Contribution of standards for developing networks, crop ontologies and a global portal to provide access to plant genetic resources

Adriana ALERCIA[1] and Michael MACKAY[1]

1. Bioversity International, Via dei Tre Denari 472/a - 00057 Maccarese (Fiumicino), Rome, Italy

Abstract

The world is continually challenged to increase crop productivity and identify new varieties better-adapted to changing climatic conditions in order to meet the needs of local communities and contribute to global food security efforts. To effectively respond to these challenges, farmers and breeders must have access to a wide range of plant genetic resources (PGR), as well as their associated characterization and evaluation data.

The value of genetic resources is dependent upon the information available to adequately and appropriately use such germplasm. Accurate characterization and evaluation data promote the use of PGR, especially if available in a standardized format. Through the production of crop standards, or 'descriptors', and in collaboration with National Agricultural Research Systems, Consultative Group on International Agricultural Research (CGIAR) Centres and crop research institutes and networks, Bioversity International aims to stimulate the characterization and evaluation of germplasm collections by providing uniform guidelines for the description and exchange of information on germplasm.

Crop descriptors and derived standards are essential for the scientific documentation of PGR. They are an important tool permitting the international community to identify and exchange information through a 'common' language. These standards have been adopted by the semantic community, including the Trait Ontology and Generation Challenge Programme Ontology and are also being promoted by global portals such as the Crop Genebank Knowledge Base (System-wide Genetic Resources Programme) and the GENESYS Portal (Bioversity International). These standards, in turn, contribute to the development of active networks focusing on PGR, by facilitating efficient data exchange and enhancing collaboration. Furthermore, such networks can be assembled to form a global partnership, whereby access to information and use of germplasm are further supported, contributing to the aims of the International Treaty on Plant Genetic Resources for Food and Agriculture.

Keywords: Standards, Ontology, Portal, Access, Information

Introduction

This paper explores how crop standards contribute to establishing databases and developing networks, thus facilitating access to plant genetic resources (PGR) information to further the utilization of germplasm. The paper also provides insight into the history of descriptor lists - scientific standards for documenting PGR - which have been developed by Bioversity International (formerly the International Plant Genetic Resources Institute, IPGRI) since 1976.

We can not manage what we do not know. In the same way, we can not use genetic resources if we do not know about them, or where they can be found. Germplasm accessions without information have little value, like a book without a title.

An immense amount of PGR is already conserved in genebanks worldwide, but the associated information is of limited use, as either relevant documentation is lacking or data are scattered across various, disconnected genebank information systems. Most genebanks have developed unique information systems according to their specific needs, yet few linkages exist between these systems. Such constraints prevent breeders and other users from effectively searching for and identifying desired genetic traits across samples held in genebanks around the world.

Accurate characterization and evaluation data both promote and increase the use of PGR, particularly if available in a standardized format; access to information is the key link between PGR conservation and use. The conservation and sustainable use of PGR for food and agriculture is crucial for ensuring global food security; especially now, when the continuous improvement of crops is essential for sustained agriculture in response to a changing climate. This relies on the use of genetic variability through breeding.

Since its inception in 1975, Bioversity International (Bioversity) has been working to support the use and conservation of PGR, while aiming to promote an international network on genetic resources to further the collecting, conservation, documentation, evaluation and utilization of plant germplasm, thereby raising the standard of living for people throughout the world. This network was essential to establish what could be called a global system of germplasm conservation.

Initially, network partners did not have the capacity to fully understand characterization and evaluation data for common crops. Characterization data refer to characteristics and traits that enable an easy and quick discrimination between phenotypes, are highly heritable, can be easily seen by the eye and are equally expressed in all environments, while evaluation data refer to those traits that depend on the environment and require special experimental designs and techniques to assess them, such as yield or agronomic performance. This hindered the exchange of information and utilization of PGR. Furthermore, no crop-specific descriptors (or crop standards) were available; they were, instead, developed through the individual experiences and observations of those persons working on a given crop. However, with many researchers working on the same crop, it was evident that there was a need to standardize the manner in which characteristics or traits were described in order to effectively review, share and compare information from different sources. Nonetheless, cases still exist where different genebanks measure the same traits for specific crops in different ways, generating totally different values for the same crop trait.

This lack of compatibility and harmonization in PGR documentation systems has seriously hindered data exchange between and among genebank collections. As such, the need to develop an internationally approved, standardized methodology for crop trait documentation and description for use among and within different countries was identified. In order to be applicable and effective, the methodology needed to unambiguously and accurately describe a specific accession to allow for differentiation between accessions in the same collection. The comprehensive and standardized description of a crop provides for better compatibility between documentation systems and facilitates information exchange thus encouraging and supporting collaboration among scientists working in various countries.. It also reflects the values attached to traits by PGR researchers and users.

The idea of developing standardized methodologies for describing germplasm accessions was first discussed around 1976, when it was apparent that such a universally applicable scheme was essential to the effective operation of networks of genebanks or a global system for conservation. At this time, Bioversity's mandate was to care for collecting, conservation, characterization and evaluation of plant genetic material. There was little emphasis on use of PGR at this stage, as use was still considered to lead to the loss of genetic diversity when landraces were replaced by improved varieties with a narrow genetic base and higher yielding capacity. However, no other international organization was then focusing its efforts on PGR, at a global scale, and Bioversity naturally took the lead in this field of research (Gotor *et al.*, 2008).

In collaboration with partners throughout the world, Bioversity aims to build the knowledge base needed to ensure effective use of diversity in order to increase sustainable agricultural production, improve livelihoods and meet the challenges presented by climate change. The development of crop descriptors and derived standards is one avenue through which Bioversity seeks to achieve its mission, as the value of conserved PGR is dependent upon the type and quality of information available to promote germplam use.

Contribution to crop description

The development of standardized descriptors began in 1976, when the establishment of a universal system was required to promote and support global efforts in PGR conservation and for crop networks and genebanks to coordinate and operate effectively. Bioversity, in collaboration with the Centres supported by the Consultative Group for International Agricultural Research (CGIAR), crop networks, research organizations and national programmes, has since developed descriptors for more than 100 crops, published in various languages, including Spanish, French, Arabic, Portuguese and Italian.

Each crop descriptor list represents an important tool for global data exchange by providing an internationally agreed format and universally understood language for PGR information, particularly with regards to characterization and evaluation data. The descriptor lists constitute the basis for a standardized documentation system and the adoption of these tools (and in some cases, a conversion method to Bioversity's format) has helped to establish an efficient and reliable instrument for information and data exchange, storage and retrieval, facilitating the optimal utilization of germplasm.

The crop descriptors published by Bioversity, thus far, represent 3.5 million accessions held in genebanks worldwide and cover 90% of the crops listed in Annex I of the Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA). They also cover 95% of the CGIAR mandate crops. Descriptors contribute to increased knowledge among users and facilitate research activities, not only on mandate crops, but also on crops receiving limited attention from the research community, which are often those crops accessible to the poor. Of Bioversity's crop descriptor lists, 35% deal with Neglected and Underutilized Crops.

In addition to being used by the majority of individuals working on PGR, Bioversity's crop standards are also considered as 'best practices' and most have been adopted by the Crop Genebank Knowledge Base, a product of the System-wide Genetic Resources Programme (SGRP, available at http://sgrp.cgiar.org/). Furthermore, initiatives such as the Trait Ontology Consortium and the Generation Challenge Programme (GCP), have adopted the majority of the text included in the Bioversity crop standards, or have used these standards as data templates.

Impact of passport standards

When germplasm collections were integrated into multicrop collections at the national level, it was evident that common descriptors were required to improve consistency across different crops. As a result, Bioversity, in partnership with the Food and Agriculture Organization of the United Nations (FAO), and with substantial contributions from European countries and CGIAR Centres, published the FAO/IPGRI List of Multi-crop Passport Descriptors (MCPD) in 2001. The MCPD standard represents a minimum set of passport descriptors which are contained in each crop-specific descriptor list.

This passport list provides an international standard across crops to facilitate information exchange on passport data among research institutions worldwide. Each descriptor includes a brief explanation on how to record crop information and provides a coding scheme to use, when relevant. It also includes a suggested field name to assist in the computerized exchange of data.

This passport standard has had an immense impact, especially in Europe where the European Plant Genetic Resources Search Catalogue (EURISCO, available at http://eurisco.ecpgr.org/), which contains passport information on *ex situ* collections maintained in Europe, was developed based on the MCPD standard. EURISCO contains germplasm information from 40 countries, representing more than 1.0 million accessions in Europe. Additionally, the MCPD standard is being used as the basis for the development of several Central Crop Databases belonging to the European Cooperative Programme for Plant Genetic Resources (ECPGR).

Further, the FAO-WIEWS Directory of Germplasm Collections includes passport data fields following this standard and nearly all CGIAR Centres follow, when possible, the format and content proposed in the MCPD, making it easier to both retrieve and exchange data.

Outside the CGIAR system, other scientific organizations are also utilizing the MCPD standard, among them the Scottish Crop Research Institute which adopted this standard to create their database (GERMINATE), which links phenotypic and genotypic data through passport data. The GCP has also adopted the standard for the creation of 'The Bioversity/FAO Multi-crop Passport Descriptor Ontology', an adaptation of the FAO/IPGRI passport standard.

Coping with marker technologies information

With increased molecular and biochemical characterization of PGR, the need arose to also define common standards for documenting information about genetic markers. In order to address this issue, 'Descriptors for Genetic Marker Technologies' were published to complement classical agro-botanical analysis (De Vicente *et al.* 2004). This descriptor list includes a minimum set of standards for documenting information about genetic markers and is targeted at researchers who use such technologies, aiming to facilitate the documentation and exchange of standardized genetic marker data by providing descriptions of content and coding schemes to assist in computerized data exchange.

Promoting farmers' knowledge

In an attempt to integrate traditional knowledge with scientific research, Bioversity developed a standard for sharing data describing farmers' knowledge about plants in 2009. This standard has been developed in order establish a *lingua franca* to capture and share information amongst farmers and scientists and to integrate biology and traditional knowledge.

The standard provides a format for the gathering, storage, retrieval and exchange of farmers' traditional knowledge of plants. It aims to capture key characteristics, uses and values of cultivated and wild plants, as described by farmers and those living in farming communities. For example, cultural, social and socioeconomic characteristics important for farmers. Many of these descriptors are not included in conventional descriptor lists. Wild and weedy plants are also included in this standard, as they often play a significant role in farming communities, from socio-economic and ecological standpoints.

This standard focusing on farmers' knowledge is a first attempt to combine a documentation system historically used in controlled environments (genebanks, breeding institutes) with an approach that involves people and their conventional knowledge 'in the field'. It is anticipated that this descriptor list, which is the result of reviewing many years of fieldwork conducted by scientists and field practitioners, will become an important tool for linking biology and traditional knowledge in order to sustain and conserve agricultural biodiversity (Bioversity and The Christensen Fund, 2009).

Providing access to crop information

Bioversity is currently developing common information standards to describe critical characteristics of genetic resources important for crop improvement, to promote effective communication among the PGR community.

Bioversity and its partners are contributing to the development of a global portal of germplasm information to facilitate data exchange on agricultural biodiversity. The global portal, containing accession-level passport, characterization and evaluation data, will increase the utilization of biodiversity worldwide, which is the key to sustained agricultural development in the face of climate change and water scarcity (Lobell *et al.*).

The global portal, GENESYS, aims to facilitate access to and use of material stored in genebanks by breeders and other users and to make available to all, through the portal, information on the accessions included. Funded in part by the Global Crop Diversity Trust (the Trust), the development of the GENESYS portal specifically addresses obstacles faced by breeders, crop researchers and others who seek to locate hard-to-find data relating to germplasm, including information on resistance to pests, diseases and other stresses that reduce productivity and yields. The project addresses the challenge of ensuring data from germplasm collections are readily available, which has been repeatedly identified as a key to increasing PGR use.

The GENESYS project includes three components: developing information standards to describe the characteristics of genetic resources of most interest to users; deploying a genebank data-management system appropriate to users in developing countries; and developing a global accession-level information system. The first component will mobilize germplasm users to agree on core subsets of characterization and evaluation standards. These will be incorporated into the state-of-the-art GRIN-Global genebank data-management system being developed by the United States Department of Agriculture through a separate project supported by the Trust. GRIN-Global will be made available to national programmes and other genebanks through the project's second component. Finally, the third project component involves the development of GENESYS, a global portal that will link up national, regional and international genebank databases, providing access to accession-level data (see Fig. 1).
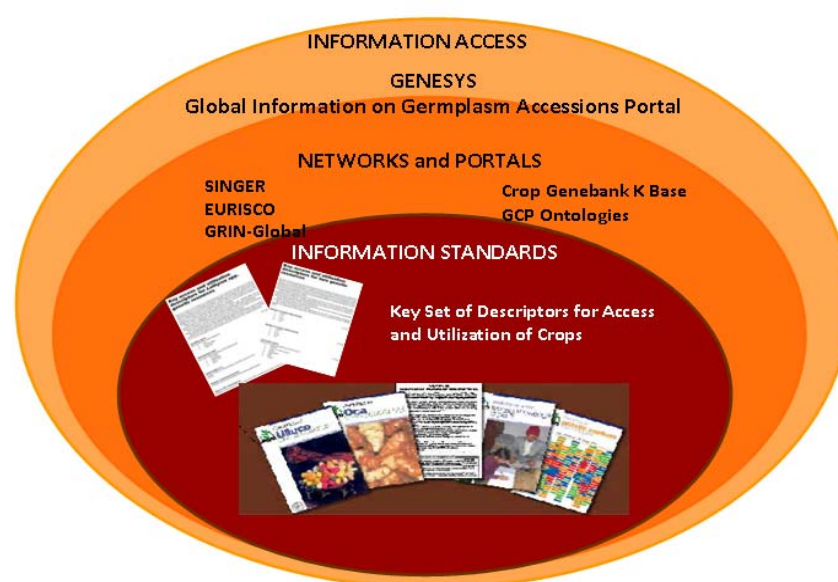

Fig. 1. Contribution of standards to developing networks and providing access to PGR utilization

Conclusion

Through the production and publication of crop descriptors, in collaboration with National Agricultural Research Systems, CGIAR centres, and crop research institutes and networks, Bioversity aims to stimulate the characterization and evaluation of PGR collections by providing uniform and unambiguous guidelines for the description and exchange of information on germplasm.

Crop descriptors and derived standards are essential for the scientific documentation of PGR. They are an instrumental tool permitting the international community to exchange information using a 'common' language, which, in turn, helps to develop active crop networks by enhancing collaboration through the easy exchange of crop-related data. Furthermore, these networks can be assembled to form a global partnership, whereby access to information and the use of germplasm are further facilitated, supporting the ITPGRFA.

Based on recent impact studies (Gotor, E., 2008), the benefits and contributions resulting from the creation of Bioversity's standards, with regards to creating networks and providing access to PGR data, can be summarized as follows:

Contributed to development of databases;
Increased uniformity of documentation;
Increased ability to work with other partners;
Facilitated data exchange;
Enabled greater efficiency in collections identifying and reducing duplication;
Helped to develop core collections; and
Increased use of accessions by helping scientists to identify relevant material.

A universal and well-defined system to catalogue and manage the PGR information flow has long been recognized as a precursor to developing an accessible data system in order to facilitate the sustainable and effective management and utilization of PGR. Bioversity, among other institutions, has established standardized procedures to define variables related to the characterization, management and evaluation of PGR. The adoption of descriptors and derived standards is integral to genetic resources activities, assuring the effective conservation of PGR and their efficient utilization in crop improvement programmes.

References

Article
GOTOR E., ALERCIA, A., RAO, R.V., WATTS J., CARACCIOLO F., 2008. The scientific information activity of Bioversity International: the descriptor lists. *Genetic Resources Crop Evolution*, 55:757-772, DOI 10.1007/s10722-008-9342-x.

Book
BIOVERSITY INTERNATIONAL and THE CHRISTENSEN FUND, 2009. *Descriptors for farmers' knowledge of plants*. Rome, Italy: 16 p. Available at: http://www.bioversityinternational.org/fileadmin/bioversity/publications/pdfs/1321.pdf

DE VICENTE M.C., METZ T. and ALERCIA, A. 2004. *Descriptors for Genetic Marker Technologies*. IPGRI, Rome, Italy. Available at: http://www.bioversityinternational.org/index.php?id=19&user_bioversitypublications_pi1[showUid]=2789

FAO/IPGRI. 2001. *FAO/IPGRI Multi-crop Passport Descriptors*. Available at: http://www.bioversit yinternational.org/index.php?id=19&user_bioversitypublications_pi1[showUid]=2192

Chapter in book
LALIBERTÉ B., WITHERS L., ALERCIA A. and T. HAZEKAMP T., 1999. Adoption of crop descriptors - IPGRI. In: *A Synthesis of Findings concerning CGIAR Case Studies on the Adoption of Technological Innovations*, L. Sechrest, M. Stewart and T. Stickle, eds. CGIAR Impact Assessment and Evaluation Group Secretariat, IAEG Secretariat, May 1999, Rome, Italy, p. 80-87.