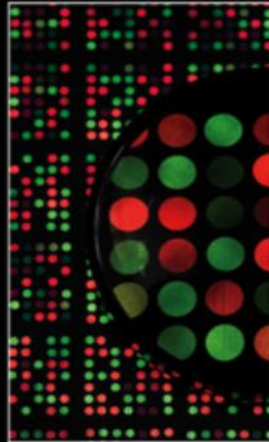


Molecular Markers in Plants



Edited by Robert J. Henry



 WILEY-BLACKWELL

Contents

[Cover](#)

[Title Page](#)

[Copyright](#)

[Contributors](#)

[Preface](#)

[Chapter 1: Evolution of DNA Marker
Technology in Plants](#)

[Introduction](#)

[Early Marker Technologies](#)

[Evolving Range of Applications of DNA Markers in
Plants](#)

[Applications](#)

[Future Developments](#)

[References](#)

[Chapter 2: Whole-Genome Sequencing
for Marker Discovery](#)

[Sequencing Strategies](#)

[Sequencing Technologies](#)

[Epigenetic Markers](#)

[Genome-Wide Selection](#)
[Data Analysis Resources](#)
[References](#)

[Chapter 3: Amplicon Sequencing for Marker Discovery](#)

[Introduction](#)
[Background](#)
[Maximizing Efficiency Through Sample Pooling](#)
[Limitations of Amplicon-Based MPS](#)
[Bioinformatics](#)
[Concluding Remarks](#)
[Acknowledgments](#)
[References](#)

[Chapter 4: Transcriptome Sequencing for Marker Discovery](#)

[Introduction](#)
[Basic Approach](#)
[Conclusions](#)
[References](#)

[Chapter 5: Molecular Markers in Plant Improvement](#)

[Introduction](#)
[Plant Domestication and Traditional Breeding](#)
[Application of Molecular Markers to Breeding](#)
[Next-Generation Approaches to QTL Discovery](#)
[Conclusion](#)

References

Chapter 6: Applications of Molecular Markers in Plant Conservation

Introduction

Traditional Approaches

The Way Forward

Conclusion

References

Chapter 7: Molecular Markers for Plant Biosecurity

Introduction

The Present—PCR for Specific Diagnosis and for DNA Barcoding

The Future—Next-Generation Sequencing

Methods to Revolutionize Plant Quarantine

Diagnostics

Conclusions

Acknowledgments

References

Chapter 8: Molecular Markers for Harnessing Heterosis

Introduction

Molecular Markers for Understanding the Genetic Basis of Heterosis

Molecular Diversity and Heterosis—Molecular Markers for Predicting Heterosis

[Conclusion](#)
[References](#)

[Chapter 9: Genetic Variant Discovery and Its Use in Genome Characterization of Agronomically Important Crop Species](#)

[Introduction](#)
[Sanger Resequencing](#)
[Single Feature Polymorphisms](#)
[Next-Generation Sequencing](#)
[High-Density Genotyping using the Illumina Golden Gate Platform](#)
[Genotyping by Sequencing](#)
[Genome Characterization and Haplotypes](#)
[Conclusions and Perspectives](#)
[References](#)

[Chapter 10: Future Prospects of Molecular Markers in Plants](#)

[Introduction](#)
[Molecular Markers: The Past](#)
[Molecular Markers: The Present](#)
[Molecular Markers: The Future](#)
[Novel Approaches or Platforms for Plant Breeding](#)
[Conclusions](#)
[Acknowledgments](#)
[References](#)

[Index](#)

Advertisements

Molecular Markers in Plants

Editor

Robert J. Henry

*Queensland Alliance for Agriculture and Food Innovation
University of Queensland
Brisbane, Queensland, Australia*

 **WILEY-BLACKWELL**

A John Wiley & Sons, Inc., Publication

This edition first published 2013 © 2013 by John Wiley & Sons, Inc.

Wiley-Blackwell is an imprint of John Wiley & Sons, formed by the merger of Wiley's global Scientific, Technical and Medical business with Blackwell Publishing.

Editorial offices:

2121 State Avenue, Ames, Iowa 50014-8300, USA
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

9600 Garsington Road, Oxford, OX4 2DQ, UK

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell.

Authorization to photocopy items for internal or personal use, or the internal or personal use of specific clients, is granted by Blackwell Publishing, provided that the base fee is paid directly to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For those organizations that have been granted a photocopy license by CCC, a separate system of payments has been arranged. The fee codes for users of the Transactional Reporting Service are ISBN-13: 978-0-4709-5951-0/2013.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book. This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the publisher is not engaged in rendering professional services. If professional advice or

other expert assistance is required, the services of a competent professional should be sought.

Library of Congress Cataloging-in-Publication Data

Molecular markers in plants / editor, Robert J. Henry.
p. cm.

Includes bibliographical references and index.

ISBN 978-0-470-95951-0 (hardback : alk. paper) 1. Plant breeding.

2. Genetic markers. 3. Plant genetics. 4. Crop improvement.

I. Henry, Robert J.

SB123.M58 2013

634.9'56--dc23

2012019771

A catalogue record for this book is available from the British Library.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Contributors

Matthew A. Campbell

DuPont Pioneer
Johnston, Iowa, United States

Stéphane Deschamps

DuPont Agricultural Biotechnology
Experimental Station - P.O. Box 80353
200 Powder Mill Road
Wilmington, DE 19880-0353

Mark Edwards

Southern Cross Plant Sciences
Southern Cross University
Lismore, Australia

Celine H. Frere

School of Agricultural and Food Sciences
University of Queensland
Brisbane, Australia

Andrew D.W. Geering

Cooperative Research Centre for National Plant Biosecurity
and the Queensland Alliance for Agriculture and Food
Innovation
The University of Queensland
Ecosciences Precinct
Brisbane, Queensland, Australia

Edward K. Gilding

School of Agricultural and Food Sciences
University of Queensland

Brisbane, Australia

Susan Gillies

Southern Cross Plant Sciences
Southern Cross University
Lismore, Australia

Ian D. Godwin

School of Agricultural and Food Sciences
University of Queensland
Brisbane, Australia

Robert J. Henry

Queensland Alliance for Agriculture and Food Innovation
University of Queensland
Brisbane, Queensland, Australia

Pavana J. Hiremath

International Crops Research Institute for the Semi-Arid
Tropics
Hyderabad, India

Gopala S. Krishnan

Division of Genetics
Indian Agricultural Research Institute
New Delhi, India

Ana Pavasovic

School of Biomedical Science
Queensland University of Technology
Brisbane, Australia

Peter J. Prentis

School of Earth, Environmental and Biological Sciences
Queensland University of Technology

Brisbane, Australia

Oscar Riera-Lizarazu

International Crops Research Institute for the Semi-Arid
Tropics
Hyderabad, India

Maurizio Rossetto

National Herbarium of NSW
Royal Botanic Gardens and Domain Trust
Sydney, Australia

Reyazul R. Mir

International Crops Research Institute for the Semi-Arid
Tropics
Hyderabad, India

Paul D. Rymer

Hawkesbury Institute for the Environment
University of Western Sydney
Richmond, Australia

Timothy R. Sexton

Department of Forest Sciences
The University of British Columbia
Vancouver, British Columbia, Canada

Frances M. Shapter

Southern Cross Plant Sciences
Southern Cross University
Lismore, Australia

A.K. Singh

Division of Genetics
Indian Agricultural Research Institute

New Delhi, India

Rajeev K. Varshney

International Crops Research Institute for the Semi-Arid
Tropics

Hyderabad, India;

CGIAR-Generation Challenge Programme

Mexico;

School of Plant Biology

Faculty of Natural and Agricultural Sciences

The University of Western Australia

Crawley, Australia

Daniel L.E. Waters

Southern Cross Plant Science

Southern Cross University

Lismore, Australia

Preface

Plants are fundamental to life, being the basis of our food production and an essential part of the global ecosystem on which life on earth depends. Plants have been used as a source of a wide range of materials, but the threat of exhaustion of fossil oil supplies has resulted in a renewed evaluation of plants as a source of energy and biomaterials. Molecular analysis of plants has found many applications in plant improvement, in the management of plant production, and the conservation of plant resources. Molecular markers are routinely used to identify plants for forensic or intellectual property applications. Molecular tools have become key contributors to the management of wild plant populations helping to conserve biodiversity. The relentless need for the continuous development of genetically improved crops to satisfy the demands of a global human population growing in number and affluence is now strongly supported by molecular marker technology.

Recent dramatic advances in DNA sequencing are now providing cost-effective options for the discovery of very large numbers of markers for any plant species. These developments significantly change the approach to marker discovery and analysis in plants and greatly expand the potential range of applications. This book outlines the technologies for molecular analysis of plants in support of plant breeding, production, and conservation. The techniques that have been used in the past are reviewed in relation to recent developments and future potential. This book updates earlier volumes on this topic featuring significant advances in both the technology and application of markers.

Robert Henry
University of Queensland

1

Evolution of DNA Marker Technology in Plants

Robert J. Henry

Contents

[Introduction](#)

[Early Marker Technologies](#)

[Evolving Range of Applications of DNA Markers in
Plants](#)

[Applications](#)

[Future Developments](#)

[References](#)

Introduction

Genetic markers are key tools for plant identification and plant improvement (Henry, 2001). Genetic marker technology has evolved rapidly with early methods based on phenotyping or isozymes being replaced by DNA-based methods of increasing sophistication. Early markers were few in number and difficult to assay. The ultimate development of the technology will provide simple methods to assess all genetic variation in the genome. This chapter provides a brief account of the development of genetic marker technology and its application to plants over the last

30 years. This perspective is provided as background and context for the accounts of the latest technologies (Henry *et al.*, 2012) to follow in subsequent chapters.

Molecular marker technology has evolved through several phases. Early methods employing non-DNA-based methods were replaced by DNA-based methods as the technologies for DNA analysis improved. Early hybridization-based methods were displaced rapidly following the development of polymerase chain reaction (PCR). PCR-based methods greatly increased the feasibility of high-throughput marker screening. Early PCR-based methods relied upon arbitrary primers because of a lack of sequence information for many species. These in turn were overtaken by the widespread adoption of more robust microsatellite or simple sequence repeat (SSR) markers. Single nucleotide polymorphisms (SNPs) have more recently replaced SSR markers as larger volumes of sequence data became available (Henry, 2008). Second-generation sequencing technologies have greatly accelerated the move to sequence-based markers. Ongoing improvements in DNA sequencing promise a continued convergence of sequencing and genotyping technologies. Third-generation sequencing promises delivery of technology for routine sequencing of even complex plant genomes enabling ready marker discovery and analysis.

Molecular markers have a wide range of applications in biological systems including plants. Molecular markers are very useful in identification of plants and in determining the relationships between plants. Plant identification may be important in plant breeding, plant production and processing, policing of intellectual property rights, and forensic applications. Determination of genetic relationships is required in evolutionary and conservation genetic analyses and in selection of germplasm in plant breeding. Plant breeding is often directly supported by marker-assisted selection.

This chapter will outline the evolution of molecular marker techniques and their applications to plants.

Early Marker Technologies

Biochemical markers (e.g., isozymes) have been widely applied to the analysis of the genetics of plant populations. This approach has provided cost-effective options for laboratories with little equipment studying poorly known biological systems. The most common of these techniques was the assay of isozymes.

Improving DNA analysis methods with greater discrimination, simplicity of analysis, and suitability for automation have largely replaced these non-DNA-based methods. However, early biochemical approaches have still been in limited use despite recent dramatic advances in DNA analysis technologies. Isozyme analysis has persisted for species for which little or no DNA sequence data was available and in laboratories not equipped with facilities for DNA analysis or where labor costs are low compared with reagent costs. Recent advances in DNA analysis tools should provide cost-effective and preferable DNA-based methods for almost all applications.

DNA-Based Methods

The development of DNA analysis methods provided an opportunity to directly analyze difference in the genome of the organism rather than rely on inference from analysis of expressed genes (as in isozyme analysis). Early DNA-based methods used hybridization of DNA to detect variation in the DNA samples. The development of restriction fragment length polymorphism (RFLP) analysis provided an approach that was widely adopted in the 1980s and became the standard approach until replaced by PCR-based methods in

the 1990s. The evolving DNA-based methods are listed in [Table 1.1](#). DNA extraction from plants is generally much more difficult than from typical animal sources because of the rigid cell wall and the high levels of secondary metabolites in plant cells. This requires tissue disruption techniques to break the cell wall that are likely to shear the DNA if too vigorous. Improved DNA extraction methods made a significant contribution to the advances in application of DNA analysis methods in plants (Thompson and Henry, 1993; Graham, Mayer, and Henry, 1994).

Table 1.1 Evolving genotyping methods.

Method	Acronym
Restriction fragment length polymorphism	RFLP
Random amplified polymorphic DNA	RAPD
Amplified fragment length polymorphism	AFLP
Diversity arrays	DArT
Sequence characterized amplified region	SCAR
Simple sequence repeat	SSR
Single nucleotide polymorphism	SNP

Restriction Fragment Length Polymorphism

RFLP analysis used one or more endonucleases (restriction enzymes) to digest genomic DNA from the sample. Differences in DNA sequence in the sample at the restriction site or between adjacent restriction sites resulted in differences in the length of the fragment in the digest. The digested DNA was separated by electrophoresis. The DNA was then transferred to a membrane for analysis of fragments. Specific probes derived from genomic DNA or cDNA were labeled and used to detect fragments in the digest containing related sequences. Labeling was with radioisotopes or later specific proteins that could be detected using antibodies. Membranes could be hybridized

with a series of different probes to explore polymorphisms in different parts of the genome. The limitations of these early methods included the need for large amounts of DNA especially for species with large genome sizes.

PCR-Based Methods

The development of the PCR in the late 1980s provided a new tool that rapidly changed approaches to DNA analysis including DNA-based markers. PCR based techniques replaced the earlier non-PCR methods because of their greater sensitivity, discrimination power and ease of automation. The use of PCR has greatly accelerated the development and application of DNA markers in plants. PCR-based marker analysis has facilitated much greater automation of analysis improving throughput and reliability of analysis. PCR also allowed the use of much smaller quantities of DNA for plant DNA analysis. This greater sensitivity together with the high specificity of PCR resulted in rapid replacement of earlier methods. Simpler DNA extraction methods could be developed for use with PCR because of the need for less DNA that did not need to be of high molecular weight. The small amount of sequence data available at the time limited the application of PCR to amplification of specific or known genetic sequences and required the development of techniques based upon arbitrarily primed PCR to generate markers.

Arbitrary Methods

A lack of DNA sequence data for most species was overcome by use of PCR primers of arbitrary sequence. These primers were used to generate a fingerprint for the genotypes. Random amplified DNA polymorphism (RAPD) was followed by amplified fragment length polymorphism (AFLP) and DNA amplification fingerprinting (DAF). The

diversity array technique method also generated large numbers of markers without sequence data.

Random Amplified DNA Polymorphism

The RAPD method used short (10 mer) primers to amplify fragments for analysis by gel electrophoresis (Williams *et al.*, 1990). This method required careful standardization of conditions for the low stringency amplifications involved and as a result was difficult to repeat in different laboratories.

Other variations on this approach included the DAF method.

Amplified Fragment Length Polymorphism

The AFLP method used restriction enzymes (Zabeau and Vos, 1993) to produce fragments. Oligonucleotides were ligated to the ends to produce priming sites for amplification of a subset of fragments. This produced complex mixtures of fragments. AFLP was widely used for DNA fingerprinting. The method persisted in use for longer than the RAPD method because of the greater reproducibility of the technique.

Diversity Array Technique

The most recent of these methods, not based on specific sequence targets, is the diversity array technique (Jaccoud *et al.*, 2001). In this method, a genomic representation of a species is arrayed for analysis by hybridization. This technique provided very large numbers of markers and has been widely applied especially in construction of genetic maps (Xia *et al.*, 2005). This method generates very large numbers of useful markers but requires significant development effort for each species.

Specific Sequence-Based PCR

As increasing amounts of sequence data became available from Sanger sequencing, PCR amplification has been widely applied to the analysis of known sequence polymorphisms (Garland *et al.*, 2000; McIntosh, Pacey-Miller, and Henry, 2005). Differences in the length of the amplified fragment or differences in the sequence of the amplified fragment can be detected by a range of methods: restriction digestion, melting temperature analysis (Shepherd *et al.*, 1998), and hybridization with a labeled probe. Differences in the primer site can be used to develop assays that only allow amplification from specific target sequences.

Sequence Characterized Amplified Region

Sequence characterized amplified region markers were often derived from RAPD, AFLP, or other markers. Sequencing of the amplified fragment was used to design a PCR with highly specific primers. This allowed conventional robust PCR to be used to assay the marker rather than continuing to rely on the more complex or less reproducible arbitrary primed PCR used in the early methods. These early sequence-based markers indicated the way that markers would develop as sequence data became more readily available.

Reverse Transcription-Polymerase Chain Reaction

Further development of PCR marker methods included the monitoring of the progress of PCR by direct detection approaches as an alternative to electrophoretic separation of the products for analysis. A range of methods employing these real-time PCR methods have been adapted for the detection of DNA polymorphisms as markers in plants.

These often employ specific probes to detect the PCR product (Kennedy *et al.*, 2006b; Bundock *et al.*, 2006).

Simple Sequence Repeat

Short repeated sequences are found widely in the genome and show great variation in most populations making them very useful markers. PCR also accelerated the analysis of these SSRs or “microsatellites.” Most SSR analysis targeted dinucleotide or trinucleotide repeats.

SSR markers have been widely applied in plants (Cordeiro, Taylor, and Henry, 2000). Analysis required that the sequences of the DNA on either side of the repeat sequence be determined to allow design of specific primers for PCR amplification (Ablett, Hill, and Henry, 2006). Analysis of variations in the length of the amplified fragment was originally achieved by gel electrophoresis but later was widely conducted using capillary electrophoresis systems including those developed for DNA fragment length separation in DNA sequencing by the Sanger method (Rossetto, McNally, and Henry, 2002). Capillary electrophoresis allowed automated data capture and facilitated analysis of much larger numbers of samples. Analysis of many loci for each sample was achieved efficiently by combining loci with very different lengths and with primers labeled with different colored fluorophores in the same capillary. Instruments with up to 96 capillaries were used to achieve high throughput.

Innovative approaches were required to discover SSR loci and obtain the flanking sequence data needed for their analysis. Early markers were based on sequencing libraries of genomic clones isolated by hybridization with repeat sequences. As more DNA sequence data have become available, these have been identified efficiently in gene sequences (e.g., EST data; Ablett, Hill, and Henry, 2006) and more recently whole-genome sequence data.

Single Nucleotide Polymorphism

The ultimate minimum unit of genetic variation is the single nucleotide. Analysis of SNP markers has increased as more sequence data has become available (Henry and Edwards, 2009). Often insertions and deletions (indels) are assayed as SNP.

Discovery

The discovery of SNP has been achieved in a variety of ways but is now probably best achieved by DNA sequencing using the latest technologies (Arai-Kichise, 2011; Barbazuk *et al.*, 2007; Deschamps and Campbell, 2010; Imelfort *et al.*, 2009; Novaes *et al.*, 2008; Trebbi *et al.*, 2011; Trick *et al.*, 2009; Yamamoto *et al.*, 2010). Early techniques relied upon a range of mutation detection methods (Cross *et al.*, 2008).

Analysis

Protocols for the analysis of SNP have evolved (Pacey-Miller and Henry, 2003) and have until recently been limited by the availability of known SNP in most organisms. This situation has changed with recent advances in DNA sequencing and we may now need to developed new SNP analysis tools to adequately explore variation in the very large numbers of SNP that can be easily discovered by sequencing.

Platforms for automated SNP analysis allow large numbers of samples and SNP to be assayed (Pattemore *et al.*, 2010; Sexton *et al.*, 2010; Kharabian-Masouleh *et al.*, 2009). Some are most effective for very large numbers of samples and others for very large numbers of SNPs. The mass array platform (Sequenom) automates the processing of samples in 384 well format with analysis of up to 40 SNP per sample by mass spectrometry of primer extension reactions designed to distinguish the SNP. Illumina platforms are

designed for automation of larger numbers of SNP. These platforms allow very large numbers of SNP to be screened with panels of SNP defined for each species or system.

Impact of Advancing DNA Sequencing Technology

DNA sequencing has been subjected to continual improvement with method based on conventional Sanger sequencing predominating until the last few years. Since then, second-generation (or next generation) sequencing technologies have dramatically increased sequencing efficiencies (Varshney and May, 2012). Further advances in DNA sequencing (Thudi *et al.*, 2012) promise to continue the trend to delivery of larger volumes of DNA sequence data at lower costs.

A range of applications of these new sequencing technologies to marker discovery in plants are described in [Table 1.2](#).

Table 1.2 Sequence-based approaches to marker discovery.

Whole-genome sequencing	Krishnan <i>et al.</i> (2012)
Organelle sequencing	Nock <i>et al.</i> (2011)
Transcriptome sequencing	Gillies, Furtado, and Henry (2012)
Enriched genome sequencing	Bundock, Casu, and Henry (2012)

Whole-Genome Sequencing

Assembly of complex plant genomes remains difficult (Souza *et al.*, 2011). Many genomes have now been assembled using next-generation sequencing data (TPGSC, 2011; Kim *et al.*, 2010; Chan *et al.*, 2010; Argout *et al.*, 2010; Lai *et al.*, 2010; Yan *et al.*, 2011). Shotgun resequencing of whole genomes of genotypes is now relatively simple for species with a high-quality reference genome sequence (Krishnan *et al.*, 2012). Short sequence

reads can be aligned to the reference genome sequence for analysis of SNP and indels.

For example, resequencing of six lines of rice identified 2 819 086 SNP, 160 478 insertions, and 163 556 deletions relative to the Nipponbare reference genome (Krishnan *et al.*, 2012). These were parents used in hybrid production allowing all polymorphisms between the parents to be defined for possible crosses within this group.

Organelle Sequencing

Sequencing of organelle genomes has widespread application in research on evolutionary relationships in plants (Zhang, Ma, and Li, 2011; Whittall *et al.*, 2010; Doorduyn *et al.*, 2011). Maternal genome markers may be found in non-nuclear genomes ([Table 1.3](#)). Sequencing of organelle genomes (mitochondria and chloroplast) has required the isolation of the organelle or the specific amplification of sequences from the organelle. Advances in DNA sequencing now allow a simpler approach of analysis of the organelle sequence using a total shotgun sequence of DNA from the sample (Nock *et al.*, 2011).

Table 1.3 Sources of DNA for marker analysis in plants.

	Genome size in rice (bp)
Nucleus	382 000 000
Mitochondria	491 000
Chloroplast	135 000

For example, sequencing of total rice DNA allowed a complete chloroplast sequence to be extracted and revealed a single base error in the published sequence for cultivar Nipponbare. This error in the reference genome was confirmed by Sanger sequencing (Nock *et al.*, 2011).

Transcriptome Sequencing

In many studies, the sequence of the genes expressed in the tissue defines the biological role of the tissue, and studies of differential expression between tissues, developmental stages, and genotypes support gene discovery (Dubey *et al.*, 2011). Sequencing of the transcriptome has been an important way to discover functional differences in plant genomes (Ablett *et al.*, 2000). EST sequencing with Sanger sequencing has been widely used to find SNP in plant genomes. More recent sequencing technologies have made this an even more useful approach to marker discovery (Hiremath *et al.*, 2011).

Sequencing of the transcriptome (Jain, 2012) or RNA-sequencing (RNA-seq) methods have become the preferred way to study differential expression of genes (Auer, Srivastava, and Doerge, 2012).

For example, Pattermore *et al.* (unpublished) sequenced the transcriptome of barley varieties to explore variation between and within cultivars. Gillies, Furtado, and Henry (2012) sequenced the transcriptome of the aleurone and starchy endosperm of developing wheat seeds to explore the differentiation of these two main tissues within the grain and potential for genetic improvement of wheat from a human nutrition and functional properties perspective.

Amplicon Sequencing

The discovery of markers in specific genes or groups of genes can be targeted by amplicon sequencing (Kharabian-Masouleh *et al.*, 2011; Sexton *et al.*, 2012). Very large numbers of genotypes and genes can be screened simultaneously by sequencing the products of amplification from large numbers of genes and or genotypes (Bundock *et al.*, 2009; Kharabian-Masouleh *et al.*, 2011; Malory *et al.*, 2011).

For example, Kharabian-Masouleh *et al.* (2011) sequenced genes of starch biosynthesis in a diverse collection of rice

genotypes to define the contribution of variation in gene sequence to starch properties and the resulting cooking and quality traits of rice.

Sexton *et al.* (2012) sequenced the genes of cell wall metabolism in eucalypts to discover the control of wood properties, identifying a specific gene that was associated with shrinkage of wood.

Enriched Genome Sequencing

The coverage of gene encoding regions of the genome can be improved by enrichment prior to sequencing (Hodges *et al.*, 2007). Temperature-controlled annealing can be used to separate the nonrepetitive sequences in the genomic DNA. Arrays and solution-based hybridization techniques can also be used.

For example, Bundock, Casu, and Henry (2012) used probes designed from the sorghum genome to enrich sugarcane genomic DNA. This allowed a 10-fold enrichment in sequences matching to genes and identified SNP in a large proportion of the genes in the whole genome.

Genotyping by Sequencing

The ultimate genotype data set is the genome sequence of the individual. As the technology to obtain a genome sequence improves, it will become an attractive alternative to the assay of a set of genetic markers as a routine tool for genotyping.

Evolving Range of Applications of DNA Markers in Plants

The recent significant advances in sequencing technology are resulting in an expanded potential for DNA markers to

be applied in plants ([Table 1.4](#)). Easier marker discovery allows application to systems and questions for which marker technology was previously considered too expensive.

Table 1.4 Applications of DNA markers in plants.

IP protection
Regulatory applications
Biosecurity
Conservation biology
Plant functional biology
Evolutionary biology
Plant improvement

Plant Identification for IP Protection

The protection of the intellectual property of plant breeders requires reliable and efficient identification of plant varieties. Molecular markers are now widely used to characterize and protect plant varieties. Policing of plant variety rights requires that evidence of a standard that can be present in a legal court is produced. DNA evidence has been successfully used to prosecute cases of infringement of plant breeders' rights. New molecular technology may make these cases easier and allow more cost-effective protection of plant variety IP.

Plant Variety Identification for Production and Quality Control

Technology developments suggest that early applications of DNA technology to protection of valuable intellectual property will continue to expand to more routine general potential applications in process and quality control in food processing and marketing.

Applications

The identification of plants using DNA markers maybe used to prosecute cases involving prohibited or restricted plants. Conventional phenotypic identification has been used to prosecute legal cases. DNA technology may allow investigations that link to the specific source of the plant material.

Biosecurity Applications

The identification of pest and diseases and the control of weeds can be aided by DNA marker methods. Speed of analysis maybe important in quarantine situations and rapid DNA analysis tools are required. DNA analysis may have special advantages when genotypes very similar in appearance pose very different biosecurity risks.

Applications in Conservation Biology

Conservation of biodiversity requires knowledge of the genetic relationships between individuals and species in wild populations. Molecular tools have been widely applied to conservation biology. However, recent advances in DNA sequencing promise to allow cost-effective extension of these molecular approaches to many more species and systems. DNA analysis has the potential to become a more routine tool in the management of biodiversity.

Application in Evolutionary Biology

Phylogenetic relationships have been increasingly based on DNA sequencing information. Much effort has been devoted to defining appropriate genetic loci for “barcoding” different groups of species. Advancing sequencing technologies (Henry, 2012) may allow complete analysis of the DNA in an organism to be used avoiding many of the difficult decision