is genetically more variable than the modern cultivars. There is evidence that the genotypic composition of the commercial product has changed significantly over the years, and assessing the extent and cause of the change is a potential research topic. Combining the accessions into one would prevent any such research, and therefore a decision has been taken to retain them as separate accessions instead of combining them (see Case Study 7).

5.3.2 Landraces

Landraces of outbreeding species are genetically more variable. They contain much continuous, multi-dimensional and hidden

Case study 8. Cabbage at CGN
(Centre for Genetic Resources Netherlands, Plant Research International, Wageningen, Netherlands)

Reducing the number of cabbage and brussels sprouts accessions

The *Brassica* collection at the Centre for Genetic Resources, Netherlands (CGN) contains many genetically similar selections of the same landraces. Since regenerating Brassicas is rather difficult and expensive, we have chosen to combine them, and by doing so have reduced the number of accessions in the collection by 80%.

The Netherlands has a long history of selection and breeding of *Brassica oleracea*. Breeders and farmers have made their own selections of landraces and old cultivars. The Dutch material in the CGN *B. oleracea* collection consists to a large extent of such selections, many of which are derived from a limited number of parental landraces, so-called 'umbrella varieties'. Since cabbage is an insect pollinated biennial crop, and regeneration is difficult and expensive, it was decided to limit the number of accessions as far as possible (Boukema and Hintum 1994). Sometimes up to 16 selections were combined to reconstruct one umbrella variety.

Material derived from the same umbrella variety was planted side by side. Assisted by *B. oleracea* experts involved in commercial plant breeding and variety registration, groups of very similar selections were composed. Other selections were kept as individual accessions. In some cases a number of groups from a single umbrella variety were created on the basis of maturity or other distinctive traits. As a result the number of accessions in the CGN collection of Dutch *B. oleracea* was reduced from 273 to 54, a reduction of 80%.

Subsequently, the process of lumping accessions was validated by an isoenzyme study, using a number of cabbage and Brussels sprouts groups. We tested the hypothesis that isoenzyme markers would correctly place a single accession in one of the groups. It appeared that most of the accessions were correctly classified. All misclassifications were within similar groups. In two cases the isoenzyme patterns suggested that the groups could have been even larger. In one of these cases this was a real option since it involved two groups made from the same umbrella variety. In the other case it involved groups with a common genetic background but a distinct identity as defined by morphology and history (Hintum *et al.* 1996).

Ietje Boukema
Curator CGN Brassica collection
For more information see http://www.plant.wageningen-ur.nl/cgn/

variation. Splitting is therefore not generally effective. Moreover, reconstituting a landrace after splitting can be difficult. If we can assume that gene frequencies in the original landrace were in Hardy-Weinberg equilibrium, then reconstitution requires a cycle of seed multiplication after re-mixing the components in the correct proportions in order to regenerate the original heterozygote frequencies. The additional economic cost and time delay involved in reconstituting a landrace from its components must be considered as part of a decision on whether to split.

Case Study 9. Collecting temperate forages at IGER
(Institute of Grassland and Environmental Research, Aberystwyth, UK)

Accepting drift to reduce shift and contamination and improve utilization

Since variable accessions and rare alleles are difficult to utilize, and since the potential loss of diversity by shift and contamination is high, the Genetic Resources Unit at IGER (GRU), has chosen to concentrate on controlling shift and contamination to maintain diversity between accessions. We decided that an increase in drift and loss of diversity within accessions is acceptable for conservation and beneficial for utilization.

Much of the variation within accessions is continuous and multi-dimensional with low correlations between characters. Therefore the conditions for splitting are not met in genebank accessions. On the other hand, splitting at the time of collecting a wild population can be appropriate. Wild populations are genetically highly variable for many characters of evolutionary and agronomic value. Some of the genetic variation within populations is associated with adaptation to easily identifiable micro-environmental heterogeneity within the site, for example a path through a pasture. We form a separate accession for plants collected from each obviously distinguishable micro-environment, to improve conservation, utilization and evolutionary research. For example, collecting separate accessions from nearby plants on and off a path through one site provides an efficient basis for conserving, identifying and studying genes for trampling tolerance.

Even in an apparently uniform site, we find much genetic variation within populations, and genetically distinct subpopulations. We also find a strong genetic shift associated with collecting seed samples of perennial forage species (Hayward and Sackville Hamilton 1997). To avoid this shift we collect vegetative samples (usually 30 adult plants per population) and produce seed for storage in isolation chambers from the original vegetative plants (Sackville Hamilton and Chorlton 1995). One consequence of collecting 30 plants per population is that our accessions have a small effective population size, we reduce genetic variation within accessions, and we do not effectively conserve rare alleles. Although this may be regarded as a disadvantage in terms of PGR conservation, it has benefits for utilization since the reduction in diversity within accessions facilitates evaluation and is essential for breeding. It replaces a cycle of pre-breeding that would otherwise be needed to reduce diversity within accessions.

The small population size also enables regeneration and storage procedures to be modified in a number of ways to reduce genetic shifts and eliminate genetic contamination, and thus conserve diversity between accessions more effectively. The set of modifications represents an integrated package, but one modification is of particular relevance to this publication. One result of the set of modifications is that it costs us almost nothing extra to store the seed of each original mother plant in a separate container while managing the population as one accession. Apart from improving conservation by reducing genetic shift during subsequent regeneration cycles, this also maintains some of the original population structure needed for population genetics studies, and maintains high correlation by descent between characters of seed from each plant. We can therefore at any time choose to split an accession by mother plant, and do so whenever research requirements indicate that it would be beneficial.

N R Sackville Hamilton
IGER GRU
For more information see http://www.igergru.bbsrc.ac.uk/

Combining similar landraces may be an option for some purposes. Case Study 8 shows an example where similar landraces have been combined to create umbrella landraces.

5.3.3 Wild populations

The principles are broadly similar to those for wild populations of inbreeding species, except that there is more genetic variation within populations and thus more heterozygosity. Wild populations can be divided into highly distinct sub-populations (Hayward and Sackville Hamilton 1997). For example, a series of studies on *Trifolium repens* (reviewed by Sackville Hamilton 1989) has shown differentiation with respect to many micro-environmental variables, even in a superficially homogeneous site. In perennial grasses, intra-population variation in reproductive output can be very high (Sackville Hamilton 1999), implying the possibility of large and rapid genetic changes during regeneration. Rapid genetic shifts in response to selection in newly sown pastures are well documented (e.g. Brougham and Harris 1967; Charles 1964; Falkner and Casler 2000). These indicate that high priority should be attached to limiting genetic shifts during regeneration.

A decision to attach high priority to limiting genetic shifts in accessions has implications for many aspects of genebank management, including not only collecting methodology but also regeneration, storage and chatacterization, is illustrated in Case Study 9.

As with wild populations of inbreeding species, a decision to split or combine must be made at the time of collecting; a decision to combine during collecting is irreversible. Splitting into a separate sample for each identifiable micro-environment within a site is always recommended. Finer splitting is necessary for more detailed population genetics studies on the structure of genetic variation within populations. Such studies are an essential component of improving our understanding of the causes and implications of genetic diversity.

Table 2. Summary of case studies presented

Box	Genebank	Crop	Crop type	Problem	Solution
1	CIP	Potato	Out-breeding clonal	High cost of conservation	Eliminate duplicate clones
2	CGN	Flax	In-breeding annual	Badly documented accessions	Combine similar accessions
3	IPK	All in-breeding crops	In-breeding	High variation within accessions	Split while regenerating
4	IPK	Barley core collection	In-breeding annual	Results not comparable across trials	Split one genotype from each variable accession
5	CGN	Lettuce	In-breeding	High cost per plant regenerated	Accept drift to increase number of accessions
6	CGN	*Lactuca serriola*	In-breeding wild	High variation within wild populations	Split or not while collecting depending on intended use
7	IGER	temperate forages	Out-breeding perennial	High cost of identifying duplicates	Archive instead of combine
8	CGN	Brassicas	Out-breeding	High cost of regeneration	Combine selections of umbrella varieties
9	IGER	temperate forages	Out-breeding perennial	High loss of genetic integrity through shift	Accept drift to improve control of shift

6 Provisional recommendations

It is not appropriate to give definitive recommendations, in part because of inadequate knowledge. We know in principle that a decision to split or lump has genetic and economic implications for conservation and for utilization, and we know in principle that we should not choose an action if its economic and genetic costs are too high relative to its economic and genetic benefits. But we have failed to find even a single case study where decisions have been based entirely on a quantitative cost-benefit analysis. All the case studies above include at least an element of qualitative subjective assessment, even those that have included some quantitative analysis. Therefore this document is intended rather to promote further consideration of the issues and to highlight the knowledge gaps.

Moreover, the curator's decisions should depend critically on the objectives of conservation and utilization, so we can do little more than to identify the tools available to help the curator choose a strategy.

In this spirit, the following may be regarded as provisional recommendations pending further research and discussion.

1. Do not split or combine unless the reasons are clear. The reasons must arise from a consideration of policy, genetic and economic implications in relation to your genebank objectives.
2. If you choose to split or combine, always retain data on the original accession(s) in your database, and always document how you split or combined—you may need that information later.
3. If you choose to split or combine, if possible also keep a low-cost backup of seed of the original accession(s) (i.e. store and document, but do nothing more with the seed or data) in case of wrong decisions. Even if your decision is right for today's technologies and objectives for conservation and utilization, it may become wrong in the future as technologies develop and objectives change.
4. If you choose to split but cannot keep a backup of the original before splitting, first determine the composition of the original sample as much as possible—you may need that information later.
5. If you keep separate base and active collections, consider making separate decisions for them; for the base collection choose what is best to meet your conservation objectives, and for the active collection choose what is best to meet your utilization objectives. For example,

- Split for the base collection (e.g. if this improves long-term conservation of genetic integrity), and lump for the active collection (e.g. if this improves usage efficiency).
- Maintain a landrace intact in the base collection (to conserve a sample of the original), and split in the active collection (if this reduces cumulative loss of genetic integrity by frequent use and regeneration).

6. If your decisions to split or lump are user-driven, remember that different users have different needs. For example, breeders may prefer landraces to be split into genetically more uniform components that are easier to incorporate into a breeding programme, whereas a traditional community requesting repatriation of its landraces will require them to be repatriated complete with their full original diversity.
7. Split under the following conditions:
   - Split while collecting, regardless of breeding system, if it improves conservation and or utilization.
   - Split while collecting as much as possible (down to separate mother plants or even separate flower heads), regardless of breeding system, if intended usage includes research into population genetics structure.
   - Split mixtures at any time (e.g. during regeneration or evaluation) for inbreeding but not outbreeding species—but only if splitting is compatible with users' objectives and with the economics of conservation and utilization.
8. Lump under the following conditions:
   - Maintenance (regeneration, storage, germination testing) or evaluation costs per accession are high.
   - Costs of identifying duplicates are relatively low.
   - Usage is more efficient.
   - The value of the accessions to be combined lies more in the diversity of genes that they jointly possess than in the specific genotypic composition of each accession. For example, if nothing is known about the origin or characteristics of an accession, there is little to be gained by investing resources in maintaining it in its original state.
   - There is no requirement to assess genetic variation between the original accessions.
9. Consider alternatives to lumping and splitting, since these are only two of the tools available to improve management efficiency. For example:
   - Reduce costs by combining the base and active collections (D. Debouck pers. comm.)—but only if this does not exacerbate conflicts between conservation and utilization.

Standard recommended practice is to maintain them separately (FAO 1994) in order to resolve such conflicts.

– Eliminate rather than combine duplicates—but only for perfect clonal duplicates.
– Reduce costs by archiving instead of combining—but only if this is compatible with intended use.
– Reduce costs of regeneration by reducing the number of parent plants—but only if this is compatible with intended use and conservation objectives.

7 Conclusions

In this document we have demonstrated that curators should not consider an accession as a fixed entity. Conserving an accession 'as is' may be regarded as an appropriate default action in the absence of good reasons to do otherwise. Nevertheless, subject to careful analysis of genetic and economic consequences in relation to genebank objectives, curators should be prepared to consider splitting or combining accessions to optimize the genebank's efficiency. Decisions may be made to reduce running costs, increase the economic and/or genetic efficiency of conservation, increase the economic and/or genetic efficiency of identifying alleles for utilization, or increase the range of types of research that can be undertaken with an accession. We have discussed the issues that need to be considered in reaching a decision, and we have presented a number of specific situations where splitting or combining is appropriate. The specific examples presented are not comprehensive. We hope, however, that we have given a sufficient variety of examples to encourage other curators to think about improving their own situation in these or in other ways.

References

Beattie A.J. and D.C. Culver 1979. Neighbourhood size in *Viola*. Evolution 33:1226-1229.

Boukema I.W. and Hintum Th J.L. van. 1994 *Brassica oleracea*, a case of an integral approach to genetic resources conservation. Pp 123-129 *in* Evaluation and Exploitation of Gen-etic Resources, Pre-Breeding. Proceedings of the Genetic Resources Section Meeting of Eucarpia.

Breese E.L. 1989. Regeneration and multiplication of germplasm resources in seed genebanks: the scientific background. International Board for Plant Genetic Resources, Rome, Italy.

Brougham R.W. and W. Harris 1967. Rapidity and extent of changes in genotypic structure induced by grazing in a ryegrass population. New Zealand Journal of Agricultural Research 3:442-453.

Brown G.M. and J.H. Goldstein. 1984. A model for valuing endangered species. Journal of Environmental Economics and Management 11:303-309.

Burstin J., M. Lefort, M. Mitteau, A. Sontot and J. Guiard. 1997. Towards the assessment of the cost of genebanks management: conservation, regeneration and character-ization. Plant Varieties and Seeds 10:163-172.

Cahalan C.M. and C. Gliddon. 1985. Genetic neighbourhood sizes in *Primula vulgaris*. Heredity 54:65-70

CBD. 1992. The Convention on Biological Diversity: Convention Text. Published online at http://www.biodiv.org/convention/articles.asp.

CGRFA. 2001. Report of the FAO Commission on Genetic Resources for Food and Agriculture, sixth extra-ordinary session, Rome 25–30 June 2001. Food and Agriculture Organization of the United Nations, Rome, 2001. Also published online at ftp://ext-ftp.fao.org/waicent/pub/cgrfa8/ex6/e6repe.pdf.

Chapman M.J. and D.L. Mulcahy. 1997. Effect of genome–plastome interaction on meiosis and pollen development in *Oenothera* species and hybrids. Sexual Plant Reproduction 10:288-292.

Charles A.H. 1964. Differential survival of plant types in swards. Journal of the British Grassland Society 19:198-204.

Crawford T.J. 1984. What is a population? Pp. 135-173 i*n* Evolutionary Ecology (B. Shorrocks, ed.). Blackwell Scientific Publications, Oxford.

Epperson J.E., D. Pachico and C.L. Guevara. 1997. A cost analysis of maintaining cassava plant genetic resources. Crop Science 37:1641-1649.

Evenson R.E. and D. Gollin. 1997. Genetic resources, inter-national organizations, and improvement in rice varieties. Economic Development and Cultural Change 45(3):471-500.

Evenson R.E., D. Gollin and V. Santaniello. 1998. Agricultural Values of Plant Genetic Resources. CABI, FAO and University of Tor Vergata, Rome, Italy.

Evenson R.E. and S. Lemarié. 1998. Crop Breeding Models and Implications for Valuing Genetic Resources. *In* Farmers, Gene Banks and Crop Breeding: Economic Analyses of Diversity in Wheat, Maize and Rice (M. Smale, ed.). Kluwer Academic, Dordrecht and International Maize and Wheat Improvement Center (CIMMYT), Mexico.

Excoffier L., P.E. Smouse and J.M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479-491.

Falkner L.K. and M.D. Casler. 2000. Genetic shifts in smooth bromegrass under grazing: changes in nutritional value and preference for surviving vs original genotypes. Grass and Forage Science 55:351-360.

FAO. 1994. Genebank Standards. Food and Agriculture Organization of the United Nations, Rome and International Plant Genetic Resources Institute, Rome, Italy.

FAO. 1996. FAO State of the World's Plant Genetic Resources for Food and Agriculture. Food and Agriculture Organiz-ation of the United Nations, Rome, Italy.

Fisher R.A. 1930. The Genetical Theory of Natural Selection. Clarendon Press, Oxford, UK.

Gollin D., M. Smale and B. Skovmand. 2000. Optimal search for traits in *ex situ* collections of wheat genetic resources. American Journal of Agricultural Economics 82(4):812-827.

Guarino L., V. Ramanatha Rao and R. Reid. 1995. Collecting Plant Genetic Diversity—Technical Guidelines. CAB International, Wallingford, UK.

Hammer K., R. Fritsch, P. Hanelt, H. Knüpffer and K. Pistrick. 1995. Collecting by the Institute of Plant Genetics and Crop Plant Research (IPK) at Gatersleben. Pp. 713–725 *in* Collecting Plant Genetic Diversity—Technical Guidelines (L. Guarino, V. Ramantha Rao and R. Reid eds.). CAB International, Wallingford, UK.

Hawkes J.G. 1980. Crop genetic Resources Field Collection Manual. IBPGR/EUCARPIA, Rome, Italy.

Hayward M.D. and N.R. Sackville Hamilton. 1997. Genetic diversity: population structure and conservation. Pp. 49–76 *in* Biotechnology and Plant Genetic Resources (B.V. Ford-Lloyd, H.J. Newbury and J.A. Callow, eds.). CAB International, Wallingford, UK.

van Hintum ThJ.L. 1993. A computer compatible system for scoring heterogeneous populations. Genetic Resources and Crop Evolution 40:133-136.

van Hintum ThJ.L. and D.L. Visser. 1995. Duplication within and between germplasm collections. II Duplication in four European barley collections. Genetic Resources and Crop Evolution 42:135-145.

van Hintum ThJ.L., I.W. Boukema and D.L. Visser. 1996. Reduction of duplication in a *Brassica oleracea* germplasm collection. Genetic Resources and Crop Evolution 43:343-349.

van Hintum ThJ.L., N.R. Sackville Hamilton, J.M.M. Engels and R. van Treuren. 2001. Accession management strategies: splitting and lumping. International Conference on Science and Technology for Managing Plant Genetic Diversity in the 21st Century, Kuala Lumpur, Malaysia, 12–16 June 2000.

Huamán Z. 1994. Management of potato and sweetpotato field genebanks. CIP circular 20(3):1-7.

Huamán Z. 1998. Collection, maintenance and evaluation of potato genetic resources. Plant Varieties and Seeds 11:29-38.

Kerster H.W. and D.A. Levin. 1968. Neighbourhood size in *Lithospermum caroliniense.* Genetics 60:577-587.

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 E.A.V., eds.), Wiley, Chichester, UK.

Krutilla J.V. 1967. Conservation reconsidered. American Economic Review 57(3):777-786.

Lehmann C.O. and R. Mansfeld. 1957. Zur Technik der Sortimentserhaltung. Kulturpflanze 5:108-138.

Levin D.A. and H.W. Kerster. 1968. Local gene dispersal in *Phlox.* Evolution 22:130-139.

Lombard V. and R. Delourme. 2001. A consensus linkage map for rapeseed (*Brassica napus* L.): construction and integration of three individual maps from DH populations. Theoretical and Applied Genetics 103(4):491-507.

Mather K. 1973. Genetical Structure of Populations. Chapman and Hall, London, UK.

Mather K. and J.L. Jinks. 1971. Biometrical Genetics. Chapman and Hall, London, UK.

Monestiez P., M. Goulard and G. Charmet. 1994. Geostatistics for spatial genetic structures: study of wild populations of perennial ryegrass. Theoretical and Applied Genetics 88:33-41.

Ortiz J.P.A., S.C. Pessino, V. Bhat, M.D. Hayward and C.L. Quarin. 2001. A genetic linkage map of diploid *Paspalum notatum*. Crop Science 41:823-830.

Pardey P.G., B. Koo, B.D. Wright, M.E. van Dusen, B. Skovmand and S. Taba. 2001. Costing the *Ex Situ* Conservation of Genetic Resources: Maize and Wheat at CIMMYT. Crop Science 41:1286-1299.

Parzies H.K., W. Spoor and R.A. Ennos. 2000. Genetic diversity of barley landrace accessions (*Hordeum vulgare* ssp. *vulgare*) conserved for different lengths of time in *ex situ* genebanks. Heredity 84:476-486.

Pearce D. and D. Moran. 1994. The Economic Value of Biodiversity. Earthscan, London, UK.

Polasky S. and A. Solow. 1995. On the Value of a Collection of Species. Journal of Environmental Economics and Management 29(3):298-303.

Rana R.S., R.L. Sapra, R.C. Agrawaland Rajeev Gambhir. 1991. Plant genetic resources documentation and information management. National Bureau of Plant Genetic Resources, New Delhi, India.

Richards A.J. and H. Ibrahim. 1978. Estimation of neighbourhood size in two populations of *Primula veris.* Pp. 165–174 *in* The Pollination of Flowers by Insects. (A.J. Richards, ed.). Linnean Society Symposia Series, 6.

del Rio A.H. J.B. Bamberg and Z. Huamán. 1997a. Assessing changes in the genetic diversity of potato gene banks. 1: Effects of seed increase. Theoretical and Applied Genetics 95:191-198.

del Rio A.H., J.B. Bamberg, Z. Huamán, A. Salas and S.E. Vega. 1997b. Assessing changes in the genetic diversity of potato gene banks. 2: *In situ* vs *ex situ*. Theoretical and Applied Genetics 95:191-198.

Sackville Hamilton N.R. 1989. Life history studies. Pp. 1–16 *in* Advances in Legume Biology (C.H. Stirton and J.L. Zarucchi, eds), vol. 28.

Sackville Hamilton N.R. 1999. The rationalization of regeneration methods: how far can we go? The example of forage grasses. *In* Implementation of the Global Plan of Action in Europe—Conservation and Sustainable Utilization of Plant Genetic Resources for Food and Agriculture (T. Gass, L. Frese, G. Begeman and E. Lipman, eds.). Proceedings of the European Symposium, 30 June–3 July 1998, Braunschweig, Germany. International Plant Genetic Resources Institute, Rome, Italy.

Sackville Hamilton N.R. and K.H. Chorlton. 1995. Collecting vegetative material of forage grasses and legumes. Pp. 467-484 *in* Collecting Plant Genetic Diversity—Technical Guidelines (L. Guarino, V. Ramanatha Rao and R. Reid, eds.). CAB International, Wallingford, UK.

Sackville Hamilton N.R. and K.H. Chorlton. 1997. Regeneration of accessions in seed collections: a decision guide. Handbook for Genebanks 5. International Plant Genetic Resources Institute, Rome, Italy.

Sackville Hamilton N.R. and M.A. Al Khawlani. 1981. Collecting in the Yemen Arab Republic, part III. Plant Genetic Resources Newsletter 45:13-16.

Schmitt J. 1980. Pollinator foraging behaviour and gene dispersal in *Senecio* (Compositae). Evolution 34:934-943.

Simpson R.D., R.A. Sedjo and J.W. Reid. 1996. Valuing biodiversity for use in pharmaceutical research. Journal of Political Economy 104(1):163-185.

Swanson T. 1996. Global Values of Biological Diversity: the Public Interest in the Conservation of Plant Genetic Resources for Agriculture. Plant Genetic Resources Newsletter 105:1-7.

van Treuren, L.J.M van Soest and Th.J.L van Hintum. 2001. Marker-assisted rationalization of genetic resources collections: a case study in flax using AFLPs. Theoretical and Applied Genetics 103:144-152.

Weitzman M. 1993. What to preserve: an application of diversity theory to crane conservation. Quarterly Journal of Economics 108:57-184.

Zohrabian A. 2000. An economic model of utilization of U.S. crop germplasm collection: The case of soybean collection and soybean cyst nematode. PhD thesis, Auburn University, Alabama, USA.

IPGRI is a Future Harvest Centre supported by the Consultative Group on International Agricultural Research (CGIAR)

ISBN 92-9043-516-X