

Biotechnology in Agriculture Series, No.19

Biotechnology and Plant Genetic Resources

Conservation and Use

Edited by

J.A. Callow,

B.V. Ford-Lloyd

and

H.J. Newbury



CAB INTERNATIONAL

BIOTECHNOLOGY AND PLANT GENETIC RESOURCES

Conservation and Use

Biotechnology and Plant Genetic Resources

Conservation and Use

Edited by

**J.A. Callow,
B.V. Ford-Lloyd**
and

H.J. Newbury
*School of Biological Sciences
University of Birmingham
UK*

CAB INTERNATIONAL

CAB International
Wallingford
Oxon OX10 8DE
UK

Tel: +44 (0)1491 832111
Fax: +44 (0)1491 833508
E-mail: cabi@cabi.org

CAB International
198 Madison Avenue
New York, NY 10016-4341
USA

Tel: +1 212 726 6490
Fax: +1 212 686 7993
E-mail: cabi-nao@cabi.org

© CAB INTERNATIONAL 1997. All rights reserved. No part of this publication may be reproduced in any form or by any means, electronically, mechanically, by photocopying, recording or otherwise, without the prior permission of the copyright owners.

A catalogue record for this book is available from the British Library, London, UK

A catalogue record for this book is available from the library of Congress, Washington DC, USA

ISBN 0 85199 142 4

Typeset in 10/12pt Palatino by Columns Design Ltd., Reading

Printed and bound in the UK by Biddles Ltd, Guildford and King's Lynn

Contents

Contributors	vii
Preface	ix
Foreword <i>G.C. Hawtin</i>	xi
1. Overview <i>J.A. Callow, B.V. Ford-Lloyd and H.J. Newbury</i>	1
2. Use of molecular marker techniques for description of plant genetic variation <i>A.L. Westman and S. Kresovich</i>	9
3. Genetic diversity – population structure and conservation <i>M.D. Hayward and N.R. Sackville Hamilton</i>	49
4. Genomic relationships, conserved syntenry and wide-hybrids <i>D.A. Laurie, G.J. Bryan and J.W. Snape</i>	77
5. Molecular markers and the management of genetic resources in seed genebanks: a case study of rice. <i>B.V. Ford-Lloyd, M.T. Jackson and H.J. Newbury</i>	103
6. <i>In vitro</i> conservation methods <i>F. Engelmann</i>	119
7. Conservation of DNA: DNA banking <i>R.P. Adams</i>	163
	v

8.	Genetic resources and plant breeding – identification, mapping and manipulation of simple and complex traits <i>M.J. Kearsey</i>	175
9.	Gene identification, isolation and transfer <i>I.D. Godwin</i>	203
10.	Importance of biotechnology for germplasm health and quarantine <i>H. Barker and L. Torrance</i>	235
11.	Biodiversity for bioindustries <i>G. Tamayo, W.F. Nader and A. Sittenfeld</i>	255
12.	Internet resources for the biologist <i>M.L. Anderson and S.W. Cartinhour</i>	281
	Index	301

Contributors

- R.P. Adams**, *Plant Biotechnology Center, Baylor University, Box 669 Gruver, TX 79040, USA.*
- M.L. Anderson**, *Nottingham Arabidopsis Stock Centre, Department of Life Science, University of Nottingham, Nottingham NG7 2RD, UK.*
- H. Barker**, *Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, UK.*
- G.J. Bryan**, *John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK.*
- J.A. Callow**, *School of Biological Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK.*
- S.W. Cartinhour**, *Crop Biotechnology Center, Texas A&M University, College Station, TX 77843, USA.*
- F. Engelmann**, *International Plant Genetic Resources Institute (IPGRI), Via delle Sette Chiese 142, 00145 Rome, Italy.*
- B.V. Ford-Lloyd**, *School of Biological Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK.*
- I.D. Godwin**, *Department of Agriculture, University of Queensland, Brisbane, Queensland 4072, Australia.*
- G.C. Hawtin**, *International Plant Genetic Resources Institute (IPGRI), Via delle Sette Chiese 142, 00145 Rome, Italy.*
- M.D. Hayward**, *Institute of Grassland and Environmental Research, Plas Gogerddan, Aberystwyth, Dyfed SY23 3EB, UK.*
- M.T. Jackson**, *International Rice Research Institute, P.O. Box 933, 1099 Manila, Philippines.*

- M.J. Kearsey**, *Plant Genetics Group, School of Biological Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK.*
- S. Kresovich**, *US Department of Agriculture, Agricultural Research Service, Plant Genetic Resources Conservation Unit, University of Georgia, 1109 Experiment Street, Griffin, GA 30223-1797, USA.*
- D.A. Laurie**, *John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK.*
- W.F. Nader**, *Instituto Nacional de Biodiversidad (INBio), Apartado Postal 22-3100, Santo Domingo, Heredia, Costa Rica.*
- H.J. Newbury**, *School of Biological Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT UK.*
- N.R. Sackville Hamilton**, *Institute of Grassland and Environmental Research, Plas Gogerddan, Aberystwyth, Dyfed SY23 3EB, UK.*
- A. Sittenfeld**, *Instituto Nacional de Biodiversidad (INBio), Apartado Postal 22-3100, Santo Domingo, Heredia, Costa Rica.*
- J.W. Snape**, *John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK.*
- G. Tamayo**, *Instituto Nacional de Biodiversidad (INBio), Apartado Postal 22-3100, Santo Domingo, Heredia, Costa Rica.*
- L. Torrance**, *Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, UK.*
- A.L. Westman**, *US Department of Agriculture, Agricultural Research Service, Plant Genetic Resources Conservation Unit, University of Georgia, 1109 Experiment Street, Griffin, GA 30223-1797, USA.*

Preface

Advances in molecular and cell biology over the past ten years have led to the development of a wide range of techniques for manipulating genomes, collectively termed 'biotechnology'. This book sets out to explore the application and impact of biotechnology on the assessment of plant genetic variation, its conservation and utilization.

Biotechnology promises to make significant contributions in the fields of health care, food safety, environmental protection, bioindustrial processing and the development of sustainable forms of agriculture. A veritable Pandora's box of applications for genome research has been opened. Although much of the focus in the plant sciences has been on the direct manipulation of plant genomes where the newer techniques form an important adjunct to more traditional methods of plant breeding, the rise of biotechnology has also catalysed a renewed emphasis on the importance of biological and genetic diversity and its conservation, an area which has, in recent years, attracted growing public and scientific interest and political support.

The methods of biotechnology now permit a more rapid and deeper understanding of both species and their genetic diversity, the mechanisms by which that variation is generated in nature, and the significance of that variation in the adaptation of plants to their environment. They allow the development of rapid methods for screening germplasm for specific characters, including the presence of disease-causing organisms. They promote more effective conservation strategies by helping to define the extent of genetic diversity thus allowing rational decision-making on the collection and management of genetic resources. Tissue culture-based techniques are

available for conserving germplasm that cannot be maintained by more traditional methods. The understanding of natural diversity at the molecular level in areas such as pest and disease resistance and secondary metabolic pathways provides novel opportunities for the improvement of crops and in developing recombinant routes to high-value plant products. The ever-improving development of technologies for direct gene manipulation in crop plants offers a virtually unlimited gene pool for the plant breeder to exploit. The information stemming from these molecular and cellular technologies is increasingly underpinned by sophisticated and comprehensive informatics systems which enable information on plant genetics and molecular biology to be cross-related to systematic, ecological and other data through international networks.

This volume attempts to capture the flavour of many of these new technologies and their application in the various areas of plant genetic resources work, and to evaluate their actual or potential impact. The chapters have been written by scientists involved in plant biotechnology and genetic resources work, and aim to give a wide coverage of the impact of molecular and other technologies on plant genetic resources work. The book is not intended to be a manual of techniques although developing methodologies have been outlined as appropriate. The chapters have been written for a broad, international readership encompassing researchers in the various national and international centres of biodiversity and germplasm conservation and crop improvement, museums and herbaria, and those who teach various relevant disciplines of genetics, taxonomy, molecular biology and ecology. We also hope that the volume will be valuable to the increasing number of policy-makers who advise on the disposition of funding for research and training in biodiversity.

J.A. Callow
B.V. Ford-Lloyd
H.J. Newbury

Foreword

G.C. Hawtin

*Director General, International Plant Genetic Resources
Institute (IPGRI), Rome, Italy*

A quarter of a century ago, the world was preoccupied with an impending food crisis. Birth rates and life expectancy were increasing, especially in the developing world. The situation was particularly menacing in more populous regions where mass starvation was predicted on an unprecedented scale. In response to the crisis, a few far-sighted scientists set out to tackle the problem of increasing food production. Their solution: through the application of genetics, to develop new varieties of the world's most important food crops – wheat and rice – that would be both higher yielding and more responsive to inputs such as fertilizer and irrigation. Supportive government policies and action allowed the new varieties to spread rapidly and they soon came to occupy vast areas, especially in Asia. Food production increased apace and, through what became known as the Green Revolution, the impending crisis was averted – or rather, delayed.

Today, there are growing indications that food production is once again struggling to maintain pace with a relentlessly expanding population. Food production must double over the next 25 years, just to keep up. The technologies that underpinned the Green Revolution, while still contributing to food production gains, appear inadequate to meet the challenges that lie ahead. The expanding population has taken an additional and serious toll on the environment: soils have eroded, forests have been cut down, and biodiversity has been lost. People living in harsh regions – environments that are marginal for agriculture – have largely been bypassed by the Green Revolution and today are among the poorest in the world.

While solutions to many of these problems are to be found in the

application of wise policies and the allocation of additional resources to development, new technologies undoubtedly also have a major part to play. Just as the Green Revolution had its origins in science and technology, and particularly in the science of genetics, so the application of new biotechnological methods could lead to a new revolution – the Gene Revolution.

New techniques offer renewed hope for solving the intricate problems that lie at the heart of today's concerns for poverty, nutrition and the environment. These techniques allow us to understand better the composition and functioning of genomes, and to transfer useful genes among widely differing groups of organisms. By such means we are learning how to more accurately tailor our food crops to meet new pest and disease challenges, and to produce crops that can better withstand the rigours of stressful environments or which provide new or improved products suited to today's complex and increasingly urban markets.

Modern plant biotechnologies can contribute not only to increasing and protecting today's agricultural production, but also to conserving the biological resources that will underpin our ability to continue to meet new challenges in the future. Techniques have been developed, for example, which allow us to understand better the nature and distribution of genetic diversity, or which provide for the long-term conservation of plant tissues or even of DNA itself.

In 1992 the largest ever meeting of Heads of State was held in Rio de Janeiro – the United Nations Conference on Environment and Development (UNCED). The Conference brought into sharp focus the growing popular recognition of the linkages between poverty and environmental degradation. The signing of the Convention on Biological Diversity at the Conference signalled to the world a determination to halt the loss of the earth's irretrievable biological resources and instead to conserve and use them as a basis for economic development. Fundamental to the implementation of the Convention is the transfer of appropriate technologies, particularly biotechnologies, to those countries and institutions that have the most to gain from their application. These countries are in large part located in the developing world, where some of the greatest problems of poverty are to be found; ironically, these countries are also home to a large share of the earth's genetic wealth. This book makes a notable contribution to a most worthwhile objective – the implementation of the Convention on Biological Diversity – by setting out in a clear, comprehensive and authoritative manner some of the latest advances in biotechnology of relevance to the conservation and use of plant genetic resources.

Overview

J.A. Callow, B.V. Ford-Lloyd and H.J. Newbury

*School of Biological Sciences, University of Birmingham,
Edgbaston, Birmingham B15 2TT, UK*

1.1. Introduction: Plant Genetic Diversity and Sustainable Agriculture

The conservation of plant genetic resources has, in recent years, attracted growing public and scientific interest and political support. There is an increasing awareness of the relevance of biological diversity and its conservation to the health of the biosphere. Particularly urgent is the need to raise agricultural output to meet the basic nutritional needs of increasing populations. The dilemma is that the required increase in food production must be obtained through 'sustainable' forms of agriculture which are less dependent on the use of modern high-yielding varieties bred for intensive, high-input systems. It seems unlikely that increased food production on this scale will come from increasing the area under cultivation. Rather, major solutions will depend upon innovations (Tribe, 1994) which exploit sophisticated methods to manipulate plant genomes and the totality of natural genetic variation.

The area of disease resistance will serve to illustrate this argument. Many authors (e.g. Hilder *et al.*, 1992) have emphasized the importance of innovation in the area of crop protection since crop losses due to pests and disease may account for between 20 and 40% of productivity worldwide. Indeed, over 50% of the research and development (R&D) spent in the plant breeding/seed industry is focused on the identification and incorporation of resistance traits (Swanson, 1996). In a situation where there is some evidence that the pool of available variation for disease resistance genes within commercial breeding lines is becoming limited (e.g. for

yellow rust in cereals; G.J. Jellis, personal communication) the search is on for novel sources of resistance. There are now many examples of biotechnology being put to the service of pest control, through the incorporation into plants of transgenes for toxins, antifeedants, enzyme inhibitors, etc. But, as Woolhouse (1992) has pointed out, such is the long coevolutionary history of pests and pathogens with their hosts that present strategies for interfering with these relationships represent only a small proportion of the likely defence mechanisms used by plants in resisting foreign organisms. This enormous diversity of mechanisms represents an arsenal of ammunition which can be used in the fight to minimize disease and predation.

Are there any hard facts which support these assertions concerning the value of genetic resources to agricultural productivity? Can we put a cost on the loss of diversity? The sources of variation used by the plant breeder range from current breeding lines to wild species and the products of direct gene manipulation. In a recent analysis of the plant breeding/seed industry, Swanson (1996) showed that, over a five-year period, 6.5% of all genetic research within this industry, resulting in a marketed innovation, was concerned with germplasm from wild species and landraces. This compares with only 2.2% of new germplasm arising from technological approaches of induced mutation. The vast majority of germplasm successfully used by the industry still arises from 'conventional' sources, i.e. the heavily exploited commercial cultivars, which at first sight might be taken to suggest that wild resources are relatively unimportant in R&D. In fact Swanson (1996) argues just the opposite, suggesting that the stock of existing commercial varieties should be seen as the information base from which bioindustries develop innovations, whereas the sources of new diversity (wild species, landraces, induced mutation) should be seen as supplying increments to the information base. In other words, R&D requires an annual injection of 'new' genetic material from natural sources, estimated as amounting to approximately 7% of the stock of material currently within the system.

Biodiversity is valuable to industries other than those that seek novel genes for crop plant improvement. Biodiversity has traditionally provided a source of compounds to the pharmaceutical, foodstuffs, crop protection and other industries and here it is possible to give more precise evaluations of its importance. As outlined in Chapter 11 of this volume, drug sales based on natural products from plants were estimated at \$US43 billion in 1985 alone (Principe, 1989) and the value of yet undiscovered pharmaceuticals in tropical rainforests is estimated to be as much as \$US147 billion to society as a whole (Mendelsohn and Balick, 1995). Vast markets exist for the replacement of synthetic biocides, many of which have associated toxicological and environmental problems, by 'natural' compounds, which may have more specific modes of action and fewer side-effects.

Even greater may be the use of genetic resources in more fundamental research where their value in the longer term may be unknown or impossible to estimate accurately at the present time. Diverse genetic material is needed and now regularly utilized to create mapping populations and to study inheritance of a wide range of traits which may or may not be of immediate practical value. Very large numbers of germplasm samples are regularly screened and evaluated for different characteristics so that the underlying molecular biology can be investigated. The intention here is that a greater understanding of fundamental biology will give rise to wealth creation in the future.

1.2. The Advent of Biotechnology

So then to the main theme of this book. Advances in molecular and cell biology over the past ten years have led to the development of a whole range of techniques for manipulating genomes, informational macromolecules and organisms, collectively termed 'biotechnology'. Biotechnology promises to make significant contributions in the fields of health care, food safety, environmental protection, bioindustrial processing and the development of sustainable forms of agriculture. The exploration and assessment of biodiversity, its conservation, and utilization in crop improvement are all aspects of the latter and form the focus for the present volume.

The application of the methodologies of biotechnology to germplasm conservation and utilization falls into three main categories (Barlow and Tzotsos, 1995).

- The use of novel molecular technologies in the assessment of biodiversity and its monitoring.
- New tools for conservation:
 - facilitating critical decisions on what should be conserved and the management of collections;
 - providing additional conservation technologies through tissue culture, DNA banks.
- More effective utilization of biodiversity, through more efficient screening of germplasm, marker-assisted selection in breeding and the use of genetic manipulation to facilitate the widest use of conserved genes.

1.3. Novel Technologies for Biodiversity Assessment and Monitoring

The analysis and characterization of genetic diversity is fundamental to any *ex situ* or *in situ* conservation strategy. Over the years the methods for

detecting genetic diversity have expanded from Mendelian analyses of discrete morphological variants, to biometrical approaches, biochemical techniques based upon protein and isozyme profiles, to methods based on DNA sequence variation. There is now a huge array of molecular marker techniques, and Chapter 2 reviews the major technologies and their appropriate use in addressing key issues in germplasm conservation

The application of molecular technologies to increase our understanding of biodiversity at several levels is exemplified by Chapter 3, which is concerned with the population level, and by Chapter 4, which is concerned with molecular approaches to genetic relatedness. Chapter 3 reviews how the collection of diverse germplasm for *ex situ* purposes and the development of effective *in situ* collections both require an understanding of the spatial and temporal aspects of population structure and associated genetic diversity and the factors that control it.

Studies with low- or single-copy markers are particularly valuable in answering questions about genetic relatedness between species. Chapter 4 illustrates how comparative molecular mapping provides new opportunities for exploring the genetical basis of phenotypic variation within and between plant species. The implication of comparative mapping is that genes controlling specific phenotypic traits will have corresponding map locations in different taxa. This provides a new framework for the identification of homologous genes in germplasm and new tools for the effective utilization of germplasm in crop plant improvement. This approach is also valuable in the study of quantitative traits controlled by several genes since flanking markers for major genes provide markers for quantitative trait loci (QTL) manipulation and selection and the map-based cloning of at least some genes controlling quantitative traits.

1.3.1. Biotechnological tools for conservation

Biotechnology can make a substantial contribution to the quality of germplasm collections and their efficient management. For example, in the management of very large germplasm collections there is a clear requirement for procedures for fast, reliable taxonomic identification of material and the elimination of redundancy. Samples characterized by molecular criteria may prove superior to traditional analyses of phenotype and examples of the use of DNA markers to guide the management of germplasm are explored in Chapter 5 by reference to one of the world's largest collections of germplasm at the International Rice Research Institute.

Not all plant species can be stored in the form of desiccated seed. So-called 'recalcitrant' species have to be kept in moist, warm conditions and even then have relatively low longevity. In addition, there are many

cultivated plants which must be maintained clonally, such as banana, which produce few seeds, or potato, which does not come true from seed. To these traditional problem species we must now add the products of biotechnology, new categories of germplasm including clones of elite genotypes, and transformed cell lines, for which storage of seed is inappropriate. Biotechnology has impact on storage technologies in two quite different ways. The first of these, reviewed in Chapter 6, concerns the various types of *in vitro* culture which allow the propagation of plant material in an aseptic environment, thus ensuring the production of disease-free stocks and simplifying quarantine procedures for the international exchange of germplasm. The development of sensitive molecular probes to detect low levels of pathogens is discussed in Chapter 10.

A more radical approach to safe storage and distribution of germplasm, and possibly the only way to significantly reduce the rate of loss in biodiverse tropical forests, within the bounds of available resources, is to develop alternatives to the costly (and inherently vulnerable) methods of *ex situ* storage of seed or tissue cultures, by conserving germplasm in the form of DNA or DNA-rich materials. The latter include those dried specimens conserved in the vast resources of the world's herbaria. The DNA Bank-Net, an international network of DNA repositories, has now been established (Chapter 7) with both storage and distribution functions. Although technologies for rapid removal, analysis, amplification and secure storage of small amounts of DNA are now well advanced, research is still needed on methods to amplify the entire genomic DNA of a species, which means that storage of DNA or DNA-rich materials should still be considered to be something of an insurance policy rather than a replacement for conventional modes of germplasm storage.

1.3.2. Biotechnology-assisted crop improvement

Effective use of genetic resources in plant breeding programmes requires methods to determine whether useful genetic variation exists, and cost-effective methods to introduce useful genes into the material. Biotechnology has made a contribution to both these aspects. The use of genetic maps based on molecular markers linked to genes of interest facilitates the more rapid selection of both simple and complex traits (Chapter 8) thus accelerating their incorporation into breeding materials. Where traits are controlled by one or a few genes, direct isolation of genes and their incorporation into elite cultivars by increasingly efficient transgenic technologies (Chapter 9) provide an additional approach to the more traditional methods of plant breeding. This prospect also challenges one of the classical concepts in plant genetic resources work, namely the 'gene pool' concept of Harlan and de Wet (1971), which separated cultivated

species and their wild relatives into different categories based upon their 'crossability'. The value of this concept in guiding the work of the plant breeder must now be re-evaluated in the light of molecular genetic technologies which have not only cast a completely new light on the relatedness of crop plant genomes (Chapter 4), but which have effectively widened the potential gene pool of a species to include all other life forms. What challenges does this imply for the traditional plant genetic resources programmes? Why is the emphasis on conserving species that are related to crops when more exotic and potentially valuable genes may be present in liverworts, or green algae, or other less studied taxa?

1.4. Bioprospecting for Novel Genes

The value of genetic resources as a source of novel compounds for various industries has been alluded to above. For centuries biodiversity has provided sources of fuels, medicines, foodstuffs, pesticides, etc. The high returns available from introducing a new drug, for example, plus the realization that biodiversity is rapidly being lost (estimated rates of extinction range from 30 to 300 species per day) has encouraged new attitudes to the exploration and exploitation of biodiversity through 'bioprospecting' (Chapter 11). How does biotechnology impact on this traditional area? A fundamental operation in any bioprospecting programme is the screening of thousands of extracts for compounds of interest. Gene technology now allows the speeding up of screening programmes for new compounds through the development of more sophisticated *in vitro* assays. For example, the genes encoding receptor proteins for certain classes of drug, or enzymes that may serve as targets for novel pesticides, may be cloned and expressed on a large scale in high-throughput *in vitro* assays into which thousands of plant extracts may be applied. Bioprospecting is also concerned with discovering novel genes and natural variants of genes or gene products that may either be used directly or to guide sequence improvement by molecular methods. Chapter 11 gives examples of amino acid sequence improvement in peptide hormones as a result of screening homologous genes from the wild, and this area of biotechnology is likely to develop in parallel with new molecular methods of generating random variation through employing the mutational potential of the polymerase chain reaction (PCR) – i.e. 'forced molecular evolution' – or recombinatorial libraries of peptides.

Whilst gene prospecting has provided a new perspective on biodiversity, the area is not without political controversy because these advances tend to favour agriculture and other industries of developed countries, and may actually displace traditional commodity production in resource-rich but underdeveloped countries. Chapter 11 also outlines

the development of strategies which integrate conservation and exploitation with economic returns and sustainable development in those countries from which the bioresources were first obtained.

1.5. The Exchange of Information and Germplasm

The value of all plant genetic resources lies in their utilization in crop improvement programmes or in other bioindustries. The vast majority of conserved germplasm is available for distribution to the global scientific community but this international distribution poses considerable risks of accidental introduction of non-indigenous plant pathogens and pests. The detection and elimination of such disease-causing agents is therefore essential. Often, such agents will be present at very low levels and may actually be symptomless. In this context then, new technologies reviewed in Chapter 10 have been of great value in the development of reliable and sensitive nucleic acid- or antibody-based diagnostic probes, and these will be of great importance in optimizing the health and value of germplasm collections.

Finally, if the biotechnology revolution has had great impact in developing novel molecular and cellular technologies, no less a revolution has affected radically the way in which the world organizes and disseminates information. Powerful computers linked by the Internet have facilitated the development of sophisticated and comprehensive informatics systems which enable information on plant genetics and molecular biology to be cross-related to systematic, ecological and other data through international networks. Chapter 11 explains how this information is organized and distributed against a background of continual development.

References

- Barlow, B. and Tzotsos, G.T. (1995) Biotechnology. In: Heywood, V.H. and Gardner, K. (eds) *Global Biodiversity Assessment*. Cambridge University Press, Cambridge, pp. 671–710.
- Harlan, J.R. and de Wet, J.M.J. (1971) Toward a rational classification of cultivated plants. *Taxon* 20, 509–517.
- Hilder, V.A., Boulter, D. and Gatehouse, A.M.R. (1992) Introduction. In: Gatehouse, A.M.R., Hilder, V.A. and Boulter, D. (eds) *Plant Genetic Manipulation for Crop Protection*. CAB International, Wallingford, pp. 1–5.
- Mendelsohn, R. and Balick, M.J. (1995) The value of undiscovered pharmaceuticals in tropical forests. *Economic Botany* 49, 223–228.
- Principe, P.P. (1989) The economic significance of plants and their constituents as drugs. In: Wagner, H., Hikino, H. and Farnsworth, N.R. (eds) *Economic and Medicinal Plant Research*, vol. 3. Academic Press, London, pp. 1–17.

- Swanson, T. (1996) Global values of biological diversity: the public interest in the conservation of plant genetic resources for agriculture. *Plant Genetic Resources Newsletter* 105, 1.
- Tribe, D. (1994) *Feeding and Greening the World: The Role of Agricultural Research*. CAB International, Wallingford.
- Woolhouse, H.W. (1992) Promoting crop protection by genetic engineering and conventional plant breeding: problems and prospects. In: Gatehouse, A.M.R., Hilder, V.A. and Boulter, D. (eds) *Plant Genetic Manipulation for Crop Protection*. CAB International, Wallingford, pp. 249–256.

Use of Molecular Marker Techniques for Description of Plant Genetic Variation

A.L. Westman and S. Kresovich

*US Department of Agriculture, Agricultural Research Service,
Plant Genetic Resources Conservation Unit, University of
Georgia, 1109 Experiment Street, Griffin, GA 30223-1797,
USA.*

2.1. Introduction

Effective conservation and use of plant genetic resources involve asking many questions about the extent and distribution of genetic variation. Only when the appropriate markers and technologies for describing this variation are accessible can such questions be adequately addressed. The most appropriate markers for a given question should: (i) be heritable; (ii) discriminate between the individuals, populations, or taxa being examined; (iii) be easy to measure and evaluate; and (iv) provide results that can be compared with results of similar studies (Hillis and Moritz, 1990).

Molecular markers and assays that meet these criteria were introduced in the 1950s and have proliferated ever since. Use of protein isozymes as genetic markers increased rapidly after their introduction (Hunter and Markert, 1957). Nucleic acid markers gained popularity in the 1970s, with the advent of DNA sequencing and restriction fragment analysis. The number of different molecular markers and assays has increased exponentially since the late 1980s, when introduction of the polymerase chain reaction (PCR) made DNA amplification and sequencing possible and affordable.

As the diversity of markers and assays increases, the cost of using them continues to decrease. At the same time, the range of taxa and tissues that can be analysed has expanded. Improved DNA extraction, amplification, and sequencing protocols allow us to analyse ancient DNA, from 13,000-year-old specimens of *Myiodon* (ground sloths) (Pääbo, 1989) and 17–20 million-year-old *Taxodium* (bald cypress) specimens (Soltis *et al.*,

1992), and samples as small as single cells or fractions of a seed (Li *et al.*, 1988; Chunwongse *et al.*, 1993; Wang *et al.*, 1993). Thus the chances are good that at least one marker assay will be appropriate for any question about variation in any taxa. Furthermore, results of one study can generally be related to results from many other studies.

Such opportunities are inevitably accompanied by challenges. None of the challenges are unique to molecular studies, but all become more critical as the number of available marker assays increases. Faced with a bewildering variety of molecular technologies (and their acronyms), choosing a marker and assay that are appropriate for a scientist's particular question, taxa of interest, and practical constraints can be overwhelming. The temptation is great to use techniques that are recommended by coworkers, are frequently used, for which the most information is available, or are the newest – without adequately evaluating the techniques to determine which is most suitable.

Because molecular markers can be measured in virtually all plant taxa and minimal tissue is required, extra care must be taken to verify that the entries in a test array are correctly identified. Since molecular markers are widely applicable and their constraints are not always understood, thoughtful planning is particularly important when designing molecular studies, to ensure that they have an appropriate range of entries, an appropriate number of entries per taxon, and conclusions that stay within the scope of a specific question and test array. In addition, the vast amount of published research using molecular markers complicates the task of thoughtfully relating one study to others with similar marker assays, research questions, and / or taxa.

Another challenge is formulating research objectives that ask critical questions about plant genetic resources. Descriptive studies that use molecular marker assays to examine variation between and within specific taxa can generate valuable information, but some plant conservation questions are best addressed by experimental studies (Drew, 1994). The wide variety of molecular marker assays available provides new ways to address such questions – if the markers are appropriately used (Milligan *et al.*, 1994).

Such challenges can be met successfully and thoughtful decisions can be made if logical frameworks are available to categorize and compare these valuable molecular tools. A solid foundation for such comparisons has been provided recently (Hillis and Moritz, 1990; Avise, 1994). This chapter builds on that foundation by reviewing both early and recently developed molecular markers and assays, citing examples of their use. The types of questions asked about plant genetic resources are described, as are guidelines for selecting appropriate markers and assays to address these questions. The literature cited here is far from comprehensive, but includes applications of each marker assay in plant genetic research. While the list of molecular techniques will soon be dated, hopefully the framework

presented will be helpful for categorizing and comparing new tools as they are made available. The chapter concludes with a perspective on new marker assays as they apply to, and may change, the study of plant genetic variation.

2.2. Types of Questions, Markers, and Marker Assays

All studies of genetic variation have four primary components: (i) the research question; (ii) the marker used to address it; (iii) the technique used to generate and measure the marker; and (iv) the method used to analyse the data produced. This chapter focuses on the first three components.

2.2.1. Types of plant genetic resource questions

Most questions asked about *ex situ* and *in situ* plant genetic variation fit into one of three categories. While boundaries between them are not always clear, these categories are useful when deciding how to address specific questions. Questions of identity are forensic in nature and ask whether two or more samples are genetically the same. These questions concern whether genebank accessions are present in duplicate, whether populations are genetically distinct or are geographical subdivisions of one population, and whether genetic change has occurred in an accession or population over time.

In contrast, questions of location and diagnostics concern the presence and location of a particular allele or nucleotide sequence in a taxon, genebank accession, *in situ* population, individual, chromosome, or cloned DNA segment. These questions are asked to locate populations or individuals that have desirable traits; determine whether traits have been lost from a population, or have been transferred from one population or plant to another; and establish the position of markers on physical or genetic maps. The differences between diagnostic and forensic questions are significant, but are not always clarified (Lander, 1989). Both types of question require qualitative information about specific loci. Yet while diagnostic studies often use single-locus markers, each with few alternative states, forensic questions are commonly addressed with information from many multiallelic loci.

Finally, relationship and structure questions examine relatedness or similarity between genotypes, the amount of genetic variation present, and how variation is distributed between individuals, populations, and taxa. Such questions are used to pinpoint gaps in genebank collections, decide which genotypes are high priority for *in situ* conservation, design germ-plasm sampling strategies and regeneration programmes, construct core

collections, study gene flow, and determine optimum sizes for conserved populations or genebank accessions. These issues are generally addressed with information from many loci. Quantitative data can be used, but qualitative data are more appropriate for some phylogenetic and parentage questions.

2.2.2. Types of molecular markers

The appropriate markers for a study can discriminate between entries in an array, but are not so polymorphic that important variation is masked by random noise (Brower and DeSalle, 1994). Molecular markers range from highly conserved to hypervariable, and can be either proteins or nucleic acids. The nucleic acids used as markers include entire genomes, single chromosomes, fragments of DNA or RNA, and single nucleotides.

A wide variety of nucleic acid fragments are utilized as markers. While some occur once in a genome, others are repeated. Many repeated sequences used as markers are noncoding; others are elements of multi-gene families. Some repeated sequences are interspersed throughout the genome, either distributed randomly or in clusters. These interspersed repeats are common in plant and animal nuclear genomes, and are found in plant (but not animal) mitochondrial genomes (Palmer, 1992; Rogowsky *et al.*, 1992). The chloroplast genome contains a large inverted repeat (IR); most angiosperm chloroplasts have two copies, separated by a short single-copy region. Repeat length and (rarely) loss of one copy can vary between taxa (Downie and Palmer, 1992).

Much research at present is focused on repeated sequences that occur in tandem. The classes of tandem repeats are distinguished by the length of the core repeat unit, the number of repeat units per locus, and the abundance and distribution of loci (Table 2.1). The names for these classes are themselves varied and have been inconsistently used, but Harding *et al.* (1992) and Tautz (1993) have clarified the nomenclature.

Tandem repeats were first reported in the literature as ‘satellites’ of DNA, detected in CsCl density gradients as fractions with different GC content than the rest of the genome (Britten and Kohne, 1968). These satellites have repeat units that are usually several hundred nucleotides long, with thousands of copies at each of several loci in the nuclear genome. These loci are usually in heterochromatin, often near centromeres. Satellite DNA is present in numerous species. For many satellites, the number of loci and number of repeat units per locus vary between species and higher taxa (Ingles *et al.*, 1973; Sibson *et al.*, 1991).

Minisatellites (often called variable number of tandem repeat loci, or VNTR loci) are widely used as markers, especially in forensics. The repeat units are usually less than 100 nucleotides long, with tens to hundreds of

Table 2.1. Classes of nucleic acid sequences used as fragment markers.

Sequence class	Genome ^a	No. of loci/genome	Coding region ^b	Repeat unit length (bp)	No. of tandem units/locus	Marker assay ^c				
						CsCl density gradient	DNA–DNA hybridization	<i>In situ</i> hybridization	Restriction site analysis	PCR amplification
Single copy	n, cp, mt	One	±	N.A.	N.A.	–	+	+	+	+
Interspersed repeat	n, mt	Variable	±	Variable	N.A.	–	+	+	+	+
Inverted repeat	cp	One or two	±	20 000–30 000	N.A.	–	–	–	+	+
Tandem repeats:										
Nuclear rRNA gene cluster	n	One to several	±	9000–11 000	10 ² –10 ⁴	–	–	+	+	+
Nuclear rRNA IGS subrepeat	n	10 ² –10 ⁴ per rDNA locus	–	100–500	≤20	–	–	–	+	+
Satellite	n	One to several	–	2–1000	10 ³ –10 ⁷	+	+	+	+	+
Minisatellite	n	10 ³	–	≤100	10–100	–	–	+	+	+
Microsatellite	n	10 ³ –10 ⁵	±	1–6	5–100	–	–	+	+	+

^an = nuclear, cp = chloroplast, mt = mitochondrial.^bmarker sequence present in coding regions (+), noncoding regions (–), either coding or noncoding regions (±).^cappropriate (+) or inappropriate (–) marker assay.

copies per locus. Thousands of loci in a genome may have similar core repeat units. The number of repeat units at a minisatellite locus can vary greatly between individuals and populations. First described in humans (Jeffreys *et al.*, 1985), minisatellites are found in numerous animal species, often near telomeres. They are also common in plants (Rogstad *et al.*, 1988), and are often associated with satellites and centromeres. The number of repeat units per locus is less variable in plants than in animals, but is still high; plant minisatellites are useful markers for variation between and within species (Rogstad, 1993; Tourmente *et al.*, 1994).

As suggested by their name, microsatellites – also called simple sequence repeats (SSRs), or simple sequence length polymorphisms (SSLPs) – have very short repeat units, no more than six nucleotides long. SSRs are more abundant than minisatellites in noncoding regions of the nuclear genome, and are present in some nuclear genes and organelle genomes (Tautz *et al.*, 1986; Wang *et al.*, 1994; Powell *et al.*, 1995). The number of repeat units per locus is lower for SSRs than for minisatellites, but can approach 100 in animals and 50 in plants (Tautz, 1993; Saghai Maroof *et al.*, 1994). The abundance and polymorphism of SSRs make them particularly valuable for describing variation between populations and individuals (Brown *et al.*, 1996).

Like minisatellites, SSRs were documented first in humans (Tautz *et al.*, 1986; Litt and Luty, 1989; Weber and May, 1989) and later in plants (Condit and Hubbell, 1991). Plants and animals differ in the abundance of specific SSR motifs in the genome. In both plant and animal genomes, chromosomal distribution of SSRs is variable. Some animal SSRs are found near heterochromatin or interspersed repeats, but most are randomly dispersed (Tautz *et al.*, 1986; Miller and Archibald, 1993). Dinucleotide SSRs are randomly distributed in *Arabidopsis thaliana* (L.) Heynh. (Bell and Ecker, 1994), but other studies have located plant SSRs near genes, highly methylated DNA, satellites, or centromeres (Sibson *et al.*, 1991; Bennetzen *et al.*, 1994; Arens *et al.*, 1995; Sharma *et al.*, 1995).

Tandemly repeated genes are also utilized as markers. Perhaps the most widely used are the nuclear genes that encode ribosomal RNA (rRNA) (Long and Dawid, 1980; Hamby and Zimmer, 1992). The three rRNA genes are separated by two internal transcribed spacer regions, generally referred to as ITS1 and ITS2. These genes and spacers form a unit that is tandemly repeated hundreds of times, at one to several loci in the genome. At each of these loci, the individual repeat units are separated by nontranscribed intergenic spacer (IGS) regions. In the middle region of each IGS are tandem copies of a short subrepeat sequence. Variation in rRNA gene clusters can be measured at several levels, each evolving at a different rate: (i) the number and location of rRNA loci, which is highly conserved; (ii) the (more variable) number of tandem gene clusters per locus; (iii) the conserved sequences of the three genes; (iv) the variable sequences of the

ITS regions; and (v) the highly variable number of subrepeats in the IGS region. These features make rRNA gene clusters versatile and informative markers for mapping and phylogenetic analyses (Maluszynska and Heslop-Harrison, 1993; Fukui *et al.*, 1994; Kranz *et al.*, 1995).

2.2.3. Types of molecular marker assays

Molecular marker assays (Table 2.2) are generally classified by whether the molecules evaluated are proteins or nucleic acids, and whether the character analysed in a nucleic acid marker assay is the entire genome, a chromosome, a fragment, or a nucleotide. Alternatively, marker assays can be categorized by the type of character measured (Avisé, 1994). Some methods measure quantitative differences between entries in an array. Others measure qualitative characters, each with two or more possible states. Marker assays also differ in the number of loci evaluated per analysis, whether multiple loci are evaluated simultaneously or sequentially, and the type and amount of information needed about the marker loci before conducting the assay.

Protein marker assays

When proteins are used as genetic markers, the assumption is made that any variation between proteins reflects heritable variation in their amino acid sequences. This assumption has long been debated, since protein phenotypes can be affected by factors such as post-translational modification, plant phenology, and environmental conditions during plant growth (Murphy *et al.*, 1990; Smith and Smith, 1992). For many tasks, however, protein marker assays continue to be useful molecular tools.

Some assays measure immunological properties of proteins. When antigenic proteins of one plant or animal species are injected into an animal (usually a rabbit), antibodies are produced. Genetic differences between entries in an array can be assessed by comparing how each entry's proteins bind to antibody serum (antiserum) from a reference species. This binding affinity is measured by the amount of antibody–antigen precipitate produced or by microcomplement fixation (MCF) (Sarich and Wilson, 1966). Each MCF solution contains antiserum from a reference sample, antigen from a test entry, and a protein complement. The complement can only bind to antibody that is bound to the antigen. The amount of unfixed complement is measured; this is used to calculate the amount of unbound antigen, which indicates the extent of variation (the immunological distance) between the entry and the reference sample. After MCF assays are conducted for each entry in an array, these immunological distances are compared between entries. MCF has long been used in vertebrate phylogenetic analysis (Leone, 1964; Maxson and Maxson, 1990).

Table 2.2. Summary of molecular marker assays used to measure plant genetic variation.

Marker assay	Type of molecule	Genomes assayed ^a	Character analysed	No. of character states ^b	No. of loci per assay ^c	Multilocus analysis ^d	Inheritance ^e
Microcomplement fixation	Protein	Total	Reactivity	Quant.	One to many	Sim.	–
Monoclonal antibody assay	Protein	Total	±Reactivity	2	One or several	Sim.	–
Protein electrophoresis	Protein	n, cp	Electromorph	<10	One to several	Seq.	Codom.
CsCl density gradient	DNA	Total	Buoyant density	Quant.	Many	Sim.	–
DNA–DNA hybridization	DNA	Total	ΔT_m	Quant.	Many	Sim.	–
Flow cytometry	DNA	n	DNA content, no. of chromosomes	Quant.	Many	Sim.	–
Chromosome banding	DNA	n	±Band	2	Many	Seq.	Codom.
Fragments electrophoresed, then detected:							
Single-copy or low-copy RFLP assay	DNA, RNA	n, cp, mt	±Restriction site	2	One to several	Seq.	Codom.
Multilocus restriction fragment assay	DNA	n, cp, mt	±Fragment	2	Many	Sim.	Dom.
Arbitrary PCR	DNA, RNA	n, cp, mt	±Fragment	2	Many	Sim.	Dom.
Designed-primer PCR	DNA, RNA	n, cp, mt	Fragment length	≥2	One to many	Seq.	Codom.

Fragments detected directly,
without electrophoresis:

Dot or slot blot hybridization	DNA, RNA	n, cp, mt	±Fragment Signal intensity	2 Quant.	Many Many	Sim. Sim.	Dom. –
<i>In situ</i> hybridization	DNA	n	±Fragment	2	One to many	Seq.	Dom.
Nucleic acid sequencing	DNA, RNA	n, cp, mt	Nucleotide	4	One	–	Dom.or codom.

^an = nuclear, cp = chloroplast, mt = mitochondrial.

^bQuant. = quantitative.

^cIn the nuclear genome; the chloroplast and mitochondrial genomes are each considered as one locus.

^dLoci analysed simultaneously (sim.) or sequentially (seq.).

^eFor loci in diploid (nuclear) genomes, dominant (dom.) or codominant (codom.) inheritance of alleles.

Until recently, immunological markers were seldom used to study plant variation, perhaps because few suitable proteins had been identified. During the past few years, however, monoclonal antibodies to seed proteins of several cereal species have been produced. Reactivity of these antibodies with antigenic proteins can be evaluated by enzyme-linked immunosorbent assays (ELISAs). Monoclonal antibodies have been used to identify cDNA library clones and describe variation between crop species and cultivars (Donovan *et al.*, 1989; Esen *et al.*, 1989; Zawistowski and Howes, 1990).

Other marker assays measure the rate of protein migration through a starch, polyacrylamide, or cellulose acetate gel in response to an electrical current. The migration rate is related to protein size, shape, isoelectric point and /or ionic charge. In these assays, protein extracts from entries in an array are loaded in adjacent gel lanes and electrophoresed. The gel is then treated with a histochemical stain that reacts with a specific marker protein. Variation between the entries is measured by the positions of their stained proteins on the gel (Murphy *et al.*, 1990). Many of the proteins used as markers are isozymes – functionally similar forms of an enzyme, encoded by different loci or different alleles at a locus. In general all alleles at a locus are expressed, so allelic isozymes can be analysed as codominant variants. Isozymes are frequently used to describe variation between and within plant populations, species, and genera (Brown, 1978; Gottlieb, 1981; Hamrick and Godt, 1989); for mapping and diagnostics; and in some cases for identification. When comparing higher taxonomic levels, the probability of genetically different electromorphs migrating to the same gel position is increased. For closely related taxa and populations, however, isozyme analysis is informative and continues to be refined (Smith and Smith, 1992; O'Neill and Mathias, 1995).

Immunoelectrophoretic assays measure both protein migration rate in gels and the antibody specificity of proteins. Antigenic proteins are electrophoretically separated, then are detected by their reactions with antibodies. This technique is valuable for diagnostics, mapping, and describing variation between plant cultivars, species, and higher taxa (Stegink and Vaughn, 1990; Hilu and Esen, 1993; Masojc *et al.*, 1993).

Nucleic acid marker assays: total genomic DNA and chromosome markers

Although proteins are useful as genetic markers, the phenotype each protein describes is determined by the genotype plus the type of tissue sampled, plant phenological stage, environment, and post-translational processing. In addition, the percentage of a genome sampled by protein markers is limited: these markers only describe coding sequences, and not all proteins can be separated and detected by established methods. Some questions are best addressed by markers assumed to be selectively neutral, but this assumption is in question for many proteins (Bachmann, 1994).

Finally, some proteins are unique to particular taxonomic groups, and thus are suitable as markers only in a limited range of taxa.

In contrast, nucleic acids are present in all organisms and are a common genetic currency for comparing virtually any taxa. Nucleic acid markers describe genotypes, not phenotypes, and can sample both coding and noncoding regions of a genome. DNA methylation and secondary structure can cause artefactual variation in some assays, disrupting the direct relationship between marker and genotype. However, these artefacts can often be detected and eliminated.

The first nucleic acid marker assays measured properties of total genomic DNA. When centrifuged in a CsCl density gradient, DNA separates into bands according to GC content. Because highly repetitive (satellite) sequences differ in GC content from the rest of the genome, these gradients can be used to isolate and quantify satellite DNA (Ingles *et al.*, 1973). This method has been utilized to compare satellite DNA content between plant species (Beridze, 1975), and is still used to detect highly repetitive DNA (Cavallini, 1993).

Another technique that separates repetitive and low-copy DNA relies on DNA reassociation kinetics (Werman *et al.*, 1990). The two complementary strands of a DNA molecule denature when heated and reassociate into duplexes when cooled. At a standard temperature and salt concentration, reassociation rate depends on the percentage of strands with similar sequences. Repetitive sequences reassociate rapidly and at high temperatures, compared to single-copy sequences. Thus highly repetitive, middle repetitive, and single-copy fractions of a genome can be isolated by fragmenting, heating, and cooling double-stranded DNA. Because reassociation rate is affected by fragment length and the pattern of sequence repetition, the type and dispersion of repeats can be measured by varying the lengths of the fragments analysed. This method has been used to characterize and compare repeat sequences between plant species and hybrids (Kumar *et al.*, 1992; Aliev, 1993).

DNA–DNA hybridization employs the same principles, but with single-copy sequences, to compare entries in an array (Werman *et al.*, 1990). When single-stranded DNA from two entries is combined, their overall sequence similarity is related to the temperature at which 50% of the strands form heteroduplexes. This T_m is compared to the homoduplex T_m for each entry; ΔT_m measures the genetic distance between entries. DNA–DNA hybridization is used to examine variation between species and higher taxa, usually with vertebrates (Britten and Kohne, 1968; Wetmur and Davidson, 1968) and less often with plants and bacteria (Egel *et al.*, 1991; Kumar *et al.*, 1992).

Variation in DNA content of nuclei and chromosomes can be measured by flow cytometry (Bennett and Leitch, 1995; Heslop-Harrison, 1995). Particles are isolated, suspended, stained with a DNA-binding

fluorochrome, and passed through a cytometer. The fluorescent signal intensity from each particle indicates its DNA content. Flow cytometry was first utilized to analyse mammalian genomes, but has since been used to compare nuclear DNA content and ploidy level between plant families, genera, species, and ecotypes and to evaluate wide hybrids (Arumuganathan and Earle, 1991; Hammatt *et al.*, 1991; McMurphy and Rayburn, 1991; Dickson *et al.*, 1992). Because chromosomes can be identified by their signal intensity, flow cytometry can be used to generate high-resolution karyotypes and sort chromosomes for mapping (Doležel and Lucretti, 1995).

Karyotypes of chromosomes prepared in mitotic metaphase squashes have long been used for diagnostics, mapping, and describing genetic variation between plant species and hybrids (Sessions, 1990). Entries in an array can be characterized by their chromosome number and morphology. In addition, chromosome spread preparations can be treated with stains specific to AT-rich regions, heterochromatin, or nucleolar organizing regions. The stain-specific banding patterns of individual chromosomes are then used as markers to characterize and compare entries. Chromosome banding is informative for mapping studies and describing variation between genebank accessions and cultivars (Zeller *et al.*, 1991; Friebe *et al.*, 1992). Banding is often used in combination with flow cytometry or other techniques (Sukias and Murray, 1990). For chromosomes that are very small or lack distinctive bands, cytogenetic techniques may have limited use. As high-resolution microscopes and more sensitive stains are developed, however, the value of these assays continues to increase (Khuong and Schubert, 1990; de Carvalho and Saraiva, 1993; Murray, 1994).

Nucleic acid marker assays: fragment markers

Nucleic acid fragments are widely used as genetic markers. As mentioned above, CsCl density gradients and DNA reassociation techniques can detect and roughly quantify satellite DNA sequences. Yet these methods do not measure discrete character differences or variation in the sequences themselves. In addition, differences in GC content between genomes can significantly affect the results (Werman *et al.*, 1990).

In contrast, fragment markers provide high-resolution qualitative information about sequence variation. Either DNA or RNA can be evaluated, and minimal amounts are required. Numerous fragment marker assays are available; all involve detecting fragments that differ in presence, size, or quantity between entries in an array.

Fragments that are produced, electrophoresed, then detected. In many assays, fragments are produced, separated by agarose or polyacrylamide gel electrophoresis, then detected. The length and conformation of a fragment determine its migration rate in a gel. Several different approaches are

taken for detecting fragments. In some assays with purified sequences, all fragments are detected and visualized. The fragments are either labelled before or stained after electrophoresis. Ethidium bromide, silver, fluorescent, and radioactive stains and labels are commonly used. In other assays, specific target fragments are detected and visualized by transfer hybridization (Southern, 1975). All fragments in an assay are electrophoresed, denatured if double-stranded, then transferred from the gel to a nylon or nitrocellulose membrane. A labelled single-stranded probe complementing the target sequence is hybridized to the membrane and anneals to the target fragments. The extent of mismatch hybridization varies with salt and formamide concentration and the temperature during hybridization (Dowling *et al.*, 1990).

A variety of probes can be used: synthesized oligonucleotides, uncharacterized clones from genomic DNA libraries, fragments isolated from other gels, RNA sequences, or total genomic DNA. The probe may be from the same species as the membrane-bound DNA or from a different species. Until recently most probes were radioactively labelled, but nonradioactive colorimetric, fluorescent, and chemiluminescent labels are increasing in popularity (Mansfield *et al.*, 1995).

Fragments produced by restriction digest. Restriction site analysis is widely used in plant genetics. The principles, protocols and data interpretation have been discussed in detail (Sambrook *et al.*, 1989; Dowling *et al.*, 1990; Ausubel *et al.*, 1994). Restriction fragments are generated by treating double-stranded DNA with restriction endonucleases (REs), which are enzymes produced by bacteria. Each RE cleaves double-stranded DNA at a specific sequence, usually four to six nucleotides in length. Several hundred REs have been characterized; they vary in recognition sequence and ability to digest methylated DNA (Roberts, 1984). DNA sequence variation between the entries in an array is detected as variation in the number and/or lengths of restriction fragments. This variation can be caused by point mutations at RE recognition sites or by the insertion, deletion or rearrangement of sequences at or between RE sites.

When single- or low-copy fragments are evaluated, variations in fragment length are referred to as restriction fragment length polymorphisms (RFLPs). In some RFLP assays of highly purified DNA sequences, all of the fragments are visualized by labelling them before or staining them after electrophoresis. In most RFLP assays, however, target fragments are detected by hybridized probes. For each entry in an array, the presence or absence of each individual fragment is recorded. These fragment data are used to infer the presence or absence of each restriction site. For diploid genomes, entries heterozygous for the presence of a restriction site can be distinguished from those that are homozygous. Although few RFLP loci are examined on each gel, every locus can be genetically defined. RFLP

analysis can be expensive, but costs per trial can be reduced by probing each membrane several times with different probes. Single- or low-copy RFLPs are used extensively in diagnostics, for mapping, and to verify interspecific hybridization and study genetic relationships and structure (Dowling *et al.*, 1990; Bates, 1992; Schwarzhacher, 1994). Probes that hybridize across species are especially useful for comparative mapping (Ahn *et al.*, 1993).

In most restriction site analyses of repetitive DNA, target fragments are detected by probing with a known repeat sequence. The fragments, not the genetically defined sites, are usually the characters analysed. Entries that are homozygous for the presence of a fragment cannot be distinguished from heterozygotes. These assays provide information about variation at many loci simultaneously, but reveal little about the type of sequence variation. Restriction site analysis is often used to study interspersed repeats (Kumar *et al.*, 1990) and satellites (Sibson *et al.*, 1991). It was utilized in the first assays of hypervariable loci (Jeffreys *et al.*, 1985), and is presently the most common method in forensics for evaluating minisatellite VNTR loci. Multilocus restriction fragment profiling of plant genomes is useful for identifying clonal cultivars and studying breeding systems (Smith and Smith, 1992; Nybom, 1993). In some studies, fragment length polymorphism can be interpreted as variation of codominant alleles at a locus. However, because inheritance of markers in multilocus profiles is seldom clearly defined, their use in phylogenetic and gene flow studies may be limited (Avisé, 1994).

Restriction site analysis can generate highly reproducible results, if all steps of the process are carefully controlled. With a variety of probe-enzyme combinations available, assays of the nuclear genome can be used to address many questions about genetic variation at or below the species level. As in isozyme analysis, the probability of genetically different restriction fragments co-migrating on a gel is increased when higher taxonomic levels are compared. Because chloroplast genomes evolve slowly, chloroplast restriction site assays are often used to study variation between higher taxa (Clegg and Zurawski, 1992; Downie and Palmer, 1992; Badenes and Parfitt, 1995). In some species, both chloroplast and nuclear genomes are biparentally inherited. In many plant species, however, chloroplasts are either maternally or paternally inherited. For these species, chloroplast restriction site analysis is very informative for studies of gene flow and interspecific hybridization, especially when used in combination with assays of biparentally inherited nuclear markers (Arnold *et al.*, 1991; McCauley, 1995).

Fragments produced by the polymerase chain reaction. Fragment analysis began in the 1970s with the introduction of RFLPs, and grew exponentially a decade later when the polymerase chain reaction (PCR) was developed

(Saiki *et al.*, 1985; Sambrook *et al.*, 1989; Ausubel *et al.*, 1994). PCR uses the principles of DNA reassociation and the action of DNA polymerase to amplify nucleic acid fragments *in vitro*. Each PCR reaction solution contains a double-stranded DNA template, short single-stranded oligonucleotide primers, thermostable DNA polymerase, enzyme cofactors, and the four deoxynucleotides (dNTPs: dATP, dCTP, dGTP, dTTP). The template is denatured by heating; the temperature is lowered, and the primers anneal to complementary sequences on the template. The solution is heated again, and the polymerase adds dNTPs to the 3' end of each annealed primer. Double-stranded fragments are produced and serve as templates for the next cycle, which generates fragments with primer sequences at each end. During subsequent cycles, these fragments are amplified at a geometric rate.

With automated thermocyclers and cheaper, improved reagents, PCR is fast becoming more reproducible and affordable (Erlich *et al.*, 1991). A wide range of known and unknown sequences can be amplified from fresh or ancient DNA and from RNA, and little template is needed; hence PCR is applicable to any taxon. However, contamination is a concern, since contaminant DNA in a reaction solution can function as a template. The specificity and reproducibility of PCR depend on several factors: the concentration and quality of the reaction ingredients; the design and GC content of the primers; and the temperature, duration and number of cycles. However, measures can be taken to minimize and detect contamination (Belák and Ballagi-Pordány, 1993) and optimize cycling protocols (Innis and Gelfand, 1990; Caetano-Anollés, 1993). The disadvantages of PCR are greatly outweighed by the tremendous opportunities it provides.

When amplified fragments are used for high-resolution mapping, diagnostics or forensics, they are generally detected by probes (Belák and Ballagi-Pordány, 1993). When used to describe variation between and within species, amplified fragments may be detected by labelling primers or dNTPs in the reaction solution, labelling PCR products, or staining gels after electrophoresis.

Numerous PCR-based marker assays have been developed and used in plant genetics. They can be categorized by whether target sequences are known prior to amplification, whether the primer sequences are designed or are arbitrary, the number of primers, and the size range of amplified products (Table 2.3). In the first PCR assays the target loci were known single-copy sequences, amplified by two primers designed to complement the regions flanking the loci (Mullis and Faloona, 1987). Primer pairs have since been designed for numerous single-copy loci in plant and animal nuclear and organelle genomes. These loci are very useful as markers for mapping, diagnostics, phylogenetic analyses, and studying hybridization (Olson *et al.*, 1989; Paran and Michelmore, 1993). In most genome mapping projects, cloned DNA sequences present at only one site in a genome are

Table 2.3. Summary of fragment marker assays.

Marker assay ^a	Target sequence ^b	Fragment production ^c	No. of primers or probes ^d			Fragment size (kb)
			A	D	C	
<i>Restriction fragment assay</i>						
Single-copy sequence	U or K	RD		Variable		≤20
Repetitive sequence	K	RD		1		≤20
<i>Amplified fragment assay</i>						
DAF	U	PCR	1			≤1
RAPD	U	PCR	1			0.5–3
Mini-hairpin DAF	U	PCR			1	≤1
Inter-repeat PCR	U	PCR		≥1		Variable
Anchored PCR	U	PCR	1	1		Variable
Designed-primer PCR, single-copy or low-copy sequence	K	PCR		2		≤5
Designed-primer PCR, tandem repeat locus	K	PCR		2		≤0.5
Nested PCR	K	PCR then PCR		2 per reaction		Variable
<i>Restricted and/or amplified fragment assay</i>						
Cleaved template	U or K	RD then PCR		Variable		Variable
Cleaved fragment	U or K	PCR then RD		Variable		Variable
Single-stranded fragment	U or K	RD or PCR then denature		Variable		Variable

^aDAF = DNA amplification fingerprint, RAPD = random amplified polymorphic DNA, PCR = polymerase chain reaction.

^bU = unknown, K = known.

^cRD = restriction endonuclease digest.

^dA = arbitrary, D = designed, C = combination of arbitrary and designed.

used as physical mapping landmarks. Primer pairs have been designed to amplify many of these landmarks as sequence tagged sites (STSs) (Olson *et al.*, 1989). This strategy is used extensively in the Human Genome Project (HGP) (Collins and Galas, 1993) and other mapping programmes.

In assays of these designed-primer PCR loci, the character evaluated may be the fragment itself, with presence dominant over absence. Often, however, inheritance can be defined. For loci in diploid genomes, differences in fragment length can then be analysed as codominant alleles.

Both interspersed and tandemly repeated sequences can be amplified by designed-primer PCR. When several interspersed repeat loci are amplified in one sample, each locus is evaluated separately. When tandem repeat loci are amplified, length polymorphism at a locus is assumed to

reflect variation in the number of tandem repeat units. Amplification of human minisatellites with designed primer pairs (Decorte *et al.*, 1990) was soon followed by amplification of plant minisatellites and SSRs (Condit and Hubbell, 1991; Akkaya *et al.*, 1992; Saghai Maroof *et al.*, 1994; Tourmente *et al.*, 1994). Amplified plant SSRs are informative for mapping, genotyping, diagnostics, and population studies (Brown *et al.*, 1996). SSR alleles may differ little in length, and discriminating them can be difficult. New electrophoresis, labelling and imaging technologies are making detection easier (Schwengel *et al.*, 1994).

Development of designed-primer PCR markers involves several steps. Target sequences are identified and isolated, usually from a genomic DNA library; the loci and flanking regions are sequenced; and primer pairs are designed and synthesized. Currently these tasks are expensive, and not all laboratories have the equipment required. Once primers are designed and synthesized, however, the costs for each analysis are not high. The initial expenses of marker development may be reduced by identifying the desired sequences from nucleic acid databases, e.g. Genbank or EMBL (Brown *et al.*, 1996). However, this is only applicable to species well represented in the databases. As another cost-cutting measure, some primers designed to amplify loci in one species can be used to amplify loci in related taxa (Schlötterer *et al.*, 1991; Kresovich *et al.*, 1995). Primers that amplify across taxa are also useful for phylogenetic analyses and studying genome evolution.

For development of designed-primer PCR markers, both target and flanking region sequences must be known. Soon after PCR was introduced, however, its ability to amplify unknown sequences was realized (Saiki *et al.*, 1988). In inter-repeat PCR, primers that complement interspersed repeats or other known sequences are used to amplify unknown sequences between the repeats (Ledbetter *et al.*, 1990; Rogowsky *et al.*, 1992). 'Universal' primers have been designed to complement conserved coding sequences in mitochondrial and chloroplast genomes; these primers can be used to amplify the noncoding regions between genes (Taberlet *et al.*, 1991; Demesure *et al.*, 1995).

Other marker assays require no prior sequence information, for either the target region or the flanking regions. These assays – randomly amplified polymorphic DNA analysis (RAPD; Williams *et al.*, 1990), arbitrarily primed PCR (AP-PCR; Welsh and McClelland, 1990), and DNA amplification fingerprinting (DAF; Caetano-Anollés *et al.*, 1991) – use one short primer that has an arbitrary sequence. This primer anneals to the template at complementary sequences in both 'sense' and 'antisense' orientation, and the regions between primers in opposite orientation are amplified. The number and length of fragments amplified by one primer can vary; short primers generally produce a large number of small fragments. In single-primer PCR, entries that are homozygous for the presence of a

fragment cannot be distinguished from heterozygotes; thus fragment presence is generally dominant over absence. Many loci can be evaluated simultaneously, although little genetic information about each locus is provided. Arbitrary PCR is particularly valuable when no sequence information is available for the taxa being studied. It has been used extensively to examine variation between and within plant species, identify genebank accessions and cultivars, and construct genetic maps (Caetano-Anollés *et al.*, 1991; Tingey and del Tufo, 1993; Kresovich *et al.*, 1994).

Many PCR assays are variations or combinations of designed-primer PCR and arbitrary PCR (Table 2.3). Most were developed to increase the reproducibility of results and / or to better discriminate between closely related genotypes. In arbitrary PCR, reproducibility and the extent of mismatch primer annealing are highly dependent on the cycling parameters and concentration of reaction components (Ellsworth *et al.*, 1993; Micheli *et al.*, 1994). One solution to this problem is to identify polymorphic fragments in arbitrary PCR assays and then isolate, clone, sequence and design primers for these fragments (Paran and Michelmore, 1993). This process generates new designed-primer PCR loci but involves many steps. Other assays with increased reproducibility combine features of single-primer and inter-repeat PCR: unknown sequences are amplified by one primer that complements an SSR, its flanking region, or both (Meyer *et al.*, 1993; Gupta *et al.*, 1994; Wu *et al.*, 1994; Kuramoto *et al.*, 1995; Weising *et al.*, 1995). Finally, 'anchored' PCR uses one designed and one arbitrary primer to amplify unknown sequences (Miller and Archibald, 1993; Wu *et al.*, 1994).

For genotyping, diagnostics, and assessing variation between close relatives, PCR-based marker assays with high resolution are needed. One such variation is nested PCR. In this technique, one amplification reaction with two designed primers is followed by a second reaction. The first PCR product is the template for the second reaction, which uses two designed primers internal to the first set (Mullis and Faloona, 1987; Livache *et al.*, 1994). Although the possibility of contamination is high (Belák and Ballagi-Pordány, 1993), nested PCR is valuable for extremely sensitive genotyping and diagnostics. At present, it is used most often in medical research. Resolution of arbitrary PCR can be increased by adding stable mini-hairpin sequences to the 5' ends of short primers. This technique greatly improved discrimination between cultivars of centipedegrass [*Eremochloa ophiuroides* (Munro) Hackel] (Caetano-Anollés and Gresshoff, 1994).

Other marker assays with improved resolution combine principles of PCR and restriction site analysis. RE digestion of PCR templates before they are amplified was first suggested by Mullis and Faloona (1987). Arbitrary PCR with cleaved templates has since been used to map plant genomes and identify genotypes (Caetano-Anollés *et al.*, 1993; Koebner, 1995). Amplified fragment length polymorphism (AFLP) analysis is a variation of cleaved-template PCR. After DNA is digested, a short adaptor

sequence is ligated to each end of the restriction fragments. The fragments are then used as templates for single-primer PCR (Zabeau and Vos, 1993). The primer contains a constant sequence (the adaptor and RE recognition site sequences) plus selective bases that determine the specificity of primer annealing. Both reproducibility and resolution are increased by using AFLPs.

Resolution can also be increased by digesting PCR products before electrophoresis (Ochman *et al.*, 1988). Digested products of both designed-primer PCR and arbitrary PCR have been used to measure variation between genebank accessions (Halward *et al.*, 1991), refine RFLP maps (Jarvis *et al.*, 1994), and study interspecific gene flow (Arnold *et al.*, 1991). Resolution of both PCR and restriction site assays can be improved by denaturing double-stranded DNA fragments before they are electrophoresed. Due to internal point mutations or differences in conformation or GC content, the two single strands from one fragment may have different migration rates on polyacrylamide gels. Both denaturing gradient gel electrophoresis (DGGE) and single-stranded conformation polymorphism (SSCP) analysis can detect differences between closely related genotypes (Myers *et al.*, 1985; Orita *et al.*, 1989; Dweikat *et al.*, 1993; McClelland *et al.*, 1994).

Fragments detected directly by probe hybridization. In some marker assays the target sequences are not separated by electrophoresis, but are detected directly by probe hybridization. The types of probes used are the same as in electrophoretic assays, and can be radioactively or nonradioactively labelled.

In dot and slot blot hybridization, single-stranded DNA fragments, genomic DNA, or RNA is immobilized on a nylon or nitrocellulose membrane. Labelled probes complementing the target sequence are hybridized to the membrane blot. All loci in the blot are sampled simultaneously; abundance of the target sequence is measured by the intensity of the hybridization signal. Many blots on one membrane can be evaluated at the same time, and the procedure is relatively simple. Because the presence of nontarget sequences in the blot can affect hybridization, precise measurement of target sequence abundance in genomic DNA blots can be difficult. When the blots are carefully prepared and hybridization conditions are controlled, however, these assays can generate valuable information for diagnostic and relationship studies. Dot and slot blots have been used to detect specific alleles of genes (Saiki *et al.*, 1986) and RAPD loci (Aitken *et al.*, 1994), estimate the copy number of satellites (Halldén *et al.*, 1987) and interspersed repeats (Aledo *et al.*, 1995), and compare gene expression between organs of a plant (Holdsworth *et al.*, 1988).

In another technique, *in situ* hybridization (ISH), probes are hybridized to the same chromosome spread preparations used for karyotypes

and banding pattern assays (Sessions, 1990). ISH is used to detect the presence and chromosomal distribution of target sequences. While banding assays are only appropriate with chromosomes that contain specific sequences, ISH can detect a wide range of sequences. Radioactive and non-radioactive probes are used, but the popularity of fluorescent *in situ* hybridization (FISH) is rapidly increasing (Lichter *et al.*, 1990).

ISH has many applications for genome analysis. These include 'painting' the genome of one species onto chromosomes of another, to study species evolution (Jellen *et al.*, 1994); painting a genome with one of its own chromosomes, to study chromosome evolution (Vega *et al.*, 1994); and characterizing the chromosomal distribution of interspersed repeats (Aledo *et al.*, 1995). ISH is often used in combination with other marker assays. In wide crosses and somatic hybrids, ISH with species-specific interspersed repeat probes can indicate the abundance and chromosomal locations of transferred sequences; restriction site analysis can provide information about structure and rearrangement of the sequences (Ananthawat-Jónsson and Heslop-Harrison, 1992; Bates, 1992). Physical mapping with ISH can complement RFLP genetic mapping studies (Maluszynska and Heslop-Harrison, 1993; Schwarzacher, 1994). Aledo *et al.* (1995) used slot blots to estimate the copy number of an interspersed repeat in the maize (*Zea mays* L.) genome, and used FISH to determine the chromosomal distribution of the repeat. With both ISH and heterochromatin-specific staining on the same chromosome preparation, Harrison and Heslop-Harrison (1995) confirmed that two *Brassica* satellite families are located near centromeres. By varying the hybridization stringency, differences between satellite sequences on different chromosomes were revealed.

ISH can be time consuming, and results can be affected by the specific methods used for chromosome preparation, denaturation, and probe hybridization and detection. However, new protocols are increasing the reproducibility of ISH. Fukui *et al.* (1994) used a thermocycler to control temperatures during denaturation, and developed accurate imaging methods for detecting and analysing fluorescent signals. Because of such innovations, the value of ISH as a molecular tool will most likely continue to increase.

Nucleic acid marker assays: nucleotide sequencing

Nucleic acid variation can be described most directly by using the nucleotides themselves as discrete characters, each with four possible states. Because nucleotides are the basic units of genetic information in all organisms, sequencing can be used to detect very low levels of variation. Nuclear and organelle genomes of virtually any taxon can be sequenced; minimal amounts of DNA or RNA are needed, and fresh or preserved DNA can be used. With most marker assays, entries in a trial are compared to each other; results are not easily compared across trials that have no standard entries in common. In contrast, sequences from one study can be

compared with sequences from any other study (Avisé, 1994). In addition, sequencing can reveal whether variation is due to nucleotide substitution or sequence rearrangement.

Two sequencing methods have been used since the late 1970s, and are described in detail by Hillis *et al.* (1990). In both methods, labelled fragments are generated from DNA or RNA that has been isolated and purified; the fragments are electrophoresed; and the sequence is 'read' from the fragment patterns. Maxam–Gilbert (chemical) sequencing is used to sequence double-stranded DNA (Maxam and Gilbert, 1977). DNA is end-labelled, then divided into four reaction solutions. Each solution contains one of four reagents, each of which cleaves DNA at specific bases. The sequence of the DNA determines the lengths of the fragments produced. *In vitro* (dideoxy chain termination) sequencing is applicable to both DNA and RNA (Sanger *et al.*, 1977). The single-stranded template is divided into four reaction solutions. Each solution also contains the following: a labelled oligonucleotide primer that complements one end of the template; the four dNTPs; and one of four dideoxynucleotides, each of which terminates extension at one of the four bases. In each reaction, the primer is extended along the template. As in chemical sequencing, the template sequence determines the lengths of the fragments produced. Chemical sequencing is preferred for some genomic regions with complex secondary structure; in general, however, the chain-termination method is more commonly used. Since it has been coupled with PCR, combined with fluorescent labelling and automation, this technique is rapidly becoming the method of choice (Yang and Youvan, 1989; Griffin and Griffin, 1993).

Sequencing can be used to address a wide range of plant genetic resource questions about virtually any taxa, but involves considerable investment of money, labour and time. Automation and large-scale production of cheaper reagents have made sequencing more cost effective, but few laboratories can afford to sequence large genome segments from multiple entries. As a result, in many studies only one locus or a few loci are sequenced for a limited test array. While many adjacent nucleotides may be evaluated, together they sample only one locus. Thus sequencing is usually not practical for mapping, multilocus forensics, population genetics, or analyses of intraspecific variation. It is often used for high-resolution diagnostics, studying specific genes, and describing variation between genera and higher taxa. In phylogenetic studies, however, inferring genome evolution from sequence variation at a single locus should be done with caution if at all (Doyle, 1991).

Many regions of plant nuclear and organelle genomes have been sequenced and used as markers. These sequences vary greatly in evolutionary rate and the taxonomic levels at which they are employed. Chloroplast genomes evolve slowly; sequences of the entire genome or its *rbcL* gene are informative as markers for describing variation between

genera and higher taxa (Clegg and Zurawski, 1992; Soltis *et al.*, 1992). Nuclear genes that have been sequenced and used as markers include the phytochrome, heat shock protein, and nuclear rRNA gene families (Soltis and Soltis, 1995). In general, the rRNA coding regions are highly conserved; internal transcribed spacers are more variable, and intergenic spacers are highly polymorphic (Hamby and Zimmer, 1992; Kranz *et al.*, 1995). Finally, sequencing is often used – in combination with other marker assays – to characterize tandem repeat loci (Jeffreys *et al.*, 1985; Sibson *et al.*, 1991).

While sequencing may be the ultimate molecular tool, caution should be taken to avoid several potential problems. Because each nucleotide is a separate character, sequence fidelity is crucial. Thus nucleic acids must be carefully isolated and purified before they are sequenced. In addition, PCR contamination and amplification errors have more serious effects for sequencing than for fragment analysis. Amplification errors can be minimized by optimizing PCR conditions and using polymerases with proof-reading capability (Saiki *et al.*, 1988; Erlich *et al.*, 1991; Tong and Smith, 1993). Mistakes during gel reading (manual or automated) must also be considered and minimized. Subjectivity is a real concern when aligning sequences before they are compared. Alignment is complicated and requires assumptions about expected amounts of base substitution and insertion / deletion (Hillis *et al.*, 1990). Despite these concerns, sequencing remains the tool of choice for many studies.

2.3. Guidelines for Selecting Marker Assays

Choosing appropriate marker assays can be challenging, but several considerations can make the task easier. Important issues are: (i) what question is being asked? (ii) what level of resolution is required? (iii) how can the results be related to characteristics of the taxa being studied? and (iv) are sufficient resources available in terms of personnel, equipment, finances, and time? (Kresovich and McFerson, 1992).

Questions of identity are best addressed with high-resolution assays and large numbers of codominant marker loci that are evaluated sequentially (Avisé, 1994). Electrophoretic assays of isozymes and nucleic acid fragments meet most of these criteria (Table 2.4). Genetically defined isozymes, amplified SSR loci, and single-copy RFLPs are very informative markers for identification. Due to practical constraints, however, multilocus assays of restriction fragment and PCR markers (VNTRs, RAPDs and AFLPs) are often used. Dot blots, sequencing and monoclonal antibody assays are also appropriate, but the costs of analysing many loci per assay may be prohibitive.

The most appropriate assays for diagnostic questions have high levels of resolution and use genetically defined, codominant markers. Because

Table 2.4. Application of molecular marker assays to plant genetic variation questions.^a

Marker assay	Identity	Diagnostic	Mapping	Relationship and structure	
				Within species	Between taxa
Protein microcomplement fixation	—	—	—	+	+
Protein monoclonal antibody assay	+P	+	—	+	—
Protein electrophoresis	+	+	+	+	—
CsCl density gradient, DNA–DNA hybridization	—	—	—	—	+
Flow cytometry	—	—	+	+	+
Chromosome banding	—	—	+	+	+
RFLP and PCR fragment electrophoresis	+	+	+	+	—
Dot and slot blot hybridization	+P	+	—	+P	+P
<i>In situ</i> hybridization	—	+	+	+	+
Nucleic acid sequencing	+P	+	+P	+P	+

^a+, appropriate use; —, inappropriate type of character or problems with resolution; P, appropriate use but possible practical constraints.

these questions can often be addressed with one or few loci, useful methods include electrophoretic assays of isozymes, single-copy RFLPs, or SSRs; dot or slot blot assays; and sequencing. For some questions, monoclonal antibody assays and ISH are also appropriate. Mapping studies often use a combination of tools: isozyme or DNA fragment electrophoresis for genetic maps, and chromosome banding, flow cytometry, ISH, and STS markers for physical mapping (Schwarzacher, 1994).

For questions of relationship and structure, important issues include the type of question asked and the range and levels of taxa evaluated.

Because evolutionary questions must be answered with qualitative data, the assays used for diagnostics and identification are appropriate. Questions of similarity can be addressed with these qualitative marker assays or with quantitative assays (microcomplement fixation, flow cytometry, CsCl density gradients, and DNA–DNA hybridization). The latter two techniques have low levels of resolution, and are generally used to compare genera and higher taxa. Interspecific and intraspecific variation is more appropriately described with isozyme and DNA fragment markers, which have higher levels of resolution.

In electrophoretic assays of isozymes and DNA fragments, co-migration of genetically dissimilar electromorphs or fragments is likely when higher taxa are compared. If resolution is too high, however, the assays may be numerically precise but contain random ‘evolutionary noise’ that masks meaningful variation (Bachmann, 1994). Researchers can often decide whether a particular marker assay is suitable for a given study by reviewing the results of previous studies using that technique.

Several characteristics of the taxa being studied are important when selecting a marker and marker assay. One of these characteristics is the mating system. Many apomictic and self-pollinated plant species contain little variation, so techniques with high resolution may be required. When deciding whether to use a chloroplast genome marker, information about chloroplast inheritance is critical. For many marker assays, polyploidy can be a concern. When polyploid taxa are evaluated, the number of loci sampled, the problem of dissimilar fragments co-migrating, and the level of data complexity are often increased. This complexity may be reduced by using genetically defined isozymes, chloroplast genes, genes that evolve slowly, or single-copy markers – rather than RAPDs or multilocus VNTR profiles (Sorrells, 1992; Bachmann, 1994).

Practical considerations are often the most important. The amount and quality of proteins or nucleic acids required can differ greatly between marker assays. Sequencing and PCR-based assays require little template and can be used to analyse ancient DNA samples, but are very sensitive to contamination. DNA–DNA hybridization requires large amounts of DNA, but small amounts of contaminant DNA are less of a problem. For some species, fresh tissue is required for isozyme electrophoresis. For some species and marker techniques, the presence of secondary compounds can complicate storage and extraction of DNA or protein, and can interfere with marker detection.

Marker assays also vary in the amount of setup work required and the extent of sequence information needed before conducting the assay. Karyotyping, ISH, and RFLP assays involve preparing chromosome spreads and/or probes. In contrast, some PCR-based assays are streamlined for high throughput and require little setup work. Sequence information is needed for probe hybridization and designed-primer PCR assays,

but not for arbitrary PCR. Finally, the resources available must be considered. Minimal equipment is needed for protein electrophoresis, but other techniques require thermocyclers, fluorescent microscopes, cytometers and radioactive labelling.

Choices between newly developed and well-documented techniques should be made in light of the expertise and personnel available. Other considerations are reproducibility, number of entries and loci sampled per analysis, and cost per unit of information. Comparative studies can make selection of a marker assay easier and reveal where costs can be reduced (Ragot and Hoisington, 1993). Costs and technologies are constantly changing, however, and the comparisons themselves may be expensive and time consuming; their value must be weighed against the resources they require and the necessity of continual updates. When comparative studies are not feasible, thoughtful consideration of the issues described here may be helpful.

2.4. Future Dimensions of Genetic Analysis

The goals and expectations for analysing plant genetic variation parallel those established across many other fields of biological research, from agriculture, ecology, and evolution to the medical sciences. In all of these fields, future genetic marker assays must incorporate methods to detect, describe, interpret, and store DNA sequence information. Molecular tools of the future are expected to be user friendly, accurate, precise, high throughput, low cost, and potentially automated.

DNA sequence information is the foundation for developing and applying genetic markers to questions of biological variation, whether *in situ* or *ex situ*. Researchers who develop and use sequence-based marker assays for diagnostic, forensic and relationship studies will continue to benefit greatly from information and technologies generated by the international Human Genome Project (HGP). Since its beginning, the HGP has endeavoured to develop genetic and physical maps and determine DNA sequences for the genomes of humans and several model organisms. As part of this effort, genetic analysis methods have been improved significantly. Notable achievements – now taken for granted by many – include new types of genetic markers (particularly those based on PCR), more efficient cloning vectors, the introduction of STS markers, and improved technology and automation for DNA sequencing. Underlying this progress is recognition that the successful development of new technologies for genetic research has been, and will continue to be, critical to many future scientific breakthroughs.

In addition to the HGP, other complementary projects also support initiatives for the development of new molecular marker assays. One such

project is the Advanced Technology Program (ATP) of the US National Institute of Standards and Technology. Among the ATP's goals is working with industry to deliver DNA diagnostic tools to a variety of users, at one-tenth to one-hundredth of the current price. Another goal is helping industry make similar reductions in DNA sequencing costs and make DNA sequencing apparatus available at about one-third of the current price (Advanced Technology Program, 1995).

Many biological questions once considered recalcitrant due to the number and cost of the assays required will be reconsidered in the light of new technological developments. The following section includes examples of progress that can be readily applied to the description of plant genetic variation. Also noted are limitations that are unique to plant-based studies, and are not presently addressed by most genetic analysis programmes.

2.5. Technology Transfer from the Human Genome Project

Technology transfer from the Human Genome Project (HGP) to the study of plant genetic resources has been very important in two areas: (i) sequence characterization and marker development; and (ii) integrating all stages in genetic analysis projects, from template preparation through data interpretation. Some of these innovations are now being used in plant genetic research; others will require some level of optimization before they are applied.

Plant genetic variation is best characterized with a large number of highly informative, easily analysed molecular markers. While the number of PCR-based markers will most likely continue to increase, the question at hand is how these markers will be detected. Manual, autoradiographic detection of amplified markers presently predominates in plant genetic research. However, ongoing programmes in human genotyping via semi-automated, fluorescent detection of SSRs are progressing at an incredible pace (Reed *et al.*, 1994; Schwengel *et al.*, 1994). Reed *et al.* (1994) reported that semi-automated technology can genotype over 100 000 samples, using approximately 250 marker loci, in less than six months. Moreover, Schwengel *et al.* (1994) found that data generated by this method were at least as accurate, more efficient, less labour intensive, and as cost effective as data generated by standard radiolabelling techniques.

Human genotyping kits (Perkin Elmer, Foster City, CA) are available for approximately 400 markers that define a ± 10 cM resolution genomic index map. Kits for plant species and families are in development (J.S. Ziegler, Perkin Elmer, Foster City, CA, 1995, personal communication). Although cost remains an issue, these marker assays are clearly capable of handling the large arrays commonly encountered in plant conservation and breeding programmes.

The costs of generating and detecting numerous marker loci for each of many individuals can be reduced by evaluating several of an individual's loci in a single lane on a gel. In the studies cited above, SSR loci from an individual were multiplexed at the electrophoresis level: several loci were amplified in separate reactions, the PCR products were pooled, and the pooled sample was electrophoresed. In contrast, several laboratories are developing methods to multiplex at the reaction level, amplifying several SSR loci in one tube. Eleven putative loci in canola (*Brassica napus* L.) have been amplified simultaneously, using 11 pairs of fluorescently labelled primers in one reaction (S.M. Mitchell and S. Kresovich, 1995, unpublished results). The loci were distinguished from each other by fragment size and fluorescent label colour. This approach is also under way with SSRs of maize and sorghum [*Sorghum bicolor* (L.) Moench] (E.C.L. Chin, Pioneer Hi-Bred International, Johnston, IA, 1995, personal communication).

Another molecular marker used in human genetic research reveals discrete changes (single nucleotide substitutions) in a specific DNA sequence. Variation of these polymorphic sequence tagged sites (pSTS) is the most frequent and widely distributed type of sequence variation in the genome (Nickerson *et al.*, 1992; Kwok *et al.*, 1994). These individual nucleotide markers are usually biallelic within a population, and are generally less polymorphic than SSRs and other tandem repeats. However, a number of closely linked pSTSs can be combined into a multilocus marker that is as informative as an SSR.

Developing and assaying pSTS markers involve several steps. Random genomic DNA clones (each 300–600 bp long) from selected plant species are sequenced, then single nucleotide differences between individuals are identified by comparing automated sequencing electrophoregrams. Biallelic pSTSs are selected, and primer pairs are designed to amplify them. After a pSTS locus is amplified, the oligonucleotide ligation assay (OLA) is used to label the PCR products and discriminate the two allelic forms (Nickerson *et al.*, 1990; Barany, 1991; Landegren, 1993). In this assay, the PCR products are denatured and then serve as templates for a ligation detection reaction (LDR). This reaction uses three 'probes': two diagnostic probes, each of which complements one of the two pSTS alleles, and one labelled probe that is common to both alleles. The diagnostic probes anneal to complementary template sequences, then are ligated to the labelled probe. The two diagnostic probes differ in length. Thus when the LDR-labelled fragments are electrophoresed, pSTS allelic variation between entries can be detected as variation in migration rate of fragments on the gel, due to differences in fragment length. As in many PCR-based marker assays, significant costs are presently incurred for pSTS marker development. Once markers are developed, however, the costs for each assay are not high.

The previously highlighted markers are usually associated with non-coding regions of the genome. Complementary efforts are under way to develop high-throughput, automated methods to partially sequence cDNA clones, then use them to identify expressed genes in many organisms, including plants (Adams *et al.*, 1991; Waterston *et al.*, 1992; Newman *et al.*, 1994). Thanks to the rapid proliferation of expressed sequence tag (EST) databases, at present the probable gene function of a cDNA clone can often be inferred solely from the similarity of nucleotide or amino acid sequences between the clone and genes of known function (Park *et al.*, 1993; Newman *et al.*, 1994). Using these techniques to compare isoforms of a gene (particularly its untranslated 3' regions) between members of a plant species or family, PCR-based markers that describe variation of this gene could be developed. When assays using these single-gene markers are compared with multilocus studies of the taxon, a more complete picture of genetic variation may be revealed.

2.6. Present Limitations on Describing Plant Genetic Variation with Molecular Tools

Human genetic research often focuses on genomic mapping and analysing specific genes. In contrast, many plant genetic resource questions are best addressed by evaluating numerous individuals from many taxa. Thus widespread use of molecular markers in plant genetic research will involve preparing and analysing large amounts of DNA template and assaying many individuals at the same time. However, techniques have not yet been developed to extract, clone, replicate or amplify DNA in nanolitre volumes (Advanced Technology Program, 1995). Also needed are techniques for pooling DNA samples from individuals in a population, assaying the pooled sample, and generating a comprehensive picture of population variation. In the future, sequence-based approaches may be useful for these tasks (Kwok *et al.*, 1994). At present, high costs make these approaches impractical for many high-throughput applications.

2.7. Thoughts on the Future

In the HGP, technological improvements unanticipated in 1990 have already changed the scope of the research and allowed for more ambitious approaches and goals (Collins and Galas, 1993). In the plant kingdom as well, progressive visions of DNA analysis will most likely change the ways in which problems related to describing genetic variation are perceived and resolved. Beyond the marker assays discussed above, another approach to automated DNA diagnostics merits mention. Sequencing by

hybridization (SBH) uses a large number of short oligonucleotide probes, immobilized in an array on a small solid base. Single-stranded, fluorescently labelled DNA of an unknown sequence is hybridized to the probes on this 'DNA chip', and the sequence is determined from the hybridization pattern (Bains and Smith, 1988; Lysov *et al.*, 1988; Pease *et al.*, 1994; Lipshutz *et al.*, 1995). SBH may be the ultimate application of 'clinical' sequencing to the description of genetic variation. Prior to widespread use of SBH, however, new approaches to synthesizing oligonucleotides, attaching them to solid surfaces, and fabricating the microchips will be required (Advanced Technology Program, 1995).

In order to better characterize plant genetic diversity and address genetic resource questions, plant scientists will need molecular marker assays that cost-effectively detect and describe DNA sequence variation over many areas of the genome (both coding and noncoding), for many individuals in a population or taxon. This goal may seem daunting, but much progress is being made. As long as plant scientists are aware of and build on innovations in other fields, a stream of new tools and approaches will be available for the challenge.

2.8. Conclusions

Many molecular markers and assays are available, and their number and diversity are increasing. This situation simultaneously presents many opportunities and several challenges for studying plant genetic variation. Not least of the challenges is how to choose and use appropriate molecular tools for a given research question. This review is an attempt to categorize and compare the markers and assays presently available, in a framework that makes such decisions easier.

When molecular technologies and their applications are reviewed, two themes become apparent. First, many innovations and new techniques were developed by augmenting features of marker assays already in use. In addition, the most informative studies generally use several approaches (molecular or nonmolecular) to address different aspects of a single question. Thus understanding and having access to a variety of marker assays can greatly benefit those who study plant genetic variation. Second, many molecular technologies described here were originally developed and are most often used in medical research and the Human Genome Project (HGP). Much research related to these marker assays is published in journals not routinely read by many plant scientists. Yet this literature is available and of use to those who recognize the advantages of technology transfer between fields and organisms. Moreover, progress in molecular biology and genetics has broadened the definition of plant genetic resources to include cloned nucleotide sequence libraries and databases of

gene, primer pair, and marker sequence information, in addition to the more traditional collections of living plants, seeds, clones, and *in vitro* cultures (Weissinger, 1990). These new genebanks can contain entries from virtually any known plant, animal, or microbe, whether extant or extinct.

Plant scientists can clearly benefit by becoming part of a much broader scientific community. This integration will facilitate the exchange of marker production and detection methods, as well as genebank entries and molecular marker sequences. Perhaps most importantly, the combined insights from many fields may generate new approaches for asking and addressing questions relevant to all taxa – questions related to the conservation and wise use of genetic resources.

Acknowledgements

The authors gratefully acknowledge S.M. Brown, D.P. Butts, R.R. Duncan, G.D. Kochert and S.E. Mitchell for their advice and encouragement during the preparation of this chapter.

References

- Adams, M.D., Kelly, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, H., Merril, C.R., Wu, A., Olde, B., Moreno, R.F., Kerlavage, A.R., McCombie, W.R. and Venter, J.C. (1991) Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* 252, 1651–1656.
- Advanced Technology Program (1995) *Tools for DNA Diagnostics White Paper*. National Institute of Standards and Technology, Gaithersburg, MD, 11 pp.
- Ahn, S., Anderson, J.A., Sorrells, M.E. and Tanksley, S.D. (1993) Homologous relationships of rice, wheat and maize genomes. *Molecular and General Genetics* 241, 483–490.
- Aitken, S.A., Tinker, N.A., Mather, D.E. and Fortin, M.G. (1994) A method for detecting DNA polymorphism in large populations. *Genome* 37, 506–508.
- Akkaya, M.S., Bhagwat, A.A. and Cregan, P.B. (1992) Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics* 132, 1131–1139.
- Aledo, R., Raz, R., Monfort, A., Vicient, C.M., Puigdomènech, P. and Martínez-Izquierdo, J.A. (1995) Chromosome localization and characterization of a family of long interspersed repetitive DNA elements from the genus *Zea*. *Theoretical and Applied Genetics* 90, 1094–1100.
- Aliev, R.T. (1993) The changes of repeated sequences fractions in plant genomes under heterosis. *Genetika* 29, 990–994.
- Anamthawat-Jónsson, K. and Heslop-Harrison, J.S. (1992) Species specific DNA sequences in the Triticeae. *Hereditas* 116, 49–54.
- Arens, P., Odinot, P., van Heusden, A.W., Lindhout, P. and Vosman, B. (1995) GATA- and GACA-repeats are not evenly distributed throughout the tomato genome. *Genome* 38, 84–90.

- Arnold, M.L., Buckner, C.M. and Robinson, J.J. (1991) Pollen-mediated introgression and hybrid speciation in Louisiana irises. *Proceedings of the National Academy of Sciences of the USA* 88, 1398–1402.
- Arumuganathan, K. and Earle, E.D. (1991) Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter* 9, 208–218.
- Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1994) *Current Protocols in Molecular Biology*, 3 vols. John Wiley, New York.
- Avise, J.C. (1994) *Molecular Markers, Natural History and Evolution*. Chapman & Hall, New York, 511 pp.
- Bachmann, K. (1994) Tansley Review No. 63. Molecular markers in plant ecology. *New Phytologist* 126, 403–418.
- Badenes, M.L. and Parfitt, D.E. (1995) Phylogenetic relationships of cultivated *Prunus* species from an analysis of chloroplast DNA variation. *Theoretical and Applied Genetics* 90, 1035–1041.
- Bains, W. and Smith, G.C. (1988) A novel method for nucleic acid sequence determination. *Journal of Theoretical Biology* 135, 303–308.
- Barany, F. (1991) Genetic disease detection and DNA amplification using cloned thermostable ligase. *Proceedings of the National Academy of Sciences of the USA* 88, 189–193.
- Bates, G.W. (1992) Molecular analysis of nuclear genes in somatic hybrids. *Physiologia Plantarum* 85, 308–314.
- Bélak, S. and Ballagi-Pordány, A. (1993) Experiences on the application of the polymerase chain reaction in a diagnostic laboratory. *Molecular and Cellular Probes* 7, 241–248.
- Bell, C.J. and Ecker, J.R. (1994) Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. *Genomics* 19, 137–144.
- Bennett, M.D. and Leitch, I.J. (1995) Nuclear DNA amounts in angiosperms. *Annals of Botany* 76, 113–176.
- Bennetzen, J.L., Schrick, K., Springer, P.S., Brown, W.E. and Sanmiguel, P. (1994) Active maize genes are unmodified and flanked by diverse classes of modified, highly repetitive DNA. *Genome* 37, 565–576.
- Beridze, T. (1975) DNA nuclear satellites of the genus *Brassica*: variation between species. *Biochimica et Biophysica Acta* 395, 274–279.
- Britten, R.J. and Kohne, D.E. (1968) Repeated sequences in DNA. *Science* 161, 529–540.
- Brower, A.V.Z. and DeSalle, R. (1994) Practical and theoretical considerations for choice of a DNA sequence region in insect molecular systematics, with a short review of published studies using nuclear gene regions. *Annals of the Entomological Society of America* 87, 702–716.
- Brown, A.H.D. (1978) Isozymes, plant population genetic structure and genetic conservation. *Theoretical and Applied Genetics* 52, 145–157.
- Brown, S.M., Szewc-McFadden, A.K. and Kresovich, S. (1996) Development and application of simple sequence repeat (SSR) loci for plant genome analysis. In: Jauhar, P.P. (ed.) *Methods of Genome Analysis in Plants*. CRC Press, Boca Raton, FL, pp. 147–159.
- Caetano-Anollés, G. (1993) Amplifying DNA with arbitrary oligonucleotide primers. *PCR Methods and Applications* 3, 85–94.

- Caetano-Anollés, G. and Gresshoff, P.M. (1994) DNA amplification fingerprinting using arbitrary mini-hairpin oligonucleotide primers. *Biotechnology* 12, 619–623.
- Caetano-Anollés, G., Bassam, B.J. and Gresshoff, P.M. (1991) DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. *Biotechnology* 9, 553–557.
- Caetano-Anollés, G., Bassam, B.J. and Gresshoff, P.M. (1993) Enhanced detection of polymorphic DNA by multiple arbitrary amplicon profiling of endonuclease-digested DNA: identification of markers tightly linked to the supernodulation locus in soybean. *Molecular and General Genetics* 241, 57–64.
- Cavallini, A. (1993) Cytophotometric and biochemical analyses of DNA variations in the genus *Aloe* L. *International Journal of Plant Sciences* 154, 169–174.
- Chunwongse, J., Martin, G.B. and Tanksley, S.D. (1993) Pre-germination genotypic screening using PCR amplification of half-seeds. *Theoretical and Applied Genetics* 86, 694–698.
- Clegg, M.T. and Zurawski, G. (1992) Chloroplast DNA and the study of plant phylogeny: present status and future prospects. In: Soltis, P.S., Soltis, D.E. and Doyle, J.J. (eds) *Molecular Systematics of Plants*. Chapman & Hall, New York, pp. 1–13.
- Collins, F. and Galas, D. (1993) A new five-year plan for the United States Human Genome Project. *Science* 262, 43–46.
- Condit, R. and Hubbell, S.P. (1991) Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. *Genome* 34, 66–71.
- de Carvalho, R.C. and Saraiva, S.L. (1993) A new heterochromatin banding pattern revealed by modified HKG banding technique in maize chromosomes. *Heredity* 70, 515–519.
- Decorte, R., Cuppens, H., Marynen, P. and Cassiman, J.-J. (1990) Rapid detection of hypervariable regions by the polymerase chain reaction technique. *DNA and Cell Biology* 9, 461–469.
- Demesure, B., Sodzi, N. and Petit, R.J. (1995) A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology* 4, 129–131.
- Dickson, E.E., Arumuganathan, K., Kresovich, S. and Doyle, J.J. (1992) Nuclear DNA content variation within the Rosaceae. *American Journal of Botany* 79, 1081–1086.
- Doležel, J. and Lucretti, S. (1995) High-resolution flow karyotyping and chromosome sorting in *Vicia faba* lines with standard and reconstructed karyotypes. *Theoretical and Applied Genetics* 90, 797–802.
- Donovan, G.R., Skerritt, J.H. and Castle, S.L. (1989) Monoclonal antibodies used to characterize complementary DNA clones expressing specific endosperm proteins. *Journal of Cereal Science* 9, 97–112.
- Dowling, T.E., Moritz, C. and Palmer, J.D. (1990) Nucleic acids II: restriction site analysis. In: Hillis, D.M. and Moritz, C. (eds) *Molecular Systematics*. Sinauer Associates, Sunderland, MA, pp. 250–317.
- Downie, S.R. and Palmer, J.D. (1992) Use of chloroplast DNA rearrangements in reconstructing plant phylogeny. In: Soltis, P.S., Soltis, D.E. and Doyle, J.J. (eds) *Molecular Systematics of Plants*. Chapman & Hall, New York, pp. 14–35.
- Doyle, J.J. (1991) Gene trees and species trees: molecular systematics as one-character taxonomy. *Systematic Botany* 17, 144–163.

- Drew, G.S. (1994) The scientific method revisited. *Conservation Biology* 8, 596–597.
- Dweikat, I., MacKenzie, S., Levy, M. and Ohm, H. (1993) Pedigree assessment using RAPD-DGGE in cereal crop species. *Theoretical and Applied Genetics* 85, 497–505.
- Egel, D.S., Graham, J.H. and Stall, R.E. (1991) Genomic relatedness of *Xanthomonas campestris* strains causing diseases of citrus. *Applied and Environmental Microbiology* 57, 2724–2730.
- Ellsworth, D.L., Rittenhouse, K.D. and Honeycutt, R.L. (1993) Artifactual variation in randomly amplified polymorphic DNA banding patterns. *BioTechniques* 14, 214–217.
- Erllich, H.A., Gelfand, D. and Sninsky, J.J. (1991) Recent advances in the polymerase chain reaction. *Science* 252, 1643–1651.
- Esen, A., Mohammad, K., Schurig, G.G. and Aycock, H.S. (1989) Monoclonal antibodies to zein discriminate certain maize inbreds and genotypes. *Journal of Heredity* 80, 17–23.
- Friebe, B., Mukai, Y. and Gill, B.S. (1992) C-banding polymorphisms in several accessions of *Triticum tauschii* (*Aegilops squarrosa*). *Genome* 35, 192–199.
- Fukui, K., Ohmido, N. and Khush, G.S. (1994) Variability in rDNA loci in the genus *Oryza* detected through fluorescence *in situ* hybridization. *Theoretical and Applied Genetics* 87, 893–899.
- Gottlieb, L.D. (1981) Electrophoretic evidence and plant populations. *Progress in Phytochemistry* 7, 2–46.
- Griffin, H.G. and Griffin, A.M. (1993) DNA sequencing – recent innovations and future trends. *Applied Biochemistry and Biotechnology* 38, 147–159.
- Gupta, M., Chyi, Y.-S., Romero-Severson, J. and Owen, J.L. (1994) Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple-sequence repeats. *Theoretical and Applied Genetics* 89, 998–1006.
- Halldén, C., Bryngelsson, T., Säll, T. and Gustafsson, M. (1987) Distribution and evolution of a tandemly repeated DNA sequence in the family Brassicaceae. *Journal of Molecular Evolution* 25, 318–323.
- Halward, T.M., Stalker, H.T., Larue, E.A. and Kochert, G. (1991) Genetic variation detectable with molecular markers among unadapted germplasm resources of cultivated peanut and related wild species. *Genome* 34, 1013–1020.
- Hamby, R.K. and Zimmer, E.A. (1992) Ribosomal RNA as a phylogenetic tool in plant systematics. In: Soltis, P.S., Soltis, D.E. and Doyle, J.J. (eds) *Molecular Systematics of Plants*. Chapman & Hall, New York, pp. 50–91.
- Hammatt, N., Blackhall, N.W. and Davey, M.R. (1991) Variation in the DNA content of *Glycine* spp. *Journal of Experimental Botany* 42, 659–666.
- Hamrick, J.L. and Godt, M.J.W. (1989) Allozyme diversity in plants. In: Brown, A.H.D., Clegg, M.T., Kahler, A.L. and Weir, B.S. (eds) *Plant Population Genetics, Breeding and Genetic Resources*. Sinauer Associates, Sunderland, MA, pp. 43–63.
- Harding, R.M., Boyce, A.J. and Clegg, J.B. (1992) The evolution of tandemly repetitive DNA: recombination rules. *Genetics* 132, 847–859.
- Harrison, G.E. and Heslop-Harrison, J.S. (1995) Centromeric repetitive DNA sequences in the genus *Brassica*. *Theoretical and Applied Genetics* 90, 157–165.
- Heslop-Harrison, J.S. (1995) Flow cytometry and genome analysis. *Probe Newsletter* 5, 14–17.

- Hillis, D.M. and Moritz, C. (eds) (1990) *Molecular Systematics*. Sinauer Associates, Sunderland, MA, 588 pp.
- Hillis, D.M., Larson, A., Davis, S.K. and Zimmer, E.A. (1990) Nucleic acids III: sequencing. In: Hillis, D.M. and Moritz, C. (eds) *Molecular Systematics*. Sinauer Associates, Sunderland, MA, pp. 318–370.
- Hilu, K.W. and Esen, A. (1993) Prolamin and immunological studies in the Poaceae. III. Subfamily Chloridoideae. *American Journal of Botany* 80, 104–113.
- Holdsworth, M.J., Schuch, W. and Grierson, D. (1988) Organization and expression of a wound-ripening-related small multigene family from tomato. *Plant Molecular Biology* 11, 81–88.
- Hunter, R.L. and Markert, C.L. (1957) Histochemical demonstration of enzymes separated by zone electrophoresis in starch gels. *Science* 125, 1294–1295.
- Ingles, J., Pearson, G.C. and Sinclair, J. (1973) Species distribution and properties of nuclear satellite DNA in higher plants. *Nature* 242, 193–197.
- Innis, M.A. and Gelfand, D.H. (1990) Optimization of PCRs. In: Innis, M.A., Gelfand, D.H., Sninsky, J.J. and White, T.J. (eds) *PCR Protocols: A Guide to Methods and Applications*. Academic Press, San Diego, CA, pp. 3–12.
- Jarvis, P., Lister, C., Szabo, V. and Dean, C. (1994) Integration of CAPS markers into the RFLP map generated using recombinant inbred lines of *Arabidopsis thaliana*. *Plant Molecular Biology* 24, 685–687.
- Jeffreys, A.J., Wilson, V. and Thein, S.L. (1985) Hypervariable ‘minisatellite’ regions in human DNA. *Nature* 314, 67–73.
- Jellen, E.N., Gill, B.S. and Cox, T.S. (1994) Genomic *in situ* hybridization differentiates between A/D-genome and C-genome chromatin and detects intergenomic translocations in polyploid oat species (genus *Avena*). *Genome* 37, 613–618.
- Khuong, N.T. and Schubert, I. (1990) Differential staining of mitotic chromosomes of *Zea mays* L. by actinomycin D, DAPI and restrictase Giemsa technique. *Biologisches Zentralblatt* 109, 271–277.
- Koebner, R.M.D. (1995) Predigestion of DNA template improves the level of polymorphism of random amplified polymorphic DNAs in wheat. *Genetic Analysis: Biomolecular Engineering* 12, 63–67.
- Kranz, H.D., Miks, D., Siegler, M.L., Capesius, I., Sensen, C.W. and Huss, V.A.R. (1995) The origin of land plants – phylogenetic relationships among charophytes, bryophytes, and vascular plants inferred from complete small subunit ribosomal RNA gene sequences. *Journal of Molecular Evolution* 41, 74–84.
- Kresovich, S. and McFerson, J.R. (1992) Assessment and management of plant genetic diversity: considerations of intra- and interspecific variation. *Field Crops Research* 29, 185–204.
- Kresovich, S., Lamboy, W.F., Li, R., Ren, J., Szewc-McFadden, A.K. and Bliet, S.M. (1994) Application of molecular methods and statistical analyses for discrimination of accessions and clones of vetiver grass. *Crop Science* 34, 805–809.
- Kresovich, S., Szewc-McFadden, A.K., Bliet, S.M. and McFerson, J.R. (1995) Abundance and characterization of simple-sequence repeats (SSRs) isolated from a size-fractionated genomic library of *Brassica napus* L. (rapeseed). *Theoretical and Applied Genetics* 91, 206–211.
- Kumar, L.S., Gupta, V.S. and Ranjekar, P.K. (1990) Identification and partial characterization of two species-specific repeat families in great millet *Sorghum vulgare* (Poaceae) genome. *Plant Systematics and Evolution* 171, 249–257.

- Kumar, L.S., Sivaraman, L. and Ranjekar, P.K. (1992) Genome turnover in great millet and related millets. *Plant Systematics and Evolution* 179, 155–165.
- Kuramoto, T., Mashimo, T. and Serikawa, T. (1995) 14 anonymous DNA markers of laboratory rats identified with arbitrarily primed polymerase chain-reaction (AP-PCR). *Experimental Animals* 44, 119–125.
- Kwok, P.Y., Carlson, C., Yager, T.D., Ankener, W. and Nickerson, D.A. (1994) Comparative analysis of human DNA variations by fluorescence-based sequencing of PCR products. *Genomics* 23, 138–144.
- Landegren, U. (1993) Ligation-based DNA diagnostics. *Bioessays* 15, 761–765.
- Lander, E.S. (1989) DNA fingerprinting on trial. *Nature* 339, 501–505.
- Ledbetter, S.A., Nelson, D.L., Warren, S.T. and Ledbetter, D.H. (1990) Rapid isolation of DNA probes within specific chromosome regions by interspersed repetitive sequence polymerase chain reaction. *Genomics* 6, 475–481.
- Leone, C.A. (1964) *Taxonomic Biochemistry and Serology*. Ronald Press, New York, 728 pp.
- Li, H.H., Gyllensten, U.B., Cui, X.F., Saiki, R.K., Erlich, H.A. and Arnheim, N. (1988) Amplification and analysis of DNA sequences in single human sperm and diploid cells. *Nature* 335, 414–417.
- Lichter, P., Ledbetter, S.A., Ledbetter, D.H. and Ward, D.C. (1990) Fluorescence *in situ* hybridization with *Alu* and L1 polymerase chain reaction probes for rapid characterization of human chromosomes in hybrid cell lines. *Proceedings of the National Academy of Sciences of the USA* 87, 6634–6638.
- Lipshutz, R.J., Morris, D., Chee, M., Hubbell, E., Kozal, M.J., Shah, N., Shen, N., Yang, R. and Fodor, S.P.A. (1995) Using oligonucleotide probe arrays to access genetic diversity. *Biotechniques* 19, 442–447.
- Litt, M. and Luty, J.A. (1989) A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American Journal of Human Genetics* 44, 397–401.
- Livache, T., Fougue, B. and Teoule, R. (1994) Detection of HIV-1 DNA in biological samples by an homogeneous assay: fluorescence measurement of double-stranded RNA synthesized from amplified DNA. *Analytical Biochemistry* 217, 248–254.
- Long, E.H. and Dawid, I.B. (1980) Repeated genes in eukaryotes. *Annual Review of Biochemistry* 49, 727–764.
- Lysov, Y.P., Florent'ev, V.L., Khorlin, A.A., Khrapko, K.R., Shik, V.V. and Mirzabekov, A.D. (1988) Determining the nucleotide sequence of DNA by hybridization with oligonucleotides: a new method. *Proceedings of the USSR Academy of Science* 303, 1508–1511.
- McCauley, D.E. (1995) The use of chloroplast DNA polymorphism in studies of gene flow in plants. *Trends in Ecology and Evolution* 10, 198–202.
- McClelland, M., Arensdorf, H., Cheng, R. and Welsh, J. (1994) Arbitrarily primed PCR fingerprints resolved on SSCP gels. *Nucleic Acids Research* 22, 1770–1771.
- McMurphy, L.M. and Rayburn, A.L. (1991) Genome size variation in maize populations selected for cold tolerance. *Plant Breeding* 106, 190–195.
- Maluszynska, J. and Heslop-Harrison, J.S. (1993) Physical mapping of rDNA loci in *Brassica* species. *Genome* 36, 774–781.
- Mansfield, E.S., Worley, J.M., McKenzie, S.E., Surrey, S., Rappaport, E. and Fortina, P. (1995) Nucleic acid detection using non-radioactive labelling

- methods. *Molecular and Cellular Probes* 9, 145–156.
- Masojc, P., Zawistowski, J., Howes, N.K., Aung, T. and Gale, M.D. (1993) Polymorphism and chromosomal location of endogenous alpha amylase inhibitor genes in common wheat. *Theoretical and Applied Genetics* 85, 1043–1048.
- Maxam, A.M. and Gilbert, W. (1977) A new method for sequencing DNA. *Proceedings of the National Academy of Sciences of the USA* 74, 560–564.
- Maxson, L.R. and Maxson, R.D. (1990) Proteins II: immunological techniques. In: Hillis, D.M. and Moritz, C. (eds) *Molecular Systematics*. Sinauer Associates, Sunderland, MA, pp. 127–155.
- Meyer, W., Mitchell, T.F., Freedman, E.Z. and Vilgalys, R. (1993) Hybridization probes for conventional DNA fingerprinting used as single primers in the polymerase chain reaction to distinguish strains of *Cryptococcus neoformans*. *Journal of Clinical Microbiology* 31, 2274–2280.
- Micheli, M.R., Bova, R., Pascale, E. and D'Ambrosio, E. (1994) Reproducible DNA fingerprinting with the random amplified polymorphic DNA (RAPD) method. *Nucleic Acids Research* 22, 1921–1922.
- Miller, J.R. and Archibald, A.L. (1993) 5' and 3' SINE-PCR allows genotyping of pig families without cloning and sequencing steps. *Mammalian Genome* 4, 243–246.
- Milligan, B.G., Leebens-Mack, J. and Strand, A.E. (1994) Conservation genetics: beyond the maintenance of marker diversity. *Molecular Ecology* 3, 423–435.
- Mullis, K.B. and Faloona, F.A. (1987) Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Methods in Enzymology* 155, 335–350.
- Murphy, R.W., Sites, J.W., Jr, Buth, D.G. and Haufler, C.H. (1990) Proteins I: isozyme electrophoresis. In: Hillis, D.M. and Moritz, C. (eds) *Molecular Systematics*. Sinauer Associates, Sunderland, MA, pp. 45–126.
- Murray, B.G. (1994) Heterochromatin and silver banding of rye (*Secale cereale*, Gramineae) chromosomes. *Plant Systematics and Evolution* 193, 243–248.
- Myers, R.M., Lumelsky, N., Lerman, L.S. and Maniatis, T. (1985) Detection of single base substitutions in total genomic DNA. *Nature* 313, 495–498.
- Newman, T., de Bruijn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, E. and Somerville, C. (1994) Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous *Arabidopsis* cDNA clones. *Plant Physiology* 106, 1241–1255.
- Nickerson, D.A., Kaiser, R., Lappin, S., Stewart, J., Hood, L. and Landegren, U. (1990) Automated DNA diagnostics using an ELISA-based oligonucleotide ligation assay. *Proceedings of the National Academy of Sciences of the USA* 87, 8923–8927.
- Nickerson, D.A., Whitehurst, C., Boysen, C., Charmley, P., Kaiser, R. and Hood, L. (1992) Identification of clusters of biallelic polymorphic sequence-tagged sites (pSTSs) that generate highly informative and automatable markers for genetic linkage mapping. *Genomics* 12, 377–387.
- Nybom, H. (1993) Applications of DNA fingerprinting in plant population studies. In: Pena, S.D.J., Chakraborty, R., Epplen, J.T. and Jeffreys, A.J. (eds) *DNA Fingerprinting: State of the Science*. Birkhäuser Verlag, Basel, pp. 293–309.
- Ochman, H., Gerber, A.S. and Hartl, D.L. (1988) Genetic applications of an inverse polymerase chain reaction. *Genetics* 120, 621–623.
- Olson, M., Hood, L., Cantor, C. and Botstein, D. (1989) A common language for

- physical mapping of the human genome. *Science* 245, 1434–1435.
- O'Neill, C.M. and Mathias, R.J. (1995) A rapid and simple method of producing high-resolution zymograms from *Brassica* and related species on isoelectric-focussing gels. *Plant Breeding* 114, 179–181.
- Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K. and Sekiya, T. (1989) Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proceedings of the National Academy of Sciences of the USA* 86, 2766–2770.
- Pääbo, S. (1989) Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. *Proceedings of the National Academy of Sciences of the USA* 86, 1939–1943.
- Palmer, J.D. (1992) Mitochondrial DNA in plant systematics: applications and limitations. In: Soltis, P.S., Soltis, D.E. and Doyle, J.J. (eds) *Molecular Systematics of Plants*. Chapman & Hall, New York, pp. 36–49.
- Paran, I. and Michelmore, R.W. (1993) Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. *Theoretical and Applied Genetics* 85, 985–993.
- Park, Y.S., Kwak, J.M., Kwon, O.-Y., Kim, Y.S., Lee, D.S., Cho, M.J., Lee, H.H. and Nam, H.G. (1993) Generation of expressed sequence tags of random root cDNA clones of *Brassica napus* by single-run partial sequencing. *Plant Physiology* 103, 359–370.
- Pease, A.C., Solas, D., Sullivan, E.J., Cronin, M.T., Holmes, C.P. and Fodor, S.P.A. (1994) Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proceedings of the National Academy of Sciences of the USA* 91, 5022–5026.
- Powell, W., Morgante, M., Andre, C., McNicol, J.W., Machray, G.C., Doyle, J.J., Tingey, S.V. and Rafalski, J.A. (1995) Hypervariable microsatellites provide a general source of polymorphic DNA markers for the chloroplast genome. *Current Biology* 5, 1023–1029.
- Ragot, M. and Hoisington, D.A. (1993) Molecular markers for plant breeding: comparisons of RFLP and RAPD genotyping costs. *Theoretical and Applied Genetics* 86, 975–984.
- Reed, P.W., Davies, J.L., Copeman, J.B., Bennett, S.T., Palmer, S.M., Pritchard, L.E., Gough, S.C.L., Kawaguchi, Y., Cordell, H.J., Balfour, K.M., Jenkins, S.C., Powell, E.E., Vignal, A. and Todd, J.A. (1994) Chromosome-specific microsatellite sets for fluorescence-based, semi-automated genome mapping. *Nature Genetics* 7, 390–395.
- Roberts, R.J. (1984) Restriction and modification enzymes and their recognition sequences. *Nucleic Acids Research* 12, r167–r204.
- Rogowsky, P.M., Shepard, K.W. and Langridge, P. (1992) Polymerase chain reaction based mapping of rye involving repeated DNA sequences. *Genome* 35, 621–626.
- Rogstad, S.H. (1993) Surveying plant genomes for variable number of tandem repeat loci. *Methods in Enzymology* 224, 278–294.
- Rogstad, S.H., Patton, J.C., II and Schaal, B.A. (1988) M13 repeat probe detects DNA minisatellite-like sequences in gymnosperms and angiosperms. *Proceedings of the National Academy of Sciences of the USA* 85, 9176–9178.
- Saghai Maroof, M.A., Biyashev, R.M., Yang, G.P., Zhang, Q. and Allard, R.W. (1994) Extraordinarily polymorphic microsatellite DNA in barley: species diversity,

- chromosomal locations, and population dynamics. *Proceedings of the National Academy of Sciences of the USA* 91, 5466–5470.
- Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A. and Arnheim, N. (1985) Enzymatic amplification of β -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230, 1350–1354.
- Saiki, R.K., Bugawan, T.K., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1986) Analysis of enzymatically amplified β -globin and HLA-DQ α DNA with allele-specific oligonucleotide probes. *Nature* 324, 163–166.
- Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239, 487–491.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn, 3 vols. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the USA* 74, 5463–5467.
- Sarich, V.M. and Wilson, A.C. (1966) Quantitative immunochemistry and the evolution of primate albumins: micro-complement fixation. *Science* 154, 1563–1566.
- Schlötterer, C., Amos, B. and Tautz, D. (1991) Conservation of polymorphic simple sequence loci in cetacean species. *Nature* 354, 63–65.
- Schwarzacher, T. (1994) Mapping in plants: progress and prospects. *Current Opinion in Genetics and Development* 4, 868–874.
- Schwengel, D.A., Jedlicka, A.E., Nanthakumar, E.J., Weber, J.L. and Levitt, R.C. (1994) Comparison of fluorescence-based semi-automated genotyping of multiple microsatellite loci with autoradiographic techniques. *Genomics* 22, 46–54.
- Sessions, S.K. (1990) Chromosomes: molecular cytogenetics. In: Hillis, D.M. and Moritz, C. (eds) *Molecular Systematics*. Sinauer Associates, Sunderland, MA, pp. 156–203.
- Sharma, P.C., Winter, P., Bünger, T., Hüttel, B., Weigand, F., Weising, K. and Kahl, G. (1995) Abundance and polymorphism of di-, tri and tetra-nucleotide tandem repeats in chickpea (*Cicer arietinum* L.). *Theoretical and Applied Genetics* 90, 90–96.
- Sibson, D.R., Hughes, S.G., Bryant, J.A. and Fitchett, P.N. (1991) Sequence organization of simple, highly repetitive DNA elements in *Brassica* species. *Journal of Experimental Botany* 42, 243–249.
- Smith, J.S.C. and Smith, O.S. (1992) Fingerprinting crop varieties. *Advances in Agronomy* 47, 85–140.
- Soltis, P.S. and Soltis, D.E. (1995) Alternative genes for phylogenetic reconstruction in plants: introduction. *Annals of the Missouri Botanical Garden* 82, 147–148.
- Soltis, P.S., Soltis, D.E. and Smiley, C.J. (1992) An *rbcL* sequence from a Miocene *Taxodium* (bald cypress). *Proceedings of the National Academy of Sciences of the USA* 89, 449–451.
- Sorrells, M.E. (1992) Development and application of RFLPs in polyploids. *Crop Science* 32, 1086–1091.
- Southern, E.M. (1975) Detection of specific sequences among DNA fragments

- separated by gel electrophoresis. *Journal of Molecular Biology* 98, 503–517.
- Stegink, S.J. and Vaughn, K.C. (1990) Immunotaxonomy of nodule-specific proteins. *Cytobios* 61, 7–20.
- Sukias, J.P.S. and Murray, B.G. (1990) Cytological examination of three cultivars of garlic (*Allium sativum* L.) by C-banding and flow cytometry. *New Zealand Journal of Crop and Horticultural Science* 18, 17–22.
- Taberlet, P., Gielly, L., Patou, G. and Bouvet, J. (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* 17, 1105–1109.
- Tautz, D. (1993) Notes on the definition and nomenclature of tandemly repetitive DNA sequences. In: Pena, S.D.J., Chakraborty, R., Epplen, J.T. and Jeffreys, A.J. (eds) *DNA Fingerprinting: State of the Science*. Birkhäuser Verlag, Basel, pp. 21–28.
- Tautz, D., Trick, M. and Dover, G.A. (1986) Cryptic simplicity in DNA is a major source of genetic variation. *Nature* 322, 652–656.
- Tingey, S.V. and del Tufo, J.P. (1993) Genetic analysis with random amplified polymorphic DNA markers. *Plant Physiology* 101, 349–352.
- Tong, X.C. and Smith, L.M. (1993) Solid phase purification in automated DNA sequencing. *DNA Sequence* 4, 151–162.
- Tourmente, S., Deragon, J.M., Lafleurriel, J., Tutois, S., Pélissier, T., Cuvillier, C., Espagnol, M.C. and Picard, G. (1994) Characterization of minisatellites in *Arabidopsis thaliana* with sequence similarity to the human minisatellite core sequence. *Nucleic Acids Research* 22, 3317–3321.
- Vega, J.M., Abbo, S., Feldman, M. and Levy, A.A. (1994) Chromosome painting in plants: *in situ* hybridization with a DNA probe from a specific microdissected chromosome arm of common wheat. *Proceedings of the National Academy of Sciences of the USA* 91, 12041–12045.
- Wang, G.-L., Wing, R.A. and Paterson, A.H. (1993) PCR amplification from single seeds, facilitating DNA marker-assisted breeding. *Nucleic Acids Research* 21, 2527.
- Wang, Z., Weber, J.L., Zhong, G. and Tanksley, S.D. (1994) Survey of plant short tandem DNA repeats. *Theoretical and Applied Genetics* 88, 1–6.
- Waterston, R., Martin, C., Craxton, M., Huynh, C., Coulson, A., Hillier, L., Durbin, R., Green, P., Shownkeen, R., Halloran, N., Metzstein, M., Hawkins, T., Wilson, R., Berks, M., Du, Z., Thomas, K., Thierrymieg, J. and Sulston, J. (1992) A survey of expressed genes in *Caenorhabditis elegans*. *Nature Genetics* 1, 114–123.
- Weber, J.L. and May, P.E. (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics* 44, 388–396.
- Weising, K., Atkinson, R.G. and Gardner, R.C. (1995) Genomic fingerprinting by microsatellite-primed PCR: a critical evaluation. *PCR Methods and Applications* 4, 249–255.
- Weissinger, A.K. (1990) Technologies for germ plasm preservation *ex situ*. In: Orians, G.H., Brown, G.M., Jr., Kunin, W.E. and Swierzbinski, J.E. (eds) *The Preservation and Valuation of Biological Resources: Workshop, Seattle WA, USA, June 12–16, 1985*. University of Washington Press, Seattle, WA, pp. 3–38.
- Welsh, J. and McClelland, M. (1990) Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research* 18, 7213–7218.

- Werman, S.D., Springer, M.S. and Britten, R.J. (1990) Nucleic acids I: DNA–DNA hybridization. In: Hillis, D.M. and Moritz, C. (eds) *Molecular Systematics*. Sinauer Associates, Sunderland, MA, pp. 204–249.
- Wetmur, J.G. and Davidson, N. (1968) Kinetics of renaturation of DNA. *Journal of Molecular Biology* 31, 349–370.
- Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. and Tingey, S.V. (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* 18, 6531–6535.
- Wu, K.S., Jones, R., Danneberger, L. and Scolnik, P.A. (1994) Detection of microsatellite polymorphisms without cloning. *Nucleic Acids Research* 22, 3257–3258.
- Yang, M.M. and Youvan, D.C. (1989) A prospectus for multispectral multiplex DNA sequencing. *Bio/Technology* 7, 576–580.
- Zabeau, M. and Vos, P. (1993) Selective restriction fragment amplification: a general method for DNA fingerprinting. EPO Patent No. 0534858A1.
- Zawistowski, J. and Howes, N.K. (1990) Production of monoclonal antibodies against high Mr avenin polypeptides. *Journal of Cereal Science* 12, 235–244.
- Zeller, F.J., Cermeno, M.-C. and Miller, T.E. (1991) Cytological analysis on the distribution and origin of the alien chromosome pair conferring blue aleurone color in several European common wheat *Triticum aestivum* L. strains. *Theoretical and Applied Genetics* 81, 551–558.

Genetic Diversity – Population Structure and Conservation

M.D. Hayward and N.R. Sackville Hamilton

Institute of Grassland and Environmental Research, Plas Gogerddan, Aberystwyth, Dyfed SY23 3EB, UK.

3.1. Introduction

The basic principle governing the conservation of any species is that the maintenance of biodiversity is an essential prerequisite for the continued production of new cultivars of current crops, for the development and exploitation of hitherto novel crops, and for the general 'well-being' of the planet. This chapter deals with the distribution of biodiversity, the mechanisms determining its distribution in space and time, and the consequences for its conservation *ex situ* and *in situ*.

In situ conservation is regarded as conservation in any habitat where the germplasm normally occurs: not just in natural habitats, but also in farms, gardens and other man-made habitats of relevant germplasm. *Ex situ* conservation refers to any collection maintained outside the normal habitat of the germplasm: not just seed collections and *in vitro* tissue culture, but also living collections in botanic gardens and collections of species with recalcitrant seed. This classification is not universally accepted or applicable, especially in garden and horticultural contexts. It is, however, sufficient for the current purpose of reviewing the implications of evolutionary phenomena for efficient conservation.

The objective of *ex situ* conservation is to maintain a collection containing as many alleles as possible, and / or as diverse a range of gene combinations (i.e. genotypes) as possible, in a form that can be readily utilized for breeding and research. For efficient utilization, genetic variation within the collection must be appropriately characterized. For efficient conservation,

ideally the collection should be as small as possible commensurate with conserving maximum diversity. In practice, in the face of rapid genetic erosion it is often necessary to collect germplasm before it can be ascertained that it contains genes or genotypes not already present in the collection: most collections are therefore considerably larger than strictly necessary for the diversity they contain.

The objective of *in situ* conservation, at least for agricultural purposes, is to conserve the maximum possible number of alleles and / or maximum possible diversity of genotypes whilst permitting continued evolution. This is of importance in generating new genes or genotypes, particularly: (i) in response to changing environments, e.g. genes for resistance to newly evolved strains of pathogens; and (ii) by continued selection of landraces by farmers or gardeners (at least where law still permits). Additional benefits include conservation of much more biodiversity – entire ecosystems – than just the targeted crop germplasm. Against this is the disadvantage that the germplasm cannot be efficiently utilized because characterized genotypes cannot be readily tracked.

Evolutionary mechanisms determine the distribution of genetic diversity in time and space. The first part of this chapter deals with these mechanisms and the resulting patterns of diversity. We will then consider the implications for efficient conservation. There are three groups of implications. First, efficient collection of diverse germplasm for *ex situ* conservation depends on having good knowledge of the spatial distribution of genetic diversity. Inevitably it is not possible to know the exact location of every genotype. Instead, a good understanding of the factors that control the distribution of genetic diversity is necessary to devise a collecting strategy that maximizes the diversity sampled.

Second, in addition to efficient construction of an *ex situ* collection, efficient maintenance of the same also depends on good understanding of the factors that control the distribution of genetic diversity – in this case to control the genetic shifts that occur whenever a population is sampled, subsampled or regenerated.

Third, efficient *in situ* conservation depends on good knowledge of the distribution of genetic diversity in space and time, and of the factors that control its distribution. In particular, an *in situ* conservation area (or network of such areas) must have a size, heterogeneity and structure that maximizes genetic variance maintained by evolution. The efficient achievement of both the primary objectives requires a knowledge of the genetic nature of variability, population structure, the distribution of diversity and the factors that control them.

3.2. Population Genetic Considerations

The genetic structure of a population determines its capacity for response to selection, both natural and artificial, and as such is of primary consideration in the formulation of strategies for the collection and conservation of biodiversity. The structure of the population is controlled by several factors such as its life form, breeding system and effective population size. These factors, which often reflect past selection pressures, all influence the nature and maintenance of genetic variation both within and between populations and in some cases may themselves be subject to genetic determination. In the conservation of biodiversity it is the underlying genetic control of variability that is of major importance in determining appropriate strategies.

3.2.1. Genetic variation

The phenotype of an organism is controlled by a multitude of genes which act both individually and in concert upon the various stages of development and are influenced to varying degrees by the environment. Their action leads, in the majority of cases, to a quantitative expression of growth forms which are continuously distributed in nature. The genetic control of these quantitative traits is by sets of genes (polygenes or quantitative trait loci (QTL)) each of small effect; although these may be difficult to identify individually, they are inherited in a Mendelian manner and show all the properties of major genes, i.e. linkage, dominance, epistasis and the effects of ploidy. The study of such traits has in the past required the application of biometrical procedures utilizing means and variances, but developments in molecular techniques for genome analysis and genetic mapping offer the prospect of more precise identification of single quantitative trait loci.

The behaviour of genes determining quantitative traits in a population is the same as that of major genes. If random mating is the mode of reproduction, at a single locus level the individual genotypes are to be found in the Hardy–Weinberg proportions, $p^2:2pq:q^2$, where p and q represent the diallelic frequencies. When extended over many loci it can be seen that the extreme homozygous classes and thus phenotypes are rare in the population. These genotypes, which represent free variability and are directly fixable by selection, have the capacity by hybridization and segregation to create all intermediate genotypic classes (see Mather, 1973). In so doing, the majority of individuals produced will be of differing homozygous/heterozygous genotypic combinations and, as such, will give rise to intermediate phenotypes. These differing genotypic classes again have the potential, by hybridization and segregation, to release variation. Here,

however, this hidden variability is in two states, the homozygotic and heterozygotic (Mather, 1973). As the number of genes controlling a trait increases, the proportion of variability exposed to the rigours of selection in the homozygous state will decrease. Recent developments in QTL analysis have shown that for many traits of agronomic importance the number of such genes is often quite high. In tomato, for example, at least six have been found to control total yield and 12 for soluble-solids content (Tanskley *et al.*, 1996).

The mode of gene action will also influence the state and proportion of exposed variability. The effects of dominance will be two-fold depending on the direction of dominance at the individual loci concerned. Firstly, if all dominant alleles are acting in the same direction the effect will be to reduce the number of phenotypic classes observed as the heterozygous classes will be indistinguishable from the dominant homozygotes. As a consequence, selection will be more difficult as potential variability will still be present in the heterozygotes and can only be revealed by progeny testing. The distribution of individuals will be very much skewed in the direction of the dominant expression. Secondly, if dominance is ambidirectional, the effect will be to increase the proportion of intermediate phenotypes in the distribution and with it the release of variability and the potential for response to selection. Nonallelic interaction will to some extent reinforce the effects of dominance in leading to a reduction in class frequencies and the mean expression of a trait in a population.

The evolution of the genetic architecture of a trait is governed by the components of the genetic system (Darlington, 1958), namely the creation of new variation by mutation, recombination and the breeding system. These, when coupled with selection and/or genetic drift, are the major determinants controlling the manner in which variability is organized within a population.

3.2.2. Genetic change – genes and chromosomes

Mutational change at the DNA level of the genome is the basis of new variation and can take several forms such as base pair deletion, duplication or rearrangement. Its effect may be detectable at the molecular level, as for example in the changes that lead to differing forms of an enzyme (allozyme), which in most cases would be neutral in its action, or it may have a gross effect on the phenotype such as in flower form or colour. Most mutational changes are considered to be deleterious in that they generally disrupt the hitherto integrated structure of the gene. However, some may be advantageous, with their subsequent survival and spread in the population being dependent on such factors as selective advantage, population size and genetic drift. If the mutation is recessive, as in most

cases, its frequency in the homozygous state will initially be very rare in outbreeding species, hence the likelihood of exposure to the rigours of selection is very low. However, mutations that affect the breeding system can be at an immediate advantage. The occurrence of a mutant incompatibility allele in a single-locus gametophytic system, such as occurs in *Trifolium repens*, would be advantageous in that it is directly exposed in the haploid phase in the pollen grain and can be effective in promoting fertilization. In that it provides a further option for cross-pollination to occur its survival in the population / species is ensured. This may well account for the very high number of incompatibility alleles that can be found in species with gametophytic systems (Fearon *et al.*, 1994; O'Donnell and Lawrence, 1984).

Adaptive change may arise in a population through alterations at the chromosomal level. This may take the form of structural or numerical change such as gross deletions, inversions, interchanges, aneuploidy and polyploidy. The mechanisms and origin of these types of change are well documented (Darlington, 1956). It is their influence on the maintenance and release of variability and the opportunity they provide for new adaptive forms to arise which are of importance from a conservation aspect. Polyploidy, for example, which may arise by the direct doubling of a chromosome set or be coupled with wide hybridization, is well known as a mechanism for maintaining heterozygous combinations of genes (Stebbins, 1950; Breese *et al.*, 1981).

3.2.3. Recombination

The role of recombination in controlling the release and distribution of variation within a population is of fundamental concern in the development of strategies for the conservation of genetic resources. It has long been established that the mechanisms controlling chromosome pairing, and the frequency and position of crossing over in the genome are under genetic control (Rees, 1961). The evolution of the *Ph* (pairing) gene on chromosome 5B of wheat (Riley and Chapman, 1958) has led to the regular diploid pairing that takes place, and with it the stability and fertility of a diploid as opposed to the instability and sterility of an allopolyploid. Selection, irrespective of whether it be natural or artificial, can lead to marked differences in the rate of recombination between populations. In the outbreeding species *Lolium perenne* and *Festuca pratensis* bred cultivars have a higher chiasma frequency than their wild counterparts (Rees and Dale, 1974). This has arisen as a correlated response to selection for variability by the breeder. Similarly the presence of B chromosomes can influence chiasma frequency (Jones and Rees, 1982). The fine-scale collinearity of cereal and grass genomes should enable strategies to be developed for

the positional cloning of the gene(s) controlling chromosome pairing in wheat (Dunford *et al.*, 1995) and the forage grasses, opening up the prospect of genetically manipulating the processes of recombination at will and, with it, the range of variation that may be extracted from a population.

3.2.4. Breeding systems

The flow of variability within a species is dependent on its mode of reproduction. Sexual species, which may be either inbreeders or outbreeders, have the capacity for recombination and as a consequence variability may be exposed to selection. Asexual forms, which reproduce either by apomixis or vegetative means, maintain a uniform genotype, which may be well adapted to present selective forces, but lack the ability to respond to changing conditions. The breeding system is often under genetic control and may be associated with specific life forms. Inbreeding, which is predominantly found in annual life forms, often at the limits of a species distribution (Stebbins, 1950), is generally achieved by mechanisms that ensure self-pollination. Pollen may be shed within closed florets (cleistogamy), as in wheat and barley, or flowers may open and be receptive when anther dehiscence occurs (chasmogamy), as in tomato. Although these mechanisms are under precise genetic control, breakdown may occur, allowing outcrossing to take place. In barley, for example, Allard and Hansche (1965) showed that up to 1% outcrossing may be found under some environmental conditions. Novel recombinants will appear offering scope for further selection and evolutionary change.

Outbreeding is generally found in the more perennial species and is often promoted by one or more genetically controlled mechanisms. These may range from timing differences in anther dehiscence and the receptivity of the stigma through to precisely controlled incompatibility systems (de Nettancourt, 1993). The consequence of such processes is the maintenance of a high level of heterozygosity within the individual and variability both between individuals within the population and between populations. This aspect of population structure will be considered in more detail in a later section.

The apomictic mode of reproduction, which involves the production of seed by asexual means, is found in many genera, predominantly of the *Gramineae* and *Rosaceae*. It is generally associated with polyploidy and can be obligate or facultative. In those cases where sexual relatives are to be found, which allow genetic analysis, it has been shown to be under simple genetic control. For example, in *Panicum maximum* it appears to be under the control of a single dominant gene (Savidan, 1983) whilst in *Citrus* several genes are involved (Cameron and Soost, 1982). Apomixis has the

attribute of maintaining well-adapted combinations of genes together but has the disadvantage that there is no flow of variability and as such the species may well be at an evolutionary dead end (Stebbins, 1950).

Truly vegetative reproduction is rare but like apomixis can lead to the widespread distribution of a species. *Spartina anglica* is now to be found all around the shores of Great Britain having spread from its origins in Southampton Water by the continual breakup of its rhizomes. It is a reproductive mode that is often exploited by humans to maintain and distribute a crop species, as in the potato.

Each of these reproductive modes can be under genetic control and thus subjected to the forces of natural selection in the same manner as the genes responsible for other traits of adaptive significance. An insight into their effect on the state of variability and structure of populations can be obtained from the numerous studies of molecular markers in plant populations (Brown *et al.*, 1990; Soltis and Soltis, 1990; Weising *et al.*, 1995).

3.2.5. Selection, drift and gene flow

The differentiation of populations depends on the three processes of selection, drift and gene flow. The forces of selection, reflecting the environmental pressures acting upon the population, are instrumental in defining the underlying genetic structure. The differing modes which it may take, such as stabilizing, disruptive or directional, each have their own effect upon the subsequent gene action and organization of the variation (Mather, 1973). Random drift, particularly in small populations, can lead to arbitrary fixation of genes. The immigration of new genes from distinctive neighbouring subpopulations can increase the extent of both selective response and drift by introducing new alleles; or it can reduce it by repeatedly introducing genes adapted to a different microenvironment and by so doing retard microadaptation to local patches. In addition, life form and the persistence of seed banks may all influence the capacity for selection to lead to local adaptation.

3.3. Characterization of Population Structure

The state of variability within and between populations can be determined by application of both biometrical and molecular procedures. Whilst the former can describe a population in terms of means and variances and the underlying mode of gene action, if an appropriate biometrical design is applied as part of the characterization of a population (see Kearsey, 1993), it is only the procedures of molecular genetics which allow a measure of population structure at the level of the gene and genome to be gained.

Various molecular tools are available for the characterization of populations (see Pérez de la Vega, 1993). The most widely used to date is electrophoresis of isozymes, but direct DNA methods, including restriction fragment length polymorphisms (RFLPs), randomly amplified polymorphic DNA and amplified fragment length polymorphism (AFLP), are now being exploited. The application of these procedures allows several different parameters of variability within and between populations to be determined. These include the percentage of polymorphic loci, average and effective numbers of alleles per locus, heterogeneity and heterozygosity indices and the various measures based upon the F statistics of Wright (1965). Detailed considerations and the interrelationships of these are presented by Brown and Weir (1983) and Weising *et al.* (1995).

3.3.1. Characterization by isozymes

A comprehensive review of the manner in which isozyme analyses can provide an insight into the various factors influencing variability is given by Hamrick and Godt (1990). The effect of the breeding system, as revealed by the major parameters characterizing variability and population diversity, is shown in Table 3.1, for the three levels: between species; within populations; and between populations. The main features to emerge from this analysis are that outbreeding species have the highest level of variability. At the population level there was no significant difference between the two breeding system classes in total genetic diversity but the diversity within populations was greater for the outbreeders than the inbreeders.

3.3.2. DNA markers

More recently the application of DNA-based technologies, particularly 'fingerprinting' (see Weising *et al.*, 1995), has provided a wealth of information on the diversity of wild populations and the relatedness of

Table 3.1. Average measures of variability at the species and population level for differing reproductive modes. (Derived from Hamrick and Godt, 1990; see also Frankel *et al.*, 1995.)

Breeding system	Average gene diversity		Proportion of total variation among populations
	Between species	Within populations	
Inbreeders	0.124	0.074	0.51
Outbreeders	0.164	0.136	0.148

cultivars. For wild species these techniques have been applied to answering specific questions particularly in relation to breeding systems. A comparison of *Plantago* spp. using RFLP analysis revealed by the M13 probe showed that the inbreeding *P. major* possessed little variation within populations but marked differentiation between populations whilst the outbreeding *P. lanceolata* possessed high variability within populations but only moderate variability between populations. Similarly, in an analysis of cultivar differentiation in three species of Bermuda grass (*Cynodon*) Caetano-Anollés *et al.* (1995), using DNA amplification fingerprinting (DAF), were able to distinguish some closely related cultivars. On the basis of this technique's discriminatory power they recommend that it be used as a method for seed certification and registration purposes. These are just two examples of the many that have been carried out on a wide range of species exploiting the 'fingerprinting' capacity of the molecular methods.

In the majority of applications of these DNA-based technologies comparisons are based upon statistical analyses of the number of 'bands shared'. Such a procedure can be fraught with problems both on a technical and a genetic level. Some of the technical problems and other sources of error which can be encountered in the preparation, running and reading of gels have been considered by Weising *et al.* (1995), who emphasize the need for technical care and caution in scoring closely spaced bands. In addition there is also the question as to whether two DNA fragments which migrate to a common position on a gel are homologous. This may only be ascertained by extraction of the bands and comparative cross-hybridization. In RAPD analysis of the *Lolium/Festuca* complex of species Stammers *et al.* (1995) showed that four out of six amplification products were homologous when tested by Southern hybridization. Although this is a small sample and ranged across genetically diverse species, it does emphasize the need for caution in assuming homology.

The use of individual isozyme loci and similar genetic markers, as measures of variability, is limited in that they give no indication of the genomic associations that are of importance in the maintenance of co-adapted gene complexes. It is only by looking at combinations of markers that such information can be ascertained. In *Avena* spp. and *Hordeum* spp., for example, Allard and his coworkers (Allard, 1990; Allard *et al.*, 1993) have shown by comparisons involving up to 14 discrete loci that populations are made up of individuals containing differing multilocus assemblages of favourable epistatic combinations of alleles. These have arisen by rare outcrossing followed by inbreeding to near homozygosity. In *A. hirtula*, for example, a majority of Spanish populations were found to be polymorphic for different multilocus genotypes, which suggests that 'interactions at the interplant level may contribute to adaptive significance' (Allard *et al.*, 1993). In addition, in this species, polyploidy is present which, in its own right, can lead to a greater allelic diversity and potential

for differing multilocus associations. Here again the breeding system reinforces the stabilization of such associations by restricting the degree of recombination that takes place.

The current procedures using molecular markers for assessing population differentiation consider both expressed and nonexpressed parts of the genome. Given the ease with which genetic maps may now be constructed, it seems likely that, in future, measures of population differentiation will take into account genome organization and will target those regions of interest. Already this problem is being addressed as a means of determining identity by descent as part of the statutory procedures in determining 'essential derivation' of cultivars (Dillmann *et al.*, 1995).

Recent developments in QTL analysis of crop plants in defined populations, such as F_2 s and recombinant inbred lines, are now bringing together the power of genome analysis at the molecular level and biometrical procedures which will eventually allow a more detailed understanding of the genetic architecture of traits of both agronomic interest and of importance to evolutionary fitness.

These various studies of population structure provide an insight into the manner in which variability is distributed within a species and some of the factors controlling that pattern. From a conservation aspect it is now necessary to consider how these mechanisms interact to determine the spatial distribution of variability and their implications for collection and conservation.

3.4. Scaling Properties of Biodiversity

Scale-dependence is an inherent property of biodiversity. Any measure of variance necessarily depends on the spatial and temporal scale used for its assessment. Quantification of biodiversity must therefore include a description of how variance varies with scale. The general trends are well established, and may be depicted in a number of ways.

One way is to record diversity in a series of nested quadrats or regions, and plot diversity against area of the region. This is most commonly applied to species richness as the measure of diversity. On a log-log scale the relationship is often linear over a remarkably large range of scales (Begon *et al.*, 1996) – from a few square centimetres to thousands of square kilometres. The same general trend might be expected for the number of distinct genotypes of one species, though the evidence on this is more difficult to obtain.

Another approach is to plot the variance between pairs of sites against the distance between the sites. In this case the variance between sites tends to increase with distance for sites close together, but usually becomes independent of distance for widely separated sites. The same trend is observed both for measures of species diversity (Levin, 1989; Burrough, 1995) and

for measures of genetic diversity within species (Monestiez *et al.*, 1994). The geographic distance above which genetic distance becomes constant is the ‘genetic patch size’.

There are often significant deviations from these overall trends. There may be patterns leading to diversity significantly below the trend line at some scales and above it at others – patterns resulting from intrinsic spatial properties of population dynamics or from extrinsic environmental influences. There are also other factors controlling diversity that are unrelated to distance; these will appear as noise on a diversity / scale plot. Our objective is to understand what determines the form of this trend and deviations from the trend, and to apply that knowledge to efficient conservation.

3.5. Genetic Efficiency in Collecting for *ex situ* Conservation

Full technical guidelines on collecting are given by Guarino *et al.* (1995). This section concentrates just on evolutionary considerations, to facilitate adaptation of general guidelines to suit individual species, objectives and constraints.

3.5.1. Objectives

The objectives of an expedition to collect germplasm from a region may be to:

1. Acquire the maximum genetic diversity of targeted taxa within the region, within the constraints of limited available resources.
2. Acquire germplasm with the maximum novelty value with respect to a collection already held *ex situ*; i.e. the greatest number and diversity of genes and genotypes that have not previously been collected.
3. Combat genetic erosion.
4. Acquire genes or genotypes most likely to benefit a particular breeding or research objective.
5. Acquire germplasm for analysis of agroecogeographic patterns of the distribution of biodiversity.

This section deals mainly with the first of these objectives. Limiting resources vary from expedition to expedition, and may be time (time to travel to a site, time to collect overall site data, time to collect each seed or plant at a site, speed of returning live plants to base); space available in the collecting vehicle; or labour and facilities to process samples at base. Resource limitations influence optimal collecting strategy in a way that depends on population structure (Marshall and Brown, 1975; Brown and Marshall, 1995; Sackville Hamilton and Chorlton, 1995).

The principles outlined in this section also apply to objective 2 above, except that additional information is needed on the diversity and origins of the pre-existing collection.

Where the primary objective is to combat genetic erosion (objective 3 above), sampling strategy can and should still be designed to satisfy objective 1. However, painstaking planning to maximize the diversity collected may be counterproductive where the rate of erosion is so high that diversity is lost whilst planning is in progress. In these circumstances, speed of undertaking a collecting expedition is of overriding importance.

Objective 4, a breeder-driven collection to support a particular breeding objective, requires a totally different sampling strategy, to locate particular genes rather than maximize diversity of genes. Nevertheless, knowledge of evolutionary patterns can aid identification of sites most likely to contain the desired genes or genotypes.

Objective 5, collection for agroecogeographic analysis, requires yet another strategy, namely an appropriately randomized sampling procedure. This fact is not sufficiently recognized, as many published analyses are based on collections made for conservation or breeding purposes. Yet any sampling strategy that aims to maximize diversity or to target specific genes can generate incorrect and misleading estimates of components of variance. For example, suppose two collections are undertaken in two different regions, both with a sampling strategy to maximize the diversity sampled within each region. A comparison of both collections will then incorrectly suggest that there is less difference between them and more variation within each region than is really the case.

A specific example of an erroneous conclusion may be the widely accepted latitudinal cline in genetic diversity of *Trifolium repens* across Europe, according to which southern populations are believed to be highly diverse and northern ones uniform (Ellis Davies and Young, 1967). This is likely to be spurious, and at least partly a consequence of collections being specifically targeted at well-managed pastures in the north but at highly diverse habitats in the south. A northern collection targeted at diverse habitats contained as much diversity as southern populations (Sackville Hamilton, 1980).

Appropriate randomization for agroecogeographic analyses need not mean full randomization. Collecting for maximum diversity can be compatible and even beneficial to agroecogeographic analysis. Seeking to collect maximum genetic diversity by targeting maximum environmental diversity improves the sensitivity of analysis of the relationship between genetic diversity and environmental diversity. This of course requires collection of all relevant environmental data so that they can be incorporated as independent variables in statistical analysis.

3.5.2. Spatial scale of biodiversity: overall distribution of sites

The general scaling properties of biodiversity (Section 3.4) have two immediate implications. First, it is important to cover as large an area as possible. Second, adjacent sites should not be further apart than the genetic patch size, since increasing the geographical distance beyond this does not increase the expected genetic distance between two populations. For this purpose, genetic patch size should be measured using neutral genes, to provide a general baseline sampling strategy that is not influenced by any particular pattern of environmental diversity. The baseline can then be refined following the principles outlined in Section 3.4.

At the lower end of the scale, the genetic neighbourhood area defines the minimum possible scale for taking distinct population samples, at least of the seed population. At smaller scales mating is random and Hardy–Weinberg equilibrium is expected, with no possibility of division into genetically distinct subpopulations. This applies only to seeds: the population of adult plants may show genetic subdivision at smaller scales if the environment is heterogeneous at smaller scales, imposing smaller-scale heterogeneity of pressures within the genetic neighbourhood. Thus there may be merit in finer-scale sampling of adult populations than seed populations.

However, the genetic neighbourhood area of most wild plant species is remarkably small, far smaller than the unit regarded as one population by the collector. For the insect-pollinated self-incompatible perennial *Trifolium repens* the reproductive genetic neighbourhood area is 2 m² (Gliddon and Saleem, 1985); for the wind-pollinated self-incompatible perennial *Lolium perenne* it is 8.4 m² (ignoring contribution of seed dispersal: calculated from pollen dispersal data of Giddings *et al.*, 1997). In practice, therefore, each sample of a wild population in an *ex situ* collection almost invariably comprises genotypes from what were originally numerous distinct genetic populations. In some species, such as *Trifolium repens*, there is good evidence for genetic differentiation between such genetic populations, each showing distinct local adaptation to microenvironmental variation within the collecting site (reviewed by Sackville Hamilton, 1989). To date it has never been considered justifiable to maintain these genetic populations as separate accessions for conservation purposes – although it is of course necessary for detailed investigations of evolutionary responses to environmental heterogeneity. Combining many genetic populations into a single accession has consequences for its maintenance *ex situ*. These consequences will be considered in Section 3.6.

The genetic neighbourhood area of crop plants is closely related to the type of farming. In primitive farming communities it is generally much smaller than for modern agriculture. Farmers in such communities usually maintain and select their own seed, with limited ‘dispersal’ (by seed

exchange) between isolated communities or even between farmers within communities. It is essential, when collecting, to determine what are the local customs in relation to seed selection and exchange, especially (i) whether a formal centralized system exists for exchanging seed, or whether exchange is informal and centralized through the market, or informal and localized to individual farmer–farmer interactions; and (ii) how much farmers rely on their own farm-saved seed, and if so whether they consciously make their own selections. Only with such local knowledge can the collector judge the probable scale of distribution of diversity.

3.5.3. *Adaptation to environment: targeting by habitat*

Targeting the maximum diversity of habitats for collection will maximize the diversity of genes contributing to adaptation to the selection pressures imposed in the environments sampled. It will also maximize diversity of genes closely linked to the adaptive genes and of pleiotropic characters. It will have no effect on the diversity of genes that are neutral for the particular environmental diversity sampled – this includes not only genes that appear neutral with respect to all known selection pressures, but also genes that are non-neutral for different types of environmental diversity.

In this context, for agricultural and horticultural species, environmental diversity and the corresponding diversity of selection pressures necessarily includes diversity of selections made by farmers and gardeners. It includes, for example, local variation in subjective preferences for flavour and appearance, as well as local variation in perception of agronomic quality.

Effective environmental targeting in this manner depends on the collector having good knowledge of environmental diversity in the region, and of the distribution of the target taxa in relation to environmental diversity. Much of the planning phase of a collection should be devoted to identifying contrasting environments, using as many sources of information as possible, preferably in map form: not only conventional geographical maps, but also maps of surface geology, soil, temperature, rainfall, vegetation and land use. Much additional information is not available in map form, and may not be readily available prior to the expedition, being in the knowledge domain of local extension scientists and farmers. Relevant local knowledge covers not only natural variation between fields but also diversity in farmer-selection pressures resulting from variation in crop usage and variation in preferred crop characteristics.

3.5.4. Phylogeny: targeting centres of diversity

It is now widely accepted that evolution does not progress at a uniform rate, but involves periods of relative stability interspersed with periods of rapid change. Exactly how and how much the rate of evolution changes is still the subject of debate. Nevertheless, for most species and genera it is possible to identify centres of diversity, associated with a phase of rapid diversification at some stage in their evolutionary history.

Centres of diversity are most strongly developed for crop species, leading to the famous pioneering work of Vavilov (1951). These centres are associated with early agricultural developments. They are attributed to disruptive selection caused by the simultaneous action of natural selection for fitness and artificial selection for agronomic value, combined with diverse artificial selections applied by different farmers in different environments, and with introgression between conspecific crop and wild relative.

By definition, a collecting expedition will obtain the greatest diversity if it is located within the centre of diversity of the target taxa. The content of *ex situ* collections should therefore contain a bias in favour of populations from the centre of diversity.

3.5.5. Phylogenetic constraints: targeting limits of species distribution

All species have a limited distribution. Ultimately, this may be attributed to phylogenetic constraints on the extent to which a species may adapt genetically to different environments. Of course, absence of a species from a location does not imply it cannot adapt to that environment – very often absence is attributable rather to dispersal barriers preventing colonization. However, even in these cases, the barriers to dispersal are environments to which the species cannot adapt. Here we are concerned with the actual boundaries to the distribution of a species: the transition from an environment in which the species can and does persist, to one in which it cannot persist. The boundaries to the distribution of a species logically define the limits to which the species can adapt genetically. Populations near the boundary therefore are likely to contain genes not present in the centre of the species range. These genes are likely to be particularly valuable in breeding to broaden the range of adaptation of a crop, particularly for improving tolerance to stresses such as cold, heat, drought, etc.

For these purposes, it is important to consider not just the geographical distribution of the species, but all dimensions of its ecological distribution: its range for pH tolerance, temperature, nutrients, light, moisture, defoliation, competition, and so on. The upper and lower limits to the species distribution on each environmental gradient will each be associated

with different genes for extreme adaptation. It is therefore important to target as many as possible of these limits.

Some of the ecological limits will be encountered geographically within the centre of diversity; in these cases targeting the centre of diversity can be combined with targeting extremes of adaptation. However, other ecological distribution limits and geographical distribution limits are geographically distant from the centre of diversity. The centre of diversity will not contain the genes for adaptation to these extremes. It is then necessary to target these extremes separately from the centre of diversity.

3.5.6. Connectivity: targeting isolated populations

Although the genetic neighbourhood area of most plants is small (see Section 3.5.2), pollen and seed dispersal are usually highly leptokurtic and small numbers can be dispersed large distances. For example, although *Lolium perenne* has a pollen neighbourhood of 8.4 m² (a standard deviation for dispersal of 2.9 m), pollen has been detected at distances over 1000 m (Hayward *et al.*, 1995). Other species, such as *Cacao*, show even greater leptokurtosis of dispersal, with occasionally seeds being dispersed thousands of kilometres over the sea.

Whilst genetic differentiation between populations is possible at any scale above the genetic neighbourhood area, the potential for differentiation increases with the degree of isolation. A completely isolated population cannot exchange genes with other populations and therefore there is no possibility for genetic divergence by drift or selection to be counteracted by the repeated sharing of genes. This is most obvious at the species level, illustrated by the high level of endemism in isolated islands – indeed, this was a major factor leading Darwin (1859) to his theory of evolution by natural selection. The same applies to genetic variation within species, both wild and cultivated.

It is therefore important to target populations that are most isolated. For crop species, isolated farmers or farming communities should be sought. For wild species, areas should be sought where their natural habitats are highly fragmented, especially if the intervening habitats form effective barriers to gene flow.

3.5.7. Drift: targeting small populations

The rate of genetic drift is inversely proportional to the number of individuals in a population. It affects all genes, but especially those that are not subject to strong selection pressures. Thus, two small populations are likely to differ through random drift in all polymorphic parts of their

genome. This is in marked contrast to variation associated with adaptation to particular environments, which generally involves relatively few genes.

Targeting small populations is therefore an efficient means of acquiring maximal nonspecific genetic variation between populations. Against this, small populations are more inbred and contain less within-population variation.

3.5.8. Scales of environmental heterogeneity: stratified sampling

The environment is a multidimensional entity. Genetic adaptation to environment is correspondingly multidimensional. Different environmental variables show different patterns of variation in space and time. Therefore genetic variation for adaptation to different environmental variables also shows different patterns. For example, Sackville Hamilton (1980) collected *Trifolium repens* from an area of high diversity of soils and grassland management but uniform climate. Relative to the global diversity present in the entire gene pool of *T. repens*, genetic diversity between populations was high for vegetative and morphological characteristics important for adaptation to soils and management but low for time of flowering. More generally, it may be, for example, that populations from adjacent fields differ mainly in genes affecting response to management; ones from nearby fields differ mainly in genes for response to aspect; ones further apart differ mainly in genes for response to soils; ones from different altitudes differ mainly for response to temperature; ones from different villages for local human preferences; ones from different latitudes for response to daylength; and so on.

Given this situation, a stratified sampling strategy will not just maximize the genetic diversity collected; it will maximize the diversity of different types of genetic diversity collected (Marshall and Brown, 1975). A particular advantage of stratified sampling is that it does not depend on prior knowledge of the different scales of heterogeneity of different environmental attributes. Although such knowledge helps, nevertheless the fact that different environmental variables show different scales of heterogeneity is itself sufficient to make a stratified sampling procedure more efficient in obtaining qualitatively different types of genetic diversity.

For some purposes, the stratification of sampling procedure should be extended to sampling individuals within sites, at least for natural populations. Certainly this will maximize within-accession diversity sampled from such populations. For example: sampling several individuals from a single genetic population will sample diversity in genes that are truly polymorphic at the genetic population level; sampling from different quadrats within a field will acquire diversity in genes responsible for microscale adaptation to patchiness of the vegetation, soil characteristics,

and microflora, microtopography, etc.; and samples from the boundaries of the field are more likely to contain immigrant genes from nearby, differently adapted populations.

Stratification of sampling procedure within a population is rarely appropriate for crop populations and market populations (Brown and Marshall, 1995). Even for natural populations it may not always be appropriate. In particular, by maximizing within-population variance, a stratified sampling procedure will invalidate agroecogeographical comparison of different populations. If this is important, a random sampling procedure is more appropriate, unless the individual plants are maintained separately.

For species where it is difficult, or even impossible as routine practice, to distinguish plants from each other – as with most perennial herbaceous species communities – it may be impossible to take a truly random sample. In these species, only inflorescences, or leaves, or some other part of the plant, can be sampled at random. This inevitably introduces a size bias into the sample in favour of those plants with the most inflorescences, leaves, etc. (Sackville Hamilton and Chorlton, 1995; Sackville Hamilton *et al.*, 1996).

3.5.9. Breeding system: adjusting sampling procedures

The breeding system has a major influence on the distribution of genetic diversity (see Section 3.2.4.). A number of mechanisms operate to fix particular variants in lines: inbreeding fixes it through homozygosity; apomixis fixes variants even in heterozygotes; some complex chromosome linkages, like those in *Oenothera*, operate to minimize recombination. All these cases reduce within-population variance, so that a correspondingly increased proportion of the total gene pool is represented by variation between populations. In contrast, outbreeders show higher within-population variance. Sampling procedures must be adjusted correspondingly, to take relatively few individuals from many populations of inbreeding, apomictic and similar species, and many individuals from each of fewer populations of outbreeding species (Marshall and Brown, 1975).

Vegetative propagation (by stolons, bulbs, rhizomes, etc.) is functionally equivalent to apomixis in that it can generate numerous genetically identical plants. However, vegetative propagation is often associated with outbreeding, generating a complex two-level population structure. There is high genetic variance among the individuals originating by sexual reproduction through different zygotes, and zero genetic variance (ignoring somatic mutations) among the vegetative progeny derived solely by mitotic division from a single zygote.

In many such vegetatively reproduced species it is impossible to know

at a glance whether two plants are derived from the same or from different zygotes. In these species the two-level population structure can present very considerable problems for collection for efficient conservation. The commonest approach is to ensure a large enough distance between sampled individuals so as to be reasonably confident that they are genetically distinct. However, a single clone of even small herbaceous species can cover hundreds or thousands of square metres (Harberd, 1963; Oinonen, 1967). The distance between adjacent samples therefore has to be undesirably large, in that it eliminates sampling the genetic diversity that is expressed at smaller scales. For many species there is no satisfactory resolution to this problem, as the only resolution may be intensive sampling followed by genetic fingerprinting to determine the genotypic composition of the population sample, which of course is unjustifiably labour intensive.

An additional problem in such species is the sampling bias referred to in Section 3.5.8. Populations of these species often show a highly skewed distribution of physical size of genotypes, with a few large genotypes and many small ones (Sackville Hamilton *et al.*, 1996). When population samples are based on a random selection of inflorescences or leaves, the sample will be strongly biased in favour of the few large genotypes.

3.5.10. Temporal scale of biodiversity

Little attention has been paid to the temporal scale of biodiversity for conservation purposes. Whilst the importance of cyclic fluctuations, chaotic changes and continuous directional shifts are all well acknowledged and documented, it has rarely if ever been considered justifiable for conservation purposes to return to the same sites for repeat collections. The only common reason for returning to a site or region is to test specific hypotheses, for example to test the extent of genetic erosion.

The problem is probably least in species with long-lived seed banks and/or adults. In these, genes for adaptation to one temporary environment may persist in the population long after that environment has been replaced by another. Indeed, in some situations most of the genetic variation in a population may be a relict of past selection pressures (Williams, 1966), and provide a broad pool of adaptability to new unpredictable events. In these situations, a single sample, if large enough, may be sufficient to include diversity of adaptation to the environments experienced by the population over many years.

Notably, species with long-lived adults are often outbreeders that maintain high levels of diversity within populations. Acknowledging that much of this diversity may represent adaptation to temporal environmental diversity adds to the importance of conserving the within-population diversity.

3.6. Genetic Efficiency in Maintenance of *ex situ* Collections

3.6.1. Drift and selection

For efficient maintenance of diversity *ex situ*, the genetic composition of each accession must be maintained intact. Yet each time an accession is subsampled or regenerated, genetic shifts occur by random drift and by selection. Drift occurs at two stages: first in the choice of seed to be used as parents of the next generation of seed; and second in the number of progeny derived from each parent plant – both as female and as male parent. Selection is rarely deliberate and artificial, except where an identifiable genetic variant is subsampled from an accession for separate conservation. However, the environment used for multiplication imposes a natural selection pressure. It is necessary to control drift and selection as far as is economically justifiable.

The breeding system is a major determinant of the nature and extent of the problem. For inbred pure lines of naturally inbreeding species, it is minor: each accession comprises just a single genotype which breeds true on multiplication, so that possibilities for drift or response to selection are minimal. For natural populations of inbreeding species, genetic variance within accessions is small but not zero. For such populations the most efficient approach is often to subdivide accessions into identifiably distinct lines and maintain them as separate accessions. A sub-numbering system to identify component accessions derived from a single original population is a common feature of *ex situ* collections of inbreeding species.

The problem is most acute for wild populations of outbreeding species (Sackville Hamilton, 1995). In these, each accession is a complex mixture of genotypes with high genetic variance and high levels of heterozygosity. The problem is exacerbated by the normal procedure of combining multiple genetic populations into a single accession (Section 3.5.2). Since the spatial arrangement of individuals in the original population is not maintained in the sample, and the accession is maintained instead as a single random-mating population, the accession will normally show higher levels of heterozygosity than was present in the original subdivided population *in situ*, generating a greater diversity of recombinants each time seed are produced for storage. Moreover, the accession will also have higher additive genetic variance and therefore higher heritability than each of the original genetic populations. The accession is therefore more responsive to selection, more likely to change its genetic composition in response to the selection pressure imposed by the particular environment used for multiplication.

Subdivision of such accessions into distinct subsamples is rarely considered justifiable: it is not usually possible to isolate distinct variants from continuous distributions, and since each subsample will itself

remain highly variable and heterozygous, there is little to be gained by subsampling.

Much emphasis has been placed, in most existing seed multiplication protocols, on minimizing drift by using a large sample size for seed multiplication. It is, however, questionable whether this should be regarded as all-important. Probably even more important is the need to avoid the loss of diversity from the collection as a whole. Whilst drift will change the genetic composition of each accession, the direction of change will not be consistent between accessions and so there will be no loss of diversity overall. The main problems leading to loss of diversity are the following: (i) loss of genetic variance within accessions by inbreeding; (ii) loss of genetic variance within accessions by uniform directional selection; (iii) loss of genetic variance between accessions by convergent selection; and (iv) loss of genetic variance between accessions by introgression.

3.6.2. Loss of genetic variance within accessions

The use of too few seed to produce a new generation of seed is likely to eliminate rare alleles from the population, and effectively result in inbreeding.

Seed multiplication is generally undertaken in a uniform environment, or at least one more uniform than the environments occupied by natural populations. The result is likely to be directional selection pressure favouring a single genotype, and therefore the progressive elimination of other genotypes and reduction in genetic variance within accessions.

3.6.3. Loss of genetic variance between accessions

Multiplying all accessions in a common environment (the environment in which the *ex situ* collection is maintained) is likely to impose convergent selection pressure on all accessions – a tendency for all accessions to change towards a common end point. The rate of convergence and extent to which they can converge depends on genetic variance within populations and the potential for transgressive segregation through recombination and new mutation.

Introgression between accessions during seed multiplication occurs if the accessions are not fully isolated, enabling gene flow to occur by pollen transfer. This reduces genetic variance between accessions by increasing the sharing of genes. Ultimately, the stable end point is that at which all accessions are genetically identical. Introgression does not *per se* reduce overall genetic variance in the collection, as the reduction in between-accession variance is offset by an increase in within-accession variance.

However, the combination of introgression with convergent selection has a more detrimental effect than either alone: introgression both increases the rate of convergence and removes any limit to the extent of convergence.

3.6.4. Possible solutions

Clearly there is a strong tendency towards a major loss of diversity both within and between accessions during seed multiplication, by cross-contamination with alien pollen and by uniform selection, directional selection and convergent selection. It is vital to eliminate these or at least control them, as far as is economically viable.

Introgression can be completely eliminated by multiplication within pollen-proof chambers. This procedure is adopted at the authors' institute, although it is relatively expensive to regenerate each accession. The majority of genebanks still multiply in the field, relying on isolation by distance and/or erecting barriers to pollen flow (such as taller crops) between adjacent accessions. Such methods are cheaper and can substantially reduce but not totally eliminate introgression.

Drift can be minimized by using a large number of seed for multiplication. If this is not feasible and a smaller number of seed must be used, more care is required to reduce drift. The component of drift that operates at the stage of selection of seed for multiplication can be reduced by careful choice of seed to encompass as much genetic diversity as possible. In theory it would be possible to characterize potential parents and select the most diverse set, but it is questionable whether this is economically justifiable. A more pragmatic procedure, adopted at the authors' institute, is as follows. When the accession is first prepared for *ex situ* conservation by producing seed from the original population sample, the progeny of each original mother plant are stored in a separate container in long-term storage. These separate containers then form the basis for selection of parents for future cycles of multiplication: at each cycle, one seed is taken from each container in long-term storage.

Extension of the same procedure also controls other causes of genetic shift. A balanced bulk is made for the active collection by taking and mixing an equal number of seed from each mother plant. This reduces to zero the genetic shifts that would occur by both drift and selection, operating through differential contribution of each plant as a female parent. The limitation of the procedure is that it does not control differential contribution of each plant as a male parent, whether the differential is random drift or selective. Nevertheless, although labour intensive, the procedure has benefits in terms of conserving diversity that are considered to be greater than the extra costs of multiplication and conservation.

An alternative approach to reduce the loss of diversity is to multiply seed of different accessions in different environments. For each accession, the objective is to use an environment that imposes the same selection pressure as the original environment of each accession. The ultimate would be to use the original environment itself: this would be ideal not only in terms of the mean but also the variance and structure of the environment, and amounts to full integration of *ex situ* with *in situ* (Section 3.7). Less extreme but more feasible is to maintain a national or international network of multiplication sites, and use the most appropriate site for each accession.

3.7. Genetic Efficiency in *in situ* Conservation

In situ conservation, as defined here, includes conservation on farm and in gardens of landraces and old varieties that are traditionally maintained on farms or in gardens. This germplasm is subject to the same mixing and farmer-selection that have been part of agriculture throughout history. It is the continuation of this traditional method of informal breeding that distinguishes *in situ* from *ex situ* conservation on farms and in gardens. However, many developed countries now have legislation protecting plant breeders' rights, which effectively prevents the continuation of traditional agriculture and therefore *in situ* conservation of crop species. This section therefore applies to *in situ* conservation of wild species, and of crop species only in those countries where legislation permits.

3.7.1. Choice of sites

Sites for inclusion in a network for *in situ* conservation must be chosen to maximize the diversity that can be maintained. This is similar in principle to choosing sites for collection to maximize diversity collected. Rather than repeat these principles in detail, only a summary is presented here, and the reader is referred to Sections 3.2 and 3.4 for elaboration of the underlying principles.

Sites should cover the entire ecological range of the species, with a bias towards its centre(s) of diversity and the ecological extremes of its distribution. There should be a stratified distribution of sites with several levels of clustering: each site should be large enough to encompass a cluster of several genetic populations; each site should be part of a cluster of several nearby sites, close enough for occasional gene flow between sites as a result of rare long-distance dispersal events; each cluster of sites should be part of a larger cluster; and so on to encompass the entire range of the species. Such stratification will serve a dual purpose of: (i) optimizing gene

flow within and between populations; and (ii) encompassing the maximum possible range of types of diversity. Within the general stratification, sites and clusters should be chosen to maximize diversity of environments and therefore of selection pressures within and between sites and clusters. As always, of course, diversity of selection pressures is taken to include artificial as well as natural selection, with diversity associated with variation in local preferences and farmers' concepts of quality and agronomic value.

The shape of sites and clusters also needs consideration (Forman, 1995). For example, linear habitats may be useful to increase connectivity between clusters with minimal increase in areas of the region set aside for conservation. For rare species, there may be a need to create new populations at sites with sufficient connectivity to existing populations to prevent loss of diversity through inbreeding.

3.7.2. Additional measures

Additional measures can be taken to increase biodiversity within the selected network of conservation sites, essentially by increasing the diversity of environments and selection pressures within and between sites.

For crop species, farmers can be actively encouraged to value the distinctiveness of the traditional farming practices of the region; and the local customer community can be actively encouraged to value the distinctiveness of local traditions and their consequent demands on local farmers. Important traditional farming practices can include factors such as conscious selection by the farmer for genetic variation within and between varieties for tolerance to disease, drought, heat, etc. These traditions are based on utilizing high diversity to provide low-cost, sustainable, low-risk protection from environmental stresses and hazards. That is, they benefit not only conservation of biodiversity but also the farm economy.

For wild and some crop species, there can also be opportunities for increasing diversity by appropriately diverse management. Emphasis is on diverse management, as many management procedures, especially mechanized ones, tend to reduce diversity. For example, cutting, liming, fertilization and control of weeds, pathogens and pests are usually applied uniformly across entire sites; in so doing they reduce environmental diversity and therefore the diversity of selection pressures and biodiversity at the scale of the site. If such procedures are also applied consistently from year to year, there will also be less temporal variation in selection pressures, again reducing biodiversity at the scale of the field. In contrast, management by grazing imposes cutting, trampling and fertilization that is spatially and temporally variable – to an extent that depends on the grazing behaviour of the selected herbivore.

Similarly, diversity of management should be encouraged at larger scales, including landscape and regional. The principal problem here relates to how to construct and implement a conservation policy. For example, a management policy may be implemented that maximizes biodiversity within a field; but if that same policy is applied to all sites, the same range of biodiversity will be promoted at all sites, reducing biodiversity at the larger landscape and regional scales. If the policy is to be centrally established and imposed, it may be economically impossible to incorporate the larger-scale variation in management necessary to maximize biodiversity at landscape and regional levels. A decentralized system is likely to be preferable, especially to incorporate regional variations in traditions.

3.8. Conclusions

In this chapter we have shown that biodiversity is a scale-dependent phenomenon, and that for its efficient conservation we need to include all scales from a few square centimetres to thousands of square kilometres. We have also shown that the distribution of genetic diversity of any species depends on its life cycle and consequent evolutionary characteristics. Efficient conservation depends on having a good knowledge of population structure and the life cycle characteristics that determine this – dispersal profiles, breeding system, longevity. The same principles apply not only to wild species but also to crop species, the major difference being that crop species have dispersal profiles determined largely by the farmer and market, and are subject to artificial selection by the farmer as well as natural selection.

References

- Allard, R.W. (1990) Future directions in plant population genetics. In: Brown, A.H.D., Clegg, M.T., Kahler, A.L. and Weir B.S. (eds) *Plant Population Genetics, Breeding, and Genetic Resources*. Sinauer Associates, Sunderland, MA, pp. 43–63.
- Allard, R.W. and Hansche, P. E. (1965) Population and biometrical genetics in plant breeding. In: Geerts, S.J. (ed.) *Genetics Today*, vol. 3, *Proceedings of XIth International Congress of Genetics, The Hague, Netherlands, 1963*. Pergamon Press, Oxford, pp. 665–668.
- Allard, R.W., Garcia, P., Sáenz-de-Miera, L.E. and Pérez de la Vega, M. (1993) Evolution of multilocus genetic structure in *Avena hirtula* and *Avena barbata*. *Genetics* 135, 1124–1139.
- Begon, M., Harper, J.L. and Townsend, C.R. (1996) *Ecology: Individuals, Populations and Communities*, 3rd edn. Blackwell Science, Oxford.

- Breese, E.L., Lewis, E.J. and Evans, G.M. (1981) Interspecies hybrids and polyploidy. *Philosophical Transactions of the Royal Society of London, Series B* 292, 487–497.
- Brown, A.H.D. and Marshall, D.R. (1995) A basic sampling strategy: theory and practice. In: Guarino, L., Ramanatha Rao, V. and Reid, R. (eds) *Collecting Plant Genetic Diversity: Technical Guidelines*. CAB International, Wallingford, UK, pp. 75–91.
- Brown, A.H.D. and Weir, B.S. (1983) Measuring genetic variability in plant populations. In: Tanksley, S.D. and Orton, T.J. (eds) *Isozymes in Plant Genetics and Breeding. Part A*. Elsevier, Amsterdam, pp. 219–329.
- Brown, A.H.D., Clegg, M.T., Kahler, A.L. and Weir, B.S. (1990) *Plant Population Genetics, Breeding, and Genetic Resources*. Sinauer Associates, Sunderland, MA, 449 pp.
- Burrough, P.A. (1995) Spatial aspects of ecological data. In: Jongman, R.H.G., ter Braak, C.J.F. and van Tongeren, O.F.R. (eds) *Data Analysis in Community and Landscape Ecology*. Cambridge University Press, Cambridge, pp. 213–251.
- Caetano-Anollés, G., Callahan, L.M., Williams, P.E., Weaver, K.R. and Gresshoff, P.M. (1995) DNA amplification fingerprinting analysis of bermudagrass (*Cynodon*): genetic relationships between species and interspecific crosses. *Theoretical and Applied Genetics* 91, 228–235.
- Cameron, J.W. and Soost, R.K. (1982) Breeding and development. In: 75 years of *Citrus* research. *California Agriculture* 36, 4–6.
- Darlington, C.D. (1956) *Chromosome Botany*. Allen & Unwin, London, 186 pp.
- Darlington, C.D. (1958) *The Evolution of Genetic Systems*. Oliver & Boyd, London, 265 pp.
- Darwin, C. (1859) *The Origin of Species by Means of Natural Selection*. John Murray, London.
- de Nettancourt, D. (1993) Self- and cross-incompatibility systems. In: Hayward, M.D., Bosemark, N.O. and Romagosa, I. (eds) *Plant Breeding: Principles and Prospects*. Chapman & Hall, London, pp. 203–210.
- Dillmann, C., Charcosset, A., Bar-Hen, A., Goffinet, B., Smith, J.S., Datté, Y. and Guiard, J. (1995) The estimation of molecular genetic distance in maize for DUS and ED protocols: optimisation of the information and new approaches of kinship. BMT/3/6. UPOV; Working Group on Biochemical and Molecular Techniques and DNA Profiling in Particular. Upov, Geneva, pp. 2–27.
- Dunford R.P., Kurata, N., Laurie, D.A., Money, T.A., Minobe, Y. and Moore, G. (1995) Conservation of fine-scale DNA marker order in the genomes of rice and the Triticeae. *Nucleic Acids Research* 23, 2724–2728.
- Ellis Davies, W. and Young, N.R. (1967) The characteristics of European, Mediterranean and other populations of white clover (*Trifolium repens* L.). *Euphytica* 16, 330–340.
- Fearon, C.H., Cornish, M.A., Hayward, M.D. and Lawrence, M.J. (1994) Self-incompatibility in ryegrass. X. The number and frequency of alleles in a natural population of *Lolium perenne* L. *Heredity* 73, 245–261.
- Forman, R.T.T. (1995) *Land Mosaics: the Ecology of Landscapes and Regions*. Cambridge University Press, Cambridge.
- Frankel, O.H., Brown, A.H.D. and Burdon, J.J. (1995) *The Conservation of Plant Biodiversity*. Cambridge University Press, Cambridge, 299 pp.

- Giddings, G.D., Sackville Hamilton, N.R. and Hayward, M.D. (1997) The release of genetically modified grasses. I: Pollen dispersal to traps in *Lolium perenne*. *Theoretical and Applied Genetics* (in press).
- Gliddon, C.J. and Saleem, M. (1985) Gene-flow in *Trifolium repens* – an expanding genetic neighbourhood. In: Jacquard, P., Heim, G. and Antonovics, J. (eds) *Genetic Differentiation and Dispersal in Plants. Proceedings of the NATO Advanced Workshop on Population Biology of Plants*. Springer-Verlag, Berlin, pp. 293–309.
- Guarino, L., Ramanatha Rao, V. and Reid, R. (1995) *Collecting Plant Genetic Diversity: Technical Guidelines*. CAB International, Wallingford, UK.
- Hamrick, J.L. and Godt, M.J. (1990) Allozyme diversity in plant species. In: Brown, A.H.D., Clegg, M.T., Kahler, A.L. and Weir B.S. (eds) *Plant Population Genetics, Breeding, and Genetic Resources*. Sinauer Associates, Sunderland, MA, pp. 43–63.
- Harberd, D.J. (1963) Observations on natural clones of *Trifolium repens* L. *New Phytologist* 62, 198–204.
- Hayward, M.D., Giddings, G.D., Thorpe, I., Sackville Hamilton, N.R. and Warren, J.M. (1995) Gene flow in pollinated forage grasses. *Department of the Environment Conference on Environmental Impact of Genetically Modified Crops, Monkswood, November 1995*. DOE.
- Jones, R.N. and Rees, H. (1982) *B Chromosomes*. Academic Press, New York.
- Kearsey, M.J. (1993) Biometrical genetics in breeding. In: Hayward, M.D., Bosemark, N.O. and Romagosa, I. (eds) *Plant Breeding: Principles and Prospects*. Chapman & Hall, London, pp. 163–183.
- Levin, S.A. (1989) Challenges in the development of a theory of community and ecosystem structure and function. In: Roughgarden, J., May, R.M. and Levin, S.A. (eds) *Perspectives in Ecological Theory*. Princeton University Press, Princeton, NJ, pp. 242–255.
- Marshall, D.R. and Brown, A.H.D. (1975) Optimum sampling strategies in genetic conservation. In: Frankel, O.H. and Hawkes, J.G. (eds) *Crop Genetic Resources for Today and Tomorrow*. Cambridge University Press, Cambridge, pp. 53–80.
- Mather, K. (1973) *Genetical Structure of Populations*. Chapman & Hall, London, 197 pp.
- Monestiez, P., Goulard, M. and Charmet, G. (1994) Geostatistics for spatial genetic structures: study of wild populations of perennial ryegrass. *Theoretical and Applied Genetics* 88, 33–41.
- O'Donnell, S. and Lawrence, M.J. (1984) The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas* IV. The estimation of the number of alleles in a population. *Heredity* 53, 495–507.
- Oinonen, E. (1967) The correlation between the size of Finnish bracken (*Pteridium aquilinum* (L.) Kuhn) clones and certain periods of site history. *Acta Forestalia Fennica* 83, 3–96.
- Pérez de la Vega, M. (1993) Biochemical characterization of populations. In: Hayward, M.D., Bosemark, N.O. and Romagosa, I. (eds) *Plant Breeding: Principles and Prospects*. Chapman & Hall, London, pp. 184–200.
- Rees, H. (1961) Genotypic control of chromosome form and behaviour. *Botanical Reviews* 27, 288–318.
- Rees, H. and Dale, P.J. (1974) Chiasmata and variability in *Lolium* and *Festuca* populations. *Chromosoma* 47, 335–351.

- Riley, R. and Chapman, V. (1958) Genetic control of the cytologically diploid behaviour of hexaploid wheat. *Nature* 182, 713–715.
- Sackville Hamilton, N.R. (1980) Variation and adaptation in wild populations of white clover (*Trifolium repens* L.) in East Anglia. PhD thesis, University of Cambridge. (Published by University Microfilms International, Ann Arbor, MI, 1982.)
- Sackville Hamilton, N.R. (1989) Life history studies. In: Stirton, C.H. and Zarucchi, J.L. (eds) *Advances in Legume Biology: Monographs in Systematic Botany of the Missouri Botanic Gardens*. Missouri Botanic Gardens, Missouri, pp. 1–16.
- Sackville Hamilton, N.R. (1995) Research projects in the United Kingdom. In: Gass, T., Sackville Hamilton, N.R., Kolshus, K. and Frison, E. (eds) *Report of a Working Group on Forages, Fifth meeting, 31 March–2 April 1995, Hissar, Bulgaria. European Cooperative Programme for Crop Genetic Resources Networks*. International Plant Genetic Resources Institute, Rome, p. 51.
- Sackville Hamilton, N.R. and Chorlton, K.H.C. (1995) Collecting vegetative material of forage grasses and legumes. In: Guarino, L., Ramanatha Rao, V. and Reid, R. (eds) *Collecting Plant Genetic Diversity: Technical Guidelines*. CAB International, Wallingford, UK, pp. 467–484.
- Sackville Hamilton, N.R., Jones, D., Cresswell, A. and Fothergill, M. (1996) Genetic diversity and sustainability of clover-based pastures: 2. Size hierarchies and sampling bias. In: Younie, D. (ed.) *Legumes in Sustainable Farming Systems. Symposium of the British Grassland Society, Aberdeen*. pp. 179–180.
- Savidan, Y. (1983) Genetics and utilization of apomixis for the improvement of guineagrass (*Panicum maximum*). In: Smith, J.A. and Hays, V.W. (eds) *Proceedings of the XIV International Grasslands Congress, Lexington, 1981*. Westview Press, Boulder, CO, pp. 2182–2184.
- Soltis, D.E. and Soltis, P.E. (1990) *Isozymes in Plant Biology*. Chapman & Hall, London, 268 pp.
- Stammers, M., Harris, J., Evans, G.M., Hayward, M.D. and Forster, J.W. (1995) Use of random PCR (RAPD) technology to analyse phylogenetic relationships in the *Lolium/Festuca* complex. *Heredity* 74, 19–27.
- Stebbins, G.L. (1950) *Variation and Evolution in Plants*. Columbia University Press, New York.
- Tanskley, S., Grandillo, S., Fulton, F.M., Zamir, D., Eshed, Y., Petiard, V., Lopez, J. and Beck-Bunn, T. (1996) Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. *Theoretical and Applied Genetics* 96, 213–224.
- Vavilov, N.I. (1951) The origin, variation, immunity and breeding of cultivated plants (translated by K.S. Chester). *Chronica Botanica* 13, 1–366.
- Weising, K., Nybom, H., Wolff, K. and Meyer, W. (1995) *DNA Fingerprinting in Plants and Fungi*. CRC Press, Boca Raton, FL, 322 pp.
- Williams, G.C. (1966) *Adaptation and Natural Selection*. Princeton University Press, Princeton, NJ.
- Wright, S. (1965) The interpretation of population structure by *F* statistics with special regard to system of mating. *Evolution* 19, 395–420.

Genomic Relationships, Conserved Synteny and Wide-hybrids

D.A. Laurie, G.J. Bryan and J.W. Snape

*John Innes Centre, Norwich Research Park, Colney, Norwich
NR4 7UH, UK*

4.1. Introduction

Phylogenetic trees relate plant species to one another in taxonomic frameworks. The progressive divergence of species from common evolutionary ancestors clearly implies that related species have many common features of genome organization and the genetical control of traits. Indeed, the basis of taxonomy is that related species are similar in morphology. However, the relationships between plant genomes have only become clearly defined at the DNA level following the development of molecular methods for genome analysis and genetic mapping. These methods have allowed the detailed comparison of genomes of relatively closely related species such as wheat (*Triticum aestivum*), rye (*Secale cereale*) and barley (*Hordeum vulgare*) and have allowed their genetic maps to be aligned. Importantly, it is now possible to compare the genomes of more distantly related species such as wheat, maize (*Zea mays*), sorghum (*Sorghum bicolor*) and rice (*Oryza sativa*). Likewise, the genetic maps of potato and tomato have been aligned, as have the maps of various *Brassica* species. *Brassica* genomes can also be aligned with maps of the model plant species *Arabidopsis thaliana* (thale cress).

Molecular methods therefore offer new opportunities for analysing genome relatedness and the genetical basis of phenotypic variation within and between plant species. This in turn provides a new framework for germplasm assessment and provides new tools for the efficient use of germplasm resources in plant improvement. This chapter deals with the

analysis and implications of genome relationships in crops. Most examples will be from cereals but reference will also be made to two other crop groups, the brassicas and legumes.

4.2. Structural Analysis of Genome Relationships

Cytogenetic analysis of somatic or meiotic preparations by conventional staining, chromosome banding and *in situ* hybridization ('chromosome painting') provides information on the overall organization of plant genomes. Meiotic chromosome pairing behaviour in natural or artificial hybrids is also an important guide to genome relationships, and until recently was the only method for assaying genome similarities. However, pairing can be strongly affected by single genes such as *Ph1* in wheat (Riley and Chapman, 1958), and the degree of pairing between related genomes may not directly reflect overall genome homology. Nevertheless, chromosome pairing has been shown to be in good agreement with the conventional taxonomy of the *Triticeae* (Kimber and Yen, 1988). Analysis of hybrids is important because allo- and autopolyploidy are common in plant evolution.

Allopolyploids include crop species such as oilseed rape (canola), oat, wheat, groundnut, upland and sea island cotton, some bananas, strawberry, arabica coffee and tobacco. The cultivated species derived from polyploid hybrids are generally thought to have combined agronomically important traits from their diploid ancestors.

4.3. Genome Size and Composition

The amount of nuclear DNA varies greatly between plant species (Bennett and Leitch, 1995). For example, the haploid genome of hexaploid wheat contains about 1.7×10^{10} bp, each of the three genomes being roughly similar in size to that of barley (5×10^9 bp). The genome of rice, in contrast, contains only about 4×10^8 bp (Moore *et al.*, 1993a). Variation in DNA amount, other than variation due to ploidy level, is probably due almost entirely to variation in the copy number of repeated DNA sequences (Flavell and Smith, 1976; Deshpande and Ranjekar, 1980; Rimpau *et al.*, 1980). Large genome species contain higher proportions of short tandem repeats, often present as blocks of heterochromatin, and a higher proportion of dispersed repeats. A major component of the latter in wheat and barley are retroelement-like sequences such as Wis-1, Wis-2 and Bis-1 (Moore *et al.*, 1991). Long-range analysis of repeat structure in large cereal genomes suggests that chromosome evolution has involved duplication of blocks of repeats (Moore *et al.*, 1993b).

Nuclear DNA amount can also vary within plant species. This variation,

and variation between species, can often be correlated with ecogeographical variables suggesting that DNA amount is of adaptive significance (Bennett, 1987). In maize, for example, DNA amount declines with increasing altitude or increasing latitude (Bennett, 1976; Rayburn *et al.*, 1985). This variation can be as much as 30% of the genome size and involves variation in the amount of C-band heterochromatin (Rayburn *et al.*, 1985; McMurphy and Rayburn, 1992) as well as variation in dispersed repeats (Rivin *et al.*, 1986).

Proposed mechanisms producing changes in repetitive sequences include unequal crossing over, transposition, retrotransposition, gene conversion and replication slippage (Flavell, 1986; Dover, 1993). They are rapid in evolutionary terms and consequently the copy number of specific sequences can vary greatly in different species or lineages. New sequence variants also arise frequently, probably because there is little or no selection constraining the DNA sequence of many classes of repeats. As a result, there are families of repeated sequences which are specific to individual grass species, or genomes, or which differ greatly in copy number between closely related species (Zhang and Dvořák, 1990; McNeil *et al.*, 1994), and in general the divergence of repetitive sequences can be used to indicate the degree of relatedness of whole genomes in diploid species and their polyploid derivatives.

Repeat sequence plasticity can be exploited for Southern blot analysis and for chromosome painting using dispersed repeats (Lapitan *et al.*, 1986) or total genomic DNA (Schwarzacher *et al.* 1989, 1992). For example, wheat chromosomes can be clearly differentiated from those of rye (Heslop-

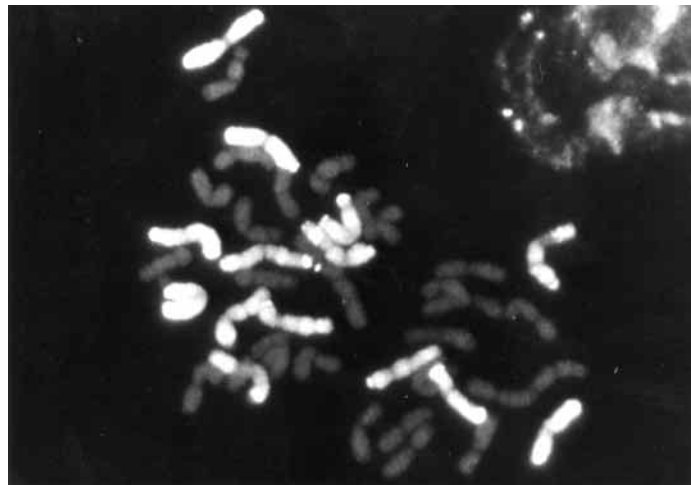


Fig. 4.1. Genomic *in situ* hybridization using digoxigenin-labelled rye DNA on root-tip metaphase chromosomes of hexaploid triticale. The 14 rye chromosomes are brightly stained. (With kind permission of Dr. Nuno Neves.)

Harrison *et al.*, 1990; Islam-Faridi and Mujeeb-Kazi, 1995) (Fig. 4.1) or barley (Mukai and Gill, 1991) in hybrids and the three genomes can be differentiated in hexaploid wheat (Mukai *et al.*, 1993). Fluorescent *in situ* hybridization (FISH) is emerging as the method of choice for such work because several probes, each labelled to give a different colour when illuminated by specific wavelengths, can be hybridized to the same preparation (Nederlof *et al.*, 1989; Leitch *et al.*, 1991; Heslop-Harrison and Schwarzacher, 1993; Jiang and Gill, 1994). *In situ* hybridization is extremely valuable for identifying hybrids and introgression lines in crop species and for studies of plant evolution in general (Bennett *et al.*, 1992; Orgaard and Heslop-Harrison, 1994).

4.4. DNA Fingerprinting and Genetic Map Construction

The rapid evolution of repeated sequences results in high levels of polymorphism that can be detected by methods based on Southern hybridization or polymerase chain reaction (PCR). Repeated sequences are, therefore, potentially useful as fingerprinting markers. Polymorphism in repeated sequences can also be used for genetic map construction. For example, it has recently been possible to map very rapidly almost 500 interspersed repetitive sequences in mouse (McCarthy *et al.*, 1995). Larger sequences, such as retrotransposons, can also be used for fingerprinting or genetic mapping (Fukuchi *et al.*, 1993). Specific PCR primers enable some repeats to be used as single locus markers. The most useful of these are the simple sequence repeats (SSRs or microsatellites), which are usually di- or tri-nucleotide repeats, since these are often flanked by unique sequence DNA. This allows the use of flanking primers for PCR detection of length polymorphism within the SSR itself (Litt and Luty, 1989; Weber and May, 1989). SSRs are highly polymorphic and are therefore useful for studies of genetic diversity as well as for mapping (Rongwen *et al.*, 1995).

Although repeated sequences can be used for genetic mapping, the genetic maps of most crop species and model organisms are based on polymorphisms of low- or single-copy sequences. Most reasonably detailed maps are based on analysis of restriction fragment length polymorphism (RFLP), which is detected by the hybridization of cloned DNA segments to genomic DNA digested with restriction enzymes. Most RFLP probes are either small fragments of genomic DNA (selected for low copy number) or cDNA clones which represent transcribed genes of the tissue from which they were isolated. In future, however, most maps will probably be based largely on amplified fragment length polymorphisms (AFLP) (Vos *et al.*, 1995), simple sequence repeats (SSRs or microsatellites) and, to a lesser extent, random amplified polymorphic DNAs (RAPD) (Williams *et al.*, 1990). These PCR methods will be favoured because of

their speed of assay, degree of polymorphism and their requirement for only small amounts of DNA from the plants of interest. This is illustrated by the analysis of the haploid megasporophytes of conifers. These yield sufficient DNA for the construction of genetic maps using RAPD markers (Binelli and Bucci, 1994; Grattapaglia and Sederoff, 1994).

Genetic maps are usually constructed to locate genes of biological or economic interest in relation to molecular markers. For single genes of large effect this is straightforward since a chromosome location can easily be established by segregational analysis. However, many traits of interest to plant breeders and researchers are controlled by several to many genes acting in combination. The availability of high-resolution genetic maps and increasingly sophisticated statistical methods has greatly increased the ability to identify the number and chromosome locations of such quantitative trait loci (QTL) (Lander and Botstein 1989; Jansen, 1994; Kearsey and Hyne, 1994).

There are two main applications for genetic maps. The first is the use of DNA markers linked to the gene or genes of interest to select individuals in breeding programmes without the need to assess the phenotype at each stage of development (Gebhardt and Salamini, 1992). Such marker-assisted selection (MAS) is already used in breeding programmes. It is particularly useful for the manipulation of quantitative traits, which are the most difficult to evaluate phenotypically in conventional breeding programmes. The second application of genetic maps is the isolation of genes by a map-based cloning approach such as chromosome walking or chromosome landing (Tanksley *et al.*, 1995). This approach has been used, for example, in the isolation of genes conferring disease resistance from tomato (Martin *et al.*, 1993).

4.5. Comparative Maps and Conserved Synteny

The cloned DNA fragments used to detect RFLP often hybridize well to DNA of related species, giving a series of common markers that enable the respective genetic maps to be aligned. This is termed comparative mapping. cDNA clones are particularly useful for inter-species work, probably because the accumulation of mutations in expressed gene sequences is limited by selection. In contrast to the plasticity of genome structure revealed by studies of repeated sequences, including nucleolus organizer regions (Leitch and Heslop-Harrison, 1992; Dubcovsky and Dvořák, 1995), the order of single-copy sequences on genetic maps of different species often shows a high degree of conservation. Conservation of marker order is termed conserved synteny, since two genes or markers are syntenic if they lie on the same piece of DNA, usually a chromosome or chromosome fragment, irrespective of whether they are genetically linked (Renwick, 1971).

4.6. Comparative Mapping in Grasses

Comparative mapping shows that the three genomes of hexaploid wheat have very similar marker orders with the exception of a few major chromosome rearrangements (Devos and Gale, 1993). Additional translocations are found when comparison is made with the rye or barley genome, but marker order is still highly conserved (Devos *et al.*, 1993; Namuth *et al.*, 1994). Comparative maps of maize and sorghum are well established (Hulbert *et al.*, 1990; Melake Berhan *et al.*, 1993; Pereira *et al.*, 1994; Pereira and Lee, 1995) and comparative maps of *Oryza* species are being developed (Jena *et al.*, 1994).

Remarkably, comparison of the genetic maps of wheat and rice (Kurata *et al.*, 1994a), rice and maize (Ahn and Tanksley, 1993) and wheat and maize (Devos *et al.*, 1994) also reveals large linkage segments that can be aligned (Moore *et al.*, 1993a; Bennetzen and Freeling, 1993). Furthermore, equivalent linkage blocks can be recognized in rice, wheat, maize, sorghum, sugar cane and millet. Thus a model of cereal genomes can be constructed using segments of the rice genetic map as the building blocks. This aligns the maps of the various species, revealing the genetic relationships between their chromosomes (Moore *et al.*, 1995). This analysis also confirms the tetraploid nature of the maize genome since the duplicated chromosome segments (Helentjaris *et al.*, 1988) can be organized into two separate genomes. An additional important result from comparative mapping is that RFLP detected by cDNA clones from single rice YAC (yeast artificial chromosome) clones either cosegregate in barley or are tightly linked in the same map order as in rice (Dunford *et al.*, 1995; Kilian *et al.*, 1995). Thus, fine-scale organization of gene order is also conserved.

Mapping studies also suggest that there are duplicated regions in diploid species, suggesting that duplication and subsequent divergence of chromosome segments has been an important feature of plant evolution even in the absence of polyploidy. Examples include regions containing dwarfing genes in rye (Plaschke *et al.*, 1995), segments of rice chromosomes 1 and 5 (Kishimoto *et al.*, 1994), duplications in sorghum (Chittenden *et al.*, 1994) and soybean (*Glycine max*) (Shoemaker *et al.*, 1992) that may reflect ancient polyploidy, and duplications in the diploids *Brassica oleracea* (Kianian and Quiros, 1992) and *B. nigra* (Truco and Quiros, 1994).

4.7. Comparative Mapping in Other Plant Groups

Comparative mapping in the *Solanaceae* showed that when RFLP maps of potato (*Solanum tuberosum*) and tomato (*Lycopersicon esculentum*) were compared, the two genomes showed a high degree of conserved synteny. Five

whole-arm inversions were identified that distinguished the two genomes (Bonierbale *et al.*, 1988; Gebhardt *et al.*, 1991; Tanksley *et al.*, 1992a).

Comparative mapping has also been reported extensively in legumes, where several studies show conserved linkage groups between species such as mungbean (*Vigna radiata*) and cowpea (*V. unguiculata*) (Menancio-Hautea *et al.*, 1993) and pea (*Pisum sativum*) and lentil (*Lens culinaris*) (Weeden *et al.*, 1992; Muehlbauer *et al.*, 1995). However, the soybean genome appears to show considerable internal duplication and rearrangement relative to other legume genomes.

Comparison of *Brassica* species shows that there are a greater number of chromosome rearrangements than commonly found in cereal species, but there are still large segments with conserved marker orders (Ferreira *et al.*, 1994; Teutonico and Osborn, 1994; Parkin *et al.*, 1995). Importantly, *Brassica* species and the model plant species *Arabidopsis thaliana* are both in the *Cruciferae* family and are sufficiently closely related for *Arabidopsis* clones to be used directly for mapping and gene isolation in *Brassica*. Comparison of the genomes of *B. oleracea* and *Arabidopsis* reveals extensive rearrangement, including at least 17 translocations and nine inversions, but homoeologous segments can still be recognized (Kowalski *et al.*, 1994). Thus, *Brassica* researchers can readily exploit the vast mapping, sequencing and genetic resources of *Arabidopsis*.

4.8. Comparative Trait Mapping

Analysis using molecular markers can be used to study the genetical control of any aspect of plant phenotype. This approach is currently being pursued in many species, including all major crops. The major implication of comparative mapping is that genes controlling specific aspects of plant phenotype will have corresponding genetic map locations (Paterson *et al.*, 1995).

Comparative trait mapping in related species should therefore identify homoeologous genes which can be defined as genes that share a common evolutionary ancestor, sequence similarity and equivalent function. This is illustrated by analyses of wheat, rye and barley which show that all three species have a major vernalization response gene in a similar map location on the long arm of the homoeologous group 5 chromosomes (Fig. 4.2). This implies that vernalization response in these three species is regulated by a common gene which forms a homoeoallelic series across the *Triticeae*. Similarly, major photoperiod response genes have been mapped in equivalent regions of the group 2 chromosomes of wheat (Worland, 1996) and barley (Laurie *et al.*, 1994).

Figure 4.2 also illustrates an example of equivalent phenotypes spanning a much greater taxonomic distance. Comparative mapping shows

that single recessive mutations conferring a liguleless phenotype are found in corresponding chromosome locations in barley, rice and maize. The group 2 chromosomes of wheat and rye have *liguleless* mutants, which have not been mapped with RFLP markers but whose genetic map location can be predicted with some certainty. The recently identified phenol reaction gene in barley may be equivalent to a gene conferring a similar phenotype in rice (Fig. 4.2).

These and other results suggest that much of the biology of cereal species is conserved. Similar results could be quoted from other plant families. Thus, traits mapped in one species should provide plant breeders and researchers with valuable information on the map locations of equivalent genes in related species. A comparative approach will therefore simplify and accelerate the genetic analysis of crop species.

Genes of major agronomic importance are likely to be cloned from crop species in the near future, particularly in rice, maize and *Brassica* (in the latter via *Arabidopsis*). Map-based cloning will not be easy in large genome species such as wheat and barley due to high amounts of repetitive DNA, but exploitation of conserved synteny with rice and other grass species offers an attractive alternative route. Once genes are identified, novel alleles could be identified in germplasm collections by PCR screening. An implication of comparative mapping is that this approach will be usable in a range of related species, even those in which the primary trait may not have been mapped. This may therefore offer a rapid means of identifying useful accessions in species where little mapping information is available.

Several authors have proposed that allelic variation may be manifest either as a qualitative major gene effect, or as a more subtle effect which can only be detected statistically, depending on the particular allele combination under study (Thompson, 1975; Robertson, 1985; Beavis *et al.*, 1991). In the latter case the effect would be termed a quantitative trait locus or QTL. Comparative mapping suggests that this analysis can be extended across species and that flanking markers for major genes will provide markers for QTL manipulation in other crosses and in other species, thus increasing selection efficiencies and the ability to construct new QTL allele combinations. Identification of qualitative differences between alleles would also allow map-based cloning of at least some genes that usually affect quantitative traits.

The added value offered by comparative mapping means that it will be highly beneficial to crop genetics if standard sets of cross-hybridizing 'anchor' probes are developed and made widely available, as these will provide the necessary basis for genetic map alignment. cDNA clones for detecting RFLP would be suitable for such anchor sets, although the analysis of RFLP is slow and cumbersome. Alternatively, it may be possible to develop PCR anchor markers as there is a high level of transferability of



Fig. 4.2. Conserved synteny in cereals. (a) Comparative maps showing the conserved location of the major vernalization response genes *Sh₂* in barley, *Vrn1* in wheat and *Sp1* in rye on the long arms of the homoeologous group 5 chromosomes. (b) Wild-type (left) and *liguleless* mutant (right) leaves of barley showing absence of the ligule (arrow) and auricles (arrowheads). (c) Comparative mapping of *liguleless* and *phenol reaction* mutants in maize, rice and barley. For clarity, only some of the mapped RFLP markers are shown in (a) and (c). References: 1, Laurie *et al.* (1995); 2, Galiba *et al.* (1995); 3, Plaschke *et al.* (1993); 4, *Maize Genetics Newsletter* vol. 67, 1993; 5, Ahn and Tanksley (1993); 6, Xiao *et al.* (1992); 7, Tanksley *et al.* (1992b); 8, Heun *et al.* (1991); 9. Pratchett and Laurie (1994); 10. Takeda and Zhang (1994).

PCR primers for cDNA sequences between grass species (G. Bryan, Scottish Crop Research Institute, unpublished data).

4.9. Experimental Advantages of Individual Cereal Species

A comparative approach to crop genetics will benefit from careful utilization of the particular experimental advantages of individual species. In cereals, for example, transposon tagging is available in maize and has been used extensively to isolate mutations (Walbot, 1992). Moreover, with the availability of cDNA sequence information it is now possible to design strategies to select for mutants affecting a specific cloned sequence, even if the sequence is of unknown function. For example, PCR screening of lines using a gene sequence primer and a primer for the *Mutator* transposon has been used to select mutant alleles of the *Anther ear1* gene containing *Mu* insertions (Bensen *et al.*, 1995). This is an extremely powerful strategy for determining the biological role of cloned sequences.

The small genome size of rice and its detailed genetic map (Kurata *et al.*, 1994b) make it attractive for chromosome walking and the assembly of complete physical genome maps using YAC and bacterial artificial chromosome (BAC) libraries. Chromosome manipulations, including single chromosome assays, are easily undertaken in hexaploid wheat, taking advantage of its tolerance of aneuploidy (Law *et al.*, 1987). Recessive mutants are easily produced in diploids like barley and rye. Wheat and barley also have simple and efficient systems for the production of doubled haploids, which makes them particularly suited to the genetical analysis of quantitative traits.

In order to identify genes controlling characters of interest, it would be valuable to have detailed knowledge of all the genes expressed in a plant. cDNA sequencing and the development of EST (expressed sequence tagged) databases can be undertaken with similar efficiency in any diploid, but an advantage of using small genome species such as *Arabidopsis* or rice is that they have comprehensive physical maps (contigs) built up of overlapping YAC and BAC clones. Hybridizing probes to such contigs will, potentially, enable all cDNAs to be mapped to specific chromosome locations without the need to identify polymorphism or to analyse segregating populations. This will greatly accelerate cDNA mapping which will, in turn, help in the identification of candidate genes for mapped traits.

4.10. Comparative Genetics in Plants as a Whole

The above sections concern gene conservation between related species. It is also important to consider what proportion of genes might be common

to plants as a whole. It is not straightforward to determine what this proportion might be, especially from comparisons of nucleotide sequence data, which may be affected by factors such as codon usage. For example, monocotyledons tend to have a more GC-rich codon usage than dicotyledons (Murray *et al.*, 1989). A more effective strategy is to compare the predicted amino-acid sequences of known function or random cDNA clones. This routinely detects significant homologies, and it is expected that a proportion of genes will be highly conserved. Examples include essential 'housekeeping' genes, cell-cycle genes and genes controlling recombination or photosynthesis.

In many cases, however, homology is restricted to part of a cloned sequence, indicating a conserved protein domain, and the biological function of the gene may remain obscure. This is typically the case for members of gene families. In such cases, homologies may be detected that span widely diverged plant groups or even the plant and animal kingdoms. For example, transcription factors of the MADS-box (Schwarz-Sommer *et al.*, 1990; Schmidt *et al.*, 1993; Purugganan *et al.*, 1995) and homeobox (Vollbrecht *et al.*, 1991; Duboule, 1994) gene families have been identified in species as diverged as yeast, *Drosophila*, plants and humans. Such gene families probably evolved early in eukaryote evolution and subsequently were recruited for many purposes by duplication and divergence.

Despite these complications, it should be possible to clone many genes from crop plants using information from other species. Furthermore, this is likely to become progressively easier as strategies for exploiting comparative genetics become refined. One of the major problems will be the need to access and collate large amounts of information from different species. This will only be possible by developing comparative databasing facilities on computer networks with sophisticated search-and-compare programs.

4.11. Comparative Mapping and Germplasm Utilization

Comparative trait mapping suggests that sets of key genes can be recognized that are predicted to have important effects on plant phenotype in many species. Genes for photoperiod response, general pathogen defence or abiotic stress response are examples. Once mapped, markers for these genes could be used to screen germplasm for intraspecific variation and for variation in the equivalent chromosome regions of other species. Where cloned genes are available, variation in the gene itself could be assayed by PCR. However, germplasm assessment requires far more than the analysis of 'key gene sets' and it is necessary to develop efficient methods for identifying, quantifying and making available the widest possible range of useful variation.

It is therefore important to consider what 'useful variation' might

mean. Genome studies, comparative mapping, and sequence database comparisons suggest that a large proportion of genes are common to many if not all plant species and that related species must therefore share a very high proportion of their genetic make-up. This suggests that biodiversity is essentially due to allelic variation rather than to the creation of novel genes. For example, the distinct morphological differences between maize and teosinte are probably due to allelic differences at a relatively small number of genes (Dorweiler *et al.*, 1993; Doebley *et al.*, 1994). Novel interactions between alleles might also be major determinants of species differences. If allelic variation is the major source of variation in germplasm collections, how can it be assayed effectively?

Relatively few genes will show easily identifiable allelic variation in any given cross. Thus, the number of accessions that can be analysed genetically is limited. Most researchers therefore feel that molecular methods that fingerprint accessions are needed to complement phenotypic analysis. Such methods would also be valuable for the development of core collections that represent the range of genetic variation and exclude duplicate, or closely related, samples (see Chapter 5).

A property of any molecular technique used for the evaluation of genetic resources is that the markers provide reliable measures of genetic variability. There must be a high correlation between variability at the molecular level and phenotypic variability since there is an implicit assumption that molecular markers can be used as indicators of genetic variability between accessions without the need for extensive crossing and mapping. To date, there is little evidence that this is so. Furthermore, until very recently, diversity studies using genetic fingerprinting methods such as RAPD have adopted a black box approach with the implicit assumption that the markers used are randomly distributed throughout the genome and do not, for example, target heterochromatin. Few attempts have been made to map RAPD markers from diversity studies, so there is little knowledge of the degree of genome coverage.

Ideally, a map-based approach should be used so that the molecular markers systematically scan the entire genome, or at least a large part of it, in order to provide an overall measure of genetic diversity. A map-based approach would also enable the genome to be partitioned into domains, giving a polymorphism profile along the chromosomes which should reflect the degree of genetic and phenotypic diversity. By selecting variation at spaced markers it may be possible to preserve polymorphism at intervening loci by linkage, especially in inbreeding species where linkage disequilibrium may result in fixed associations between markers and QTL through founder effects (see also Chapter 5).

It is theoretically possible, using PCR marker technology such as AFLP, to fingerprint a collection of material and simultaneously assess the level of polymorphism across the genome by relation to markers mapped

in one or more segregating populations. In practice it remains to be proven to what extent markers with dominant modes of inheritance, such as RAPD and AFLP, can be used in this way. RFLPs, though cumbersome to use, have the advantage of 'transportability', due to their usually codominant inheritance and their strong tendency to transcend species barriers. PCR primers derived from the cloned DNA fragments used as RFLP probes (sequence-tagged-site or STS markers) provide a valuable compromise between PCR and RFLP methods (Inoue *et al.*, 1994).

PCR methods will probably become the methods of choice because of their high throughput and requirement for small amounts of starting material (Rafalski and Tingey, 1993). PCR methods also have the distinct advantage of allowing the retrieval of the amplified sequence for cloning and sequencing. PCR-based approaches may therefore be especially valuable for comparative analysis of specific genic regions. Conventional PCR can only amplify target sequences of limited size, but recent developments in 'long PCR' allow amplification of larger sequences (Barnes, 1994). This may allow the isolation of coding sequences of genes from different species together with their promoters and other flanking regions.

Whatever the methods used, a map-based screening method would result in more rational decision-making concerning the maintenance of genetic polymorphism within a collection. In grasses, the use of markers to detect specific genetic 'building blocks' is a logical approach to germplasm screening, since these appear to form discrete segments of genomes. If so, markers for specific segments may be useful even in species for which detailed maps have not been developed since there would be a high likelihood of good coverage of the genome.

4.12. Genome Relationships and Wide-hybridization

Domestication of crops is likely to have captured only part of the allele diversity present in the wild progenitor gene pools, leaving many desirable traits absent from the crop. In some cases, such as barley $\times$ *Hordeum spontaneum* or maize $\times$ teosinte, crops can be crossed easily with wild relatives and there are few barriers to the incorporation of novel alleles since the genomes are essentially homologous. In contrast, recently evolved polyploids such as hexaploid wheat will have had few opportunities to enlarge the gene pool. This is reflected by RFLP studies of wheat, barley and maize which show that wheat is less polymorphic (Chao *et al.*, 1989).

The frequency with which allopolyploid species occur in nature shows that the processes of pollen germination, correct pollen tube growth, fertilization of the egg cell, and embryo development are often conserved in related species. This can be exploited to synthesize new combinations of genomes. Such interspecific or intergeneric hybrids are

often referred to as wide-hybrids or wide-crosses. Wide-hybrids can be used to enlarge crop gene pools for any character but are most valuable where suitable genetic variation does not exist within a species, or where screening for a desired character would be difficult, time consuming or expensive. Disease resistance or abiotic stress tolerances are the characters most frequently targeted for transfer.

4.12.1. Barriers to wide-hybridization

The ease of producing wide-hybrids, and the ease with which genetic material may be transferred by backcrossing to a crop parent, varies greatly from cross to cross. Successful gene transfer may be restricted by prefertilization, postfertilization or chromosome instability barriers. These are not necessarily directly related to the degree of genetic divergence between two species and the genetical control of hybridization barriers may be simple.

Prefertilization barriers usually involve failure of pollen germination or inhibition of pollen tube growth, which may be at the surface of the stigma, in the stigma, or in the transmitting tissue of the ovary wall. The best studied incompatibility reactions affecting cereal wide-hybrids are due to the *Kr* genes of wheat where dominant alleles inhibit the growth of pollen tubes of other species at the base of the style. These genes are best studied in crosses with rye and *Hordeum bulbosum* (Jalani and Moss, 1980; Sitch and Snape, 1987). No method is known for overcoming the effect of these genes, so wide-hybrids involving bread or durum wheat usually use genotypes with recessive, noninhibiting alleles. In grasses there are, of course, well-characterized genes for self-incompatibility (the *S* and *Z* loci) (Franklin *et al.*, 1995) but it is not known whether there is homology between these and the *Kr* loci or whether a similar mechanism of inhibition operates.

Postfertilization barriers leading to seed abortion are common in wide-hybrids. In cereals, this is frequently due to defective development and subsequent collapse of the endosperm. Embryos, however, are often viable, and hybrid plants can be recovered if the embryos are excised from developing seeds and cultured on suitable nutrient agar. In the *Triticeae*, using these techniques, a huge catalogue of wide-hybrids have been produced (Maan and Gordon, 1988). For example, hexaploid, tetraploid and diploid wheats and their wild relatives can be intercrossed but embryo rescue is usually unavoidable, especially if different ploidy levels are crossed. Similarly, ovule or embryo culture has been used to recover hybrid plants from *Brassica* crosses (Diederichsen and Sacristan, 1994).

In some crosses, for example between bread and durum wheats, hybrid necrosis kills the developing seedling. This is due to the epistatic

interaction of independent genes in the respective species. Crosses within the separate gene pools are unaffected, probably because selection has eliminated necrotic combinations of alleles. Incompatibilities between nuclear genomes and the cytoplasm donated by the female parent may also contribute to lethality or sterility in certain crosses (Anderson and Maan, 1995).

In vitro fertilization using isolated gametes, and the subsequent recovery of plants, is a means by which prefertilization and some postfertilization barriers might be overcome. Currently, *in vitro* fertilization has only been used successfully in maize (Kranz and Lörz, 1993; Faure *et al.*, 1993, 1994), but such systems have potential for wide-crosses.

A further barrier to the production of wide-hybrids is the phenomenon of chromosome elimination. In such cases fertilization results in the formation of a hybrid zygote, but the complete chromosome complement of one of the parents is lost during subsequent mitotic divisions, leading to the production of a haploid embryo. The best characterized system is the cross between barley and its wild relative *Hordeum bulbosum* (Bennett *et al.*, 1976). However, the phenomenon occurs in a wide range of other cross combinations in the *Triticeae* including wheat $\times$ *H. bulbosum* (Barclay, 1975; Sitch and Snape, 1986), wheat $\times$ barley and wheat $\times$ maize, sorghum or pearl millet (Laurie and Bennett, 1987, 1988, 1989; Laurie, 1989). The mechanism of chromosome elimination is not understood but involves the inability of the chromosomes of one genome to successfully attach to the mitotic spindle. In some crosses, the occurrence of chromosome elimination is affected by the parental genomes since some barley $\times$ *H. bulbosum* crosses produce a high frequency of barley haploids while others give a high frequency of relatively stable hybrids (Pickering, 1983). The genetical control of this regulation is probably fairly simple (Ho and Kasha, 1975). Although chromosome elimination is detrimental to gene transfer, efficient haploid production methods benefit breeding and genetic analysis programmes and are extensively used by workers in these areas.

4.12.2. Utilization of hybrid plants

Once hybrid plants have been produced there are usually problems of sterility due to the inability of the hybrid to form viable gametes, particularly male gametes. Usually, this is due to a lack of chromosome pairing and chiasma formation at meiosis between the respective genomes and the subsequent formation of unbalanced gametes from random combinations of univalents. This can occur even in cases where close homoeology between the genomes can be shown by comparative chromosome and genetic analysis. Clearly the mechanisms controlling the pairing of chromosomes can be so precise as to prevent homoeologous pairing between

different genomes. Such mechanisms may be genetically simple, as in wheat where pairing between the A, B and D genomes of bread wheat is restricted by a few pairing control genes of which the most powerful is *Ph1* on chromosome 5B (Riley and Chapman, 1958). Fertility can be partially or fully restored in many hybrids by doubling the chromosome number with colchicine or an equivalent spindle inhibitor as this provides identical pairing partners for each chromosome. In such cases new allopolyploid 'species' can be established.

The major motivation for creating wide-hybrids is the introgression of useful genes into cultivated species. Generally, there are three levels at which the transfer of genetic material can be considered. Firstly, a whole genome can be introduced into an existing crop species to create a new allopolyploid species. For example, the introduction of a diploid rye genome into durum or bread wheat to create hexaploid or octaploid triticales, respectively, has been highly successful. New genome combinations such as *Tritordeum* (the amphiploid between wheat $\times$ *Hordeum chilense*) are also being investigated as crops for specific environments.

Secondly, whole chromosomes can be introduced to create substitution or addition lines. Thirdly, translocations can be induced between homoeologous or nonhomoeologous chromosomes so that only individual arms or small segments are introduced. The starting point for introgression is to backcross a wide-hybrid to the cultivated species. In many cases this is possible using the hybrid as the female parent, even if the hybrid is highly sterile, since a proportion of egg cells contain a nonreduced or otherwise viable chromosome complement. If there is homoeologous pairing between genomes, translocation products can be identified using *in situ* hybridization or molecular markers and stabilized by further backcrossing. Backcrosses can also be tested for the presence of the desired trait or traits in the 'alien' chromosome segment. Combinations of cytological and marker techniques can then be used to characterize suitable lines.

In some hybrids pairing between homoeologous genomes does not occur, or occurs too infrequently to be useful. In wheat, this can be overcome by using the *ph1* mutant or plants lacking chromosome 5B (carrying the *Ph1* pairing control gene). This allows higher levels of homoeologous chromosome pairing. Alternatively radiation treatments can be used to induce translocations. In such circumstances it is usually more effective to develop single chromosome addition or substitution lines prior to selecting recombinants, particularly as such stocks can confirm that the desired trait is carried by a single chromosome.

Plant breeders are generally interested in maintaining the characteristics of the cultivated species while adding one or more useful genes from the alien source. Generally the aim is to reduce the size of the alien chromosome segment as much as possible since donor species may have many alleles that are undesirable in the crop. This is greatly assisted by the use of

molecular markers which define the alien segment. *In situ* hybridization using total genomic DNA or species-specific repeats is also valuable because it reveals the number and size of the alien segments (Schwarzacher *et al.*, 1992). *In situ* methods are also increasingly able to locate low-copy sequences on plant chromosomes (Abbo *et al.*, 1993; Lehfer *et al.*, 1993; Leitch and Heslop-Harrison, 1993; Jiang and Gill, 1994), thereby linking the molecular and physical mapping.

Comparative mapping information is important for the manipulation of alien chromosome segments because of the importance of producing plants with genetically balanced genomes. For example, substitution of rye chromosome 1 for a wheat group 1 chromosome is successful because these chromosomes show conserved synteny throughout their length. In contrast, rye chromosome 7 is translocated with respect to the wheat chromosomes of groups 2, 4 and 5 (Devos *et al.*, 1993) and substitution of rye chromosome 7 for a wheat group 7 chromosome produces a plant with reduced doses of group 7 segments and extra doses of group 2, group 4 and group 5 segments. Comparative mapping therefore provides a framework for the assembly of new chromosome segments in genetically balanced combinations. This principle also applies to other crop groups such as *Brassica*.

4.13. Conclusion

Molecular analyses of plant genomes enable the genetical control of phenotype to be understood and can assist the assessment and efficient utilization of germplasm. However, molecular analysis must be seen as complementary to the analysis of phenotypes in germplasm collections, not as a replacement. It is useful to have described germplasm collections, but the best long-term policy would be to preserve habitats. This will ensure that a wide range of plant genotypes are maintained, irrespective of whether they are currently thought to be useful.

References

- Abbo, S., Dunford, R.P., Miller, T.E., Reader, S.M. and King, I.P. (1993) Primer mediated *in situ* detection of the B-hordein gene cluster on barley chromosome 1H. *Proceedings of the National Academy of Sciences of the USA* 90, 11821–11824.
- Ahn, S. and Tanksley, S.D. (1993) Comparative linkage maps of the rice and maize genomes. *Proceedings of the National Academy of Sciences of the USA* 90, 7980–7984.
- Anderson, J.A. and Maan, S.S. (1995) Interspecific nuclear-cytoplasmic compatibility controlled by genes on group 1 chromosomes in durum wheat. *Genome* 38, 803–808.
- Barclay, I.R. (1975) High frequencies of haploid production in wheat (*Triticum aestivum* L.). *Genetics* 79, 1–12.

- tivum*) by chromosome elimination. *Nature* 256, 410–411.
- Barnes, W.M. (1994) PCR amplification of up to 35-kb DNA with high fidelity and high yield from λ bacteriophage templates. *Proceedings of the National Academy of Sciences of the USA* 91, 5695–5699.
- Beavis, W.D., Grant, D., Albertsen, M. and Fincher, R. (1991) Quantitative trait loci for plant height in four maize populations and their association with qualitative genetic loci. *Theoretical and Applied Genetics* 83, 141–145.
- Bennett, M.D. (1976) DNA amount, latitude and crop plant distribution. In: Jones, K. and Brandham, P.E. (eds) *Current Chromosome Research*. Elsevier / North Holland Biomedical Press, Amsterdam, pp. 151–158.
- Bennett, M.D. (1987) Variation in genomic form in plants and its ecological implications. *New Phytologist* 106, 177–200.
- Bennett, M.D. and Leitch, I.J. (1995) Nuclear DNA amounts in angiosperms. *Annals of Botany* 76, 113–176 (includes references to three additional lists of DNA amounts).
- Bennett, M.D., Finch, R.A. and Barclay, I.R. (1976) The time rate and mechanism of chromosome elimination in *Hordeum* hybrids. *Chromosoma* 54, 175–200.
- Bennett, S.T., Kenton, A.Y. and Bennett, M.D. (1992) Genomic *in situ* hybridization reveals the allopolyploid nature of *Milium montianum* (Gramineae). *Chromosoma* 101, 420–424.
- Bennetzen, J.L. and Freeling, M. (1993) Grasses as a single genetic system: genome composition, colinearity and compatibility. *Trends in Genetics* 9, 259–261.
- Bensen, R.J., Johal, G.S., Crane, V.C., Tossberg, J.T., Schnable, P.S., Meeley, R.B. and Briggs, S.P. (1995) Cloning and characterization of the maize *An1* gene. *Plant Cell* 7, 75–84.
- Binelli, G. and Bucci, G. (1994) A genetic linkage map of *Picea abies* Karst., based on RAPD markers as a tool in population genetics. *Theoretical and Applied Genetics* 88, 283–288.
- Bonierbale, M.W., Plaisted, R.L. and Tanksley, S.D. (1988) RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. *Genetics* 120, 1095–1103.
- Chao, S., Sharp, P.J., Worland, A.J., Warham, E.J., Koebner, R.M.D. and Gale, M.D. (1989) RFLP-based genetic maps of wheat homoeologous group 7 chromosomes. *Theoretical and Applied Genetics* 78, 495–504.
- Chittenden, L.M., Schertz, K.F., Lin, Y.R., Wing, R.A. and Paterson, A.H. (1994) A detailed RFLP map of *Sorghum bicolor* $\times$ *S. propinquum*, suitable for high-density mapping, suggests ancestral duplication of *Sorghum* chromosomes or chromosomal segments. *Theoretical and Applied Genetics* 87, 925–933.
- Deshpande, V.G. and Ranjekar, P.K. (1980) Repetitive DNA in three Gramineae species with low DNA content. *Hoppe-Seyler's Zeitschrift für Physiologische Chemie* 361, 1223–1233.
- Devos, K.M. and Gale, M.D. (1993) The genetic maps of wheat and their potential in plant breeding. *Outlook in Agriculture* 22, 93–97.
- Devos, K.M., Atkinson, M.D., Chinoy, C.N., Francis, H.A., Harcourt, R.L., Koebner, R.M.D., Liu, C.J., Masojc, P., Xie, D.X. and Gale, M.D. (1993) Chromosomal rearrangements in the rye genome relative to that of wheat. *Theoretical and Applied Genetics* 85, 673–680.
- Devos, K.M., Chao, S., Li, Q.Y., Simonetti, M.C. and Gale, M.D. (1994) Relationship

- between chromosome 9 of maize and wheat homoeologous group 7 chromosomes. *Genetics* 138, 1287–1292.
- Diederichsen, E. and Sacristan, M.D. (1994) The use of ovule culture in reciprocal hybridization between *B. campestris* L. and *B. oleracea* L. *Plant Breeding* 113, 79–82.
- Doebley, J., Bacigalupo, A. and Stec, A. (1994) Inheritance of kernel weight in two maize–teosinte hybrid populations: implications for crop evolution. *Journal of Heredity* 85, 191–195.
- Dorweiler, J., Stec, A., Kermicle, J. and Doebley, J. (1993) Teosinte glume architecture 1: a genetic locus controlling a key step in maize evolution. *Science* 262, 233–240.
- Dover, G.A. (1993) Evolution of genetic redundancy for advanced players. *Current Opinion in Genetics and Development* 3, 902–910.
- Dubcovsky, J. and Dvořák, J. (1995) Ribosomal RNA multigene loci: Nomads of the Triticeae genomes. *Genetics* 140, 1367–1377.
- Duboule, D. (1994) *Guidebook to the Homeobox Genes*. Oxford University Press, New York, 284 pp.
- Dunford, R.P., Kurata, N., Laurie, D.A., Money, T.A., Minobe, Y. and Moore, G. (1995) Conservation of fine-scale DNA marker order in the genomes of rice and the Triticeae. *Nucleic Acids Research* 23, 2724–2728.
- Faure, J.E., Mogensen, H.L., Dumas, C., Lörz, H. and Kranz, E. (1993) Karyogamy after electrofusion of single egg and sperm cell protoplasts from maize: cytological evidence and time course. *Plant Cell* 5, 747–755.
- Faure, J.E., Digonnet, C. and Dumas, C. (1994) An *in vitro* system for adhesion and fusion of maize gametes. *Science* 263, 1598–1600.
- Ferreira, M.E., Williams, P.H. and Osborn, T.C. (1994) RFLP mapping of *Brassica napus* using doubled haploid lines. *Theoretical and Applied Genetics* 89, 615–621.
- Flavell, R.B. (1986) Repetitive DNA and chromosome evolution in plants. *Philosophical Transactions of the Royal Society of London Series B*, 312, 227–242.
- Flavell, R.B. and Smith, D.B. (1976) Nucleotide sequence organisation in the wheat genome. *Heredity* 37, 231–252.
- Franklin, F.C.H., Lawrence, M.J. and Franklin-Tong, V.E. (1995) Cell and molecular biology of self-incompatibility in flowering plants. *International Review of Cytology* 158, 1–64.
- Fukuchi, A., Kikuchi, F. and Hirochika, H. (1993) DNA fingerprinting of cultivated rice with rice retrotransposon probes. *Japanese Journal of Genetics* 68, 195–204.
- Galiba, G., Quarrie, S.A., Sutka, J., Morgounov, A. and Snape, J.W. (1995) RFLP mapping of the vernalization (*Vrn1*) and frost resistance (*Fr1*) genes on chromosome 5A of wheat. *Theoretical and Applied Genetics* 90, 1174–1179.
- Gebhardt, C. and Salamini, F. (1992) Restriction fragment length polymorphism analysis of plant genomes and its application to plant breeding. *International Review of Cytology* 135, 210–237.
- Gebhardt, C., Ritter, E., Barone, A., Debener, T., Walkemeier, B., Schachtschabel, U., Kaufmann, H., Thompson, R.D., Bonierbale, M.W. and Ganal, M.W. (1991) RFLP maps of potato and their alignment with the homoeologous tomato genome. *Theoretical and Applied Genetics* 83, 49–57.
- Grattapaglia, D. and Sederoff, R. (1994) Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. *Genetics* 137, 1121–1137.

- Helentjaris, T., Weber, D. and Wright, S. (1988) Identification of the genomic locations of duplicate nucleotide sequences in maize by analysis of restriction fragment length polymorphisms. *Genetics* 118, 353–363.
- Heslop-Harrison, J.S., Leitch, A.R., Schwarzacher, T. and Anamthawat-Jónsson, K. (1990) Detection and characterization of 1B/1R translocations in hexaploid wheat. *Heredity* 65, 385–392.
- Heslop-Harrison, J.S. and Schwarzacher, T. (1993) Molecular cytogenetics – biology and applications in plant breeding. In: Sumner, A.T. and Chandley, A.C. (eds) *Chromosomes Today, Vol. 11*. Chapman & Hall, London, pp. 191–198.
- Heun, M., Kennedy, A.E., Anderson, J.A., Lapitan, N.L.V., Sorrells, M.E. and Tanksley, S.D. (1991) Construction of a restriction fragment length polymorphism map for barley (*Hordeum vulgare*). *Genome* 34, 437–447.
- Ho, K.M. and Kasha, K.J. (1975) Genetic control of chromosome elimination during haploid formation in barley. *Genetics* 81, 263–275.
- Hulbert, S.H., Richter, T.E., Axtell, J.D. and Bennetzen, J.L. (1990) Genetic mapping and characterization of sorghum and related crops by means of maize DNA probes. *Proceedings of the National Academy of Sciences of the USA* 87, 4251–4255.
- Inoue, T., Zhong, H.S., Miyao, A., Ashikawa, I., Monna, L., Fukuoka, S., Miyadera, N., Nagamura, Y., Kurata, N., Sasaki, T. and Minobe, Y. (1994) Sequence-tagged sites (STSs) as standard landmarks in the rice genome. *Theoretical and Applied Genetics* 89, 728–734.
- Islam-Faridi, M.N. and Mujeeb-Kazi, A. (1995) Visualization of *Secale cereale* DNA in wheat germ plasm by fluorescent *in situ* hybridization. *Theoretical and Applied Genetics* 90, 595–600.
- Jalani, B.S. and Moss, J.P. (1980) The site of action of crossability genes (*Kr1*, *Kr2*) between *Triticum* and *Secale*. I. Pollen germination, pollen tube growth, and number of pollen tubes. *Euphytica* 29, 571–579.
- Jansen, R.C. (1994) Controlling type I and type II errors in mapping quantitative trait loci. *Genetics* 138, 871–881.
- Jena, K.K., Khush, G.S. and Kochert, G. (1994) Comparative RFLP mapping of a wild rice, *Oryza officinalis*, and cultivated rice, *O. sativa*. *Genome* 37, 382–389.
- Jiang, J. and Gill, B.S. (1994) Nonisotopic *in situ* hybridization and plant genome mapping: the first 10 years. *Genome* 37, 717–725.
- Kearsey, M.J. and Hyne, V. (1994) QTL analysis: a simple marker regression approach. *Theoretical and Applied Genetics* 89, 698–702.
- Kianian, S.F. and Quiros, C.F. (1992) Generation of a *Brassica oleracea* composite RFLP map: linkage arrangements among various populations and evolutionary implications. *Theoretical and Applied Genetics* 84, 544–554.
- Kilian, A., Kudrna, D.A., Kleinhofs, A., Yano, M., Kurata, N., Steffenson, B. and Sasaki, T. (1995) Rice–barley synteny and its application to saturation mapping of the barley *Rpg1* region. *Nucleic Acids Research* 23, 2729–2733.
- Kimber, G. and Yen, Y. (1988) Analysis of pivotal-differential evolutionary patterns. *Proceedings of the National Academy of Sciences of the USA* 85, 9106–9108.
- Kishimoto, N., Higo, H., Abe, K., Arai, S., Saito, A. and Higo, K. (1994) Identification of the duplicated segments in rice chromosomes 1 and 5 by linkage analysis of cDNA markers of known functions. *Theoretical and Applied Genetics* 88, 722–726.

- Kowalski, S.P., Lan, T.H., Feldmann, K.A. and Paterson, A.H. (1994) Comparative mapping of *Arabidopsis thaliana* and *Brassica oleracea* chromosomes reveals islands of conserved organization. *Genetics* 138, 499–510.
- Kranz, E. and Lörz, H. (1993) *In vitro* fertilization with isolated, single gametes results in zygotic embryogenesis and fertile maize plants. *Plant Cell* 5, 739–746.
- Kurata, N., Moore, G., Nagamura, Y., Foote, T., Yano, M., Minobe, Y. and Gale, M.D. (1994a) Conservation of genome structure between rice and wheat. *Bio/Technology* 12, 276–278.
- Kurata, N., Nagamura, Y., Yamamoto, K., Harushima, Y., Sue, N., Wu, J., Antonio, B.A., Shomura, A., Shimizu, T., Lin, S.Y., Inoue, T., Fukuda, A., Shimano, T., Kuboki, Y., Toyama, T., Miyamoto, Y., Kirihara, T., Hayasaka, K., Miyao, A., Monna, L., Zhong, H.S., Tamura, Y., Wang, Z.X., Momma, T., Umehara, Y., Yano, M., Sasaki, T. and Minobe, Y. (1994b) A 300 kilobase interval genetic map of rice including 883 expressed sequences. *Nature Genetics* 8, 365–372.
- Lander, E.S. and Botstein, D. (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121, 185–199.
- Lapitan, N.V.L., Sears, R.G., Rayburn, A.L. and Gill, B.S. (1986) Wheat–rye translocations. *Journal of Heredity* 77, 415–419.
- Laurie, D.A. (1989) The frequency of fertilization in wheat $\times$ pearl millet crosses. *Genome* 32, 1063–1067.
- Laurie, D.A. and Bennett, M.D. (1987) The effect of the crossability loci *Kr1* and *Kr2* on fertilization frequency in hexaploid wheat $\times$ maize crosses. *Theoretical and Applied Genetics* 73, 403–409.
- Laurie, D.A. and Bennett, M.D. (1988) Cytological evidence for fertilization in hexaploid wheat $\times$ sorghum crosses. *Plant Breeding* 100, 73–82.
- Laurie, D.A. and Bennett, M.D. (1989) The timing of chromosome elimination in hexaploid wheat $\times$ maize crosses. *Genome* 32, 953–961.
- Laurie, D.A., Pratchett, N., Bezant, J.H. and Snape, J.W. (1994) Genetic analysis of a photoperiod response gene on the short arm of chromosome 2(2H) of *Hordeum vulgare* (barley). *Heredity* 72, 619–627.
- Laurie, D.A., Pratchett, N., Bezant, J.H. and Snape, J.W. (1995) RFLP mapping of five major genes and eight quantitative trait loci controlling flowering time in a winter $\times$ spring barley (*Hordeum vulgare* L.) cross. *Genome* 38, 575–585.
- Law, C.N., Snape, J.W. and Worland, A.J. (1987) Aneuploidy in wheat and its uses in genetic analysis. In: Lupton, F.G.H. (ed.) *Wheat Breeding: Its Scientific Basis*. Chapman & Hall, London, pp. 71–108.
- Lehfer, H., Busch, W., Martin, R. and Herrmann, R.G. (1993) Localization of the B-hordein locus on barley chromosomes using fluorescence *in situ* hybridization. *Chromosoma* 102, 428–432.
- Leitch, I.J. and Heslop-Harrison, J.S. (1992) Physical mapping of the 18S-5.8S-26S rRNA genes in barley by *in situ* hybridization. *Genome* 35, 1013–1016.
- Leitch, I.J. and Heslop-Harrison, J.S. (1993) Physical mapping of four sites of 5S rDNA sequences and one site of the α -amylase-2 gene in barley (*Hordeum vulgare*). *Genome* 36, 517–523.
- Leitch, I.J., Leitch, A.R. and Heslop-Harrison, J.S. (1991) Physical mapping of plant DNA sequences by simultaneous *in situ* hybridization of two differently labelled fluorescent probes. *Genome* 34, 329–333.
- Litt, M. and Luty, J.A. (1989) A hypervariable microsatellite revealed by *in vitro*

- amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American Journal of Human Genetics* 44, 397–401.
- Maan, S.S. and Gordon, J. (1988) Compendium of alloplasmic lines and amphiploids in the Triticeae. In: Miller, T.E. and Koebner, R.M.D. (eds) *Proceedings of the Seventh International Wheat Genetics Symposium*. Institute of Plant Science Research, Cambridge, pp. 1325–1369.
- McCarthy, L., Hunter, K., Schalkwyk, L., Riba, L., Anson, S., Mott, R., Newell, W., Bruley, C., Bar, I., Ramu, E., Housman, D., Cox, R. and Lehrach, H. (1995) Efficient and high-resolution genetic mapping of mouse interspersed repetitive sequence PCR products, toward integrated genetic and physical mapping of the mouse genome. *Proceedings of the National Academy of Sciences of the USA* 92, 5302–5306.
- McMurphy, L.M. and Rayburn, A.L. (1992) Chromosomal and cell size analysis of cold tolerant maize. *Theoretical and Applied Genetics* 84, 798–802.
- McNeil, D., Lagudah, E.S., Hohmann, U. and Appels, R. (1994) Amplification of DNA sequences in wheat and its relatives of the Dgas44 and R350 families of repetitive sequences. *Genome* 37, 320–327.
- Martin, G.B., Brommonschenkel, S.H., Chunwongse, J., Frary, A., Ganai, M.W., Spivey, R., Wu, T., Earle, E.D. and Tanksley, S.D. (1993) Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science* 262, 1432–1436.
- Melake Berhan, A., Hulbert, S.H., Butler, L.G. and Bennetzen, J.L. (1993) Structure and evolution of the genomes of *Sorghum bicolor* and *Zea mays*. *Theoretical and Applied Genetics* 86, 598–604.
- Menancio-Hautea, D., Fatokun, C.A., Kumar, L., Danesh, D. and Young, N.D. (1993) Comparative genome analysis of mungbean (*Vigna radiata* L. Wilczek) and cowpea (*V. unguiculata* L. Walpers) using RFLP mapping data. *Theoretical and Applied Genetics* 86, 797–810.
- Moore, G., Cheung, W., Schwarzacher, T. and Flavell, R. (1991) BIS 1, a major component of the cereal genome and a tool for studying genomic organization. *Genomics* 10, 469–476.
- Moore, G., Gale, M.D., Kurata, N. and Flavell, R.B. (1993a) Molecular analysis of small grain cereal genomes: current status and prospects. *Bio/Technology* 11, 584–589.
- Moore, G., Abbo, S., Cheung, W., Foote, T., Gale, M.D., Koebner, R.M.D., Leitch, A., Leitch, I., Money, T., Stancombe, P., Yano, M. and Flavell, R. (1993b) Key features of cereal genome organization as revealed by the use of cytosine methylation-sensitive restriction endonucleases. *Genomics* 15, 472–482.
- Moore, G., Devos, K.M., Wang, Z. and Gale, M.D. (1995) Grasses, line up and form a circle. *Current Biology* 5, 737–739.
- Muehlbauer, F.J., Kaiser, W.J., Clement, S.L. and Summerfield, R.J. (1995) Production and breeding of lentil. *Advances in Agronomy* 54, 283–332.
- Mukai, Y. and Gill, B.S. (1991) Detection of barley chromatin added to wheat by genomic *in situ* hybridization. *Genome* 34, 448–452.
- Mukai, Y., Nakahara, Y. and Yamamoto, M. (1993) Simultaneous discrimination of the three genomes in hexaploid wheat by multicolor fluorescence *in situ* hybridization using total genomic and highly repeated DNA probes. *Genome* 36, 489–494.

- Murray, E.E., Lotzer, J. and Eberle, M. (1989) Codon usage in plant genes. *Nucleic Acids Research* 17, 477–498.
- Namuth, D.M., Lapitan, N.V.L., Gill, K.S. and Gill, B.S. (1994) Comparative RFLP mapping of *Hordeum vulgare* and *Triticum tauschii*. *Theoretical and Applied Genetics* 89, 865–872.
- Nederlof, P.M., Robinson, D., Abuknesha, R., Wiegant, J., Hopman, A.H.N., Tanke, H.J. and Raap, A.K. (1989) Three-color fluorescence *in situ* hybridization for the simultaneous detection of multiple nucleic acid sequences. *Cytometry* 10, 20–27.
- Orgaard, M. and Heslop-Harrison, J.S. (1994) Investigations of genome relationships between *Leymus*, *Psathyrostachys* and *Hordeum* inferred by genomic DNA : DNA *in situ* hybridization. *Annals of Botany* 73, 195–203.
- Parkin, I.A.P., Sharpe, A.G., Keith, D.J. and Lydiate, D.J. (1995) Identification of the A and C genomes of amphidiploid *Brassica napus* (oilseed rape). *Genome* 38, 1122–1131.
- Paterson, A.H., Lin, R.Y., Li, Z., Schertz, K.F., Doebley, J.F., Pinson, S.R.M., Liu, S.C., Stansel, J.W. and Irvine, J.E. (1995) Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. *Science* 269, 1714–1718.
- Pereira, M.G. and Lee, M. (1995) Identification of genomic regions affecting plant height in sorghum and maize. *Theoretical and Applied Genetics* 90, 380–388.
- Pereira, M.G., Lee, M., Bramel-Cox, P., Woodman, W., Doebley, J. and Whitkus, R. (1994) Construction of an RFLP map in sorghum and comparative mapping in maize. *Genome* 37, 236–243.
- Pickering, R.A. (1983) The influence of genotype on doubled haploid barley production. *Euphytica* 32, 863–876.
- Plaschke, J., Börner, A., Xie, D.X., Koebner, R.M.D., Schlegel, R. and Gale, M.D. (1993) RFLP mapping of genes affecting plant height and growth habit in rye. *Theoretical and Applied Genetics* 85, 1049–1054.
- Plaschke, J., Korzun, V., Koebner, R.M.D. and Börner, A. (1995) Mapping the GA₃-insensitive dwarfing gene *ct1* on chromosome 7 in rye. *Plant Breeding* 114, 113–116.
- Pratchett, N. and Laurie, D.A. (1994) Genetic map location of the barley developmental mutant *liguleless* in relation to RFLP markers. *Hereditas* 120, 35–39.
- Purugganan, M.D., Roundsley, S.D., Schmidt, R.J. and Yanofsky, M.F. (1995) Molecular evolution of flower development: Diversification of the plant MADS-box regulatory gene family. *Genetics* 140, 345–356.
- Rafalski, J.A. and Tingey, S.V. (1993) Genetic diagnosis in plant breeding: RAPDs, microsatellites and machines. *Trends in Genetics* 9, 275–280.
- Rayburn, A.L., Price, H.J., Smith, J.D. and Gold, J.R. (1985) C-band heterochromatin and DNA content in *Zea mays*. *American Journal of Botany* 72, 1610–1617.
- Renwick, J.H. (1971) The mapping of human chromosomes. *Annual Review of Genetics* 5, 81–120.
- Riley, R. and Chapman, V. (1958) Genetic control of the cytologically diploid behaviour of hexaploid wheat. *Nature* 182, 713–715.
- Rimpau, J., Smith, D.B. and Flavell, R.B. (1980) Sequence organisation in barley and oat chromosomes revealed by interspecies DNA / DNA hybridisation. *Heredity* 44, 131–149.

- Rivin, C.J., Cullis, C.A. and Walbot, V. (1986) Evaluating quantitative variation in the genome of *Zea mays*. *Genetics* 113, 1009–1019.
- Robertson, D.S. (1985) A possible technique for isolating genic DNA for quantitative traits in plants. *Journal of Theoretical Biology* 117, 1–10.
- Rongwen, J., Akkaya, M.S., Bhagwat, A.A., Lavi, U. and Cregan, P.B. (1995) The use of microsatellite DNA markers for soybean genotype identification. *Theoretical and Applied Genetics* 90, 43–48.
- Schmidt, R.J., Veit, B., Mandel, M.A., Mena, M., Hake, S. and Yanofsky, M.F. (1993) Identification and molecular characterization of *ZAG1*, the maize homologue of the *Arabidopsis* floral homeotic gene, *AGAMOUS*. *Plant Cell* 5, 729–737.
- Schwarzacher, T., Leitch, A.R., Bennett, M.D. and Heslop-Harrison, J.S. (1989) *In situ* localization of parental genomes in a wide hybrid. *Annals of Botany* 64, 315–324.
- Schwarzacher, T., Ananthawat-Jónsson, K., Harrison, G.E., Islam, A.K.M.R., Jia, J.Z., King, I.P., Leitch, A.R., Miller, T.E., Reader, S.M., Rogers, W.J., Shi, M. and Heslop-Harrison, J.S. (1992) Genomic *in situ* hybridization to identify alien chromosomes and chromosome segments in wheat. *Theoretical and Applied Genetics* 84, 778–786.
- Schwarz-Sommer, Z., Huijser, P., Nacken, W., Saedler, H. and Sommer, H. (1990) Genetic control of flower development by homeotic genes in *Antirrhinum majus*. *Science* 250, 931–936.
- Shoemaker, R.C., Guffy, R.D., Lorenzen, L.L. and Specht, J.E. (1992) Molecular genetic mapping of soybean: Map utilization. *Crop Science* 32, 1091–1098.
- Sitch, L.A. and Snape, J.W. (1986) Doubled haploid production in winter wheat and triticale genotypes, using the *Hordeum bulbosum* system. *Euphytica* 35, 1045–1051.
- Sitch, L.A. and Snape, J.W. (1987) Factors affecting haploid production in wheat using the *Hordeum bulbosum* system. 1. Genotypic and environmental effects on pollen grain germination, pollen tube growth and the frequency of fertilization. *Euphytica* 36, 483–496.
- Takeda, K. and Zhang, C.L. (1994) Inheritance and geographical distribution of a phenol reaction-less mutant of barley. *Barley Genetics Newsletter* 24, 32–34.
- Tanksley, S.D., Ganai, M.W., Prince, J.P., de Vicente, M.C., Bonierbale, M.W., Broun, P., Fulton, T.M., Giovannoni, J.J., Grandillo, S. and Martin, G.B. (1992a) High density molecular linkage maps of the tomato and potato genomes. *Genetics* 132, 1141–1160.
- Tanksley, S.D., Causse, M., Fulton, T., Ahn, N., Wang, Z., Wu, K., Xiao, J., Yu, Z., Second, G. and McCouch, S. (1992b) A high density molecular map of the rice genome. *Rice Genetics Newsletter* 9, 111–115.
- Tanksley, S.D., Ganai, M.W. and Martin, G.B. (1995) Chromosome landing: a paradigm for map-based cloning in plants with large genomes. *Trends in Genetics* 11, 63–68.
- Teutonico, R.A. and Osborn, T.C. (1994) Mapping of RFLP and qualitative trait loci in *Brassica rapa* and comparison to the linkage maps of *B. napus*, *B. oleracea*, and *Arabidopsis thaliana*. *Theoretical and Applied Genetics* 89, 885–894.
- Thompson, J.N. (1975) Quantitative variation and gene number. *Nature* 258, 665–668.
- Truco, M.J. and Quiros, C.F. (1994) Structure and organization of the B genome

- based on a linkage map in *Brassica nigra*. *Theoretical and Applied Genetics* 89, 590–598.
- Vollbrecht, E., Veit, B., Sinha, N. and Hake, S. (1991) The developmental gene *Knotted-1* is a member of a maize homeobox gene family. *Nature* 350, 241–243.
- Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau M. (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23, 4407–4414.
- Walbot, V. (1992) Strategies for mutagenesis and gene cloning using transposon tagging and T-DNA insertional mutagenesis. *Annual Review of Plant Physiology and Plant Molecular Biology* 43, 49–82.
- Weber, J.L. and May, P.E. (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics* 44, 388–396.
- Weeden, N.F., Muehlbauer, F.J. and Ladizinsky, G. (1992) Extensive conservation of linkage relationships between pea and lentil genetic maps. *Journal of Heredity* 83, 123–129.
- Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. and Tingey, S.V. (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* 18, 6531–6535.
- Worland, A.J. (1996) The influence of flowering time genes on environmental adaptability in European wheats. *Euphytica* 89, 49–57.
- Xiao, J., Fulton, T., McCouch, S., Tanksley, S.D., Kishimoto, N., Ohsawa, R., Ukai, Y. and Saito, A. (1992) Progress in integration of the molecular maps of rice. *Rice Genetics Newsletter* 9, 124–128.
- Zhang, H.B. and Dvořák, J. (1990) Characterization and distribution of an interspersed repeated nucleotide sequence from *Lophopyrum elongatum* and mapping of a segregation distortion factor with it. *Genome* 33, 927–936.

Molecular Markers and the Management of Genetic Resources in Seed Genebanks: a Case Study of Rice

B.V. Ford-Lloyd¹, M.T. Jackson² and H.J. Newbury¹

¹*School of Biological Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK;* ²*International Rice Research Institute, P.O. Box 933, 1099 Manila, Philippines*

5.1. Introduction

Plant germplasm has been accumulated over many decades, and is now stored in genebanks in countries around the world. The United States National Plant Germplasm System reputedly holds more than 380 000 different accessions of some 8700 species of plant, within which, for example, there are more than 110 000 accessions of wild and cultivated cereals held in the Small Grains Collection in Idaho (National Research Council, 1991). The Consultative Group on International Agricultural Research (CGIAR) centres conserve about 500 000 germplasm samples of more than 30 crop species and wild relatives, mostly as *ex situ* seed collections, but also as field genebanks and *in vitro* collections. The International Rice Genebank at the International Rice Research Institute (IRRI) has a collection of more than 80 000 samples of rice alone from more than 110 countries (Jackson and Huggan, 1993) comprising landrace varieties, breeding lines and commercial varieties of *Oryza sativa*, landrace varieties of *O. glaberrima*, and all 20 wild species of the genus *Oryza*. Since 1973 over 740 000 packets of rice have been distributed throughout the world for use in basic and applied research, and this germplasm has contributed to improvements in many characteristics of new rice varieties (Jackson, 1994). Pressure for germplasm distribution will increase over the next 30 years as plant scientists strive to meet the demands for a 70% increase in rice production by the year 2025.

The management of such collections is difficult due simply to their vast size, and there is a clear requirement for the development of

procedures which utilize fast and reliable methods for the identification of material, the measurement of diversity, or for the determination of redundancy, in order to facilitate the organization and prioritization of germplasm (Virk *et al.*, 1995a). Even more important may be the need to rapidly and efficiently identify the most appropriate rice germplasm for research and crop improvement.

In recent years there has been an explosion of new DNA-based marker methods such as restriction fragment length polymorphism (RFLP) analysis, and those utilizing the polymerase chain reaction (PCR) such as RAPD (random amplification of polymorphic DNA) (Williams *et al.*, 1990). RAPD technology has been used successfully for measuring diversity in plants, and the patterns of variation observed have been shown to closely resemble those obtained using more classical characters (Howell *et al.*, 1994; Virk *et al.*, 1995a). A range of other PCR-based techniques suitable for the measurement of diversity have been developed over the last few years. Some of these target repeat regions of the genome to produce markers. Single primers complementary to minisatellite repeats provide multilocus markers (Matsuyama *et al.*, 1993; Neuhaus *et al.*, 1993). More widespread is the use of pairs of primers complementary to sequences flanking microsatellite repeats. This strategy produces single locus markers which, because of the hypervariability of microsatellite repeat lengths, is an efficient method for the detection of polymorphisms (Morgante and Olivieri, 1993; Senior and Heun, 1993; Cregan *et al.*, 1994). Yang *et al.* (1994) used PCR technology to identify microsatellite polymorphism across landraces and cultivars of rice, whilst Wu and Tanksley (1993) have reported the identification of microsatellite alleles that are specific to either indica or japonica rices. Recently, an additional PCR-based marker technique – amplified fragment length polymorphism (AFLP) analysis – has been developed, which results in the selective amplification of restriction fragments from within a total digest of genomic DNA to yield typically 50–100 dominant marker bands per polyacrylamide gel track (Vos *et al.*, 1995). This method seems certain to find wide application in the study of plant diversity.

Germplasm collections in many parts of the world, like that at IRRI, are faced with challenges related to the various activities which must take place within the genebank, and these are exacerbated by the size of some of the collections and the number of accessions which are conserved. It is now clear that some of the constraints could be alleviated by the application of the molecular marker technology which is becoming more and more readily available. Examples in which molecular markers may be suitably employed to assist genebank management, organization and the way that material is accessed include:

- The accurate identification of germplasm.
- The routine maintenance of germplasm, which is a continuous process

involving seed viability testing, rejuvenation and replenishment of stocks, which will be streamlined by the identification of duplicates and the development of core collections.

- The selection of germplasm for safety storage at other genebanks.
- The choice of germplasm for use by breeders and other researchers involved in making crosses, and mapping, identifying and isolating genes of interest.

Molecular markers are, of course, being used very successfully for the assessment of genetic diversity amongst genetic resources of an increasing number of crop species and wild relatives, and this may take place either before germplasm is accepted into a genebank for storage, or after storage has been initiated. This molecular assessment of 'biodiversity' is dealt with extensively in other chapters.

5.2. Operations Within a Seed Genebank

Confining the discussion only to the primary concerns of a genebank manager, we can identify four principal areas of activity. First is the acquisition of seeds, which involves receiving material as a result of exploration and collection activities or exchange, its preparation for storage, documentation and quarantine procedures. Secondly, seeds are conserved in the form of active/working or base collections, or duplicate safety collections. Thirdly, there are seed management activities including multiplication, regeneration, viability testing, characterization and seed distribution. Finally various activities may take place related to the utilization of germplasm such as detailed evaluation or enhancement.

Taking a more detailed example, the International Rice Genebank at IRRI has as a primary aim the conservation and continued availability of genetic resources for rice improvement worldwide. The germplasm is freely available on request, and is used to continually restore valuable material which has been lost in the country of origin. Base and active collections are maintained, and provision is also made for duplicate 'black box' storage of germplasm at the National Seed Storage Laboratory, Fort Collins, USA. Incoming germplasm is examined by the Seed Health Unit at IRRI, and viability testing is carried out. As and when necessary, germplasm is rejuvenated and multiplied to produce high-quality seed for long-term conservation (Kameswara Rao and Jackson, 1996a,b) and for distribution in response to requests. Wild species receive particular attention and are always grown in pots in a quarantine screenhouse, and perennial species are often maintained as living plants for prolonged periods if seeds are difficult to produce. Seed samples – 10 g each for cultivated rices, but only 10 seeds for the wild rices – are routinely sent out on

request, with more than 740 000 since 1973. When no specific samples are requested, judgements often have to be made as to which material is most appropriate and will most effectively satisfy the demands of the recipient.

No genebank should be a 'museum collection'. The material conserved must be characterized not only to distinguish species, taxonomic groupings and varieties, but also to facilitate preliminary selection of germplasm by end-users. To this end, morphological and agronomic characters are scored in small field plots at IRRI using standard descriptors which will allow the rational choice of material for exchange and distribution, while the maintenance of passport data permits selection of germplasm on an ecogeographical basis. In addition to such routine characterization, IRRI scientists have also screened thousands of accessions for resistance to pests and diseases, and tolerance for different abiotic stresses, an evaluation process which requires more or less elaborate testing of germplasm in the laboratory or field trials.

5.3. Taxonomic Identification as part of Germplasm Characterization

The accurate identification of material held in any genebank is arguably the most essential part of the germplasm characterization process, for without such information breeders will have no means of selecting material for crosses and entry into breeding programmes. Taxonomic identification is an essential first step to determine whether any germplasm is part of the primary, secondary or tertiary gene pool of the crop concerned. While such identification may be undertaken using traditional taxonomic characters, this is not always possible or indeed accurate. A very useful summary of examples of the way molecular markers have contributed to our understanding of crop gene pools is given by Gepts (1995).

In Asian rice (*Oryza sativa*), six crossability groups have been recognized comprising the bulk of the primary gene pool. It is of particular concern to breeders that unambiguous identification of indica and japonica rices can be achieved, as these are currently the focus of plant breeders' attention for crossing and for the development of the 'new plant type', despite the fact that there is difficulty in hybridization and recombination between these two types. Breeders attempting to utilize the considerable variation represented by these groups face increasing difficulty in distinguishing material from these groups, and have regularly used isozymes to help make the necessary identifications (Glaszmann, 1987, 1988). For example, the lines Azucena and PR 304 have been classified as indica using morphological characters, whereas they behave as japonica types in crossing studies (Gurdev Khush, personal communication).

When analysed using RAPD they are clearly revealed as japonica rices (Virk *et al.*, 1995a). Such discrepancies were apparent in another experiment by Virk *et al.* (1995a). Forty-four rice accessions which had been previously classified as indica or japonica on morphological grounds were found by cluster analysis of RAPD data to divide into two major groups. All 31 accessions of one group had been classified as indica; however, eight of the other group had been designated as either japonica or javanica, while the other five had been classified as indica. Clearly the RAPD classifications do not always correlate exactly with classifications based on morphology, but they do accord well with the classifications of rice based upon crossability and isozyme data.

Six *O. sativa* isozyme groups can now be identified using RAPD; these groups reflect precisely those defined by crossability. The same can also be achieved by other molecular marker strategies including RFLP (Wang and Tanksley, 1989; Zhang *et al.*, 1992; Zheng *et al.*, 1994), and microsatellites (Wu and Tanksley, 1993). The ease with which this discrimination can be made using RAPD markers is illustrated in Fig. 5.1.

Of equal importance to genebank management is the ease with which germplasm of closely related species can be identified. In *Oryza*, and particularly within the *Oryza sativa* complex of AA genome species, there is often uncertainty with regard to the allocation of germplasm to several of the species. *O. rufipogon* and *O. nivara* from Asia, and even *O. glumaepatula*, which does seem to have a fairly distinct geographical distribution in South America, have all posed problems of identification using morphological characters. Much more precise identification can be achieved using RAPD markers (Fig. 5.2). RAPD markers have even revealed the true identity of material entering the genebank with the designation *O. meridionalis* that had been misidentified when compared to holotype material (Martin *et al.*, 1997).

5.4. Identifying Duplicates

The race against genetic erosion of crop gene pools has yielded many thousands of accessions safely stored in genebanks. However, genebanks have a finite capacity, and it is apparent that they often conserve more than one sample of the same genotype. In other words, plant genetic resources collections contain duplicate materials, yet the scale of the problem in seed genebanks cannot be determined with certainty. For many vegetatively propagated species, this situation is much more easily addressed.

From a purely management point of view, there are distinct advantages in identifying duplicate accessions, and thereby focusing most effort on unique genetic materials for conservation. Until now the identification

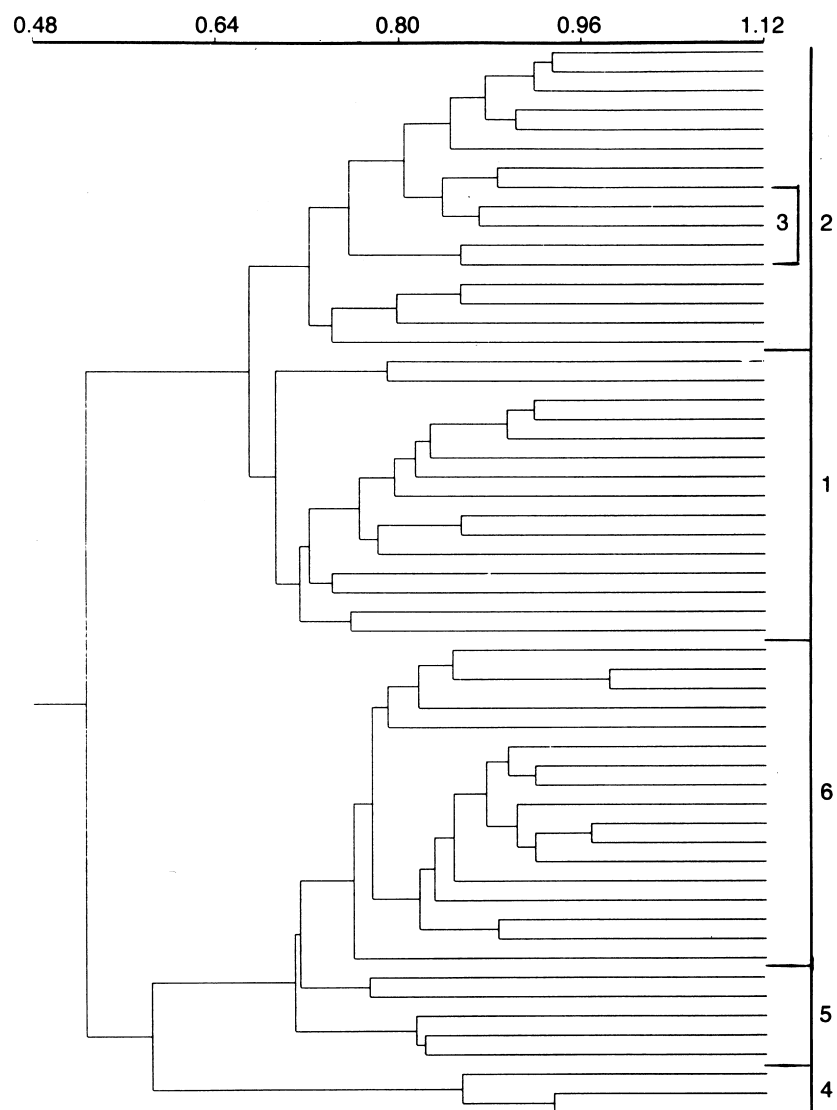


Fig. 5.1. Clustering of *Oryza sativa* accessions according to crossability group, based upon RAPD data. 1–6: rice crossability groups I–VI.

of duplicate accessions has had to rely on comparison of morphological characters, some of which are subject to environmental variation, together with passport data including (amongst others) variety name and origin. Identification of duplicates of vegetatively propagated species, such as potato, is more straightforward than for seed-propagated crops such as

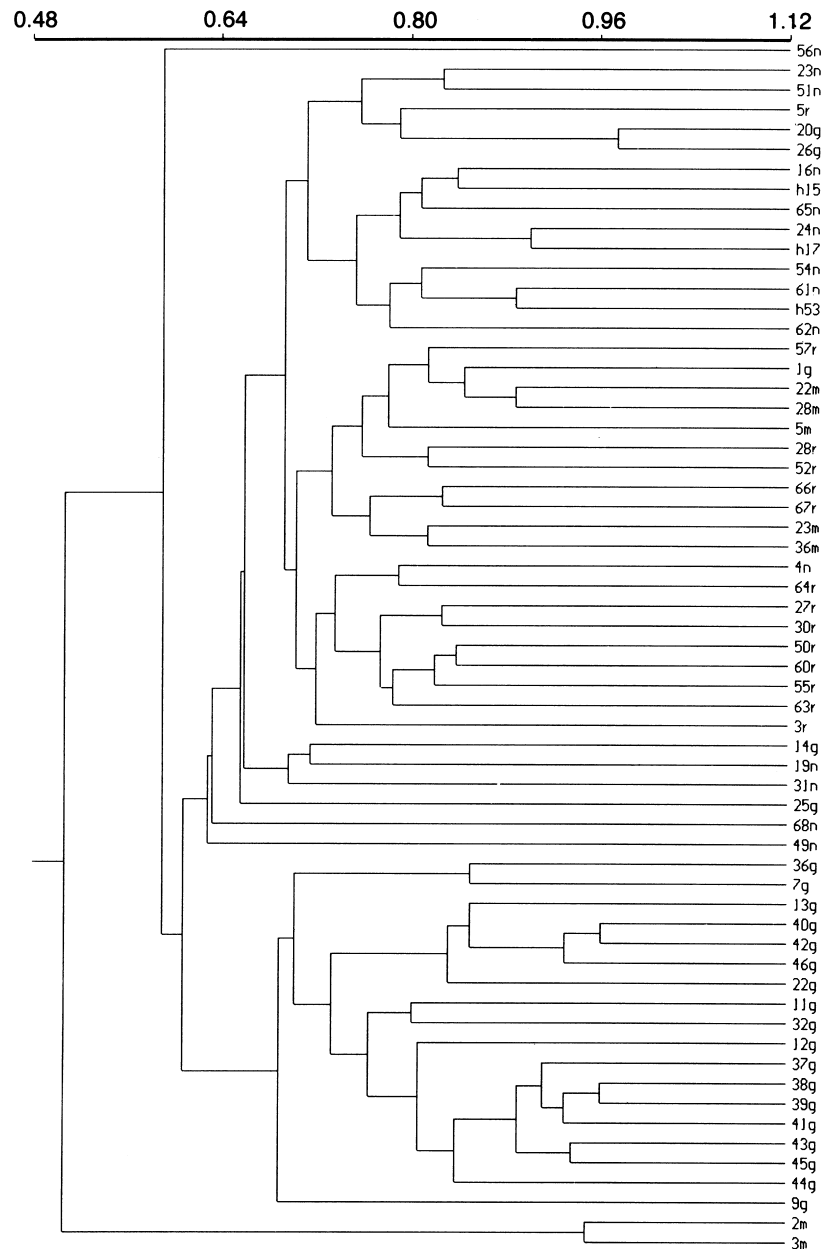


Fig. 5.2. Identification of wild rice species using cluster analysis (simple matching coefficient and UPGMA clustering) of RAPD data. (n = *Oryza nivara*; r = *O. rufipogon*; g = *O. glumaepatula*; m = *O. meridionalis*). Some accessions (e.g. 5, 22, 23, 28 and 36) were subsequently found to have been misidentified.

rice. At the International Potato Center (CIP) duplicate accessions have been routinely identified for some time by comparison of tuber proteins separated in polyacrylamide gels, and complemented by field observations of morphological characters. Duplicate clones have been eliminated from the collection of Andean potato varieties, reducing its size to more manageable proportions.

In more recent work, germplasm samples of rice from IRRI including known and suspected duplicates, as well as closely related germplasm, have been subjected to molecular analysis by Virk *et al.* (1995b). Their results demonstrate that an accurate discrimination of these categories of germplasm samples, including the identification of true and suspected duplicates, can be achieved. Two procedures have been proposed for identifying such duplicates. These differ in the method used in the initial stages, and the relative merits of either method would have to be balanced by those persons charged with conserving a collection. Local situations will vary considerably with regard to the relative costs of field work compared to molecular biology, and the expertise available to carry these out.

Procedure 1

- Select potential duplicate accessions from the collection following examination of available 'passport' data.
- Undertake initial morphological characterization of the suspected duplicates.
- Undertake a full molecular scrutiny of those germplasm samples that cannot be separated using these data.
- Designate as duplicates germplasm samples that cannot then be discriminated.

Procedure 2

- Select potential duplicates from the collection following examination of available 'passport' data as in Procedure 1.
- Carry out a pre-screen of pairs of suspected duplicates using a small number of molecular markers instead of morphological evaluation.
- Undertake a full molecular analysis of those germplasm samples which then cannot be separated.
- Designate as duplicates germplasm samples that cannot be discriminated.

Such results may provide information useful in the design of procedures that permit the routine identification of duplicates within a germplasm collection. Discussions concerning the number of marker bands that it is necessary to score before the designation of duplicates can take place are complex. It will never be possible to prove that two accessions are genetically identical without sequencing their entire genomes. Given that this is a practical impossibility, a decision must be made about the amount

of testing that will be performed before two accessions are accepted as (or 'designated' as) duplicates. This decision must be influenced by the number of potential duplicates that are to be tested. However, the results of Virk *et al.* (1995b) indicate that in rice, for one very similar pair of accessions, we can be 99% confident of detecting a difference between them if we examine a total of 86 RAPD markers. It would clearly be possible to use other types of DNA-based markers for this purpose, although it would be important to ensure that the variation defined using alternative markers was biologically valid in terms of taxonomy and genetics. Moreover, the number of bands to be scored may differ depending upon the sequence types represented by the markers.

The reliable identification of duplicate accessions will provide management options for the germplasm curator. Whether it will lead to the reduction in size of germplasm collections is debatable. In the case of CGIAR centres, their germplasm collections are held in trust, and many accessions are actually intentional duplicate materials of existing national germplasm collections. Clearly the CGIAR centres have an obligation to continue to conserve those germplasm accessions already accepted for safety storage. With the acquisition of new germplasm accessions, however, the situation is potentially different. The study of Virk *et al.* (1995b) suggests a novel procedure which would allow the level of certainty of identifying duplicate samples to be set before those samples became part of a germplasm collection, and before they were assigned unique accession numbers. This option is one which could have a significant impact on germplasm management, provided the PCR-based marker technology can be easily and economically utilized by germplasm curators.

5.5. Core Collections

A major issue in genetic resources for some time has been the size of germplasm collections in relation to their effective management and use. The large size of collections in genebanks may severely constrain many genebank practices. One answer to these problems is to develop a core collection, originally envisaged by Frankel (1984) as being a subgroup of accessions of any germplasm collection which would incorporate, with minimum redundancy, the genetic diversity of a crop species and its relatives. To all intents and purposes the core would form the 'active collection' for germplasm evaluation and distribution, with the remainder of the germplasm being kept as a 'reserve' collection. More specifically there are perhaps two rational and practical motives to develop core collections (Mackay, 1995). The first is to facilitate germplasm management, and the second to increase the use of germplasm by breeders. These two objectives may not complement each other exactly, and indeed the germplasm

curator may well be biased towards the former. However, both objectives could be achieved by the use of molecular markers.

The core collection approach has already been taken for barley, cassava, sorghum, wheat, coffee and *Phaseolus* (Hodgkin *et al.*, 1995). In rice, where the problem of collection size is as great as any, various steps have been taken towards the core approach which aim at fulfilling both of the objectives referred to above. A principal objective of the International Rice Genebank is to establish a core which can help in the safe duplication of accessions representing the broad diversity of the genus *Oryza*, in several locations around the world (Vaughan and Jackson, 1995). The more straightforward the development of this core, the better from the point of view of the germplasm curator. In addition, current knowledge of rice diversity based upon geographic, morphological, agronomic, biochemical and molecular characteristics has resulted in the development of a small core of about 270 accessions of *O. sativa* which represents the known diversity of rice (Glaszmann, 1987; Bonman *et al.*, 1990). Similarly, Vaughan (1991a,b) has designated a core collection of wild rices to enable researchers to evaluate this germplasm efficiently. Continued collecting and biodiversity studies mean that the composition of these core collections must be updated periodically. Molecular and biochemical markers can be used to determine the degree of differentiation that actually exists between wild species themselves and between them and *O. sativa*. Questions about precisely where allelic richness can be found and whether wild species really are sources of distinct alleles can be addressed in order to determine whether new accessions should be added to the core, or whether existing accessions in the core are largely redundant because they contribute little which is genetically unique. The use of the core collection in combination with marker data at IRRI has also enabled rapid identification of germplasm possessing some much sought-after traits. For instance, studies of allozyme data enabled accessions of the small rayada group of rices to be pinpointed for resistance to leaf scald (a seed-borne disease caused by *Rhynchosporium oryzae*).

How then is the choice of core material to be made? Brown (1989a) has identified stratified sampling as being more efficient in establishing a core than purely random sampling. This relatively simple procedure involves dividing the collection into nonoverlapping groups, and then taking samples from each group. The way in which the groups are established will probably vary by crop species, but will depend upon taxonomy, passport data and ecogeographical information. In the face of uneven distribution of diversity and differentiation of accessions, this method will ensure that the allelic richness of a core will be maximized. One constraint to this approach, however, is the dearth of accurate passport data that unfortunately typifies the situation in many germplasm collections worldwide.

Schoen and Brown (1995) have gone further than this by undertaking a set of simulations. Utilizing allozyme marker data to demonstrate how core collections might be established, they identified how allelic richness could be achieved in cores developed by six different strategies, two of which (H and M) were described for the first time. All six strategies invoked stratified sampling from designated geographical groups, but the H and M strategies differed in that they utilized genetic marker data to guide sampling from within groups. When compared for allele retention the six strategies gave differing results, with the two strategies guided by marker data (M and H) consistently performing best, and the simplest approach, involving only random sampling, the worst.

To what extent molecular markers will be employed in the future via strategies of the H and M categories is arguable. Schoen and Brown clearly indicate that the way forward must be through stratified sampling of an entire collection using passport and other data to develop the groups to be sampled. With the increased use of different molecular marker techniques the limitations of allozyme information will be overcome, so that it will be increasingly possible to improve the selection process of material for the core from within the stratified groups of accessions. Selecting more accessions from groups of high marker gene diversity (H strategy), or targeting particular accessions that are both high in allelic richness and well differentiated (M strategy) can offer the possibility of further improvement over what can be achieved by simple stratified sampling.

In terms of the actual size of any core collection, statistical theory rather than any practical use of marker data has been used principally to determine what needs to be done. Brown (1989a) has argued that the core should consist of about 10% of the whole collection, up to a maximum of about 3000 accessions, for each species. He estimates that, at this level of sampling, the core will generally contain over 70% of the alleles present in the whole collection (Brown, 1989b). This seems to be a rule of thumb which many germplasm curators are adopting.

However, an alternative approach to this problem is provided by the work of Lawrence *et al.* (1995a,b), who considered the size of sample required to capture at least one copy of each allele at each of a number of independently inherited loci at a given probability. Their calculations indicate that, provided the sample size chosen gives a very high probability of conserving the alleles of a single locus, this size is also sufficient to give a high probability of conserving at least one copy of each allele at all other loci. Calculations based upon assumptions that the average genome of a species contains 40 000 structural loci (Nei, 1987), and that 40% of these loci are polymorphic (Hamrick, 1989), indicate that, even if the species is predominantly inbreeding, a sample size of only 172 will give a very high probability (>0.999999828) of conserving all of the alleles at all the polymorphic loci, even if the frequency of one allele at each locus is only 0.05.

In practice, therefore, it should be relatively easy to conserve all or very nearly all of the alleles of a population in a random sample of 172 plants!

Whether core collections are made up of samples taken from 172 plants or 3000, it would seem to be clear that the sampling strategy chosen to select the material which will make up those numbers is still of overriding importance and can be substantially assisted by using stratification and molecular markers.

5.6. Facilitating the Use of Germplasm

PCR-based molecular markers are increasingly being used to assess genetic diversity in germplasm collections. Sometimes, morphological data including those for quantitative traits of economic importance are also available from the genebank. There is an increasing desire to utilize this wealth of useful information, somehow to use molecular markers to assist identification of useful characteristics amongst conserved germplasm, and therefore to narrow the gap between genebank managers and plant breeders. One of the successes of the CGIAR conservation effort over three decades has been the close linkage between conservation and exploitation of germplasm. This has led to many outstanding examples of varietal release to alleviate hunger in developing countries.

Molecular markers are increasingly being used in marker-assisted selection programmes (Stomberg *et al.*, 1994). Both theoretical and experimental studies have shown that marker-assisted selection can be highly effective for producing improved genotypes. However, the success of such selection programmes is largely taken to depend on genetic linkage between markers and the relevant gene loci. Whilst such studies are invariably based on materials derived from planned crosses, could similar principles be applied to genetic resources held in germplasm collections? The work of Virk *et al.* (1996) has gone some way to achieving this by using multiple regression analysis to predict the performance of germplasm accessions of rice given the molecular marker genotypes of those accessions. From a large number of markers it was possible to pinpoint a handful which are significantly associated with a particular trait of interest. Subsequently they were able to make accurate predictions of field performance in a range of agronomic traits such as plant height, culm number, and days to flowering in a particular environment. It is known that associations can exist because of linkage disequilibrium as well as linkage (Hastings, 1990). It is also clear that for marker-assisted procedures to work for prediction or selection of complex inherited traits, it is of benefit that a high level of linkage disequilibrium must exist, particularly if the genome is not well saturated with markers (Stuber, 1990). This appears to be the case in the study of Virk *et al.* (1996), and if true, it appears that

associations between alleles at quantitative trait loci (QTLs) and at marker loci has been conserved throughout the period of diversification of rice germplasm in South and Southeast Asia.

One obvious benefit of obtaining information about molecular markers and quantitative traits would be the more efficient selection of putative parents for producing populations to map QTLs for a particular trait. Also, the procedure could be used as an initial screening method for the identification of QTLs. The established method for this is the selection of two parents that differ markedly in a particular quantitative character, and then the determination of associations between markers and that character in F_2 or backcross progeny. The apparent advantages of using diverse germplasm instead are: (i) that this could allow the detection of QTLs that vary across a wide spectrum of biodiversity rather than just between two parental lines; and (ii) that QTLs for any quantitative trait can be studied in the same investigation.

Regardless of the underlying causes of the associations which have been detected, the use of molecular markers, which are more or less randomly distributed across the genome, coupled with multiple regression analysis could substantially change and improve the way in which crop biodiversity is used in the future. The combination of techniques should allow the prediction of what a plant will look like in terms of quantitative agronomic traits prior to elaborate field trials. If a diverse test array of germplasm is scored for important traits requiring specialized assessment conditions (such as stress tolerances, for example) then marker data could provide an efficient means of predicting the value of additional germplasm for such characteristics. Such results may demonstrate the value of *ex situ* plant germplasm collections not just as repositories of useful genes, but also as sources of information about phenotypic characters.

5.7 Conclusion

One of the major criticisms regularly levelled at genetic resource conservationists over the last 40 years has been that they have frequently been unable to provide appropriate material for crop improvement programmes. However, with appropriate organization of conserved material and the application of current DNA-based marker technology, genebanks can more easily counter these criticisms and become much more valuable interfaces between the activities of conservationists on the one hand and those wishing to exploit germplasm for the benefit of humankind on the other. Often the separation of those who conserve germplasm from those who wish to use it is a barrier to effective exploitation of this valuable germplasm. The advent of molecular characterization and evaluation of germplasm opens another chapter in genetic conservation, and one which

will fundamentally change our perspectives on the nature, structure and value of crop gene pools.

References

- Bonman, J.M., Mackill, A.Q. and Glaszmann, J.C. (1990) Resistance to *Gerlachia oryzae* in rice. *Plant Disease* 74, 306–309.
- Brown, A.H.D. (1989a) The case for core collections. In: Brown, A.H.D., Marshall, D.R., Frankel, O.H. and Williams, J.T. (eds) *The Use of Plant Genetic Resources*. Cambridge University Press, Cambridge, pp. 136–151.
- Brown, A.H.D. (1989b) Core collections: a practical approach to genetic resources management. *Genome* 31, 818–824.
- Cregan, P.B., Akkaya, M.S., Bhagwat, A.A., Lavi, U. and Rogwen, J. (1994) Length polymorphisms of simple sequence repeat (SSR) DNA as molecular markers in plants. In: Gresshof, P.M. (ed.) *Plant Genome Analysis*. CRC Press, London, pp. 47–56.
- Frankel, O.H. (1984) Genetic perspectives of germplasm conservation. In: Arber, W.K., Llimensee, K., Peacock, W.J. and Starlinger, P. (eds) *Genetic Manipulation: Impact on Man and Society*. Cambridge University Press, Cambridge, pp. 161–170.
- Gepts, P. (1995) Genetic markers and core collections. In: Hodgkin, T., Brown, A.H.D., van Hintum, Th.J.L., and Morales, E.A.V. (eds) *Core Collections of Plant Genetic Resources*. Wiley, Chichester, pp. 127–146.
- Glaszmann, J.C. (1987) Isozymes and classification of Asian rice varieties. *Theoretical and Applied Genetics* 74, 21–30.
- Glaszmann, J.C. (1988) Geographic pattern of variation among Asian native cultivars (*Oryza sativa* L.) based on fifteen isozyme loci. *Genome* 30, 782–792.
- Hamrick, J.L. (1989) Isozymes and the analysis of genetic structure in plant populations. In: Soltis, D.E. and Soltis, P.S. (eds) *Isozymes in Plant Biology*. Chapman & Hall, London, pp. 87–105.
- Hastings, A. (1990) The interaction between selection and linkage in plant populations. In: Brown, A.H.D., Clegg, M.T., Kahler, A.L. and Weir, B.S. (eds) *Plant Population Genetics, Breeding, and Genetic Resources*. Sinauer, Sunderland, MA, pp. 163–180.
- Hodgkin, T., Brown, A.H.D., van Hintum, Th.J.L. and Morales, E.A.V. (eds) (1995) *Core Collections of Plant Genetic Resources*. Wiley, Chichester.
- Howell, E.C., Newbury, H.J., Swennen, R.L., Withers, L.W. and Ford-Lloyd, B.V. (1994) The use of RAPD for identifying and classifying *Musa* germplasm. *Genome* 37, 328–332.
- Jackson, M. (1994) Preservation of rice strains. *Nature* 371, 470.
- Jackson, M. and Huggan, R. (1993) Sharing the diversity of rice to feed the world. *Diversity* 9, 22–25.
- Kameswara Rao, N. and Jackson, M.T. (1996a) Seed longevity of rice cultivars and strategies for their conservation in genebanks. *Annals of Botany* 77, 251–260.
- Kameswara Rao, N. and Jackson, M.T. (1996b) Seed production environment and storage longevity of japonica rices (*Oryza sativa* L.). *Seed Science Research* 6, 17–21.

- Lawrence, M.J., Marshall, D.F. and Davies, P. (1995a) Genetics of conservation. I. Sample size when collecting germplasm. *Euphytica* 84(2), 89–99.
- Lawrence, M.J., Marshall, D.F. and Davies, P. (1995b) Genetics of conservation. II. Sample size when collecting seed of cross-pollinated species and the information that can be obtained from the evaluation of material held in gene banks. *Euphytica* 84, 101–107.
- Mackay, M.C. (1995) One core collection or many? In: Hodgkin, T., Brown, A.H.D., van Hintum, Th.J.L. and Morales, E.A.V. (eds) *Core Collections of Plant Genetic Resources*. Wiley, Chichester, pp. 199–210.
- Martin, C., Juliano, A., Newbury, H.J., Lu, B.-R., Jackson, M.T. and Ford-Lloyd, B.V. (1997) The use of RAPD markers to facilitate the identification of *Oryza* species within a germplasm collection. *Genetic Resources and Crop Evolution* 44, 175–183.
- Matsuyama, T., Oyama, M. and Akiyama, T. (1993) DNA fingerprinting in *Citrus* cultivars. In: Hayashi, T., Omura, M. and Scott, N.S. (eds) *Techniques on Gene Diagnosis and Breeding in Fruit Trees*. Fruit Tree Research Station, Ibaraki, pp. 26–30.
- Morgante, M. and Olivieri, A.M. (1993) PCR-amplified microsatellites as markers in plant genetics. *Plant Journal* 3, 175–182.
- National Research Council (1991) *Managing Global Genetic Resources. The US National Plant Germplasm System*. National Academy Press, Washington.
- Nei, M. (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- Neuhaus, D., Kuhl, H., Kohl, J.G., Dorfel, P. and Borner, T. (1993) Investigation on the genetic diversity of *Phragmites* stands using genomic fingerprinting. *Aquatic Botany* 45, 357–364.
- Schoen, D.J. and Brown, A.H.D. (1995) Maximising genetic diversity in core collections of wild relatives of crop species. In: Hodgkin, T., Brown, A.H.D., van Hintum, Th.J.L. and Morales, E.A.V. (eds) *Core Collections of Plant Genetic Resources*. Wiley, Chichester, pp. 55–76.
- Senior, M.L. and Heun, M. (1993) Mapping maize microsatellites and polymerase chain reaction confirmation of the targeted repeats using a CT primer. *Genome* 36, 883–889.
- Stomberg, L.D., Dudley, J.W. and Rufener, G.K. (1994) Comparing conventional early generation selection with molecular marker assisted selection in maize. *Crop Science* 34, 1221–1225.
- Stuber, C.W. (1990) Molecular markers in the manipulation of quantitative characters. In: Brown, A.H.D., Clegg, M.T., Kahler, A.L. and Weir, B.S. (eds) *Plant Population Genetics, Breeding, and Genetic Resources*. Sinauer, Sunderland, MA, pp. 334–350.
- Vaughan, D.A. (1991a) Choosing rice germplasm for evaluation. *Euphytica* 54, 147–154.
- Vaughan, D.A. (1991b) Core collections: Providing access to wild *Oryza* germplasm. In: *Rice Genetics II*. IRRI, Manila.
- Vaughan, D.A. and Jackson, M.T. (1995) The core as a guide to the whole collection. In: Hodgkin, T., Brown, A.H.D., van Hintum, Th.J.L. and Morales, E.A.V. (eds) *Core Collections of Plant Genetic Resources*. Wiley, Chichester, pp. 229–239.

- Virk, P.S., Ford-Lloyd, B.V., Jackson, M.T. and Newbury, H.J. (1995a) Use of RAPD for the study of diversity within plant germplasm collections. *Heredity* 74, 170–179.
- Virk, P.S., Newbury, H.J., Jackson, M.T. and Ford-Lloyd, B.V. (1995b) The identification of duplicate accessions within a rice germplasm collection using RAPD analysis. *Theoretical and Applied Genetics* 90, 1049–1055.
- Virk, P.S., Ford-Lloyd, B.V., Jackson, M.T., Pooni, H.S., Clemeno, T.P., and Newbury, H.J. (1996) Predicting quantitative variation within rice germplasm using molecular markers. *Heredity* 76, 296–304.
- Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23, 4407–4414.
- Wang, Z.Y. and Tanksley, S.D. (1989) Restriction fragment length polymorphism in *Oryza sativa* L. *Genome* 32, 1113–1118.
- Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. and Tingey, S.V. (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* 18, 6231–6235.
- Wu, K.-S. and Tanksley, S.D. (1993) Abundance, polymorphism and genetic mapping of microsatellites in rice. *Molecular and General Genetics* 241, 225–235.
- Yang, G.P., Saghai Maroof, M.A., Xu, C.G., Zhang, Q. and Biyashev, R.M. (1994) Comparative analysis of microsatellite DNA polymorphism in landraces and cultivars of rice. *Molecular and General Genetics* 24, 187–194.
- Zhang, Q., Saghai Maroof, M.A., Lu, T.Y., Shew, B.Z. (1992) Genetic diversity and differentiation of *indica* and *japonica* rice detected by RFLP analysis. *Theoretical and Applied Genetics* 83, 495–499.
- Zheng, K., Qian, H., Shen, B., Zhuang, J., Lin, H. and Lu, J. (1994) RFLP-based phylogenetic analysis of wide compatibility varieties in *Oryza sativa* L. *Theoretical and Applied Genetics* 88, 65–69.

***In Vitro* Conservation Methods**

F. Engelmann

*International Plant Genetic Resources Institute (IPGRI), Via
delle Sette Chiese 142, 00145 Rome, Italy*

6.1. Introduction

The most widely used method of conserving plant genetic resources depends on seed. Many species produce seeds that can be dried to low moisture contents and stored at low temperature. Their longevity can be extended by reducing their moisture content and decreasing their storage temperature. Seeds that can tolerate extensive desiccation and can be stored in this way are termed orthodox (Roberts, 1973).

However, several categories of crops present problems with regard to seed storage. A number of species, predominantly tropical or subtropical, such as coconut (*Cocos nucifera*), cacao (*Theobroma cacao*) and many tree and shrub species, have seeds which do not undergo maturation drying and are shed at relatively high moisture contents (Chin and Pritchard, 1988). These seeds are unable to withstand much desiccation and are often sensitive to chilling, and therefore, cannot be stored dry at low temperature. These so-called recalcitrant seeds (Roberts, 1973) have to be kept in moist, relatively warm conditions, and even when stored in an optimal manner, their longevity is limited to weeks, occasionally months. Recent investigations have identified species which exhibit an intermediate form of seed storage behaviour (Ellis *et al.*, 1990). These seeds can tolerate desiccation to fairly low moisture content but the dry seeds are injured by low temperature. In comparison to truly recalcitrant seeds, the storage life of these seeds can be prolonged by some drying, but long-term conservation, i.e. comparable to orthodox seeds, remains unattainable. This category of seeds includes economically important

species such as coffee (*Coffea* spp.) and oil palm (*Elaeis guineensis*) (Ellis *et al.*, 1990, 1991).

Some crop species have genotypes which do not produce seeds and others, such as potato (*Solanum tuberosum*), yam (*Dioscorea* spp.), cassava (*Manihot* spp.), sweet potato (*Ipomea batatas*) and sugar-cane (*Saccharum* spp.), have either sterile genotypes or produce orthodox seeds which are highly heterozygous and are therefore of limited interest for the conservation of particular gene combinations. These species are mainly propagated vegetatively to maintain clonal genotypes. Other crop species such as banana and plantain (*Musa* spp.) do not produce seeds and are thus multiplied vegetatively.

At present, the most common method to preserve the genetic resources of these problem crop species is as whole plants in the field. There are, however, several serious problems with field genebanks (Withers and Engels, 1990). The collections are exposed to natural disasters and attacks by pests and pathogens; moreover, labour costs and the requirement for technical personnel are very high. In addition, distribution and exchange from field genebanks is difficult because of the vegetative nature of the material and the greater risks of disease transfer.

Until now, most plant genetic resources conservation has focused on crop species. However, the conservation of rare and endangered plant species has also become an issue of concern (Fay, 1994).

Finally, the development of biotechnology has led to the production of a new category of germplasm, including clones obtained from elite genotypes, cell lines with special attributes, and genetically transformed material (Engelmann, 1994). This new germplasm is often of high added value and very difficult to produce. The development of efficient techniques to ensure its safe conservation is therefore of paramount importance.

During the last 20 years, *in vitro* culture techniques have been extensively developed and applied to more than 1000 different species (Bigot, 1987). Tissue culture techniques are of great interest for the collecting, multiplication and storage of plant germplasm (Engelmann, 1991a). Tissue culture systems allow propagation of plant material with high multiplication rates in an aseptic environment. Virus-free plants can be obtained through meristem culture in combination with thermotherapy, thus ensuring the production of disease-free stocks and simplifying quarantine procedures for the international exchange of germplasm. The miniaturization of explants allows reduction in space requirements and consequently labour costs for the maintenance of germplasm collections.

However, the high multiplication rates which can be achieved using *in vitro* culture procedures lead to the regular production of large amounts of plant material. This creates problems for the management of large *in vitro* collections. In addition, risks of losing material through contamination or human error are present at each subculture. More importantly, the risks of

losing the genetic integrity of the plant material through somaclonal variation increase with time in culture (Scowcroft, 1984). Storage techniques which reduce the burden placed upon all *in vitro*-based procedures and preserve the genetic integrity of the plant material are urgently needed.

Different *in vitro* conservation methods are employed, depending on the storage duration required (Engelmann, 1991a). For short- and medium-term storage, the aim is to reduce growth and to increase the intervals between subcultures. For long-term storage, cryopreservation – i.e. storage at ultra-low temperature, usually that of liquid nitrogen (-196°C) – is the only current method. At this temperature, all cellular divisions and metabolic processes are stopped. The plant material can thus be stored without alteration or modification for a theoretically unlimited period of time. Moreover, cultures are stored in a small volume, protected from contamination, requiring a very limited maintenance. *In vitro* collecting, slow growth and cryopreservation techniques are described and analysed in the following sections. The sections describing *in vitro* and slow growth techniques have been adapted from a recent review by Withers and Engelmann (1997).

6.2. *In vitro* Techniques for Collection and Exchange of Germplasm

6.2.1. *In vitro* collecting

Collectors are faced with various problems when collecting germplasm of recalcitrant seed and vegetatively propagated plant species. Collecting missions often require travelling for relatively long periods in remote areas. It is thus necessary to keep the material collected in good state for some days/weeks before it can be placed in optimal growth or storage conditions. There are thus great risks that recalcitrant seeds either germinate or deteriorate before they are brought back to the genebank (Allen and Lass, 1983). In addition, many recalcitrant seeds have considerable weight and bulk, which increases the volume of material to handle and induces additional costs, if an adequate sample of the population is to be collected. With vegetatively propagated species, the material collected will consist of stakes, pieces of budwood, tubers, corms or suckers. Not only will most of these explants not be adapted to survival once excised from the parent plant, but they will also present health risks due to their vegetative nature and contamination with soil-borne pathogens (Peacock, 1987; Withers, 1987).

Difficulties can also be encountered when collecting germplasm of orthodox seed-producing species. Even with careful planning of the time

of a collecting mission, there might be no or little seed available for all or part of the germplasm to be collected, or seeds might not be at the optimal developmental stage, or might be shed from the plant or eaten by grazing animals (IBPGR, 1984; Guarino *et al.*, 1995; Withers, 1995).

The problems described previously can be overcome if it is realized that the seed is not the only material which can be collected: zygotic embryos, or vegetative tissues such as pieces of budwood, shoots or apices can be sampled, transported and grown successfully if placed under adequate conditions.

Following an expert meeting organized by IBPGR in 1984 and sponsorship of various research programmes, *in vitro* collecting techniques have been developed for different materials including embryos of coconut (Assy-Bah *et al.*, 1987, 1989), cacao, avocado, *Citrus*, vegetative tissues of cacao (Yidana *et al.*, 1987), *Musa*, coffee (CATIE, 1997), *Prunus*, grape (Elias, 1988), cotton (Altman *et al.*, 1990) and several forage grasses (Ruredzo, 1989).

The critical points to consider for the development of *in vitro* collecting techniques have been synthesized and analysed by Withers (1995). The techniques developed are very simple and flexible, as illustrated below with coconut and cacao.

In the case of coconut, the *in vitro* field collecting technique developed by Assy-Bah *et al.* (1987, 1989) consisted of extracting from the nut with a corkborer a plug of endosperm containing the embryo. After surface sterilization with calcium hypochlorite or commercial bleach, the embryos were dissected on the spot under the shelter of a wooden box and inoculated onto semisolid medium, or the endosperm plugs were transported to the laboratory where the embryos were dissected and inoculated onto semisolid medium in aseptic conditions. This technique is very efficient since after approximately 6 to 9 months in culture in the laboratory under standard conditions, an average of 75% of the embryos collected developed into plantlets which could be successfully transferred to the nursery and then to the field. In addition, embryos inoculated *in vitro* in the field could be kept in the open for two months before being grown in the laboratory, without influencing their further development.

This technique has been modified by other researchers towards higher (Sossou *et al.*, 1987) or lower sophistication (Rillo and Paloma, 1991). In its more sophisticated version, inoculation of the embryos onto sterile medium is performed in an inflatable glove box. The simpler procedure consists of sterilizing the endosperm plugs containing the embryos and of placing them in a cool box in sterile plastic bags for transportation. Upon arrival in the laboratory, a second disinfection is performed and the embryos are dissected and inoculated onto culture medium inside the laminar airflow cabinet.

In the case of cacao, an *in vitro* collecting method was developed for

budwood (Yidana *et al.*, 1987). Considering that absolute sterility would be difficult to achieve in the field and would not necessarily be essential for robust, woody material, the aim was to place the samples collected in conditions which would suppress or delay deterioration. Stem nodal cuttings were disinfected with boiled water in which drinking-water sterilizing tablets had been dissolved and containing fungicides, then inoculated onto semisolid medium supplemented with fungicide and antibiotics. Explants could be maintained in a relatively clean (though not necessarily completely sterile) condition for up to 6 weeks.

6.2.2. In vitro germplasm exchange

In vitro culture techniques have been used extensively for the international exchange of germplasm because of their obvious advantages over *in vivo* material, notably their reduced weight and volume, and phytosanitary condition. Indeed, meristem culture, alone or employed in combination with thermotherapy, can eliminate viral pathogens (Karthi, 1986). Plant material can thus be multiplied, stored and exchanged in a disease-free state. In particular, because aseptic culture conditions are used and provided that the plant material has been cleansed of internal contaminants, problems with the international movement of germplasm are considerably reduced. Most countries will accept batches of plants *in vitro* with a phytosanitary certificate, without requiring a rigorous quarantine period (Fay, 1994). However, an important caution in this area is that transfer to *in vitro* culture does not confer disease-free status (Withers and Engelmann, 1997). Indexing, using one or a combination of the various techniques available, which include symptomatology, grafting/inoculation on indicator plants, ELISA and molecular techniques such as dsRNA detection, is the only sure way of making this judgement.

Routine procedures for the international exchange of *in vitro* cultures have been developed, notably for potato, cassava, yam and *Musa* (e.g. Espinoza *et al.*, 1992). Samples are usually placed in glass, or preferably plastic, test tubes on the standard culture medium, possibly with the gelling agent modified to increase its firmness and a reduction in the carbon source to limit growth. Heat-sealable polythene bags are sometimes employed instead of more fragile plastic or glass test tubes (Reed, 1991). With species such as potato and sweet potato, where it is possible to induce their formation *in vitro*, tubers are also used for germplasm exchange, since they are often easier to transport and to handle than plantlets (Dodds, 1992; Espinoza *et al.*, 1992).

Encapsulation of plant material in alginate beads has been suggested recently as a possible way of germplasm exchange in the case of banana (Rao *et al.*, 1993). Using encapsulated apices instead of *in vitro* plantlets

would result in a further volume reduction. In addition, encapsulated apices might be either grown *in vitro* after their receipt, or directly sown in the soil as seeds where they would develop into plantlets (Mathur *et al.*, 1989). The encapsulation technique has also been employed with nodal segments of yam (Hasan and Takagi, 1995). In a germplasm exchange simulation experiment, encapsulated nodal segments were conserved for two weeks in the dark in cryotubes containing culture medium, without any detrimental effect on their further growth.

6.3. Slow-growth Storage

6.3.1. Classical techniques

Standard culture conditions can be used for medium-term storage with species which have a naturally slow-growing habit only. A wide range of *Coffea* species are thus conserved on standard medium at 27°C without subculturing for durations varying from 6 to 12 months, depending on the species (Bertrand-Desbrunais, 1991). However, for most species, growth reduction is achieved by modifying the environmental conditions and/or the culture medium. The most widely applied technique is temperature reduction, which can be combined with a decrease in light intensity or culture in the dark.

Temperatures in the range of 0–5°C are employed with cold-tolerant species. Strawberry (*Fragaria* × *ananassa*) plantlets have been stored at 4°C in the dark and kept viable for 6 years with the regular addition of a few drops of liquid medium (Mullin and Schlegel, 1976). Apple (*Malus domestica*) and *Prunus* shoots survived 52 weeks at 2°C (Druart, 1985).

Tropical species are often cold-sensitive and have to be stored at higher temperatures, which depend on the cold sensitivity of the species. Kiwi fruit shoots can be conserved at 8°C (Monette, 1986) and taro (*Colocasia esculenta*) tolerates 3 years of storage at 9°C (Staritsky *et al.*, 1986). *Musa in vitro* plantlets can be stored at 15°C without transfer for up to 15 months (Banerjee and De Langhe, 1985). Other tropical species such as oil palm and cassava are much more cold-sensitive: oil palm somatic embryos and plantlets do not withstand even a short exposure to temperatures lower than 18°C (Corbineau *et al.*, 1990). Cassava shoot cultures have to be conserved at temperatures higher than 20°C (Roca *et al.*, 1984).

Various modifications can be made to the culture medium in order to reduce growth. Embryogenic cultures of carrot could be conserved on a medium without sucrose for two years, and reproducted if a sucrose solution was supplied (Jones, 1974). Kartha *et al.* (1981) could conserve coffee plantlets on a medium devoid of sugar and with only half of the

mineral elements of the standard medium. Replacement of sucrose by ribose allowed the conservation of banana plantlets for 24 months (Ko *et al.*, 1991). The addition of osmotic growth inhibitors (e.g. mannitol) or hormonal growth inhibitors (e.g. abscisic acid) is also employed successfully to reduce growth (Westcott, 1981a,b; Staritsky *et al.*, 1986; Ng and Ng, 1991; Viterbo and Rabinowitch, 1994; Vysotskaya, 1994).

The type of explant as well as its physiological state when entering storage can influence the duration of storage achieved. Roxas *et al.* (1995) indicate that, in the case of chrysanthemum, nodal segments showed higher survival rates than apical buds. The presence of a root system generally increases the storage capacities, as observed by Kartha *et al.* (1981) with coffee plantlets. Microtubers can be successfully employed as storage propagules, as demonstrated with potato (Kwiatowski *et al.*, 1988). Preconditioning the explants by exposing them briefly to temperature and light conditions intermediate between standard and storage conditions was favourable for *Nephrolepis* and *Cordyline* cultures (Hvoslef-Heide, 1992). Higher survivals were obtained with shoot cultures of wild cherry, chestnut and oak if they were kept for 10 days under standard conditions after the last subculture before their transfer to the cold storage chamber (Janeiro *et al.*, 1995).

The type of culture vessel, its volume as well as the type of closure of the culture vessel can greatly influence the survival of stored cultures (Engelmann, 1991a; Withers, 1992). Roca *et al.* (1984) indicate that cassava shoot cultures could be stored for longer periods in a better condition by increasing the size of the storage containers. White spruce embryogenic tissues withstood a one-year storage period in hermetically sealed serum-capped flasks (Joy *et al.*, 1991). Replacing cotton plugs by polypropylene caps, thus reducing the evaporation of the culture medium, increased the survival rate of *Rauvolfia serpentina* during storage (Sharma and Chandel, 1992). The use of heat-sealable polypropylene bags instead of glass test tubes or plastic boxes was beneficial for the storage of several strawberry varieties (Reed, 1991, 1992).

At the end of a storage period, cultures are transferred onto fresh medium and usually placed for a short period in optimal conditions to stimulate regrowth before entering the next storage cycle.

6.3.2. Alternative techniques

Alternative techniques include modification of the gaseous environment of cultures, and desiccation and/or encapsulation of explants. Growth reduction can be achieved by reducing the quantity of oxygen available to the cultures. The simplest method consists of covering the explants with paraffin, mineral oil or liquid medium. This technique was first developed

by Caplin (1959), who could store carrot calluses for 5 months under a layer of paraffin oil. Augereau *et al.* (1986) and Moriguchi *et al.* (1988) applied this technique to calluses of *Catharanthus* and grape, respectively. Florin (1989) showed that 50% of a collection of 313 callus lines survived after storage under mineral oil for 12 months. However, Mannonen *et al.* (1990) reported that conservation under mineral oil did not preserve the productivity of *Catharanthus* and *Panax ginseng* callus cultures for more than 6 months.

Similar experiments performed with shoot cultures of various species led to contradictory results (Withers and Engelmann, 1997). After 4 months of storage under mineral oil, regrowth of surviving coffee shoot cultures was very slow (Jouve *et al.*, 1991) and pear microcuttings did not survive (Chatti-Dridi, 1988). In contrast, several ginger species could be conserved for up to two years under mineral oil with high viability (Dekkers *et al.*, 1991).

Reduction of the quantity of oxygen can also be achieved by decreasing the atmospheric pressure of the culture chamber or by using a controlled atmosphere. Tobacco and chrysanthemum plantlets could be stored under low atmospheric pressure (with 1.3% oxygen) for 6 weeks (Bridgen and Staby, 1981). Oil palm polyembryogenic cultures were conserved for 4 months at room temperature in a controlled atmosphere with 1% oxygen (Engelmann, 1990). Dorion *et al.* (1994) showed that hypoxic regimes at standard or intermediate temperature could replace low-temperature storage for shoot cultures of peach. Rice calluses could be stored for 12 weeks in a CO₂ or N₂ saturated atmosphere (Watanabe *et al.*, 1991).

Desiccation of cultures as a means of achieving medium-term conservation was first reported by Nitzsche (1980), who stored dehydrated carrot callus for one year at 15°C under 25% relative humidity. Increasing interest has been paid recently to this technique, with the development of so-called synthetic seeds for various plant species (Gray and Purohit, 1991; Attree and Fowke, 1993). Synthetic seed is a generic term for a somatic embryo delivery system used as a means of clonal propagation (Janick *et al.*, 1993). The aim is to use such somatic embryos as true seeds: embryos, encapsulated or not in alginate gel, could be stored after partial dehydration and sown directly *in vivo*. After progressive dehydration using saturated salt solutions, naked alfalfa somatic embryos could be conserved with 10–15% moisture content at room temperature for one year and showed only a 5% decrease in their conversion rate at the end of the storage period (Senaratna *et al.*, 1990). Carrot somatic embryos were stored for 8 months at 4°C without viability loss (Lecouteux *et al.*, 1992). Shorter storage durations were achieved with encapsulated material. Encapsulated carrot somatic embryos survived after 3 months in liquid medium at low temperature (Shigeta *et al.*, 1993) and encapsulated shoot tips of *Valeriana wallichii* could be conserved for over 6 months at 4–6°C (Mathur *et al.*,

1989). According to Redenbaugh *et al.* (1991), the rapid dehydration of the encapsulating matrix limits the respiration of the encapsulated material, which is the cause for the rapid survival loss generally observed.

6.4. Cryopreservation

6.4.1. History

The development of cryopreservation for plant cells and organs has followed the advances made with mammalian species, albeit several decades later. The first report on survival of plant tissues exposed to ultra-low temperatures was made by Sakai in 1960 when he demonstrated that very hardy mulberry twigs could withstand freezing in liquid nitrogen after dehydration mediated by extra-organ freezing. More than a decade later it was shown that cultured cells of flax could resist freezing to -50°C after pretreatment with dimethylsulphoxide (DMSO) (Quatrano, 1968). This study was followed by a similar one in which cell cultures of carrot were shown to survive after freezing in liquid nitrogen (Latta, 1971). The methodology employed in the above experiments followed the classical procedures which had been successful with other living systems (Ashwood-Smith and Farrant, 1980; Grout and Morris, 1987): chemical cryoprotection, slow dehydrative cooling followed by rapid immersion in liquid nitrogen, storage in liquid nitrogen, rapid thawing, washing and recovery.

The development for plant tissue cultures of so-called classical techniques which took place in the 1970s and 1980s is based on the above sequence of treatments (Kartha, 1985). In recent years, several new cryopreservation techniques have been developed which allow cryopreservation to be applied to a larger range of tissues and organs, in different technical environments (Kartha and Engelmann, 1994; Withers and Engelmann, 1997).

6.4.2. Dehydration and freezing injury

Most of the experimental systems employed in cryopreservation (cell suspensions, calluses, shoot tips, embryos) contain high amounts of cellular water and are thus extremely sensitive to freezing injury since most of them are not inherently freezing-tolerant. Cells have thus to be dehydrated artificially to protect them from the damage caused by the crystallization of intracellular water into ice (Meryman, 1966; Mazur, 1969, 1970, 1984). The techniques employed and the physical mechanisms upon which they

are based are different in classical and new cryopreservation techniques (Withers and Engelmann, 1997). Classical techniques involve freeze-induced dehydration, whereas new techniques are based on vitrification. Vitrification can be defined as the transition of water directly from the liquid phase into an amorphous phase or glass, whilst avoiding the formation of crystalline ice (Fahy *et al.*, 1984).

Classical cryopreservation techniques involve slow cooling down to a defined prefreezing temperature, followed by rapid immersion in liquid nitrogen. With temperature reduction during slow cooling, the cells and the external medium initially undergo supercooling; this is followed by ice formation in the medium (Mazur, 1969). The cell membrane acts as a physical barrier and prevents the ice from seeding the cell interior, so the cells remain unfrozen but supercooled. As the temperature is further decreased, an increasing amount of the extracellular solution is converted into ice, thus resulting in the concentration of extracellular solutes. Since cells remain supercooled and their aqueous vapour pressure exceeds that of the frozen external compartment, cells equilibrate by loss of water to external ice. In optimal conditions, most or all intracellular freezable water is removed, thus reducing or avoiding detrimental intracellular ice formation upon subsequent immersion of the specimen in liquid nitrogen. Rewarming should be as rapid as possible to avoid the phenomenon of recrystallization, in which ice melts and re-forms at a thermodynamically favourable, larger and more damaging crystal size (Meryman and Williams, 1985).

New techniques are vitrification-based procedures. In such procedures, the freezable intracellular water is removed prior to freezing by exposing the samples to highly concentrated cryoprotective media and/or air desiccation. Dehydration is followed in most cases by rapid cooling. As a result, the internal solutes vitrify and deleterious intracellular ice formation is avoided. Glass transitions (i.e. changes in the structural conformation of the glass) during cooling and rewarming have been recorded with various materials using thermal analysis (Sakai *et al.*, 1990; Dereuddre *et al.*, 1991a; Tannoury *et al.*, 1991; Niino *et al.*, 1992a).

6.4.3. Classical procedures

For each new material to be cryopreserved, optimal conditions have to be defined for each of the successive steps of the protocol.

The physiological state of the material can affect its survival. Cell suspensions are more likely to withstand freezing when they are employed during their exponential growth phase, i.e. when they are small and have a relatively low water content (Withers and Street, 1977). Survival of carnation shoot tips after freezing decreased progressively with their rank on

the shoot apex, starting from the terminal meristem, thus reflecting physiological and developmental differences found in explants sampled on the same plant (Dereuddre *et al.*, 1988). Harding *et al.* (1991) indicated that a long-term period in tissue culture before cryopreservation significantly reduced the ability of potato apices to survive freezing. A pregrowth period before the cryoprotective treatment on medium with osmotically active compounds such as mannitol or sorbitol can increase freeze-tolerance (Seitz, 1987).

Samples are submitted to a cryoprotective treatment before freezing. It is carried out using various cryoprotective substances such as dimethylsulphoxide, sorbitol, mannitol, sucrose or polyethylene glycol. Additional information on the role and mechanisms of action of cryoprotectants can be found in several reviews (e.g. Finkle *et al.*, 1985; Meryman and Williams, 1985; Kartha and Engelmann, 1994). Cryoprotectants are employed alone or in binary or ternary mixtures. Mixtures of cryoprotectants have proved especially effective with cell suspensions (Withers, 1985).

Classical cryopreservation procedures include a two-step freezing: slow, controlled cooling to a defined prefreezing temperature followed by rapid immersion of samples in liquid nitrogen. Freezing rates of between 0.5 and 2°C min⁻¹ down to prefreezing temperatures around -40°C generally give satisfactory results in most cases (Dereuddre and Engelmann, 1987). However, while for some materials, such as oil palm somatic embryos, a wide range of freezing rates can be employed without modification of the recovery rate (Engelmann and Dereuddre, 1988), other materials (e.g. grape cell suspensions) require very precise freezing parameters to achieve survival (Dussert *et al.*, 1991). Most classical freezing procedures require the use of expensive programmable freezing devices which achieve precise freezing conditions. However, sophisticated apparatus is not always necessary to obtain high survival. Withers and King (1980) have successfully cryopreserved various cell suspensions with an improvised and simple apparatus that can offer reproducible but nonlinear slow cooling. More recently, several authors have successfully employed domestic or laboratory deep-freezers to perform the slow cooling step (Lecouteux *et al.*, 1991; Sakai *et al.*, 1991; Nishizawa *et al.*, 1992; Engelmann *et al.*, 1994a; Tessereau *et al.*, 1994).

Removal of cryoprotectants by washing after thawing, which was widely adopted in the past, was found to be deleterious for a number of cell cultures. A much less stressing technique for removing cryoprotectants, developed by Chen *et al.* (1984), consists of moving cultures on a filter paper through a series of Petri dishes of solid culture medium. This technique was successfully applied to various cell suspensions and calluses (Kartha and Engelmann, 1994). In some cases, standard culture conditions have to be modified to improve recovery of cryopreserved cultures. Recovery of cell suspensions is generally improved if they are

cultured for some days on solid medium before being transferred to liquid medium (Dussert *et al.*, 1992). Cultures are often placed in the dark or under reduced light intensity to avoid photooxidation phenomena which are harmful to the material (Benson, 1990). Finally, the culture medium can be transitorily altered by modifying its hormonal balance, mineral composition or by incorporating activated charcoal (Engelmann *et al.*, 1985; Kuriyama *et al.*, 1989).

Classical freezing techniques have been assessed with different materials including cell suspensions, calluses, apices and embryos. They have proven to be highly successful with most cell suspensions and calluses, i.e. culture systems which consist of small units of relatively uniform morphology. However, apart from exceptions like carnation apices (Dereuddre *et al.*, 1988), these techniques are not suitable for the cryopreservation of larger units comprising a mixture of cell sizes and types, such as apices and zygotic and somatic embryos.

6.4.4. New techniques

New cryopreservation techniques offer practical advantages in comparison to classical ones (Steponkus *et al.*, 1992; Sakai, 1995). Vitrification-based procedures are operationally less complex than classical ones since they do not require the use of a programmable freezer. In addition, since ice formation is avoided during freezing, they are more adapted for freezing complex organs such as apices or embryos which contain a variety of cell types, each with unique requirements under conditions of freeze-induced dehydration (Withers and Engelmann, 1997). Finally, they have greater potential for broader applicability than classical techniques since only minor modifications are required for various cell types.

Seven different vitrification-based procedures can be identified: (i) encapsulation-dehydration; (ii) a procedure actually termed vitrification; (iii) encapsulation-vitrification; (iv) desiccation; (v) pregrowth; (vi) pre-growth-desiccation; and (vii) droplet freezing. Table 6.1 presents a list of species to which these new freezing techniques have been applied.

Encapsulation-dehydration

The encapsulation-dehydration technique is based on the technology developed for the production of synthetic seeds where somatic embryos are encapsulated in a bead of hydrosoluble gel (Redenbaugh *et al.*, 1991). This technique has been applied mostly to apices of more than ten species of temperate and tropical origin, and to somatic embryos of several crop species as well as to a *Catharanthus* cell suspension (Table 6.1).

Before cryopreservation, specimens can be submitted to conditioning treatments which increase their survival potential. In the case of

Table 6.1. List of plant species for which new cryopreservation techniques (encapsulation-dehydration, vitrification, encapsulation-vitrification, pregrowth-desiccation, pregrowth, desiccation) have been developed using different types of specimens.

<i>Species</i>	<i>Specimen</i>	<i>No. of accessions</i>	<i>Technique</i>	<i>Reference</i>
<i>Aesculu</i>				
<i>hypocastanea</i>	Zygotic embryo	1	Desiccation	Pence, 1990
<i>Allium sativum</i>	Apex	12	Vitrification	Niwata, 1995
<i>Allium wakegi</i>	Apex	7	Vitrification	Kohmura <i>et al.</i> , 1994
<i>Arachis hypogaea</i>	Zygotic embryo	6	Desiccation	Runthala <i>et al.</i> , 1993
<i>Araucaria excelsa</i>	Zygotic embryo	1	Desiccation	Pritchard and Prendergast, 1986
<i>Armoracia rusticana</i>	Hairy root culture	1	Encaps./dehydr.	Phunchindawan <i>et al.</i> , 1994
		1		Hirata <i>et al.</i> , 1995
<i>Artocarpus</i>				
<i>heterophyllus</i>	Zygotic embryo	?	Desiccation	Krishnapillay, 1989
<i>Asparagus officinalis</i>	Stem segment	1	Pregrowth/desicc.	Uragami <i>et al.</i> , 1990
	Somatic embryo	1	Vitrification	Uragami <i>et al.</i> , 1989
	Cell suspension	1	Vitrification	Uragami <i>et al.</i> , 1989
		1		Nishizawa <i>et al.</i> , 1993
<i>Baccaurea motleyana</i>	Zygotic embryo	1	Desiccation	Normah and Marzalina, 1995
<i>Baccaurea polyneura</i>	Zygotic embryo	1	Desiccation	Normah and Marzalina, 1995
<i>Beta vulgaris</i>	Apex	2	Encaps./dehydr.	Vandenbussche and De Proft, 1995
<i>Brassica campestris</i>	Cell suspension	1	Vitrification	Langis <i>et al.</i> , 1989
<i>Brassica napus</i>	Microspore	1	Pregrowth/desicc.	Uragami <i>et al.</i> , 1993
	Embryo	1	Encaps/dehydr.	Uragami <i>et al.</i> , 1993
<i>Bromus inermis</i>	Cell suspension	1	Vitrification	Ishikawa <i>et al.</i> , 1994
<i>Calamus manan</i>	Zygotic embryo	1	Desiccation	Marzalina <i>et al.</i> , 1992

Continued

Table 6.1. Continued

<i>Species</i>	<i>Specimen</i>	<i>No. of accessions</i>	<i>Technique</i>	<i>Reference</i>
<i>Camellia sinensis</i>	Apex	12	Vitrification	Kuranuki and Sakai, 1994
	Zygotic embryo	ns	Desiccation	Chaudhury <i>et al.</i> , 1991
		1		Wesley-Smith <i>et al.</i> , 1992
<i>Capsella bursa-pastoris</i>	Zygotic embryo	1	Pregrowth	Monnier and Leddet, 1978
<i>Carva</i>	Zygotic embryo	1	Desiccation	Pence, 1990
<i>Castanea</i>	Zygotic embryo	1	Desiccation	Pence, 1990
<i>Catharanthus</i>	Cell suspension	1	Encaps./dehydr.	Phunchindawan <i>et al.</i> , 1994
<i>Centaurium rigualii</i>	Nodal segment	1	Vitrification	Gonzalez-Benito and Perez, 1994a
<i>Chicorium intybus</i>	Apex	1	Encaps./dehydr.	Vandenbussche <i>et al.</i> , 1993
<i>Chrysanthemum morifolium</i>	Apex	1	Vitrification	Schneibel-Preikstas <i>et al.</i> , 1992a
<i>Citrus aurantifolia</i>	Zygotic embryo	1	Desiccation	Normah and Nordaini, 1994
<i>Citrus halimii</i>	Zygotic embryo	1	Desiccation	Normah and Hamida, 1992
<i>Citrus mitis</i>	Zygotic embryo	1	Desiccation	Normah and Nordaini, 1994
<i>Citrus sinensis</i>	Nucellar cell suspension	1	Vitrification	Sakai <i>et al.</i> , 1990
		4		Sakai <i>et al.</i> , 1991
<i>Cocos nucifera</i>	Zygotic embryo	4	Pregrowth/desicc.	Assy-Bah and Engelmann, 1992a,b
<i>Coffea</i> spp.	Zygotic embryo	1	Desiccation	Normah and Vengadasalam, 1992
		3		Abdelnour-Esquivel <i>et al.</i> , 1992
		1		Hor <i>et al.</i> , 1993
	Apex	2	Encaps./dehydr.	Engelmann <i>et al.</i> , 1994b
	Somatic embryo	1		Hatanaka <i>et al.</i> , 1994
	Somatic embryo	1	Desiccation; pregrowth/desicc.	Mycok <i>et al.</i> , 1995
<i>Corylus avellana</i>	Zygotic embryo	2	Desiccation	Gonzalez-Benito and Perez, 1994b
		1		Normah <i>et al.</i> , 1995
		1		Reed <i>et al.</i> , 1994

Continued

Table 6.1. Continued

Species	Specimen	No. of accessions	Technique	Reference
<i>Cucumis melo</i>	Somatic embryo	1	Pregrowth/desicc.	Shimonishi <i>et al.</i> , 1991
<i>Daucus carota</i>	Somatic embryo	1	Encaps./dehydr.	Dereuddre <i>et al.</i> , 1991b
<i>Dianthus caryophyllus</i>	Apex	1	Encaps./dehydr.	Tannoury, 1993
	Nodal segment	1	Encaps./dehydr.	Fukai <i>et al.</i> , 1994a
		1	Encaps./vitrif.	Tannoury <i>et al.</i> , 1991
		1	Vitrification	Langis <i>et al.</i> , 1990
<i>Durio zibenthinus</i>	Zygotic embryo	1	Pregrowth/desicc.	Hor <i>et al.</i> , 1990
<i>Elaeis guineensis</i>	Polyembryogenic culture	1	Pregrowth/desicc.	Engelmann <i>et al.</i> , 1985
		7		Dumet <i>et al.</i> , 1993
		75		Dumet, 1994
	Zygotic embryo	1	Desiccation	Grout <i>et al.</i> , 1983
		1		Engelmann <i>et al.</i> , 1995a
		1	Encaps./dehydr.	Engelmann <i>et al.</i> , 1995b
<i>Eucalyptus gunnii</i>	Apex	1	Encaps./dehydr.	Poissonier <i>et al.</i> , 1992
<i>Euphoria longan</i>	Zygotic embryo	1	Desiccation	Fu <i>et al.</i> , 1990
<i>Fagus</i>	Zygotic embryo	1	Desiccation	Pence, 1990
<i>Fraxinus excelsior</i>	Zygotic embryo	1	Desiccation	Brearley <i>et al.</i> , 1995
<i>Hevea brasiliensis</i>	Zygotic embryo	1	Desiccation	Normah <i>et al.</i> , 1986
<i>Hopea odorata</i>	Zygotic embryo	1	Desicc./slow cooling	Chai <i>et al.</i> , 1994
<i>Howea fosteriana</i>	Zygotic embryo	1	Desiccation	Chin <i>et al.</i> , 1988
<i>Ipomea batatas</i>	Apex	1	Vitrification	Schneibel-Preikstas <i>et al.</i> , 1992c
	Apex	2		Towill and Jarret, 1992
<i>Juglans regia</i>	Zygotic embryo	1	Desiccation	Pence, 1990
	Somatic embryo	1	Encaps./dehydr.	de Boucaud <i>et al.</i> , 1994
<i>Landolphia kirkii</i>	Zygotic embryo	1	Desiccation	Vertucci <i>et al.</i> , 1991
<i>Lansium domesticum</i>	Zygotic embryo	1	Desiccation	Normah and Jamilah, 1992
<i>Lilium</i>	Apex	4	Encaps./vitrif.	Sakai and Matsumoto, 1995
		6	Vitrification	Matsumoto <i>et al.</i> , 1994a
<i>Livistonia chinensis</i>	Zygotic embryo	1	Desiccation	Normah and Marzalina, 1995

Continued

Table 6.1. Continued

<i>Species</i>	<i>Specimen</i>	<i>No. of accessions</i>	<i>Technique</i>	<i>Reference</i>
<i>Malus</i> spp.	Apex	3	Encaps./dehydr.	Niino and Sakai, 1992
		5		Paul, 1994
		5	Vitrification	Niino <i>et al.</i> , 1992a
<i>Manihot esculenta</i>	Zygotic embryo	1	Desiccation	Marin <i>et al.</i> , 1990
	Apex	2	Encaps./dehydr.	Benson <i>et al.</i> , 1992
		3		Engelmann <i>et al.</i> , 1994b
	Somatic embryo	1	Desiccation; pregrowth/desicc.	Mycok <i>et al.</i> , 1995
<i>Menta</i>	Apex	1	Vitrification	Towill, 1990
<i>Morus</i>	Apex	1	Encaps./dehydr.	Niino and Sakai, 1992
		13	Vitrification	Niino <i>et al.</i> , 1992b
<i>Musa</i> spp.	Zygotic embryo	2	Desiccation	Abdelnour-Esquivel <i>et al.</i> , 1992
	Apex	1	Encaps./dehydr.	Panis, 1995
	Meristematic clump	5	Pregrowth	Panis, 1995
<i>Nephelium lappaceum</i>	Zygotic embryo	1	Desiccation	Hor <i>et al.</i> , 1990
<i>Nicotiana tabacum</i>	Cell suspension	1	Vitrification	Takano and Tamura, 1992
		1		Reinoud <i>et al.</i> , 1994
<i>Olea europaea</i>	Zygotic embryo	1	Desiccation	Gonzalez-Rio <i>et al.</i> , 1994
<i>Oryza sativa</i>	Cell suspension	1	Vitrification	Watanabe and Steponkus, 1994
<i>Panax ginseng</i>	Hairy root culture	1	Vitrification	Yoshimatsu <i>et al.</i> , 1994
<i>Phaseolus vulgaris</i>	Zygotic embryo	1	Pregrowth	Zavala and Sussex, 1986
<i>Phoenix dactylifera</i>	Somatic embryo	1	Desiccation; pregrowth/desicc.	Mycok <i>et al.</i> , 1995
<i>Pisum sativum</i>	Zygotic embryo	1	Pregrowth/desicc.	Mycok <i>et al.</i> , 1989
	Somatic embryo	1	Desiccation; pregrowth/desicc.	Mycok <i>et al.</i> , 1995
<i>Poncyus trifoliata</i>	Zygotic embryo	1	Desiccation	Radhamani and Chandel, 1992
<i>Prunus amygdalus</i>	Zygotic embryo	1	Desiccation	Chaudhury and Chandel, 1995
<i>Prunus persica</i>	Zygotic embryo	1	Desiccation	De Boucault and Brison, 1991
<i>Prunus</i> spp.	Apex	2	Vitrification	Brison <i>et al.</i> , 1995
<i>Ptychosperma macarthurii</i>	Zygotic embryo	1	Desiccation	Normah and Marzalina, 1995

Continued

Table 6.1. Continued

<i>Species</i>	<i>Specimen</i>	<i>No. of accessions</i>	<i>Technique</i>	<i>Reference</i>
<i>Pyrus</i> spp.	Apex	1	Encaps./dehydr.	Dereuddre <i>et al.</i> , 1990
		1		Niino and Sakai, 1992
		1		Scottez <i>et al.</i> , 1992
		11		Scottez, 1993
<i>Quercus</i>	Zygotic embryo	8	Vitrification	Niino <i>et al.</i> , 1992a
<i>Ribes</i>	Apex	1	Desiccation	Pence, 1990
		3	Encaps./dehydr.	Reed and Yu, 1995
		?	Vitrification	Reed, 1992
<i>Saccharum</i>	Apex	7?	Encaps./dehydr.	Paulet <i>et al.</i> , 1993
		14		Engelmann <i>et al.</i> , 1994b
<i>Secale cereale</i>	Protoplast	1	Vitrification	Langis and Steponkus, 1990
<i>Shorea leprosula</i>	Zygotic embryo	1	Desicc., slow cooling	Chai <i>et al.</i> , 1994
		1	Encaps./dehydr.	Engelmann <i>et al.</i> , 1995c
<i>Shorea macrophylla</i>	Zygotic embryo	1	Desicc./slow cooling	Marzalina, 1995
<i>Shorea ovalis</i>	Zygotic embryo	1	Desicc./slow cooling	Normah and Marzalina, 1995
<i>Shorea parvifolia</i>	Zygotic embryo	1	Desicc./slow cooling	Normah and Marzalina, 1995
<i>Solanum</i> spp.	Apex	1	Encaps./dehydr.	Fabre and Dereuddre, 1990
		6		Benson <i>et al.</i> , 1995
		1	Vitrification	Schneibel-Preikstas <i>et al.</i> , 1992c
		5		Lu and Steponkus, 1994
		100	Droplet freezing	Schäfer-Menuhr, 1994, 1995
<i>Swietenia macrophylla</i>	Zygotic embryo	1	Encaps./dehydr.	Marzalina, 1995
<i>Theobroma cacao</i>	Zygotic embryo	1	Pregrowth/desicc.	Pence, 1991
<i>Trifolium repens</i>	Apex	3	Vitrification	Yamada <i>et al.</i> , 1991
<i>Triticum aestivum</i>	Zygotic embryo	1	Pregrowth	Zavala and Sussex, 1986
<i>Veitchia merrillii</i>	Zygotic embryo	1	Desiccation	Chin <i>et al.</i> , 1988
<i>Vitis vinifera</i>	Apex	1	Encaps./dehydr.	Plessis <i>et al.</i> , 1991
		6		Plessis <i>et al.</i> , 1993
<i>Wasabia japonica</i>	Apex	4	Vitrification	Matsumoto <i>et al.</i> , 1994b
		5	Encaps./vitrif.	Matsumoto <i>et al.</i> , 1995
<i>Zea mays</i>	Zygotic embryo	1	Pregrowth	Delvallée <i>et al.</i> , 1989

cold-tolerant species, *in vitro* mother plants or apices can be placed at low temperature for several weeks (Niino and Sakai, 1992; Scottez *et al.*, 1992; Paul, 1994). Mulberry apices are transferred daily on media with progressively increasing sucrose concentrations to initiate dehydration (Niino and Sakai, 1992).

Specimens are usually encapsulated in 3% calcium alginate gel. They are then submitted to the following successive steps: pregrowth, dehydration, freezing, thawing and recovery. Pregrowth is performed in liquid medium enriched with sucrose (0.3 to 1.5 M) for periods varying between 16 hours (Niino and Sakai, 1992) and 10 days (Fabre and Dereuddre, 1990). Replacement of sucrose with other sugars did not improve survival of grape apices (Plessis *et al.*, 1993). Progressive increase in sucrose concentration generally overcomes sensitivity to direct exposure to high sucrose levels which is encountered with some species such as eucalyptus, grape and coffee (Plessis *et al.*, 1991; Poissonier *et al.*, 1992; Engelmann *et al.*, 1994b).

Desiccation is performed using either the air current of a laminar airflow cabinet or silica gel. The latter method is generally preferred because it ensures more precise and reproducible desiccation rates (Paulet *et al.*, 1993). The water content of the beads allowing optimal survival rates is around 20% (fresh weight basis). If pregrowth conditions have been well defined, only limited loss is observed after desiccation with most species. However, banana apices were found to be highly sensitive to both sucrose pregrowth treatment and to even limited desiccation, which induced a drastic survival drop (Panis, 1995).

Desiccated samples are usually frozen by direct immersion in liquid nitrogen. However, modifications in the freezing rate can have different consequences on the survival of different materials. Survival of encapsulated grape apices was higher after slow cooling down to -100°C (Plessis *et al.*, 1991). In contrast, survival of sugar-cane apices was higher after rapid freezing than after slow, controlled cooling (Gonzalez-Arno *et al.*, 1993a,b). Freezing encapsulated carnation apices at cooling rates of between 0.5 and $200^{\circ}\text{C min}^{-1}$ had no effect on survival (Tannoury *et al.*, 1994).

Samples are usually stored in liquid nitrogen at -196°C . Scottez (1993) and Tannoury (1993) showed that survival of pear and carnation apices, respectively, was not modified after 2 and 3 years of storage at -196°C . Several authors have demonstrated that samples can be frozen and conserved in deep-freezers provided that the storage temperature is below that of ice recrystallization (-50 to -70°C). Pear apices were conserved at -75°C for one year (Scottez, 1993) and apple, pear and mulberry apices were stored for 5 months at -135°C (Niino and Sakai, 1992).

Samples are usually placed directly under standard conditions for recovery. However, transitorily modified conditions can enhance recovery

in some cases. Sugar-cane apices are placed in the dark for one week on a medium supplemented with growth hormones (Paulet *et al.*, 1993). Addition of the antioxidant ascorbic acid significantly improved the recovery of sugar-beet apices, which are extremely sensitive to oxidation (Vandenbussche and De Proft, 1995). Extraction of apices from the beads was necessary to allow regrowth of pear and grape apices (Plessis *et al.*, 1991; Scottez *et al.*, 1992). Regrowth of material frozen using the encapsulation-dehydration technique is usually direct and rapid, without callus formation. This is due to the fact that, contrary to what is observed after classical freezing, where many cells are destroyed and callusing is frequently observed during recovery, encapsulation-dehydration preserves the structural integrity of most cells. Therefore, regrowth usually originates from the whole meristematic zone, as observed notably in the case of sugar-cane (Gonzalez-Arnan *et al.*, 1993a).

There are presently four crops (pear, apple, sugar-cane and potato) on which the encapsulation-dehydration technique has been successfully extended to several genotypes or varieties (Table 6.1). In all cases, even though genotypic variations were noted, results were sufficiently high to envisage routine application of the cryopreservation protocols developed.

Vitrification

Vitrification involves treatment ('loading') of samples with cryoprotective substances, dehydration with highly concentrated vitrification solutions, rapid freezing and thawing, removal of cryoprotectants ('unloading') and recovery. Vitrification procedures have been developed for more than 20 different species using protoplasts, cell suspensions, apices and somatic embryos (Table 6.1).

The physiological state of the explants can influence their survival potential. In some cases, the plant material is thus submitted to various treatments before the cryopreservation procedure itself. *In vitro* mother plants can be cultured at low temperature for several weeks (Niino *et al.*, 1992a,b). Explants can be placed on a medium with cryoprotectants (Towill, 1990; Brison *et al.*, 1995) and/or cultured at low temperature for a short period (Yamada *et al.*, 1991). Pepó (1994) indicated that axillary shoot tips are much more sensitive to vitrification than apical shoots.

Loading consists of placing the explants in liquid medium containing cryoprotective substances (sucrose, glycerol, ethylene glycol) for a short period, varying between 5 and 90 min, depending on the material. This reduces the sensitivity of the material to the highly concentrated vitrification solutions.

Vitrification solutions are complex mixtures of cryoprotectants which have been formulated for their high ability to vitrify (i.e. form an amorphous glassy structure). Most solutions employed are derived from ones elaborated by either of two groups: that of Sakai's group (Sakai *et al.*, 1990),

which consists of 22% glycerol, 15% ethylene glycol, 15% polypropylene glycol, 7% DMSO and 0.5 M sorbitol; and that of Steponkus' group (Langis *et al.*, 1989), which comprises 40% ethylene glycol, 15% sorbitol and 6% bovine serum albumin. The duration of contact between explants and the vitrification solutions is a critical parameter, in view of their high toxicity. The dehydration period generally increases with the size of the explants used. Rye protoplasts are dehydrated for 60 s only (Langis and Steponkus, 1990), whereas the optimal dehydration duration of apices of pear and apple is 80 min (Niino *et al.*, 1992b).

Performing the dehydration step at 0°C instead of room temperature reduces the toxicity of vitrification solutions and thus broadens the window of exposure durations ensuring survival of samples. This also allows the manipulation of a large number of samples at the same time. Survival of asparagus embryogenic cell suspensions dropped rapidly after 5 min of dehydration at 25°C but high survival was obtained for dehydration periods between 5 and 60 min if it was performed at 0°C (Nishizawa *et al.*, 1993).

Specimens are then cooled rapidly by direct immersion in liquid nitrogen in order to achieve vitrification of internal solutes. Reduction in the volume of cryoprotectants and the use of plastic straws (500 µl–1 ml) which have a small diameter and a large surface area of contact with the exterior allow an increase in the cooling rate, reaching 990°C min⁻¹ in the case of asparagus cell suspensions frozen in 50 µl of medium in a 500 µl straw (Uragami *et al.*, 1989). Apices of mint and sweet potato were frozen without cryoprotective medium, thus reaching a cooling rate of 4800°C min⁻¹ (Towill, 1990; Towill and Jarret, 1992).

Rewarming of samples is performed as rapidly as possible to avoid devitrification, which would lead to the formation of ice crystals detrimental to cellular integrity. Samples are immersed in a water-bath or liquid medium held at 20–40°C.

Unloading aims at removing progressively the vitrification solution in order to reduce the osmotic shock. Liquid medium containing 1.2 M sucrose or sorbitol is added progressively to dilute the vitrification solution. Explants are then transferred to standard conditions.

Vitrification procedures generally achieve high survival rates. Recovery is usually direct and rapid, even though a few authors have reported callusing and/or abnormal plant development (Towill, 1990; Gonzalez-Benito and Perez, 1994a). Vitrification experiments involving a large range of genotypes have been performed with several species including *Allium*, tea, *Citrus*, apple, mulberry, grape, potato and wasabi (Table 6.1).

Encapsulation-vitrification

This technique is a combination of encapsulation-dehydration and vitrification procedures. Samples are encapsulated in alginate beads, then

subjected to freezing by vitrification. It has been developed by Tannoury *et al.* (1991) with carnation apices, and applied recently by Sakai's group to apices of *Armoracia*, lily and wasabi (Phunchindawan *et al.*, 1994; Matsumoto *et al.*, 1995; Sakai and Matsumoto, 1995).

In the case of carnation, encapsulated apices were pregrown for 16 h with progressively more concentrated sucrose solutions, then incubated for 6 h in a vitrification solution containing ethylene glycol and sucrose, and frozen either rapidly or slowly. Maximum survival was 100% and 92% after rapid and slow cooling, respectively (Tannoury *et al.*, 1991).

The technique employed with wasabi apices was slightly different. Apices were encapsulated in beads containing various loading solutions (glycerol or ethylene glycol and sucrose), then placed in the vitrification solution either at 25°C for 30 min or at 0°C for 70–100 min before rapid freezing. Ninety-five percent of frozen apices recovered growth within 3 days.

Even though this technique has been used with a limited number of species only, it possesses great potential both in terms of efficiency and practicality. Matsumoto *et al.* (1995) mention that the recovery rate of apices frozen using the encapsulation-vitrification technique was 30% higher than with the encapsulation-dehydration technique. A reason for this might be that the alginate capsule reduces the toxicity of the vitrification solution. From a practical point of view, the number of manipulations is reduced as well as the total duration of the procedure.

Desiccation

Desiccation is a very simple procedure since it consists of dehydrating the plant material before rapid freezing by direct immersion in liquid nitrogen. This technique has been applied mainly to zygotic embryos or embryonic axes of various species, including numerous tropical forest trees (see Normah and Marzalina, 1995, for a review), as well as to somatic embryos of several species and to shoot tips of mulberry (Table 6.1).

Various parameters, such as the developmental stage of the embryos at the time of harvest, can greatly influence survival. Mature coffee embryos displayed a higher survival rate than immature ones (Abdelnour-Esquivel *et al.*, 1992). High variability was observed in the survival of embryos of several recalcitrant-seed-producing tree species harvested at different periods (Pence, 1992).

Desiccation is usually performed by placing the embryos in the air current of a laminar airflow cabinet. However, more precise and reproducible desiccation conditions are achieved by placing the embryos in a stream of compressed air (Pammenter *et al.*, 1991) or in an airtight container with silica gel. Optimal survival rates are obtained when the water content of the embryos is around 10–20% (fresh weight basis) (Engelmann, 1992).

Freezing is usually rapid but positive results have been obtained using slow cooling with several tropical forest tree species (Chai *et al.*, 1994; Normah and Marzalina, 1995).

Regrowth of the frozen material usually takes place in standard conditions even though modification of the hormonal balance of the recovery medium proved beneficial with coffee zygotic embryos (Abdelnour-Esquivel *et al.*, 1992; Normah and Vengadasalam, 1992). Direct development of the embryos into plantlets is common but abnormal development patterns have been noted in some cases, such as the nondevelopment of the haustorium in the case of *Howea* and *Veitchia* (Chin *et al.*, 1988), or callusing and / or incomplete development with *Hevea* (Normah *et al.*, 1986), *Castanea* and *Quercus* (Pence, 1992), and oil palm (Engelmann *et al.*, 1995a).

Pregrowth

This technique consists of cultivating the plant material for different durations (hours to weeks, depending on the material) in the presence of cryoprotectants, then freezing rapidly by direct immersion in liquid nitrogen. It was first developed by Monnier and Leddet (1978) with zygotic embryos of *Capsella bursa-pastoris*, then applied to embryos of bean, wheat and maize (Zavala and Sussex, 1986; Delvallée *et al.*, 1989), and more recently to meristematic clumps of banana (Panis, 1995).

Culture for 3 days up to 3 weeks on solid medium with high sucrose concentration (0.3 to 0.9 M) was employed with *Capsella*, maize and banana whereas a short treatment with liquid medium containing polyethylene glycol, glucose and DMSO was sufficient to ensure survival of bean and wheat zygotic embryos.

The pregrowth technique was successfully applied to five banana cultivars, resulting in survival rates ranging between 6 and 42.5% (Panis, 1995). Modifications in the sucrose concentration and the duration of the pregrowth treatment should lead to improvements in the survival rate.

Pregrowth-desiccation

Pregrowth-desiccation is a combination of the two techniques described immediately above. In this technique, samples are treated with cryoprotectants, partially desiccated, cooled and rewarmed rapidly. This technique has been applied to a limited number of specimens only (Table 6.1): somatic embryos of oil palm, date palm, coffee, pea and melon; stem segments of *in vitro* plantlets of asparagus; microspore embryos of rapeseed; and zygotic embryos of coconut.

Culture of samples on media containing cryoprotectants usually takes place before desiccation. However, coconut zygotic embryos are dehydrated before preculture with cryoprotectants (Assy-Bah and Engelmann, 1992a). Sugars (sucrose, glucose) are generally employed for preculture. The duration of the preculture varies between 20 h for coconut (Assy-Bah

and Engelmann, 1992a) and 7 days for oil palm (Dumet *et al.*, 1993).

Various methods are employed for desiccation: coconut embryos are placed in the air current of a laminar airflow cabinet (Assy-Bah and Engelmann, 1992a); asparagus stem segments and oil-palm somatic embryos are dehydrated using silica gel (Uragami *et al.*, 1990; Dumet *et al.*, 1993); melon somatic embryos are placed over a saturated salt solution (Shimonishi *et al.*, 1991); and date palm, coffee and pea somatic embryos are dehydrated using a stream of compressed air (Mycock *et al.*, 1995). The optimal water content (fresh weight basis) varies between 11.8% for melon and 25–30% for oil-palm somatic embryos.

Desiccated samples are frozen rapidly by direct immersion in liquid nitrogen. Specimens are usually stored in liquid nitrogen. However, Dumet *et al.* (1994) have demonstrated that oil-palm somatic embryos could be conserved for 6 months in a deep freezer at -80°C , i.e. below the glass transition temperature, without any modification in their recovery rate.

After rewarming, samples are usually placed directly under standard culture conditions. However, oil-palm somatic embryos are cultured on media with progressively reduced sucrose concentration, with transitory addition of 2,4-dichlorophenoxyacetic acid (2,4-D) to stimulate reproliferation (Dumet *et al.*, 1993).

Recovery of samples is usually rapid and direct. Alterations in the regrowth pattern of cryopreserved specimens were observed with coconut and oilseed rape microspore embryos. The haustorium of cryopreserved coconut embryos browned rapidly and did not develop further (Assy-Bah and Engelmann, 1992a). Only half of frozen rapeseed microspore embryos developed directly into plantlets, whereas the other half produced callus and/or secondary embryos (Uragami *et al.*, 1993).

Pregrowth-desiccation has been tested with four varieties of coconut, giving recovery rates ranging between 33 and 93% (Assy-Bah and Engelmann, 1992a). This technique is routinely employed for the long-term storage of 80 clones of oil-palm somatic embryos (Dumet, 1994).

Droplet freezing

The droplet freezing technique has presently been applied to potato apices only (Schäfer-Menuhr, 1994). Apices are pretreated for 2–3 h in liquid medium supplemented with DMSO, placed on an aluminium foil in 5 μl droplets of cryoprotective medium and frozen rapidly by direct immersion in liquid nitrogen. This procedure is adapted from the classical procedure developed by Kartha *et al.* (1982) with cassava where apices placed in droplets of liquid medium were frozen slowly in a programmable freezer.

The droplet freezing method has been successfully applied to 100 varieties of potato with recovery rates ranging between 5 and 100% (Schäfer-Menuhr, 1995). Around 25 000 apices are stored for the long term in liquid nitrogen at DSM, Braunschweig, Germany.

Genetic stability of in vitro conserved material

Genetic conservation is based on the assumption that the material is stored under conditions ensuring genetic stability. However, there are various factors linked with *in vitro* culture / conservation procedures which can be a source of variation. These factors, as well as the various approaches for assessing the genetic stability of plants recovered from *in vitro* culture, have been reviewed recently by Harding (1995). Genetic variation can pre-exist in the collected material, possibly linked to its genetic structure, such as in sugar-cane or banana where polyploids are more prone to instability than diploids. It can also arise from somaclonal variation, particularly if the propagation system includes a dedifferentiated phase or from stresses imposed by the conservation procedures. The information available on the genetic stability of plant material conserved *in vitro* is presented below.

Slow growth

Different genotypes do not grow at the same rate when placed in culture. This induces a risk of selection when plants are placed under stress conditions, such as slow growth. It is thus important to minimize the risk of selection. One of the most efficient ways is to use, whenever possible, differentiated, organized structures such as shoot tips or embryos for which the risks of variation are minimal (D'Amato, 1985).

A limited number of experiments only have been performed to assess the genetic stability of plants during slow-growth storage. The most exhaustive study has been performed in the framework of a joint project between CIAT and IPGRI which aimed at monitoring the genetic stability of cassava shoot cultures stored *in vitro* (CIAT-IPGRI, 1994). Isozyme and DNA analysis as well as examination of the morphology of plants regrown in the field did not reveal any modification after 10 years of storage under slow growth. In contrast, growth of potato shoot tips on mannitol-supplemented medium, which caused morphological changes in the material, was correlated with DNA hypermethylation, which may be an adaptive response to conditions of high osmotic stress (Harding, 1994). This clearly illustrates the importance of defining storage conditions which are likely to cause as little stress as possible to the plant material.

Cryopreservation

Cryopreservation involves a series of stresses which might lead to modifications in the recovered cultures and regenerated plants. It is therefore necessary to verify that the genetic stability of the cryopreserved material is not altered before using this technique routinely for the long-term storage of germplasm. There is now increasing evidence that, provided the cryopreservation technique applied ensures the greatest possible maintenance of the integrity of the frozen specimen, there will be no modification at the

phenotypic, biochemical, chromosomal or molecular level attributable to cryopreservation; none has been reported to date.

It has been shown that cell suspensions of numerous species maintain their biosynthetic and morphogenic potential after cryopreservation (Withers, 1985; Kartha, 1987; Seitz, 1987). The only exception concerns lavender cell suspensions exposed to successive freeze-thaw cycles, for which the number of colonies recovered from cryopreserved cells increased with the number of freeze-thaw cycles (Watanabe *et al.*, 1985). However, no modifications were noted in the biosynthetic capacities of cryopreserved cells, suggesting a change in population structure rather than genetic change. A similar observation was made by Bercetche *et al.* (1990), who noted that cryopreserved *Picea abies* embryogenic calluses recovered faster than nonfrozen controls; nonembryogenic tissues were killed by freezing, thus leading to the production of a more homogeneous population, with a higher embryogenic potential. This suggests that cryopreservation could be used as a tool to 'rejuvenate' cultures when their proliferation capacities are decreasing.

Plants regenerated from cryopreserved apices of strawberry and cassava were phenotypically normal (Kartha *et al.*, 1980; Bajaj, 1983). No differences were noted in the vegetative and floral development of several hundreds of oil palms regenerated from control and cryopreserved somatic embryos (Engelmann, 1991a). In contrast, Fukai *et al.* (1994b) showed that a high percentage of plants regenerated from apices of a periclinally chimeric chrysanthemum cultivar frozen using a classical protocol had an altered flower colour. These apices had been severely harmed during freezing, leading to callusing during recovery. Regenerated plants had thus an adventitious origin which explained the disturbance in the initial chimeric structure. The same group compared the effect of freezing carnation apices using a classical protocol and the encapsulation-dehydration technique (Fukai *et al.*, 1994a). Encapsulation-dehydration ensured 100% recovery after freezing, and regrowth was rapid and direct. In contrast, after slow, controlled freezing, the recovery rate was 50% only, and callusing was observed during regrowth. The two above examples underline the importance of selecting the most appropriate freezing technique for any given material, not only as regards recovery rates but also as regards recovery pattern and genetic stability.

Electrophoretic profiles of two enzymatic systems were comparable in plants regenerated from control and cryopreserved apices of sugar-cane (Paulet *et al.*, 1994) and sweet orange somatic embryos (Marin *et al.*, 1993). The ploidy level of plants regenerated from oilseed rape somatic embryos and sensitive dihaploids of potato was not modified by cryopreservation (Uragami *et al.*, 1993; Ward *et al.*, 1993). Restriction fragment length polymorphism (RFLP) patterns of plants regenerated from sugar-cane embryogenic cell suspensions were identical to those of unfrozen controls

(Chowdhury and Vasil, 1993). Several molecular types were uncovered in plants recovered from both control and cryopreserved sugar-cane apices, thus indicating that the variation was not due to freezing but was pre-existing among the *in vitro* mother plants (Glaszmann *et al.*, 1996).

Finally, DNA analysis showed that a gene integrated in the genome of a navel orange cell suspension was maintained after freezing and one-year storage in liquid nitrogen (Kobayashi *et al.*, 1994).

Comparative studies

A limited amount of work has been published on the comparative effects of slow growth and cryogenic storage on the stability of plant material. Mannonen *et al.* (1990) showed that the production of secondary metabolites by cell cultures of *Panax ginseng* and *Catharanthus roseus* decreased drastically after 6 months of culture either in standard conditions (i.e. with weekly subcultures) or in slow growth under mineral oil, whereas the productivity of cell suspensions cryopreserved and stored in liquid nitrogen during the same period was identical to that of the original cultures.

Modifications in the RFLP pattern were observed in potato plants stored for 6 months under slow growth on a medium supplemented with mannitol, whereas no such modifications were noted in plants regenerated from cryopreserved apices (Harding, 1991).

6.5. Present use of *in vitro* conservation techniques

Classical *in vitro* conservation techniques have been developed for a wide range of species, including temperate woody plants (Aitken-Christie and Singh, 1987), fruit trees (Withers, 1992), horticultural species (Engelmann, 1991b), as well as numerous tropical species (Dodds, 1991; Engelmann, 1991a; Zakri *et al.*, 1991).

However, a recent FAO survey (Anonymous, 1994) indicated that only 37 600 accessions are conserved *in vitro* worldwide. Slow-growth conservation is used routinely for the conservation of only a few species, including banana, potato and cassava, in regional and international germplasm conservation centres such as INIBAP, CIP, CIAT and IITA.

Alternative medium-term conservation techniques are still at the experimental stage. Low-oxygen storage might be interesting for storing cold-sensitive tropical species since it allows growth reduction without decreasing the temperature. However, complementary experiments involving additional species and longer storage periods are still required.

Medium-term storage of desiccated somatic embryos will be mainly applicable for the management of large-scale production of elite genotypes. Encapsulated apices might be employed for medium-term genetic

resources conservation. Research is still needed to refine the protocols and extend the storage periods.

Cryopreservation procedures have been developed for around 100 different plant species cultured in various ways including cell suspensions, calluses, apices, and zygotic and somatic embryos (Withers, 1992; Kartha and Engelmann, 1994; Engelmann *et al.*, 1995b; Withers and Engelmann, 1997). Most of this work has been performed in the framework of academic studies and has involved only one or a few genotypes. However, there are recent reports of large-scale experimentation of classical freezing techniques involving cell suspensions from several hundred genotypes (Cyr *et al.*, 1994; Park *et al.*, 1994). In the case of organized structures such as apices or embryos, the development of new cryopreservation procedures has also allowed experiments involving a relatively wide range of genotypes / varieties (Fukai *et al.*, 1991a,b). In the case of potato, apices from more than 100 different varieties have been successfully frozen and around 25 000 apices are presently stored in liquid nitrogen (Schäfer-Menuhr, 1994).

There is an increasing number of cases where cryopreservation techniques can be considered operational. However, their routine application is mostly restricted to the conservation of cell lines in research laboratories (Withers, 1985). The only example of routine application of cryopreservation to another type of material is oil palm, where 80 clones of somatic embryos are stored in liquid nitrogen and samples thawed upon request for plant production (Dumet, 1994).

6.6. Conclusion

Significant advances have been made during the past decade in the development of *in vitro* techniques for the conservation of plant genetic resources. *In vitro* collecting, slow-growth and cryopreservation procedures are now available for a wide range of plant species. There is an increasing number of cases where slow growth and cryopreservation could be used to improve the conservation of genetic resources of problem species. Obviously, slow-growth techniques are in a more advanced state of development and should become more widely applied once their flexibility, simplicity and practicality are clearly demonstrated.

For most undifferentiated cultures, such as cell suspensions and calluses, cryopreservation procedures can be set up rapidly with only slight adaptations of already published procedures. For differentiated materials such as embryos or apices, even though the situation has considerably improved with the emergence of new (vitrification-based) freezing procedures in recent years, the development of a cryopreservation protocol for any new material still requires considerable inputs. However, it is time to

move the experimentation from the research laboratory to the genebank, or at least to research units which have a conservation mandate. This will allow the development of methodologies which meet the criteria required in the genebank context and the incorporation into *in vitro* conservation research and development the knowledge of the genebank staff in relation to genetic stability and genetic characterization (Withers and Engelmann, 1997).

Finally, it is important to stress that *in vitro* conservation should not be seen as a replacement for conventional *in situ* and *ex situ* approaches. *In vitro* conservation offers genebank curators a set of additional tools to allow them to improve the conservation of germplasm collections for which they are responsible.

References

- Abdelnour-Esquivel, A., Villalobos, V. and Engelmann, F. (1992) Cryopreservation of zygotic embryos of *Coffea* spp. *Cryo-Letters* 13, 297–302.
- Aitken-Christie, J., Singh, A.P. (1987) Cold storage of tissue cultures. In: Bonga, J.M. and Durzan, D.J. (eds) *Cell and Tissue Culture in Forestry*. Martinus Nijhoff, Dordrecht, pp. 285–304.
- Allen, J.B. and Lass, R.A. (1983) London Cocoa Trade Amazon Project: Final Report, Phase 1. *Cocoa Growers Bulletin* No. 34.
- Altman, D.W., Ryxell, P.A., Koch, S.D. and Howell, C.R. (1990) *Gossypium* germplasm conservation augmented by tissue culture techniques for field collecting. *Economic Botany* 44, 106–113.
- Anonymous (1994) *Survey of Existing Data on ex situ Collections of Plant Genetic Resources for Food and Agriculture*. Food and Agriculture Organization, Rome.
- Ashwood-Smith, M.J. and Farrant, J. (1980) *Low Temperature Preservation in Medicine and Biology*. Pitman Medical, Tunbridge Wells.
- Assy-Bah, B. and Engelmann, F. (1992a) Cryopreservation of mature embryos of coconut (*Cocos nucifera* L.) and subsequent regeneration of plantlets. *Cryo-Letters* 13, 117–126.
- Assy-Bah, B. and Engelmann, F. (1992b) Cryopreservation of immature embryos of coconut (*Cocos nucifera* L.). *Cryo-Letters* 13, 67–74.
- Assy-Bah, B., Durand-Gasselín, T. and Pannetier, C. (1987) Use of zygotic embryo culture to collect germplasm of coconut (*Cocos nucifera* L.). *FAO/IBPGR Plant Genetic Resources Newsletter* 71, 4–10.
- Assy-Bah, B., Durand-Gasselín, T., Engelmann, F. and Pannetier, C. (1989) Culture *in vitro* d'embryons zygotiques de cocotier (*Cocos nucifera* L.). Méthode, révisée et simplifiée, d'obtention de plants de cocotiers transférables au champ. *Oléagineux* 44, 515–523.
- Attree, S.M. and Fowke, L.C. (1993) Embryogeny of gymnosperms: advances in synthetic seed technology of conifers. *Plant Cell, Tissue and Organ Culture* 35, 1–35.
- Augereau, J.M., Courtois, D. and Pétiard, V. (1986) Long term storage of callus cultures at low temperatures or under mineral oil layer. *Plant Cell Reports* 5, 372–376.

- Bajaj, Y.P.S. (1983) Cassava plants from meristem cultures freeze-preserved for three years. *Field Crop Research* 7, 161–167.
- Banerjee, N. and De Langhe, E. (1985) A tissue culture technique for rapid clonal propagation and storage under minimal growth conditions of *Musa* (banana and plantain). *Plant Cell Reports* 4, 351–354.
- Benson, E.E. (1990) *Free Radical Damage in Stored Plant Germplasm*. International Board for Plant Genetic Resources, Rome, 128 pp.
- Benson, E.E., Chabrilange, N. and Engelmann, F. (1992) A comparison of cryopreservation methods for the long-term *in vitro* conservation of cassava. In: *Proceedings of SLTB Autumn Meeting, Stirling, UK, 1992*.
- Benson, E.E., Wilkinson, M., Donally, A., Ekuere, U. and Lyon, J. (1995) Histological and morphogenetic assessment of plants regenerated from cryopreserved apices of potato. *Cryo-Letters* 16, 66.
- Bercetche, J., Galerne, M. and Dereuddre, J. (1990) Augmentation des capacités de régénération de cals embryogènes de *Picea abies* (L.) Karst après congélation dans l'azote liquide. *Comptes Rendus de l'Académie des Sciences, Paris, Série III*, 310, 357–363.
- Bertrand-Desbrunais, A. (1991) La conservation des ressources génétiques des caféiers. Thèse d'Université, Université Paris 6, 259 pp.
- Bigot, C. (1987) *In vitro* manipulation of higher plants: some achievements, problems and perspectives. In: Boccon-Gibod, J., Benbadis, A. and Shont, K.C. (eds) *Proceedings of IAPTC French-British Meeting, 8–9 October, Cell Culture Techniques Applied to Plant Production and Plant Breeding, Angers, France*, pp. 5–17.
- Brearley, J., Henshaw, G.G., Davey, C., Taylor, N.J. and Blakesley, D. (1995) Cryopreservation of *Fraxinus excelsior* zygotic embryos. *Cryo-Letters* 16, 215–218.
- Bridgen, M.P. and Staby, G.L. (1981) Low pressure and low oxygen storage of *Nicotiana tabacum* and *Chrysanthemum* × *Morifolium* tissue cultures. *Plant Science Letters* 22, 177–186.
- Brison, M., de Boucaud, M.T. and Dosba, F. (1995) Cryopreservation of *in vitro* grown shoot tips of two interspecific *Prunus* rootstocks. *Plant Science* 105, 235–242.
- Caplin, S.M. (1959) Mineral oil overlay for conservation of plant tissue cultures. *American Journal of Botany* 46, 324–329.
- CATIE (1997) *In vitro Collecting*. Manual Based on IBPGR-CATIE Training Course, FAO, Rome.
- Chai, Y.M., Marzalina, M. and Krishnapillay, B. (1994) Effect of sucrose as a cryoprotective agent in the cryopreservation of some Dipterocarp species. *Proceedings of National Biotechnology VI*.
- Chatti-Dridi, B. (1988) Expériences préliminaires sur la conservation à court terme et l'amélioration de la micropropagation *in vitro* dans le cas du pêcher (*Prunus persica*). Thesis, Mémoire de fin d'études, ENSH Versailles.
- Chaudhury, R. and Chandel, K.P.S. (1995) Cryopreservation of embryonic axes of almond (*Prunus amygdalus* Batsch.) seeds. *Cryo-Letters* 16, 51–56.
- Chaudhury, R., Radhamani, J. and Chandel, K.P.S. (1991) Preliminary observations on the cryopreservation of desiccated embryonic axes of tea (*Camellia sinensis* (L.) O. Kuntze) seeds for genetic conservation. *Cryo-Letters* 12, 31–36.
- Chen, T.H.H., Kartha, K.K., Leung, N.L., Kurz, G.W.G., Chatson, K.B. and

- Constabel, F.C. (1984) Cryopreservation of alkaloid-producing cell cultures of periwinkle (*Catharanthus roseus*). *Plant Physiology* 73, 726–731.
- Chin, H.F. and Pritchard, H.W. (1988) *Recalcitrant Seeds, a Status Report*. International Board for Plant Genetic Resources, Rome.
- Chin, H.F., Krishnapillay, B. and Alang, Z. (1988) Cryopreservation of *Veitchia* and *Howea* palms embryos: non-development of the haustorium. *Cryo-Letters* 9, 372–379.
- Chowdhury, M.K.U. and Vasil, I.K. (1993) Molecular analysis of plants regenerated from embryogenic cultures of hybrid sugarcane cultivars (*Saccharum* spp.). *Theoretical and Applied Genetics* 86, 181–188.
- CIAT-IPGRI, (1994) *Establishment and Operation of a Pilot in vitro Active Genebank*. A co-publication of CIAT, Cali, Colombia, and IPGRI, Rome, Italy.
- Corbineau, F., Engelmann, F. and Côme, D. (1990) Ethylene production as an indicator of chilling injury in oil palm (*Elaeis guineensis* Jacq.). *Plant Science* 71, 29–34.
- Cyr, D.R., Lazaroff, W.R., Grimes, S.M.A., Quan, G., Bethune, T.D., Dunstan, D.I. and Roberts, D.R. (1994) Cryopreservation of interior spruce (*Picea glauca engelmanni* complex) embryogenic cultures. *Plant Cell Reports* 13, 574–577.
- D'Amato, F. (1985) Cytogenetics of plant cell and tissue cultures and their regenerates. In: Conger, B.V. (ed.) *Critical Review in Plant Sciences*. CRC Press, Boca Raton, FL, pp. 73–112.
- de Boucaud, M.T., Brison, M. and Negrier, P. (1994) Cryopreservation of walnut somatic embryos. *Cryo-Letters* 15, 151–160.
- De Boucault, M.T. and Brison, M. (1991) Cryopreservation of embryonic axes of *Prunus persica* L. Batsch. In: *Proceedings of the Annual Symposium of the Society for Cryobiology*. *Cryo* 91, Leuven, Belgium, 7–12 July 1991, p. 102.
- Dekkers, A.J., Rao, A.N. and Goh, C.J. (1991) *In vitro* storage of multiple shoot cultures of ginger at ambient temperatures. *Scientia Horticulturae* 47, 157–167.
- Delvallée, I., Guillaud, J., Beckert, M. and Dumas, C. (1989) Cryopreservation of immature maize embryos after freeze-hardening in the ear and *in vitro*. *Plant Science* 60, 129–136.
- Dereuddre, J. and Engelmann, F. (1987) The use of cryopreservation for setting up banks of plant germplasm. In: Boccon-Gibod, J., Benbadis, A. and Shont, K.C. (eds) *Proceedings of Coll. Franco-Britannique IAPTC, 8–9 October, Angers, France*, pp. 48–78.
- Dereuddre, J., Fabre, J. and Bassaglia, C. (1988) Resistance to freezing in liquid nitrogen of carnation (*Dianthus caryophyllus* L. var. Eolo) apical and axillary shoot tips excised from differently aged *in vitro* plantlets. *Plant Cell Reports* 7, 170–173.
- Dereuddre, J., Scottez, C., Arnaud, Y. and Duron, M. (1990) Résistance d'apex caulinaires de vitroplants de poirier (*Pyrus communis* L. cv. Beurré Hardy), enrobés dans l'alginate à une déshydratation puis à une congélation dans l'azote liquide: effet d'un endurcissement préalable au froid. *Comptes Rendus de l'Académie des Sciences, Paris Série III*, 310, 317–323.
- Dereuddre, J., Hassen, M., Blandin, S. and Kaminski, M. (1991a) Resistance of alginate-coated somatic embryos of carrot (*Daucus carota* L.) to desiccation and freezing in liquid nitrogen: 2. Thermal analysis. *Cryo-Letters* 12, 135–148.
- Dereuddre, J., Hassen, N., Blandin, S. and Kaminski, M. (1991b) Resistance of

- alginate-coated somatic embryos of carrot (*Daucus carota* L.) to desiccation and freezing in liquid nitrogen. 1. Effects of preculture. *Cryo-Letters* 12, 125–134.
- Dodds, J.H. (1991) *In vitro Methods for Conservation of Plant Genetic Resources*. Chapman & Hall, London.
- Dodds, J.H. (1992) Biotechnological techniques applied to potato and sweet potato improvement for developing countries. In: Sasson, A. and Costarini, V. (eds). *Plant Biotechnologies for Developing Countries. Proceedings of Symposium CTA and FAO, Luxembourg, 26–30 June 1989*. Chayce, UK, pp. 221–227.
- Dorion, N., Regnard, J.L., Serpette, I. and Bigot, C. (1994) Effects of temperature and hypoxic atmosphere on preservation and further development of *in vitro* shoots of peach ('Armking') and peach × almond ('GF 677'). *Scientia Horticulturae* 57, 201–213.
- Druart, P. (1985) *In vitro* germplasm preservation technique for fruit trees. In: Schäfer-Menuhr, A. (ed.) *In vitro Techniques – Propagation and Long-term Storage*. Nijhoff/Junk for CEC, Dordrecht, pp. 167–171.
- Dumet, D. (1994) Cryoconservation des massifs d'embryons somatiques de palmier à huile (*Elaeis guineensis* Jacq.) par déshydratation-vitrification. Etude du rôle du saccharose pendant le prétraitement. Thèse d'Université. Université Paris 6, 115 pp.
- Dumet, D., Engelmann, F., Chabrilange, N. and Duval, Y. (1993) Cryopreservation of oil palm (*Elaeis guineensis* Jacq.) somatic embryos involving a desiccation step. *Plant Cell Reports* 12, 352–355.
- Dumet, D., Engelmann, F., Chabrilange, N. and Duval, Y. (1994) Effect of desiccation and storage temperature on the conservation of cultures of oil palm somatic embryos. *Cryo-Letters* 15, 85–90.
- Dussert, S., Mauro, M.C., Deloire, A., Hamon, S. and Engelmann, F. (1991) Cryopreservation of grape embryogenic cell suspensions. 1. Influence of pre-treatment, freezing and thawing conditions. *Cryo-Letters* 12, 287–298.
- Dussert, S., Mauro, M.C. and Engelmann, F. (1992) Cryopreservation of grape embryogenic cell suspensions. 2. Influence of post-treatment conditions and application to different strains. *Cryo-Letters* 13, 12–22.
- Elias, K. (1988) *In vitro* culture and plant genetic resources, a new approach: *in vitro* collecting. *Lettere d'Informazione Istituto Agronomico Mediterraneo, Valenzano* 3, 33–34.
- Ellis, R.E., Hong, T. and Roberts, E.H. (1990) An intermediate category of seed storage behaviour? I. Coffee. *Journal of Experimental Botany* 41, 1167–1174.
- Ellis, R.E., Hong, T., Roberts, E.H. and Soetisna, U. (1991) Seed storage behaviour in *Elaeis guineensis*. *Seed Science Research* 1, 99–104.
- Engelmann, F. (1990) Use of atmospheres with low oxygen contents for the storage of oil palm (*Elaeis guineensis* Jacq.) somatic embryo cultures. *Comptes Rendus de l'Académie des Sciences, Paris Série III*, 310, 679–684.
- Engelmann, F. (1991a) *In vitro* conservation of tropical plant germplasm – review. *Euphytica* 57, 227–243.
- Engelmann, F. (1991b) *In vitro* conservation of horticultural species. *Acta Horticulturae* 298, 327–334.
- Engelmann, F. (1992) Cryopreservation of embryos. In: Dattée, Y., Dumas, C. and Gallais, A. (eds) *Reproductive Biology and Plant Breeding*. Springer Verlag, Berlin, pp. 281–290.

- Engelmann, F. (1994) Cryopreservation for the long-term conservation of tropical crops of commercial importance. In: *Proceedings of the International Symposium on the Application of Plant In vitro Technology*. Universiti Pertanian Malaysia, Selangor, 16–18 November 1993, pp. 64–77.
- Engelmann, F. and Dereuddre, J. (1988) Cryopreservation of oil palm somatic embryos: importance of the freezing process. *Cryo-Letters* 9, 220–235.
- Engelmann, F., Duval, Y. and Dereuddre, J. (1985) Survie et prolifération d'embryons somatiques de palmier à huile (*Elaeis guineensis* Jacq.) après congélation dans l'azote liquide. *Comptes Rendus de l'Académie des Sciences, Paris Série III*, 301, 111–116.
- Engelmann, F., Dambier, D. and Ollitrault, P. (1994a) Cryopreservation of cell suspensions and embryogenic calluses of *Citrus* using a simplified freezing process. *Cryo-Letters* 15, 53–58.
- Engelmann, F., Benson, E.E., Chabrillange, N., Gonzalez-Arno, M.T., Mari, S., Michaux-Ferrière, N., Paulet, F., Glaszmann, J.C. and Charrier, A. (1994b) Cryopreservation of several tropical plant species using encapsulation / dehydration of apices. In: *Proceedings of VIIIth IAPTC Meeting, Firenze, Italy*.
- Engelmann, F., Chabrillange, N., Dussert, S. and Duval, Y. (1995a) Cryopreservation of zygotic embryos and kernels of oil palm (*Elaeis guineensis* Jacq.). *Seed Science Research* 5, 81–86.
- Engelmann, F., Dumet, D., Chabrillange, N., Abdelnour-Esquivel, A., Assy-Bah, B., Dereuddre, J. and Duval, Y. (1995b) Cryopreservation of zygotic and somatic embryos from recalcitrant and intermediate-seed species. *IPGRI/FAO Newsletter* 103, 27–31.
- Engelmann, F., Chandel, K.P.S., Krishnapillay, D. and Hor, Y.L. (1995c) Interest of cryopreservation for the long-term storage of recalcitrant-seed species. In: *Abstracts of the 24th ISTA Congress, Copenhagen, Denmark, 7–16 June*.
- Espinoza, N., Lizarraga, R., Siguenas, C., Buitron, F., Bryan, J. and Dodds, J.H. (1992) *Tissue Culture: Micropropagation, Conservation and Export of Potato Germplasm*. CIP Research Guide 1. International Potato Center, Lima, Peru, 19 pp.
- Fabre, J. and Dereuddre, J. (1990) Encapsulation-dehydration: a new approach to cryopreservation of *Solanum* shoot-tips. *Cryo-Letters* 11, 413–426.
- Fahy, G.M., MacFarlane, D.R., Angell, C.A. and Meryman, H.T. (1984) Vitrification as an approach to cryopreservation. *Cryobiology* 21, 407–426.
- Fay, M.F. (1994) In what situations is *in vitro* culture appropriate to plant conservation? *Biodiversity and Conservation* 3, 176–183.
- Finkle, B.J., Zavala, M.E. and Ulrich, J.M. (1985) Cryoprotective compounds in the viable freezing of plant tissues. In: Kartha, K.K. (ed.) *Cryopreservation of Plant Cells and Organs*. CRC Press, Boca Raton, Florida, pp. 75–114.
- Florin, B. (1989) Etude de différentes voies de conservation d'embryons, de tissus et de cellules de végétaux cultivés *in vitro*: applications de l'hypoxie et de la cryoconservation. PhD Thesis, Université François Rabelais, Tours.
- Fu, J.R., Zhang, B.Z., Wang, X.P., Qiao, Y.Z. and Huang, X.L. (1990) Physiological studies on desiccation, wet storage and cryopreservation of recalcitrant seeds of three fruit species and their excised embryonic axes. *Seed Science and Technology* 18, 743–754.
- Fukai, S., Goi, M. and Tanaka, M. (1991a) Cryopreservation of shoot tips of

- Chrysanthemum morifolium* and related species native to Japan. *Euphytica* 54, 201–204.
- Fukai, S., Goi, M. and Tanaka, M. (1991b) Cryopreservation of shoot tips of Caryophyllaceae ornamentals. *Euphytica* 56, 149–153.
- Fukai, S., Togashi, M. and Goi, M. (1994a) Cryopreservation of *in vitro* grown *Dianthus* by encapsulation-dehydration. *Technical Bulletin of Faculty of Agriculture, Kagawa University* 46, 101–107.
- Fukai, S., Goi, M. and Tanaka, M. (1994b) The chimeric structure of the apical dome of chrysanthemum (*Dendranthema grandiflorum* (Ramat.) Kitam.) is affected by cryopreservation. *Scientia Horticulturae* 57, 347–351.
- Glaszmann, J.C., Rott, P. and Engelmann, F. (1996) Role of *in vitro* maintenance of sugarcane for germplasm conservation and exchange. In: Croft, B.J., Piggitt, C.M., Wallis, E.S. and Hogarth, O.M. (eds) *Sugarcane Germplasm Conservation and Exchange*. ACIAR Proceedings No. 67, 67–70.
- Gonzalez-Arnan, M.T., Engelmann, F., Huet, C. and Urra, C. (1993a) Cryopreservation of encapsulated apices of sugarcane: effect of freezing procedure and histology. *Cryo-Letters* 14, 303–308.
- Gonzalez-Arnan, M.T., Engelmann, F., Urra, C. and Lynch, P. (1993b) Cryoconservacion de meristemas apicales de plantas *in vitro* de cana de azucar mediante el metodo de encapsulacion/deshidratacion. *Biotechnologia Aplicada* 10, 225–228.
- Gonzalez-Benito, M.E. and Perez, C. (1994a) Cryopreservation of *Centaurea* shoot tips by vitrification. *Botanical Gardens Micropropagation Newsletter* 1, 82–84.
- Gonzalez-Benito, M.E. and Perez, C. (1994b) Cryopreservation of embryonic axes of two cultivars of hazelnut (*Corylus avellana* L.). *Cryo-Letters* 15, 41–46.
- Gonzalez-Rio, F., Gurriaran, M.J., Gonzalez-Benito, E. and Revilla, M.A. (1994) Desiccation and cryopreservation of olive (*Olea europaea* L.) embryos. *Cryo-Letters* 15, 337–342.
- Gray, D.J. and Purohit, A. (1991) Somatic embryogenesis and development of synthetic seed technology. *Critical Review in Plant Science* 10, 33–61.
- Grout, B.W.W. and Morris, G.J. (1987) *The Effects of Low Temperatures on Biological Systems*. Edward Arnold, London.
- Grout, B.W.W., Shelton, K. and Pritchard, H.W., (1983) Orthodox behaviour of oil palm seed and cryopreservation of the excised embryo for genetic conservation. *Annals of Botany* 52, 381–384.
- Guarino, L., Rao, R. and Reid R. (1995) *Collecting Plant Genetic Diversity, Technical Guidelines*. CAB International, Wallingford, Oxon, UK.
- Harding, K. (1991) Molecular stability of the ribosomal RNA genes in *Solanum tuberosum* plants recovered from slow growth and cryopreservation. *Euphytica* 55, 141–146.
- Harding, K. (1994) The methylation status of DNA derived from potato plants recovered from slow growth. *Plant Cell, Tissue and Organ Culture* 37, 31–38.
- Harding, K. (1995) Approaches to assess the genetic stability of plants recovered from *in vitro* culture. In: Normah, M.N. and Narimah, M.K. (eds) *Proceedings of the International Workshop on In vitro Conservation of Plant Genetic Resources, 4–6 July, 1995, Kuala Lumpur*. Plant Biotechnology Laboratory, Universiti Kebangsaan Malaysia, Bangi, Serdang (in press).
- Harding, K., Benson, E.E. and Smith, H. (1991) The effects of *in vitro* culture period

- on the recovery of cryopreserved shoot-tips of *Solanum tuberosum*. *Cryo-Letters* 12, 17–22.
- Hasan, S.M.Z. and Takagi, H. (1995) Alginate-coated nodal segment of yam (*Dioscorea* spp.) for germplasm exchange and distribution. *IPGRI/FAO Plant Genetic Resources Newsletter* 103, 32–35.
- Hatanaka, T., Yasuda, T., Yamaguchi, T. and Sakai, A. (1994) Direct regrowth of encapsulated somatic embryos of coffee (*Coffea canephora*) after cooling in liquid nitrogen. *Cryo-Letters* 15, 47–52.
- Hirata, K., Phunchindawan, M., Kanemoto, M., Miyamoto, K. and Sakai, A. (1995) Cryopreservation of shoot primordia induced from horseradish hairy root cultures by encapsulation and two-step dehydration. *Cryo-Letters* 16, 122–127.
- Hor, Y.L., Stanwood, P.C. and Chin, H.F. (1990) Effects of dehydration of freezing characteristics and survival in liquid nitrogen of three recalcitrant seeds. *Pertanika* 13, 309–314.
- Hor, Y.L., Stanwood, P.C. and Chin, H.F. (1993) Cryopreservation of *Coffea liberica* seeds and embryos following desiccation and freezing treatments. *Pertanika* 16(2), 75–80.
- Hvoslef-Heide, A.K. (1992) Effect of pre-storage conditions on storage of *in vitro* cultures of *Nephrolepis exaltata* (L.) Schott and *Cordyline fruticosa* (L.) A. Chev. *Plant Cell, Tissue and Organ Culture* 28, 167–174.
- International Board for Plant Genetic Resources (IBPGR) (1984) *The Potential for Using In vitro Techniques for Germplasm Collection*. IBPGR, Rome.
- Ishikawa, M., Tandon, P., Suzuki, M. and Yamagishi-Ciampi, A. (1994) Cryopreservation of Bromegrass (*Bromus inermis* Leyss) suspension cultured cells with slow prefreezing method and with vitrification method. *Cryobiology* 31, 578–579.
- Janeiro, L.V., Vieitez, A.M. and Ballester, A. (1995) Cold storage of wild cherry, chestnut and oak. *Annales des Sciences Forestières* 52, 287–293.
- Janick, J., Kim, Y.H., Kitto, S. and Saranga, Y. (1993) Desiccated synthetic seed. In: Redenbaugh, K. (ed.) *Synseeds, Applications of Synthetic Seeds to Crop Improvement*. CRC Press, Boca Raton, FL, pp. 11–33.
- Jones, L.H. (1974) Long term survival of embryos of carrot (*Daucus carota* L.). *Plant Science Letters* 2, 221–224.
- Jouve, L., Engelmann, F. and Charrier, A. (1991) The effects of hypoxia and temperature on the *in vitro* storage of leafy shoots of *Coffea arabica* L. *Café, Cacao, Thé* 35, 205–210.
- Joy IV, R.W., Kumar, P.K. and Thorpe, T.A. (1991) Long-term storage of somatic embryogenic white spruce tissue at ambient temperature. *Plant Cell, Tissue and Organ Culture* 25, 53–60.
- Kartha, K.K. (1985) *Cryopreservation of Plant Cells and Organs*. CRC Press, Boca Raton, FL.
- Kartha, K.K. (1986) Production and indexing of disease-free plants. In: Withers, L. and Alderson, P.G. (eds) *Plant Tissue Culture and its Agricultural Applications*. Butterworths, London, pp. 219–238.
- Kartha, K.K. (1987) Advances in the cryopreservation technology of plant cells and organs. In: *Plant Tissue and Cell Culture*. Alan R. Liss, New York, pp. 447–458.
- Kartha, K.K. and Engelmann, F. (1994) Cryopreservation and germplasm storage.

- In: Vasil, I.K. and Thorpe, T.A. (eds.) *Plant Cell and Tissue Culture*. Kluwer, Dordrecht, pp. 195–230.
- Kartha, K.K., Leung, N.L. and Pahl, K. (1980) Cryopreservation of strawberry meristems and mass propagation of plantlets. *Journal of the American Society of Horticultural Science* 105, 481–484.
- Kartha, K.K., Mroginski, L.A., Pahl, K. and Leung, N.L. (1981) Germplasm preservation of coffee (*Coffea arabica* L.) by *in vitro* culture of shoot apical meristems. *Plant Science Letters* 22, 301–307.
- Kartha, K.K., Leung, N.L. and Mroginski, L.A. (1982) *In vitro* growth response and plant regeneration from cryopreserved meristems of cassava (*Manihot esculenta* Crantz). *Zeitschrift für Pflanzenphysiologie* 107, 133–140.
- Ko, W.H., Hwang, S.C. and Ku, F.M. (1991) A new technique for storage of meristem-tip cultures of “Cavendish” banana. *Plant Cell, Tissue and Organ Culture* 25, 179–183.
- Kobayashi, S., Sakai, A., Ohgarawa, T. and Nakamura, Y. (1994) Stable maintenance of an integrated gene in nucellar cells of navel orange (*Citrus sinensis* Osb.) under storage in LN₂. *Journal of Japanese Society for Horticultural Science* 63, 553–558.
- Kohmura, H., Ikeda, Y. and Sakai, A. (1994) Cryopreservation of apical meristem of Japanese shallot (*Allium wakegi* A.) by vitrification and subsequent high plant regeneration. *Cryo-Letters* 15, 289–298.
- Krishnapillay, B., (1989) Towards the development of a protocol for cryopreservation of embryos of a recalcitrant seed (*Artocarpus heterophyllus* Lam.). PhD Thesis, Universiti Pertanian Malaysia.
- Kuranuki, Y. and Sakai, A. (1994) Cryopreservation of *in vitro*-grown shoot tips of tea (*Camellia sinensis*) by vitrification. *Cryobiology* 31, 581.
- Kuriyama, A., Watanabe, K., Ueno, S. and Mitsuda, H. (1989) Inhibitory effect of ammonium ion on recovery of cryopreserved rice cells. *Plant Science* 64, 231–235.
- Kwiatowski, S., Martin, M.W., Brown, C.R. and Sluis, C.J. (1988) Serial microtuber formation as a long-term conservation method for *in vitro* germplasm. *American Potato Journal* 65, 369–375.
- Langis, R. and Steponkus, P. (1990) Cryopreservation of rye protoplasts by vitrification. *Plant Physiology* 92, 666–671.
- Langis, R., Schnabel-Preikstas, B., Earle, E.D. and Steponkus, P.L. (1989) Cryopreservation of *Brassica campestris* L. suspensions by vitrification. *Cryo-Letters* 10, 421–428.
- Langis, R., Schnabel-Preikstas, B.J., Earle, E.D. and Steponkus, P. (1990) Cryopreservation of carnation shoot tips by vitrification. *Cryobiology* 27, 657–658.
- Latta, R. (1971) Preservation of suspension cultures of plant cells by freezing. *Canadian Journal of Botany* 49, 1253–1254.
- Lecouteux, C., Florin, B., Tessereau, H., Bollon, H. and Pétiard, V. (1991) Cryopreservation of carrot somatic embryos using a simplified freezing process. *Cryo-Letters* 12, 319–328.
- Lecouteux, C., Tessereau, H., Florin, B., Courtois, D. and Pétiard, V. (1992) Preservation of somatic embryos of carrot (*Daucus carota* L.) by dehydration. *Comptes Rendus de l'Académie des Sciences, Paris Série III*, 314, 423–428.
- Lu, S. and Steponkus, P.L. (1994) Cryopreservation of *Solanum* shoot-tips by vitrification. *Cryobiology* 31, 569.

- Mannonen, L., Toivonen, L. and Kauppinen, V. (1990) Effects of long-term preservation on growth and productivity of *Panax ginseng* and *Catharanthus roseus* cell cultures. *Plant Cell Reports* 9, 173–177.
- Marin, M.L., Mafla, G., Roca, W.M. and Withers, L.A. (1990) Cryopreservation of cassava zygotic embryos and whole seeds in liquid nitrogen. *Cryo-Letters* 11, 257–264.
- Marin, M.L., Gogorcena, Y., Ortiz, J. and Duran-Vila, N. (1993) Recovery of whole plants of sweet orange from somatic embryos subjected to freezing thawing treatments. *Plant Cell, Tissue and Organ Culture* 34, 27–33.
- Marzalina, M. (1995) Pemyimpanan biji nehi mahogani (*Swietenia macrophylla* King). PhD Thesis, University Kebangsaan Malaysia.
- Marzalina, M., Krishnapillay, B. and Aziah, M.Y. (1992) Cryostorage of bamboo seeds and rattan embryos with a view to long term storage of germplasm of these species. In: *Proceedings of National Biotechnology IV*.
- Mathur, J., Ahuja, P.S., Lal, N. and Mathur, A.K. (1989) Propagation of *Valeriana wallichii* DC. using encapsulated apical and axial shoot buds. *Plant Science* 60, 111–116.
- Matsumoto, T., Sakai, A. and Yamada, K. (1994a) Cryopreservation of *in vitro*-grown apical meristems of lily (*Lilium* L.) by vitrification. *Cryobiology* 31, 579–580.
- Matsumoto, T., Sakai, A. and Yamada, K. (1994b) Cryopreservation of *in vitro*-grown apical meristems of wasabi (*Wasabia japonica*) by vitrification and subsequent high plant regeneration. *Plant Cell Reports* 13, 442–446.
- Matsumoto, T., Sakai, A., Takahashi, C. and Yamada, K. (1995) Cryopreservation of apical meristems of wasabi (*Wasabia japonica*) by encapsulation-vitrification method. *Cryo-Letters* 16, 189–196.
- Mazur, P. (1969) Freezing injury in plants. *Annual Review of Plant Physiology* 20, 419–448.
- Mazur, P. (1970) Cryobiology: the freezing of biological systems. *Science* 168, 939–949.
- Mazur, P. (1984) Freezing of living cells: mechanisms and applications. *American Journal of Physiology* 247, 125–142.
- Meryman, H.T. (1966) *Cryobiology*. Academic Press, London.
- Meryman, H.T. and Williams, R.J. (1985) Basic principles of freezing injury to plant cells: natural tolerance and approaches to cryopreservation. In: Kartha, K.K. (ed.) *Cryopreservation of Plant Cells and Organs*. CRC Press, Boca Raton, FL, pp. 13–48.
- Monette, P.L. (1986) Cold storage of kiwi shoot tips *in vitro*. *Hortscience* 21, 1203–1205.
- Monnier, M. and Leddet, C. (1978) Sur l'acquisition de la résistance au froid des embryons immatures de *Capsella bursa-pastoris*. *Comptes Rendus de l'Académie des Sciences, Paris Série D*, 287, 615–618.
- Moriguchi, T., Kozaki, I., Matsuta, N. and Yamaki, S. (1988) Plant regeneration from grape callus stored under a combination of low temperature and silicone treatment. *Plant Cell, Tissue and Organ Culture* 15, 67–71.
- Mullin, R.H. and Schlegel, D.E. (1976) Cold storage maintenance of strawberry meristem plantlets. *Horticultural Science* 11, 100–101.
- Mycock, D.J., Berjak, P., Blakeway, F. and Watt, P. (1989) Development of a model

- system for cryostorage: applicability to recalcitrant material. In: *Abstracts of the International Conference: The Impact of Biotechnology in Agriculture, Amiens, France, 10–12 July 1989*, p. CP1.
- Mycock, D.J., Wesley-Smith, J. and Berjak, P. (1995) Cryopreservation of somatic embryos of four species with and without cryoprotectant pretreatment. *Annals of Botany* 75, 331–336.
- Ng, S. and Ng, N.Q. (1991) Reduced-growth storage of germplasm. In: Dodds, J.H. (ed.) *In vitro Methods for Conservation of Plant Genetic Resources*. Chapman & Hall, London, pp. 11–39.
- Niino, T. and Sakai, A. (1992) Cryopreservation of alginate-coated *in vitro*-grown shoot tips of apple, pear and mulberry. *Plant Science* 87, 199–206.
- Niino, T., Sakai, A., Yakuwa, H. and Nojiri, K. (1992a) Cryopreservation of *in vitro*-grown shoot tips of apple and pear by vitrification. *Plant Cell, Tissue and Organ Culture* 28, 261–266.
- Niino, T., Sakai, A., Enomoto, S., Magosi, J. and Kato, K. (1992b) Cryopreservation of *in vitro*-grown shoot tips of mulberry by vitrification. *Cryo-Letters* 13, 303–312.
- Nishizawa, S., Sakai, A., Amano, Y. and Matsuzawa, T. (1992) Cryopreservation of asparagus (*Asparagus officinalis* L.) embryogenic suspension cells and subsequent plant regeneration by a simple freezing method. *Cryo-Letters* 13, 379–388.
- Nishizawa, S., Sakai, A., Amano, Y. and Matsuzawa, T. (1993) Cryopreservation of asparagus (*Asparagus officinalis* L.) embryogenic suspension cells and subsequent plant regeneration by vitrification. *Plant Science* 91, 67–73.
- Nitzsche, W. (1980) One year storage of dried carrot callus. *Zeitschrift für Pflanzenphysiologie* 100, 267–271.
- Niwata, E. (1995) Cryopreservation of apical meristems of garlic (*Allium sativum* L.) and high subsequent plant regeneration. *Cryo-Letters* 16, 102–107.
- Normah, M.N. and Hamidah S. (1992) Cryopreservation of *Garcinia mangostana* L. and *Citrus halimii* embryos. In: *Reproductive biology and plant breeding book of poster abstracts. XIIth Eucarpia Congress, July 6–11, 1992, Angers, France*, pp. 423–424.
- Normah, M.N. and Jamilah, S. (1992) Viability studies on *Lansium* seeds and embryos. In: *Symposium Abstracts XXIII International Seed Testing Congress, Buenos Aires, Argentina, 2–4 November 1992*.
- Normah, M.N. and Marzalina, M. (1995) Achievements and prospects for forest tree germplasm. In: *International Workshop on In vitro Conservation of Plant Genetic Resources*. Universiti Kebangsaan Malaysia and International Plant Genetic Resources Institute, Kuala Lumpur, Malaysia, p. 20.
- Normah, M.N. and Nordaini, O.L. (1994) Cryoexposure behaviour in several citrus species. In: Chong-Lek Koh (ed.) *Proceedings of the First National Genetic Congress, Kuala Lumpur*, pp. 209–211.
- Normah, M.N. and Vengadasalam, M. (1992) Effect of moisture content on cryopreservation of *Coffea* and *Vigna* seeds and embryos. *Cryo-Letters* 13, 199–208.
- Normah, M.N., Chin, H.F. and Hor, Y.L. (1986) Desiccation and cryopreservation of embryonic axes of *Hevea brasiliensis* Muel. Arg. *Pertanika* 9, 299–303.
- Normah, M.N., Reed, B.M. and Yu, X. (1995) Seed storage and cryoexposure behaviour in hazelnut (*Corylus avellana* L. cv. Barcelona). *Cryo-Letters* 15, 315–322.

- Pammenter, N.W., Vertucci, C.W. and Berjak, P. (1991) Homeohydrous (recalcitrant) seeds: dehydration, the state of water and viability characteristics in *Landolphia kirkii*. *Plant Physiology* 96, 1093–1098.
- Panis, B. (1995) Cryopreservation of banana (*Musa* spp.) germplasm. Dissertationes de Agricultura, Katholieke Universiteit Leuven, Belgium, 201 pp.
- Park, Y.S., Pond, S.E. and Bonga, J.M. (1994) Somatic embryogenesis in white spruce (*Picea glauca*): genetic control in somatic embryos to storage, maturation treatments, germination, and cryopreservation. *Theoretical and Applied Genetics* 89, 742–750.
- Paul, H. (1994) La régénération par embryogenèse somatique et la cryoconservation d'apex axillaires de vitroplants et d'embryons somatiques chez le Pommier (*Malus* × *domestica* Borkh.). Thèse d'Université, Université de Picardie Jules Verne.
- Paulet, F., Engelmann, F. and Glaszmann, J.C. (1993) Cryopreservation of apices of *in vitro* plantlets of sugarcane (*Saccharum* sp. hybrids) using encapsulation / dehydration. *Plant Cell Reports* 12, 525–529.
- Paulet, F., Eksomtramage, T., Lu, Y.H., D'Hont, A. and Glaszmann, J.C. (1994) Callus and apex cryopreservation for sugarcane germplasm management. In: *Proceedings of ISSCT Sugar Breeding Workshop, Montpellier, France*. CIRAD/ISSCT, Montpellier, France.
- Peacock, W.J. (1987) Molecular biology and genetic resources. In: Brown, A.D.H., Marshall, D.R., Frankel, O.H. and Williams, J.T. (eds) *The Use of Plant Genetic Resources Collections*. Cambridge University Press, Cambridge, pp. 363–376.
- Pence, V.C. (1990) Cryostorage of embryo axes of several large-seeded temperate tree species. *Cryobiology* 27, 212–218.
- Pence, V.C. (1991) Cryopreservation of immature embryos of *Theobroma cacao*. *Plant Cell Reports* 10, 144–147.
- Pence, V.C. (1992) Desiccation and the survival of *Aesculus*, *Castanea*, and *Quercus* embryo axes through cryopreservation. *Cryobiology* 29, 391–399.
- Pepó, P. (1994) Krioprotektív anyagok hatása a burgonya (*Solanum tuberosum* L.) merisztémák túlérére. *Növénytermelés* 43, 469–476.
- Phunchindawan, M., Hirata, K., Kobayashi, S., Miyamoto, K. and Sakai, A. (1994) Cryopreservation of shoot primordia induced from horseradish hairy root cultures by an encapsulation-dehydration method. *Cryobiology* 31, 580–581.
- Plessis, P., Leddet, C. and Dereuddre, J. (1991) Resistance to dehydration and to freezing in liquid nitrogen of alginate coated shoot-tips of grape vine (*Vitis vinifera* L. cv. Chardonnay). *Comptes Rendus de l'Académie des Sciences, Paris Série III*, 313, 373–380.
- Plessis, P., Leddet, C., Collas, A. and Dereuddre, J. (1993) Cryopreservation of *Vitis vinifera* L. cv. Chardonnay shoot tips by encapsulation-dehydration: effect of pretreatment, cooling and postculture conditions. *Cryo-Letters* 14, 309–320.
- Poissonnier, M., Monod, V., Paques M. and Dereuddre, J. (1992) Cryopreservation in liquid nitrogen of *Eucalyptus gunnii* shoot tips grown *in vitro* following encapsulation and dehydration. *Annales de Recherches Sylvicoles, AFOCEL* 5–23.
- Pritchard, H.W. and Prendergast, F.G. (1986) Effects of desiccation on the *in vitro* viability of embryos of the recalcitrant seed species *Araucaria hunstenii* K. Schum. *Journal of Plant Physiology* 37, 1388–1397.

- Quatrano, R.S. (1968) Freeze-preservation of cultured flax cells utilizing DMSO. *Plant Physiology* 43, 2057–2061.
- Radhamani, J. and Chandel, K.P.S. (1992) Cryopreservation of axes of trifoliolate orange (*Poncirus trifoliata* (L.) Raf.). *Plant Cell Reports* 11, 204–206.
- Rao, P.S., Ganapathi, T.R., Suprasanna, P. and Bapat, V.A. (1993) Encapsulated shoot tips of banana: a new propagation and delivery system. *InfoMusa* 2, 4–5.
- Redenbaugh, K., Fujii, J.O. and Slade, D. (1991) Synthetic seed technology. In: Vasil I.K. (ed.) *Cell Culture and Somatic Cell Genetics of Plants*. Academic Press, San Diego, CA, pp. 35–74.
- Reed, B.M. (1991) Application of gas-permeable bags for *in vitro* cold storage of strawberry germplasm. *Plant Cell Reports* 10, 431–434.
- Reed, B.M. (1992) Cold storage of strawberries *in vitro*: a comparison of three storage systems. *Fruit Varieties Journal* 46, 98–102.
- Reed, B.M. and Yu, X. (1995) Cryopreservation of *in vitro* grown gooseberry and currant meristems. *Cryo-Letters* 16, 131–136.
- Reed, B.M., Normah, M.N. and Xiaoling, Yu. (1994) Stratification is necessary for successful cryopreservation of axes from stored hazelnut seed. *Cryo-Letters* 15, 377–384.
- Reinhoud, P., Kars, I., Iren Van, F. and Kijne, J. (1994) Vitrification of tobacco cells: Cryoprotection by proline? *Cryobiology* 31, 579.
- Rillo, E.P. and Paloma, M.B.F. (1991) Storage and transport of zygotic embryos of *Cocos nucifera* L. for *in vitro* culture. *FAO/IBPGR Plant Genetic Resources Newsletter* 86, 1–4.
- Roberts, H.F. (1973) Predicting the viability of seeds. *Seed Science and Technology* 1, 499–514.
- Roca, W.M., Reyes, R. and Beltran, J. (1984) Effect of various factors on minimal growth in tissue culture storage of cassava germplasm. In: *Proceedings of Sixth Symposium of the International Society for Tropical Root Crops, 21–26 February, Lima, Peru*, pp. 441–446.
- Roxas, N.J.L., Tashiro, Y., Miyasaki, S., Isshiki, S. and Takeshita, A. (1995) *In vitro* preservation of Higo chrysanthemum (*Dendranthema* × *grandiflorum* (Ramat.) Kitam.). *Journal of the Japanese Horticultural Society* 63, 863–870.
- Runthala, R., Jana, M.K. and Mohanan, K. (1993) Cryopreservation of groundnut (*Arachis hypogaea* L.) embryonic axes for germplasm conservation. *Cryo-Letters* 14, 337–340.
- Ruredzo, T.J. (1989) *Progress Report on IBPGR-ILCA Tissue Culture Project*. IBPGR Report Number 89/11. IBPGR, Rome.
- Sakai, A. (1960) Survival of the twigs of woody plants at -196°C . *Nature* 185, 392–394.
- Sakai, A. (1995) Cryopreservation for germplasm collection in woody plants. In: Jain, S., Gupta, P. and Newton, R. (eds) *Somatic Embryogenesis in Woody Plants*, vol. 1. Kluwer, Dordrecht, pp. 293–315.
- Sakai, K. and Matsumoto, T. (1995) A novel cryogenic procedure for cryopreservation of *in vitro* grown meristems of temperate crops – encapsulation / vitrification. In: *International Workshop on In vitro Conservation of Plant Genetic Resources*. Universiti Kebangsaan Malaysia and International Plant Genetic Resources Institute, Kuala Lumpur, Malaysia, p. 33.
- Sakai, A., Kobayashi, S. and Oiyama, I. (1990) Cryopreservation of nucellar cells of

- navel orange (*Citrus sinensis* Osb. var. *brasiliensis* Tanaka) by vitrification. *Plant Cell Reports* 9, 30–33.
- Sakai, A., Kobayashi, S. and Oyama, I. (1991) Cryopreservation of nucellar cells of naval orange (*Citrus sinensis* Osb.) by a simple freezing method. *Plant Science* 74, 243–248.
- Schäfer-Menuhr, A. (1994) *Refinement of Cryopreservation Techniques for Potato, Project Progress Report*. Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, (DSMZ), Institut für Pflanzenbau Bundesforschungsanstalt für Landwirtschaft (FAL), Germany.
- Schäfer-Menuhr, A. (1995) *Refinement of Cryopreservation Techniques for Potato, Project Progress Report*. Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, (DSMZ), Institut für Pflanzenbau Bundesforschungsanstalt für Landwirtschaft (FAL), Germany.
- Schneibel-Preikstas, B., Earle, E.D. and Steponkus, P.L. (1992a) Cryopreservation of chrysanthemum shoot tips by vitrification. *Cryobiology* 29, 739.
- Schneibel-Preikstas, B., Earle, E.D. and Steponkus, P.L. (1992b) Cryopreservation of sweet potato shoot tips by vitrification. *Cryobiology* 29, 738–739.
- Schneibel-Preikstas, B., Earle, E.D. and Steponkus, P.L. (1992c) Cryopreservation of potato shoot tips by vitrification. *Cryobiology* 29, 747.
- Scottez, C. (1993) Cryoconservation après encapsulation-déshydratation d'apex de poirier (*Pyrus communis* L. cv. Beurre Hardy) cultivé *in vitro*: effets d'un endurcissement par les basses températures et étude du métabolisme des lipides. Thèse d'Université, Université Paris 6.
- Scottez, C., Chevreau, E., Godard, N., Arnaud, Y., Duron, M. and Dereuddre, J. (1992) Cryopreservation of cold-acclimated shoot tips of pear *in vitro* cultures after encapsulation-dehydration. *Cryobiology* 29, 691–700.
- Scowcroft, W.R. (1984) *Genetic Variability in Tissue Culture: Impact on Germplasm Conservation and Utilization*. IBPGR, Rome.
- Seitz, U. (1987) Cryopreservation of plant cell cultures. *Planta Medica* 4, 311–314.
- Senaratna, T., McKersie, B.D. and Bowley, S.R. (1990) Artificial seeds of alfalfa (*Medicago sativa* L.). Induction of desiccation tolerance in somatic embryos. *In vitro Cellular and Development Biology* 26, 85–90.
- Sharma, N. and Chandel, K. (1992) Low-temperature storage of *Rauvolfia serpentina* Benth. ex Kurz.: an endangered, endemic medicinal plant. *Plant Cell Reports* 11, 200–203.
- Shigeta, J., Mori, T. and Sato, K. (1993) Storage of encapsulated somatic embryos of carrot. *Biotechnology Techniques* 7, 165–168.
- Shimonishi, K., Ishikawa, M., Suzuki, S. and Oosawa, K. (1991) Cryopreservation of melon somatic embryos by desiccation method. *Japanese Journal of Breeding* 41, 347–351.
- Sossou, J., Karunaratne, S. and Kovoor, A. (1987) Collecting palm: *In vitro* explanting in the field. *FAO/IBPGR Plant Genetic Resources Newsletter* 69, 7–18.
- Staritsky, G., Dekkers, A.J., Louwaars, N.P. and Zandvoort, E.A. (1986) *In vitro* conservation of aroid germplasm at reduced temperatures and under osmotic stress. In: Withers, L.A. and Alderson, P.G. (eds) *Plant Tissue Culture and its Agricultural Applications*. Butterworths, London, pp. 277–284.
- Steponkus, P.L., Langis, R. and Fujikawa, S. (1992) Cryopreservation of plant tissues by vitrification. In: Steponkus, P.L. (ed.) *Advances in Low Temperature Biology*. JAI Press, Hampton Hill, UK.

- Takano, M. and Tamura, S. (1992) Vitrification of cultured cells of *Nicotiana tabacum*. *Cryobiology* 29, 739.
- Tannoury, M. (1993) Cryoconservation d'apex d'Oeillet (*Dianthus caryophyllus* L.) et d'embryons somatiques de Carotte (*Daucus carota* L.) par les procédés d'enrobage-déshydratation et d'enrobage-vitrification. PhD Thesis, Université Paris VI.
- Tannoury, M., Ralambosoa, J., Kaminski, M. and Dereuddre, J. (1991) Cryopreservation by vitrification of coated shoot-tips of carnation (*Dianthus caryophyllus* L.) cultured *in vitro*. *Comptes Rendus de l'Académie des Sciences, Paris Série III*, 313, 633–638.
- Tannoury, M., Vintejou, C. and Dereuddre, J. (1994) Cryoconservation of encapsulated carnation shoot tips: comparative study of different experimental protocols. *Cryobiology* 31, 581.
- Tessereau, H., Florin, B., Meschine, M.C., Thierry, C. and Pétiard, V. (1994) Cryopreservation of somatic embryos: a tool for germplasm storage and commercial delivery of selected plants. *Annals of Botany* 74, 547–555.
- Towill, L.E. (1990) Cryopreservation of isolated mint shoot tips by vitrification. *Plant Cell Reports* 9, 178–180.
- Towill, L.E. and Jarret, R.L. (1992) Cryopreservation of sweet potato (*Ipomoea batatas* L. Lam.) shoot tips by vitrification. *Plant Cell Reports* 11, 175–178.
- Uragami, A., Sakai, A., Nagai, M. and Takahashi, T. (1989) Survival of cultured cells and somatic embryos of *Asparagus officinalis* cryopreserved by vitrification. *Plant Cell Reports* 8, 418–421.
- Uragami, A., Sakai, A. and Magai, M. (1990) Cryopreservation of dried axillary buds from plantlets of *Asparagus officinalis* L. grown *in vitro*. *Plant Cell Reports* 9, 328–331.
- Uragami, A., Lucas, M.O., Ralambosa, M.R., Renard, M. and Dereuddre, J. (1993) Cryopreservation of microspore embryos of oilseed rape (*Brassica napus* L.) by dehydration in air with or without alginate encapsulation. *Cryo-Letters* 14, 83–90.
- Vandenbussche, B. and De Proft, M.P. (1995) Cryopreservation of encapsulated sugar beet shoot-tips (*Beta vulgaris* L.) using rapid freezing. *Cryo-Letters* 16, 67.
- Vandenbussche, B., Demeulemeester, M.A.C. and De Proft, M.P. (1993) Cryopreservation of alginate-coated *in vitro* shoot-tips of chicory (*Chicorium intybus* L.) using rapid freezing. *Cryo-Letters* 14, 259–266.
- Vertucci, C.W., Berjak, P., Pammenter N.W. and Crane J. (1991) Cryopreservation of embryonic axes of an homeohydrous (recalcitrant) seed in relation to calorimetric properties of tissue water. *Cryo-Letters* 12, 339–350.
- Viterbo, A.A. and Rabinowitch, H.D. (1994) *In vitro* propagation and germplasm cold-storage of fertile and male-sterile *Allium trifoliatum* subsp. *hirsutum*. *Genetic Research and Crop Evolution* 41, 87–98.
- Vysotskaya, O.N. (1994) Long-term maintenance of strawberry plants *in vitro*. *Russian Journal of Plant Physiology* 41, 827–833.
- Ward, A.C.W., Benson, E.E., Blackhall, N.W., Cooper-Bland, S., Powell, W., Power, J.B. and Davey, M.R. (1993) Flow cytometric assessments of ploidy stability in cryopreserved dihaploid *Solanum tuberosum* and wild *Solanum* species. *Cryo-Letters* 14, 145–152.

- Watanabe, K. and Steponkus, P.L. (1994) Vitrification of cell suspensions of *Oryza sativa*. *Cryobiology* 31, 568.
- Watanabe, K., Yamada, Y., Ueno, S. and Mitsuda, H. (1985) Change of freezing resistance and retention of metabolic and differentiation potentials in cultured green *Lavandula vera* cells which survived repeated freeze-thaw procedures. *Agricultural and Biological Chemistry* (Tokyo) 49, 1727–1731.
- Watanabe, K., Kawai, F., Kanamori, M. and Mitsuda, H. (1991) Preservation of rice callus by gas-replacement. *Plant Tissue Culture Letters* 8, 36–38.
- Wesley-Smith, J., Vertucci, C.W., Berjak, P., Pammenter, N.W. and Crane, J. (1992) Cryopreservation of desiccation-sensitive axes of *Camellia sinensis* in relation to dehydration, freezing rate and the thermal properties of tissue water. *Journal of Plant Physiology* 140, 596–604.
- Westcott, R.J. (1981a) Tissue culture storage of potato germplasm. 1. Minimal growth storage. *Potato Research* 24, 331–342.
- Westcott, R.J. (1981b) Tissue culture storage of potato germplasm. 2. Use of growth retardants. *Potato Research* 24, 343–352.
- Withers, L.A. (1985) Cryopreservation of cultured plant cells and protoplasts. In: Kartha, K.K. (ed.) *Cryopreservation of Plant Cells and Organs*. CRC Press, Boca Raton, pp. 243–267.
- Withers, L.A. (1987) *In vitro* methods for germplasm collecting in the field. *FAO/IBPGR Plant Genetic Research Newsletter* 69, 2–6.
- Withers, L.A. (1992) *In vitro* conservation. In: Hammerschlag, F. and Litz, R.E. (eds) *Biotechnology of Perennial Fruit Crops*. CAB International, Wallingford, Oxon, pp. 169–200.
- Withers, L.A. (1995) Collecting *in vitro* for genetic resources conservation. In: Guarino, L., Rao, R. and Reid, R. (eds) *Collecting Plant Genetic Diversity*. CAB International, Wallingford, Oxon, pp. 511–515.
- Withers, L.A. and Engelmann, F. (1997) *In vitro* conservation of plant genetic resources. In: Altman, A. (ed) *Biotechnology in Agriculture*. Marcel Dekker, New York (in press).
- Withers, L.A. and Engels, J.M.M. (1990) The test tube genebank – a safe alternative to field conservation. *IBPGR Newsletter for Asia and the Pacific* 3, 1–2.
- Withers, L.A. and King, P.J. (1980) A simple freezing-unit and routine cryopreservation method for plant cell cultures. *Cryo-Letters* 1, 213–220.
- Withers, L.A. and Street, H.E. (1977) Freeze-preservation of cultured plant cells: III. The pregrowth phase. *Physiologia Plantarum* 39, 171–178.
- Yamada, T., Sakai, A., Matsumura, T. and Higuchi, S. (1991) Cryopreservation of apical meristems of white clover (*Trifolium repens* L.) by vitrification. *Plant Science* 78, 81–87.
- Yidana, J.A., Withers, L.A. and Ivins, J.D. (1987) Development of a simple method for collecting and propagating cocoa germplasm *in vitro*. *Acta Horticulturae* 212, 95–98.
- Yoshimatsu, K., Yamaguchi, H. and Shimomura, K. (1994) Cryopreservation of *Panax ginseng* hairy roots by vitrification method. *Cryobiology* 31, 580.
- Zakri, A.H., Normah, M.N., Karim, A.A.G. and Senawi, M.T. (1991) Conservation of plant genetic resources through *in vitro* methods. In: *Proceedings of MNCPGR/CSC International Workshop on Tissue Culture for the Conservation of*

-
- Biodiversity and Plant Genetic Resources, Kuala Lumpur, Malaysia*. Forest Research Institute Malaysia, Kuala Lumpur, p. 270.
- Zavala, M.E. and Sussex, I.M. (1986) Survival of developing wheat embryos and bean axes following cryoprotection and freezing in liquid nitrogen. *Journal of Plant Physiology* 122, 193–197.

Conservation of DNA: DNA Banking

R.P. Adams

*Plant Biotechnology Center, Baylor University, Box 669
Gruver, TX 79040, USA*

7.1. Introduction

The loss of native plant diversity also means a loss of genetic diversity present in and available to our current and potential crop species. Cultivated crops are extremely inbred for factors such as yield, uniform flowering and height, and cosmetic features of products. This narrow genetic base has resulted in several disastrous crop failures. Ireland's potato famine of 1846, which resulted in the emigration of a quarter of the country's human population, occurred because the potato crop had no resistance to the late blight fungus, *Phytophthora infestans* (Mont.) DeBary (Plucknett *et al.*, 1987). This can be traced to the lack of genetic diversity in Irish potatoes, which had been multiplied using clonal materials from just two separate South American introductions, to Spain in 1570 and to England in 1590 (Hawkes, 1983).

A more recent example is the epidemic of southern corn leaf blight caused by the fungus *Helminthosporium maydis* Nisik. & Miy., in 1970 in the USA. Because almost all of the corn (maize: *Zea mays* L.) in the USA was of hybrid origin and contained the Texas cytoplasmic male sterile line, the fields of corn presented an unlimited extremely narrow gene base habitat for the fungus. By the late summer, 1970, plant breeders were scouring corn germplasm collections in Argentina, Hungary, Yugoslavia and the USA for resistant sources (Plucknett *et al.*, 1987). Nurseries and seed fields were used in Hawaii, Florida, the Caribbean, and Central and South America to incorporate the resistance into hybrid corn in time for planting in the spring of 1971 (Ullstrup, 1972). Without these genetic resources this technological feat would not have been possible.

Since the first plant-to-plant gene transfer in 1983 (Murai *et al.*, 1983), genes have been transferred to plants from viruses (Nelson *et al.*, 1988), bacteria (Barton *et al.*, 1987; Della-Cioppa *et al.*, 1987; Fischhoff *et al.*, 1987), and even from mammals to plants (Lefebvre *et al.*, 1987; Maiti *et al.*, 1988). Genetic transfers are being performed in order to attain insect, bacterial, viral and fungal resistance, a more nutritionally balanced protein, more efficient photosynthesis, nitrogen fixation, and salt and heavy metal tolerance, to name a few. These kinds of gene transfers from one unrelated organism to another indicate that we must now view the world's genetic resources (i.e. genes, or DNA) from a horizontal perspective in which gene transfers will cut across species, genus and family boundaries.

For example, a strain of cowpea (*Vigna unguiculata* (L.) Walp.) discovered in a market in Ilorin, Nigeria, contains a protein that inhibits trypsin digestion by insects (Redden *et al.*, 1984). This gene has been moved to tobacco (*Nicotiana*) where the trypsin-inhibiting gene is expressed and offers tobacco the same resistance against insects as in cowpea (Newmark, 1987). It is interesting to note that although a very active form of the gene has been found in a Nigerian cowpea, scarcely 100 of the world's 13 000 legume species have been examined for this gene. Yet the tropical legumes, one of the most promising groups for the evolution of natural insecticides, will certainly be subject to considerable germplasm loss in the next decade.

The novel insecticides, biocides, medicines, etc. that may exist in nature are innumerable. Yet the principal areas of diversity among plants, the lowland tropical forests, will have been felled or severely damaged within the next 20 years (Raven, 1988). The Amazon River system, for example, contains eight times as many species as the Mississippi River system (Shulman, 1986). Raven (1988) estimated that as many as 1.2 million species would become extinct in the next 20 years. The loss of plant species will mean a loss of potential plant-derived pharmaceuticals, estimated at \$US2 billion/year in the USA alone (US Congress, 1987).

The National Cancer Institute (NCI) has spent \$US8 million in the last five years for a massive plant-collecting effort in the tropics to find anti-cancer and anti-AIDS virus compounds (Booth, 1987). The plant collectors will gather leaves and/or bark and air-dry the material for shipment to NCI where it will be extracted and assayed against 100 cancer cell lines and the AIDS virus. Yet, *no* genetic resources are being collected! When a promising compound is found, the plants will have to be recollected. For extensive testing (as well as commercial utilization), plantations will have to be established in the tropics to provide material.

7.2. Initiation of DNA Bank-Net

Concurrent with the advancements in gene cloning and transfer has been the development of technology for the removal and analysis of DNA. DNAs from the nucleus, mitochondrion and chloroplast are now routinely extracted and immobilized onto nitrocellulose sheets where the DNA can be probed with numerous cloned genes. In addition, the rapid development of polymerase chain reaction (PCR) now means that one can routinely amplify specific oligonucleotides or genes from the entire mixture of genomic DNA. These advances, coupled with the prospect of the loss of significant plant genetic resources throughout the world, have led to the establishment of DNA Bank-Net, an international network of DNA repositories for the storage of genomic DNA.

The organizational meeting involved a group of 18 scientists who met at the Royal Botanic Gardens, Kew, London, in April 1991, to share national and institutional experiences using *in vitro* biotechnology and particularly cryostorage of DNA and DNA-rich materials (Adams and Adams, 1991). A second meeting was held at the Missouri Botanical Garden in 1993 (Adams *et al.*, 1994). Relatively few scientists were interested in a 'genetic insurance policy' when the idea of banking genomic DNA from plants was first proposed (Adams, 1988, 1990). However, currently there are 40 institutions, representing 25 nations and every inhabited continent, that have expressed interest in DNA Bank-Net (Fig. 7.1).

The conserved DNA will have numerous uses: molecular phylogenetics and systematics of extant and extinct taxa; production of previously characterized secondary compounds in transgenic cell cultures; production of transgenic plants using genes from gene families; *in vitro* expression

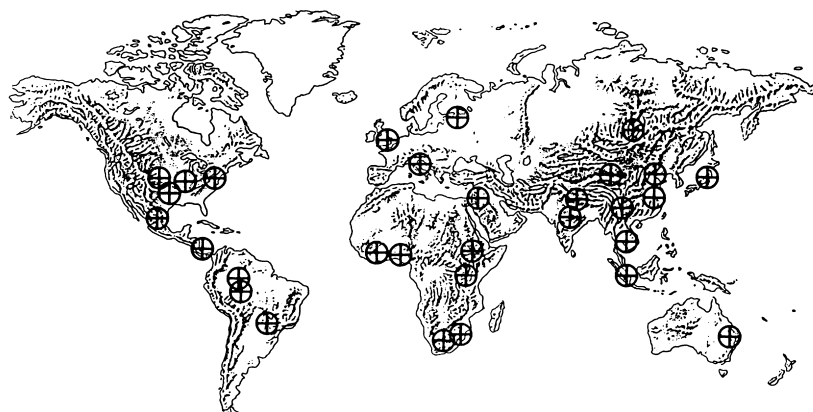


Fig. 7.1. Map of individuals/institutions currently interested in DNA Bank-Net.

and study of enzyme structure and function; and genomic probes for research laboratories.

7.3. Structure and Operation of DNA Bank-Net

At the organizational meeting of DNA Bank-Net, a task force was convened to define the functions of working (DNA-dispensing) and reserve (base) nodes in the DNA Bank network. The group recommended the following functions (Adams and Adams, 1991):

Working (DNA-dispensing) nodes:

1. Collection of plant material by taxonomists. This may be the primary function of a particular node or be in association with other organizations such as universities, botanic gardens, etc.
2. DNA extraction by molecular biologists or trained staff.
3. Long-term preservation of DNA-rich materials and /or extracted DNA in liquid nitrogen.
4. DNA analysis / gene replication by molecular biologists or trained staff.
5. Distribution of DNA (genes, gene segments, oligonucleotides, etc.).

Reserve (base) nodes:

1. Long-term DNA preservation in liquid nitrogen and monitoring of potential DNA degradation.
2. Act as genetic reserve buffer for working nodes.
3. Replenishment of DNA if a working node experiences the catastrophic loss of storage parameters and DNA.

Figure 7.2 depicts the relationship between working and reserve nodes. Note the projected flow of plant materials and DNA through the working (DNA-dispensing) node. It is likely that some of the working nodes would be actively acquiring and /or dispensing DNA from some geographic area (e.g. Africa), yet maintain separate cryovats, functioning as a reserve (base) node for another area (e.g. South America).

7.4. General Requirements for Nodes in the DNA Bank-Net

The task group recommended (Adams and Adams, 1991) that the following were the minimum requirements for nodes:

Working (DNA dispensing) nodes:

- Personnel: taxonomists / collectors, biochemists / molecular biologists, technicians for practical work, capable administration.
- Equipment: storage facilities (liquid nitrogen, cryovats), extraction

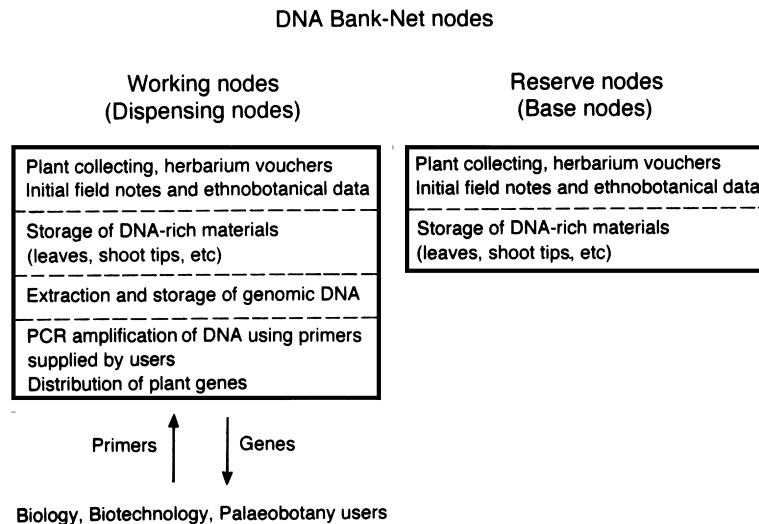


Fig. 7.2. Schematic representation of the flow of materials and the relationship between working (DNA-dispensing), reserve (base) nodes and users.

facilities (centrifuges, gel electrophoresis, UV spectrophotometer, etc.), DNA analyses and PCR duplication (PCR thermal cycler, microcentrifuges, etc.), distribution system (packaging and mailing supplies), computer (database for inventory and correspondence).

Reserve (base) nodes:

- Personnel: technicians, capable administration.
- Equipment: storage facilities (liquid nitrogen, cryovats), computer (data base for inventory and correspondence).

Each DNA collection should be split initially into at least two or three portions. One sample (DNA-rich material or extracted DNA) should be stored at a working (DNA-dispensing) node and another portion(s) should be stored in one, but preferably two, back-up reserve (base) nodes. The reserve nodes should be in different countries and if possible on different continents to safeguard the DNA samples against various natural and man-made catastrophes. An example is shown in Fig. 7.3, where three replicate samples are collected and taken to Medellin, with replicates then sent to the Vavilov Institute and the Missouri Botanical Garden. Plant materials (in silica gel) could be stored in a freezer until the identification and other documentation have been accomplished and then shipped in quantity with other samples in off-season periods. No doubt other strategies will be developed with experience.



Fig. 7.3. Hypothetical example of a triplicate collection in Colombia. The DNA-rich materials are placed in silica gel or Drierite for interim preservation and taken to the working node (A, Medellin). From Medellin (A), two replicates would be sent to reserve nodes at (B) Vavilov Institute and (C) the Missouri Botanical Garden.

Several general recommendations came from the task groups (Adams and Adams, 1991) and these include:

1. DNA should be extracted from cryopreserved DNA-rich materials only when the DNA is needed. Delaying the extraction has the advantage of letting technology catch up, so advanced techniques can be used as they become available.
2. Working nodes should generally be an existing organization with adequate biochemical expertise and an associated herbarium. Although on-site herbarium is not required, a very close, local (same city) association with a recognized herbarium (Holmgren *et al.*, 1990) is required.
3. For the working as well as reserve nodes, it is necessary to have a strong institutional commitment, not just a personal commitment, in order that the collection be maintained in perpetuity and not just for the lifetime of one person who has committed himself/herself to the idea.
4. Consideration should be given to the availability of dependable supplies of electricity and liquid nitrogen in determining the feasibility of establishing a node.
5. Considerable interest was shown in the concept of storing composite DNA samples (e.g. a composite of DNA from all the legumes in a region, to be used for screening or retrieval of unusual genes).
6. The need for computer and database compatibility was expressed. Given the number of flat file and relational databases that are compatible with dBASE, it would seem that dBASE compatibility would be desirable. No consensus was reached in regard to this or on the use of a flat file

versus relational database. It was felt that the critical issue at present was to begin collecting DNA-rich materials.

7.5. Scope of Plant Collections

The task group given this assignment felt that there is a need for an initial focus rather than random collections and that economically useful plants should be given some priority (Adams and Adams, 1991). However, this priority would not include the major crop plants of commercial usage that are widely cultivated (e.g. maize, rice, wheat, etc.), but rather those indigenous species that are tended and /or otherwise used by local people.

One problem with giving a priority to species is that field collecting then becomes 'plant hunting' trips, which tend to be very expensive. It would seem that the cheapest and most practical way to preserve the largest percentage of plant genes would be to utilize the current (and additional) floristic collectors (such as those of the Missouri Botanical Garden, Royal Botanic Gardens, etc.), who are already in the field and are familiar with vegetation of the region. The collections of DNA-rich material (leaves) could be done with little additional effort when specimens are collected.

7.6. Collecting Procedures

DNA collectors should be considered the same as all other plant collectors. Consequently they should (Adams and Adams, 1991):

1. Voucher all collections in recognized herbaria (i.e. ones listed in *Index Herbariorum*, 8th edn; see Holmgren *et al.*, 1990).
2. Provide proper label information as to the locality, habitat, etc. for each plant collected.
3. Follow all procedures concerning permits, convenios, and deposition of duplicate vouchers in the country of origin.
4. Collect leaf samples and pack them in desiccants (see Adams *et al.*, 1991) immediately (the same day). Leaves themselves provide simple long-term storage.
5. In the case of legumes, samples of root nodules should be taken if possible, but kept as a separate accession.
6. If a chemical treatment is used in the field, information should be provided concerning the method, and some untreated leaves must be stored in desiccant (see no. 4 above).
7. When possible, fossil material should be included in DNA Bank-Net. In this case, when destruction of the source material occurs, documentation via photographs and fragments is necessary.

8. Some material may be accessioned from herbarium specimens under control of local curators using current methods of DNA extraction. Herbarium sheets should be marked if sampled for DNA. Herbarium specimens are limited in supply and their utility appears to be limited to material collected without chemical preservation. Material may be sampled directly from the sheet or the attached specimen envelope if it contains sufficient leaf material ($\approx 0.1\text{--}0.5$ g dry wt.) for DNA extraction.

7.7. Interim Field Storage of Specimens

The problems associated with transporting fresh or frozen materials can generally be overcome by specialists (e.g. the worldwide collections of fresh foliage of *Juniperus* for analysis of essential oils and DNA undertaken by the author). However, botanists doing floristic research will likely collect many of the specimens from rare and endangered tropical species. They often collect specimens from scores of different species in a single day. The sheer bulk of the materials that they have to process and ship requires that any protocol for the collection of samples for specialized needs (e.g. DNA storage/analyses) must be quick, simple and trouble-free. The generalist collector, working in tropical areas, cannot be expected to preserve hundreds or thousands of samples for months under tropical conditions and then arrange transport through customs, all the while keeping the individual specimens frozen.

Fortunately, at least as far as DNA preservation is concerned, interim preservation in silica gel or Drierite is an effective way to keep plant materials in the field and/or in transit for several months at ambient temperatures (Adams *et al.*, 1991).

7.8. Protocol for Field Preservation of Foliage

Drierite has a water capacity of 10 to 14%, but above 6.6% the capacity varies inversely with temperature (W.A. Hammond, Drierite Co., personal communication). One would not want to risk possible rehydration of leaves, so storage ratios should be based on the 6.6% capacity. In laboratory tests, silica gel absorbed 8.85% of its weight of water after exposure to 100% humidity for 16 h at 22°C (Adams *et al.*, 1991). We have found that plant materials contain as much as 92% moisture (Adams *et al.*, 1991), so a useful approximation would be to assume the plant is mostly water and use 16 to 20 times the fresh leaf weight for the Drierite or silica gel component.

Now that inexpensive (\$US100) battery-powered, portable balances are available, one could take a supply of jars that hold, e.g., 100 g of silica

gel and then weigh out 5 g of fresh leaf material and add it to the jar along with silica gel (or Drierite). We have found that air-dried leaves (suitable for herbarium vouchers) generally contain 10 to 15% water. Using a robust value of 20% water for air-dried leaves, one can weigh out 5 g of these leaves ($5 \text{ g} \times 20\% = 1 \text{ g}$ water) or 1 g fresh leaves per 20 g of silica gel. This procedure may seem time consuming, but in practice we merely do a quick check on the leaf area needed to give approximately 1 g (fresh leaves) or 5 g (dried leaves) and then just use that amount of leaf area. For example, for spinach, a $2 \text{ cm} \times 4 \text{ cm}$ fresh leaf area weighs about 1 g. So, one can just cut the leaves into roughly $2 \text{ cm} \times 4 \text{ cm}$ squares and add one square to 20 g of silica gel. For succulent leaves, a slightly different protocol may be used. Liston *et al.* (1990) removed succulent leaf material after 24 h in Drierite and placed it in fresh Drierite.

A note of caution is necessary concerning field drying of specimens for subsequent silica gel/Drierite storage. We have experienced difficulty obtaining DNA from leaves dried at temperatures higher than about 55°C . In very rainy conditions where high drying temperatures (from butane stoves, for example) are used to dry specimens, it would seem advisable to merely blot leaves free of surface moisture and then place the fresh leaf material directly into silica gel or Drierite. Liston *et al.* (1990) took 2–5 g of plant tissue and wrapped it in tissue paper to prevent it from fragmenting, then placed it in a 125 ml Nalgene bottle, one-third prefilled with Drierite (with blue indicator crystals), and then filled the bottle (two-thirds) with additional Drierite.

Clear plastic bottles are probably preferable to glass, to avoid breakage in transit. Using clear jars allows one to check the indicating crystals without opening the jar. The lids should be sealed with vinyl tape to prevent moisture leakage. The use of parafilm to seal containers is not recommended, as we have found it to come loose at 37°C (and of course, at tropical temperatures!).

Silica gel and Drierite do differ in one characteristic that may be a consideration. We have found that silica gel can be dried (recharged) at 100°C for 24 h, but Drierite must be dried at much a higher temperature (200°C). In addition, we could easily dry (recharge) silica gel, but were unable to dry (recharge) Drierite in a microwave oven. If the desiccant gets wet before use, silica gel appears to be much easier to dry. Silica gel is used in large quantities for flower drying and, thus, may be cheaper, depending on the source. Both Drierite and silica gel could be recharged for reuse on subsequent trips, but one should be very careful to remove any leaf fragments. If materials are to be checked through customs, it is useful to have a small container of silica gel/Drierite that you can open and show customs agents. A demonstration that the blue indicator crystals will turn pink when you breathe on or moisten them is helpful in convincing the customs officials *not* to open your sealed specimen jars.

7.9. Future Considerations

The vast resources of dried specimens in the world's herbaria may hold considerable DNA that would be suitable for polymerase chain reaction (PCR). It seems likely that the integrity of DNA would decrease with the age of specimens. Because there are many types of herbarium storage environments, preservation and collections, there is a need for systematic investigations of the effect of modes of preparation, collection and storage on the integrity of DNA in the world's major holdings.

One of the major concerns in storing DNA from extinct species is the limited amount of DNA available for distribution. A general process is needed by which the DNA could be immobilized, and then specific genes or oligonucleotides amplified. Genomic DNA immobilized onto nylon might be used, as described by Kadokami and Lewis (1990) for cDNA from spiders. Amplification would then involve removing the membrane with the bound DNA from cryostorage and amplifying the desired gene, washing away the primers and placing the bound DNA back into cryostorage. Although Kadokami and Lewis (1990) reported successful PCR amplification of membrane-bound cDNA, we have not been able to extend their work to genomic plant DNA. Additional research is needed in this area.

Research is also needed to amplify the entire genomic DNA of a species. Some modification of the GAWTS type (Genomic Amplification with Transcript Sequencing; Sommer *et al.*, 1990) protocol needs to be developed for eventual supplementation of DNA reserve stocks and to obviate the need for replenishment from outside sources.

Technical workshops need to be conducted in order to bring experts together and develop specific techniques and protocols for DNA extraction, amplification and storage.

Institutions in developing countries continue to report difficulties in obtaining liquid nitrogen (both supplies and sustained funds). There is a need for the development of DNA storage at ambient temperatures. These procedures would also be of great benefit when a freezer is lost or liquid nitrogen is not available for brief periods.

Funding for initial start-up is modest for the reserve nodes but would require substantial funds for working nodes. Working nodes will need to be added to an existing molecular biology laboratory. Operating funds would be modest for a reserve node, depending on the cost of liquid nitrogen. Several international funding organizations may consider support for the establishment of nodes, but funding for operations will likely need to be borne by the institution or national government.

Acknowledgements

DNA Bank-Net has been supported by funds from the Conservation, Food and Health Foundation, the Helen Jones Foundation and the Wallace Genetic Foundation.

References

- Adams, R.P. (1988) The preservation of genomic DNA: DNA Bank-Net. *American Journal of Botany* 75, 156 (abst., suppl.).
- Adams, R.P. (1990) The preservation of Chihuahuan plant genomes through *in vitro* biotechnology: DNA Bank-Net, a genetic insurance policy. In: Powell, A.M., Hollander, R.R., Barlow, J.C., McGillivray, W.B. and Schmidly, D.J. (eds) *Third Symposium on Resources of the Chihuahuan Desert Region*. Printech Press, Lubbock, TX, pp. 1–9.
- Adams, R.P. and Adams, J.E. (eds) (1991) *Conservation of Plant Genes: DNA Banking and In Vitro Biotechnology*. Academic Press, New York.
- Adams, R.P., Do, N. and Chu Ge-lin (1991) Preservation of DNA in plant specimens by desiccation. In: Adams, R.P. and Adams, J.E. (eds) *Conservation of Plant Genes: DNA Banking and In Vitro Biotechnology*. Academic Press, New York, pp. 135–152.
- Adams, R.P., Miller, J.S., Golenberg, E.M. and Adams, J.E. (1994) *Conservation of Plant Genes II: Utilization of Ancient and Modern DNA*. Missouri Botanical Garden Press, St Louis, MI, 276 pp.
- Barton, K.A., Whiteley, H.R. and Yang, N. (1987) *Bacillus thuringiensis*-endotoxin expressed in transgenic *Nicotiana tabacum* provides resistance to Lepidopteran insects. *Plant Physiology* 85, 1103–1109.
- Booth, W. (1987) Combing the earth for cures to cancer, AIDS. *Science* 237, 969–970.
- Della-Cioppa, G., Bauer, S.C., Taylor, M.L., Rochester, D.E., Klein, B.K., Shah, D.M., Fraley, R.T. and Kishore, G.M. (1987) Targeting a herbicide resistant enzyme from *Escherichia coli* to chloroplasts of higher plants. *Bio/Technology* 5, 578–584.
- Fischhoff, D.A., Bowdish, K.S., Perlak, F.J., Marrone, P.G., McCormick, S. M., Niedermeyer, J.G., Dean, D.A., Kusano-Kretzmer, K., Mayer, E.J., Rochester, D.E., Rogers, S.G. and Fraley, R.T. (1987) Insect tolerant transgenic tomato plants. *Bio/Technology* 5, 807–813.
- Hawkes, J.G. (1983) *The Diversity of Crop Plants*. Harvard University Press, Cambridge, MA, 184 pp.
- Holmgren, P.K., Holmgren, N.H. and Barnett, L.C. (1990) Index herbariorum. Part I: The herbaria of the world, 8th edn. *Regnum Vegetabile* 120, 1–693.
- Kadokami, Y. and Lewis, R.V. (1990) Membrane bound PCR. *Nucleic Acids Research* 18, 3082.
- Lefebvre, D.D., Miki, B.L. and Laiberte, J.F. (1987) Mammalian metallothionein functions in plants. *Bio/Technology* 5, 1054–1056.
- Liston, A., Rieseberg, L.H., Adams, R.P., Do, N. and Zhu, Ge-lin. (1990) A method for collecting dried plant specimens for DNA and isozymes analyses, and the

- results of a field test in Xinjiang, China. *Annals of the Missouri Botanical Garden* 77, 859–863.
- Maiti, I.B., Hunt, A.G. and Wagner, G.J. (1988) Seed-transmissible expression of mammalian metallothionein in transgenic tobacco. *Biochemical and Biophysical Research Communications* 150, 640–647.
- Murai, N., Sutton, D.W., Murray, M.G., Slightom, J.L., Merlo, D.J., Reichert, N.A., Sengupta-Gopalan, C., Stock, S.A., Barker, R.J., Kemp, J.D. and Hall, T.C. (1983) Phaseolin gene from bean is expressed after transfer to sunflower via tumor-inducing plasmid vectors. *Science* 222, 476–482.
- Nelson, R.S., McCormick, S.M., Delannay, X., Dube, P., Layton, J., Anderson, E.J., Kaniewska, M., Proksch, R.K., Horsch, R.B., Rogers, S. G., Farley, R.T. and Beachy, R.N. (1988) Virus tolerance, plant growth, and field performance of transgenic tomato plants expressing coat protein from tobacco mosaic virus. *Bio/Technology* 6, 403–409.
- Newmark, P. (1987) Trypsin inhibitor confers pest resistance. *Bio/Technology* 5, 426.
- Plucknett, D.L., Smith, N.J.H., Williams, J.T. and Anishetty, N.M. (1987) *Gene Banks and the World's Food*. Princeton University Press, Princeton, NJ.
- Raven, P.H. (1988) Tropical floristics tomorrow. *Taxon* 37, 549–560.
- Redden, R.J., Singh, S.R. and Luckefahr, M.J. (1984) Breeding for cowpea resistance to Bruchids at IITA. *Protection Ecology* 7, 291–303.
- Shulman, S. (1986) Seeds of controversy. *BioScience* 36, 647–651.
- Sommer, S.S., Sarkar, G., Koeberl, D.D., Bottema, C.D.K., Buerstedde, J-M., Schowalter, D.B. and Cassady, J.D. (1990) Direct sequencing with the aid of phage promoters. In: Innis, M.A., Gelfand, D.H., Sninsky, J.J. and White, T.J. (eds) *PCR Protocols*. Academic Press, San Diego, CA, pp. 197–205.
- Ullstrup, A.J. (1972) The impacts of the southern corn leaf blight epidemics of 1970–1971. *Annual Review of Phytopathology* 10, 37–50.
- US Congress (1987) *Technologies to Maintain Biological Diversity*. Office of Technology Assessment, Washington DC.

Genetic Resources and Plant Breeding

(Identification, mapping and manipulation of simple and complex traits)

M.J. Kearsey

*Plant Genetics Group, School of Biological Sciences,
University of Birmingham, Edgbaston, Birmingham B15 2TT,
UK*

8.1. Introduction

In order for genetic resources to be efficiently utilized in plant breeding programmes, it is first necessary to determine whether useful genetic variation exists in the material and secondly to develop the most cost-effective method of introducing the potentially useful genes into commercially acceptable material. Until recently, genetic analysis has relied on conventional segregation analysis in controlled crosses for qualitative traits determined by major genes, and on standard biometrical methods for quantitative traits controlled by polygenes. Now, however, the use of DNA techniques has opened up new possibilities for genetic analysis and breeding, and in this chapter the general principles and applications of these molecular marker-based procedures will be discussed.

In its specific sense, marker-assisted or marker-aided selection (MAS) is the use of appropriate easily recognizable genetic markers to facilitate or accelerate the selection of linked genes controlling useful agronomic traits. The nature of the marker loci themselves is not important and they are used simply to indicate the presence of useful alleles of genes of commercial or practical importance. However, the term has also come to include the wider use of markers to improve breeding programmes.

8.2. What are the Uses of Marker-aided Selection?

In an ideal world, selection should be aimed directly at those genes which control the trait to be improved, but this assumes that genotypes at these target genes can be recognized easily, unambiguously and at an appropriate time without impeding the breeding programme. There are several reasons for using marker genes to improve selection efficiency.

Difficulties in identifying trait genotype

The genotypes at a single-trait locus may be difficult to recognize because the desired alleles may have poor penetrance, be recessive, or interact with other genes or the environment. Environmental variation itself may impede accurate genotyping by causing the phenotype of different genotypes to overlap; this is particularly a problem with polygenic traits. Some phenotypes such as resistance to a particular pest (Schön *et al.*, 1993), pathogen or abiotic factor, like drought or salinity, may only appear under particular conditions which are difficult to define or control. These difficulties may be more severe during the early stages of a breeding programme when plant numbers are too few to allow adequate replication or the breeder does not want to place his valuable yet scarce breeding material at hazard.

Earlier selection

Another major reason why it may be valuable to use marker or indicator genes is to reduce the time from sowing to selection. Many traits of economic importance are only apparent in the mature plant and so may not occur for months or even years after sowing. This is particularly a problem with biennial species and long-lived crops such as trees (King *et al.*, 1991). Indicator genes, on the other hand, may be detectable in seedlings within days of sowing, if not in the seeds themselves, so avoiding the waste of valuable resources involved in raising and harvesting plants, most of which may prove to be of no value to the breeding programme.

More intense selection

Selection at a juvenile stage, particularly among seedlings, may also allow much larger populations to be studied and hence more intense selection to be applied. Indeed, selection may be possible in tissue culture without raising plants or running trials.

Nondestructive scoring

Many characters are scored before maturity and the act of scoring may prevent seed being collected. Imposition of selection for pests and diseases may result in plants which are less able to produce seed. The use of certain types of marker loci may only require the removal of small quantities of leaf or other material to permit full genotyping of the plant.

Linkage drag

Conventionally plant breeders have used backcrossing with selection to introduce a useful allele, such as one conferring disease resistance, from one strain into another. The source of the allele is often an alien or exotic species, but may be a related cultivar. The F_1 is backcrossed to the recurrent parent, and backcrossing is repeated for five to ten generations, the intention being to introduce the desired donor allele whilst returning to the recurrent parent's genotype at all other loci. At every backcross generation, the breeder selects for the phenotype of the allele to be introgressed whilst simultaneously selecting for the phenotype of the recurrent parent for all other traits.

It was shown by Stam and Zeven (1981) that this procedure results in the leaving of very large amounts of donor genome associated with the selected alien allele (Fig. 8.1). On average, about 20% of an average chromosome will consist of a donor segment even after ten backcross generations, while in any given line the actual size could be much more. Thus, as well as introducing useful alleles, the breeder will also be introducing very many undesirable alleles from the donor and so the gain in one trait may be at the expense of losses elsewhere. The use of marker loci can enable the breeder to reduce the size of this unwanted associated region as well as to accelerate the speed of return to the desired genotype of the recurrent parent.

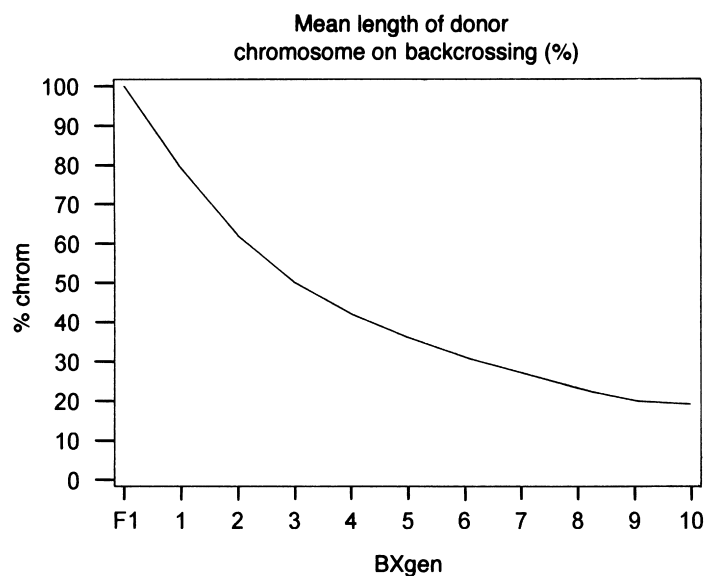


Fig. 8.1. Linkage drag. The average amount of chromosome surrounding a selected gene following backcrossing. BXgen = generation of backcrossing from F_1 .

Heterosis potential

Heterosis, or hybrid vigour, is due mainly to the presence of dispersed dominant alleles in the parental cultivars. Breeders concerned with developing lines for hybrid production are often concerned to identify lines which differ at as many gene loci as possible. It is often assumed that this implies that the lines also have very different genotypes at marker loci and hence such loci can be used to assess the genetic distance between potential parents (Dudley *et al.*, 1991; Melchinger *et al.*, 1991), but this can be a dangerous assumption unless a wide range of different types of markers are used.

Quality control

A major problem faced by most conservationists and breeders is to maintain the integrity of the material they are working with and to keep it free from contamination. Contamination can occur in various ways. It may arise through errors due to carelessness in labelling or sorting seed or plant material, mistakes in maintaining the pedigree, or as a result of outcrossing due to stray pollen from unrelated material. Such errors are difficult to recognize when only the gross phenotype is available as a guide, but genetic markers can provide a clear genetic fingerprint which can be used at any time to confirm the origin of the material (Welsh and McClellan, 1990). It can also be used to confirm that the intended cross or self has occurred, or to separate the desired progeny from a particular parent which might be a mixture of selfs and crosses (Torres *et al.*, 1982).

Other uses

Marker information can be used in other situations also: distinguishing haploids, diploids and aneuploids, for example, particularly following wide interspecific crosses or tissue culture. Doubled haploid lines produced by microspore or ovule culture can be recognized in this way and possible aneuploids arising during chromosome doubling can be identified and removed (Colby and Peirce, 1988). Not only can genetic markers identify contaminant material but they also allow closely related or identical material to be identified in genebank collections, so reducing unnecessary duplication.

8.3. Types of Genetic Markers

Up to the mid-1960s the only easily usable genetic markers were those that produced clearly visible effects on the plant's phenotype such as colour, size, shape, disease resistance, etc. (Qualset *et al.*, 1965). All these were the result of mutations to alleles which had, obviously, a major effect on the development of the plant. Such mutants were rare and generally not

desirable in commercial material although they may be selected by breeders of decorative plants for their novelty value. As genetic markers, however, they have many disadvantages. They often affect the fitness of the individual and could well have effects on important agronomic traits, i.e. they exhibit pleiotropy. Because they are rare, it would be necessary to bring several such mutants together in the same material in order to use them, and they would not normally be present in commercial cultivars or, indeed, in wild material. This not only takes time, but the combined effects of many such alleles would mitigate against their value as a genetic tool.

In the 1960s through to the 1980s, naturally occurring genetic polymorphisms at the protein level came to be easily recognized (Gebts, 1990; Hamrick and Godt, 1990). Variation in enzymes and storage proteins was detectable on an electrophoretic gel for a very high proportion of gene loci which could be studied in this way and, unlike major mutants, such naturally occurring allelic variants had much smaller effects on fitness and hence were more common and existed naturally in populations (Tanksley *et al.*, 1982). The number of such polymorphisms was, however, insufficient to provide the coverage of the genome desired by breeders and conservationists.

The development of techniques during the late 1980s up to the present (Phillips and Vasil, 1994) revealed the vast amounts of genetic polymorphism which exists at the DNA level and has revolutionized genetic analysis, opening up a whole new range of genetic tools (Beckmann and Soller, 1983; Soller and Beckmann, 1983; Tanksley, 1993; McCouch and Doerge, 1995; Stuber, 1995). As described elsewhere in this book (see Chapter 5), this variation arises through the existence of occasional base changes in the DNA which can be recognized by restriction enzymes or by primers used in analysis involving the polymerase chain reaction (PCR). Much of the DNA, possibly as much as 90% in many species, is noncoding and hence reliance on morphological or biochemical variants was only using the restricted variation that plants could tolerate within the coding parts of their DNA. Clearly, there are limits to the tolerance of such variation. Variation in the noncoding regions, whether in the intergenic regions or within introns, is probably under far less constraint by natural selection, and hence the number of polymorphic sites is potentially enormous. Because most of these sites are outside coding genes they are referred to as loci rather than genes; genes are coding regions whilst most marker loci are not.

Depending on how the DNA polymorphism is studied, various types of marker are used. Restriction fragment length polymorphisms (RFLPs) have been the most widely used until now (Gebhart and Salamini, 1992). They have the advantage of being codominant and hence all genotypes in a cross can be identified. On the other hand they require quite large quantities of DNA, and hence plant material, and generally rely on radioactive

probes, although fluorescent techniques are becoming more popular. As a result, RFLPs are expensive to use. The use of radioactive probes incurs safety considerations, as well as taking longer to visualize the bands. PCR-based markers are now being used more widely. Randomly amplified polymorphic DNAs (RAPDs), for example, require far less material because only small amounts of DNA are needed and the required sequences are then amplified. They do not involve radioactive probes and are both quick and cheap. On the other hand RAPDs are dominant although it is possible to distinguish heterozygotes from homozygotes by the amount of product. Amplified fragment length polymorphisms (AFLPs) offer even cheaper and more easily identified and extensive PCR-based polymorphism. Because of the number of polymorphisms which exist on individual gels and the speed of production, computer software is necessary to scan the banding patterns on the gels and to analyse and assimilate the information that is generated. Other popular markers are microsatellites (short sequence repeats) and cleaved amplified polymorphisms (CAP)-based primers (Talbert *et al.*, 1994).

All the genetic variants discussed above will be subsumed under the broad title of marker loci. They may be of little interest in their own right, and this is particularly true of the molecular polymorphisms; their value lies in their use as markers of more useful genes.

8.4. Gene Mapping

An important, though by no means an essential, step in genetic analysis is to produce genetic maps of the marker loci. Such maps are often referred to as 'framework maps' because they provide a framework within which important genes can be located, as well as providing a means of comparing chromosome organization in other closely or distantly related species. There are two stages to mapping. Firstly, to arrange the markers in a linear sequence separated by an appropriate map distance, i.e. to construct a linkage map (Bailey, 1961; Griffiths *et al.*, 1996; Kearsey and Pooni, 1996). The second is to relate the linkage maps to particular recognizable chromosomes. The latter is often the most difficult and generally not essential for breeding or conservation work. We will, therefore, concentrate on the former.

Chromosomes contain a single linear molecule of DNA and hence the markers on that chromosome occur at particular positions along that molecule. Typically, each chromosome contains 10^7 to 10^8 base pairs (bp), i.e. 10^4 to 10^5 kilo-base pairs (kbp) of DNA, while a typical structural gene, coding for a polypeptide chain, would be between 1 and 2 kbp long. Assuming only 10% of the genome is coding DNA, then a chromosome probably contains something of the order of 1000 to 10 000 genes.

Fortunately, the fact that the chromosome is a linear molecule means that the genes need to be mapped in only one dimension.

Currently, the only useful method of gene mapping, at least as far as breeders and conservationists are concerned, relies on recombination between homologous chromosomes resulting from genetic exchange, which can be seen as chiasmata at diplotene through first metaphase of meiosis. A single chiasma on a particular chromosome results in half the gametes from such a meiosis being recombinant for that chromosome: half the gametes will contain a copy of that chromosome that is part maternal and part paternal in origin, while the remaining gametes will be entirely parental. Very few plant species have a sufficiently low number of clearly recognizable chromosomes that the number of chiasmata on a particular chromosome can be compared in different nuclei. However, in many species the total number of chiasmata in each diplotene nucleus can be counted, at least in pollen meiosis, and the average number of chiasmata per nucleus calculated. The number and position of chiasmata on a particular pair of homologous chromosomes will vary from nucleus to nucleus; it is as though there are a large number of potential sites of exchange, but that in any given meiosis only a few sites are actually involved. The longer the chromosome the more potential sites there might be. There will invariably be one chiasma and there may be two, three, four or more; a typical set of results is shown in Table 8.1a. A chromosome that has one chiasma on average is said to be 50 centimorgans (cM) long – a map length unit named after the American geneticist Thomas Hunt Morgan. This relates to the fact mentioned above that one chiasma results in 50% recombinant chromosomes. By extension, a chromosome with an average of 2.5 chiasmata is said to be 125 cM (i.e. 2.5×50 cM) long. It appears that, as a rough rule of thumb, each chromosome has on average two chiasmata and so is 100 cM long. It follows, therefore, that providing the haploid chromosome number is known (n), the total map length will be approximately $2 \times n \times 50$ cM. The average number of chiasmata per nucleus has been calculated in many species and this rule generally holds (Table 8.1b).

Table 8.1a. Chiasma frequency and mapping. Chiasma frequency in chromosomes of *Secale cereale*.

Number of chiasmata per chromosome	Frequency	Percentage unrecombined chromosomes
1	0.267	50.0
2	0.716	25.0
≥ 3	0.017	12.5
Mean	–	22.0

Table 8.1b. Relationship between DNA content, chiasma frequency and map length in various species.

	<i>Arabidopsis thaliana</i>	Rice	<i>Brassica oleracea</i>	Tomato	Maize	Bread wheat
Chromosome number per gamete	5	10	9	12	10	21
Total genome size (kbp)*	1.5×10^5	4.4×10^5	6.1×10^5	1.3×10^6	3.0×10^6	1.6×10^7
Mean chiasma frequency per nucleus (<i>n</i>)	10	—	22	22	27	55
Expected map length ($2 \times n \times 50$ cM)	500	—	1100	1100	1350	2750
Current observed map length (cM)	501	1575	1080	1400	1350	2575
kbp per cM genome size/observed map length	299	279	565	929	2222	6214

* Arumaganathan and Earle, 1991.

This rough rule is important because it gives the geneticist an idea of the total genetic map length that should be expected as more genes or markers are mapped. It provides a guide to the extent to which the currently mapped markers cover the full genetic map. Because the map is based in chiasma units, it is not the same as a physical map. The distribution of chiasmata is not uniform over a given chromosome in a species and will vary with the chromosome and the species.

Although a knowledge of the number of chromosomes and, ideally, the mean chiasma frequency per nucleus for the species, indicates the total map length, the actual map has to be constructed from examining the frequencies of progeny in crosses; i.e. maps are constructed from information based on the consequences of chiasmata, namely recombination of genetic markers. Recombination of genetic markers is identified by the presence of gametes that contain recombined markers, and such gametes can be recognized from the phenotypes of progeny of a cross. The frequency of such gametes is an estimate of the recombination frequency between the two markers concerned (Griffiths *et al.*, 1996; Kearsey and Pooni, 1996). The closer two markers are to each other on the chromosome, the less likely a chiasma will occur between them and the less likely they are to recombine. Providing the two markers are sufficiently close that either none or one chiasma occurs between them but never more, then the recombination frequency as a percentage is equal to the map distance in centimorgans as described above. The problem arises when the markers are sufficiently far apart for two or more chiasmata to occur between them in some meioses. It can be shown that not only does a single chiasma between a pair of markers result in 50% recombination between them, but so also do two, three or more chiasmata, on average. It is necessary to say 'on average', because there are a variety of possible consequences with two, three or more chiasmata in any given meiosis but, in practice, when one is looking at the progeny of a cross or self, each progeny will be the result of gametes from different meioses.

This fact implies that although map distance increases linearly with the number of chiasmata, the same is not true with the frequency of recombination, which can never be more than 50% for a pair of markers even though they might be 100 cM apart at opposite ends of a chromosome. These relationships are illustrated in Fig. 8.2. In order to overcome this problem, mapping functions have been devised to correct for multiple chiasmata. The most common of these are the Haldane (1919) and Kosambi (1943) mapping functions. Haldane's assumes that the probability of none, one, two or more chiasmata in a given interval follows a Poisson distribution, i.e. that the chiasmata are independent, random events. Kosambi's method allows for a certain degree of nonindependence in chiasma occurrences. It has long been known by cytogeneticists that chiasma interference occurs over quite large regions of a chromosome (Mather, 1933).

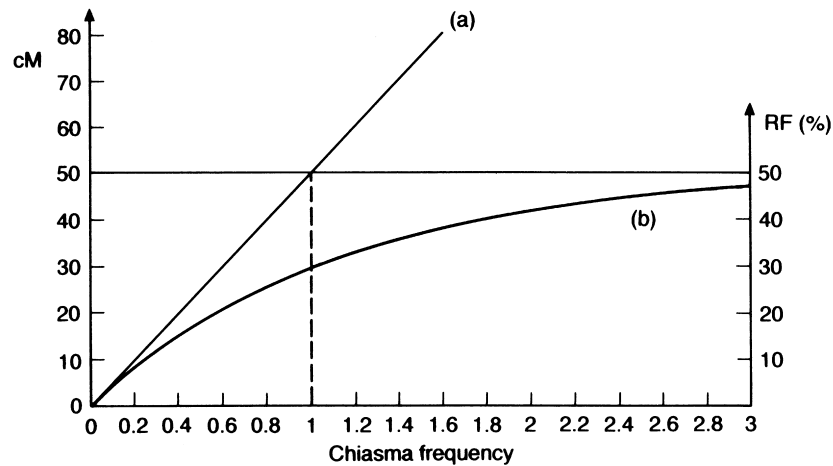


Fig. 8.2. The relationship of chiasma frequency with (a) genetic map length and (b) recombination frequency.

Interference means that if a chiasma occurs at a particular point on the chromosome, the next one will never occur closer than some fixed distance, the interference distance, from it. Beyond this distance, there is a short distance in which the interference disappears and the next chiasma can then occur randomly outside this range. Observational data in various species of plants are confirming this and suggest that the interference distance is generally between 15 and 20 cM, i.e. 15 to 20% of a typical chromosome on either side of the chiasma. This has important and useful consequences for genetic and breeding work as will be shown later.

Recombination frequencies between pairs of markers can be scored in a wide variety of different crosses. The simplest to use are generations derived from an F_1 because only two alleles are segregating, and their distribution in the chromosomes of the parents of the F_1 can be determined. The generations that can be used are F_2 , backcrosses (Bc), recombinant inbred lines (RILs) or doubled haploid (DH) lines, with F_2 s being the most informative. General formulae for calculating recombination frequencies are given by Allard (1956). In outbreeding species where F_1 s are not available, a given individual can be considered as an F_1 and its selfed progeny as being an F_2 , even though the inbred parents do not exist. If it cannot be selfed but controlled crossing to another individual is possible, then again those genes segregating in the cross can be mapped although the situation is more complex. There could be from two to four alleles segregating at each locus and, with respect to any one locus, the cross could represent an F_2 , Bc1 or Bc2. Moreover, the distribution of alleles in the two parents cannot be known and has to be inferred from the progeny. This complexity is

compounded by the large number of marker loci that may be segregating in any given cross.

Fortunately, a range of software packages are available to estimate these recombination frequencies, identify linkage groups, assign the markers to the most likely order and space them in map units (cM) on these linkage groups; examples are MAPMAKER (Lander *et al.*, 1987) and JOIN-MAP (Stam, 1993). Ideally the number of linkage groups should be equal to the chromosome number in the gametes but, unless the markers available provide a good coverage of the genome, it is likely that those markers on a given chromosome may appear as two or more separate linkage groups simply because the subsets are not sufficiently close for them to be recognized as being together on one chromosome. A typical marker framework map is shown in Fig. 8.3 for the chromosomes of *Brassica oleracea*.

Clearly these maps provide a very detailed coverage of the chromosome. Many of the markers are very close, i.e. less than 10 cM, and over this distance it matters little which of the two mapping functions is used. With the more widely spaced markers, i.e. recombination frequency >15%, Haldane's function will exaggerate their distance apart because it will allow for more double crossovers than actually occur. As more markers are mapped, the total length should converge on the value predicted by the chiasma frequency, i.e. approximately $2 \times n \times 50$ cM (Table 8.1b).

It is important to remember that genetic map distances have standard errors which are dependent on the size and type of population used to construct the map, the mapping population. The map obtained is that which best fits, given the data and the assumptions underlying it, but there may well be other maps which also fit the data well, though are slightly

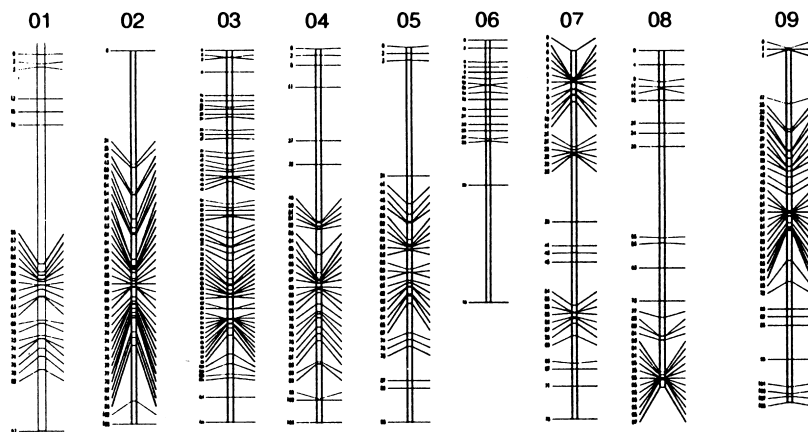


Fig. 8.3. Typical molecular marker framework map. The nine linkage groups of *Brassica oleracea* based on RFLP, isozyme and morphological markers.

less likely. Any particular map always develops a greater aura of respectability once it is published, frequently without stating its reliability, and subsequent yet different maps are treated with suspicion. A recombination frequency of true value p has a standard error of $\sqrt{p(1-p)/N}$, where N is the size of the gamete population. For example, if p is 0.1 (i.e. RF = 10%) and $N = 100$, the standard error is 0.03 or 3%. In other words the 95% confidence interval of the estimate lies approximately between 4% and 16%.

There are other reasons, apart from statistical sampling, why maps could be wrong. The most obvious is the accuracy with which the data are scored and recorded. Autoradiographs and banding patterns on gels can easily be misread and research workers are loath to discard data even though its interpretation is ambiguous. All data should be checked independently by two or more people and any ambiguities removed. Wrong scores will bias the data and often exaggerate map lengths as they suggest double recombination; wherever double recombination appears to have occurred in two short, adjacent intervals, the data should be checked. Changes in methylation rather than scoring errors could cause a restriction site to appear or disappear so giving the false impression of double crossing-over, or it could simply be an error.

Where maps appear to have major inconsistencies in different crosses, different chromosomal structural arrangements may be responsible, such as translocations or inversions. There is considerable evidence that chiasma frequencies, and hence recombination frequencies, may be quite different in male and female meioses even on the same plant (Lagercrantz and Lydiate, 1995; Kearsey *et al.*, 1996) although in other studies this may not be so (Wang *et al.*, 1995). Thus although the order of markers on a linkage group may remain the same, the relative distances between them may be quite different if the map derives from male versus female meioses. For example, in *Brassica* it appears that the genetic map obtained from a back-cross where the F_1 is the female parent is 60% longer than when the F_1 is the male parent (Kearsey *et al.*, 1996). It is also clear that environmental factors during meiosis, particularly temperature, can affect the number and distribution of chiasmata, and this could be very critical if the crosses are set up at different times of the year.

Two or more different populations will almost certainly be segregating for different combinations of polymorphic markers, although some at least should be in common. These common ones can be used to overlay the two or more maps and a consensus map produced from an amalgamation of these. Again software is available to produce these consensus maps (Stam, 1993). Despite the various error-causing factors discussed above a considerable degree of consensus can be produced from such populations and, as is shown in Chapter 4, considerable similarity of chromosome sequence is conserved between even distantly related species, allowing the

potential for cross-species genetic transfer (Ahn and Tanksley, 1993; Moore *et al.*, 1995a,b).

It was stated above that the total map length should be dictated by the chiasma frequency, in so far as this can be accurately measured. When the first few genes were initially put onto genetic maps in the first 40 years of this century, they only covered a small part of the genome – except in well-studied species such as *Drosophila* and maize. As more and more genes were placed on the map, so the total map lengths increased towards the asymptote dictated by the chiasma frequency. However, during the first decade following the use of molecular markers, the map lengths of some species appeared to exceed that expected, and it was even argued by some that the chiasma theory of recombination might be wrong (Nilsson *et al.*, 1993). No one would wish to argue that chiasma frequency data are free from error but there would have had to have been quite excessive under-scoring to result in the disparity which was occasionally found. It would appear, however, that the excessive lengths were due in part to errors in scoring and to the use of small mapping populations, because subsequent map sizes have shown a progressive decrease towards the predicted asymptote. Scoring errors bias estimates of RF generally upwards, while RFs have to be large to be detectable in small populations. Our own experience with mapping in cereals and brassicas has also shown a progressive decrease in estimated map length as more markers can be found to fill some of the large gaps in the map and as errors are removed from the mapping data.

8.5. Locating Genes of Major Effect

The use of extensive molecular and other markers as described above provides a general framework map. Although this has value in its own right for comparing linkage groups in different species as well as possible structural variation within a species (Moore *et al.*, 1995b), the main value lies in providing a set of markers to locate genes of economic or special scientific interest.

Locating individual major genes is a straightforward development of the procedures above. Let us assume that a cultivar is identified which contains an allele of interest which changes the phenotype in a clearly recognizable fashion and which appears to segregate as a single gene in crosses to other cultivars. Using a knowledge of the positions of existing marker loci, a small subset of markers can be chosen which provide a reasonably even coverage of the genome, say every 20 to 30 cM, i.e. approximately five per chromosome. These then have to be shown to differ between the two cultivars and any which are monomorphic replaced by polymorphic markers.

A rough position of the gene of interest can then be obtained by bulked segregant analysis (Giovannoni *et al.*, 1991; Michelmore *et al.*, 1991; Churchill *et al.*, 1993). This involves taking the mapping population, e.g. a Bc, and dividing the plants into two groups depending on whether or not they show the phenotype of interest. Bulk samples of DNA are taken from each group, and the two samples assayed for the marker systems chosen as in the paragraph above. Any marker which is unlinked to the gene to be located will produce similar banding patterns in both samples because it will be segregating independently of the target gene. Conversely, should one of the markers be very closely linked to the target gene, it will co-segregate and the two samples will show quite different patterns for that marker. In practice, complete co-segregation with any one marker is not very likely but, given a reasonable spread of markers over the genome, one or two are likely to show a strong association with the two groups. Thus although bands associated with both marker alleles will be found in each sample, their relative intensities will be quite different because of the association with the target gene.

Having located the chromosome and the region on the chromosome, the actual position can be identified in a segregating population using the principal marker associated in the bulked segregant analysis together with other polymorphic markers situated about 10 cM on either side of the principal marker. The target gene is then mapped with respect to the three markers. This approach can be used for a wide variety of populations and traits. Even if the trait has poor penetrance, the affected group will clearly show the marker banding pattern even if the nonaffected group consists of a mixture due to misclassification (Shalom *et al.*, 1996)

8.6. Mapping Genes Controlling Quantitative Traits

Quantitative traits such as yield, quality, height and flowering time, which are controlled by several genes and are greatly influenced by the environment, create particular difficulties for gene mapping. The problems arise because the genotype for the trait concerned can never be clearly identified from the phenotype; many different genotypes could produce the same or very similar phenotypes whilst the same genotype can result in different phenotypes depending on what may be elusive factors in the environment. Whereas with a single gene difference controlling a major effect, two or more Mendelian phenotypic classes might be recognized in ratios of 3:1, etc., quantitative traits resulting from the joint segregation of many genes show a continuous, often normally distributed range of phenotypes (Fig. 8.4). The total phenotypic variation in an F_2 , for example, V_p , is made up of genetic and environmental components, V_G and V_E (Falconer and Mackay, 1996; Kearsey and Pooni, 1996). The genes underlying such traits

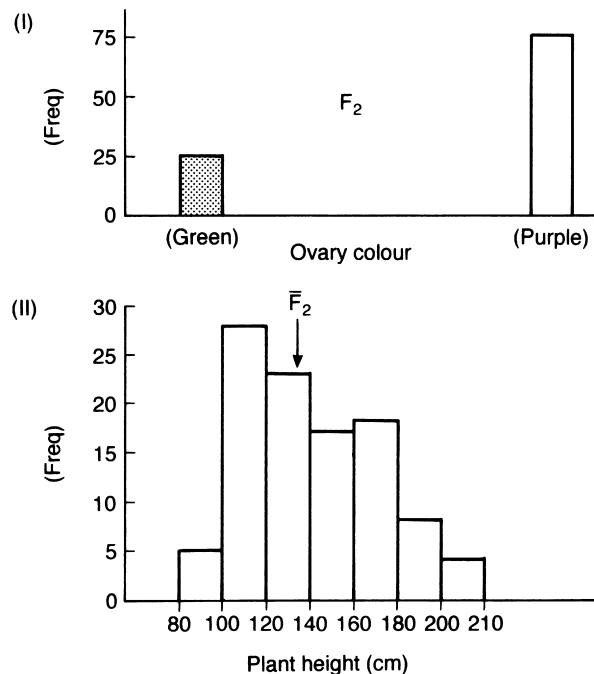


Fig. 8.4. Contrast between distributions of a qualitative and quantitative trait in an F_2 . Data for *Nicotiana rustica*, ovary colour and height. (Supplied by Dr H.S. Pooni.)

are variously called polygenes, effective factors or, more recently, quantitative trait loci or QTL (Gelderman, 1975).

However, it is still possible to map and measure the effects of the QTL by techniques analogous to those used for single major genes, i.e. by looking for the effects of co-segregation of particular marker loci with differences in the quantitative trait. Various mapping populations can be used as before, F_2 , Bc, RIL, DH or open-pollinated full sib (FS) or half sib (HS) families, and the individuals or families are scored both for their phenotype for the quantitative trait(s) and for their genotype at the marker loci. The type of population that is used will depend on the traits studied and the breeding system of the plant species but, other things being equal, an F_2 population will normally provide the most information for a given size (Soller and Brody, 1976). Unfortunately, most traits of economic importance such as yield, quality and disease resistance are not usefully scored in an F_2 population because one is interested in how the material performs in something similar to the high-density monoculture that it will encounter in agricultural practice. Measuring yield in an F_2 population with spaced plants is straightforward and statistically very

powerful but it may well not provide any insight into how yield is controlled in agricultural practice. It is therefore necessary to use a plot structure for most traits in an attempt to simulate commercial conditions. For this reason most plant breeders will prefer to work with genotypes that can be extensively and easily replicated such as RILs, DHs or F_3 s. It is also easier to explain the principles underlying QTL mapping using a population of DH lines, and so for both reasons the methodology will be illustrated with DH lines, although the same principles apply to all populations. Because several trait loci are concerned, probably on different chromosomes, bulk segregant analysis is of limited applicability, at least initially.

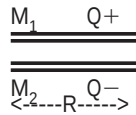
8.7. Theory of QTL Mapping in Doubled Haploid Lines

Doubled haploid (DH) lines can be produced parthenogenetically from an F_1 by microspore or ovule culture and, in some species, such as wheat and barley, by wide outcrossing (Snape *et al.*, 1986). Each original DH plant is derived from a single gamete, which subsequently becomes homozygous and disomic either naturally or by treatment with the drug colchicine, which inhibits spindle formation, and hence chromosome disjunction, at metaphase of mitosis. Selfed seeds from such a plant produce a DH line of identical individuals so the line becomes effectively immortal.

If we consider an F_1 that is heterozygous for a marker locus M_1/M_2 and for a linked QTL locus Q^+/Q^- , where the $+/-$ indicates the allele with an increasing or decreasing effect on the trait, this will produce the gametes and hence the DH lines shown in Table 8.2. If the DH lines are scored for some trait for which the Q^+/Q^+ homozygotes increase the trait above the overall mean, m , by a units while the Q^-/Q^- homozygotes decrease the mean by a , then the means of the four gametic types are as shown in Table 8.2, where R is the recombination frequency between the QTL and the marker. From this it follows that the mean trait score of all those DH lines which are M_1/M_1 ($\overline{M_1M_1}$) is $m + a(1 - 2R)$, while the mean of all those DH lines which are M_2/M_2 ($\overline{M_2M_2}$) is $m - a(1 - 2R)$. In other words, half the difference between these two means ($\overline{M_1M_1} - \overline{M_2M_2}$) is $a(1 - 2R)$. Clearly, if the marker is so close to the QTL that it never recombines, R will be 0 and the marker difference will represent the QTL effect, a . Conversely if the two loci are unlinked, i.e. $R = 0.5$, then the difference will be zero. This effect is, of course, the basis of bulked segregant analysis. In general, the magnitude of the marker effect will decline linearly with the distance of the marker from the QTL in terms of recombination frequency, and so with several linked markers, their individual trait effects will be as shown in Fig. 8.5. Similar arguments apply to other populations such as F_2 s and RILs.

Table 8.2. 2 QTL model for marker analysis in doubled haploid (DH) lines from an F_1 . (a) The F_1 constitution; M = Marker and Q = QTL genotype. (b) DH genotypes, frequencies and genetic values. (c) Expectations of marker means and differences. R = recombination frequency between M and Q; m and a are the mean and additive genetic effects of the QTL; δ is the additive genetic deviation at the marker locus

(a)



(b)

	F ₁ gametes			
	M ₁ Q ⁺	M ₁ Q ⁻	M ₂ Q ⁺	M ₂ Q ⁻
Frequency	$s(1 - R)$	$s(R)$	$s(R)$	$s(1 - R)$
DH genotype	M ₁ M ₁ Q ⁺ Q ⁺	M ₁ M ₁ Q ⁻ Q ⁻	M ₂ M ₂ Q ⁺ Q ⁺	M ₂ M ₂ Q ⁻ Q ⁻
Genetic value	$m + a$	$m - a$	$m + a$	$m - a$

(c)

$$\begin{aligned}\overline{M_1 M_1} &= s(1 - R)(m + a) + s(R)(m - a) \\ &= m + (1 - R)a \\ \overline{M_2 M_2} &= s(R)(m + a) + s(1 - R)(m - a) \\ &= m - (1 - R)a\end{aligned}$$

$$\begin{aligned}\delta &= (\overline{M_1 M_1} - \overline{M_2 M_2})/2 \\ &= (1 - 2R)a\end{aligned}$$

If the trait means of the markers and their map are known very accurately, and only one QTL existed on the chromosome, then as Fig. 8.5 shows, it would be relatively easy to locate the QTL. Figure 8.6 shows the distribution of flowering time among a number of doubled haploid lines of *Brassica oleracea* containing alternative alleles at a marker located very close to a QTL. In practice neither the map positions of the markers nor their means are known with very great accuracy because the number of DH lines and replicated plots are generally few and the heritability of quantitative traits is generally low. Moreover, linked QTLs create difficulties because their individual effects can combine either to create the impression of an intermediate, ghost QTL, or to conceal their effects altogether. Various analytical procedures have been developed to maximize the efficiency and accuracy of locating individual QTLs and to separate linked QTLs (Lander and Botstein, 1989; Knapp *et al.*, 1990; Haley and Knott, 1992; Luo and Kearsey, 1992; Martinez and Curnow, 1992; Jansen, 1993; Kearsey and Hyne, 1994; Zeng, 1994; Kruglyak and Lander, 1995). They all rely on the relationships described in the previous paragraph and

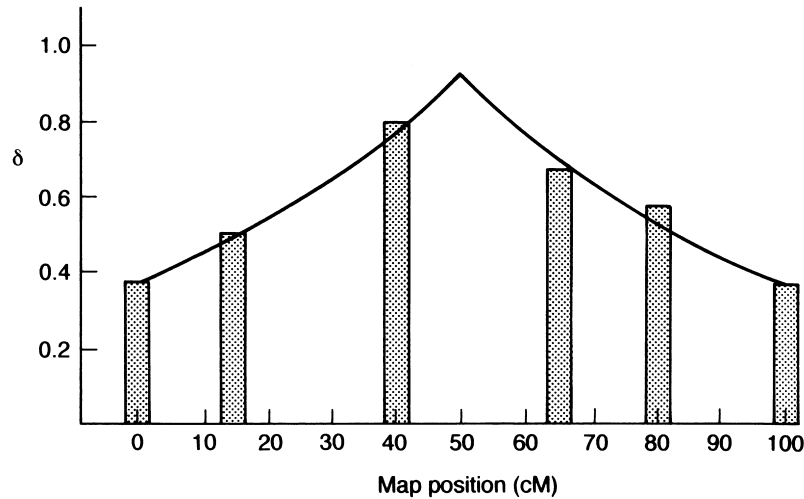


Fig. 8.5. Quantitative trait effects (δ) associated with marker loci. The bars indicate the observed effects, while the curve illustrates the expected relationship if a QTL of $\delta = 1.0$ existed at 50 cM.

use either maximum likelihood or weighted least squares regression procedures to estimate the QTL locations and effects that best fit the observed trait scores for each genotype. The main difficulty is that there are two unknown parameters with respect to each marker score; the QTL effect, a , and map position. This means that it is necessary, in all methods, to use an iterative approach which involves trying all possible QTL locations along each chromosome and identifying the position or positions of the QTL which best fit the observed data as indicated by the size of the likelihood or the residual regression variance. As a result, a very large number of statistical tests are performed with the concomitant risk of false positive results.

All methods provide estimates of QTL locations and effects with similar precision in terms of confidence intervals (Haley and Knott, 1992; Hyne *et al.*, 1995), and this precision drops rapidly as the heritability of individual QTLs declines, i.e. the estimates are accurate only when there are just two or three unlinked QTLs with large effects (Darvasi *et al.*, 1993; Hyne *et al.*, 1995). For example, Hyne *et al.* (1995) showed that the 95% confidence interval for the location of a QTL contributing 10% to the phenotypic variance of an F_2 could be as large as 35 cM. However, conservationists and breeders who are concerned with introducing QTL into commercial cultivars from other cultivars or more distantly related species, would normally only be interested in those cases where a few QTLs of major effect are to be manipulated. There is no evidence that dense maps are required for initial

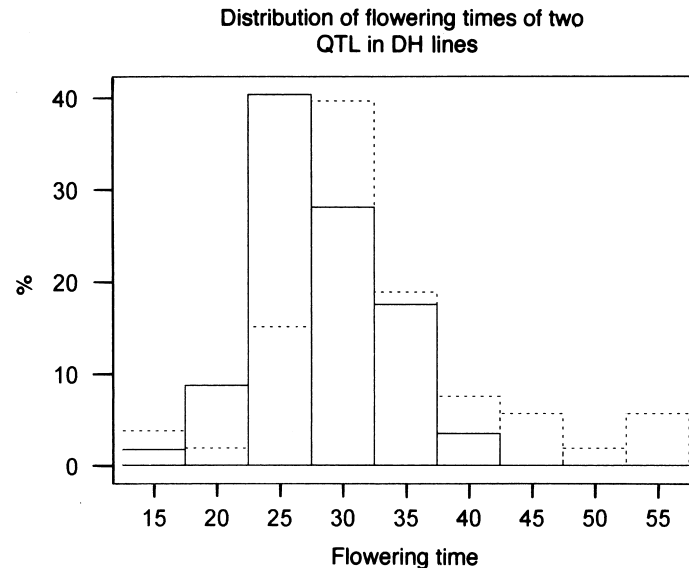


Fig. 8.6. Distribution of flowering times among doubled haploid (DH) lines in *Brassica oleracea* containing alternative alleles of a marker linked to a QTL. The two sets have significantly different means indicating a QTL with an effect of approximately ± 2 days.

QTL location (Hyne *et al.*, 1995). It is better to have large populations and a few (5–10) markers per chromosome.

There is evidence that many of the larger QTLs located so far map closely to previously known major genes. This is often used as evidence that there are, in fact, very few QTLs, but this may well be misleading for several reasons. The QTL map locations are sufficiently imprecise that they could well appear spuriously close to candidate loci. Only QTLs of large effect will stand much chance of being located and, by definition, they are likely to be major genes. It is likely that many cases where QTLs have very large effects could be due to the chance association of alleles of like effect at several QTLs along a chromosome, which cannot be separated as individual effects. Much effort has been devoted to increasing the precision of QTL location, in particular in improving the power of the test to reduce the failure to detect genuine QTLs (Jansen, 1994; Jansen and Stam, 1994).

A long-established problem with breeding for quantitative traits in plants is the genotype $\times$ environment interaction. Using QTL location techniques, it is now possible to explore these effects at the level of individual QTLs (Paterson *et al.*, 1991; Hayes *et al.*, 1993) although it is often difficult to know whether genotype $\times$ environment or poor repeatability is responsible (van Ooijen, 1992).

8.8. Alien Gene Transfer

Once useful genes have been located on a framework map, it is then possible to use the molecular markers to facilitate the transfer of the useful genes between species and strains in an efficient manner. As stated earlier, recombination is not a frequent event along a chromosome and, furthermore, two recombination events rarely occur within a distance of 15 to 20 cM of each other. It therefore follows that if a useful target gene (a major gene or a QTL) is known to be located within a region of chromosome flanked by two markers that are no more than 15 cM apart, then these markers can be used to follow and control the progression of the target gene through successive stages of a breeding programme (Fig. 8.7). Clearly, the alleles at the marker loci have to differ in the donor and recipient cultivars, but any chromosome that contains A_1 and B_1 at the marker loci will also contain T_1 , the target allele to be transferred. Should single recombination occur close to T , then the chromosome will contain A_1 or B_1 but not both. Providing the markers are less than 15 cM apart double recombination resulting in $A_1T_2B_1$, and hence loss of T_1 , can be safely ignored. If the location of T is not accurately known, then three or more marker loci may be needed to be sure of safely bracketing the region containing T without fear of double recombination.

There are conflicting requirements in marker-assisted gene transfer: the need to keep the bracketed region large enough to be sure of holding the gene to be transferred whilst not having it so large that too many other linked but undesirable alleles are transferred with it. Let us now consider the basic procedure of gene transfer.

Traditionally, breeders have introduced a useful allele into their cultivars by backcrossing the original F_1 to the commercial cultivar whilst selecting at each generation for the desired allele. As was shown earlier, this can result in a very large region of chromosome around the target gene surviving even after ten generations of backcrossing – a phenomenon known as linkage drag. With markers, however, it is possible to reduce linkage drag considerably whilst simultaneously reducing the number of backcrossing generations to two or three instead of the normal six to eight. Moreover, the amount of plant material raised at each generation can be reduced because much of the molecular genotyping can be achieved at the juvenile stage. The aim is to hold the target allele, T_1 , heterozygous throughout the backcrossing process, by selecting for the flanking markers, whilst simultaneously selecting for the genotype of the recurrent, commercial cultivar at all other loci. The latter can readily be achieved by having as few as four or five well-spaced markers on all chromosomes and selecting for recurrent parent alleles at each generation. Depending on the chromosome number, some individuals in the first backcross generation will be homozygous for several whole chromosomes whilst still being

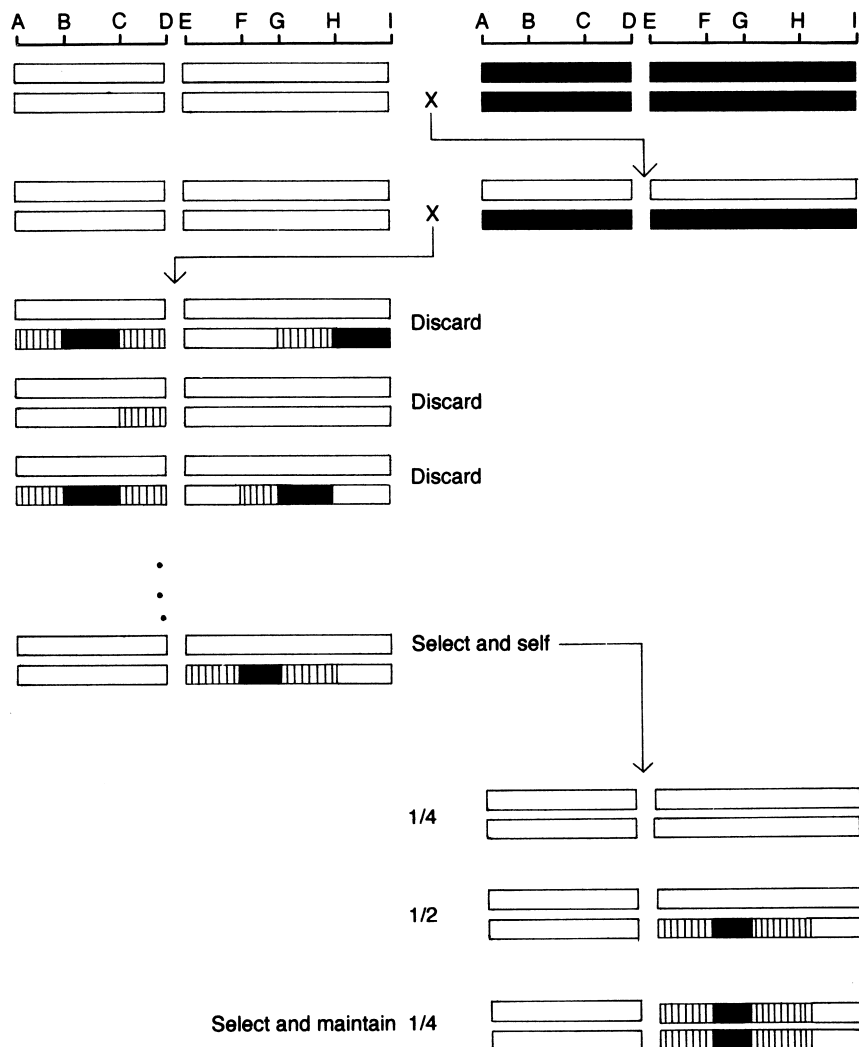


Fig. 8.7. The use of molecular markers to facilitate gene introgression.

heterozygous for the target gene. Those which have most of the recurrent parent chromosomes are selected and backcrossed again, and if necessary the process is repeated for a third generation. For each generation, fewer markers need to be screened as those found to be homozygous in the previous backcross no longer have to be checked (Howell *et al.*, 1996; Ramsay *et al.*, 1996).

Unless the position of the target gene is very accurately defined, it is prudent to have several markers covering its likely position, again with the proviso that they need to be separated by no more than 15 cM. At the final stage of the introgression process, a backcross individual which is homozygous for most of the recurrent parent alleles is selfed or, if this is not possible, intercrossed with a similar genotype. From among the progeny, individuals are chosen which have different combinations of markers bracketing the target region as shown in Fig. 8.7. These can be selfed again and homozygotes for the different sequences identified and multiplied for reassessment for the target trait. Some will fail to show the target trait because of recombination, but of those that do, the lines with the shortest sequence of donor markers are selected for further trials. The approach can obviously be adapted to the simultaneous introgression of several genes: no new principles are involved but the screening process becomes more elaborate.

It is always possible at a later stage to reduce the length of the introgressed region around T by looking at other, more closely linked, markers. Similarly, it is wise to check the origins of the other chromosomes to avoid having introgressed unwanted regions through missing double recombinants in the central regions of individual chromosomes or single recombinants towards their ends. If such errors are found they can easily be corrected by another round of backcrossing to the recurrent parent. When attempting to reduce the length of donor chromosome surrounding the target gene, it is important to realize that this should be done in two successive generations. If the sites of the required recombination events are less than 15 cM apart, then the simultaneous double event will not occur. Even if they are far enough apart for the double event to occur, the probability of such an event may be too low to be practicable. If two recombination events each have a probability of 0.05 (i.e. 5%), their combined probability is at best 0.0025. Therefore, it would be necessary to genotype approximately 1200 individuals to be 95% sure of having at least one double recombinant. If it was tackled in two stages then it would require only the genotyping of 86 (= 28 in the first round plus 58 in the second; less being needed in the first round because the recombination could occur on either side of T), a very considerable saving in effort of 93%, or 1114 plants!

Marker-aided selection has been successfully tried for quantitative traits in several crops (Stuber and Sisco, 1991; Hospital *et al.*, 1992; Stuber, 1995) while the efficiency of the procedure has been examined by Lande and Thompson (1990).

8.9. Map-based Gene Cloning

Providing a target gene can be located to within a pair of marker loci, it should be possible to use standard cloning techniques to walk along the chromosome between the markers and to find the target gene. The feasibility of this depends on the distance between the markers and the ease with which the target gene can be recognized from among the clones (Wicking and Williamson, 1991).

Depending on the species, a genetic length of 1 cM can, on average, consist of as few as 280 kbp of DNA in rice, or 2220 kbp in maize (Table 8.1b). It is not always easy to locate a target gene with sufficient accuracy to be able to say that it is between two markers as close as 1 cM, and hence the actual lengths of DNA between them can often be much greater. With QTLs one can seldom achieve anything approaching this accuracy by existing methods of conventional mapping. Furthermore, these are average distances per centimorgan; some chromosomal regions could be much longer or, indeed, less.

Increases in mapping reliability can be obtained by using what are called near isogenic lines (NILs), which are lines obtained by selfing or backcrossing and are known to differ from some standard genotype by just a short defined section of chromosome. Providing this section is delineated by accurately mapped flanking markers, and the NILs which do and do not contain this region can clearly be shown to differ for the target gene, then the target gene has to be in that region. The genetic and physical proximity of the markers can be established by conventional mapping and by gene cloning.

Map-based gene cloning involves starting from one of the markers and walking to the next by means of a series of partially overlapping cosmid clones. Cosmid clones are necessary because they can be up to 50 kb in length and hence require fewer cloning steps in order to cover the DNA between the markers (280 to 2200 kb per cM in the examples cited above).

Identifying the target gene is more difficult. Clearly, if the gene is between the two flanking markers it has to be on one of the cloned regions and could possibly be on two overlapping clones. It could be identified by transforming plants with each clone separately and identifying which transformant expresses the effect. If the gene concerned exhibits classical dominance it would be necessary to transform the homozygous recessive with the dominant allele or the transformant would not be recognized. Alternatively, the gene has to be introduced into a species that does not normally express the gene. Once the effective cloned fragment is identified, the actual location of the gene within the fragment could be identified by successively transforming sub-fractions of the clone or, following sequencing, by looking for potential open reading frames. In those cases where the target gene produces a known product, cDNA clones derived

from tissue likely to be rich in the appropriate mRNA of the target gene can be used to identify likely sequences in the previous overlapping clones.

Once the gene has been identified by one of these methods, its structure and activity can be studied in detail. Disease resistance genes are obvious candidates for such studies as their action is very specific and their location on the genome of many species is well documented. However, unless suitable candidate loci are available, identifying QTLs in this way may prove more difficult.

8.10. Summary

It is clear that molecular markers have opened up a wide range of new opportunities for analysing and utilizing plant genetic resources as well as exploiting them more efficiently in the future. One should, however, maintain a sense of proportion when contemplating the use of molecular techniques. Conventional selection procedures, if applied efficiently, still have enormous potential in a wide range of breeding situations involving quantitative as well as qualitative traits. Marker technology will prove most useful for traits of major importance in high-value crops.

References

- Ahn, S.N. and Tanksley, S.D. (1993) Comparative linkage maps of the rice and maize genomes. *Proceedings of the National Academy of Sciences of the USA* 90, 7980–7984.
- Allard, R.W. (1956) Formulas and tables to facilitate the calculation of recombination frequencies in heredity. *Hilgardia* 24, 235–277.
- Arumaganathan, K. and Earle, E.D (1991) Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter* 9(3), 208–218
- Bailey, N.T.J. (1961) *Introduction to the Mathematical Theory of Linkage*. Oxford University Press, Oxford.
- Beckmann, J.S. and Soller, M. (1983) Restriction fragment length polymorphisms in genetic improvement: methodologies, mapping and costs. *Theoretical and Applied Genetics* 67, 35–43.
- Churchill, G.A., Giovannoni, J.J. and Tanksley, S.D. (1993) Pooled-sampling makes high-resolution mapping practical with DNA markers. *Proceedings of the National Academy of Sciences of the USA* 90, 16–20.
- Colby, L.W. and Peirce, L.C. (1988) Using an isozyme marker to identify doubled haploids from anther culture in asparagus. *Horticultural Science* 23, 761–763.
- Darvasi, A., Weintreb, A., Minke, V., Weller, J. and Soller, M. (1993) Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. *Genetics* 134, 943–951.

- Dudley, J.W., Saghai Maroof, M.A. and Rufener, G.K. (1991) Molecular markers and grouping of parents in maize breeding programs. *Crop Science* 31, 718–723.
- Falconer, D.S. and Mackay, T. (1996) *Introduction to Quantitative Genetics*, 4th edn. Longman, Harlow, UK.
- Gebhardt, C. and Salamini, F. (1992) Restriction fragment length polymorphism analysis of plant genomes and its application to plant breeding. *International Review of Cytology* 135, 201–237.
- Gebts, P. (1990) Genetic diversity of seed storage proteins in plants. In: Brown, H.D., Clegg, M.T., Kahler, A.L. and Weir, B.S. (eds) *Plant Population Genetics, Breeding and Genetic Resources*. Sinauer Associates, Sunderland, MA, pp. 64–82.
- Gelderman, H. (1975) Investigation on inheritance of quantitative characters in animals by gene markers. I. Methods. *Theoretical and Applied Genetics* 46, 300–319.
- Giovannoni, J.J., Wing, R.A., Ganai, M.W. and Tanksley, S.D. (1991) Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. *Nucleic Acids Research* 19, 6553–6558.
- Griffiths, A.J.F., Miller, J.H., Gebhart, W.M., Lewontin, R.C. and Suzuki, D.T. (1996) *An Introduction to Genetic Analysis*, 5th edn. W.H. Freeman, New York.
- Haldane, J.B.S. (1919) The combination of linkage values and the calculation of distances between the loci of linked factors. *Journal of Genetics* 8, 299–309.
- Haley, C.S. and Knott, S.A. (1992) A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* 69, 315–324.
- Hamrick, J.L. and Godt, M.J. (1990) Allozyme diversity in plants. In: Brown, H.D., Clegg, M.T., Kahler, A.L. and Weir, B.S. (eds) *Plant Population Genetics, Breeding and Genetic Resources*. Sinauer Associates, Sunderland, MA, pp. 43–63.
- Hayes, P.M., Liu, B.H., Knapp, S.J., Chen, F., Jones, B., Blake, T., Franckowiak, J., Rasmusson, D., Sorrells, M., Ullrich, S.E., Wesenberg, D. and Kleinhofs, A. (1993) Quantitative trait locus effects and environmental interaction in a sample of North American barley germ plasm. *Theoretical and Applied Genetics* 87, 392–401.
- Hospital, F., Chevalet, C. and Mulsant, P. (1992) Using markers in gene introgression breeding programs. *Genetics* 132, 1199–1210.
- Howell, P.M., Marshall, D.F. and Lydiate, D.J. (1996) Towards developing inter-varietal substitution lines in *Brassica napus* using marker-assisted selection. *Genome* 39, 348–358.
- Hyne, V., Kearsey, M.J., Pike, D.J. and Snape, J.W. (1995) QTL analysis: unreliability and bias in estimation procedures. *Molecular Breeding* 1, 273–282.
- Jansen, R.C. (1993) Interval mapping of multiple quantitative trait loci. *Genetics* 135, 205–211.
- Jansen, R.C. (1994) Controlling the type I and type II errors in mapping quantitative trait loci by using molecular markers. *Genetics* 138, 871–884.
- Jansen, R.C. and Stam, P. (1994) High resolution of quantitative traits into multiple loci via interval mapping. *Genetics* 136, 1447–1455.
- Kearsey, M.J. and Hyne, V. (1994) QTL analysis: a simple ‘marker regression’ approach. *Theoretical and Applied Genetics* 89, 698–702.
- Kearsey, M.J. and Pooni, H.S. (1996) *The Genetical Analysis of Quantitative Traits*. Chapman & Hall, London.
- Kearsey, M.J., Ramsay, L.D., Jennings, D., Lydiate, D.J., Bohuon, E. and Marshall, D.F. (1996) Higher recombination frequencies in female compared to male

- meioises in *Brassica oleracea*. *Theoretical and Applied Genetics* 92, 363–367.
- King, G.J., Alston, F.A., Batlle, I., Chevreau, E., Gessler, C., Janse, J., Lindhout, P., Manganaris, A.G., Sansavini, S., Schmidt, H. and Tobutt, K. (1991) The European Apple Genome Mapping Project – developing a strategy for mapping genes coding for agronomic characters in tree species. *Euphytica* 56, 89–94.
- Knapp, S.J., Bridges, W.C. and Birkes, D. (1990) Mapping quantitative trait loci using molecular marker linkage maps. *Theoretical and Applied Genetics* 79, 583–592.
- Kosambi, D.D. (1943) The estimation of map units from recombination values. *Annals of Eugenics London* 12, 172–175.
- Kruglyak, L. and Lander, E.S. (1995) A nonparametric approach for mapping quantitative trait loci. *Genetics* 139, 1421–1428.
- Lagercrantz, U. and Lydiate, D.J. (1995) RFLP mapping in *Brassica nigra* indicates differing recombination rates in male and female meiosis. *Genome* 38, 255–264.
- Lande, R. and Thompson, R. (1990) Efficiency of marker assisted selection in the improvement of quantitative traits. *Genetics* 124, 743–756.
- Lander, E.S. and Botstein, D. (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121, 185–199.
- Lander, E.S., Green, P., Abrahamson, J., Barlow, A., Daly, M.J., Lincoln, S.E. and Newburg, L. (1987) MAPMAKER; an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1, 174–181.
- Luo, Z. and Kearsey, M.J. (1992) Interval mapping of quantitative trait loci in an F_2 population. *Heredity* 69, 236–242.
- McCouch, S.R. and Doerge, R.W. (1995) QTL mapping in rice. *Trends in Genetics* 11, 482–487.
- Martinez, O. and Curnow, R.N.C. (1992) Estimating the locations and sizes of the effects of quantitative trait loci using flanking markers. *Theoretical and Applied Genetics* 85, 401–488.
- Mather, K.M. (1933) The relation between chiasmata and crossing over in diploid and triploid *Drosophila melanogaster*. *Journal of Genetics* 27, 243–259.
- Melchinger, A.E., Mesmer, M.M., Lee, M., Woodman, W.L. and Lamkey, K.R. (1991) Diversity and relationship among U.S. maize inbreds revealed by restriction fragment length polymorphisms. *Crop Science* 31, 669–678.
- Michelmore, R.W., Paran, W.I. and Kesseli, R.V. (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis; a rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the National Academy of Sciences of the USA* 88, 9828–9832.
- Moore, G., Devos, K., Wang, Z. and Gale, M.D. (1995a) Grasses line up and form a circle. *Current Biology* 5, 737–739.
- Moore, G., Foote, T., Helentjaris, T., Devos, K., Kurata, N. and Gale, M.D. (1995b) Was there a single ancestral cereal chromosome? *Trends in Genetics* 11, 81–82.
- Nilsson, N.O., Sall, J. and Bengtsson, B.O. (1993) Chiasma data and recombination in plants – are they compatible? *Trends in Genetics* 9, 344–348.
- Paterson, A.H., Damon, S., Hewitt, J.D., Zamir, D., Rabinowitch, H.D., Lincoln, S.E., Lander, E.S. and Tanksley, S.D. (1991) Mendelian factors underlying quantita-

- tive traits in tomato: comparison across species, generations and environments. *Genetics* 127, 181–197.
- Phillips, R.L. and Vasil, I.K. (1994) DNA-based markers in plants. In: Phillips, R.L. and Vasil I.K. (eds) *Advances in Cellular and Molecular Biology of Plants*, vol. 1. Kluwer Academic Publishers, Dordrecht.
- Qualset, C.O., Schaller, C.W. and Williams, J.C. (1965) Performance of isogenic lines of barley as influenced by awn length, linkage blocks and environment. *Crop Science* 5, 489–494.
- Ramsay, L.D., Jennings, D.E., Bohuon, A.E., Lydiate, D.J., Arthur, A.E., Kearsey, M.J. and Marshall, D.F. (1996) The construction of a substitution library of recombinant backcross lines in *Brassica oleracea* for the precision mapping of quantitative trait loci (QTL). *Genome* 39, 558–567.
- Schön, C.C., Lee, M., Melchinger, A.E., Guthrie, W.D. and Woodman, W.L. (1993) Mapping and characterisation of quantitative trait loci affecting resistance against second generation European corn borer in maize with the aid of RFLPs. *Heredity* 70, 648–659.
- Shalom, A., Darvasi, A., Barendse, W., Cheng, H. and Soller, M. (1996) Single parent segregant pools for allocation of markers to a specified chromosomal region in outcrossing species. *Animal Genetics* 27, 9–17.
- Snape, J.W., de Buyser, J., Henry, Y. and Simpson, E. (1986) A comparison of methods of haploid production in a cross in wheat (*Triticum aestivum*). *Zeitung Pflanzenzuchtung* 96, 320–330.
- Soller, M. and Beckmann, J.S. (1983) Genetic polymorphism in varietal identification and genetic improvement. *Theoretical and Applied Genetics* 67, 25–33.
- Soller, M. and Brody, T. (1976) On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. *Theoretical and Applied Genetics* 47, 35–39.
- Stam, P. (1993) Construction of integrated genetic linkage maps by means of a computer package: Joinmap. *The Plant Journal* 5, 739–744.
- Stam, P. and Zeven, A.C. (1981) The theoretical proportion of the donor genome in near isogenic lines of self-fertilisers bred by backcrossing. *Euphytica* 30, 227–238.
- Stuber, C.W. (1995) Mapping and manipulating quantitative traits in maize. *Trends in Genetics* 11, 477–481.
- Stuber, C.W. and Sisco, P.H. (1991) Marker facilitated transfer of QTL alleles between elite inbred lines and responses in hybrids. In: *Proceedings of the 46th Annual Corn and Sorghum Industry Research Conference, American Seed Trade Association* 46, 104–113.
- Talbert, L.E., Blake, N.K., Chee, P.W., Blake, T. and Magyar, G.M. (1994) Evaluation of 'sequence-tagged-site' PCR products as molecular markers in wheat. *Theoretical and Applied Genetics* 87, 789–794.
- Tanksley, S.D. (1993) Mapping polygenes. *Annual Review of Genetics* 27, 205–233.
- Tanksley, S.D., Medina-Filho, H. and Rick, C.M. (1982) Use of naturally occurring enzyme variation to detect and map genes controlling quantitative traits in an interspecific backcross of tomato. *Heredity* 49, 11–25.
- Torres, A.M., Soost, R.K. and Mau-Lastovicka, T. (1982) Citrus isozymes. Genetics and distinguishing nucellar from zygotic seedlings. *Heredity* 73, 335–339.
- van Ooijen, J.W. (1992) Accuracy of mapping quantitative trait loci in autogamous

- species. *Theoretical and Applied Genetics* 84, 803–811.
- Wang, G., Hyne, V., Chao, S., Henry, Y., De Buyser, J., Gale, M.D. and Snape, J.W. (1995) A comparison of male and female recombination frequency in wheat using RFLP maps of homoeologous group 6 and 7 chromosomes. *Theoretical and Applied Genetics* 91, 744–746.
- Welsh, J. and McClellan, M. (1990) Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research* 18, 7213–7218.
- Wicking, C. and Williamson, B. (1991) From linked marker to genes. *Trends in Genetics* 7, 288–293.
- Zeng, Z.-B. (1994) Precision mapping of quantitative trait loci. *Genetics* 136, 1457–1468.

Gene Identification, Isolation and Transfer

I.D. Godwin

*Department of Agriculture, University of Queensland,
Brisbane, Queensland 4072, Australia*

9.1. Introduction

The conservation of plant genetic resources has long been espoused as an essential part of plant genetic improvement. Without genetic diversity it would not have been possible to make the genetic advances seen in crops worldwide this century. Indeed the very basis of the 'Green Revolution' lay in the ability to harness the benefits of plant breeding and crop physiology, and germplasm collections of important species such as wheat, rice and maize underpinned these improvements. However, with the Green Revolution came a growing awareness that the use of high-yielding, genetically uniform cultivars across the world led to increasing genetic vulnerability (Ford-Lloyd and Jackson, 1986). Much of the genetic gain, in terms of yield and/or adaptation to biotic and abiotic stresses, has been realized in the major crop species, most notably the cereals, grain legumes, oilseeds and, to a lesser extent, fibres and horticultural crops. The quality of many crops has altered such that for a number of species, the produce seen on modern Western supermarket shelves bears little resemblance to the commodities of the late nineteenth century. Some would say this has been to the detriment of crops such as tomatoes, the most commonly complained-about fresh produce (Van Bruna, 1992).

The source of most of this genetic diversity has generally come from within the cultivated species. Nevertheless there are many notable instances where extremely useful traits have originated from wild relatives, both at the interspecific and intergeneric levels. This is exemplified by crops such as wheat, with introgression from within and beyond the

genus, most notably from rye, with the 1B/1R translocation furnishing resistance genes to fungal diseases such as powdery mildew, and stem and leaf rust. This translocation is widely used in many European wheat cultivars (Mettin *et al.*, 1973). Similarly, stripe rust resistance from *Agropyron elongatum* was transferred to wheat with a radiation-induced translocation (Knott, 1961) and has been used extensively in Australian wheat cultivars (Driscoll, 1981).

Many of the disease resistance genes in tomato originate from the wild species of South America including *Lycopersicon hirsutum* (early blight resistance) (Saccardo *et al.*, 1975), *L. pimpinellifolium* (*Fusarium* wilt resistance) (Bohn and Tucker, 1940) and *L. peruvianum* (TMV resistance) (Alexander, 1963) and introgressed into cultivated tomato (Kalloo, 1991). Many cultivars of banana and plantain are naturally occurring triploid and tetraploid interspecific hybrids of the Southern Asian and Southeast Asian species, *Musa acuminata* $\times$ *M. balbisiana* (Simmonds and Shepherd, 1955). There are edible diploid forms of *M. acuminata* (A genome) but not of *M. balbisiana* (B genome). All sugar-cane cultivars are complex aneuploid interspecific hybrids, and although based on the noble cane, *Saccharum officinarum*, cultivars contain genetic material from *S. spontaneum* and *S. robusta*.

A number of very important crops are allopolyploids of existing diploids. In some cases, such as oilseed rape or canola (*Brassica napus*), the allopolyploid is a naturally occurring hybrid of two cultivated species, namely cabbage (*B. oleracea*) and turnip (*B. campestris*) (U, 1935), which arose from spontaneous chromosome doubling or the fusion of unreduced gametes. More, however, are the result of some historical hybridization involving species which are of little or no commercial importance. Cotton (*Gossypium hirsutum*), for example, is an allotetraploid derived from the diploid progenitors *G. herbaceum* and *G. thurberi*. As a result, the diploid progenitors are potentially useful sources of genetic variation, as long as sexual introgression of wild genes from this secondary gene pool is possible.

One particularly illustrative example of the importance of germplasm conservation is the alpine weed species, *Arabidopsis thaliana*. Although of no intrinsic economic value, this species has played a central role in the advances made in plant molecular genetics in the past decade. The species has great attractions as a model plant species: it has a very small genome, with little repeat sequence DNA; it is rapid cycling, with only 6–8 weeks generation turnover time; it takes up little space, even for highly replicated experiments; and it can be crossed / selfed and therefore is amenable to many experimental designs (Redei, 1975; Meinke and Sussex, 1979).

In 1983, the first report of a genetically engineered plant was published. A transgenic tobacco plant had been produced by utilizing genetically manipulated strains of *Agrobacterium tumefaciens*, introducing a bacterial gene encoding neomycin phosphotransferase, which rendered

the plant resistant to a group of antibiotics, including kanamycin (Zambryski *et al.*, 1983). Shortly after, similar plants were produced using a direct gene transfer system based on protoplast DNA uptake (Paszkowski *et al.*, 1984). In the decade or so since, a plethora of different methods and modifications of methods have been applied to produce transgenic plants of most important crop species.

With the development of plant genetic engineering technology, it is necessary to reassess the basis for plant genetic resource conservation. No longer is the gene pool for a species such as tomato within the *Lycopersicon* genus or even the *Solanaceae*. Genes may now be transferred to tomato not only across the family barrier to include all plant species, but from sources outside the plant kingdom including bacteria, fungi, viruses and animals. Indeed transgenic tomatoes have been produced with sequences from viruses (virus coat protein-mediated resistance to TMV; Nelson *et al.*, 1988), bacteria (herbicide resistance to phosphinothricin from *Salmonella typhimurium*; DeBlock *et al.*, 1987), antisense sequences of existing tomato genes, specifically designed to downregulate the expression of an endogenous gene (downregulation of the fruit-ripening enzyme polygalacturonase; Smith *et al.*, 1988) and fish (improved cold tolerance with a gene encoding antifreeze proteins; Hightower *et al.*, 1991). Hence, the tertiary gene pool for any transformable plant species can now be regarded as taking in all life forms.

9.2. Target Traits

Conventional plant breeders have, depending on species of interest, targeted a wide range of traits for improvement, including yield, resistance or tolerance to biotic and abiotic stresses, adaptation to edaphic and climatic conditions, and product quality. With the possible exception of yield, there are examples of all other traits which have been manipulated using genetic engineering techniques. Hence it can be generally assumed that virtually all traits which are being manipulated via conventional breeding may be altered using molecular methods. However, with the wider gene pool accessible with genetic engineering, other traits not able to be manipulated via conventional means can be considered.

9.3. Genetic Engineering Technology

Genetic manipulation for biotechnological plant improvement relies on the ability to modify genes as DNA sequences. This requires utilization of standard, well-documented techniques. Enzymes are the basic tool kit of the genetic engineer, and can be used for basic functions of manipulating

nucleic acids in such a way as to copy (DNA polymerase), transcribe (RNA polymerase), cut at specific sites (restriction endonucleases), cut and degrade in a nonspecific nature (nucleases) and join pieces together (DNA ligase).

DNA libraries are another important resource for genetic manipulation as they represent a means of storing and screening sequences from an entire organism. The two basic types of library are the cDNA library and the genomic library. A cDNA (complementary DNA) library is based on expressed sequences. Such a library is made by extracting mRNA from a particular plant tissue of interest, then making first single-stranded DNA (via reverse transcriptase) then double-stranded DNA (via DNA polymerase). A cDNA library will then represent only those genes which are expressed in the tissue of interest (or in response to a stress of some sort), and will be devoid of introns and regulatory sequences.

A genomic library represents the entire genome (often including the organellar genomes), and will include all sequences whether they are expressed in the donor tissue or not, and even sequences which are not expressed at all. Genomic libraries will therefore include highly repeated sequences such as rRNA genes and nonexpressed simple sequence repeats (SSRs), as well as genes and their regulatory sequences and introns. Genomic libraries may be cloned and stored in lambda bacteriophage, cosmids or larger fragments which will often include multiple genes in bacterial artificial chromosomes (BACs; Shizuya *et al.*, 1992) and yeast artificial chromosomes (YACs; Burke *et al.*, 1987). BAC library inserts may average 150 kb, whereas YAC inserts are larger on average, at up to 1 Mb of DNA.

Overall, the ability to genetically engineer a plant relies on:

1. A genetic transformation system for transferring DNA into plant cells.
2. A plant regeneration system to produce plants from the transgenic cells.
3. A gene construct which can modify the plant phenotype in a desirable and controlled manner.

9.4. Genetic Transformation Techniques

A number of genetic transformation techniques have been developed since the first report of a transgenic plant. With few exceptions, these methods rely on the ability to transfer DNA to a single cell or small group of cells, such as nondifferentiated leaf tissue or meristems. Then some form of tissue culture is required to regenerate whole, fertile transgenic plants for the process to be completed. Therefore, the first requirement for plant transformation is a reliable plant regeneration system using explants/tissues which are amenable to the DNA delivery system.

This requirement was, more than any other factor, the greatest impediment to genetic engineering of members of the cereal and grain legume families – the two most important crop groups. Regeneration in most of these species is difficult, particularly for species like soybean (Ghazi *et al.*, 1986) and chickpea (Adkins *et al.*, 1995). At best, regeneration is extremely genotype-dependent for species such as rice, where the *japonica* rices are easier to manipulate *in vitro* (Chu and Croughan, 1990; Kunanuvatchaidach *et al.*, 1995) than the *indica* rices, which are more important on a worldwide basis. However, international effort in most of the important cereal species has largely overcome the problem, although there are still genotypes of a number of species which remain recalcitrant to regeneration.

The cereals and pulses (grain legumes) proved somewhat recalcitrant to genetic transformation until the late 1980s, when rice became the first cereal to become reliably transformed at high frequency (Toriyama *et al.*, 1988; Shimamoto *et al.*, 1989), with soybean first successfully transformed via *Agrobacterium* in 1988 (Hinchey *et al.*, 1988), although repeatable transformation was only achieved later using a microprojectile approach (Christou *et al.*, 1990).

There are three basic transformation systems being routinely used in most genetic engineering laboratories internationally. These are:

1. *Agrobacterium*-mediated transformation.
2. DNA uptake via protoplasts.
3. Microprojectile bombardment.

9.4.1. *Agrobacterium*-mediated transformation

Agrobacterium tumefaciens, the causal agent of crown gall disease, transfers a portion of its own DNA (T-DNA) to the plant, where it is stably incorporated into the plant genome (Chilton *et al.*, 1977). The T-DNA is delimited by specific 25 bp repeat sequences, known as the border sequences (Thomashow *et al.*, 1980). None of the T-DNA is actually required for transfer, and whatever DNA is bound by the borders will be transferred (Garfinkel *et al.*, 1981), hence the pathogenic DNA can be effectively replaced with genes of interest, allowing delivery of these desirable genes to cells from which whole plants can be regenerated (Zambryski *et al.*, 1983). The ability of *Agrobacterium* to transfer genes to plants is known as virulence, and is largely controlled by the *vir* genes on the Ti plasmid, and to a lesser extent, the chromosomal virulence (*chv*) genes (Klee *et al.*, 1983). The *vir* genes are induced by the plant's wound response, in particular the production of phenolic compounds such as acetosyringone (Stachel *et al.*, 1986). The Vir proteins then facilitate the identification, copying, transfer and targeting to the host nucleus, and

incorporation on a plant chromosome of the T-DNA. Generally the simplest manner by which to replace the T-DNA is to remove it, leaving a 'disarmed' Ti plasmid with the *vir* genes, then introducing a binary plasmid which contains the gene(s) of interest within the border sequences (Bevan, 1984). Hence, with the induction of virulence, the Vir proteins will effect the transfer of the engineered T-DNA.

Most dicotyledonous species are hosts to *Agrobacterium* and as such can be genetically engineered in this manner. *Agrobacterium*-mediated transformation remains the method of choice for most species of the *Solanaceae* and *Brassicaceae*, and other top 20 most important plants such as grape, orange, cotton, sugarbeet and apple (Table 9.1). While there are a number of reports of successful transformation of important grain legumes including soybean (Hinchey *et al.*, 1988), chickpea (Fontanna *et al.*, 1993) and *Phaseolus* bean (Russell *et al.*, 1993), there is generally great genotype-dependence of such methods. It has been shown repeatedly that plant genotype $\times$ *Agrobacterium* strain interactions can vary from complete absence of transformation to 100% transformation within many species, including soybean (Owens and Smigocki, 1988), pea (Puonti-Kaerlas *et al.*, 1989) and sugarbeet (Godwin *et al.*, 1992).

Most important monocotyledonous species were once regarded as being outside the host range of *Agrobacterium* (DeCleene and DeLey, 1976). In the late 1980s it became evident that some monocot species, including asparagus (Bytebier *et al.*, 1987), yam (Schafer *et al.*, 1987) and onion (Dommissie *et al.*, 1990), were transformable using *Agrobacterium*. Early reports of *Agrobacterium*-mediated transformation of rice (Raineri *et al.*, 1990), maize (Gould *et al.*, 1991) and sorghum (Godwin and Chikwamba, 1993) remained largely unrepeatable in these and other labs, until the first report of a high-frequency, repeatable, and painstakingly substantiated system was published for three *japonica* rice cultivars (Hiei *et al.*, 1994).

Chan *et al.*, (1993) first reported stable transformation of rice using *A. tumefaciens*. Progeny analysis and Southern hybridization demonstrated that it was possible to generate transgenic rice via *Agrobacterium*, albeit at low frequency. Hiei *et al.* (1994) improved the efficiency greatly by co-cultivating callus originating from scutella of immature embryos with a super binary vector then selecting on hygromycin, and produced over 400 primary transgenics of three cultivars. Detailed analyses of 29 plants and their progeny revealed that stable transformation was possible, and in most cases marker genes segregated in the expected Mendelian fashion. In addition, these authors sequenced T-DNA / plant DNA junctions and found that incorporation occurred at similar sites to those found in transgenic tobacco plants. The 'super binary' plasmid contained *virB* and *virG* genes from the hypervirulent pTiBo542 plasmid (Hood *et al.*, 1986), and, although it enhanced transformation frequencies, was not an absolute requirement for transformation.

Table 9.1. Transformation methods used to produce fertile transgenic whole plants of the top 20 crops worldwide, ranked on total weight of produce recorded for 1993. Source of rankings was FAO (1994). Where more than one method is currently used, or the first successful method has been superseded, all methods are given in chronological order of publication. Reports of methods other than electroporation, microprojectiles or *Agrobacterium* are not included.

Species (commodity)	Transformation method	Reference
1. Maize	Electroporation of protoplasts	Shillito <i>et al.</i> (1989)
	Microprojectile bombardment	Gordon-Kamm <i>et al.</i> (1990)
2. Rice	Electroporation of protoplasts	Shimamoto <i>et al.</i> (1989)
	Microprojectile bombardment	Christou <i>et al.</i> (1991)
	<i>Agrobacterium</i> -mediated	Hiei <i>et al.</i> (1994)
3. Wheat	Microprojectile bombardment	Vasil <i>et al.</i> (1992)
4. Potato	<i>Agrobacterium</i> -mediated	DeBlock (1988)
5. Barley	Microprojectile bombardment	Wan and Lemaux (1994)
6. Cassava	Not yet reported	
7. Soybean (seed)	<i>Agrobacterium</i> -mediated	Hinchee <i>et al.</i> (1988)
	microprojectile bombardment	Sato <i>et al.</i> (1993)
8. Sweet potato	Microprojectile bombardment	Prakash and Varadarajan (1992)
9. Sugar-cane (raw sugar)	Microprojectile bombardment	Bower and Birch (1992)
10. Banana/plantain	Microprojectile bombardment	Sagi <i>et al.</i> (1995)
	<i>Agrobacterium</i> -mediated	May <i>et al.</i> (1995)
11. Tomato	<i>Agrobacterium</i> -mediated	McCormick <i>et al.</i> (1986)
12. Sorghum	Microprojectile bombardment	Casas <i>et al.</i> (1993)
13. Orange	<i>Agrobacterium</i> -mediated	Moore <i>et al.</i> (1992)
14. Grape	<i>Agrobacterium</i> -mediated	Mullins <i>et al.</i> (1990)
15. Cotton (seed + lint)	<i>Agrobacterium</i> -mediated	Umbeck <i>et al.</i> (1987)
16. Sugarbeet (raw sugar)	<i>Agrobacterium</i> -mediated	Lindsay and Gallois (1990)
17. Apple	<i>Agrobacterium</i> -mediated	James <i>et al.</i> (1993)
18. Coconut	Not yet reported	
19. Oat	Microprojectile bombardment	Somers <i>et al.</i> (1992)
20. Rapeseed/canola	<i>Agrobacterium</i> -mediated	Pua <i>et al.</i> (1987)

Basic requirements for *Agrobacterium*-mediated transformation are host recognition, virulence gene induction, the ability to select and proliferate transgenic cells, and the ability to regenerate whole, fertile plants from these cells. Host recognition and virulence gene induction require the correct carbon source and pH, and the presence of virulence-inducing molecules (Godwin *et al.*, 1992; see Fig. 9.1). Many wounded plant tissues produce signal molecules such as acetosyringone (Stachel *et al.*, 1986), which set in train the events of T-DNA transfer and integration into the host nuclear genome (for details of the process, see Zambryski, 1992).

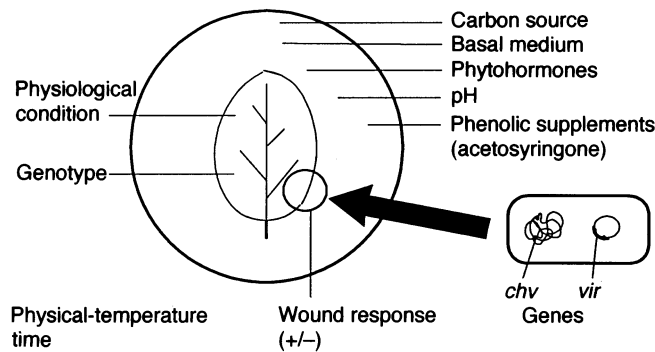


Fig. 9.1. A summary of the plant, co-cultivation medium, bacterial, and physical factors which affect the efficiency of virulence induction and *Agrobacterium*-mediated gene transfer (after Godwin *et al.*, 1992).

9.4.2. Protoplasts

Direct gene transfer has been the method of choice for monocotyledons, predominantly because most of these species have proved recalcitrant to *Agrobacterium*-mediated transformation. Protoplasts are cells with the cell wall enzymatically removed. In the absence of a cell wall, macromolecules such as proteins and nucleic acids may be taken up by such cells. The efficiency of uptake of DNA is greatly enhanced by the application of physiological shock, such as an ionic, osmotic or electrical current, which causes reversible permeabilization of the cell membrane. DNA has been shown to transfer to cells of most major cereals using the electric shock, or electroporation, methodology. If the DNA enters the nucleus and is stably incorporated into the plant genome, the cell is transgenic. This has been achieved with all major cereals (e.g. sorghum; Battraw and Hall, 1991). One of the advantages of this methodology is that a large number of theoretically competent cells can be exposed to DNA at one time.

Regeneration of whole, fertile plants from cereal protoplasts has remained the greatest hurdle to electroporation-mediated transformation. Even when protoplasts are regenerable and DNA transfer is effected, it is difficult to combine the two. The physiological shock necessary for DNA uptake may often be enough to prevent subsequent cell wall synthesis or substantially reduce regenerability. As protoplast cultures usually require greater time *in vitro*, there may also be an increased frequency of somaclonal variation (Larkin and Scowcroft, 1981). Success has been achieved with regeneration of rice, wheat and maize but protoplasts remain a widely used method for rice only, with microprojectile-mediated transformation overtaking it in efficiency.

9.4.3. Microprojectile bombardment

The potential for genetic transformation with microprojectiles was first reported by Klein *et al.* (1987), who achieved strong transient expression of a reporter gene in intact onion tissues. The first report of fertile transgenic plants and their progeny described stably transformed tobacco plants with Mendelian inheritance among selfed progenies. However, much of the focus with microprojectile bombardment has concentrated on cereals, due to the apparent inability of *Agrobacterium* to transfer genes to most cereal species, and the recalcitrance and genotype-specificity of cereals (with the exception of rice) to regeneration of fertile plants from transformed protoplasts.

The microprojectile bombardment, or 'Biolistic' (commercial description used by Bio Rad Labs), approach is a simple concept, based on the introduction of DNA-coated small inert projectiles into plant cells, not unlike the action of a microscale shotgun. The system is a modification of one used by virologists, whereby high-velocity microprojectiles are used to facilitate viral infection (MacKenzie *et al.*, 1966). To adapt this for transformation, the major alterations required were to reduce microprojectile size to reduce the extent of cell damage, to be able to perform the procedure under sterile conditions to exclude contamination by bacteria or fungi, and most importantly, to ensure that the target tissue be regenerable to produce whole, fertile plants. Here lies one of the greatest advantages of the microprojectile approach over protoplast and *Agrobacterium*-mediated techniques. Effectively, any target tissue may be used for bombardment, so the most regenerable tissues can be used. In addition, there does not appear to be any genotype specificity in the ability to deliver DNA, which is demonstrably the case with *Agrobacterium*-mediated transformation (e.g. Owens and Smigocki, 1988). The genotype dependence of plant regeneration remains the greatest hurdle in most species, although this is by no means the barrier that protoplast regeneration presents. The microprojectiles, which are commonly an inert metal such as tungsten or gold, range in diameter from 1 to 4 μm ; a velocity of approximately 250 m s^{-1} is required to penetrate cell walls and membranes (Franks and Birch, 1991). For successful stable transformation, the DNA must be delivered intact to the nucleus where the DNA will then be incorporated onto a chromosome. This requires that the velocity of microprojectiles be sufficient to penetrate cells, but not cause excessive damage.

For most of the cereals, the target tissue of choice is the immature zygotic embryo, particularly the scutellum, or embryogenic callus cultures, usually derived from zygotic embryos. Other targets such as embryogenic suspension cultures have also been used. As can be seen (Table 9.1), the most important cereal species have been recently transformed using this

approach. In addition, transgenic oats (Somers *et al.*, 1992) and rye (Castillo *et al.*, 1994) have been produced with microprojectiles.

Early microprojectile apparatus was based on a gunpowder charge for particle acceleration. Improvements in safety and reliability were made by using electrical discharge or helium gas to provide acceleration, and most laboratories currently use helium-based guns, such as the commercially available Biolistic gun (Russell-Kikkert, 1993). Tissue sterility is maintained by performing all operations in a laminar flow cabinet with bombardment actually taking place in a sealed chamber under a slight vacuum. One of the drawbacks of the microprojectile approach is the cost of the Biolistic gun, hence other laboratories have developed simpler apparatus, such as the particle inflow gun (PIG), which relies on a solenoid to regulate helium gas inflow (Finer *et al.*, 1992; Vain *et al.*, 1993). The great advantage of the PIG is that it can be built for around US\$1000. There are potential questions about the use of such apparatus and whether there is any contravention of international patents.

9.4.4. Alternative transformation methods

While the great majority of transgenic plants worldwide are being produced via *Agrobacterium*, microprojectile bombardment and, to a lesser extent, protoplasts, a number of alternative transformation methods have been trialled. Although not intended to be an inclusive list, these include:

1. Virus-mediated transformation.
2. Pollen pathway or intact inflorescence methods.
3. Microinjection.
4. Silicon carbide fibres.
5. Electroporation of intact tissues.
6. Electrophoresis of intact tissues.

While it is not possible to discuss the success, advantages and disadvantages of these methods, the subject has been recently reviewed by Songstad *et al.* (1995).

The pollen pathway has long been recognized as a potentially useful means of introducing genes to the germline, with its greatest advantage being the avoidance of tissue culture and the need to regenerate undifferentiated tissues such as callus or protoplasts (Hess, 1975, 1987). The absence of any time in plant tissue culture should also totally avoid the problem of somaclonal variation. There have been many difficulties with this approach, however, and there are no substantiated instances of genetic transformation via the pollen pathway.

More promising methods are the use of silicon carbide fibres and electroporation / electrophoresis of intact tissues. Stable tissue transformation

of tobacco and corn (maize) suspension cultures has been achieved following agitation with silicon carbide fibres (Kaeppeler *et al.*, 1992). Stable cereal transformants have been produced by electroporating germinating rice seeds (Li *et al.*, 1991) and immature embryos of corn (D'Halluin *et al.*, 1992), and by electrophoresis of germinating barley seeds (Ahokas, 1989).

However, until some of these alternative systems can be made more reliable and efficient, it would appear that the two most widely used transformation systems will remain *Agrobacterium* and, in the case of cereals, microprojectile bombardment.

9.5. Cloning of Plant Traits

Plant gene cloning is currently limited to simply inherited single gene traits. The transfer of genes to plants is limited to single genes or, at most, two to three genes encoding a particular, well-characterized biochemical pathway. Manipulation of multigenic traits is currently only feasible within the existing gene pool, sometimes made more efficient by marker-assisted selection, as outlined in Chapter 8 of this volume.

When traits are regarded in the molecular genetic sense, often the coding sequence is the major consideration. Obviously, the coding sequence is of major importance as this is the template for determining the amino acid sequence of the final protein or enzyme which determines the phenotype. However, a major factor in controlling the phenotype is the presence of regulatory sequences surrounding the coding sequence in the form of upstream (promoters and enhancers) and downstream (terminator) regulatory sequences and, in many eukaryotic genes, introns. When traits are being manipulated via genetic engineering, the role of these regulatory sequences cannot be trivialized. Put simply, these sequences control such things as the level of gene expression, tissue / organ specificity, developmental expression and response to particular abiotic / biotic stimuli.

The regulatory sequences that have been subjected to greatest attention in plants are the upstream promoters. For proper expression of a seed storage protein, for example, the gene of interest must be under the control of a promoter which expresses in developing endosperm (monocots) or cotyledons (dicots). Expression of other genes may be required only in anthers, root hairs or photosynthetic tissues, or in response to a wound such as that caused by a chewing insect or penetration of a fungal infection structure.

In the earliest cases of transgenic plants, introduced genes were placed under the control of constitutive promoters of bacterial (octopine or nopaline synthetase) or viral (cauliflower mosaic virus (CaMV) 35S) origin. These promoters have been found to give adequate expression of selectable marker and reporter genes in dicotyledonous species. However,

expression is commonly not of a sufficient level in monocots, especially cereals, for good selection with antibiotic or herbicide resistance genes (Chamberlain *et al.*, 1994). And, as already discussed, such a constitutive pattern of gene expression is commonly not desirable for most traits. For cereals, a number of other promoters have been developed to achieve high-level constitutive expression of selectable markers or reporter genes. These include sequences cloned directly from cereal tissues such as the ubiquitin (Christensen *et al.*, 1992) and rice actin 1 promoters (McElroy *et al.*, 1991) and synthetic promoters involving various spliced sequences from a range of sources such as the Emu promoter (Last *et al.*, 1991). However, to properly control gene expression such that the introduced character is most efficiently targeted without other pleiotropic effects, it is generally found that regulatory sequences of plant origin are desirable. Further to this, expression is usually best controlled when a cereal promoter is used in a cereal or a dicot promoter in a dicot. The most widely used means of cloning such regulatory sequences is by using cDNA libraries from specific tissue types and looking for abundant signals. Naturally, cDNA libraries will not contain the regulatory sequences, hence the abundant cDNA clones must then be used to screen a genomic library so that the regulatory sequences can be identified and cloned.

9.6. Methods for Cloning Plant Traits

9.6.1. Cloning based on knowledge of protein structure/sequence

Many of the first plant traits to be cloned were those in which some degree of biochemical understanding preceded any work with DNA. The understanding may have come from plant systems or from outside the plant kingdom. Most of the cloned genes in these cases were enzymes which were part of a well-characterized biochemical pathway or had a protein structure which was otherwise well known, such as seed storage proteins.

Cloning genes based on a knowledge of the protein sequence can be achieved by working backwards from the gene product (amino acid sequence) to the coding sequence or the DNA template. Such an approach has been used for cloning many seed storage proteins, such as the 2S sulphur-rich proteins from Brazil nut (Altenbach *et al.*, 1992).

Genes of plant origin have been used in transgenic plants to improve resistance to insect pests. One of the first to be characterized was the cowpea trypsin inhibitor, a protease inhibitor which acts as an antifeedant to a range of insect pests. The protein is naturally produced in seeds of some genotypes of cowpea (*Vigna unguiculata*), which was shown to confer lepidopteran insect resistance when expressed constitutively in tobacco (Hilder *et al.*, 1987). Another such report from the same group at Durham

University described the use of a snowdrop lectin gene to confer aphid resistance in transgenic tobacco (Hilder *et al.*, 1995), the first report of genetically engineered resistance to a sucking insect in plants. The lectin from snowdrop (*Galanthus nivalis*), known as GNA, has been demonstrated to be effective against two major rice pests, the brown planthopper and the green leafhopper (Powell *et al.*, 1993), as well as the peach potato aphid. When expressed at 0.1% of total leaf protein in tobacco, GNA is antimetabolic to homopteran pests. These levels of protein were accumulated in leaf tissues when placed under the control of the CaMV 35S promoter. However, Hilder *et al.* (1995) postulate that better control would be achieved with the use of a phloem-specific promoter, such as the sucrose synthase 1 promoter from rice (Wang *et al.*, 1992).

Another significant group of genes which have been cloned based on prior knowledge of the end product are the seed storage genes, including both seed storage proteins and lipids. In both cases there was considerable biochemical knowledge of the pathways involved in biosynthesis and storage. For seed storage proteins, the genes can be cloned based on the knowledge of the amino acid sequence of the mature protein and targeting signal or transit peptide.

Molecular genetic manipulation of vegetable oils is a major area of research in the developed world. While approximately 90% of vegetable oils are used for human consumption (Röbbelen, 1988), there are emerging markets for industrial applications including lubricants, paints, cosmetics, plasticizers and soaps (Thierfelder *et al.*, 1992). With increasing emphasis on quality and dietary considerations, there is a desire in developed countries to substantially alter the types and balances of fatty acids found in the major oil crops. Six crop species (soybean, oil palm, rapeseed / canola, sunflower, cottonseed and peanut) account for 84% of the world's vegetable oil production, with the major fatty acids produced being palmitic, stearic, linoleic, linolenic and oleic acids. However, almost 200 different fatty acids are produced by plants (Kishore and Sommerville, 1993), mostly occurring in nondomesticated plant species. Hence, genetic manipulation of the major crop species, either via conventional breeding or genetic engineering, has the potential to modify the fatty acid types produced and stored by seeds such as soybean and canola.

The biochemistry of fatty acid biosynthesis is well established for many of the major pathways, with enzymes identified which control the chain length of fatty acids, and the degree of saturation, or number of double bonds in the fatty acid chain. The oils stored in seeds have been altered by downregulation of particular enzymes using antisense RNA techniques. The oleic acid content of rapeseed oil has been increased to 83% (from 62% in conventionally bred cultivars) by downregulating the enzyme $\Delta 12$ -oleate desaturase using antisense RNA (Yadav *et al.*, 1993). A similar result has been achieved using co-suppression, which has

produced rapeseed oil with 87% oleic acid content (cited in Töpfer *et al.*, 1995).

Genetic engineering approaches have also been used to introduce totally novel fatty acids to major oilseed crops. Petroselinic acid, an isomer of oleic acid, is found in abundance in the spice plant, coriander (*Coriandrum sativum*). This oil has uses in cosmetics and pharmaceuticals, and its oxidation leads to the formation of lauric acid, used in detergents, and adipic acid, which is used in nylon production. Hence this is a potentially valuable oil which may enable soybean and canola crops – i.e. renewable resources – to replace some of the nonrenewable carbon fuels which are currently used for many industrial polymer and lubrication applications (Gunstone *et al.*, 1994). A cDNA clone from coriander encoding an acyl-ACP desaturase has been genetically engineered into tobacco, resulting in the accumulation of petroselinic acid (Cahoon *et al.*, 1992). Oil crops such as rapeseed and soybean could be genetically modified to produce this fatty acid on a commercial scale. Hence, here is an example whereby a minor crop plant has become a donor of genetic diversity which will enable the commercial production of a very useful array of products for food and industrial purposes.

9.6.2. Cloning genes known only by phenotype

However, a more challenging task has been to move beyond proteins and pathways about which sufficient biochemistry is known into cloning genes known by phenotype only. By far the majority of genes fit into this category, which includes many of the most important plant traits such as disease resistance, adaptation to abiotic stresses, phenology and flowering behaviour and fertility.

Only in the past few years have genes for traits known only by phenotype been cloned. One of the most interesting areas of plant biology is the understanding of plant–pathogen interactions at the molecular level.

The major methods used for cloning genes known only by phenotype are:

1. Transposon tagging.
2. Map-based or positional cloning.
3. T-DNA tagging.

There have been a number of plant genes cloned using all these methods or modifications of them. There are also examples where a combination of methods has been used to expedite the rate of identification of the unknown sequence.

Transposon tagging

The proposal that specific pieces of DNA are able to autonomously excise from a particular locus and insert, theoretically at random, into another linked or nonlinked locus was first proposed by Barbara McClintock (1949, 1950). Transposable elements are now accepted as ubiquitous (Walbot, 1992), and are particularly well characterized in organisms such as *Escherichia coli*, *Drosophila* and maize.

In two plant species, maize and the snapdragon (*Antirrhinum majus*), transposable elements are sufficiently well characterized to be used for mutagenesis and gene tagging. The best understood system is the maize *Ac/Ds* (Activator/Dissociation) family of elements (Table 9.2). The behaviour of the *Ac* element is such that it can excise autonomously from its locus, under the control of its own transposase, and insert into another chromosomal position. Where the *Ac* element is in or near a gene, that gene is inactivated (i.e. not expressed), and will appear phenotypically as a recessive mutation. When *Ac* transposes (excises and reinserts at another locus), the gene is commonly returned to normal function, and hence is termed a revertant. Should this *Ac* element then insert into another gene (transcribed or regulatory sequence), then a new recessive mutation will manifest. Transposition can occur either germinally, which results in an altered phenotype for the whole plant, or somatically, which results in a mosaic mutation, and can often be detected as a sectoring of plant or seed colour (Walbot, 1992). The *Ds* element is actually a group of elements derived from defective *Ac* elements, with a common feature being their lack of effective *Ac* transposase. Hence these elements are nonautonomous, and will only transpose in the presence of an *Ac* element capable of producing transposase.

Equipped with sequence and map information of an *Ac* element in maize, geneticists have been able to clone important maize genes using the transposon-tagging approach. When a gene is inactivated by an *Ac* element, proof of which can be demonstrated by a low frequency of trait reversion, the *Ac* element can be used as a probe for 'fishing-out' the tagged gene. A tagged gene can be cloned by performing Southern blot analysis with the *Ac* element used as probe, and any polymorphism represents an excision / insertion event. The technique is described more fully in the review by Walbot (1992).

Table 9.2. Plant transposable elements which have been used for transposon-tagging.

Element	Species	Type of element	Size (kb)	Reference
<i>Ac</i>	Maize	Autonomous	4.6	Federoff <i>et al.</i> (1983)
<i>Ds</i>	Maize	Nonautonomous	0.5–30	Federoff <i>et al.</i> (1983)
<i>Tam1</i>	Snapdragon	Autonomous	15	Bonas <i>et al.</i> (1984)
<i>Tam2</i>	Snapdragon	Nonautonomous	5	Bonas <i>et al.</i> (1984)
<i>En/Spm</i>	Maize	Autonomous	8.3	Gierl <i>et al.</i> (1985)

The *Ac* system has proven extremely useful for cloning maize genes for traits such as plant pigment genes (*bronze*, Federoff *et al.*, 1984; the anthocyanin pathway, Paz-Ares *et al.*, 1986), starch biosynthesis (*opaque2*, Motto *et al.*, 1988), and abscisic acid (AbA) insensitivity (McCarty *et al.*, 1989). Other maize transposable elements such as *Mutator* (*Mu*, Robertson, 1978) and the *En/Spm* system (O'Reilly *et al.*, 1985) have been used to clone maize genes, as has the *Tam* system in snapdragon (Martin *et al.*, 1985). As a result, the search began for transposable element systems in other species so that a similar approach could be taken.

However, it was soon demonstrated that the well-characterized *Ac/Ds* system could be genetically engineered into other plant species including tobacco (Baker *et al.*, 1986), *Arabidopsis* (Van Sluys *et al.*, 1987), tomato (Yoder *et al.*, 1988) and petunia (Haring *et al.*, 1989). It was subsequently shown that the *Ac* element acted in a similar manner in these heterologous species, undergoing germinal and somatic excision and insertion, and maintaining the propensity to reinsert at a linked site (usually within 4 cM; Wessler, 1988). As a result, the frequency of mutation in the surrounding area approaches 10^{-4} (Walbot, 1992). It was not until 1993 that a plant gene was cloned using the heterologous system, that of the *Ph6* gene in petunia, which is involved in floral pigmentation (Chuck *et al.*, 1993).

Using the transposon-tagging approach, a number of plant disease resistance genes (R genes) have now been cloned. The first of these was the tomato *Cf-9* gene, which confers resistance to race 9 of the fungal pathogen *Cladosporium fulvum*, causal agent of leaf mould. The gene was tagged using a modified *Ac/Ds* system (Jones *et al.*, 1994). The system relied on the use of a stabilized *Ac* element (sAc), which was effectively a disabled *Ac* element that nevertheless produced *Ac* transposase (Scofield *et al.*, 1992) and hence enables *Ds* transposition. The two elements are maintained in separate lines, and hence are both stable until controlled crossing is performed. This gives a significant advantage in that *Ds* insertion mutants can be maintained indefinitely in the homozygous or hemizygous form, and revertants can be isolated by crossing lines with sAc tomatoes.

The transposon-tagging approach used to isolate the *Cf-9* gene, which was known by phenotype only, did have considerable advantages in that much was known of the pathogen. The hypersensitive response elicited by the virulent strain fitted into the gene-for-gene hypersensitive response proposed by Flor (1971), with a dominant avirulence gene (*Avr9*) interacting with the *Cf-9* gene to form an incompatible or hypersensitive response. *Avr9* had previously been cloned and fully characterized (Kan *et al.*, 1991), and was known to specify a 28 amino acid peptide which alone could elicit a necrotic response in resistant tomatoes.

Crosses were made to develop a transgenic tomato line homozygous for *Cf-9*, but heterozygous for *Ds*. Crossing was performed as outlined in Fig. 9.2. As can be seen, most seedlings died from systemic hypersensitive

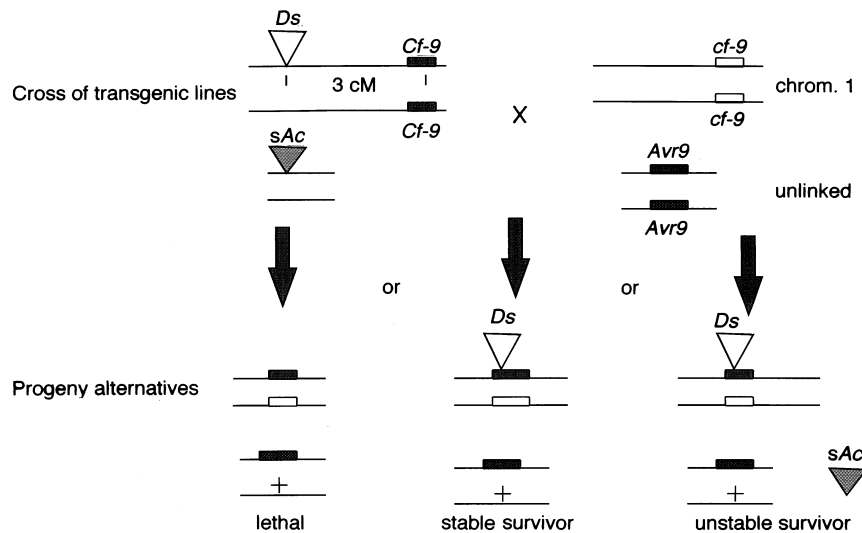


Fig. 9.2. Graphical representation of the crossing strategy of transgenic *Ac/Ds* tomato lines used by Jones *et al.* (1994) to transposon tag the *Cf-9* gene of tomato.

response, with *Avr9* and *Cf-9* present in all cells. The only survivors would be individuals where *Ds* had inserted into the *Cf-9* gene. Where this happened in plants carrying *sAc*, the seedlings were variegated for the hypersensitive response (somatic excision), and in the absence of *sAc*, the survivors were stable (germinal excision).

This work was a major undertaking, involving the germination of approximately 160 000 progeny, yielding 118 survivors, representing 63 independent mutations. After considerable effort to characterize these, the *Cf-9* gene was cloned and characterized, with an open reading frame encoding an 863 amino acid protein.

Map-based cloning

The previously described transposon-tagging scheme utilized to clone *Cf-9* had an element of the map-based approach by utilizing the phenomenon of *Ac/Ds* transposition to linked sites, with a *Ds* element within 3 cM of the locus of interest, which greatly enhanced the likelihood of insertional mutagenesis. However, map-based cloning requires a much tighter genetic linkage (in centimorgans or recombination fraction), which in turn needs to be converted into a physical distance (kbp of DNA sequence).

The map-based, or positional cloning concept relies on building a well-saturated genetic map of DNA sequences of the organism in question. The genetic map is built up in a polymorphic mapping population, commonly an F_2 , backcross (BC) or recombinant inbred, with markers such as

RFLPs (restriction fragment length polymorphisms). The first RFLP map to be made was of human (Botstein *et al.*, 1980). Based on such maps, important human genes responsible for disorders such as cystic fibrosis (Rommens *et al.*, 1989) and muscular dystrophy (Kenwrick *et al.*, 1987) have been cloned in this manner.

Genetic maps of plant species have been made with DNA markers in virtually all important plant species, particularly the major cereals, grain legumes, oilseeds and important vegetable crops including tomato and potato. The RFLP marker system is based on Southern blot analysis (Southern, 1975), while many of the newer technologies use the polymerase chain reaction (PCR) (Ehrlich *et al.*, 1988). These include random amplified polymorphic DNA (RAPDs) (Williams *et al.*, 1990), amplified fragment length polymorphisms (AFLPs) (Zabeau and Vos, 1994; Vos *et al.*, 1995), and a number of different marker systems based on simple sequence repeats (SSRs), which may use PCR, such as inter simple sequence repeat (ISSR)-PCR (Zietkiewicz *et al.*, 1994; Salimath *et al.*, 1995), or Southern analysis (Akkaya *et al.*, 1992). The applications of DNA markers to plant genetic analysis and breeding are many, and the subject has been reviewed by various authors (see Tanksley *et al.*, 1989).

The population for saturation mapping should be segregating in a Mendelian manner for the trait of interest. This requires a lot of tedious but demanding work, and it may be that particular populations such as near isogenic lines, or DNA pooling strategies, such as bulked segregant analysis (Michelmore *et al.*, 1991), will allow the mapping effort to be more targeted to the region of interest, hence improving the rate of map saturation. It is desirable to obtain extremely close genetic linkage to the locus of interest (<1 cM), with flanking markers if possible. It can be of considerable advantage to identify a marker which co-segregates (say within 0.1 cM)

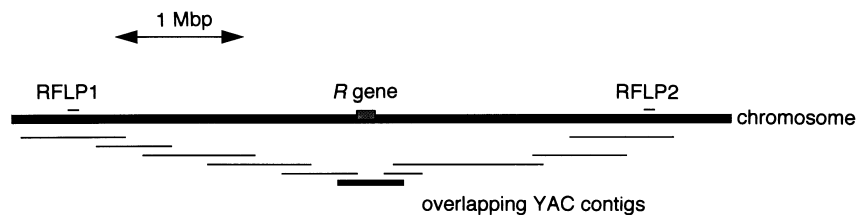


Fig. 9.3. Generalized strategy of chromosome walking used in map-based gene cloning. The gene of interest (in this illustration, *R*) is flanked by markers such as RFLPs or SSRs. A series of overlapping YAC contigs is used to cover the chromosomal region containing the gene. The YAC which contains the complete gene is marked in bold. This is then subcloned or matched to a cDNA library to identify the gene(s) contained in the region, and ultimately identify the *R* gene.

with the trait. The next step is to convert the genetic map into a physical map of the chromosome region, translating distances from centimorgans to nucleotides.

Physical mapping then involves the use of YAC or BAC clones which span the region of interest. These YAC/BAC contigs will span the region between two flanking DNA marker sequences as shown in Fig. 9.3, a process known as chromosome walking. In this example, two YAC/BAC clones contain the gene of interest. Methodology is being developed to enable BAC clones to be directly transformed into plants, to find which one contains the gene of interest (in the example, the *R* gene). Alternatively, the YAC/BAC must be subcloned (for example as a cDNA library), and these subclones can be put into the susceptible plant and tested for the *R* phenotype.

The first *R* gene to be cloned in plants via map-based cloning was the *Pto* gene, which confers resistance to *Pseudomonas syringae* pv. *lycopersici*, the causal agent of bacterial speck in tomato. The starting points for the work were:

1. A high-density RFLP map of tomato (Tanksley *et al.*, 1992).
2. A tomato YAC library (Martin *et al.*, 1992).

The *Pto* locus had been previously mapped to tomato chromosome 5 (Martin *et al.*, 1991) and one RFLP marker co-segregated with *Pto*. Hence this was a useful starting point to screen the YAC library. A 400 kb YAC was identified, and via inverse PCR, some end-specific probes were generated. The left arm of the YAC was 1.8 cM from *Pto*, with the right arm perfectly co-segregating. This YAC was then used to screen a 920 000 plaque leaf cDNA library, from which 30 were selected. One of these was shown to co-segregate with *Pto*. This cDNA was then transformed into a susceptible tomato cultivar under the control of the CaMV 35S promoter, where it conferred resistance. Hence the gene was identified, and could then be sequenced and characterized (Martin *et al.*, 1993a,b).

The success of this approach was partly due to the small genome size of tomato (950 Mbp), and a relatively low level of repeat sequence. Such an approach is much more difficult in species with larger genomes such as maize (2300–2700 Mbp) and wheat (16 000 Mbp) (Arumaganathan and Earle, 1991). One such example of this problem is the 79 000 clone maize YAC library produced by Zeneca Seeds. As reported by K. Edwards (cited by Young and Phillips, 1994), only 15% of the YACs have unique ends, with the remainder terminating in repeat sequences, which makes chromosome walking very difficult. The synteny of genomes is a major advantage in this instance. It is now well established that there is a high degree of synteny of genome organization, as seen in the co-linearity of maps of the most important grass genomes (Bennetzen and Freeling, 1993). This has resulted in the great interest in using this co-linearity to facilitate

cloning of genes in large genomes such as maize and wheat, by starting with the smaller genomes, particularly rice, and to a lesser extent, sorghum. One particular study of maize YACs demonstrated that a single YAC with the *Adh1* sequence contained 36 different repetitive sequences (Springer *et al.*, 1994), whereas these repetitive sequences were largely absent in corresponding areas of the rice and sorghum genomes, a pattern seen also for the *a1-sh2* region in all three genomes (J. Bennetzen, Purdue University, personal communication). In a similar manner, many genes of interest for *Brassica* improvement are being identified in the smaller *Arabidopsis thaliana* genome.

Paterson *et al.* (1995) have recently demonstrated the potential of this approach, by identifying quantitative trait loci (QTLs) for seed mass and phenology which were co-linear on molecular marker maps of maize, sorghum and rice. They proposed that, as the maize genome is four and six times larger than the sorghum and rice genomes, respectively, the genes of interest in maize may be cloned in sorghum /rice more expeditiously. It must be remembered that QTLs are subject to genotype $\times$ environment interaction, which will be a major complicating factor in attempts to clone and characterize a QTL (Jansen *et al.*, 1995).

T-DNA tagging

When T-DNA is inserted into the plant genome, it does so via illegitimate recombination (Mayerhofer *et al.*, 1991). In the same way as a transposon, the insertion of T-DNA into a gene will inactivate it, which phenotypically will appear as a mutation. The comparative advantage of such a method is that stable mutations will be conferred. This can be a disadvantage, however, as reversions, such as those made possible by the *Ac/Ds* system are not possible. It is also the case that each independent transformation event will result in just one mutation event (or more if multiple independent insertions are made), whereas the heterologous transposon approach allows multiple mutation events to take place from a single initial transformant. Hence, mutated genes with T-DNA insertions can be cloned, in much the same manner as those with transposon insertions, by using part of the T-DNA as a probe in Southern blots.

The procedure for selecting out the tagged gene is quite straightforward, particularly when the plant contains a single T-DNA copy. However, this means that a large number of independent transformants must be generated. This can be achieved with species such as *Arabidopsis*, where Feldmann (1991) demonstrated that a seed-based transformation system using *Agrobacterium* can be quite efficient. Feldmann (1991) demonstrated that by inoculating approximately 1000 seeds, approximately 300 000 selfed T2 seed can be collected. They estimated that approximately 8000 transformants could be recovered from an experiment of this scale and, using this approach, cloned a gene involved in trichome development (Marks and

Feldmann, 1989) and a dwarf gene (Feldmann *et al.*, 1989) from *Arabidopsis*. A regulatory gene in the ethylene biosynthesis pathway of *Arabidopsis* has also been cloned using the T-DNA tagging strategy (Kieber *et al.*, 1993).

With modification, T-DNA tagging methodology can be used to clone plant promoter sequences. Koncz *et al.* (1989) designed a vector which contained a promoterless antibiotic selection gene (aminoglycoside phosphotransferase II) and transformed this into *Nicotiana* and *Arabidopsis*. In this way, promoters could be 'trapped' as plants that were regenerated on selective antibiotic were antibiotic resistant because the T-DNA was inserted downstream of a promoter such that the gene was expressed. The most serious limitation to the T-DNA tagging approach is that large numbers of independent transformation events must be generated, which is not currently possible with most species. This is really only feasible with a system such as the seed transformation technique developed for *Arabidopsis*, or possibly the 'whole-leaf' transformation system developed for tobacco (Fisher and Gultinan, 1995).

9.7. Summary

Most important plant species are now transformable and, for those that are not, it is probably because little or no effort has gone into developing a transformation system. For many species, efficiency of transformation is very low and there will certainly be major improvements over the next decade. Gene cloning is theoretically possible from all plant species, although due to differences in genome size, reproductive behaviour and physical size, some species such as *Arabidopsis* and rice will give specific advantages to the speed of research. Now that methodologies have been developed which allow the cloning of genes known only by phenotype, the number of genes which become available is expanding rapidly. Methodologies will continue to become faster and less labour intensive, particularly as automation becomes more accessible. Your species of interest can now be considered as having all other plant species (and indeed all life) as its tertiary gene pool, which serves as a powerful illustration of the importance of conserving the earth's biodiversity.

References

- Adkins, A.L., Godwin, A.L. and Adkins, S.W. (1995) An efficient *in vitro* regeneration system of Australian-grown chickpea (*Cicer arietinum*) cultivars. *Australian Journal of Botany* 43, 491–497.
- Ahokas, H. (1989) Transfection of germinating barley seed electrophoretically with exogenous DNA. *Theoretical and Applied Genetics* 77, 469–472.

- Akkaya, M.S., Bhagwat, A.A. and Cregan, P.B. (1992) Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics* 132, 1131–1139.
- Alexander, L.J. (1963) Transfer of a dominant type of resistance to the known Ohio pathogenic strains of tobacco mosaic virus (TMV) from *Lycopersicon peruvianum* to *L. esculentum*. *Phytopathology* 53, 896.
- Altenbach, S.B., Kuo, C.C., Staraci, L.C., Pearson, K.W., Wainwright, C., Georgescu, A. and Townsend, J. (1992) Accumulation of a Brazil nut albumin in seeds of transgenic canola results in enhanced levels of seed protein methionine. *Plant Molecular Biology* 18, 235–245.
- Arumaganathan, K. and Earle, E.D. (1991) Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter* 9, 208–218.
- Baker, B., Schell, J., Lörz, H. and Federoff, N.V. (1986) Transposition of the maize controlling element *Activator* in tobacco. *Proceedings of the National Academy of Sciences of the USA* 83, 4844–4848.
- Battraw, M. and Hall, T.C. (1991) Stable transformation of *Sorghum bicolor* protoplasts with chimeric neomycin phosphotransferase II and β -glucuronidase genes. *Theoretical and Applied Genetics* 82, 161–168.
- Bennetzen, J.L. and Freeling, M. (1993) Grasses as a single genetic system: Genome composition, colinearity and compatibility. *Trends In Genetics* 9, 259–261.
- Bevan, M.W. (1984) Binary *Agrobacterium* vectors for plant transformation. *Nucleic Acids Research* 22, 8711–8721.
- Bohn, G.W. and Tucker, C.M. (1940) Studies on *Fusarium* wilt of tomato. I. Immunity in *Lycopersicon pimpinellifolium* Mill. and its inheritance in hybrids. *Missouri Agriculture Experiment Station Research Bulletin* 311, 1–82.
- Bonas, U., Sommer, H. and Saedler, H. (1984) The 17 kb *Tam1* element of *Antirrhinum* induces a 3 bp duplication upon integration into the chalcone synthase gene. *EMBO Journal* 3, 1015–1021.
- Botstein, D., White, R.L., Skolnick, M. and Davis, R.W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* 32, 314–331.
- Bower, R. and Birch, R.G. (1992) Transgenic sugarcane plants via microprojectile bombardment. *The Plant Journal* 2, 409–416.
- Burke, D.T., Carle, G.F. and Olson, M.V. (1987) Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* 230, 806–812.
- Bytebier, B., Deboeck, F., De Greve, H., Van Montagu, M. and Hernalsteens, J-P. (1987) T-DNA organization in tumour cultures and transgenic plants of the monocotyledon *Asparagus officinalis*. *Proceedings of the National Academy of Sciences of the USA* 84, 5345–5349.
- Cahoon, E.B., Shanklin, J. and Ohlrogge, J.B. (1992) Expression of a coriander desaturase results in petroselinic acid production in transgenic tobacco. *Proceedings of the National Academy of Sciences of the USA* 89, 11184–11188.
- Casas, A.M., Kononowicz, A.K., Zehr, U.B., Tomes, D.T., Axtell, J.D., Butler, L.G., Bressan, R.A. and Hasegawa, P.M. (1993) Transgenic sorghum plants via microprojectile bombardment. *Proceedings of the National Academy of Sciences of the USA* 90, 11212–11216.
- Castillo, A.M., Vasil, V. and Vasil, I.K. (1994) Rapid production of fertile transgenic plants of rye (*Secale cereale*). *Bio/Technology* 12, 1366–1371.

- Chamberlain, D.A., Brettel, R.I.S., Last, D.I., Witrzens, B., McElroy, D., Dolferus, R. and Dennis, E.S. (1994) The use of the Emu promoter with antibiotic and herbicide resistance genes for the selection of transgenic wheat callus and rice plants. *Australian Journal of Plant Physiology* 21, 95–112.
- Chan, M.T., Chang, H.H., Ho, S.L., Tong, W.F. and Yu, S.M. (1993) *Agrobacterium*-mediated production of transgenic rice plants expressing a chimeric α -amylase promoter/ β -glucuronidase gene. *Plant Molecular Biology* 22, 491–506.
- Chilton, M.-D., Drummond, M.H., Merlo, D.J., Sciaky, D., Montoya, A.L., Gordon, M.P. and Nester, E.W. (1977) Stable incorporation of plasmid DNA into higher plants: the molecular basis of crown gall tumorigenesis. *Cell* 11, 263–271.
- Christensen, A.H., Sharrock, R.A. and Quail, P.H. (1992) Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Molecular Biology* 18, 675–689.
- Christou, P., McCabe, D.E., Martinell, B.J. and Swain, W.F. (1990) Soybean genetic engineering – commercial production of transgenic plants. *Trends in Biotechnology* 8, 145.
- Christou, P., Ford, T.L. and Kofron, M. (1991) Production of transgenic rice (*Oryza sativa* L.) plants from agronomically important indica and japonica varieties via electric discharge particle acceleration of exogenous DNA into immature zygotic embryos. *Bio/Technology* 9, 957–962.
- Chu, Q.R. and Croughan, T.P. (1990) Genetics of plant regeneration in immature panicle culture of rice. *Crop Science* 30, 1194–1197.
- Chuck, G., Robbins, T., Nijjar, C., Ralston, E., Courtney-Gutterson, N. and Dooner, H.K. (1993) Tagging and cloning of a petunia flower colour gene with the maize transposable element *Activator*. *The Plant Cell* 5, 371–378.
- DeBlock, M. (1988) Genotype-independent leaf disc transformation of potato (*Solanum tuberosum*) using *Agrobacterium tumefaciens*. *Theoretical and Applied Genetics* 76, 767–774.
- DeBlock, M., Botterman, J., van de Wiele, M., Dockx, J., Thoen, C., Gossele, V., Rao Movre, N., Thompson, C., van Montagu, M. and Leemans, J. (1987) Engineering herbicide resistance in plants by expression of a detoxifying enzyme. *EMBO Journal* 6, 2513–2518.
- DeCleene, M. and DeLey, J. (1976) The host range of crown gall. *Botanical Reviews* 42, 389–466.
- D'Halluin, K., Bonne, E., Bossut, M., De Beuckeleer, M. and Leemans, J. (1992) Transgenic maize plants by tissue electroporation. *The Plant Cell* 4, 1495–1505.
- Domisse, E.M., Leung, D.W.M., Shaw, M.L. and Conner, A.J. (1990) Onion is a monocotyledonous host for *Agrobacterium*. *Plant Science* 69, 249–257.
- Driscoll, C.J. (1981) New approaches to wheat breeding. In: Evans, L.T. and Peacock, W.J. (eds) *Wheat Science Today and Tomorrow*. Cambridge University Press, Cambridge, pp. 97–106.
- Ehrlich, H.A., Gelfand, D.H. and Saiki, R.K. (1988) Specific DNA amplification. *Nature* 331, 461–462.
- FAO (1994) *Quarterly Bulletin of Statistics*, vol. 7. United Nations Food and Agriculture Organization, Rome.
- Federoff, N.V., Wessler, S. and Shure, S. (1983) Isolation of the transposable maize controlling elements Ac and Ds. *Cell* 35, 235–240.

- Federoff, N.V., Furtek, D.B. and Nelson, O.E. (1984) Cloning of the *bronze* locus in maize by a simple and generalizable procedure using the transposable element Activator (Ac). *Proceedings of the National Academy of Sciences of the USA* 81, 3825–3829.
- Feldmann, K.A. (1991) T-DNA insertion mutagenesis in *Arabidopsis*: Mutational spectrum. *The Plant Journal* 1, 71–82.
- Feldmann, K.A., Marks, M.D., Christianson, M.L. and Quatrano, R.S. (1989) A dwarf mutant of *Arabidopsis* generated by T-DNA tagging. *Science* 243, 1351–1354.
- Finer, J.J., Vain, P., Jones, M.V.V. and McMullen, M.D. (1992) Development of the particle inflow gun for DNA delivery to plant cells. *Plant Cell Reports* 11, 323–328.
- Fisher, D.K. and Guiltinan, M.J. (1995) Rapid, efficient production of homozygous transgenic tobacco plants with *Agrobacterium tumefaciens*: a seed-to-seed protocol. *Plant Molecular Biology Reporter* 13, 278–289.
- Flor, H.H. (1971) Current status of the gene-for-gene hypothesis. *Annual Review of Phytopathology* 9, 275–301.
- Fontanna, G.S., Santini, L., Caretto, S., Frugis, G. and Mariotti, D. (1993) Genetic transformation in the grain legume *Cicer arietinum* L. (chickpea). *Plant Cell Reports* 12, 194–198.
- Ford-Lloyd, B. and Jackson, M. (1986) *Plant Genetic Resources: An Introduction to their Conservation and Use*. Edward Arnold, London, 146 pp.
- Franks, T. and Birch, R.G. (1991) Microprojectile techniques for direct gene transfer into intact plant cells. In: Murray, D.R. (ed.) *Advanced Methods in Plant Breeding and Biotechnology*. CAB International, Wallingford, UK, pp. 103–127.
- Garfinkel, D.J., Simpson, R.B., Ream, L.W., White, F.F., Gordon, M.P. and Nester, E.W. (1981) Genetic analysis of crown gall: fine structure map of T-DNA by site-directed mutagenesis. *Cell* 27, 143–153.
- Ghazi, T.D., Cheema, H.V. and Nabors, M.W. (1986) Somatic embryogenesis and plant regeneration from embryogenic callus of soybean, *Glycine max* L. *Plant Cell Reports* 5, 452–456.
- Gierl, A., Schwarz-Sommer, Z. and Saedler, H. (1985) Molecular interactions between the components of the En-1 transposable element system of *Zea mays*. *EMBO Journal* 4, 579–583.
- Godwin, I.D. and Chikwamba, R.K. (1993). Fertile transgenic sorghum (*Sorghum bicolor* L. Moench.) plants via *Agrobacterium*. In: Imrie, B.C. and Hacker, J.B. (eds) *Focussed Plant Improvement: Towards Responsible and Sustainable Agriculture. Proceedings of the 10th Australian Plant Breeding Conference*, vol. 1. Australian Convention and Travel Service, Canberra, pp. 253–257.
- Godwin, I.D., Ford-Lloyd, B.V. and Newbury, H.J. (1992) *In vitro* approaches to extending the host-range of *Agrobacterium* for plant transformation. *Australian Journal of Botany* 40, 751–763.
- Gordon-Kamm, W.J., Spencer, T.M., Mangano, M.L., Adams, T.R., Daines, R.J., Start, W.G., O'Brien, J.V., Chambers, S.A., Adams, Jr. W.R., Willetts, N.G., Rice, T.B., Mackey, C.J., Krueger, R.W., Kausch, A.P. and Lemaux, P.G. (1990) Transformation of maize cells and regeneration of fertile transgenic plants. *The Plant Cell* 2, 603–618.
- Gould, J., Devey, M., Hasegawa, O., Ulian, E.C., Paterson, G. and Smith, R.H. (1991)

- Transformation of *Zea mays* L. using *Agrobacterium tumefaciens* and the shoot apex. *Plant Physiology* 95, 426–434.
- Gunstone, F.D., Harwood, J.L. and Padley, F.D. (1994) *The Lipid Handbook*. Chapman & Hall, London.
- Haring, M.A., Gao, J., Volbeda, T., Rommens, C.M.T., Nijkamp, J.J. and Hille, J. (1989) A comparative study of *Tam3* and *Ac* transposition in transgenic tomato and petunia plants. *Plant Molecular Biology* 13, 189–201.
- Hess, D. (1975) Uptake of DNA and bacteriophage into pollen and genetic manipulation. In: Ledoux, L. (ed.) *Genetic Manipulation with Plant Material*. Plenum Press, New York, pp. 519–537.
- Hess, D. (1987) Pollen based technique and genetic manipulation. *International Review of Cytology* 107, 169–190.
- Hiei, Y., Ohta, S., Komari, T. and Kumashiro, T. (1994) Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. *The Plant Journal* 6, 271–282.
- Hightower, R., Baden, C., Penzes, E., Lund, P. and Dunsmuir, P. (1991) Expression of antifreeze protein in transgenic plants. *Plant Molecular Biology* 17, 1013–1021.
- Hilder, V.A., Gatehouse, A.M.R., Sheerman, S.E., Barker, R.F. and Boulter, D. (1987) A novel mechanism of insect resistance engineered into tobacco. *Nature* 330, 160–163.
- Hilder, V.A., Powell, K.S., Gatehouse, A.M.R., Gatehouse, J.A., Gatehouse, L.N., Shi, Y., Hamilton, W.D.O., Merryweather, A., Newell, C.A., Timans, J.C., Peumans, W.J., van Damme, E. and Boulter, D. (1995) Expression of snowdrop lectin in transgenic tobacco plants results in added protection against aphids. *Transgenic Research* 4, 18–25.
- Hinchee, M.A.W., Connor-Ward, D.V., Newell, R.E., McDonnell, R.E., Sato, S.J., Gasser, C.S., Fischhoff, D.A., Re, D.B., Fraley, R.T. and Horsch, R.B. (1988) Production of transgenic soybean plants using *Agrobacterium*-mediated DNA transfer. *Bio/Technology* 5, 263.
- Hood, E.E., Helmer, G.L., Fraley, R.T. and Chilton, M.-D. (1986) The hypervirulence of *Agrobacterium tumefaciens* A281 is incoded in a region of pTiBo542 outside of the T-DNA. *Journal of Bacteriology* 168, 1291–1301.
- James, D.J., Uratsu, S., Cheng, J., Negri, P., Viss, P. and Dandekar, A.M. (1993) Acetosyringone and osmoprotectants like betaine or proline synergistically enhance *Agrobacterium*-mediated transformation of apple. *Plant Cell Reports* 12, 559–563.
- Jansen, R.C., Van Ooijen, J.W., Stam, P., Lister, C. and Dean, C. (1995) Genotype-by-environment interaction in genetic mapping of multiple quantitative trait loci. *Theoretical and Applied Genetics* 91, 33–37.
- Jones, D.A., Thomas, C.M., Hammond-Kosack, K.E., Balint-Kurti, P.J. and Jones, J.D.G. (1994) Isolation of the tomato Cf-9 gene for resistance to *Cladosporium fulvum* by transposon tagging. *Science* 266, 789–793.
- Kaeppeler, H.F., Somers, D.A., Rines, H.W. and Cockburn, A.F. (1992) Silicon carbide fiber-mediated stable transformation of plant cells. *Theoretical and Applied Genetics* 84, 560–566.
- Kalloo, G. (1991) Interspecific and intergeneric hybridization in tomato. In: Kalloo, G. (ed.) *Monographs on Theoretical and Applied Genetics 14: Genetic Improvement of Tomato*. Springer-Verlag, Berlin, pp. 73–82.

- Kan, J.A.L. van, Ackerveken, G.F.J.M. van den, and de Wit, P.J.G.M. (1991) Cloning and characterization of cDNA of avirulence gene *avr9* of the fungal pathogen *Cladosporium fulvum*, causal agent of tomato leaf mold. *Molecular Plant-Microbe Interactions* 4, 52–59.
- Kenwick, S., Patterson, M., Speer, A., Fischbeck, K. and Davies, K. (1987) Molecular analysis of the Duchenne muscular dystrophy region using pulsed field gel electrophoresis. *Cell* 48, 351–357.
- Kieber, J.J., Rothenburg, M., Roman, G., Feldmann, K.A. and Ecker, J.R. (1993) *CTR1*, a negative regulator of the ethylene response pathway in *Arabidopsis*, encodes a member of the Raf family of protein kinases. *Cell* 72, 427–441.
- Kishore, G.M. and Sommerville, C.R. (1993) Genetic engineering of commercially useful biosynthetic pathways in transgenic plants. *Current Opinion in Biotechnology* 4, 152.
- Klee, H.J., White, F.F., Iyer, V.N., Gordon, M.P. and Nester, E.W. (1983) Mutational analysis of the virulence region of an *Agrobacterium tumefaciens* Ti plasmid. *Journal of Bacteriology* 153, 878–883.
- Klein, T.M., Wolf, E.D., Wu, R. and Sanford, J.C. (1987) High-velocity microprojectiles for delivering nucleic acids into living cells. *Nature* 327, 70–73.
- Knott, D.R. (1961) The inheritance of rust resistance. VI. The transfer of stem rust resistance from *Agropyron elongatum* to common wheat. *Canadian Journal of Plant Science* 41, 109–123.
- Koncz, C., Martini, N., Mayerhofer, R., Koncz-Kalman, Z. and Körber, H. (1989) High-frequency T-DNA-mediated gene tagging in plants. *Proceedings of the National Academy of Sciences of the USA* 86, 8467–8471.
- Kunananvachaidach, R., Godwin, I.D. and Adkins, S.W. (1995) High efficiency plant regeneration from callus induced on mature *Indica* rice caryopses. *Australian Journal of Botany* 43, 337–348.
- Larkin, P.J. and Scowcroft, W.R. (1981) Somaclonal variation: a novel source of genetic variability from cell cultures for plant improvement. *Theoretical and Applied Genetics* 60, 197–214.
- Last, D.I., Brettel, R.I.S., Chamberlain, D.A., Chaudhury, A.M., Larkin, P.J., Marsh, E.L., Peacock, W.J. and Dennis, E.S. (1991) pEMU: An improved promoter for gene expression in cereal cells. *Theoretical and Applied Genetics* 81, 581–588.
- Li, B.J., Xu, X.P., Shi, H.P. and Ke, X.Y. (1991) Introduction of foreign genes into the seed embryo cells of rice by electro-injection and the regeneration of transgenic rice plants. *Science China* 34, 923–930.
- Lindsay, K. and Gallois, P. (1990) Transformation of sugarbeet (*Beta vulgaris* L.) by *Agrobacterium tumefaciens*. *Journal of Experimental Botany* 41, 529–536.
- McCarty, D.R., Carson, C.B., Stinard, P.S. and Robertson, D.S. (1989) Molecular analysis of *viviparous-1*: an abscisic acid insensitive mutant of maize. *Plant Cell* 1, 523–532.
- McClintock, B. (1949) Mutable loci in maize. *Carnegie Institute of Washington Year Book* 48, 142–154.
- McClintock, B. (1950) The origin and behaviour of mutable loci in maize. *Proceedings of the National Academy of Sciences of the USA* 36, 344–355.
- McCormick, S., Niedermeyer, J., Fry, J., Barnason, A., Horsch, R. and Fraley, R. (1986) Leaf disc transformation of cultivated tomato (*Lycopersicon esculentum*) using *Agrobacterium tumefaciens*. *Plant Cell Reports* 5, 81–84.

- McElroy, D., Blowers, A.D., Jenes, B. and Wu, R. (1991) Construction of expression vectors based on the rice actin 1 (*Act1*) 5' region for use in monocot transformation. *Molecular and General Genetics* 231, 150–160.
- MacKenzie, D.R., Anderson, P.M. and Wernham, C.C. (1966) A mobile air blast inoculator for plot experiments with maize dwarf mosaic virus. *Plant Disease Reporter* 50, 363–367.
- Marks, M.D. and Feldmann, K.A. (1989). Trichome development in *Arabidopsis thaliana* I. T-DNA tagging of the *GLABROUS1* gene. *The Plant Cell* 1, 1043–1050.
- Martin, C., Carpenter, R., Sommer, H., Saedler, H. and Coen, E.S. (1985) Molecular analysis of instability in flower pigmentation of *Antirrhinum majus*, following isolation of the *pallida* locus by transposon tagging. *EMBO Journal* 4, 1625–1630.
- Martin, G.B., Williams, J.G.K. and Tanksley, S.D. (1991) Rapid identification of markers linked to *Pseudomonas syringae* resistance in tomato using random primers and near isogenic lines. *Proceedings of the National Academy of Sciences of the USA* 88, 2336–2340.
- Martin, G.B., Ganai, M.W. and Tanksley, S.D. (1992) Construction of yeast artificial chromosome library to tomato and identification of cloned segments linked to two disease resistance loci. *Molecular and General Genetics* 233, 25–32.
- Martin, G.B., de Vicente, M.C. and Tanksley, S.D. (1993a) High-resolution linkage analysis and physical characterization of the *Pto* bacterial resistance locus in tomato. *Molecular Plant–Microbe Interactions* 6, 26–32.
- Martin, G.B., Brommonschenkel, S.H., Chunwongse, J., Frary, A., Ganai, M.W., Spivey, R., Wu, T., Earle, E.D. and Tanksley, S.D. (1993b) Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science* 262, 1432–1436.
- May, G.D., Afza, R., Mason, H.S., Wiecko, A., Novak, F.J. and Arntzen, C.J. (1995) Generation of transgenic banana (*Musa acuminata*) plants via *Agrobacterium*-mediated transformation. *Bio/Technology* 13, 486–492.
- Mayerhofer, R., Koncz-Kalman, Z., Nawrath, C., Bakkeren, G. and Cramer, A. (1991) T-DNA integration: a mode of illegitimate recombination in plants. *EMBO Journal* 10, 697–704.
- Meinke, D.W. and Sussex, I.M. (1979) Embryo-lethal mutants of *Arabidopsis thaliana*: A model system for genetic analysis of plant embryo development. *Developmental Biology* 72, 50–61.
- Mettin, D., Bluthner, W.D. and Schlegel, W.G. (1973) Additional evidence on spontaneous 1B/1R wheat-rye substitutions and translocations. In: *Proceedings 4th International Wheat Genetics Symposium, Columbia, Missouri*, pp. 179–184.
- Michelmore, R., Paran, I. and Kessell, R. (1991) Identification of markers linked to disease-resistance genes by bulked segregation analysis: A rapid method to detect markers in specific genome regions by using segregating populations. *Proceedings of the National Academy of Sciences of the USA* 88, 9828–9832.
- Moore, G.A., Jacono, C.C., Neidigh, J.L., Lawrence, S.D. and Cline, K. (1992) *Agrobacterium*-mediated transformation of *Citrus* stem segments and regeneration of transgenic plants. *Plant Cell Reports* 11, 238–242.
- Motto, M., Maddaloni, M., Ponziani, G., Brembilla, M. and Marotti, R. (1988) Molecular cloning of the *o2-m5* allele of *Zea mays* using transposon marking. *Molecular and General Genetics* 212, 488–494.

- Mullins, M.G., Tang, F.C.A. and Facciotti, D. (1990) *Agrobacterium*-mediated genetic transformation of grapevines: transgenic plants of *Vitis rupestris* Scheele and buds of *Vitis vinifera* L. *Bio/Technology* 8, 1041–1045.
- Nelson, R.S., McCormick, S.M., Delannay, X., Dube, P., Layton, J., Anderson, E.J., Vanieska, M., Proksch, P.K., Horsch, R.B., Rogers, G.G., Fraley, R.T. and Beachy, R.N. (1988) Virus tolerance, plant growth and field performance of transgenic tomato plants expressing coat protein from tobacco mosaic virus. *Bio/Technology* 6, 403–409.
- O'Reilly, C., Sheperd, N.S., Periera, A., Schwarz-Sommer, Z., Bertram, I. and Saedler, H. (1985) Molecular cloning of the *a1* locus of *Zea mays* using the transposable elements *En* and *Mu*. *EMBO Journal* 4, 877–882.
- Owens, L.D. and Smigocki, A.N. (1988) Transformation of soybean cells using mixed strains of *Agrobacterium tumefaciens* and phenolic compounds. *Plant Physiology* 88, 570–573.
- Paszkowski, J., Shillito, R.D., Saul, M., Maudak, V., Hohn, T., Hohn, B. and Potrykus, I. (1984) Direct gene transfer to plants. *EMBO Journal* 3, 2717–2722.
- Paterson, A.H., Lin, Y.-R., Li, Z., Schertz, K.F., Doebley, J.F., Pinson, S.R.M., Liu, S.-C., Stansel, J.W. and Irvine, J.E. (1995) Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. *Science* 269, 1714–1718.
- Paz-Ares, J., Wienand, U., Peterson, P.A. and Saedler, H. (1986) Molecular cloning of the *c* locus of *Zea mays*: a locus regulating the anthocyanin pathway. *EMBO Journal* 8, 1315–1321.
- Powell, K.S., Gatehouse, A.M.R., Hilder, V.A. and Gatehouse, J.A. (1993) Antimetabolic effects of plant lectins and plant and fungal enzymes on the nymphal stages of two important rice pests, *Nilaparvata lugens* and *Nephotettia cinctipes*. *Entomological Experimental Applications* 66, 119–126.
- Prakash, C.S. and Varadarajan, U. (1992) Genetic transformation of sweet potato by particle bombardment. *Plant Cell Reports* 11, 53–57.
- Pua, E.C., Mehra-Polta, A., Nagy, F. and Chua, N.H. (1987) Transgenic plants of *Brassica napus* L. *Bio/Technology* 5, 815–817.
- Puonti-Kaerlas, J., Stabel, P. and Eriksson, T. (1989) Transformation of pea (*Pisum sativum* L.) by *Agrobacterium tumefaciens*. *Plant Cell Reports* 8, 321–324.
- Raineri, B.M., Bottini, P., Gordon, M.P. and Nester, E.W. (1990) *Agrobacterium*-mediated transformation of rice (*Oryza sativa* L.) *Bio/Technology* 8, 33–38.
- Redei, G.P. (1975) *Arabidopsis* as a genetic tool. *Annual Review of Genetics* 9, 111–127.
- Röbbelen, G. (1988) Development of new industrial oil crops. In: Applewhite, T.H. (ed.) *Proceedings of the World Conference on Biotechnology for the Fats and Oils Industry*. AOCS, Champaign, Illinois, p. 78.
- Robertson, D.S. (1978) Characterization of a *mutator* system in maize. *Mutation Research* 5, 21–28.
- Rommens, J.M., Iannuzzi, M.C., Kerem, B., Drumm, M.L., Melmer, G., Dean, M., Rozmahel, R., Cole, J.L., Kennedy, D., Hidaka, N., Zsiga, M., Buchwald, M., Riordan, J.R., Tsui, L.-C. and Collins, F.S. (1989) Identification of the cystic fibrosis gene: Chromosome walking and jumping. *Science* 245, 1059–1065.
- Russell, D.R., Wallace, K.M., Bathe, J.H., Martinell, B.J. and McCabe, D.E. (1993) Stable transformation of *Phaseolus vulgaris* via electric-discharge mediated particle acceleration. *Plant Cell Reports* 12, 165–169.

- Russell-Kikkert, J. (1993) The Biolistic PDS-1000 / He device. *Plant Cell, Tissue and Organ Culture* 33, 221–226.
- Saccardo, F., Tomarchio, L. and Monti, L.M. (1975) *Lycopersicon hirsutum* as a source of resistance to *Phytophthora infestans* in tomatoes. *Genetica Agraria* 29, 215–222.
- Sagi, L., Panis, B., Reny, S., Schoofs, H., DeSmet, K., Swennen, R. and Cammue, B.P.A. (1995) Genetic transformation of banana and plantain (*Musa* spp.) via particle bombardment. *Bio/Technology* 13, 481–485.
- Salimath, S.S., de Oliveira, A.C., Godwin, I.D. and Bennetzen, J.L. (1995) Assessment of genome origins and genetic diversity in the genus *Eleusine* with DNA markers. *Genome* 38, 757–763.
- Sato, S., Newell, C., Kolacz, K., Tredo, C., Finer, J. and Hinchee, M. (1993) Stable transformation via particle bombardment in two different soybean regeneration systems. *Plant Cell Reports* 12, 408–413.
- Schafer, W., Gorz, A. and Kahl, G. (1987) T-DNA integration and expression in a monocot crop plant after induction of *Agrobacterium*. *Nature* 327, 529–532.
- Scofield, S.R., Harrison, K., Nurrish, S.J. and Jones, J.D.G. (1992) Promoter fusions to the *Activator* transposase gene cause distinct patterns of *Dissociation* excision in tobacco cotyledons. *The Plant Cell* 4, 573–582.
- Shillito, R.D., Carswell, G.K., Johnson, C.M., DiMaio, J.J. and Harms, C.T. (1989) Regeneration of fertile plants from protoplasts of elite inbred maize. *Bio/Technology* 7, 581–587.
- Shimamoto, K., Terada, R., Izawa, T. and Fujimoto, H. (1989) Fertile transgenic rice plants regenerated from transformed protoplasts. *Nature* 338, 274–276.
- Shizuya, H., Birren, B., Kim, U.J., Mancino, V., Slepak, T., Tachiiri, Y. and Simon, M. (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proceedings of the National Academy of Sciences of the USA* 89, 8794–8797.
- Simmonds, N.W. and Shepherd, K. (1955) The taxonomy and origin of the cultivated bananas. *Journal Linnean Society of Botany* 55, 302–312.
- Smith, C.J.S., Watson, C.F., Ray, J., Bird, C.R., Morris, P.C., Schuch, W. and Grierson, D. (1988) Antisense RNA inhibition of polygalacturonase gene expression in transgenic tomatoes. *Nature* 334, 724–726.
- Somers, D.A., Rines, H.W., Gu, W., Kaeppler, H.F. and Bushnell, W.R. (1992) Fertile, transgenic oat plants. *Bio/Technology* 10, 1589–1594.
- Songstad, D.D., Somers, D.A. and Griesbach, R.J. (1995) Advances in alternative DNA delivery techniques. *Plant Cell, Tissue and Organ Culture* 40, 1–15.
- Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* 98, 503–517.
- Springer, P.S., Edwards, K.J. and Bennetzen, J.L. (1994) DNA clone organization on maize *ADH1* yeast artificial chromosomes. *Proceedings of the National Academy of Sciences of the USA* 91, 863–876.
- Stachel, S.E., Nester, E.W. and Zambryski, P. (1986) A plant cell factor induces *Agrobacterium tumefaciens* vir gene induction. *Proceedings of the National Academy of Sciences of the USA* 83, 379–383.
- Tanksley, S.D., Young, N.D., Paterson, A.H. and Bonierbale, M.W. (1989) RFLP mapping in plant breeding: New tools for an old science. *Bio/Technology* 7, 257–264.

- Tanksley, S.D., Ganai, M.W., Prince, J.C., de Vicente, M.C., Bonierbale, M.W., Broun, P., Fulton, T.M., Giavononni, J.J., Grandillo, S., Martin, G.B., Messegeur, R., Miller, J.C., Paterson, A.H., Pineda, O., Roder, M.S., Wing, R.A., Wu, W. and Young, N.D. (1992) High density molecular linkage maps of the tomato and potato genomes. *Genetics* 132, 1141–1160.
- Thierfelder, A., Luhs, W. and Friedt, W. (1992) Breeding of industrial oil crops with the aid of biotechnology: a review. *Proceedings 1st European Symposium on Industrial Crops and Products*. Elsevier, Amsterdam, pp. 261–271.
- Thomashow, M.F., Panagopolous, C.G., Gordon, M.P. and Nester, M.W. (1980) Host range of *Agrobacterium tumefaciens* is determined by the Ti plasmid. *Nature* 283, 794–796.
- Töpfer, R., Martini, N. and Schell, J. (1995) Modification of plant lipid synthesis. *Science* 268, 681–685.
- Toriyama, K., Arimoto, Y., Uchimiya, H. and Hinata, K. (1988) Transgenic rice plants after direct gene transfer into protoplasts. *Bio/Technology* 6, 1072–1074.
- U, N. (1935) Genome analysis in *Brassica* with special reference to experimental formation of *B. napus* and peculiar mode of fertilization. *Japanese Journal of Genetics* 7, 389–452.
- Umbeck, P., Johnson, G., Barton, K. and Swain, W. (1987) Genetically transformed cotton (*Gossypium hirsutum* L.) plants. *Bio/Technology* 5, 263–266.
- Vain, P., Keen, N., Murillo, J., Rathus, C., Nemes, C. and Finer, J.J. (1993) Development of the particle inflow gun. *Plant Cell, Tissue and Organ Culture* 33, 237–246.
- Van Bruna, J. (1992) The battle of engineered tomatoes. *Bio/Technology* 10, 748.
- Van Sluys, M.A., Tempe, J. and Federoff, N.V. (1987) Transposition of the maize Activator element in *Arabidopsis thaliana* and *Daucus carota*. *EMBO Journal* 6, 3881–3889.
- Vasil, V., Castillo, A.M., Fromm, M.E. and Vasil, I.K. (1992) Herbicide resistant fertile transgenic wheat plants obtained by microprojectile bombardment of regenerable embryogenic callus. *Bio/Technology* 10, 667–674.
- Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. and Zabeau, M. (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23, 4407–4414.
- Walbot, V. (1992) Strategies for mutagenesis and gene cloning using transposon tagging and T-DNA insertional mutagenesis. *Annual Review of Plant Physiology and Plant Molecular Biology* 43, 49–82.
- Wan, Y. and Lemaux, P.G. (1994) Generation of large numbers of independently transformed fertile barley plants. *Plant Physiology* 104, 37–48.
- Wang, M.B., Boulter, D. and Gatehouse, J.A. (1992) A complete sequence of the rice sucrose synthase-1 (*Rss1*) gene. *Plant Molecular Biology* 19, 881–885.
- Wessler, S.R. (1988) Phenotypic diversity mediated by the maize transposable elements *Ac* and *Ds*. *Science* 242, 399–405.
- Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. and Tingey, S.V. (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* 18, 6531–6535.
- Yadav, N., Wierzbicki, A., Knowlton, S., Pierce, J., Ripp, K., Hitz, W., Aegerter, M. and Browse, J. (1993) Genetic manipulation to alter fatty acid profiles of oilseed crops. *Current Topics in Plant Physiology* 9, 60–66.

- Yoder, J.I., Palys, J., Alpert, K. and Lassner, M. (1988) Ac transposition in transgenic tomato plants. *Molecular and General Genetics* 213, 219–296.
- Young, N.D. and Phillips, R.L. (1994) Meeting Report: Cloning genes known only by phenotype. *The Plant Cell* 6, 1193–1195.
- Zabeau, M. and Vos, P. (1994) Selective restriction fragment amplification: a general method for DNA fingerprinting. *European Patent Application* 92402629.7 (Publication number: 0 534 858 A1).
- Zambryski, P.C. (1992) Chronicles from the *Agrobacterium*-plant cell DNA transfer story. *Annual Review of Plant Physiology and Plant Molecular Biology* 43, 465–490.
- Zambryski, P., Joos, H., Genetello, C., Leemans, J., van Montagu, M. and Schell, J. (1983) Ti plasmid vector for the introduction of DNA into plant cells without alteration of their normal regeneration capacity. *EMBO Journal* 2, 2143–2150.
- Zietkiewicz, E., Rafalski, A. and Labuda, D. (1994) Genome fingerprinting by simple sequence repeats (SSR) anchored PCR amplification. *Genomics* 20, 176–183.

Importance of Biotechnology for Germplasm Health and Quarantine

H. Barker and L. Torrance

Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, UK

10.1. Introduction

Germplasm collections for particular crops have been made and are conserved at many International Agricultural Research Centres. There are also many private and publicly funded germplasm collections in developed countries and increasing awareness of the importance of *in situ* conservation of landraces in centres of crop diversity (Brush, 1995). The value of all these genetic resources lies in their utilization in crop improvement programmes. The original plants, frequently acquired from a noncultivated environment, are selected for their plant breeding potential and not for their freedom from plant pests and pathogens. Therefore, it is not surprising that germplasm collections may contain plants infected with pathogens endemic to the geographical area in which they were collected. The type of collection dictates the approach to pathogen detection and elimination. Thus, for collections in International Centres, systematic pathogen elimination can be practised, whereas for landrace collections, control may be very difficult or impossible. There are also differences in the strategy of pathogen detection and elimination depending on the type of germplasm maintained. Vegetatively propagated species tend to be maintained in *in vitro* culture or as field or greenhouse-grown plants, whereas seed-propagated and some root and tuber crops are maintained through seed. However, there are several instances where germplasm is maintained by a combination of several methods. Seed genebanks present fewer problems because there are relatively fewer seed-borne pathogens.

The international distribution of germplasm poses the risk of

accidental introduction of nonindigenous plant pathogens along with the host material. The detection and elimination of such pathogens is essential to prevent the introduction of disease problems into a breeding programme with the potential for subsequent release in new cultivars. In recent years the Food and Agriculture Organization (FAO) of the United Nations and the International Plant Genetic Resources Institute (IPGRI) have launched a collaborative programme to generate a series of crop-specific technical guidelines (for an example, see Diekmann and Putter, 1995) that provide relevant information on disease indexing and other procedures that will help to ensure phytosanitary safety when germplasm is moved internationally. Pathogens that are often symptomless, such as viruses and viroids, pose the greatest risk to germplasm collections, and this chapter focuses on the increasing role and future potential of biotechnology in their detection and elimination from quarantine and germplasm collections.

10.2. Strategy of Virus Detection and Elimination in Germplasm Collections

10.2.1. Rationale for virus elimination

There are a number of reasons for pathogen elimination in germplasm collections. The *raison d'être* of germplasm collections is to disseminate new sources of genetic diversity to breeding programmes but without spreading pathogens. It is important that breeders can see and utilize the full potential of a collection without being influenced by unfavourable characteristics caused by virus infection. Infection may cause poor growth, loss of productivity and ultimately affect survival of members of the collection which would act to decrease diversity. If a collection is known to contain infected material it may be avoided by breeders and therefore its full potential diversity will not be exploited. Such a situation occurred with a *Pisum sativum* germplasm collection, many accessions of which were known to be infected with pea seed-borne mosaic virus (Hampton *et al.*, 1993).

10.2.2. Germplasm and types of tissue at risk

Crops propagated by vegetative means can be afflicted by severe disease problems because, if unchecked, viruses can accumulate and persist in propagated material. With only few exceptions, viruses and viroids are transmitted very efficiently in vegetatively propagated tissues. Thus, in the

exchange of germplasm, the greatest risk is posed by the use of vegetative propagules. Most viruses are eliminated during seed formation but there are many exceptions and for seed-borne viruses the proportion of infected seed is not fixed and depends on many environmental and host-specific factors. Viruses that are seed-borne are also frequently transmitted via pollen. There are several valuable reference sources containing descriptions and detection / diagnosis protocols for plant viruses, including Brunt *et al.* (1996), FAO / IPGRI technical notes (Diekmann and Putter, 1995) and the AAB / CMI Virus Descriptions now published by the Association of Applied Biologists.

10.2.3. Obtaining virus-free nuclear stock plants

Virus-free source plants for use in a germplasm collection can be identified by indexing a number of individual plants in the hope that a proportion are healthy. Once identified, healthy plants can be used for propagation with continued indexing to ensure that reinfection does not occur and that latent viruses do not appear. A frequently used alternative method is to remove apical meristems, which contain little or no virus, and culture these to develop virus-free plantlets. If virus-free plants cannot be obtained by these means, it is possible to free many plant species from virus infection by thermotherapy. For example Kunkel (1936) showed that peach plants could be freed from yellows disease agent by growing them at 34–36°C for several weeks. About half the known viruses of vegetatively propagated horticultural species can be eliminated from plants by this treatment. The technique relies on holding actively growing plant tissues at 35–40°C for several weeks. Low-temperature thermotherapy has been found to be successful in the elimination of viroids (Paduch-Cichal and Kryczynski, 1987). There are a few chemical compounds, e.g. purine and pyrimidine base analogues, that have known antiviral effects and these have been used for viral elimination (Griffiths *et al.*, 1990). Frequently, culture of apical meristems combined with thermotherapy or chemotherapy is used to obtain virus- or viroid-free plants.

10.2.4. Development of biotechnology-based virus detection methods

The majority of plant viruses consist of a nucleic acid genome surrounded by a protein coat, although there are some notable exceptions such as viroids, some satellite sequences and nonencapsidated virus strains, e.g. NM forms of tobnaviruses (Harrison and Robinson, 1981) which have no coat protein. Apart from techniques which rely on microscopy and bioassay, detection techniques can be divided broadly into two groups: those

that rely on serological detection of the capsid protein, and those that rely on detection of the viral nucleic acid. Serological techniques originated in the 1940s, initially depending on precipitation reactions in liquid or agar gel media. These early techniques were not very sensitive and used relatively large amounts of antisera. However, with the development of enzyme-linked immunosorbent assay (ELISA) for plant viruses in the mid-1970s and its more recent variants, great increases in sensitivity have been obtained. Techniques in current use include ELISA in polystyrene microtitre plates (Clark and Adams, 1977), dot-ELISA on nitrocellulose membranes (Weidemann, 1988) and magnetic microsphere enzyme-linked immunoassay (Banttari *et al.*, 1991).

Techniques which rely on detection of the viral nucleic acid include those in which target molecules, frequently immobilized on nitrocellulose support membranes, are hybridized with cloned nucleic acid sequences linked either to radioactive or nonradioactive reporter groups. Detection of nucleic acid by polymerase chain reaction (PCR) is a technique which has found great utility because of the enormous sensitivity that PCR can provide.

The choice of detection technique may be dictated by a variety of circumstances: for example, whether or not an antiserum is available for serological testing or the existence of sequence information for the design of PCR primers. For nonencapsidated viruses and viroids the detection technique must be based on nucleic acid detection. Costs, sensitivity of detection and suitability for testing large numbers of samples also play a part in selecting the optimum technique. Thus ELISA may be simple, cheap and suitable for testing very large sample numbers because it is easy to automate. On the other hand, greater sensitivity may be obtained by PCR detection but it is likely to impose greater cost and labour inputs and, at present, is unsuitable for large sample numbers. However, recent improvements could aid the adoption of PCR for more routine and automated use (Bariana *et al.*, 1994; Hataya *et al.*, 1994).

The kinds of tests used will also depend on a number of other factors, such as: timescale for release of material from quarantine; range of potential contaminating pathogens; nature of germplasm; facilities and budget available for quarantine tests; the ease with which nonindigenous pathogens, which may not be well known in the country of import, can be detected. Biotechnological approaches should enable a more rapid throughput of material with a consequent decrease in the requirement for expensive screenhouse or heated glasshouses if plants need to be grown during quarantine. The tests should aim to be sensitive, cheap and laboratory-based rather than relying on a biological test in which transmission to an indicator host plant could increase the risk of accidental release. Indeed, for some crops such as potatoes, cereals and legumes there is potential to achieve this because many of the viruses of these hosts are well studied

and new biotechnological detection methods have been developed. However, in certain perennial crops such as tree fruit, small fruit, and less-studied crops, we recognize that biological testing will continue to be an essential requirement to intercept diseases of unknown aetiology. Another advantage of biotechnology is that tests can use standardized reagents and protocols necessary for accreditation of testing procedures in order to give an internationally recognized standard of performance. This should lead to greater confidence among breeders and plant health authorities for the safe movement of germplasm.

10.3. Virus Detection Methods

10.3.1. Bioassay

Traditional methods for bioassay include mechanical inoculation of tissue extracts or grafting test plants to susceptible test plants. Such methods can be very sensitive but have the disadvantage that suitable test plants must be available whenever they are required. Furthermore, it may not be possible or economic to arrange a supply of every test species required to distinguish all likely target viruses. Although a typical bioassay may reveal that an infectious agent is present, it may not provide a precise diagnosis but could provide sufficient reason to eliminate a particular plant. Nevertheless, bioassay can be hazardous because it involves deliberate propagation of the very organisms for which attempts are being made to intercept and eliminate. Even if such activity is confined to a containment or quarantine facility, there is the potential for inadvertent dispersal.

10.3.2. Detection by microscopy

Virus detection using the electron microscope (EM) has been used for many years and has several advantages over other techniques. For some viruses, a simple examination of tissue extracts may provide sufficient information for diagnosis on the basis of a characteristic particle morphology. Although many viruses share a common morphology which can invalidate a simple diagnosis, as with bioassay, the recognition of virus particles can provide sufficient reason to eliminate a plant. However, simple EM techniques are relatively insensitive although the use of virus-specific antiserum-coated grids (immunosorbent electron microscopy – ISEM) can improve sensitivity and aid diagnosis (Roberts, 1986).

10.3.3. Nucleic acid-based detection

Double-stranded RNA (dsRNA)

Healthy plants should not, under normal circumstances, contain any large double-stranded RNA (dsRNA) because host plant RNA is transcribed from DNA. Thus, the detection of high-molecular mass ($>1.0 \times 10^6$ daltons) dsRNA in plants indicates some abnormality, the most likely being infection by an agent with an RNA genome (Dodds *et al.*, 1984; Jones, 1992). Once extracted, techniques for detection and analysis of dsRNAs comprise gel electrophoresis and detection by staining with ethidium bromide and/or silver nitrate. Care must be taken in the use of dsRNA analysis and directly comparable controls should be included to ensure that false positive results do not occur. For example, virus infection is common in arthropods that infest plants, including those that are considered not to be vectors of plant viruses, and infected individuals can be a source of contamination to virus-free plants (Jones, 1992). Furthermore, dsRNA species of apparently plant host origin can be detected in normal-looking plants (Gould and Francki, 1981; Wakarchuk and Hamilton, 1985).

Although most plant viruses have single-stranded RNA (ssRNA), a few have double-stranded RNA (dsRNA) genomes, including members of the families *Reoviridae* and *Partitiviridae*. The concentration of dsRNA found in plants infected with these viruses is relatively large and easy to detect. The use of dsRNA analysis to detect the so-called cryptoviruses (members of the family *Partitiviridae*), which include white clover cryptic virus and beet cryptic virus, is noteworthy. Such viruses are associated with latent infections in usually symptomless plants, are present in low concentrations, are not transmissible by insect vector, graft or mechanical means, but are transmitted through seed and pollen. These characteristics have made them difficult to detect by other means. However, cryptovirus infections have been readily detected by dsRNA analysis of *Vicia faba* plants (Abou-Elnasr *et al.*, 1985), brassicas (Jones *et al.*, 1986) and cucumber (Nameth and Dodds, 1985).

Viruses with ssRNA genomes replicate in plants by the production of an RNA strand complementary to its genome which forms a base-paired dsRNA molecule termed replicative form (RF). Although ssRNA viruses can usually be readily detected by other means, there are some instances when dsRNA analysis is useful despite the fact that considerably less dsRNA is found in plants infected with viruses having ssRNA genomes, than with those having dsRNA genomes. For example, Hu *et al.* (1991) indicated that because of the involvement of different viruses and unavailability of a wide range of antisera, dsRNA assay is particularly useful for testing grapevines for leafroll disease. Indeed, such viral agents were readily detected in polyacrylamide gel electrophoresis (PAGE) by Habili *et al.* (1992), who concluded that dsRNA assay may be used as a faster and less

expensive method than biological indexing in assessing the elimination of the viral agent causing grapevine leafroll.

Detection of nucleic acid by hybridization

Nucleic acid hybridization techniques for detection of viroids and viruses have been in use since the late 1970s. The technique involves the use of labelled complementary DNA or RNA probes prepared from purified viroid / viral nucleic acid or a cloned copy of such nucleic acid. When the probe is incubated with viroid or viral nucleic acid extracts, double-stranded hybrids form by specific base pairing which can be detected by the presence of labelled nucleotides incorporated into the nucleic acid probe. Initially, attention with this technique was focused on the detection of viroid RNA to provide a means of detection because of the lack of serological tests and to provide a faster test than time-consuming biological tests. Thus, Palukaitis *et al.* (1981) reported that it took just a few days to detect avocado sunblotch viroid by using a cDNA probe, whereas up to two years was required for indexing by biological methods. Hybridization studies were done initially in liquid reactions, but the technique was greatly improved when Owens and Diener (1981) developed the nucleic acid spot hybridization (NASH) test (sometimes known as dot-blot hybridization assay). In the NASH test, samples are blotted onto a nitrocellulose membrane support which is then incubated in a solution of labelled probe.

Until recently, the most common method of probe labelling involved incorporation of ^{32}P - or ^{35}S -labelled dNTPs into the cDNA or cRNA. However, the use of radioisotopes involves special procedures and facilities for handling and a requirement for the reliable delivery of radiochemicals, which frequently have short half-lives. Such facilities may be expensive and / or, in some countries, difficult to obtain. Furthermore, radioactive decay limits the time a probe can be used – e.g. ^{32}P -labelled cDNA probes have a maximum useful lifespan of about two weeks. By comparison, nonradioactive probes avoid many of the problems associated with handling radioactive probes. The use of biotin, and more recently the digoxigenin (DIG) system developed by Boehringer Mannheim, for non-radioactive labelling of RNA and DNA probes has been refined so that sensitivity is comparable with radioactive probes (Kanematsu *et al.*, 1991).

The DIG system uses digoxigenin, a steroid hapten in the form of DIG-11-dUTP, to label cloned DNA or RNA. Probes can be produced by a number of methods, including PCR and random priming. The process of hybridization and detection is relatively simple and resembles a conventional NASH procedure. Membrane-bound nucleic acids are hybridized with labelled probes, which are detected using an alkaline phosphatase-conjugated anti-digoxigenin antibody and a chemiluminescent substrate. Chemiluminescent substrates can be visualized by exposure to X-ray film,

thus providing a permanent record which is directly comparable to results that would be obtained using radioactive probes. Luminescent detection is fast and sensitive, and membranes can easily be stripped and reprobed.

DIG probes have several advantages over radioactive probes. For example, using chemiluminescent substrates, detection can be completed with a 15–30 min exposure to X-ray film compared with 1–3 days required for most ^{32}P -labelled probes. Another advantage is that probes can be stored at 4°C or –20°C for many years with no apparent loss in sensitivity. For example, radioactive and nonradioactive cDNA probes to tobacco rattle virus RNA were prepared from a sample of the same plasmid preparation used to make a DIG probe nearly three years earlier and which had been stored at –20°C (Webster and Barker, 1997). In a comparative test, the stored DIG-labelled probe was as sensitive for detection as the freshly prepared DIG and radioactive probes. DIG-labelled cDNA probes have been used successfully for detection of a number of pathogens including potato spindle tuber viroid in plants (Kanematsu *et al.*, 1991) and true potato seedlings (Borkhardt *et al.*, 1994); potato virus Y in dormant potato tubers (Singh and Singh, 1995); and the satellite RNA of groundnut rosette umbravirus in groundnut (Blok *et al.*, 1995).

Polymerase chain reaction

The polymerase chain reaction (PCR) was first reported to be useful for detection of a plant virus, bean yellow mosaic virus (BYMV) in gladiolus leaf tissue, by Vunsh *et al.* (1990). Detection of viral RNA sequences is accomplished by first synthesizing cDNA using reverse transcriptase (RT step). These cDNA molecules are subsequently used as a template for the PCR, which is capable of amplification of a specific sequence of DNA in a cyclic process. The whole process for detection of RNA genome viruses is known as RT-PCR. The amplification reaction uses two oligonucleotide primers flanking the region of DNA to be amplified and hybridizing to opposite strands. The annealed primers are orientated with their 3' ends facing each other such that synthesis by DNA polymerase extends across the region of the original DNA template between the primers. Since each primer is complementary to one of the newly synthesized strands, each new strand can participate as a template in subsequent cycles of primer extension and amplification. Therefore, each cycle of strand denaturation, primer annealing, and enzymatic extension doubles the amount of DNA from the previous cycle. Each primer becomes incorporated into the amplification products, so that the product that accumulates exponentially is a discrete fragment whose length is a function of the distance between the 5' ends of each primer.

PCR is capable of the selective amplification of a particular DNA sequence by a factor of 10^6 – 10^7 , which enables detection of a single or few copies of DNA (Wong *et al.*, 1987; Saiki *et al.*, 1988), and has the advantage

over cDNA hybridization techniques that the reaction product can be detected easily in electrophoresis gels by ethidium bromide staining. An increase in sensitivity can be obtained by blotting the PCR products with a labelled complementary nucleic acid probe (Vunsh *et al.*, 1990).

Potato germplasm is particularly prone to virus infection and whilst most potato viruses can be detected readily by ELISA of growing foliage, virus is more difficult to detect in tuber tissue. Therefore, a sensitive test would be needed if potato germplasm was imported in the form of tubers and propagation was to be avoided. In recent years RT-PCR has been used successfully on dormant tuber tissue to detect the following viruses: potato virus Y (Barker *et al.*, 1993); potato leafroll virus (Spiegel and Martin, 1993); potato mop-top virus (Arif *et al.*, 1994); and tobacco rattle virus (Robinson and Dale, 1994).

Since the PCR method was first devised, a number of modifications have been developed. For example, the transcriptional amplification system (TAS) (Kwoh *et al.*, 1989) involves the PCR generation of cDNA containing the bacteriophage T7 transcription promoter using appropriately designed primers. The cDNA is then transcribed with T7 RNA polymerase to produce a dsRNA amplification product, up to 100 dsRNA copies per DNA template molecule giving an additional 10- to 100-fold increase in sensitivity. Rosner *et al.* (1992) improved the signal amplification of RT-PCR for BYMV by using this protocol and obtained a positive signal for BYMV in some gladiolus corms in which virus could not be detected by the standard PCR protocol.

Another notable modification for RT-PCR detection of encapsidated viruses is the method of immunocapture PCR (Jansen *et al.*, 1990). In this technique viral particles are trapped and partially purified from plant sap components in an antibody-coated tube. Trapped virus particles can be used for direct RNA extraction and RT-PCR in the same tubes. The use of antibody to trap virus particles simplifies the RNA extraction procedure and enables an increase in signal amplification. Thus Wetzel *et al.* (1992) found that immunocapture PCR, when compared to direct PCR, molecular hybridization with radioactive cRNA probes and ELISA, gave 250-, 625- and 5000-fold increases in sensitivity for detecting plum pox virus.

The application of virus detection by PCR in seed tissue would be of great benefit to the safe movement of germplasm. Sáiz *et al.* (1994) have devised an RT-PCR method to detect bean common mosaic virus in bean seed tissues. Primers have been designed which enable serotype-specific detection, and by using multiplex RT-PCR, serotype A and B isolates were amplified from seed tissues. The PCR system described was approximately 10^3 – 10^5 times more sensitive than ELISA of seed extracts (Sáiz *et al.*, 1994). Seed-borne viruses are of great concern in legume germplasm collections. Bariana *et al.* (1994) have developed an RT-PCR detection method for five seed-borne legume viruses. The RT-PCR assay is over 10^5 times more

sensitive than ELISA and can identify all five seed-borne viruses in a single tube reaction in one multiplex RT-PCR test. Primers have been designed so that the size of the RT-PCR product is indicative of the virus amplified, and conserved regions of sequence have been given preference as primer targets (Bariana *et al.*, 1994).

10.3.4. Antibody-based detection

Antisera have been used to identify plant viruses and provide information on serological relationships since the early days of plant virus research (Matthews, 1991). Nowadays, there are two major kinds of antibody preparations used for virus detection and diagnosis: (i) those obtained from the sera of immunized animals, often called polyclonal antisera; and (ii) monoclonal antibodies. There are several advantages and disadvantages associated with the different kinds of antibody preparation.

Polyclonal antisera

These are generally easier and quicker to produce. A protocol of two or three subcutaneous injections of purified virus preparations over a period of one month followed by successive bleedings over the following 4–6 months is generally sufficient to produce large quantities of sera for routine use. However, a major disadvantage is that the sera sometimes contain antibodies to healthy plant proteins (co-purified with the virus) which can interfere with sensitive tests such as ELISA, producing unacceptably high background reactions. Also, although large quantities of antisera can be produced from one rabbit, the supply is finite and at some point new sera must be made, necessitating further virus preparations; moreover, the new sera may have different titre and binding characteristics.

Monoclonal antibodies

Unlike the diverse population of antibodies found in polyclonal antisera, a preparation of monoclonal antibodies (MAbs) consists of homogeneous antibody molecules, all having the same specificity and affinity for an antigenic determinant or epitope. MAbs are produced *in vitro* from a clonal population of hybridoma cells, which are the product of a somatic fusion between a myeloma cell and a single spleen cell from an immunized animal, usually a mouse (Goding, 1983; Campbell, 1984; Harlow and Lane, 1988). MAbs have been widely used in many areas of the medical, veterinary and agricultural sciences because they confer the advantages of defined specificity and the ready availability of unlimited quantities of a standardized reagent.

Since the early 1980s, monoclonal antibodies (MAbs) have been produced to more than 60 plant viruses in 20 different genera (van

Regenmortel and Dubs, 1993). There are some recent reviews discussing the use of MAbs in routine virus detection (Jordan, 1992; Torrance, 1992a; van Regenmortel and Dubs, 1993). The key points can be summarized as follows: MAbs confer a very high degree of specificity for an epitope and therefore can distinguish epitopes unique to individual virus strains or common to a virus group. Furthermore, the hybridoma cells secreting MAbs remain viable after low-temperature preservation for long periods. Stored cells can be revived and cultured to produce more MAb and fresh cells for preservation. Thus, the MAb is 'immortalized' and, in theory, can be produced indefinitely and in unlimited quantities. Use of MAbs can ensure uniformity and standardization between tests and therefore their use in diagnostic testing is highly desirable.

Some disadvantages of MAbs are as follows: thorough screening of the MAbs against a wide range of virus strains and serotypes is required to ensure the selected MAbs detect appropriate epitopes – the epitope recognized by the MAb must be stable under the conditions of the test so that false negative results do not occur. Also, some MAbs are sensitive to environmental conditions (pH, salt, freezing) or do not work in certain assay formats (e.g. lose activity when used to coat microtitre plates).

It generally takes 6 to 12 months to produce and characterize a panel of MAbs. Therefore, production of a MAb (or MAbs) with all of the desired properties, correct binding specificity as well as the necessary robustness to be used in different test formats, is time consuming and expensive. In purely commercial terms, the expense may be justified for MAbs used in large-scale routine testing, e.g. those raised against viruses of crops such as potatoes, tree fruit and bulbous ornamentals where the health status is guaranteed by certification schemes; or tests for virus infection of seed lots or micropropagated plants, or breeding lines for virus resistance, where very large numbers of tests are done annually. There are other kinds of tests that may need to be done on a smaller scale and where the expense incurred by incorporation of MAbs could not be justified on purely commercial grounds. For example, screening of germplasm by plant quarantine authorities, and the provision of plant passports for movement of plants between different countries within the European Union (EU) will necessitate screening for many different viruses. However, use of standardized MAb reagents of known quality to provide accredited testing would inspire confidence among individual member states.

Antibody-based test methods

Tests on germplasm must be reliable, sensitive, capable of screening large as well as small numbers of samples, and should be relatively unaffected by the composition of the sample. Enzyme immunoassays are well suited to such tests, and numerous formats, enzymes and substrates are employed (Cooper and Edwards, 1986; Nakamura *et al.*, 1986; Gow and

Williams, 1989; Torrance, 1992b). The format first introduced to plant virus detection by Clark and Adams (1977) was the direct double antibody sandwich (DAS) ELISA, and although it is probably the format still most commonly used, there are many variants. Standard DAS-ELISAs are sensitive enough to detect nanogram quantities of virus particles. In the original DAS-ELISA format using polyclonal antisera, antigen was trapped by antibody coated to microtitre plate wells and the antigen detected by virus-specific antibody conjugated to enzyme. In tests with MAbs, a triple antibody sandwich (TAS-ELISA) is usually used in which antibody-trapped antigen reacts with the MAb, which is detected by an anti-mouse (or rat) enzyme conjugate. However, MAbs may be directly coupled to biotin or enzyme (van Regenmortel and Dubs, 1993). Also, plates may or may not be precoated with virus-specific antibodies. The time taken per assay can be reduced by incubating enzyme-conjugated virus-specific antibodies and sample together in the wells of the plate (van Vuurde and Maat, 1985; van den Heuvel and Peters, 1989). Squash or dot blot tests can be done on individual leaf samples (taken in the field) by pressing leaf sap onto a support membrane (nitrocellulose, nylon or even notepaper) (Heide and Lange, 1988; Weidemann, 1988; Mitchell *et al.*, 1990).

The key point with all enzyme immunoassays is to optimize the tests for each pathogen, i.e. to determine the best dilutions of antibodies, enzyme conjugates, extraction buffers, blocking solutions and incubation times to ensure greatest absorbance values from the virus-containing samples with minimum absorbance from nonvirus control wells. The results achieved differ markedly with the quality of antibody, the solid phase and the test protocol used (Hewings and D'Arcy, 1984).

ELISA incorporating a MAb was shown to be the most accurate immunoassay for detecting bean common mosaic virus in *Phaseolus* seed of the US Department of Agriculture (USDA) germplasm collection (Klein and Wyatt, 1992). DAS-ELISA was used to assist detection and elimination of pea seed-borne mosaic virus (PSbMV) from the United States Department of Agriculture, Agricultural Research Service (USDA-ARS) *Pisum* germplasm collection (Hampton *et al.*, 1993), and was sensitive enough to detect PSbMV at 1/1000 of the concentration normally encountered in pea seeds. A nondestructive assay of soybean seeds has been devised to detect soybean mosaic virus (SMV) by ELISA of 3 mm cores of seed tissue (Higley *et al.*, 1993). This test was successful in part because SMV is detected in all parts of the seed, but other tests on dormant seed may be less reliable, for example when virus is confined to different tissues (seed coat or embryo).

Four viruses infecting potatoes grown in the Andes are known to be transmitted through true potato seed (TPS); potato virus T (Bolivian strain), tobacco ringspot virus, Andean potato calico strain (TRSV),

arracacha virus B (AVB), and Andean potato latent virus (Jones, 1982). TRSV was detected in individual dormant seeds produced on infected plants by DAS-ELISA using polyclonal antisera, although in similar tests AVB was only reliably detected in groups of ten seeds (Newton, 1989). Therefore in tests on TPS, the incorporation of MAb together with more sensitive substrates in ELISA, or application of immunocapture RT-PCR tests would probably give more reliable results.

Methods to increase sensitivity of ELISA

The introduction of specific, high-affinity MAbs exhibiting little or no non-specific binding should help to produce more sensitive and reproducible ELISAs because antibody specificity and affinity are probably the most significant factors affecting the sensitivity of a sandwich immunoassay (Porstmann *et al.*, 1985; Siddle, 1985). MAbs are especially useful in labelled-antibody assays because the labelled-antibody is present in excess and it is essential to have little or no nonspecific binding for maximum sensitivity (Siddle, 1985). Incorporation of MAbs together with the use of better enzyme conjugation methods (or better-quality second antibody conjugates) and fluorescent or chemiluminescent enzyme reaction products can speed up the assay and make it several orders of magnitude more sensitive (Ishikawa *et al.*, 1983; Porstmann *et al.*, 1985; Siddle, 1985; Jeanson *et al.*, 1988; Bronstein and McGrath, 1989; Gow and Williams, 1989). The microtitre plate ELISA can also be enhanced by amplifying the antibody-labelled alkaline phosphatase reaction with a second set of enzymes (Johannsson and Bates, 1988); this technique detected less than 100 000 molecules of thyroid-stimulating hormone per well (Johannsson and Bates, 1988). Amplified ELISA increased the sensitivity of detection of barley yellow dwarf virus (BYDV; Torrance, 1987) and potato leafroll virus (PLRV; van den Heuvel and Peters, 1989). For PLRV, the increased sensitivity over conventional direct ELISA (approximately 15-fold, detecting 50–100 pg per well) was achieved without any other modification such as the use of MAbs or different enzyme conjugation procedures, so the sensitivity could theoretically be improved further.

Production of novel antibody-like proteins by recombinant DNA technology

The production of MAbs can be a time-consuming and a relatively inefficient process. Sometimes many fusion experiments must be done before stable hybridomas secreting MAbs of the desired specificity are obtained. In future, the production of antibodies may be done by alternative methods utilizing recombinant DNA technology. This is possible because the advances in knowledge of antibody gene sequences together with PCR-assisted cloning have enabled the expression of fragments of antibody genes in bacterial systems (Orlandi *et al.*, 1989). In addition, functional antibody fragments can be expressed fused to the surface of filamentous

phage (phage-display; Clackson *et al.*, 1991; Marks *et al.*, 1992a). The fragments comprise the heavy and light chain variable (antigen-binding) domains of the antibody molecules linked by a short peptide to form a single polypeptide chain (scFv). Phage-display thus allows a specific scFv to be selected from a population of phages carrying many different scFvs by binding to and then eluting from the antigen. The eluted phage preparation can then be enriched for clones exhibiting high binding capacity by reinfecting *Escherichia coli* and repeating the procedure. In this way genetically pure populations of phage which encode the scFv can be obtained after several repeated cycles. Moreover, diverse synthetic antibody expression libraries have been constructed and used to obtain antibody fragments specific for a wide range of antigens (Hoogenboom and Winter, 1992; Nissim *et al.*, 1994) without the need to immunize animals.

Using a synthetic antibody gene library without recourse to immunization of experimental animals, Ziegler *et al.* (1995) have obtained an scFv specific for cucumber mosaic cucumovirus. Although the scFv did not perform as well as the polyclonal antiserum, giving a weaker signal when used to probe Western blots and in ELISA, the scFv was obtained after only four rounds of selection from a human synthetic phage display library and was used without any further modification. Monovalent scFvs obtained from such phage libraries are known to have moderate binding affinities, and several strategies have been suggested to improve affinity, such as mutation and chain shuffling (Marks *et al.*, 1992b; Winter *et al.*, 1994). In addition, the production of bivalent molecules has been shown to improve the avidity of an scFv (Holliger *et al.*, 1993; Pack *et al.*, 1993). It is now possible to obtain scFvs which bind with high affinity, by direct selection from very large libraries (Vaughan *et al.*, 1996). Furthermore, it is possible to obtain scFv genetically fused to reporter molecules such as alkaline phosphatase, or to biotin (biotin carboxy-carrier protein), hexahistidine or other peptide tags for detection and purification (Ducancel *et al.*, 1993; Weiss *et al.*, 1994). These improvements offer the prospect of producing recombinant antibody reagents much more cheaply in bacterial culture than by conventional hybridoma culture. Furthermore, recombinant technology can produce enzyme-linked antibodies without chemical coupling, which can be detrimental to activity. Therefore, production of recombinant antibodies by selection from large expression libraries has the potential to replace conventional methods of antibody production involving animal immunization and culture of hybridoma cells.

10.4. Conclusions

Biotechnology can provide a range of rapid, sensitive methods for detecting and diagnosing viruses without the need for bioassays and continued

propagation of tissue. Future advances should aim to develop broad-spectrum tests capable of detecting the majority of viruses in a group. Some advances have been made in this direction, for example the use of degenerate PCR primers for potyviruses (Langeveld *et al.*, 1991) and antibodies which can detect group-specific epitopes, e.g. potyvirus MAbs (Jordan and Hammond, 1991). These tests may be useful in detecting viruses which are undescribed but related to other well-characterized viruses in a group. However, broad-spectrum detection methods may only be applicable to certain virus groups and there are likely to be many viruses for which specific detection methods must be developed. Indeed, although laboratory-based methods can give confidence in the detection of well-characterized viruses and those related to established groups, other less well-known viruses may continue to pose a risk. Importers should monitor material after release from quarantine for unusual or transient symptoms, while bearing in mind that some virus-like symptoms may have a physiological basis.

It is becoming evident that the greater application of biotechnology to virus detection will eventually result in rapid, more uniform, standardized tests that all importers can have confidence in. Thus biotechnology has much to offer for the maintenance, utilization and health of germplasm collections.

References

- Abou-Elnasr, M.A., Jones, A.T. and Mayo, M.A. (1985) Detection of dsRNA in particles of vicia cryptic virus and in *Vicia faba* tissues and protoplasts. *Journal of General Virology* 66, 2453–2460.
- Arif, M., Torrance, L. and Reavy, B. (1994) Improved efficiency of detection of potato mop-top furovirus in potato tubers and in the roots and leaves of soil-bait plants. *Potato Research* 37, 373–381.
- Banttari, E.E., Clapper, D.L., Hu, S.-P., Daws, K.M. and Khurana S.M.P. (1991) Rapid magnetic microsphere enzyme immunoassay for potato virus X and potato leafroll virus. *Phytopathology* 81, 1039–1042.
- Bariana, H.S., Shannon, A.L., Chu, P.W.G. and Waterhouse, P.M. (1994) Detection of five seedborne legume viruses in one sensitive multiplex polymerase chain reaction test. *Phytopathology* 84, 1201–1205.
- Barker, H., Webster, K.D. and Reavy, B. (1993) Detection of potato virus Y in potato tubers: a comparison of polymerase chain reaction and enzyme linked immunosorbent assay. *Potato Research* 36, 13–20.
- Blok, V.C., Ziegler, A., Scott, K., Dangora, D.B., Robinson, D.J. and Murrant, A.F. (1995) Detection of groundnut rosette umbravirus infections with radioactive and non-radioactive probes to its satellite RNA. *Annals of Applied Biology* 127, 321–328.
- Borkhardt, B., Vongsasitorn, D. and Albrechtsen, S.E. (1994) Chemiluminescent

- detection of potato spindle tuber viroid in true potato seed using a digoxigenin labelled DNA probe. *Potato Research* 37, 249–255.
- Bronstein, I. and McGrath, P. (1989) Chemiluminescence lights up. *Nature* 338, 599–600.
- Brunt, A.A., Crabtree, K., Dallwitz, M.J., Gibbs, A.J. and Watson, L. (1996) *Viruses of Plants*. CAB International, Wallingford, UK, 1488 pp.
- Brush, S.B. (1995) In situ conservation of landraces in centers of crop diversity. *Crop Science* 35, 346–354.
- Campbell, A.M. (1984) *Monoclonal Antibody Technology*. Elsevier, Amsterdam, 265 pp.
- Clackson, T., Hoogenboom, H.R., Griffiths, A.D. and Winter, G. (1991) Making antibody fragments using phage display libraries. *Nature* 352, 624–628.
- Clark, M.F. and Adams, A.N. (1977) Characteristics of the microplate method of enzyme-linked immunosorbent assay for the detection of plant viruses. *Journal of General Virology* 34, 475–483.
- Cooper, J.I. and Edwards, M.L. (1986) Variations and limitations of enzyme-amplified immunoassays. In: Jones, R.A.C. and Torrance, L. (eds) *Developments and Applications in Virus Testing*. Association of Applied Biologists, Wellesbourne, pp. 139–154.
- Diekmann, M. and Putter, C.A.J. (eds) (1995) *FAO/IPGRI Technical Guidelines for the Safe Movement of Germplasm* (e.g. No. 14: *Small Grain Temperate Cereals*). Food and Agriculture Organization of the United Nations, Rome/International Plant Genetic Resources Institute, Rome.
- Dodds, J.A., Morris, T.J. and Jordan, R.L. (1984) Plant viral double-stranded RNA. *Annual Review of Phytopathology* 22, 151–168.
- Ducancel, F., Gillet, D., Carrier, A., Lajeunesse, E., Menez, A. and Boulain, J-C. (1993) Recombinant colorimetric antibodies: construction and characterization of a bifunctional F(ab)₂/alkaline phosphatase conjugate produced in *Escherichia coli*. *Bio/Technology* 11, 601–605.
- Goding, J.W. (1983) *Monoclonal Antibodies: Principles and Practice*. Academic Press, London, 276 pp.
- Gould, A.R. and Francki, R.I.B. (1981) Immunochemical detection of ds-RNA in healthy and virus-infected plants and specific detection of viral ds-RNA by hybridization to labelled complementary DNA. *Journal of Virological Methods* 2, 277–286.
- Gow, I.F. and Williams, B.C. (1989) Immunoassays for antigens. *Current Opinion in Immunology* 1, 940–947.
- Griffiths, H.M., Slack, S.A. and Dodds, J.H. (1990) Effect of chemical and heat therapy on virus concentrations in *in vitro* potato plantlets. *Canadian Journal of Botany* 68, 1515–1521.
- Habili, N., Krake, L.R., Barlass, M. and Rezaian, M.A. (1992) Evaluation of biological indexing and dsRNA analysis in grapevine virus elimination. *Annals of Applied Biology* 121, 277–283.
- Hampton, R.O., Kraft, J.M. and Muehlbauer, F.J. (1993) Minimizing the threat of seedborne pathogens in crop germ plasm: elimination of pea seedborne mosaic virus from the USDA-ARS germ plasm collection of *Pisum sativum*. *Plant Disease* 77, 220–224.

- Harlow, E. and Lane, D. (1988) *Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 726 pp.
- Harrison, B.D. and Robinson, D.J. (1981) Tobraviruses. In: Kurstak, E. (ed.) *Handbook of Plant Virus Infections and Comparative Diagnosis*. Elsevier / North Holland, Amsterdam, pp. 515–540.
- Hataya, T., Inoue, A.K. and Shikata, E. (1994) A PCR-microplate hybridization method for plant virus detection. *Journal of Virological Methods* 46, 223–236.
- Heide, M. and Lange, L. (1988) Detection of potato leafroll virus and potato viruses M, S, X, and Y by dot immunobinding on plain paper. *Potato Research* 31, 367–373.
- Hewings, A.D. and D'Arcy, C.J. (1984) Maximizing the detection capability of a beet western yellows virus ELISA system. *Journal of Virological Methods* 9, 131–142.
- Higley, P.M., McGee, D.C. and Burris, J.S. (1993) Development of methodology for non-destructive assay of bacteria, fungi and viruses in seeds of large-seeded field crops. *Seed Science and Technology* 21, 399–409.
- Holliger, P., Prospero, T.D. and Winter, G. (1993) Diabodies: small bivalent and bispecific antibody fragments. *Proceedings of the National Academy of Sciences of the USA* 90, 6444–6448.
- Hoogenboom, H.R. and Winter, G. (1992) By-passing immunization. Human antibodies from synthetic repertoires of germline VH gene segments rearranged *in vitro*. *Journal of Molecular Biology* 227, 381–388.
- Hu, J.S., Gonsalves, D., Boscia, D., Maixner, M. and Golino, D. (1991) Comparison of rapid detection assays for grapevine leafroll disease associated closteroviruses. *Vitis* 30, 87–95.
- Ishikawa, E., Imagawa, M., Hashida, S., Yoshitake, S., Hamaguchi, Y. and Ueno, T. (1983) Enzyme-labeling of antibodies and their fragments for enzyme immunoassay and immunohistochemical staining. *Journal of Immunoassay* 4, 209–327.
- Jansen, R.W., Siegl, G. and Lemon, S.M. (1990) Molecular epidemiology of human hepatitis A virus defined by an antigen-capture polymerase chain reaction method. *Proceedings of the National Academy of Sciences of the USA* 87, 2867–2871.
- Jeanson, A., Cloes, J.-M., Bouchet, M. and Rentier, B. (1988) Comparison of conjugation procedures for the preparation of monoclonal antibody–enzyme conjugates. *Journal of Immunological Methods* 111, 261–270.
- Johannsson, A. and Bates, D.L. (1988) Amplification by second enzymes. In: Kemeny, M. and Challacombe, S.J. (eds) *ELISA and Other Solid Phase Immunoassays*. John Wiley, Chichester, pp. 85–106.
- Jones, A.T. (1992) Application of double-stranded RNA analysis of plants to detect viruses, virus-like agents, virus satellites and subgenomic viral RNAs. In: Duncan, J.M. and Torrance, L. (eds) *Techniques for the Rapid Detection of Plant Pathogens*. Blackwell Scientific Publications, Oxford, pp. 115–128.
- Jones, A.T., Abou-Elnasr, M.A., Mayo, M.A. and Hodgkin, J.R.T. (1986) Detection of viruses in brassicas. In: *Report of the Scottish Crop Research Institute for 1985*. SCRI, Dundee, pp. 169–170.
- Jones, R.A.C. (1982) Tests for transmission of four potato viruses through potato true seed. *Annals of Applied Biology* 100, 315–320.

- Jordan, R. (1992) Potyviruses, monoclonal antibodies, and antigenic sites. *Archives of Virology* suppl. 5, 81–95.
- Jordan, R. and Hammond, J. (1991) Comparison and differentiation of potyvirus isolates and identification of strain-, virus-, subgroup-specific and potyvirus group-common epitopes using monoclonal antibodies. *Journal of General Virology* 72, 25–36.
- Kanematsu, S., Hibi, T., Hashimoto, J. and Tsuchizaki, T. (1991) Comparison of nonradioactive cDNA probes for detection of potato spindle tuber viroid by dot-blot hybridization assay. *Journal of Virological Methods* 35, 189–197.
- Klein, R.J. and Wyatt, S.D. (1992) Comparative immunoassays of bean common mosaic virus in individual bean (*Phaseolus vulgaris*) seed and bulked bean seed samples. *Plant Disease* 76, 57–59.
- Kunkel, L.O. (1936) Heat treatments for the cure of yellows and other virus diseases of peach. *Phytopathology* 26, 809–830.
- Kwoh, D.Y., Davis, G.R., Whitfield, K.M., Chappelle, H.L., DiMichele, L.J. and Gineras, T.R. (1989) Transcription-based amplification system and detection of amplified human immunodeficiency virus type 1 with bead-based sandwich hybridization format. *Proceedings of the National Academy of Sciences of the USA* 86, 1173–1177.
- Langeveld, S.A., Dore, J.-M., Memelink, J., Derks, A.F.L.M., van der Vlugt, C.I.M., Asjes, C.J. and Bol, J.F. (1991) Identification of potyviruses using the polymerase chain reaction with degenerate primers. *Journal of General Virology* 72, 1531–1541.
- Marks, J.D., Hoogenboom, H.R., Griffiths, A.D. and Winter, G. (1992a) Molecular evolution of proteins on filamentous phage. *Journal of Biological Chemistry* 267, 16007–16010.
- Marks, J.D., Griffiths, A.D., Malmqvist, M., Clackson, T.P., Bye, J.M. and Winter, G. (1992b) By-passing immunization: building high affinity human antibodies by chain shuffling. *Bio/Technology* 10, 779–783.
- Matthews, R.E.F. (1991) *Plant Virology*, 3rd edn. Academic Press, London, 835 pp.
- Mitchell, D.H., Howell, P.J. and Rose, D.G. (1990) Squash blots – A novel device for sampling plant viruses. *Potato Research* 33, 144–145.
- Nakamura, R.M., Voller, A. and Bidwell, D.E. (1986) Enzyme immunoassays: heterogeneous and homogeneous systems. In: Weir, D.M., Herzenberg, L.A., Blackwell, C. and Herzenberg, L.A. (eds) *Handbook of Experimental Immunology*, vol. 1. Blackwell Scientific Publications, Oxford, pp. 27.1–27.20.
- Nameth, S.T. and Dodds, J.A. (1985) Double-stranded RNAs detected in cucurbit varieties not inoculated with viruses. *Phytopathology* 75, 1293.
- Newton, E.L. (1989) Studies towards the control of viruses transmitted through true potato seed. PhD Thesis, University of Birmingham, 212 pp.
- Nissim, A., Hoogenboom, H.R., Tomlinson, I.M., Flynn, G., Midgley, C., Lane, D. and Winter, G. (1994) Antibody fragments from a ‘single pot’ phage display library as immunochemical reagents. *EMBO Journal* 13, 692–698.
- Orlandi, R., Güssow, D.H., Jones, P.T. and Winter, G. (1989) Cloning immunoglobulin variable domains for expression by the polymerase chain reaction. *Proceedings of the National Academy of Sciences of the USA* 86, 3833–3837.
- Owens, R.A. and Diener, T.O. (1981) Sensitive and rapid diagnosis of potato spindle tuber viroid disease by nucleic acid hybridisation. *Science* 213, 670–672.

- Pack, P., Kujau, M., Schroeckh, V., Knüpfer, U., Wenderoth, R., Riesenberger, D. and Plückthun, A. (1993) Improved bivalent miniantibodies, with identical avidity as whole antibodies, produced by high cell density fermentation of *Escherichia coli*. *Bio/Technology* 11, 1271–1276.
- Paduch-Cichal, E. and Kryczynski, S. (1987) A low temperature therapy and meristem-tip culture for eliminating four viroids from infected plants. *Journal of Phytopathology* 118, 341–346.
- Palukaitis, P., Rakowski, A.G., Alexander, D.McE. and Symons, R.H. (1981) Rapid indexing of the sunblotch disease of avocados using a complementary DNA probe to avocado sunblotch viroid. *Annals of Applied Biology* 98, 439–449.
- Porstmann, B., Porstmann, T., Nugel, E. and Evers, U. (1985) Which of the commonly used marker enzymes gives the best results in colorimetric and fluorimetric enzyme immunoassays: horseradish peroxidase, alkaline phosphatase or β -galactosidase? *Journal of Immunological Methods* 79, 27–37.
- Roberts, I.M. (1986) Practical aspects of handling, preparing and staining samples containing plant virus particles for electron microscopy. In: Jones, R.A.C. and Torrance, L. (eds) *Developments and Applications in Virus Testing*. Association of Applied Biologists, Wellesbourne, pp. 213–243.
- Robinson, D.J. and Dale, M.F.B. (1994) Susceptibility, resistance and tolerance of potato cultivars to tobacco rattle virus infection and spraing disease. *Aspects of Applied Biology* 39, 61–66.
- Rosner, A., Stein, A. and Levy, S. (1992) Transcription amplification of polymerase chain reaction products of bean yellow mosaic virus RNA extracted from gladioli corms. *Annals of Applied Biology* 121, 269–276.
- Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239, 487–491.
- Sáiz, M., Castro, S., De Blas, C. and Romero, J. (1994) Serotype-specific detection of bean common mosaic potyvirus in bean leaf and seed tissue by enzymatic amplification. *Journal of Virological Methods* 50, 145–154.
- Siddle, K. (1985) Properties and applications of monoclonal antibodies. In: Collins, W.P. (ed.) *Alternative Immunoassays*. John Wiley, Chichester, pp. 13–37.
- Singh, M. and Singh, R.P. (1995) Digoxigenin-labelled cDNA probes for the detection of potato virus Y in dormant potato tubers. *Journal of Virological Methods* 52, 133–143.
- Spiegel, S. and Martin, R.R. (1993) Improved detection of potato leafroll virus in dormant potato tubers and microtubers by the polymerase chain reaction and ELISA. *Annals of Applied Biology* 122, 493–500.
- Torrance, L. (1987) Use of enzyme amplification in an ELISA to increase sensitivity of detection of barley yellow dwarf virus in oats and in individual vector aphids. *Journal of Virological Methods* 15, 131–138.
- Torrance, L. (1992a) Developments in methodology of plant virus detection. *Netherlands Journal of Plant Pathology* 98 (suppl. 2), 21–28.
- Torrance, L. (1992b) Serological methods to detect plant viruses: production and use of monoclonal antibodies. In: Duncan, J.M. and Torrance, L. (eds) *Techniques for the Rapid Detection of Plant Pathogens*. Blackwell Scientific Publications, Oxford, pp. 7–33.
- van den Heuvel, J.F.J.M. and Peters, D. (1989) Improved detection of potato leafroll

- virus in plant material and in aphids. *Phytopathology* 79, 963–967.
- van Regenmortel, M.H.V. and Dubs, M.-C. (1993) Serological procedures. In: Matthews, R.E.F. (ed.) *Diagnosis of Plant Virus Diseases*. CRC Press, London, pp. 159–214.
- van Vuurde, J.W.L. and Maat, D.Z. (1985) Enzyme-linked immunosorbent assay (ELISA) and disperse-dye immunoassay (DIA): comparison of simultaneous and separate incubation of sample and conjugate for the routine detection of lettuce mosaic virus and pea early-browning virus in seeds. *Netherlands Journal of Plant Pathology* 91, 3–13.
- Vaughan, T.J., Williams, A.J., Pritchard, K., Osbourn, J.K., Pope, A.R., Earnshaw, J.C., McCafferty, J., Hodits, R.A., Wilton, J. and Johnson, K.S. (1996) Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library. *Bio/Technology* 14, 309–314.
- Vunsh, R., Rosner, A. and Stein, A. (1990) The use of the polymerase chain reaction (PCR) for the detection of bean yellow mosaic virus in gladiolus. *Annals of Applied Biology* 117, 561–569.
- Wakarchuk, D.A. and Hamilton, R.I. (1985) Cellular double-stranded-RNA in *Phaseolus vulgaris*. *Plant Molecular Biology* 5, 55–63.
- Webster, K.D. and Barker, H. (1997) Detection and quantification of transcript RNA in transgenic plants using digoxigenin-labelled cDNA probes. In: Foster, G. and Taylor, S. (eds) *Plant Virology Protocols: From Virus Isolation to Transgenic Resistance (Methods in Molecular Biology)*. Humana Press, (in press).
- Weidemann, H.-L. (1988) Rapid detection of potato viruses by dot-ELISA. *Potato Research* 31, 485–492.
- Weiss, E., Chatellier, J. and Orfanoudakis, G. (1994) *In vivo* biotinylated recombinant antibodies: construction, characterization, and application of a bifunctional Fab-BCCP fusion protein produced in *Escherichia coli*. *Protein Expression and Purification* 5, 509–517.
- Wetzel, T., Candresse, T., Macquaire, G., Ravelonandro, M. and Dunez, J. (1992) A highly sensitive immunocapture polymerase chain reaction method for plum pox potyvirus detection. *Journal of Virological Methods* 39, 27–37.
- Winter, G., Griffiths, A.D., Hawkins, R.E. and Hoogenboom, H.R. (1994) Making antibodies by phage display technology. *Annual Review of Immunology* 12, 433–455.
- Wong, C., Dowling C.E., Saiki, R.K., Higuchi, R.G., Erlich, H.A. and Kazazian, H.H. (1987) Characterization of β -thalassaemia mutations using direct genomic sequencing of amplified single copy DNA. *Nature* 330, 384–386.
- Ziegler, A., Torrance, L., Macintosh, S.M., Cowan, G.H. and Mayo, M.A. (1995) Cucumber mosaic cucumovirus antibodies from a synthetic phage display library. *Virology* 214, 235–238.

Biodiversity for Bioindustries

G. Tamayo, W.F. Nader and A. Sittenfeld

*Instituto Nacional de Biodiversidad (INBio), Apartado Postal
22-3100, Santo Domingo, Heredia, Costa Rica*

11.1. Biodiversity and its Benefits

Biodiversity refers to the variety and variability of living material and ecological complexes in a given area and comprises species, genetic and ecosystem diversity. Biodiversity is not only the basis of life on earth, but also provides the goods and services essential to support every type of human endeavour. Accordingly, biodiversity enables societies to adapt to different needs and situations (US National Research Council, 1992).

Biodiversity generates economic value in different ways. Populations are interconnected, for instance where predators and disease organisms control populations of their prey, or when pollinators and seed dispensers promote the growth of plant populations. Thus agriculture directly benefits from a functioning ecosystem, allowing the extensive use of agrochemicals to be avoided. Biodiversity also generates economic value from extractable products obtained from individual species (Wilson, 1992). For centuries biodiversity has provided fuels, medicines, materials for shelter, food and energy. The use of compounds, genes and species is essential to meet industry needs. Furthermore, ecosystems contribute to climate regulation, maintenance of hydrological cycles and nitrification of soils. In addition, recreation, science and education also figure among the vast array of social, ethical, spiritual, cultural and economic goods and services provided by biodiversity that are recognized as fundamental for human livelihoods and aspirations.

Bioprospecting links biodiversity and industry. Previously, this activity generated benefits almost exclusively for industry, leaving biodiversity

conservation and source countries to generate benefits and returns elsewhere. The rapid loss of biological diversity, with the extinction of 30 to 300 species per day (Japan Economic Newswire, 1995), has initiated a new attitude towards the exploration of natural resources. Costa Rica's Instituto Nacional de Biodiversidad (the National Biodiversity Institute, INBio) has pioneered a new concept of bioprospecting that integrates product discovery with financial and intellectual returns to 'nature'. INBio's Biodiversity Prospecting Department links the understanding and non-damaging exploration of biodiversity to conservation activities and economic development of the countries where bioresources were first obtained (Sittenfeld and Villers, 1994). The exploration and conservation of the world's biotic resources require an approach involving bioindustries, research centres and developing countries, all collaborating towards a common goal, each participant benefiting from the relationship. Presently, a natural resource conservation strategy based on the three overlapping steps – saving, knowing and using biodiversity – is paving the way towards implementing joint activities for the benefit of industry, biodiversity conservation and source countries (*Global Biodiversity Strategy*, 1992; Janzen and INBio, 1992).

Gene technology opens a new dimension for bioprospecting. However, because in its current stage of development it represents more of a threat than a benefit to the primarily agricultural societies of the developing world, new strategies must be implemented to combine benefits to the biotech industry with biodiversity conservation and the development of biodiversity-rich nations.

11.2. Industry and Biodiversity

11.2.1. Development of drugs and pesticides

Up until the present, the primary beneficiaries of biodiversity have been the pharmaceutical and agricultural industries. Sales of drugs based on natural products from plants were estimated at \$US43 billion in 1985 worldwide, accounting for approximately 40% of the total drug market (Principe, 1989). Earnings from a single new successful pharmaceutical on the market can be in the range of a billion dollars. The value of yet undiscovered pharmaceuticals in tropical forests is estimated at \$US3–4 billion for a private pharmaceutical company, and as much as \$US47 billion to society as a whole (Mendelsohn and Balick, 1995).

On the other hand, sales of pesticides in the US by 18 major suppliers were estimated at \$US6.5 billion in 1992 (Frost & Sullivan Inc., 1992). Although humans have traditionally used plant products like rotenone or nicotine as agricultural insecticides, most industrial pesticides on the

market have not been derived from natural compounds (Wink, 1993). However, the toxicity and ecological hazards created by most of these chemicals have initiated intensive research for more specific and biodegradable pesticides, including new biological pest control agents based on natural compounds and microorganisms. The first results of these screening efforts are already on the market. For example, avermectins are macrolides derived from *Streptomyces avermitilis* and possessing insecticidal properties (Babu, 1988). Dehydration of avermectins yields the even more efficient ivermectins, which generate approximately \$US1 billion in annual sales, and are useful not only for the treatment of infestations of parasitic worms and insects in livestock, but also in humans. Pyrethroids, on the other hand, were derived by chemical synthesis from pyrethrins as lead structures as natural insecticides found in chrysanthemum flowers (Wink, 1993). Pyrethrins are more stable and active than their natural precursors (by a factor of 1000), but also produce more side-effects in mammals. These side-effects have led Germany to consider banning pyrethrins from the market (Wink, 1993). Still undergoing development is a nematocide, the pyrrolidine alkaloid 2*R*,5*R*-dihydroxymethyl-3*R*,4*R*-dihydroxypyrrolidine from *Lonchocarpus* species, developed through a collaboration between the British Technology Group, the Royal Botanical Garden at Kew and INBio (Janzen *et al.*, 1990; Birch *et al.*, 1993).

Many of the technical aspects of screening and developmental processes involved in creating new drugs or pesticides are related to more general issues like biodiversity, technology transfer and biodiversity conservation, and will be described in more detail below.

Sources for new drugs or pesticides

Three major sources for the screening of new compounds, suitable for drug or pesticide development, are available: (i) molecule libraries created by combinatorial chemistry; (ii) fermentation broths of microorganisms; and (iii) plant and animal extracts. Modern automated methods have created high-throughput screening, allowing thousands of substances to be tested for their biological activity rapidly and inexpensively. Therefore, access to these three sources can be regarded as a prerequisite rather than a privilege. The goal is the discovery of a 'leading structure' that will guide the development of new drugs, pesticides or fine chemicals.

Although combinatorial chemistry plays an increasingly important role, half of the ten best-selling drugs are derived from secondary metabolites originally isolated from microorganisms or plants (O'Neill and Lewis, 1993). Obviously, organic chemistry has not yet caught up with the capacity of nature to create new structures with a complex molecular diversity (for review see Ecker and Crooke, 1995).

Screening for natural compounds has traditionally concentrated on plants and microorganisms, identifying a huge variety of pharmacologically

active alkaloids, terpenoids, aromatics and glycosides. Fungi and bacteria can easily be isolated from soil samples and other sources. Once in a strain collection, microorganisms and their products are readily accessible. Plants, on the other hand, are more difficult to collect, but offer a higher molecular complexity and diversity. Marine organisms yield new structures with high molecular diversity and important biological activities (König *et al.*, 1994). For example, briostatin and didemnin B are compounds found in molluscs and shown to have strong antitumour activity; they have reached preclinical and clinical trials, respectively. Shark and tunicate alkaloids are also currently undergoing intensive investigations (Moore *et al.*, 1993; Research Foundation of the State University of New York, 1994). In contrast, insects, spiders and other invertebrates have mainly been examined for their potential to produce bioactive peptides and proteins rather than small molecules (see below). The same applies to vertebrates like frogs, snakes and bats. Certain alkaloids obtained from the skin of frogs (see below), and insect hormones and pheromones useful for pest control are the exceptions. Nevertheless, drug research is also heading in this direction. In collaboration with Merck & Co. and Bristol-Myers Squibb (within an International Cooperative Biodiversity Group together with Cornell University), INBio is screening insects for small bioactive compounds (Sittenfeld and Lovejoy, 1995).

However, years of research may fail if the initial collection and documentation of biological material are not done properly. Problems may arise if further material from the same species or subspecies, from the same environment or even from the same location, is not available for later investigation. Calanolides, for example, were isolated from *Calophyllum lanigerum* var. *austrocoriaceum* (Kashman *et al.*, 1992). These compounds inhibit HIV *in vitro*. Although fully characterized (Cardellina *et al.*, 1993a) material from the same tree is now unavailable because the tree was subsequently cut down. There is a lesson to be learned from this event as other specimens from the same area did not yield even trace amounts of the desired compounds.

The collection strategy

Natural compounds can be accessed through ethnobiological information, evaluation of chemotaxonomic relationships, random sampling or biorational observations.

Approximately 3000 million people use traditional medicines (Balick, 1994). Many of today's multinational giants in the pharmaceutical industry made their first millions with products derived from ethnobotany. Prior to Bayer's aspirin, American and Eurasian peoples treated fevers, inflammation and pain with salicin-containing plants like willows and poplars. The worldwide market for phytomedicines derived from ethnobotany is estimated at \$US12.4 billion, headed by products derived from ginseng, ginkgo, garlic, horse chestnut and echinacea (Grünwald, 1995). Today

ethnobotanical information is readily accessible through Internet's databases, e.g. NAPRALERT or AGIS. But most companies avoid the search for new compounds on the basis of ethnobiological information. An evaluation of 'hits' during the Natural Products Drug Discovery Program at the National Cancer Institute revealed no appreciable differences between samples collected at random and those screened on the basis of ethnobotanical leads (Cragg *et al.*, 1994). These results were obtained after dereplication, which eliminates a substantial number of substances. Interpretation of illness descriptions from the 'native pharmacologists' is also problematic. However, ethnobotanical information can be extremely useful when applied to diseases that *can* be translated into the language of Western medicine, e.g. diabetes, skin infections, wounds, etc. Aspects of intellectual property rights in relation to this approach are discussed in more detail in a recent study of the Rural Advancement Foundation International (1994).

Chemotaxonomy is based on the assumption that related species from the same genus or family produce the same type of secondary metabolites. For example, most members of the *Asteraceae* produce sesquiterpenelactones. An example is artemisinin, isolated from *Artemisia annua* on the basis of ethnobotanical information. Likewise, members of the *Rubiaceae* may produce alkaloids (quina alkaloids, for example), and species of the genus *Taxus* may contain taxanes, etc.

Most large pharmaceutical companies, not limited by their screening capacity, collect biological material randomly. They may show a preference for certain plant families, but in general only the taxonomic identification and exact documentation of the collection site of the sample are required.

The biorational approach requires the systematic study of interactions between organisms within the ecosystem. The leaf-cutter ant, *Atta cephalotes*, for example, avoids feeding the symbiotic fungus found in its nest with leaves from *Hymenaea courbaril* (Harborne, 1989). The tree contains a terpenoid (caryophyllene epoxide) that inhibits the growth of the fungus. Observations like this may provide decisive information leading to the discovery of antifungal or formicide compounds. Evaluation of larvae or adult insect feeding preferences may lead to the discovery of new insecticides.

Technological aspects

The amount of material that can be taken from a particular environment without causing damage is a primary concern for biodiversity conservation. Bioassays can be performed on 10–100 mg of a compound mixture, the usual yield from 10 g of dry plant material. For confirmatory assays and further fractionation and isolation, amounts in the range 100–1000 g must be collected. Depending on a given compound's yield, preclinical trials to evaluate toxic side-effects and effectiveness in animals may require

large amounts (tonnes) of dried plant material if the compound is too complex to be synthesized. There is no doubt that the survival of certain species has been threatened in the past by drug researchers. The exploitation of pilocarpine, for example, threatened the survival of *Pilocarpus pignatifolios*, *P. microphylla* and *P. jaborandi* species in South America (Balick, 1994). In another case, clinical trials with taxol have affected the survival of *Taxus brevifolia* in its natural habitat.

When considering the possibility of countering this danger through sustainable breeding or planting, one must keep in mind that cultivation of a desired species may lead to a loss of the desired compound. The isolation and identification of epibatidine from *Epipedobates tricolor*, a frog used by indigenous people to poison arrowheads, has been described (Sapnde *et al.*, 1992). Epibatidine is a remarkably simple, but highly efficient alkaloid that is 200 to 500 times more potent than morphine in analgesic assay systems. However, investigation of the alkaloid required the skins of 750 frogs collected from the wild. Attempts to breed these frogs in an artificial environment led to a loss of epibatidine in the skin of the second generation of frogs. Therefore, the compound is probably produced by complex interactions with other organisms in the environment.

Extraction of plant material is a philosophy in its own right. An efficient protocol for organic extraction was developed at the National Cancer Institute (McCloud *et al.*, 1988). The decision whether to use raw extracts, prepurified fractions or even randomly isolated pure compounds for bioassay depends on a drug company's philosophy. Depending on the bioassay, natural compounds like polysaccharides, tannins, saponins or fatty acid esters must be removed prior to testing. This is of particular importance for bioassays involving cell cultures (Cardellina *et al.*, 1993b).

In recent years, drug bioassays have become increasingly specific, rapid, reproducible and sensitive. At the same time, they have become less susceptible to matrix and other effects that tend to cause false positive or negative reactions. Even during the 1970s, companies were screening with receptors isolated from cell cultures, a technique that yielded various top-selling drugs. Today, gene technology allows the cloning and expression of receptors, enzymes and other proteins important for signal transduction, metabolic conversions, or cell or viral structures. Once produced on a large scale, they can be immobilized on ELISA (enzyme-linked immunosorbent assay) plates and integrated into a chromogenic assay to measure ligand-receptor interactions. The end result is the introduction of a huge variety of new assays, which has increased the industry's screening activities and the overall demand for new compounds. Nevertheless, the development of drugs against diseases not investigated down to the molecular level still requires complex cell culture assays or even screening in animal models. This is the case for most types of anticancer drug.

Drug development, from the collection and taxonomic classification of

biological material to compound extraction and fractionation, bioassays, structure elucidation and final clinical testing, is usually not performed by a single drug company, but by various research partners, working under contract. Therefore, the transfer of related technology to developing-world source countries is a possible, and even logical, approach for the technological development of these regions. Collection, taxonomic classification and extraction are not trivial tasks, but require a high standard of documentation and reliable and reproducible work that can easily be conducted in Southern nations with the aid of technology transfer. Drug companies would benefit from carrying out further bioassaying and chemical structure elucidation directly in the source countries, not only because of developmental issues and cheaper labour costs, but also because it would ensure these companies gained more direct access to important markets.

11.2.2. Prospecting for genes

Genetic engineering opens a totally new dimension for bioprospecting. The search for new genes and their applications is the primary objective of the biotech industry. Today's biotech products, already on the market, are based on genes from humans, domesticated animals and cultivated plants. Examples are human cytokines and growth factors like interferons, colony-stimulating factors, erythropoietin and bovine somatotropin, and also Calgene's Sav-R-Flavr tomato. The market value of erythropoietin alone amounts to several billion dollars per year worldwide. Hence the tremendous appetite of this new industry for novel genes; indeed, the hunt for them in tropical rainforests is already on. In contrast to random screening in natural compound research (see above), gene technology allows a more straightforward approach, as illustrated by the following examples.

The pharmaceutical biotech industry and biodiversity

Biodiversity and protein engineering. Minor changes to the amino acid sequences of pharmaceutical proteins and peptides (biologics) and enzymes may lead to new or improved activities. Computer simulation in combination with site-directed mutagenesis are the basis for a new technology, called protein engineering, which creates its own 'molecular diversity'. Nevertheless, the first successful examples of amino acid sequence improvement are the result of screening genes from the wild. Calcitonin is a peptide hormone that inhibits the release of calcium ions and phosphate from the bones, and has therapeutic uses for osteoporosis (MacIntyre *et al.*, 1987). The investigation of related hormones from animals revealed that the calcitonin from salmon is more active and has a longer half-life within the human body than the human peptide structure

(Epand *et al.*, 1986). Today, chemically synthesized 'salcatonin' is on the market under tradenames including Calsynar and Miacalcic.

Industrial enzymes, used to catalyse chemical processes, can be improved to increase their heat stability, activity and specificity. Naturally thermophilic bacteria have become a useful source of industrial enzymes. Hydantoinase, for example, catalyses the conversion of chemically synthesized hydantoins to precursors of D-amino acids. D-Amino acids, like D-hydroxyphenylglycine and D-phenylglycine, are needed to derive amoxicillin and ampicillin from penicillin, just as D-serine is a precursor for pesticide production. The first enzymes to be used for this conversion on an industrial scale were isolated from common soil bacteria. The characteristics of these enzymes limited the maximal temperature for the catalysis to 40°C (Kanegefuchi Co., 1978; Yamada *et al.*, 1978), conditions which do not permit hydantoins to dissolve well in water. Researchers at BASF screened thermophilic bacteria from Yellowstone geysers and found two new hydantoinases with much improved heat stability and specificity characteristics (BASF AG, 1987). Through recombinant DNA technology, these enzymes are now produced in *Escherichia coli* and already on the market. The catalytic process can be performed more efficiently and more competitively at temperatures reaching 75°C.

Protein engineering based on computer simulation is doubtless a very powerful tool. However, the screening of natural products for improved principles still has a higher success rate, proving again that the computer cannot yet rival Mother Nature.

Animal defence and attack mechanisms as a source for biologics. For several decades now spider, snake, frog and bee venoms and squid and leech salivas have been investigated for pharmaceutically active peptides and proteins. Leeches (e.g. *Hirudo medicinalis*) have been used in traditional medicine to treat thrombosis since ancient times. The active principle from their saliva, the protein hirudin is now an ingredient of numerous ointments and gels and thus used against varicosis and haemorrhoids. Hirudin was one of the first proteins isolated from wild biodiversity. Recombinant hirudin has now been produced in *Escherichia coli* (Fortkamp *et al.*, 1986). Other leech species are currently under investigation to discover new hirudin variants with improved therapeutic applications (Sacheri *et al.*, 1993). Promotion of blood clotting during wound healing can be achieved using proteins from snake venom (e.g. from the Egyptian sand viper) that induce platelet aggregation (Baheer *et al.*, 1995).

The mammalian enzyme tissue plasminogen activator (tPA) dissolves thrombotic blood clots. Recombinant human tPA, developed and patented by Genentech, has been approved as a therapeutic agent against heart attack in the USA and Europe. Researchers at Schering AG found four similar proteins in the saliva of *Desmodus rotundus*, the common vampire bat,

that are more efficient and safer for therapeutic application than their human counterparts (Schleuning *et al.*, 1992). These products are presently undergoing preclinical studies.

Eledoisin is a hendecapeptide isolated from the salivary gland of certain squids (*Eledone* spp.). Its physiological action resembles that of other tachykinins. The peptide stimulates extravascular smooth muscles; it is a potent vasodilator and hypotensive agent (Pisano, 1968), and has potential therapeutic use to counter dry-eye syndrome.

Through combining common biological knowledge with simple observation and commonsense, a new biotech. company with millions of dollars in venture capital may be formed. In the following case, common knowledge of frogs and the Gram-negative bacteria found in wet environments was sufficient to launch a successful search for antibiotics. Although frogs live in ponds infested with Gram-negative bacteria, they rarely become infected by these pathogens. Based on their research, the biotech company Magainin Sciences Inc. is named after a peptide (magainin) that is highly effective against Gram-negative bacteria and occurs naturally in the skin of frogs. Effective antibiotics against this type of bacteria are rare and as a result this peptide is currently undergoing clinical trials (Jacob and Zasloff, 1994). This same way of thinking led to the discovery of the steroid squalamine, which has antibiotic properties and occurs in the stomach, liver and other organs of the shark (Moore *et al.*, 1993).

These few examples indicate that there is a whole new world to be found in wild biodiversity, accessible by gene technology, and merely awaiting exploration by the pharmaceutical biotech industry.

The agricultural biotech industry and biodiversity

Recombinant genes found in wild biodiversity may be even more important for agriculture than for the pharmaceutical industry. Classical breeding has quite successfully used genes from wild ancestors of cultivated plants to promote pest resistance and develop new and improved crop variations. Gene technology now enables humans to integrate revolutionary new properties into cultured plants through interspecific gene transfer. As with recombinant pharmaceuticals, research and development does not require random screening, but is rather a product-orientated engineering approach.

Exploiting natural defence and attack mechanisms for pest control. Plants protect themselves against pathogens through various enzymes, enzyme inhibitors and lectins. For example, the basic chitinases from rice and the acidic β -1,3-glucanase from alfalfa are directed against the cell walls of fungi. Transfer of the genes encoding these compounds into tobacco has yielded resistance against these pests (Zhu and Lamb, 1991; Maher *et al.*, 1994; Zhu *et al.*, 1994). The α -amylase inhibitor in the common bean makes

the starch of the seed indigestible for insects, and this property can be transferred into the garden pea (*Pisum sativum*) (Shade *et al.*, 1994).

The transfer into plants of genes from viruses, bacteria and animals is becoming a standard procedure in crop protection. Numerous successful field trials with recombinant crops expressing genes for the δ -endotoxins of *Bacillus thuringiensis* prompted intensive screening of bacterial species with insecticidal proteins (e.g. Koziel *et al.*, 1993). Researchers at Monsanto found a cholesterol oxidase in a streptomycete which lyses the midgut epithelium of pest insects; expression of the gene in transgenic plants could promote insect resistance (Corbin *et al.*, 1994).

Spider, scorpion and mite venoms contain neurotoxic peptides which specifically kill insects. The expression of their respective genes in plants also leads to resistance against insect pests (FMC Corporation, 1993).

Resistance against bacterial infections can be achieved by the production of peptides with antibiotic activities, such as cecropin B produced by wounded silk moths (Florack *et al.*, 1995). The production of viral coat or nonstructural proteins in plants protects the plant not only against infection from that same virus, but also against related virus types (Murray *et al.*, 1993). Even mammalian enzymes can be useful. For instance, 2'-5'-oligoadenylate synthetase from rats, when produced in potatoes, protects the plant against virus X under field conditions (Truve *et al.*, 1993).

Finally, resistance against herbicides can also be achieved through heterologous gene expression. A detoxification pathway for 2,4-dichlorophenoxyacetic acid, an agonist of indoleacetic acid, can be created in plants through the expression of a monooxygenase gene from the bacterium *Alcaligenes eutrophus* (Lyon *et al.*, 1989).

Although only few of these examples deal with the transfer of genes from wild biodiversity, there is no doubt that tropical environments, especially tropical rainforests, engender a multitude of defence and attack mechanisms among their inhabitants. This might lead us to suspect that these survival mechanisms must be highly sophisticated, and may represent a rewarding resource for the genetic engineer. It can be expected that the investigation of novel defence mechanisms will increase dramatically in the future as pest resistance develops to counter the first generation of recombinant plant variations. However, DNA sequences coding for defence proteins and peptides can be patented, and this may give cause for socioeconomic problems in the source countries, many of which are developing countries. This issue will be discussed below.

Engineering of metabolic pathways. Expression of recombinant genes in cultivated plants is under intensive investigation to improve oils, proteins, starch and other polymers for the food industry. A comprehensive overview of the state of the art is given by Beck and Ulrich (1993). However, it is not solely the food industry that stands to benefit: paper,

packaging and chemical industries will also be greatly affected. For example, Zeneca's method (Zeneca Ltd., 1995) to suppress cinnamyl alcohol dehydrogenase through antisense technology in trees facilitates the removal of lignin from cellulose, and will therefore have an impact on the aforementioned industries. In most cases, new plant variations are engineered by interspecific transfer of genes, coding for enzymes, which alter metabolic pathways. Examples based on genes from the wild biodiversity are still rare, but already quite impressive. For the production of soaps, chocolate and candies, medium-chain fatty acids found in coconut and palm kernel oil have great economic importance. In 1992, the US alone imported 600 000 tons of these oils, which contain up to 50% of trilaurin, a medium-chain dodecanoic unsaturated fatty acid. Recently, the USDA approved a high-laurate canola oil developed by the US ag-biotech company Calgene (PR Newswire, 1994). Calgene researchers investigated the synthesis of lauric acid in certain plant families and found a thioesterase which prematurely hydrolyses the growing acyl thioester of the fatty acid with an acyl-carrier protein in the wild Californian bay (*Umbellularia californica*). This 12:0-acyl-carrier protein thioesterase from the bay's developing oilseeds was expressed in transgenic *Arabidopsis thaliana* and *Brassica napus* ssp. *napus*, with the result that laurate and stearate became the most abundant types of fatty acid found in the oil of these plants (Voelker *et al.*, 1992). Three of Calgene's patents cover the purified enzyme, the recombinant nucleic acid construct with the gene for the enzyme, and a method to produce laurate in recombinant *Brassica* seeds (Calgene Inc., 1994a,b,c). The socioeconomic consequences of these patents will be discussed below.

Long-chain wax esters are required for a variety of industrial applications including pharmaceuticals, cosmetics, detergents, plastics and lubricants. Such products, especially long-chain wax esters, have previously been available from endangered species such as the sperm whale, or more recently, from the desert shrub, jojoba (*Simmondsia chinensis* or *S. californica*). Waxes are fatty alcohol and fatty acid esters, and their synthesis requires a fatty acid reductase as well as a synthase. The jojoba genes for the wax synthase (fatty acyl-CoA:fatty alcohol acyltransferase) and the fatty acyl reductase have been cloned and patented by Calgene (Calgene Inc., 1995a,b), and as a result, wax may be produced in rape seed in the future.

New biodegradable plastics produced from the bacterial storage compound polyhydroxybutyrate have been developed and are already on the market. However, production of the compound through fermentation is not cost efficient. In recent development, the transfer of the entire anabolic pathway, consisting of three enzymes from the bacterium *Alcaligenes eutrophus*, into *Arabidopsis thaliana* (Nawrath *et al.*, 1994) may transform the farm field into a chemical factory for plastics in the near future. Examples for pathway engineering do not yet involve genes from the tropical rainforest, but a thorough and product-orientated survey of oils, fats, waxes and

other polymers (formerly not of commercial interest because of low productivity or abundance) from tropical plants, animals and microorganisms may lead to new compounds for industrial applications. These new compounds would have the advantage of being both biodegradable and available for farm production through genetic engineering.

Gene technology will also lead to the more efficient production of natural compounds presently used for pharmaceuticals and pesticides. The chrysanthemyl diphosphate synthase gene from *Chrysanthemum cinerariaefolium* was patented by Agridyne Technologies Inc. (1995) and promises to open new paths for producing insecticidal pyrethrins, pyrethroids and their derivatives with greater efficiency and higher purity levels in cultivated plants.

Metabolic engineering of medicinal plants has been performed successfully with *Atropa belladonna* to produce the alkaloid scopolamine. The naturally occurring alkaloid in this plant is hyoscyamine, known for its anticholinergic activity, and an active ingredient in eye drops, antidotes and spasmolytics (Yun *et al.*, 1992). Scopolamine, found throughout the *Solanaceae*, is an epoxy-derivative of hyoscyamine but with a broader therapeutic spectrum, including, e.g., antiemetics and hypnotics. Expression of the hyoscyamine 6- β -hydroxylase gene from *Hyoscyamus niger* in *Atropa* led to an almost exclusive accumulation of scopolamine in the plant's leaf and stem. Flux through a pathway to a plant secondary product can be elevated by genetic engineering. For example, over-expression of the yeast ornithine decarboxylase gene in transgenic roots of *Nicotiana rustica* led to enhanced nicotine accumulation (Hamill *et al.*, 1990). The same effect can be obtained to produce sterols in plants by over-expressing 3-hydroxy-3-methylglutaryl CoA reductase, which catalyses the production of the sterol building block mevalonate (Amoco Corporation, 1994).

11.3. Modern Bioprospecting: Linking Industry, Biodiversity Conservation and Developing Country Technology Acquisition

The rapid loss of biological diversity – indicated by the extinction of an estimated 30 to 300 species per day (Japan Economic Newswire, 1995) – together with the potential opportunities and threats of gene technology, led to the United Nations Convention on Biological Diversity (UNEP, 1992). In light of the vast potential of biotic materials and the need to ensure their survival, in addition to measures taken to improve biodiversity conservation activities, it is imperative that industries move from the passive role of simple users to the more active one of reinvesting part of their revenues into conservation efforts. Companies should be aware that

they are among the first to lose as a consequence of species extinction, and indeed such an awareness is growing.

The principle of this modern approach to bioprospecting may be simple, but the link between biodiversity conservation and its sustainable use requires a careful design and strategic planning. The goals are to maximize those uses which generate information, and to reinvest part of the benefits obtained from bioproducts into acquiring knowledge and improving biological resource management. As a consequence, wildland biodiversity can be developed as part of a country's national economy at the same time as its preservation into perpetuity is guaranteed. The bioindustries are thereby encouraged to initiate relationships with partners in biodiversity-rich countries. Following the guidelines of the Biodiversity Convention such partnership can facilitate sustainable and nondamaging biological and genetic resource use for research and development, while taking care to share economic and intellectual benefits with the owners of the biological resources.

The process of collecting bioresources, extracting and testing constituents (either chemicals or genes) for biological activity, and the further development of a product is long and expensive (Reid *et al.*, 1993). Based on the research and development costs of 93 randomly selected new chemical entities during the period 1970 to 1982, the development of a single drug to the point of market approval was estimated at \$US114 million for the USA (DiMasi *et al.*, 1991). Although rewards might be high, the chances of failure are equally so: of 10 000 different products tested, only one will make it to the market (Farnsworth, 1994).

However, the real challenge for this new generation of bioprospectors is to find a way to capture part of the financial revenues for the source country's biodiversity conservation efforts and economic development. As an example of this innovative approach to prospecting activities, INBio is negotiating agreements with scientific research centres, universities and private enterprise that are mutually beneficial to all parties (Sittenfeld and Lovejoy, 1994). These pioneering agreements provide significant returns to Costa Rica while simultaneously assigning economic value to natural resources, and providing a new source of income to support the maintenance and development of the country's Conservation Areas (Sittenfeld and Lovejoy, 1995).

11.3.1. Bioprospecting frameworks

Modern biodiversity prospecting requires the creation of appropriate frameworks and the cooperation and involvement of governments, intermediary institutions, private enterprise, academia, and local communities and entities. This activity also requires the involvement of lawyers,

lawmakers, scientists, managers and economists from developing and developed countries (Sittenfeld and Lovejoy, 1995).

The fundamental point of departure for a biodiversity prospecting framework is *macro-policy*, the set of governmental and international regulations, laws and economic incentives that determine land use patterns, access to and control of biological resources, intellectual property rights regimes, technology promotion, and industrial development. Macro-policies are formed on the international, national and social levels. On the international level, agreements, conventions and other mechanisms establish the relationships and protocols for sharing biological resources between countries. Documents considered important in providing the guidelines and regulations for biological resource use include: the Biodiversity Convention, the Trade Related Intellectual Property Rights (TRIPs) of the General Agreement on Tariffs and Trade (GATT), the Draft Declaration on Indigenous Rights of the United Nations Working Group on Indigenous Populations, and also subregional agreements, such as the North American Free Trade Agreement (NAFTA), the Amazonian Treaty, and the Pacto Andino.

Nevertheless, conventions, agreements and organizations still leave open the major responsibilities of designing adequate legislation and regulations regarding land ownership, land tenure rights, the creation of protected areas, the use of biological resources, nationally recognized intellectual property rights, the definition of public-domain resources, and the creation of market incentives or deterrents for private enterprise and research investments to each individual country. Such legislation and regulations promote stability and manoeuvrability of in-country partners, characteristics considered attractive to private industry and academic research counterparts.

Deterrents, such as national policy vacuums or legislation drafted outside the framework of the Biodiversity Convention, still exist in many countries and continue to create obstacles to establishing collaborations with academic and industrial research partners. In general, changes in laws and policies governing the ownership of and access to genetic resources are needed as well as changes in the way bio-business has evolved to date. The importance of favourable national policies, regulations and laws becomes obvious when considering international intellectual property rights. Drug research within the source country itself is an important step towards national economic development, but will only be attractive to the industrial partner if results can be patented. It is for this reason more than any other that international patent laws should be recognized by national law.

At the same time, there is concern within the industrial sector that countries, spurred by the Biodiversity Convention, may promulgate new laws restricting access to biological and genetic resources, and reducing

renewed enthusiasm for natural products (Putterman, 1994). Yet the recognition by the Convention of sovereign rights of nations over their genetic resources is intended to encourage world trade in genetic resources, since it commits countries to facilitate access, based on mutually agreed terms (Putterman, 1994). National governments should implement rules, regulations and policies that take advantage of Articles 15 and 16 of the Convention. These Articles encourage source-country participation in researching their own biological resources, transferring technologies to utilize these resources and reaping a fair share of the benefits from their commercial exploitation (Sittenfeld and Lovejoy, 1995).

Finally, on the social level, heavy investment in education and other social services has created a scientific environment of qualified institutions, researchers and educated personnel in Costa Rica. Such an environment is a prerequisite for research collaborations with private enterprise and is essential for integrating biodiversity into economic development (Sittenfeld and Lovejoy, 1995).

11.3.2. Inventories, business development and technology access

Supported by a favourable international and national macro-policy, three basic elements guide the rational and productive use of biological resources in prospecting agreements: (i) biodiversity inventories and information handling; (ii) business development; and (iii) technology access.

Inventories and information management

As pointed out in Section 11.2.1, screening for drugs and pesticides will only be successful through the development and management of biological, ecological, taxonomic, and related systematic information on living species and systems. Even with these data, further information is required for the more systematic screening approach used in gene technology. For example, biochemical data must be evaluated for, e.g., the occurrence of certain biopolymers, metabolic pathways, enzymes and defence or attack mechanisms. Biodiversity inventories create catalogues of available resources and their location. They prevent damage to ecosystems, areas, species and populations by indicating what resources are available, and where they can be collected without damaging the environment (Raven and Wilson, 1992). Simultaneously, the source-country collaborator becomes a more attractive, knowledgeable and reliable business partner because the inventory-generated information reduces the uncertainties of collecting further material should this prove necessary.

Business development

Building upon inventory-generated knowledge, business development defines markets, market needs, major players, and national capacities in

science and technology as well as institutional strategies and goals. Important requirements include knowledge of one's assets and drawbacks, market surveys and evaluation of conservation needs. The key to business development is interacting with international industry in order to approach the market in a realistic and practical way. Because bioprospecting should promote source-country economic development, business development must encourage the sustainable use of biodiversity by local entrepreneurs. However, this is a considerable challenge in developing countries where industry normally cannot take the financial risk of applying innovative technologies, let alone those that are sustainable and non-damaging to the environment.

Technology transfer

One of the major issues discussed in the Biodiversity Convention refers to technology transfer, allowing source countries to convert raw biological materials into products of greater value in exchange for access to their biodiversity (Putterman, 1994). This issue is of tremendous importance, particularly in a decade of patentable genes, and will be discussed in more detail below. In the near future, genes isolated from tropical biodiversity may provide the farmers in developed countries with advantages over the farmers of the source countries. These advantages stem from the bioindustries' historical development and their physical proximity to the developed agricultural economies of the North. This may lead biotechnological research and development to concentrate solely on improving the properties of crop and livestock in the North (for discussion, see Shand, 1993). Technology transfer may enable source countries to keep pace with the developed countries, and avoid being left out of important agricultural developments (Lesser and Krattiger, 1993). This scenario is realistic because gene technology, in contrast to natural compound chemistry, does not particularly rely on expensive investments in laboratory equipment, and would therefore be easier to implement.

11.3.3. Contract negotiations

In general, contract negotiation is divided into three basic sets of issues: scientific issues, business issues and legal issues. To negotiate, an organization must have a good sense of its own fundamental needs and those of its potential collaborator. The typical source-country needs are: the generation of income to support protected areas and conservation management activities through direct contributions as well as royalties; the transfer of processing technologies and a guaranteed future profit-sharing if commercial products are forthcoming. Sampling must be done under best ecological

practice without damaging the ecosystem. For bilateral contracts with industrial partners, exclusivity and time limitations are further requirements.

In summary, modern bioprospecting requires that the source country:

1. Creates an infrastructure guaranteeing a reliable supply of natural products (including correct taxonomic identification, quality control, full support from government and adherence to national or local regulations on access to resources).
2. Acquires technology that adds value to natural products wherever possible (from extracts to partially purified or pure compounds or gene sequences).
3. Takes advantage of local capabilities using all types of organisms as biological resources attractive to industry (from plants and microbial resources through marine or freshwater life forms to arthropods).
4. Develops a reputation as a reliable business partner over the time.
5. Reinvests part of the revenues in improving biodiversity management and conservation.

In exchange for access to biological resources, the industrial partner must agree to:

1. The fair and equitable sharing of benefits, both in intellectual and monetary terms.
2. The implementation of collection and production methods with minimum effects on biodiversity.
3. The use of equitable bioprospecting practices for further research on tropical diseases and problems specifically associated with developing countries.

11.4. How to Face the New Challenge of Gene Technology

With few exceptions classical bioprospecting for drugs has not proved economically beneficial for developing nations, but nor has it directly damaged these economies. In contrast, bioprospecting for genes may soon pose a real threat to the economic survival of these biodiversity-rich countries. The farmers of the North are currently suffering under low prices and overproduction of certain traditional crops and livestock. As a result they are looking for new markets and products.

Modern biotechnology promises to aid Northern farmers in this endeavour, eliminating or displacing traditional export commodities from developing countries and transferring production or substitutes from the farm fields of the South to those of the North. Quite possibly the transfer may even skip the Northern farms and jump straight into bioreactors. In Africa alone, US\$ 10 billion in exports are vulnerable to industry-induced

changes in raw material prices and requirements (Shand, 1993). Most developments in plant biotechnology have been achieved with crops cultivated mainly in industrialized countries. This denies the farmers of developing countries the chance to benefit from the new agricultural opportunities of gene technology.

As mentioned above, Calgene's high-laurate canola oil may displace coconut and palm kernel oil, posing a threat to the economic survival of millions of farm families in the South. In the Golfito region of Costa Rica the government has started a programme to grow oil palms on banana fields deserted by the US fruit multinationals. All these efforts, which are partially financed with Northern developmental aid, may vanish into thin air as a result of Calgene's new rape seed. Thaumatin, a sweet-tasting basic protein from the tropical plant *Thaumatococcus*, has been traditionally used in West Africa as a sweetener. With the collection of *Thaumatococcus* fruits for the British food industry, the population in this region earned a large part of its income. However, the thaumatin gene has now been cloned and the sweetening protein can be produced by large-scale fermentation of brewer's yeast at low cost (Lee *et al.*, 1988). The same applies to natural compounds like indigo, which can now be produced by the fermentation of *Escherichia coli* engineered with genes from a toluol-degrading subspecies of the soil bacterium *Pseudomonas putida* (Ensley *et al.*, 1983). Products like vanilla, pyrethrum and rubber may follow this same path.

Along the same lines, the extension of patent laws to the developing nations through the GATT could mean that the biotech industry obtains a monopoly on genetically engineered livestock and crops, which farmers in developing countries must cultivate under constraint in order to remain competitive. For example, US-Patent No. 5 159 135 of Agracetus (a subsidiary of W.R. Grace & Co.) covers all genetically engineered cotton. This patent is a warning of potential future problems, and has already caused an outcry in developing countries like India (Kidd and Dvorak, 1994). If the biotech industry continues developing without adequate controls, the consumers and farmers of developing countries may even end up paying royalties on biotech products that were originally developed from their very own resources and knowledge.

As a step in the right direction, the US Patent and Trademark Office reversed its decision to grant the Agracetus patent at the end of 1994, but primarily as a result of pressure from the US biotech industry which argued that patents like this will inhibit research and development (AGWEEK, 1994). Such issues bring important questions to light: How can developing nations be motivated to conserve their biodiversity under these threatening circumstances? Are the regulations and tools of modern bioprospecting, as described above, sufficient to face this challenge?

The Biodiversity Convention attempts to address this new threat by requiring that access to biodiversity's genetic potential be combined with

the biotechnology transfer to the South in order for those countries to develop their own methods of sustainable biodiversity utilization. Nevertheless, the Convention suffers from three major drawbacks.

1. The treaty was not ratified by the USA, the leading country in biotechnological research, development and application.
2. It specifically excludes (under US pressure) *ex situ* genebank material collected before the enactment of the treaty (Shand, 1993). As a result, huge stocks of germplasm collected by the North, mostly in tropical and subtropical countries, are not restricted by the Convention. The recent transfer by the Consultative Group on International Agricultural Research of its 12 genebanks to the auspices of the UN must be the first step in keeping the access of developing countries open to their own resources (Madeley, 1994).
3. It still remains nearly impossible to control the illegal transfer of genetic material into the North. Genes can be cloned from minute amounts of DNA or RNA and isolated from biological material that easily fits into an airmail envelope. Genes do not have tags designating their country of origin, and once they are cloned, they are no longer controlled by their source country. This is quite different from the isolation of natural compounds from plants, where larger amounts of plant material must be collected and, for the process of isolation and structure elucidation, must be re-collected. In this last case, controlling the flow of biological material is possible simply because industry will eventually require sample resupply at a given point, and for this the industry needs reliable partners in countries of origin.

These issues must be approached in an active and more aggressive manner than traditionally used. During collaborations with traditional pharmaceutical and biotech companies like Merck & Co., Bristol-Myers Squibb and the British Technology Group, INBio used such an approach, proving to industry that fair partnerships are mandatory and conducive to success. More importantly, INBio demonstrated that reliable applied research is possible in a developing country, and that technology transfer to acquire necessary know-how and equipment works to the advantage of the industrial partner as well.

The same applies to the biotech industry. In this case, the transfer of gene technology is not critical, in contrast to natural compound chemistry, because it does not require million-dollar investments in laboratory infrastructure. Rather, the industry relies on the know-how already existing in many source countries. Gene technology also represents a promising development opportunity for countries that do not have large research budgets at their disposal. Moreover, gene technology is a very straightforward approach, relying on natural history observations (e.g. whether certain plants show natural resistance to pathogens), and not involving

the automated random screening of thousands of samples. Biodiversity inventories, which are already in place in countries like Costa Rica, are a reasonable and advantageous prerequisite for successful 'gene prospecting'.

Costa Rica's aggressive strategy to foster collaborations with the international industry and academic institutions in drug research, gene technology and agriculture actively seeks to develop and patent natural compounds, proteins and genes in Costa Rica based on a foundation of national research. Collaborations of this nature will help launch Costa Rica onto a scientific and technological plane that offers services and goods that are both competitive and compatible with those of industrial nations.

Simultaneously, INBio works within this strategy to increase knowledge about Costa Rican biodiversity in general, access revenues for further conservation efforts, and to assign biodiversity a higher value than it has had in the past. There is little doubt that such activities will encourage society's willingness to preserve biodiversity for future generations, by making it worthwhile for the Costa Rican population to maintain tropical forests and other ecosystems on their own.

Acknowledgements

We thank Mrs Annie Lovejoy for critical reading, comments and major improvements of the manuscript. This work is supported in part by a grant (to Dr Nader) of the Centre for International Migration and Development, Frankfurt, Germany, and by grant No. 5 U01 TW/CA00312 from the National Institute of Health (Fogarty International Center).

References

- Agridyne Technologies Inc. (1995) Chrysanthemyl diphosphate synthase, corresponding genes and use in pyrethrin synthesis. US Patent No. 5 443 978.
- AGWEEK (1994) Plant patent dispute reaches standoff. December 12.
- Amoco Corporation (1994) Method and composition for increasing sterol accumulation in higher plants. US Patent No. 5 365 017.
- Babu, J.R. (1988) Avermectins: biological and pesticidal activities. In: Cutler, H.G. (ed.) *Biological Active Natural Products: Potential Use in Agriculture*. American Chemical Society, Washington DC, pp. 91–108.
- Baheer, A.R., el-Asmar, M.F. and Soslau, G. (1995) Characterization of a potent platelet aggregation inducer from *Cerastes cerastes* (Egyptian sand viper) venom. *Biochimica et Biophysica Acta* 1250, 97–109.
- Balick, M. (1994) Ethnobotany, drug development and biodiversity conservation – exploring the linkages. In: Chadwick, P.J. and Marsh, J. (eds) *Ethnobotany and*

- the Search for New Drugs. Ciba Foundation Symposium 185.* John Wiley, Chichester, pp. 4–18; discussion pp. 19–24.
- BASF AG (1987) Verfahren zur Herstellung von mesophilen Mikroorganismen, die eine bei höherer Temperatur aktive D-Hydantoinase enthalten. German Patent No. 3535987.
- Beck, C.I. and Ulrich, T.H. (1993) Biotechnology in the food industry. *Bio/Technology* 11, 895–902.
- Birch, A.N.E., Robertson, W.M., Geoghegan, I.E., McGavin, W.J., Alphey, T.J.W., Phillips, M.S., Fellows, L.E., Watson, A.A., Simmonds, M.S.J. and Porter, E.A. (1993) DMDP – A plant-derived sugar analogue with systematic activity against plant parasitic nematodes. *Nematologica* 39, 517–530.
- Calgene Inc. (1994a) Plant thioesterase having preferential hydrolase activity toward C12. US Patent No. 5 304 481.
- Calgene Inc. (1994b) Plant medium-chain-preferring acyl-ACP thioesterases and related methods. US Patent No. 5 298 421.
- Calgene Inc. (1994c) Plant thioesterases. US Patent No. 5 344 771.
- Calgene Inc. (1995a) Jojoba wax biosynthesis gene. US Patent No. 5 445 947.
- Calgene Inc. (1995b) Fatty acyl reductase. US Patent No. 5 411 879.
- Cardellina, J.H., Gustafson, K.R., Beutler, J.A., McKee, T.C., Hallock, Y.F., Fuller, R.W. and Boyd, M. (1993a) National Cancer Institute intramural research on human immunodeficiency virus inhibitory and antitumor plant natural products. In: Kinghorn, D.A. and Balandrin, M.F. (eds) *Human Medicinal Agents from Plants*. American Chemical Society, Washington DC, pp. 218–227.
- Cardellina, J.H., Munro, M.H.G., Fuller, R.W., Manfredi, K.P., McKee, T.C., Tischler, M., Bokesch, H.R., Gustafson, K.R., Beutler, J.A., and Boyd, M. (1993b) A chemical screening strategy for the dereplication and prioritization of HIV-inhibitory aqueous natural products extracts. *Journal of Natural Products* 56, 1123–1129.
- Corbin, D.R., Greenplate, J.T., Wong, E.Y. and Purcell, J.P. (1994) Cloning of an insecticidal cholesterol oxidase gene and its expression in bacteria and in plant protoplasts. *Applied and Environmental Microbiology* 60, 4239–4244.
- Cragg, G., Boyd, M.R., Cardellina, J.H., Newman, D.J., Snader, K.M. and McCloud, T.G. (1994) Ethnobotany and drug discovery: the experience of the US National Cancer Institute. In: Chadwick, D.J. and Marsh, J. (eds) *Ethnobotany and the Search for New Drugs. Ciba Foundation Symposium 185.* John Wiley, Chichester, pp. 178–190; discussion pp. 191–196.
- DiMasi, J.A., Hansen, R.W., Grabowski, H.G. and Lasagna, L. (1991) Cost of innovation in the pharmaceutical industry. *Journal of Health Economics* 10, 107–142.
- Ecker, D.J. and Crooke, S.T. (1995) Combinatorial drug discovery: which method will produce the greatest value. *Bio/Technology* 13, 351–361.
- Ensley, B.D., Ratzkin, B.J., Osslund, T.D., Simon, M.J., Wackett, L.P. and Gibson, D.T. (1983) Expression of naphthalene oxidation genes in *Escherichia coli* results in the biosynthesis of indigo. *Science* 222, 167–169.
- Epand, R.M., Epand, R.F., Orlowski, R.C., Seyler, J.K. and Colescott, R.L. (1986) Conformational flexibility and biological activity of salmon calcitonin. *Biochemistry* 25, 1964–1968.
- Farnsworth, N.R. (1994) Ethnopharmacology and drug development. In: Chadwick, D.J. and Marsh, J. (eds) *Ethnobotany and the Search for New Drugs. Ciba Foundation Symposium 185.* John Wiley, Chichester, pp. 42–59.

- Florack, D., Allefs, S., Bollen, R., Bosch, D., Visser, B. and Stiekema, W. (1995) Expression of giant silkworm cecropin B genes in tobacco. *Transgenic Research* 4, 132–134.
- FMC Corporation (1993) Insecticidally effective peptides. US Patent No. 5 441 934.
- Fortkamp, E., Rieger, M., Heisterberg-Moutsers, G., Schweitzer, S. and Sommer, R. (1986) Cloning and expression in *Escherichia coli* of a synthetic DNA for hirudin, the blood coagulation inhibitor in the leech. *DNA* 5, 511–517.
- Frost & Sullivan Inc. (1992) *The U.S. Market for Chemical Pesticides*. Frost & Sullivan Inc., New York, p. 8.
- Global Biodiversity Strategy: Guidelines for Action to Save, Study, and Use Earth's Biotic Wealth Sustainably and Equitably* (1992) WRI/IUCN/UNEP, USA.
- Grünwald, J. (1995) The European phytomedicines market: figures, trends, analyses. *HerbalGram* (Austin, Texas) 34, 60–65.
- Hamill, J.D., Robins, R.J., Parr, A.J., Evans, D.M., Furze, J.M. and Rhodes, M.J. (1990) Over-expressing a yeast ornithine decarboxylase gene in transgenic roots of *Nicotiana rustica* leads to enhanced nicotine accumulation. *Plant Molecular Biology* 15, 27–38.
- Harborne, J. (1989) *Introduction to Ecological Biochemistry*, 4th edn. Academic Press, London, UK, pp. 171–172.
- Jacob, L. and Zasloff, M. (1994) Potential therapeutic applications of magainins and other antimicrobial agents of animal origin. In: *Ciba Foundation Symposium 186*. John Wiley, Chichester, pp. 197–216; discussion pp. 216–223.
- Janzen, D. and INBio (1992) A north–south perspective on science in the management, use and development of biodiversity. In: Sandlund, O.T., Hindart, K. and Brown, A.H.D. (eds) *Conservation of Biodiversity for Sustainable Development*. Scandinavian University Press, Oslo, pp. 27–54.
- Janzen, D.H., Fellows, L.E. and Waterman, P.G. (1990) What protects *Lonchocarpus* (Leguminosae) seeds in a Costa Rican dry forest? *Biotropica* 22(3), 272–285.
- Japan Economic Newswire (1995) Biodiversity report proposes government action. March 29.
- Kanegafuchi Company (1978) Verfahren zur Herstellung von D-N-Carbamoyl-a-Aminosäuren. German Patent No. 2757980.
- Kashman, Y., Gustafson, K.R., Fuller, R.W., Cardellina, J.H., McMahon, J.B., Currens, M.J., Buckneit, R.W., Hughes, S.H., Cragg, G.M. and Boyd, M.R. (1992) The calanolides, a novel HIV-inhibitory class of coumarin derivatives from the tropical rainforest tree, *Calophyllum lanigerum*. *Journal of Medicinal Chemistry* 35, 2735–2743.
- Kidd, G. and Dvorak, J. (1994) Agracetus' cotton patent draws opposition. *Bio/Technology* 12, 659.
- König, G., Wright, A.D., Sticher, O., Angerhofer, C.K. and Pezzuto, J.M. (1994) Biological activities of selected marine natural products. *Planta Medica* 60, 532–537.
- Kozziel, M.G., Beland, G.L., Bowman, C., Carozzi, N.B., Crenshaw, R., Crossland, L., Dawson, J., Desai, N., Hill, M., Kadwell, S., Laitis, K., Lewis, K., Maddox, D., McPherson, K., Meghji, M.R., Merlin, E., Rhodes, R., Warren, G.W., Wright, M. and Evola, S.V. (1993) Field performance of elite transgenic maize plants expressing an insecticidal protein derived from *Bacillus thuringiensis*. *Bio/Technology* 11, 194–200.

- Lee, J.H., Weickmann, J.L., Koduri, R.K., Gosh-Dastidar, P., Saito, K., Blair, L.C., Date, T., Lai, J.S., Hollenberg, S.M. and Kendall, R.L. (1988) Expression of synthetic thaumatin genes in yeast. *Biochemistry* 27, 5101–5107.
- Lesser, W. and Krattiger, A.F. (1993) *Facilitating South–North and South–South Technology Flow Processes for “Genetic Technology”*. Working paper R7W. International Academy of the Environment – Biodiversity / Biotechnology Programme, Geneva.
- Lyon, B.R., Llewellyn, D.J., Huppatz, J.L., Dennis, E.S. and Peacock, W.J. (1989) Expression of a bacterial gene in transgenic tobacco plants confers resistance to the herbicide 2,4-dichlorophenoxyacetic acid. *Plant Molecular Biology* 13, 533–540.
- McCloud, T.G., Nemec, J., Mushik, G., Sheffield, H.G., Quesenberry, P., Suffness, M., Cragg, G. and Thompson, J. (1988) Extraction of bioactive molecules from plants. In: *Proceedings of the International Congress on Natural Products Research, July 17–21, Park City, Utah*.
- MacIntyre, I., Alevizaki, M., Bevis, P.J. and Zaidi, M. (1987) Calcitonin and the peptides from the calcitonin gene. *Clinical Orthopedics* 43, 45–55.
- Madeley, J. (1994) International plant genebanks come under UN auspices. *Financial Times* October 26.
- Maher, E.A., Lamb, C.J. and Dixon, R.A. (1994) Stress response in alfalfa (*Medicago sativa* L.). XVII. Identification of multiple hydrolyases and molecular characterization of an acidic glucanase. *Physiological and Molecular Plant Pathology* 43, 329–342.
- Mendelsohn, R. and Balick, M.J. (1995) The value of undiscovered pharmaceuticals in tropical forests. *Economic Botany* 49, 223–228.
- Moore, K.S., Wehrli, S., Roder, H., Rogers, M., Forrest, J.N., McCrimmon, D. and Zasloff, M. (1993) Squalamine: an aminosterol antibiotic from the shark. *Proceedings of the National Academy of Sciences of the USA* 90, 1354–1358.
- Murray, L.E., Elliott, L.G., Capitant, S.A., West, J.A., Hanson, K.K., Scarafia, L., Johnston, S., DeLuca-Flaherty, C., Nichols, S., Cunanan, D., Dietrich, P.S., Mettler, I.J., Dewald, S., Warnick, D.A., Rhodes, C., Sinibaldi, R.M. and Brunke, K. (1993) Transgenic corn plants expressing MDMV strain B coat protein are resistant to mixed infections of maize dwarf mosaic virus and maize chlorotic mottle virus. *Bio/Technology* 11, 1559–1564.
- Nawrath, C., Poirier, Y. and Somerville, C. (1994) Targeting of the polyhydroxybutyrate biosynthetic pathway to the plastids of *Arabidopsis thaliana* results in high levels of polymer accumulation. *Proceedings of the National Academy of Sciences of the USA* 91, 12760–12764.
- O'Neill, M.J. and Lewis, J.A. (1993) The renaissance of plant research in the pharmaceutical industry. In: Kinghorn, D.A. and Balandrin, M.F. (eds) *Human Medicinal Agents from Plants*. American Chemical Society, Washington DC, pp. 48–55.
- Pisano, J.J. (1968) Vasoactive peptides in venomes. *Federation Proceedings* 27, 58–62.
- Principe, P.P. (1989) The economic significance of plants and their constituents as drugs. In: Wagner, H., Hikino, H. and Farnsworth, N.R. (eds) *Economic and Medicinal Plant Research*, vol. 3. Academic Press, London, pp.1–17.
- PR Newswire (1994) Calgene receives USDA approval to produce laurate canola. November 1.

- Putterman, D. (1994) Trade and the biodiversity convention. *Nature* 371, 553–554.
- Raven, P. and Wilson, E.O. (1992) A fifty-year plan for biodiversity surveys. *Science* 258, 1099–1100.
- Reid, W.V., Laird, S., Meyer, C. A., Gámez, R., Sittenfeld, A., Janzen, D.H., Gollin, M.A. and Juma, C. (1993) *Biodiversity Prospecting: Using Genetic Resources for Sustainable Development*. World Resources Institute, Washington DC, 341 pp.
- Research Foundation of the State University of New York (1994) Biological applications of alkaloids derived from the tunicate *Eudistoma* sp. US Patent No. 5 278 168.
- Rural Advancement Foundation International (1994) *Conserving Indigenous Knowledge: Integrating Two Systems of Innovation*. United Nations Development Programme, New York.
- Sacheri, F., Nitti, G., Valsasina, B., Orsini, G., Visco, C., Ferrera, M., Sawyer, R.T. and Sarmientos, P. (1993) Novel hirudin variants from the leech *Hirudinaria manilensis*. Amino acid sequence, cDNA cloning and genomic organization. *European Journal of Biochemistry* 214, 295–304.
- Sapnde, T.F., Garraffo, H.M., Edwards, M.W., Yeh, H.J.C., Pannell, L. and Daly, J.W. (1992) Epibatidine: A novel (chloropyridyl)azabicycloheptane with potent analgesic activity from an Ecuadoran poison frog. *Journal of the American Chemical Society* 114, 3475–3478.
- Schleuning, W.D., Alagon, A., Boidol, W., Bringmann, P., Petri, T., Kratzschmar, J., Haendler, B., Langer, G., Baldus, B. and Witt, W. (1992) Plasminogen activators from the saliva of *Desmodus rotundus* (common vampire bat): unique fibrin specificity. *Annals of the New York Academy of Sciences* 667, 395–403.
- Shade, R.E., Schroeder, H.E., Pueyo, J.J., Tabe, L.M., Murdock, L.L., Higgins, T.J.V. and Chrispeels, M.J. (1994) Transgenic pea seeds expressing the alpha-amylase inhibitor of the common bean are resistant to bruchid beetles. *Bio/Technology* 12, 793–796.
- Shand, H. (1993) Agbio and third world development. *Bio/Technology* 11, S13.
- Sittenfeld, A. and Lovejoy, A. (1994) Biodiversity prospecting. *Our Planet (UNEP: Magazine for Sustainable Development)* 6(4), 20–21.
- Sittenfeld, A. and Lovejoy, A. (1995) Biodiversity prospecting frameworks: The INBio experience in Costa Rica. In: McNeely, J. and Guruswamy, L. (eds) *Their Seed Preserve: Strategies for Protecting Global Biodiversity*. Duke University Press.
- Sittenfeld, A. and Villers, R. (1994) Costa Rica's INBio: Collaborative biodiversity research agreements with the pharmaceutical industry. Case Study 2. In: Meffe, G.K., Carroll, C.R. and contributors. *Principles of Conservation Biology*. Sinauer Associates, Sunderland, MA, 500–504.
- Truve, E., Aaspollu, A., Horikanen, J., Puska, R., Mehto, M., Hassi, A., Teeri, T.H., Kelve, M., Seppänen, P. and Saarma, M. (1993) Transgenic potato plants expressing mammalian 2'-5' oligoadenylate synthetase are protected from potato virus X infection under field conditions. *Bio/Technology* 11, 1048–1052.
- UNEP (1992) *Convention on Biological Diversity*. United Nations Environmental Programme, Nairobi, Kenya.
- US National Research Council (Panel of the Board on Science and Technology for International Development) (1992) *Conserving Biodiversity: A Research Agenda for Development Agencies*. National Academic Press, Washington DC.
- Voelker, T.A., Worrell, A.C., Anderson, L., Bleibaum, J., Fan, C., Hawkins, D.J.,

- Radke, S.E. and Davies, H.M. (1992) Fatty acid biosynthesis redirected to medium chains in transgenic oil seed plants. *Science* 257, 72–74.
- Wilson, E.O. (1992) Resolution. In: Wilson, E.O. (ed.) *The Diversity of Life*. Harvard University Press, Cambridge, MA, pp. 311–342.
- Wink, M. (1993) Production and application of phytochemicals from an agricultural perspective. In: van Beek, T.A. and Breteler, H. (eds) *Proceedings of the Phytochemical Society of Europe: Phytochemistry and Agriculture*. Clarendon Press, Oxford.
- Yamada, H., Takahashi, S., Kii, Y. and Kumagai, H. (1978) Distribution of hydan-toin hydrolyzing activity in microorganisms. *Journal of Fermentation Technology* 56, 484–491.
- Yun, D.J., Hashimoto, T. and Yamada, Y. (1992) Metabolic engineering of medicinal plants: transgenic *Atropa belladonna* with an improved alkaloid composition. *Proceedings of the National Academy of Sciences of the USA* 89, 11799–11803.
- Zeneca Ltd. (1995) Modification of lignin synthesis in plants. US Patent No. 5 451 514.
- Zhu, Q. and Lamb, C.J. (1991) Isolation and characterization of a rice gene encoding basic chitinase. *Molecular and General Genetics* 226, 289–296.
- Zhu, Q., Maher, E.A., Masoud, S., Dixon, R.A. and Lamb, C.J. (1994) Enhanced protection against fungal attack by constitutive co-expression of chitinase and glucanase genes in transgenic tobacco. *Bio/Technology* 12, 807–812.

Internet Resources for the Biologist **12**

M.L. Anderson¹ and S.W. Cartinhour²

¹Nottingham Arabidopsis Stock Centre, Department of Life Science, University of Nottingham, Nottingham NG7 2RD, UK; ²Crop Biotechnology Center, Texas A&M University, College Station, TX 77843, USA

12.1. Introduction

In the last few years biotechnology has had a significant impact on the ways in which genetic resources are maintained, characterized and used. At the same time, another technology has revolutionized the methodology for organizing and disseminating information: the use of computers and the international network known as the Internet. The Internet has become a vital resource, allowing access to information spanning the disciplines of biodiversity and systematics to molecular biology and biotechnology. It is difficult to overstate the relevance the Internet now has to all biologists, many of whom rely as a matter of course on their ability to access electronic mail (email) and a growing number of on-line data resources.

On-line information relating to genetic resources is media-rich, encompassing photographs of specimens, electron micrographs, molecular data in the form of autoradiographs of restriction fragment length polymorphisms (RFLPs), randomly amplified polymorphic DNA (RAPD) analysis, DNA fingerprinting, transformation vector and sequence information, mapping populations, genetic maps, physical maps, bibliographies and colleague information. Information from one resource can be integrated with data at another resource to synthesize a new 'view' that is available from neither one alone. This is true whether or not the information being integrated is provided by the same or geographically distant sites. Given the complexity and quantities of data being generated by biological research, it is certain that more and more information will be available

exclusively in electronic form. Network access is therefore no longer a luxury but is fast becoming a necessity.

This chapter will deal with how information can be organized and distributed, both from the view of data consumers and publishers. However, it should be understood immediately that it is impossible to compile an exhaustive list of on-line resources or to evaluate their contents. The Internet is unlike a telephone exchange or library; no central authority offers a comprehensive index. Moreover, what is available changes so rapidly that any detailed record would soon be out of date. Nonetheless we hope to provide a sense of what is available to biologists and will include a number of useful starting places from which exploration can begin.

It is also hoped that readers will understand the value and feasibility of making their own data available. The Internet, with its decentralized structure, is an opportunity to distribute information, including that which might be otherwise difficult to publish by traditional means, but which will be useful to colleagues around the world.

Several books are available which deal with the technical details of how to connect to the Internet (Krol, 1994), how to access the various services that the Internet carries (Krol, 1994) and how to set up your own Internet facilities (Liu *et al.*, 1994). Once you have made the initial connection to the Internet, the Internet itself becomes a reservoir of information about how to navigate, search and retrieve data. The Internet is the most current source of information and this should be borne in mind as you read the chapter.

Although security is beyond the scope of this chapter, readers should be reminded that the confidentiality of any data accessible via the Internet can be compromised unless appropriate measures are taken (Garfinkel and Spafford, 1995; Zimmerman, 1995). In addition, while the Internet is global, legal protections for privacy can change abruptly when borders are crossed.

12.2. Electronic Information Dissemination Methods

12.2.1. Electronic mail, mailing lists, bulletin boards and newsgroups

Perhaps the simplest form of communication is electronic mail, or email. Email is commonly used to exchange information between individuals, including simple messages, news, data and software. Many academics access email through a free or inexpensive connection provided by their institution; however, commercial connections (such as via CompuServe) are widely available. Although email is sometimes criticized as frivolous, it greatly simplifies collaborations in which manuscripts and grant pro-

posals must be exchanged, and reduces the frustrations of communicating with colleagues in different time zones.

Email can also be used to contact members of a group automatically from a central mailing address and establish what is essentially a small electronic community. Group mailers are simple to set up and tend to be used for communication between small numbers of users. Access to messages is controlled by a distribution list. For example, the Biodiversity and Ecosystems Network (BENE) (58) maintains a mailing list carrying discussions concerning conservation, ecosystem protection, and management. A mailing list need not be permanent; it can be used temporarily to coordinate any group working together on a project.

Larger interest groups tend to communicate through bulletin boards and/or newsgroups. These are public forums for posting to interested parties, many of whom may not be known to each other. Thousands of specialized newsgroups exist and can be subscribed to free of charge. A significant number are of interest to biologists. Further information including how to access a list of the 'bionet' newsgroups can be obtained from the bionet Frequently Asked Questions (FAQ) list (12). There is no 'typical' newsgroup but messages dealing with technical questions, meetings, job openings, and protocols are common. Messages are archived and indexed and can be queried easily (59) and thus need not consume precious space on your local computer. It is possible to subscribe to a bionet newsgroup and receive its postings as part of your regular email. Many users, however, prefer to use newsreader software which allows easy browsing, message selection, and posting. The creation of a bionet newsgroup is a formal process, involving a proposal, discussion and voting, all of which occur on-line. Nonetheless it is not difficult to establish a new group and this is the logical next step when a community outgrows its mailing list.

12.2.2. File Transfer Protocol (FTP)

FTP is a protocol for depositing and retrieving data to and from remote machines. It is the method of choice for distributing free software as well as large files that would be impossible to send via email. For example, all of GenBank is available from the National Center for Biological Information (NCBI) (7) and a variety of software for biologists can be found at the European Bioinformatics Institute (EBI) (46). Catalogues, manuscripts, photographs and documentation are also distributed in this fashion. FTP sites come and go and so the Internet itself is the place to find the most up-to-date lists. FTP has a flexible set of commands but in practice only a few are needed (Krol, 1994). Nearly all FTP servers will respond with primitive command documentation if 'help' or '?' is typed at the FTP prompt.

Establishing an FTP site of your own may be worthwhile if you find

yourself responding to large numbers of requests for information from your laboratory or project. In this case FTP will relieve some of the burden and may in fact increase the audience for the data. For example, a germplasm resource centre can distribute a seed catalogue via FTP and avoid the costs incurred by printing and mailing (10).

12.2.3. Telnet

Another basic communication procedure used on the Internet is telnet, which makes it possible to log in to a remote host and conduct an interactive session. A username and password are required (these may be obtained either by previous arrangement, or may be widely advertised). Although the use of telnet for information distribution has become less common, some sites offer it to provide specialized or proprietary services which cannot be accessed via Gopher or the World Wide Web (WWW; see below), or which require password protection. For example, the Bath Information and Data Services (BIDS) (35), the bibliographic data service for higher education and research, is accessible through telnet and the Arabidopsis Information Management System (61) maintains telnet access to its database at Michigan State University, although Gopher and WWW interfaces are also available.

12.2.4. Gopher

Gopher, originally developed at the University of Minnesota, is a lookup tool which facilitates navigation through the Internet regardless of the type of file or the geographical location of the information (Krol, 1994). Information at a Gopher site is presented in a hierarchy of files and directories which are accessed via menus. Like all of the protocols we shall discuss it works on a client-server basis. The server, the computer on which the information resides, responds via server software to queries issued by client software, which is installed on the user's computer. This allows for fast and easy access of information since the intensive tasks of storing and sorting large data collections are all handled by the server. Client and server software are freely available and can be obtained via FTP from many different sites (Krol, 1994).

Although in principle the information distributed via a Gopher server could be made available by FTP, there are two significant differences from the user's point of view. The first is convenience. Gopher navigation does not involve arcane commands and often requires nothing more than the use of the arrow and ENTER keys, or the mouse. Second, and of even greater importance, is that one Gopher site can 'point to' another. A given

menu may therefore list resources from related sites all over the world. The result is an environment in which the user sees information 'space' as richly interconnected. As will be discussed below, this theme is an essential part of the World Wide Web, which has largely superseded Gopher as the major information distribution mechanism on the Internet.

Gopher is still an important tool for biologists. Information available via more sophisticated methods is often distributed in parallel via Gopher to ensure that users with slow Internet connections, or text-only terminals, are not disenfranchised. For example GrainGenes (28), a database for the *Triticeae*, has active Gopher (Fig. 12.1) and WWW sites, which distribute information from the database, plus additional information about wheat cultivars, oat pedigrees, etc. Many libraries, including the British Library (36), offer Gopher access to their catalogues and other resources.

Gopher servers often provide access to two search tools. Veronica

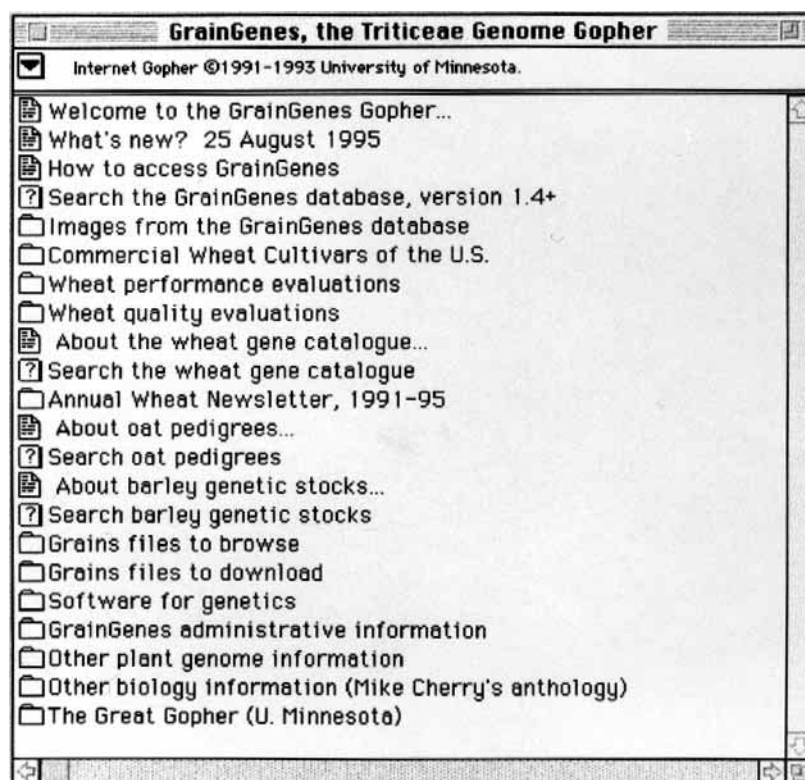


Fig. 12.1. The GrainGenes project Gopher server. Information is presented as a hierarchy of files and directories. Text documents are indicated by an icon of a sheet of paper, a folder indicates a directory of files and the box containing a '?' indicates that this is a searchable index.

(very easy-rodent orientated net-wide index to computerized archives), is accessed using the Gopher client software and can be used to search for resources by looking for words that have been used in titles of documents or directories on Gopher servers, WWW servers, usenet archives, and telnet-accessible information services. Veronica constantly searches these resources, indexes this information and stores the location of the title within cyberspace in a large database. When a keyword query is made using veronica, the combined index is scanned within the database and the various locations of the information are presented in the form of Gopher directories which can be selected as in browsing. Veronica has its own FAQ (13) and on-line help for constructing queries (14). Wide Area Information Service (WAIS) is also a tool for searching through indexed material and is ideal for working with collections of datasets or databases. WAIS can be used to create an index of every word in a document, which makes this a very powerful tool as any word can then be used in a query. Generalized queries can be made through the use of wild cards (*), or more refined queries can be made by the use of booleans like AND, OR and NOT. A queryable index of WAIS sources is available on-line (15). Many databases covering plant genome information have been indexed for searching. Other information resources, such as abstracts of grants that have been funded at USDA, DOE, NIH and NSF, can also be searched in this way (39).

Setting up a Gopher server is useful if you have a well-organized set of documents to offer. For example, a collection of protocols, abstracts, photographs, tables or genetic maps are compatible with presentation via Gopher. WAIS can be used to support simple searches.

12.2.5. World Wide Web (WWW)

The World Wide Web represents the state of the art in terms of information dissemination through the Internet, and is the primary means by which most providers distribute information today. This system was developed at CERN (European Laboratory for Particle Physics), Geneva, Switzerland (74). One of its major features is that it is possible to access Gopher, FTP, WAIS, and newsgroups through the WWW using the same client software, which also functions as a browser for 'native' WWW documents. In this sense the WWW is a prototype for a uniform user interface for Internet services. Many different client applications are available for the WWW but arguably the current standard is Netscape Navigator (9), developed by Netscape Communications Corporation, which is available as freeware to educational establishments and can be bought very reasonably by commercial organizations. Netscape has produced an introduction to the Internet and a series of on-line tutorials about how to navigate the Internet which are an excellent starting point for information. Each file on the

WWW has its own identifier or address, the universal (uniform) resource locator, or URL, which conforms to the following address structure:

`http://machine address:port/directory/filename`

where http stands for the hypertext transfer protocol. When using a WWW browser, URLs are also used to access Gopher, FTP, etc. In this case the address changes and http is exchanged for whichever protocol is being used. For example to access a Gopher server the URL will have the structure gopher://machine address/directory/filename.

As mentioned earlier, a major feature offered by Gopher servers is the ability to reference different resources, at the same or different network sites, on a single menu. The WWW extends this idea to include references from *within* documents. A single document can contain a large number of such 'hypertext links' to other documents, which themselves may be linked in various ways. The links need not involve other text documents but can reference sounds, images and movies. The structure of hypertext documents can be simple and the underlying markup code (hypertext markup language, or html) is quickly learned (78). Many HTML editors are now available, either as freeware, shareware or commercially. Products of this kind are going to increase in quality and number. A current review of the latest editors for Unix, Macintosh and PCs can be accessed on-line (76).

Hypertext has revolutionized the way that information providers can think about how information can be linked and displayed. Using Gopher and WAIS, information is displayed as plain text; images are presented separately. On the WWW images can be embedded in a text document. In addition, text can be formatted to be italicized or bold, fonts can be chosen and font size selected. The appearance of the final product, as viewed by the user, however, depends on how the client software interprets HTML as well as on the actual HTML codes inserted during markup. Information can now be published in a professional style akin to that previously found only in hard-copy publications. More importantly, not only can flat files, such as an on-line catalogue, be published electronically, but also entire databases can be presented and networked through this medium and users can interact with them in a variety of ways. These features are in large part responsible for attracting large numbers of commercial interests to the Internet, which prior to the WWW was mainly the domain of researchers and students.

Several search engines have been developed to allow for easy access of information through the WWW. A review of these can be accessed on-line (42). The various search facilities have been designed to search at different levels, i.e. some search directory titles, others in file titles, or in the files themselves. The way the result of a query is presented also differs. Some return just a link to the various resources which contain the

keywords used in the search, whilst others will return the link, plus a short résumé about the link that has been found.

12.3. Resources on the WWW

There is currently a wealth of information concerning floras, herbaria, journals, molecular databases, plant genome databases, genetic stock resources, botanical gardens, professional societies, conference proceedings, biotechnology resources and companies distributed through the WWW. It is impossible to review them all here. Key starting points to finding information in this subject area are given in the appendix and a few examples will be discussed below.

In the area of taxonomy and biodiversity, a good starting point is the Biodiversity and Biological Collections web server (79), which carries an array of links to biodiversity databases, listserv archives, taxonomic authority files and software. Another good resource is the Biodiversity and Ecosystems server (80), which provides a wide array of information spanning both the plant and animal kingdom. The Herbaria Type Specimen Data Database (HUHP) (81), developed at Harvard, contains information relating to the 80 000 type specimens of vascular plants in the Harvard Herbaria; access is also available to the 300 000 vascular plant names carried on the Gray Herbarium Card Index. Many botanic gardens distribute information relating to their research activities as well as on-line garden tours. In particular, the Royal Botanic Gardens at Kew (5) and the Missouri Botanical Gardens (6) are good sites to visit.

The Germplasm Resources Information Network (GRIN) server (66) provides germplasm information about plants, animals, microbes and invertebrates within the National Genetic Resources Program of the US Department of Agriculture's (USDA) Agricultural Research Service (ARS). Country Plant Genetic Resources Information can also be accessed from the FAO International Conference and Programme on Plant Genetic Resources (ICPPGR) WWW server (64). UK Plant Genetic Centres can be accessed through the UK Plant Genetic Resources Group WWW server (70).

Several resource centres, including the Nottingham Arabidopsis Stock Centre (NASC) (69) and the Maize Genetics Cooperation Stock Center (67) have WWW servers which distribute on-line catalogues. Forms are used to allow customers to communicate subscription information, information regarding seed donations, or as a mechanism to ask questions or place orders. Commercial culture collections such as the ATCC (62) and biotechnology companies (3) also advertise through the use of on-line catalogues, signifying the advent of electronic merchandising.

Many plant genome specific databases are being developed as part of



Plant Genome Databases at NAL

[[What's New?](#) (12 September 1995) | [Services](#) | [How to...](#) | [Resources](#)]

Search Databases

Plant Genome Databases [[query](#) | [about](#)]

- AAtDB--*Arabidopsis* [[browse](#) | [query](#) | [about](#)]
- Alfagenes--alfalfa (*Medicago sativa*) [[browse](#) | [query](#) | [about](#)]
- BeanGenes--*Phaseolus* and *Vigna* [[browse](#) | [query](#) | [about](#)]
- ChlamyDB--*Chlamydomonas reinhardtii* [[browse](#) | [query](#) | [about](#)]
- CottonDB--*Gossypium hirsutum* [[browse](#) | [query](#) | [about](#)]
- GrainGenes--wheat, barley, rye and relatives [[browse](#) | [query](#) | [about](#)]
- MaizeDB--maize [[browse](#) | [query](#) | [about](#)]
- RiceGenes--rice [[browse](#) | [query](#) | [about](#)]
- SolGenes--Solanaceae [[browse](#) | [query](#) | [about](#)]
- SorghumDB--*Sorghum bicolor* [[browse](#) | [query](#) | [about](#)]
- SoyBase--soybeans [[browse](#) | [query](#) | [about](#)]
- TreeGenes--forest trees [[browse](#) | [query](#) | [about](#)]

- Mendel--plant-wide gene names [[browse](#) | [query](#) | [about](#)]

Documents

- [Newsletters, Journals and other Publications](#)
- [ACEDB Documentation](#)
- [Collaborators](#)

To the [[Animal Genome](#) | [Other Genome](#) | [Biological](#) | [All](#)] Databases

Genome Informatics Group / 1 September 1995

Fig. 12.2. The NAL Agricultural Genome Information Service (AGIS) page to Plant Genome Databases. Items which are underlined are hypertext links to further information. From this central service each database can be individually browsed or searched, in any combination.

the respective genome programmes and are distributed through Gopher and/or the WWW by the development sites (see appendix for individual addresses). Moreover, as part of the USDA's project for genome analysis these resources, which include 12 plant databases, have been centralized for distribution at the National Agricultural Library (NAL) (23) (Fig. 12.2). Here it is possible to cross-query the various plant databases. Importantly other genome databases are also distributed from NAL and so can also be

included in the querying process. Relevant documentation and publications from the Cucurbita Genetics Cooperative, the Tomato Genetics Cooperative, the Rice Genetics Newsletter and *Weeds World*, the electronic *Arabidopsis* newsletter, and abstracts from previous plant genome conferences are all distributed from this site.

A good starting point for information relating to DNA, RNA or proteins is either EMBL/EBI (47/48) or GenBank (52), who collaborate with the DNA Database of Japan (DDBJ) (45) to provide all the publicly available nucleic acids information, as well as comprehensive links to the other major molecular biology databases (Table 12.1). These databases can be

Table 12.1. Molecular biology databases available through the European Bioinformatics Institute (EBI).

3D-ALI	Structure-based sequence alignments
ALU	ALU sequences and alignments
BERLIN	5S rRNA sequences
BLOCKS	Protein Blocks Database
CPGISLE	CpG islands database
CUTG	Codon usage tabulated from GenBank
DSSP	Secondary structure digests of PDB files
EMBL	Nucleotide sequence database
ECDC	<i>E. coli</i> database collection
ENZYME	Database of EC nomenclature
EPD	Eukaryotic promoter database
FANS-REF	Functional analysis bibliography
FLYBASE	<i>Drosophila</i> genetic map database
HAEMB	Haemophilia B database of mutations
HLA	HLA class I and II sequence database
HSSP	Homology-derived protein structures
Kabat	Proteins of immunological interest
LIMB	Listing of molecular biology databases
METHYL	Site-specific methylation
PDB	Brookhaven protein 3D structures
PKCDD	Protein kinase catalytic domains
PROSITE	Protein pattern database
RELIB	Restriction enzyme library
REBASE	Restriction enzyme database
RLDB	Reference Library Database
RRNA	Small subunit rRNA sequences
SEQANALREF	Sequence analysis bibliography
SMALLRNA	Small RNA sequences
SRP	Signal recognition particle database
SWISS-PROT	Protein sequence database
TFD	Transcription Factor Database
TRANSTERM	Translation termination signals
TRNA	tRNA sequences

accessed from the EBI, the home of the EMBL data library, through the WWW, exploiting the concept of 'one stop shopping', even though the databases are held at various sites around the world. These databases can be cross-queried at the EBI through the use of the Sequence Retrieval System (SRS) searching tool (54). Similarly, GenBank has developed the WWW Entrez browser to allow cross-querying amongst their services, including the Bibliographic Database MedLine (55). On-line tutorials are provided for each of these search methods. In addition to search facilities, it is possible to analyse data at these two sites, e.g. the Basic Local Alignment Search Tool (BLAST) suite of tools (44) allows protein predictions to be made on a DNA sequence. Many other molecular biology servers exist, including the ExPASy Molecular Biology server at Geneva (48), which is an excellent repository for documents concerning molecular biology-orientated services, software, email servers, and FTP sites.

12.4. Providing Information via the WWW

Setting up a WWW server is appropriate if you wish to offer anything from the simplest documents to an interface to a database. Encouragingly, most data providers would agree that the server software is easier to understand and maintain than that for Gopher. WWW servers can be run from Unix, Macs and PCs. Information on how to establish servers is available on-line (77). If maintaining an in-house server is not practical, commercial organizations can establish and maintain for-fee servers.

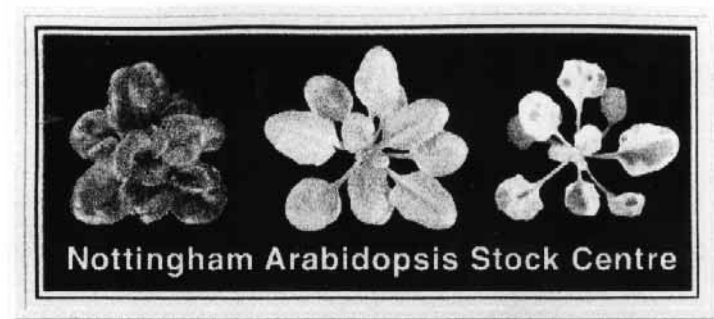
Before establishing such a service several issues should be considered. One issue concerns the range of services. Will information be accessible only through the WWW (note that WWW access alone may prevent access by some users), or will text versions be made available for downloading by FTP or Gopher? Will WAIS indexing be used for simple querying? Another issue concerns organization: how the information will be structured. This is a complex matter involving taste as well as ease of use and maintenance. Will the material be broken down into related documents and how will they be linked? Will a database be used? Will the resource grow without limit or will it be static? Will it contain links to other information either on or off site and other related sites? If images are to be distributed how will they be displayed and linked? Will the resource offer interactive features such as on-line ordering and question asking, or have electronic donation forms for information submission? To answer these questions satisfactorily will require a fair amount of prototyping but this can make the difference between a site that is useful versus one that is not. It is recommended that part of the design process involve on-line visits to a variety of WWW sites, with each reviewed pragmatically for adoptable features.

12.4.1. Flat files – an on-line germplasm catalogue

When all the issues are considered, even a simple on-line catalogue can be very sophisticated in the way information is displayed but very easy for users to access the information. The WWW site maintained by the Nottingham Arabidopsis Stock Centre (69) is an example of a resource which has been designed to provide the functionality of a traditional hard-copy catalogue, with a series of flat files describing stocks. The individual stock descriptions, however, are more expansive and detailed than those carried in the hard copy, and are linked via hypertext to other documents within the catalogue. The information is therefore split into a series of 'layers' through which the user can navigate. For example, to browse the catalogue, the user selects 'Browse the Complete Seed List' on the home page (Fig. 12.3), which opens the contents page of the catalogue. This lists the various categories of lines that are available. From the options listed it is possible to select 'Biochemical mutants', which opens a new page to the list of biochemical mutants, giving the Stock Number, the locus name and map position. Selecting the Stock Number opens a detailed description of the individual line with links to the background, mutagen, donor, etc. From this page it is possible to select a link which opens the Order Form, which can be filled in and 'posted' on-line. Other fill-in forms (built in to html) are employed for users to communicate with the Centre either to obtain or provide information. Furthermore, linked into the home page for the Centre are links to other *Arabidopsis* resources, which invite users to navigate amongst the other relevant information. There are many advantages to distributing the catalogue in this way. The information is constantly updateable and available to users; it is possible to distribute colour images of stocks linked to very detailed descriptions, and the audience that can be reached is much larger than that of a hard copy. This adds up to a very cost-effective and user-friendly information dissemination mechanism.

12.4.2. Databases – genome information servers

For some kinds of information a flat-file representation is impractical. For example, the scope of the typical plant genome database includes details about loci, genetic maps, germline resources, DNA sequences, physical maps, bibliographic references, etc., all interlinked and searchable, with individual items being modified as new information is made available to the curators. In this case, the maintenance of a copy of the database as a large set of individual hypertext files with one file per item would be a significant burden, although technically possible. Fortunately another solution exists. A database can be connected to the WWW in a way that eliminates the need to maintain objects as pre-existing text files. These



Welcome to the [Nottingham Arabidopsis Stock Centre \(NASC\) World Wide Web](#) server for *Arabidopsis* information. From here, you can browse the seed list, order seed stocks, ask questions, register with the Stock Centre, and access other *Arabidopsis* information servers -- all with the click of a button.

Nottingham Arabidopsis Stock Centre on-line catalogue

- [Browse the Complete Seed List](#)
- [Text search the Seed List \(through October 94\)](#)
- [New Releases November 94](#)
- [Browse the Picture Book](#)
- [Print out your own copy of the Seed List](#) - formatted for the Mac **NEW**
- [Print out your own copy of the Seed List](#) - as plain text **NEW**
- [ORDER SEEDS](#)
- [Request a hard-copy Seed List](#)
- [Register my name with the Stock Centre](#)
- [Send a question to NASC](#)
- [Seed Donation Form](#) **NEW**

Fig. 12.3. The Nottingham Arabidopsis Stock Centre on-line catalogue. A section of the homepage for the NASC WWW server. A list of hypertext links detail the options on how to access the information carried in the catalogue.

systems deliver information on demand by extracting it directly from the database and formatting it on the fly. In doing so they not only replace the flat-file representation, they provide the power of a database management system as well, including the ability to submit queries.

Several database management systems can be configured in this way, including commercial relational databases such as Sybase and Oracle (e.g. MaizeDB uses Sybase; see 29). Noncommercial software can be used as well and many providers of genome information for both plants and animals have chosen a free object-orientated system known as ACEDB (Durbin and Thierry-Mieg, 1991). This option is discussed here.

Sequence: A08599

A08599 Quit

Title Synthetic nucleotide sequence 1kb
fragment containing the Arabidopsis
thaliana 2S albumin gene

Type DNA **A08599** 1028

Structure Length 1028

Subsequence **NCBI gi: 411667** 432 926

Library **GenBank** A08599 A08599

Sequence_group Other_sequence

Origin Date 08-SEP-1993

General Gene **Arabidopsis thaliana 2S albumin**

Remark Patent: WO 8903887-A 18 05-MAY-1989;
None Artificial sequences.



Database: AAtDB

[*Arabidopsis thaliana* | 3 Aug 1995]**Neighbours** --- [Link to other classes](#)**Class Sequence A08599**

Title Synthetic nucleotide sequence 1kb fragment containing the Arabidopsis
thaliana 2S albumin gene

Type DNA **A08599** 1028

Structure Length 1028

Subsequence **NCBI gi: 411667** 432 926 ([GenBank](#))

Library **GenBank** A08599 ([AGIS](#)) ([BLAST](#)) ([EMBL](#)) ([GenBank](#)) ([GeneBase](#))

Sequence_group Other_sequence

Origin Date 08-SEP-1993

General Gene **Arabidopsis thaliana 2S albumin**

Remark Patent: WO 8903887-A 18 05-MAY-1989;
None Artificial sequences.

Entrez Reports

NCBI	Entrez	Report	Fasta	AGIS	MEDLINE	Protein
DNA	Save	MacSave	Help			

LOCUS A08599 1028 bp DNA PAT 08-SEP-1993

DEFINITION Synthetic nucleotide sequence 1kb fragment containing the
Arabidopsis thaliana 2S albumin gene.

ACCESSION A08599

KEYWORDS .

SOURCE unidentified.

ORGANISM unidentified.

REFERENCE 1 (bases 1 to 1028)

AUTHORS .

JOURNAL Patent: WO 8903887-A 18 05-MAY-1989;

COMMENT NCBI gi: 411666

FEATURES

source Location/Qualifiers

1..1028

/organism="Artificial sequences"

432..926

/gene="Arabidopsis thaliana 2S albumin"

/note="NCBI gi: 411667"

/codon_start=1

/translation="MANKLFLVCAALALCFLLTNASIYRTVVEFEEDDATNPICPKMR
KCRKEFQKEQHLRACQQLMLQQARQGRSDEFDFEDDMENPQQQQEQQLFQQCCNELR
QEEPDCVCPTLNQAAKAVRLQQQHQPMQVRRIYQTAKHLPNVCDIPQVDVCPFNIPSF
PSFY"

The tasks involved in configuring an ACEDB database for a new species are covered in detail by on-line documentation (71). Although configuring any database can be a daunting affair, ACEDB does not require advanced computer expertise or the ability to program. The major issues facing the new curator are more likely to revolve around data acquisition and organization. Once data are loaded correctly, ACEDB can be distributed as a stand-alone database for Unix or Macintosh systems; built in to the software are graphical displays which handle genetic maps, physical maps, DNA sequences and other data.

From the user's point of view, the major characteristic of ACEDB is that it is a hypertext system. In nearly every context, including the graphical displays, a mouseclick on an item retrieves additional information. ACEDB therefore translates naturally to the WWW environment, which uses the same strategy to link related data (Fig. 12.4). The connection of an ACEDB database to the WWW involves establishing a 'gateway interface' (75) to allow the ACEDB software to communicate with the WWW server software. Two such gateways are available (72, 73) and many sites worldwide now offer ACEDB databases in this form. As mentioned earlier, plant data are particularly well covered by the site maintained by the National Agricultural Library, USDA (30).

Once a database is configured as a 'web server' it is straightforward to link its information to other network resources. For example, in the case of the plant databases, a germplasm record containing a GRIN accession number can be linked directly with the relevant record available from GRIN's WWW server, or a DNA sequence can be linked via its accession number to a sequence record at NCBI's GenBank (Fig. 12.4b). Even if the database does not contain the full URL, this can be inserted 'on the fly' before the data leave the server. As a corollary, as soon as a database is connected in this fashion, other resources can point back to it from their own records, creating a two-way link (Fig. 12.4c). Such links add value by connecting related information and making it easier to find. This simple but powerful strategy is representative of a trend which may someday see on-line databases integrated to a much larger degree.

Fig. 12.4. (opposite) Connection of an ACEDB database to the WWW. (a) A text window from a stand-alone database, showing part of a sequence record. Items in bold are hypertext links to other information in the database. (b) The same data viewed on the WWW at the Agricultural Genome Information Server (AGIS). The original links, now underlined, are preserved. Additional links have been added 'on the fly' to servers maintained by DNA and protein sequence repositories. (c) A sequence record as delivered from GenBank. The 'AGIS' button at the top is a link back to the AGIS sequence.

12.5. The Future

The pace of progress in computer technology makes predictions difficult, even over the short term, except for the most general: computers will become more powerful, software more feature-laden, and networks faster. Biologists will benefit from new methods for representing data. For example, Sun Microsystems is developing a new programming language called Java (40), designed to allow WWW browsers to download and run software. This greatly extends the possible ways in which users can interact with and analyse data.

For the biologist it is not the technical specifics that are important but rather their consequences: more and more information will be made available in electronic form. This is a relentless trend, driven by the cost advantages of new technologies and the inability of older methods to handle vast amounts of data. Since electronic dissemination offers timeliness and (in some cases) exclusive access, biologists who exploit it appropriately will enjoy an advantage.

Appendix of Internet Addresses

Biotechnology resources

- (1) Biotechnology sources “<http://ra.cs.ohiou.edu/gopher/non-academic/tto/biotech.html>”
- (2) Biotech Education Home Page
“http://biotech.zool.iastate.edu/Biotech_Public_Ed.html”
- (3) Biotechnology Companies “<http://www.informatik.uni-rostock.de/HUM-MOLGEN/biotech/companies/>”
- (4) Biotech Company list “<http://155.41.115.114/jumper/wfgo.html>”

Botanical gardens

- (5) Kew Gardens “<http://www.rbgekew.org.uk/>”
- (6) Missouri Botanic Gardens
“<http://www.mobot.org/welcome.html>”

FTP sites

- (7) GenBank “ncbi.nlm.nih.gov in the genbank directory”
- (8) Gopher “<gopher://boombox.micro.umn.edu:70/11/gopher>”
- (9) Netscape Navigator “<http://home.netscape.com/>”
- (10) Nottingham Arabidopsis Stock Centre Seed List “nasc.nott.ac.uk in the/pub/SeedList directory”
- (11) WAIS “<http://ls6-www.informatik.uni-dortmund.de/freeWAIS-sf/>”

Frequently asked questions (FAQ) files and archives

- (12) BioSci FAQ “gopher://gopher.bio.net:70/00/doc/biosci.FAQ”
- (13) Veronica FAQ
“gopher://veronica.scs.unr.edu/00/veronica/veronica-faq”
- (14) Veronica queries
“gopher://veronica.scs.unr.edu:70/00/veronica/how-to-query-veronica”
- (15) Index to WAIS sources “gopher://gopher-gw.micr.umn.edu:70/1/WAISes/Everything”

General starting points (links to many sites)

- (16) Harvard Biological Labs Home Page “http://golgi.harvard.edu/”
- (17) USGS Index of Biological Servers
“http://info.er.usgs.gov/network/science/biology/index.html”
- (18) ANU Bioinformatics Hypermedia Service
“http://life.anu.edu.au/”
- (19) Pedro’s Research Tools
“http://www.public.iastate.edu/~pedro/research_tools.html”
- (20) WWW sites of interest to botanists (very comprehensive)
“http://meena.cc.uregina.ca/~liushus/bio/botany.html”
- (21) A collection of botany-related URLs – an excellent listing
“http://www.helsinki.fi/~rlampine/botany.html”
- (22) Yahoo Hierarchical Hotlist Summary
“http://akebono.stanford.edu/yahoo/all.html”

Genome servers

- (23) Agricultural Genome Information Server (AGIS)
“http://probe.nalusda.gov:8000/index.html”
- (24) AtDB – the *Arabidopsis thaliana* database, Stanford
“http://genome-www.stanford.edu/Arabidopsis/”
- (25) CottonDB Data Collection Site, TAMU
“http://algodon.tamu.edu/”
- (26) Dendrome – a genome database for forest trees, USDA
“http://s27w007.pswfs.gov/”
- (27) GrainGenes – a database for small grains and sugar-cane, USDA
“http://wheat.pw.usda.gov/graingenes.html”
- (28) GrainGenes Gopher server “gopher://greengenes.cit.cornell.edu/”
- (29) Maize Genetic Database, Missouri
“http://www.agron.missouri.edu/top.html”
- (30) Plant Genome Databases at AGIS
“http://probe.nalusda.gov:8300/index.html”
- (31) Plant Genome and Data Information Center, NAL
“http://www.nalusda.gov/answers/info_centers/pgdic/pgdic.html”
- (32) Rice Genome Research Program (RGP), Japan

[“http://www.staff.or.jp/”](http://www.staff.or.jp/)

(33) Soybase, ISU [“http://mendel.agron.iastate.edu:8000/main.html”](http://mendel.agron.iastate.edu:8000/main.html)

(34) Australian National Genomic Information Service (ANGIS)

[“http://morgan.angis.su.oz.au/”](http://morgan.angis.su.oz.au/)

Libraries

(35) Bath Information and Data Services (BIDS) [“telnet://bids.ac.uk”](telnet://bids.ac.uk)

(36) British Library [“gopher://portico.bl.uk”](gopher://portico.bl.uk)

Locating colleagues

(37) Yellow Pages – Menu [“http://www.index.org/”](http://www.index.org/)

(38) White Pages [“http://home.netscape.com/home/internet-white-pages.html”](http://home.netscape.com/home/internet-white-pages.html)

Miscellaneous

(39) Abstracts of grants

[“http://probe.nalusda.gov:8000/otherdocs/projects.html”](http://probe.nalusda.gov:8000/otherdocs/projects.html)

(40) Hot Java [“http://java.sun.com/”](http://java.sun.com/)

(41) Index to plant databases

[“http://probe.nalusda.gov:8300/plant/index.html”](http://probe.nalusda.gov:8300/plant/index.html)

(42) WWW search engines review [“http://cuiwww.unige.ch/meta-index.html”](http://cuiwww.unige.ch/meta-index.html)

Molecular and informatics servers

(43) Bio Resources, Alphabetical Listing “

[http://www.library.wisc.edu/Biotech/resource/cumulative.html”](http://www.library.wisc.edu/Biotech/resource/cumulative.html)

(44) BLAST tools [“http://www.ncbi.nlm.nih.gov/BLAST/index.html”](http://www.ncbi.nlm.nih.gov/BLAST/index.html)

(45) DNA Database of Japan (DDBJ) [“http://www.nig.ac.jp/”](http://www.nig.ac.jp/)

(46) European Bioinformatics Institute (EBI)

[“http://www.ebi.ac.uk:80/”](http://www.ebi.ac.uk:80/)

(47) European Molecular Biology Laboratory (EMBL)

[“http://www.embl-heidelberg.de/”](http://www.embl-heidelberg.de/)

(48) ExPASy Molecular Biology Server [“http://expasy.hcuge.ch/”](http://expasy.hcuge.ch/)

(49) GenoBase Database Gateway, NIH

[“http://specter.dcrn.nih.gov:8004/”](http://specter.dcrn.nih.gov:8004/)

(50) Johns Hopkins University Bioinformatics Server

[“http://www.gdb.org/hopkins.html”](http://www.gdb.org/hopkins.html)

(51) Keck Center for Genome Informatics

[“http://keck.tamu.edu/cgi/cgi.html”](http://keck.tamu.edu/cgi/cgi.html)

(52) National Center for Biotechnology Information (NCBI) – GenBank

[“http://www.ncbi.nlm.nih.gov/”](http://www.ncbi.nlm.nih.gov/)

(53) National Center for Genome Resources (NCGR) – GSDB

[“http://www.ncgr.org/”](http://www.ncgr.org/)

(54) Sequence Retrieval System (SRS) searching tool

["http://www.ebi.ac.uk/srs/srsc"](http://www.ebi.ac.uk/srs/srsc)

(55) WWW Entrez ["http://www3.ncbi.nlm.nih.gov/Entrez/"](http://www3.ncbi.nlm.nih.gov/Entrez/)

Newsgroups and listservers

(56) List of all the news groups ["news:*"](#)

(57) List of all newsgroups relating to biology ["news://bionet.*"](#)

(58) Biodiversity and Ecosystems Network

["http://straylight.tamu.edu/bene/home/bene.email.html"](http://straylight.tamu.edu/bene/home/bene.email.html)

(59) BioSci archive of mail messages

["gopher://gopher.bio.net/77/.wais-sources/biosci/"](http://gopher.bio.net/77/.wais-sources/biosci/)

Resource centres

(60) Arabidopsis Biological Resource Center (ABRC), MSU

["telnet://aims.cps.msu.edu"](http://aims.cps.msu.edu)

(61) Arabidopsis Information Management System (AIMS) – database to ABRC ["telnet://genesys.cps.msu.edu"](http://genesys.cps.msu.edu)

(62) American Type Culture Collection (ATCC)

["http://www.atcc.org/"](http://www.atcc.org/)

(63) Culture Collections ["http://www.hgmp.mrc.ac.uk/Public/culture-collections.html"](http://www.hgmp.mrc.ac.uk/Public/culture-collections.html)

(64) FAO International Conference and Programme on Plant Genetic Resources (ICPPGR) ["http://web.icppgr.fao.org/"](http://web.icppgr.fao.org/)

(65) Genome Centres ["http://www.hgmp.mrc.ac.uk/Public/genome-centres.html"](http://www.hgmp.mrc.ac.uk/Public/genome-centres.html)

(66) Germplasm Resources Information Network (GRIN)

["http://www.ars-grin.gov/"](http://www.ars-grin.gov/)

(67) Maize Genetics Cooperation – Stock Center, UIUC

["http://w3.ag.uiuc.edu/maize-coop/mgc-home.html"](http://w3.ag.uiuc.edu/maize-coop/mgc-home.html)

(68) National Plant Germplasm System (NPGS) ["http://www.ars-grin.gov/"](http://www.ars-grin.gov/)

(69) Nottingham Arabidopsis Stock Centre ["http://nasc.nott.ac.uk/"](http://nasc.nott.ac.uk/)

(70) UK Plant Genetic Resources Group (a link to many centres)

["http://nasc.nott.ac.uk:8200/"](http://nasc.nott.ac.uk:8200/)

Resource creation documentation

(71) ACEDB set up documentation

["http://probe.nalusda.gov:8000/acedocs/index.html"](http://probe.nalusda.gov:8000/acedocs/index.html)

(72) ACEDB Moulon server

["http://moulon.inra.fr/acedb_conf_eng.html"](http://moulon.inra.fr/acedb_conf_eng.html)

(73) ACEDB WWW interface

["http://probe.nalusda.gov:8000/tools/acelib.html"](http://probe.nalusda.gov:8000/tools/acelib.html)

(74) CERN ["http://www.w3.org/"](http://www.w3.org/)

(75) Common Gateway Interface ["http://hoohoo.ncsa.uiuc.edu/cgi/"](http://hoohoo.ncsa.uiuc.edu/cgi/)

(76) HTML editors

“<http://www.w3.org/hypertext/WWW/Tools/Overview.html>”

(77) How to create web services

“<http://home.netscape.com/home/how-to-create-web-services.html>”

(78) Learning html “<http://www.ncsa.uiuc.edu/demoweb/html-primer.html>”

Systematics and taxonomy servers

(79) Biodiversity and Biological Collections WWW Server

“<http://muse.bio.cornell.edu/welcome.html>”

(80) Biodiversity and Ecosystems network

“<http://straylight.tamu.edu/bene/home/bene.information.html>”

(81) The Herbaria Type Specimen Data Database (HUHP) Harvard

“gopher://huh.harvard.edu/11/collections_info/huh”

(82) International Organization for Plant Information (IOPI)

“<http://life.anu.edu.au/biodiversity/iopi/iopi.html>”

(83) Systematics and Taxonomy

“<http://straylight.tamu.edu/bene/systematics.html>”

References

- Durbin, R. and Thierry-Mieg, J. (1991–) A *Caenorhabditis elegans* database. Documentation, code and data available from anonymous FTP servers at lirmm.lirmm.fr, cele.mrc-lmb.cam.ac.uk, and ncbi.nlm.nih.gov.
- Garfinkel, S. and Spafford, G. (1995) *Practical UNIX Security*. O'Reilley & Associates, Sebastapol, CA, 512 pp.
- Krol, E. (1994) *The Whole INTERNET User's Guide and Catalog*, 2nd edn. O'Reilley & Associates, Sebastapol, CA, 543 pp.
- Liu, C., Peek, J., Jones, R., Buus, B. and Nye, A. (1994) *Managing INTERNET Information Services*. O'Reilley & Associates, Sebastapol, CA, 630 pp.
- Zimmerman, P.R. (1995) *The Official PGP User's Guide*. MIT Press, Cambridge, MA, 127 pp.

Index

- Ac/Ds 217–219
ACEDB 293, 295 *see also* database
active / working collections 105, 111
adaptation 62, 65
Advanced Technology Program (ATP) 34, 36, 37
AFLP *see* amplified fragment length polymorphisms
AGIS 289
Agrobacterium 204, 207–213, 222
agroecogeographic analysis, 60
Agropyron elongatum 204
Alcaligenes eutrophus 264, 265
alfalfa 126, 263
alkaloid 257–260, 266
allele diversity 36, 88, 89, 112–113, 177, 179
alleles 49, 52
Allium 138 *see also* garlic, onion
allopolyploid 78, 89, 92 *see also* polyploid
amplified fragment length polymorphism (AFLP) 26, 27, 30, 104, 80, 88, 89, 180, 220
ancient DNA 9, 23
antibiotics 205, 263
antibody 15 *see also* monoclonal antibodies
antifreeze proteins 205
antifungal 259
antigen 15
Antirrhinum 217
antisense RNA *see* RNA
antiserum 15 *see* polyclonal antisera
apple *see* *Malus domestica*
Arabidopsis 14, 77, 83, 84, 182, 204, 218, 222–223, 265, 288, 290, 292–293
Armoracia 139
Artemisia annua 259
asparagus 138, 140, 141, 209
Asteraceae 259
Atropa belladonna 266
Avena 57, 208, 212, 285
avocado 122, 241

B-chromosomes 53
bacterial artificial chromosome (BAC) 86, 206, 221
Bacillus thuringiensis 264
banana 5, 120, 123, 125, 136, 140, 142, 144, 204, 208, 272 *see also* Musa
base collections 105
barley *see* *Hordeum*
bean 140, 242, 243, 263 *see also* mungbean, *Phaseolus*, *Vicia faba*
beet *see* sugarbeet
Bermuda grass 57
bioassays 260
Biodiversity Convention 268, 270, 272
biodiversity 58–59, 61, 67, 72–73, 88, 255, 256, 261, 263, 266–269, 271, 272, 283
biolistic *see* microprojectile
bioprospecting 6, 255–256, 261–269, 271

- Brassica* 28, 35, 77, 82–84, 90, 93, 186, 204, 222, 240
Brassica napus 35, 140, 141, 143, 204, 208, 215, 216, 265
Brassica oleracea, 182, 185, 191, 193, 204
Brazil nut 214
bread wheat 182
breeding system 51, 52, 54, 56, 57, 66, 73
bulk segregant analysis 188, 190, 220
- cabbage *see Brassica oleracea*
cacao *see Theobroma cacao*
Calophyllum lanigerum 258
canola *see Brassica napus*
CAP *see* cleaved amplified polymorphism
Capsella bursa-pastoris 140
carnation 128, 136, 139
carrot 126, 127
cassava *see Manihot*
Castanea 125, 140
Catharanthus 126, 130, 144
centipedegrass 26
centres of diversity 63, 235
centromeres 12, 14, 28
Cf-9 gene 218, 219 *see also* resistance genes
chemotherapy 237
chestnut *see Castanea*
chiasma 53, 181–184, 186–187
chickpea 207, 209
chloroplast 12, 13, 17, 22, 25, 29, 32
chromosome
 bacterial artificial chromosome (BAC) 86, 206, 221
 B-chromosomes 53
 chromosome banding 16, 20, 31
 chromosome landing 81
 chromosome painting 28
 chromosome pairing 78, 92, 181–184, 197, 219–222
 chromosome walking 81, 86
 yeast artificial chromosome (YAC) 82, 86, 206, 220–222
 see also centromeres and telomeres
Chrysanthemum 126, 143, 257, 266
Citrus 54, 122, 138
 sweet orange 143
 navel orange 144
 orange 208, 209
Cladosporium fulvum 218
classification 104, 260, 261
cleaved amplified polymorphism (CAP) 24, 180
clover *see Trifolium repens*
co-linearity *see* comparative mapping
co-suppression 215
- coconut *see Cocos nucifera*
Cocos nucifera 119, 122, 140–141, 208, 265, 272
Coffea 112, 120, 122, 124, 126, 136, 139, 140, 141
coffee *see Coffea*
colchicine 190
collections
 active / working collections 105, 111
 base collections 105
 collecting procedures 59, 169, 258
 core collections 11, 105, 111–114
 duplicate safety collections 105 *see also* safety storage
ex situ collections,
 conservation and genebanks 3–5, 11, 33, 49, 59, 68–70, 103, 115, 146, 273
 see also seeds
in situ conservation 3–4, 11, 33, 49, 50, 71, 146, 235
in vitro collections 103, 121, 122
 plant collections 169
 see also genebanks
Colocasia esculenta 124
comparative mapping *see* mapping
conservation 267, 271
Cordyline 125
core collections 11, 105, 111–114
coriander 216
cotton 122, 204, 208, 209
cottonseed 215
cowpea *see Vigna*
crossability 6, 106, 107, 108
cryopreservation 121, 127–143, 145, 165, 168, 172
 see also preservation in liquid nitrogen
Cucurbita 290
Cynodon 57
- DAF *see* DNA amplification fingerprint
database 25, 36, 292, 295 *see also* ACEDB
date palm 140, 141
dehydration 127–128, 130–137, 138, 139–140, 141
denaturing gradient gel electrophoresis (DGGE) 27
desiccation *see* dehydration
Desmodus rotundus 262
digoxigenin (DIG) 241, 242
Dioscorea 120
 see also yam
disease resistance genes 218–219, 221
 see also resistance genes
DNA
 DNA amplification fingerprint (DAF) 24, 25, 57

- see also random amplified
 polymorphic DNA
 DNA banks 3, 5, 163–72
 DNA content 19–20, 78, 181
 DNA–DNA hybridization 32
 DNA fingerprinting 80, 88, 281
 DNA libraries 21, 25, 206, 214, 221
 DNA markers 56, 81
 DNA reassociation 19, 20
 DNA repeated sequences 12–13, 19, 22,
 24, 27, 78–79, 81, 222
 DNA satellites 12, 14, 19, 20, 22, 27, 28
 DNA sequencing 9, 17, 28–30, 31, 32, 33,
 34, 110, 292, 295
 sequencing by hybridization (SBH)
 36, 37
 T-DNA 207–209
 see also Tm
 dot and slot blot hybridization 17, 27, 28,
 30–31, 241, 246
 double antibody sandwich (DAS) ELISA see
 ELISA
 double-stranded RNA (dsRNA) see RNA
 doubled haploid 178, 184, 190–191, 193
 drift see genetic drift
Drosophila 87, 187, 217
 duplicate safety collections 105
 see also safety storage
 duplicate accessions 11, 105, 107, 108, 110,
 111, 112

Echinacea 258
 ecogeographical 78
 ecological distribution 63
 effective population size 51
Elaeis guineensis 120, 124, 126, 129, 140, 141,
 143, 145, 215
 electronic mail 282–283
 electrophoresis 16–18, 20–21, 25, 31, 212,
 213, 240
 electroporation 210, 212
 ELISA 247, 248, 260
 double antibody sandwich (DAS)
 246–247
 enzyme-linked immunosorbent assay
 18, 238, 243–246
 triple antibody sandwich (TAS) 246
 EMBL 25, 290
 embryos
 embryo culture 90
 embryo rescue 90
 somatic embryo 126, 129, 130, 137, 139,
 140–141, 143, 144–145
 zygotic embryos 131–135, 139, 140
 encapsulation 130, 136, 137, 138–139

 environmental diversity 62
 enzyme-linked immunosorbent assay see
 ELISA
Epipedobates tricolor 260
 epistasis 90
Escherichia coli 217, 262, 272
 EST see expressed sequence tag
 ethnobotany 258–259
 eucalyptus 136
 evolution 50, 63–64, 77, 78, 80, 87
 ex situ collections,
 conservation and genebanks 3–5, 11, 33,
 49, 59, 68–70, 103, 115, 146, 273
 see also seeds
 expressed sequence tag (EST) 36, 86
 extinction 6

 fatty acids 215–216 see also oleic acid and
 petroselinic acid
Festuca pratensis 53
 field genebanks 120
 File Transfer Protocol (FTP) 283–284
 fingerprinting see DNA fingerprinting
 flow cytometry 16, 19–20, 31–32
 fluorescent in situ hybridization (FISH) see
 hybridization
 fluorochrome 19
Fragaria 124, 143
 Frequently Asked Questions 283
 FTP see File Transfer Protocol
 fungi
Cladosporium fulvum 218
Helminthosporium maydis 163
Phytophthora infestans 163
Rhynchosporium oryzae 112
 yeast 87

 garlic 258
 see also *Allium*
 GenBank 25, 290
 gene
 flow 55, 64, 69, 71
 gene-for-gene interaction 218
 see also resistance genes
 herbicide resistance genes 205, 214
 major genes 187, 193
 polygenes 175, 189
 pool 5, 6, 66, 89, 91, 106, 107, 213
 prospecting 6, 261
 see also bioprospecting
 rbcL gene 29
 transfer 164, 203–223, 263
 virulence genes 207–210
 see also marker-assisted gene transfer

- genebank 121, 146
 genetic
 genetic diversity 49, 50, 51, 59–62, 65–66, 72–73, 163
 genetic drift 11, 64, 68, 70
 genetic erosion 59–60
 genetic mapping 77, 80–85, 281, 292, 295
 see also linkage groups
 genetic stability 142
 genetic variation 51, 72
 genome analysis 77
 see also genetic mapping
 genome 78
 see also DNA
 geographical distribution 63
 germplasm
 catalogue 292
 exchange 123
 utilization 87
 ginger 126
 ginkgo 258
 ginseng *see* *Panax ginseng*
 gladiolus 242
 glycosides 257
 Gopher 284–287
 Gramineae 54
 grape 122, 126, 129, 136, 137, 138, 208, 209, 240, 241
 Green Revolution 203
 GRIN 288, 295
 groundnut 242

 Hardy-Weinberg 51, 61
 heat shock protein 30
Helminthosporium maydis 163
 herbicides 260
 herbicide resistance and resistance genes 205, 214
 heterochromatin 12, 14, 20
 heterosis 178
 heterozygosity 26, 54, 180
Hevea 140
Hirudo medicinalis 262
Hordeum 57
 barley 54, 80, 82, 84, 86, 89, 91, 112, 208, 213, 247
 Hordeum bulbosum 90, 91
 Hordeum chilense 92
 Hordeum spontaneum 89
 Hordeum vulgare 77
 horse chestnut 258
Howea 140
 human 25, 35
 Human Genome Project (HGP) 24, 33, 34, 36, 37

 human genotyping kits 34
 hybridization
 DNA–DNA hybridization 13, 16, 19, 21, 27, 32, 241
 dot and slot blot hybridization 17, 27, 28, 30, 31, 241, 246
 fluorescent *in situ* hybridization (FISH) 28, 80
 in situ hybridization (ISH) 13, 17, 27, 28, 31, 32, 93
 nucleic acid spot hybridization (NASH) 241
 see also DNA sequencing by hybridization and DNA reassociation kinetics
 sexual hybridization 80, 89–92
Hymenaea courbaril 259
Hyoscyamus niger 266

 IGS *see* ribosomal RNA
 immunocapture PCR *see* polymerase chain reaction
 immunocapture RT-PCR *see* polymerase chain reaction
 immunoelectrophoretic assays 18
 immunosorbent electron microscopy-ISEM 239

in situ
 conservation 3–4, 11, 33, 49, 50, 71, 146, 235
 hybridization (ISH) *see* hybridization
 population 11
in vitro
 assays 6
 collections 103, 121, 122
 conservation 119–146
 culture 5, 38, 120, 142, 207, 235 *see also* tissue culture
 fertilization 91
 incompatibility 53
 informatics 7, 281–296
 information management 269
 insect resistance 214, 215
 insecticidal 257, 264, 266
 intergenic spacers (IGS) *see* ribosomal RNA
 internal transcribed spacers (ITS) *see* ribosomal RNA
 Internet 7, 281–282
 inter-simple sequence repeat 220
 inter-repeat PCR *see* polymerase chain reaction
 introgression 177, 194–196, 203
 introns 213
 Ipomea batatas 120, 123, 138, 208
 isozyme 4, 9, 18, 30–32, 52, 56, 57, 106, 112, 113, 142, 185

- ITS *see* ribosomal RNA internal transcribed spacers
- JOINMAP 185
- jojoba 265
- Juniperus* 170
- karyotypes 20, 27, 32
- Kiwi fruit 124
- landraces 2, 50, 71, 103, 235
- lectin 215
- Lens culinaris* (lentil) 83
- lily 139
- linkage
- linkage disequilibrium 88, 114
 - linkage drag 177, 194, 195
 - linkage groups, 180, 185, 186, 187
see also genetic mapping
- Lolium* 53, 57, 61, 64
- Lycopersicon* 52, 54, 77, 81, 82, 182, 203, 204, 205, 208, 218–221, 290
- MAbs *see* monoclonal antibodies
- maintenance 68
- maize *see* *Zea mays*
- major genes *see* genes
- Malus domestica* 124, 136, 138, 208, 209
- Manihot* 112, 120, 123, 124–125, 142, 143, 144, 208
- mapping
- comparative mapping 4, 22, 81–85, 87, 93, 222
 - JOINMAP 185
 - map-based cloning 81, 84, 197, 216, 219–221
 - map length 181, 182, 183, 184, 186, 187, 197
 - MAPMAKER 185
 - mapping functions 183, 185
 - physical mapping 28, 31, 93, 220–222, 281, 292, 295
 - see also* genetic mapping and linkage groups
- marker-assisted
- marker-assisted selection 175–176, 196, 213
 - marker-assisted gene transfer 194
 - marker-assisted prediction of performance 114–115
- melon 140, 141
- microcomplement fixation (MCF), 15, 16, 31, 32
- microprojectile 207, 208, 211–212, 213
- microsatellites *see* simple sequence repeats
- microspore or ovule culture, 178
- millet 82
- minisatellites *see* variable number of tandem repeats
- mint 138
- mitochondrial 13, 17, 25
- molecular diversity 258
- molecular markers 81, 83, 88
- see* amplified fragment length polymorphism
 - see* cleaved amplified polymorphism
 - see* DNA amplification fingerprint
 - see* expressed sequence tag
 - see* inter-simple sequence repeats
 - see* random amplified polymorphic DNA
 - see* restriction fragment length polymorphism
 - see* sequence tagged sites
 - see* simple sequence repeat
 - see* simple sequence length polymorphisms
 - see* single-stranded conformation polymorphism
 - see* variable number of tandem repeats
- monoclonal antibodies 16, 18, 30, 31, 244–245, 247
- see also* antibodies
- morphine 260
- mulberry 127, 136, 138
- multiplex 35
- mungbean 83
- Musa* 120, 122, 123, 124
- see also* banana and plantain
- Myiodon (ground sloths) 9
- navel orange *see* *Citrus*
- near isogenic lines (NILs) 197, 220
- nematocide 257
- Nephrolepis* 125
- nested PCR *see* polymerase chain reaction
- Netscape Navigator 286
- newsgroups 283
- Nicotiana*
- Nicotiana rustica* 189, 266
 - Nicotiana tabacum* 126, 204, 211, 213, 214, 215, 216, 218, 223, 263
- nucleic acid spot hybridization (NASH) *see* hybridization
- nucleolar organizing regions (NORs) 20
- see also* ribosomal RNA
- oak *see* *Quercus*

- oats *see* *Avena*
Oenothera 66
oil-palm *see* *Elaeis guineensis*
oilseed rape *see* *Brassica napus*
oleic acid 215
oligonucleotide ligation assay (OLA) 35
onion 209, 211
 see also *Allium*
orange *see* *Citrus*
orthodox seeds *see* orthodox seeds
Oryza sativa 4, 77, 82, 84, 86, 103–112,
 114–115, 126, 182, 197, 207, 208, 209,
 210, 211, 215, 222, 223, 263
- palm oil 265, 272
Panax ginseng 126, 144, 258
Panicum maximum 54
passport data 106, 108, 110, 112–113
pathogens 50, 263
PCR *see* polymerase chain reaction
pea *see* *Pisum sativum*
peach 126
peanut 215
pear 136–137, 138
pearl millet 91
pesticide 256–257, 262
petroselinic acid 216
phage-display 248
pharmaceutical 2, 164, 256, 263
Phaseolus 112, 209, 246
 see also bean
phenotypic variation 4, 188, 192, 193
phylogeny 63, 77, 165
physical mapping *see* mapping
phytochrome 30
Phytophthora infestans 163
Picea abies 143
pilocarpine 259
Pilocarpus pignatifolios 259
Pisum sativum 83, 140, 141, 236, 246, 263
plant breeding 2, 5, 175
plant collections 169
Plantago 57
plantain 120, 204
 see also *Musa*
pleiotropic effects 179, 214
plum 243
 see also *Prunus*
polyclonal antisera 244, 247 *see* antiserum
polygalacturonase 205
polygenes 175, 189
polygenic traits 176
polymerase chain reaction (PCR) 9, 13, 16,
 22, 23–27, 29, 31, 32, 80, 88, 104, 238
 anchored PCR 24, 26
 arbitrarily primed PCR (AP-PCR) 25
 immunocapture PCR 243
 immunocapture RT-PCR 247
 inter-repeat PCR 24, 26
 nested PCR 24, 26
 RT-PCR 242–244
 see also Random amplified
 polymorphic DNA, DNA
 amplification fingerprint and
 amplified fragment length
 polymorphism
 polyploidy 53, 54, 57, 178, 204
 see also allopolyploid
 population structure 49
 potato *see* *Solanum tuberosum*
 preservation in liquid nitrogen *see*
 cryopreservation
 protein engineering 261
 protoplasts 135, 210, 212
Prunus 122, 124
 see also plum and wild cherry
Pseudomonas 221, 272
Pto locus 221
 see also resistance genes
pyrethroids 257
- quality control 178
quantitative trait loci (QTLs) 4, 51–52, 58,
 81, 84, 88, 115, 189–194, 197, 222
quantitative traits 4, 51, 114, 175, 176, 188,
 191–192, 196, 213
Quercus 125, 140
- random amplified polymorphic DNA
 (RAPD) 24, 25, 27, 30, 32, 57, 88, 89,
 104, 107–109, 111, 180, 220, 281
 see also DNA amplification fingerprinting
rape *see* *Brassica napus*
Rauvolfia serpentina 125
rubisco large subunit (rbcL) gene 29
recalcitrant seeds *see* seeds
recombinant inbred lines 184, 190
recombination 53, 66, 181, 186, 222
 recombination frequency 183–187,
 190–191, 194
regeneration in culture 206–207, 209–210
regulatory sequences 213–214, 223
repeated DNA sequences *see* DNA
resistance 1, 106, 112
 resistance genes 204
 see also Cf9, Pto, disease resistance
 genes, herbicide resistance genes
 and insect resistance genes
restriction 21–22, 24, 26, 30, 186

- endonucleases 206
 fragment analysis 9
 fragment length polymorphism (RFLP)
 16, 21–22, 28, 30–32, 57, 80, 81, 84, 89,
 104, 107, 143, 144, 179, 180, 185, 220,
 281
Rhynchosporium oryzae 112
 ribosomal RNA *see* RNA
 rice *see* *Oryza*
 RNA
 antisense RNA 215, 264
 double-stranded RNA (dsRNA) 240
 ribosomal RNA 13–15, 30
 intergenic spacers (IGS) 13, 14, 15, 30
 internal transcribed spacers (ITS) 30
 nucleolar organizing regions (NORs)
 20
Rosaceae 54
 rRNA *see* RNA
Rubiaceae 259
 rye *see* *Secale cereale*

Saccharum 82, 120, 142, 143–144, 204, 208
 safety storage 105
 see also duplicate safety collections
 sampling procedures 61, 66, 67
 satellite DNA *see* DNA
Secale cereale 77, 79, 86, 90, 93, 138, 181, 204,
 212
 secondary metabolites 259
 see also alkaloid, terpenoid, pyrethroid
 and pilocarpine
 seeds
 ex situ seed collections 103
 see also ex situ conservation
 seed multiplication 69, 70
 orthodox seeds 119, 120, 121
 recalcitrant seeds 41, 119, 121
 synthetic seed 126
 see also encapsulation
 selection 53, 55, 68
 sequence tagged sites (STS) 24, 31, 33, 35,
 89
 sequencing *see* DNA sequencing
Simmondsia 265
 simple sequence length polymorphisms
 (SSLPs) 14
 simple sequence repeat (SSRs) 13, 14, 25,
 30, 31, 34, 35, 80, 104, 107, 180, 206,
 220
 single-stranded conformation
 polymorphism (SSCP) 27
 slot blots 28
 see also dot blot
 slow-growth storage 124

 snowdrop 215
Solanum tuberosum 5, 55, 77, 82, 108, 110,
 120, 123, 129, 138, 141, 143, 144–145,
 163, 208, 215, 220, 238, 242, 243, 245,
 246–247, 264
 somaclonal variation 121, 210, 212
 somatic embryo *see* embryos
 sorghum 35, 77, 82, 91, 112, 208, 209, 210,
 222
 soybean 82, 207, 208, 209, 215, 216
Spartina anglica 55
 species diversity 58
 SSLP *see* simple sequence length
 polymorphism
 SSCP *see* single-stranded conformation
 polymorphism
 SSRs *see* simple sequence repeat
 STS *see* sequence tagged sites
 steroid 263
 stratified sampling 65–66, 112, 113
 strawberry *see* *Fragaria*
Streptomyces avennitis 257
 sugar-beet 137, 208, 209
 sugar-cane *see* *Saccharum*
 sunflower 215
 super binary plasmid 209
 sustainable 272
 sweet orange *see* Citrus
 sweet potato *see* Ipomea batatas
 synteny *see* comparative mapping
 synthetic seed *see* seeds

 T-DNA tagging 216, 222–223
 T-DNA 207–209
 taro 124
Taxodium (bald cypress) 9
 taxol 260
 taxonomic identification 4, 106, 107, 109
Taxus 259–260
 tea 138
 technology transfer 270
 Telnet 284
 telomeres 14
 teosinte 88, 89
 see also *Zea mays*
 terpenoids 257
Thaumatococcus 272
Theobroma cacao 64, 119, 122
 thermotherapy 237
 tissue culture 5, 38, 120, 142, 235
 cryopreservation 121, 127–143, 145, 165,
 168, 172
 embryo culture 90
 embryo rescue 90
 in vitro assays 6

- tissue culture *continued*
in vitro collections 103, 121, 122
in vitro fertilization 91
microspore or ovule culture, 178
protoplasts 135, 210, 212
regeneration in culture 206–207, 209, 210, 211
slow-growth storage 124
somaclonal variation 121, 210, 212
somatic embryo 126, 129, 130, 137, 139, 140, 141, 143, 144, 145
vitricification 128, 130, 131–135, 137, 138, 139, 145
- Tm 19
- tobacco *see Nicotiana*
- tomato *see Lycopersicon*
- transcription factors 87
- transformation 165, 204–213, 222, 223
- transgenic plants *see* transformation, DNA transfer and *Agrobacterium*
- transposon tagging 86, 216, 217, 219
see also Ac/Ds
- Trifolium repens* 60, 61, 65, 240
- triple antibody sandwich (TAS-ELISA) *see* ELISA
- Triticeae 285
- Triticum aestivum* 54, 77, 79, 80, 82, 86, 89, 90, 91, 92, 93, 112, 140, 203, 204, 208, 210, 221, 222, 285
- Tritordeum* 92
- trypsin inhibitor 214
- turnip *see Brassica oleracea*
- Umbellularia californica* 265
- universal (uniform) resource locator (URL) 287
- utilization 91
- variable number of tandem repeats (VNTR) 12, 13, 14, 22, 25, 30, 32, 104
- Valeriana wallichii* 126
- Veitchia* 140
- Vicia faba* 240
see also bean
- Vigna* 83, 164, 214
- viroids *see* viruses
- virulence genes 207–210
- virus 236–249
bioassays 239
resistance 205
viroids 236, 237, 241, 242
virus-free plants 120
- vitricification 128, 130, 131–135, 137–139, 145
- VNTR *see* variable number of tandem repeats
- wasabi 138, 139
- Western blots 248
- wheat *see Triticum*
- white spruce 125
- wild cherry 125
see also Prunus
- World Wide Web (WWW) 285–296
- YAC *see* yeast artificial chromosome
- yam 120, 123, 209
see also Dioscorea
- yeast 87
- yeast artificial chromosome (YAC) 82, 86, 206, 220–222
- Zea mays* 28, 35, 77, 79, 82, 84, 86, 88, 89, 91, 140, 163, 182, 187, 197, 208, 209, 210, 213, 217, 221, 222, 288, 293
see also teosinte
- zygotie embryos *see* embryos