B.14. Information system and data management

B.14.1. Overview

Why is data critical to landrace conservation and use?

It is widely accepted within the PGRFA conservation and user community that one major factor hindering effective conservation and use of PGRFA is the lack of easy access to data, as well as obstacles to information exchange due to the many different approaches in managing data. If we are to inventory and build a National management plan for LR conservation, then consistent data collation and management is required.

Historically as noted above there have been many obstacles to information exchange between projects involved in the inventory and establishment of national management plan for LR conservation, the few projects that have addressed these tasks have developed stand-alone information systems to manage their LR related data. However in recent years the adoption of data collection and information management standards has been achieved to a large degree for the management of ex situ collections data using standard data descriptors such as the FAO/IPGRI Multi-crop Passport Descriptors (MCPD) version 2 published in June 2012 (http://www.bioversityinternational.org/index.php?id=19&user_bioversitypublications_pii[showUid]=6901). But even these standards do not adequately cater for the full range of data types that are of relevance to landrace conservation and use.

The EC funded PGR Secure project (see http://www.pgrsecure.org/) has as its dual goal agrobiodiversity conservation and the promotion of its sustainable use. One element of which is to develop (i) Europe-wide LR inventory, (ii) Exemplar national LR inventories and (iii) European LR conservation and use strategy. Each of these three deliverables requires extensive data management and intra- and inter-project data exchange. Thus significant progress was required and a set of minimum descriptors for the documentation of on-farm conservation and management activities have been developed, Descriptors for web-enabled national in situ landrace inventories (see http://www.pgrsecure.bham.ac.uk/sites/default/files/documents/helpdesk/LRDESCRIPTORS_PGRSECURE.pdf)\(^\text{195}\). The published descriptor list includes fields related to the inventory identification, taxon identification, landrace/population identification, site and location identification, landrace characteristics and finally fields concerning conservation and monitoring actions to be taken in favour of the

\(^{195}\) Negri et al. (2012)
landrace diversity maintenance. These descriptors have been designed to record the landrace(s) present on-farm, as well as to describe aspects of farm management practices (e.g., agricultural system, cropping management and farm labour division by gender). Descriptors to describe the seed supply system, the farmer’s criteria for distinguishing landraces, selection criteria, seed storage practices and crop uses, amongst others, are included. PGR Secure will within the context of the national LR inventories that are planned in Finland, Italy and the United Kingdom will test and refine the descriptors, but the methodology used for data collation and the descriptors are deliberately generic so that they will have applicability globally.

B.14.2. Methodology
Information on landraces is available from wide range of sources, but retrieving it presents a number of challenges. Firstly, in existing databases, such as those managed by plant gene banks, landrace accessions are generally not distinguished from modern varieties, although this issue should not arise if the FAO/IPGRI Multi-crop Passport Descriptors are used as the SAMPSTAT descriptor allows for the distinction between LR and other types of collection sample. Secondly, different scientists use different definitions of LR, so what is a LR to one is not to another. Thirdly, the crop variety name can sometimes be used to guide decisions as to whether a variety is a LR (for example, if the name of a LR is directly associated with a particular geographic location), but this is not a reliable method because modern varieties can also be given similar names. Furthermore, obtaining information about varieties that people grow for business purposes can be hindered by issues of commercial sensitivity, concerns about the potential legal repercussions associated with national listing of unregistered varieties and insufficient time and resources available to the business to respond. These challenges are not insurmountable but they do demand a carefully considered and tested approach (particularly with regard to obtaining information from commercial enterprises) and a considerable amount of time.

LR Data were collated from various sources, including LR, maintainers, PGR experts, governmental documents, NGOs, commercial companies, gene banks, websites and the literature. The types of data collated will fall into four basic types:

• Ecogeographic data (taxonomic, ecological, geographic and genetic: passport),
• Field population data (passport),
• Conservation management data (curatorial),
• Characterization and evaluation data (descriptive).
Each of these data types are collated using some type of standard descriptor. A descriptor may be defined as “any attribute referring to a population, accession or taxon which the conservationist uses for the purpose of describing, conserving and using this material”. Descriptors are abstract in a general sense, and it is the descriptor states that conservationists actually record and utilise. Standard descriptors for ecogeographic, field and conservation management data are included in the Descriptors for web-enabled national in situ landrace inventories\textsuperscript{196}, while formal characterization and evaluation descriptors are associated with various standardized ‘Crop descriptor lists’ published by FAO, Bioversity, UPOV (see http://www.bioversityinternational.org/publications.html). It is important to stress that standard lists of descriptors should be used when they are available. The use of well-defined, tested and rigorously implemented descriptor lists for scoring descriptors considerably simplifies all operations concerned with data recording, such as updating and modifying data, information retrieval, exchange, data analysis and transformation. When data are recorded, they should be classified and interpreted with a pre-defined list of descriptors and descriptor states to consult. This clearly saves a considerable amount of time and effort associated with data entry. The use of lists ensures uniformity, while reducing errors and problems associated with text synonyms.

B.14.3. Examples and applied use

There are few examples of data management within the context of the production of a National management plan for LR conservation. However, one reported example is the Vegetable landrace inventory of England and Wales will be made available via the UK’s Information Portal on Genetic Resources for Food and Agriculture (http://grfa.org.uk/\textsuperscript{197}. The methodology applied\textsuperscript{198} involved:

\textit{Experts’ meeting}

An experts’ meeting was called involving all those stakeholders with knowledge or interest in LR conservation and use to discuss the general project strategy and to share existing knowledge of how to obtain information on UK vegetable landraces, how to make contact with landrace maintainers, and a possible strategy for obtaining germplasm samples for ex situ conservation. The specific objectives of the meeting were to:

1. Provide an introduction to the project and discuss the proposed project strategy, including the following specific objectives:

\textsuperscript{196} Negri et al. (2012)
\textsuperscript{197} Kell et al. (2009)
\textsuperscript{198} Maxted et al. (2009)
a. Review official government documentation and scientific/popular literature
b. Review NGO and commercial company knowledge and holdings of landrace diversity
c. Review ex situ seed bank holdings of landraces
d. Discuss LR diversity with LR maintainers.

2. Share knowledge of how to achieve each of the above objectives (e.g., specific contacts, literature sources, government documents, relevant NGOs, commercial companies and seed banks).

3. Discuss a procedure for obtaining germplasm samples for ex situ conservation and outline a strategy for ensuring sufficient material is duplicated in the appropriate seed banks.

4. Provide examples of existing successful on-farm vegetable LR conservation projects in the UK (or elsewhere) that can be used for reference purposes when formulating conservation recommendations for other vegetable LR.

5. Provide examples of the use of LR germplasm in formal crop improvement programmes that can be used for reference purposes in the final report to Defra.