THEMATIC BACKGROUND STUDY

Indicators of Genetic Diversity, Genetic Erosion and Genetic Vulnerability for Plant Genetic Resources for Food and Agriculture

Anthony H D Brown
Honorary Research Fellow, CSIRO Plant Industry, Canberra, ACT 2601, Australia
and Honorary Research Fellow, Bioversity International, Rome, Italy

September 2008

Disclaimer

The content of this document is entirely the responsibility of the authors, and does not necessarily represent the views of the Food and Agriculture Organization of the United Nations (FAO), or its Members. The designations employed and the presentation of material do not imply the expression of any opinion whatsoever on the part of FAO concerning legal or development status of any country, territory, city or area or of its authorities or concerning the delimitation of its frontiers or boundaries. The mention of specific companies or products of manufacturers, whether or not these have been patented, does not imply that these have been endorsed by FAO in preference to others of a similar nature that are not mentioned.

CONTENTS

EXECUTIVE SUMMARY 5

1. INTRODUCTION 7

1.1 Indicators 7
Interpretation 7
Sampling 7
Aggregation 7
Comparability 8
1.2 Recent attempts or processes 8

2. GENETIC DIVERSITY 8

Use of indicators 8
Genetic diversity 9
Diversity richness and evenness 9
2.1 Diversity and management indicators by resource category 9
Resource category 9
Indicators by category, numbers and diversity 9
2.2 Wild relatives 10
Lack of species equivalence 10
2.3 Cultivated species germplasm collections 11
2.4 Variety diversity *in situ* 12

3. GENETIC EROSION 13

4. GENETIC VULNERABILITY 14

5. ROLE OF MOLECULAR TECHNIQUES 16

5.1 Genetic precision 17
5.2 Limitations 17
5.3 Roles 17

6. RECOMMENDATIONS 18

ACKNOWLEDGMENTS 19

REFERENCES 20

BOX 1
RECENTLY SUGGESTED INDICATORS 25

BOX 2
RICHNESS AND EVENNESS OF DIVERSITY 26

EXECUTIVE SUMMARY

Summary measures of genetic diversity in cultivated plants and their wild relatives are needed to guide managerial decisions, to monitor progress and to warn of emerging problems in agricultural production.

Such indicators are needed to measure the genetic diversity currently present in agricultural populations on farm and held in germplasm collections, and to detect genetic erosion, or serious loss of diversity in time, and warn of vulnerability due to adverse deployment of genetic diversity in space.

Several international bodies has already made considerable efforts to formulate indicators appropriate for the management of biodiversity at all levels. Such indicators must meet a number of criteria to be usefully implemented.

While diversity itself encompasses many concepts, richness of diversity – the number of different kinds of individuals regardless of their frequencies – is the most important theme, followed by evenness – how similar the frequencies of the different variants are.

Many variables are plausible as indicators of diversity. The more practical are based on number of individuals or area occupied *in situ* and on the number of accessions and the number of gene banks *ex situ*.

Genetic erosion is measurable as the proportion of richness of genetic diversity no longer existing in current populations when compared with the crop a decade previously or predicted to be lost in the next decade without remedial action.

Genetic vulnerability is inversely related to richness of diversity that is present locally, particularly if it is known to possess adaptation to exotic or new mutant pathotypes or insect strains or environments.

Census information forms the primary data but should be supplemented and validated using more-direct assays at the DNA level with molecular techniques.

In conclusion, several recommendations are made for projects to develop the proposals and ideas contained in the body of this report, test them on suitable existing sets of data and prepare protocols for measuring the variables.

1. INTRODUCTION

Crucial decisions that concern the genetic diversity of crop plants and their wild related species continue to be made at many levels – local, national and international. The issues behind such decisions range from the agricultural to the economic and the political. It has become increasingly evident that indicators – summary measures of diversity – are needed to guide these decisions, to monitor progress towards improving the genetic resources available to farmers and the conservation of these resources, and to give early warning of problems. Examples of important indicators in other fields of human endeavour are global CO^2 levels, global mean annual temperature, economic indicators such as the rate of inflation or the cost of living. The significance and the use in human affairs of such example indicators are widely known.

In the case of plant genetic resources for food and agriculture (PGRFA), three types of measure are needed. The first type addresses the current state of PGRFA, or the standing **genetic diversity**, including that existing *in situ* in fields or natural areas, and that stored away from its site of origin *ex situ* in orchards, seed banks or gene banks. The second type is aimed at measuring changes in the status quo over time, in particular as indicators of the rate of loss of diversity or **genetic erosion**. The third type has to do with deployment of diversity in space, in particular indicators of **genetic vulnerability**. Such vulnerability arises when genetic homogeneity, or the lack of diversity, renders the crop growing in a region liable to decimation if a new biotype of pest or pathogen were to invade it. Already substantial efforts have been made to provide indicators in some of these areas. This report is a discussion and distillation of these efforts and provides a list of suggested indicators to meet these needs more fully.

Complexity and simplicity: The development of indicators for PGRFA faces a fundamental dilemma. The world's agro-ecosystems are highly complex, and many and varied forces act to threaten gene pools. These forces include environmental factors and the decisions of individual farmers, scientist, breeders, communities, industries, conservation agencies and governments. Consequently an equally diverse set of parameters of varied complexity, scale and cost might be nominated as important variables to monitor and therefore as proxy indicators of diversity. However, such a suite will attract few users as indicators; there would be too many to follow, to interpret and to act upon. A reductionist approach is inescapable.

1.1 Indicators

Table 1 lists the desirable properties of indicators for managing genetics resources, drawn from similar published lists (Brown and Brubaker 2002). Of key importance are that indicators should be valid, comprehensible, aggregative and readily implemented. The tendency of a committee to nominate a list of genetic diversity indicators that includes every interest of that committee must be avoided.

Interpretation

Deciding which indicators to use is only the first step. The interpretation of the actual estimates presents further challenges. One procedure is to ascribe a meaning to a specific **'benchmark' value** by having absolute standards (e.g. a minimum number of varieties that should underpin crop production in a given area, or the minimum value of germination for gene-bank accessions). Alternatively the purpose may be to monitor **trends over time**, with desirable or acceptable rates of change specified. The action values require both scientific and stakeholder input, such that meaningful outcomes are assured. Even so, there is need for a process to confirm that the indicator actually measures the quantity intended.

Sampling

Because of the constraints of costs, virtually all indicators involve a sampling process to measure their current values. Sampling is a key step that determines the avoidance of bias and the validity of up-scaling. Stratified random sampling is a basic technique that allows the aggregation of values for heterogeneous strata, and of data from finer scales. In addition, stratified sampling has the advantage that the overall statistics can be disaggregated, to recover the values for contributing strata if targeted action is required.

Aggregation

Aggregation is a common process in obtaining and using numerical values for indicators. Aggregation is the combining of values for component regions, or time periods, or species. Hence the property is listed as highly desirable in Table 1. For example, Hamrick and Godt (1989) summed estimates of diversity over different species, categorized by breeding

system, to obtain overall estimates of diversity in plants. Averaging over unlike entities raises a general problem: should the contributing entities be counted equally, or weighted according to some factor? The weighting factor for each component might be some function of its relative size, frequency, quality, productive capacity, or importance, for example in weighting the components of a sustaining diet. Alternatively an appropriate weighting factor might be a relative measure of the economic value of the component. For example, the member crops of a suite of fruit species might be weighted according to their total market value. For studies of trends over time it will be important to retain the component diversity values, or unweighted composite value, particularly if weighting factors themselves change over time.

Comparability

A second pitfall in making comparisons of averages based on heterogeneous elements is the failure to base comparisons on common elements. An extreme example of this problem would be changes in proportions of traditional varieties when estimates for, say, horticultural crops are included in some but not all averages. Any changes in overall patterns could be due to differences in the composition of the averages.

1.2 Recent attempts or processes

Within the last decades specialist committees of several agencies have proposed approaches to the framing of indicators of genetic diversity, usually within a broad context such as the whole environment, agro-ecosystem or the biodiversity related to agriculture. Measures for genetic diversity from two recent examples are summarized in Box 1.

The Organization for Economic Co-operation and Development (OECD) has a process for developing environmental indicators for the performance of agriculture in member countries. The task is not only to nominate variables, but to provide protocols and to use and interpret existing data (OECD 2001). Since the emphasis is on the practical properties of indicators as discussed above, OECD indicators tend to be broad brush. For example, the number of registered varieties of the major crops provides a very limited indication of extant diversity. OECD documents discuss these limitations and point out that registered variety number may not reflect diversity at the gene level. In addition, for developing countries the number of registered varieties omits the diversity contained in traditional or unimproved varieties.

The Streamlining of European Biodiversity Indicators (SEBI) process aims to provide indicators as a European response to the challenge set by the United Nations Convention on Biological Diversity (CBD) to reduce or halt the loss of biodiversity by 2010. The table in Box 1 presents a selection of the 26 indicators devised to report against particular focal areas of the convention (EEA 2007). The box selection is of the indicators more relevant to agricultural diversity. Again, these indicators are quite broad and do not address issues such as the nature of intra-specific genetic diversity, its erosion or its deployment to render agricultural production less vulnerable to changes such as climate change. Further, some choices are questionable at best. Thus to address the need for a 'headline' indicator for "trends in genetic diversity of domesticated animals plants and fish species of major socio-economic importance", the choice was 'Livestock genetic diversity'. The reason given was the greater quality and availability of data for animals than for crops, trees and fish (EEA 2007). This cluster of indicators was reviewed in the EASAC report (2005) in terms of the existing data and desirable properties of indicators. Thus, the EEA indicator measures diversity in terms of animal breeds (local native breeds versus introduced breeds per country) rather than full genetic diversity. Such an approach is similar to that of Jarvis *et al*. (2008) for crop plant diversity as discussed below.

2. GENETIC DIVERSITY

At the outset, the task of devising a limited set of variables to measure the amount of genetic diversity seems to be straightforward. A manager or decision-maker simply wishes to be able to report, for example, that the genetic diversity in a set of plant species in a specific region has increased by say 20%, or that it has been held at a constant value under the current stewardship. In this way the indicator functions to monitor any change in genetic diversity, or to reflect managerial achievement.

Use of indicators

The questions to be answered and the purpose or use of the data must be clearly in mind in proposing indicators of genetic diversity. For PGRFA, the main purposes are comparative ones. For example one may wish to compare the

variability status of different crop species, e.g. is sugar beet, which is a recent domesticate, more genetically diverse than pea, an ancient crop? Other sorts of comparison include kinds of crops (fruit trees versus field crops), kinds of breeding system, changes in variability status with time, geographic and ecological comparisons, diversity *in situ* versus that in collections, kinds of characters assessed. Comparisons within species are also of critical importance. For example, where, if any, are the 'hot spots' of diversity?

Genetic diversity

Genetic diversity arises primarily as variants in the linear sequence of nucleotides in DNA. Mutations can happen in the coding region of genes, or the spacer regions within and between genes, in the number of copies of genes, the linkage relation between several genes or indeed in whole chromosomes. A small portion of these changes translates into protein variation, into marker polymorphisms, characters, physiological and morphological variation in agronomic characters and ultimately into varieties given different names by farmers. Some correlation exists between the variation expressed at these different levels, but, even so, the choice of which is the best level at which to assess diversity is unclear. It is clear that we cannot rely solely on any one level, and that it will be important to cross-check major trends in diversity over several levels.

Diversity richness and evenness

The appropriate statistical measures of diversity to use have long been a matter of discussion (Magurran 2003). A contention here is that indicators of diversity should account for two basic concepts of diversity, namely richness and evenness. Box 2 elaborates on these two concepts and shows how the so-called evenness index (h) of diversity, which is the complement of the Simpson index (=D = 1 – h), relates to both richness and evenness.

2.1 Diversity and management indicators by resource category

Several early attempts to devise a set of indicators of genetic diversity have included not only direct estimates of diversity (such as number of taxa), but also measures of processes that are known or likely to influence diversity (see Brown and Brubaker, 2002 for review). This is particularly the case for resources directly managed, such as *ex situ* collections (e.g. accession viability), but also for resources *in situ* (e.g. proportion of a species present in a protected reserve). These indicators are not strictly measures of genetic diversity as such, but rather are indicators of effective conservation practices intended to have a major effect on diversity.

Resource category

In order to devise a set of indicators to measure progress toward the sustainable management of plant genetics resources, Brown and Brubaker (2002) delineated four categories of resource based on two kinds of gene pool and two conservation strategies (*in situ* or *ex situ*). The two kinds of gene pool are broadly distinct: cultivated species with populations that have been deliberately planted; and wild species belonging to the same genus as cultivated species. These gene pools correspond largely with Harlan and de Wet's (1971) primary versus secondary gene pools although their categories were genetic, relating to ease and fertility of crossing between crops (primary pool) and their wild progenitors of (secondary pool).

Indicators by category, numbers and diversity

Table 2 adapts Brown and Brubaker's (2002) suggestions for indicators of biodiversity based on the resource categories. The lead indicators for each of the four categories in Table 2 are in essence based on numbers. This reflects the fact the total genetic diversity within a taxon broadly tends to increase with increasing population size, increasing area occupied or increasing total numbers. Thus monitoring a change in numbers of populations or numbers of individuals of one species over time usually indicates a trend in the level of genetic diversity they harbour.

Comparisons among species are less clear cut; abundant species may not always be more diverse than rarer species of the same genus. Yet several panels of researchers and policy-makers have suggested **number** as a key surrogate indicator of genetic diversity, usually supplemented with subsidiary diversity measures using genetic techniques (see Brown and Brubaker 2002 for references). Research is needed to test and confirm the reliability of the relationship between numbers and diversity at and below species level and to identify the major attributes of species that affect or predict this relationship.

Logarithm transformation

As mentioned above, aggregation is a key feature of indicators and numbers lend themselves readily to summation. However the value for the more numerous species will clearly dominate the total of numbers of entities (accessions, individuals, populations or subspecies) of different species.

Two individuals from the same population (or species) are more likely on average to share the same gene than are two individuals that come from different populations (or species), because their most recent common ancestor is likely to be closer in time. To reduce this effect, a logarithmic transformation should be applied; the aggregation should be the sum of the log of numerical values for each entity, and the sum converted back to the numerical scale. There is a theoretical sampling basis for such a logarithmic transformation in a hierarchical system for sampling neutral alleles (Brown and Hardner 2000).

The logarithmic transformation has the virtue of being straightforward, and well known in ecology. Although theoretical distributions or empirical data are generally lacking to establish equivalences among aggregating categories, it is tempting to speculate that the log transform could be extended to each higher level in a hierarchy. Thus, for example, to aggregate values from populations of different sizes, one would use as weights the logarithm of those sizes. Then aggregating species within a genus can be based on the logarithm on the number of populations per species, and in like manner for genera within families.

2.2 Wild relatives

Lack of species equivalence

While we may treat the wild species related to cultivated plants as entities distinct from crop species, they themselves do not form a single homogenous class. The main sources of problems are as follows:

- The number of taxa involved can be very large. For example, crop wild relatives (CWR) are said to number 20 000 species in Europe alone (Flor *et al*. 2006).
- The taxa differ greatly in likely importance for the improvement of their related crop, and indeed in their significance to science.
- The number and conservation status of the subspecific entities, such as ecotypes, morphotypes, outliers, etc. vary widely among genera.
- The taxa within any one genus differ greatly in their distribution, their population numbers and sizes and the likely viability of their populations.

The oat genus, *Avena*, is a typical example. Species of this genus range from being some of the world's worst and most abundant weeds to rare and endangered taxa restricted to a few islands. In a simple sum of all wild oat populations, the rare and interesting taxa would be swamped. Autogamous or apomictic species can multiply relatively few genotypes over large areas. The population sizes of such species could mislead as indicators of their standing genetic diversity. For aggregation, we need to build on formally defined genetypic differences within species (subspecies, morphotypes, ecotypes, etc), despite problems in their recognition. For example, to count the number of morphotypes of the species *Glycine clandestina* (Pfeil *et al*. 2001) as an indicator of managed diversity is more instructive than knowing the total number of populations of this species complex extant.

Management versus diversity

Brown and Brubaker's (2002) consideration of indicators for wild relatives focused too heavily on two aspects of the management of these resources, and too little on diversity *per se*. Their first indicator was a crisis-based approach applied to populations *in situ*, and addressed only the rare or endangered elements of wild crop relatives. It borrowed the experience of natural conservation agencies in codifying their 'red lists'. The management indicator for *in situ* resources was simply the proportion of such elements that were comparatively safe in that they occurred in protected areas such as natural reserves. The second aspect was invoked for samples held *ex situ*, and emphasized the actual use (use in its broadest sense), or the number of requests to gene banks for wild resources, and was applied to wild samples *ex situ*. This too is a resource managerial indicator that aims to display the importance of collections and the need for their continued support. Like the proportion of endangered species or subspecies that is conserved *in situ*, statistics summarizing use are not measures of genetic diversity.

Numbers

A better approach to measuring diversity builds on the basic positive relationship between number (the size of a population or sample) and genetic diversity. Such an approach uses as indicator the number of recognisable subspecific taxa or, conceivably, the number of organisms comprising the sample. The subspecific taxa could go beyond the formally described subspecies and include ecotypes, morphotypes, ecogeographic fragments of the full species range, or any reasonably distinct group within the whole species sample. For *ex situ* collections this would amount to a species or subspecies list together with the total number of accessions for each taxon.

2.3 Cultivated species germplasm collections

Numbers

The obvious indicator for the management of crop genetic resources *ex situ* is some function of the number, the spread among countries and the size of germplasm collections (Brown and Brubaker 2002). The disposition of collections among countries is included because it is desirable to have backup, and to have a diversity of agencies and cultures involved. One attractive feature of this measure is that considerable historic data are available both nationally and globally. Working with collection numbers as an indicator thus affords the chance to exemplify the benefits and pitfalls of indicators. Interpretation can focus on the reliability of the data and the role that subsidiary variables might play to improve interpretation. Considerable thought has been given to the assessment of collections. Holden *et al*. (1993) have detailed how variables that describe the state of a germplasm collection can be combined to yield a 'score' to attach to each accession. The International Plant Genetic Resources Institute (IPGRI, now Bioversity International) and FAO have published standards for gene bank management that provide variables and benchmark values for indicators (FAO and IPGRI 1994).

Problems in using number and as a measure of diversity

Broadly, two major problems are of concern in using the simple number of accessions as an indicator of diversity in *ex situ* collections. The first is redundancy – the amount of repetition including the level of planned backup duplication within and between collections, and of inadvertent redundancy between very similar or identical samples of an accession. The second is viability and security of accessions. This includes the quality of accessions, especially the viability of propagating material, the regeneration frequency and strategy, and the housing, staffing, security and long-term sustainability of the whole collection and the institution that houses it.

Supporting indicators

In principle, each of the collection variables can be handled as weighting or adjusting factors (Holden *et al*. 1993). Using fractional weights at the level of the accession, the effective size of a collection can be adjusted for variation in viability, estimated from subsamples of accessions, and taking account of the age of seed from the date of accessioning and known shape of viability curves as a function of seed age. Redundancy can be estimated as a probability of 'identity' for name or origin when two random accessions are compared. This could be refined using such techniques as molecular fingerprinting with an arbitrary level of divergence (e.g. 10% of fragments different).

Aggregating subspecies or species taxa

This leads us to discuss to what extent collection size is a reasonable surrogate measure of genetic diversity present in that collection. Surely the size of a germplasm collection has much to do with the significance of the crop species.

The very large global collections of wheat, maize and rice are not a measure of the inherent diversity of these crops. Hence in Table 2, the lead indicator is the number of recognizable taxa, which is an echo of that for wild diversity *ex situ* as discussed above. Yet the number of accessions of a particular taxon is indicative of the intraspecific diversity collected, assuming that extreme biases of amplification are absent or can be corrected for.

The fact that the number of wheat accessions stored globally exceeds 107 whereas that for rye is likely to be less than 106 is indicative of their comparative levels of stored diversity. This order-of-magnitude difference supports the suggestion that logarithm transformation should be used for combining sizes over species, regions, countries, etc from the sizes of heterogeneous units.

Breeding system and numbers

A question of general interest is the effect of the breeding system of a crop species on the assessment of total collection size as an indicator of diversity. In particular, it might be assumed that germplasm collections of inbreeding (self-pollinating) crops contain much less diversity than collections of outbreeding species that are of equivalent total size. However, at the level of comparing different individual accessions, the reduction in effective size due to close inbreeding may not be as marked as implied by the true-breeding tendencies within a line (Frankel *et al*. 1995). Whereas the individual seeds within an accession are likely to share the same highly homozygous parentage, the seed from different accessions may be unrelated or related through deliberate hybridization in a breeding pedigree. In other words, the accession is the product of a round of deliberate outcrossing.

Overall, self-pollination reduces effective size to some degree (theoretically a halving) and thus reduces genetic diversity, but not by an order of magnitude unless accompanied by severe bottlenecks.

2.4 Varietal diversity *in situ*

What are the meaningful indicators of genetic diversity for populations of crop species growing *in situ* on farm, particularly applicable to traditional varieties or landraces? A complete and detailed census of all extant populations of a crop species under study is almost invariably impossible. Instead we must depend on estimates from a carefully chosen sample of farms, chosen so that can be reliably up-scaled.

Varietal data gathering

The steps in the process are:

1. Specify the crop species, the region and the communities, as the basic source from which ideally a random or structured random sample of households is drawn for survey. The number and structure for the farms and the area cultivated is recorded.
2. Define the units of genetic diversity to be assessed, for example so-called 'farmer managed unit of diversity' or named varieties. This step requires participatory techniques that ask community groups of farmers to agree on their managed units.
3. Census the sampled communities and farms for these defined varieties and estimate the area under each variety.
4. Compute the summary statistics, e.g. landrace richness, evenness and divergence.

A recent synthesis of disparate data on diversity in traditional varieties of 17 field and horticultural crops (27 species) growing in eight countries (Jarvis *et al*. 2008) illustrates the compilation of simple diversity indicators. Table 3 is an extract of these data, specifically for rice landraces in Nepal, and, by way of aggregation, the overall estimates for all crops and communities in the study.

The data for rice in Nepal were based on three contrasting communities directly representing over 1500 ha of rice-fields (line 3). The communities differed in degree of dependence on traditional cultivars (4 and 5), and rice-field size (7). For the rice fields in this study, the richness of diversity at the level of the individual farm (line 9) exceeded one landrace per household, and was very high at Kaski. The evenness index (h) (line 9) was appreciable – two random plants on one farm were almost as likely to belong to different varieties as to the same variety. Substantial differences were evident at the community level (lines 10 and 11).

Overall perspectives on crop landrace diversity *in situ*

Most of these variables were readily aggregated to more crops and to higher scales to yield very interesting overall summary measures. The remarkable features to emerge were that the majority of farmers who grew landraces were likely to grow more than one such distinct variety, and that farmers in the same community tended to adopt divergent varietal strategies. Two trends significant for developing indicators were: (1) a close relationship between richness and evenness index (correlations exceeding 0.90); and (2) an appreciable positive relationship between farm field area (log scale) and diversity. These results are important for two reasons. First, farm field area (or population size) within crops, culture and environments is a valuable, albeit surrogate, comparative indicator for on-farm genetic diversity. Second, the evenness index is a good estimator of richness of diversity. The evenness index (h) is assessable in relatively small samples because it converges with the true underlying population value, whereas richness does not reach its population value until the whole of the population is counted.

Farmer-named varieties and diversity

Statistics based on farmer-named variety are questionable as valid measures of genetic diversity. Some authors question their validity when they detect minor discrepancies at the DNA level between variety names and the genes they contain. This matter is discussed further in Section 5, below. Here we make the obvious point that variety names are assessable rapidly over a wide sampling base, enabling broad hypotheses for the distribution of diversity. Many farmer managerial decisions are made at the varietal level, as do many modes of selection (such as climate, soil, elevation, maturity time) operate at the whole field. By planting a reputedly tolerant variety in a stressful situation, farmers reinforce the attributes of the varieties they recognize as units of diversity. They directly benefit from wise decisions based on names, and suffer the consequences of poor ones.

3. GENETIC EROSION

Genetic erosion is a process that refers to a change in genetic diversity over time, and as such is difficult to specify in an index or indicator. To monitor changes in the rate of genetic erosion strictly requires directly comparable if not identical measures of the state of a system at several points in time. Alternatively, it is possible to measure the major agents of erosion (e.g. deterioration or destruction of habitat due to urbanization, land clearing, overgrazing, salinization, drought, climate change, etc). However, such indirect measures are very broad and have other and possibly more profound impacts than causing loss of diversity.

Defining genetic erosion

Maxted and Guarino (2006) define genetic erosion as follows: "Genetic erosion is the permanent reduction in richness (or evenness) of common local alleles, or the loss of combinations of alleles over time in a defined area." This is helpful, in that it draws attention on the aspect of local adaptation. However, it is not clear why a definition should specify reductions in either richness or evenness. The problem with taking too broad a definition for the purpose of constructing an indicator comes in the summing up, the aggregation. Neutral or trivial changes could mask critical changes when summed over loci, genotypes, populations or species. A temporal indicator should reveal and be most sensitive to the changes of concern and not be swamped by relatively unimportant changes. For example, the loss of a few alleles at a highly polymorphic microsatellite locus is likely to be of trivial or no importance compared with the loss of disease resistance alleles. An additional problem lies in stressing combinations of alleles; in sexual species all multilocus genotypes are unique and ephemeral. Thus when a claim is made that some percentage of distinct clones or genotypes have been lost from a region or a species, this is not necessarily genetic erosion. The life of each genotype is finite in sexually reproduced species, although vegetative reproduction might prolong that life (such as in named cultivars of fruit trees). A reduction in population size, and not increased recombination, is the primary agent of erosion. Thus an inclusive concept and definition of genetic erosion such as Maxted and Guarino's may be theoretically rigorous, but it does not readily lead to practical ways of monitoring the key issues of the phenomenon.

Appraising erosion

For indicators, it is more important to focus on the loss of genes or genotypes of concern within specified regions or production systems, rather than working with inclusive concepts and measures of the whole dynamics of diversity in the full geographic context. Fluctuations in the diversity of all rare gene combinations over time and in particular patches of a spatial distribution can be a distraction, unless they are indirectly measuring the loss of important components of the genome. Far more critical is the loss of highly localized alleles, locally adapted complexes or unique specific uses, if they cannot be replaced by recombination of genes from other populations. Even if we had fully detailed inventories of genotypes in space and at two time points we would still require expert assessment of gene-pool changes to be in a sound position to speak about significant genetic erosion.

Erosion in retrospect or in prospect

A consequence of this reasoning is that relevant measures of genetic erosion will inevitably include some subjective assessment based on expertise and local knowledge on the significance of any loss. The inclusion of such evaluative information in measuring erosion is desirable. The challenge is to format it in such a way that at least a tentative quantitative treatment is possible. The FAO survey and database of reported instances of genetic erosion has the potential to provide the basic information for constructing such a measure (Diulgheroff 2004). Many of the records so far assembled are in descriptive, narrative style of local expert opinion. Summing these stories over crops or regions or time periods requires

their conversion to quantitative estimates, which is a significant challenge.

We should adopt a procedure that could look back (retrospective) or look forward (prospective). In the former case the researcher has before him or her a gene pool containing some variation and asks the question as to what proportion remains of diversity that was known or assumed to have been present a decade ago. Initially one could work with a richness concept of diversity. The estimate of what was previously extant should rely on as much evidence as possible.

Alternatively a predictive or prospective view could be appropriate. In this case two quantities are essential for any reported instance. These are:

1. A measure of the significance if the gene pool in question were to become extinct. This is approached by estimating the extent of the total similar diversity at risk. This could in turn be based on the area cultivated or the number of varieties or populations with a factor of 0.20 as an estimate of the proportion of all diversity (in this case allelic richness) that is locally common (Brown and Hardner 2000). Suppose 20% of the area or of the varieties are deemed to be at risk. Then this amounts to 0.2 x 20% = 4% of the species genetic diversity imperiled.
2. A category of the likelihood of loss under the current situation, with no intervention (in some time period such as one decade) Classes: C = Almost certain (P >90%); L = likely (P>50%), U = unlikely but threat still real (P<50%), V = very unlikely (P<10%). This may be affected by the area growing these varieties.

Both these are subjective estimates, but ideally could be based on local knowledge of the specific crop and threats to it. Any existing survey data can be used within this framework to support the estimates. While individual estimates and predictions may be prone to error, the framework is a way to codify the best opinion and the averages will converge to give a trend. Finally the predicted erosion is estimable as the proportion of the resource under threat of erosion multiplied by the estimated probability of loss.

From narrative to estimate

The basic task in this procedure is to convert a series of descriptive narratives of the state of a variety of gene pools into number that can be compared in time and among cases. Table 4 gives some examples from the FAO database of this transformation. In addition to the two erosion variables there are several parameters to specify the geographic sampling space and three categories of aggregation: the kind of management (cultivated versus wild used versus wild unused), the taxonomic level of loss and the major kinds of threat. As date-marked reports accumulate in the database over time it should be feasible to summarize trends in estimated rates of realized erosion or predicted rates of erosion in prospect for various categories of crops within decade intervals or due to various agents. The essential statistic is the proportion of variants (alleles, genotypes, or populations) lost or likely to be lost in a given time period (for example, a decade). Such estimates can be combined as weighted or unweighted averages. Note that this statistic is based on richness of diversity only, a choice that is debatable but probably the best or at least the most obvious and practical way forward currently. Other concepts of diversity might be developed later.

The four essential elements of the procedure are:

1. Specifying the sample basis that is the subject of the inferences and that guides the aggregation of estimates.
2. Estimating the diversity previously present.
3. Estimating the extent or fraction of the diversity that is at risk.
4. Estimating the likelihood of the loss occurring.

The key assumptions and problems of this model are that:

- diversity is uniformly spread (but overall at risk 'hot spots' probably balance very safe ones);
- the likelihood of loss cannot be estimated retrospectively as the taxon is known to be present today. Past erosion rates will require guesses about what has disappeared; and
- the fraction of diversity that is 'localized' will increase as the proportion of threatened resource increases.

4. GENETIC VULNERABILITY

Whereas genetic erosion is a key aspect of the dynamics of diversity in time, the phenomenon of genetic vulnerability arises from patterns of deployment or impoverishment of genetic diversity in space. Populations of a crop species are said to be genetically vulnerable if they lack the diversity necessary to adapt to a biotic challenge or to an abiotic stress that is likely to intensify. The concept of vulnerability implies a lack or low level of genetic diversity, most graphically realized when vast areas of a region are a monoculture of a single variety. If one plant succumbs to a newly arriving

disease, to a new biotype or to a new extreme of climatic stress, all the fields of the region respond similarly because of their shared genetic heritage particularly for the genes involved in the host plant's susceptible (or 'compatible') response. The concept of 'vulnerability' could apply to a whole range of adverse situations arising from the precariousness of living systems. It is arguable that for vulnerability to be 'genetic' requires that other varieties or populations exist elsewhere that contain resistance or tolerance genes that would have moderated the loss in yield if they had been present. Thus the concept of genetic vulnerability should go beyond mere genetic uniformity *per se*. Ideally, genetic vulnerability should add the notion of genotype × environment interaction, i.e. not all genotypes (and in particular not all populations or varieties from other regions) succumb as readily as the home population to the new threat to yield. Indicators of genetic vulnerability should therefore include:

- a measure of the lack of genetic diversity, particularly for resistance genes affecting host-plant response to major likely diseases; and
- a measure of lowered diversity of host-pathogen interactions and differential responses to different biotypes, with some spatial structure.

Here, we first consider indicators for genetic vulnerability to biotic challenges, and then assess the extension of this framework to indicators for vulnerability to abiotic stresses such as climate change.

Kinds of vulnerability

Table 5 lists four kinds of genetic vulnerability upon which indicators can be framed. The first of these is genetic homogeneity. Increasing diversity in the current cropping region lowers vulnerability. Strictly, the diversity should refer to the genes determining plant response to disease. It is insufficient to have a large number of named varieties as a hedge against crop failure if they share the same genes for resistance. This was the case in the USA, where male-sterile yet disease-susceptible cytoplasmic DNA was shared among many maize hybrid varieties, resulting in them all being vulnerable to the southern corn leaf blight. However, knowledge of the comparative resistance structure of the varieties available to farmers is generally lacking, so that a census of variety names may be the only readily obtainable information.

Richness and evenness of varieties as indicators of genetic vulnerability

The indicator for the initial concept of genetic vulnerability is varietal diversity measured as both richness (the number varieties per crop, reduced if any are known to be closely related) and evenness (as measured by the evenness index). Computing the latter requires estimates of the area planted to each variety. High scores of richness imply there are many future varietal options near at hand and that seed is available for increase if needed. High richness implies insurance against pathogen evolution. In some cases, richness is high but much of the region is planted to a single dominant variety. When the dominant variety succumbs to a new disease biotype, losses will be incurred for a few seasons until more resistant varieties are multiplied and deployed. On the other hand, high evenness (lack of dominance) implies resistance diversity is already deployed to meet a new stress, and could save the farmer from severe immediate loss. It is therefore arguable that a high value for evenness diversity (i.e. low dominance) is a better indicator of low genetic vulnerability than is a high richness score.

Mutational vulnerability

The second type of vulnerability listed in Table 5, mutational vulnerability, specifically aims to conceptualize vulnerability to a new virulence mutation in a pest organism. Strictly speaking, the pathogenic properties of a future new virulent mutant are unknowable. One approach to a quantitative measure is to test the response of the present cultivar(s) to a random sample of distinct isolates or defined pathotypes. From these data it is possible to compute the probability of infection or the average level of damage caused by non-local isolates. The scores for each pathotype are not weighted by the pathotype frequency of occurrence. The indicator is thus the probability of disease (or the measured adverse effect caused by the disease) in non-local environments. Clearly this indicator requires experimental measurement, essentially the assessment of the performance of a representative sample of local material in alien stress-prone environments. Many breeders routinely conduct trials for many crop-disease or pest situations, but the data are dispersed and rarely synthesized. The summing of averages of individual variety scores, weighted by the current frequency of the varieties on farm in a given region, would provide a synthetic overview of mutational vulnerability. Technical consistency of approach is obviously necessary for the comparison of estimates over time and over different locations.

Migrational vulnerability

The idea behind recognizing migrational vulnerability as distinct from mutational vulnerability is to divide future risks into two categories. Defining the specific actual agent of risk in the mutational case is virtually impossible. The nature of a new mutant pathotype of a disease (its virulence spectrum or aggressiveness) in the future cannot be known for certain. Therefore we cannot test specifically for genetic diversity to meet such a possible future challenge. The only strategy for unknowable risks is to retain as much diversity as possible. On the other hand, migrational vulnerability refers to pressures that are currently absent from a certain home environment, but are foreseeable as inevitably arising from an alien source at some future date if unchecked, e.g. the Ug99 pathotype of wheat stem rust (Singh *et al*. 2006).

Environmental vulnerability

Abiotic environmental stresses that arise from prolonged unidirectional changes in the physical environment, such as global warming, increasing regional aridity or increasing climatic variability, resemble the threat to crop production from the invasion of pest organisms of known virulent strains (e.g. Ug99). Once again the degree of vulnerability to such a future threat can be measured experimentally by the performance or response of a local sample of varieties to specific pressure. The values of the likely impact of several separate risks on productivity can then be integrated, weighting by an estimate of the likely probability of each threat.

Although this fourth type of vulnerability resembles migrational vulnerability in Table 5 it is worthwhile to recognise that it merits developing separate indicators because of the topicality of climate change, the marked difference in spatial scales, in how the stresses increase and in how agencies will respond to such data. Plant ecologists (e.g. Gómez-Mendoza. and Arriaga, 2007) are developing approaches to model changes in the natural geographic distribution of species under various scenarios of future climate. These authors used current distributions to predict decreases of between approximately 1% and 50% for different species of *Pinus* and *Quercus* in Mexico as a result of climate change. They use these estimates as measure of differential species vulnerability and recommend conservation priorities.

Off-site testing – pursuing measurement of G × E

It may seem to be overly problematic, unduly complex and impractical to attempt a systematic, detailed risk and genetic remediation analysis to derive measures of vulnerability. The need to attempt such computation arises from the limitation of relying solely on estimates of varietal richness diversity alone. Such counts lack a test of relevance of that diversity, i.e. whether it will help cope with future threats to productivity. As mentioned at the outset of this section, the unifying concept underlying reduced genetic vulnerability is the provision of a diversity of interactions. Whether this can be measured satisfactorily by the tools of genotype × environment (G × E) analysis in plant breeding remains to be investigated. In this case 'genotype' represents the suite of available varieties and 'environment' the different pathogen populations or abiotic stress levels. Situations of low genetic vulnerability obtain when the G × E component of variance accounts for a large fraction of the overall performance variance, particularly when different cultivars are resistant or perform better in different stress states. Another indicator is the character of the variance–covariance matrix of performance across environments. Situations of low risk are associated with negative covariance values. This result is analogous with modern investment portfolio theory of market economics, in which risk (i.e. vulnerability) is minimized when the total investment is made over a diversity of the stocks whose performance patterns in the past feature negative covariances. A portfolio of stocks that have responded differentially provide the best hedge against risk.

5. ROLE OF MOLECULAR TECHNIQUES

We live in an era when technology is evolving at breakneck speed, opening up new possibilities almost faster than the possibilities are foreseen. Full genomic sequencing of all crop species and many of their wild relatives is now feasible. Further, it is possible to accumulate large amount of sequence data on a growing number of diverse genes and on relatively large numbers of individuals. This technical capacity impacts on all areas of biological inquiry. What then is the role of such techniques and the data they generate in the formulation of indicators for genetic diversity, genetic erosion and genetic vulnerability? Notwithstanding the amazing growth in capacity of these new techniques, they are still costly in time, technical and financial resources. Brown and Brubaker (2002) considered the question of what established and emerging molecular techniques offer in devising, implementing or improving indicators for sustainable management of plant genetic resources. The relation of their findings to the task of measuring genetic diversity is extracted and revised below.

5.1 Genetic precision

Molecular techniques give the power to monitor genetic variation at the elemental level of DNA sequences. It is now possible to compare organisms from the genome level (using, for example, fluorescent *in situ* hybridization [FISH] and genomic *in situ* hybridization [GISH]) down to the level of single nucleotides (DNA sequencing and single nucleotide polymorphisms [SNPs]). In theory, such data could increase the validity and credibility of indicators by increasing the clarity of interpretation.

The immediate and obvious benefit is the flexibility and precision by which genetic diversity can be assayed. Marker systems can be tailored to specific organisms to accommodate differences in breeding systems and relative levels of genetic diversity, and can be scaled depending on the number of accessions to be screened, how many loci are needed and which sequences in the genome are to be sampled. Furthermore, many anonymous DNA markers (e.g. restriction fragment length polymorphisms [RFLPs] and amplified fragment length polymorphisms [AFLPs]) can be mapped in the genome, guaranteeing an even sample of the genome. With sequence tagged sites (STSs) developed from expressed sequence tags (ESTs) it is even possible to use expressed genes specific to life history stages, rather than anonymous sequence differences, to assay genetic differences among accessions. Because database comparisons can often identify the functional product of an EST, the gene bank manager not only gets an indicator of genetic diversity and relationships among accessions but an increase in the information content of the sampled accession.

There is a fundamental gain in genetic knowledge; not only is it possible to prove that two individuals or two gene copies differ, but they can be placed in a phylogenetic hierarchy of relationships based on their recency of a shared ancestor. Once this is done, the phylogenetic diversity of the collection can be estimated (Crozier 1997). Calculating the phylogenetic diversity of a collection allows the extension and improvement of core collections of the gene pools of wild relatives. Phylogenetic diversity measures can also be used to identify a subset of related wild species that maximizes the genetic information content of the collection.

5.2 Limitations

There are clear benefits to the greater use of these more precise measures of genetic variation. However, such techniques cannot be considered essential in the development of indicators. First, they are costly in human and financial resources. They can only be employed in a few collections. Therefore, the selection of which species, which genes and which samples is inescapable and crucial. Since the aim is to obtain the maximum amount of useful information from a limited sample, the use of core collections is an obvious approach. The designation of core collections should use all the data available to ensure their entries are representative of genetic diversity. The basic procedure is to recognize groups of related or similar accessions within the collection and to sample from each group. DNA sequence analysis provides the opportunity to measure how different these empirically derived groups are and to test for relationship between them.

Second, the techniques themselves are rapidly evolving in accuracy and in capacity to handle larger samples. This complicates the role of indicators in tracking diversity over time. The ideal is to maintain the technical capacity to carry out comparable assays of diversity. However, this is difficult and costly to do as knowledge advances and fashions or priorities change.

Third, while it is feasible to assay known functional sequences, much of the present data refer to anonymous or more neutral regions of the genome. There is need to get a representative view of the genome, that different kinds of gene sequences evolve at different rates and march to different evolutionary rules. The challenge is to assay the genes that matter for performance and benefits to farmers and not have that information lost in summary measures that are dominated by largely trivial variation.

5.3 Roles

Molecular techniques therefore have a secondary, but nonetheless important, role in indicator development. They enable a deeper appreciation of the recognition of taxa and hence provide a ground-truthing of the diversity units monitored at the phenotype level. Sequence changes introduce a temporal perspective (coalescent theory) of evolving relationships and the measurement of evolutionary processes such as migration and breeding systems.

A key assumption in the proposals for indicator of genetic diversity made in earlier sections is the reliable, consistent recognition of identities of types (subspecies, variety names) and differences not only within communities, but at broader

spatial and temporal scales. Molecular techniques have a role in testing the limits of that assumption. Likewise, one issue for indicators of genetic erosion is the matter of identifying locally common alleles that are important for adaptation. Molecular techniques have a possible role in assessing the uniqueness of such alleles in sample samples. In addition, molecular fingerprinting of a current and a past sample of varieties could in principle measure proportionate declines in genomic diversity. Such an approach requires the benchmarking of the significance of observed decay rates of molecular diversity. For genetic vulnerability, it is important to add data on performance in assays of biotic and abiotic stress to measures of varietal homogeneity. Molecular techniques already play a significant role in identifying a very restricted representative set of standard isolates needed for such bio-assays.

6. RECOMMENDATIONS

This survey of the conceptual basis of indicators of genetic diversity, genetic erosion and genetic vulnerability leads the following specific recommendations to advance their development:

- **Develop supporting indicators:** Having decided the key indicators of diversity, particularly those that use numbers and size as surrogates of genetic diversity, develop a set of subsidiary indicators, methods of measuring and associated research to improve interpretation of changes in values of the primary indicators:

  e.g. number of accessions – level of duplication

- **Research diversity indicators:** Research the relationship between population size and diversity and between diversity measured in different ways. Explore major categories of crop types in relation to diversity patterns to determine the scope for up-scaling, to see whether the predictive power of the categories can be strengthened. The large databases of *ex situ* seed banks and associated data (e.g. EURISCO, SINGER) will be important as exemplars for computing diversity.
- **Research genetic erosion indicators:** Compile data from current databases (e.g. FAO database) of reports of genetic erosion and develop the process of codifying such reports. Develop categories of erosion situations as a basis for up-scaling more detailed studies. Develop and exemplify methods of incorporating more subjective, qualitative data and expert opinion into measures of erosion.
- **Research genetic vulnerability indicators:** Develop a survey to garner data on the response of varieties of major crops currently deployed to exotic strains and biotic pressures. Investigate the setting up of networks of participating labs that test the response of varieties to a sample of pathogen strains. Discover and exploit current cooperative international testing of varieties for resistance. Develop case studies with good collaboration for various kinds of crops as part of methodology. Develop measurement of abiotic environmental vulnerability that combines predictions of climate change with varietal diversity of sensitivity and other demands for genetic diversity such as agronomic, social or market pressures.
- **Develop and test protocols:** Since the wide implementation of indicators is the raison d'être of this field, develop the necessary tool-kit of instructions and examples to assist partners and collaborative efforts.

Caution should be used when using biodiversity indices, as they are merely attempts at simplifying complex systems and may often misrepresent what they are meant to simplify. Yet major management decisions have to be made, and indeed are being made. Such decisions can either invoke diversity criteria, e.g. saving endangered gene pools, or will be made on grounds other than the biological well-being of the system. Our task is to decide on the best simplified measures, which may be less than desirable but still ensure the most important outcomes.

A clear need and golden opportunity exist for research to develop the ideas outlined here and the indicators proposed, and to test them with suitable databases. Particularly for *ex situ* collections, there are a growing number of synthetic databases where data have been brought into comparable format.

ACKNOWLEDGMENTS

I am grateful for discussions with Drs L. Collette and S. Diulgheroff (FAO), Dr D.R. Marshall, L. Guarino (Global Crop Diversity Trust), Drs T. Hodgkin and E. Dulloo (Bioversity International) Dr J.-L. Pham (Institut de recherché pour le développement [IRD]) for comments on previous drafts of the report.

REFERENCES

Brown, A.H.D. and Brubaker C.L. 2002. Indicators for sustainable management of plant genetic resources: How well are we doing? In JM.M. Engels, V. Ramanatha Rao, A.H.D. Brown and M.T. Jackson, eds. Managing Plant Genetic Diversity, pp. 249–262. CABI Publishing, Wallingford, UK.

Brown, A.H.D. and Hardner, C.M. 2000. Sampling the gene pools of forest trees and *ex situ* conservation. In A. Young, T.J.B. Boyle and D. Boshier, eds. Forest Conservation Genetics Principles and Practices, pp. 185–196. CSIRO Publishing, Collingwood, VIC, Australia.

Brown, A.H.D. and Hodgkin, T. 2007. Measuring, managing, and maintaining crop genetic diversity on farm. In D.I. Jarvis, C. Padoch and H.D. Cooper, eds. Managing Biodiversity in Agricultural Ecosystems, pp. 13–33. Columbia University Press, New York, USA.

Brown, A.H.D. and Weir, B.S. 1983. Measuring genetic variation in plant populations. In S.D. Tanksley and T.J. Orton, eds. Isozymes in Plant Genetics and Breeding Part A, pp. 219–239. Elsevier Science Publishing Company, New York, USA.

Crozier, R.H. 1997. Preserving the information content of species: genetic diversity, phylogeny and conservation worth. Ann. Rev. Ecol. Syst. 28: 243–268.

Diulgheroff, S. 2004. A global overview of assessing and monitoring genetic erosion of crop wild relatives and local varieties using WIEWS and other elements of the FAO global system of PGR. In B.V. Ford-Lloyd, S.R. Dias and E. Bettencourt, eds. Genetic Erosion and Pollution Assessment Methodologies. Proceedings of PGR Forum Workshop 5, Terceira Island, Autonomous Region of the Azores, Portugal, 8–11 September 2004, pp. 6–14. Published on behalf of the European Crop Wild Relative Diversity Assessment and Conservation Forum, by Bioversity International, Rome, Italy. 100 pp. Available at http://www.bioversityinternational.org/fileadmin/bioversity/publications/pdfs/1171.pdf (accessed 11 January 2010).

EASAC. 2005. A user's guide to biodiversity indicators. Holbrooks Printers, Portsmouth, UK. Available at http://www.leopoldina-halle.de/cms/fileadmin/user_upload/leopoldina_downloads/biodiversity.pdf (accessed 11 January 2010).

EEA. 2007. Halting the loss of biodiversity by 2010: proposal for a first set of indicators to monitor progress in Europe. EEA Technical report No. 11. European Environment Agency, Copenhagen, Denmark. Available at http://www.eea.europa.eu/publications/technical_report_2007_11/at_download/file (accessed 11 January 2010).

FAO and IPGRI. 1994. Genebank Standards. FAO, Rome, Italy, and IPGRI, Rome, Italy.

Flor, A., Bettencourt, E., Arriegas, P.I. and Dias, S.R. 2006. European crop wild relative conservation criteria – indicators for the CWR species' list prioritization. In B.V. Ford-Lloyd, S.R. Dias and E. Bettencourt, eds. Genetic Erosion and Pollution Assessment Methodologies, pp. 83–87. Bioversity International, Rome, Italy.

Frankel, O.H., Brown, A.H.D. and Burdon, J.J. 1995. The Conservation of Plant Biodiversity. Cambridge University Press, UK. 313 pp.

Gómez-Mendoza, L. and Arriaga, L. 2007 Modeling the effect of climate change on the distribution of oak and pine species of Mexico. Conserv. Biol. 21: 1545–1555.

Hamrick, J.L. and Godt, M.J.W. 1989. Allozyme diversity in plant species. In A.H.D. Brown, M.T. Clegg, A.L. Kahler and B.S. Weir, eds. Plant population Genetics, Breeding, and Genetic Resources, pp. 43–63. Sinauer Associates Inc., Sunderland, MA, USA.

Harlan, J.R. and de Wet, J.M.J. 1971 Toward a rational classification of cultivated plants. Taxon 20: 509–517.

Holden, J., Peacock, J. and Williams, T. 1993. Genes, Crops and the Environment. Cambridge University Press, Cambridge, UK.

Jarvis, D.I., Brown, A.H.D., Hung Cuong, P., Collado-Panduro, L., Latournerie-Moreno, L., Gyawali, S., Tanto, T., Sawadogo, M., Mar, I., Sadiki, M., Thi-Ngoc Hue, N., Arias-Reyes, L., Balma, D., Bajracharya, J., Castillo, F., Rijal, D., Loubna Belqadi, Rana, R., Saidi, S., Ouedraogo, J., Zangre, R., Rhrib, K., Chavez, J.L., Schoen, D., Sthapit, B., De Santis, P., Fadda, C. and Hodgkin, T. 2008. A global perspective of the richness and evenness of traditional crop-variety diversity maintained by farming communities. P. Natl Acad. Sci. USA 105: 5326–5331.

Magurran, A.E. 2003. Measuring Biological Diversity. Blackwell, Oxford, UK.

Maxted, N. and Guarino, L. 2006. Genetic erosion and genetic pollution of crop wild relatives. In B.V. Ford-Lloyd, S.R. Dias and E. Bettencourt, eds. Genetic Erosion and Pollution Assessment Methodologies. Proceedings of PGR Forum Workshop 5, Terceira Island, Autonomous Region of the Azores, Portugal, 8–11 September 2004, pp. 35–45. Published on behalf of the European Crop Wild Relative Diversity Assessment and Conservation Forum, by Bioversity International, Rome, Italy. 100 pp. Available at http://www.bioversityinternational.org/fileadmin/bioversity/publications/pdfs/1171.pdf (accessed 11 January 2010).

OECD. 2001. Environmental Indicators for Agriculture. Volume 3: Methods and results. Organization for Economic Co-operation and Development, Paris, France.

Pfeil, B.E., Tindale, M.D. and Craven, L.A. 2001. A review of the Glycine clandestina species complex (Fabaceae: Phaseolae) reveals two new species. Aust. Syst. Bot. 14: 891–900.

Singh, R.P., Hodson, D.P., Jin, Y., Huerta-Esopino, J., Kinyua, M., Wanyera, R., Njua, P. and Ward, R.W. 2006. Current status, likely migration and strategies to mitigate the threat to wheat production from race Ug99 (TTKS) of stem rust pathogen. CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources 1, No. 054, 13 pp.

TABLE 1

Optimal properties of indicators of genetics resources. Properties shown in italics are highly desirable

Intrinsic properties of the indicator
• *scientifically valid and credible simple and cheap to assay, easily implemented* • *simple and cheap to assay, easily implemented* • *aggregative (capable of meaningful summation over items and scales)* • easy to understand, unambiguous and robust • based on accepted and well–documented methods • cost-effective (i.e. an expensive indicator should yield greater information) • adaptable for use at a range of spatial scales • capable of showing trends over time • sustainable in the long-term, unlikely to become obsolescent
Properties that managers of PGR appreciate
• *render progress evident* • relevant to management objectives and fit in a policy framework • part of the management cycle, and not an end in itself • focused on the use of information rather than the gaining of it • kept under review and refined when necessary
Properties that users (i.e. breeders, scientists, farmers) of PGR require
• *have been developed with all people involved: stakeholders, monitors* • reflects an *essential, fundamental and highly valued element* of the object being monitored • *warns of emerging issues* or problems early socially and politically acceptable
Properties that users (i.e. breeders, scientists, farmers) of PGR require
• partnerships between communities, governments, companies and research agencies setting up and running the process and sharing information • the provision of adequate resources (time, expertise, funds) • the commitment to collect new reliable data • continuing research to improve indicators, interpretation and determine cause and effect

Source: Brown and Brubaker (2002)

TABLE 2

Indicators of genetic diversity in four categories of plant genetic resources for food and agriculture, managed in a particular region or country for conservation and use

Gene pool	*In situ*	*Ex situ*
Cultivated	• **Number and frequency of landraces, and proportion of the area planted that is growing them** • Environmental amplitude of crop area • Number of farmer selection criteria, and evolution of farmer management • *Security of traditional knowledge*	• **Number of crop species, subspecies or geographic categories adequately sampled in gene banks** • **Number of accessions held in the gene bank** • **Number of collections or gene banks** • Country distribution of seed gene banks • Coverage in collections of crop diversity • *Backup duplication provisions* • *Extent of usage and representation in core collections* • *Collection health, accession viability* • *Documentation and evaluation of collection*
Wild	• **Number of species, subspecies or geographic subdivisions of taxa distributed in protected areas, that cover the species environmental range** • Abundance as population numbers and sizes, particularly of rare wild crop relatives • Gene diversity, divergence and distribution	• **Number of wild species, subspecies or geographic subdivisions of taxa related to crops adequately sampled in the gene bank** • Coverage of species range • Evolutionary relationships and taxonomic resolution • Accession viability, documentation and duplication • *Number and frequency of accessions used* • *'Prebreeding' activities, including evaluation*

The **lead** or primary indicator is in **bold**; the secondary or support indicators are measures to aid interpretation the values of the primary variables. Measures of processes that affect diversity are shown in italics.

Source: adapted from Brown and Brubaker (2002).

TABLE 3

Estimates of sampling and diversity variables in rice in the three communities in Nepal

Crop species	**Rice**			**17 Crop species**
Site or community	**Bara**	**Kaski**	**Jumla**	**25 communities in 8 countries**
The sampling base – the total area of a specific crop(s) in community or communities and countries (ha)	1 034	460	88	63 610[1]
The number of modern varieties available to the community	20	6	0	0.45[2]
Proportion of the farm area growing landraces	27	76	100	92[3]
Number of farms or households sampled	89	161	180	4 074[1]
Area of traditional varieties crop per farm (sq. m)	3 256	3 500	1 100	4 186[4]
Varietal diversity				
Farm (or household) landrace richness	1.51	3.79	1.09	1.82
Farm evenness (h)	0.15	0.46	0.03	0.26
Community richness	28	63	21	14
Community evenness	0.88	0.93	0.60	0.70
Divergence (between /total %)	83	51	95	63

[1] Grand total, unweighted;
[2] Antilog (i.e. exponent) of the average over farms of the log (1 + number of introduced varieties), unweighted;
[3] Weighted average over crops where the weights were the log of the total area for each crop;
[4] Exp unweighted average of log of farm areas.

Source: extracted from Jarvis *et al.* (2008).

TABLE 4

Measuring genetic erosion: Illustrative examples of quantitative estimates of potential erosion or the rate of erosion, based on survey reports

Variable	Definition or description	Examples			
Year	**Year of observation**	**1998**	**2001**	**2001**	**Unknown**
Region	Sensible groupings of countries		Pacific Islands	Caucasus	Pacific Islands
Country		Ecuador	Fiji	Azerbaijan	Fiji
Area	Geographic region of observation (name / km2)	3 provinces	Most	5	Most
Crop group and management type[1]	Cereal, pulse, fruit tree, root, vegetable, harvested wild, unused wild	H	R	F	F
Taxon	Name of taxon	*Vasconcella*	*Colocasia*	*Prunus avium*	*Cocos nucifera*
Threatened entity or taxon[2]	Genus, species, ssp., cultivars or populations (number)	2 spp.	28 cvs	2 cvs	4 cvs
Fraction threatened[3]	Proportion of the total, e.g. in first case 2 spp. threatened of a total of 7 spp.	2 of 7 = 0.29	28 of 112 = 0.26	2 of 8 = 0.25	4 of 14 = 0.29
Likelihood of loss (cf. IUCN species categories) [4]	Probability of loss under the current situation, with no intervention (in one decade)	0.95	0.50	0.95	0.05
Predicted erosion[5]	Proportion of resource × probability of loss	0.28	0.13	0.24	0.015
Kinds of threat[6]	New varieties; other species; major abiotic change; major biotic change; loss of farming area or wild habitat	A	NV	NV	NV
Data source	FAO database	*	†	‡	†

[1] Management class – C= cereal, R=root, F=fruit tree, H=wild harvested or used, W=wild and unused by humans. Example categories for aggregation.
[2] Level of potential loss or extinction and category for aggregation.
[3] Proportion of the total number of kinds of the higher category – order of magnitude is sufficient.
[4] Category of estimated likelihood of loss: Classes: Almost certain (P >90%); Likely (P>50%), Unlikely but the threat still real (<50% but >10%), Very unlikely (<10%).. We adopted the most conservative value for each class.
[5] Predicted erosion = proportion of resource × probability of loss × 0.20 (locally common genes).
[6] Kind of threat: NV= New Varieties; OS= Other species; C=Major abiotic change; D=Major biotic change; A=Loss of farming area.
* http://www.pgrfa.org/gpa/ecu/quesreport.jspx?quesno=1&rowno=4&instid=S-58-2&tablename=xmlanswers&iterationno=1 and R. Morales (pers. comm.)
† http://www.pgrfa.org/gpa/fji/quesreport.jspx?quesno=1&rowno=3&instid=S-66-8&tablename=xmlanswers&iterationno=1 and T. Kete (pers. comm.)
‡ http://www.pgrfa.org/gpa/aze/quesreport.jspx?quesno=1&rowno=4&instid=S-52-6&tablename=xmlanswers&iterationno=1 and Z. Akparov (pers. comm.)

TABLE 5

Indicators of genetic vulnerability

Concept of genetic vulnerability	Theoretical measure	Indicator
1) Genetic homogeneity – The standing crop consists of a single genotype or few varieties or genotypes.	The diversity of resistances in host population. Richness diversity represents diversity near at hand that could be deployed. Evenness diversity or low dominance indicates diversity deployed to meet the current pathogen population.	- The number of varieties per crop present on farm or in a region. - Evenness index – more important for disease vulnerability.
2) Mutational vulnerability – The standing crop consists of genotypes that require a single mutation in the pathogen for virulence.	The fraction of non-local pathotypes that can attack a random plant.	Probability of disease (or quantitative adverse effect) when tested with a set of distinct experimental isolates.
3) Migrational vulnerability – The standing crop consists of locally resistant genotypes that are susceptible to a new migrant strain of a pathogen or pest.	The probability that a random migrant pathogen propagule will succeed in causing disease on a random healthy plant in the population in question. This assumes the environment is favourable to the pathogen, and is calculated by integrating the frequency of particular compatible (diseased) interactions between alien disease strains on local crop genotypes, and could be distance-weighted. Ideally the statistic is also weighted by the relative frequency of pathotypes.	Proportion of plants that become diseased when grown in other disease-prone environments.
4) Environmental vulnerability – The standing crop consists of genotypes that are adapted to the current abiotic environment (climate, soil) but lack adaptation to environmental stresses that are intensifying with time.	The stress-induced yield depletion of current varieties relative to the performance of resistant non-local varieties that exhibit stress tolerance adjusted for the likelihood of degrees of stress, and for the frequency of local variety occurrence.	- Relative sensitivity of local varieties when grown in clines of increasing stress. - Proportional loss of cropping area for specific varieties following the increase of regions inhospitable due to climate change

Concept 1 is a crop-plant diversity concept; concepts 2 and 3 are defined on host–parasite interactions; and concept 4 deals with the physical abiotic environment.

BOX 1
Recently suggested indicators

OECD 2001 Indicators for agricultural biodiversity

At the gene level

- For main crops or livestock species – the total number of crop varieties or breeds registered and certified for marketing.
- The share of crop varieties in the total marketed production.
- The share of livestock breeds in total numbers of animals.
- The number of national crop varieties that are endangered.

Related information:
Crop gene banks – the number of gene banks and the number of accessions held.

At the species level

- Wild species – trends in population distributions and numbers of wild species related to agricultural species.
- Non-native species – trends in population distributions and numbers of species that threaten agricultural production or agro-ecosystems.

SEBI2010 Indicators that are relevant to agricultural genetic diversity

CBD designated focal area		Proposed indicator
Status & trends in components of biological diversity	1	Abundance and distribution of selected species
	2	Red List Index for European species – changes in status
	4	Ecosystem coverage
	6	Livestock genetics resources
	7	National protected areas
Threats to biodiversity	10	Invasive alien species
Sustainable Use	20	Agricultural area that is managed sustainably to support biodiversity
Access and benefit sharing	24	Patent applications based on genetic resources

This list is essentially the same is that from EASAC (2005)

BOX 2
Richness diversity and evenness diversity

Fundamental to the measurement of diversity is an understanding of the different concepts or meanings we might give to the expression "Population A is more diverse that population B". The first concept is that population A might harbour many more recognizable, distinct types than does population B. This we call **richness diversity**, which refers to the number of different kinds of individuals regardless of their frequencies. Population A is richer in diversity when it contains more types than population B. Another related but distinct concept of diversity is the **evenness** in frequency of the types in population A compared with population B. Evenness measures how similar the frequencies of the different variants are, with low evenness indicating the dominance by one or two types. If the frequencies of the different types in A are very similar, the variance in their frequency is lower compared with that in B.

The measure of richness is, straightforwardly, the number (k; k = 1, 2, 3 …) of types in a sample. Evenness, on the other hand, is less obvious. A standard, conceptual parameter for measuring variation in biology is the coefficient of variation of the frequencies of types, where the coefficient of variation (CV[pi]) is the square root of the variance divided by the mean frequency (p = 1/k). If all the types in the population are equally frequent, then the variance of their frequencies is very low or zero, and the evenness diversity is high. The evenness index ($h = 1 - \Sigma\, pi^2$; $0 \leq h \leq 1.0$) is also called the genetic diversity index and is the complement of the Simpson index of dominance (D = 1 – h) in ecology. The symbol **h** signifies the close parallel with expected heterozygosity in population genetics. Despite these potentially confusing names, **h** is perhaps the most understandable measure of evenness diversity. This is because **h** is simply the chance that two members of the population, drawn at random, are different.

Because of close parallel with expected heterozygosity for a single gene polymorphism in a random-mating population, we use the symbol h. It is known that the evenness index (h) is a simple function of the evenness and richness measures:

$$h = 1 - \{1 + CV^2[p]\}/k$$

This formula shows that this evenness index (h) increases as the richness (k) increases, and as the coefficient of variation of the frequency of types decreases. Numerically, **h** is largely determined by the frequency of the most frequent, or dominant type. (Hence the Simpson Index is sometimes called the dominance index.) In general, **h** is more a measure of evenness than it is of richness.

There are in theory other additional concepts and measures of genetic diversity (Brown and Weir 1983; Brown and Hodgkin 2007) that could serve as indicators. However, the two measures (k and h) discussed here are the most useful and readily understandable, and these two concepts are fundamental to the present discussion.