Molecular Plant Breeding

_____

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 pathbreaking 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

www.cabi.org

Contents

The colour plate section can be found following p. XXX

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 testcrosses 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 geno-typing, selective and pooled DNA analysis, and chip-based large-scale germplasm evalua-tion, 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 sup-port 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 encour-agement 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 indispensible 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 addi-tion, 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, par-ticularly 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 laure-ate 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 indispensible, any remaining errors remain my sole responsibility. The names and affiliations of the reviewers (alphabetically) are:

Raman Babu (Chapters 7 and 9), CIMMYT, Mexico
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 Caroll (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

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

D<small>R</small> N<small>ORMAN</small> E. B<small>ORLAUG</small>

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 × 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 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 refer-
ences. Even many of the tables are packed with information and references. A good rep-
resentation of international and domestic breeding is reflected through many examples.
The importance of G × 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, includ-
ing its requirements and advantages, along with the multitude of quantitative trait locus
(QTL) analysis methods. Transformation technologies leading to the extensive use of trans-
genic 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.