Chapter 6: Strategies for the collecting of wild species

Mauricio Parra-Quijano
Departamento de Biología Vegetal
Universidad Politécnica de Madrid, Ciudad Universitaria s/n
28040 Madrid, Spain
E-mail: mauricio.parra@agrobiodiversidad.org

José María Iriondo
Área de Biodiversidad y Conservación, Departamento de Biología y Geología
Universidad Rey Juan Carlos, Tulipán s/n
28933 Móstoles, Madrid, Spain
E-mail: jose.iriondo@urjc.es

Elena Torres Lamas
Departamento de Biología Vegetal
Universidad Politécnica de Madrid, Ciudad Universitaria s/n
28040 Madrid, Spain
E-mail: elena.torres@upm.es

Abstract

Crop wild relatives (CWR) have received increasing interest from the plant genetic resources community since the first Global Plan of Action in 1996 called attention on its poor representation in germplasm collections around the world. Collecting has become a strategic activity to improve the *ex situ* representation of wild germplasm, but improvements are now more focused on qualitative issues (i.e., introducing novel genetic diversity, useful traits, environmental adaptations) than being simply quantitative (i.e., aimed only at the number of accessions). Since the publication of these Technical Guidelines in 1995, there have been some significant technical and methodological advances in novel CWR collecting and evaluation of past collecting efforts.

In the last decade, several tools and methodologies that have been successfully tested in related disciplines (such as plant biology, ecology and biodiversity conservation) have been gradually introduced and/or popularized in plant genetic resources, such as the use of geographic information systems, ecogeographical analysis or species distribution models. Thus, current scientific knowledge and technical capabilities have allowed collecting missions for CWR following research or breeding interests to be carried out with “surgical” precision. Collecting for genetic diversity and/or conservation is the area where the new advances have been used more intensively, allowing for the design of effective and efficient collecting missions, the detection of collection biases in existing germplasm collections, the assessment of their representativeness and, therefore, the need for additional germplasm collecting.

The original version of this chapter (authored by von Bothmer and Seberg) provides an adequate framework for the integration and discussion of these new advances, and the identification of future

This chapter is a synthesis of new knowledge, procedures, best practices and references for collecting plant diversity since the publication of the 1995 volume *Collecting Plant Diversity: Technical Guidelines*, edited by Luigi Guarino, V. Ramanatha Rao and Robert Reid, and published by CAB International on behalf of the International Plant Genetic Resources Institute (IPGRI) (now Bioversity International), the Food and Agriculture Organization of the United Nations (FAO), the World Conservation Union (IUCN) and the United Nations Environment Programme (UNEP). The original text for Chapter 6: Strategies for the Collecting of Wild Species, authored by R. von Bothmer and O. Seberg, has been made available online courtesy of CABI. The 2011 update of the Technical Guidelines, edited by L. Guarino, V. Ramanatha Rao and E. Goldberg, has been made available courtesy of Bioversity International.
CHAPTER 6: STRATEGIES FOR THE COLLECTING OF WILD SPECIES

challenges. This update presents the most significant contributions on CWR collecting that have taken place in the last 15 years. However, instead of keeping the focus on the tribe Triticeae of the Poaceae family as in the original chapter, it provides examples of a broad taxonomic range.

Introduction

Wild plant species, in a broad sense, are a relatively modern target in the history of plant genetic resources for food and agriculture (PGRFA). However, for those interested in the conservation and use of agricultural biodiversity, only some of the wild species that are genetically related to domesticated or cultivated species deserve major attention. They are referred as crop wild relatives (CWR) and include progenitors and other species belonging to the primary, secondary and tertiary genepools of crop species (Maxted et al. 2008b).

CWR are a valuable source of traits for resistance to and tolerance of biotic and abiotic stresses – traits that contributed to increases in yield and quality of about USD$350 million per year in the USA alone in the mid-1980s (Prescott-Allen and Prescott-Allen 1986). It can be assumed that this value has further increased in the last decades as we headed towards the current scenarios of global change. CWR became a matter of concern at a global scale when, at the end of the 1990s, the scientific community circulated several warnings about the risk of extinction of many CWR species and the consequent loss of useful genetic material. FAO’s Global Plan of Action (1996) provided one of the most important calls for attention to the low representation of CWR in ex situ germplasm collections and the need to collect them. Consequently, many organizations interested in PGRFA conservation have focused their efforts on improving the representation of CWR species in genebanks.

As the Second Report on the State of the World’s Plant Genetic Resources for Food and Agriculture (SoWPGR-2) notes, 240,000 new accessions were collected over the period 1996–2007, and approximately 40% of them were CWR (FAO 2010). This means that CWR were the predominant type of germplasm collected over the last 15 years. At the same time, SoWPGR-2 underlined the fact that most CWR germplasm was collected by crop genebanks as a matter of local or national interest. Some botanical gardens also operate seed banks and, therefore, collect seeds of wild plants (not only CWR), but there is no global data (on number of species and accessions) available about their conservation efforts. In addition, botanical gardens focus their attention mainly on threatened, rare or endemic wild plants (Bacchetta et al. 2008) and tend to represent species richness more than intraspecific genetic variation (Maunder et al. 2004).

Despite the increasing importance of CWR, information on CWR collecting is scarce and poorly represented in the scientific media. Scientific reports about collecting missions of plant genetic resources (PGR) are often undervalued and published just as “short communications” in most cases. Furthermore, most local/national initiatives involved in the collecting of CWR germplasm do not produce easily accessible outputs, and the results remain as grey literature or internal reports. As a result, CWR collecting expeditions are difficult to follow and quantify.

Fortunately, unlike the reports of collecting missions, some relevant collecting methodologies for PGR have been published in recent years, many of them focused on CWR. New, available methodologies are the result of previous joint developments, including recent advances in geographical information system (GIS) tools and data accessibility and sharing (through the internet). Progress in this area comes from individual and collaborative efforts involved in projects on CWR conservation (Heywood et al. 2007).

Most new advances on CWR collecting have considered representativeness as a main goal. In genetic terms, representativeness in ex situ conservation refers to the attempt to capture the genetic variability of the species to the greatest extent possible. However, until a complete view of the existing genetic variability across species is available, it is hardly possible to determine gaps in genetic representativeness (GR). The assessment of the genetic variability of all genebank samples of a particular species might be affordable but it would be difficult to achieve this for all the populations of the species. For this reason, most advances have focused on assessing ecogeographical representativeness (ER) as a surrogate of GR. Because there is a
link between genotype or phenotype and environmental conditions (because of the selective pressures and local adaptation processes that affect wild populations) the use of ecogeographical data as a surrogate of genetic data can be justified. These relationships have been addressed by several authors, such as Steiner (1999), and supported by a constant number of publications over time (see, for example, the review for *Triticum* wild relative species by Nevo [2011]). In addition to this, the advantages of using ER are linked to the low information requirements about the target species’ populations. Georeferenced collecting sites are the baseline of any ecogeographical study in PGR, and geographic coordinates are the most appropriate data to define unequivocally any collecting site (Hijmans et al. 1999). When collecting is carried out following ER parameters, the principle is simple: sampling CWR of distinct ecogeographical areas should ensure sampling of representative intraspecific genetic variation (Greene and Hart 1999; Jones et al. 1997; Parra-Quijano et al. 2008; Steiner and Greene 1996).

Collecting CWR following representativeness criteria is obviously carried out within the framework of “collecting for genetic diversity study and conservation”, as established in the original chapter in 1995. For other collecting alternatives (i.e., collecting for specific use in breeding programmes or for taxonomic/phylogenetic/biosystematics research), the most significant innovations also come from similar methodologies and tools (GIS, ecogeographical approach or species distribution models).

**Current status**

In this section, the most remarkable advances in CWR collecting methods published in the last 15 years are briefly described, following the structure of the original 1995 chapter.

**Collecting for taxonomic, phylogenetic and biosystematics research**

There are few publications on expeditions for CWR collecting following these interests, although some remarkable examples can be cited, such as the continued efforts on collecting wild relatives of *Solanum tuberosum* by Spooner et al. (1992, 1994, 2000). These articles describe exploration criteria, routes, taxonomic and/or phylogenetic interests and collecting results in Ecuador, Bolivia and Mexico in detail.

On the other hand, van Treuren et al. (2011) nicely showed how to use passport data from PGR documentation systems and other sources to determine collection gaps in crop and wild leafy vegetable species based on the genepool structure of the target crop species.

Collecting for taxonomic research frequently demands drawing the distribution of known species and assessing species richness in the territory of study. In this sense, Hijmans and Spooner (2001) established the geographic distribution of about 199 wild potato species and determined the areas with the highest species richness. This work was supported by GIS tools. Similarly, Sawkins et al. (1999) analysed the case of some *Stylosanthes* species, where, in addition to using GIS tools, they introduced a predictive distribution model (based on Principal Component Analysis) to estimate the potential distributions of these taxa.

In other cases, the objective (besides mapping species distribution) has been to identify environmental adaptations and/or climatic ranges for sets of CWR species. Segura et al. (2003) linked the distributions of five *Passiflora* species to several temperature and rainfall variables through GIS tools and species distribution models. With this information about species adaptation, they planned *ex situ* and *in situ* conservation strategies. Similarly, Ferguson et al. (2005) used GIS tools, species distribution models and climatic variables to study the biogeography of 69 wild *Arachis* species. These kinds of studies are also helpful for meeting the objectives of the types of collecting interests that are presented next.

**Collecting for genetic diversity study and conservation**

“Capture of genetic diversity” commonly refers to between- and within-population sampling procedures within a species. Molecular methods for characterizing genetic diversity developed enormously since 1995, and techniques such as ISSRs, microsatellites or AFLPs, which were not mentioned in the 1995 version of this chapter, have been extensively applied (Spooner et al. 2005). Nevertheless, at this level, collecting strategies in wild species is still a matter of determining the size of the sample per site.
(population/subpopulation) and the number of sites that need to be sampled to ensure appropriate allelic retention (Brown and Briggs 1991). Sampling theory on collecting PGR was extensively detailed by Brown and Marshall (1995) in chapter 5 of "Collecting Plant Genetic Diversity: Technical Guidelines", as well as by von Bothmer and Seberg in the 1995 version of this chapter. Since then, there have been few new developments. For instance, optimum sample size has been documented for some species, e.g., wild populations of Phaseolus lunatus (Zoro Bi et al. 1998) or Zizania texana (Richards et al. 2007), while other studies have explored different models to determine minimum sample size under specific conditions (e.g., ploidy level, matting system or multiple loci) (Sapra et al. 1998; Lockwood et al. 2007).

In spite of the extraordinary advances of molecular methods for characterizing genetic diversity, when it comes to planning the collection of seeds of target CWR for PGR conservation, the norm is that there are no previous genetic diversity data published on the target species or that published data are inadequate. Therefore, in most cases, it is not possible to rely on such data to design the sampling strategy. Even if the genetic diversity of the species has been previously studied, it should be noted that neutral genetic variation, as characterized using current widespread molecular techniques, does not necessarily correlate with adaptive genetic variation, which is the type of genetic variation that is of greatest interest for breeding purposes. In this regard, the studies performed by Storfer (1996), Pfrender et al. (2000) and Bekessy et al. (2003), among others, show that currently used molecular markers are not associated with genes of adaptive value and that the correlation between the indicators of diversity obtained in the two types of genes is usually low.

Concerning the use of ecogeographical data as a proxy for data on genetic diversity, in chapter 15 of the 1995 edition of "Collecting Plant Genetic Diversity: Technical Guidelines", Guarino (1995b) mentioned the potential of GIS tools in PGR collecting activities for the first time, and Greene et al. (1999a, 1999b) described the first case of the use of GIS tools in sampling intraspecific variation in CWR species of Trifolium, Lotus and Medicago. Afonin and Greene (1999) published a detailed description of the use of GIS in the design of CWR collecting. They focused their attention on the possible uses of GIS to (1) determine where to search through the use of predictive models, (2) determine where to search to collect germplasm adapted to abiotic stresses, (3) define a sampling framework and (4) assess the anthropogenic influence on target sampling areas.

Another important advance, also based on GIS techniques, was presented by Hijmans et al. (2000), concerning the identification of collecting biases in wild potato samples from Bolivia conserved in the world’s six main wild potato genebanks. Four types of bias were described: species, species-area, hotspot and infrastructure bias. Knowledge about the type and magnitude of collecting biases is relevant information that should always be taken into account when improving the representativeness of any genebank and designing new collecting expeditions.

Gap analysis is a comparative method that can be used to improve biodiversity conservation (Maxted et al. 2008a) and, specifically, to improve the efficiency in CWR collecting activities. Recently, Ramírez-Villegas et al. (2010) described the use of gap analysis to prioritize geographic areas for collecting wild Phaseolus species. This methodology has been used to determine “potential” gaps in several CWR species (see http://gisweb.ciat.cgiar.org/GapAnalysis). While potential gaps come from comparisons between collecting sites of germplasm that is already conserved (genebank data) and species distribution models, a “classic” gap analysis compares genebank collecting sites with data on species occurrence from other sources, such as herbaria, botanical databases and other types of presence data (see Parra-Quijano et al. 2008). Another example of the use of species distribution models in collecting CWR can be found in the case of Vasconcellea in Venezuela (Rodriguez et al. 2005).

The basis of ER studies in CWR ex situ germplasm collections is described by Parra-Quijano et al. (2008). The ER studies use gap-analysis techniques and ecogeographical characterization data to determine which areas should be explored to find and collect populations that grow in underrepresented environments. Ecogeographical characterization of the collection sites (Steiner and Greene 1996) or the territory where the target taxon occurs (Parra-Quijano et al. 2011a) is essential to ER analysis in genebanks. Ecogeographical variables are the first element to be considered. There are currently hundreds of ecogeographical variables available (in GIS layer format) for most countries, but only some of them might have a significant influence
on the abiotic adaptation of the target species. Thus, Bennett and Bullita (2003) determined the most influential ecogeographical variables in six species of *Trifolium* collected in Sardinia.

Based on these premises, an optimized PGR collecting method has been described by Parra-Quijano et al. (2011b). This method was developed and validated in six *Lupinus* species (five CWR and one cultivated) in Spain. Optimized collection attempts to make collecting activities and field exploration as efficient as possible, combining spatial and ecogeographical gap analysis, species distribution models and GIS tools. Thus, optimized collecting designs will guide explorers in finding and collecting populations in underrepresented environments (from spatial and ecogeographical points of view). At this point, two possible cases of optimized collections that follow representativeness criteria can be identified: (1) collecting a CWR species that is already represented in the genebank and (2) collecting a new species that has no previous accessions in the genebank. The *Lupinus* example mentioned above belongs to the first case, since the objective there was to improve the ER of the Spanish *Lupinus* collection held by the Plant Genetic Resources Centre of the National Institute of Agriculture and Food Research (CRF-INIA). Based on similar ER principles, Ghamkhar et al. (2007) determined ecogeographical regions and subregions within the Mediterranean and Irano-Turanian ecoregions related to *Trifolium spumosum* adaptations and then identified areas to explore to fill gaps in the collection of the Australian Trifolium Genetic Resources Centre (ATGRC).

Some CWR species may stand out because of their rarity or threat degree, and seed collecting may be called for as part of the effort to conserve the species. This is the case of *Capsicum flexuosum*, a rare and endangered species that has low representation in *Capsicum* genebanks. In this context, Jarvis et al. (2005) developed a successful collecting strategy for sampling *C. flexuosum* in some areas of Paraguay, based on GIS and species distribution models.

**Collecting for immediate use in a breeding programme**

Reports for this type of collecting are scarce and difficult to find. This is in part due to the fact that breeding programmes are often conducted by private companies, which are zealous about not disclosing their activities and sources of germplasm. However, it is important to note that collecting expeditions that try to capture genetic diversity are frequently biased towards breeders’ interests since the ultimate goal of PGR conservation is the utilization of its resources.

The most novel strategy for collecting CWR for immediate use in a breeding programme is emerging from an extension of the focused identification germplasm strategy (FIGS) that originated as a way of selecting subsets of germplasm from genetic resource collections in such a way as to maximize the likelihood of capturing a specific trait at a higher frequency than if the subset were selected at random. This approach was initially presented as a core collection sampling strategy (Mackay 1986, 1990), but later, it was further developed (Mackay and Street 2004) to meet the needs of a plant breeder approaching a genebank for sources of genetic diversity for crop breeding programmes aimed at a specific trait. If we look at CWR populations in their natural habitats, the FIGS approach can be used to select the populations where seeds should be collected in order to maximize the likelihood of capturing the desired trait. In essence, this method can guide field explorations to populations exposed to specific environmental stresses of interest to breeders. This methodology is currently being tested in the context of the PGR Secure project (www.pgrsecure.org).

**Future challenges/needs/gaps**

Future predictable advances and their consequences in regard to CWR collecting activities can be listed as follows:

1. Presence data on CWR species will become more abundant, detailed and easily available, particularly data from hotspot areas and centres of crop origins and diversity. They will be useful for performing more accurate CWR gap analyses and for implementing optimized collecting activities more efficiently.
CHAPTER 6: STRATEGIES FOR THE COLLECTING OF WILD SPECIES

2. Ecogeographical data in GIS layer format will improve in quality and resolution, and will be more easily available. This will allow ecogeographical representativeness studies to be performed – even at the within-population level – in the coming years.

3. Optimized methods of spatial interpolation of genetic diversity data will allow this information to be used as conventional GIS layers and will help improve the collection of genetic diversity in the field.

4. The development and popularization of non-neutral genetic markers linked to adaptive traits will enable a direct measure of the type of genetic variation that is sought for conservation in genebanks and use in breeding programmes. This will allow a much better assessment of the genetic diversity already stored in genebanks and will facilitate the design of more-efficient seed collecting strategies. It will also help calibrate the link with ecogeographical variables.

**Conclusion**

A significant change since the Technical Guidelines were published in 1995 is that CWR species conservation is now one of the highest priorities in plant genetic resources. Fifteen years ago, CWR was barely a rising topic. Several reasons for this huge change can be argued: the need for novel genes, the need of crop adaptation to extreme conditions, the breakdown of barriers to introgression, the increasing pressure on wild species populations brought about by global changes, and a growing awareness of the false sense of getting the job done with cultivated species. Although the effective conservation of CWR is the goal, lessons from cultivated plants have led to a focus on how to collect the most representative and useful germplasm of wild species.

Wild plant species are a target shared with other disciplines, such as biological conservation or plant ecology, so some of their tested tools have been gradually transferred to make possible the development of efficient methodologies for collecting CWR. As a result, a new range of methodological approaches based on ecogeographical data and plant adaptation is now available. GIS tools have provided a solid vehicle for facilitating the implementation and popularization of these methodologies. Their potential role in collecting plant genetic resources was highlighted by Guarino (1995a) in chapter 16 of these Technical Guidelines. Although genotypic and phenotypic approaches have also undergone significant changes (i.e., development of genomics, proteomics, new molecular markers, etc.), their effects have not been highly influential in the development of new strategies for collecting wild species.

**References**


CHAPTER 6: STRATEGIES FOR THE COLLECTING OF WILD SPECIES


Internet resources*

AEGRO website: http://aegro.jki.bund.de/aegro

Bioversity Crop Wild Relatives Global Portal: www.cropwildrelatives.org


European Crop Wild Relative Diversity Assessment and Conservation Forum: http://pgrforum.org

FAO Facilitating Mechanism for the implementation of the Global Plan of Action, Portal for Plant Genetic Resources for Food and Agriculture: www.globalplanofaction.org


GapAnalysis: http://gisweb.ciat.cgiar.org/GapAnalysis

Genetic Reserve Information System (GenResIS): www.agrobiodiversidad.org/aegro

Global Biodiversity Information Facility (GBIF): www.gbif.org

PGR Secure: www.pgrsecure.org

*Note that additional internet resources related to this chapter can be found in the update of chapter 15/16.