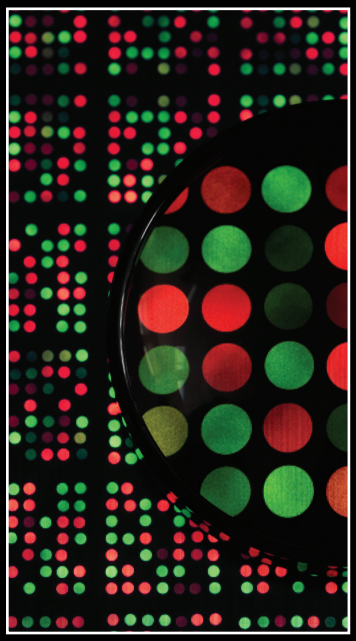
Molecular Markers in Plants

Edited by Robert J. Henry



WILEY-BLACKWELL

Molecular Markers in Plants

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

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.

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.

Cover design by Modern Alchemy LLC

Set in 9.5/12.5pt Palatino by Aptara® Inc., New Delhi, India

1 2013

Contents

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

Molecular Markers in Plants

1 Evolution of DNA Marker Technology in Plants

Robert J. Henry

Contents

Molecular Markers in Plants, First Edition. Edited by Robert J. Henry.

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 though 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 makers 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 makers 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