Kasi Azhakanandam · Aron Silverstone
Henry Daniell · Michael R. Davey *Editors*

# Recent Advancements in Gene Expression and Enabling Technologies in Crop Plants

Springer

# Recent Advancements in Gene Expression and Enabling Technologies in Crop Plants

Kasi Azhakanandam • Aron Silverstone
Henry Daniell • Michael R. Davey
Editors

Foreword by Mary-Dell Chilton, PhD, Syngenta Biotechnology, Inc.

# Recent Advancements in Gene Expression and Enabling Technologies in Crop Plants

Springer

*Editors*
Dr. Kasi Azhakanandam
Syngenta Biotechnology Inc.
3054 Cornwallis Rd
Research Triangle Park, NC 27709, USA

Dr. Aron Silverstone
Syngenta Biotechnology Inc.
3054 Cornwallis Rd
Research Triangle Park, NC 27709, USA

Dr. Henry Daniell
University of Pennsylvania
School of Medicine
240 South 40th Street
Philadelphia PA 19104, USA

Dr. Michael R. Davey
University of Nottingham
School of Biosciences
Loughborough LE12 5RD,
United Kingdom

ISBN 978-1-4939-2201-7 ISBN 978-1-4939-2202-4 (eBook)
DOI 10.1007/978-1-4939-2202-4

Library of Congress Control Number: 2015931208

Springer New York Heidelberg Dordrecht London


Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

*This book is dedicated to people who have died of starvation*

# Foreword

In the following pages, some of the world's most renowned researchers take a look at the state of the art and science of introducing novel genes into plant cells and plants. The various chapters deal with a wide range of products, from genetically modified seeds and plants to commodities made by such transgenic plants, including enzymes or vaccines. One important consideration is where and how the new genes are integrated into the host plant. The donor DNA may be inserted into the plant chromosome at random places or targeted to a specific location, by recombination or by employing site-specific nucleases. A future targeting technology may employ a minichromosome, an artificial vector assembled from parts of a normal chromosome (Chapter 13). A minichromosome is actually a megavector, which will be especially attractive for the introduction of a block of genes, for example those encoding an entire biochemical pathway for production of a valuable metabolite. At the other extreme of size, free replicons such as a (modified) plant DNA viral genome might be the most useful vector for some traits. Whatever the form and location of the vector, the DNA construct itself must mimic the plant's strategy for dictating quantity, timing, and location for the encoded protein to be made. In Chapter 2, Dr. Nuccio et al. provides a wellspring of information on plant trait gene design and approaches that have worked.

This book addresses many of these issues and will be useful to the plant genetic engineer, whether student or accomplished professional. I found new ideas and information in each chapter. I skipped around as my curiosity led me, and was excited to discover how many different types of challenges plant genetic engineering has posed, and how many creative solutions have been devised. I found the book quite readable for a technical work, with a refreshing honesty about the sometimes halting progress of scientific research.

While we are on the topic of honesty, I must confess to a motive underlying my writing of this foreword. I wanted to reach you, readers of this book, with one more message. Let me begin with a brief story: When my sons were quite young, we subscribed to a journal about the environment called Ranger Rick. One month it carried a story about insect galls, describing how the mother insect uses chemical signals to stimulate growth of the plant cells into a gall at the site where she deposits her eggs. When the insect larvae hatch, the gall serves her babies as a nice source

of food. By coincidence, my colleagues and I at the University of Washington had recently begun a research project on crown gall tumors, induced by *Agrobacterium* in plants. The insect gall story, aimed at children, made me think. Crown galls were known to produce new metabolites—octopine or nopaline, depending on the *Agrobacterium* strain that incited the gall. Could octopine and nopaline be baby food for *Agrobacterium*? When it was my turn to talk at our weekly research group meeting, I reported on the Ranger Rick article, and proposed that *Agrobacterium*, like the mother insect, might be producing the crown gall as a means of feeding its progeny. I can well recall the laughter and ridicule that ensued. The concept was named the Ranger Rick Hypothesis, and I was teased mercilessly about it for many months, until our competitors in France, Australia, and Belgium announced this very same concept as the "rationale of the gall" (in three languages). It became a respectable idea, eventually supported by increasing amounts of evidence.

There are several potential morals to this story, and I invite you to consider any of them that interest you. For me, the moral is that *Agrobacterium* truly was a genetic engineer before my colleagues and I ever thought of the possibility. The process that we now use to make genetically modified plants, the topic of this volume, is a natural one at core, invented first by a microbe and only refined by *Homo sapiens*. *Agrobacterium* worked out a way to transfer its desirable genes to the host plant cells, genes that caused abundant growth (the gall) and delicious (we suppose) meals for future generations. I hope that you who take a serious interest in the contents of this book will take equally seriously the need to inform the public that gene transfer is a natural and normal process. The products made by genetic modification of plants are more precise and predictable than those made by plant breeding, especially plant breeders use of wide crosses for introduction of new traits from wild relatives of crop plants.

By the year 2050, the world's population is expected to grow from its current 7 billion to 9 billion, a 30 % increase in the number of people. A distressing number of our present population is already hungry, even starving. Biotechnology alone cannot solve this problem, but it certainly has the potential to be an important part of the solution. Unless people accept foods produced through biotechnology, progress in food security will be slow. I believe that the principal risk of genetically modified crops is public perception, not the safety of the products themselves, which are thoroughly tested. If you share my view, I hope that you will not keep it a secret. Seek opportunities to speak to school children, garden clubs, church groups, or anyone who will listen. Tell them that there is nothing unnatural about gene transfer to plants by *Agrobacterium*. I believe that the success of genetically modified plant products depends upon the efforts of scientists like you and me to communicate to the public the safety and sanity of biotech plants.

Mary-Dell Chilton
Research Triangle Park, NC, USA

# Preface

When we decided to edit a book on gene expression in plants, we realized that the most valuable contribution would be to combine reports from the biotech industry, and academic and research institutes that would focus on gene expression studies with economically important crops and related enabling technologies. Such a volume should be useful for students and researchers at all levels. Tremendous progress has been made in introducing novel genes and traits into plant genomes since the first creation of transgenic plants 30 years ago, and the first commercialization of genetically modified maize in 1996. Consequently, cultivation of biotech crops with useful traits has increased more than 100-fold from 1.7 million ha in 1996 to over 175 million ha globally in 2013. This achievement has been made possible by continued advances in understanding the basic molecular biology of regulatory sequences to modulate gene expression, enhancement of protein synthesis, and new technologies for transformation of crop plants.

In this book, authors who are experts in their fields describe current advances on commercial crops and key enabling technologies that will underpin future advances in biotechnology. They discuss state-of-the-art discoveries as well as future challenges. This book has three parts that encompass knowledge on genetically modified (GM) food crops that are currently used by consumers, those that are anticipated to reach the market place in the near future and enabling technologies that will facilitate the development of next generation GM crops. Part I focuses only on genetically modified maize and soybean (three chapters each), while Part II discusses the GM food crops rice, wheat, sorghum, vegetables, and sugarcane. Part III covers exciting recent developments in several novel enabling technologies, including gene targeting, minichromosomes, and *in planta* transient expression systems.

In the first chapter, Lu et al. provide a detailed overview of fascinating aspects of maize protein expression. This chapter reviews current understanding and future perspectives on key aspects that affect recombinant protein expression in this crop. These authors have summarized various factors that control gene expression, including promoters, subcellular targeting, and different regulatory elements, including introns, 5′ and 3′ untranslated regions (UTRs), spacers and insulators. In Chapter 2, Nuccio et al. present a detailed understanding on transgene design with

plant trait gene expression cassette design. The authors characterized several native maize promoters, and used the structure of these promoters to design constructs that deliver high-level gene expression/accumulation in maize. Chapter 3 is also devoted to maize. Howard and Hood review different strategies to maximize recombinant protein expression in kernels and discuss the characteristics that make maize a popular choice for recombinant protein production. These authors also assess various factors that contribute to high-level expression of heterologous proteins, together with examples of successful approaches.

In Chapter 4, Ramachandra et al. outline the breeding and biotech approaches to improve yield in soybean. The use of transgenes to complement traditional breeding through "gene stacking" will be important to further increase soybean yield and overcome biotic and abiotic stresses. One of the most successful innovations of biotech that had a major impact on farming is the introduction of herbicide tolerance in plants. Consequently, Huang et al. in Chapter 5 discuss the details of genes/traits, which have been exploited to make plants tolerant to herbicides. Tolerance to broad-spectrum herbicides makes weed control more efficient, which greatly assists the farming community. However, the increase of resistant weeds is creating new challenges for the biotech industry. In order to address this concern, authors discuss the use of trait stacking to manage hard-to-control and resistant weeds. They also describe the development of a new herbicide trait system for dicamba tolerance. Herman and Schmidt (Chapter 6) have focused on modification of soybean seeds for their use as protein bioreactors. Soybean seeds have high protein content and are used as a protein source in animal feed. These authors present the success and limitations of different approaches to produce heterologous proteins in seeds. They describe a protein rebalancing approach that increases expression of a model protein (green fluorescent protein) from 1.5 to 8% of the total soy seed protein.

Significant progress has been made in cereal biotechnology. Many traits have been engineered into the rice genome to protect against biotic and abiotic stress or to improve grain and nutritional quality. In Chapter 7, Nandi and Khush review strategies to increase heterologous protein expression in rice grains. These authors summarize key factors responsible for controlling expression, including regulatory sequences, translational efficiency, posttranslational modifications, and compartmentalization of foreign proteins. They also discuss strategies to down-regulate endogenous protein expression in order to boost heterologous protein accumulation. In Chapter 8, Jones summarizes current advances in wheat biotechnology, particularly methods adopted for wheat transformation. He also summarizes progress in enhancing tolerance to biotic stress and to improve quality traits such as those for breadmaking. Biotechnology plays an important role in meeting the global demand for wheat, which is anticipated to increase more than 50% by 2050. Recent advances in sorghum biotechnology are outlined by Do and Zhang (Chapter 9), with the challenges related to the tissue culture and transformation of this crop. The biotech approaches for insect pest management in vegetable crops are featured in Chapter 10 by Sreevathsa et al. The Bt protein was tested in vegetable crops to control insect pests, with discussion of different promoters used to achieve high-level expression, conferring greater resistance against target pests. The authors also discuss other

strategies, including the use of inhibitors of insect digestive enzymes, or engineering secondary metabolism of volatile communication compounds to combat pests. In recent years, there has been more biotechnology research directed to sugarcane not only for sugar production, but also for its use as biofuels. In Chapter 11, Wu discusses techniques for boosting sugar content through genetic engineering, including the expression of novel sugars.

As the opportunities of biotechnology increase, more complex tools are needed to deliver desired targets. In addition, newly acquired plant genomes' sequences provide a wealth of data that can be exploited. A key to understanding the functions of specific genes is the ability to rapidly overexpress or turn them off. Part III explores these enabling technologies. In Chapter 12, Petolino et al. describe gene targeting in plants by using Zinc Finger Nucleases (ZFNs). These authors explain how ZFNs are exploited for target mutagenesis, gene deletion and site-specific transgene integration. They also discuss other nuclease technologies, such as TALENs, meganucleases, and CRISPRs, as well as the relative advantages and limitations of these procedures. Minichromosomes combine native chromosome structural elements, like centromeres, along with transgenes for introduction into crop plants. Birchler (Chapter 13) reviews the status of "Minichromosome" technology in plants. One of the key advantages of artificial chromosomes is that multiple genes of interest could be stacked into plant genomes as a single entity without linkage to other chromosomes. Birchler also discusses both the challenges and opportunities associated with this novel technology.

Studies on gene function(s) utilizing stable transformation is time consuming and expensive. However, *in planta* transient sytems, using viral vectors developed in recent years, make it possible to study gene function by knocking down target genes or overexpression of genes of interest, although this approach has been limited to small genes (< 1.5 kb) in crop plants. There are efforts to build viral vectors, which can accommodate larger inserts. In Chapter 14, Lee et al. review various *in planta* transient expression systems for both RNAi-mediated down-regulation and over expression of target genes in monocotyledonous plants. These authors discuss the increasing use of transient *in planta* expression systems, such as virus-induced gene silencing (VIGS), virus-mediated overexpression (VOX), and cell culture-based transient approaches, as well as the advantages and disadvantages associated with each transient system. Chapter 15 by Whitham et al. presents recombinant plant viruses that are capable of carrying genetic payloads of whole genes or gene fragments that provide convenient platforms as vectors for transient gene expression and silencing in soybean. These authors focus on seven viral vector systems that have been used in this leguminous crop for VOX and/or VIGS applications. They discuss key features of the viral genomes, and future prospects to exploit viral vectors for soybean improvement.

In summary, this volume highlights a wide range of research tools, current methods, and future enabling technologies to improve crop plants to meet the ever increasing global demand for food, feed, and fuel. The editors believe that this book will be an excellent reference source for the scientific community interested in extending model plant systems into valuable applications in crop plants. We sincerely

thank all the authors for their hard work and valuable contributions, and colleagues at Springer for the invitation to edit this unique contribution to the literature for the scientific community.

Kasi Azhakanandam
Aron Silverstone
Henry Daniell
Michael R. Davey

# Abstract

In the past two decades, agricultural biotechnology has had a major impact on farming, with genetically modified (GM) crops grown on more than 175 million ha globally. Although plant biotechnology has exploited model systems to gain fundamental knowledge, parallel research on field-grown plants has facilitated the development of GM crops that are used by consumers today. Biotechnology has also helped to create a rich pipeline of future products. This volume focuses on the innovations in both applied and basic research that are advancing our ability to deliver more complex multigene traits into plants. Although much of the work to date has been done on corn and soybean, other plants that are the subject of active transgenic development include rice, wheat, sorghum, sugarcane, and vegetable crops. There is a progression from the use of constitutive promoters and single traits to gene stacking, the design of transgene cassettes to more resemble native genes, the subcellular location of recombinant proteins, and manipulating storage tissues to achieve optimal performance. Herbicide tolerance and insect control have been and will continue to be highly desired traits. The future holds promise for novel modes of action to overcome current limitations. Targets for engineered recombinant proteins go beyond agronomic traits and focus on industrial or pharmaceutical uses, yield, and nutritional enhancement. Undoubtedly, future farming will advance from food/feed to industrial products, making crops more rewarding with value-added traits. Soon, even more sophisticated tools, including precision insertion or editing of genes and building novel chromosomes, will increase our ability to overcome current barriers in gene expression technology and facilitate rapid regulatory approval. The use of transient expression systems for crop plants will facilitate rapid evaluation of transgenes in crop plants. This book highlights a wide range of current research tools and enabling technologies to improve crop plants, with special emphasis on next generation approaches for engineering complex traits and value-added products that will revolutionize the future of agriculture to meet the ever increasing global demand for food, feed, fuel, and industrial products.

# Acknowledgements

The editors would like to thank Rongda Qu, Todd J. Jones, Michael Nuccio, Albert Lu, Bernard Vernooij, Jared Conville, Sandeep Kumar, Robert Thomas Gaeta, P. Dayanandan, Brian Hague, and Vidya Rajan for reviewing several of the chapters.

# Contents

# Contributors

**James A. Birchler** Division of Biological Sciences, University of Missouri, Columbia, MO, USA

**Xi Chen** Syngenta Biotechnology, Inc. East Cornwallis Road, Research Triangle Park, NC, USA

**R. V. Chowda-Reddy** Department of Plant Pathology and Microbiology, Iowa State University, Ames, IA, USA

**Mark Cigan** DuPont Pioneer, Johnston, IA, USA

**Jared Conville** Syngenta Biotechnology, Inc., Research Triangle Park, NC, USA

**Scott Diehn** DuPont Pioneer, Johnston, IA, USA

**Phat T. Do** Plant Transformation Core Facility, Division of Plant Sciences, 007A, Sears Plant Growth Facility, University of Missouri, Columbia, MO, USA

**Alan L. Eggenberger** Department of Plant Pathology and Microbiology, Iowa State University, Ames, IA, USA

**Christine Ellis** Monsanto Company, Chesterfield, MO, USA

**Kim E. Hammond-Kosack** Wheat Pathogenomics Team, Plant Biology and Crop Science Department, Rothamsted Research, Harpenden, UK

**Brian Hauge** Monsanto Company, Chesterfield, MO, USA

**Eliot M. Herman** School of Plant Sciences, BIO5 Institute Room No. 249, University of Arizona, Tucson, AZ, USA

**John H. Hill** Department of Plant Pathology and Microbiology, Iowa State University, Ames, IA, USA

**Elizabeth E. Hood** Biosciences Institute, Arkansas State University, State University, AR, USA

**John A. Howard** Cal Poly Technology Park, San Luis Obispo, CA, USA

**Jintai Huang** Monsanto Company, Chesterfield, MO, USA

**Jones Huw D.** Rothamsted Research, Harpenden, Hertfordshire, UK

**Kostya Kanyuka** Wheat Pathogenomics Team, Plant Biology and Crop Science Department, Rothamsted Research, Harpenden, UK

**Balasulojini Karunanandaa** Monsanto Company, Chesterfield, MO, USA

**Gurdev S. Khush** Global HealthShare® Initiative, Department of Plant Science, University of California, Davis, CA, USA

**P. Ananda Kumar** Institute of Biotechnology, ANGRAU, Hyderabad, India

**Wing-Sham Lee** Wheat Pathogenomics Team, Plant Biology and Crop Science Department, Rothamsted Research, Harpenden, UK

**Xiaomei Liu** Syngenta Biotechnology, Inc., Research Triangle Park, NC, USA

**Paul Loida** Monsanto Company, Chesterfield, MO, USA

**Albert Lu** DuPont Pioneer, Johnston, IA, USA

**Savitha Madappa** Monsanto Company, Bangalore, India

**Kathleen M. Martin** United States Department of Agriculture, Agricultural Research Service, Corn Insects and Crop Genetics Research Unit, Ames, IA, USA

**Somen Nandi** Global HealthShare® Initiative, Department of Molecular and Cellular Biology, University of California, Davis, CA, USA

**Michael Nuccio** Syngenta Biotechnology, Inc., Research Triangle Park, NC, USA

**Joseph F. Petolino** Dow AgroSciences, Indianapolis, IN, USA

**Jonathan Phillips** Monsanto Company, St. Louis, MO, USA

**Youlin Qi** Monsanto Company, Chesterfield, MO, USA

**Dhanalakshmi Ramachandra** Monsanto Company, Bangalore, India

**J. Pon Samuel** Dow AgroSciences, Indianapolis, IN, USA

**Lakshmi Sastry-Dent** Dow AgroSciences, Indianapolis, IN, USA

**Monica A. Schmidt** School of Plant Sciences, BIO5 Institute Room No. 303, University of Arizona, Tucson, AZ, USA

**Amolkumar U. Solanke** National Research Centre for Plant Biotechnology, New Delhi, India

**Rohini Sreevathsa** National Research Centre for Plant Biotechnology, New Delhi, India

**Marguerite J. Varagona** Biotechnology-Agronomic Traits, Monsanto Company, Chesterfield, MO, USA

**Steven A. Whitham** Department of Plant Pathology and Microbiology, Iowa State University, Ames, IA, USA

**Luguang Wu** School of Agriculture and Food Science, Faculty of Science, The University of Queensland, St. Lucia, Qld, Australia

**Chunquan Zhang** Department of Agriculture, Alcorn State University, Lorman, MS, USA

**Zhanyuan J. Zhang** Plant Transformation Core Facility, Division of Plant Sciences, 007A, Sears Plant Growth Facility, University of Missouri, Columbia, MO, USA

**Ailing Zhou** Syngenta Biotechnology, Inc., Research Triangle Park, NC, USA

# About the Editors

**Dr. Kasi Azhakanandam** earned his Bachelor, Master, and MPhil degrees in Biology from Madras Christian College, the University of Madras, India and a PhD in Plant Biotechnology from the University of Nottingham, UK. He worked as a Guest Lecturer at Madras Christian College for a short period before joining Mahyco, India, as a Deputy Chief Scientist/Principal Investigator, where he established a crop transformation laboratory. He led a team, which established transformation in commercial Indica rice, Indian cotton varieties, and six different vegetable crops, including Bt eggplant; these are waiting for approval for commercial cultivation in India while the Bt eggplant is approved for commercial cultivation in Bangladesh. He also successfully produced marker-free rice and vegetable crops. Following his postdoctoral work related to vaccine production for cervical cancer at North Carolina State University, Dr. Azhakanandam joined Syngenta Biotechnology, Inc., at Research Triangle Park, NC as a Staff Scientist III. He has worked on a range of projects to improve crops through genetic engineering, and currently leads a technical team for developing new traits for corn.

**Dr. Aron Silverstone** gained a Bachelor's degree in Biology from Harvard University, and a PhD in Plant Physiology from the University of California, Davis. He conducted his postdoctoral research at Duke University's Department of Botany, studying gibberellin biosynthesis and response. Following his postdoctoral work, Dr. Silverstone joined Syngenta Biotechnology, Inc., at Research Triangle Park, NC as a Staff Scientist I. He has worked on several projects in corn, soy and sugarcane to improve crops through genetic engineering. Dr. Silverstone is currently working on protecting plants from abiotic stresses.

**Dr. Henry Daniell** received his education in India, and is currently a Professor and Director of Translational Research at the University of Pennsylvania. He is a Fellow of the American Association for the Advancement of Science and a foreign member of the Italian National Academy of Sciences (14th American to be inducted in the past 230 years). He is the editor-in-chief of the *Plant Biotechnology Journal*, Oxford, UK. Dr. Daniell is the recipient of several awards, including the American Diabetes Association Award, Bayer Hemophilia Global Award, and Bill

and Melinda Gates Foundation Award, for his scientific contributions. He is recognized for pioneering chloroplast genetic engineering as a new platform to produce and deliver orally low-cost vaccines and biopharmaceuticals bioencapsulated in plant cells. His invention was ranked by Nature Biotechnology among the top ten inventions of the past decade and among Biomed Central's Hot 100 authors in the world. He has more than 150 published patents and over 200 scientific publications.

**Dr. Michael R. Davey** has a BSc Honours degree in Botany from University College, Swansea, Wales, and a PhD from the University of Leicester, UK. In 1970, he was appointed to a research position at the University of Nottingham, UK where he continued his work on plant ultrastructure and gene transfer techniques. He has published extensively on plant cell culture and genetic engineering, and holds an Honorary Lectureship in the School of Biosciences, University of Nottingham.

# Part I
# Corn and Soybean

# Chapter 1
# Maize Protein Expression

**Albert Lu, Scott Diehn and Mark Cigan**

## Introduction and Perspectives

Maize has been and will continue to be an important global food source with 857 million metric tons of corn produced in 2012–2013 for human and livestock consumption (USDA 2013). In addition to food and feed, industrial applications for maize extend into biofuel and starch production. Approximately 88 % of the maize acreage in the USA is transgenic, with insect resistance (IR), and/or herbicide resistance (HR) being the most prominent traits (Table 1.1). These traits improve yield and yield stability as a result of reducing stresses to the plant due to insect feeding or competition for essential nutrients by weeds. As a result of this success, companies involved in agricultural biotechnology, such as DuPont Pioneer, Monsanto, Syngenta, Bayer, and Dow AgroSciences, continue to perform research and develop new traits directed at maize crop improvement with the objective to increase grower's productivity and sustainably produce food to help feed a growing world population. In addition to productivity gains offered by transgenic traits, transgenic maize has been deployed as a cost-effective platform for expression of recombinant proteins on an agricultural scale (Table 1.2). The success of these applications is dependent on the ability to express effectively a single or multiple proteins in transgenic events.

Today's generation of transgenic maize events involves a routine process utilizing either particle bombardment- or *Agrobacterium*-based technologies. In either

A. Lu (✉)
DuPont Pioneer, 7300 NW 62nd Ave., 50131 Johnston, IA, USA
e-mail: Albert.L.Lu@cgr.dupont.com

S. Diehn · M. Cigan
DuPont Pioneer, 7300 NW 62nd Ave. Johnston, IA 50131, USA
e-mail: Scott.Diehn@Pioneer.com

M. Cigan
e-mail: Mark.Cigan@Pioneer.com



K. Azhakanandam et al. (eds.), *Recent Advancements in Gene Expression and Enabling Technologies in Crop Plants,* DOI 10.1007/978-1-4939-2202-4_1

**Table 1.1** Proteins expressed in commercial maize events for insect resistance and/or herbicide resistance[a]

| Event name | Year | Protein(s) expressed | Trait | Expression elements | | | Gene design |
|---|---|---|---|---|---|---|---|
| | | | | Promoter | 5′ UTR/intron | Terminator/3′ UTR | |
| MON810 | 1996 | Cry1Ab | IR | E35S CaMV | Hsp70 intron | no terminator (due to deletion) | Optimized |
| BT11 | 1996 | Cry1Ab | IR | 35S CaMV | ZM-ADH1gene IVS6 intron | NOS | Optimized |
| TC1507 | 2001 | Cry1Fa2 | IR | ZM Ubi | Zm-Ubi 5′ UTR-intron1 | ORF25 | Optimized |
| MON863 | 2002 | Cry3Bb1 | IR | 35S CaMV | | Tahsp17 | Optimized |
| 59122 | 2005 | Cry34Ab1 | IR | ZM Ubi | Zm-Ubi 5′ UTR-intron1 | PINII | Optimized |
| | | Cry35Ab1 | | Ta Peroxidase | | PINII | Optimized |
| MON88017 | 2005 | Cry3Bb1 | IR | 2XE35S CaMV | | Tahsp17 | Optimized |
| MIR604 | 2007 | mCry3Aa | IR | MTL | | NOS | Optimized |
| MON89034 | 2008 | Cry1A.105 | IR | 35S CaMV | Cab-5UTR OsActin1 intron | Tahsp17 | Optimized |
| | | ZM-RBC SSU-CTP-Cry2Ab | | FMV | hsp70 intron | NOS | Optimized |
| MIR162 | 2010 | VIP3Aa20 | IR | ZM Ubi | Zm-Ubi 5′ UTR-intron1 | 35S | Optimized |
| 5307 | 2013 | eCry3A.1Ab | IR | CMP | | NOS | Optimized |
| DP4114 | 2013 | Cry1Fa2 | IR | ZM Ubi | Zm-Ubi 5′ UTR-intron1 | ORF25 | Optimized |
| | | Cry34Ab1 | | ZM Ubi | Zm-Ubi 5′ UTR-intron1 | PINII | Optimized |
| | | Cry35Ab1 | | Ta Peroxidase | | PINII | Optimized |
| T25 | 1995 | PAT | HR | 35S CaMV | | 35S | Optimized |
| GA21 | 1997 | ZM-RBC-SSU CTP-MEPSPS | HR | OsActin | OsActin intron | NOS | Native |
| NK603 | 2000 | AtCTP2-CP4EPSPS | HR | OsActin | OsActin-intron | NOS | Native |

[a] Compiled from GM crop database (CERA 2012)

*IR* insect resistance, *HR* herbicide resistance, *CaMV* cauliflower mosaic virus, *FMV* figwort mosaic virus, *MTL* maize metallothionein, *PIN* II protease inhibitor II

**Table 1.2** Examples of industrial and nonpharmaceutical applications in transgenic maize using constitutive promoters

| Protein expressed | Expression elements | | | Targeting | Gene design | Reference |
|---|---|---|---|---|---|---|
| | Promoter | Intron | Terminator | | | |
| E1 endo-glucanase (*Acidothermus cellulolyticus*) | 35S CaMV | | Nos | PR1A SS | Native | Biswas et al. 2006 |
| Avidin (chicken) | Zm-Ubi | Ubi intron | Pin II | | Optimized | Hood et al. 1997 |
| Beta- gluc-uronidase (*E. coli*) | Zm-Ubi | Ubi intron | Pin II | | Native | Witcher et al. 1998 |
| Aprotinin (Bovine) | Zm-Ubi | Ubi intron | Pin II | BAA SS | Optimizes | Zhong et al. 1999 |
| Mn per-oxidase (*Pha-nerochaete chrysospo-rium*) | Zm-Ubi | Ubi intron | Pin II | +/- BAA SS | Native | Clough et al. 2006 |
| Laccase I (*Trametes versicolor*) | PGNpr1 | Ubi intron | Pin II | +/- BAA SS; KDEL | Native | Hood et al. 2003 |
| Xylanase bsx (*Bacillus sp.* NG-27) | Rubi3 | Rubi intron | Nos | BAASS | Optimized | Gray et al. 2011 |
| Xylanase xynB (*Clostridium stercorarium*) | Rubi3 | Rubi intron | Nos | BAASS | Optimized | Gray et al. 2011 |

*BAASS* Barley alpha amylase signal peptide, *Pin* II protease inhibitor II, *CaMV* cauliflower mosaic virus

case, transgenic events result from the integration of the foreign DNA that contains a gene or genes of interest to be expressed, as well as a marker gene (such as an herbicide resistance gene) for selection and identification of transgenic events. The components or genetic elements within the integrated DNA can originate from multiple and diverse sources such as different plant and microbial species; all of which can be engineered to function in combination to contribute to effective expression of those genes and accumulation of the gene products within the correct tissue, at the right level, and at the right developmental stage(s) in maize plants. The fact that the genes to be expressed or genetic elements involved in expression may come from different species, genera, or even kingdoms, also presents a major challenge for finding ways to ensure that these elements work effectively together in a different host organism that results in the required level of protein expression. In this area, optimization of the coding region, choice of promoter, and other regulatory

elements (such as introns, untranslated regions, and terminators) can contribute to successful protein expression. Subcellular targeting can also be beneficial to protein expression by sequestering the protein in compartments where the turnover rate of the protein may be reduced, or the protein is prevented from exerting an effect that negatively impacts agronomic performance due to high expression of the foreign protein. In agricultural production, yield parity between nontransgenic and transgenic products plays a role in trait development, whereas cost and high protein yield is more of a factor in those applications where transgenic maize is used as a recombinant protein production vehicle.

In addition to the importance of genetic elements to protein expression, both integration site and the copy number of the insert can influence the level and consistency of protein expression. Generally, integration of the foreign DNA is difficult to control and genome-based effects may have significant impacts on expression levels. Efforts to target DNA to very precise locations in the maize genome are being developed to reduce positional effects, and the discovery of genetic elements that can buffer integrated DNA from surrounding influence has provided strategies that may ensure more consistent (maybe even more predictable) expression in maize.

The efforts to develop transgenic maize for input traits and as platforms for recombinant protein expression have resulted in the development of strategies to maximize transgene expression. This chapter explores the influence on, and contribution of, several of these strategies to the optimization of transgenic maize protein expression as well as providing knowledge of elements that have been tested or developed for this purpose.

## Applications for Proteins Expressed in Maize

### *Insect Resistance and Herbicide Resistance*

Commercial events expressing insecticidal proteins and/or enzymes conferring resistance to herbicides account for a large percentage of the transgenic acreage for maize. A summary of those events and their traits can be found in Table 1.1 along with the details of the various expression elements that were used to achieve levels of expression needed for trait efficacy.

Maize events with insect-resistance traits express one or more insecticidal proteins that are derived from the soil bacterium *Bacillus thuringiensis (Bt)*. *Bt* has been exploited not only as a natural pest control agent but also as a source of insecticidal proteins that can be expressed in maize (and other crops) for the purpose of plant protection against a spectrum of lepidopteran and coleopteran insects (Szekacs and Darvas 2012) that can damage plants and reduce yield without chemical pesticide intervention. Since 1996, when the first commercial product was approved, nine maize events have been authorized by US regulatory agencies and eight of those continue to be available commercially in the USA. Recently, Event

5307 (Agrisure® Duracade™) and DP4114 maize have been deregulated by the United States Department of Agriculture (USDA; APHIS 2013). The experience gained by the process of optimization involved in the commercialization of insecticidal and herbicide traits has facilitated current understanding of what strategies may be important for protein expression in maize.

Expression of insecticidal genes derived from *Bt* in different crop species has been challenging due to the significant differences in GC nucleotide content between *Bt* and plant species. However, gene optimization to reduce the AT nucleotide content of *Bt* genes has been a contributing factor that may allow *Bt* genes to be expressed successfully at levels sufficient for plant protection in maize (Koziel et al. 1993; De la Riva and Adang 1996). An increase in GC content (with a concomitant reduction in AT content) generally reduces the presence of known or cryptic processing or instability signals that are AT-rich by nature, allowing for improved *in planta* expression (see gene optimization section). From Table 1.1, all IR transgenic events express *Bt* proteins that have been modified from their native (*Bt*) coding sequences for improved expression as indicated by "optimized" in the gene design column. Consistent with the strategy used for *Bt* gene expression, successful use of the phosphinothricin N-acetyltransferase (PAT) gene from the bacterium *Streptomyces viridochromogenes* to confer herbicide resistance to glufosinate (T25) required plant optimization of the coding sequence. In contrast, glyphosate resistance was achieved in maize through the use of essentially the native (plant) versions of the maize 5-enolpyruvyl shikimate-3-phosphate synthase (EPSPS) gene with specific amino acid mutations (GA21), or the EPSPS gene from *Agrobacterium tumefaciens* CP4 (NK603). Today, commercial products often express one or more IR and HR trait genes, increasing the complexity of the optimization process required to provide expression levels to meet commercial trait efficacy.

Promoter selection is also a factor that contributes to the ability to express genes at efficacious levels in the necessary tissues at the correct developmental stages in maize. Promoters that are seed-specific, for example, are preferred for expression of proteins that have pharmaceutical and industrial value when using maize as protein production platforms. These promoters allow for high and stable accumulation of functional protein in the natural storage organs, kernels, of maize (Stroger et al. 2002; see also Chap. 3). Promoters that facilitate strong constitutive expression of proteins throughout different developmental stages of maize are useful for IR and HR applications. In these cases, high levels of protein expression of IR or HR genes are needed for protection against insect pests at multiple feeding sites (e.g., leaf, sheath, stalk, root, silk, and ears), or in the tissues that are sensitive to the action of herbicides, respectively. Most commercial events expressing *Bt* genes have used either the maize polyubiquitin 1 promoter (Ubi-1;Christensen and Quail 1996) or a plant viral promoter derived from the caulimovirus family (35S of cauliflower mosaic virus or figwort mosaic virus; Odell et al. 1985; Bhattacharyya et al. 2002). Root-preferred promoters such as a maize metallothionein (MTL) or a wheat peroxidase (Ta-Peroxidase) have been used to express corn rootworm insecticidal proteins in MIR604 and 59122 (Table 1.1). Resistance to the herbicide glyphosate in GA21 and NK603 has been achieved by constitutive expression of EPSPS genes

using a rice actin (Os-Actin) promoter (McElroy et al. 1990). The inclusion of a native intron that is naturally associated with the promoter, or the introduction of a heterologous plant intron within the 5′ untranslated leader sequence (UTR) of a gene, is a common strategy that has been used to enhance maize protein expression. (See section in this chapter on intron-mediated enhancement of gene expression.) This strategy has been effective particularly in combination with plant viral promoters such as c*auliflower mosaic virus* 35S promoter (CaMV 35S) and figwort mosaic virus (Table 1.1).

## *Industrial Enzymes and Nonpharmaceutical Protein Reagents Produced in Maize*

Several proteins with industrial or reagent-based applications have been expressed in maize due to the competitive opportunity for large-scale protein production (Table 1.2). The advantages of using maize as a plant-based platform for protein production include a well-established system for genetic transformation, an established toolbox of regulatory elements, and targeting signals to help maximize transgene expression and accumulation, high yield in the field, infrastructure for field production and harvest, and relatively large grain size compared to other plant species (Ramessar et al. 2008). Maize as an expression platform can provide for the correct folding of complex proteins such as antibodies, posttranslational modification, scale of expression, and absence of human pathogens (Naqvi et al. 2011). The ability to express proteins in selective tissues like kernels offers flexibility for storage over long periods of time before protein extraction without significant loss in protein activity. Kernels may also be a means for delivery in feed applications. Grain size is an important factor when considering the often successful strategy of accumulation of recombinant protein in grain. Ramessar et al. (2008) and Hood and Howard (2009) provided an excellent overview of the range and purpose of proteins expressed in maize plants (particularly using seed-specific promoters) and are not covered extensively in this chapter.

Strategies that improve the expression and accumulation of heterologous proteins in maize for recombinant protein expression platforms have been developed with the emphasis on maximizing the yield of recombinant proteins per unit biomass to be as economically feasible as possible. Reduction of any potential negative impact of high protein expression on plant health, agronomics, and yield is also desirable. The need to satisfy both high yield per unit biomass and minimize effects on yield and agronomics has led to one strategy that combines the use of strong constitutive promoters, such as maize ubiquitin (Ubi1), rice ubiquitin (rUBi3), and CaMV 35S, in combination with subcellular targeting (Table 1.2).

High expression of proteins throughout the plant can be achieved by the use of these strong constitutive promoters. However, in several cases, aberrant plant phenotypes have been observed such as early senescence, male sterility, and low/no seed set (Clough et al. 2006), stunting and plant mortality (Hood et al. 2003),

and stunting, reproductive development problems, and shriveled grain (Gray et al. 2011). In some cases, constitutive expression resulted in high expression and normal plant phenotype (Hood et al. 1997; Witcher et al. 1998; Zhong et al. 1999). Whether a protein has an effect on plant health can be related to a combination of the properties of the overexpressed protein (e.g., enzyme, solubility, capability of interaction with plant proteins) and how well maize cells or tissues tolerate its expression. In several cases, depending on the types of genes that were expressed, subcellular targeting signals designed to sequester the proteins in different subcellular compartments (e.g., cell wall, endoplasmic reticulum, vacuole, cytoplasm) have been used to achieve high expression without observable aberrant plant phenotypes (Zhong et al. 1999; Hood et al. 2003). In other cases, confining expression of the heterologous protein to kernels using seed-specific promoters has been an effective strategy (see Chap. 3 in this book from Howard and Hood).

## Influence of Gene Optimization on Protein Expression Levels

One factor in the successful expression of proteins in maize (and any other heterologous expression system) is the coding sequence. The nucleotide sequence can impact expression due to multiple factors that may affect how well a gene is expressed and translated in plant cells. In the majority of commercial products, gene optimization is a part of the process to maximize expression of heterologous proteins for different applications (see Table 1.1 and 1.2), especially if the gene is derived from phylogenically different sources (e.g., bacteria, animals). The increasingly low cost of gene synthesis provides the opportunity to back translate a protein and modify its nucleotide coding sequence to optimize expression without changing the protein sequence. In fact, many gene synthesis companies independently provide codon optimization services based on different algorithms that have been designed to improve expression. Most of these algorithms adapt the codon usage of a gene of interest to the typical codon usage of the intended host as one component of the design process, and generally take into account several other parameters including mRNA secondary structure.

A benefit of the genomics revolution has been the exposure of codon biases for many different plant species. This has led to codon counting to decipher which codons are favored in high expressing genes from an organism of interest. Adaptation of codon bias (Sharp and Li 1987; Carbone et al. 2003; Jansen et al. 2003) is usually a primary consideration for gene optimization in plants with the intention of mimicking a well-expressed host gene. Selecting the most frequently used codon for each amino acid allows the use of the most abundant tRNAs and minimizes effects on expression due to the presence of rare codons. The Codon Adaptation Index (CAI; Sharp and Li 1987) is one of several statistical approaches that have been developed that compares a designed gene with host codon bias. Genes that

maximize the CAI have expressed well in many instances, although the tested gene set is small. Maize has an overall G+C content of about 55% (Nakamura 2000; www.kazusa.or.jp/codon) with a preference for a G or a C nucleotide in the third or wobble position of the codon (Fennoy and Bailey-Serres 1993; Liu et al. 2010). Koziel et al. (1993) constructed a synthetic version of a Cry1Ab gene for transgenic maize expression by increasing G+C content to 65% that reflected a maize-preferred codon usage. This study reported expression of Cry1Ab protein in transgenic maize events at levels insecticidal to European corn borer. Improvements in the expression of heterologous genes as a result of maize codon optimization have been reported for blue fluorescent protein (BFP), green fluorescent protein (GFP), yellow fluorescent protein (YFP; Sattarzadeh et al. 2010) and xylanase bsx (Gray et al. 2011). Whether improved expression is due directly to the codon bias, or to other factors is difficult to differentiate. Increasing G+C content may inherently remove potential elements such as cryptic splicing sites, premature polyadenylation sites, RNA instability motifs (Murray et al. 1991; van Aarssen et al. 1995; Christov et al. 1998; Diehn et al. 1998; De Rocher et al. 1998), and other elements that may lead to reduced transcriptional and translational efficiency. The intentional elimination of several polyadenylation signals and instability motifs improved expression of a *Bt* gene in maize (De la Riva and Adang 1996).

Frequently, gene optimization is performed in the absence of experimentally testing expression of the native gene sequence in maize. This is done *a priori* based on a general assumption that an improvement in expression will be the likely outcome (Hood et al. 1997; Zhong et al. 1999; Gray et al. 2011). Optimization may be particularly beneficial if a gene to be expressed in maize originates from a bacterial species such as *B. thuringiensis* (Table 1.1) where its G+C content (35.5%) is significantly lower compared to maize (55%; De la Riva and Adang 1996). The lower G+C content increases the probability that multiple deleterious sequence motifs may be present since several of these sequence motifs (described above) frequently contain A+T rich sequences. A very low preference for G+C (24.6%) at the wobble position, in the case of *B. thuringiensis* genes, compared to maize (64%; www.kazusa.or.jp/codon) may result in the presence of maize rare codons in the native sequence.

Optimization may not be necessary to achieve good expression for every heterologous gene. There are several examples of native genes from fungi and animals that express well in maize and achieve their intended functionality (Hood et al. 2003; Woodard et al. 2003; Clough et al. 2006; Biswas et al. 2006) (Tables 1.1 and 1.2). In these cases, the genes have maize-like characteristics. Overall G+C content and preference for G+C in the wobble position is comparable to, or greater than, maize and deleterious sequences such as cryptic splicing sites, premature polyadenylation sites, and RNA instability motifs are rare or absent. The presence of rare codons is also minimal in these sequences. However, strict adherence to these characteristics may not always be required to obtain desired expression levels.

## Control of Protein Expression

### *Promoters*

A consideration to achieving the desired levels of expression in maize and other plant species is the choice of promoter. Promoters direct expression of transgenes in plants quantitatively, spatially, and temporally. Proper selection of a promoter is reflected by the specific end-use application of the transgene, most commonly recombinant protein production or crop protection. In transgenic maize plants generated for the purpose of recombinant protein expression, the latter may be targeted specifically in the seed. Applications directed toward IR or HR commonly focus on constitutive expression throughout most developmental stages of the plant. In both cases, optimizing protein expression and accumulation can require a balance between maximizing expression in the tissues of interest and minimizing negative impacts on the plant in the form of agronomic or yield penalties. How well a gene can be expressed (e.g., gene design), how potent the gene product is (e.g., efficacy, enzymatic activity), and the inherent level of plant toxicity caused by overexpression of the recombinant protein influences promoter selection. In most cases, optimization of expression (and phenotype) will be empirical, requiring the careful evaluation of multiple promoters to identify those that function effectively to achieve a desired outcome. This empirical approach requires the availability of alternative promoter choices that can be tested with each transgene.

The need for alternative promoters also plays a role in the ability to effectively coexpress multiple genes in a molecular stack configuration (Peremarti et al. 2010). Multigene transformation continues to increase in plant biotechnology in order to generate complex trait stacks or pyramids that satisfy future needs for transgenic maize products. These products may include different trait package combinations of HR, IR, improved agronomic characteristics, improved nutritional value, and recombinant protein production. The versatility to deploy different promoters can be beneficial for coexpression of multiple genes but also can increase construct integrity and reduce the potential for gene silencing. In the last 5 years, about 120 maize promoters have been patented (Fig. 1.1) primarily by commercial entities to provide promoters with different strengths and specificities to help meet the expression challenges needed for various transgenic applications.

### *Application of Expression Profiling Technology to Promoter Discovery*

Previous methods used to identify promoters with desirable expression patterns relied primarily on information generated from the libraries of expressed sequence tags (ESTs) and microarrays which identified promoter candidates based on the