



# DNA banks – providing novel options for genebanks?

**M.C. de Vicente and M.S. Andersson (editors)**



# **DNA banks—** providing novel options for genebanks?

**M.C. de Vicente and M.S. Andersson (editors)**

**The International Plant Genetic Resources Institute (IPGRI)** 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 Future Harvest 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. IPGRI 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 Improving Livelihoods in Commodity-based Systems.

The international status of IPGRI 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 IPGRI'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 IPGRI's Annual Reports, which are available in printed form on request from [ipgri-publications@cgiar.org](mailto:ipgri-publications@cgiar.org) or from IPGRI's Web site ([www.ipgri.cgiar.org](http://www.ipgri.cgiar.org)).

The System-wide Genetic Resources Programme (SGRP) joins the genetic resources programmes and activities of the Future Harvest Centres in a partnership whose goal is to maximize collaboration, particularly in five thematic areas. The thematic areas—policy, public awareness and representation, information, knowledge and technology, and capacity-building—relate to issues or fields of work that are critical to the success of genetic resources efforts. The SGRP contributes to the global effort to conserve agricultural, forestry and aquatic genetic resources and promotes their use in ways that are consistent with the Convention on Biological Diversity. IPGRI is the Convening Centre for SGRP. The Inter-Centre Working Group on Genetic Resources (ICWG-GR), which includes representatives from the Centres and the Food and Agriculture Organization of the United Nations, is the Steering Committee.

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 IPGRI, SGRP 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:** de Vicente MC and Andersson MS (eds). 2006. DNA banks—providing novel options for genebanks? Topical Reviews in Agricultural Biodiversity. International Plant Genetic Resources Institute, Rome, Italy

**Cover photo:** Preparing samples prior to storage in the tissue collection. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Cali, Colombia. Photo: Carlos Andres Tovar/IPGRI.

ISBN-13: 978-92-9043-702-4

ISBN-10: 92-9043-702-2

IPGRI

Via dei Tre Denari 472/a

00057 Maccarese

Rome, Italy

# Contents

<b>Foreword</b>	v
<b>Acknowledgements</b>	vii
<b>I. Introduction and overview</b> <i>M. Carmen de Vicente, Meike S. Andersson and Jan Engels</i>	1
<b>II. State of the art of DNA storage: results of a worldwide survey</b> <i>Meike S. Andersson, Eloina Mesa Fuquen and M. Carmen de Vicente</i>	6
<b>III. DNA storage as a complementary conservation strategy</b> <i>Ehsan Dulloo, Yoshiaki Nagamura and Oliver Ryder</i>	11
<b>IV. Platforms for DNA banking</b> <i>Christina Walters and Robert Hanner</i>	25
<b>V. The role of bioinformatics in coordinating conservation efforts</b> <i>Theodore Kisha and Oliver Ryder</i>	36
<b>VI. DNA banks: a primary resource for conservation research</b> <i>Nicole Rice, Robert Henry and Maurizio Rossetto</i>	41
<b>VII. Tissue collections as a means of storing DNA: a contribution to the conservation of Colombian biodiversity</b> <i>Juan Diego Palacio-Mejía</i>	49
<b>VIII. DNA banking of animal genetic resources</b> <i>Jinggong Xiangyu and Ya-ping Zhang</i>	56
<b>IX. Opportunities, limitations and needs for DNA banks</b> <i>Andreas W. Ebert, Jawahir L. Karihaloo and Márcio Elías Ferreira</i>	61
<b>X. A model for DNA banking to enhance the management, distribution and use of <i>ex situ</i> stored PGR</b> <i>Andreas Graner, Meike S. Andersson and M. Carmen de Vicente</i>	69
<b>Acronyms and abbreviations</b>	77
<b>List of authors</b>	81



# Foreword

The rapidly expanding study of DNA in so many areas of science has created an odd surplus: the DNA itself. Reasonably easy and inexpensive to store, with established techniques for almost infinite multiplication, the samples of DNA created in laboratories around the world have almost by accident become an important resource for future research. Participants at an IPGRI-organized meeting on The Evolving Role of Genebanks in the Fast-Developing Field of Molecular Genetics (see de Vicente MC, editor. The Evolving Role of Genebanks in the Fast-developing Field of Molecular Genetics. Issues in Genetic Resources No. 11. IPGRI, Rome, Italy; 2004) discussed informally how a more organized and collaborative approach to DNA banking might benefit the conservation and use of genetic resources and other areas of research. Those discussions prompted IPGRI to undertake a worldwide survey, which in turn resulted in this book.

As will be seen from the individual chapters, DNA banking is not at present widespread (fewer than 20% of the institutions responding to a survey bank DNA) but would be welcomed (more than half of the respondents expressed interest in banking DNA). The book also presents considerable evidence of the ways in which DNA banking, by permitting relatively easy transfer and analysis of genetic material, can improve genebank management, plant breeding, conservation, population studies, and many other subjects. Taken together, these suggest that there is indeed a need for practitioners to cooperate to develop the kind of collaborative system envisaged in this book.

Many issues remain to be sorted out. Some are extremely practical: how should DNA banks be physically arranged and should they store both DNA and the tissues from which it was extracted? How should DNA samples be distributed? Others hinge on technical issues: what are the best bioinformatic methods to tie together the various strands of information and the DNA samples from which the information was derived? Yet others involve high-level policy decisions: what rules and procedures will govern the exchange of DNA samples and sequence information, and how will access and benefit sharing, including the benefits of further research on the samples, be governed? By pulling together so many individual experiences and perspectives this book sets out several signposts that will help to guide further discussion and elaboration of DNA banking.

It should be stressed that none of the contributors sees DNA banking as a substitute for the conservation of genetic resources. DNA banks complement conservation strategies that make use of *ex situ* and *in situ* conservation, and they can help to ensure the optimal use of plant and animal populations. DNA banks are probably not going to permit the

resurrection of extinct populations. But their potential to enhance so many aspects of the use and conservation of genetic resources makes them an important topic for further discussion and elaboration, to which we hope this book will make a useful and distinct contribution.

*Emile Frison*

Director General

International Plant Genetic Resources Institute

November 2005

# Acknowledgements

The idea for this publication was conceived during an expert consultation meeting in November 2002 in León, Spain, organized by IPGRI for the System-wide Genetic Resources Programme (SGRP). The meeting brought together scientists from different CGIAR centres, national agricultural research programmes, universities and the private sector. IPGRI would like to thank all the participants of that meeting. Their names appear in 'The evolving role of genebanks in the fast-developing field of molecular genetics' (*Issues in Genetic Resources No. 11*, IPGRI, Rome, Italy).

As a result of this meeting, IPGRI decided to conduct a survey on DNA banking among international and national plant genetic resources research institutes, botanical gardens, universities and the private sector. IPGRI is grateful to all those who responded to the survey. The respondents provided significant input to this publication through comments that expressed their concerns and needs. Furthermore, the editors greatly appreciate the efforts of Coosje Hoogendoorn, who, believing in the relevance of this work for IPGRI, encouraged the conduct of the survey and the preparation of this publication.

IPGRI owes special thanks to the contributors of the chapters in this publication: E. Dulloo, A.W. Ebert, J. Engels, M.E. Ferreira, A. Graner, R. Hanner, R. Henry, J.L. Karihaloo, T. Kisha, E. Mesa F., Y. Nagamura, J.D. Palacio-Mejía, N. Rice, M. Rossetto, O. Ryder, C. Walters, J. Xiangyu and Y.P. Zhang.

Thanks are also due to IPGRI's Communications Services, which helped bring the publication into being, and Kathleen Matics, who helped with the production of this document.





# I. Introduction and overview

M. Carmen de Vicente<sup>1</sup>, Meike S. Andersson<sup>1</sup> and Jan Engels<sup>2</sup>

<sup>1</sup>International Plant Genetic Resources Institute (IPGRI-Americas), Cali, Colombia

<sup>2</sup>International Plant Genetic Resources Institute (IPGRI), Rome, Italy

## Background

Over the last thirty years or so, awareness of the important value of genetic diversity has contributed significantly to the conservation of plant genetic resources (PGR). The strength and weaknesses of the individual *ex situ* and *in situ* approaches have ensured that both approaches are widely applied.

*In situ* (=on-site) conservation refers to the maintenance of plant populations in the habitats where they naturally occur and evolved. In the case of crop species, conservation can be achieved in home gardens or on-farm, and in the case of wild species through the designation of protected areas such as national parks and nature reserves.

*Ex situ* (=off-site) conservation of germplasm takes place outside the natural habitat or outside the production areas, in facilities specifically created for this purpose. Depending on the biological nature of the species to be conserved, different *ex situ* conservation methods may be appropriate. The storage of PGR as seeds in genebanks at subzero temperatures is the most widely applied method. Approximately 90% of more than 6 million accessions stored *ex situ* worldwide are maintained as seeds (FAO 1996; Engelmann and Engels 2002). Other methods include maintenance of whole plants in field genebanks, as tissue culture (different tissues) in *in vitro* genebanks, and immersion of tissues, embryos or seeds in liquid nitrogen (i.e. 'cryopreservation'). The latter two methods are particularly used for the conservation of vegetatively propagated species and for species with recalcitrant seeds.

All the conservation approaches and methods mentioned above have advantages and disadvantages; for this reason, a complementary approach to safely conserve the genetic diversity belonging to the gene pool of a given crop is strongly recommended (Bowen 1999). Information is available about the seed storage behaviour of more than 7000 species (Hong et al. 1998), and a number of methods for complementary conservation schemes are well developed for a wide range of species. A variety of manuals, guidelines and protocols have been published providing technical recommendations and information about opportunities and limitations (e.g. Cromarty et al. 1982; Breese 1989; Stalker and Chapman 1989; FAO/IPGRI 1994; Engelmann 1999; Engelmann and Takagi 2000; Mohd Said Saad and Rao 2001).

PGR conservation is, however, not restricted to the collecting and conservation of germplasm. It also includes their efficient management and utilization and it is necessary to document, characterize and evaluate the genetic variation available in gene pools of cultivated crops and their wild relatives in order to maximize the use of these genetic resources in crop breeding (FAO 1996).

Besides the traditional documentation and characterization/evaluation techniques, some biotechnologies offer important tools for the effective conservation and use of PGR, particularly molecular markers. Their usefulness for the management of genebanks has been widely applied and discussed, e.g. for the characterization of genotypes, assessment of genetic diversity, estimation of genetic relationships within collections, identification of duplicates, establishment of core collections as well as monitoring genetic stability and integrity (e.g. Anderson and Fairbanks 1990; Hodgkin and Debouck 1992; Karp et al. 1997; Karp 2002; de Vicente 2004; de Vicente et al. 2004).

Molecular marker technologies are based on DNA analysis and total DNA as crude extracts will be generated for all target samples. Moreover, samples may also remain as amplification products as a result of applying technologies based on the polymerase chain reaction, e.g. random amplified polymorphic DNA (RAPD) or amplified fragment length polymorphism (AFLP). Other technologies, prior to the identification of the markers, require the construction of libraries, i.e. collections of segments of DNA (genomic, c-DNA) containing several copies of the genome of a given species. It is now routine to keep leftover DNA samples or libraries, even after a particular research project has been finished, because they constitute a valuable reserve for research and are a way to continue capitalizing on the investment (time and funds) already made to develop knowledge and/or products. DNA resources can be maintained at  $-20^{\circ}\text{C}$  for short- and mid-term storage (i.e. up to 2 years), and at  $-70^{\circ}\text{C}$  or in liquid nitrogen for longer periods. Although these DNA storage activities generally are not planned, being spin-off activities of varied ongoing projects, this situation has prompted the consideration of DNA collections as a genetic resource (Ramanatha Rao and Riley 1994).

In some instances, however, the establishment of DNA 'banks' may be a planned activity and/or the main objective of a research project. Then, DNA storage is undertaken with a specific objective, e.g. the extraction of genetic material to be made readily available for molecular applications. These experiments normally aim to obtain knowledge to improve the efficiency of some conservation activities or to scientifically inform decisions related to the conservation of germplasm. Other objectives of the creation of DNA banks may be related to training or distribution to scientists with an interest in different areas of biology.

In none of these cases however, are DNA banks assembled as a means to replace traditional methods of conserving genetic resources. This is important to note as conservation of genome fragments or individual genes is quite a different situation from the conservation of entire genotypes, as living organisms, for their future use.

## Existing DNA banks

One of the oldest DNA collections of its kind and, to the best of our knowledge, the only DNA bank with an exclusive conservation objective, is the 'Frozen Zoo<sup>®</sup>', hosted by the Zoological Society of San Diego. The Frozen Zoo<sup>®</sup> was initiated over 25 years ago (Benirschke 1984) and today preserves the genetic material of approximately 7000 endangered or threatened animal species (mammals, birds and reptiles) in the form of frozen tissue samples for the extraction of DNA for its subsequent use in conservation research ([http://www.sandiegozoo.org/conservation/zooproject\\_frozenzoo.html](http://www.sandiegozoo.org/conservation/zooproject_frozenzoo.html)).

Among the few institutions assembling and documenting PGR DNA specimens, the following maintain the most important collections:

- The Royal Botanic Gardens, Kew, UK currently house the world's largest and most comprehensive PGR DNA bank, consisting of over 20 000 DNA specimens representative of all plant families (<http://www.kew.org/data/dnaBank/homepage.html>).
- The US Missouri Botanical Garden holds a collection of 20 000 plant tissue samples with the exclusive objective of supplying researchers with the raw material for the extraction of DNA for its subsequent use in conservation research. A database of the available taxa is accessible through ([http://www.mobot.org/MOBOT/research/applied\\_research/](http://www.mobot.org/MOBOT/research/applied_research/)).
- The Australian Plant DNA Bank was launched at the Southern Cross University in June 2002 to conserve representative genetic information from the entire Australian flora (<http://www.dnabank.com.au/>).
- The DNA bank at the Leslie Hill Molecular Systematics Laboratory of the National Botanical Institute (NBI) in Kirstenbosch, South Africa, in collaboration with the Royal Botanic Gardens

Kew, preserves genetic material of the South African flora (443 out of 1932 genera already in the DNA bank) (<http://www.nbi.ac.za/research/dnabank.htm>).

- The DNA bank at the National Institute of Agrobiological Sciences (NIAS), Ibaraki, Japan, together with the Ministry of Agriculture, Forestry and Fisheries (MAFF) in Tokyo, is responsible for the preservation of DNA and molecular information of agricultural organisms, and presently maintains DNA samples from rice and pig specimens (<http://www.dna.affrc.go.jp/bank/index.html>).

## Rationale for this publication

The groundwork for this publication was a discussion held among germplasm experts on the future role of genebanks in the modern era of molecular genetics and genomics (de Vicente 2004). The primary objective of genebanks is and will continue to be the conservation of plant genetic resources for their future use in plant improvement, to contribute to sustainable agriculture systems, and ultimately thereby helping to eradicate hunger and poverty.

Some of the new applications of molecular biology tools and methods contribute to the routine conservation operations of germplasm in genebanks. However, these technologies are most important as the basis of major advances in genomics, and genebanks and their products are attracting a considerable number of new clients, such as molecular biologists and geneticists alongside the traditional plant breeders. An important challenge for genebanks is to satisfy the needs of this wide range of users, usually with a great diversity of expectations.

These considerations were the basis for circulating a questionnaire to gather up-to-date information on PGR DNA banking activities and outlooks (a summary is presented in Chapter II). The results of this survey stimulated the preparation of this publication with the aim of establishing a point of departure rather than providing tailor-made answers. A wealth of valuable information has been assembled concerning different aspects related to DNA banking in general, and of PGR in particular.

Chapter III discusses the importance of DNA storage for the conservation of species, particularly for gene and genome conservation, and analyzing the advantages and disadvantages of DNA banking as part of an *ex situ* conservation strategy using the cases of the Frozen Zoo® (San Diego, USA) and of the rice genome research in Japan as examples.

Chapter IV reviews best practices and experiences with the preservation of DNA samples, recent advances in technical aspects related to DNA storage (DNA bar-coding), as well as issues that still need to be resolved.

The amount of data and information produced by molecular applications is growing exponentially and the rapid advances in molecular as well as in information management technologies require special attention to genetic resources documentation. Chapter V presents some of the bioinformatics applications already employed and calls for the implementation of a standardized, universal molecular database.

Chapter VI reviews the usefulness of DNA banks as a primary resource for conservation research, in view of the fact that they provide the raw material for molecular genetic applications that facilitate the assessment and analysis of diversity (e.g. genetic diversity, molecular evolution), thus combining research and conservation.

The following two chapters (VII and VIII) deal with DNA banking of other than PGR. DNA banking efforts for these other genetic resources sectors are, in many instances, far more advanced and, therefore, the experiences accumulated might be transferable and useful for PGR DNA storage. Chapter VII reviews the use of tissue collections as a means of storing DNA, based on the experiences of the Colombian National Tissue Collection, hosted by the Alexander von Humboldt Biological Resources Institute. Chapter VIII summarizes the experiences with

DNA banking as a tool for animal genetic resources conservation and demonstrates its utility for conservation research by describing recent studies on endangered animal species.

Chapter IX analyzes the particular opportunities, limitations and needs for DNA banking from the perspective of developing countries. This bears in mind that most of the unique genetic diversity has originated in these countries and can still be found there. At the same time, these countries frequently do not have the required capacity and resources to maintain this important heritage and to benefit from its use as developed countries do.

In the final chapter (Chapter X), a model is presented for linking DNA storage with traditional *ex situ* germplasm conservation in order to enhance the management, distribution and use of *ex situ*-stored PGR. The establishment of a network of DNA banks is proposed to be linked to the current global genebank system under the auspices of the FAO. This would build on the model network of working and reserve nodes as developed and proposed by Adams and Adams (1992). The chapter also includes capacity building aspects and the possible role of IPGRI.

## References

- Anderson WR and Fairbanks DJ. 1990. Molecular markers: important tools for plant genetic resources characterization. *Diversity* 6:51–53.
- Benirschke K. 1984. The frozen zoo concept. *Zoological Biology* 3:325–328.
- Bowen BW. 1999. Preserving genes, species, or ecosystems? Healing the fractured foundations of conservation policy. *Molecular Ecology* 8:S5–S10.
- Breese EL. 1989. Regeneration and multiplication of germplasm resources in seed genebanks: the scientific background. International Board for Plant Genetic Resources, Rome, Italy. 69 p.
- Cromarty AS, Ellis RH and Roberts EH. 1982. The design of seed storage facilities for genetic conservation. *Handbooks for Genebanks No. 1*. International Plant Genetic Resources Institute (IPGRI), Rome, Italy. 100 p.
- de Vicente MC, editor. 2004. The evolving role of genebanks in the fast-developing field of molecular genetics. *Issues in Genetic Resources No. 11*. International Plant Genetic Resources Institute, Rome, Italy.
- de Vicente MC, Metz T and Alercia A. 2004. Descriptors for genetic markers technologies. International Plant Genetic Resources Institute (IPGRI), Rome, Italy.
- Engelmann F. 1999. Management of field and *in vitro* germplasm collections. Proceedings of a consultation meeting, 15–20 January 1996, CIAT, Cali, Colombia. International Plant Genetic Resources Institute (IPGRI), Rome, Italy. 165 p.
- Engelmann F and Engels JMM. 2002. Technologies and strategies for *ex situ* conservation. *In: Engels JMM, Rao VR, Brown AHD and Jackson MT, editors. Managing plant genetic diversity*. CABI Publishing and International Plant Genetic Resources Institute (IPGRI), Wallingford, UK and Rome, Italy. pp 89–104
- Engelmann F and Takagi H. 2000. Cryopreservation of tropical plant germplasm, current research progress and applications. International Plant Genetic Resources Institute (IPGRI), Rome, Italy. 496 p.
- FAO. 1996. Report on the state of the world's plant genetic resources for food and agriculture. Food and Agriculture Organization of the United Nations (FAO), Rome, Italy. 75 p.
- FAO/IPGRI. 1994. Genebank Standards. International Plant Genetic Resources Institute (IPGRI), Rome, Italy. 13 p.
- Hodgkin T and Debouck DG. 1992. Some possible applications of molecular genetics in the conservation of wild species for crop improvement. *In: Adams RP and Adams JE, editors. Conservation of plant genes. DNA banking and in vitro biotechnology*. Academic Press, San Diego, California, USA. pp. 135–152

- Hong TD, Linington S and Ellis RH 1998. Compendium of information on seed storage behaviour, Vol. 1 and 2. Royal Botanical Gardens Kew, Richmond, UK. 400 pp.
- Karp A. 2002. The new genetic era: will it help us in managing genetic diversity? *In*: Engels JMM, Rao VR, Brown AHD and Jackson MT, editors. Managing plant genetic diversity. CABI Publishing and International Plant Genetic Resources Institute (IPGRI), Wallingford, UK and Rome, Italy. pp 43–56
- Karp A, Kresovich S, Bhat KV, Ayad WG and Hodgkin T. 1997. Molecular tools in plant genetic resources conservation: a guide to the technologies. International Plant Genetic Resources Institute (IPGRI), Rome, Italy.
- Mohd Said Saad and Rao VR. 2001. Establishment and management of field genebank, a training manual. IPGRI-APO, Serdang, Malaysia. 122 p.
- Ramanatha Rao V and Riley KW. 1994. The use of biotechnology for conservation and utilization of plant genetic resources. Plant Genetic Resources Newsletter 97:3–19.
- Stalker HT and Chapman C. 1989 Scientific management of germplasm: characterization, evaluation and enhancement. International Board of Plant Genetic Resources (IBPGR), Rome, Italy. 194 p.

## II. State of the art of DNA storage: results of a worldwide survey

*Meike S. Andersson<sup>1</sup>, Eloina Mesa Fuquen<sup>2</sup> and M. Carmen de Vicente<sup>1</sup>*

<sup>1</sup>International Plant Genetic Resources Institute (IPGRI-Americas), Cali, Colombia

<sup>2</sup>Escuela de Ingeniería Industrial y Estadística, Universidad del Valle, Cali, Colombia

A worldwide survey about DNA storage and use was carried out in 2004 by the International Plant Genetic Resources Institute (IPGRI) among international and national plant genetic resources research institutes, botanical gardens, universities and the private sector, focussing on: 1) the inventory of the current state of knowledge and extent of use of DNA storage; 2) the documentation of the prevailing practices to process and store DNA; and 3) legal issues related to DNA banking.

### Questionnaire development and assemblage of contacts

The survey consisted of two parts. Part one addressed those institutions that store DNA, with the objective to find out why they store DNA; the procedures they use; if DNA is being supplied to others and how; the legal issues related to DNA storage and exchange; the kind of problems encountered; and the kind of improvements proposed. Part two was directed to those institutions that do not store DNA. It tried to find out if they had ever considered DNA storage and why; if so, the reasons why it was decided not to apply it; and whether consideration to storing DNA would be given in the future and under which conditions.

The first part of the survey consisted of 50 questions and the second of 19, the majority of which were simple or multiple-choice. The survey was prepared in English and Spanish and distributed by email and fax. It was sent to 816 institutions in 134 countries assembled from the IPGRI Germplasm Collection Directories (Frison and Servinsky 1995; Knudsen 2000), the FAO Directory 2001 of Correspondents to the Global Network of the World Information and Early Warning System on Plant Genetic Resources for Food and Agriculture (FAO 2001) and by Internet search. Special attention was paid to geographical representation, seeking to include institutions from as many different countries as possible, with a balance between developed and developing countries. The regions were classified as 'developing' and 'developed' according to the status of the majority of the countries constituting them (IMF 2004).

More detailed information about the survey and results of the statistical analysis are presented in the technical report (Andersson 2004, available on request).

### Results

Two hundred and forty-three (30%) valid replies from 77 countries were obtained. Of these, 51 respondents (21%) stored DNA and 192 (79%) did not. The frequency and relative proportion of respondents in the different regions surveyed and according to the type of institution are shown in Tables 1 and 2, respectively.

The proportion of people replying to the survey was similar in developed countries (7.0 to 28.4%) and developing countries and countries in transition (4.5 to 21.4%) (Table 1). The sample studied can thus be considered as balanced with respect to the relationship between developed and developing countries, without bias. The result is noteworthy given the fact that the distribution of the questionnaire was mainly by electronic mail.

**Table 1.** Number and frequency of respondents to the survey per region

Region	Store DNA (No.)	Do not store DNA (No.)	Total respondents (No.)	Responses (%)
<i>Developed countries</i>				
North America and Canada	10	14	24	9.9
Western Europe	13	56	69	28.4
Australia and New Zealand	3	14	17	7.0
<i>Developing countries</i>				
Eastern Europe and Russia	5	29	34	14.0
Asia, Pacific, Oceania	8	16	24	9.9
Central and South America	10	42	52	21.4
West Asia and North Africa	2	10	12	4.9
Sub-Saharan Africa	11	11	11	4.5
Total	51	192	243	100

**Table 2.** Number and frequency of respondents to the survey per type of institution

Type of Institution	Store DNA (No.)	Do not store DNA (No.)	Total respondents (No.)	Responses (%)
National research institutes	31	103	134	55.1
Universities	8	43	51	21.0
Private institutions	4	16	20	8.2
International research institutes	6	12	18	7.4
Others	2	18	20	8.2
Total	51	192	243	100

## Current state of knowledge and extent of use of DNA storage

Analysis of the results showed that in general DNA storage is not a common practice in the Plant Genetic Resources (PGR) community. In fact, out of all the replies received, only 20% stored DNA. The survey revealed that at present the majority (79%) of the institutions do not store DNA due to budget problems, insufficient equipment, as well as constraints in personnel and training. However, more than half of them (57%) would consider DNA storage if base level long-term funding were available to ensure stability of the collections, and funding for equipment and training, and standardized protocols were available.

The results indicated a substantial lack of information about DNA storage and conservation. Eighty-four per cent of the institutions that store DNA expressed a need for more or better information about DNA storage and conservation; 42% of those institutions not storing DNA mentioned the lack of information as an important reason for not doing so. Respondents agreed that information is needed for logistics (i.e. storage procedures, etc.), protocols, legal issues related to DNA exchange, the uses of DNA storage and funding opportunities to support it. The need for coordination and standardization of DNA banking activities as well as the creation of web-based databases with information about available samples worldwide were other issues frequently mentioned by survey respondents. In particular, respondents from developing countries expressed their concern about the possible applications of DNA banking, and the implications, responsibilities and costs involved.



Virtually all institutions (98%) that stored DNA did so with the objective of ensuring its availability for future research activities, whereas gene/genome conservation (29%) and safety duplicates (8%) were only of secondary importance.

Interestingly, DNA storage was mentioned as an ongoing activity with similar frequency in developed and developing countries, with more than half (55%) of the DNA banks located at national research institutes. However, research institutes in both developed and developing countries suffered from inadequate equipment and supplies, e.g. commercial kits for DNA extraction (see below). The USA was the country most engaged in PGR DNA storage (eight institutions), followed by the UK (four), Colombia, Germany and Japan (three each), and Australia, Canada, the Czech Republic, India and Israel (two each).

### **Prevailing practices in processing and storing DNA**

Virtually all institutions that store DNA (98%) extract the DNA themselves, using standard manual extraction protocols or modifications of these. Roughly one-third of the respondents (39%) also resort to commercial kits for DNA extraction, and 4% use commercial kits exclusively. Not surprisingly, the use of commercial extraction kits was considerably higher in developed countries than in developing countries. None of the institutions in developing countries could afford to use commercial kits exclusively, and the institutions in Africa and the Middle East depended 100% on manual extraction. In North America and Western Europe on the other hand, half or more of the institutions (50% and 62%, respectively) had access to kits, and nearly every fifth respondent (20% and 15%, respectively) used them exclusively.

Replies showed that DNA is generally stocked for both medium (6 months to 2 years, at -20 °C or -70 °C) and long (more than 2 years, at -70 °C) periods. One third of the respondents keep it only for short periods (less than 6 months). Although this is usually stored as total DNA, often DNA fragments are also kept (39%).

Almost half of the institutions (45%) that store DNA supply DNA samples to others, mainly for scientific purposes but in some cases also for commercial purposes (private institutions). The DNA is mostly provided for free (65% of the cases), but 69% of the respondents agreed that DNA supply should have a cost covering the material and transport expenses of the institute. The institutions also concurred that, apart from basic passport information, ideally as much additional information as possible should be provided, e.g. quantity and quality of the DNA, the extraction protocol used, molecular marker data, sequence information, references to works or publications on the respective species, original donor, pedigree information, links to genebank data, and names of other institutions that have requested the respective accessions.

Three of the institutions that supply DNA samples have implemented web-based databases providing information about the available accessions:

- USDA Oregon, USA: <http://www.ars-grin.gov/cor>
- NIAS, Japan: <http://www.dna.affrc.go.jp>
- Australian DNA Bank: <http://www.dnabank.com.au> and three others are in the process of developing them:
- CIRAD-Montpellier, France: <http://www.cirad.fr>
- RIFP, Poland: <http://www.insad.pl>
- IRRI, Philippines: <http://www.irri.org/grc>

### **Legal issues related to DNA banking**

Nearly half of the respondents (43%) that supply DNA to others stated that their institutions had not dealt with legal issues related to ownership and the international movement of the

DNA samples (e.g. Convention on Biological Diversity [CBD], International Treaty on Plant Genetic Resources for Food and Agriculture [IT], International Property Rights [IPR]). Only every fourth institution (25%) had established official policies, mostly in the form of material transfer agreements (MTA). MTAs reflect national and international regulations and require signature prior to any germplasm exchange, in order to facilitate the exchange of PGR and to protect them from claims about IPR. MTAs usually include clauses regulating the IPR of the requested material and related information, conditions of its use and distribution to third parties, benefit-sharing, furnishing of the dispatching institute with any data, information or research results obtained by using the genetic material concerned. However, in most cases these agreements are specifically adapted to the exchange of seeds or plant tissue, and do not consider issues that are of importance in the case of exchanging DNA (e.g. IPR and patenting). Examples of MTAs that make explicit reference to DNA are those used by CATIE (Costa Rica), NIAS (Japan), the Missouri Botanical Garden (USA) and the Royal Botanic Gardens, Kew (UK).

Indeed, the legal situation with respect to the collection, exchange and patenting of PGR in general and of plant DNA in particular is very complex, and international regulations are just being developed. An overview about the most important aspects related to this issue is given in Chapter IX.

## Conclusion

Respondents agreed that DNA banking is a complementary conservation strategy on the one hand, and an effective tool for the more efficient management of genebanks and for conservation research on the other. However, the survey showed that DNA storage of plant genetic resources is not a common practice. Particularly in developing countries only a few public sector institutions are using advanced research techniques and most are only in the early stages of developing the capacity for DNA storage. Private sector research has been virtually absent in these countries, in contrast to the developed world where over 70% of DNA storage is performed by private firms. This demonstrates that technical capacity and financial resources for DNA banking are currently very limited in developing countries.

More and better information about the possible uses of DNA banking, and a broader discussion among the scientific community of the opportunities and limitations is crucial. Moreover, there are many questions related to DNA exchange, and IPR and patent protection are still unresolved and legally in a state of flux. Therefore, there is an urgent need to conduct policy research in this field in order to propose solutions for closing these gaps.

IPGRI was requested to initiate and coordinate efforts on a) the standardization of species-specific protocols for DNA extraction, amplification and storage and b) the development of a DNA banking manual, providing the theoretical background and information on the different uses of DNA storage, limitations and opportunities, legal issues related to DNA exchange and patenting, and technical procedures and methodologies. These suggestions are further elaborated in the last chapter of this publication.

## References

- Andersson M. 2004. Feasibility study for DNA storage and use. DNA banks—providing novel options for genebanks? A worldwide survey. Technical Report. International Plant Genetic Resources Institute (IPGRI), Cali, Colombia. 42 p.
- FAO. 2001. Correspondents to the Global Network of the World Information and Early Warning System on Plant Genetic Resources for Food and Agriculture. CGRFA/WG-PGR-1/01/Inf.4. Information Document No. 4 of the First Session of the Commission on Genetic Resources for Food and Agriculture, Working Group on Plant Genetic Resources for Food and Agriculture,

Rome, 2–4 July 2001. Available from: <http://www.fao.org/ag/AGP/AGPS/PGR/ITWG/pdf/P1I4E.pdf>

Frison EA and Servinsky J, editors. 1995. ECP/GR Directory of European Institutions Holding Crop Genetic Resources Collections (4th ed.). International Plant Genetic Resources Institute (IPGRI), Rome, Italy.

IMF. 2004. World Economic and Financial Surveys. World Economic Outlook Database, April 2004. Available from: <http://www.imf.org/external/pubs/ft/weo/2004/01/data/index.htm>

Knudsen H, editor. 2000. Directorio de colecciones de germoplasma en América Latina y el Caribe (1st ed.). International Plant Genetic Resources Institute (IPGRI), Rome, Italy.

### III. DNA storage as a complementary conservation strategy

Ehsan Dulloo<sup>1</sup>, Yoshiaki Nagamura<sup>2</sup> and Oliver Ryder<sup>3</sup>

<sup>1</sup> IPGRI, Rome, Italy

<sup>2</sup> Head of DNA Bank, NIAS, Japan

<sup>3</sup> Zoological Society of San Diego, California, USA

DNA may be a cost effective form for conserving germplasm depending on the objective of the conservation and the type of use to which it would be applied. For many species that are difficult to conserve by conventional means (either as seeds or vegetatively) or that are highly threatened in the wild, DNA storage may provide the ultimate way to conserve the genetic diversity of these species and their populations in the short term, until effective methods can be developed. A survey on the feasibility of DNA storage and use carried out by the International Plant Genetic Resources Institute (IPGRI) in 2004 showed 1) that DNA storage is still not a common practice as a conservation method and 2) that institutions involved in germplasm conservation would apply DNA storage only if a number of constraints are resolved (Andersson 2004; see also Chapter II, this issue). In the present chapter we look at the role that DNA storage can play in the overall conservation strategy of target species, and discuss the concept of complementary conservation strategies and its application to conservation of plant germplasm.

It has long been recognised that no single conservation technique applied alone for conservation of plant genetic resources adequately conserves the full range of genetic diversity of a target species or gene pool (Hoyt 1988; Maxted et al. 1997). The two approaches of conservation, namely *ex situ* and *in situ*, are both important in the conservation and use of genetic diversity and should be regarded as complementary to each other (Maxted et al. 1997; Dulloo et al. 1998; Rao 1999; Engels and Wood 1999). The Convention on Biological Diversity (CBD) also alludes to the complementary nature of *ex situ* conservation to *in situ* conservation in both its preamble and Article 9 of the convention (UNCED 1992). The Global Strategy for Plant Conservation that was adopted in April 2002 by the Conference of the Parties (COP) to the CBD includes a recommendation to “employ *in situ* conservation measures as the primary approach for conservation complementing them where necessary with *ex situ* measures...”.

The concept of a complementary conservation approach has been around for a long time. Different people have used different terminologies to describe it. Nature conservationists prefer to call it ‘integrated strategies’ (Falk 1987; 1990) while plant genetic resources conservationists mostly use the term ‘complementary conservation strategies’ (Maxted et al. 1997, Damania 1996). But for all intents and purposes they refer to the same concept.

A complementary conservation strategy can be defined as “the combination of different conservation actions, which together lead to an optimum sustainable use of genetic diversity existing in a target genepool, in the present and future” (Dulloo et al. 2005). This concept recognizes that the conservation of genetic diversity is not only for the sake of conservation, but also for utilization. The form in which genetic diversity is used is an important one, and would, in consequence, determine the way in which it should be conserved so that its utilization can be facilitated.

It is clear that a complementary conservation strategy is necessary to ensure optimum sustainable use of genetic diversity of any target species. Developing such a strategy involves doing a number of things. It is beyond the scope of the present chapter to discuss in detail how a complementary conservation strategy is formulated. In brief, a number of different issues,

including consideration of biological characteristics, identification of conservation objectives, methodologies available, socio-economic factors and organizational and funding issues, need to be taken into account. Specific case studies have been documented on complementary conservation strategies in relation to a number of crops including musa (Sharrock and Engels, 1997), coffee (Dulloo et al. 1998), coconuts (Dulloo et al. 2005), Asian sweet potatoes (Nissilä et al. 1999) and tropical fruit trees (Drew and Ashmore 2004). Dulloo et al. (2005) provides a framework and decision-making guideline for developing a complementary conservation strategy for coconuts that can be used as a model to develop complementary conservation strategies for other crops or species.

There are various techniques for conserving germplasm, each with their own advantages and disadvantages (Maxted et al. 1997, Maunder et al. 2004). As progress in biotechnology advances, DNA storage is regarded as one of the emerging *ex situ* techniques for germplasm conservation. The next section describes briefly the range of *ex situ* techniques available for germplasm conservation.

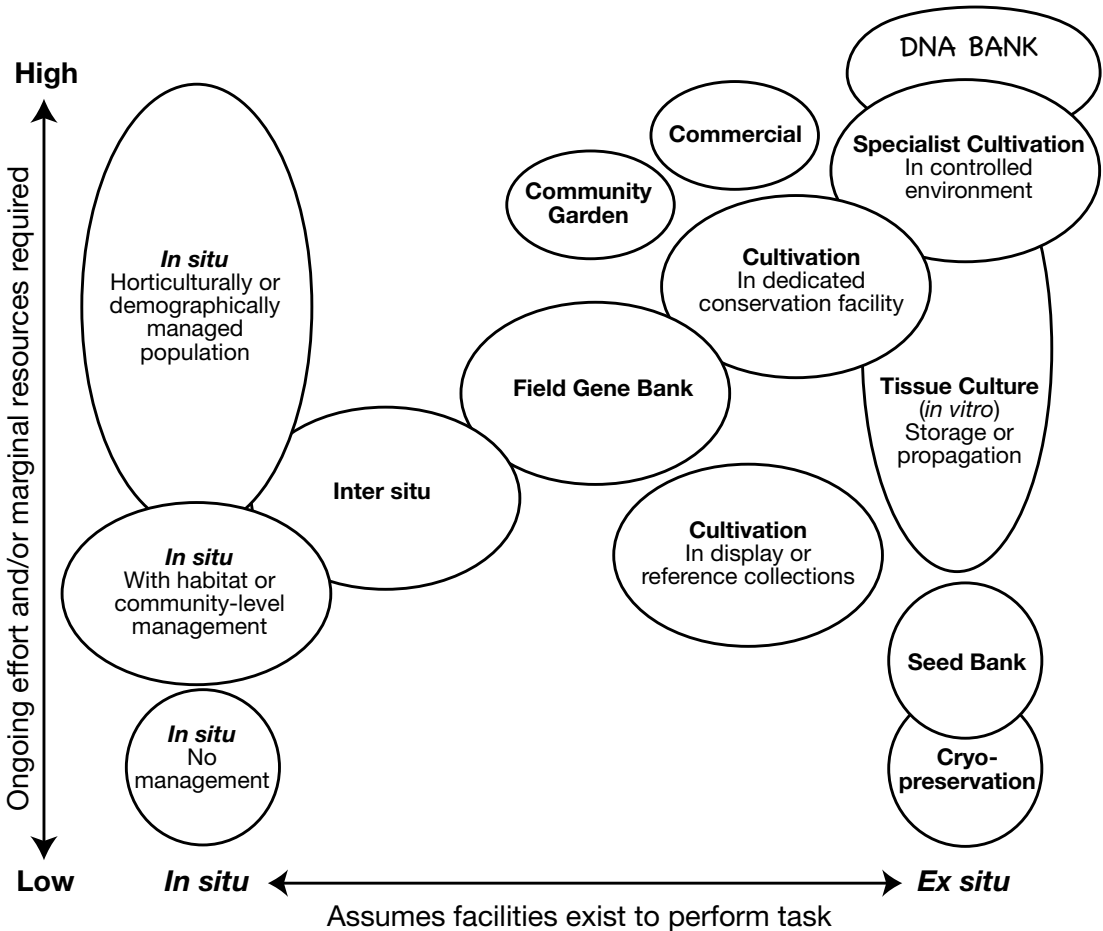
## **Ex situ conservation techniques**

The CBD defines *ex situ* conservation as “the conservation of components of biological diversity outside their natural habitats” (UNCED 1992). There exists a wide range of *ex situ* conservation techniques. The choice of techniques for germplasm storage is determined by a number of factors depending upon amenability of the target species to the particular technique in question (Thormann et al. 2005). The most widely used technique of germplasm conservation is in the form of seeds. This is usually achieved, conventionally, in seed banks for species with orthodox seed, which can be stored at low temperatures and low moisture content (Roberts 1973), and 90% of the 6 million accessions held in over 1500 genebanks globally (FAO 1998; Imperial College 2003) are conserved in this way. However, many species produce seeds that are desiccation sensitive and/or cold sensitive (Roberts 1973; Pammenter and Berjak 1999) or are usually propagated vegetatively. In such cases, they cannot be kept in seed banks. Conventionally they have been conserved as living plants in field genebanks (Reed et al. 2004) and botanic gardens. Botanic Gardens are estimated to cultivate four million accessions of plant taxa globally, but have very little representation of genetic diversity. With the advent of biotechnology, new methods have been developed using *in vitro* slow growth and cryopreservation techniques. *In vitro* conservation provides a useful aseptic method of conservation and solves the problems associated with pests and diseases and environmental hazards that are a major limiting factor for field collections. At the same time, it increases their availability for distribution if the materials are maintained virus free (Reed et al. 2004). Cryopreservation is another option now being regarded as one of the most promising techniques for conservation for the so called ‘problem’ or ‘difficult’ species, i.e. those that are vegetatively propagated and species with recalcitrant seeds (Engelmann and Engels 2002). Although protocols for several hundreds of species of temperate and tropical origins have been developed using different types of explants (cell suspension, calluses, apices, zygotic and somatic embryos) (Engelmann and Dussert 2000), there is still a rather limited number of cases where cryopreservation is currently used in genetic resources conservation. Cryopreservation is more advanced for vegetatively propagated crops than for recalcitrant seed species. There is much more work to be done in understanding the recalcitrant behaviour of some species in order to be able to develop cryopreservation protocols.

DNA banking is an emerging technique in genetic resources conservation. DNA extracts, DNA or RNA sequences are considered as genetic resources and they are now routinely extracted and conserved in DNA banks (Adams 1997). It is often argued that DNA banking is an unrealistic option for germplasm conservation given that DNA cannot resurrect extinct species

(Maunder et al. 2004). DNA storage should still be regarded as an insurance policy rather than a replacement for conventional modes for germplasm storage (Callow et al. 1997). One of the most frequent forms of germplasm requested from genebanks is DNA material (de Vicente, 2004). If such is a need for provision of DNA material for molecular genetic research, then DNA banks have their *raison d'être*.

Maunder et al. (2004) provides a good illustration showing the relative relationship between the different conservation techniques (Figure 1). In this framework, we have included the position of DNA storage to show its relationship to the other techniques of conservation.



**Figure 1.** Illustration showing the relative relationship between the different conservation techniques (after Maunder et al., 2004; from *Ex situ* conservation by Edward O. Guerrant et al., editors. Copyright © 2004 Island Press. Reproduced by permission of Island Press, Washington, DC, USA).

## ***In situ* conservation and DNA storage**

The biodiversity of plants and animals, including species diversity and intra-specific genetic variation, are best preserved in natural environments. This has often been considered as a preferred method for conservation. *In situ* conservation is defined as “the conservation of ecosystems and natural habitats and the maintenance and recovery of viable populations of species in their

natural surroundings and, in the case of domesticated or cultivated species, in the surroundings where they have developed their distinctive properties.” (UNCED 1992, Article 2). Thus, *in situ* conservation is regarded as conserving germplasm not only in the natural habitat where the target species is found, but also in man-made habitats such as farms and home gardens. The latter applies particularly to cultivated plants and their cultivars, landraces and weedy forms, and not so much to wild relatives of crops. Maintaining viable populations is another key phrase in the definition because it ensures that species do not become static but maintain their evolutionary potential and ecological functions (Kennedy 2004).

In contrast to *ex situ* conservation, *in situ* methods offer the best opportunities for the conservation of multiple species, particularly recalcitrant species, while allowing the dynamic evolutionary processes to continue, especially in producing new resistance to pests and diseases. Provided *in situ* reserves are designed optimally, they may conserve a broader range of variability of the target taxa than is possible in *ex situ* collections, particularly for recalcitrant seed species. However, the difficulties are that the materials are not easily accessible for use and they may be vulnerable to natural and man-made calamities and other biotic interferences such as invasive alien plants. In fact, for both plants and animals, loss of habitat and altered environmental conditions have led to the decline and disappearance of some species and the endangerment of others. It is a fact that no *in situ* conservation programme would ever be able to protect all the populations and the full range of genetic variability, except for a limited number of highly threatened species. Some priority setting will have to be made in designing *in situ* protected areas (FAO, DFSC and IPGRI 2001; Heywood and Dulloo 2005).

For these reasons, *ex situ* conservation should be regarded as an important complementary approach for conservation and use of the biological diversity, targeting those populations that are difficult to conserve *in situ*. Collecting representative samples of the wild populations is crucial and there has been much debate on how much to collect and from how many populations to ensure the maximum of genetic diversity (Brown and Briggs 1991; Neel and Cummings 2003). In case of critically endangered species where very few populations or even individuals exist, the entire genepool can be collected and maintained in *ex situ* collection as safety duplication to the *in situ* populations (see the example on the California condor below). Making available the diversity in these collections is also of critical importance. Depending on the type of use of the material, conserving the material in the form of DNA can be very advantageous and will help to facilitate the use of such material. Here DNA banking can have a major role to play. However, the utilisation of these materials should be well regulated to comply with the provision of the access and benefit sharing of the Convention on Biological Diversity and the International Treaty on Plant Genetic Resources for Food and Agriculture (FAO 2004).

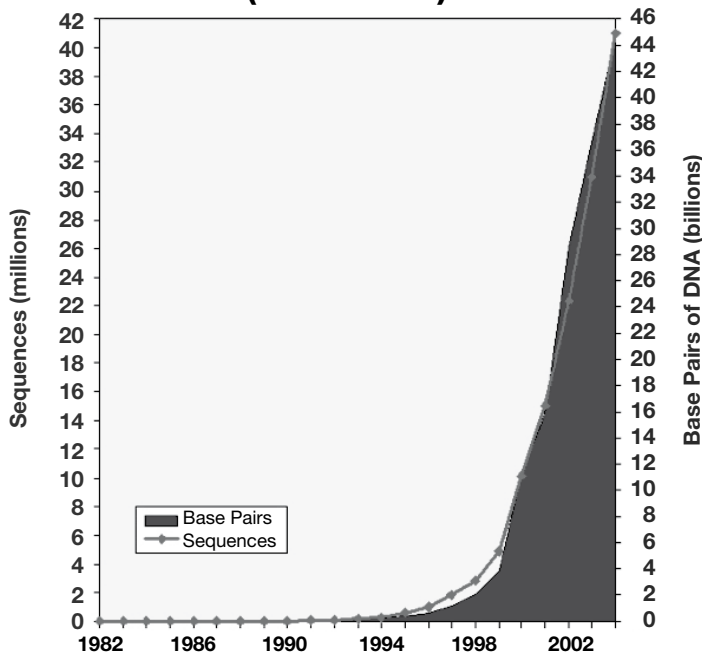
## Complementary nature and importance of DNA storage

One of the major contributions of *ex situ* genetic resources is the ready provision of documented specimens for use in studies to assess, monitor and manage biological diversity across taxonomic levels, trophic levels, and ecosystems. Increasingly, knowledge of the evolutionary ecological, physiological and behavioural biology of species is amenable to elucidation through genetic studies for which access to small amounts of DNA is crucial. Genetic analyses are becoming a *sine qua non* for systematic and strategic conservation efforts. Accordingly, storage and curation of documented DNA samples serve a crucial infrastructure need in conservation science and complements efforts to preserve biological diversity both *in situ* and *ex situ*.

Recent advances in DNA sequencing technology and reduction in the cost of sequencing reagents have brought remarkable progress in genome analysis of various species. At present (as of February 2005), the genomes of 1389 species are being analyzed by various research

groups all over the world. Among them, 258 species have been completely sequenced and the rest, including 655 prokaryotic species and 474 eukaryotic species, are in various phases of genome sequence analysis. The nucleotide sequence data generated from these efforts are submitted and registered in the three major international sequence databases, namely NCBI/GenBank, DDBJ (DNA Data Bank of Japan) and EMBL/EBI (European Bioinformatics Institute). Researchers worldwide can have direct access to these data and use them for various analyses such as homology search, etc. As of February 2005, the registered base sequence information in NCBI/GenBank has reached more than 40 580 000 entries corresponding to 44 billion nucleotide base pairs (Figure 2).

### Growth of GenBank (1982-2004)



**Figure 2.** Growth in the number of registered base sequence information in NCBI/GenBank. Source: (<http://www.ncbi.nlm.nih.gov/Genbank/genbankstats.html>).

In addition to the sequence information, researches associated with genome analysis generally produced biological materials such as cDNA clones, RFLP (Restriction Fragment Length Polymorphism) markers on the genetic map, cosmid clones, PAC (P1-derived Artificial Chromosome) clones, BAC (Bacterial Artificial Chromosome) clones, YAC (Yeast Artificial Chromosome) clones, etc. These genome resources are indispensable tools for future post-genome research, such as physiological and morphological characterization of a species, functional analysis of genes, comparative genomics and plant breeding. Therefore, it is necessary to maintain an efficient system for conservation and management of DNA materials.

Genetic resource collections in the form of frozen tissues, purified DNA samples, frozen, viable cell cultures and derivatives such as RNA, cDNA and genomic libraries all represent valuable components of comprehensive storage strategy. Initially, frozen tissues and viable seed collections were developed. Purification of DNA for a large number of samples is laborious and time consuming. Furthermore, DNA may be more stable in frozen tissues than in aqueous solution. Lyophilization of DNA to provide long-term storage of purified material adds an

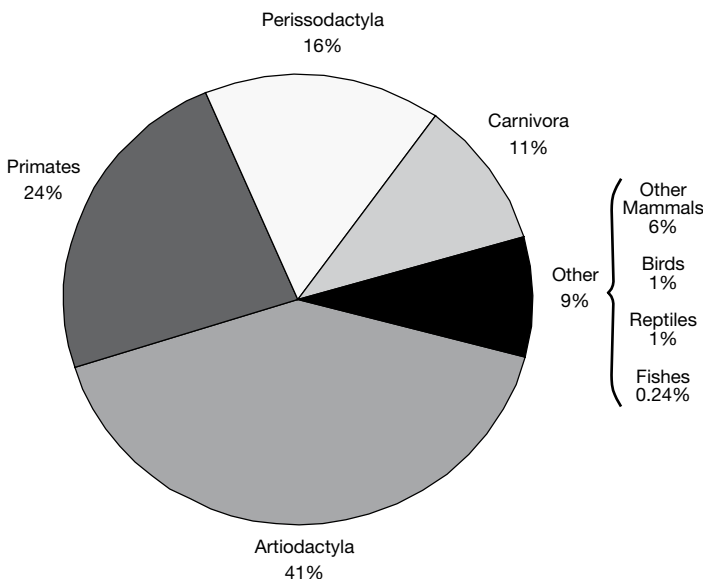


additional step in the preparation process that is costly when large numbers of samples are considered. However, tissue storage in ultra cold freezers is notoriously unreliable; many valuable collections have been lost for lack of vigilant freezer surveillance and contingency plans for the inevitable freezer failures. Liquid nitrogen storage is more expensive but has been successfully developed as a long-term conservation strategy for both plant and animal genetic resources. However, there are no long-term liquid nitrogen storage facilities for DNA samples.

Because DNA storage collections have a relatively more recent history and more specialized use than do frozen tissues and germplasm collections, we can anticipate that recurrent evaluation and strategic planning will continue to shape thinking and guide policy for DNA banking efforts. However, it is of considerable importance to point out that, as biodiversity declines continue, no better opportunities than the present are available to collect and incorporate extant biodiversity as part of the efforts to document and conserve a legacy for future generations. Efforts to establish DNA banks have been progressing worldwide. Two examples of this kind of initiative for animal and plant diversity are provided below.

### The Frozen Zoo®

The Zoological Society of San Diego's initiative called 'Conservation and Research for Endangered Species' (CRES) has established a unique repository for frozen cell cultures, frozen tissues, gametes, embryos and DNA samples. CRES began by preserving biological specimens with the formation of the Frozen Zoo® in 1975. The Frozen Zoo® currently holds viable fibroblast cell lines and tissue pieces (the living tissue collection) and/or nucleic acid preparations from approximately 7300 individual mammals, birds, and reptiles. Artiodactyls, perissodactyls, primates and carnivores are very heavily represented (Figure 3), together comprising greater than 85% of the total collection. The emphasis of the collection in the past was on specimens from mammalian taxa. Now 20 of the 26 mammalian orders (Wilson and Reeder 1993) are represented. More than 95 species within the primate order alone are held in the collection. Another focus was to seek significant representation in the collection of samples from Africa, Asia, and South America. More recently, this aim has been expanded in the effort to obtain samples from other regions, including San Diego's own local species.



**Figure 3.** Samples in the Frozen Zoo® by major taxonomic division.

Of major importance are animals listed as endangered and threatened. Of the approximately 500 species/subspecies in the collection, more than 200 are listed in the IUCN Red List of Threatened Animals. An equally important concern from the onset was the interaction of CRES conservation efforts with those of the Species Survival Plan (SSP) groups under the auspices of the American Zoo and Aquarium Association (AZA). Seventy-two of the 74 SSP mammalian taxa are now represented in the CRES repository, and close cooperation with several SSP groups has enabled the continuing collection and storage of specimens from all captive individuals of certain targeted species.

## Rice genome research in Japan

As an example of current efforts for a commercially important plant species, the mapping and sequencing of the rice genome is informative. Rice is one of the major cereal crops and is the staple food for over 50% of the world population mainly concentrated in Asia. The rice plant originated from domestication of wild rice which resulted in various cultivars through adaptation to a wide range of environmental conditions and continuous breeding for over 10 000 years. However, despite a continuous increase in world population, rice production has actually declined in recent years due to drastic environmental changes. Thus it is necessary to develop innovative approaches to breed new rice varieties with high productivity in a shorter period of time than it takes by conventional breeding methods. This can be easily achieved through identification of the genes that control high yield such as flowering time, seed grain density etc. at the DNA level. For this reason, it is necessary to fully characterize the structure of the rice genome and determine the function of all genes that comprise the rice plant.

The Ministry of Agriculture, Forestry and Fisheries (MAFF) in Japan started a comprehensive analysis of the rice genome in 1991, using a japonica rice cultivar 'Nipponbare'. As fundamental research for clarifying all genome information of the rice plant, the objectives of the first stage of the Rice Genome Research Program (RGP) were: 1) to establish a catalogue of all expressed genes; 2) to construct a high-density linkage map; and 3) to generate a YAC-based physical map of the rice genome. The results and resources obtained were then used in the second phase of RGP (1998–2005), which focused on the sequencing of the entire genome. The other objectives of the rice genome project included: 1) analysis of full-length cDNA sequences; 2) analysis of gene function by insertional mutagenesis using an endogenous retrotransposon; and 3) isolation of genes for important agronomic traits in rice. The rice genome sequencing is a collaboration of publicly funded laboratories from 10 countries, namely Brazil, China, France, India, Japan, Korea, Taiwan, Thailand, the UK and the USA, which participated in the International Rice Genome Sequencing Project (IRGSP). This consortium accelerated the genome sequencing and achieved its goal of accurately deciphering the complete nucleotide sequence of the japonica cultivar 'Nipponbare' by the end of 2004.

Major new initiatives to understand plant and animal biology are closely integrated with DNA storage collections, although perhaps not to the extent that the future scientific community will consider desirable. As plant and animal genomes are sequenced, access to DNA samples from the individual organisms whose genomes were sequenced is required for many continuing studies. Similarly, large insert libraries, such as bacterial artificial chromosome (BAC) and fosmid libraries, play a crucial role in initial and ongoing investigations of genome organization and sequencing in a variety of species. To advance basic genomic studies and facilitate comparative studies, DNA library resources must be expanded and a functional system for ensuring long-term access to existing libraries must be put in place.

## Rice DNA resources

The rice express sequence tag (EST) project generated a total of 30 040 cDNA clones with partial sequence from the 5' and 3' end (Sasaki et al. 1994; Yamamoto and Sasaki 1997). These clones were derived from cDNA libraries constructed from various tissues and organs such as green shoot, etiolated shoot, root, panicle and callus under various culture conditions. The cDNA libraries for the rice full-length cDNA project were constructed by the biotinylated cap-trapper and oligo-capping methods. A total of 175 642 cDNA clones were partially sequenced and subjected to clustering. Then the full sequence of 32 127 non-redundant clones was analyzed (The Rice Full-Length cDNA Consortium 2003). All rice ESTs and full-length cDNAs are available for distribution.

For accurate analysis of the genome, a high-density linkage map (genetic map) was constructed by RFLP analysis using cDNA and genomic clones. As a result, 3599 DNA markers were mapped on the 12 rice chromosomes. A set of land markers representing specific positions in the 12 chromosomes is available for distribution (NIAS DNA BANK).

Rice genome sequencing analysis was carried out by hierarchical genome shotgun sequencing method based on the accurate genetic and physical map. For physical mapping, a yeast artificial chromosome (YAC) library with 350 Kb average insert size, a bacterial artificial chromosome (BAC) library and a P1-derived artificial chromosome (PAC) library with 150 Kb average insert size were constructed. In particular, the PAC and BAC clones were used to construct a sequence-ready physical map that serves as template for genome sequencing.

Insertional mutagenesis is an efficient strategy for characterizing the function of genes in rice. It has been proved that the activation of rice retrotransposon (Tos17) cause tissue culture induced mutations in rice. From 50 000 mutant lines produced by insertional mutagenesis using Tos17, about 5000 lines have been analyzed for phenotype and flanking sequence of insertion sites (Hirochika 1997). Seeds of these Tos17 mutant lines are available (Rice Genome Resource Centre).

Most agronomically useful traits such as plant height, panicle number, heading date or the so-called quantitative trait loci (QTLs) are difficult to analyze because they are controlled by multiple genes. However, research materials for genetic analysis, such as back-cross inbred lines (BILS), nearly-isogenic lines (NILs) and chromosome segment substitution lines (CSSLs), have been developed using the DNA markers and genome sequence data. These materials facilitate comprehensive analysis and positional cloning of target QTLs (Yano et al. 2000; Kojima et al. 2002).

Detailed information of the rice genome resources, such as nucleotide sequence, chromosomal position, etc. can be accessed at the following Web sites:

- Rice EST — <http://bank.dna.affrc.go.jp/%7Eqxrice/hiho>
- Rice Full Length cDNA — <http://cdna01.dna.affrc.go.jp/cDNA>
- Rice RFLP marker — <http://rgp.dna.affrc.go.jp/publicdata/caps/index.html>
- Rice YAC, BAC, PAC — <http://rgp.dna.affrc.go.jp/giot/INE.html>
- Rice Tos17 Insertion Lines — <http://tos.nias.affrc.go.jp>
- Rice Genetic analysis materials — <http://www.rgrc.dna.affrc.go.jp/stock.html>

## Access to rice DNA resources

These DNA materials are valuable resources that have been developed from comprehensive analysis of the rice genome over a long period of time, using enormous budget and considerable human effort. They are useful research tools that can be utilized in various research fields, such as breeding, isolation of genes for useful traits, functional analysis of the genes, comparative genome analysis, etc (Sasaki et al. 2005). These genome resources are particularly valuable because they are accompanied by relevant information, such as an accurate nucleotide sequence, functional information, positional information on the chromosome, etc. The National Institute of Agrobiological Sciences (NIAS), through the DNA Bank and the Rice Genome Resource Centre,

facilitates the preservation and management of the biological materials generated from the rice genome project (Antonio et al. 2003). It also provides researchers from all over the world with access to these materials to facilitate efficient use of these resources in plant science research.

The DNA clones, such as ESTs, full-length cDNAs, BACs, PACs and YACs, are fundamentally stored and maintained in 96-well microplates or 384-well microplates stored in  $-80^{\circ}\text{C}$  ultra-deep freezers. These clones are preserved in duplicate or triplicate in order to prevent the risk that may be caused by troubles arising from mechanical failures and other disturbances. For cDNA clones, plasmid DNA is extracted from the host *Escherichia coli*, and stored at  $-30^{\circ}\text{C}$  in 1.5 ml Eppendorf tubes labelled with a two-dimensional bar code. This facilitates a system for inventory and accelerates the distribution of the samples to users. It also minimizes human error that may occur during sample preparation. Seeds of 50 000 Tos17 mutant lines are also stored and labelled with bar codes.

Recently, the NIAS, in collaboration with RIKEN Institute in Japan, constructed a DNA book for rice as an alternative method for maintenance and distribution of DNA materials (Kawai and Hayashizaki 2003). The 'Rice Full-Length cDNA Encyclopedia DNABook™' contains 32 000 clones printed or fixed on special paper and bound as a book. The DNA book has the following advantages: 1) efficient maintenance of a large number of DNA materials in small space and under normal conditions; 2) easy recovery of DNA from paper and extraction using PCR; 3) less labour-intensive and less time-consuming as compared with storage in freezer; and 4) cheap distribution using ordinary mail.

## DNA collections in support to biodiversity evaluations

Current initiatives to produce phylogenies for plants and animals that involve greater taxonomic sampling, incorporation of larger datasets of molecular and morphological characters and combinatorial analytical methods are resulting in unprecedented collaborations producing increasingly large datasets. A key factor in undertaking these new initiatives is the availability of samples for appropriate taxon sampling. It is also important that molecular analysis be available from the same or a closely related taxon to those used in morphological analyses of extant taxa.

The 'tree of life' that is generated provides a strong framework for interpreting biological phenomena. Accordingly, it will be very useful to future investigators to have access to specimens that yielded the data producing the resultant phylogenies. Although such a need can be anticipated, it is not possible to say that there are significant initiatives underway to establish and make available such a framework resource of biomaterials for further studies in biotic evolution.

It should be noted that some of the species that constitute key taxa for resolving phylogenies may be threatened or endangered, suggesting that the availability of samples to future investigators cannot be assumed unless specific steps are taken to include these species in genetic resource banks from which they will be broadly and equitably shared. For example, it is estimated that approximately one-fourth of all mammals and as many as one-third of all primates may face extinction in the next two or three decades.

Efforts to collect samples from endangered species and incorporate them into DNA banking serve broader interests than evolutionary and comparative studies. Increased knowledge of the biology of endangered species may be crucial for their long-term survival and it is likely that rigorous scientific analysis of DNA data will require the availability of a larger set of samples than are currently being incorporated into DNA banks. Both the number of individuals sampled and the amount of DNA available may become limiting resources for conducting rigorous scientific studies. Furthermore, it is difficult to anticipate which species, now represented by only small amounts of DNA, in DNA banks may be sought after by future investigators. Establishing cell

cultures through cryopreservation of tissue fragments, from which cell cultures can later be established, provides a more renewable resource for supplying DNA. Additionally, cell cultures also provide intact cellular constituents, facilitate cytological studies, and may be an important source of cellular RNAs for analysis and construction of cDNA libraries.

While DNA banks exist in many countries and in different regions of the world, there is no effective coordination among DNA banks nor are there any databases that can be queried to provide details regarding DNA collections. However, some efforts are underway and, although these are in their beginning stages, represent steps toward facilitating access to DNA specimens or investigations relating to basic science and conservation biology that are non-commercial in nature. INPRIMAT is a European initiative focusing on making available existing primate DNA resources and developing and providing DNA libraries and various reagents (such as DNA microsatellite primers) for primates. In the USA the Integrated Primate Biomaterials and Information Resource (IPBIR) is an NSF-sponsored programme that aims to provide high quality DNA samples of known provenance for research studies. IPBIR relies heavily upon the production of cell lines as renewable resources for providing DNA and also includes associated data to confirm the species identity of individual samples and other voucher information.

As whole genome sequencing efforts are being undertaken for more species, including endangered ones such as the chimpanzee (*Pan troglodytes*), Savannah elephant (*Loxodonta africana*), the Western lowland gorilla (*Gorilla gorilla gorilla*) and the Sumatran orangutan (*Pongo pygmaeus abelii*), DNA sequencing efforts for these taxa will not produce the same extent of coverage as for model organisms such as human, mouse, drosophila, maize, *Arabidopsis* etc. Thus, there will continue to be calls for DNA from these species that are not now readily available (except in the case of the great apes through IPBIR and INPRIMAT).

An example of a DNA banking project that contributes to the understanding of a major taxonomic group (Aves) and has a decided conservation application is the effort in progress to produce a genetic map for the endangered California condor, *Gymnogyps californianus*. The California condor was on the verge of extinction, with only eight birds remaining in the wild and a small captive population, when the crucial decision to bring all the birds into captivity was made. Since then, DNA samples have been saved from every California condor and a complete pedigree established. DNA analysis is now routinely applied for sexing the monomorphic bird (Chemnick et al. 2000). Elucidated by multi-locus probe DNA fingerprinting techniques, studies of shared genetic variation have identified three closely related groups (clans) of California condors that constitute the remnants of the gene pool of the species that once distributed throughout much of North America. Both a California condor BAC library and an enriched microsatellite library are available. A segregating hereditary disorder, chondrodystrophy, has been identified in the California condor (Ralls et al. 2000), using the banked DNA samples (including those from affected chicks). A physical map of the California condor can be constructed taking advantage of the advances in genomic biology that have been made for the chicken (International Chicken Genome Sequencing Consortium 2004).

Another example from the plant world, which illustrates the value of DNA banking to our understanding of plant's taxonomy and phylogeny and in helping develop conservation and population management strategies, is the Tropical Plant DNA Bank located in the Fairchild Centre for Tropical Plant Conservation. This centre was established to serve as a permanent repository for genetic materials from Fairchild's plants collection and in particular their palm collection. The Fairchild Centre has the world's largest collection of palm DNA, mostly from palms and cycads cultivated at the Centre. Their goal is to have every palm and cycad species in the living collection represented in the DNA bank. It allows harvesting research material (DNA) from relatively transient *ex situ* collections and serves to carry out phylogenetic work on palms and cycads, population management including recovery planning, and security mark-

ing for high value cycads (M. Maunder personal communications). The bank also includes samples collected directly from the wild. This is important because it provides the opportunity of analyzing data from the wild and helps in the development of conservation strategies. For example, samples from wild populations of the princess palm (*Dictyosperma album* (Bory) H.A. Wendl. & Drude ex Scheffer), a highly endangered species from Mascarene islands in the Indian ocean, have been compared with cultivated populations in order to develop a new conservation strategy for *Dictyosperma* (C. Lewis, personal communications).

## Future prospects for DNA banking of economically important species and global biodiversity

As the importance of collecting and providing access to samples of biological diversity for describing, understanding, and conserving the living world becomes more recognized and ingrained, we envision that efforts similar to the Frozen Zoo® and the Rice genome research project will foster a much-needed collaborative system for sharing information and samples. However, currently the collective effort is incompletely described and its global extent not fully recognized.

Rice is considered a model plant because it has the smallest genome size among the major cereal crops. The accurate sequence of the rice genome will be invaluable in understanding the structure of other cereal genomes. In addition, various genome resources such as 32 000 rice full-length cDNA clones with full sequence information (structural information), about 50 000 Tos17 mutant lines and genetic analysis materials are also available. In the future, these molecular research materials will become an important tool of the plant life science research, such as functional genome analysis, comparative genome analysis, microarray development, etc. These DNA materials will be useful in rice breeding as well. As we can see in cultivated rice, only traits that are useful for humans have been selected during the breeding process and many other useful traits may have been deleted. The genome sequence information and accompanying biological materials may provide the key for unravelling many important genetic resources that remain undiscovered in wild rice species. For these reasons, valuable DNA materials must be stably maintained and conserved in the long term since they will be indispensable tools for new discoveries in breeding in order to provide enough rice to feed the growing population. This is the mission of DNA banking. We hope that the molecular and genetic resources will be efficiently utilized by the scientific community to facilitate innovative researches in both functional and applied genomics of rice and other cereal crops.

## References

- Adams RP. 1997. Conservation of DNA: DNA banking. *In* Callow JA, Ford-Lloyd BV, Newbury HJ, editors. *Biotechnology and Plant Genetic Resources: Conservation and Use*. Biotechnology in Agriculture Series 19. CAB International, Wallingford, UK. Pp. 163–174.
- Andersson M. 2004. Feasibility study for DNA storage and use. DNA banks—providing novel options for genebanks? A worldwide survey. Technical Report. International Plant Genetic Resources Institute (IPGRI), Cali, Colombia.
- Antonio BA, Miyao A, Nagamura Y, Sasaki T. 2003. The Rice Genome Resource Center as an outlet for distribution of biological materials from the Rice Genome Project. *Rice Genet. Newsl.* 20:10–12.
- Brown AHD, Briggs JD. 1991. Sampling strategies for genetic variation in *ex situ* collections of endangered plant species. *In* Falk DA, Holsinger KE, editors. *Genetics and conservation of rare plants*. Oxford University Press, New York, USA. pp. 99–119

- Callow JA, Ford-Lloyd BV, Newbury HJ, editors. 1997. *Biotechnology and Plant Genetic Resources: Conservation and Use*. Biotechnology in Agriculture Series 19. CAB International, Wallingford, UK.
- Chemnick LG, Kumamoto AT, Ryder OA. 2000. Genetic analyses in support of conservation efforts for the California condor *Gymnogyps californianus*. *Int. Zoo Yearbook* 37:330–339.
- Damania AB. 1996. Biodiversity conservation: a review of options complementary to standard *ex situ* methods. *Plant Genet. Resour. Newsl.* 107:1–18.
- de Vicente MC. 2004. The Evolving Role of Genebanks in the Fast-developing Field of Molecular Genetics. *Issues in Genetic Resources* No. 11. International Plant Genetic Resources Institute (IPGRI), Rome, Italy.
- Drew R, Ashmore S. 2004. Complementary Conservation strategy of tropical fruit trees. Consultancy report to International Plant Genetic Resources Institute, Rome, Italy. 69 p.
- Dulloo ME, Rao VR, Engelmann F, Engels JMM. 2005. Complementary Conservation strategy for coconuts. *In* Batugal P, Rao VR, editors. *Coconut Genetic Resources*. International Plant Genetic Resources Institute (IPGRI), Rome, Italy.
- Dulloo ME, Guarino L, Engelmann F, Maxted N, Newbury HJ, Attere F, Ford-Lloyd BV. 1998. Complementary conservation strategies for the genus *Coffea*: a case study of Mascarene *Coffea* species. *Genet. Resour. Crop Ev.* 45:565–579.
- Engelmann F, Dussert S. 2000. Current development and use of cryopreservation for the conservation of plant genetic resources. *Cah. Ag.* 9:237–245.
- Engelmann F, Engels JMM. 2002. Technologies and strategies for *ex situ* conservation. *In* Engels JMM, Rao VR, Brown AHD, Jackson MT, editors. *Managing Plant Genetic Diversity*. CABI Publishing, Wallingford, UK. pp. 89–103.
- Engels JMM, Wood D. 1999. Conservation of agrobiodiversity. *In* Wood D, Lenné JM, editors. *Agrobiodiversity Characterisation, Utilisation and Management*. CABI Publishing, Wallingford, UK. pp. 355–385.
- FAO. 2004. *The International Treaty on Plant Genetic Resources for Food and Agriculture*. Food and Agriculture Organization of the United Nations (FAO), Rome, Italy.
- FAO. 1998. *The state of the world's plant genetic resources for food and agriculture*. Food and Agriculture Organization of the United Nations (FAO), Rome, Italy. 510 p.
- FAO, DFSC, IPGRI. 2001. *Forest Genetic Resources Conservation and Management Vol.2: In managed natural forests and protected areas (in situ)*. International Plant Genetic Resources Institute (IPGRI), Rome, Italy.
- Falk DA. 1987. Integrated conservation strategies for endangered plants. *Nat. Area J.* 7:118–123.
- Falk DA. 1990. Integrated strategies for conserving plant genetic diversity. *Ann. Missouri Bot. Gard.* 77:38–47.
- Heywood VH, Dulloo ME. 2005. *In situ* conservation of wild plant species. A critical global review. Technical Bulletin No. 11. International Plant Genetic Resources Institute (IPGRI), Rome, Italy.
- Hirochika H. 1997. Retrotransposons of rice: their regulation and use for genome analysis. *Plant Mol. Biol.* 35:231–240.
- Hoyt E. 1988. *Conserving wild relatives of crops*. World Conservation Union (IUCN), Gland, Switzerland.
- Imperial College. 2003. *Crop Diversity at risk. The case for sustaining crop collections*. Imperial College of Science and Technology and Medicine, London, UK.
- International Chicken Genome Sequencing Consortium. 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* 432:695–716.
- Kawai J, Hayashizaki Y. 2003. *DNA Book*. *Genome Res.* 13:1488–1495.

- Kennedy KL. 2004. The role of Federal Guidance and State and Federal partnership in *ex situ* plant conservation in the United States. In Guerrant jr EO, Havens K, Maunder M, editors. *Ex situ* plant conservation supporting species survival in the wild. Island Press, Washington DC, USA. pp. 67–83.
- Kojima S, Takahashi Y, Kobayashi Y, Monna L, Sasaki T, Araki T, Yano M. 2002. *Hd3a*, a rice ortholog of the *Arabidopsis FT* gene, promotes transition to flowering downstream of *Hd1* under short-day condition. *Plant Cell Physiol.* 43:1096–1105.
- Maunder M, HavensK, Guerrant Jr EO, Falk DA. 2004. *Ex situ* methods a vital but underused set of conservation resources. In Guerrant Jr EO, Havens K, Maunder M, editors. *Ex situ* Plant conservation supporting species survival in the wild. Island Press, Washington DC, USA. pp. 3–20.
- Maxted N, Ford-Lloyd BV, Hawkes JG. 1997. *Plant genetic conservation: the in situ approach.* Chapman and Hall, London, UK.
- Neel MC, Cummings MP. 2003. Effectiveness of conservation targets in capturing genetic diversity. *Conserv. Biol.* 17:219–229.
- Nissilä EAJ, Rao VR, Engelmann F, Riley KW. 1999. *Ex situ* strategies for complementary conservation of Asian sweet potatoes. *Plant Genet. Resour. Newsl.* 117:1–11.
- NIAS DNA Bank, National Institute of Agrobiological Sciences, Ibaraki, Japan. Available from: <http://www.dna.affrc.go.jp/>
- Pammenter NW, Berjak P. 1999. A review of recalcitrant seed physiology in relation to desiccation-tolerance mechanisms. *Seed Sci. Res.* 9:13–37.
- Ralls K, Ballou JD, Rideout BA, Frankham R. 2000. Genetic management of chondrodystrophy in California condors. *Anim. Conserv.* 3:145–153.
- Rao VR. 1999. Complementary conservation strategy. In Bhag Mal, Mathur PN, Rao R, editors. South Asia Network on Plant Genetic Resources (SANPGR), Proceedings of Fourth Meeting, 1–3 September, 1998, Kathmandu, Nepal. IPGRI Office for South Asia, New Delhi, India. pp. 139–150.
- Reed B, Engelmann F, Dulloo ME, Engels JMM. 2004. Technical Guidelines on management of field and *in vitro* germplasm collections. Handbook for Genebanks No.7. International Plant Genetic Resources Institute (IPGRI), Rome, Italy.
- Rice Genome Resource Centre, Japan. Available from: <http://www.rgrc.dna.affrc.go.jp/index.html.en>
- Roberts EH. 1973. Predicting the viability of seeds. *Seed Sci. Technol.* 1:499–514.
- Sasaki T, Matsumoto T, Antonio BA, Nagamura Y. 2005. From mapping to sequencing, post-sequencing and beyond (review). *Plant Cell Physiol.* 46:3–13.
- Sasaki T, Song J, Koga-Ban Y, Matsui E, Fang F, Higo H, Nagasaki H, Hori M, Miya M, Murayama-Kayano E, Takiguchi T, Takasuga A, Niki T, Ishimaru K, Ikeda H, Yamamoto Y, Mukai Y, Ohta I, Miyadera N, Havukkala I, Minobe Y. 1994. Toward cataloguing all rice genes: large scale sequencing of randomly chosen rice cDNAs from a callus cDNA library. *Plant J.* 6:615–624.
- Sharrock S, Engels J. 1997. Complementary conservation. In Networking Banana and Plantain: INIBAP Annual Report 1996. International Network for the Improvement of Banana and Plantain, Montpellier, France. pp. 6–9.
- The Rice Full-Length cDNA Consortium. 2003. Collection, Mapping, and Annotation of Over 28,000 cDNA Clones from *japonica* Rice. *Science* 301:376–379.
- Thormann I, Dulloo ME, Engels JMM. 2005. Techniques of *ex situ* plant conservation. In Henry R, editor. *Plant Conservation Genetics.* Centre for Plant Conservation Genetics, Southern Cross University, Lismore, Australia. The Haworth Press Inc, Binghampton, NY, USA.
- UNCED. 1992. Convention on Biological Diversity. United Nations Conference on Environment and Development, Geneva, Switzerland.



- Wilson DE, Reeder DM, editors. 1993. *Mammal Species of the World: a taxonomic and geographic reference*. 2nd ed. Smithsonian Institution Press. Washington DC, USA. 1206 p.
- Yamamoto K, Sasaki T. 1997. Large-scale EST sequencing in rice. *Plant Mol. Biol.* 35:135–144.
- Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, Sasaki T. 2000. Hd1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the arabidopsis flowering time gene CONSTANS. *Plant Cell* 12:2473–2484.

## IV. Platforms for DNA banking

Christina Walters<sup>1</sup> and Robert Hanner<sup>2</sup>

<sup>1</sup> Plant Germplasm Preservation Research Unit, USDA–ARS National Center for Genetic Resources Preservation, Fort Collins, CO, USA

<sup>2</sup> Biodiversity Institute of Ontario, Department of Integrative Biology, University of Guelph, Guelph, Ontario, Canada

The accelerated use of genomic technology in biological research portends a changing pattern in the use of genetic resource collections and fuels the desire among researchers to establish DNA banks (e.g. Savolainen and Reeves 2004; Edwards et al. 2005). DNA sequencing of herbarium sheet specimens, cell cultures and germplasm, harboured at various institutions for various purposes, offers a new approach to studies of gene function, evolution, taxonomy and epidemiology among plants. A catalogue of unique genetic sequences or ‘DNA barcodes’ can conceptually unite diverse assemblages of repositories under a common registry of sequence accessions. This will enhance online access to information about species, populations or individuals and enable a broadly applicable reference database that is essential for performing DNA-based identifications on samples of unknown origin. The International Plant Genetic Resources Institute (IPGRI) offers a useful ontology for standardizing genetic information and tracking associated information about the origin, identification and subsequent uses of plant DNA (de Vicente et al. 2004).

One of the major benefits of genomic technology is fast, reliable and accurate characterization that is broadly comparable across all life stages and species. DNA banks will become the libraries for reference sequences that can be applied to estimate genetic diversity and relative change within populations or to identify samples when phenotypes confound morphologically-based identification (e.g. Olson et al. 1991; Tanksley and McCouch 1997; Avise 2004). It is anticipated that the current demand for DNA samples will keep pace with the exponentially increasing knowledge about genomes. It may also be assumed that questions about cellular differentiation, genome size, structure and gene function will become increasingly sophisticated, requiring nucleic acids unadulterated during handling and storage. The major limitation of genetic analysis involves specimen quality, which dictates the type and degree of characterization that can be successfully undertaken.

The challenge for present-day DNA repositories is to provide a product that is amenable for uses anticipated in the future, and also balancing user needs for genetic resources with institutions’ abilities to efficiently provide these materials from *ex situ* collections. The timescale for the usefulness of the sample is an important consideration, as it dictates the stringency for collection, preparation and documentation. Timescales for usefulness will also determine how large the sample must be, how (if) it can be augmented or regenerated and when (if) it can be decommissioned. Many genebanks distinguish between short-term collections, where newly obtained material is characterized and distributed based on current goals, and long-term collections, where samples are placed in safe-keeping for retrospective or population screening studies or because there is a high risk that the material will not be available in the future. Genebanks typically employ both strategies by sub-sampling specimens to meet both short-term (characterization) and long-term (archival) goals.

Long-term needs will have a decisive impact on operational organization and repository design. The first of a series of ‘Best Practices for Repositories’ focuses on human biological materials, but offers useful guidelines for the management of all collections (ISBER 2005). DNA

banks should review these general guidelines and also consider actual case studies of existing collections developed in the genomic and proteomic era (e.g. Eiseman et al. 2003). Design criteria for DNA banks should follow guidelines considered more generally for genebanks and repositories, including:

- the source of the material
- the anticipated duration of its utility
- the capacity of the bank and the total volume of the stored material and
- the institutional commitment to provide infrastructure for storage, processing and distribution.

In addition, DNA banks must delineate responsibilities for extracting DNA. For example, should this be a service by the collector, a curatorial function where the source material is stored, or a job for the investigator who uses the material? The roles of various members of the repository team will determine how DNA banks integrate into existing repositories and whether providing DNA becomes an additional service or a new mandate. Who bears the cost for the activity will be a deciding factor in how DNA banks evolve.

## Preserving DNA samples

The inherent stability of DNA and the number of pressing questions that can be addressed concerning genetic information justifies a coordinated effort to bank DNA. The quality of the DNA (molecular weight and authenticity of sequences) is the major determinant of its value in genomic surveys. The curator of the DNA bank must, therefore, preserve the material in a form that is close to the original state and document the risks and actual changes within the prescribed timescale for archival storage. Institutional policies designed to standardize DNA sample collection, extraction, characterization, distribution and archival storage will allow for a more careful evaluation of DNA quality and the effect of different storage platforms on DNA stability over time. This information is necessary for DNA banking strategies to evolve for an ever-increasing array of potential applications.

DNA is an unusually stable biomolecule that often outlasts the organism it encodes. The discovery that DNA from ancient specimens may be sufficiently intact to clone specific genes or reveal evolutionary detail is a testament to the resistance of the molecule to degradation (e.g. Goloubinoff et al. 1993; Jaenicke-Després et al. 2003). Models of degradation kinetics predict that fully hydrated DNA will depolymerize into small fragments in about 10 000 years at room temperature (Lindahl 1993). Other models suggest that a background radiation dose equivalent to 32 000 years is required to lethally destabilize chromosomes in cryogenically-stored organisms (Ashwood-Smith and Grant 1977). Practical experience tells us that DNA degrades much more quickly if it is not protected from ionizing radiation, activated oxygen or nucleases.

Non-enzymatic degradation of the primary (sequence) and secondary (helical) structure of DNA is usually initiated by cleavage of the N-glycosidic bond between the sugar backbone and the base and is followed by oxidative reactions that fragment the molecule and make the bases susceptible to deamination (Britt 1996). As the DNA template becomes progressively fragmented, it is increasingly difficult to amplify large molecular weight products through polymerase chain reactions (PCR) because primers hybridize poorly with shorter template segments and modified bases (Pääbo et al. 2004; Halliwell and Gutteridge 1999).

The quality of DNA extracted from plant specimens is dependent on the condition of the specimen before storage, the storage environment and the duration of storage. General guidelines on how to collect, store and transport samples from the field that are intended for molecular analyses can be found elsewhere (Prendini et al. 2002). The physical appearance of the sample does not necessarily indicate that DNA can be extracted and amplified, although senescing tissues consistently give low yields of DNA that are difficult to amplify using standard techniques

(Taylor and Swann 1994; Rogers and Bendich 1985). Conversely DNA fragments from badly charred seeds collected from ancient burial sites appear to be more amenable to PCR amplification (e.g. Goloubinoff et al. 1993). Studies that use DNA extracted from herbarium specimens are highly revealing about the stability of DNA with time under relatively controlled storage conditions. Valuable sequence information could be obtained from samples stored for over 100 years; however, researchers generally reported low yield and low molecular weight of DNA extracted from aged herbarium specimens (Rogers and Bendich 1985; Taylor and Swann 1994; Ristaino et al. 2001; DeCastro and Menale 2004). Extraction of DNA using extraction kits is highly efficient, although procedures usually need to be modified to accommodate compositional differences among species and tissue types or when samples are highly degraded.

Stability of DNA in leaf and seed tissues has not been compared directly. DNA is more commonly extracted from leaf herbarium samples of 150 years or less or mummified seed samples of 500 years or more. Our experience extracting DNA from seeds stored for about 130 years in an attic was similar to reports using aged herbarium samples: low yield and highly fragmented DNA (Walters et al. in prep). However, DNA extracted from dead seeds that were stored for 70 years, mostly under ambient laboratory conditions, had high molecular weight and provided a good template for PCR amplifications of the 650 bp segment coding for the internal transcribed spacer (ITS region) of the nuclear ribosomal RNA gene family (Walters et al. submitted). Most species of seeds are expected to survive for 100 years if stored at  $-18^{\circ}\text{C}$  (Walters et al. 2005a), and it is expected that viable seeds will yield high quality DNA, although there is some evidence that desiccation and storage induce minor lesions that can be repaired when surviving organisms are rehydrated (Roberts 1988; Phillips et al. 2002; Osborne et al. 2002).

The stability of DNA molecules within dead cells and deteriorating tissues presents unforeseen consequences and unique opportunities for genebanks. The increased use of PCR-based DNA sequence analyses has led to a regrettable decline of meticulous field collection and archival practices associated with an earlier era of protein or germplasm-based studies that required both greater quantity and higher quality starting materials (Edwards et al. 2005). On the other hand, accessions that were deemed worthless because viability was lost (seed collections) or freezers failed (tissue samples) can now be salvaged for molecular studies (e.g. Hanner et al. 2005).

DNA preserved in DNA banks will be stored either within cells and extracted upon retrieval from storage or extracted from cells and purified before storage. The approach that is used to preserve genetic material depends on the goals of the DNA bank. When the object is to isolate DNA from cells, the yield, purity, molecular weight and amplification efficiency are the major considerations. Storage of cells and tissues, rather than purified DNA, provide added value for diagnostic tests, epidemiological studies and identification of gene products (i.e. Austin et al. 1996; Ristaino et al. 2001; Edwards et al. 2005), and should be considered for rare specimens that produce germplasm that is difficult to preserve or grow. 'Immortalized' cells should be considered if biochemical and molecular studies of living cells accompany the genomic studies, if DNA supplies need to be replenished occasionally, or if the sample is rare (e.g. Austin et al. 1996; Visvikis et al. 1998).

Preservation stresses (drying, freezing or time) cause some damage to DNA, but most chromosomal aberrations are repaired in surviving cells after a few cell divisions. Preserving immortalized cells requires that normal genome function is retained. Remarkable stability of chromosome structure and function was demonstrated by the ability of mouse spermatozoa to generate normal offspring, even though the spermatozoa had been killed by cryo-exposure or freeze-drying (Kusakabe et al. 2001). However, most reports of somatic cells show some chromosomal aberrations that are symptomatic of an overall loss in vigour in response to preservation (e.g. Phillips et al. 2002; Osborne et al. 2002). When cells are killed by preservation stresses, enzymes that rapidly degrade DNA are released. In some cases, cells initially appear

to survive preservation stresses, but then fail to fully recover and eventually dwindle or lyse. This type of morbidity has been linked to programmed cell death (PCD), a process analogous to apoptosis in stressed mammal cells (Earnshaw et al. 1999). Non-specific damage to DNA or PCD has been demonstrated following cryo-exposure of animal cells (e.g. Baust 2003; Men et al. 2003) and has been indicted in cryo-exposed vegetative cells (Benson 1999).

Hydrated cells must be cryogenically stored to maintain viability and cell structure. 'Best practices' recommend storage below a critical temperature (ISBER 2005), which is usually interpreted to be the glass transition temperature for pure water ( $\sim$ -140 °C) (Franks 1985). Below the glass transition temperature, molecular motions that promote ice formation or deteriorating reactions are severely limited on a practical timescale. Cells that are partially dehydrated by air drying, freeze desiccation or by the addition of cryo-protectants will likely have higher (less negative) glass transition temperatures, but the critical temperature for storage has not been studied. The fate of DNA at post-thaw stages has been correlated with pre-exposure conditions (rate of drying or entry into liquid nitrogen), storage duration and overall tolerance of cells to stress (e.g. Benson 1999; Men et al. 2003; Baust 2003)

Preservation treatments of hydrated samples can be less stringent if full genome function is not required. The quality of DNA declines within days in hydrated samples held at room temperature or in refrigerators. Drying the sample or storing it in freezers or liquid achieves better preservation of DNA molecular size. Most of the information available on the stability of DNA in hydrated cells comes from studies of red blood cells, which require some processing to remove cations and inhibitors of PCR reactions (Austin et al. 1996; Visvikis et al. 1998). There is little information on the long-term stability of DNA during frozen storage, but most repositories consider several years to decades as realistic (e.g. Visvikis et al. 1998). Because intracellular ice formation does not appear to affect DNA yield or molecular weight, the critical temperature for frozen materials is the melting temperature (usually -3 to -6 °C), not the glass transition temperature that is considered the standard when maintaining cell viability. Storage at temperatures less than the critical value will likely provide longer shelf-life; however, the relationship between subzero temperature and the kinetics of DNA degradation is unexplored and so predictions of the extent to which shelf-life is prolonged are speculative.

Decisions about using conventional freezers (-18 °C), -80 °C freezers and liquid nitrogen storage are usually based on space, containers, inventory needs, frequency of access and importance of preventing contamination and temperature fluctuations. Liquid nitrogen (LN<sub>2</sub>) freezers offer the advantage of resiliency because they have no mechanical compressors to fail (in case of power outages or excessive wear), but such infrastructure is more costly to implement and is primarily suited for long-term archival storage of hydrated samples. Multiple freeze-thaw cycles tend to degrade DNA in frozen mammalian cells (Visvikis et al. 1998), probably because nucleases are released during thawing. Consequently, most tissue repositories aliquot samples before they are put in storage and many laboratories put samples that are accessed frequently at -18 °C and place reserve archival samples in -80 °C storage. DNA should be extracted immediately from samples retrieved from frozen storage using protocols that inhibit nuclease activity. Procedures for handling, storing and extracting DNA in large-scale tissue banks (the literature reports predominantly on banks used for human specimens) are not standardized and require optimization (Visvikis et al. 1998). 'Best practices' manuals provide basic guidelines for repository operations (ISBER 2005).

Dried cells are becoming popular storage platforms because of the flexibility they provide for handling, distribution and storage. Archiving dried samples intended for genetic analyses began during World War II when Guthrie cards (a blood sample dried onto filter paper) were used to identify soldiers lost in battle via enzymatic assays. DNA stores well at low relative humidity, as demonstrated by the continued, extensive use of blood samples stored on Guthrie

cards (e.g., Verlingue et al. 1994; Kline et al. 2002) and the successful use of dried herbarium specimens in molecular evolution studies (Ristaino et al. 2001; Drábková et al. 2002; DeCastro and Menale 2004). Blotting tissues onto filter paper that has been treated with nuclease inhibitors, antioxidants and detergents to break up cells are suitable for DNA storage (Kline et al. 2002) and purportedly extend the shelf-life of DNA in dried specimens (see manufacturer's descriptions for FTA paper (Whatman, Florham Park, NJ) and Isocode paper (Schleicher & Schuell, Keene, NH)). Tissue blots on specialized filter paper should be stored at 30% relative humidity to maximize the shelf-life of samples (Genvault, Carlsbad, CA), which is also considered nearly optimum for the long-term survival of desiccation-tolerant organisms (Walters et al. 2005b). The filter paper platform was designed for storage at ambient temperatures; however, storage at lower temperatures may increase shelf life according to the same principles that govern stability of dry organisms (Walters et al. 2005b).

Literature on the stability of purified DNA dissolved in buffer suggests that the overall fragment size decreases with storage time, and that the usefulness of the specimen for PCR-based assays may be 1–2 years when stored at 4 °C, 4–7 years when stored at -18 °C and greater than 4 years when stored at -80 °C (Madisen et al. 1987; Visvikis et al. 1998). Extracted DNA, stored in buffer at -18 or -80 °C, may be appropriate for short-term goals, but there is little enthusiasm for its use in long-term projects (e.g. Madisen et al. 1987; Visvikis et al. 1998; Drábková et al. 2002). Purification procedures, which are needed to remove degrading agents and PCR inhibitors, inevitably shear DNA into smaller fragments and remove proteins that stabilize tertiary structure. Fragment sizes of 20 to 50 Kb are typical for routine DNA isolations and do not interfere with many of today's applications, although physical mapping projects and studies of DNA damage usually require larger DNA molecules of about 100–200 Kb. Methods to efficiently extract high molecular weight DNA from plants are available (Bennett et al. 2001), but the long-term stability of these molecules is unexplored. Extracted DNA shows greater stability when it is dried (e.g. Madisen et al. 1987; Trapmann et al. 2004), presumably because degrading reactions are slowed.

In summary, even though DNA is an intrinsically stable molecule, it can be easily degraded during extraction and storage. Most DNA banks store cells or tissues and extract DNA upon request. Standard storage protocols can be used, and the choice of temperature usually depends on the moisture level within the sample.

## Characterizing DNA samples

DNA, the basic code for all life forms, can be the substance that unifies biological collections of all sorts and for all purposes. Exchanging tissue samples or purified DNA for genomic analyses provides researchers with immediate access to information that may not be readily available because of the restrictions that limit germplasm exchange or prevent the spread of invasive organisms across national boundaries. DNA banking as a formal repository activity can, therefore, unify research programmes by providing a source of genetic information for indexing and retrieval.

Access to DNA that has been linked to expertly-identified biological samples can be used to better characterize and identify species, and is likely to be the most immediate need for plant DNA banks. These samples will serve as references to create a stable classification scheme that integrates resource collections with the modern information technology necessary to establish a unified and stable taxonomy (Godfray 2002). An attractive approach to species' molecular diagnostics, known as 'DNA bar-coding', proposes employing a common set of standard markers to accelerate the identification of all species (Hebert et al. 2003). Using sequence database accession numbers to cross-index genebank holdings with voucher collection databases and the

published literature, it is possible to develop a virtual platform capable of reconciling specimens and their associated histories with valid scientific names derived from authoritative sources. DNA banks are thus fundamental to organizing a comprehensive 'Encyclopaedia of Life', as envisioned by E.O. Wilson (2003). Ultimately, this research agenda derives from a need to obtain stable classifications that make information retrieval about particular taxa possible and further informs predictions about taxa based on evolutionary propinquity.

Reference sequence projects will evolve from a coarse and rather typological species identification tool to a more refined tool that places individual haplotypes within a phylogeographic framework. DNA banks will provide information on frequencies of appropriate genetic markers within populations to assess population structure and demographic trends and to assign individuals to populations using probability-based models. Stored DNA samples will also be used as reference samples for analyses of parentage or verification of specific genotypes.

Standardization of the DNA source and quality across investigators is required for the collaborative inventory of biological materials that will enhance the value of existing collections and unite their diverse missions and geographic specialties. The quality of DNA provided by DNA banks should parallel the molecular diagnostic tools available, and steps to obtain and verify that quality are needed. A major role of DNA banks will be to provide the quality standard. Because DNA banks are a new venture, there are few guidelines to help establish quality assurance standards for the physical integrity of banked samples. However, different qualities of DNA may result from different storage platforms and may undermine goals for standardization that are required to justify institutional investments. Degradation of DNA is most typically evaluated by the molecular weight of extracted DNA, which can be characterized using gel electrophoresis and reported in a structure consistent with guidelines from de Vicente et al. (2004). Efficiency of amplifying template by PCR and assessments for purity using Absorbance at 260/280 nm may be other criteria of quality (Visvikis et al. 1998). Interestingly, electrophoresis proved as informative as PCR, reverse transcription-PCR (RT-PCR) and Northern blot analysis in determining the quality of nucleic acids extracted from banked human tissue (Jewell et al. 2002). Laboratory and procedural guidelines to reduce risks of contaminating DNA and artefacts in aged specimens during PCR may also be appropriate for DNA banks (Cooper and Poinar 2000; Pääbo et al. 2004).

Extracting DNA from non-living cells will lead to the eventual exhaustion of highly requested specimens. Whole genome amplification (WGA) is a method to regenerate samples of purified DNA and extend the use of the sample (Lasken and Egholm 2004; Hughes et al. 2005). Kits for WGA are available commercially, although some evidence of allelic dropout has been observed in WGA products relative to the template DNA, particularly when the starting template is of low quality or when less than 10 ng of the template is used in the WGA process. However, a new DNA amplification technology, restriction and circularization-aided rolling circle amplification (RCA-RCA), appears promising for degraded samples, including those treated with formalin (Wang et al. 2004). Such breakthroughs will likely continue to aid in the analysis of historic specimens, but should not be viewed as a substitute for careful field collection and preservation protocols involving new material.

Poor data quality associated with both the provenance of deposited specimens and the information derived from them is a serious concern for DNA banks. Some DNA banks, such as the Integrated Primate Biomaterials and Information Resource, conduct baseline genetic characterization of their specimens as a means of quality assurance, using sequence data to help verify the taxonomic identification obtained with the specimen from the submitter (Lorenz et al. 2005). Recipients of DNA bank specimens can subsequently use this baseline information for independent verification of the sample. Moreover, collections that genetically characterize their holdings and disseminate the information in genetic databases are creating digitized taxonomic information about the specimen. The effort contributes to the reference sequence database

that will facilitate identification of biomaterials of unknown origin and also raises the profile of collections (Cambon-Thomsen 2003). For example, the Consortium for the Barcode of Life (CBOL) is a growing international collaboration to construct a comprehensive database of DNA sequences that are associated with morphological voucher specimens representing described species. Hosted by the Smithsonian Institution's National Museum of Natural History (see: <http://www.barcoding.si.edu>), this effort involves museums, herbaria, biological repositories and biodiversity inventory sites, working together with academic experts in genomics, taxonomy and computer science. Voucher specimens at participating sites become increasingly valuable because they are the only practical way to address known errors in genetic databases (Ruedas et al. 2000; Bridge et al. 2003; Harris 2003).

The 'DNA Bar-coding' initiative represents the first global effort to standardize the use of specific genetic markers and is a significant advance to calibrate molecular identifications. Bar-coding relies on short orthologous sequences for rapid, accurate, cost-effective and universally accessible species identification (Hebert et al. 2003a; 2003b). While animal species can typically be discerned using the 5' end of the mitochondrial cytochrome C oxidase (COI) gene, this locus evolves too slowly to be of such use in plants. However, recent work has shown that a combination of the nuclear internal transcribed spacer (ITS) region and the plastid *trnH-psbA* intergenic spacer are potentially useful DNA regions for applying bar-coding to flowering plants (Kress et al. 2005). ITS is the most commonly sequenced locus in species level phylogenetic investigations in plants and *trnH-psbA* is the most variable plastid region in angiosperms. In this respect, the combination of a nuclear and plastid gene for bar-coding plants will allow efficient detection of hybridization, thereby overcoming one of the major limitations of using a single maternally inherited marker as in animal bar-coding.

The National Center for Biotechnology Information (NCBI) supports the aims of CBOL and is developing taxonomic and collection support for global barcode projects, including a Web-based tool to facilitate the bulk submission of barcode-flagged data records to GenBank®. Barcode-labelled records in GenBank® should include latitude-longitude data, collection dates, collector details, and specimen vouchers that will be displayed as structured comments. Links among museums, herbaria and other collections with specialized databases that maintain collateral natural history records, photographs and descriptions are also planned. As of spring 2005, the genomic database collaborators—the DNA Data Bank of Japan (DDBJ), the European Molecular Biology Laboratory (EMBL) and GenBank®—agreed to adopt and exchange additional DNA sequence qualifiers suggested for bar-code entries (but also available for all sequence submissions) as follows:

- lat\_lon
- collection\_date
- collected\_by
- identified\_by
- fwd\_primer
- rev\_primer
- fwd\_primer\_name
- rev\_primer\_name

Electropherograms associated with barcode sequences can now be archived by the NCBI so that quality scores can be computed for inferred bases. Although not a requirement, CBOL encourages the deposition of trace files in the NCBI trace archives regardless of where the initial sequence information is deposited.

In summary, quality assurance/quality control procedures need to be developed to ensure that DNA provided by DNA banks meets standards for quality and authenticity. Associated information on the provenance of the DNA will be invaluable to its usefulness.



## Hosting information associated with DNA samples

Most repositories have databases to link information about the sample with the deposited sample. This information is crucial to the utility of the repository and to the goals of documenting the distribution of biological diversity in time and across environmental gradients (e.g. Gimaret-Carpentier et al. 2003). The vision of unifying repositories through DNA banks requires that information networks are also unified, globally accessible and updated as research results become available. The GenBank® platform is an important component of this network but, as an archival database, it rarely provides updates about specimen or locality data. The Global Biodiversity Information Facility (GBIF) was established in Copenhagen to promote standardization and aggregation of repository data via a network of participating providers (Edwards 2004). GBIF advocates a structured data ontology for biospecimen collections (e.g. Noy and McGuinness 2005) and either the simplified Darwin Core or the ABCD Schema for data exchange (<http://www.gbif.org>) to provide confidence when accessing records on a taxon, date and location. Characterizing DNA specimens using standardized markers will facilitate retrieving files for authoritative taxonomies (e.g. Index Herbariorum) and published information through linkages with GenBank®, PubMed and the NCBI Taxonomy Browser (de Vicente et al. 2004).

## Conclusions

A DNA bank is a particular type of genetic resource bank that preserves and distributes the DNA molecule and provides associated information. Existing DNA banks usually support human medical applications and provide a wealth of experience on various procedures and storage platforms. Exchanging DNA of other species will be increasingly important as specialized techniques in genomics become more taxonomically widespread and global projects, such as DNA bar-coding, seek to provide a unified information structure. While DNA banks focus primarily on the DNA molecule, resources for more labile molecular targets such as replication, transcription and translation machinery will be needed in the post-genomic era. The genebank operator must, therefore, balance future needs against present research trends.

It is therefore imperative that the method of preservation, both in the field and at the repository, maximizes the potential utility of collected samples. There are numerous ways to preserve DNA structure; however, preserving chromosome function and patterns of gene expression usually requires maintaining cell viability. Even though there is little information that documents the long-term stability of DNA and other cellular components in different storage platforms, it appears that there are several good methods available to store whole or slightly processed cells that can provide high quality DNA for many years.

DNA banks provide an accessible means for genetic characterization and electronic integration of the many existing biomaterial collections. Thus, DNA banks are most useful when paired with traditional herbarium collections and frozen and/or living tissue collections. Standardization of DNA quality and characterization is crucial to a broadly applicable system of DNA-based information on taxonomy, species distribution and quantitative traits. Linkages among genetic resource collections and information help to justify the continued maintenance and expansion of all biorepositories.

## References

- Ashwood-Smith MJ, Grant E. 1977. Genetic stability in cellular systems stored in the frozen state. *In* Elliot K and Whelan J, editors. *Freezing of Mammalian Embryos*. Ciba Foundation Symposium 52. Elsevier, Amsterdam, The Netherlands. pp. 251–272.

- Austin MA, Ordovas JM, EckfeldtTJH., Tracy R, Boerwinkle E, Lalouel JM, Printz M. 1996. Guidelines of the National Heart, Lung, and Blood Institute working group on blood drawing, processing, and storage for genetic studies. *Am. J. Epidemiol.* 144:437–441.
- Avisé JC. 2004. *Molecular Markers, Natural History and Evolution*. 2nd ed. Sinauer Associates, Sunderland, MA, USA.
- Baust JM. 2003. Molecular mechanisms of cellular demise associated with cryopreservation failure. *Cell Preserv. Technol.* 1:17–31.
- Bennett PV, Hada M, Hidema J, Lepre AM, Pope LC, Quaitte FE, Sullivan JH, Takayanagi S, Sutherland JC, Sutherland BM. 2001. Isolation of high molecular length DNA: Alfalfa, pea, rice, sorghum, soybean, and spinach. *Crop Sci.* 41:167–172.
- Benson EE. 1999. Cryopreservation. In Benson EE, editor. *Plant Conservation Biotechnology*. Taylor and Francis, London, UK. Pp. 83–95.
- Bridge PD, Roberts PJ, Spooner BM, Panchal G. 2003. On the unreliability of published DNA sequences. *New Phytol.* 160:43–48.
- Britt AB. 1996. DNA damage and repair in plants. *Ann. Rev. Plant Phys.* 47:75–100.
- Cambon-Thomsen A. 2003. Assessing the impact of biobanks. *Nature Genet.* 34:25–26.
- Cooper A, Poinar HN. 2000. Ancient DNA: Do it right or not at all. *Science* 289:1139b.
- DeCastro O, Menale B. 2004. PCR amplification of Michele Tenore's historical specimens and facility to utilize and alternative approach to resolve taxonomic problems. *Taxon* 53:147–151.
- de Vicente C, Metz T, Alercia A. 2004. Descriptors for genetic marker technologies. International Plant Genetic Resources Institute (IPGRI), Rome, Italy.
- Drábková L, Kirschner J, Vlček C. 2002. Comparison of seven DNA extraction and amplification protocols in historical herbarium specimens of Juncaceae. *Plant Mol. Biol. Rep.* 20:161–175.
- Earnshaw WC, Martins LM, Kaufmann SH. 1999. Mammalian caspases: structure, activation, substrates and functions during apoptosis. *Ann. Rev. Biochem.* 68:383–424.
- Edwards JL. 2004. Research and societal benefits of the Global Biodiversity Information Facility. *BioScience* 54:485–486.
- Edwards SV, Birks S, Brumfield RT, Hanner R. 2005. Future of avian genetic resources collections: archives of evolutionary and environmental history. *Auk* 122:14–19.
- Eiseman E, Bloom G, Brower J, Clancy N, Olmsted SS. 2003. Case studies of existing human tissue repositories: 'best practices' for a biospecimen resource for the genomic and proteomic era. Rand Corporation, Santa Monica, CA, USA.
- Franks F. 1985. *Biophysics and Biochemistry at Low Temperature*. Cambridge University Press, Cambridge, UK. 210 p.
- Gimaret-Carpentier C, Dray S, Pascal J-P. 2003. Broad-scale biodiversity pattern of the endemic tree flora of the Western Ghats (India) using canonical correlation analysis of herbarium records. *Ecography* 26:429–444.
- Godfray HCJ. 2002. Challenges for taxonomy. *Nature* 417:17–19.
- Goloubinoff P, Pääbo S, Wilson AC. 1993. Evolution of maize inferred from sequence diversity of an *Adh2* gene segment from archaeological specimens. *Proc. Natl. Acad. Sci. USA* 90:1997–2001.
- Halliwell B, Gutteridge JMC. 1999. *Free Radicals in Biology and Medicine*. 3rd ed. Oxford University Press, Oxford, UK.
- Hanner R, Corthals A, Dessauer HC. 2005. Salvage of genetically valuable tissues following a freezer failure. *Mol. Phylogenet. Evol.* 34:452–455.
- Harris JD. 2003. Can you bank on GenBank? *Trends Ecol. Evol.* 18:317–319.
- Hebert PDN, Cywinska A, Ball SL, DeWaard JR. 2003a. Biological identifications through DNA barcodes. *P. Roy. Soc. Lond. B Bio.* 270:313–321.
- Hebert PDN, Ratnasingham S, DeWaard JR. 2003b. Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *P. Roy. Soc. Lond. B Bio.* 270:S596–S599.

- Hughes S, Arneson N, Done S, Squire J. 2005. The use of whole genome amplification in the study of human disease. *Prog. Biophys. Mol. Bio.* 88:173–189.
- ISBER. 2005. Best practices for repositories I: Collection, storage and retrieval of human biological materials for research. *Cell Preserv. Technol.* 3:5–48.
- Jaenicke-Després V, Buckler ES, Smith BD, Gilbert MTP, Cooper A, Doebley J, Pääbo S. 2003. Early allelic selection in maize as revealed by ancient DNA. *Science* 302:1206–1208.
- Jewell SD, Srinivasan M, McCart LM, Williams N, Grizzle WH, LiVolsi V, MacLennan G, Sedmak DD. 2002. Analysis of the molecular quality of human tissues: an experience from the Cooperative Human Tissue Network. *Am. J. Clin. Pathol.* 118:733–741.
- Kline MC, Duewer DL, Redman JW, Butler JM, Boyer DA. 2002. Polymerase chain reaction amplification of DNA from aged blood-stains: Quantitative evaluation of the ‘suitability for purpose’ of four filter papers as archival media. *Anal. Chem.* 74:1863–1869.
- Kress JW, Wurdack KJ, Zimmer EA, Weight LA, Janzen DH. 2005. Use of DNA barcodes to identify flowering plants. *P. Natl Acad. Sci. USA* 102(23):8369–8374.
- Kusakabe H, Szczygiel MA, Whittingham DG, Yanagimachi R. 2001. Maintenance of genetic integrity in frozen and freeze-dried mouse spermatozoa. *P. Natl Acad. Sci. USA* 98:13501–13506.
- Lasken RS, Egholm M. 2004. Whole genome amplification: Abundant supplies of DNA from precious samples or clinical specimens. *Trends Biotechnol.* 21:531–535.
- Lindahl T. 1993. Instability and decay of the primary structure of DNA. *Nature* 362:709–715.
- Lorenz JG, Jackson WE, Beck JC, Hanner R. 2005. The problems and promise of DNA barcodes for species diagnosis of primate biomaterials. *Phil. Trans. R. Soc. B* 360:1869–1877.
- Madisen L, Hoar DI, Holroyd CD, Crisp M, Hodes ME. 1987. DNA banking: the effects of storage of blood and isolated DNA on the integrity of DNA. *Am. J. Med. Genet.* 27:379–390.
- Men H, Monson RL, Parrish JJ, Rutledge JJ. 2003. Detection of DNA damage in bovine metaphase II oocytes resulting from cryopreservation. *Mol. Reprod. Dev.* 64:245–250.
- Noy NF, McGuinness DL. 2005. Ontology development 101: A guide to creating your first ontology. Available from: [http://protege.stanford.edu/publications/ontology\\_development/ontology101-noy-mcguinness.html](http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html)
- Olson RR, Runstadler JA, Kocher TD. 1991. Whose larvae? *Nature* 351:357–358.
- Osborne DJ, Boubriak I, Leprince O. 2002. Rehydration of dried systems: Membranes and the nuclear genome. *In* Black M, Pritchard HW, editors. *Desiccation and Survival in Plants: Drying without Dying*. CABI Publishing, Wallingford, UK. Pp. 343–364.
- Pääbo S, Poinar H, Serre D, Jaenicke-Després V, Hebler J, Rohland N, Kuch M, Krause J, Vigilant L, Hofreiter M. 2004. Genetic analyses from ancient DNA. *Ann. Rev. Genet.* 38:645–679.
- Phillips JR, Oliver MJ, Bartels D. 2002. Molecular genetics of desiccation and tolerant systems. *In* Black M, Pritchard HW, editors. *Desiccation and Survival in Plants: Drying without Dying*. CABI Publishing, Wallingford, UK. Pp. 319–342.
- Prendini L, Hanner R, DeSalle R. 2002. Obtaining, storing and archiving specimens for molecular genetic research. *In* DeSalle R, Giribet G, Wheeler W, editors. *Techniques in Molecular Systematics and Evolution. Methods and Tools in Biosciences and Medicine*. Birkhauser, Basel, Switzerland. Pp. 176–248.
- Ristaino JB, Groves CT, Parra GR. 2001. PCR amplification of the Irish potato famine pathogen from historic specimens. *Nature* 411:695–697.
- Roberts E. 1988. Seed Aging: The genome and its expression. *In* Nooden LD, Leopold AC, editors. *Senescence and Aging in Plants*. Academic Press., New York, USA. Pp. 456–498.
- Rogers SO, Bendich AJ. 1985. Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. *Plant Mol. Biol.* 5:69–76.
- Ruedas LA, Salazar-Bravo J, Dragoo JW, Yates TL. 2000. The importance of being earnest: what, if anything, constitutes a “specimen examined?”. *Mol. Phylogenet. Evol.* 17:129–132.

- Savolainen V, Reeves G. 2004. A plea for DNA banking. *Science* 304:1445.
- Stoeckle M. 2003. Taxonomy, DNA, and the bar code of life. *BioScience* 53:2–3.
- Tanksley SD, McCouch SR. 1997. Seed banks and molecular maps: unlocking the genetic potential from the wild. *Science* 277:1063–1066.
- Taylor JW, Swann EC. 1994. DNA from herbarium specimens. *In* Herrmann B, Hummell S, editors. *Ancient DNA*. Springer-Verlag, New York, USA. Pp. 166–181.
- Trapmann S, Catalani P, Hoorfar J, Prokisch J, van Iwaarden P, Schimmel H. 2004. Development of a novel approach for the production of dried genomic DNA for use as standards for qualitative PCR testing of food-borne pathogens. *Accredit. Qual. Assur.* 9:695–699.
- Verlingue C, Mercier B, Lecoq I, Audrezet MP, Laroche D, Travert G, Ferec C. 1994. Retrospective study of the cystic-fibrosis transmembrane conductance regulator (CFTR) gene-mutations in Guthrie cards from a large cohort of neonatal screening for Cystic-Fibrosis. *Hum. Genet.* 93:429–434.
- Visvikis S, Schlenck A, Maurice M. 1998. DNA extraction and stability for epidemiological studies. *Clin. Chem. Lab. Med.* 36:551–555.
- Walters C, Wheeler LM, Grotenhuis JM. 2005a. Longevity of seeds stored in a genebank: species characteristics. *Seed Sci. Res.* 15:1–20.
- Walters C, Hill LM, Wheeler LJ. 2005b. Dying while dry: Kinetics and mechanisms of deterioration in desiccated organisms. *Integ. Comp. Biol.* 45:751–758.
- Walters C, Reilley AA, Baszczak J, Richards CM. The utility of aged seeds in DNA banks (submitted to *Seed Sci. Res.*).
- Wang G, Maher E, Brennan C, Chin L, Leo C, Kaur M, Zhu P, Rook M, Wolfe JL, Makrigiorgos GM. 2004. DNA amplification method tolerant to sample degradation. *Genome Res.* 14:2357–2366.
- Wilson EO. 2003. The encyclopaedia of life. *Trends Ecol. Evol.* 18:77–80.

# V. The role of bioinformatics in coordinating conservation efforts

Theodore Kisha<sup>1</sup> and Oliver Ryder<sup>2</sup>

<sup>1</sup> Western Regional Plant Introduction Station (WRPIS), USDA-ARS National Plant Germplasm System, Pullman, WA, USA

<sup>2</sup> Conservation and Research for Endangered Species (CRES), Zoological Society of San Diego, CA, USA

To date, bioinformatics needs in the context of conserving germplasm and genetic resources have received little attention. This is in spite of the efforts applied to the development of genome analysis tools and the expansion of Web-based data query structures (led by GenBank®; Benson et al. 2003). Even though the special needs resulting from agricultural plants, their wild relatives and endangered animals (Ryder 2005) have yet to be addressed, it is hoped that this will be done soon in a manner that will facilitate the rapid advancement of conservation bioinformatics.

## The situation of endangered animal species

One of the aims of managers of conserved endangered animal species is the retention of genetic diversity as it exists in nature. Inter-specific hybridization is discouraged (even if fertile offspring arise) and subspecies mixing using captive populations (*ex situ*) to support *in situ* populations is unsuitable, except under dire circumstances. Genetic tools have been used to support small population management of endangered animal species for at least thirty years (Ryder and Fleischer 1996). Now, this is a rapidly growing area of conservation biology. Much of the impetus for this development has come from the human genome project and plant genome sequencing projects. As a consequence, the list of species for which efforts in light sequencing coverage or intensive genome sequencing are being undertaken is expanding quickly.

Genetic studies of endangered species may benefit from this expanding collection of data and the comparative studies they facilitate. For example, studies on endangered canids (including foxes and wolves) have been facilitated by information derived from domestic dog genome sequencing efforts.

In these circumstances, annotation of genome sequences becomes an ever-increasing challenge even as the cost of sequencing projects decreases and technical obstacles are routinely handled. For example, GenBank® accepts sequences as submitted, but their use requires a level of scrutiny by investigators that may not always be practicable. In addition, collections without sufficient identifiers and associated data often lose their value and cannot be used, or are used and provide misleading data; for example, ascribing DNA sequence information to another species.

New uses for stored samples, such as is now being realized through the increasing use of DNA technology and data generated, often require new data linkages to combine information from separate sources. An initial study of elephant mitochondrial cytochrome B sequences resulted in the deposition of a sequence in GenBank®. Subsequent studies found that this sequence was an outlier. Because it was possible to return to the original sample and a sample from another tissue collected at the same time, re-sequencing determined that the original sequence had been erroneous (Fleischer et al. 2001).

It is now quite common for investigators to analyze samples from multiple individuals. However, to draw accurate information from the results, knowledge of where and when the sample was taken is crucial, and accurate assessments of whether the individuals that were

sampled are related to one another should be undertaken. Thus, for many DNA samples of endangered species for which access is limited (and often only available through a zoological or botanical institution) knowledge of kinship, as is available through stud-book records, is essential. In one instance, a study of ten individuals of a species of ape involved four samples from the same individual that had been housed at four different locations.

As a result, DNA bar-coding is receiving increasing attention (see also previous chapter). As part of the nucleotide sequence data record, documentation sufficient to relate the sequence to other, independent, systematic identifiers (e.g. museum accession information, digital images, stud-book records, and provenance information) needs to be included in the respective databases.

Some protocols are now emerging for Web-based resources to allow specialists who are not genome biologists per se to be able to assemble, analyze and evaluate data and to hypothesize with genomic sequences and gene expression data, as in the case for primates, for example. Ultimately, this facilitates efforts to annotate the human genome but, not coincidentally, provides new tools for characterizing the genetic organization and variation within the genome of endangered primates. The current estimate that up to 30% of primate taxa may be vulnerable or endangered foretells the utilization of this type of resource, when fully available, for research applications of primate adaptation, population genetics and demographic analyses.

## The challenge for using plant genetic resources

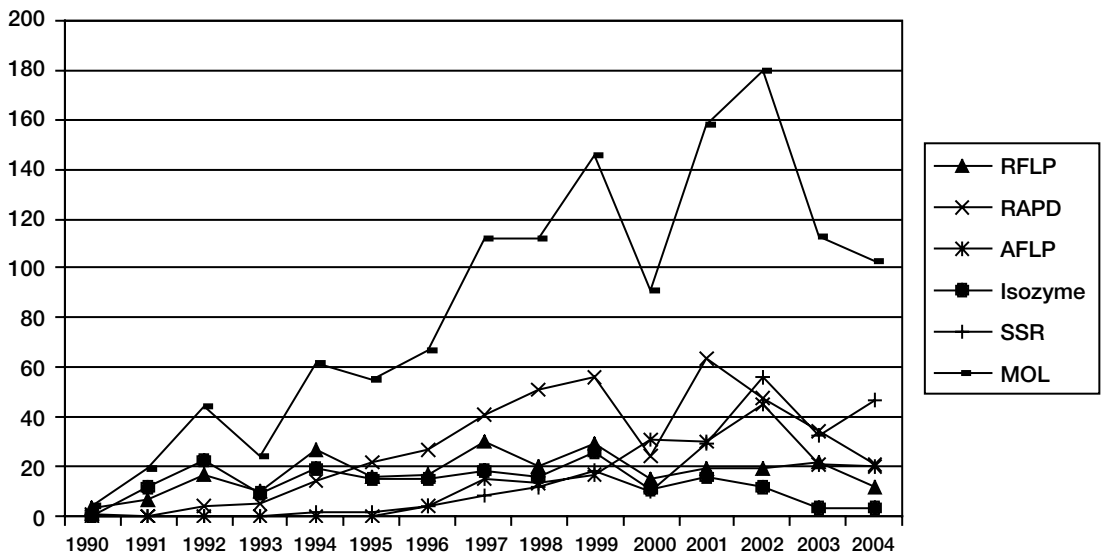
The *ex situ* conservation of plant genetic resources has received worldwide attention for many years. Agriculture in the USA, as well as in many developed countries, is primarily based on crop plants native to other parts of the world. Improvement of major crops for yield, quality and resistance to biotic and abiotic stresses must rely on new genes found in old or wild germplasm. Additionally, greater sensitivity to environmental issues has increased farmers' dependence on host plant resistances as opposed to the use of pesticides to fight diseases and pest damage. Despite some useful breakthroughs in biotechnology allowing the tapping of tertiary gene pools for genes with specific purposes, primary and secondary gene pools are still the most important sources of genetic variation for plant breeders. Germplasm collections worldwide provide genes for today's breeding efforts, while preserving other genes for future needs.

Availability of genetic diversity is of limited use, however, without the identification and characterization of that diversity, so it can be exploited and applied in an efficient manner. Transgressive segregation for quantitative traits, such as yield, in crop plants relies on the recombination of many different genes positively affecting that trait. Given the potential number of genetically distinct progeny from a single cross and the number of parents available for crossing, knowledge of parental characteristics and their relationship to one another is imperative. This is especially true when searching collections for useful traits such as pathogen resistance.

The Western Regional Plant Introduction Station (WRPIS), located on the campus of Washington State University in Pullman, Washington, USA, is part of the USDA-ARS National Plant Germplasm System (NPGS) with the responsibility of maintaining seed and clonal germplasm. Currently, WRPIS maintains over 70 000 accessions from over 2600 plant species from 376 genera. The Germplasm Resources Information Network of the NPGS (GRIN, <http://www.barc.usda.gov/psi/ngri/dbmu.htm>) documents 33 repositories in the United States, comprising 462 253 accessions from 10 768 species and 1669 genera (as of 30 Jan 2004), and these numbers increase daily. The System-wide Information Network for Genetic Resources (SINGER, <http://singer.grinfo.net/>) lists collections maintaining over 600 000 accessions from 12 of 15 international agricultural research centres of the Consultative Group on International Agricultural Research (CGIAR). Knowledgeable and effective exploitation of such large collections is difficult and their great size is a barrier to their use.

Molecular markers have become an accepted and widely used tool for the measurement of genetic diversity. The Centre for Genetic Resources (CGN) in the Netherlands (2000) collated a bibliography representing the many applications of genetic markers in managing germplasm collections. Molecular marker technology can be used to characterize the extent of diversity within a collection and for the development of collection management strategies, which may include the establishment of core collections, identification of redundancies or contamination, guidance for future collection efforts, and identification of gaps of ancestral crop relatives. In addition, analysis of worldwide genetic diversity can identify areas suited for the establishment of *in situ* conservation sites.

The amount of research dedicated to the analysis of genetic diversity using molecular markers grows every year. A search of the AGRICOLA database (<http://agricola.nal.usda.gov/>) by year using search terms 'diversity' and each respective molecular marker type is represented in Figure 1. The advent of new marker technologies is soon followed by its application to diversity analysis, but old technology continues to be exploited.



**Figure 1.** The number of manuscripts identified in the AGRICOLA database characterizing genetic diversity by marker type. (RFLP, restriction fragment length polymorphism; RAPD, random amplified polymorphic DNA; AFLP, amplified fragment length polymorphism; SSR, simple sequence repeats (microsatellite markers); and MOL, total for all molecular markers)

## Storage of molecular data

Few, if any, studies in genetic diversity can be directly compared or compiled. One of the reasons may be due to the fact that most studies are limited to a few accessions or to accessions from a limited area of interest. As a consequence, after publication, the marker data may be lost or forgotten.

Years ago, the same was true for DNA sequence data. Researchers all over the world were sequencing DNA, and there was no central data repository that scientists could use to compare results or to measure diversity. In 1988, the National Center for Biotechnology Information (NCBI; <http://www.ncbi.nih.gov/Sitemap/index.html>) was created as a division of the National Library of Medicine to create and maintain databases to enhance biomedical research. The NCBI assumed

responsibility for GenBank® in 1992 (Benson et al. 2003), and the number of DNA sequences has since risen from a few thousand to over 40 million at the end of 2004 (NCBI 2005). Moreover, the number of links and related databases at the NCBI increases as sequence data are analyzed and new uses for the data emerge. For example, phylogenetic studies based on sequence similarity can now be run from the existing database. 'Virtual' cluster analyses are now possible using tools available at the NCBI Web site, without having to duplicate costly sequence information.

However, the state of molecular marker data, fundamental to the study of genetic biodiversity, is lagging behind. Virtual cluster analyses based on the comparison of new accessions to a complete database of accrued marker information would result in savings of both time and money. Relationship queries can be adjusted to filter data based on geographical regions, environments, latitude, etc., if the database is formatted to contain such information.

GenBank®, and the organization of the NCBI Web site is a model for international cooperation among agencies and institutions. A molecular marker database would have to follow the example of that model. It would need to be a curated database because of the somewhat imprecise nature of naming markers based on fragment size. A centre or collaborating centres within the network should be responsible for a particular species and define a core set of primers for each marker type, covering the genome randomly and uniformly. A number of 'reference' accessions with defined markers should be included, so that virtual analysis could be anchored, and an image defined of the expected marker pattern with monomorphic and polymorphic bands to ease the construction of an input file with which to conduct a virtual cluster analysis and for submission to the database. As with sequence submissions, the file would be linked to any publications resulting from the analysis.

The database should compile individual, marker-centred accession files akin to sequence-centred GenBank® flat files. Formatting should include the species, marker type (information should be available for comparison with any technology currently in use), primer(s) used for RAPD or AFLP studies, or primer sequences and locus names for microsatellite studies, as well as any passport data or descriptors deemed valuable for filtered studies (i.e. inclusion/exclusion based on origin, or other descriptors). Data for inbred lines would contain loci defined by primer, fragment size and '0' or '1' to indicate the presence or absence of the marker respectively, while data for synthetic populations, obligate out-crossers, or land races would contain allele frequencies.

## Conclusion

The number of benefits for the conservation and use of genetic resources that can be drawn from available molecular data are almost limitless. However, at present a common platform for storage and analysis of genetic resources marker data is not available. The construction of a universal molecular database may seem like a daunting task, but it can come to fruition by the construction of locally created databases developed through collaborative efforts among members of germplasm conservation centres. The first step is to come to a consensus on the format of such locally generated databases for eventual linkage to a worldwide network.

## References

- Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostelland J, Wheeler DL. 2003. GenBank. *Nucleic Acids Res.* 31:23–27.
- Centre for Genetic Resources, The Netherlands. 2000. Applications of Genetic Markers (van Treuren R, editor). Available from: <http://www.cgn.wageningen-ur.nl/pgpr/research/molgen/applicts.htm>



- Fleischer RC, Perry EA, Muralidharan K, Stevens EE, Wemmer CM. 2001. Phylogeography of the Asian elephant (*Elephas maximus*) based on mitochondrial DNA. *Evolution* 55(9):1882–92.
- National Center for Biotechnology Information. 2005. GenBankStats. Available from: <http://www.ncbi.nlm.nih.gov/Genbank/genbankstats.html>
- Ryder OA. 2005. Conservation genomics: applying whole genome studies to species conservation efforts. *Cytogenet. Genome Res.* 108:6–15.
- Ryder OA, Fleischer RC. 1996. Genetic research and its application in zoos. *In* Kleiman DG, Allen ME, Thompson KV, Lumpkin S, Harris H, editors. *Wild Mammals in Captivity: Principles and Techniques*. The University of Chicago Press, Chicago, IL, USA. Pp. 255–262.

# VI. DNA banks: a primary resource for conservation research

Nicole Rice,<sup>1</sup> Robert Henry<sup>2</sup> and Maurizio Rossetto<sup>3</sup>

<sup>1</sup> Australian Plant DNA Bank, Southern Cross University, Lismore NSW, Australia

<sup>2</sup> Centre for Plant Conservation Genetics, Southern Cross University, Lismore NSW, Australia

<sup>3</sup> National Herbarium of NSW, Botanic Gardens Trust, Sydney NSW, Australia

The conservation of biodiversity is a global concern and an increasing amount of attention is focused on measuring the extent of diversity and monitoring the potential fate of these natural resources (Orr 2003). There are many different approaches to the conservation of biodiversity, including combinations of both *in situ* and *ex situ* techniques, but before it can be conserved, biodiversity needs to be adequately catalogued. Conservation activities are expensive and time-consuming. However, a balance of all available methods is advisable to ensure that the long-term goals are achieved (Bowen 1999).

The advancement and development of techniques in molecular biology and genetics has seen an increase in the ability to quantify and qualify diversity, and in the use of this information for the purpose of conservation. Molecular analyses are particularly useful in determining bottlenecks and historical events, population structure and phylogenetic and phylogeographic relationships (Frankham et al. 2002). As a result many research groups, through their use of molecular analyses, are developing large archives of extracted genomic DNA.

In the late 1980s, efforts were made to encourage groups and institutions to establish repositories of genomic DNA in an attempt to complement other conservation efforts and by 1992, 31 institutions were participating in a network of DNA banks (Adams 1997). A number of DNA banks have now been established (Table 1).

While the previous Chapters IV and V discussed technical aspects of DNA banking, this chapter will focus on the potential application of plant DNA banks as a primary resource for conservation research.

## Plant DNA Banks in context

Traditional plant conservation activities include the management of *ex situ* living collections in seed banks and botanic gardens, and specimen collections in herbaria and museums. It could be argued that these can serve as sources of DNA for the purpose of molecular analyses. In the case of seed collections, these germplasm stores are often referred to as 'gene banks' or 'gene libraries' (Given 1994; Miglani 1998). However, these same terms have more recently been used to describe collections of 'extracted genomic DNA' (Mattick et al. 1992; Adams 1997). This discussion will refer to plant DNA banks as collections containing extracted plant genomic DNA (Rice in press), plant tissues for DNA extraction and products of molecular-biology-based research.

A DNA bank, therefore, conserves genomic DNA or tissue for the purpose of extracting genomic DNA and the collection cannot be readily used to re-create the species using conventional methods of plant breeding. It has been suggested that the rapid advancement of biotechnology will, in the future, enable the use of *ex situ* collections in transformation and cloning (Godwin 1997; Rabiya 2000).

The approach to store plant tissues and genomic extracts is suitable for enabling the storage of large numbers of samples securely, efficiently and cheaply (Brown et al. 1997). These large collections of plant materials and genomic extracts can then be used to support molecular-based

**Table 1.** Plant DNA banks and their Web sites

DNA Bank	Location	Web site
Missouri Botanic Gardens	Missouri, USA	<a href="http://www.mobot.org/MOBOT/research/diversity/dna_banking.htm">http://www.mobot.org/MOBOT/research/diversity/dna_banking.htm</a>
Kew Botanic Gardens	Surrey, Great Britain	<a href="http://www.kew.org/data/dnaBank/homepage.html">http://www.kew.org/data/dnaBank/homepage.html</a>
South Africa	Capetown, South Africa	<a href="http://www.nbi.ac.za/">http://www.nbi.ac.za/</a>
Brazil	Rio, Brazil	<a href="http://www.jbrj.gov.br/pesquisa/div_molecular/bancodna/index_ing.htm">http://www.jbrj.gov.br/pesquisa/div_molecular/bancodna/index_ing.htm</a>
Australian Plant DNA Bank	Lismore, Australia	<a href="http://www.biobank.com">http://www.biobank.com</a>

plant conservation research. It might be argued that if a DNA bank only stores material to be used for DNA extraction, then we may be doing little more than duplicating existing collections such as herbaria and museum specimens. However, the ultimate aims of herbaria and DNA banks are different. Herbaria maintain specimens for the purpose of identification and classification based on morphology, whereas a DNA bank keeps specimens for the purpose of DNA extraction and use in molecular based research. Although it is sometimes possible to extract genomic DNA from herbarium specimens this is not the primary purpose of the collection, and if the herbarium specimens have been treated during preservation this can make the extraction of high quality DNA difficult (Adams 1997). The advantages of storing plant material as part of a plant DNA bank collection are three-fold (Adams 1997):

- The collection provides a constant supply of material for future genomic extractions.
- Extraction of genomic DNA can be prioritized and the collection can still grow through the incorporation of new plant samples.
- If the plant material is to be stored with the aim of being used for extraction of DNA, then it will be preserved to facilitate downstream applications and may allow for improvements in techniques.

In contrast, preserving only isolated DNA, rather than combinations of plant tissue and DNA, is also more space efficient and less subject to long-term degradation brought about by initial processing approaches. However, it is not feasible to consider that a collection of plant DNA extracts will exist without the original plant material from which the genomic DNA was derived. Ideally, to confirm its validity, the material lodged within a DNA bank needs to be backed up and supported by the relevant herbarium accessions (or equivalent database and easily accessible plant samples). It may be that stronger collaborative efforts are needed among institutions such as herbaria, germplasm collections, universities and museums to ensure the rationalization of linked samples to the centres with particular expertise, e.g. genomic DNA in a DNA bank and associated seed in a seed bank with cross-references to each collection in their respective database.

## The value of plant DNA banks as a resource for conservation research

### Plant DNA banks as a conservation tool

Plant DNA banks have an enormous potential to collect and conserve a great deal of information about global plant genetic diversity. It has already been highlighted that preserving the

DNA of species is in essence about the storage of information (Mattick et al. 1992). Mattick et al. (1992) estimated that the genetic variation in different species is in the range of 1 million to 10 million base pairs and that a DNA bank holding the 10 million species on earth would have an associated 1015 pieces of information contained in the database. Despite this, little effort has been made to collect and document DNA as a genetic resource (Ryder et al. 2000). However, there are individual groups that do aim to share their collections, e.g. the Australian Plant DNA Bank and the DNA banks at Missouri Botanic Gardens (USA), Kew Botanic Gardens (UK), Kirstenbosch Botanic Garden (South Africa) and Rio de Janeiro Botanic Garden (Brazil).

The Australian Plant DNA Bank focuses on the collection of Australian native flora and aims not only to include a representative of each known species but also to capture the within species genetic variation. Similarly, the South African and Brazilian plant DNA banks also focus on the diversity within their own native flora. In contrast to this, plant DNA banks at larger institutions such as Kew Botanic Gardens and Missouri Botanic Gardens aim to include representation from as many species as possible across all countries and do not always cover the within-species variation. All of these collections support phylogenetic studies and those that have captured the within-species genetic variation can be used in population genetics.

In addition to these DNA banks, many private collections exist and it would be good if a coordinated effort could be made to share the numerous private archives of DNA that have developed through large biotechnology programmes or molecular-based research. For example, the Australian Plant DNA Bank works with local conservation groups to ensure that very rare individuals are stored so that the genetic information will be available for future generations (Rice in press).

## DNA banks as a tool to support conservation based research

It is well known how access to DNA from species and molecular analyses can enhance the conservation of the species. On the whole, the assessment of genetic variation within plant populations has made valuable contributions to conservation genetics of species (Byrne 2005). For example the Australian Plant DNA Bank in Lismore now holds samples of extracted genomic DNA that represent the entire known genetic diversity for a range of local rare species, such as *Fontainea oraria*, *Elaeocarpus williamsianus*, *Eidothea hardeniana*, and well over 50% of the known diversity of other rare species, such as *Elaeocarpus* sp 'Rocky Creek'. The availability of this stored DNA has already enabled the development of a series of conservation-focused studies on these extremely rare species (Rossetto et al. 2000; Rossetto et al. 2004a,b; Rossetto and Kooyman 2005).

In the Australian genus *Acronychia*, DNA-based techniques have been used to clarify its taxonomic status (Rossetto 2005). This also has important implications for the use of this molecular data for the development of conservation strategies. A study on *Fontainea oraria* used random amplified polymorphic DNA (RAPD) to elucidate the relationships between individuals within a population (Rossetto et al. 2000). This study was particularly important in relation to the development of conservation strategies since the molecular data illustrated that only a few adults were contributing to the next generation of seedlings.

Molecular biology is also useful in estimating the levels of clonality (Rossetto et al. 2004a), and in assessing the effects of habitat fragmentation on dispersal (Rossetto et al. 2004b). Knowledge and application of DNA based microsatellite markers developed in closely related economic species can often be used to establish the effects and extent of habitat fragmentation on population structure. For example, microsatellite loci developed from grape (*Vitis vinifera*) have been used in molecular studies in the closely related Australian members of the Vitaceae (*Cissus hypoglauca* and *Cissus sterculifolia*) (Arnold et al. 2002).

It cannot be overlooked that genomic DNA samples will also be valuable if a species becomes extinct or its numbers significantly decline. For species that have declined since the collection of DNA material, the available samples will allow access to information on pre-decline levels of genetic diversity (Rabiya 2000), as long as the collections are representative and good records of origin are kept. It could be argued that other collections already represent the snapshot of historical diversity, but often the amount of material available for DNA extraction is limited and the preservation process may render it unsuitable (Adams 1997).

Plant DNA banks do have an important role to play in aiding the conservation of biodiversity by providing facilitated access to the raw material for molecular research. They also preserve genetic information for future research, i.e. they provide an insurance policy to ensure that the full extent of the information about a species or plant family is not lost after they become extinct. It has been argued that the lack of emphasis placed on the value of DNA banks as a conservation tool may be appropriate at the current time, but the lack of knowledge of their existence and scope might also be a hindrance and perhaps the current generation does not value the legacy they would leave for the future (Ryder et al. 2000). One potential of DNA collections is that they may assist with easing the pressure placed on natural ecosystems from disturbances that multiple collecting trips may incur (Ryder et al. 2000). Similarly, accessing plant material from distant, poorly serviced areas or areas of social unrest can be costly, difficult and dangerous (especially for the researchers). A central repository of material can avoid wastage of resources that can otherwise be used for conservation research and practices. The information held in the DNA samples can support research into plant evolution and diversity. Extinction of species of interest will not necessarily hinder the research if the DNA is available.

### **Conservation of genetic resources of importance to food and agriculture**

The need to conserve species for food and agriculture has been central to many conservation efforts. Molecular analyses play an important role in the management of germplasm collections and in the selection of material for breeding programmes, which also includes crop wild relatives (Richards 2004). Recent examples of DNA analysis of the diversity of plant germplasm associated with the Australian Plant DNA Bank include studies of sugar-cane (Cordeiro et al. 2003), barley (Bundock and Henry 2004), sorghum (Dillon et al. 2005) and pine trees (Shepherd et al. 2005). Even though protection of crop wild relatives has received little attention, they represent extremely valuable resources that are critical to world food security (Henry 2005). Domestication of crop species has led to a loss of genetic diversity in the cultivated varieties leaving a gene pool with little potential to provide resistance to biotic and abiotic stresses (Nevo 2005). This is particularly relevant to current changes in environmental pressures and predicted climatic shifts. The results of molecular studies on wild barley genetic resources (*Hordeum spontaneum*) from three countries indicate that this material is genetically more diverse than landraces (National Herbarium of NSW, Botanic Gardens Trust, Sydney NSW, Australia evo 2005).

In the example of the Australian tea tree (*Melaleuca alternifolia*; Myrtaceae), which is grown to produce pharmaceutical grade oil, a comprehensive genetic study showed that although there is a great deal of genetic diversity across the species, the majority of the variation occurs within single populations (Rossetto et al. 1999). In other words, between-population genetic variation is low, thus simplifying potential germplasm collection exercises for the purpose of identifying material for breeding programmes.

By using DNA to determine phylogenetic relationships it is also possible to identify the closest potential sources for new genes in plant breeding programmes. Molecular studies in the Australian Vitaceae identified that *Cissus hypoclauca* and *Cissus sterculifolia* are closely related to

cultivated grapes and, as a result, are potentially suitable for improvement of grapes (Rossetto et al. 2001). Similarly, Australian native *Sorghum* species have been identified as a potentially good source of new genes for improvement of *Sorghum bicolor* (Dillon et al. 2001).

## **DNA bar-codes and DNA banks: is this the way forward for DNA-based conservation research?**

DNA bar-coding is the use of one or a few reference genes to screen samples in order to identify species and/or to enhance the discovery of new species (Moritz and Cicero 2004). In particular, the catch phrase 'DNA bar-codes' refers to the use of a single mitochondrial gene cytochrome c oxidase (COI) (Moritz and Cicero 2004). The terminology has been described as a bad fit as it infers that each species has a fixed characteristic like a bar-code on a supermarket product (Moritz and Cicero 2004). Although most of the preliminary work has been done on animals, recent work suggests that the same approach can be followed for plants, albeit based on the ITS region and the *trnH-psbA* spacer rather than a single region (Kress et al. 2005). The use of DNA-sequence-based fingerprinting of one or more loci has been highlighted below as a potentially important standardized tool for DNA banks. These approaches to large-scale sequencing will proceed rapidly if diverse collections of genomic DNA are available through DNA banks. A DNA bank would then be able to house the reference bar-code samples and keep track of the relevant databases. The DNA bar-code approach needs further investigation to determine its suitability as a technique for identifying individuals and as a technique for assessing genetic diversity (Moritz and Cicero 2004).

## **The challenges for Plant DNA Banks**

For DNA banks to be a useful resource for conservation, it is necessary that:

- a coordinated approach be taken, i.e. through international networks that aim to develop standardized protocols and identify priorities
- a commitment of both financial and human resources from the host institution
- appropriate collaborative networks are developed and the necessary funds are made available to ensure that research on banked DNA progresses efficiently
- each sample must have a herbarium voucher and adequate information maintained in an appropriate database and
- legal issues related to sample exchange are investigated and documented.

## **Standardized protocols for DNA banking**

In the development of standardized protocols for DNA banking, it is evident that methods for monitoring the quality of stored DNA need to be developed. In conjunction, the development of protocols for DNA-sequence-based fingerprinting of appropriate loci might enable the direct comparison of samples held in DNA banks throughout the world. An example for plant species has shown that the granule bound starch synthase I (GBSSI) is useful for the identification of individual species within the Poaceae family (McIntosh et al. 2005). Research into methods for storage of DNA requires further investigation and it is hoped that this research will explore feasible alternatives to storing frozen extracts. There is an imminent need for the ability to store extracted DNA at room temperature. This would provide the advantage of reducing the need for specialized facilities and the costs of hosting collections (Adams 1997). The other problem that should be addressed in the management of DNA banks is that the regeneration of DNA samples may no longer be possible through repeating the extraction from fresh or preserved

leaf material. Extinction of plant species is a reality and the diversity held in collections of DNA may not be represented in living *ex situ* collections, such as botanic gardens, or in herbaria. In this case, as the DNA samples are used, regeneration methods for whole genome amplification will need to be adopted to maintain the collection. However, these techniques will require further evaluation (Brown et al. 1997).

### **Databases and sample documentation**

In particular, clearly defined guidelines must be developed in regard to the amount of associated information kept with each sample. The lack of accompanying information with physical samples is a well-documented problem often associated with *ex situ* collections (Given 1994). A lack of well-defined objectives and planning can also lead to the downfall of projects like DNA banks (Given 1994; Adams 1997). Another issue is the correct taxonomic identification of samples—a reason why DNA banks need to coexist with other institutions like herbaria, museums and zoos so that the necessary expertise can be accessed (Given 1994; Powledge 1995). An obvious step towards achieving these goals is to develop databases in collaboration with those institutions that already have successful databasing procedures. Databases are central in allowing the research community to search collections for samples of interest. A Web-based database has the advantage of allowing researchers to search the collection for samples of interest and order them online. The value of bioinformatics for DNA banks and conservation has been discussed in Chapter V.

### **Legal issues**

Legislation surrounding the access and use of genetic resources is another concern that may have an influence on the extent of DNA collections and, as a result, determine their value as a conservation resource. The implementation of the United Nations Convention on Biological Diversity (CBD) by governments has seen the access and use of genetic resources change. The convention is important to prevent over-exploitation of resources, to protect ecosystems and to ensure financial returns from commercialization flow to the original owners. This has resulted in each institution developing its own material transfer agreement to ensure that the transfer of material to third parties does not leave the DNA bank's host institution liable for misuse of the samples. The Food and Agriculture Organization of the United Nations (FAO) is also in the process of implementing the International Treaty for Food and Agriculture, which has been developed from the underlying principles of the CBD. Unlike the CBD, the International Treaty covers a list of species of importance to food and agriculture and aims to monitor their use and access to them. The issues surrounding how these international agreements effect the use of genetic resources are largely unresolved and it is not uncommon for countries that feel they have not received a fair share of benefits in the past to prevent access to their genetic resources (Kaplan 1998).

### **Conclusions**

The full value of DNA banks as a resource for conservation is yet to be realized. DNA banks do complement existing *ex situ* and *in situ* conservation efforts. However, they need to be developed in conjunction with these other activities. A coordinated effort would aid the development of DNA collections and ensure that the species included represent an accurate snapshot of the genetic diversity at the time of collection. To further increase the value of these collections as research and conservation tools, DNA banks should develop clear aims in relation to the scope of their collections.

## References

- Adams RP. 1997. Conservation of DNA: DNA banking. *In* Callow JA, Ford-Lloyd BV, Newbury HJ, editors. *Biotechnology and Plant Genetic Resources: Conservation and Use*. CABI Publishing, Wallingford, UK. Pp. 163–174.
- Arnold C, Rossetto M, McNally J, Henry RJ. 2002. The potential application of SSRs characterised for grape (*Vitis vinifera*) to conservation studies in Vitaceae. *Am. J. Bot.* 89:22–28.
- Bowen BW. 1999. Preserving genes, species, or ecosystems? Healing the fractured foundations of conservation policy. *Mol. Ecol.* 8:S5–S10.
- Brown AHD, Brubaker CL, Grace JP. 1997. Regeneration of germplasm samples: wild versus cultivated plant species. *Crop Sci.* 37(1):7–13.
- Bundock P, Henry RJ. 2004. Single nucleotide polymorphism, haplotype diversity and recombination in the *isa* gene of barley. *Theor. Appl. Genet.* 109:543–551.
- Byrne CA. 2005. Genetic variation in plant populations: assessing cause and pattern. *In* Henry RJ, editor. *Plant Diversity and Evolution*. CABI Publishing, Wallingford, UK. Pp. 139–165.
- Cordeiro G, Pan YB, Henry RJ. 2003. Sugarcane microsatellites for the assessment of genetic diversity in sugarcane germplasm. *Plant Sci.* 165:181–189.
- Dillon SL, Lawrence PK, Henry RJ. 2001. The use of ribosomal ITS to determine phylogenetic relationships within *Sorghum*. *Plant Syst. Evol.* 230:97–110.
- Dillon SL, Lawrence PK, Henry RJ. 2005. The new use of *Sorghum* bicolor-derived SSR markers to evaluate genetic diversity in 17 Australian *Sorghum* species. *Plant Genet. Res.* 3:19–28.
- Frankham R, Ballou JD, Briscoe DA. 2002. *Introduction to Conservation Genetics*. Cambridge University Press, Cambridge, UK.
- Given DR. 1994. *Principles and Practices of Plant Conservation*. Chapman and Hall, London, UK.
- Godwin ID. 1997. Gene identification, isolation and transfer. *In* Callow JA, Ford-Lloyd BV, Newbury HJ, editors. *Biotechnology and Plant Genetic Resources: Conservation and Use*. CABI Publishing, Wallingford, UK. Pp. 203–234.
- Henry RJ. 2005. Conserving genetic diversity in plants of environmental, social or economic importance. *In* Henry RJ, editor. *Plant Diversity and Evolution*. CABI Publishing, Wallingford, UK. Pp. 317–325.
- Kaplan JK. 1998. Conserving the world's plants. *Agr. Res.* 46(9):4–9.
- Kress WJ, Wurdack KJ, Zimmer EA, Weight LA, Janzen DH. 2005. Use of barcodes to identify flowering plants. *PNAS* 102:8369–8374.
- Mattick JS, Ablett EM, Edmonson DL. 1992. The Gene Library: Preservation and analysis of genetic diversity in Australasia. *In* Adams RP, Adams JE, editors. *Conservation of Plant Genes: DNA Banking and In Vitro Biotechnology*. Academic Press, San Diego, USA. Pp. 15–35.
- McIntosh SR, Pacey-Miller T, Henry RJ. 2005. A universal protocol for identification of cereals. *J. Cereal Sci.* 41(1):37–46.
- Miglani GS. 1998. *Dictionary of Plant Genetics and Molecular Biology*. The Haworth Press, Binghamton, NY, USA.
- Moritz C, Cicero C. 2004. DNA Barcoding: Promise and pitfalls. *PLoS Biol.* 2(10):e354.
- Nevo E. 2005. Genomic diversity in nature and domestication. *In* Henry RJ, editor. *Plant Diversity and Evolution*. CABI Publishing, Wallingford, UK. Pp. 287–316.
- Orr D. 2003. Diversity. *Conserv. Biol.* 17(4):948–951.
- Powledge F. 1995. The food supply's safety net. If a global agricultural crises occurred, could the international germplasm community survive a run on its genebanks? *Bioscience* 45(4):235–243.
- Rabiya S. 2000. DNA Banks Noah's Ark at -200°C. HMS Beagle 84.
- Rice NF. Conservation of plant genes and the role of DNA banks. *In* Henry RJ, editor. *Plant Conservation Genetics*. Haworth Press, Binghamton, NY, USA (in press).



- Richards C. 2004. Molecular technologies for managing and using genebank collections. In de Vicente MC, editor. *The Evolving Role of Genebanks in the Fast-developing Field of Molecular Genetics*. Issues in Genetic Resources No. 11. International Plant Genetic Resources Institute (IPGRI), Rome, Italy. Pp. 13–18.
- Rossetto M. 2005. A simple molecular approach for identifying a rare *Acronychia* (Rutaceae) provides new insights on its multiple hybrid origins. *Biol. Conserv.* 121(1):35–43.
- Rossetto M, Kooyman RM. 2005. The tension between dispersal and persistence regulates the current distribution of rare paleo-endemic rainforest flora: a case study. *J. Ecol.* 93(5):906–917.
- Rossetto M, Slade RW, Baverstock PR, Henry RJ, Lee LS. 1999. Microsatellite variation and assessment of genetic structure in tea tree (*Melaleuca alternifolia* - Myrtaceae). *Mol. Ecol.* 8:633–643.
- Rossetto M, Jackes BR, Scott KD, Henry RJ. 2001. Intergeneric relationships in the Australian Vitaceae: new evidence from cpDNA analysis. *Genet. Res. Crop Evol.* 48:307–314.
- Rossetto M, Gross CL, Jones R, Hunter J. 2004a. The impact of clonality on an endangered tree (*Elaeocarpus williamsianus*) in a fragmented rainforest. *Biol. Conserv.* 117(1):33–39.
- Rossetto M, Jones R, Hunter J. 2004b. Genetic effects of rainforest fragmentation in an early successional tree (*Elaeocarpus grandis*). *Heredity* 93(6):610–618.
- Rossetto M, McNally J, Henry RJ, Hunter J, Matthes M. 2000. Conservation genetics of an endangered rainforest tree (*Fontainea oraria* - Euphorbiaceae) and implications for closely related species. *Conserv. Genet.* 1:217–229.
- Ryder OA, McLaren A, Brenner S, Zhang Y-P, Benirschke K. 2000. DNA banks for endangered animal species. *Science* 288(5464):275–277.
- Shepherd M, Cross M, Egger P, Henry RJ, Williams G. 2005. Structural and recombinational concordance in *Australes* (Genus *Pinus*; subgenus *Pinus*). *Mol. Ecol.* (in press).

# **VII. Tissue collections as a means of storing DNA: a contribution to the conservation of Colombian biodiversity**

*Juan Diego Palacio-Mejía*

Alexander von Humboldt Biological Resources Research Institute, Cali, Colombia

Traditionally, our biological knowledge has been based on what we have perceived with our senses, leading to an immense sensorial history that makes up the base of our accumulated knowledge. With the development of genetics and, later, of molecular techniques, the possibility arose to study diversity from a different perspective. Genetic studies using molecular techniques are being employed to study diversity in different fields, such as ecology, systematics, phylogeny, biogeography and population genetics (Hansson and Westerberg 2002). This has significantly increased our knowledge about biodiversity, and its conservation and use (Haig 1998). In some cases, molecular information has answered questions that could not be addressed with traditional methods or, in other cases, it has rendered a greater level of resolution.

Knowledge about genetic diversity is important for a number of reasons. For example, in the scope of international law, the documentation and molecular characterization of biodiversity is a prerequisite for a country to exercise national sovereignty over its genetic resources. The assessment of biodiversity can contribute to a more coherent application of agreements and policies in relation to the environment. Furthermore, national conservation efforts aiming at the management and use of natural populations may also benefit from molecular studies, since the viability of species depends not only on demographic factors but also on genetic ones (Rhymer and Simberloff 1996; Avise 1994; Lynch 1996).

There is thus a need to include the genetic dimension in biodiversity studies. This is usually done by means of the analysis of nucleic acids and proteins. Nucleic acids allow a wider range of applications. They are classified into deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). Plant DNA, which is the carrier of the genetic information, can be isolated from the cell nucleus, mitochondria or chloroplasts.

Advances in molecular applications have increased the interest of researchers in obtaining tissues for DNA extraction. In order to obtain DNA of good quality it is necessary that biological samples be appropriately collected, transported and stored (Prendini et al. 2002). The best way to ensure high-quality biological samples for DNA extraction is to collect them in the field. If this is not a viable option, alternative approaches need to be used (Dessauer et al. 1996). The dependence on field collections every time material is needed implies inconveniences, such as the investment of considerable costs and time, amongst others. Furthermore, considering the increased threats most of our ecosystems are facing, there is no guarantee that on return to a collection site it still remains unchanged.

Thanks to developments in forensic medicine, there are now DNA isolation techniques suitable for many different types of organic samples, independent of their age and state of conservation. This has broadened the range of options for obtaining DNA from samples stored in herbaria, botanical museums, zoos and botanical gardens. Nevertheless, in spite of advanced DNA extraction protocols, these options are sub-optimal alternatives. However, if a continuous source of material is needed to provide high-quality DNA, a properly developed biological collection must exist. These circumstances have led to the establishment of tissue and DNA collections (Palacio-Mejía 2003).

Traditionally, tissue collections were the tissues left over from research studies after the initial project had finished. Over time, the number of residual tissue samples grew and researchers were forced to identify exclusive locations for the storage of the samples. Often, these small collections were hosted in biological collections already established. Samples were commonly inadequately labelled and preserved without any particular consideration of management and storage conditions, thus generating a series of small and unorganized tissue collections.

With increased awareness of the value of these samples, some institutions, such as natural history museums, decided to centralize and organize their tissue collections and give them a proper status, independent from the traditional collections. Other institutes, lacking these legacy collections, established tissue and DNA collections designed to preserve biological material for future use (Palacio-Mejía 2003).

Apart from being a reservoir of raw material for DNA extraction, tissue collections provide an opportunity to preserve biodiversity without the need to distinguish between biological groups. They are long term, require little space and are cost-efficient. Tissue collections, therefore, are a conservation strategy complementing traditional *ex situ* and *in situ* conservation strategies.

## The Colombian experience

### Background

When the Alexander von Humboldt Biological Resources Research Institute (AvHI) started its activities in 1995, it first conducted an inventory of the state-of-the-art of biodiversity research in Colombia. The aim was to identify gaps and define a strategy to fill them, complementing continuing research efforts in the country. One of the gaps identified was the study of biodiversity from a genetic and evolutionary perspective (Instituto Humboldt 1998).

Generally, biodiversity is associated with diversity at the species level. In Colombia, biodiversity studies have traditionally focused on biodiversity at the ecosystem and species levels and, in some detailed studies, on the population level. The genetic level of biodiversity, however, has rarely been investigated in Colombia, with the exception of some isolated karyotype and molecular studies. The reasons are mainly the high costs of molecular techniques, the lack of skilled personnel and the lack of laboratory infrastructure and of biological collections for the provision of tissue samples (Instituto Humboldt 1999).

In 1998 the AvHI started a research programme dedicated to the study of the molecular genetics of biodiversity. The goal was two-fold: 1) to fill the gaps mentioned above and 2) to stimulate the acquisition of knowledge from the application of molecular techniques for biodiversity studies. As part of this programme, and with financial support of the Colombian Ministry of Environment and technical assistance of the Laboratory of Molecular Systematics of the Smithsonian Institution in Washington DC, USA, a banking facility was set up to host the national collection of tissues and its associated DNA samples from Colombia (Instituto Humboldt 1999).

Because the intention was to build an active working collection, in the same year a laboratory for molecular biology was established with support of the science funding agency of Colombia, Colciencias, and the technical assistance of the Biotechnology Unit of the International Centre for Tropical Agriculture (CIAT). The laboratory was equipped for the extraction and purification of DNA of all biological groups and to perform a number of molecular techniques based on the polymerase chain reaction. With this infrastructure, the tissue collection and the molecular biology laboratory, the AvHI incorporated the genetic dimension into its

studies of Colombian biodiversity, to the service of the national and international scientific communities. Now Colombia has the opportunity to contribute to the study of its biodiversity from a genetic perspective.

### **Management of the collection**

Different options for preserving the tissue samples were evaluated, including storage at room temperature in different types of solutions, in alcohol and dehydrated. These methods only allow short-term conservation and, therefore, the use of low temperatures was considered in order to extend the periods of preservation. Here, too, several alternatives were evaluated, including conventional refrigerators (4°C) for short-term storage (a few days) and freezers (-20 °C) for medium-term storage (up to several months). For long-term storage, the options of ultra-cold freezers (-70 to -80 °C) and cryopreservation in liquid nitrogen (-196 °C) were evaluated. In the end, given that the objective was long-term storage, cryopreservation in liquid nitrogen was chosen as the most appropriate method.

Cryopreservation in liquid nitrogen protects the tissue sample from biological and biochemical degradation. The tissue is deep-frozen from the moment it enters into the liquid nitrogen, entering a state of suspended animation. This system allows the safe storage during long periods of time and does not require electric energy (Sytsma et al. 1993). Tanks are equipped with monitoring and alarm devices, rendering this conservation method perhaps the safest option available at present.

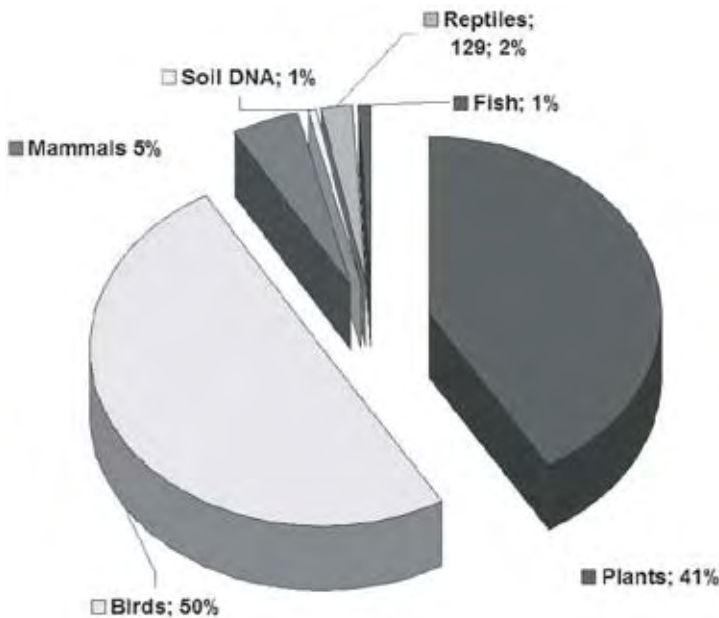
Two tanks with a capacity of 76 000 samples each were acquired, one of which was put into operation while the second serves as an emergency back up and for future expansion plans of the collection. Here, the only factor that needs to be monitored is the level of nitrogen, which has to be replenished every 15 days, costing approximately US\$ 5000 annually. It has been estimated that there are between 45 000 and 55 000 plant species in Colombia (Instituto Humboldt 1998); thus, one tank could hold a copy of the entire Colombian flora with space left for other biological groups.

### **Composition of the collection**

In the AvHI collection, tissues samples are preserved instead of DNA, although in some cases both are stored. The main interest at present is in DNA samples. This does not preclude prioritization of other types of molecules or metabolic paths in the near future, depending on scientific advances. The aim must be to conserve as much of today's genetic resources as safely as possible so that in the future, when we have access to new technologies, we will still have the biological resources to apply them. Tissue collections thus constitute *ex situ* collections with a great potential now and for future use; their consolidation is of the utmost importance, particularly in a mega-diverse country such as Colombia (Instituto Humboldt 2001).

The objective of the AvHI tissue collection is to establish a reference collection for the conservation of Colombian biodiversity from a genetic and evolutionary perspective. As such, priority was given to interspecific diversity (i.e. as many species as possible), rather than intraspecific diversity (i.e. many individuals per species). This kind of collection is very useful for systematic and phylogenetic studies. However, the need to include samples of various populations within each species was quickly recognized, especially for users interested in studies of population genetics or microevolution. Therefore, the collection has been expanded to contain a wide range of samples from different populations of each species, covering a much of the each species' geographic distribution.

At present, the AvHI collection contains approximately 6000 samples, with an average of 1000 new entries per year. Most (90%) of the collection is made up of birds and plants, while the remaining samples correspond to mammals (5%), reptiles (3%), fishes (1%) and DNA soil samples (Figure 1). Most of the samples were collected by AvHI researchers, with only a few samples being gathered by other researchers in Colombia. This might be due to the fact that few researchers in Colombia outside AvHI have been trained in collection, transport and preservation of tissue samples for molecular purposes. Another reason might be the lack of interest in molecular studies among Colombian researchers, caused by a lack of knowledge about the utility of these techniques and the opportunities offered by the country in this field (Palacio-Mejía 2003).



**Figure 1.** Species composition of the tissue collection of the Alexander von Humboldt Institute, Cali, Colombia (6100 total samples).

In order to correct this, the AvHI has developed a training manual that provides an overview of the different procedures to collect, transport and preserve tissue samples for molecular purposes. The manual gives researchers information to decide which of these procedures will be best to meet their needs, depending on the biological group sampled, logistic and economic resources, and research objectives. The manual aims at encouraging collectors to collect tissue samples for molecular purposes during 'conventional' field trips, and shows how this can be done at low cost and with minimal extra effort.

### Services provided by the AvHI Tissue Bank

Owing to the costs involved in establishing and maintaining a cryopreserved tissue collection, the AvHI collection has been established as the principal tissue bank in Colombia. The management of the collection, therefore, needs to be flexible and agile, with strategies that permit researchers of other institutions:

- to deposit their samples in trust, with the possibility of choosing among different models, according to the preferences of the respective researcher or institution, and
- to solicit samples deposited in the collection for research purposes, when complying with established requirements.

Thus, the AvHI provides the following services:

- reception of tissue samples for conservation in trust, and
- access to stored samples for research purposes.

Policies have been established defining the terms and conditions for the acceptance and safekeeping of samples, and for their access and exchange.

#### a) Criteria for receiving samples

Since the storage space at the AvHI is limited and the costs of conservation in liquid nitrogen are high, and considering the immense biological diversity present in Colombia, clear policies regulating the entry of samples in the tissue collection were defined. These were based on the recognition that biological collections need to avoid redundancy and maximize their diversity (Braun 1997). In general, priority is given to samples collected in regions of difficult access, in threatened or strategically important ecosystems, from biological groups threatened by extinction, from taxa unknown to science, as well as from those recorded for the first time at a particular site or not recently recorded at that site (Palacio-Mejía 2000a).

Because the value of a tissue sample is directly proportional to the quality and amount of information associated with it, samples are only accepted if they have complete passport data. Correct and complete passport data including the collection date, collection site, species name, morphological characteristics etc., are particularly important because tissue characteristics are not sufficient to allow their identification (as opposed to, for example, herbarium specimens, which are collected for the purpose of facilitating identification). Samples are also rejected if the material was inappropriately collected and transported and arrives in a deteriorated state. Duplicates of samples already in the collection and samples without an associated voucher are also rejected.

#### b) Safekeeping of samples

'Safekeeping' refers to the keeping in trust of samples under safe conditions for a defined period of time and granting the access to samples according to terms of custody previously agreed upon. Palacio-Mejía (2000b) described different models for the safekeeping of samples at AvHI:

- *Donation*: the samples are entered into the collection and access to them is regulated according to the general criteria established by AvHI for the exchange of samples (see below).
- *Safekeeping with notification*: AvHI may permit access to the samples, notifying the depositors each time one of their samples is requested.
- *Exclusive safekeeping*: the depositor has exclusive access to the samples. If a third party requests access to the sample, the depositor decides whether or not access will be given. This model applies for a period of a maximum of five years after samples enter the collection. After five years, the samples become donations (model 1) and belong to the general AvHI collection. This model applies mainly to samples collected for a particular research project that requires the storage of tissues or genetic material while the project is carried out.

#### c) Exchange of samples and material transfer agreement (MTA)

The tissue collection provides limited access to tissue and genetic samples to researchers for scientific purposes. Access to samples will only be provided if certain requirements are met, to compensate for the investments (i.e. costs and time) made in the collection, documentation, management and storage of tissue and genetic samples by the AvHI and the collectors (Palacio-Mejía 2000c).

In contrast to other strategies of *ex situ* conservation, such as seed, field or *in vitro* collections, tissue collection samples cannot be regenerated or replenished and thus are used up over time (Braun 1997). Therefore, based on experiences of other tissue collections, the AvHI has established terms of utilization in the form of a material transfer agreement (MTA; available on request).

Requests for access to the collection should be directed to the curator of the tissue collection in written form. Students need to be authorized by their thesis supervisor or laboratory leader, who will assume the responsibility for compliance with the regulations for access and utilization of the tissue samples (Palacio-Mejía 2000c).

## Conclusions

Tissue collections are useful for the preservation of biological samples for genetic studies, contributing information towards a better understanding of biodiversity and decisions related to its conservation and use. It is, therefore, necessary to identify low-cost long-term options for *ex situ* storage of biological materials that guarantee optimum quality and integrity of the samples for their future use. In Colombia, this task has been assumed by the Alexander von Humboldt Biological Resources Research Institute (AvHI) through the establishment of a tissue collection for the conservation of the biological diversity and for the service of the scientific community.

## References

- Avise JC. 1994. *Molecular Markers, Natural History and Evolution*. Chapman and Hall, New York, USA. Pp. 361–401.
- Braun M. 1997. Sample request policy, Avian Genetic Resources Collection, National Museum of Natural History (USNM), Smithsonian Institution, Washington DC, USA. Available from: <http://www.lms.si.edu/tissuepolicy.html>
- Dessauer HC, Cole CJ, Hafner MS. 1996. Collection and storage of tissue. *In* Hillis DM, Moritz C, Mable BK, editors. *Molecular Systematics*. Sinauer Associates, MA, USA. Pp. 29–41.
- Haig SM. 1998. Molecular contributions to conservation. *Ecology* 79(2):413–425.
- Hansson B, Westerberg L. 2002. On the correlation between heterozygosity and fitness in natural populations. *Mol. Ecol.* 11:2467–2474.
- Instituto Humboldt. 1998. Informe 1997. Instituto Humboldt, Santafé de Bogotá, Colombia.
- Instituto Humboldt. 1999. Informe anual 1998. Instituto Humboldt, Santafé de Bogotá, Colombia. 84 p.
- Instituto Humboldt. 2001. IV Congreso Interno. Instituto Humboldt, Villa de Leyva, Colombia. 35 p.
- Lynch M. 1996. A quantitative-genetic perspective on conservation issues. *In* Avise J, Hamrick J, editors. *Conservation Genetics: Case Histories from Nature*. Chapman and Hall, New York, USA. Pp. 471–501.
- Palacio-Mejía JD. 2000a. Criterios para la conservación de muestras en el banco de tejidos. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Palmira, Colombia. 2 p.
- Palacio-Mejía JD. 2000b. Condiciones para la custodia de especímenes. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Palmira, Colombia. 5 p.
- Palacio-Mejía JD. 2000c. Política y acuerdo de transferencia de muestras y ADN del banco de tejidos. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Palmira, Colombia. 4 p.

- Palacio-Mejía JD. 2003. La biología molecular y los bancos de tejidos y su papel en el tema de áreas protegidas. *In* Arango N, redactor. Memorias V congreso interno Instituto Alexander von Humboldt. Instituto Humboldt, Bogotá, Colombia. Pp. 145–158.
- Prendini L, Hanner R, DeSalle R. 2002. Obtaining, storing and archiving specimens and tissue samples for use in molecular studies. *In* DeSalle R, Giribet G, Wheeler W, editors. *Techniques in Molecular Systematics and Evolution*. Birkhauser Verlag, Basel, Switzerland. Pp. 176–248.
- Rhymer JM, Simberloff D. 1996. Extinction by hybridization and introgression. *Ann. Rev. Ecol. Syst.* 27:83–109.
- Sytsma KJ, Givnish TJ, Smith JF, Hahn WJ. 1993. Collection and storage of land plant samples for macromolecular comparisons. *In* Zimmer EA, White TJ, Cann RL, Wilson AC, editors. *Molecular Evolution: Producing the Biochemical Data*. *Methods in Enzymology* 224. Academic Press, San Diego, CA, USA. Pp. 28–37.



# VIII. DNA banking of animal genetic resources

Jinggong Xiangyu<sup>1</sup> and Ya-ping Zhang<sup>2</sup>

<sup>1</sup>Laboratory of Cellular and Molecular Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences (CAS), Kunming, China

<sup>2</sup>Molecular Biology of Domestic Animals, Kunming Institute of Zoology, Chinese Academy of Sciences (CAS), Kunming, China

Current species' extinction rates are 50–500 times higher now than for the last 65 million years (Woodruff 2001). It is estimated that one million species will be under the threat of extinction by 2050 (Thomas et al. 2004a), including many domestic animal breeds. Real species declines have been documented for British plants, butterflies and birds (Thomas et al. 2004b) and globally for amphibians (Stuart et al. 2004). Often, the reasons for the decline are not fully understood, for example, 48% of the rapidly declining amphibians are threatened due to unknown reasons in addition to habitat loss or over-utilization (Stuart et al. 2004). Genetic studies can help us to better understand the reasons for species decline and guide our efforts to preserve them in the future.

The establishment of nature reserves and national parks in as many parts of the world as possible is necessary for the *in situ* conservation of animal diversity. *Ex situ* conservation approaches, such as living collections in zoological parks, are additional important alternatives, as well as the collection and documentation of museum specimens. Considering that the extinction rates of animal species and breeds are the highest in the history of the Earth and that our understanding of the underlying reasons is limited, increasing efforts are focusing on preserving DNA samples of animal genetic resources, especially of endangered animals and animal breeds for current and future research. In this chapter, we will discuss the objectives, applications, problems and needs of animal DNA banks.

## Applications of animal DNA storage

### Phylogeny/evolutionary history

DNA banks have facilitated the efforts to reconstruct phylogeny and evolutionary history of many animal groups. For example, extensive comparative DNA sequence data identified four primary super-ordinal clades, thereby providing new insight into the pattern of early placental mammal radiation (Murphy et al. 2001). Within the bear family Ursidae, combined mtDNA and nuclear DNA data showed that the spectacled bear was the earliest divergence. The brown and polar bears and Asiatic and American black bears clustered as sister groups respectively (Yu et al. 2004). Within the big cats genus *Panthera*, analyses of nuclear and mtDNA data revealed that *P. tigris* diverged first, followed by *P. onca*, *P. leo* and the two sister species, *P. pardus* and *P. uncia* (Yu and Zhang 2005). Similarly, molecular studies may help to clarify the so far inconsistent phylogeny of modern placental mammals.

### Assessment of genetic diversity and population structure

Knowledge about genetic diversity and population structure is the basis for our understanding of the population history of species, and hence is very helpful for the design of conservation strategies for endangered species. For example, based on allozyme data from populations of the golden snub-nosed monkey (*Rhinopithecus roxellana*), it was concluded that a bottleneck had

occurred within the last 15 000 years, and that the effective population size was about 1000 at that time. In addition, fossil evidence suggests that the species might have undergone another bottleneck during the late Pleistocene. Considering that the golden snub-nosed monkey has relatively large population sizes (10 000–20 000) compared with other snub-nosed monkeys, the present population sizes seem to have recovered after the bottlenecks (Li et al. 2003).

Similarly, only 15 polymorphic sites were found when partly sequencing the mitochondrial D-loop region of 32 individuals from four regional giant panda populations. The low genetic diversity suggests that the current panda populations underwent a severe bottleneck about 43 000 years ago. The population sizes of giant panda could increase to some extent if their habitat was well protected (Zhang et al. 2002).

### **Species identification and origin of species**

Assessment of the genetic diversity of endangered animal species can help to clarify subspecies classification and define evolutionary significant units for conservation (Ryder 1986). Traditionally, based on morphological characters and geographical distribution, the tigers (*Panthera tigris*) were divided into eight subspecies. Data from mtDNA and nuclear microsatellites showed six clades corresponding to six taxonomic units or subspecies. These findings will be helpful for the improved management and conservation of tiger populations (Luo et al. 2004).

Sequence information of the wild relatives can also provide insights about the origin of domestic animal breeds. It is well known that domestic dogs derived from the wolf, but it is still unknown precisely when and where they came from them. Savolainen et al. (2002) examined mtDNA sequences among 654 domestic dogs covering all major dog breeds worldwide. Their results suggest that domestic dogs originated from East Asia about 15 000 years ago.

## **Problems of animal DNA banks**

### **Animal recovery from preserved DNA**

Fresh animal tissues can be preserved at -80 °C, allowing simple DNA extraction. Animal species can be regenerated from cryopreserved embryos and cell lines, but not from preserved DNA, due to limitations of the technology. Hence, the main objective of animal DNA banks should be the preservation of animal genetic resources as a complementary conservation strategy. In fact, some animal DNA banks preserve germplasm in the form of extracted DNA as well as cell lines. For example, the Conservation Genome Resource Bank for Korean Wildlife (CGRB) preserves tissue, DNA, cells and semen from Korean animal species (CGRB 2005). The Southwest China Germplasm Bank of Wild Species, under construction in Yunnan province, will comprise an animal germplasm bank, a DNA bank and other sections (Kunming Institute of Botany, CAS 2005).

### **Source of DNA samples**

Fresh tissues are ideal resources for DNA extraction by common methods. Hairs are another resource for DNA extraction, but the yield of DNA can be very low and quickly degrade after extraction. Museum samples are often not suitable for the extraction of high-molecular-weight DNA due to degradation of the genetic material. It is expected, though, that rapid technological advances will make hairs, faeces and museum samples useful resources for animal DNA banks in the near future. In fact, it is already possible to amplify more than 78% of the whole genome from a single cell using random primers (Zhang et al. 1992) and there is another method that allows the amplification of degraded samples (about 200bp) for whole genome analysis (Wang et al. 2004).

## **Preservation of animal DNA samples**

Animal DNA banks preserve purified high-molecular-weight DNA and biological samples for DNA extraction, such as fresh blood, muscle, kidney, heart and hairs, among others. Fresh tissues are preserved at  $-80\text{ }^{\circ}\text{C}$  and DNA at  $4\text{ }^{\circ}\text{C}$ ,  $-20\text{ }^{\circ}\text{C}$  or  $-80\text{ }^{\circ}\text{C}$ . The storage of tissue samples rather than of isolated DNA is considered to be a better preservation means to reduce DNA damage. However, dry purified DNA may be preserved at room temperature for many years (Ryder et al. 2000). Frequent monitoring of the quality of preserved DNA samples is required, and research on the conditions of long-term preserved DNA should be a major goal of DNA banks. Chapter IV presents a detailed review of DNA preservation techniques.

## **Economic issues**

Costs for the establishment and maintenance of animal DNA banks vary depending on the geographic location where the DNA bank is established and consequently are difficult to determine with accuracy. Ryder et al. (2000) estimated that the annual costs per site in the United States were about US\$ 50 000–80 000 for unprocessed specimens. Saving tissues of all tropical rainforest species in liquid nitrogen would cost about US\$ 20 million per year (Benford 1992). Clearly, high expenses involved in these activities hamper the establishment of DNA bank facilities worldwide, but especially in developing countries.

## **Exchange of DNA samples and legal issues**

Compared with the importation and exportation of animals, transporting DNA samples is far more convenient since often neither quarantine tests nor other time-consuming and costly procedures are required. This is in addition to the transportation fee that is considerably lower for samples. The exchange of DNA samples could thus facilitate genetic and genomic research.

Some plant DNA banks share their DNA samples with researchers. For example, the Royal Botanic Gardens at Kew (UK) maintain a Web-based DNA database and provide plant DNA samples for non-commercial purposes after signature of a material transfer agreement (MTA) (Royal Botanic Gardens, Kew 2005). At present, no wild animal DNA bank provides a similar service. Animal scientists generally exchange DNA samples rather by personal communication, not by posting information on dedicated Web sites. This situation certainly restricts the progress on animal conservation and genetic research among teams that have similar interests. It would be highly desirable if a Web-based database, similar to that established by Kew, could be implemented as soon as possible to facilitate the exchange of animal DNA samples.

## **Infrastructures and capacity-building needs for animal DNA banks**

### **Basic laboratory infrastructure**

An animal DNA bank should include facilities for the preparation of voucher specimens and for the preservation of tissue and DNA samples, a laboratory for DNA extraction and molecular analyses, and an administrative office. Separate rooms are required for DNA extraction, molecular analyses (e.g. PCR), and the preservation of tissue and DNA samples, respectively, to avoid contamination of the DNA samples. The office should be equipped with at least two computers, one for common use and the other as a server for the DNA database and sample information.

## Capacity building

Establishing and operating DNA banks need several kinds of professionals, such as a genetic resources manager, taxonomists, molecular biologists, data analysts and Web-managers. It is difficult for DNA banks in developing countries to hire all the necessary expertise and support from developed countries and the international scientific community is needed to adequately train DNA bank personnel.

It is also necessary to hold regular international meetings to share information on how to obtain, extract and preserve animal DNA, and to discuss legal issues and intellectual property rights relating to the exchange of DNA samples. In this way, global efforts will benefit the preservation of the world's biodiversity now and in the future.

## Conclusions

The examples of applications above show how conservation and evolutionary research can benefit from genetic materials stored in DNA banks, and how research projects stimulate sample collection. For example, population genetic studies of endangered species require sampling covering as much of the species' distribution range as possible. Genetic materials collected in the course of such research projects would well represent the genetic diversity of the species. Therefore, the conjunction of sample collection with research projects is a good approach for DNA banking.

Although many research groups hold considerable quantities of extracted animal DNA, samples are not always shared with other conservation biologists. They are often considered as 'private belongings' rather than as a public resource. This increases the cost of efforts on conservation biology by individual teams. Establishing a Web-based facility to register animal DNA banks would be a very cost-efficient approach to promote collaborative research and advance the conservation and use of animal genetic resources. The register should include additional information about the DNA samples, for example the taxonomic status, gender, collecting site and other data. Such a register would facilitate the exchange of DNA samples among researchers and greatly benefit and enhance our understanding on conservation and evolutionary biology.

## Acknowledgements

This work is supported by the Chinese Academy of Sciences, Natural Science Foundation of China, and Ministry of Science and Technology of China.

## References

- Benford G. 1992. Saving the 'library of life'. *Proc. Natl. Acad. Sci. USA* 89:11098–11101.
- Conservation Genome Resource Bank (CGRB) for Korean Wildlife. 2005. Available from: [http://www.cgrb.org/index\\_e.htm](http://www.cgrb.org/index_e.htm)
- Kunming Institute of Botany, CAS. 2005. Southwest China Germplasm Bank of Wild Species. Available from: <http://www.kib.ac.cn/KIBEnglish/soft/91.htm>
- Li HP, Meng SJ, Men ZM, Fu YX, Zhang YP. 2003. Genetic diversity and population history of golden monkeys (*Rhinopithecus roxellana*). *Genetics* 164:269–275.
- Luo SJ, Kim JH, Johnson WE, van der Walt J, Martenson J, Yuhki N, Miquelle DG, Uphyrkina O, Goodrich JM, Quigley HB, Tilson R, Brady G, Martelli P, Subramaniam V, McDougal C, Hean S, Huang SQ, Pan W, Karanth UK, Sunquist M, Smith JL, O'Brien SJ. 2004. Phylogeography and genetic ancestry of tigers (*Panthera tigris*). *PLoS Biol.* 2:e442.

- Murphy WJ, Eizirik E, Johnson WE, Zhang YP, Ryder OA, O'Brien SJ. 2001. Molecular phylogenetics and the origins of placental mammals. *Nature* 409:614–618.
- Royal Botanic Gardens, Kew. 2005. DNA Bank Database. Available from: <http://www.rbgkew.org.uk/data/dnaBank/homepage.html>
- Ryder OA. 1986. Species conservation and systematics: the dilemma of subspecies. *Trends Ecol. Evol.* 1:9–10.
- Ryder OA, McLaren A, Brenner S, Zhang YP, Benirschke K. 2000. Preservation of DNA from endangered species. *Science* 288:725–727.
- Savolainen P, Zhang YP, Luo J, Lundeberg J, Leitner T. 2002. Genetic evidence for an East Asian origin of domestic dogs. *Science* 298:1610–1613.
- Stuart SN, Chanson JS, Cox NA, Young BE, Rodrigues AS, Fischman DL, Waller RW. 2004. Status and trends of amphibian declines and extinctions worldwide. *Science* 306:1783–1786.
- Thomas CD, Cameron A, Green RE, Bakkenes M, Beaumont LJ, Collingham YC, Erasmus BF, De Siqueira MF, Grainger A, Hannah L, Hughes L, Huntley B, Van Jaarsveld AS, Midgley GF, Miles L, Ortega-Huerta MA, Peterson AT, Phillips OL, Williams SE. 2004a. Extinction risk from climate change. *Nature* 427:145–148.
- Thomas JA, Telfer MG, Roy DB, Preston CD, Greenwood JJ, Asher J, Fox R, Clarke RT, Lawton JH. 2004b. Comparative losses of British butterflies, birds, and plants and the global extinction crisis. *Science* 303:1879–1881.
- Wang G, Maher E, Brennan C, Chin L, Leo C, Kaur M, Zhu P, Rook M, Wolfe JL, Makrigiorgos GM. 2004. DNA amplification method tolerant to sample degradation. *Genome Res.* 14:2357–2366.
- Woodruff DS. 2001. Declines of biomes and biotas and the future of evolution. *Proc. Natl. Acad. Sci. USA* 98:5471–5476.
- Yu L, Li QW, Ryder OA, Zhang YP. 2004. Phylogeny of the bears (Ursidae) based on nuclear and mitochondrial genes. *Mol. Phylogenet. Evol.* 32:480–494.
- Yu LL, Zhang YP. 2005. Phylogenetic studies of pantherine cats (Felidae) based on multiple genes, with novel application of nuclear  $\beta$ -fibrinogen intron 7 to carnivores (in press).
- Zhang L, Cui XF, Schmitt K, Hubert R, Navidi W, Arnheim N. 1992. Whole genome amplification from a single cell: implications for genetic analysis. *Proc. Natl. Acad. Sci. USA* 89:5847–5851.
- Zhang YP, Wang XX, Ryder OA, Li HP, Zhang HM, Yong YG, Wang PY. 2002. Genetic diversity and conservation of endangered animal species. *Pure Appl. Chem.* 74:575–584.

# IX. Opportunities, limitations and needs for DNA banks

Andreas W. Ebert<sup>1</sup>, Jawahir L. Karihaloo<sup>2</sup> and Márcio Elías Ferreira<sup>3</sup>

<sup>1</sup>Tropical Agricultural Research and Higher Education Centre (CATIE), Turrialba, Costa Rica

<sup>2</sup>National Research Centre on DNA Fingerprinting, NBPGR, New Delhi, India

<sup>3</sup>EMBRAPA Genetic Resources and Biotechnology, Brasília, Brazil

Concern over the dramatic loss of agrobiodiversity in farmers' fields and in nature has led to efforts to conserve plant germplasm in several *ex situ* collections maintained by national and international research centres throughout the world. Apart from conservation, routine operations of genebanks include germplasm collection and acquisition (enrichment), characterization, evaluation, documentation and supply of germplasm to users and other genebanks. Many national and regional genebank collections, however, have deficits in characterization and evaluation data of the accessions. This lack of information drastically reduces the possibility of utilizing the conserved germplasm, by both plant breeders and farmers alike. It also limits the correct judgement of the representativeness of a genebank. Characterization and, where possible, evaluation of accessions is, therefore, central to strategic germplasm conservation. DNA banks might represent an option to speed up the characterization, evaluation and utilization of genetic resources. In fact, molecular characterization programmes pursued by genebanks could become catalysts for the creation of DNA banks since these routinely generate excess DNA samples that are often stored for future reference and research.

Opportunities for DNA storage in genebanks, especially in connection with molecular characterization, are becoming more and more apparent. There are a few areas where one could already foresee an impact of DNA banks in the next few years, including its use as: 1) a resource for high throughput germplasm characterization and improved genebank management; 2) a substrate for association genetics and marker-assisted selection (MAS); 3) a promoter of germplasm information exchange, including novel services such as exchanging DNA samples and sequence information; 4) a reference basis for evolutionary and comparative genomic studies; and 5) a complementary preservation procedure aiming at gene and genome conservation. Some of these opportunities are discussed below.

## Opportunities for DNA storage in genebanks

### High throughput germplasm characterization and improved genebank management

Molecular or genetic markers are seen as descriptors that offer reproducible complementary information to the classical morphological descriptors and phenotypic data used in the characterization and agronomic evaluation of genebank accessions (de Vicente et al. 2004). Detailed genetic information of the accessions in a given collection improves genebank management in several aspects as described by Karp et al. (1997): 1) it allows for the detection of gaps in the collection, guiding new collection missions and the exchange of germplasm; 2) it provides valuable knowledge concerning molecular diversity and genetic relationships through systematic genetic fingerprinting, within and between genebanks; 3) it allows for the identification of unique genotypes of special importance to the genebank, as well as the identification of dupli-

cate accessions by genetic fingerprinting (Duplicate accessions can be bulked to prevent loss of alleles [Sackville-Hamilton et al. 2002]); 4) it also allows for monitoring the genetic stability and integrity of the accessions, detecting genetic drift, natural selection and contamination during regeneration cycles; and 5) it benefits the potential users, allowing them to identify valuable traits and types quickly based on the genetic information on the accessions.

Large genebank collections are often poorly characterized, making it difficult for plant breeders to access and exploit such collections. Using marker technologies, the genetic diversity in such collections can be systematically assessed. This information is then used to establish core collections (Hodgkin et al. 1995). These core collections or subsets of large collections contain a limited number of accessions that capture most of the genetic variability present in the entire collection. These core subsets facilitate the management of the collection and contribute to an increased utilization of the germplasm. DNA banks as hubs of molecular marker application could, therefore, play a prominent role in germplasm characterization, facilitating genebank management and germplasm use.

### **Association genetics and marker-assisted selection**

The application of molecular tools to identify genes controlling specific traits in accessions of cultivated species and their wild relatives constitutes an important new role of genebanks to enhance germplasm utilization. The introgression of identified genes into genotypes with a more desirable genetic background using marker-assisted selection (MAS) has been a component of this approach (Tanksley and Nelson 1996; Xiao et al. 1996; Ortiz and Engels 2004). The MAS scheme discussed by Causse et al. (2001) for the transfer of the five most important quantitative trait loci (QTLs) involved in the organoleptic fruit quality of tomato is an example of such an approach.

The analysis of genetic variation in germplasm collections and the proper documentation of the number and types of useful polymorphisms offer genebank curators the opportunity to estimate the value of the conserved accessions and enable them to offer specific accessions with desired characteristics to plant breeders and users in general. These users can then make an informed choice and select only those genotypes that best fit their objectives (Ortiz and Engels 2004). In a paper discussing the evolving role of genebanks, de Vicente (2004) stressed the increasing demand for specific genes and alleles at QTLs.

The rate of discovery of nucleotide variation at QTLs contributing to phenotypic variation of complex traits is expected to increase with the adoption of linkage disequilibrium and candidate gene strategies for fine mapping and cloning of QTLs (Rafalski 2002; Morgante and Salamini 2003). This approach would eliminate the requirement for structured segregating populations and genetic studies could be directly performed on the accessions deposited in the genebank (Graner et al. 2004). The feasibility of association mapping between DNA markers and agronomic traits has been successfully demonstrated in a genebank collection of 600 potato varieties. Highly significant association with QTLs for resistance to late blight and plant maturity was detected with PCR markers, specific to a major gene for resistance to late blight (Gebhardt et al. 2004). Thus, in the near future, genebanks might be asked to provide not only seed, but also DNA samples and the corresponding information on both marker and sequence data.

### **Exchange of genetic information and DNA samples**

The exchange of DNA samples will certainly facilitate genetic or genomic studies on accessions in a given genebank. It will be a lot easier to exchange DNA samples, rather than seed or vegetative propagules. Transboundary movement of seed and other plant material requires seed inspection, phytosanitary certificates and quarantine testing to ensure it is free of undesir-

able diseases and pests. Exchanging DNA samples instead of seeds avoids the need for these time-consuming and costly procedures. The risk of pathogen contamination will be simply circumvented. Moreover, transportation costs of DNA samples would be much lower than that of seed or vegetative material.

The transfer of genetic material in the form of DNA samples would, therefore, be preferable in programmes focusing on genetic and genomic studies, rather than agronomic performance. Upon the conclusion of a study based on the DNA samples and having identified a specific germplasm accession possessing a gene of interest, the user could request only that accession, thereby reducing risk, cost and time.

### **Reference basis for evolutionary and comparative genomic studies**

DNA markers can be used for the taxonomic determination of plant genetic resources for any given genus to at least the species level. In some cases, the classical taxonomy had to be revised based on new DNA and sequence data (Graner et al. 2004). Aligned nucleotide sequences can be used to make inferences about the ancestral relations between them in molecular phylogenetics (Hartl 2000). These studies shed light on patterns of species evolution at the molecular level due to natural and artificial selection. In studies with cultivated wheat using microsatellites, Khlestkina et al. (2004) demonstrated that modern wheat breeding caused a qualitative shift in genetic diversity over the past 50 years, rather than a quantitative one. They concluded that it is necessary to maintain the existing *ex situ* collections and to collect new material, in order to exploit the whole range of allelic variation.

Genetic diversity can be utilized beyond the species-boundary of the primary gene pool using genomic approaches. The development of new technologies, such as microarrays and libraries of expressed sequence tags (ESTs) facilitate the detection of genes of special interest within germplasm collections (Richards 2004). The information available on one intensively studied species or crop can be used to predict and locate genes in a completely different species. The application of comparative genomics in cereal crops of the *Poaceae* family has shown that there is a considerable conservation in gene content and gene order over 60 million years of evolution in species such as wheat, maize and rice (Richards 2004). This homeology suggests that the available sequence and linkage data in one species can be useful for mapping orthologous genes in other species. Comparative genomics also offers the possibility of exploiting the potential of wild relatives of crops in genebanks more efficiently.

### **DNA banks as a complementary conservation strategy**

DNA banking could constitute a complementary conservation strategy for safeguarding the genetic diversity of a crop's gene pool, especially if combined with *in vitro* conservation or cryopreservation. DNA banks can also serve as backup or safety duplicates of the physical seed, field or *in vitro* collections, in case of catastrophic losses. Although it is not (yet) possible to recover a plant from a DNA sample, the storage of entire genomes (total DNA) or genome fragments (genomic libraries) would permit the preservation of its valuable genetic information, thus, contributing to the objective of gene or genome conservation (Andersson 2004). With the impressive advances in molecular genetics, these preserved genes or genomes might be of high relevance in the future.

Genome conservation could play a major role for species that are currently under threat of extinction. This applies in particular to those in densely populated tropical regions that are under severe threat. Of the approximately 17 000 vascular plant species reported to exist in India, nearly 15% are under threat (Ahmedullah 1999). This is aggravated by the high level of endemism existing



in several ecologically vulnerable regions of the world. In Sikkim state of India, 60% of the plant species are endemic (Myers 1988), while 33% of the total Indian flora is endemic to the country (TERI 2004). In Brazil, only 8% of the original Atlantic Forest, a highly biodiverse tropical biome that covered the coast from north to south, has been preserved. This area has been subjected to intense human pressure for the last 500 years. Central America and Mexico together form one of the Vavilov centres of origin of cultivated crops. According to Zeven and de Wet (1982), 225 domesticated plant species have their origin in this centre of diversity, representing roughly 9% of the total number of 2489 domesticated species worldwide. Alarming high deforestation rates in the range of 2.1% per year have been reported for Central America by FAO (1993). This single factor presents a major threat to the wealth of economically important species in this region. It is unlikely that the current *in situ* or *ex situ* conservation efforts will be adequate to guarantee the survival of all these vulnerable species. DNA conservation, on the other hand, at least offers the possibility of ensuring the availability of the genome long after the plant has become extinct.

## Limitations for DNA storage in genebanks

### Plant recovery from stored DNA

A major limitation of DNA banking is the fact that technologies to regenerate plants from stored DNA are not (yet) available. Hence, DNA conservation cannot be considered as a substitute for conventional conservation strategies, but can only be seen as a complementary strategy. In general, the inability to recover living plants from stored DNA may discourage the curators of genetic resources to support investment in DNA conservation. Furthermore, if the national programmes do not envision comprehensive DNA-based research, justification of DNA conservation would not be apparent.

### Plant DNA extraction procedures

Several plant species are recalcitrant to the commonly used DNA extraction protocols. Those with high concentrations of polysaccharides, proteins, tannin and lipids need special treatment. Thus, there is a need to refine and improve existing protocols to overcome these limitations. It is possible that general extraction protocols applied to specific taxonomic groups would be refined to serve as a reference to groups of species.

### Life-span of stored DNA

Another limitation of DNA banking is the relatively short life-span of stored DNA, making it necessary to replace the DNA at frequent intervals. However, the life-span can be easily extended if various factors are optimized, such as extraction and purification procedures and the management during storage. A life-span of 9 years has been observed with DNA extracted from coffee germplasm at the Tropical Agricultural Research and Higher Education Centre (CATIE) in Costa Rica (C. Astorga, personal communication). The Royal Botanical Gardens, Kew, which has in storage over 13 000 samples of DNA from a diverse range of plant species, has been largely successful in terms of long-term storage (Royal Botanical Gardens, Kew 2004). Their extraction protocols and storage procedures could serve as a basis for the standardization and improvement of applied protocols.

Long-term conservation could also be achieved using a solid medium, such as paper, instead of a solution for DNA storage. The use of cellulose-based cards has been an efficient method of long-term storage of human blood cells. Its use for the conservation of plant DNA is also an efficient

means of inactivating pathogens and protecting plant DNA from degradation. DNA can be stored directly in the treated paper after plant tissue disruption and transference of the substrate to the paper. It could also be stored as extracted DNA, after submitting a plant tissue to an extraction protocol and transferring the nucleic acid to the paper. The DNA maintained in a conservation paper can be stored at room temperature in a silica-gel-dried container. Identification is facilitated with the use of bar-coded tags that allows for a complete recovery of the sample information. DNA stored in paper would greatly facilitate sample exchange among institutions.

## Resource and policy constraints

Besides the above-mentioned technical limitations, there are a number of cost and policy considerations that might limit the application of DNA storage as a genebanking option in some countries. Molecular marker techniques, and also DNA extraction per se, are generally quite expensive, especially in developing countries that depend on imported reagents and materials. Import prices can easily double for DNA extraction kits, enzymes and genetic markers, if all additional costs, such as customs handling and storage fees of goods are considered. In the case of importation of dangerous laboratory reagents to Central American countries, such as acetic acid, ethanol, chloroform, ethidium bromide and silver nitrate, additional costs of US\$ 100 per product are levied by the airline carriers in the context of the new Bio-terrorism Act implemented by the United States (C. Astorga, personal communication). Additional problems are the bureaucratic hurdles to get imported perishable items such as enzymes out of customs without interruption of the cooling chain. In general, the costs of establishing a DNA laboratory and operating it may be quite prohibitive for many small, resource-poor developing countries. Though the consumable costs have been declining lately, DNA conservation may not be affordable for most national genebanks located in the tropics.

The availability of liquid nitrogen might also be a limitation in some countries and an important additional cost factor. Manufacturers of liquid nitrogen are often based in major cities, thus requiring transportation over long distances, if the genebank or laboratory is located in a remote place. Uninterrupted power supply is another major concern for many developing countries in Africa, Asia and Latin America, because the provision of energy often does not keep pace with the steadily increasing consumption due to rapid economic and population growth. To guarantee the safety of the DNA samples, it will be necessary to install backup power supply sources. A further frequent limitation in several developing countries is the unavailability or restricted availability of molecular geneticists to genebank curators and/or the lack of connection between them (Hamon et al. 2004). Staff members in genebanks are usually specialized in phenotypic characterization of accessions and have difficulties in adopting new genomic technologies due to a lack of training and, sometimes, lack of vision. Limited access to relevant information may also prevent global application of DNA conservation technology on a sustainable basis.

It is obvious that the use of marker technologies in genebank management requires significant additional funding. Many genebanks in developing countries are struggling to survive, have insufficient human and financial resources at their disposal to provide adequate germplasm management and simply cannot afford to invest in these new technologies. This might deepen the gap between rich and poor genebanks (Graner et al. 2004). It has to be stressed that the conservation of genetic resources and their proper management continues to be a very important task of genebanks, not only for the provision of basic material for the work of geneticists but also for the provision of phenotypic data, which are essential for the utilization of genetic information in breeding programmes. It is important to stimulate and assure the existence of genebanks in diverse environments and countries for the sake of plant diversity conservation and for the good of humanity.

## Intellectual property and legal issues

The exchange of seed and other planting materials usually requires the signature of a material transfer agreement (MTA) by the beneficiary prior to the shipment. The MTAs regulate the intellectual property rights (IPR) of the requested material and related information, the conditions of its use and distribution to third parties, as well as benefit sharing. The MTAs usually ask for the submission of information and research results obtained with the genetic material to the dispatching genebank. In most cases, the MTAs are specifically designed for the exchange of seed or vegetative propagules and do not consider IPR issues in the event that DNA samples are exchanged. According to Andersson (2004), the following institutions make explicit reference to the exchange of DNA in their MTA: CATIE, Costa Rica; the National Institute of Agrobiological Sciences (NIAS), Japan, the Missouri Botanical Garden, USA; and the Royal Botanic Gardens, Kew, UK. However, even if the MTA covers the exchange of DNA samples, there are still different interpretations concerning the question whether this precludes the patenting of specific genes or not. Therefore, legal issues related to DNA exchange and transfer require immediate attention.

There are concerns in the developing world about the protection of its intellectual property rights on genetic resources accessed by the developed countries. Despite the provision of benefit-sharing in the Convention on Biological Diversity and the International Treaty on Plant Genetic Resources for Food and Agriculture, there are only a few examples of this having been effected. It remains to be seen how effective the MTAs are in enabling benefit-sharing between the genetic resources donor countries and the agencies and companies utilizing them for commercial benefits.

## Infrastructure and capacity-building needs for DNA storage in genebanks

### Basic laboratory infrastructure

A well-lit laboratory with working tables, storage space, running water and assured power supply is required for establishing a DNA bank. For DNA extraction, purification and storage, laboratory equipment should include a fume hood, laminar flow, reverse osmosis-based water purification system, hot air oven, microwave oven, horizontal gel electrophoresis unit with power supply, UV transilluminator, UV spectrophotometer, liquid nitrogen container, refrigerator, -20 °C freezer, -80 °C deep freezer and an autoclave. It is advisable to have access to a herbarium for the storage of reference samples with passport information and, if feasible, for replenishment of DNA stocks.

For the integration of fingerprinting data on accessions into the corresponding genebank documentation system and for the exchange of compatible data sets in networks, a computer with a printer and Internet connection and a skilled database administrator are absolute musts. There is, of course, also a need for technical staff trained in DNA extraction and purification and in preparation of herbarium samples.

### Capacity building

Well-trained personnel in molecular biology and computer applications are necessary to organize and run the DNA bank. Comprehensive short-term (4–6 months) training workshops in well-known laboratories or genebanks applying long-term conservation of DNA and modern genebank management would fill this gap. The training content would comprise principles and applications of molecular biology, modern genetic resources management, genomic sci-

ences, bioinformatics, experimental design, data analysis, interpretation and utilization, and intellectual property rights of DNA exchange. Educational and training materials could also be made available through the Internet, although this would not be a substitute for practical experience and laboratory training. A Web-based newsletter covering information on current developments in DNA technologies would be very useful.

While some globally reputed laboratories involved in DNA studies would be centres of choice for such training, it would be desirable to develop strategically located regional training centres in developing regions. Selected laboratories in these regions that are already involved in DNA research could be identified for catering to the training needs of neighbouring countries.

As these new technologies are very cost-intensive and require a high level of specialization, regional collaboration and networking of countries should be encouraged. The CGIAR Centres are in a good position to promote the appropriate integration of modern biotechnology tools into genebank management. Efforts in this direction, supported by national programmes, will certainly enhance the collection (enrichment), conservation, characterization, documentation and, above all, the utilization of plant genetic resources for the benefit of humanity (de Vicente 2004).

## Conclusions

DNA banks may evolve as a strategic component of modern genebanks providing the basis for improved genebank management and facilitating high throughput germplasm characterization, association genetics and marker-assisted selection. DNA samples can be exchanged much more easily and at lower costs than can living plant materials, without the inherent risk of spreading diseases and pests. DNA banks serve as reference basis for evolutionary and comparative genomic studies and may offer a complementary conservation strategy for species under threat of extinction. They can also serve as safety duplicates for the physical seed, field or *in vitro* collections. However, there are also several limitations of a technical and legal nature to DNA storage in genebanks. The major limitations include the lack of technologies to regenerate plants from stored DNA and the relatively short life-span of stored DNA samples. Some limitations apply specifically to genebanks in resource-poor developing countries, which require support to build up and maintain the necessary infrastructure and to train human resources in DNA and molecular techniques. Regional collaboration and networking among countries is imperative, given the costs and the level of specialization required for these new technologies.

## References

- Ahmedullah A. 1999. Country report on red listing of plants in India. *In* Using the IUCN red list criteria at the national level: A regional consultative workshop for South and South Asia. Available from: <http://www.biodiversityasia.org/ebook/html>. Pp. 20–22.
- Andersson MS. 2004. Feasibility study for DNA storage and use: DNA banks—providing novel options for genebanks? A worldwide survey. IPGRI Report No. 6900/D10A. International Plant Genetic Resources Institute (IPGRI), Rome, Italy. 38 p.
- Causse M, Leconte L, Baffert N, Duffé P, Hospital F. 2001. Marker-assisted selection for the transfer of QTLs controlling fruit quality traits into tomato elite lines. Proceedings of an International Symposium on Molecular Markers. *Acta Hort* 546:557–564.
- de Vicente MC, editor. 2004. The Evolving Role of Genebanks in the Fast-developing Field of Molecular Genetics. Issues in Genetic Resources No. 11. International Plant Genetic Resources Institute (IPGRI), Rome, Italy.
- de Vicente C, Metz T, Alercia A. 2004. Descriptors for genetic markers technologies. International Plant Genetic Resources Institute, Rome, Italy. 24 p.

- FAO. 1993. Forest resources assessment, 1990. Tropical countries. Food and Agricultural Organization of the United Nations Forestry Paper 112. FAO, Rome, Italy.
- Gebhardt C, Ballvora A, Walkemeier B, Oberhagemann P, Schüler K. 2004. Assessing genetic potential in germplasm collections of crop plants by marker-trait association: a case study for potatoes with quantitative variation of resistance to late blight and maturity type. *Mol. Breed.* 13:93–102.
- Graner A, Dehmer KJ, Thiel T, Bömer A. 2004. *In de Vicente MC*, editor. The Evolving Role of Genebanks in the Fast-developing Field of Molecular Genetics. *Issues in Genetic Resources No. 11*. International Plant Genetic Resources Institute (IPGRI), Rome, Italy. Pp. 26–32.
- Hamon S, Frison E, Navarro L. 2004. Connecting plant germplasm collection and genomic centres: how to better link curators, molecular biologists and geneticists? *In de Vicente MC*, editor. The Evolving Role of Genebanks in the Fast-developing Field of Molecular Genetics. *Issues in Genetic Resources No. 11*. International Plant Genetic Resources Institute (IPGRI), Rome, Italy. Pp. 33–45.
- Hartl DL. 2000. A Primer of Population Genetics. 3rd ed. Sinauer Associates, Sunderland, MA, USA. 221 p.
- Hodgkin T, Brown AHD, van Hintum ThJL, Morales EAV. 1995. Core Collections of Plant Genetic Resources. John Wiley and Sons, Chichester, UK.
- Karp A, Kresovich S, Bhat KV, Ayad WG, Hodgkin T. 1997. Molecular tools in plant genetic resources conservation: a guide to the technologies. Technical Bulletin No. 2. International Plant Genetic Resources Institute (IPGRI), Rome, Italy.
- Khlestkina EK, Huang XQ, Quenum FJ-B, Chebotar S, Röder MS, Börner A. 2004. Genetic diversity in cultivated plants—loss or stability? *Theor. Appl. Genet.* 108:1466–1472.
- Morgante M, Salamini F. 2003. From plant genomics to breeding practice. *Curr. Opin. Plant Biotechnol.* 14:214–219.
- Myers N. 1988. Threatened biotas: 'hotspots' in tropical forestry. *The Environmentalist* 8:1–20.
- Ortiz R, Engels J. 2004. Genebank management and the potential role of molecular genetics to improve the use of conserved genetic diversity. *In de Vicente MC*, editor. The Evolving Role of Genebanks in the Fast-developing Field of Molecular Genetics. *Issues in Genetic Resources No. 11*. International Plant Genetic Resources Institute (IPGRI), Rome, Italy. Pp. 19–25.
- Rafalski A. 2002. Applications of single nucleotide polymorphisms in crop genetics. *Curr. Opin. in Plant Biol.* 5:94–100.
- Richards C. 2004. Molecular technologies for managing and using genebank collections. *In de Vicente MC*, editor. The Evolving Role of Genebanks in the Fast-developing Field of Molecular Genetics. *Issues in Genetic Resources No. 11*. International Plant Genetic Resources Institute (IPGRI), Rome, Italy. Pp. 13–18.
- Royal Botanic Gardens, Kew. 2004. DNA Bank Database. Available from: <http://www.kew.org/data/dnaBank/homepage.html>
- Sackville-Hamilton NR, Engels JMM, van Hintum ThJL, Koo B, Smale M. 2002. Accession management: combining or splitting accessions as a tool to improve germplasm management efficiency. Technical Bulletin No. 5. International Plant Genetic Resources Institute (IPGRI), Rome, Italy.
- Tanksley SD, Nelson JC. 1996. Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theor. Appl. Genet.* 92:191–203.
- TERI. 2004. Hot spots in India. Available from: <http://www.teriin.org/biodiv/hotspot.htm>
- Xiao J, Grandillo S, Ahn SN, McCouch SR, Tanksley SD, Li J, Yuan L. 1996. Genes from wild rice improve yield. *Nature* 384:223–224.
- Zeven AC, de Wet MJM. 1982. Dictionary of cultivated plants and their regions of diversity; excluding ornamentals, forest trees and lower plants. CAPD, Wageningen, The Netherlands.

# **X. A model for DNA banking to enhance the management, distribution and use of *ex situ* stored PGR**

*Andreas Graner*<sup>1</sup>, *Meike S. Andersson*<sup>2</sup> and *M. Carmen de Vicente*<sup>2</sup>

<sup>1</sup> Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany

<sup>2</sup> International Plant Genetic Resources Institute (IPGRI–Americas), Cali, Colombia

The application of biotechnologies is advancing rapidly, generating exponentially increasing amounts of DNA resources, including their related information. Standardized strategies and procedures for maintaining DNA of the type already developed for seed, tissue culture samples and field collections are urgently needed to ensure that resources and information are indeed fully used. Issues to be addressed are: how to store DNA, the amount of DNA stored, criteria and minimum standards to assure DNA quality and its maintenance over time, how to document DNA samples and how to exchange them, amongst others.

In the previous chapters, DNA banking is considered from different perspectives, such as general considerations about the role of DNA banks as a complementary *ex situ* conservation strategy, technical aspects and best practices, needs and limitations from the point of view of developing countries, and the role of bioinformatics in coordinating conservation efforts. The role of DNA banks as a complement to conventional *ex situ* conservation strategies has been demonstrated, highlighting their usefulness for efficient genebank management, conservation research and diversity assessment. The advantages and disadvantages of storing only the isolated DNA or in conjunction with tissue samples for subsequent DNA extraction have been discussed. In this scenario, the missing link is that of proposing a management structure for DNA banking that would take adequate account of conservation and use issues.

At the end of the 1980s, attempts were made to create the plant 'DNA Bank-Net' (Adams 1988; Adams and Adams 1992). In its design, plant DNA banks would have acted as reserve (base) and working (dispense) nodes and constituted a coordinated network of standardized plant DNA banking worldwide, facilitating the distribution and exchange of DNA samples. The first organizational meeting of DNA Bank-Net was held at the Royal Botanic Gardens, Kew, UK in 1991 and was attended by 18 invited scientists. A year later the network already comprised 30 institutions from all continents (Adams and Adams 1992). Unfortunately, the DNA Bank-Net initiative has remained inactive for several years (R.P. Adams, personal communication 2004).

Recently, DNA banking attracted publicity as a result of the initiatives of Ryder et al. (2000) and Savolainen and Reeves (2004), who called for a coordinated worldwide attempt to store samples of DNA for endangered species and for the establishment of a Web-based method to register DNA banks. In 2000 and 2001, international scientific meetings were held in San Diego and New York with participants from Australia, China, England, France, Scotland and South Africa that dealt with this subject (Morin 2000; Morin et al. 2000). Some progress has been made since then and the concept is still robust, but the original vision of an openly accessible Web-based list of DNA banks with accompanying information about the holdings, type of samples stored, policies for access etc. has not been realized, largely due to an unfavourable economic climate and scarce funding opportunities (O. Ryder, personal communication 2005).

Some stepping-stones are in place, including descriptors for the standardization of molecular marker information (de Vicente et al. 2004) and existing data banks supporting the central stor-

age and coordination of DNA sequence information. In this sense, one of the major databases is GenBank®, a genetic sequence database at the National Center for Biotechnology Information, National Institute of Health (NIH–NCBI), Bethesda, Maryland, USA (<http://www.ncbi.nlm.nih.gov/Genbank/index.html>) that hosts all publicly available DNA sequences (Benson et al. 2003). It is part of the International Nucleotide Sequence Database Collaboration comprised of the DNA Data Bank of Japan (DDBJ), Mishima, Japan (<http://www.ddbj.nig.ac.jp/>), the European Molecular Biology Laboratory (EMBL–EBI) Nucleotide Sequence Database, Hinxton, UK (<http://www.ebi.ac.uk/>) and GenBank®. These three organizations exchange data on a daily basis.

In this chapter, the implementation of a pilot project for a global network linking DNA resources for conservation and research is proposed. This model represents the ‘optimum’ case, i.e. the framework expected if DNA banking were already well established and fully operational. However, the implementation of such a comprehensive network would not be accomplished easily, and a first step towards the realization of the model could include a Web-based DNA information portal for registering DNA resources and available associated information. For a start, exchange of these resources could be managed on an informal basis, by direct contact between the requesting person/institution and the respective owner institution of the DNA samples. Standards for the documentation, preservation and distribution of the DNA samples could then be agreed upon and implemented in subsequent stages, as well as the formal establishment of the DNA bank consortium.

## Pilot DNA bank network

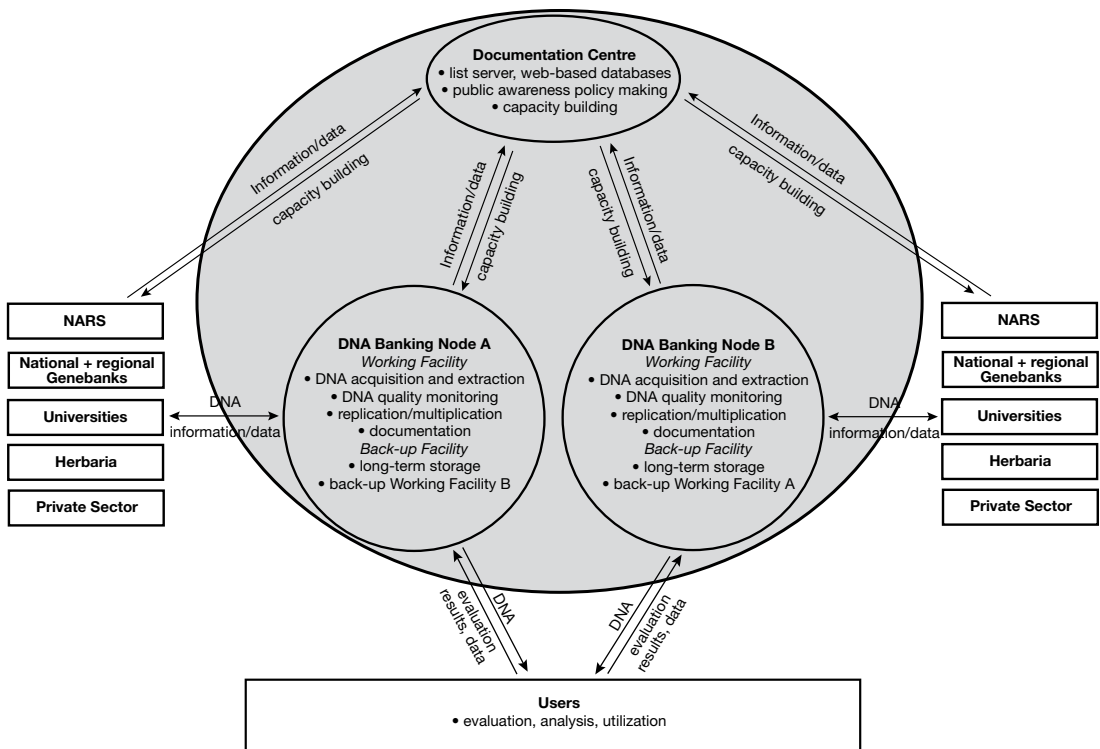
The final model would be based on a small-scale network of DNA banks, linked to the current system of *ex situ* collections in the international agricultural research centres (IARC) of the Consultative Group of International Agricultural Research (CGIAR) under the auspices of the Food and Agriculture Organization of the United Nations (FAO). It should focus on the 35 crops and crop complexes listed in Annex 1 of the International Treaty on Plant Genetic Resources for Food and Agriculture (IT) (FAO 2004). Linking such a DNA bank model to the CGIAR system is reasonable, since the CGIAR recognizes that “there is a need to apply modern and efficient means for the collection, long-term conservation, evaluation and characterization of germplasm”. Moreover, one of the specific goals among the new CGIAR priorities for the next 10 years is the assemblage of DNA banks to conserve and develop new genetic stocks of priority crops and wild relatives and to distribute them to producers and consumers (CGIAR Science Council 2005; Working Document on CGIAR 2005–2015 System Priorities. Preliminary draft. Available from: <http://www.sciencecouncil.cgiar.org/publications/pdf/SP5-15.pdf>. p. 20)

Furthermore, the DNA bank model network could and should take advantage of other synergies, for example:

- One of the primary goals of the Generation Challenge Program (<http://www.generationcp.org>) is enhancing the use of public genetic resources by promoting the application of genetic and genomic tools for crop improvement (GCP 2005). As a result, significant amounts of DNA samples are being generated that could be kept for further public use. Data belonging to these resources are being properly stored together with passport and morphological information.
- Another important initiative is the Global Conservation Trust (<http://www.startwithaseed.org>), which aims to provide a permanent source of funding for crop diversity collections around the world. The Global Conservation Trust, a campaign in partnership between IPGRI (for the CGIAR) and the FAO, might thus be interested in supporting the establishment of a DNA bank network.

- The UK government's Darwin Initiative (<http://www.defra.gov.uk/environment/darwin/>) promotes biodiversity conservation and sustainable use of resources around the world. It assists countries that are rich in biodiversity but poor in financial resources to implement the Convention on Biological Diversity (CBD) through the funding of collaborative projects that draw on UK biodiversity expertise. Amongst other projects, the Darwin Initiative supported the establishment of the DNA Bank of South African Flora at the National Botanical Institute (NBI), Kirstenbosch, Cape Town, South Africa (<http://www.nbi.ac.za/frames/researchfram.htm>).

Three different scenarios of such a DNA bank model—inspired by the model of Working and Reserve Nodes developed by Adams and Adams (1992)—are proposed below. In all cases, the DNA bank network would consist of the following components (Figure 1):



**Figure 1.** Model of a pilot DNA bank network linked to the current system of International Agricultural Research Centres (IARC) of the CGIAR (Consultative Group of International Agricultural Research) (grey circle).

- The Documentation Centre (e.g. IPGRI)  
The role of the Documentation Centre would include the documentation and dissemination of the genetic information held by the DNA banks, capacity building with respect to DNA collection, molecular analysis and documentation, and policy-making in order to facilitate the development of an adequate legal framework for the issues related to DNA storage and exchange. As mentioned earlier, a Web-based service facility readily accessible and searchable using appropriate informatics tools would be crucial for the documentation and coordination of the vast amount of data already generated and expected in the future (see also Chapter V). This Web-based facility would operate in close collaboration with other global initiatives



such as the Global Biodiversity Information Facility (GBIF 2005; Edwards et al. 2000), and take advantage of informatics tools that are already at hand (Bisby 2000). We suggest that IPGRI take over the role of the Documentation Centre, similar to its task to safeguard and document the PGR maintained in the other centres of the CGIAR.

- Various DNA banking nodes, each consisting of a physically separated working and reserve facility
- The working facilities would assume the acquisition of DNA and biological samples for DNA extraction (plant tissue, seeds, herbarium specimen), either by actively collecting in the field, or through exchange with national agricultural research systems (NARS), universities, herbaria, other institutions and the private sector. The working facilities would also be in charge of DNA extraction and documentation. They would be responsible for the replenishment of DNA samples when finished and would provide DNA samples to users on request for research.
- The reserve facilities would take charge of the long-term preservation of the DNA as reference samples. Each reserve facility would store a duplicate DNA collection of another working node as back-up reserve in case of catastrophic loss.
- Associated institutions, e.g. NARS and Genebanks, universities, herbaria and the private sector
- Users

### Model 1

In this model, each CGIAR centre would constitute a DNA banking node (A, B, ...; Figure 1), establishing a physically separated working and back up facility. Different species out of the 35 crops and crop complexes covered by the IT would be assigned to each DNA banking node, including the centre's respective mandate crops as well as some IT crops outside its mandate. Each IARC would extract DNA from these assigned species and preserve samples of isolated DNA as well as of the associated biological samples from which it was derived.

When acquiring biological samples for DNA extraction and/or isolated DNA from third parties, such as NARS, universities, herbaria and the private sector, standards previously agreed upon would need to be followed, with respect to the samples (minimum amount of associated information, DNA quality and quantity). Two possibilities present themselves:

- The samples (plant tissue, seeds and/or extracted DNA) are consigned to the DNA banking node for their holding in trust.
- The respective owner institution preserves the samples.

If the owner institution preserves the DNA samples, the associated biological samples, or both, it will need to comply with standards concerning the preservation and monitoring of the samples. The availability of samples for distribution to users would need to be ensured, thus this option would also be more demanding in terms of the documentation required.

When distributing DNA samples to users, standardized material transfer agreements (MTAs) would bind the user to provide the sender (DNA banking node or associated institution [NARS, universities etc.]) with the research results obtained by utilization of the DNA samples to supplement the existing information. The DNA banking nodes and the associated institutions would permanently provide the Documentation Centre with this information to keep the databases updated.

### Model 2

The Model 2 scenario follows Model 1, but only four to five IARCs would be designated to act as DNA banking nodes. The species assigned to each of them would include the respective centre's mandate crops, the mandate crops of one or two other IARCs not designated as DNA banking nodes, as well as some of the IT crops that are non-CGIAR mandate crops.

### Model 3

The Model 3 scenario is similar to Model 2, with the only difference that the DNA banking nodes would store only extracted DNA, and not the associated biological samples from which it was derived.

### Legal issues

The legal situation with respect to the collection, exchange and patenting of PGR in general and of plant DNA in particular is very complex, and international regulations are just being developed. Legally, DNA is a genetic resource as defined in 1992 by the Convention on Biological Diversity (CBD), Article 2 (UNEP, 1992). However, with respect to its transfer and exchange, quarantine authorities consider DNA of chemical and not of biological nature, which means that, unlike animal and plant tissue, it is not subject to restrictions due to disease risks.

The FAO Global Plan of Action for the Conservation and Sustainable Utilization of Plant Genetic Resources for Food and Agriculture (PGRFA), which was adopted at the 4th International Technical Conference on PGR, can be seen as a tool for implementing strategies from Agenda 21 and the CBD with regard to PGRFA. Its main aims are to ensure the conservation and sustainable use of PGR as a basis for food security, to promote a fair and equitable sharing of the benefits arising from their use, and to assist countries and institutions responsible for conserving these resources in identifying priorities for action and building programmes to attain these purposes.

In 1994, an agreement was signed by the IARCs of the CGIAR, placing their plant germplasm collections under the auspices of the FAO. The ownership and thus the Intellectual Property Rights (IPR) of the germplasm held in trust as seeds or *in vitro* in the CGIAR genebanks remains with the country of origin, based on the CBD, which recognizes national sovereignty over biological resources found within a country's borders and defines access to these resources as being dependent on prior informed consent and on mutually agreed terms (UNEP 1992).

In this paper, a similar system is suggested to be developed for the DNA held in trust in a DNA Bank Network as proposed above, with the ownership and IPR of the stored DNA remaining with the country of origin of the genetic resource (i.e. the isolated DNA or the biological sample [tissue, seed] from which the DNA has been extracted). A more complex issue, which needs further research, is the ownership, IPR and benefit-sharing of products derived from research on these DNA samples, i.e. sequence information, isolated genes etc.

The CBD encourages the sharing of benefits with countries of origin of genetic resources. Based on the convention, several countries have adopted legislation regulating access to their genetic resources by a concept of negotiated bilateral or regional agreements (see the example of Colombia in Chapter VII).

Recently, the IT adopted by the FAO in 2001 has entered into force (FAO 2004). At this point, it is the principal international legal instrument governing the exchange of PGRFA. The treaty establishes a multilateral system for access and benefit-sharing that applies to 35 crops (and crop complexes) and a number of forages. Facilitated access will be granted to materials within that multilateral system. The IT is the first legally binding international treaty regarding PGR that formulates conditions and standards of access, transfer and exchange of PGR, including regulations concerning intellectual and other property rights and the implementation of standardized material transfer agreements (MTAs).

However, some important terms are not defined and this has led to confusion and controversy (Fowler et al. 2003; Fowler 2004). Particularly, Article 12.3(d) that sets out the conditions under

which facilitated access is granted, i.e. “Recipients shall not claim any intellectual property or other rights that limit the facilitated access to the PGRFA, or their genetic parts or components, in the form received from the Multilateral System”. This gives rise to controversial interpretations. For example, the phrase ‘in the form received’ is subject to discussions about whether it precludes the patenting of specific genes or not. Some countries claim that the patented gene is the same as that received, while others argue that the isolated and purified DNA molecule for which a function is identified is different from ‘the form received’.

With respect to patent issues, the Trade Related Aspects of Intellectual Property Rights (TRIPs) agreement of the World Trade Organization (WTO) from 1994 specifies that WTO members shall provide patent protection for any inventions, whether products or processes, in all fields of technology (Leskien and Flitner 1997). Koo et al. (2004) give an illustrative overview of the variety of plant-related IP protection systems implemented by the United States, Canada, Europe and the Andean Community for plants and plant parts.

The TRIPs Agreement seems to be very clear and precise as to the patentable subject matter, but it, too, leaves numerous crucial questions open. It is, for example, not clear whether members are allowed to exclude only plants and animals as such from patentability or also parts of plants and animals, such as DNA sequences or genes. Also, there is no requirement to involve or consult with local communities or governments nor to provide benefit-sharing, when patenting a compound that is based on a natural product from that country. Many developing countries, therefore, believe that TRIPs is in conflict with the spirit of the CBD (Masood 1998). The relationship between IPR, in particular the TRIPs Agreement, and obligations resulting from other instruments of international law, in particular the CBD and the IT, has been and is still under debate (Feyt 2003), including the controversial question whether DNA should be patentable or not (Doll 1998; Resnik 2001; Beachy 2003; Gepts 2004).

## Conclusions

Undoubtedly, the availability of DNA resources may initiate a quantum leap in the deployment of genetic resources. It will stimulate not only genome-driven research projects, but provide a basis for the development of improved procedures for the management of collections and a more efficient genome- or gene-based selection of plant genetic resources to be deployed in breeding programmes. In this context, the present paper represents an attempt to define and describe the organizational framework for the establishment of DNA banks and some of the legal issues emerging from such an undertaking. Owing to the complexity of the issue, it is not possible to deliver final guidelines, but we intend to stimulate and intensify further discussion. The successful establishment of DNA banks and their use will critically depend on the development and the sustained curation of adequate information systems as well as on the availability of sufficient amounts of high-quality DNA. Regarding the first, IPGRI could offer its expertise in the field of information management and the establishment of appropriate networks. Given an estimated number of 6 million *ex situ* PGR worldwide (which will mainly include Annex 1 accessions), the generation of the DNA resource will require major investments, since the extraction of high-quality DNA represents an expensive undertaking and the selection of samples, notably from large and complex DNA banks, may require sophisticated logistics and robotic equipment. Although it would be desirable to have available as soon as possible the DNA samples from as many of the available *ex situ* PGR as possible, the proposed approach could be easily adapted to an incremental development of the corresponding resources. During an initial phase, this would allow for the implementation of technical improvements and could provide opportunities to join the DNA bank consortium at a later stage.

## References

- Adams RP. 1988. The preservation of genomic DNA: DNA Bank-Net. AIBS Annual Meeting 1988. University of California, Davis, CA, USA.
- Adams RP, Adams JE, editors. 1992. Conservation of Plant Genes. DNA banking and *in vitro* biotechnology. Academic Press, San Diego, CA, USA.
- Beachy RN. 2003. IP policies and serving the public. *Science* 299:473.
- Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL. 2003. GenBank. *Nucleic Acids Res* 31:23–27.
- Bisby FA. 2000. The quiet revolution: biodiversity informatics and the Internet. *Science* 289:2309–2312.
- CGIAR Science Council. 2005. Working Document on CGIAR 2005–2015 System Priorities, preliminary draft. 20 p. Available from: <http://www.sciencecouncil.cgiar.org/publications/pdf/SP5-15.pdf>
- de Vicente MC, Metz T, Alercia A. 2004. Descriptors for genetic markers technologies. International Plant Genetic Resources Institute (IPGRI), Rome, Italy.
- Doll JJ. 1998. Biotechnology: the patenting of DNA. *Science* 280:689–690.
- Edwards JL, Lane MA, Nielsen ES. 2000. Interoperability of biodiversity databases: biodiversity information on every desktop. *Science* 289:2312–2313.
- FAO. 2004. The International Treaty on Plant Genetic Resources for Food and Agriculture. Food and Agriculture Organization of the United Nations (FAO), Rome, Italy.
- Feyt H. 2003. Intellectual property rights on varieties and access to plant genetic resources: towards a coherent and an ethically based approach. *Plant Genet. Res.* 1:97–102.
- Fowler C. 2004. Accessing genetic resources: international law establishes multilateral system. *Genet. Res. Crop Evol.* 51:609–620.
- Fowler C, Moore G, Hawtin G. 2003. The International Treaty on Plant Genetic Resources for Food and Agriculture: a primer for the Future Harvest Centres of the CGIAR. International Plant Genetic Resources Institute (IPGRI), Rome, Italy.
- GBIF (Global Biodiversity Information Facility) [online]. 2005. Available from: <http://www.gbif.org>
- GCP (Generation Challenge Programme) [online]. 2005. Available from: <http://www.generationcp.org>
- Gepts P. 2004. Who owns biodiversity, and how should the owners be compensated? *Plant Physiol.* 134:1295–1307.
- Koo B, Nottenburg C, Pardey PG. 2004. Plants and intellectual property: an international appraisal. *Science* 306:1295–1297.
- Leskien D, Flitner M. 1997. Intellectual property rights and plant genetic resources: options for a *sui generis* system. *Issues in Genetic Resources No. 6*. International Plant Genetic Resources Institute (IPGRI), Rome, Italy.
- Masood E. 1998. Social equity versus private property: striking the right balance. *Science* 392:537.
- Morin PA. 2000. Genetic resources: opportunities and perspectives for the new century. *Conserv. Genet.* 1:271–275.
- Morin PA, Poinar HN, Eglinton G. 2000. Preservation of DNA from endangered species. *Science* 289:725–727.
- Resnik DB. 2001. DNA patents and scientific discovery and innovation: assessing benefits and risks. *Sci. Eng. Ethics* 7:29–62.
- Ryder OA, McLaren A, Brenner S, Zhang Y-P, Benirschke K. 2000. DNA banks for endangered animal species. *Science* 288:275–277.
- Savolainen V, Reeves G. 2004. A plea for DNA banking. *Science* 304:1445.
- UNEP. 1992. Convention on Biological Diversity. United Nations Environment Programme (UNEP), Rio de Janeiro, Brazil.



# Acronyms and abbreviations

AFLP	amplified fragment length polymorphism
AvHI	Alexander von Humboldt Biological Resources Research Institute, Cali, Colombia
AZA	American Zoo and Aquarium Association ( <a href="http://www.aza.org/">http://www.aza.org/</a> )
BAC	bacterial artificial chromosome
BILs	back-cross inbred lines
CAS	Chinese Academy of Sciences, Kunming, P.R. China ( <a href="http://english.cas.cn/">http://english.cas.cn/</a> )
CATIE	Centro Agronómico Tropical de Investigación y Enseñanza, Costa Rica ( <a href="http://www.catie.ac.cr/">http://www.catie.ac.cr/</a> )
CBD	Convention on Biological Diversity (of UNEP) ( <a href="http://www.biodiv.org/">http://www.biodiv.org/</a> )
CBOL	Consortium for the Barcode of Life, Smithsonian Institution's National Museum of Natural History, Washington DC, USA ( <a href="http://www.barcoding.si.edu/">http://www.barcoding.si.edu/</a> )
cDNA	complementary DNA
CGIAR	Consultative Group on International Agricultural Research ( <a href="http://www.cgiar.org/">http://www.cgiar.org/</a> )
CGN	Centre for Genetic Resources, The Netherlands ( <a href="http://www.cgn.wageningen-ur.nl/">http://www.cgn.wageningen-ur.nl/</a> )
CGRB	Conservation Genome Resource Bank for Korean Wildlife, Seoul, Korea ( <a href="http://www.cgrb.org/">http://www.cgrb.org/</a> )
CIAT	Centro Internacional de Agricultura Tropical, Cali, Colombia ( <a href="http://www.ciat.cgiar.org/">http://www.ciat.cgiar.org/</a> )
CIRAD	Centre de Coopération Internationale en Recherche Agronomique pour le Développement, France ( <a href="http://www.cirad.fr/">http://www.cirad.fr/</a> )
CPCG	Centre for Plant Conservation Genetics, NSW, Australia
CRES	Conservation and Research for Endangered Species, CA, USA ( <a href="http://cres.sandiegozoo.org/">http://cres.sandiegozoo.org/</a> )
CSSLs	chromosome segment substitution lines
DDBJ	DNA Data Bank of Japan, Mishima, Japan ( <a href="http://www.ddbj.nig.ac.jp/">http://www.ddbj.nig.ac.jp/</a> )
DNA	deoxyribonucleic acid
EBI	European Bioinformatics Institute, Cambridge, UK ( <a href="http://www.ebi.ac.uk/">http://www.ebi.ac.uk/</a> )
EMBL	European Molecular Biology Laboratory, Cambridge, UK ( <a href="http://www.embl.org/">http://www.embl.org/</a> )
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária, Brazil ( <a href="http://www.embrapa.br/">http://www.embrapa.br/</a> )
EST	expressed sequence tag

FAO	Food and Agriculture Organization of the United Nations, Rome, Italy ( <a href="http://www.fao.org/">http://www.fao.org/</a> )
GAWTS	genomic amplification with transcript sequencing
GBIF	Global Biodiversity Information Facility, Copenhagen, Denmark ( <a href="http://www.gbif.org/">http://www.gbif.org/</a> )
GBSSI	granule bound starch synthesis I
GenBank®	genetic sequence database hosted at the NIH-NCBI, MD, USA ( <a href="http://www.ncbi.nlm.nih.gov/Genbank/index.html">http://www.ncbi.nlm.nih.gov/Genbank/index.html</a> )
GCP	Generation Challenge Programme [ <a href="http://www.generationcp.org">http://www.generationcp.org</a> ]
IARC	International Agricultural Research Center (of the CGIAR)
GRIN	Germplasm Resources Information Network, USDA-ARS, USA ( <a href="http://www.ars-grin.gov/">http://www.ars-grin.gov/</a> )
IFPRI	International Food Policy Research Institute, WA, USA ( <a href="http://www.ifpri.org/">http://www.ifpri.org/</a> )
IMF	International Monetary Fund, Washington DC, USA ( <a href="http://www.imf.org/">http://www.imf.org/</a> )
INPRIMAT	European network to promote primate molecular biology, Barcelona, Spain ( <a href="http://www.inprimat.org/">http://www.inprimat.org/</a> )
IPBIR	Integrated Primate Biomaterials and Information Resources ( <a href="http://www.ipbir.org/">http://www.ipbir.org/</a> )
IPGRI	International Plant Genetic Resources Institute, Rome, Italy ( <a href="http://www.ipgri.cgiar.org/">http://www.ipgri.cgiar.org/</a> )
IPK	Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany ( <a href="http://www.ipk-gatersleben.de/">http://www.ipk-gatersleben.de/</a> )
IPR	intellectual property rights
IRGSP	International Rice Genome Sequencing Project ( <a href="http://rgp.dna.affrc.go.jp/IRGSP/">http://rgp.dna.affrc.go.jp/IRGSP/</a> )
IRRI	International Rice Research Institute, Los Baños, Philippines ( <a href="http://www.irri.org/">http://www.irri.org/</a> )
ISBER	International Society for Biological and Environmental Repositories, MD, USA ( <a href="http://www.isber.org/">http://www.isber.org/</a> )
ISNAR	International Service for National Agricultural Research (of IFPRI), Addis Ababa, Ethiopia ( <a href="http://www.ifpri.org/divs/isnar.htm">http://www.ifpri.org/divs/isnar.htm</a> )
IT	International Treaty on Plant Genetic Resources for Food and Agriculture (of FAO) ( <a href="http://www.fao.org/ag/cgrfa/itpgr.htm">http://www.fao.org/ag/cgrfa/itpgr.htm</a> )
ITS region	internal transcribed spacer region
IUCN	The World Conservation Union (International Union for Conservation of Nature and Natural Resources), Gland, Switzerland ( <a href="http://www.iucn.org/">http://www.iucn.org/</a> )
LN2	liquid nitrogen
MAFF	Ministry of Agriculture, Forestry and Fisheries, Tokyo, Japan ( <a href="http://www.maff.go.jp/">http://www.maff.go.jp/</a> )
MAS	marker-assisted selection
MTA	material transfer agreement
MOL	molecular markers
mtDNA	mitochondrial DNA
NARS	national agricultural research systems
NBI	National Botanical Institute, Cape Town, South Africa ( <a href="http://www.nbi.ac.za/">http://www.nbi.ac.za/</a> )
NBPGR	National Bureau of Plant Genetic Resources, New Delhi, India ( <a href="http://nbpgr.delhi.nic.in/">http://nbpgr.delhi.nic.in/</a> )
NCBI	National Center for Biotechnology Information, MD, USA ( <a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a> )
NCGRP	National Center for Genetic Resources Preservation, CO, USA
NIAS	National Institute of Agrobiological Sciences, Ibaraki, Japan ( <a href="http://www.dna.affrc.go.jp/">http://www.dna.affrc.go.jp/</a> )

NIH	National Institute of Health, MD, USA ( <a href="http://www.nih.gov/">http://www.nih.gov/</a> )
NILs	nearly-isogenic lines
NPGS	National Plant Germplasm System, USDA-ARS
NSF	National Science Foundation, VI, USA ( <a href="http://www.nsf.gov/">http://www.nsf.gov/</a> )
PAC	P1-derived artificial chromosome
PCD	programmed cell death
PCR	polymerase chain reaction
PGR	plant genetic resources
PGRFA	plant genetic resources for food and agriculture
QTLs	quantitative trait loci
RAPD	random amplified polymorphic DNA
RCA-RCA	restriction and circularization-aided rolling circle amplification
RFLPs	restriction fragment length polymorphisms
RGP	Rice Genome Research Program ( <a href="http://rgp.dna.affrc.go.jp/">http://rgp.dna.affrc.go.jp/</a> )
RH	relative humidity
RIKEN	RIKEN Institute, Japan ( <a href="http://www.riken.jp/engn/">http://www.riken.jp/engn/</a> )
RIPF	Research Institute of Pomology and Floriculture, Skierniewice, Poland ( <a href="http://www.isk.skierniewice.pl">http://www.isk.skierniewice.pl</a> )
RNA	ribonucleic acid
RT-PCR	real-time polymerase chain reaction
SINGER	System-wide Information Network for Genetic Resources ( <a href="http://singer.grinfo.net/">http://singer.grinfo.net/</a> )
SSP	Species Survival Plan® Program (of AZA)
SSR	simple sequence repeats (microsatellite markers)
TE	Tris-HCl EDTA buffer
TRIPs	Trade Related Aspects of Intellectual Property Rights (WTO) ( <a href="http://www.wto.org/english/tratop_e/trips_e/trips_e.htm">http://www.wto.org/english/tratop_e/trips_e/trips_e.htm</a> )
UNCED	United Nations Conference on Environment and Development
UNEP	United Nations Environment Programme, Nairobi, Kenya ( <a href="http://www.unep.org/">http://www.unep.org/</a> )
USDA	United States Department of Agriculture, Washington DC, USA ( <a href="http://www.usda.gov/">http://www.usda.gov/</a> )
WGA	whole genome amplification
WRPIS	Western Regional Plant Introduction Station, USDA-ARS, Pullman, WA, USA ( <a href="http://www.ars.usda.gov/pwa/pullman/wrpis">http://www.ars.usda.gov/pwa/pullman/wrpis</a> )
WTO	World Trade Organization, Geneva, Switzerland ( <a href="http://www.wto.org/">http://www.wto.org/</a> )
YAC	yeast artificial chromosome clone





# List of authors

**Andersson, Meike S.**  
Conservation Biologist  
IPGRI  
A.A. 6713  
Cali, Colombia  
Tel.: 57-2-4450048  
Fax: 57-2-4450096  
m.andersson@cgiar.org

**Dulloo, Ehsan**  
Senior Scientist, Germplasm Conservation  
IPGRI  
Via dei Tre Denari, 472/a  
00057 Maccarese, Rome, Italy  
Tel.: 39-06-6118206  
Fax: 39-06-61979661  
e.dulloo@cgiar.org

**Ebert, Andreas W.**  
Coordinator, PGR and Biotechnology  
Tropical Agricultural Research and Higher Education Center (CATIE)  
Apartado Postal 01-7170 Turrialba, Cartago, Costa Rica  
Tel.: 506-5582401  
Fax: 506-5566480  
awebert@catie.ac.cr

**Engels, Jan**  
Genetic Resources Management Adviser, IPGRI  
Via dei Tre Denari, 472/a  
00057 Maccarese, Rome, Italy  
Tel.: 39-06-6118222  
Fax: 39-06-61979661  
j.engels@cgiar.org

**Ferreira, Márcio Elías**  
Plant Genetics and Molecular Marker Analysis  
EMBRAPA-CENARGEN Genetic Resources and Biotechnology  
Parque Estação Biológica  
Av. W/5 Norte (Final)  
70770-900 Brasília DF, Brazil  
Tel.: 55-61-4484792  
Fax: 55-61-3403624  
ferreira@cenargen.embrapa.br

**Graner, Andreas**

Head, Department of Genebank  
Institute of Plant Genetics and Crop Plant Research (IPK)  
Corrensstrasse 3  
D-06466 Gatersleben, Germany  
Tel.: 49-39482-5521  
Fax: 49-39482-5155  
graner@ipk-gatersleben.de

**Hanner, Robert**

Associate Director, Canadian Barcode of Life Network  
Past President, International Society for Biological and Environmental Repositories (ISBER)  
Biodiversity Institute of Ontario  
Department of Integrative Biology  
University of Guelph  
Guelph, Ontario N1G 2W1, Canada  
Tel.: 1-519-8244120 Ext 53479  
Fax: 1-519-7671656  
rhanner@uoguelph.ca

**Henry, Robert**

Director  
Centre for Plant Conservation Genetics (CPCG)  
Southern Cross University  
Lismore NSW 2480, Australia  
Tel.: 61-2-66203010  
Fax: 61-2-66222080  
rhenry@scu.edu.au

**Karihaloo, Jawahir Lal**

Project Director  
NRC on DNA Fingerprinting  
National Bureau of Plant Genetic Resources (NBPGR)  
Pusa Campus  
New Delhi, 110012, India  
Fax: 91-11-25849459  
jlk@nbpgr.delhi.nic.in

**Kisha, Theodore**

Plant Geneticist  
USDA-ARS  
59 Johnson Hall, Washington State University  
Pullman, WA 99164-6402, USA  
Tel.: 1-509-3356898  
Fax: 1-509-3356654  
tkisha@wsu.edu

**Mesa Fuquen, Eloina**

Associate Professor

Universidad del Valle, Escuela de Ingeniería Industrial y Estadística

Cali, Colombia

Tel.: 57-2-3334903

eloimesa@univalle.edu.co

**Nagamura, Yoshiaki**

Head of DNA Bank

National Institute of Agrobiological Sciences (NIAS)

2-1-2, Kannondai

Tsukuba, Ibaraki 305-8602, Japan

Tel.: 81-29-8387456

Fax: 81-29-8387408

nagamura@dna.affrc.go.jp

**Palacio-Mejía, Juan Diego**

Curator, National Tissue Collection

Alexander von Humboldt Biological Resources Research Institute

A.A. 6713

Cali, Colombia

Tel.: 57-2-4450000 Ext 3174

Fax: 57-2-4450073

jdpalacio@humboldt.org.co

**Rice, Nicole**

Curator

Australian Plant DNA Bank

Southern Cross University

Lismore NSW 2480, Australia

Tel.: 61-2-66203409

Fax: 61-2-66222080

nrice@scu.edu.au

**Rossetto, Maurizio**

Conservation Geneticist

National Herbarium of NSW

Botanic Gardens Trust

Sydney NSW 2000, Australia

Tel.: 61-2-92318337

Fax: 61-2-9251231

maurizio.rossetto@rbgsyd.nsw.gov.au

**Ryder, Oliver**

Kleberg Genetics Chair and Head, Genetics Division  
Conservation and Research for Endangered Species (CRES)  
Zoological Society of San Diego  
15600 San Pasqual Valley Road  
Escondido, CA 92027-7000 USA  
Tel.: 1-760-291-5452  
Fax: 1-760-291-5451  
oryder@ucsd.edu

**de Vicente, M. Carmen**

Senior Scientist, Plant Molecular Genetics  
IPGRI  
A.A. 6713  
Cali, Colombia  
Tel.: 57-2-4450029  
Fax: 57-2-4450096  
c.devicente@cgiar.org

**Walters, Christina**

Research Leader  
Plant Germplasm Preservation Research Unit  
National Center Genetic Resources Preservation (NCGRP), USDA-ARS  
1111 South Mason Street  
Fort Collins, CO 80521-4500, USA  
Tel.: 1-970-4953202  
Fax: 1-970-2211427  
chrisv@lamar.colostate.edu

**Xiangyu, Jingong**

Laboratory of Cellular and Molecular Evolution  
Kunming Institute of Zoology  
Chinese Academy of Sciences (CAS)  
Kunming 650223, China  
xiangyuj@yahoo.com.cn

**Zhang, Ya-ping**

Molecular Biology of Domestic Animals  
Kunming Institute of Zoology  
Chinese Academy of Sciences (CAS)  
Kunming 650223, China  
Tel.: 86-871-5190761  
Fax: 86-871-5195430  
zhangyp1@263.net.cn



## FUTURE HARVEST

IPGRI is  
a Future Harvest Centre  
supported by the  
Consultative Group on  
International Agricultural  
Research (CGIAR)

ISBN-13: 978-92-9043-702-4

ISBN-10: 92-9043-702-2