

Molecular Plant Breeding



Mixed Sources

Product group from well-managed
forests and other controlled sources
www.fsc.org Cert no. SA-COC-1565
© 1996 Forest Stewardship Council

In Memoriam

Norman Ernest Borlaug
(25 March 1914–12 September 2009)

Norman Borlaug was one of the greatest men of our times – a steadfast champion and spokesman against hunger and poverty. He dedicated his 95 richly lived years to filling the bellies of others, and is credited by the United Nations' World Food Program with saving more lives than any other man in history.

An American plant pathologist who spent most of his years in Mexico, it was Dr Borlaug's high-yielding dwarf wheat varieties that prevented wide-spread famine in South Asia, specifically India and Pakistan, and also in Turkey. Known as the 'Green Revolution', this feat earned him the Nobel Peace Prize in 1970. He was instrumental in establishing the International Maize and Wheat Improvement Center, known by its Spanish acronym CIMMYT, and later the Consultative Group of International Agricultural Research (CGIAR), a network of 15 agricultural research centres.

Dr Borlaug spent time as a microbiologist with DuPont before moving to Mexico in 1944 as a geneticist and plant pathologist to develop stem rust resistant wheat cultivars. In 1966 he became the director of CIMMYT's Wheat Program, seconded from the Rockefeller Foundation. His full-time employment with the Center ended in 1979, although he remained a part-time consultant until his death. In 1984 he began a new career as a university professor and went on to establish the World Food Prize, which honours the achievements of individuals who have advanced human development by improving the quality, quantity or availability of food in the world. In 1986, he joined forces with former US President Jimmy Carter and the Nippon Foundation of Japan, under the chairmanship of Ryoichi Sasakawa, to establish Sasakawa Africa Association (SAA) to address Africa's food problems. Since then, more than 1 million small-scale African farmers in 15 countries have been trained by SAA in improved farming techniques.

Dr Borlaug influenced the thinking of thousands of agricultural scientists. He was a path-breaking wheat breeder and, equally important, his stature enabled him to influence politicians and leaders around the world. His legacy and his work ethic – to get things done and not mind getting your hands dirty – influenced us all and remain CIMMYT guiding principles.

We will honor Dr Borlaug's memory by carrying forward his mission and spirit of innovation: applying agricultural science to help smallholder farmers produce more and better-quality food using fewer resources. At stake is no less than the future of humanity, for, as Borlaug said: 'The destiny of world civilization depends upon providing a decent standard of living for all'. His presence will never really leave CIMMYT; it is embedded in our soul.

Thomas A. Lumpkin
Director General, CIMMYT

Marianne Bänziger
Deputy Director General for Research and Partnerships, CIMMYT

Hans-Joachim Braun
Director for Global Wheat Program, CIMMYT

Molecular Plant Breeding

Yunbi Xu

*International Maize and Wheat Improvement Center (CIMMYT)
Apdo Postal 6-641
06600 Mexico, DF
Mexico*



CABI is a trading name of CAB International

CABI Head Office
Nosworthy Way
Wallingford
Oxfordshire OX10 8DE
UK

CABI North American Office
875 Massachusetts Avenue
7th Floor
Cambridge, MA 02139
USA

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

Tel: +1 617 395 4056
Fax: +1 617 354 6875
E-mail: cabi-nao@cabi.org

©CAB International 2010. 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.

Library of Congress Cataloging-in-Publication Data

Xu, Yunbi.

Molecular plant breeding / Yunbi Xu.

p. cm.

ISBN 978-1-84593-392-0 (alk. paper)

1. Crop improvement. 2. Plant breeding. 3. Crops--Molecular genetics. 4. Crops--Genetics. I. Title.

SB106.147X8 2010

631.5'233--dc22

20009033246

ISBN: 978 1 84593 392 0

Typeset by SPi, Pondicherry, India.

Printed and bound in the UK by MPG Books Group.

Contents

Preface	ix
Foreword by Dr Norman E. Borlaug	xv
Foreword by Dr Ronald L. Phillips	xvii
1 Introduction	1
1.1 Domestication of Crop Plants	1
1.2 Early Efforts at Plant Breeding	3
1.3 Major Developments in the History of Plant Breeding	4
1.4 Genetic Variation	9
1.5 Quantitative Traits: Variance, Heritability and Selection Index	10
1.6 The Green Revolution and the Challenges Ahead	16
1.7 Objectives of Plant Breeding	17
1.8 Molecular Breeding	18
2 Molecular Breeding Tools: Markers and Maps	21
2.1 Genetic Markers	21
2.2 Molecular Maps	43
3 Molecular Breeding Tools: Omics and Arrays	59
3.1 Molecular Techniques in Omics	59
3.2 Structural Genomics	68
3.3 Functional Genomics	81
3.4 Phenomics	91
3.5 Comparative Genomics	93
3.6 Array Technologies in Omics	100
4 Populations in Genetics and Breeding	113
4.1 Properties and Classification of Populations	113
4.2 Doubled Haploids (DHs)	116
4.3 Recombinant Inbred Lines (RILs)	131
4.4 Near-isogenic Lines (NILs)	138
4.5 Cross-population Comparison: Recombination Frequency and Selection	145
5 Plant Genetic Resources: Management, Evaluation and Enhancement	151
5.1 Genetic Erosion and Potential Vulnerability	152
5.2 The Concept of Germplasm	155

5.3	Collection/Acquisition	161
5.4	Maintenance, Rejuvenation and Multiplication	166
5.5	Evaluation	171
5.6	Germplasm Enhancement	186
5.7	Information Management	188
5.8	Future Prospects	192
6	Molecular Dissection of Complex Traits: Theory	195
6.1	Single Marker-based Approaches	197
6.2	Interval Mapping	202
6.3	Composite Interval Mapping	205
6.4	Multiple Interval Mapping	209
6.5	Multiple Populations/Crosses	214
6.6	Multiple QTL	217
6.7	Bayesian Mapping	219
6.8	Linkage Disequilibrium Mapping	223
6.9	Meta-analysis	233
6.10	<i>In Silico</i> Mapping	237
6.11	Sample Size, Power and Thresholds	239
6.12	Summary and Prospects	247
7	Molecular Dissection of Complex Traits: Practice	249
7.1	QTL Separating	249
7.2	QTL for Complicated Traits	258
7.3	QTL Mapping across Species	262
7.4	QTL across Genetic Backgrounds	264
7.5	QTL across Growth and Developmental Stages	270
7.6	Multiple Traits and Gene Expression	274
7.7	Selective Genotyping and Pooled DNA Analysis	277
8	Marker-assisted Selection: Theory	286
8.1	Components of Marker-assisted Selection	288
8.2	Marker-assisted Gene Introgression	293
8.3	Marker-assisted Gene Pyramiding	308
8.4	Selection for Quantitative Traits	318
8.5	Long-term Selection	327
9	Marker-assisted Selection: Practice	336
9.1	Selection Schemes for Marker-assisted Selection	337
9.2	Bottlenecks in Application of Marker-assisted Selection	339
9.3	Reducing Costs and Increasing Scale and Efficiency	344
9.4	Traits Most Suitable for MAS	350
9.5	Marker-assisted Gene Introgression	356
9.6	Marker-assisted Gene Pyramiding	363
9.7	Marker-assisted Hybrid Prediction	367
9.8	Opportunities and Challenges	378
10	Genotype-by-environment Interaction	381
10.1	Multi-environment Trials	383
10.2	Environmental Characterization	386
10.3	Stability of Genotype Performance	394
10.4	Molecular Dissection of GEI	402

10.5	Breeding for GEI	410
10.6	Future Perspectives	414
11	Isolation and Functional Analysis of Genes	417
11.1	<i>In Silico</i> Prediction	419
11.2	Comparative Approaches for Gene Isolation	426
11.3	Cloning Based on cDNA Sequencing	431
11.4	Positional Cloning	435
11.5	Identification of Genes by Mutagenesis	441
11.6	Other Approaches for Gene Isolation	454
12	Gene Transfer and Genetically Modified Plants	458
12.1	Plant Tissue Culture and Genetic Transformation	458
12.2	Transformation Approaches	461
12.3	Expression Vectors	468
12.4	Selectable Marker Genes	473
12.5	Transgene Integration, Expression and Localization	480
12.6	Transgene Stacking	487
12.7	Transgenic Crop Commercialization	492
12.8	Perspectives	499
13	Intellectual Property Rights and Plant Variety Protection	501
13.1	Intellectual Property and Plant Breeders' Rights	502
13.2	Plant Variety Protection: Needs and Impacts	504
13.3	International Agreements Affecting Plant Breeding	509
13.4	Plant Variety Protection Strategies	518
13.5	Intellectual Property Rights Affecting Molecular Breeding	524
13.6	Use of Molecular Techniques in Plant Variety Protection	535
13.7	Plant Variety Protection Practice	541
13.8	Future Perspectives	543
14	Breeding Informatics	550
14.1	Information-driven Plant Breeding	550
14.2	Information Collection	554
14.3	Information Integration	562
14.4	Information Retrieval and Mining	568
14.5	Information Management Systems	572
14.6	Plant Databases	579
14.7	Future Prospects for Breeding Informatics	595
15	Decision Support Tools	599
15.1	Germplasm and Breeding Population Management and Evaluation	600
15.2	Genetic Mapping and Marker–Trait Association Analysis	605
15.3	Marker-assisted Selection	613
15.4	Simulation and Modelling	615
15.5	Breeding by Design	621
15.6	Future Perspectives	623
	References	627
	Index	717

The colour plate section can be found following p. 270

This page intentionally left blank

Preface

The genomics revolution of the past decade has greatly enhanced our understanding of the genetic composition of living organisms including many plant species of economic importance. Complete genomic sequences of *Arabidopsis* and several major crops, together with high-throughput technologies for analyses of transcripts, proteins and mutants, provide the basis for understanding the relationship between genes, proteins and phenotypes. Sequences and genes have been used to develop functional and biallelic markers, such as single nucleotide polymorphism (SNP), that are powerful tools for genetic mapping, germplasm evaluation and marker-assisted selection.

The road from basic genomics research to impacts on routine breeding programmes has been long, windy and bumpy, not to mention scattered with wrong turns and unexpected blockades. As a result, genomics can be applied to plant breeding only when an integrated package becomes available that combines multiple components such as high-throughput techniques, cost-effective protocols, global integration of genetic and environmental factors and precise knowledge of quantitative trait inheritance. More recently, the end of the tunnel has come in sight, and the multinational corporations have ramped up their investments in and expectations from these technologies. The challenge now is to translate and integrate the new knowledge from genomics and molecular biology into appropriate tools and methodologies for public-sector plant breeding programmes, particularly those in low-income countries. It is expected that harnessing the outputs of genomics research will be an important component in successfully addressing the challenge of doubling world food production by 2050.

What does *Molecular Plant Breeding* include?

The term ‘molecular plant breeding’ has been much used and abused in the literature, and thus loved or maligned in equal measure by the readership. In the context of this book, the term is used to provide a simple umbrella for the multidisciplinary field of modern plant breeding that combines molecular tools and methodologies with conventional approaches for improvement of crop plants. This book is intended to provide comprehensive coverage of the components that should be integrated within plant breeding programmes to develop crop products in a more efficient and targeted way.

The first chapter introduces some basic concepts that are required for understanding fundamentally important issues described in subsequent chapters. The concepts include crop domestication, critical events in the history of plant breeding, basics of quantitative genetics (variance, heritability and selection index), plant breeding objectives and molecular breeding goals. Chapters 2 and 3 introduce the key genomics tools that are used in molecular breeding programmes, including molecular markers, maps, ‘omics’ technologies and arrays. Different types of molecular markers are compared and construction of molecular maps is discussed. Chapter 4 describes common types of populations that have been used in genetics and plant breeding, with a focus on recombinant inbred lines, doubled haploids and near-isogenic lines. Chapter 5 provides an overview of marker-assisted germplasm evaluation, management and enhancement. Chapters 6 and 7 discuss the theory and practice, respectively, of using molecular markers to dissect complex traits and locate quantitative trait loci (QTL). Chapters 8 and 9 cover the theory and practice, respectively, of marker-assisted selection. Genotype-by-environment interaction (GEI) is discussed in Chapter 10, including multi-environment trials, stability of genotype performance, molecular dissection of GEI and breeding for optimum GEI. Chapter 11 provides a summary of gene isolation and functional analysis approaches, including *in silico* prediction of genes, comparative approaches for gene isolation, gene cloning based on cDNA sequencing, positional cloning and identification of genes by mutagenesis. Chapter 12 describes the use of isolated and characterized genes for gene transfer and the generation of genetically modified plants, focusing on the vital elements of expression vectors, selectable marker genes, transgene integration, expression and localization, transgene stacking and transgenic crop commercialization. Chapter 13 is devoted to intellectual property rights and plant variety protection, including plant breeders’ rights, international agreements affecting plant breeding, plant variety protection strategies, intellectual property rights affecting molecular breeding and the use of molecular techniques in plant variety protection. The last two chapters (14 and 15) discuss supporting tools that are required in molecular breeding for information management and decision making, including data collection, integration, retrieval and mining and information management systems. Decision support tools are described for germplasm and breeding population management and evaluation, genetic mapping and marker-trait association analysis, marker-assisted selection, simulation and modelling, and breeding by design.

Intended audience and guidance for reading and using this book

This book is intended to provide a handbook for biologists, geneticists and breeders, as well as a textbook for final year undergraduates and graduate students specializing in agronomy, genetics, genomics and plant breeding. Although the book has attempted to cover all relevant areas of molecular breeding in plants, many examples have been drawn from the genomics research and molecular breeding of major cereal crops. It is hoped that the book can also serve as a resource for training courses as described below. As each chapter covers a complete story on a special topic, readers can choose to read chapters in any order.

Advanced Course on Quantitative Genetics: Chapters 1, 2, 4, 6, 7, 10 and 14, which cover all molecular marker-based QTL mapping, including markers, maps, populations, statistics and genotype-by-environment interaction.

Comprehensive Course on Marker-assisted Plant Breeding: Chapters 1, 2, 3, 4, 5, 8, 9, 10, 13, 14 and 15, which cover basic theories, tools, methodologies about markers, maps, omics, arrays, informatics and support tools for marker-assisted selection.

Short Course on Genetic Transformation: Chapters 1, 11, 12 and 13, which provide a brief introduction to gene isolation, transformation techniques, genetic-transformation-related intellectual property and genetically modified organism (GMO) issues.

Introductory Course on Breeding Informatics: Chapters 1, 2, 3, 4, 5, 10, 14 and 15, which cover bioinformatics, focusing on plant breeding-related applications, including basic concepts in plant breeding, markers, maps, omics, arrays, population and germplasm management, environment and geographic information system (GIS) information, data collection, integration and mining, and bioinformatics tools required to support molecular breeding. Additional introductory information can be found in other chapters.

History of writing this book

This book has been almost a decade in preparation. In fact, the initial idea for the book was stimulated by the impact from my previous book *Molecular Quantitative Genetics* published by China Agriculture Press (Xu and Zhu, 1994), which was well received by colleagues and students in China and used as a textbook in many universities. Preliminary ideas related to the book were developed in a review article on QTL separation, pyramiding and cloning in *Plant Breeding Reviews* (Xu, 1997). Much of the hopeful thinking described in this paper has fortunately come true during the following 10 years, and the manipulation of QTL has been revolutionized and become mainstream. As complete sequences for several plant genomes have become available and with more anticipated, as shown by numerous genes and QTL that have been separated and cloned individually, some of them have been pyramided for plant breeding through genetic transformation or marker-assisted selection.

I started making tangible progress on this book while working as a molecular breeder for hybrid rice at RiceTec, Inc., Texas (1998–2003). This experience shaped my thinking about how an applied breeding programme could be integrated with molecular approaches. With numerous QTL accumulating for a model crop, taking all the QTL into consideration becomes necessary. Initial thoughts on this were described in ‘Global view of QTL...’, published in the proceedings on quantitative genetics and plant breeding, which considered various genetic background effects and genotype-by-environment interaction (Xu, Y., 2002). Hybrid rice breeding, which involves a three-line system, requires a large number of test-crosses in order to identify traits that perform well in seed and grain production. My experience in development of marker-assisted selection strategies for breeding hybrid rice was then summarized in a review article in *Plant Breeding Reviews* (Xu, Y., 2003), which also covered general strategies for other crops using hybrids.

Moving on to research at Cornell University with Dr Susan McCouch helped me to better understand how molecular techniques could facilitate breeding of complex traits such as water-use efficiency, which is a difficult trait to measure and requires strong collaboration among researchers across many disciplines. In addition, this experience with rice as a model crop raised the issue of how we can use rice as a reference genome for improvement of other crops, which was discussed in an article published in a special rice issue of *Plant Molecular Biology* (Xu *et al.*, 2005).

With over 20 years’ experience in rice, I decided to shift to another major crop by working for the International Maize and Wheat Improvement Center (CIMMYT) as the principle maize molecular breeder. CIMMYT has given me exposure to an interface connecting basic research with applied breeding for developing countries and the resource-poor. Comparing public- and private-sector breeding programmes has given me an intense understanding of the importance of making the type of breeding systems that have been working well for the private sector a practical reality for the public sector, particularly in developing countries. This has been addressed in a recent review paper published in *Crop Science* (Xu and Crouch, 2008), which discussed the critical issues for achieving this translation. My most recent research has focused on the development of various molecular breeding

platforms that can be used to facilitate breeding procedures through seed DNA-based genotyping, selective and pooled DNA analysis, and chip-based large-scale germplasm evaluation, marker–trait association and marker-assisted selection (see Xu *et al.*, 2009b for further details). Thus, my career has evolved alongside the transition from molecular biology research to routine molecular plant breeding applications and I strongly believe that now is the right time for a mainstream publication providing comprehensive coverage of all fields relevant for a new generation of molecular breeders.

Acknowledgements

Assistance and professional support

The dream of writing this book could not have become reality without the wonderful support of Dr Susan McCouch at Cornell University and Dr Jinhua Xiao, now at Monsanto, who have both fully supported my proposal since 2002. Their support and consistent encouragement has greatly motivated me throughout the process. While working with Susan, she allowed me so much flexibility in my research projects and working hours so that I could continue to make progress on the writing of this book. At the same time the Cornell libraries were an indispensable source of the major references cited throughout the book. Susan's encouragement provided the impetus to keep working on the book through a very difficult time in my life. I also extend my appreciation to Dr Jonathan Crouch, the Director of the Germplasm Resources Program at CIMMYT, where I received his full understanding and support so that I could complete the second half of the book. Jonathan's guidance and contribution to my research projects and publications while at CIMMYT has significantly impacted the preparation of the book.

I would also like to thank the chief editors of the three journals for which I have served on the editorial boards during the preparation of this book: Dr Paul Christou for *Molecular Breeding*, Dr Albrecht Melchinger for *Theoretical and Applied Genetics*, and Dr Hongbin Zhang for *International Journal of Plant Genomics*. I thank them for their patience, support and flexibility with my editorial responsibilities during the preparation of the book. In addition, Drs Christou and Melchinger also reviewed several chapters in their respective fields.

My appreciation also goes to Yanli Lu (a graduate student from Sichuan Agricultural University of China) and Dr Zhuanfang Hao (a visiting scientist from the Chinese Academy of Agricultural Sciences) who helped prepare some figures and tables during their work in my lab at CIMMYT, Mexico. I would like to give special thanks to Dr Rodomiro Ortiz at CIMMYT for his consistent information sharing and stimulating discussions during our years together at CIMMYT. Finally, I would like to thank my colleagues at CIMMYT, particularly Drs Kevin Pixley, Manilal William, Jose Crossa and Guy Davenport, who provided useful discussions on various molecular breeding-related issues.

Forewords

I am greatly indebted to Dr Norman E. Borlaug, visioned plant breeder and Nobel laureate for his role in the Green Revolution, and Dr Ronald L. Phillips, Regents Professor and McKnight Presidential Chair in Genomics, University of Minnesota, who each contributed a foreword for the book. Their contributions emphasized the importance of molecular breeding in crop improvement and the role that this book will play in molecular breeding education and practice.

Reviewers

Each chapter of the book has undergone comprehensive peer review and revision before finalization. The constructive comments and critical advice of these reviewers have greatly improved this book. The reviewers were selected for their active expertise in the field of the respective chapter. Reviewers come from almost all continents and work in various fields including plant breeding, quantitative genetics, genetic transformation, intellectual property protection, bioinformatics and molecular biology, many of whom are CIMMYT scientists and managers. Considering that each chapter is relatively large in content, reviewers had to contribute a lot of time and effort to complete their reviews. Although these inputs were indispensable, any remaining errors remain my sole responsibility. The names and affiliations of the reviewers (alphabetically) are:

Raman Babu (Chapters 7 and 9), CIMMYT, Mexico
Paul Christou (Chapter 12), Lleida, Spain
Jose Crossa (Chapter 10), CIMMYT, Mexico
Jonathan H. Crouch (Chapters 13 and 15), CIMMYT, Mexico
Jedidah Danson (Chapters 7 and 9), African Center for Crop Improvement, South Africa
Guy Davenport (Chapter 14), CIMMYT, Mexico
Yuqing He (Chapter 8), Huazhong Agricultural University, Wuhan, China
Gurdev S. Khush (Chapter 1), IRRI, Philippines
Alan F. Krivanek (Chapter 4), Monsanto, Illinois, USA
Huihui Li (Chapter 6), Chinese Academy of Agricultural Sciences, China
George H. Liang (Chapter 12), San Diego, California, USA
Christopher Graham McLaren (Chapter 14), GCP/CIMMYT, Mexico
Kenneth L. McNally (Chapter 5), IRRI, Philippines
Albrecht E. Melchinger (Chapter 8), University of Hohenheim, Germany
Rodomiro Ortiz (Chapters 12, 13 and 15), CIMMYT, Mexico
Edie Paul (Chapter 14), GeneFlow, Inc., Virginia, USA
Kevin V. Pixley (Chapters 1, 4 and 5), CIMMYT, Mexico
Trushar Shah (Chapter 14), CIMMYT, Mexico
Daniel Z. Skinner (Chapter 12), Washington State University, USA
Debra Skinner (Chapter 11), University of Illinois, USA
Michael J. Thomson (Chapters 2 and 3), IRRI, Philippines
Bruce Walsh (Chapters 1, 6 and 8), University of Arizona, USA
Marilyn L. Warburton (Chapter 5), USDA/Mississippi State University, USA
Huixia Wu (Chapter 12), CIMMYT, Mexico
Rongling Wu (Chapter 1), University of Florida, Gainesville, USA
Weikai Yan (Chapter 10), Agriculture and Agri-Food Canada, Ottawa, Canada
Qifa Zhang (Chapters 8 and 12), Huazhong Agricultural University, Wuhan, China
Wanggen Zhang (Chapter 12), Syngenta, Beijing, China
Yuhua Zhang (Chapter 12), Rothamsted Research, UK

Publishers and development editors

Several editors at CABI have been working with me over the years: Tim Hardwick (2002–2006), Sarah Hulbert (2006–2007), Stefanie Gehrig (2007–2008), Claire Parfitt (2008–2009), Meredith Carroll (2009) and Tracy Head (2009). These editors and their associates have done a superb job of converting a series of manuscripts into a useable and coherent book. I thank them for their effort, consideration and cooperation.

Research grants

During the preparation of the book, my research on genomic analysis of plant water-use efficiency at Cornell University was supported by the National Science Foundation (Plant Genome Research Project Grant DBI-0110069). My molecular breeding research at CIMMYT has been supported by the Rockefeller Foundation, the Generation Challenge Programme (GCP), Bill and Melinda Gates Foundation and the European Community, and through other attributed or unrestricted funds provided by the members of the Consultative Group on International Agricultural Research (CGIAR) and national governments of the USA, Japan and the UK.

Family

It is difficult to imagine writing a book without the full support and understanding of one's family. My greatest thanks go to my wife, Yu Wang, who has given me her wholehearted and unwavering support, and to my sons, Sheng, Benjamin and Lawrence, who have retained great patience during this long adventure. And finally to my parents, for their love, encouragement and vision that unveiled in me from my earliest years the desire to thrive on the challenge of always striving to reach the highest mountain in everything I do.

Foreword

DR NORMAN E. BORLAUG

The past 50 years have been the most productive period in world agricultural history. Innovations in agricultural science and technology enabled the 'Green Revolution', which is reputed to have spared one billion people the pains of hunger and even starvation. Although we have seen the greatest reductions in hunger in history, it has not been enough. There are still one billion people who suffer chronic hunger, with more than half being small-scale farmers who cultivate environmentally sensitive marginal lands in developing countries.

Within the next 50 years, the world population is likely to increase by 60–80%, requiring global food production to nearly double. We will have to achieve this feat on a shrinking agricultural land base, and most of the increased production must occur in those countries that will consume it. Unless global grain supplies are expanded at an accelerated rate, food prices will remain high, or be driven up even further.

Spectacular economic growth in many newly industrializing developing countries, especially in Asia, has spurred rapid growth in global cereal demand, as more people eat better, especially through more protein-heavy diets. More recently, the subsidized conversion of grains into biofuels in the USA and Europe has accelerated demand even faster. On the supply side, a slowing in research investment in the developing world and more frequent climatic shocks (droughts, floods) have led to greater volatility in production.

Higher food prices affect everyone, but especially the poor, who spend most of their disposable income on food. Increasing supply, primarily through the generation and diffusion of productivity-enhancing new technologies, is the best way to bring food prices down and secure minimum nutritional standards for the poor.

Today's agricultural development challenges are centred on marginal lands and in regions that have been bypassed during the Green Revolution, such as Africa and resource-poor parts of Asia, and are experiencing the ripple effects of food insecurity through hunger, malnutrition and poverty.

Despite these serious and daunting challenges, there is cause for hope. New science and technology – including biotechnology – have the potential to help the world's poor and food insecure. Biotechnologies have developed invaluable new scientific methodologies and products for more productive agriculture and added-value food. This journey deeper into the genome to the molecular level is the consequence of our progressive understanding of the workings of nature. Genomics-based methods have enabled breeders greater precision in selecting and transferring genes, which has not only reduced the time needed to

eliminate undesirable genes, but has also allowed breeders to access useful genes from distant species.

Bringing the power of science and technology to bear on the challenges of these riskier environments is one of the great challenges of the 21st century. With the new tools of biotechnology, we are poised for another explosion in agricultural innovation. New science has the power to increase yields, address agroclimatic extremes and mitigate a range of environmental and biological challenges.

Molecular Plant Breeding, authored by my CIMMYT colleague Yunbi Xu, is an outstanding review and synthesis of the theory and practice of genetics and genomics that can drive progress in modern plant breeding. Dr Xu has done a masterful job in integrating information about traditional and molecular plant breeding approaches. This encyclopedic handbook is poised to become a standard reference for experienced breeders and students alike. I commend him for this prodigious new contribution to the body of scientific literature.

Foreword

DR RONALD L. PHILLIPS

The New Plant Breeding Roadmap

The road is long from basic research findings to final destinations reflecting important applications – but it is a road that can ultimately save time and money. There may be obstacles along the way that delay building that road but they are generally overcome by careful thought and timely considerations. A new road may involve the former road but with some widening and the filling in of certain potholes. We seldom look back and think that the improvements were not useful.

The road to improved varieties by traditional plant breeding has and continues to serve society well. That approach has been based on careful observation, evaluation of multiple genotypes (parents and progenies), selection at various generational levels, extensive testing and the sophisticated utilization of statistical analyses and quantitative genetics. About 50% of the increased productivity of new varieties is generally attributed to genetic improvements, with the remaining 50% due to many other factors such as time of planting, irrigation, fertilizer, pesticide applications and planting densities.

The statistical genetics associated with traditional plant breeding can now be supplemented by extensive genomic information, gene sequences, regulatory factors and linked genetic markers. We can now draw on a broader genetic base, the identification of major loci controlling various traits and expression analyses across the entire genome under various biotic and abiotic conditions. One can anticipate a future when the networking of genes, genotype-by-environment ($G \times E$) interactions, and even hybrid vigour will be better understood and lead to new breeding approaches. The importance of *de novo* variation may modify much of our current interpretation of breeding behaviour; *de novo* variation such as mutation, intragenic recombination, methylation, transposable elements, unequal crossing over, generation of genomic changes due to recombination among dispersed repeated elements, gene amplification and other mechanisms will need to be incorporated into plant breeding theory.

This book calls for an integration of approaches – traditional and molecular – and represents a theoretical/practical handbook reflecting modern plant breeding at its finest. I believe the reader will be surprised to find that this single-authored book is so full of information that is useful in plant genetics and plant breeding. Students as well as established researchers wanting to learn more about molecular plant breeding will be

well-served by reading this book. The information is up-to-date with many current references. Even many of the tables are packed with information and references. A good representation of international and domestic breeding is reflected through many examples. The importance of $G \times E$ interactions is clearly demonstrated. Various statistical models are provided as appropriate. The importance of defining mega-environments for varietal development is made clear. The role of core germplasm collections, appropriate population sizes, major databases and data management issues are all integrated with various plant breeding approaches. Marker-assisted selection receives considerable attention, including its requirements and advantages, along with the multitude of quantitative trait locus (QTL) analysis methods. Transformation technologies leading to the extensive use of transgenic crops are reviewed along with the increased use of trait stacking. The procurement of intellectual property that, in part, is driving the application of molecular genetics in plant breeding provides the reader with an understanding of why private industry is now more involved and why some common crops represent new business opportunities.

Molecular Plant Breeding is not like other plant breeding books. The interconnecting road that it depicts is one where you can look at the beautiful new scenery and appreciate the current view, yet see the horizon down the road.

1

Introduction

Several definitions of plant breeding have been put forward, such as ‘the art and science of improving the heredity of plants for the benefit of humankind’ (J.M. Poehlman), or ‘evolution directed by the will of man’ (N.I. Vavilov). Bernardo (2002), however, offers the most universal description: ‘Plant breeding is the science, art, and business of improving plants for human benefit.’

Plants are employed in the manufacture of a multitude of products for domestic (cosmetics, medicines and clothing), industrial (manufacture of rubber, cork and engine fuel) and recreational uses (paper, art supplies, sports equipment and musical instruments) and plant breeders have therefore been driven by the challenges of meeting the ever increasing demands of the manufacturers of these products. Lewington has described the diverse uses of plants in his book *Plants for People* (2003).

Plant breeding began by the domestication of crop plants and has become ever more sophisticated. New developments in molecular biology have now led to an increasing number of methods which can be used to enhance breeding effectiveness and efficiency. This chapter includes a brief history of plant breeding together with breeding objectives and some background information relevant to the theories

and technologies discussed in the following chapters of this book.

1.1 Domestication of Crop Plants

The earliest records indicate that agriculture developed some 11,000 years ago in the so-called Fertile Crescent, a hilly region in south-western Asia. Agriculture developed later in other regions. Archaeologists suggest that plant domestication began because of the increasing size of populations and changes in the exploitation of local resources (see <http://www.ngdc.noaa.gov/paleo/ctl/10k.html> for further details). Domestication is a selection process carried out by man to adapt plants and animals to their own needs, whether as farmers or consumers. Successive selection of desirable plants changed the genetic composition of early crops. Primitive farmers, knowing little or nothing about genetics or plant breeding, accomplished much in a short time. They did so by unconsciously altering the natural process of evolution. Indeed, domestication is nothing more than directed evolution; as a result, the process of evolution is accelerated. The key to domestication is the selective advantage of rare mutant alleles, which are desirable for successful cultivation,

but unnecessary for survival in the wild. The process of selection continues until the desired mutant phenotype dominates the population. There are three important steps in the domestication process. Man not only planted seeds, but also: (i) moved seeds from their native habitat and planted them in areas to which they were perhaps not as well adapted; (ii) removed certain natural selection pressures by growing the plants in a cultivated field; and (iii) applied artificial selection pressures by choosing characteristics that would not necessarily have been beneficial for the plants under natural conditions. Cultivation also creates selection pressure, resulting in changes in allele frequency, gradations within and between species, fixation of major genes, and improvement of quantitative traits. By the end of the 18th century, the informal processes of selection practised by farmers everywhere led to the worldwide creation of thousands of different cultivars or landraces for each major crop species.

More than 1000 species of plants have been domesticated at one time or another, of which about 100–200 are now major components of the human diet. The 15 most important examples can be divided into the following four groups:

1. Cereals: rice, wheat, maize, sorghum, barley.
2. Roots and stems: sugarbeet, sugarcane, potato, yam, cassava.
3. Legumes: bean, soybean, groundnut.
4. Fruits: coconut, banana.

Certain characteristics may have been selected deliberately or unwittingly. When farmers set aside a portion of their harvest for planting in the next season, they were selecting seeds with specific characteristics. This selection has resulted in profound differences between crop plants and their progenitors. For example, many wild plants have a seed dispersal mechanism that ensures that seeds will be separated from the plants and distributed over as large an area as possible, while modern crops have been modified by selection against seed dispersal. The absence of seed dormancy mechanisms in some

domesticated plants is another example. For further information see <http://oregon-state.edu/instruct/css/330/index.htm> and Swaminathan (2006).

It is generally believed that domestication of crop plants was undertaken in several regions of the world independently. The Russian geneticist and plant geographer N.I. Vavilov, collected plants from all over the world and identified regions where crop species and their wild relatives showed great genetic diversity. In 1926 he published 'Studies on the origin of cultivated plants' in which he described his theories regarding the origins of crops. Vavilov concluded that each crop had a characteristic primary centre of diversity which was also its centre of origin. He identified eight areas and hypothesized that these were the centres from which all our modern major crops originated. Later, he modified his theory to include 'secondary centres of diversity' for some crops. These 'centres of origin' included China, India, Central Asia, the Near East, the Mediterranean, East Africa, Mesoamerica, and South America. From these foci, agriculture was progressively disseminated to other regions such as Europe and North America. Subsequently, others including the American geographer Jack Harlan, challenged Vavilov's hypothesis because many cultivated plants did not fit Vavilov's pattern, and appeared to have been domesticated over a broad geographical area for a long period of time.

In recent years, variation in DNA fractions and other approaches have been used to study the diversity of crop species. In general, these studies have not confirmed Vavilov's theory that the centres of origin are the areas of greatest diversity, because while centres of diversity have been identified, these are often not the centres of origin. For some crops there is little connection between the source of their wild ancestors, areas of domestication, and the areas of evolutionary diversification. Species may have originated in one geographic area, but domesticated in a different region and some crops do not appear to have centres of diversity, thus a continuum of evolution-

ary activity is perceived rather than discrete centres.

In 1971, Jack Harlan described his own views on the origins of agriculture. He proposed three independent systems, each with a centre and a 'concentre' (larger, diffuse areas where domestication is thought to have occurred): Near East + Africa, China + South-east Asia, and Mesoamerica + South America.

Evidence gathered since that time suggests that these centres are also more diffuse than he had envisioned. After the initial phases of evolution, species spread out over large, ill-defined areas. This is probably due to the dispersal and evolution of crops associated with itinerant populations. Regional and/or multiple areas of origin may prove to be more accurate than the hypothesis of a unique, localized origin for many crops. However, the probable geographic origin of many crops is listed in Table 1.1.

1.2 Early Efforts at Plant Breeding

For thousands of years selective breeding has been employed to re-engineer plants to produce traits or qualities that were considered to be desirable to consumers. Selective breeding began with the early farmers, ranchers and vintners who

selected the best plants to provide seed for their next crop. When they found particular plants that fared well even in bad weather, were especially prolific, or resisted disease that had destroyed neighbouring crops, they naturally tried to capture these desirable traits by crossbreeding them into other plants. In this way, they selected and bred plants to improve their crop for commercial purposes. Although unbeknown to them, farmers have been utilizing genetics for centuries to modify the food we eat by selecting and growing seeds which produce a healthier crop that has a better flavour, richer colour and stronger resistance to certain plant diseases.

Modern plant breeding started with sedentary agriculture and the domestication of the first agricultural plants, cereals. This led to the rapid elimination of undesirable characters such as seed-shattering and dormancy and we can only speculate on how much foresight or what kind of planning based on experience was used by the first selectors of non-shattering wheat and rice, compact-headed sorghum, or soft-shelled gourds. For 10,000 years man has consciously been moulding the phenotype (and so the genotype) of hundreds of plant species as one of the many routine activities in the normal course of making a living (Harlan, 1992). Over long periods of time there was a transition from the collection of

Table 1.1. Probable geographic origins for crops.

Region	Crops
Near East (Fertile Crescent)	Wheat and barley, flax, lentils, chickpea, figs, dates, grapes, olives, lettuce, onions, cabbage, carrots, cucumbers, melons; fruits and nuts
Africa	Pearl millet, Guinea millet, African rice, sorghum, cowpea, groundnut, yam, oil palm, watermelon, okra
China	Japanese millet, rice, buckwheat, soybean
South-east Asia	Wet- and dryland rice, pigeon pea, mung bean, citrus fruits, coconut, taro, yams, banana, breadfruit, coconut, sugarcane
Mesoamerica and North America	Maize, squash, common bean, lima bean, peppers, amaranth, sweet potato, sunflower
South America	Lowlands: cassava; Mid-altitudes and uplands (Peru): potato, groundnut, cotton, maize

See <http://agronomy.ucdavis.edu/gepts/pb143/lec10/pb143i10.htm> for a thorough presentation on the geographic origins of crops.

wild plants for food to the selection of those to be cultivated which began to guide the evolutionary process. Now plant breeders accelerate the evolution of major crop species through skilful manipulation of breeding procedures. High-input agriculture emerged as a result of voyages of discovery and modern science.

Many traits important to early agriculturists were heritable and, therefore, could be reliably selected. However, this phase of breeding was empirical and generally not considered scientific in the modern sense because changes in these plant and animal populations were not analysed in an attempt to explain biological phenomena. At this stage of agriculture, the focus was on the practical goal of producing food rather than finding rational explanations for nature (Harlan, 1992). Ideas about heredity during the period when many early crops were domesticated ranged from mythological interpretations to near-scientific notions of trait transmission. In his Presidential Address to the American Society for Horticulture Science in 1987, Janick (1988) stated:

The origin of new information in horticulture derives from two traditions: empirical and experimental. The roots of empiricism stem from efforts of prehistoric farmers, Hellenic root diggers, medieval peasants, and gardeners everywhere to obtain practical solutions to problems of plant growing. The accumulated successes and improvements passed orally from parent to child, from artisan to apprentice, have become embodied in human consciousness via legend, craft secrets, and folk wisdom. This information is now stored in tales, almanacs, herbals, and histories and has become part of our common culture. More than practices and skills were involved as improved germplasm was selected and preserved via seed and graft from harvest to harvest and generation to generation. The sum total of these technologies makes up the traditional lore of horticulture. It represents a monumental achievement of our forbears – unknown and unsung.

Large-scale breeding activities began very early in Europe, often under the auspices

of commercial seed production enterprises. Besides selecting plants with useful characteristics, breeders also arrange ‘marriages’ between plants with different traits in the hope of producing fertile offspring carrying both traits. The use of artificial crosses in pre-Mendelian breeding is exemplified by the case of *Fragaria* × *ananassa* developed in the botanical garden of Paris by Duchesne, in the 17th century by crossing *Fragaria chiloense* with *Fragaria virginiana*. In England, at about the same time new cultivars of fruits, wheat and peas were being obtained by artificial hybridization (Sánchez-Monge, 1993).

Hybridization combined with selection was adopted by Patrik Sheireff in 1819 in wheat and rice where the new selections were grown along with cultivars for comparative purposes. He speculated that introduction and hybridization to be the important sources of new cultivars and stressed crossing of carefully selected parents to meet the aims of new cultivars. Although the essential elements of plant breeding were known by this time, there was still a lack of knowledge regarding the scientific basis of variation among plants. For example, the first generation of crossed materials were mistakenly expected to inevitably produce new cultivars but instead took several generations to stabilize. Many historical examples of successful plant breeding can be found in the literature, although there were still many important discoveries to be made before it could be called a technology (Chahal and Gosal, 2002).

1.3 Major Developments in the History of Plant Breeding

Plant breeders of today use various methods to accelerate the evolutionary process in order to increase the usefulness of plants by exploiting genetic differences within a species. This has been made possible by the determination of the genetic basis for developing crop breeding procedures and this in turn has a long history.

1.3.1 Breeding and hybridization

The role of reproduction in plants was first reported in 1694 by Camerarius who noticed the difference between male and female reproductive organs in maize and produced the first artificial hybrid plant. He established that seed could not be produced without the participation of pollen produced in male reproductive organs of plants. The first hybridization experiment was carried out on wheat by Fairchild in 1719 and the current technique of hybridization is largely based on the work of Kölreuter (1733–1806), a French researcher who carried out his experiments in the 1760s. Hybridization freed the breeder from the severe constraints of working within a limited population, enabled him to bring together useful traits from two or more sources, and allowed specific genes to be introduced.

By understanding the reproductive capacities of plants, plant breeders can manipulate these crosses to produce fertile offspring which carry traits from both parents. Crossing has been very valuable to plant breeders, because it allows some measure of control over the phenotype of a plant. Nearly all modern plant breeding involves some use of hybridization.

1.3.2 Mendelian genetics

It was Gregor Johann Mendel, a Moldavian monk, who in 1865 discovered the basic rules that govern heredity as a result of a series of experiments in which he crossed two cultivars of pea plants. By studying the inheritance of all-or-none variation in peas, Mendel discovered that inherited traits are determined by units of material that are transferred from one generation to another. Mendel was probably ahead of his time as other biologists of that era took 35 years to appreciate his work and plant breeding remained deprived of the deliberate application of the law of genetics until 1900 when Hugo de Vries, Carl Correns and Erich von Tschermak-Seysenegg rediscovered Mendel's work.

1.3.3 Selection

In 1859 Darwin proposed in *The Origin of Species* that natural selection is the mechanism of evolution. Darwin's thesis was that the adaptation of populations to their environments resulted from natural selection and that if this process continued for long enough, it would ultimately lead to the origin of new species. Darwin's 'Theory of Evolution through Natural Selection' hypothesized that plants change gradually by natural selection operating on variable populations and was the outstanding discovery of the 19th century with direct relevance to plant breeding.

1.3.4 Breeding types and polyploidy

Other historical developments in plant breeding include, pedigree breeding, back-cross breeding (Harlan and Pope, 1922) and mutation breeding (Stadler, 1928). Natural and artificial polyploids also offered new possibilities for plant breeding. Blakeslee and Avery (1937) demonstrated the usefulness of colchicine in the induction of chromosome doubling and polyploidy, enabling plant breeders to combine entire chromosome sets of two or more species to evolve new crop plants.

1.3.5 Genetic diversity and germplasm conservation

The importance of genetic diversity in plant breeding was recognized by the 1960s and Sir Otto Frankel coined the term 'genetic resources' in 1967 to highlight the relevance and need to consider germplasm as a natural resource for the long-term improvement of crop plants. The potentially harmful effects of genetic uniformity became apparent with the epidemic of southern corn leaf blight in the USA in 1970 which destroyed about 15% of US maize in just 1 year. The National Academy of Sciences, USA, released the results of its study *Genetic Vulnerability*

in *Major Crops* that brought into focus the causes and levels of genetic uniformity and its consequences. It was a turning point in the history of germplasm resources and the International Board for Plant Genetic Resources (IBPGR) was established in 1974, and was later renamed the International Plant Genetic Resources Institute (IPGRI) and now Biodiversity International, to collect, evaluate and conserve plant germplasm for future use.

1.3.6 Quantitative genetics and genotype-by-environment interaction

Quantitative genetics is the study of the genetic control of those traits which show continuous variation. It is concerned with the level of inheritance of these differences between individuals rather than the type of differences, that is quantitative rather than qualitative (Falconer, 1989). Several important books have been published which document the major developments in quantitative genetics and these include *Animal Breeding Plans* (Lush, 1937), *Population Genetics and Animal Improvement* (Lerner, 1950), *Biometrical Genetics* (Mather, 1949), *Population Genetics* (Li, 1955), *An Introduction to Genetics Statistics* (Kempthorne, 1957) and *Introduction to Quantitative Genetics* (Falconer, 1960).

Many of the misconceptions regarding the inheritance of quantitative traits, which include most of the economically important characters, were corrected by the classical work of Fisher (1918) who successfully applied Mendelian principles to explain the genetic control of continuous variation. He divided the phenotypic variance observed into three variance components: additive, dominance and epistatic effects. This approach has been substantially refined and applied to the improvement of the efficiency of plant breeding. Fisher also laid the foundations for scientific crop experimentation by developing the theory of experimental designs that is an essential part of any plant breeding programme. Quantitative genetics has however evolved considerably in the past two decades because of the development

of plant genomics, particularly molecular markers, and other molecular tools that can be used to dissect complex traits into single Mendelian factors (Xu and Zhu, 1994; Buckler *et al.*, 2009; Chapters 6 and 7).

Genotype-by-environment interaction (GEI) and its importance to plant breeding were first recognized by Mooers (1921) and Yates and Cochran (1938). Since then, various statistical methods have been developed for the evaluation of GEI using joint linear regression, heterogeneity of variance and lack of correlation, ordination, clustering, and pattern analysis. As an important field in quantitative genetics, GEI has been receiving more attention in recent years and is covered in Chapter 10 along with molecular methods for GEI analysis.

1.3.7 Heterosis and hybrid breeding

Although early botanists had observed increased growth when unrelated plants of the same species were crossed, it was Charles Darwin who carried out the first seminal experiments. In 1877, he showed that crosses of related strains did not exhibit the vigour of hybrids. He observed heterosis, i.e. the tendency of cross-bred individuals to show qualities superior to those of both parents, in crops like maize and concluded that cross-fertilization was generally beneficial and self-fertilization injurious. In 1879, William Beal demonstrated hybrid vigour in maize by using two unrelated cultivars. The best combinations yielded 50% more than the mean of the parents. Reports by Sanborn in 1890 and McClure in 1892 confirmed Beal's earlier reports and extended the generality of the superiority of hybrids over the average of the parental forms.

1.3.8 Refinement of populations

Several different 'population breeding methods' can be used: (i) bulk; (ii) mass selection; and (iii) recurrent selection. One

of the methods used for managing large populations of segregates was the 'bulk method' proposed by Harlan *et al.* (1940) for multi-parent crosses. This concept changed the breeding methodologies for self-pollinated species. Mass selection is a system of breeding in which seeds from individuals selected on the basis of phenotype are bulked and used to grow the next generation. Mass selection is the oldest breeding method for plant improvement and was employed by early farmers for the development of cultivated species from their ancestral forms.

The enhancement of open-pollinated populations of crops such as rye, maize and sugarbeet, herbage grasses, legumes, and tropical trees such as cacao, coconut, oil palm, and some rubber, depends essentially on changing the gene frequencies so that the favourable alleles are fixed, while maintaining a high (but far from maximal) degree of heterozygosity. Recurrent selection is a method of plant breeding associated with quantitatively inherited traits by which the frequencies of favourable genes are increased in populations of plants. The methodology is cyclical with each cycle encompassing two phases: (i) selection of genotypes that possess the favourable or required genes; and (ii) crossing among the selected genotypes. This leads to a gradual increase in the frequencies of the desired alleles. While recurrent selection is often successful it also has potential limitations in closed populations and this has led to numerous modifications and alternative schemes (see Hallauer and Miranda, 1988). Recurrent selection breeding methods have been applied to a wide range of plant species, including self-pollinated crops.

1.3.9 Cell totipotency, tissue culture and somaclonal variation

The discovery of auxins, by Went and colleagues, and cytokinins, by Skoog and colleagues, preceded the first success of *in vitro* culture of plant tissues (White, 1934; Nobécourt, 1939).

All the genes necessary to make an entire organism can be induced to function in the correct sequence from a living cell isolated from a mature tissue (called totipotency). Regeneration of whole plants from single cells is an important new source of genetic variability for refining the properties of plants because when somatic embryos derived from single cells are grown into plants, the plants' characteristics vary somewhat. Larkin and Scowcroft (1981) coined the term 'somaclonal variation' to describe this observed phenotypic variation among plants derived from micro-propagation experiments. When it was recognized as a genuine phenomenon, somaclonal variation was considered to be a potential tool for the introduction of new variants of perennial crops that can be asexually propagated (e.g. banana). Somaclonal variation has also been exploited by plant breeders as a new source of genetic variation for annual crops.

1.3.10 Genetic engineering and gene transfer

The discovery of the structure of DNA by Watson and Crick has enhanced traditional breeding techniques by allowing breeders to pinpoint the particular gene responsible for a particular trait and to follow its transmission to subsequent generations. Enzymes that cut and rejoin DNA molecules allow scientists to manipulate genes in the laboratory. In 1973 Stanley Cohen and Herbert Boyer spliced the gene from one organism into the DNA of another to produce recombinant DNA which was then expressed normally and this formed the basis of genetic engineering. The goal of plant genetic engineers is to isolate one or more specific genes and introduce these into plants. Improvement in a crop plant can often be achieved by introducing a single gene, and genes can now be transferred to plants using the natural gene transfer system of a promiscuous pathogenic soil bacterium, *Agrobacterium tumefaciens*. DNA can also be introduced into cells by bombardment with DNA-coated particles or by electroporation. Transgenic breeding

has the potential to decrease or increase the environmental impact of agricultural practices.

The initial successes in plant genetic engineering marked a significant turning point in crop research. In the 1990s in particular, there was an upsurge of private sector investment in agricultural biotechnology. Some of the first products were plant strains capable of synthesizing an insecticidal protein encoded by a gene isolated from the bacterium *Bacillus thuringiensis* (*Bt*). *Bt* cotton, maize, and other crops are now grown commercially. There are also crop cultivars which are tolerant to or capable of degrading herbicides. Proponents stress the value of these crops in conserving tillage soil, reducing the use of harmful chemicals and reducing the labour and costs involved in crop production.

1.3.11 DNA markers and genomics

During the 1980s and 1990s, various types of molecular markers such as restriction fragment length polymorphism (RFLP) (Botstein *et al.*, 1980), randomly amplified polymorphic DNA (RAPD) (Williams *et al.*, 1990; Welsh and McClelland, 1990), microsatellites and single nucleotide polymorphism (SNP) were developed. Because of their abundance and importance in the plant genome, molecular markers have been widely used in the fields of germplasm evaluation, genetic mapping, map-based gene discovery and marker-assisted plant breeding. Molecular marker technology has become a powerful tool in the genetic manipulation of agronomic traits.

Initiated by the complete sequencing of the *Arabidopsis* genome in 2000 (The *Arabidopsis* Genome Initiative, 2000) and the rice genome in 2002 (Goff *et al.*, 2002; Yu *et al.*, 2002), the genomes of an increasing number of plants have been or are being sequenced. Technological developments in bioinformatics, genomics and various omics fields are creating substantial data on which future revolutions in plant breeding can be based.

1.3.12 Breeding efforts in the public and private sectors

Agricultural research has mainly been the responsibility of a national and/or state government department. To accelerate progress in food production especially in developing countries, international agricultural research centres were established with major emphasis on the development of high yielding cultivars. Two centres, International Rice Research Institute (IRRI), Philippines, and Centro Internacional de Mejoramiento de Maiz y Trigo (CIMMYT), Mexico, established in the 1960s, made phenomenal contributions to food production by developing shorter and higher-yielding rice, wheat and maize cultivars. Encouraged by the astonishing success of these centres and two others which were established later, the Consultative Group on International Agricultural Research (CGIAR) was established in 1971. The CGIAR now has 15 international agricultural research centres, of which eight concentrate on specific crop plants and one on genetic resources with a mission to contribute towards sustainable agriculture for food security especially in developing countries. The breeding materials developed at these centres are distributed to public and private sector research programmes for utilization in the development of locally adapted cultivars. Through National Agricultural Research Systems (NARS), these centres work in close coordination with public and private breeding programmes in each country and share their breeding technologies and stocks of germplasm.

In the USA, crop breeding, with the exception of cotton, began largely as a tax-supported endeavour with breeding programmes taking place in most State Agricultural Experimental Stations and in the United States Department of Agriculture (USDA). This pattern changed with the advent of hybrid maize when inbred lines were initially developed by public institutions and utilized to produce hybrids by private companies. With the implementation of a Plant Variety Protection Act in the USA in 1974, private breeding was expanded to

include forages, cereals, soybean, and other crops. The activities of private companies contributed to the total crop breeding effort and offered a large number of cultivar options for farmers and consumers. In the USA and other industrialized countries today, the new life-science companies notably the big multinationals such as Dow, DuPont and Monsanto, dominate the application of biotechnology to agriculture, and have developed many proprietary products.

1.4 Genetic Variation

The creation of new alleles and the mixing of alleles through recombination give rise to genetic variation which is one of the forces behind evolution. Natural selection favours one phenotype over another and these phenotypes are conditioned by one or more alleles. Genetic variation is fundamental for selection, by which progress in plant breeding can be made. There are various sources of genetic variation and those described in this section are largely based on the information provided at the following web sites: <http://www.ndsu.nodak.edu/instruct/mcclean/plsc431/mutation> and <http://evolution.berkeley.edu/evosite/evo101/IIICGeneticvariation.shtml>.

1.4.1 Crossover, genetic drift and gene flow

Chromosomal crossover takes place during meiosis and results in a chromosome with a completely different chemical composition from the two parent chromosomes. During the process, two chromosomes intertwine and exchange one end of the chromosome with the other. The mechanism of crossing over is the cytogenetic base for recombination.

Gene flow refers to the passage of traits or genes between populations to prevent the occurrence of large numbers of mutations and genetic drift. In genetic drift, random variation occurs in small populations

leading to the proliferation of specific traits within that population. The degree of gene flow varies widely and is dependent on the type of organism and population structure. For example genes in a mobile population are likely to be more widely distributed than those in a sedentary population, resulting in high and low rates of gene flow, respectively.

1.4.2 Mutation

A mutation is any change in the sequence of the DNA encoding a gene which leads to a change in the hereditary material when an organism undergoes DNA replication. During the process of replication, the nucleotides of a chromosome are altered, so rather than creating an identical copy of DNA strands, there are chemical variations in the replicated strands. The alteration on the chemical composition of DNA triggers a chain reaction in the genetic information of an individual. The effect of a mutation depends on its size, location (intron or exon etc.), and the type of cell in which the mutation occurs. Large changes involve the loss, addition, duplication or rearrangement of whole chromosomes or chromosome segments. Most DNA polymerases have the ability to proofread their work to ensure that the unaltered genetic material is transferred to the next generation. There are many types of mutation and the most common are listed below.

1. Point mutations represent the smallest changes where only a single base is altered. For example, a single nucleotide change may result in the change of an amino acid (aa) codon into a stop codon and thus produce a change in the phenotype. Point mutations do not usually benefit the organism as most occur in recessive genes and are not usually expressed unless two mutations occur at the same locus.

2. In synonymous or silent substitutions the aa sequence of the protein is not changed because several codons can code for the same aa, and in non-synonymous

substitutions changes in the aa sequence may not affect the function of the protein. However, there have been many cases where a change in a single nucleotide can create serious problem, e.g. in sickle cell anaemia.

3. Wild-type alleles typically encode a product necessary for a specific biological function and if a mutation occurs in that allele, the function for which it encodes is also lost. The general term for these mutations is loss-of-function mutations and they are typically recessive. The degree to which the function is lost can vary. If the function is entirely lost, the mutation is called a null mutation. It is also possible that some function may remain, but not at the level of the wild-type allele, these are known as leaky mutations.

4. A small number of mutations are actually beneficial to an organism providing new or improved gene activity. In these cases, the mutation creates a new allele that is associated with a new function. Any heterozygote containing the new allele along with the original wild-type allele will express the new allele. Genetically this will define the mutation as a dominant. This class of mutation is known as gain-of-function mutations.

5. A substitution is a mutation in which one base is exchanged for another. Such a substitution could change: (i) a codon to one that encodes a different aa thus causing a small change in the protein produced; (ii) a codon to one that encodes the same aa resulting in no change in the protein produced; or (iii) an aa-coding codon to a single 'stop' codon resulting in an incomplete protein (this can have serious effects since the incomplete protein will probably not be functional).

6. Insertions/deletions (indels) produce changes by deleting or inserting sections of DNA into the 'parental' DNA sequence. Because it is usually impossible to say whether a sequence has been deleted from one plant or inserted into another, these differences are called indels. Obviously the deletion of part of a gene can seriously affect the phenotype of organisms. Insertions can be disruptive if they insert themselves into the middle of genes or regulatory regions.

7. A mutation in which one nucleotide is changed causing all the codons to its right to be altered is known as frame-shift mutation. Since protein-coding DNA is divided into codons of three bases long, insertions and deletions of a single base can alter a gene so that its message is no longer correctly parsed. As a result, a single base change can have a dramatic effect on a polypeptide sequence.

8. Mutations which occur in germ line cells including both the gametes and the cells from which they are formed are known as germinal mutations. A single germ line mutation can have a range of effects: (i) no phenotypic change; mutations in junk DNA are passed on to the offspring but have no obvious effect on the phenotype; (ii) small (or quantitative) phenotypic changes; and (iii) significant phenotypic change.

9. Mutations in somatic cells which give rise to all non-germ line tissues, only affect the original individual and cannot be passed on to the progeny. To maintain this somatic mutation, the individual containing the mutation must be cloned.

In general, the appearance of a new mutation is a rare event. Most mutations that were originally studied occurred spontaneously. Such spontaneous mutations represent only a small number of all possible mutations. To genetically dissect a biological system further, induced mutations can be created by treating an organism with a mutagenic agent.

1.5 Quantitative Traits: Variance, Heritability and Selection Index

Recent advances in high-throughput technologies for the quantification of biological molecules have shifted the focus in quantitative genetics from single traits to comprehensive large-scale analyses. So-called omic technologies have now enabled geneticists to determine how genetic information is translated into biological function (Keurentjes *et al.*, 2008; Mackay *et al.*, 2009). The ultimate goal of quantitative genetics in the era of omics is to link genetic variation

to phenotypic variation and to identify the molecular pathway from gene to function. The recent progress made in humans by combining linkage disequilibrium mapping (Chapter 6) and transcriptomics (Chapter 3) holds great promises for high-resolution association mapping and identification of regulatory genetic factors (Dixon *et al.*, 2007). Information from omics research will be integrated with our current knowledge at the phenotypic level to increase the effectiveness and efficiency of plant breeding.

1.5.1 Qualitative and quantitative traits

In general, qualitative traits are genetically controlled by one or a few major genes, each of which has a relatively large effect on the phenotype but is relatively insensitive to environmental influences. Trait distribution in a typical segregating population such as an F_2 shows multi-peak distribution, although individuals within a category show continuous variation. Each individual in the population can be classified unambiguously into distinct categories that correspond to different genotypes so that they can be studied using Mendelian methods.

Quantitative traits are genetically controlled by many genes, each of which has a relatively small effect on the phenotype, but is largely influenced by environmental factors (Buckler *et al.*, 2009). Trait distribution in an F_2 population usually shows a normal or bell-shape distribution and as a result, individuals cannot be classified into phenotypic categories that correspond to different genotypes thus making the effects of individual genes indistinguishable. Quantitative genetics is traditionally described as the study of all these genes as a whole and the total variation observed in a population results from the combined effects of genetic (polygenes as a whole) and environmental factors. However, quantitative variation is not due solely to minor allelic variation in structural genes as regulatory genes no doubt also contribute to this variation. We expected polygenes to show all the typical properties of chromosomal genes

both in terms of action and in transmission through meiosis.

1.5.2 The concept of allelic and genotypic frequencies

A biological population is defined genetically as a group of individuals that exist together in time and space and that can mate or be crossed to each other to produce fertile progeny. Statistically, this group is called a 'population'. Breeding populations are created by breeders to serve as a source of cultivars that meet specific breeding objectives.

At the population level, genetics can be characterized by allelic and genotypic frequencies. The allele frequencies refer to the proportion of each allele in the population, while the genotypic frequencies refer to the proportion of individuals (plants) in the population that have a particular genotype. A gene may have many allelic states. Some of the alleles of a given gene may have such marked effects as to be clearly recognized as a classical major mutant. Other alleles, though potentially separable at the DNA level, may well cause only minor differences at the level of the external phenotype. For example, one allele at a locus involved with growth hormone production could be inactive and result in a dwarf plant, while others may simply reduce or increase height by a few percent.

Allele and genotypic frequencies can be calculated by simple counting in the population. For a gene with n alleles, there are $n(n + 1)/2$ possible genotypes. The relationship between allele frequency and genotypic frequency for a single gene at the population level can be used to infer the genetic status of the gene in that population, relative to the expected equilibrium under some assumed mating system. Allele frequencies are generally not an issue in breeding populations created from non-inbred parents or from three or more inbred parents. But breeding populations in both self-pollinated and cross-pollinated crops are often created by crossing two inbred individuals.

1.5.3 Hardy–Weinberg equilibrium (HWE)

A population is in equilibrium if the allele and genotypic frequencies are constant from generation to generation. A collection of pure selfers is also at equilibrium if all are completely selfed, with $P_{A_1A_1} = p$ and $P_{A_2A_2} = q$. This implies that the allele frequency and genotypic frequency share a simple relationship:

$$\begin{aligned} P_{A_1A_1} &= p^2 \\ P_{A_1A_2} &= 2pq \\ P_{A_2A_2} &= q^2 \end{aligned}$$

or

$$(p + q)^2 = p^2 + 2pq + q^2$$

With one generation of random mating, i.e. an individual in the population that is equally likely to mate with any other individual, the above simple relationship will be obeyed. However, HWE represents idealized populations and breeders routinely use procedures that cause deviations from HWE. These procedures include the lack of random mating, the use of small population sizes, assortative mating, selection, and inbreeding during the development of progenies. Some of these procedures, such as inbreeding and the use of small population sizes, affect all loci in the population while others affect only certain loci. Suppose that two traits are controlled by different sets of loci, and a change in one trait does not affect the other. If selection occurs only for the first trait, the loci affecting that trait may deviate from HWE, but the loci for the other trait will remain in equilibrium. In large natural populations, migration, mutation, and selection are the forces that can change allelic frequencies from generation to generation.

1.5.4 Population means and variances

Theoretically, a population can be described by its parameters such as the mean and variance which depend on the probability distribution of the population. The arithmetic

mean, μ , also known as the first moment about the origin, is a parameter used to measure the central location of a frequency distribution. The population variance, σ^2 , also known as the second moment about the mean, provides measures of the dispersion of the distribution. If the yield trait for a cultivar that is genetically homogenous is taken as an example, the genetic effect for this cultivar population is a constant. The yield for all individuals should also be a constant provided that environmental factors do not affect the yield which is equal to the population mean. However, the yield for each individual is affected not only by its genotype but also by environmental factors such as temperature, sunlight, water, and various nutrients. As a result, individuals may have different phenotypic values, in this case yield, resulting in continuous variation among individuals. Therefore, the individual yield measures vary either positively or negatively around the population mean so that they are either higher or lower than the population mean by a certain number which is determined by its variance.

1.5.5 Heritability

The response of traits to selection depends on the relative importance of the genetic and non-genetic factors which contribute to phenotypic variation among genotypes in a population, a concept referred to as heritability. The heritability of a trait has a major impact on the methods chosen for population improvement, inbreeding, and selection. Selection for single plants is more efficient when the heritability is high. The extent to which replicated testing is required for selection depends on the heritability of the trait.

The question of whether a trait variation is a result of genetic or environmental variation is meaningless in practice. Genes cannot cause a trait to develop unless the organism is growing in an appropriate environment, and, conversely, no amount of manipulation will cause a phenotype to

develop unless the necessary gene or genes are present. Nevertheless, the variability observed in some traits might result primarily from difference in the numbers and the magnitude of the effect of different genes, but that variability in other traits might stem primarily from the differences in the environments to which various individuals have been exposed. It is therefore essential to identify reliable measures to determine the relative importance of not only the numbers and magnitude of the effects of the genes involved, but also of the effects of different environments on the expression of phenotypic traits (Allard, 1999).

Heritability is defined as the ratio of genetic variance to phenotypic variance:

$$h^2 = \frac{\sigma_G^2}{\sigma_P^2} = \frac{\sigma_G^2}{\sigma_G^2 + \sigma_E^2}$$

where σ_P^2 is phenotypic variance, which has two components, genetic variance σ_G^2 and environment variance σ_E^2 . σ_E^2 can be estimated by the phenotypic variances of non-segregating populations such as inbred lines and F_1 s because individuals in such a population have the same genotypes and thus, phenotypic variation in these populations can be attributed to environmental factors. σ_G^2 can be estimated using segregating populations such as F_2 and backcrosses where variance components can be obtained theoretically.

1.5.6 Response to selection

Genetic variation forms the basis for selection in plant breeding. Selection results in the differential reproduction of genotypes in a population so that gene frequencies change and, with them, genotypic and phenotypic values (mean and variance) of the trait being selected. Response to selection, or advance in one generation of selection, is measured by the difference between the selected population and their offspring population, which is denoted as R . Response to selection has been referred to by several different names, including genetic progress, genetic advance,

genetic gain, and predicted progress or gain, and has been denoted as R , GS , G and ΔG .

Starting with a parental population of mean, μ , a subset of individuals is selected. The selected individuals have a mean \bar{x} , while the offspring of the selected population has a mean \bar{y} . The difference between the selected population and the original population is defined as the selection differential, and denoted by S , i.e.

$$S = \bar{x} - \mu$$

The response to selection, R , can be written as

$$R = \bar{y} - \mu$$

The relationship between S and R is determined by heritability,

$$R = h^2 S$$

How much of the selection differential is realized in the offspring population depends on the heritability of the trait. The heritability, h^2 , in the formula can be either h_N^2 or h_B^2 (depending on whether the offspring are produced by sexual or asexual reproduction, respectively). From the above formula, $\bar{y} = \mu + h^2 S$.

The population mean of the offspring derived from the selected individuals is equal to the parental population mean plus the response to selection (Fig. 1.1). When $h^2 = 1$, the selection differential will be fully realized in the offspring population so that its mean will deviate from the parental population by S . When $h^2 = 0$, the selection differential cannot be realized so the offspring population mean will regress to its parental population. When $0 < h^2 < 1$, the selection differential is partially realized so that the mean of the offspring population will deviate from the parental population by $h^2 S$. It is very useful to predict the response before selection is undertaken and details of the mathematical derivation of these predictions together with the various complications encountered can be found in Empig *et al.* (1972), Hallauer and Miranda (1988) and Nyquist (1991).

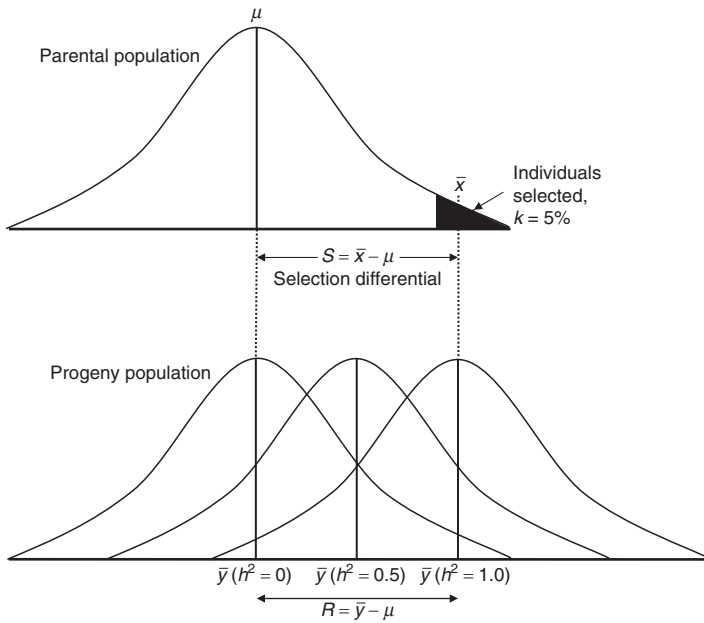


Fig. 1.1. Distribution of parental and progeny populations with a selection intensity of 5%. Because the phenotypic values of the selected plants include both a genetic and an environmental component, the progeny means depend on the heritability of the trait selected.

1.5.7 Selection index and selection for multiple traits

In most plant breeding programmes, there is a need to improve more than one trait at a time. For example, a high-yielding cultivar susceptible to a prevalent disease would be of little use to a grower. Recognition that improvement of one trait may cause improvement or deterioration in associated traits serves to emphasize the need for the simultaneous consideration of all traits which are important in a crop species. Three selection methods, which are recognized as appropriate for the simultaneous improvement of two or more traits in a breeding programme, are index selection, independent culling, and tandem selection. Independent culling requires the establishment of minimum levels of merit for each trait. An individual with a phenotype value below the critical culling level for any trait will be removed from the population. That is, only individuals meeting requirements for all traits will be selected.

With tandem selection, one trait is selected until it is improved to a satisfactory level or a critical phenotypic value. Then, in the next generation or programme, selection for a second trait is carried out within the population selected for the first trait, and so on for the third and subsequent traits. A selection index is a single score which reflects the merits and demerits of all target traits. Selection among individuals is based on the relative values of the index scores. Selection indices provide one method for improving multiple traits in a breeding programme. The use of a selection index in plant breeding was originally proposed by Smith (1936) who acknowledged critical input from Fisher (1936). Subsequently, methods of developing selection indices were modified, subjected to critical evaluation, and compared to other methods of multiple trait selection.

It is generally recognized that a selection index is a linear function of observable phenotypic values of different traits. There are a number of forms of the equations avail-

able from index selection for multiple traits in grain. To construct a selection index, the observed value of each trait is weighted by an index coefficient,

$$I = b_1x_1 + b_2x_2 + \dots + b_nx_n$$

where I is an index of merit of an individual, x_i represents the observed phenotypic value of the i th trait, and $b_1 \dots b_n$ are weights assigned to phenotypic trait measurements represented as $x_1 \dots x_n$. The \mathbf{b} values are the products of the inverse of the phenotypic variance–covariance matrix, genotypic variance–covariance matrix, and a vector of economic weights. A number of variations of this index, most changing the manner of computing the \mathbf{b} values, have been developed. These include the base index of Williams (1962), the desired gain index of Pesek and Baker (1969), and retrospective indexes proposed by Johnson *et al.* (1988) and Bernardo (1991). The emphasis in the retrospective index developments is on quantifying the knowledge experienced breeders have obtained. Baker (1986) summarized all select indexes in plant breeding developed before that time.

1.5.8 Combining ability

Combining ability is a very important concept in plant breeding and it can be used to compare and investigate how two inbred lines can be combined together to produce a productive hybrid or to breed new inbred lines. Selection and development of parental lines or inbreds with strong combining ability is one of the most important breeding objectives, no matter whether the goal is to create a hybrid with strong vigour or develop a pure-line cultivar with improved characteristics compared to their parental lines. In maize breeding, Sprague and Tatum (1942) partitioned the genetic variability among crosses into effects due to primarily either additive or non-additive effects, which correspond to two categories of combining ability, general combining ability (GCA) and special combining ability (SCA). The relative importance of GCA and SCA depends

on the extent of previous testing of the parents included in the crosses. Although these concepts were developed for breeding maize, an open-pollinated crop, they are generally applicable to self-pollinated crops.

The GCA for an inbred line or a cultivar can be evaluated by the average performance of yield or other economic traits in a set of hybrid combinations. The SCA for a cross combination can be evaluated by the deviation in its performance from the value expected from the GCA of its two parental lines. If the crosses among a set of inbred lines are made in such a way that each line is crossed with several other lines in a systematic manner, the total variation among crosses can be partitioned into two components ascribable to GCA and SCA. The mean performance of a cross (\bar{x}_{AB}) between two inbred lines A and B can be represented as

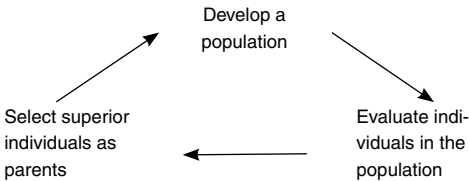
$$\bar{x}_{AB} = GCA_A + GCA_B + SCA_{AB}$$

The GCA_A and GCA_B are the GCA of the parents A and B, respectively, and the cross of A \times B is expected to have a performance equal to the sum ($GCA_A + GCA_B$) of the GCA of their parents. The actual performance of the cross, however, may be different from the expectation by an amount equivalent to the SCA. Sprague and Tatum (1942) interpreted these combining abilities in terms of type of gene action. The differences due to GCA of lines are the results of additive genetic variance and additive by additive interaction whereas SCA is a reflection of non-additive genetic variances.

1.5.9 Recurrent selection

Recurrent selection can be broadly defined as the systematic selection of desirable individuals from a population followed by recombination of the selected individuals to form a new population. The basic feature of recurrent selection methods is that they are procedures conducted in a repetitive manner, or recycling, including development of a base population with which to begin selection, evaluation of individuals from

the population, and selection of superior individuals as parents that can be crossed to produce a new population for the next cycle of selection, as shown below:



A cycle of selection is completed each time a new population is formed. The initial population that is developed for a recurrent selection programme is referred to as the base, or cycle 0, population. The population formed after one cycle of selection is called the cycle 1 population; the cycle 2 population is developed from the second cycle of selection, and so on.

Recurrent selection procedures are conducted for primarily quantitatively inherited traits. The objective of recurrent selection is to improve the mean performance of a population of plants by increasing the frequency of favourable alleles in a consistent manner in order to enhance the value of the population and to maintain the genetic variability present in the population as effectively as possible. In addition, separation of the genetic and environmental effects is an important facet of effective recurrent selection methods. The improved populations can be used as a cultivar *per se*, as parents of a cultivar-cross hybrid and as a source of superior individuals that can be used as inbred lines, pure-line cultivars, clonal cultivars, or parents of a synthetic line. Successful recurrent selection results in an improved population that is superior to the original population in mean performance and in the performance of the best individuals within it. Ideally, the population will be improved without its genetic variability being significantly reduced so that additional selection and improvement can occur in the future. Recurrent selection is complementary to inbred development procedures; in fact the concept of recurrent selection was developed, particularly

for outcrossing crops, to rectify limitations in inbred development by continuous selfing that rapidly leads to inbreeding and allele fixation and thus inadequate opportunity for selection. There are two ways by which recurrent selection address this limitation in inbred development (Bernardo, 2002). First, recurrent selection increases the frequency of favourable alleles in the population by repeated cycles of selection. Secondly, recurrent selection maintains the degree of genetic variation in the population to allow sustained progress from subsequent cycles of selection. Genetic variation is maintained by recombining a sufficiently large number of individuals to reduce random fluctuations in allele frequencies, i.e. genetic drift.

Since the late 1950s, extensive research has been conducted to determine the relative importance of different genetic effects on the inheritance of quantitative traits for most cultivated plant species. As indicated by Hallauer (2007), quantitative genetic research has provided extensive information to assist plant breeders in developing breeding and selection strategies. Directly and/or indirectly, the principles for the inheritance of quantitative traits are pervasive in developing superior cultivars to meet the worldwide food, feed, fuel and fibre demands. The principles of quantitative genetics will have continued importance in the future.

1.6 The Green Revolution and the Challenges Ahead

The application of science and technology to crop production in the second half of the 20th century resulted in significant yield improvements for rice, wheat and maize in the developed countries, and the final result of these efforts was the Green Revolution which led to a new type of agriculture – high-input or chemical-genetic agriculture – which replaced the more traditional system. Countries involved in the ‘Green Revolution’, a term coined by Borlaug (1972), included Japan, Mexico, India and China among others.

By production and acreage yardsticks, agriculture has been very successful. The application of scientific knowledge to agriculture has resulted in greatly increased yields per unit land area for many of our important crops as exemplified by the 92% increase in cereal production in the developing world between 1961 and 1990. The sharp increase in human populations has been paralleled by the increase in food supply. However, yield growth rates are stagnating in some areas and, in a few cases, falling. A slowdown in the rate of yield increase of major cereals raises concern because increased yields are expected to be the source of increased food production in the future (Reeves *et al.*, 1999). On the other hand, increased national wealth resulting from economic development is not necessarily correlated with a decrease in the rate of population growth. Widespread hunger persists in a world that produces enough food.

There are many reasons for being concerned about meeting future food demands (Khush, 1999; Swaminathan, 2007). Expansion of the planet's population creates an increased demand for food and income. Other issues such as the cost of food, which may represent 60% of income in the developing world, the 800 million people who are food insecure, the 200 million children who are malnourished, and the continuing decline of available land for farming and water to irrigate crops, all indicate the need to use all the technologies available to increase productivity, assuming they can be employed in harmony with the environment. Plant breeding has generally accounted for one-half of the increases in productivity of the major crops and the future will continue to depend on advances in plant breeding. The increase in productivity has meant that large areas of land can be saved as wildlife habitats or used for purposes other than agriculture. As the availability of land and water is decreasing and populations are increasing in size, the 50% increase in food production predicted to be required over the next 25 years, poses an obvious challenge.

The danger of population growth overtaking food supplies was predicted by Malthus in 1817. The dire predictions of

Malthus were forestalled, at least temporarily, by the extensive cultivation of new land and by the development of a modern agricultural science which enabled food crops to be produced at far higher yields than Malthus could have ever anticipated. However, the production of food has still not been optimized in all areas of the world.

Weather and climate profoundly affect crop production and natural events can disrupt normal climate cycles and affect agriculture. In addition, human-induced climatic change is set to accelerate during this century and this will also impact on crop production. Much of the arable land has been used for industrial purposes and land-use patterns indicate an increase in intensive farming which, however, must be sustainable.

Agricultural products are affected by abiotic and biotic stresses and one of the major challenges to the future of plant breeding is the development of cultivars and hybrids with multiple resistances or tolerances to these stresses.

The security of the food supply for an increasing world population largely depends on the availability of water for agriculture. Increasing the efficiency of water use for our major crop species is an important target in agricultural research, particularly in light of the increasing competition for limited supplies of fresh water in many parts of the world.

There are four prerequisites for greater productivity (Poehlman and Quick, 1983): (i) an improved farming system; (ii) instruction of farmers; (iii) optimization of the supply of water and fertilizers; and (iv) availability of markets. To increase crop productivity planting high-yielding cultivars must be combined with improved practices of irrigation, fertilization and pest control. Maximum crop yield will only be achieved if the improved crop cultivar receives and responds to the optimum combination of water, fertilizer and cultural practices.

1.7 Objectives of Plant Breeding

The aim of plant breeders is to reassemble desirable inherited traits to produce crops

with improved characteristics. Thus far, plant breeders have mainly been concerned with bringing about a continuous improvement in the productivity of that part of the plant which is of economic importance, the stability of production through in-built resistance to pests and diseases and nutritive and organoleptic or other desired quality characters.

Many parameters and selection criteria should be included as breeding objectives. According to Sinha and Swaminathan (1984) and other sources, the major objectives of plant breeders can be summarized by the following list:

1. High primary productivity and efficient final production for each unit of cultivation and solar energy invested: to ensure that all the light that falls on a field is intercepted by leaves and that photosynthesis itself is as efficient as possible. Greater efficiency in photosynthesis could perhaps be achieved by reducing photorespiration.
2. High crop yield: plants must be selected which invest a large proportion of their total primary productivity into those areas which are commercially desirable, e.g. seeds, roots, leaves or stems.
3. Desirable nutritional value, organoleptic properties and processing qualities: the proportion of essential amino acids and the total protein in cereal grains, for example, should be increased to improve their nutritional quality.
4. Biofortifying crops with essential mineral elements that are frequently lacking in the human diet such as Fe and Zn, vitamins and amino acids (Welch and Graham, 2004; White and Broadley, 2005; Bekaert *et al.*, 2008; Mayer *et al.*, 2008; Ufaz and Galili, 2008; Naqvi *et al.*, 2009; Xu *et al.*, 2009a).
5. Modifying crop plants to generate plant-derived pharmaceuticals to supply low-cost drugs and vaccines to the developing world (Ma *et al.*, 2005).
6. Adaptation to cropping systems: including breeding for contrasting cropping, inter-cropping, and sustainable cropping systems (Brummer, 2006).
7. More extensive and efficient nitrogen fixation: breeding cereals that encourage the

growth of increased numbers of nitrogen-fixing microorganisms around their roots to reduce the need for nitrogen fertilizer.

8. More efficient use of water whether there is a plentiful supply or dearth of water.
9. Stability of crop production by resilience to weather fluctuations, resistance to the multiple alliance of weeds, pests and pathogens, and tolerance to various abiotic stresses such as heat, cold, drought, wind, and soil salinity, acidity or aluminium toxicity.
10. Insensitivity to photoperiod and temperature: selection of crop cultivars that are insensitive to photoperiod or temperature and characterized by a high per-day biomass production would allow the development of contingency cropping patterns to suit different weather probabilities.
11. Plant architecture and adaptability to mechanized farming: the number and positioning of the leaves, branching pattern of the stem, the height of the plant, and the positioning of the organs to be harvested are all important to crop production and often determine how well plants can be harvested mechanically.
12. Elimination of toxic compounds.
13. Identification and improvement of hardy plants suitable for sources of biomass and renewable energy.
14. Multiple uses of a single crop.
15. Environmentally-friendly and stable across environments.

In conclusion, plant breeding has many breeding objectives and each of the objectives can be addressed in a specific breeding programme. A successful breeding programme consists of a series of activities as Burton (1981) summarized in six words: variate, isolate, evaluate, intermate, multiply and disseminate.

1.8 Molecular Breeding

By 2025, the global population will exceed seven billion. In the interim per-capita availability of arable land and irrigation water will decrease from year to year as biotic and abiotic stresses increase. Food security, best defined as economic, physical

and social access to a balanced diet and safe drinking water will be threatened, with a holistic approach to nutritional and non-nutritional factors needed to achieve success in the eradication of hunger. Science and technology can play a very important role in stimulating and sustaining an Evergreen Revolution leading to long-term increases in productivity without any associated ecological harm (Borlaug, 2001; Swaminathan, 2007). The objectives of the plant breeder can be realized through conventional breeding integrated with various biotechnology developments (e.g. Damude and Kinney, 2008; Xu *et al.*, 2009c).

Plant breeding can be defined as an evolving science and technology (Fig. 1.2). It has gradually been evolving from art to science over the last 10,000 years, starting as an ancient art to the present molecular design-based science. With the development of molecular tools which will be discussed further in Chapters 2 and 3, plant

breeding is becoming quicker, easier, more effective and more efficient (Phillips, 2006). Plant breeders will be well equipped with innovative approaches to identify and/or create genetic variation, to define the genetic feature of the genes related to the variation (position, function and relationship with other genes and environments), to understand the structure of breeding populations, to recombine novel alleles or allele combinations into specific cultivars or hybrids, and to select the best individuals with desirable genetic features which enable them to adapt to a wide range of environments.

Sequencing data for many plants is now readily available and the GenBank database is doubling every 15 months. Over 20 plant species including many important crops are in the process of being sequenced (Phillips, 2008). The next challenge is to determine the function of every gene and eventually how genes interact to form the basis of complex traits. Fortunately, DNA chips and other technologies are being developed to study the expression of multiple or even all genes simultaneously. High throughput robotics and bioinformatics tools will play an essential role in this endeavour.

New information about our crop species is expanding our capabilities to use molecular genetics. For example, we did not previously realize how similar broadly related species are in terms of their gene content and gene order. Since these species cannot usually be crossed, there was no means of assessing their relatedness. With the advent of DNA-based molecular markers, the extensive genetic mapping of chromosomes became readily possible for a variety of species. We learned that the genomes were highly similar and that this similarity allowed the prediction of gene locations among species. For example, rice has become the model or reference species for the cereals as many of the gene sequences on the rice chromosomes are shared with other cereals such as maize, sorghum, sugarcane, millet, oats, wheat and barley (Xu *et al.*, 2005). Knowing the complete DNA sequence of a model or reference genome allows genes/traits from this

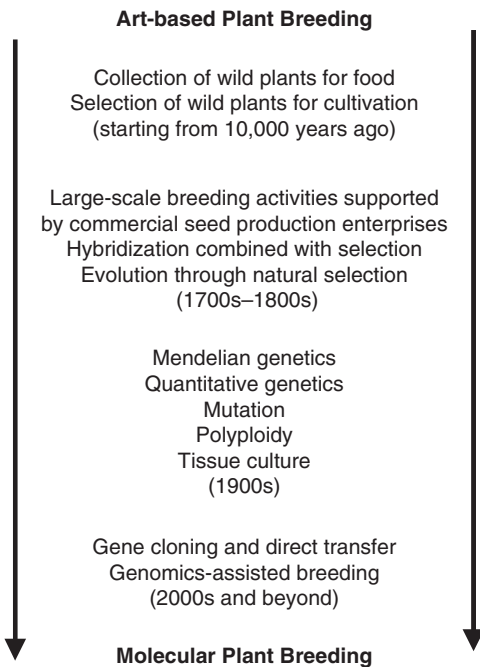


Fig. 1.2. The steps of evolution of 'plant breeding'. With the availability of more sophisticated tools, the art of plant breeding became science-based technology, molecular plant breeding.

model to be tracked to other genomes. We have come to realize that the differences between species of plants are not due to novel genes, but to novel allelic specifications and interactions.

Since many fundamental aspects of current plant breeding procedures are not well understood, further data relating to the genetics of crop species may help to shed light on the genetic gains obtained from plant breeding. For example, in successful plant breeding programmes, the genetic base often becomes narrower rather than broader. 'Elite by elite' crosses may be the rule in these programmes. Molecular genetic markers have been widely employed to identify cryptic and novel genetic variation among cultivars and related species and used to increase the efficiency of selection for agronomic traits and the pyramid of genes from different genetic backgrounds.

Long-term selection programmes would be expected to lead to genetic fixation, however this has not been found to be the case so far and variation is still observed. Several mechanisms for *de novo* variation have been described, including intragenic recombination, unequal crossing over among repeated elements, transposon activity, DNA methylation, and paramutation. Another important feature in plant breeding whose molecular basis is not understood is heterosis although it is used as the basis for many seed-producing industries. Genomics and particularly transcriptomics are now being used to identify the heterotic genes responsible for increasing crop yields. Comprehensive quantitative trait locus-based phenotyping (phenomics) combined with genome-wide expression analysis, should help to identify the loci controlling heterotic phenotypes and thus

improve the understanding of the role of heterosis in evolution and the domestication of crop plants (Lippman and Zamir, 2007), and finally to make it possible to predict hybrid performance.

Messenger RNA transcript profiling is an obvious candidate for functional genomic application to plant breeding. Although direct selection at the gene transcript level using microarray or real-time PCR may be a long-term goal, other genomic tools can be used to achieve shorter term goals with more practical applications (Crosbie *et al.*, 2006). Genetic modification of crops today involves the interfacing of molecular biology, cell and tissue culture, and genetics/breeding. The transfer of genes by cellular and molecular means will increase the available gene pool and lead to second generation biotechnology plant products such as those with a modified oil, protein, vitamin, or micronutrient content or those that have been engineered to produce compounds that can be used as vaccines or anti-carcinogens.

While all these new innovations have been useful, practical plant breeding continues to be based on hybridization and selection with little change in the basic procedures. A more complete understanding of the mechanisms by which genetic and environmental variation modify yield and composition is needed so that specific quantitative and qualitative targets can be identified. To achieve this aim, the expertise of plant genomics (including various omics), physiology and agronomy, as well as plant modelling techniques must be combined (Wollenweber *et al.*, 2005) and many logistic and genetic constraints also need to be resolved (Xu and Crouch, 2008).