Diversity and Distributions, (Diversity Distrib.) (2008) **14**, 1018–1030

BIODIVERSITY RESEARCH

Gap analysis: a tool for complementary genetic conservation assessment

Nigel Maxted[1]*, Ehsan Dulloo[2], Brian V. Ford-Lloyd[1], Jose M. Iriondo[3] and Andy Jarvis[4,5]

[1]*School of Biosciences, University of Birmingham, Birmingham B15 2TT, UK,* [2]*Bioversity International, Via dei Tre Denari 472/a, 00057, Maccarese (Fiumicino), Roma, Italy,* [3]*Area de Biodiversidad y Conservación ESCET, Universidad Rey Juan Carlos, c/Tulipán s/n E-28933 Móstoles, Madrid, Spain,* [4]*International Centre for Tropical Agriculture (CIAT), Cali, Colombia,* [5]*Bioversity International, Via dei Tre Denari 472/a, 00057, Maccarese (Fiumicino), Roma, Italy*

*Correspondence: Nigel Maxted, School of Biosciences, University of Birmingham, Birmingham B15 2TT, UK, and Bioversity International, Via dei Tre Denari 472/a, 00057, Maccarese (Fiumicino), Roma, Italy.
E-mail: nigel.maxted@dial.pipex.com

ABSTRACT

Aim Gap analysis is a well-established conservation technique that identifies areas in which selected elements of biodiversity are represented and through comparison with existing in situ protected area networks identifies habitats or ecosystems that need additional protection. We aim to demonstrate that gap analysis may be extended to encompass both *in situ* and *ex situ* genetic diversity conservation strategies.

Location Global, with exemplar case study from sub-Saharan Africa.

Methods An extended methodology of gap analysis is proposed that involves the following steps: (1) circumscription of target taxon and target area; (2) assessment of natural diversity through a review of intrinsic taxonomic, genetic and ecogeographical diversity combined with threat assessment; (3) assessment of current complementary *in situ* and *ex situ* conservation strategies; and (4) reformulation of the conservation strategy through analysis of the differences between the pattern of natural, intrinsic diversity and the elements of that diversity already effectively represented by existing *in situ* and *ex situ* conservation actions.

Results To illustrate the gap analysis approach proposed, the methodology was applied to the conservation of African *Vigna* species (cowpea *Vigna unguiculata* (L.) Walp. and its wild relatives) and indicated: (1) genetic reserves should be established at the southern tip of Lake Tanganyika, the coastal area of Sierra Leone and between Lake Victoria and the other Great Lakes, and (2) 14 taxa and several countries should be targeted for further seed collection.

Main conclusions The robust nature of the extended methodology for gap analysis has been demonstrated and indicates that its scope as an effective conservation tool may be expanded to fully address the need for a more comprehensive and complementary conservation strategy that encompasses both *in situ* and *ex situ* applications. However, it should be stressed that the methodology is applicable for any form of biodiversity (wild or cultivated), where the conservation of genetic diversity is the prime goal.

Keywords
Agrobiodiversity, crop wild relative, *ex situ* conservation, gap analysis, genetic conservation, *in situ* conservation, protected areas.

INTRODUCTION

The current threats to biodiversity from genetic erosion and extinction were recognized by the CBD's Global Strategy for Plant Conservation (CBD, 2002) which in Target 9 called for '70% of the genetic diversity of crops and other major socio-economically valuable plant species conserved' and further the 2010 Biodiversity Target committed the parties 'to achieve by 2010 a significant reduction of the current rate of biodiversity loss at the global, regional and national level as a contribution to poverty alleviation and to the benefit of all life on earth'. To address this target, it is necessary to focus conservation efforts and ensure that investment is targeted where it is most needed. The major goal of plant genetic conservation is to maximize the proportion of the gene pool of the target taxon which is conserved, whether *in situ* or *ex situ* in a complementary manner and to make it

DOI: 10.1111/j.1472-4642.2008.00512.x

available for potential or actual utilization (Maxted *et al*., 1997). Historically, the goal of plant genetic conservation as stated by Marshall & Brown (1975) is to conserve '95% of all the alleles at a random locus occurring in the target population with a frequency greater than 0.05'. They estimated that to achieve this goal, it was necessary to randomly sample 50 individuals from 50 populations, although subsequently Lawrence *et al*. (1995) calculated that for any one species just 172 randomly sampled plants would suffice. The important question in reviewing current species conservation status is whether 50 individuals from 50 populations or 172 randomly sampled plants, or both, have been achieved for any species.

Both approaches appear at first sight to be simple to achieve, but it may not necessarily fully address conservation needs. The reasons for this are twofold. Effective conservation of any one population of plants *in situ* is governed by population genetic and demographic parameters that dictate the minimum viable population size, regardless of the total number of alleles that the species possesses. It is also important to understand how the genetic diversity of the target taxon is distributed across its range. The other reason is only becoming clearer with the development of knowledge of genomics and gene networks; it may be necessary to conserve genotypes as much as possible rather than individual alleles so that important 'adaptive gene complexes' are conserved.

Notwithstanding, if world *ex situ* holdings for major crops (FAO, 1998) are considered, then it would seem likely that the Marshall and Brown criterion may have been met for wheat, rice and maize, with 788,654, 420,341 and 261,584 accessions held *ex situ*, but how likely is it that it has been met for the minor crops, let alone wild species? Even for the other major crops, such as barley, sorghum, millet and beans, although in some cases extensive collections have been made do we know that the accessions adequately represent the taxon's ecogeographical range or adaptive amplitude, let alone its genetic diversity? The only way to answer this question would be to genetically analyse representative samples of a target taxon gene pool and estimate allele frequencies in many populations; total species genetic diversity could then be compared with that which is currently being conserved *in situ* or *ex situ*. In practice this has not been systematically undertaken for any crop or wild species and it is unlikely to be achieved without substantial financial resources being made available. Even though the costs of genetic analysis are decreasing substantially, the costs of sampling the genetic material required from throughout the taxon's range alone is still likely to be prohibitive. However, this does not negate the need to undertake some form of current conservation status assessment in order to formulate future conservation priorities. If, when assessing for example a complex of species closely related to a crop, the assessment indicates that one particular species is thought to be effectively conserved using a combination of *in situ* and *ex situ* techniques and a second species is less effectively conserved, then conservation priority will be given to the second species. Therefore, if the assessment of current conservation status indicates gaps in conserved materials, whether *in situ* or *ex situ*, then further conservation action is likely to be required. In the absence of detailed knowledge of the pattern of genetic diversity for a target taxon, including comprehensive information on allele and genotype frequencies and distributions, a series of secondary actions can be undertaken to estimate the effectiveness of current conservation actions by way of gap analysis.

The comparison of patterns of diversity with the proportion that is conserved and the identification of 'gaps' in conserved diversity is generally referred to as 'gap analysis'. This concept was put forward as a conservation evaluation technique that identifies new areas in which selected elements of biodiversity should exist (Margules, 1989) and has largely been applied to indigenous forests and habitats. There is now an extensive literature associated with gap analysis as a conservation evaluation technique that essentially identifies areas in which selected elements of biodiversity are underrepresented (Margules *et al*., 1988; Margules, 1989; Margules & Pressey, 2000; Balmford, 2003; Brooks *et al*., 2004; Dietz & Czech, 2005; Riemann & Ezcurra, 2005). However, in each of these cases, the focus is ensuring habitat diversity conservation within protected area networks, yet the concept of gap analysis can equally be applied to document taxonomic and genetic diversity and its distribution among existing wild populations, and to develop strategies for their genetic conservation. Burley (1988) identified four steps in gap analysis: (1) to identify and classify biodiversity, (2) to locate areas managed primarily for biodiversity, (3) to identify biodiversity that is underrepresented in those managed areas, and (4) to set priorities for new conservation action. Here we attempt to show how the original ecosystem-based conservation gap analysis methodologies can be adapted from the original habitat diversity focus for application in the plant genetic conservation context.

METHODS

Assessment of conservation efficiency and representativeness involves the comparison of 'total' natural plant diversity with the diversity already actively conserved either *in situ* or *ex situ*. In essence it will require answers to two fundamental questions: what level of diversity naturally exists *in situ*, and does the conserved diversity adequately represent that natural diversity? This is the basis for gap analysis, which can be divided into four consecutive steps.

Step 1: circumscription of target taxon and target area

First, the taxonomic (e.g. genus, section or species) and geographical (e.g. global, regional, country or province) breadth of the analysis must be established.

Step 2: assessment of natural diversity

The level of diversity occurring within the target taxon must be defined at the taxonomic, genetic or ecogeographical levels, i.e. how many taxa occur in the circumscribed taxon, but also the inherent genetic diversity within those taxa. As such, natural diversity assessment is likely to involve subordinate assessments:

(a) Taxonomic diversity assessment

This will list the taxa encompassed by the taxonomic circumspection, which is likely to involve identifying the accepted classification for the target taxon from consultation of specialist publications, taxon experts and searches of online sources of information.

(b) Genetic diversity assessment

Ideally, having established the list of taxa covered, the next step would be to collate existing data or generate new data on the inherent genetic diversity within those taxa. However, this may not be realistic as knowledge of inherent patterns of genetic diversity is often limited and it may be too resource-intensive to collate *de novo*. Consequently, genetically based approaches to conservation assessment either in terms of 'richness', the total number of genotypes or alleles present regardless of frequency, or 'evenness' of the frequencies of different alleles or genotypes, can therefore only be applied to the most highly prioritized taxa. However, proxy or surrogate measures of genetic diversity may be applied in the absence of novel estimation of genetic diversity and it can be argued that steps (a) and (c) are adequate proxies. Also the intended uses of conservation objectives for the target taxon may dictate the type of molecular analysis to be used if actually undertaken, as neutral markers may not detect potentially important diversity in high priority traits (such as pest/disease resistance, environmental adaptation, etc.). The range of molecular marker techniques that can be applied is extensive (see Nybom, 2004; Spooner *et al.*, 2005).

Whichever approach is taken, the assessment will need to determine if the genetic diversity that exists in taxa or populations is represented by samples held in gene banks, or by populations in protected areas or genetic reserves.

(c) Ecogeographical diversity assessment

As knowledge of natural patterns of genetic diversity is lacking for the vast majority of species, it may be necessary to employ some form of proxy measure such as ecogeographical diversity. This involves the collation of secondary information on the ecology and geography of the species under study (Maxted *et al.*, 1995; Maxted & Guarino, 2003). Ecological and geographical data sets can be obtained from existing online data bases, literature, herbarium specimen label data, gene bank data bases or even from new field studies if resources permit.

The basic assumption underlying this now routinely applied methodology is that ecogeographical diversity among sites is correlated to patterns of genetic diversity among the populations found at those sites. Therefore, it is assumed that sampling populations from distant locations and diverse habitats will provide a representative sample of the genetic diversity for the taxon (Maxted *et al.*, 1995). As population sampling by both botanists and conservationists is unlikely to be truly systematic, bias associated with practical constraints such as civil strife, accessibility of location or political boundaries means that the sampling is unlikely to provide the ideal representation of the ecogeographical breadth of the taxon. However, it is important when undertaking gap analysis to ensure that any bias is minimized and both populations from taxon hotspots and ecogeographical margins are sampled and included in the analysis. Recent studies (Ferguson *et al.*, 1998; S. Hargreaves, R. Hirano, M. Abberton, B.V. Ford-Lloyd & N. Maxted, unpublished data) conclude that the correlation between genetic diversity with geographical and ecological diversity is not always present and is species dependent; it is notable that species likely to introgress with cultivated forms are less likely to maintain ecogeographical related genetic distinction.

(d) Threat assessment

This is an important facet of gap analysis, as it facilitates the relative assessment of conservation priorities; those taxa most threatened will have higher conservation priority compared to those less threatened. Threat assessment is now routinely carried out through the application of the IUCN Red List categories version 3.1 (IUCN, 2001). Using herbarium or gene bank accession associated data as a basis for the assessment, the most likely criteria to be used are criterion B (geographical range in the form of either extent of occurrence or area of occupancy) and D (very small or restricted population). Each year more taxa are included in global, regional and national Red Data lists, but still a relatively small number of plant taxa have been assessed. The lack of population data is also a limitation in more universal IUCN Red List assessment.

In the absence of sufficiently detailed population data to undertake IUCN Red List assessment, a relatively large ecogeographical data set can be used to make a more tentative threat assessment. Maxted *et al.* (2005) proposed and applied a technique referred to as Taxon Vulnerability Assessment, in situations where there was insufficient data to permit formal IUCN Red Listing. Vulnerability to genetic diversity loss and even extinction is assessed by compounding seven criteria as follows: (1) Rarity is estimated from the total number of herbarium specimens and gene bank accessions of each taxon in the ecogeographical data base. It is assumed that in most cases this will provide a true indicator of actual occurrence, unless there is contrary evidence or the taxon is cultivated or very rare, both of which cases may lead to relative oversampling by collectors. (2) Distributional range is calculated by taking a given radius around each collecting locality and then merging the resulting circles, providing an approximation of the overall species range using the methodology described by Hijmans & Spooner (2001). (3) Representation in *ex situ* collections compared to herbarium collections can provide a relative estimate of whether a species' gene pool is sufficiently sampled *ex situ*. (4) The relative geographical coverage of *ex situ* collections compared to the geographical breadth based on *ex situ* conserved accessions and herbarium samples. (5) Intraspecies coverage of *ex situ* collections can be used comparatively for species that have multiple infraspecific categories to estimate if each infraspecific taxon is adequately represented in *ex situ* collections. (6) Usage potential of a species is particularly relevant for plant genetic resource conservation, where there will

be a particular incentive for conserving those species with the highest use potential and it may also be the case that species with high use potential are more likely to be threatened due to excessive utilization. (7) Taxon extinction assessment can be estimated by applying Solow's equation (Solow, 1993) as proposed by Burgman *et al.* (1995), which uses a combination of collection timing, frequency and specimen numbers. Each of these seven criteria is assessed for each species and a numerical score recorded. These are then summed to establish relative taxon vulnerability.

Step 3: assessment of current conservation strategies

The diversity occurring *in situ* can be compared to the diversity currently conserved in order to assess the efficiency of both *in situ* and *ex situ* conservation techniques.

In situ *conservation assessment*

Within the context of plant conservation, the definition of *in situ* conservation provided by the CBD (1992) includes two distinct conservation techniques: protected area (genetic reserve) conservation for wild species and on-farm conservation in the case of traditional crop varieties, widely known as landraces. Genetic reserve conservation maintains wild species in their natural surroundings, usually within an existing protected area where the site has been selected and is managed and monitored to maintain the genetic diversity of the target taxa.

- *Genetic reserve/protected area assessment:* This involves a review of existing protected areas and the species within them that are being actively managed for conservation. As few centralized data bases detail which species are being actively conserved in the world's protected areas, obtaining detailed knowledge of the protected areas in the target area is likely to involve contacting protected area managers to ascertain if particular species are present and being actively managed and monitored. Increasingly it is possible to use GIS techniques to compare the various protected area GIS layers from the World Database on Protected Areas (http://sea.unep-wcmc.org/wdbpa/) with species distributional data collated in Step 2 above to predict which priority species are found in which protected areas, but having matched these data sets there would still be a need to contact specific protected area managers to confirm the species predicted to be present are indeed present.
- *On-farm conservation assessment:* Similarly, the on-farm conservation of landraces requires reviewing existing on-farm conservation projects and the crop species included. The review of on-farm conservation is likely to be simpler than the review of protected areas due to the more limited number of on-farm conservation projects and the relative ease of discovering which crop species are included.

Ex situ *conservation assessment*

Information on current *ex situ* holdings in gene banks is available in Bioversity International's Germplasm Holdings Data base and associated directories, though these can become outdated (Bettencourt *et al.*, 1989). Other sources of what material is currently being conserved can be obtained from botanic gardens and gene banks, as well as from national and international catalogues, data bases and web sites (such as the SINGER data base of all CGIAR-based germplasm collections – http://singer.grinfo.net/, the European *ex situ* PGR information system EURISCO – http://eurisco.ecpgr.org/and the US Genetic Resources Information Network (GRIN) – http://www.ars-grin.gov/npgs/acc/acc_queries.html). Other initiatives such as that of the European Native Seed Conservation Network are now in the process of providing additional data – http://www.ensconet.eu/Database.htm.

When using ecogeographical distribution as a proxy for knowledge of genetic diversity, the ideal *ex situ* collection would contain samples from geographically diverse sites spread throughout the entire range of distribution of the crop or species. Such a proxy can be calculated using herbarium and gene bank collection data, and the circular area statistic (CA) (Hijmans *et al.*, 2001). The circular area statistic is calculated by assigning a circle of set radius (r) around each collection, and the area of those circles for all collections calculated (counting overlapping regions only once). For collections that are highly geographically clumped, the CA is relatively low compared to a set of collections geographically distributed over a wide region (due to greater overlap in clumped collections). This statistic can be used to compare germplasm collections with all collections (germplasm and herbarium collection data) in order to identify how geographically representative the germplasm collection is. Germplasm collections whose geographical distribution are representative should have a CA statistic similar to that of the entire collection. Conversely, germplasm collections that poorly represent the geographical distribution would have a low CA compared to the entire collection, due to concentrated *ex situ* collecting in regions representing only a subset of the wider distribution of the species.

Step 4: reformulation of conservation strategy

Assessment of the effectiveness of current conservation coverage in relation to natural *in situ* diversity identifies the element of diversity that is underconserved, i.e. the 'gaps' in the existing conservation strategy, and helps refocus the strategy to conserve the maximum diversity and fill these gaps. The revised priorities are likely to require complementary *in situ* and *ex situ* conservation actions to ensure the comprehensive conservation of the target taxon's gene pool.

In situ *conservation priorities*

- *Genetic reserve/protected area* – The location and establishment of genetic reserves within existing protected areas avoid resource expenditure on purchasing new sites for protection and can be based on relative taxon concentration. Therefore, new genetic reserves are often established within existing protected areas and the highest concentrations of actual and predicted

target taxa within protected areas indicate the areas where *in situ* activities should be focused.

• *On-farm conservation priorities* – Similarly, species richness can be used to indicate priority sites for *in situ* conservation for the on-farm conservation of landraces. Areas that have either a high concentration of landraces of multiple crops or landraces of individual crops will be the desired sites to establish on-farm conservation projects.

Ex situ *conservation priorities*

Areas that have been undersampled in *ex situ* collections are highlighted as priorities for future collection and *ex situ* conservation.

RESULTS AND DISCUSSION

To illustrate the proposed methodology for plant genetic diversity gap analysis, a case study for cowpea (*Vigna unguiculata* (L.) Walp.) and its wild relatives (Maxted *et al.*, 2005) is reviewed. Cultivation of various bean species from the genus *Vigna* Savi plays an increasing role in subsistence agriculture in sub-Saharan Africa and Asia (Ehlers & Hall, 1997; FAOSTAT, 2007). Yet until recently, we have remained largely ignorant of their exploitation potential or how these important genetic resources might be most effectively conserved for use by future generations. Although cowpea production itself has increased tenfold in the last 20 years (www.iita.org), further exploitation has been hampered by a lack of taxonomic, genetic and ecogeographical knowledge, lack of *in situ* and *ex situ* conserved material that could be easily exploited by breeders and a lack of coordinated national, regional or international conservation strategies for *Vigna* diversity (Maxted *et al.*, 2005). African *Vigna* species are therefore an obvious target for crop plant genetic diversity gap analysis, which is outlined below.

Step 1: circumscription of target taxon and target area

This was defined by the International Board for Plant Genetic Resources (now Bioversity International) in the project commission statement. The target taxon was set as all taxa of the legume genus *Vigna* Savi and the target area was the continent of Africa.

Step 2: natural diversity assessment

Taxonomic diversity assessment

The first step was to select a classification for African *Vigna* to identify the taxa encompassed. As there was no current comprehensive classification of the genus, an amended version was used of the most recent comprehensive classification of Maréchal *et al.* (1978) which incorporated subsequently described taxa; Pasquet's (2001) concept of *V. unguiculata* and the Tomooka *et al.* (2002) conception of the subgenus *Ceratotropis* were applied. Based on this amended classification, 61 species and 63 infraspecific African *Vigna* taxa were identified.

Genetic diversity assessment

The magnitude and distribution of genetic diversity could only be assessed for *V. unguiculata* (cowpea) itself, as other African *Vigna* species have not been sufficiently studied. The cowpea gene pool is unusually large, with eleven subspecies plus several botanical varieties (Pasquet, 2001). Coulibaly *et al.* (2002) found that the wild annual cowpea had greater genetic diversity than the domesticated cowpea using amplified fragment length polymorphism markers. Mithen (1987) suggests that while domestication of cowpea occurred in West Africa, considerable genetic and biochemical diversity within *V. unguiculata* is found in southern Africa. Even now following the recent establishment of a core collection for cowpea (Mahalakshmi *et al.*, 2007) the representation of genetic diversity even within *V. unguiculata* remains relatively unclear and the broader gene pool is truly opaque, so comparison of levels of conserved genetic diversity with natural diversity is not possible because of insufficient knowledge of the magnitude and structure of the natural diversity.

Ecogeographical diversity assessment

The ecogeographical data base used to assess the geographical and ecological diversity of African *Vigna* contained data on 7289 herbarium specimens and 1802 gene bank accessions representing all 61 African *Vigna* species (Maxted *et al.*, 2005). The data set was analysed using DIVA-GIS (http://diva-gis.org). Figure 1 illustrates the breadth and density of the sites from which herbarium specimens or gene bank accessions were collected. The genus *Vigna* has a broad distributional range across the African continent south of the Sahara and north of 30° south and is most species rich in subtropical latitudes around 10° N and S. Localities have been found from sea level to nearly 4500 m, though 99% are below 3000 m, and the greatest species richness is found below 1500 m. Three hotspots of high species richness were identified, notably around the Great Lakes, the southern tip of Lake Tanganyika, and the Cameroon Highlands using inverse distance weighting (Fig. 2). Application of a climate envelope model, which assumes that the climate at the points of observation and/or collection of a species is representative of the environmental range of the organism (see Jones *et al.*, 1997, 2002), highlighted other areas of potential high species richness where the amount of collection was limited (Fig. 3).

Threat assessment

Much of the African savannah has been modified by unsustainable harvesting, logging, burning, desertification and climate change and, indirectly, by overgrazing of cattle, sheep and goats (Stuart & Adams, 1990; Stuart & Stuart, 1995). It is difficult to show a specific correlation between these general factors and *Vigna* species, although Laghetti *et al.* (1998) cite the destabilizing action of humans as the single most important factor threatening the survival of wild species of *Vigna* in Africa. In Ethiopia, drastic changes in land use have been cited as the most important factor (Engels & Hawkes, 1991), while afforestation, mainly by conifer

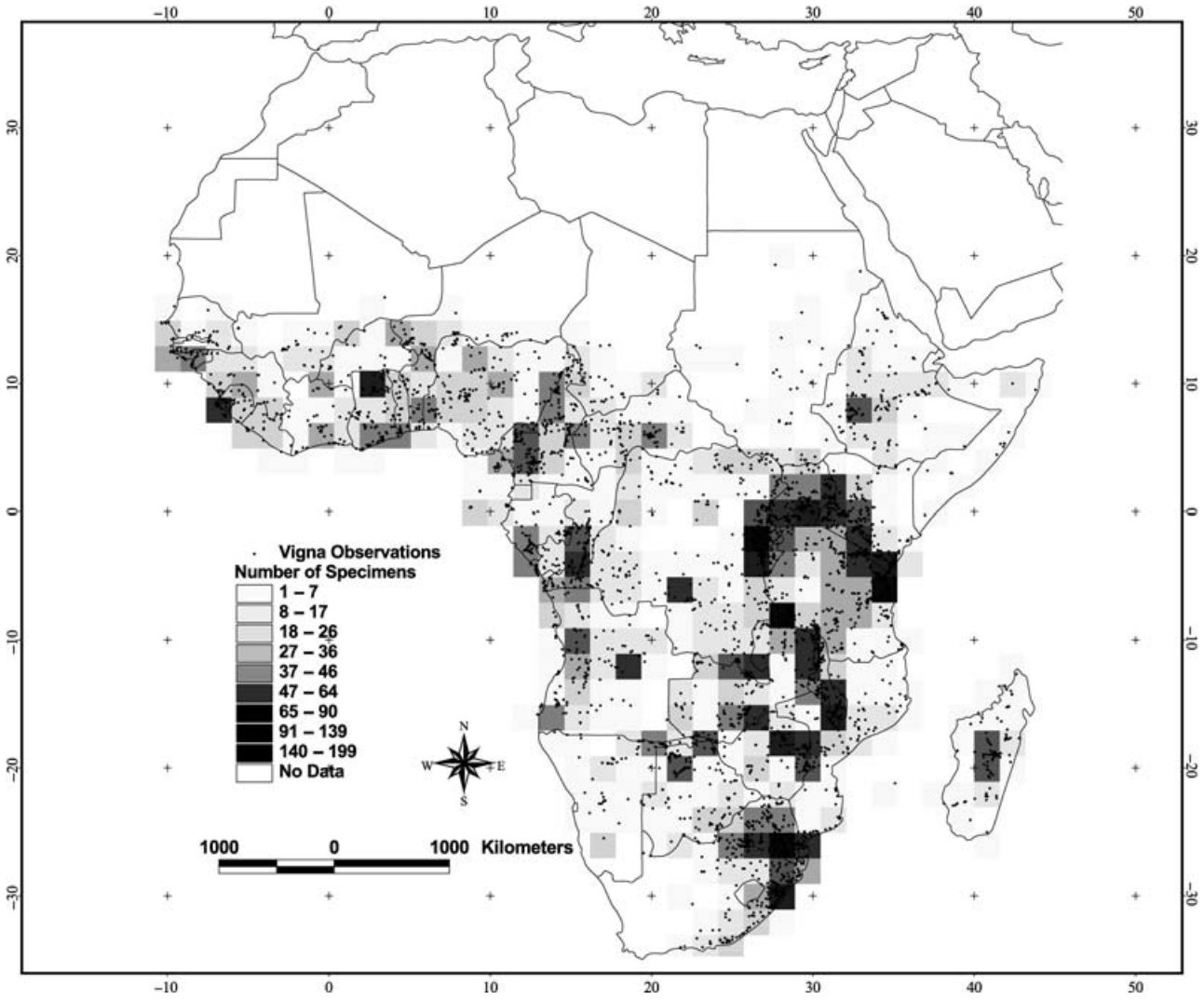

Figure 1 Herbarium specimen and gene bank accession collections in 200 km × 200 km grid cells for African *Vigna* indicating density of sampling (Maxted *et al*., 2005).

plantations, is the most important factor in Zimbabwe (Mithen, 1987). In relation to southern African *Vigna,* Pienaar (1992) cites habitat destruction for development and settlement schemes as the main threats. The few plant collectors who have recorded threats to wild *Vigna* populations cite overgrazing as an important threat factor and less often flooding or drought.

Due to the lack of sufficiently accurate population data, African *Vigna* species have not been thoroughly assessed using the IUCN Red List categories version 3.1 (IUCN, 2001). In fact the only taxon previously assessed was *Vigna comosa* ssp. *abercornensis* given a status of Vulnerable D2 in Zambia by Bingham & Smith (2002) using the 1994 Red List Categories (IUCN, 1994). This implies that the subspecies is located in a very restricted area (typically less than 100 km^2) or that the number of locations (typically five or fewer) is such that it is prone to the effects of human activities or stochastic events within a very short time period. Maxted *et al*. (2005) produced a provisional assessment for each African *Vigna* taxon using the IUCN Red List Categories and Criteria version 3.1 (IUCN, 2001). Six species were found to be critically endangered, eight endangered, ten vulnerable, five near threatened and four species data deficient (see Appendix S1 in Supporting Information). A Taxon Vulnerability Assessment was also carried out for African *Vigna* (see Appendix S2 in Supporting Information).

Step 3: assessment of current conservation strategies

In situ *conservation assessment*

Genetic reserve/protected area. Although there is passive protected area conservation of African *Vigna* species, there are currently no known reserves specifically established to actively conserve African *Vigna* species *per se* or where they are priority taxa within the protected area management plan. However, the majority of species is widely distributed in grassland, along roadsides, field margins and open primary forest, and therefore national parks

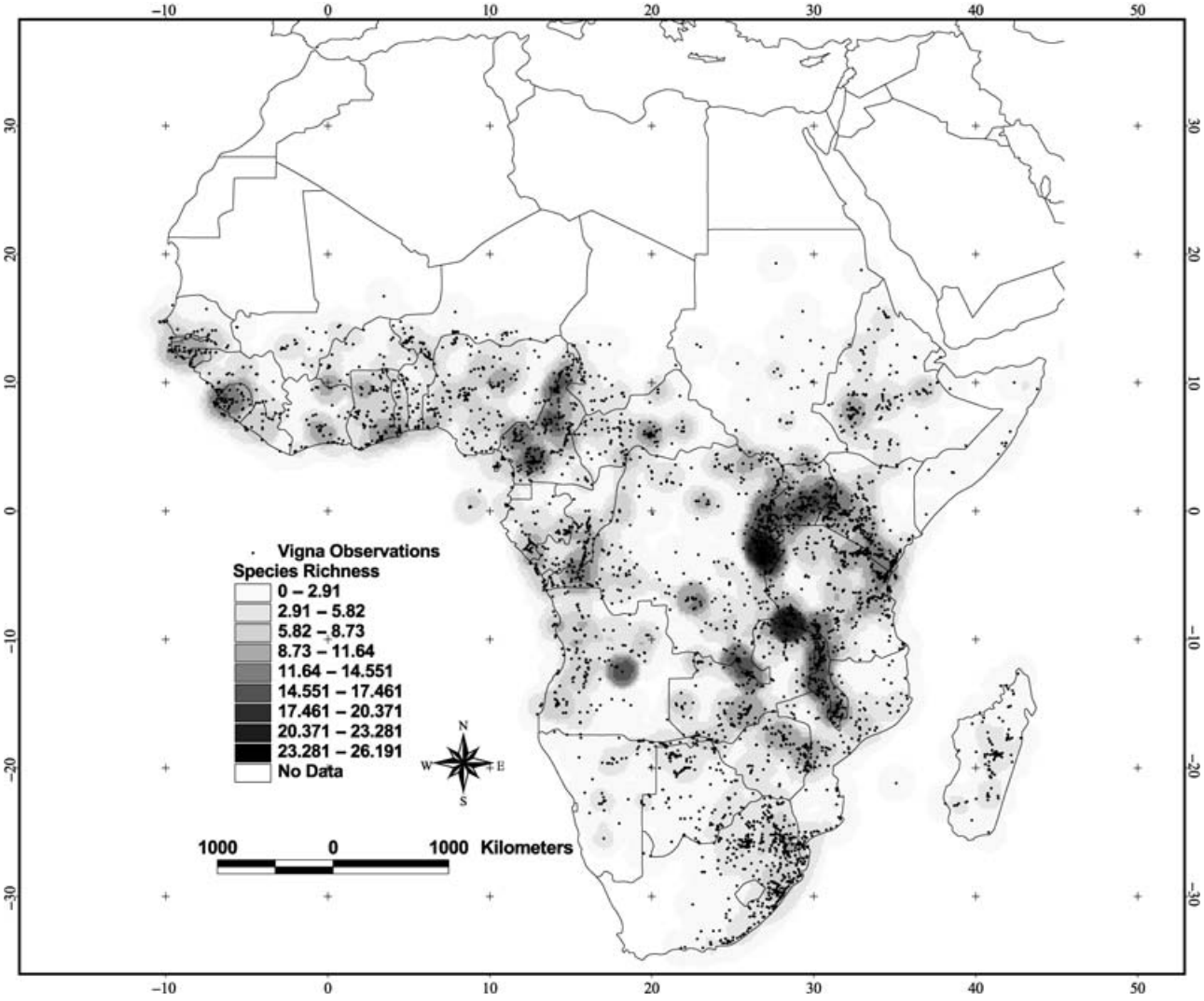

Figure 2 Species richness of *Vigna* taxa in 20 km × 20 km grid cells smoothed using inverse distance weighting and a window of 200 km radius (Maxted *et al*., 2005).

and other protected area networks will almost certainly contain many *Vigna* species. Comparison of the location data for herbarium specimens or gene bank accessions with existing African protected areas found that 54% of wild African *Vigna* species have populations predicted to be present in at least one protected area (Maxted *et al*., 2005). The real figure for population presence in protected areas is likely to be higher because the data set only refers to populations that have been sampled for herbarium specimens or germplasm within protected areas, so unsampled populations will not be included. This broader distribution of African *Vigna* species in protected areas was estimated using the species prediction distribution model (Fig. 3) and 47 of the 48 *Vigna* species studied were predicted to be found in existing protected areas (WDPA, 2003), though many species had small predicted ranges within a protected area.

On-farm conservation. African farmers have been conserving cultivated *Vigna* species for millennia via annual cycles of planting, cultivating, harvesting, selecting seed and planting. However, there are no on-farm projects in Africa that explicitly focus on *Vigna* species, although there are several projects that include *Vigna* species along with other native crops. Cowpea (*V. unguiculata*) is included in Bioversity International's on-farm conservation project in Burkina Faso (Jarvis & Ndungú-Skilton, 2000), the Shea project in Uganda includes Bambara groundnut (*Vigna subterranea*) (http://www.pnumen.com/covaol/onfarm.html) and the Community Technology Development Trust project in Zimbabwe, which is looking at the relationship between socio-economic factors and on-farm crop diversity of several crops, including *V. subterranea* and *V. unguiculata* (Odero, 2001 – http://www.cbdcprogram.org).

Ex situ *conservation assessment*

The largest germplasm collection of African *Vigna* is held in trust at the International Institute of Tropical Agriculture (IITA),

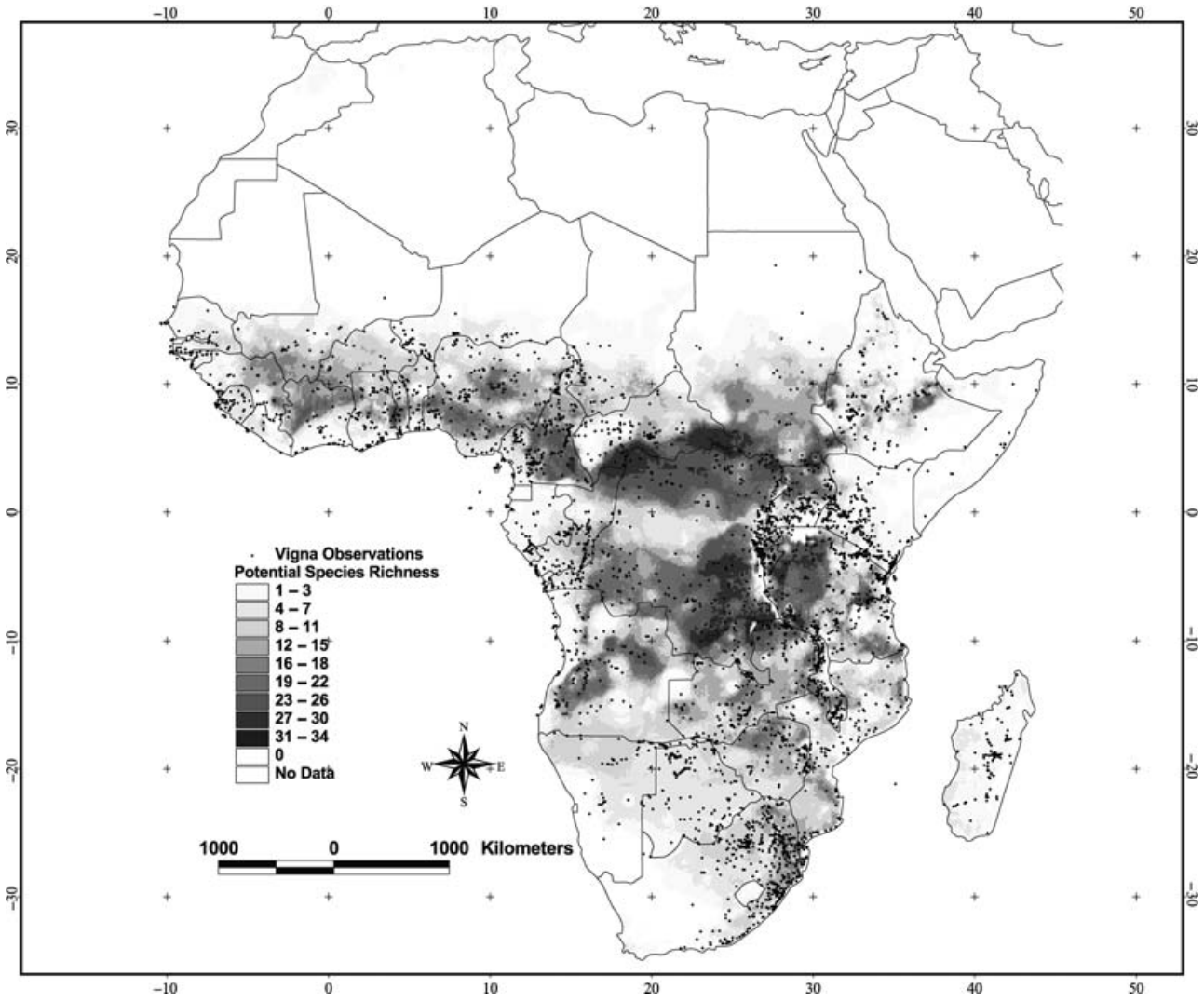

Figure 3 Predicted distribution of species richness based on climate envelope modelling for *Vigna* taxa (Maxted *et al*., 2005).

Ibadan, Nigeria, which holds the CGIAR world mandate for cowpea, Bambara groundnut and their wild relatives. It holds 18,686 accessions, of which 14,887 are cultivated *V. unguiculata* ssp. *unguiculata*, 553 are wild *V. unguiculata*, and 2032 are *V. subterranea*. Other gene banks that hold accessions of *Vigna* within the CGIAR system include International Livestock Research Institute and Centro Internacional de Agricultura Tropical, with 276 and 842 *Vigna* accessions, respectively (http://singer2.cgiar.org). The largest collection of *Vigna* outside the CGIAR system is held by the Phaseolinae gene bank at The National Botanical Gardens of Belgium (Jardin Botanique National de Belgique) with 507 accessions of primarily wild *Vigna* species. Other gene banks, such as the National PGR Laboratory in the Philippines, Asian Vegetable Research and Development Centre in Taiwan and the National Bureau of Plant Genetic Resources in India, also have large collections of cultivated material (FAO, 1998). Although there are large seed collections of cultivated and wild *Vigna* species, 25 wild African *Vigna* species have no *ex situ* conserved germplasm and the majority of the remaining wild species are represented by few collections in a single gene bank (Maxted *et al*., 2005). Figure 4 illustrates absolute species richness based on germplasm collections alone and provides an interesting comparison to Fig. 2 based on herbarium specimens only. Comparison of these two maps identifies geographical gaps; particularly near the *Vigna* hotspots identified close the great lakes regions, Democratic Republic of the Congo and Madagascar. These are clear regions of priority for increasing the efficiency of germplasm coverage. Maxted *et al*. (2005) conclude that the broad African *Vigna* gene pool is not adequately conserved in *ex situ* gene banks.

Herbarium and gene bank collection data were used to calculate the circular area statistic (CA) (Hijmans *et al*., 2001). The assumption is that herbarium collections are more complete and representative of species distribution than germplasm collections, which is the case in the *Vigna* gene pool. Figure 5 shows the relative circular area (RCA, no. complete circles, radius 50 km) for *Vigna* germplasm collections plotted against relative circular area for both germplasm and herbarium collections, performed

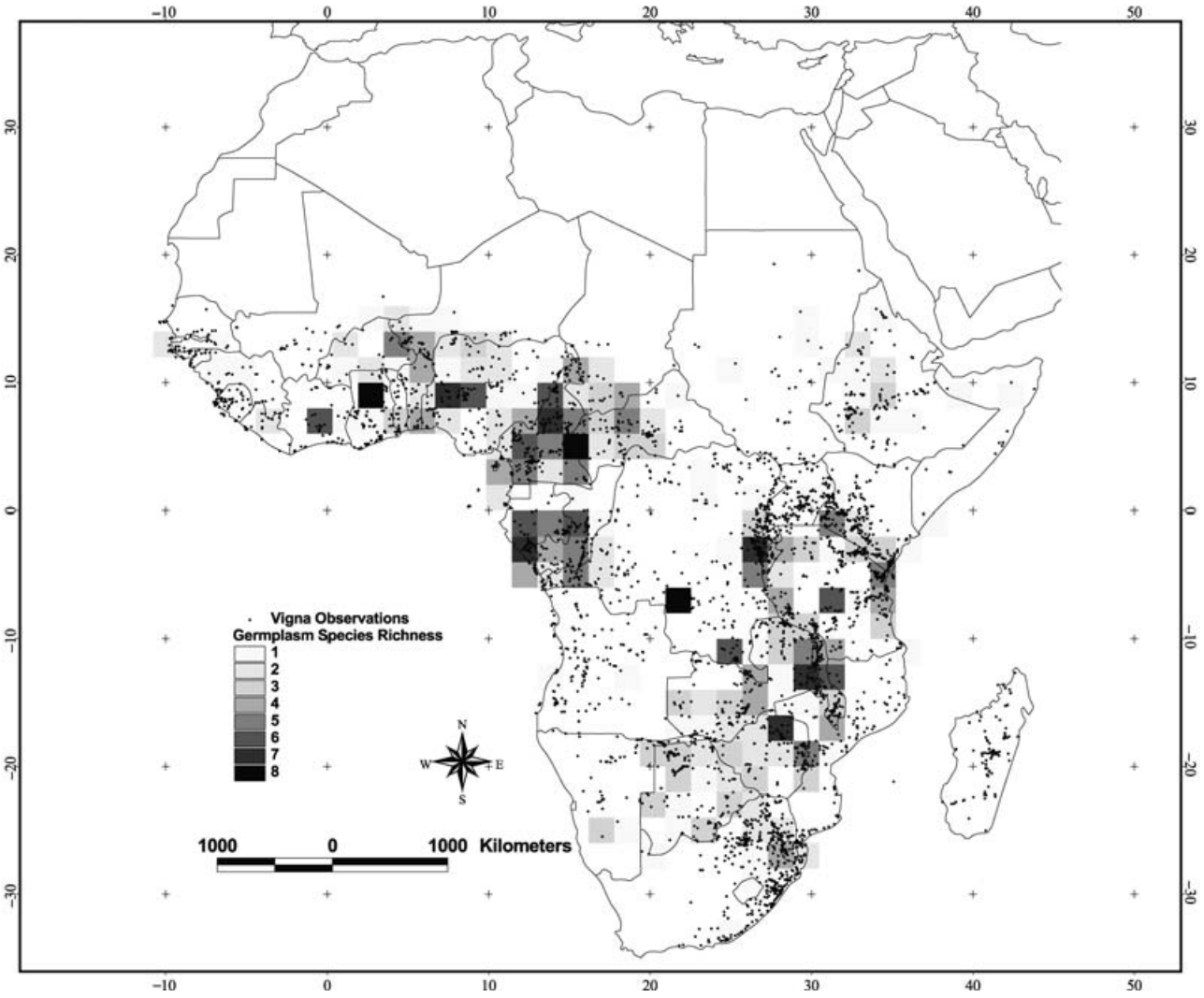

Figure 4 Absolute species richness based on *Vigna* germplasm collections alone in 200 km × 200 km grid cells (Maxted *et al.*, 2005).

for each individual species. *V. nigritia*, *V. radicans*, *V. reticulata* and *V. unguiculata* each has appreciably higher than average geographical coverage in the germplasm collection, while species such as *V. frutescens*, *V. luteola* and *V. vexillata* have appreciably lower geographical coverage in germplasm collections.

Step 4: reformulating conservation strategy

In situ *conservation priorities*

Genetic reserve/protected area – Actual and predicted species richness for African *Vigna* species is shown in Figs 2 and 3, respectively, and clearly indicates African *Vigna* hotspots. However, if resources are sufficient for the establishment of more than one genetic reserve, then species richness alone may not be an efficient means of selecting locations for genetic reserves. It would be preferable to select sites that contain complementary taxa rather than duplicating existing taxa already conserved *in situ* simply because they are found in sites with high species richness (Kirkpatrick & Harwood, 1983; Margules *et al.*, 1988; Pressey & Nicholls, 1989; Pressey *et al.*, 1993; Rebelo, 1994). Figure 6 illustrates the priority sites to establish *in situ* reserves for *Vigna* using complementarity analysis via DIVA-GIS (http://www.diva-gis.org). Based on the analysis, twenty-three 100 km^2 areas are required to capture all 61 species in the *Vigna* genus. However, three grid cells contain 37 species (54% of all species in the genus) and these are the southern tip of Lake Tanganyika (23 species), the coastal area of Sierra Leone (eight new species), and between Lake Victoria and the other Great Lakes (three new species) (Fig. 7). These locations would need to be surveyed to determine which is the most appropriate for establishing the genetic reserve. Following completion of reserve site surveying, target taxon hotspots can be identified and final recommendations made. For *Vigna* one genetic reserve within each priority area would be appropriate to conserve a significant proportion of *Vigna* genetic diversity.

On-farm conservation priorities – For African *Vigna*, 23 of the 61 species are utilized in Africa and many of these species

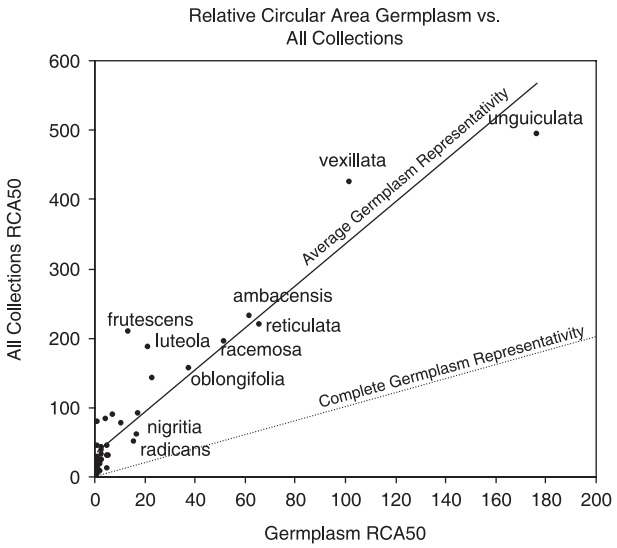

Figure 5 Relative circular area of germplasm collections of *Vigna* species compared with total relative circular area of all collections (germplasm and herbarium).

have multiple uses within subsistence agriculture, which means there is likely to be high levels of intrinsic landrace diversity. With the rapid industrialization of African agriculture to meet expanding food demand and the associated loss of intracrop diversity, there is a clear need to establish on-farm conservation projects for *Vigna* to ensure long-term food security. The most immediate priority is to better conserve the two most widely cultivated grain legume species, *V. subterranea* and *V. unguiculata*, and to establish on-farm projects in locations with the highest landrace diversity. However, there is currently no systematic documentation of the geographical and genetic patterns of landrace diversity in these two species to indicate where on-farm projects might most appropriately be established. Nor is there an accurate picture of *Vigna* crop landrace erosion or extinction. So before any on-farm conservation projects for *Vigna* are established, further research would be necessary.

Ex situ *conservation priorities*

Priorities for *ex situ* conservation action, based on the numbers of existing *Vigna* accessions held in gene banks, can be viewed in terms of taxa and countries.

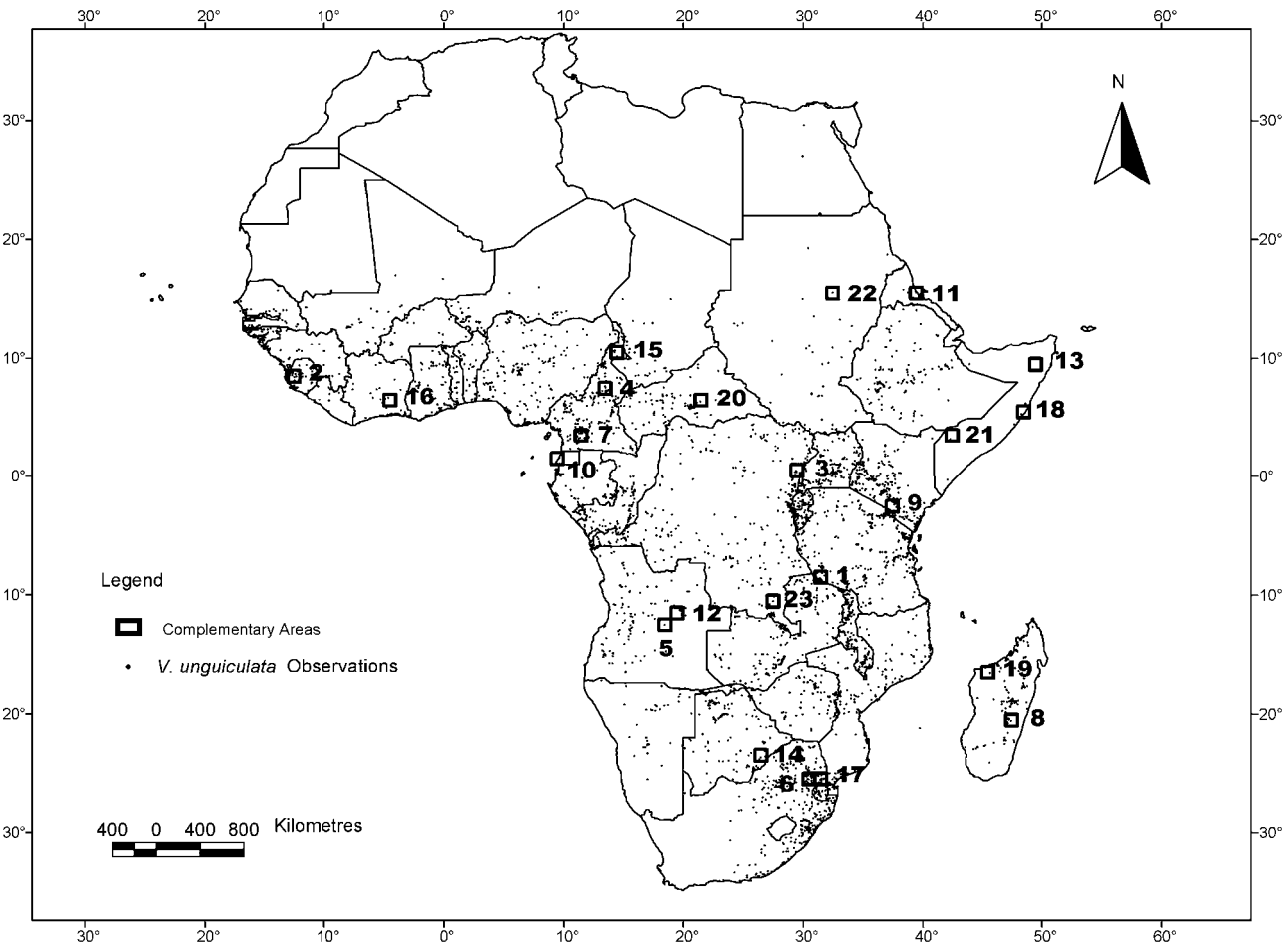

Figure 6 Complementarity analysis of grid cells in order of priority for establishment of *in situ* reserve conservation (Maxted *et al.*, 2005).

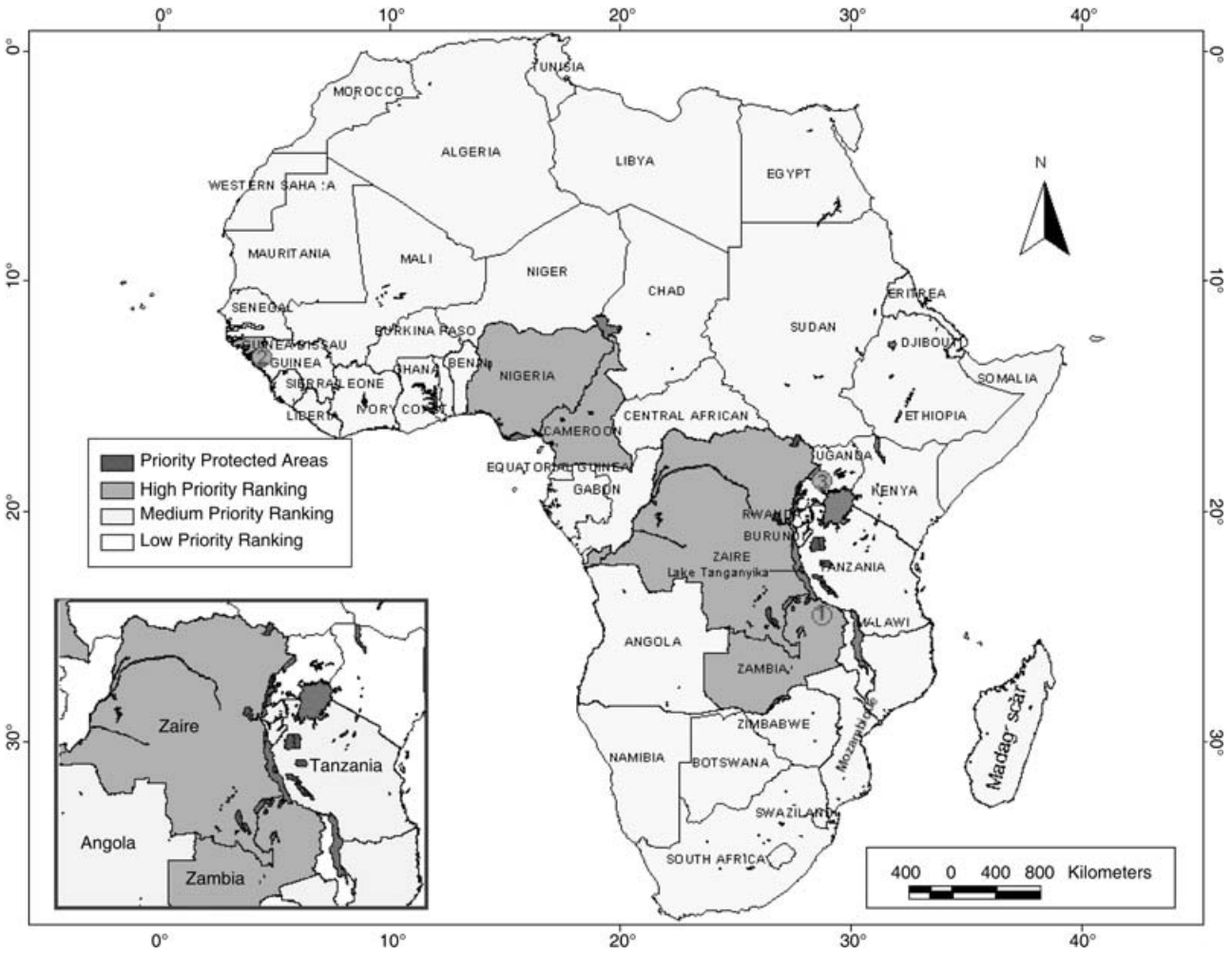

Figure 7 Three priority areas where *Vigna* conservation action is required in Africa (Maxted *et al.*, 2005).

• *Taxon-based priorities* – *Vigna* taxa prioritized for *ex situ* conservation include *V. dolomitica*, *V. haumaniana* var. *pedunculata*, *V. monantha*, *V. nuda*, *V. richardsiae*, *V. somaliensis*, *V. stenophylla*, *V. subterranea* var. *spontanea*, *V. unguiculata* ssp. *unguiculata* var. *spontanea*, *V. unguiculata* ssp. *aduensis*, *V. unguiculata* ssp. *baoulensis*, *V. unguiculata* ssp. *burundiensis*, *V. vexillata* var. *dolichonema* and *V. virescens*. Although the cultivated forms of *V. unguiculata* and *V. vexillata* appear well-conserved *ex situ*, collection of the wild taxa should continue in order to ensure full representation of all infraspecific taxa and ecological variants. Furthermore, in the light of failure of interspecific crosses involving *V. unguiculata*, it is of paramount importance that as wide a range of accessions of wild forms of the species as possible is collected and evaluated in order to identify material that may be useful for cowpea improvement.

• *Country-based priorities* – As defined by the number of *Vigna* taxa included and current *ex situ* representation, there is a need for further collecting and conservation of germplasm *ex situ* in the following countries: Angola, Benin, Burundi, Cameroon, Cote d'Ivoire, the Democratic Republic of the Congo, Djibouti, Eritrea, The Gambia, Guinea, Guinea Bissau, Liberia, Madagascar, Mozambique, Nigeria, Rwanda, Sierra Leone, Somalia, Tanzania and Zambia. Of these Cameroon, Democratic Republic of the Congo, Guinea Bissau, Nigeria and Zambia are the highest priorities.

CONCLUSIONS

We have demonstrated how the scope of an effective conservation tool, gap analysis, might be expanded to fully address the need for a more comprehensive and complementary conservation strategy that encompasses both *in situ* and *ex situ* applications. It has also shown how the study of herbarium and germplasm accessions' passport collection data coupled with ecogeographical analyses can quantify the completeness of current *in situ* and *ex situ* conservation actions and identify gaps in conservation diversity at both the taxon and the geographical level that in turn helps in the prioritization of future conservation actions. This basic methodology has been successfully applied to assist the development of national conservation strategies for crop wild relative diversity in the UK (Maxted *et al.*, 2007), Portugal (Magos Brehm *et al.*, 2007) and Ireland (H. Fitzgerald, S. Waldren & N. Maxted, unpublished data), but here the methodology was demonstrated

using the African genus *Vigna* showing that even for taxa where there is a less than complete genetic and ecogeographical knowledge, the methodology is sufficiently robust to yield useful results that can bolster conservation efficiency. The methodology is applicable to any facet of biodiversity, but especially useful in the context of conserving genetic diversity of useful species. If the conservation community is to meet the challenge of the CBD 2010 Biodiversity Target (CBD, 2002) and achieve by 2010 and beyond a significant reduction in the rate of biodiversity loss, then tools such as the gap analysis methodology described are likely to prove essential.

ACKNOWLEDGEMENTS

This study was supported by Bioversity International (formerly the International Board for Plant Genetic Resources), Centro Internacional de Agricultura Tropical (CIAT), the Association of Commonwealth Universities and University of Birmingham. We thank Lori De Hond for providing assistance in manuscript production.

REFERENCES

Balmford, A. (2003) Conservation planning in the real world: South Africa shows the way. *Trends in Ecology and Evolution*, **18**, 435–438.

Bettencourt, E., Konopka, J. & Damania, A.B. (1989) *Directory of crop germplasm collections. 1. I. Food legumes*. International Board for Plant Genetic Resources, Rome, Italy.

Bingham, M.G. & Smith, P.P. (2002) Zambia. *South African plant red data lists* (ed. by J.S. Golding), pp. 135–156. SABONET, Pretoria, South Africa.

Brooks, T.M., Bakarr, M.I., Boucher, T., Da Fonesca, G.A.B., Hilton-Taylor, C. & Hoekstra, J.M. (2004) Coverage provided by the global protected area system: is it enough? *Bioscience*, **54**, 1081–1091.

Burgman, A., Grimson, R.C. & Ferson, S. (1995) Inferring threat from scientific collections. *Conservation Biology*, **9**, 923–929.

Burley, F.W. (1988) Monitoring biological diversity for setting priorities in conservation. *Biodiversity* (ed. by E.O. Wilson and F.M. Peter), pp. 227–230. National Academy Press, Washington DC.

Convention on Biological Diversity (1992) *Convention on biological diversity: text and annexes*. Secretariat of the Convention on Biological Diversity, Montreal, Quebec, Canada.

Convention on Biological Diversity (2002) *Global strategy for plant conservation*. Secretariat of the Convention on Biological Diversity, Montreal, Quebec, Canada.

Coulibaly, S., Pasquet, R.S., Papa, R. & Gepts, P. (2002) AFLP analysis of the phenetic organisation and genetic diversity of *Vigna unguiculata* L. Walp reveals extensive gene flow between wild and domesticated types. *Theoretical and Applied Genetics*, **104**, 358–366.

Dietz, R.W. & Czech, B. (2005) Conservation deficits for the continental United States: an ecosystem gap analysis. *Conservation Biology*, **19**, 1478–1487.

Ehlers, J.D. & Hall, A.E. (1997) Cowpea (*Vigna unguiculata* L. Walp.). *Field Crops Research*, **53**, 187–204.

Engels, J.M.M. & Hawkes, J.G. (1991) The Ethiopian gene centre and its genetic diversity. *plant genetic resources of Ethiopia* (ed. by J.M.M. Engels, J.G. Hawkes and M. Worede), pp. 23–41. Cambridge University Press, Cambridge, UK.

Food and Agriculture Organisation of the United Nations (FAO) (1998) *The state of the world's plant genetic resources for food and agriculture*. FAO, Rome, Italy.

Food and Agriculture Organisation of the United Nations (FAOSTAT) (2007) *FAOSTAT database*. FAOSTAT, Rome, Italy.

Ferguson, M.E., Ford-Lloyd, B.V., Robertson, L.D., Maxted, N. & Newbury, H.J. (1998) Mapping the geographical distribution of genetic variation in the genus Lens for the enhanced conservation of plant genetic diversity. *Molecular Ecology*, **7**, 1743–1755.

Hijmans, R.J., Cruz, M., Rojas, E. & Guarino, L. (2001) DIVA-GIS Manual, Version 1.4. *A geographic information system for the management and analysis of genetic resources data*. International Potato Center (CIP), Lima, Peru.

Hijmans, R. & Spooner, D. (2001) Geographic distribution of wild potato species. *American Journal of Botany*, **88**, 2101–2112.

IUCN (1994) *IUCN red list categories*. IUCN Species Survival Commission. International Union for Conservation of Nature and Natural Resources, Gland, Switzerland.

IUCN (2001) *IUCN red list categories*, Version 3.1. IUCN Species Survival Commission. International Union for Conservation of Nature and Natural Resources, Gland, Switzerland.

Jarvis, D. & Ndungú-Skilton, J. (2000) IPGRI *in situ* project: research and institutions supporting local management of agrobiodiversity, *Encouraging diversity: the conservation and development of plant genetic resources*. (ed. by C. Almekinders and W. de Boef), pp. 134–141. Intermediate Technology Productions, London.

Jones, P.G., Beebe, S.E., Tohme, J. & Galwey, N.W. (1997) The use of geographical information systems in biodiversity exploration and conservation. *Biodiversity and Conservation*, **6**, 947–958.

Jones, P., Guarino, L. & Jarvis, A. (2002) Computer tools for spatial analysis of plant genetic resources data: 2. FloraMap. *Plant Genetic Resources Newsletter*, **130**, 1–6.

Kirkpatrick, J.B. & Harwood, C.E. (1983) Conservation of Tasmanian macrophyte wetland vegetation. *Proceedings of the Royal Society of Tasmania*, **117**, 5–20.

Laghetti, G., Pienaar, B.L., Padulosi, S. & Perrino, P. (1998) The ecogeographical distribution of *Vigna* Savi in Southern Africa and some areas of the Mediterranean basin. *Plant Genetic Resources Newsletter*, **115**, 6–12.

Lawrence, M.J., Marshall, D.F. & Davies, P. (1995) Genetics of genetic conservation. I. Sample size when collecting germplasm. *Euphytica*, **84**, 89–99.

Magos Brehm, J., Maxted, N., Ford-Lloyd, B.V. & Martins-Loução, M.A. (2007) National inventories of crop wild relatives and wild harvested plants: case-study for Portugal. *Genetic Resources and Crop Evolution*, doi: 10.1007/s10722-007-9283-9.

Mahalakshmi, V., Ng, Q., Lawson, M. & Ortiz, R. (2007) Cowpea [*Vigna unguiculata* (L.) Walp.] core collection defined by geographical, agronomical and botanical descriptors. *Plant Genetic Resource: Characterization and Utilization*, **5**, 113–119.

Maréchal, R., Mascherpa, J. & Stainier, F. (1978) Etude taxonomique d'un groupe complexe d'espèces des genres *Phaseolus* et *Vigna* (Papilionaceae) sur la base de donées morphologiques et polliniques, traitées par l'analyse informatique. *Boissiera*, **28**, 1–273.

Margules, C.R. (1989) Introduction to some Australian developments in conservation evaluation. *Biological Conservation*, **50**, 1–11.

Margules, C.R., Nicholls, A.O. & Pressey, R.L. (1988) Selecting networks of reserves to maximise biological diversity. *Biological Conservation*, **43**, 63–76.

Margules, C.R. & Pressey, R.L. (2000) Systematic conservation planning. *Nature*, **405**, 243–253.

Marshall, D.R. & Brown, H.D. (1975) Optimum sampling strategies in conservation. *Crop genetic resources for today and tomorrow* (ed. by O.H. Frankel and J.G. Hawkes), pp. 53–80. Cambridge University Press, Cambridge, UK.

Maxted, N., Ford-Lloyd, B.V. & Hawkes, J.G. (eds) (1997) *Plant genetic conservation: the* in situ *approach*. Chapman and Hall, London.

Maxted, N. & Guarino, L. (2003) Planning plant genetic conservation. *Seed conservation: turning science into practice* (ed. by R.D. Smith, J.B. Dickie, S.H. Linington, H.W. Pritchard and R.J. Probert), pp. 37–78. Royal Botanic Gardens, Kew, UK.

Maxted, N., Mabuza-Dlamini, P., Moss, H., Padulosi, S., Jarvis, A. & Guarino, L. (2005) *An ecogeographic survey: African* Vigna. Systematic and Ecogeographic Studies of Crop Genepools 10. International Plant Genetic Resources Institute, Rome, Italy.

Maxted, N., Scholten, M.A., Codd, R. & Ford-Lloyd, B.V. (2007) Creation and use of a national inventory of crop wild relatives. *Biological Conservation*, **140**, 142–159.

Maxted, N., Van Slageren, M.W. & Rihan, J. (1995) Ecogeographic surveys. *Collecting plant genetic diversity: technical guidelines* (ed. by L. Guarino, V. Ramanatha Rao and R. Reid), pp. 255–286. CAB International, Wallingford, UK.

Mithen, R. (1987) The African genepool of *Vigna*. I. *V. nervosa* and *V. unguiculata* from Zimbabwe. *Plant Genetic Resources Newsletter*, **70**, 13–19.

Nybom, H. (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology*, **13**, 1143–1155.

Odero, K.K. (2001) *Socioeconomic factors determining on-farm agricultural biodiversity in Zimbabwe*. International Union for Conservation of Nature and Natural Resources, Zimbabwe.

Pasquet, R.S. (2001) *Vigna Savi, Flora Zambesiaca*, **3**, 121–156. Royal Botanic Gardens, Kew, UK.

Pienaar, B.J. (1992) A taxonomic revision of the genus *Vigna* Savi (Fabaceae) in Southern Africa. MSc Thesis. University of Pretoria, Pretoria, South Africa.

Pressey, R.L., Humphries, C.R., Vane-Wright, R.I. & Williams, P.H. (1993) Beyond opportunism: key principles for systematic reserve selection. *Trends in Ecology and Evolution*, **8**, 124–128.

Pressey, R.L. & Nicholls, A.O. (1989) Efficiency in conservation evaluation: scoring versus iterative approaches. *Biological Conservation*, **50**, 199–218.

Rebelo, A.G. (1994) Iterative selection procedures: centres of endemism and optimal placement of reserves. *Botanical diversity in southern Africa* (ed. by B.J. Huntley), pp. 231–257. National Botanical Institute, Pretoria, South Africa.

Riemann, H. & Ezcurra, E. (2005) Plant endemism and natural protected areas in the peninsula of Baja California, Mexico. *Biological Conservation*, **122**, 141–150.

Solow, A.R. (1993) Inferring extinction from sighting data. *Ecology*, **74**, 962–964.

Spooner, D., van Treuren& de Vicente, M.C. (2005) *Molecular markers for genebank management*. *IPGRI technical bulletin no. 10*. International Plant Genetic Resources Institute, Rome, Italy.

Stuart, S.N. & Adams, R.J. (1990) *Biodiversity in Sub-Saharan Africa and its Islands: conservation, management and sustainable use*. International Union for Conservation of Nature and Natural Resources, Gland, Switzerland.

Stuart, C. & Stuart, T. (1995) *Africa: a natural history*. Swan Hill Press, Shrewsbury.

Tomooka, N., Vaughan, D.A., Moss, H. & Maxted, N. (2002) *The Asian Vigna: genus Vigna subgenus Ceratotropis genetic resources*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

WDPA (2003) *World database on protected areas*. http://www.unep-wcmc.org/wdpa.

Editor: Andrew Lowe

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Major data used in IUCN Red List Category assessment (Maxted *et al.*, 2005).

Appendix S2 Taxon vulnerability assessment scores for *Vigna* species (Maxted *et al.*, 2005).

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.