BIOVERSITY TECHNICAL BULLETIN NO. 13

Bioversity
International

Developing crop descriptor lists

Bioversity Technical Bulletins are published by Bioversity International with the intention of putting forward definitive recommendations for techniques in genetic resources. They are specifically aimed at National Programmes.

Previous titles in this series:

A protocol to determine seed storage behaviour
T.D. Hong and R.H. Ellis
IPGRI Technical Bulletin No. 1, 1996.

Molecular tools in plant genetic resources conservation: a guide to the technologies
A. Karp, S. Kresovich, K.V. Bhat, W.G. Ayad and T. Hodgkin
IPGRI Technical Bulletin No. 2, 1997.

Core collections of plant genetic resources
Th.J.L. van Hintum, A.H.D. Brown, C. Spillane and T. Hodgkin
IPGRI Technical Bulletin No. 3, 2000.

Design and analysis of evaluation trials of genetic resources collections
Statistical Services Centre and University of Reading
IPGRI Technical Bulletin No. 4, 2001.

Accession management: combining or splitting accessions as a tool to improve germplasm management efficiency
N.R. Sackville Hamilton, J.M.M. Engels, Th.J.L. van Hintum, B. Koo and M. Smale
IPGRI Technical Bulletin No. 5, 2002.

Forest tree seed health
J.R. Sutherland, M. Diekmann and P. Berjak
IPGRI Technical Bulletin No. 6, 2002.

***In vitro* collecting techniques for germplasm conservation**
V.C. Pence, J.A. Sandoval, V.M. Villalobos A. and F. Engelmann
IPGRI Technical Bulletin No. 7, 2002.

Análisis Estadístico de datos de caracterización morfológica
T.L. Franco y R. Hidalgo
IPGRI Technical Bulletin No. 8, 2002.

A methodological model for ecogeographic surveys of crops
L. Guarino, N. Maxted and E.A. Chiwona
IPGRI Technical Bulletin No. 9, 2005.

Molecular markers for genebank management
D. Spooner, R. van Treuren and M.C. de Vicente
IPGRI Technical Bulletin No. 10, 2005.

***In situ* conservation of wild plant species a critical global review of good practices**
V.H. Heywood and M.E. Dulloo
Bioversity Technical Bulletin No. 11, 2006

Crop genetic diversity to reduce pests and diseases on-farm. Participatory diagnosis protocols. Version 1.
D.I. Jarvis and D.M. Campilan
Bioversity Technical Bulletin No. 12, 2007

Copies can be obtained in PDF format from Bioversity's Web site (www.bioversityinternational.org) or in printed format by sending a request to bioversity-publications@cgiar.org.

Developing crop descriptor lists

Guidelines for developers

Bioversity International is an independent international scientific organization that seeks to improve the well-being of present and future generations of people by enhancing conservation and the deployment of agricultural biodiversity on farms and in forests. It is one of 15 centres supported by the Consultative Group on International Agricultural Research (CGIAR), an association of public and private members who support efforts to mobilize cutting-edge science to reduce hunger and poverty, improve human nutrition and health, and protect the environment. Bioversity has its headquarters in Maccarese, near Rome, Italy, with offices in more than 20 other countries worldwide. The Institute operates through four programmes: Diversity for Livelihoods, Understanding and Managing Biodiversity, Global Partnerships, and Commodities for Livelihoods.

The international status of Bioversity is conferred under an Establishment Agreement which, by January 2006, had been signed by the Governments of Algeria, Australia, Belgium, Benin, Bolivia, Brazil, Burkina Faso, Cameroon, Chile, China, Congo, Costa Rica, Côte d'Ivoire, Cyprus, Czech Republic, Denmark, Ecuador, Egypt, Greece, Guinea, Hungary, India, Indonesia, Iran, Israel, Italy, Jordan, Kenya, Malaysia, Mali, Mauritania, Morocco, Norway, Pakistan, Panama, Peru, Poland, Portugal, Romania, Russia, Senegal, Slovakia, Sudan, Switzerland, Syria, Tunisia, Turkey, Uganda and Ukraine.

Financial support for Bioversity's research is provided by more than 150 donors, including governments, private foundations and international organizations. For details of donors and research activities please see Bioversity's Annual Reports, which are available in printed form on request from bioversity-publications@cgiar.org or from Bioversity's Web site (www.bioversityinternational.org).

The geographical designations employed and the presentation of material in this publication do not imply the expression of any opinion whatsoever on the part of Bioversity or the CGIAR concerning the legal status of any country, territory, city or area or its authorities, or concerning the delimitation of its frontiers or boundaries. Similarly, the views expressed are those of the authors and do not necessarily reflect the views of these organizations.

Mention of a proprietary name does not constitute endorsement of the product and is given only for information.

Citation
Bioversity International. 2007. Guidelines for the development of crop descriptor lists. Bioversity Technical Bulletin Series. Bioversity International, Rome, Italy. xii+72p.

ISBN: 978-92-9043-792-1

Bioversity encourages the use of material from this publication for educational or other non-commercial purposes without prior permission from the copyright holder. Acknowledgement of Bioversity's material is required. This publication is available to download in portable document format from URL: <http://www.bioversityinternational.org>.

Bioversity International
Via dei Tre Denari, 472/a
00057 Maccarese
Rome, Italy
© Bioversity International, 2007

Contents

Acknowledgements **v**
Foreword **vii**
Published Descriptor lists **ix**

BACKGROUND **1**

A brief history of descriptors **1**
Bioversity's role **3**
Impact of descriptors **3**

INTRODUCTION **5**

1. The Concept of descriptor lists **5**
1.1 Descriptor definitions 5
1.2 Descriptor elements 6
1.2.1 Descriptor names *6*
1.2.2 Descriptor states *7*
1.2.3 Descriptor methods *9*

2. Descriptors and derived standards **11**
2.1 Crop-specific descriptors 12
2.2 Multi-crop passport descriptors (MCPD) 12
2.3 Descriptors for genetic marker technologies 12

3. Crop-specific descriptors **13**
3.1 Crop descriptor categories 13
3.1.1 Passport category *14*
3.1.2 Management category *15*
3.1.3 Environment and site category *15*
3.1.4 Characterization category *16*
3.1.5 Evaluation category *17*

4. Similarities to and differences from other technical guidelines **18**
4.1 Union Internationale pour la Protection des Obtentions Végétales 19
4.2 Council for Mutual Economic Aid 19
4.3 United States Department of Agriculture Genetic Resources Information Network (USDA-GRIN) 20

GATHERING AND ANALYSING DATA **21**

5. Types of data **21**
5.1 Qualitative data 21
5.2 Quantitative data 25

6. Types of Scales **26**
6.1 Scales for qualitative characteristics 27
6.1.1 Qualitative descriptors using nominal scale *28*
6.1.2 Qualitative descriptors using ordinal scale *29*
6.1.3 Qualitative descriptors using a binary scale *30*
6.2 Scales for quantitative characteristics 30
6.2.1 Quantitative descriptors using a continuous scale *30*
6.2.2 Quantitative descriptors using a discrete scale *31*

7. Resolution, complexity and costs **33**
7.1 Level of precision required 33
7.2 Complexity of the measurement or observation 34
7.3 Cost per measurement 34

8. Methods for collecting data **35**

9. Experimental design **35**

10. Scoring, coding and recording of descriptors **36**
10.1 Recording heterogeneous data 38

11. Numeric versus alphanumeric coding schemes **41**

MAKING DESCRIPTORS WORK **43**

12. The development process **43**

13. A few basic rules **43**

14. Step-by-step checklist for defining descriptors **44**

LOOKING FORWARD **47**

BIBLIOGRAPHY AND FURTHER REFERENCES **49**

Appendix I – Contributors, Coordinators and Reviewers **51**

Appendix II – Standard reference sources **53**
A2.1 References and standards used in developing crop descriptor lists 53
A2.2 Base units – Système International d’Unités (SI) 54

Appendix III – FAO/IPGRI Multi-Crop Passport Descriptors **55**

Appendix IV – List of standard descriptors for site environment **59**

Appendix V – Example Collecting form **70**

Acknowledgements

Bioversity International wishes to express its sincere thanks to the numerous scientists and researchers around the world who contributed directly or indirectly to the development of the Guidelines for the development of Crop Descriptor Lists.

Ms Adriana Alercia coordinated and managed the production of this document and provided technical and scientific advice. Ms Audrey Chaunac provided assistance during the production process. Ms Patrizia Tazza prepared the cover and layout.

Special thanks are due to Dr Ramanatha Rao and Bioversity scientists for their comprehensive technical and scientific advice given during the development process for this publication.

Foreword

The world faces a continual need to increase crop productivity, and to develop new varieties more adapted to changing environmental and biological challenges or to meet the evolving needs of local communities. To meet these needs and challenges, farmers and breeders not only must have access to a wide range of plant genetic resources but also must have access to the essential information about those plant genetic resources that will allow effective use to be made of them. These guidelines have been developed to assist genebank curators, breeders, plant scientists, national programmes, networks and users of genetic resources working with specific crops and gene pools to develop their own descriptor lists in order to characterize their material and make information available to others in a systematic and unambiguous form.

In order to increase international exchange of material, a minimum element of uniformity is critical in data collection, recording, storage and retrieval. Developing standards for documentation and for exchanging information is essential for ensuring that the vast amount of data on crop species and varieties is available to countries to improve their capacity to store, manage and share information about biodiversity. The development of descriptor lists will assist in the systematic and objective recording and exchange of information such as passport, characterization and evaluation data, which in turn will increase utilization of germplasm so that people can make better use of biodiversity.

Descriptors have been developed by Bioversity International and its predecessors, the International Board for Plant Genetic Resources (IBPGR) and the International Plant Genetic Resources Institute (IPGRI), for almost 100 crops in collaboration with scientists and international research organizations. However, there is a high demand for new descriptor lists to be developed for many species and new crops, including neglected crops, crops of regional or local importance, and forest species.

The guidelines presented here have been produced based on experience gained from a wide range of crop studies and collaboration with many scientists, national programmes and crop networks. Various drafts of this Guide were circulated at different times to a number of Bioversity scientists and this publication is the consolidated result of those consultations (see Appendix I — Contributors). The development process was coordinated by Adriana Alercia, with scientific guidance from Dr Ramanatha Rao.

These guidelines provide background information, set objectives and give insights into the structure, elements and methodology used by Bioversity to develop descriptor lists. They also provide a step-by-step checklist for defining characterization and evaluation descriptors, which can serve as a quick reference guide when developing new descriptor lists.

Bioversity is thankful for the scientific advice and suggestions contributed by many scientists during the development of these guidelines.

Published Descriptor Lists

Allium (E, S)	2000
Almond (revised)* (E)	1985
Apple* (E)	1982
Apricot* (E)	1984
Avocado (E, S)	1995
Bambara groundnut (E, F)	2000
Banana (E, S, F)	1996
Barley (E)	1994
Beta (E)	1991
Black pepper (E, S)	1995
Brassica and *Raphanus* (E)	1990
Brassica campestris L. (E)	1987
Buckwheat (E)	1994
Capsicum* (E, S)	1995
Cardamom (E)	1994
Carrot (E, S, F)	1999
Cashew* (E)	1986
Chenopodium pallidicaule (S)	2005
Cherry* (E)	1985
Chickpea (E)	1993
Citrus (E,F,S)	1999
Coconut (E)	1992
Coffee (E, S, F)	1996
Cotton (Revised)* (E)	1985
Cowpea* (E)	1983
Cultivated potato* (E)	1977
Echinochloa Millet* (E)	1983
Eggplant (E, F)	1990
Faba bean* (E)	1985
Fig (E)	2003
Finger millet* (E)	1985
Forage grass* (E)	1985
Forage legumes* (E)	1984
Grapevine (E, S, F)	1997
Groundnut (E, S, F)	1992
Jackfruit (E)	2000
Kodo millet* (E)	1983
Lathyrus spp. (E)	2000
Lentil* (E)	1985
Lima bean* (E)	1982
Litchi	2002
Lupin* (E, S)	1981
Maize (E, S, F, P)	1991
Mango (Revised) (E)	2006
Mangosteen (E)	2003
Medicago (Annual)* (E, F)	1991
Melon (E)	2003
Mung bean* (E)	1980
Oat* (E)	1985
Oca* (S)	2001
Oil palm (E)	1989
Palmier dattier (F)	2005
Panicum miliaceum and *P. sumatrense* (E)	1985
Papaya (E)	1988
Peach* (E)	1985
Pear* (E)	1983
Pearl millet (E, F)	1993
Pepino (E)	2004
Phaseolus acutifolius (E)	1985
*Phaseolus coccineus** (E)	1983
Phaseolus lunatus (P)	2001
*Phaseolus vulgaris** (E, P)	1982
Pigeonpea (E)	1993
Pineapple (E)	1991
Pistacia (excluding *P. vera*) (E)	1998
Pistachio (E, F, A, R)	1997
Plum* (E)	1985
Potato varieties* (E)	1985
Quinua* (S)	1981
Rambutan (E)	2003
Rice* (E)	2006
Rocket (E,I)	1999
Rye and Triticale* (E)	1985
Safflower* (E)	1983
Sesame* (E)	2004
Setaria italica and *S. pumila* (E)	1985
Shea tree (E)	2006
Sorghum (E, F)	1993
Soyabean* (E, C)	1984
Strawberry (E)	1986
Sunflower* (E)	1985

Sweet potato (E, S, F)	1991
Taro (E, F, S)	1999
Tea (E, S, F)	1997
Tomato (E, S, F)	1996
Tropical fruit* (E)	1980
Ulluco (S)	2003
Vigna aconitifolia and *V. trilobata* (E)	1985
Vigna mungo and *V. radiata* (Rev.)* (E)	1985
Walnut (E)	1994
Wheat (Revised)* (E)	1985
Wheat and *Aegilops** (E)	1978
White Clover (E)	1992
Winged Bean* (E)	1979
*Xanthosoma** (E)	1989
Yam (E, S, F)	1997

Bioversity International's publications are available free of charge to the libraries of genebanks, university departments, research institutions, etc., in the developing world. E, F, S, C, P, I, R and A indicate English, French, Spanish, Chinese, Portuguese, Italian, Russian and Arabic respectively. Titles marked with an asterisk are out of print, but are available as Adobe Acrobat portable document format (PDF) on request (send email to: **bioversity-publications@cgiar.org**). Organizations in the developed world and individuals requiring personal copies can order copies of Bioversity's publications from EarthPrint.com (**www.earthprint.com**).

Background

One of the main reasons for the under-utilization of germplasm, according to curators, breeders and other users of plant genetic resources, is the lack of adequate passport, characterization and evaluation data: people cannot use genetic resources that lack essential information. In addition, such information is necessary for proper management of the resources in the genebanks by genebank managers. Therefore, the accurate documentation of information about the origin, characterization and performance of germplasm is essential for effective conservation and use. To this end, Bioversity has been promoting the documentation of plant genetic resources and data exchange by providing collection curators with uniform guidelines to document their information through the production of 'descriptor lists' to describe effectively diversity, which allowed better communication between scientists and institutions, resulting in increased use of conserved genetic resources.

Exchange of data and information between national programmes for plant genetic resources can help to increase these programmes' efficiency by minimizing unnecessary duplication of activities and facilitating priority setting for germplasm collecting, regeneration of accessions and other activities.

In order to exchange data, it is necessary to have compatible documentation systems. This can only be achieved through common standards for information exchange. Descriptor lists provide such well-established standards. Even though different programmes or institutions use different documentation systems (i.e. hardware and software) or languages, if they use a common descriptor system the exchange of information is facilitated.

The crop descriptor lists have an internationally accepted format and have been developed by and are shared among scientists worldwide. The utilization of a 'universal language' in well defined and thoroughly-tested descriptor lists for characterizing germplasm simplifies data recording, updating, modification, retrieval, exchange and analysis. Germplasm conservers and users adopting the same descriptors at different locations are able easily to exchange and interpret each other's data.

A brief history of descriptors

Descriptor lists have been an important element of Bioversity's germplasm documentation activities almost since the establishment of IBPGR in the 1970s and the production of the first descriptor list in 1977, although the concept, process and format have evolved substantially over the years.

- **Minimum descriptors**. The original aim of descriptor lists was to provide a *minimum* number of characteristics to describe a particular crop (e.g. Soyabean). One problem with these initial descriptor lists was that several useful additional descriptors lacked the appropriate internationally accepted definitions and descriptor states needed for consistent recording and to be able to communicate with other institutions. This lack of compatibility in documentation systems for describing plant genetic resources seriously hampered data exchange between collections.
- **Comprehensive lists of descriptors.** The idea of minimum lists was revisited in 1990, and a new approach was developed. *Comprehensive* lists of descriptors were produced including all descriptors for characterization and evaluation (e.g. Descriptors for Sweet Potato/ Descripteurs pour la Patate Douce/Descriptores de la Batata, developed in collaboration with AVRDC and CIP in 1991). The comprehensive descriptor lists also included a number of standard detailed sections (e.g. site environment and management) that were common across different crop descriptor lists and that provided users with options to choose from. This improved compatibility between documentation systems and the ease of information exchange.
- **Highly discriminating descriptors for international harmonization**. It was recognized that each curator utilized only those descriptors that were useful for the maintenance and management of their collection. Consequently, the descriptor lists were further revised in 1994 in order to provide users with more comprehensive lists but at the same time containing a minimum set of *highly discriminating descriptors*, which were flagged in the text with asterisks (*) (e.g. in Descriptors for Barley (*Hordeum vulgare* L.) [1994]).

The asterisked descriptors are those that have potential to discriminate between accessions and are important for the international harmonization of plant genetic resources data documentation. These highly discriminating descriptors also provide basic indicators of diversity within a collection. Curators and others involved in characterization and evaluation of germplasm can complement them with additional descriptors from descriptor lists, depending on the specific objectives of the collection.

Nowadays, descriptor lists tend to be comprehensive, providing an internationally recognized reference for most, if not all, known descriptors for a particular crop or gene pool. This does not mean that every curator need use all the descriptors listed, but can instead select those considered relevant to the collection. For example, a fruit tree collection maintained for the purpose of representing a broad diversity of ornamental forms would probably emphasize a set of descriptors different from those used for a collection representing diversity in general. Likewise, for temperate and tropical species, different descriptors and descriptor states might be used to describe environmental conditions at the site of collection, regeneration or evaluation.

Bioversity’s role

A major reason for the success of the descriptors developed by Bioversity and its partners is that they are well researched and are the result of extensive collaboration among scientists worldwide. Bioversity’s role is to act as international facilitator and coordinator, ensuring that the full extent of global knowledge and expertise is reflected in the crop descriptors, that a standard format is maintained in the face of potentially conflicting opinions, and that the final results have the broad support and consensus of the majority of experts.

When dealing with mandate crops of the Consultative Group on International Agricultural Research (CGIAR), Bioversity seeks the scientific advice of the relevant CGIAR centre and collaborates with it in the production of the descriptors. This is crucial for descriptor development since these centres have the expertise needed to elaborate a high quality product as they conserve and work with large and diverse collections of specific crops.

Bioversity has taken the lead in the documentation of plant genetic resources. The international status of the descriptor lists, as well established guidelines for documentation, is illustrated by its collaboration with international and national organizations, such as The International Union for the Protection of New Varieties of Plants (UPOV), Organisation internationale de la vigne et vin (OIV), The World Vegetable Centre (AVRDC), CGIAR Centres, Instituto Nacional de Investigación Agropecuaria (INIA), French Agricultural Research Centre for International Development (CIRAD), Institut national de la recherche agronomique (INRA), and a number of universities and research organizations. Some indications of the impact of descriptors that have been drawn in collaboration with these sources are discussed below.

Impact of descriptors

An indication of the impact of descriptors can be seen in the recommendations made during the Second Technical Meeting of Focal Points for Documentation in East European Genebanks (Radzikow, Poland, 1995):

> *The Meeting concluded that standardization of evaluation and characterization descriptors is not desirable on a multi crop basis. The crop-specific descriptor lists, compiled in consultation with relevant crop experts worldwide, provide the crop-specific standards (van Hintum et al. 1995).*

This was also confirmed during the preparatory process for the International Conference and Programme for Plant Genetic Resources (ICPPGR) organized at Leipzig, Germany in 1996. The country reports provided a further useful indication of the extensive use of these descriptors; their use is also cited by many countries in the State of the World Report on Genetic Resources (FAO 1996):

> *To derive an indication of the use of descriptor lists, 152 country reports were analysed in IPGRI HQ for the Leipzig Conference. The results showed that descriptor lists have a high degree of penetration and use at the international level. Of the countries surveyed, 102 are undertaking characterization or evaluation and specified which descriptors they are using. Of these countries, 93 (or 91%) use IPGRI descriptor lists alone or in combination with other lists.*

In 1999, the CG Secretariat published a 'Synthesis of findings concerning CG Case Studies on the Adoption of Technological Innovations' (Laliberté et al., 1999). The impact study was related to the adoption of the crop descriptors developed for three different crops. Some key conclusions could be drawn from the results of this case study:

- IPGRI (now Bioversity International) descriptors are well known international standards for the detailed description of crop specific resources and are used by the majority of germplasm collection managers.
- Users consider the descriptors to **be very useful** for a range of applications, such as characterization, standardization of information, the establishment of databases, documentation of accessions, creation of core collections, and data exchange.

Of the 143 germplasm collection managers responding to the above survey regarding the use of descriptors, 80% used descriptors in general and 69% used Bioversity International (ex-IPGRI) descriptors, while the remaining 11% used their own descriptors or those developed by UPOV and COMECON.

Introduction

One of the aims of developing international standards such as descriptor lists is to make the most of biodiversity by describing crops and species in a *lingua franca* and increasing access to this diversity by the users of plant genetic resources, who in turn will benefit from their economic and social values.

The crop descriptors more recently published include an 'Introduction to the crop' section, which aims to promote a specific crop to an audience not very familiar with it. Such a section in the descriptors is particularly useful because it highlights different values of crops, from nutrition to income generation, which might not be apparent to user communities. It is also important because it suggests a method of characterization and evaluation that can be used to demonstrate potentials and benefits, as in the case of *Descriptors for Rocket*. Another example is the list of *Descriptors for Date Palm*, which lists the full spectrum of benefit opportunities (e.g. use of dates to make vinegar or jams), which are not necessarily well known among scientists and other users.

1. The concept of descriptor lists

When a species name is identified and listed along with its accession number, when different shapes of a fruit are described, when the length of the leaf is measured or the number of accessions in the collection of a crop species are recorded, observations are being made on the specific attributes of a particular plant, and each characteristic is called a '*descriptor*'.

Descriptor lists include key attributes, characteristics or traits of a crop, and set out the method used to measure and document them, along with the relevant registration data.

Descriptor lists therefore aim to include information and data that are relevant for different types of genebank operations for a specific crop or gene pool, from initial registration, through characterization, evaluation and management, to their eventual use.

1.1 Descriptor definitions

Within the plant genetic resources community, a ***descriptor*** is defined as an attribute, characteristic or measurable trait that is observed in an accession of a genebank. It is used to facilitate data classification, storage, retrieval, exchange and use.

Example #01

Accession number
Flower colour
Plant height

A ***descriptor list*** is a set of individual descriptors used for the description of germplasm of a particular crop or species (e.g. list of *Descriptors for Pistachio*).

1.2 Descriptor elements

Each descriptor consists of a *descriptor* ***name***, a *descriptor* ***state,*** and a *descriptor* ***method*** explaining how the descriptor should be measured and recorded. A descriptor state could be a quality, measurable attribute or code.

Example #02		
Stem pubescence		*(descriptor name)*
Observed at the stem base		*(descriptor method)*
3	Sparse	*(descriptor state)*
5	Intermediate	*(descriptor state)*
7	Dense	*(descriptor state)*

In Example 02, 'stem pubescence' is the descriptor name; 'observed at the stem base' is the descriptor method, and 'sparse; intermediate; dense' are the descriptor states, with corresponding numbering codes (3, 5, 7) assigned to descriptor states for ease of documentation.

1.2.1 Descriptor names

The descriptor should have a full ***name*** that is *descriptive*, *unambiguous* and as *compact* as possible.

Descriptor names are frequently composed of an *object* or item, and a characteristic or *attribute* name.

Example #03

Accession *number*
Species *name*
Leaf *colour*
Flowering *habit*
Soil *fertility*

When choosing a descriptor name, it is essential to verify that the technical terms are correct and that they are generally accepted and understood by other users. The use of a glossary of botanical terms is recommended (see Bibliography and further reference sources listed in Appendix II).

1.2.2 Descriptor states

For a number of qualitative and quantitative descriptors, a ***descriptor state*** is a clearly definable state of expression to define a characteristic and harmonize descriptions. It represents the variation in the observations or measurements made on a particular descriptor. Each descriptor is allocated a corresponding numeric code for ease of data recording and exchange.

Example #04

Leaf shape		
1	Cordate	*(descriptor state)*
2	Oblong	*(descriptor state)*
3	Ovate	*(descriptor state)*

Reference materials can be used to help define the various states of expression of traits, and recommended resources include drawings, check cultivars, colour charts, phenological scales, illustrations, and lists of possible values or codes (if applicable). Examples of such reference material are given in the following sections.

1.2.2.1 Drawings

Since collections of the plant genetic resources of a crop could be sited anywhere, and large in number, a collection may not have access to a standard reference (see Section 1.2.2.2), so simple line drawing or pictures of stem branching, for example, are easier to refer to and will help users to selecting states of expression of a trait, avoiding confusion with environmental effects.

Figure captions should be brief, but complete, and should contain the name of the relevant descriptor. If a figure is taken from or based on another source, a full bibliographic reference to the source should be included in an appendix of the descriptor list.

Example #05

Stem branching		See Figure 1 (overleaf)
1	Opposite	
2	Alternate	
3	Ternate	
4	Mixed	

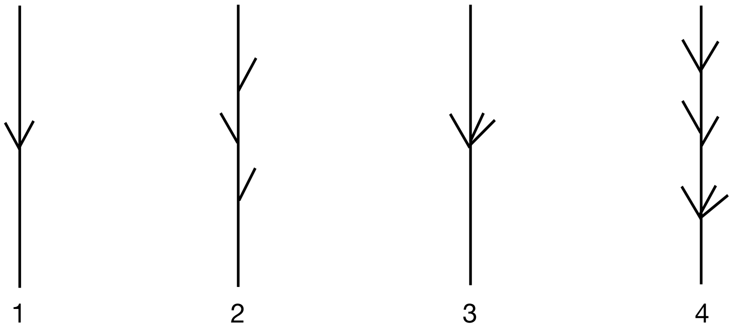

Figure 1. Stem branching

1.2.2.2 Reference standards and parameters

When a descriptor state is open to interpretation or difficult to explain, reference standards or specific parameters can be used to clarify it.

Reference standards provide an objective baseline against which measurements and comparisons can be made. They provide the means to make observations more consistent and comparable.

Often, a common cultivar is used as a standard and the standard reference is then used as a check. Check cultivars and standard references also provide useful corollary information to gauge the performance of the accessions being tested. Check cultivars should be widely available and known.

Example #06

Blade shape of mature leaf		Standard reference
1	Cordate	*Vitis cordifolia*
2	Wedge-shaped	*Vitis riparia*
3	Pentagonal	cv. Chasselas blanc
4	Circular	cv. Clairette
5	Reniform	cv. Rupestris du Lot

1.2.2.3 Colour charts

A fruit colour descriptor that describes different shades of a colour would benefit greatly from the use of a colour chart or reference standard, if available. Without a reference for comparison, descriptor states such as ‘light green’, ‘green’ and ‘dark green’ can not be scored consistently and objectively.

Example #07

Fruit colour		RHS colour code (RHS, 1986)
1	Light green	145A
2	Green	146A
3	Dark green	147A

1.2.2.4 Parameters

It is strongly recommended to use actual measurements (cm, g, mm) for making good use of quantitative data (i.e. continuous variation) for genetic diversity analysis. Actual measured values can also give us statistical data to assess variation within an accession. Character states as listed below should be used only when measuring is very difficult.

For instance, a fruit length descriptor should specify relevant ranges of measurements to avoid misinterpretation by different users. Without these ranges, descriptor states cannot be scored consistently or objectively.

Example #08

	(WRONG)		**(RIGHT)**	
	Fruit length		**Fruit length**	
1	Very short	1	Very short	(<2 cm)
2	Very short to short	2	Very short to short	(>2 – 4 cm)
3	Short	3	Short	(>4 – 6 cm)
4	Short to intermediate	4	Short to intermediate	(>6 – 8 cm)
5	Intermediate	5	Intermediate	(>8 – 10 cm)
6	Intermediate to long	6	Intermediate to long	(>10 – 12 cm)
7	Long	7	Long	(>12 – 14 cm)
8	Long to very long	8	Long to very long	(>14 – 16 cm)
9	Very long	9	Very long	(>16 cm)

1.2.3 Descriptor methods

A ***descriptor method*** describes in detail how and under what conditions a descriptor is measured or scored. The description method facilitates accurate interpretation of results and provides a protocol to be universally and consistently applied.

Example #09

Plant height [cm]

Recorded at maturity, measured from ground level to the top of spike, excluding awns. Average of 5 randomly selected plants

It is important to use technically correct terminology in descriptions. If possible, record any bibliographical references consulted and list them in an appendix to the descriptor list. This will allow others to verify the terminology and methodology. References commonly used in the development of descriptors are listed in Appendix II.

Descriptor method elements comprise:

- an *Object*;
- a *Condition*; and
- a *Sampling procedure*.

These are considered more fully below.

Object

This defines the exact part(s) of the plant to be observed or measured. A measurement of plant height that does not specify exactly between which points the measurement should be taken is incorrect, because different people may use different measuring points. In the case of quantitative descriptors, a unit of measurement should be defined. It is recommended to use only the Système International d'Unités (SI) (See Appendix II) and to include the units to be applied in square brackets following the descriptor name.

Example #10

Leaf lamina length [mm]

Recorded at the widest point. Average of 10 fully developed leaves taken from three different adult trees. Use apical leaflet in the case of compound leaf.

Condition

This defines the conditions under which the observation is made, such as duration, plant growth stage, phenological condition, temperature, humidity, 'priming' (pre-observation treatments), and specifications of particular equipment if required. In the above example (Example 10), 'fully developed' is the condition.

Sampling

The number of samples on which the observation is based should be given, thus providing an indication of data accuracy within the method. The type of method used for sample selection (random, stratified, etc.) should also be indicated. When variation of a characteristic within the accession is prevalent, it is essential to describe how the samples are to be selected and how many samples are needed.

2. Descriptors and derived standards

Exchange of information requires compatibility of documentation systems. Documentation systems can be fully compatible even if different hardware or software is used, but this implies consensus regarding standards for exchange of information, and consistency in the implementation of those standards.

Bioversity has developed three types of standard (see Figure 2):

- Crop descriptors,
- Multi-crop passport descriptors (MCPD) (FAO/IPGRI), and
- Descriptors for genetic marker technologies.

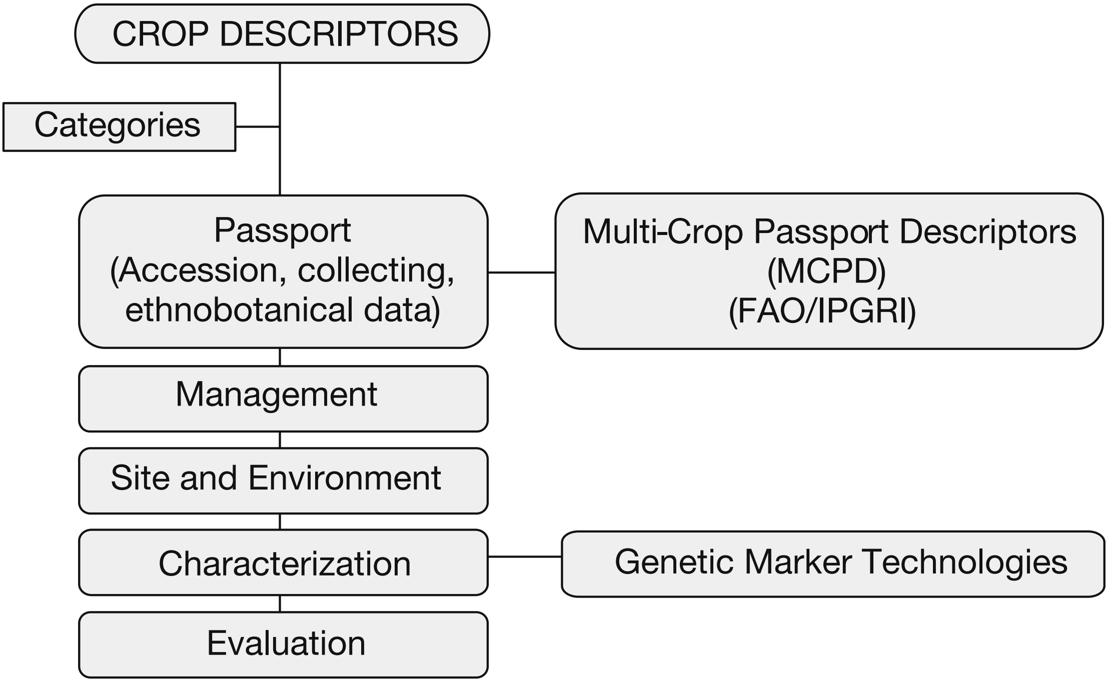

Figure 2. Descriptors and derived standards.

2.1 Crop-specific descriptors

The descriptors lists are targeted at curators, breeders, scientists and others managing crop genetic resource collections. It is an important tool in standardizing documentation systems, providing as it does an international format and a universally understood 'language' for plant genetic resources data (see Section 3).

2.2 Multi-crop passport descriptors (MCPD)

With the integration of collections at the national level into multicrop collections, it became evident that common descriptors needed to be more consistent across crops. As a result, Bioversity and FAO, with substantial contributions from European countries through the European Cooperative Programme for Crop Genetic Resources Network (ECPGR Network) and CG Centres through the System-wide Information Network for Genetic Resources (SINGER system), developed a subset of passport descriptors: the FAO/IPGRI List of Multi-crop Passport Descriptors (MCPD) (Alercia et al. 2001).

The MCPD list is a reference tool that provides international standards to facilitate germplasm passport information exchange across crops. These descriptors are compatible with the Bioversity crop descriptor lists, with the descriptors used for the FAO World Information and Early Warning System (WIEWS) on plant genetic resources, with CG Centres and with European countries through the EURISCO Catalogue. In 2005, the MCPD list was fully adopted in the Passport Module for the development of the GERMINATE database (which integrates genotypic and phenotypic information for plant genetic resources collections) in the Generation Challenge Programme, and many other initiatives. The MCPD list has had a very positive effect on the establishment of central crop databases, especially in Europe.

2.3 Descriptors for genetic marker technologies

This list of descriptors defines a minimum set of data needed to describe accessions using biochemical and molecular markers, and defines community standards for documenting information about genetic markers. The document, which was originally based on some of the descriptors listed in the traditional evaluation category of the crop descriptors, is targeted at researchers using genetic marker technologies, to facilitate the generation and exchange of standardized genetic marker data. It also provides descriptions of content and coding schemes that can assist in computerized data exchange.

3. Crop-specific descriptors

The crop descriptor lists provide the plant genetic resources community with internationally recognized guidelines for the standardized description of accessions of different crops. As noted earlier, this series of descriptor lists has had a major impact in the global plant genetic resources community and in the management of plant genetic resources.

The purpose of this standardization is to manage genetic resources, enhance the exchange of information, and increase the efficiency of communication among germplasm scientists and users of plant genetic resources. An additional purpose is to facilitate the use of germplasm resources by the plant genetic resources community.

3.1 Crop descriptor categories

To facilitate the maintenance, retrieval and updating of information on accessions, it is advisable to organize descriptors into practical sets. Bioversity has classified them into five main categories, and encourages the collection of data for all five types of categories.

These categories are:

- Passport
- Management
- Environment and site
- Characterization
- Evaluation

Data from the first four categories should be available for every accession. The number of descriptors selected in each of the categories will depend on the crop or species and its character. Descriptors listed under *Evaluation* allow for a more extensive description of an accession, but generally require replicated trials over time.

It is recommended that information be produced by closely following the descriptor list with regard to ordering and numbering of descriptors, using the descriptors specified, and using the recommended descriptor states.

In general, passport, environment and site categories are similar for all crops. Recent Bioversity descriptor lists and the MCPD list can be used as reference sources. For the definition of Management descriptors, descriptor lists of crops with similar maintenance regimes can be consulted. However, Characterization and Evaluation descriptors are very crop specific, and must be defined specifically for every new crop. A checklist that can be used in the development of such descriptors can be found in 'Making descriptors work' (Sections 12 to 14, pp. 43–44).

3.1.1 Passport category

Passport data provide the basic information used for the general management of an accession (including registration at the genebank and other identification information), and describe parameters to be observed when the accession is originally collected. They constitute a crucial element in the registration process when the sample is registered in the genebank as an accession.

Passport category is usually divided into two sections:

- accession descriptors; and
- collecting descriptors.

Accession descriptors are identification data related to the registration of the sample at the genebank. These descriptors are fundamental to the documentation system since descriptors such as accession number, genus and species can be related to different accession-specific data.

Example #11

Accession number
Genus
Species
Donor number

A large part of the passport descriptors category for each sample is recorded during germplasm collecting. These data describe in detail the environment from which the germplasm originates. Ethnobotanical descriptors, such us ethnic group, local vernacular name, plant uses and parts of the plant used, form an increasing proportion of these descriptors.

Example #12

Collecting institute
Country where collected
Collecting date
Collecting site

In order to help users during data collection in the field, Bioversity has developed a collecting form (see Appendix V for an example of a collecting form for *Allium* spp.), which is usually included as an appendix to the crop descriptor lists.

It is important that passport data are as complete as possible from the beginning, since it is often difficult or even impossible to fill in gaps at a later stage. Passport descriptors are to a large extent applicable to all crops and species. To facilitate international access and exchange of information, it is strongly recommended that the MCPD list be used as a reference (Alercia et al. 2001). The full MCPD list can be found in Appendix III.

3.1.2 Management category

Management descriptors provide the basis for the day to day management of accessions in a genebank and assist with their multiplication and regeneration. The genebank curator must ensure that these descriptors are recorded during multiplication, storage, maintenance or regeneration of each accession. Management descriptors vary according to crops or gene pools. However, collections with similar management regimes that are kept as seed, in the field, *in vitro* or as cryopreserved collections often have many descriptors in common. Recent descriptor lists of similar crops can be used as a reference basis for collections that are managed in field genebanks (e.g. *Descriptors for Rambutan*, and *Descriptors for Jackfruit*).

Typically, the management category is divided into two sections, according to the crop being described:

- plant or seed management descriptors; and
- multiplication or regeneration descriptors.

Management descriptors provide information on the amount of seed available and the viability of the seed, along with the date of the germination test. It also provides information on the number of replicates of an accession in a field genebank. The location of an accession in the genebank and the places to where it has been distributed are usually provided, as well as the location of an accession in a field genebank. When relevant, cryopreservation and *in vitro* descriptors are also included.

Example #13

Sowing date	[YYYYMMDD]
Harvest date	[YYYYMMDD]
Seed germination at storage	[%]

3.1.3 Environment and site category

Descriptors in this category describe environmental and site-specific parameters, particularly in association with characterization and evaluation trials. They are important for interpreting the results of those trials due to genotype × environment interaction. However, the level of detail needed for describing the site and environment of the characterization and evaluation will vary according to the crop, and users will select only those relevant to the crop being described.

Environment and site category are usually divided into two sections:

- characterization and/or evaluation site descriptors; and
- site environment descriptors.

Descriptors for the *characterization or evaluation site* are the same for all crops, and any recent descriptor lists can be used for reference (e.g. *Descriptors for Citrus*). They typically include the country of characterization and/or evaluation, the location (latitude and longitude), elevation, and planting and harvest dates of the characterization or evaluation trials.

Example #14

Country of characterization and/or evaluation

Site

Latitude

Longitude

Elevation

The descriptors listed under *Site environment* will be useful during collecting activities in order to describe the collection or sampling source environment, and the characterization and evaluation site environments. They include standard descriptors for soil [matrix colour, depth, pH, texture class(es)], topography, slope and climatic information such as rainfall and temperature. Site environment descriptors can be found in Appendix IV.

Example #15

Topography

Higher-level landform

Slope

Slope aspect

3.1.4 Characterization category

Describing plants is one of the most important ways that plant genetic resources users can contribute to germplasm utilization and conservation efforts.

Descriptors in this category are observations about plant characteristics that can be used for diagnostic purpose to describe the plants or trees of an accession and differentiate them from those belonging to another accession. Therefore, data gathered during characterization are used for distinguishing accessions. They provide information on the type of plants that are in a collection, and information potentially useful in crop development. They may also provide a tool to evaluate claims of novelty (helpful for variety protection or plant patents, as in the case of the UPOV Technical Guidelines).

Descriptors included under characterization can be considered as the basis for taxonomic classification, since they are mostly related to botanical characteristics. Nevertheless, some of them have agro-economic significance as well, for example colour of mango fruit is an important market trait.

Characterization descriptors pertain to those traits that tend to be highly heritable traits (i.e. traits that do not change with different environments, that means they show none or low genotype × environment interactions). The characters scored are visible to the naked eye, allow for quick and easy discrimination between accessions, and are generally controlled by major genes. They often provide additional information that assists in the identification or maintenance of the material (e.g. growth habits, leaf shapes, seed shapes).

Characterization descriptors may also include a limited number of additional traits considered to be desirable by a consensus of users of a particular crop.

3.1.5 Evaluation category

Evaluation descriptors are of great interest to plant breeders and are useful in crop improvement and the domestication of new crops. They include descriptors such as yield, agronomic and other economically important traits, biochemical traits (content of specific chemical compounds, dry matter content, etc.), and reaction to biotic and abiotic stresses. It should be noted that, until 2004, crop descriptors included a biochemical and molecular section describing the basic methods most commonly used. With the release of the list of descriptors for *Genetic Markers Technologies* (see http://www.bioversityinternational.org/Publications/pubfile.asp?id_pub=913) in February, 2004, those sections are no longer included in the crop-specific descriptors. The user is encouraged to follow the most recent list of descriptors for markers published on the Bioversity Web site.

The expression of many characteristics in the evaluation category is subject to genotype and environment interactions and usually shows high genotype × environment interactions reflecting the influence of the environment in which they are grown on the expression of gene(s), and are usually multigenic, involving minor genes, (where the genetic control of a trait results in the phenotypic expression varying from place to place and over seasons and years).

Observation of these descriptors may involve the use of simple to complex techniques and equipment.

To score evaluation descriptors, it is often necessary to use appropriate experimental designs and statistical analyses trials that are much more complex and resource intensive than is necessary for characterization. Since evaluation descriptors are influenced by environmental conditions, curators, breeders and researchers conduct replicated trials over years to obtain objective results for these traits. Evaluation is generally carried out as part of a breeding programme and in collaboration with other crop improvement scientists, such as pathologists and entomologists, and also with farmers (various participatory methods of crop improvement), where germplasm undergo evaluation for specific traits. Ideally, these results are then fed back to the originating genebank in order to complement existing data.

Descriptors listed under biotic stress susceptibilities (pests and diseases) should include both specific and common names.

Example #16

Fungi

Alternaria sesami	Leaf spot and blight
Cercospora sesami	Leaf spot
Colletotrichum spp.	Anthracnose

Recently published descriptor lists are comprehensive, with asterisked characterization and evaluation descriptors indicating the minimum set of highly discriminating characteristics that should be recorded for each accession. Asterisked descriptors are particularly useful indicators of diversity in collections and for international harmonization of documentation systems.

4. Similarities to and differences from other technical guidelines

Over the years, different guidelines for plant genetic resources documentation have been developed by UPOV, COMECON, USDA-GRIN and others, in addition to those developed by Bioversity and its predecessors. In addition, several national programmes have been developing descriptors for crops of national interest and for which internationally accepted lists were not available. Below is a short summary and description of the more commonly used descriptor lists.

4.1 Union Internationale pour la Protection des Obtentions Végétales

The Technical Guidelines developed by the Union Internationale pour la Protection des Obtentions Végétales (UPOV) have been developed specifically for testing the distinctness, uniformity and stability (DUS) of new cultivars of crops (UPOV 1989, 1993). DUS traits are central to the breeder's work since they are necessary to obtain legal protection for a bred variety.

UPOV descriptor lists are constructed with the thoroughness of legal documents. Requirements for the minimum amount of seed, number of vegetation periods, minimum number of plants, and maximum number of aberrant plants are defined. Precise rules for scoring are given, along with example varieties for each trait and level of manifestation. Many countries have adopted the UPOV guidelines for identifying and registering new plant varieties.

UPOV lists contain data that correspond to Bioversity categories of characterization and preliminary evaluation. The objective and interpretable scoring of traits is a clear advantage of this system (van Hintum et al. 1995). Nevertheless, use of the UPOV Technical Guidelines by the plant genetic resources community is limited due to the high number of standard cultivars used and the fact that these standards are based on modern cultivars, making comparison with exotic material or wild species difficult.

4.2 Council for Mutual Economic Aid

In 1977, the member countries of the Council for Mutual Economic Aid (COMECON) joined forces to develop descriptor lists for crops of primary economic importance. By 1990, 48 bilingual (Russian/English) descriptor lists had been published.

The passport category of these descriptor lists contained 13 fields. In addition to descriptive data, the characterization category (six fields) contained detailed geographical information on the location of collections in eight COMECON countries: Bulgaria, Cuba, Czechoslovakia, Hungary, Mongolia, Poland, Romania and USSR.

The characterization descriptors included data on morphology, biology, disease and pest resistance, chemical composition, economic utilization and other descriptors; botanical keys were also included.

An international databank was planned that would be accessible to all plant germplasm users within the framework of COMECON activities. After the abolition of COMECON, the N.I. Vavilov All-Russian Scientific Research Institute of Plant Industry (VIR) used the experience to develop a databank for the worldwide collections conserved at the institute.

4.3 United States Department of Agriculture Genetic Resources Information Network (USDA-GRIN)

The US National Plant Germplasm System (NPGS) has developed descriptor lists for many major food plants. Descriptor lists allow NPGS curators to enter plant trait data into the Genetic Resources Information Network (GRIN) database (see USDA-GRIN, no date, in references).

These lists match the Bioversity International format for characterization and evaluation categories and for use of the descriptors. They are linked to accession numbers, but do not contain descriptors for passport data (accession and collecting). They also omit management, site and environment data.

Gathering and analysing data

When defining descriptors, there are a range of criteria that can be applied to select the most appropriate or practical descriptors. Among these, the following should be considered:

- Type of data.
- Type of scales.
- Resolution, complexity and cost.
- Methods for collecting data.
- Experimental design.
- Scoring, coding and recording of descriptors.
- Numerical versus alphanumeric coding schemes.

5. Types of data

When determining how a certain characteristic or trait should be scored (that is, the different states of expression a descriptor can take), it is useful to classify them into two broad categories:

- qualitative data; or
- quantitative data.

It is sometimes desirable or necessary to convert or transform the data between the two categories, such as for statistical analysis purposes. However, it should be noted that when quantitative data are converted to qualitative data, commonly used information items such as shape, colour and texture are lost, yet these are the most commonly used descriptors. Studies conducted on scoring show that scoring and conversion methods using qualitative data may be problematic and can create bias (except for experts). As such, readers should be cautious when converting data.

5.1 Qualitative data

Qualitative data, such as type of sample received, are not computable by arithmetic calculations and are expressed in discontinuous states. They are self-explanatory and independently meaningful labels or names that determine the class or category in which an individual, object or process falls. All possible states are necessary to describe the full range of a characteristic, and every form of expression can be described by a single descriptor state; the order of states is not important.

Example #17

Type of material received

1	Seed
2	Seedling
3	Fruit
4	Shoot
5	Pollen

For some *qualitative* descriptors, such as colour descriptors, it is important to know whether:

- a finite number of states exist;
- all states need to be separately recorded; or
- all states can be ranked in a meaningful way that will merge a group of states under one name.

In other cases, the range of expression is at least partly continuous, but varies in more than one dimension:

Example #18

Leaf colour

1	Green
2	Yellow
3	Brown
4	Bluish green

Example #19

Seed coat texture

1	Smooth	
2	Partially rough	
3	Radially rough	
4	Partially radially rough	
5	Reticulately rough	
6	Partially reticulately rough	
99	Other	(specify in descriptor Remarks)

In the examples above, where the level of detail can be open to interpretation by different users and can complicate future statistical analysis, it is recommended to carefully select the most representative states or include colour chart codes, reference standards or drawings, as in the examples below.

Including the most representative states of expressions and splitting the descriptor:

In some situations, it may also be possible to split the qualitative characteristic into a *qualitative* and a *pseudo-quantitative* characteristic, such as in the following situation:

Example #20

Flower colour	
1	Light yellow
2	Medium yellow
3	Dark yellow
4	Green
5	Light pink
6	Medium pink
7	Dark pink

It is advisable to split this descriptor into the following characteristics:

Example #21

Flower colour	
1	Yellow
2	Green
3	Pink
Intensity of flower colour	
3	Weak (Light)
5	Medium
7	Strong (Dark)

Including RHS colour chart codes:

Example #22

Leaf colour		**(RHS colour chart code)**
1	Green yellow	(145-B)
2	Yellow green	(150-A)
3	Green	(128-A)
4	Bluish green	(120-B)
5	Dark green	(135-B)

Including reference standards:

Example #23

Leaf colour		**(Reference standard)**
1	Green yellow	Deglet nour
2	Yellow green	Mejhoul
3	Green	Besser Helou
4	Bluish green	Ammari, Menakher
5	Dark green	Ghars

Including drawings:

Example #24

Seed coat texture		**(See Figure 3)**
1	Smooth	
2	Partially rough	
3	Radially rough	
4	Partially radially rough	
5	Reticulately rough	
6	Partially reticulately rough	
99	Other (specify in descriptor Remarks)	

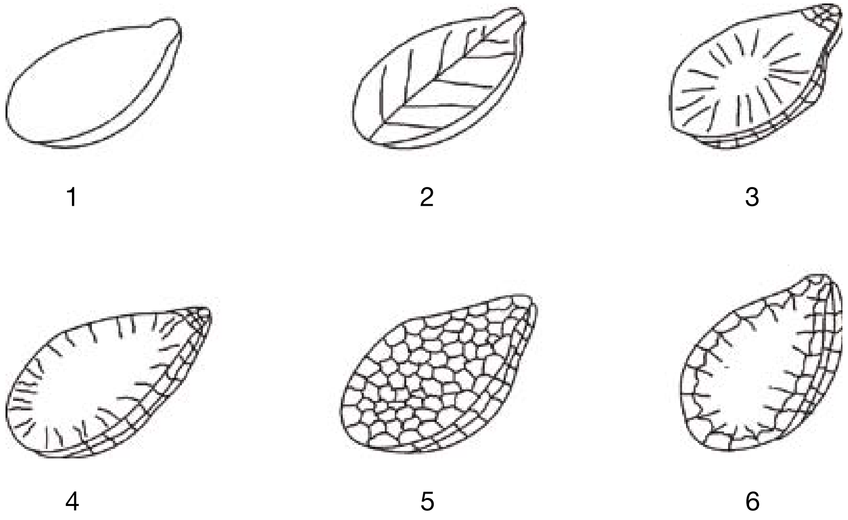

Figure 3. Seed coat texture.

Decisions on whether descriptor states need to be recorded separately or can be merged into a group will depend on the importance of the state in describing (within and between accessions), the diversity of the crop or gene pool. In general, the descriptor states of qualitative characteristics are given consecutive numbers starting with '1' and often have no upper limit.

With emerging imaging techniques, it is possible to convert the qualitative data (where the range of expression is continuous) to quantitative data, such as in colour and texture descriptors. In this case, the data can also be analysed quantitatively.

5.2 Quantitative data

Quantitative data consist of measures or counts that use numerical values, allowing statistical analyses, for which descriptions such as means and standard deviations are meaningful.

Quantitative descriptors are those in which the expression covers the full range of variation from one extreme to the other. Different states of expression of quantitative data can be recorded using *discrete* (countable data, such as "number of plants"), or *continuous* (measurable data, such as plant height, weight, length) scales.

Many quantitative characters that are continuously variable are recorded on a 1 to 9 scale, in which '1' stands for 'very short' or 'very low', and '9' corresponds to the highest expression such as 'very high' or 'very long' (see Section 10).

Nevertheless, the use of exact measurements is highly recommended, especially for easily measurable characteristics such as length or width, because there is no room for subjective interpretation. The use of 'classes' can prevent or complicate considerably statistical analysis of data. Such analysis is becoming increasingly important as many genebanks are now focusing on the use of conserved germplasm and on measuring within-accession (within-population) variation. It is easy to re-classify after measurements have been made, but it is not possible to translate those classes back to the exact data.

Sometimes only a selection of the states is described (e.g. 1, 3, 5, 7 and 9) for such descriptors. Where this occurs, the full range of codes is available for use by extension of the given codes or by interpolation between them.

Example #25

(A full scale)

Length of peduncle

1	Very short	(<3 cm)
2	Very short to short	(3-5 cm)
3	Short	(6-8 cm)
4	Short to intermediate	(9-10 cm)
5	Intermediate	(11-13 cm)
6	Intermediate to long	(11-13 cm)
7	Long	(14-16 cm)
8	Long to very long	(17-19 cm)
9	Very long	(>19 cm)

Example #26

(A reduced scale)

Length of peduncle

1	Very short	(<3 cm)
3	Short	(6-8 cm)
5	Intermediate	(11-13 cm)
7	Long	(14-16 cm)
9	Very long	(>19 cm)

6. Types of scales

There are several types of scales used for creating scoring methods. The use of a particular scale will depend on the type of data (qualitative or quantitative) to be recorded. The scale level will depend on the different states of expression of the characteristic, and on how they are recorded.

The most commonly used scales in the descriptors series are: nominal, ordinal (discrete), continuous and binary scales (See Figure 4).

The new techniques becoming available for converting qualitative to quantitative data based on computer-image analysis have the potential to exert a significant influence on selecting the type of scale to use in particular contexts.

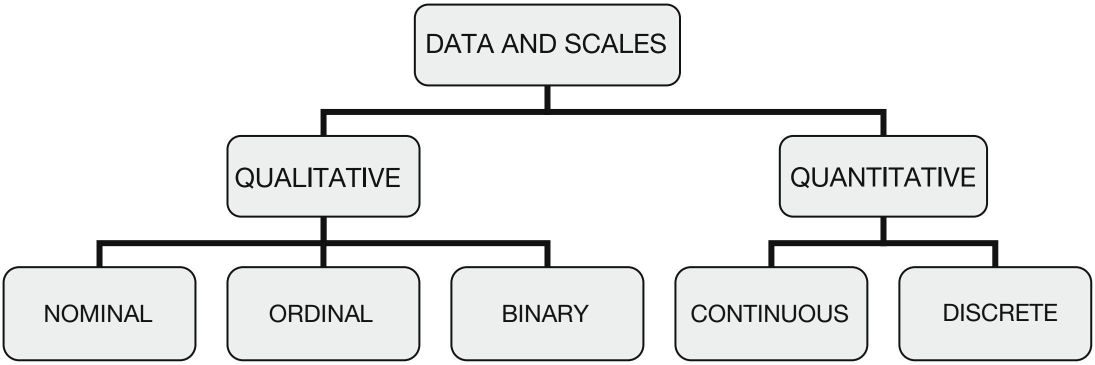

Figure 4. Types of data and scales.

6.1 Scales for qualitative characteristics

Qualitative characteristics can be measured using *nominal*, *ordinal* or *binary* scales. Within the continuum view, some descriptors, such as shape and texture, can be measured (quantitatively) using continuous scales. Their states of expression are often coded with sequential numbers, whereby special groups, such as ‘Others’, are given a special (’99’) value to set them apart from other descriptor states in order to accommodate new descriptor states if new germplasm is collected that exhibits a new form of a particular characteristic. The descriptor state ‘Others’ (listed last) is usually listed when it is presumed that other states may exist in other collections; further information may also be added in the ‘Remarks’ field. In the following example, room has been left under ‘99’ to accommodate further shapes, currently unknown but that might exist in other collections.

Example #27

Fruit shape

1	Round
2	Ovate
3	Oblong
4	Elliptic
99	Other (specify in the descriptor Remarks)

If a new shape is discovered in new germplasm collected (such as ‘obovate’), a new sequential code number should be assigned (see the example below, where ‘99’ might be used as the code for potential new shapes until they are confirmed).

Example #28

Fruit shape

1	Round
2	Ovate
3	Oblong
4	Elliptic
5	Obovate
99	Other (specify in the descriptor Remarks)

The states can be better defined by adding standards or check cultivars to the descriptor states to clarify the different states of expression of each trait.

6.1.1 Qualitative descriptors using nominal scale

Nominal scales provide code numbers for traits that are defined by text (names or labels). They do not follow a numerical or logical order or ranking sequence and the codes are arbitrary numbers (e.g. pubescence, colour, shape).

Example #29

Fruit colour

1	Yellow
2	Orange
3	Red
4	Brown
5	Purple
99	Other (specify in the descriptor Remarks)

Example #30

Leaf type

1	Tendril
2	Phyllody
3	Simple (lamina not bifurcated into leaflets)
4	Bipinnate
5	Multipinnate
99	Other (specify in the descriptor Remarks))

6.1.2 Qualitative descriptors using ordinal scale

These scales are similar to nominal scales, but have an order (e.g. data values are ranked in a numerically meaningful way). Ordinal scales rank traits from low to high. They result from visually assessed quantitative traits.

Example #31

Intensity of anthocyanin coloration

1	Low
2	Intermediate
3	Strong

Example #32

Plant growth habit

Recorded at the beginning of flowering period. (See Figure 5)

1	Erect
2	Semi-erect
3	Spreading
4	Prostrate

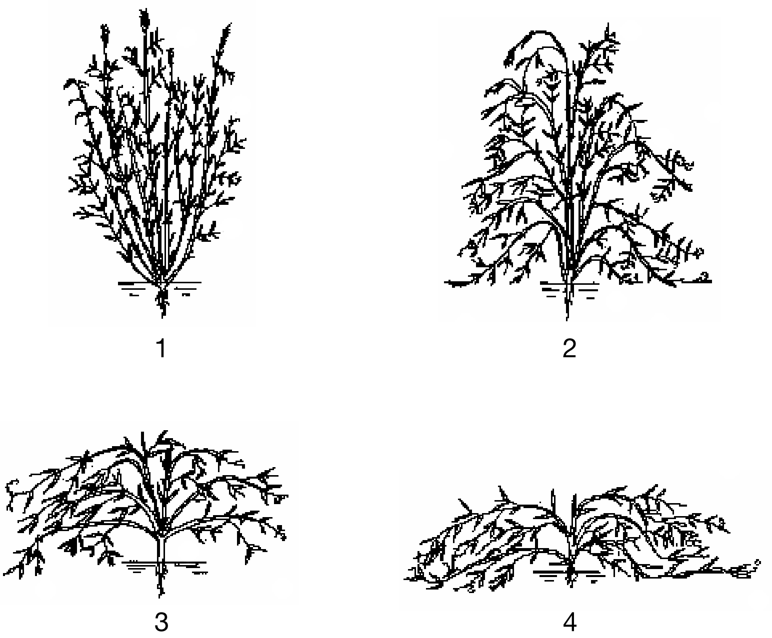

Figure 5. Plant growth habit.

6.1.3 Qualitative descriptors using a binary scale

Qualitative characteristics with only two categories (absent vs. present) are described by a special form of nominal scale. They can be scored on a binary scale (yes/no; absent/present) and the following standard coding should be used:

Example #33	
Leaf colour variegation	
0	Absent
1	Present

Example #34	
Fruit cracking	
0	No
1	Yes

Example #35	
Leaf glands	
0	Absent
1	Present

Here, '0' is used to indicate the absence of the characteristic or attribute, or that the trait is not observed.

6.2 Scales for quantitative characteristics

Quantitative characteristics are recorded by measuring, counting or weighing, and can be recorded using *continuous* or *discrete* scales.

6.2.1 Quantitative descriptors using a continuous scale

Continuous scales refer to the exact measurement of a trait and can assume an infinite number of real values. The best way of recording quantitative descriptors is by scoring the measurement in exact units, using the international unit system (Système International d'Unités – SI) (see Appendix II).

Example #36
Peduncle length [cm]

Quantitative data for such descriptors are measured on a continuous scale with well-defined units of measurement. As noted earlier, quantitative data recorded on continuous scales have a greater potential for statistical analysis than those recorded on discrete scales. Conversion from a continuous scale to a discrete scale is possible, but it is not possible to convert from a discrete scale to a continuous scale.

Example #37		
Peduncle length	30 cm	= very small
Peduncle length	very small	= ????

6.2.2 Quantitative descriptors using a discrete scale

Quantitative descriptors can be scored on a discrete scale and can take a finite or countable number of values. This is useful when one wants to use bar or pie charts to describe the variation found. This also allows the visualization of variation in a very simple way. This system is only used when precision is not required and the objective is only visualization of variation present. In these cases, a certain range of (continuous) values is grouped into discrete classes. These descriptor states, representing discrete classes, are a good measure for describing diversity within a crop or gene pool.

Example #38
Number of stolons

The discrete quantitative data of this descriptor are assessed by counting rather than measuring using a metric scale. An accession's value is determined by comparing it with a set of values assigned to that trait.

For some descriptors, the fact that they can be ordered from 'very low' to 'intermediate' to 'very high' is sufficient, without exactly defining the distances between the classes. However, these states are only possible after counting and knowing what is, for example, the minimum and maximum number of stolons.

In a common coding scheme for this type of scale, the descriptor state is scored between '1' (weakest expression) and '9' (strongest expression), as shown below. As a general rule, descriptor states are formed in such a way that for comparing expressions, a reasonable word pair is chosen (e.g. Weak/Strong; Short/Long; Small/Large).

Example #39

Quantitative descriptors using a discrete scale	
1	Very low
2	Very low to low
3	Low
4	Low to intermediate
5	Intermediate
6	Intermediate to high
7	High
8	High to very high
9	Very high

Sometimes this list is abbreviated by listing only states 3, 5, and 7. Where this has occurred, the full range of codes is available for use by extension of the codes given or by interpolation between them.

Example #40

Density of oil glands on fruit surface		**[no. per sq. cm]**
3	Low	(< 40/cm^2)
5	Intermediate	(50 – 80/cm^2)
7	High	(>90 cm^2)

The validity of exact measurements should be emphasized again because of their potential for data analysis and because transformation is minimal compared with discrete scale data.

There is also a 'limited' range of states comprising a 1 to 5 scale. It is used where the range of expression of a trait is physically limited at both ends and it is not appropriate to divide the expression into more than three intermediate states.

Example #41

Stem growth habit	
1	Erect
3	Semi-erect
5	Prostrate

When defining descriptors for different intensities of the same colour hue, the descriptor and descriptor states may be presented as pseudo-quantitative data (if they fulfil the requirements for a quantitative characteristic).

Example #42	
Intensity of green colour	
3	Light
5	Intermediate
7	Dark

Example #43	
Intensity of anthocyanin coloration	
3	Weak
5	Intermediate
7	Strong

Although this type of scheme allows a ranking of scores, it is not as precise and objective as measurement on a continuous scale. This is especially true when dealing with observations made by different observers, as individual bias leads to differences in scoring of traits.

More definition may be added by providing examples or standards to define each category in more detail, such as standard *references*.

Example #44		
Number of lobes in mature leaf		**Standard reference**
3	Few	Chardonnay
5	Intermediate	Chasselas blanc
7	Many	Hebron

7. Resolution, Complexity and Costs

7.1 Level of precision required

The preferred way of describing, for example, plant height, is by actual values measured in metres, cm or mm. It is essential to specify the unit of measurement in all cases.

Plant height can also be described in terms of discrete classes:

Example #45		
Plant height (range)		
3	Short	(<20 cm)
5	Intermediate	(30 – 50 cm)
7	Long	(>60 cm)

Quantitative data on a continuous scale has a greater potential for allowing statistical analysis than quantitative data measured on a discrete scale, and the use of exact measurements avoids differences in interpretation by different users. This does not mean that characteristics described using only discrete scale are less valuable; they are important in being diagnostic in nature, but could complicate future statistical analysis. The unit of measurement is also an indication of the level of resolution that is required.

If diversity of a specific trait (e.g. plant type) can be described by two very distinct states (dwarf type, tall type), a visual scoring of the two descriptor states could be sufficient and may be preferable to measuring every accession and recording the plant height in cm; this will save time and work.

It is good practice to keep observations and measurements as simple as possible. The objective of measurements is to determine how the trait of a specific accession compares with the diversity of the collection. When developing descriptors, one should remember that specialist knowledge and specialist equipment could be readily available at a particular institution, but this might not be the case for other institutions. In addition, methodologies that might be executed at one institution without problems, might present extraordinary logistical problems for institutions dealing with different combinations of crops or climatic environments (e.g. an institution working with one crop in comparison with a multicrop institution).

Where a term is open to interpretation, it is best to try to make a direct comparison with a well known standard or to use an absolute measurement. It is also essential to evaluate the trait in a number of randomly selected plants or a representative sample to ensure that the full range of variation present is described.

7.2 Complexity of the measurement or observation

The complexity of measurement or observation is dependent on extent of priming, special equipment or specific expertise required to execute a method of measurement. The more complex a procedure, the greater the chance of making mistakes during its execution, requiring greater care and attention to detail.

7.3 Cost per measurement

It is recommended that the costs for each measurement be carefully analysed, in terms of both staff time and materials. To have comprehensive minimum characterization data, the cost-effectiveness of observing descriptors is an important consideration. However, it must be noted that the value of data recorded is in the accuracy of its recording. Hence, it may be best to focus on a few priority measurements when funds are a major limitation.

8. Methods for collecting data

As a general rule, when developing descriptors, it is important to determine which order or method should be followed. As a rule of thumb, the order of descriptors should follow a botanical or a chronological order.

Botanical order

- Seedling (e.g. hypocotyl colour, pubescence)
- Plant (e.g. growth habit, crown shape)
- Root (e.g. shape, surface, flesh colour, system)
- Stem (length, pubescence)
- Leaf (blade, petiole, stipule)
- Inflorescence
- Flower (calyx, sepal, corolla, petal, stamen, pistil)
- Fruit
- Seed/grain

Chronological order (order of development)

- Seedling stage
- Vegetative phase
- Reproductive phase
- Pre-harvest
- Post-harvest

Within these alternative ordering methods, it is suggested to follow the most immediately visible characteristic, such as:

- Attitude
- Colour
- Shape
- Individual parts of the organ, such as base shape, apex shape and margin
- Height
- Length
- Width
- Other characteristics

9. Experimental design

For a widely applicable, uniform and more meaningful information system, the data need to be standardized, not just with regard to terminology but also in terms of measurement. This encompasses measurement techniques, data recording, units of measurement and encoding methods – all of which are relevant to the experimental design – when assessing the descriptors and their diversity, and will vary according to characterization or evaluation trials.

A useful guide for genebank managers who undertake evaluation trials on their genetic resources collections is found in Technical Bulletin No. 4 (IPGRI, 2001). The manual covers the stages involved in an experimental programme, from the determination of objectives for each trial, to the methods used for analysis, and provides general guidelines for managers to adapt to specific crops. It specifically focuses on evaluation of large collections.

The main issues to consider in the design of an experiment are:

- Set precise objectives (including background and justification for the proposed research).
- Experimental design.
- Decide on analysis strategy.
- Select treatments (number of accessions per trial, control treatments).
- Specify sites (plots and layout).
- Indicate levels of measurement (individual plant, plot or trial level).
- Collect data and analyse.

10. Scoring, coding and recording of descriptors

In defining individual descriptors, many aspects need to be considered. The following internationally accepted norms for the scoring, coding and recording of descriptor states are promoted worldwide by Bioversity International.

- The Système International d'Unités (SI) units should be used; the units to be applied for each descriptor are given in square brackets following the descriptor name.
- The Royal Horticultural Society (RHS) Colour Charts for colour descriptors is strongly recommended for all ungraded colour characters (the precise chart should be specified in the section where it is used). The observation of colour by eye may compromise accuracy in determining exact colours, depending on each individual user; a colour chart is a useful tool for assigning relevant colour codes to different colour states.
- For all quantitative descriptors, it is recommended to use actual measurements. Where resources are insufficient to take actual measurements, quantitative characters that are continuously variable can be recorded on a 1–9 scale, as follows:

Example #46	
Continuously variable attribute	
0	Absent
1	Very low
2	Very low to low
3	Low
4	Low to intermediate
5	Intermediate
6	Intermediate to high
7	High
8	High to very high
9	Very high

These descriptor states (or states of expression) and corresponding descriptor codes are provided to define each characteristic and to harmonize descriptions. Each state is allocated a corresponding numerical code for ease of data recording and for consistency in the production and exchange of the descriptions.

When a descriptor is scored using a 1–9 scale, such as in Example 46, '0' would be scored when the character is not expressed or a descriptor is not applicable. In the following example, '0' will be recorded if an accession does not have panicles:

Example #47		
Panicle number per plant		
3	Low	(5 – 10)
5	Intermediate	(15 – 20)
7	High	(25 – 30)

- Absence or presence of characters is scored as a simple binary 0 or 1:

Example #48	
Terminal leaflet	
0	Absent
1	Present

- Blanks are used for information not yet available.
- Dates should be expressed numerically in the format YYYYMMDD, where:
  - YYYY = 4 digits to represent the year

- MM = 2 digits to represent the month
- DD = 2 digits to represent the day

If the month and/or day are missing, this should be indicated with hyphens. Leading zeros are required (i.e. 197506--, or 1975----).The date format listed above follows the ISO international format for the representation of dates and times (see Bibliography and Further Reference Sources).

10.1 Recording heterogeneous data

Landraces and wild populations are not as uniform as commercial varieties. Many genebanks will therefore mainly handle accessions that are heterogeneous for many traits (each accession is not genetically uniform and contains a certain amount of variation). Recording the average or most frequently occurring state does not express the extent of variation nor its range within each accession. To a certain extent, this can complicate documentation, since special provision must be made to record the diversity. If statistical studies are foreseen, the best approach is to record actual measurements.

Nevertheless, several other approaches have been proposed to address this issue. In crop descriptor lists, references are made to different methodologies, as discussed below.

- For accessions that are not generally uniform for a descriptor (e.g. a mixed collection, or genetic segregation), the mean and standard deviation can be reported where the descriptor is continuous. Where the descriptor is discontinuous, several codes in order of frequency could be recorded.
- Another approach is the method developed by van Hintum (1993), which has the following rules:
  - record the scores in decreasing order of size; and
  - add an equals sign (=) after the score if there is only one fraction.

  In this system, homogeneous populations are followed by the '=' sign. For heterogeneous populations, the ratio between two adjacent fractions is taken. If between 1.5 and 5.0, one 'x' sign is placed between the two fractions; if the ratio is higher than 5.0, two 'x' signs are placed.
- The method developed by Sapra and Singh (Rana et al. 1991), proposes the use of numbering codes from 0 to 9 based on frequency encountered (1=very low, 9=very high, with the same scale used for quantitative characters):
  - Three codes are placed in decreasing order of frequency
  - The frequency codes are placed after each descriptor code

  - For homogeneous populations, '9' is placed after the first descriptor code and '0' (indicating absence) after the other two codes.

Example #49

Flower colour

1	White
2	Purple
3	Red

For a population with only white flowers, the scores would be 192030. For a population with few (very low) white flowers and many (very high) red flowers, the scores would be 381120.

For descriptors with single digit states (e.g. 1–9 scale), the systems for recording heterogeneity differ and are not completely compatible. At this time, it is unclear to what extent these systems are applied by genebanks.

- Another method, proposing how heterogeneity could be documented for specific traits of the crop, has recently been under development by A. Alercia and co-workers, and is outlined below.

  If an accession shows high variation between plots and within plants, such as the flower colour of *Lathyrus*, the method suggested is to use different columns for each colour, estimate for each plot the percentage of colours present, and record the average. Assign numbering codes for each colour in order of frequency (increasing order).

Example #50

Table 1. Flower colour

Plot No.	Accession No.	Yellow	White	Pink	Red
1	10123	50	10	15	25
2	10123	20	40	30	30
3	10123	30	20	25	25
4	10123	40	30	10	30
5	10123	70	10	10	10
Total		240	110	90	150
%		48%	22%	18%	30%

Assign numbering codes in increasing order using a 1–9 scale (low to high scale), as follows:

Example #51	
1	(18%)
2	(22%)
3	(30%)
4	(48%)

- The descriptor will now will appear as:

Example #52

Flower colour

1	Pink	(18%)
2	White	(22%)
3	Red	(30%)
4	Yellow	(48%)

- If, in addition, flower characteristics also show variation between them, it is recommended that they be recorded separately. For example, in Lathyrus species there is a huge variation between accessions, replications and even flowers in the same plot. When this occurs, it is recommended that flower element data be recorded separately:

Example #53	
(i)	Standard colour on upper side
(ii)	Standard colour on lower side
(iii)	Keel colour
(iv)	Wing colour
(v)	Vein colour

- Sometimes, variation may be found in the distribution of colour on standards. In this case, it is suggested to record the variation as follows:

Example #54

Distribution of colour in the standards

1	Colour present on margins
2	Colour present in the centre
3	Colour present along the veins

It is recommended that the Royal Horticultural Society Colour Chart be used for colour descriptors. However, in case RHS charts are not available or unaffordable (for example, small collections), using any other standard colour charts or generating one using computer tools is acceptable. When standards different from RHS charts are used, the standards used should be clearly mentioned along with information on colour and in some instances when the charts are uncommon, the chart itself may have to be provided along with the information.

11. Numeric versus alphanumeric coding schemes

In its descriptor lists, Bioversity promotes the use of numeric coding systems for descriptors rather than alphanumeric systems. The main reason for this choice is that crop descriptors are aimed at an international audience, and numerical codes do not need translation. Additionally, the use of numerical codes substantially facilitates the simple and accurate scoring of descriptors; updating and modification are easier and quicker as well.

The use of alphanumeric coding or short-hand codes may lead to confusion, making data retrieval and exchange extremely difficult (different users have different approaches for the coding and interpretation of data).

Example #55

Table 2. Numeric codes linked to multilingual descriptor states

Numeric Code	Fruit: apex	Fruit: sommet	Fruto: ápice
1	Indented	Déprimé	Hundido
2	Rounded	Arrondi	Redondeado
3	Pointed	Pointu	Puntiagudo
99	Other	Autre	Otro

With numeric scoring, different language versions score traits in a compatible and consistent way. This makes the use of numeric codes a convenient way to exchange data internationally.

Making descriptors work

12. The development process

Bioversity works closely with crop experts in the development of descriptor lists to ensure that the final product provides optimal support to all who work on particular crops.

Bioversity coordinates and manages the production of its descriptor lists through a number of activities, and the discussion below summarizes the process of developing a crop-specific descriptor list.

The drafting of crop descriptors is led by an expert, who is the main author of the characterization and evaluation descriptors. Bioversity then prepares a draft version applying its internationally accepted format for descriptor lists. A larger group of scientists from different countries is then invited to provide scientific advice and comments to ensure that the full extent of knowledge and expertise is reflected in the draft. At this stage, Bioversity scientists are also consulted for their technical comments and scientific advice. Relevant and substantiated comments provide input in developing the draft list to achieve a consensus document, after which the main author receives the text for final approval. Finally, the list is formatted for publication, and any necessary figures are drawn and positioned.

One of the main reasons for the success of the descriptor programme is that the process involves broad consultation, enabling Bioversity International to take into consideration comments from various regions and experts. Producing these lists to the required standard is time consuming, but cannot be compromised.

There are many aspects of crop descriptors for which the individual drafter's experience and knowledge are essential in preparing the first draft. This includes the selection of appropriate terminology, experimental design, the identification of characteristics and the selection of check cultivars, if required. In such situations, general guidance and experience are provided by Bioversity through a series of guidelines and steps, which are presented below.

13. A few basic rules

There are a number of options when someone is making a decision on which descriptors to include and how to define them. Below are a few basic guidelines:

- Keep the descriptors as simple as possible so that they are understood and can easily be used by a wide range of users.

- Use images and drawings to support textual descriptions and to clarify complex descriptors.
- Provide clear definitions of descriptors to enable others to apply them.
- Analyse carefully the unit costs per measurement (or set of measurements), in terms of both staff time and materials.
- Specify, when relevant, the unit of measurement.
- Avoid ambiguities. If a descriptor appears to be ambiguous, include references on methodology or standards. Colour is an attribute that benefits from an absolute comparator value, such as a standard colour chart value.
- Have the experimental design ready prior to assessing descriptors and their diversity.
- Consult widely among crop and genebank specialists in order to achieve a comprehensive and understandable list of descriptors that can be widely accepted.

Once the above simple rules have been followed, crop descriptor lists can function as tools to assist genebanks and other collection curators in documenting germplasm in a consistent manner and ensuring the continued use of the germplasm by the plant genetic resources community.

14. Step-by-step checklist for defining descriptors

Step 1*:* Investigate the range of diversity for the trait

- What kind and level of diversity is known for this trait?
- Is it relevant to discriminating accessions?

  Next Step:
  - If no diversity is known, no descriptor is necessary!
  - Otherwise, go to Step 2.

Step 2: Provide a name for the descriptor

- A descriptor name should be:
  - Descriptive
  - Unambiguous, and
  - Compact.
- Descriptor names are frequently composed of:
  - An object or item, plus a trait name.

Example #56

Accession number

Leaf colour

Soil fertility

Next Step:

– Go to Step 3

Step 3: Is the descriptor dealing with one or more traits?

- Avoid developing descriptors that describe more than one characteristic at the same time.

Example #57

Leaf-blade colour

1	White and striped
2	Pink and mottled

In this example, it would be better to split this descriptor up into 'leaf-blade colour' and 'leaf-blade variegation'.

Next Step:

– If more than one characteristic is involved, split the descriptor into separate descriptors, and go back to Step 2.

– If only a single characteristic is described, go to Step 4.

Step 4: Decide how to record the descriptor

- Define how the trait should be recorded (visual assessment, measured), what tools are needed and under what conditions it should be recorded.
- Add images if necessary to support the description.

Next Step:

– If the characteristic is qualitative, go to Step 5.

– If the characteristic is quantitative, go to Step 6.

Step 5: Define qualitative descriptors

- List the distinct descriptor states you want to use. Determine whether states can be merged or grouped.
- If possible, add references or standards that illustrate the different descriptor states.
- Number the states starting with '1'. If there is a need for space to list potential further states, add '99 Other'.
- Use image analysis to eventually convert qualitative data to quantitative data.

Next Step:

– Go to Step 9.

Step 6: Define quantitative descriptors

- For quantitative traits, determine whether the values should be scored on a continuous or a discrete scale.

Next Step:
- For continuous scales, go to Step 7.
- For discrete scales, go to Step 8.

Step 7: Use continuous scales

- Determine the level of resolution needed: whenever possible include the actual value of measurements and specify the units to be used (SI units).
  Next Step:
  - Go to Step 9.

Step 8: Define states on discrete scales

- Define the descriptor states you need and number these sequentially starting from '1'.
- States on a 1 to 9 scale (from weak to strong expression of the trait). Include if necessary more defined references or standards.
  Next Step:
  - Go to Step 9.

Step 9: References and standards used

- If reference is made to a specific method or a reference system, provide the relevant bibliographical references and give a full citation in an appendix.
- If check cultivars or reference varieties are included, they should be available and widely known.
  Next Step:
  - Go to Step 10.

Step 10: Is the descriptor highly discriminatory?

If the descriptor is discriminatory, mark it with an asterisk (*). This will indicate that this descriptor is particularly useful as an indicator of diversity in the collection, but also for international harmonization of documentation systems, if any.

Looking forward

The increasing attention given to molecular and biochemical characterization is reflected in the latest descriptor lists. Bioversity International recognizes the complementarity of innovative approaches with classical agro-botanical analysis. The List of Descriptors for Genetic Marker Technologies developed by IPGRI, now Bioversity International, in consultation with international experts, CGIAR Centres and research organizations was published in February 2004. Standardizing this information will facilitate the development of data exchange encoding formats, such as Extensible Markup Language (XML) and Document Type Definitions (DTD), for dissemination of information on markers and the creation of a global registry containing a full and accurate inventory of species-specific reference markers already published.

Because of the high demand from Bioversity's partners for descriptor standards, some ideas to expand the work on descriptor development are listed below:

- Develop standards for new areas (*in situ* collections, including farmer descriptors related to indigenous and traditional knowledge; forest species).
- Emphasize conservation for use ('conservation *and* use' of plants) instead of simply conservation.
- Develop descriptors that benefit people (such as the use of crops to combat desertification or soil erosion).
- Develop descriptors that benefit environment (such as the use of crops for bioenergy production).
- Apply the emerging imaging techniques for the description of complex traits, such as those of shape and texture.
- Future descriptor lists may also incorporate refinements based on the methodologies used by the biological ontology research community (Plant Ontology Consortium, see http://www.plantontology.org/), such as the assignment of a term accession identifier to each descriptor definition and placement of the terms into a structured ontology.
- Web-enabling the available descriptors.
- Electronic descriptor system. The next stage is to develop databases based on descriptors and on-line consultation in the development of new and revised descriptors. This is discussed below.

Electronic-descriptor system: a new development process

In addition to the electronic versions of descriptors (PDF and HTML files) listed on Bioversity's Web site, there is a new production process for descriptor development. Testing is underway, and a number of issues still need to be analysed.

In the future, this system will allow for a moderated real-time electronic consultation in descriptor development by an expert community, with instant feedback.

Some of this system's features include:

- on-line descriptor development;
- electronic delivery and reuse;
- representation of descriptors in XML;
- format extended to include database field definition and XML element name; and
- generation of electronic forms and database definitions.

The system will have an impact on:

- standardization of descriptor structures;
- harmonization of descriptors across crops;
- harmonization of descriptor states;
- support to communities of experts; and
- number of descriptors produced.

Bibliography and further references

Alercia A, Diulgheroff S, Metz T. 2001. List of Multi-crop Passport Descriptors. FAO/IPGRI, Rome. Available from: http://www.bioversityinternational.org/Publications/pubfile.asp?ID_PUB=124

Bartz AE. 1988. Some thoughts on measurement. In: Basic statistical concepts (3rd ed.). MacMillan, New York, NY, USA. pp. 1–21.

CAB International. 1999. Crop Protection Compendium. CD-ROM. CAB International, UK.

De Vicente C, Metz T, Alercia A. 2004. Descriptors for Genetic Markers Technologies. IPGRI, Rome, Italy. Available at: http://www.bioversityinternational.org/Publications/pubfile.asp?id_pub=913

FAO [Food and Agriculture Organization of the United Nations]. 1990. Guidelines for Soil Profile Description (3rd edition rev.). International Soil Reference Information Centre (ISRIC), Land and Water Development Division, FAO, Rome, Italy.

FAO. 2006. Guidelines for soil description. (4th edition, rev.). World Soil Resources Report. 110 p. FAO Doc. no. A0541/E. Available from: ftp://ftp.fao.org/docrep/fao/009/a0541e/a0541e00.pdf

FAO. 1996. Report on the State of the World's Plant Genetic Resources for Food and Agriculture, prepared for the International Technical Conference on Plant Genetic Resources, Leipzig, Germany, 17-23 June 1996. Food and Agriculture Organization of the United Nations, Rome, Italy.

FAO. No date. List of Institute Codes. These codes are available from http://apps3.fao.org/wiews/ for registered WIEWS users. From the Main Menu select: "PGR" and "Download". If new Institute Codes are required, they can be generated online by national WIEWS administrators, or by the FAO WIEWS administrator [Stefano.Diulgheroff@fao.org].

Plant Ontology Consortium. Available at: www.plantontology.org/

USDA-GRIN. No date. US Department of Agriculture Genetic Resources Information Network database. Available at: www.ars-grin.gov.

Harris JG, Harris MW. 1994. Plant Identification Terminology: An Illustrated Glossary. Spring Lake Publishing, Spring Lake, Utah, USA.

Henderson IF. 1989. Henderson's Dictionary of Biological Terms (10th ed.). Edited by E. Lawrence. Longman Scientific & Technical, Harlow, Essex, England, UK.

Greuter W, Mcneill J, Barrie FR, Burdet HM, Demoulin V, Filgueiras TS, Nicolson DH, Silva PC, Skog JE, Trehane P, Turland NJ, Hawksworth DL. (Editors & Compilers). 2000. International Code of Botanical Nomenclature (Saint Louis Code) adopted by the Sixteenth International Botanical Congress, St. Louis, Missouri, July–August 1999. Koeltz Scientific Books, Königstein. (Regnum Vegetabile, 138). See: http://www.bgbm.org/iapt/nomenclature/code/default.htm

IPGRI [International Plant Genetic Resources Institute]. 2001. The design and analysis of evaluation trials of genetic resources collections. A guide for genebank managers. IPGRI Technical Bulletin, No. 4.

ISO 8601:2000. Data elements and interchange formats – Information interchange – Representation of dates and times.

ISO 6709:1983 Standard representation of latitude, longitude and altitude for geographic point locations.

ISO 3166 Codes for the representation of names of countries. The ISO 3166-1: Code List can be found at: http://www.un.org/Depts/unsd/methods/m49alpha.htm

ISO 3166-3:1999. Codes for the representation of names of countries and their subdivisions – Part 3: Code for formerly used names of countries. This List is not available on line, but can be obtained from IPGRI [ipgri-mcpd@cgiar.org].

Kornerup A, Wanscher JH. 1984. Methuen Handbook of Colour. 3rd edition. Methuen, London, UK.

Laliberté B, Withers L, Alercia A, Hazekamp T. 1999. Adoption of Crop Descriptors – IPGRI. In: A Synthesis of Findings concerning CGIAR Case Studies on the Adoption of Technological Innovation. IAEG Secretariat, May 1999.

Munsell Color. 1975. Munsell Soil Color Chart. Munsell Color, Baltimore, MD, USA.

Munsell Color. 1977. Munsell Color Charts for Plant Tissues, 2nd edition, revised. Munsell Color, Macbeth Division of Kollmorgen Corporation, Baltimore MD, USA.

Painting KA, Perry MC, Denning RA, Ayad WA. 1995. Guidebook for Genetic Resources Documentation. IPGRI, Rome, Italy.

Purdy LH, Loegering WQ, Konzak CF, Peterson CJ, Allen RE. 1968. A proposed standard method for illustrating pedigrees of small grain varieties. Crop Science 8:405–406.

Rana R.S, Sapra RL, Agrawal RC, Rajeev Gambhir. 1991. Plant Genetic Resources. Documentation and Information Management. National Bureau of Plant Genetic Resources (Indian Council of Agricultural Research), New Delhi, India.

RHS [The Royal Horticultural Society]. 1966, 1986, 1995. R.H.S. Colour Chart (edn. 1, 2, 3). The Royal Horticultural Society, London.

SI Base Units (Système International d'unités). No date, See: http://www.bipm.fr/enus/3_SI/base_units.html

Stearn WT. 1995. Botanical Latin. Fourth Edition. David & Charles Publishers, Newton Abbot, UK.

UPOV [Union Internationale pour la Protection des Obtentions Végétales]. 1989. Guidelines for the Conduct of Tests for Distincness, Homogeneity and Stability. Triticale (×Triticosecale Witt.). Geneva, Switzerland.

UPOV. 1993. Draft Guidelines for the Conduct of Tests for Distinctness, Homogeneity and Stability. Wheat (Triticum aestivum L. emend. Fiori et Paol.). Geneva, Switzerland.

van Hintum ThJL. 1993. A computer compatible system for scoring heterogeneous populations. Genetic Resources and Crop Evolution 40:133–136.

van Hintum ThJL, Jongen MWM, Hazekamp Th. (editors). 1995. Standardization in Plant Genetic Resources Documentation. Report of the Second Technical Meeting of Focal Points for Documentation in East European Genebanks. Centre for Genetic Resources, The Netherlands (CGN), Wageningen, The Netherlands.

Appendix I – Contributors

Coordinators

Ms Adriana ALERCIA
Germplasm Information Specialist
Bioversity International HQ
Via dei Tre Denari 472/a
00057 Maccarese (Rome)
Italy
E-mail: a.alercia@cgiar.org

Dr Ramanatha RAO
Project Coordinator, Facilitating Use of Genetic Resources
Bioversity International
P.O.Box 236, UPM Post Office, Serdang
43400 Selangor Darul Ehsan
Malaysia
E-mail: r.rao@cgiar.org

Reviewers

Dr Abdallah BARI
Scientist, PGR Information/Data Management and Analysis Methodologies
CWANA
Bioversity International c/o ICARDA,
PO Box 5466, Aleppo
Syria
E-mail: a.bari@cgiar.org

Dr Jan ENGELS
Genetic Resources Management Advisor
Bioversity International HQ
Via dei Tre Denari 472/a
00057 Maccarese (Rome)
Italy
E-mail: j.engels@cgiar.org

Mr Samy GAIJI
Project Coordinator, UMBP
Bioversity International HQ
Via dei Tre Denari 472/a
00057 Maccarese (Rome)
Italy
E-mail: s.gaiji@cgiar.org

Mr Luigi GUARINO
Honorary Research Fellow
Secretariat of the Pacific Community, Suva
Fiji
E-mail: luigiG@spc.int

Ir Tom HAZEKAMP
Consultant
Bioversity International HQ
Via dei Tre Denari 472/a
00057 Maccarese (Rome)
Italy
E-mail: t.hazekamp@cgiar.org

Dr Annie LANE
Global Project Coordinator
Crop Wild Relative Project
Bioversity International HQ
Via dei Tre Denari 472/a
00057 Maccarese (Rome)
Italy
E-mail: a.lane@cgiar.org

Dr Stefano PADULOSI
Senior Scientist, Integrated Conservation Methodologies and Uses
Bioversity International HQ
Via dei Tre Denari 472/a
00057 Maccarese (Rome)
Italy
E-mail: s.padulosi@cgiar.org

Dr Paul Quek
Scientist, Documentation/Information
Bioversity International c/o Stesen Kuarantin Lepas Masuk
Jabatan Pertanian,Bangunan JKR (P) 1746
P.O.Box 236, UPM Post Office, Serdang
43400 Selangor Darul Ehsan
Malaysia
E-mail: p.quek@cgiar.org

Dr Xavier Schelderman
Scientist, Conservation and Use of Neotropical PGR
Bioversity International c/o CIAT
Apartado Aereo 6713, Cali
Colombia
E-mail: x.schelderman@cgiar.org

Appendix II – Standard reference sources

A2.1 References and standards used in developing crop descriptor lists

Category of descriptors	References
Passport	FAO/IPGRI List of Multi-crop Passport Descriptors (2001) Country names: — ISO 3166 Codes for the representation of names of countries, particularly the ISO 3166-1:1999 Code List — ISO 3166-3 Codes for the representation of names of countries and their subdivisions – Part 3: Code for formerly used names of countries. Institute codes: FAO codes should be used. These codes are available from http://apps3.fao.org/wiews/ for registered WIEWS users. If new Institute Codes are required, they can be generated online by national WIEWS administrators, or by the FAO WIEWS administrator. Date format – use ISO date format (ISO 8601).
Site and Environment	FAO. 1990. Guidelines for Soil Profile Description. FAO. 2006. Guidelines for Soil Description. Munsell Color. 1975. Munsell Soil Color Chart. UNESCO System for Classifying Vegetation See: http://www-eosdis.ornl.gov/source_documents/unesco.html
Characterization and Evaluation	Methuen Handbook of Colour (Kornerup A, and Wanscher JH. 1984) Munsell Color Charts for Plant Tissues (Munsell Color, 1977). Documentation and Information Management. Plant Genetic Resources. NBPGR (ICAR) Royal Horticultural Society Colour Chart (RHS, 1986) SI Units (Système Internationale d'Unités) (see http://www.bipm.fr/enus/3_SI/base_units.html) A computer compatible system for scoring heterogeneous populations (Van Hintum 1993). A proposed standard method for illustrating pedigrees of small grain varieties (Purdy et al. 1968).
Glossary of terms	Plant Identification Terminology; An Illustrated Glossary. (Harris and Harris 1994) Henderson's dictionary of biological terms- 10th edition. (Henderson 1998) Botanical Latin. (4th edition) (Stearn 1995)

A2.2 Base units – Système International d'Unités (SI)

Physical quantity	Base unit [symbol]
Length	metre [m]
Mass	gram [g]
Time	second [s]
Substance	mole [mol]
Temperature	kelvin [K] or degree Celsius [ºC]*
Electrical current	ampere [A]
Luminous intensity	candela [Cd]

Note: (1) *degree Celsius is common referred to as Centigrade (= K -273)
(2) These base units can be prefixed by factors such as kilo (k), centi (c) or milli (m) to scale the units.

Appendix III – FAO/IPGRI Multi-Crop Passport Descriptors

The list of multi-crop passport descriptors (MCPD) was developed jointly by IPGRI and FAO to provide international standards to facilitate germplasm passport information exchange. These descriptors aim to be compatible with IPGRI [Bioversity International] crop descriptor lists and with the descriptors used for the FAO World Information and Early Warning System (WIEWS) on plant genetic resources (PGR).

For each multicrop passport descriptor, a brief explanation of content, coding scheme and *suggested* field name (in parentheses) is provided to assist in the computerized exchange of this type of data. It is recognized that networks or groups of users may want to further expand this MCPD List to meet their specific needs. As long as these additions allow for an easy conversion to the format proposed in the multi-crop passport descriptors, basic passport data can be exchanged worldwide in a consistent manner.

General comments

- If a field allows multiple values, these values should be separated by a semicolon (;) without space(s), (e.g. Accession name: Rheinische Vorgebirgstrauben;Emma;Avlon).
- A field for which no value is available should be left empty e.g. Elevation). If data are exchanged in ASCII format for a field with a missing numeric value, it should be left empty. If data are exchanged in a database format, missing numeric values should be represented by generic NULL values.
- Dates are recorded as YYYYMMDD. If the month and/or day are missing, this should be indicated with hyphens. Leading zeros are required (i.e. 197506--, or 1975----).
- Latitude and longitude are recorded in an alphanumeric format. If the minutes or seconds are missing, this should be indicated with hyphens. Leading zeros are required.
- Country names: Three letter ISO codes are used for countries. The ISO 3166-1: Code List and the Country or area numerical codes added or changed are not available on-line, but can be obtained from IPGRI [ipgri-mcpd@cgiar.org]
- For institutes, the codes from FAO should be used. These codes are available from http://apps3.fao.org/wiews/ for registered WIEWS users. From the Main Menu select: 'PGR' and 'Download'. If new Institute Codes are required, they can be generated online by national WIEWS administrators, or by the FAO WIEWS administrator [at the time of writing: <Stefano.Diulgheroff@fao.org>].

- The preferred language for free-text fields is English (i.e. 'Location of collecting site' and 'Remarks').

MULTI-CROP PASSPORT DESCRIPTORS

1. Institute code (INSTCODE)
Code of the institute where the accession is maintained. The codes consist of the 3-letter ISO 3166 country code of the country where the institute is located plus a number. The current set of Institute Codes is available from the FAO website (http://apps3.fao.org/wiews/).

2. Accession number (ACCENUMB)
This number serves as a unique identifier for accessions within a genebank collection, and is assigned when a sample is entered into the genebank collection.

3. Collecting number (COLLNUMB)
Original number assigned by the collector(s) of the sample, normally composed of the name or initials of the collector(s) followed by a number. This number is essential for identifying duplicates held in different collections.

4. Collecting institute code (COLLCODE)
Code of the institute collecting the sample. If the holding institute has collected the material, the collecting institute code (COLLCODE) should be the same as the holding institute code (INSTCODE). Follows INSTCODE standard.

5. Genus (GENUS)
Genus name for taxon. Initial uppercase letter required.

6. Species (SPECIES)
Specific epithet portion of the scientific name in lowercase letters. The following abbreviation is allowed: 'sp.'

7. Species authority (SPAUTHOR)
Provide the authority for the species name.

8. Subtaxa (SUBTAXA)
Subtaxa can be used to store any additional taxonomic identifier. The following abbreviations are allowed: 'subsp.' (for subspecies); 'convar.' (for convariety); 'var.' (for variety); 'f.' (for form).

9. Subtaxa authority (SUBTAUTHOR)
Provide the subtaxa authority at the most detailed taxonomic level.

10. Common crop name (CROPNAME)
Name of the crop in colloquial language, preferably English (i.e. 'malting barley', 'cauliflower', or 'white cabbage')

11. Accession name (ACCENAME)
Either a registered or other formal designation given to the accession. First letter uppercase. Multiple names separated with semicolon without space. For example: Rheinische Vorgebirgstrauben;Emma;Avlon

12. Acquisition date [YYYYMMDD] (ACQDATE)
Date on which the accession entered the collection where YYYY is the year, MM is the month and DD is the day. Missing data (MM or DD) should be indicated with hyphens. Leading zeros are required.

13. Country of origin (ORIGCTY)
Code of the country in which the sample was originally collected. Use the 3-letter ISO 3166-1 extended country codes.

MULTI-CROP PASSPORT DESCRIPTORS

14. Location of collecting site (COLLSITE)
Location information below the country level that describes where the accession was collected. This might include the distance in kilometres and direction from the nearest town, village or map grid reference point, (e.g. 7 km south of Curitiba in the state of Parana).

15. Latitude of collecting site (see note 1) (LATITUDE)
Degrees (2 digits), minutes (2 digits) and seconds (2 digits) followed by N (north) or S (south) (e.g. 103020S). Every missing digit (minutes or seconds) should be indicated with a hyphen. Leading zeros are required (e.g. 10- - - -S; 011530N; 4531- -S).

16. Longitude of collecting site (see note 1) (LONGITUDE)
Degrees (3 digits), minutes (2 digits) and seconds (2 digits) followed by E (east) or W (west) (e.g. 0762510W). Every missing digit (minutes or seconds) should be indicated with a hyphen. Leading zeros are required (e.g. 076- - - -W).

17. Elevation of collecting site [masl] (ELEVATION)
Elevation of collecting site expressed in metres above sea level. Negative values are allowed.

18. Collecting date of sample [YYYYMMDD] (COLLDATE)
Collecting date of the sample where YYYY is the year, MM is the month and DD is the day. Missing data (MM or DD) should be indicated with hyphens. Leading zeros are required.

19. Breeding institute code (BREDCODE)
Institute code of the institute that has bred the material. If the holding institute has bred the material, the breeding institute code (BREDCODE) should be the same as the holding institute code (INSTCODE). Follows INSTCODE standard.

20. Biological status of accession (SAMPSTAT)
The coding scheme proposed can be used at 3 different levels of detail: either by using the general codes (in boldface) such as 100, 200, 300, 400, or by using the more specific codes such as 110, 120, etc.

100) Wild
    110) Natural
    120) Semi-natural/wild
200) Weedy
300) Traditional cultivar/landrace
400) Breeding/research material
    410) Breeder’s line
        411) Synthetic population
        412) Hybrid
        413) Founder stock/base population
        414) Inbred line (parent of hybrid cultivar)
        415) Segregating population
    420) Mutant/genetic stock
500) Advanced/improved cultivar
999) Other (Elaborate in REMARKS field)

21. Ancestral data (ANCEST)
Information about either pedigree or other description of ancestral information (parent variety in case of mutant or selection). For example, a pedigree ‘Hanna/7*Atlas//Turk/8*Atlas’ or a description ‘mutation found in Hanna’, ‘selection from Irene’ or ‘cross involving amongst others Hanna and Irene’.

MULTI-CROP PASSPORT DESCRIPTORS

22. Collecting/acquisition source (COLLSRC)
The coding scheme proposed can be used at 2 different levels of detail: either by using the general codes (in boldface) such as 10, 20, 30, 40, or by using the more specific codes such as 11, 12, etc.

- 10) Wild habitat
  - 11) Forest/woodland
  - 12) Shrubland
  - 13) Grassland
  - 14) Desert/tundra
  - 15) Aquatic habitat
- 20) Farm or cultivated habitat
  - 21) Field
  - 22) Orchard
  - 23) Backyard, kitchen or home garden (urban, peri-urban or rural)
  - 24) Fallow land
  - 25) Pasture
  - 26) Farm store
  - 27) Threshing floor
  - 28) Park
- 30) Market or shop
- 40) Institute, Experimental station, Research organization, Genebank
- 50) Seed company
- 60) Weedy, disturbed or ruderal habitat
  - 61) Roadside
  - 62) Field margin
- 99) Other (Elaborate in REMARKS field)

23. Donor institute code (DONORCODE)
Code for the donor institute. Follows INSTCODE standard.

24. Donor accession number (DONORNUMB)
Number assigned to an accession by the donor. Follows ACCENUMB standard.

25. Other identification (numbers) associated with the accession (OTHERNUMB)
Any other identification (numbers) known to exist in other collections for this accession. Use the following system: INSTCODE:ACCENUMB;INSTCODE:ACCENUMB;… INSTCODE and ACCENUMB follow the standard described above and are separated by a colon. Pairs of INSTCODE and ACCENUMB are separated by a semicolon without space. When the institute is not known, the number should be preceded by a colon.

26. Location of safety duplicates (DUPLSITE)
Code of the institute where a safety duplicate of the accession is maintained. Follows INSTCODE standard.

27. Type of germplasm storage (STORAGE)
If germplasm is maintained under different types of storage, multiple choices are allowed, separated by a semicolon (e.g. 20;30). (Refer to FAO/IPGRI Genebank Standards (1994) for details on storage type.)

- 10) Seed collection
  - 11) Short-term
  - 12) Medium-term
  - 13) Long-term
- 20) Field collection
- 30) *In vitro* collection (Slow growth)
- 40) Cryopreserved collection
- 99) Other (elaborate in REMARKS field)

28. Remarks (REMARKS)
The remarks field is used to add notes or to elaborate on descriptors with value 99 or 999 (=Other). Prefix remarks with the field name they refer to and a colon (e.g. COLLSRC:roadside). Separate remarks referring to different fields are separated by semicolons without space.

Note 1: To convert from longitude and latitude in degrees (º), minutes (‘), seconds (‘’) and a hemisphere (North or South; East or West) to decimal degrees, the following formula should be used:
dº m’ s”=h *(d+m/60+s/3600) where h= +1 for the Northern and Eastern hemispheres, and h= –1 for the Southern and Western hemispheres, i.e. 30º30’0” S= –30.5 and 30º15’55” N=30.265.

Appendix IV – List of standard descriptors for site environment

A4.1 Site environment

A4.1.1 Topography

This refers to the profile in elevation of the land surface on a broad scale. The reference is FAO (1990).

Numeric code	Descriptor state		
1	Flat	0 – 0.5%	
2	Almost flat	0.6 – 2.9%	
3	Gently undulating	3 – 5.9%	
4	Undulating	6 – 10.9%	
5	Rolling	11 – 15.9%	
6	Hilly	16 – 30%	
7	Steeply dissected	>30%	moderate elevation range
8	Mountainous	>30%	great elevation range (>300 m)
99	Other		specify in the appropriate section's Notes

A4.1.2 Higher-level landform (general physiographic features)

The landform refers to the shape of the land surface in the area in which the collecting site is located (adapted from FAO, 1990).

Numeric code	Descriptor state
1	Plain
2	Basin
3	Valley
4	Plateau
5	Upland
6	Hill
7	Mountain

A4.1.3 Land element and position

Description of the geomorphology of the immediate surroundings of the collecting site (adapted from FAO, 1990). (See Figure A4.1)

Numeric code	Descriptor state	Numeric code	Descriptor state
1	Plain level	17	Interdunal depression
2	Escarpment	18	Mangrove
3	Interfluve	19	Upper slope
4	Valley	20	Midslope
5	Valley floor	21	Lower slope
6	Channel	22	Ridge
7	Levee	23	Beach
8	Terrace	24	Beachridge
9	Floodplain	25	Rounded sumit
10	Lagoon	26	Summit
11	Pan	27	Coral atoll
12	Caldera	28	Drainage line (bottom position in flat or almost-flat terrain)
13	Open depression		
14	Closed depression	29	Coral reef
15	Dune		
16	Longitudinal dune	99	Other (specify in appropriate section's Notes)

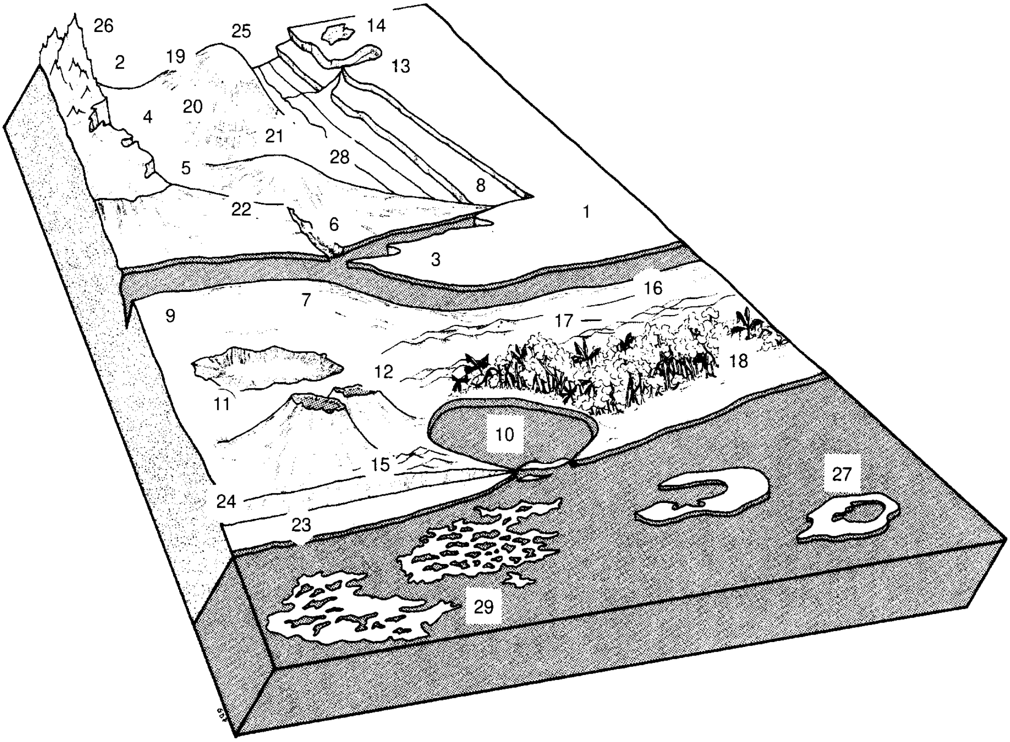

Figure A4.1. Land element and position

A4.1.4 Slope [°]

Estimated slope in degrees of the collecting site.

A4.1.5 Slope form

It refers to the general shape of the slope in both vertical and horizontal directions (FAO 1990).

Numeric code	Descriptor state
1	Straight
2	Concave
3	Convex
4	Terraced
5	Complex (irregular)

A4.1.6 Slope aspect

The direction that the slope on which the accession was collected faces. Describe the direction with symbols N, S, E, W (e.g. a slope that faces a south-western direction has an aspect of SW).

A4.1.7 Crop agriculture (Adapted from FAO 1990)

Numeric code	Descriptor state
1	Annual Field cropping
2	Perennial field cropping
3	Tree and shrub cropping

A4.1.8 Overall vegetation at and surrounding the collecting site (Adapted from FAO 1990)

Numeric code	Descriptor state	
1	Herbaceous	
	1.1 Grassland	
	1.2 Forbland	
2	Closed forest	Continuous tree layer, crowns overlapping, large number of tree and shrub species in distinct layers
3	Woodland	Continuous tree layer, crowns usually not touching, understorey may be present
4	Shrub	
5	Dwarf shrub	
99	Other	Specify in appropriate section's Notes

A4.1.9 Soil parent material

Two lists of examples of parent material and rock are given below. The reliability of the geological information and the knowledge of the local lithology will determine whether a general or a specific definition of the parent material can be given. Saprolite is used if the *in situ* weathered material is thoroughly decomposed, clay-rich but still showing rock structure. Alluvial deposits and colluvium derived from a single rock type may be further specified by that rock type.

A4.1.9.1 Unconsolidated material (Adapted from FAO 1990)

Numeric code	Descriptor state	Numeric code	Descriptor state
1	Aeolian deposits (unspecified)	10	Volcanic ash
2	Aeolian sand	11	Loess
3	Littoral deposits	12	Pyroclastic deposits
4	Lagoonal deposits	13	Glacial deposits
5	Marine deposits	14	Organic deposits
6	Lacustrine deposits	15	Colluvial deposits
7	Fluvial deposits	16	*In situ* weathered
8	Alluvial deposits	17	Saprolite
9	Unconsolidated (unspecified)	99	Other (specify in appropriate section's Notes)

A4.1.9.2 Rock type (Adapted from FAO 1990)

Numeric code	Descriptor state	Numeric code	Descriptor state
1	Acid igneous/ metamorphic rock	17	Dolomite
2	Granite	18	Sandstone
3	Gneiss	19	Quartzitic sandstone
4	Granite/gneiss	20	Shale
5	Quartzite	21	Marl
6	Schist	22	Travertine
7	Andesite	23	Conglomerate
8	Diorite	24	Siltstone
9	Basic igneous/ metamorphic rock	25	Tuff
10	Ultra-basic rock	26	Pyroclastic rock
11	Gabbro	27	Evaporite
12	Basalt	28	Gypsum rock
13	Dolerite		
14	Volcanic rock	99	Other (specify in appropriate section's Notes)
15	Sedimentary rock		
16	Limestone	0	Not known

A4.1.10 Stoniness/rockiness/hardpan/cementation

Numeric code	Descriptor state
1	Tillage unaffected
2	Tillage affected
3	Tillage difficult
4	Tillage impossible
5	Essentially paved

A4.1.11 Soil drainage (Adapted from FAO 1990)

Numeric code	Descriptor state
3	Poorly drained
5	Moderately drained
7	Well drained

A4.1.12 Soil salinity (dissolved salts)

Numeric code	Descriptor state
1	<160 ppm
2	161 – 240 ppm
3	241 – 480 ppm
4	481 – 800 ppm
5	>800 ppm

A4.1.13 Groundwater quality

Numeric code	Descriptor state
1	Saline
2	Brackish
3	Fresh
4	Polluted
5	Oxygenated
6	Stagnating

A4.1.14 Soil depth to groundwater table (Adapted from FAO 1990)

The depth to the groundwater table, if present, as well as an estimate of the approximate annual fluctuation, should be given. The maximum rise of the groundwater table can be inferred approximately from changes in profile colour in many, but not all, soils.

Numeric code	Descriptor state
1	0 - 25 cm
2	25.1 - 50 cm
3	50.1 - 100 cm
4	100.1 - 150 cm
5	> 150 cm

A4.1.15 Soil moisture

Moisture conditions prevailing in the soil at the time of collecting should be given together with the depth. Attention should be paid to unusual moisture conditions caused by unseasonal weather, prolonged exposure of the profile, flooding, etc. (from FAO 1990).

Numeric code	Descriptor state
1	Dry
5	Slightly moist
7	Moist
9	Wet

A4.1.16 Soil matrix colour (Adapted from FAO 1990)

The colour of the soil matrix material in the root zone around the accession is recorded in the moist condition (or both dry and moist condition, if possible) using the notations for hue, value and chroma as given in the Munsell Soil Color Charts (Munsell Color 1975). If there is no dominant soil matrix colour, the horizon is described as mottled and two or more colours are given and should be registered under uniform conditions. Early morning and late evening readings are not accurate. Provide depth of measurement [cm]. If a colour chart is not available, the following states may be used:

Numeric code	Descriptor state	Numeric code	Descriptor state
1	White	9	Yellow
2	Red	10	Reddish yellow
3	Reddish	11	Greenish, green
4	Yellowish red	12	Grey
5	Brown	13	Greyish
6	Brownish	14	Blue
7	Reddish brown	15	Bluish black
8	Yellowish brown	16	Black

A4.1.17 Soil organic matter content

Numeric code	Descriptor state
1	Nil (as in arid zones)
3	Low (as in long-term cultivation in a tropical setting)
5	Medium (as in recently cultivated but not yet much depleted)
7	High (as in never cultivated, and in recently cleared forest)
9	Peaty

A4.1.18 Soil pH

Actual pH value of the soil around the accession

A4.1.18.1 Root depth [cm]

Indicate the root depth at which the soil pH is being measured

A4.1.19 Soil erosion

Numeric code	Descriptor state
3	Low
5	Intermediate
7	High

A4.1.20 Rock fragments (Adapted from FAO 1990)

Large rock and mineral fragments (>2 mm) are described according to their abundance by soil volume.

Numeric code	Descriptor state
1	0 – 2%
2	2.1 – 5%
3	5.1 – 15%
4	15.1 – 40%
5	40.1 – 80%
6	> 80%

A4.1.21 Soil texture classes (Adapted from FAO 1990)

For convenience in determining the texture classes of the following list, particle size classes are given for each of the fine earth fraction listed below (Figure A4.2):

Numeric code	Descriptor state	Numeric code	Descriptor state
1	Clay	12	Coarse sandy loam
2	Loam	13	Loamy sand
3	Clay loam	14	Loamy very fine sand
4	Silt	15	Loamy fine sand
5	Silty clay	16	Loamy coarse sand
6	Silty clay loam	17	Very fine sand
7	Silt loam	18	Fine sand
8	Sandy clay	19	Medium sand
9	Sandy clay loam	20	Coarse sand
10	Sandy loam	21	Sand, unsorted
11	Fine sandy loam	22	Sand, unspecified

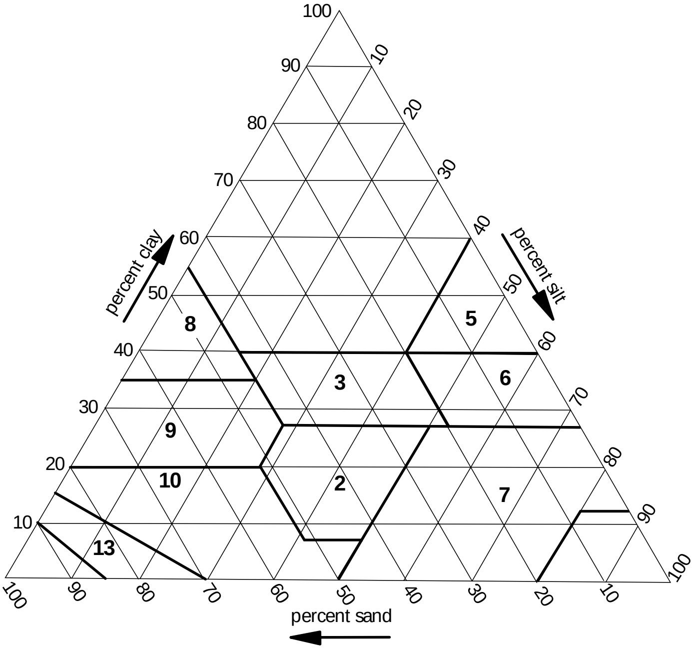

Figure A4.2. Soil texture classes

A4.1.22 Soil particle size classes (Adapted from FAO 1990)

Numeric code	Descriptor state	
1	Clay	<2 µm
2	Fine silt	3 – 20 µm
3	Coarse silt	21 – 63 µm
4	Very fine sand	64 – 125 µm
5	Fine sand	126 – 200 µm
6	Medium sand	201 – 630 µm
7	Coarse sand	631 – 1250 µm
8	Very coarse sand	1251 – 2000 µm

A4.1.23 Soil taxonomic classification

As detailed a classification as possible should be given. This may be taken from a soil survey map. State class (Alfisols, Spodosols, Vertisols, etc.).

A4.1.24 Water availability

Numeric code	Descriptor state
1	Rain-fed
2	Irrigated
3	Flooded
4	River bank
5	Sea coast
99	Other (specify in appropriate section’s Notes)

A4.1.25 Soil fertility

General assessment of the soil fertility, based on existing vegetation.

Numeric code	Descriptor state
3	Low
5	Moderate
7	High

A4.1.26 Climate of the site

Should be assessed as close to the site as possible (state number of recorded years)

- Temperature [ºC]
  Provide either the monthly or the annual mean.

- Rainfall [mm]
  Provide either the monthly or the annual mean (state number of recorded years).

- Wind
  Annual average (state number of years recorded)
  - Frequency of typhoons or hurricane force winds

Numeric code	Descriptor state
3	Low
5	Intermediate
7	High

  - Date of most recent typhoon or hurricane force wind [YYYYMMDD]
  - Annual maximum wind velocity [m/s]

- Frost
  - Date of most recent frost [YYYYMMDD]
  - Minimum temperature [ºC]
    Specify seasonal average and minimum survival temperature
  - Duration of temperature below 0ºC [days]

- Relative humidity
  - Relative humidity diurnal range [%]
  - Relative humidity seasonal range [%]

- Light

Numeric code	Descriptor state
1	Shady
2	Sunny

- Day length [hours]
  Provide either the monthly (mean, maximum, minimum) or the seasonal (mean, maximum, minimum).

Appendix V – Example Collecting form

COLLECTING FORM for *Allium* spp.

SAMPLE IDENTIFICATION

COLLECTING INSTITUTE(S) (2.1):

COLLECTING No. (2.2): PHOTOGRAPH No. (2.16):

COLLECTING DATE OF SAMPLE [YYYYMMDD] (2.3):

SPECIES (1.7): SUBTAXA(1.8):

COMMON NAME (1.11):
1. Dry bulb onion 2. Shallot 3. Japanese bunching onion/Welsh onion 4. Garlic 5. Leek
6. Kurrat 7. Great-headed garlic/elephant garlic 8. Chive 9. Rakkyo
10. Chinese chive/Oriental garlic/Nira 99. Other (specify)

COLLECTING SITE LOCATION

COUNTRY OF ORIGIN (2.4):

LOCATION (2.5): km: direction: from:

LATITUDE (2.6): LONGITUDE (2.7): ELEVATION (2.8): m asl

COLLECTING SITE ENVIRONMENT

COLLECTING / ACQUISITION SOURCE (2.9):
10. Wild habitat 20. Farm or cultivated habitat 30. Market or shop
40. Institute, Exp. Station, Research Org., Genebank 50. Seed company 60. Weedy, disturbed
99. Other (specify): or ruderal habitat

HIGHER LEVEL LANDFORM (6.1.2):

1. Plain 2. Basin 3. Valley 4. Plateau 5. Upland 6. Hill 7. Mountain

SLOPE [°] (6.1.4): SLOPE ASPECT (6.1.5; code N,S,E,W):

SAMPLE

BIOLOGICAL STATUS OF ACCESSION (2.12):
100. Wild 200. Weedy 300. Traditional cultivar/Landrace 400. Breeding/research material
500. Advanced/improved cultivar 999. Other (specify):

TYPE OF SAMPLE (2.13):
1. Vegetative 2. Seed 99. Other (specify)

NUMBER OF PLANTS SAMPLED (2.14): PREVAILING STRESSES (2.15.7):
Mention the types of major stresses, i.e. abiotic (drought), biotic (pests, diseases, etc.)

ETHNOBOTANICAL DATA

ETHNIC GROUP (2.15.1):

LOCAL/VERNACULAR NAME (2.15.2):

Plant uses (2.15.3)
1. Food uses
1.1 Raw salad
1.2 Fresh cooked
1.3 Stored/cooked/bottled/canned
1.4 Freezing
1.5 Pickling
1.6 Dehydrated
2. Medicinal
3. Ornamental
4. Forage
99. Other (specify)

PARTS OF THE PLANT USED (2.15.4)
1. Seed
2. Root/rhizome
3. Bulb/clove
4. Leaf sheath/pseudostem
5. Leaf
6. Scape
7. Flower/inflorescence
99. Other (specify)

CHARACTERIZATION

PLANT DESCRIPTORS
Foliage colour (7.1.1):
1. Light green
2. Yellow green
3. Green
4. Grey-green
5. Dark green
6. Bluish green
7. Purplish-green
99. Other (specify)

LEAF LENGTH [cm] (7.1.2):

FOLIAGE ATTITUDE (7.1.5):
3. Prostrate
5. Intermediate
7. Erect

CROSS-SECTION OF LEAF (7.1.7):
1. Circular
2. Semi-circular
3. Square
4. Pentagonal
5. V-shaped
99. Other (specify)

DEGREE OF LEAF WAXINESS (7.1.8)
3. Weak
5. Medium
7. Strong

SHAPE OF MATURE DRY BULBS (7.1.11):
1. Flat
2. Flat globe
3. Rhomboid
4. Broad oval
5. Globe
6. Broad elliptic
7. Ovate
8. Spindle
9. High top
99. Other (specify)

BULB SKIN COLOUR (7.1.15):
1. White
2. Yellow
3. Yellow and light brown
4. Light brown
5. Brown
6. Dark brown
7. Green (chartreuse)
8. Light violet
9. Dark violet
10. Mixed population
99. Other (specify)

NATURE OF STORAGE ORGAN (7.1.11)
1. Bulb, single large
2. Bulbs, several small
3. Rhizomes
4. Cloves
5. Foliage leaf bases
99. Other (specify)

INFLORESCENCE/FRUIT
Ability to flower (7.2.1):
0. No
1. Yes

GENERAL FERTILITY (7.2.5):
1. Sterile
2. Male sterile
3. Female sterile
4. Fertile

Flower number in umbel (7.2.6):
0. Absent
1. Few (<30)
2. Many (>30)

Date of 50% flowering [YYYYMMDD] (7.2.8):

COLLECTOR'S NOTES:

IPGRI and INIBAP
operate under the name
Bioversity International

Supported by the CGIAR

ISBN: 978-92-9043-729-1