BURLEIGH DODDS SERIES IN AGRICULTURAL SCIENCE

Genome editing for precision crop breeding

Edited by Dr Matthew R. Willmann, Cornell University, USA

burleigh dodds SCIENCE PUBLISHING

Contents

Series list xi
Acknowledgements xviii
Introduction xix

Part 1 Genome editing techniques

1 Using TALENs for genome editing in plants 3
Hilal Betul Kaya, Manisa Celal Bayar University, Turkey; Rhitu Rai, ICAR-National Institute for Plant Biotechnology, India; and Adam J. Bogdanove, Cornell University, USA

1 Introduction 3
2 How TALENs work 5
3 How have TALENs been used? 7
4 How to use TALENs 12
5 Conclusion 20
6 Acknowledgements 20
7 References 20

2 Double strand break (DSB) repair pathways in plants and their application in genome engineering 27
Natalja Beying, Carla Schmidt and Holger Puchta, Karlsruhe Institute of Technology, Germany

1 Introduction 27
2 DSB repair via non-homologous end joining 29
3 DSB repair using homologous sequences 40
4 Perspectives 48
5 Where to look for further information 48
6 References 49

3 Advances in the generation of insertion-based genome edits in plants 63
Baike Wang and Juan Wang, Institute of Horticulture Crops – Xinjiang Academy of Agricultural Sciences, China; Shaoyong Huang, Institute

of Horticulture Crops – Xinjiang Academy of Agricultural Sciences and College of Forestry and Horticulture – Xinjiang Agricultural University, China; and Yaping Tang, Ning Li, Shengbao Yang, Tao Yang and Qinghui Yu, Institute of Horticulture Crops – Xinjiang Academy of Agricultural Sciences, China

1 Introduction 63
2 Outcomes of knock-ins 66
3 The manner of targeting the integration of DNA fragments 70
4 Methods to increase the efficacy and accuracy of knock-ins 76
5 Conclusion and future trends 80
6 Where to look for further information 81
7 References 85

4 Viruses as vectors for the delivery of gene-editing reagents 97
Evan E. Ellison, James C. Chamness and Daniel F. Voytas, University of Minnesota, USA

1 Introduction 97
2 DNA viruses: replicon vectors for efficient gene editing through homology-directed repair (HDR) 98
3 Geminiviral replicons (GVRs): deconstructed geminiviruses that serve as replicating vectors in plants 98
4 Replicon vectors for efficient homology-directed repair 103
5 Case studies: GVR-based editing strategies in dicots and monocots 104
6 GVR-based editing strategies: summary and future trends 108
7 RNA viruses: mobile vectors broadly applicable for gene-editing reagent delivery 109
8 Case studies: use viral vectors to create permanent genetic changes 111
9 RNA viral vectors: conclusion and future trends 116
10 Where to look for further information 117
11 References 117

5 Progress in precise and predictable genome editing in plants with base editing 123
Sabine Fräbel, Shai J. Lawit, Jingyi Nie, David G. Schwark, Thomas J. Poorten and Nathaniel D. Graham, Pairwise Plants, USA

1 Introduction 123
2 Progress in mammalian systems 125
3 Cytosine base editing in plants 127
4 Adenine base editing in plants 132
5 Broadening protospacer adjacent motif (PAM) accessibility 132
6 Sequence diversification in crops 134
7 Off-target base editing 135

8 Current applications of base editors in crops 136
9 Conclusion 137
10 Where to look for further information 139
11 References 139

6 Advances in guide RNA design for editing plant genomes using CRISPR-Cas systems 147
Kaiyuan Chen, Hao Liu and Kabin Xie, Huazhong Agricultural University, China; Muhammad Tahir ul Qamar and Ling-Ling Chen, Huazhong Agricultural University, and Guangxi University, China

1 Introduction 147
2 Designing sgRNAs 149
3 Bioinformatics programs for designing sgRNAs 155
4 Predicting the outcome of CRISPR-mediated genome editing 166
5 Conclusions 169
6 Where to look for further information 169
7 References 169

7 Advances in assembling gRNA/Cas9 constructs in genome editing of plants 175
Marta Vázquez Vilar, Sara Selma, Asun Fernández del Carmen and Diego Orzáez, Instituto de Biología Molecular and Celular de Plantas de Valencia (CSIC-UPV), Spain

1 Introduction 175
2 Modular cloning tools for assembling CRISPR/Cas constructs 177
3 The basics of CRISPR/Cas9 cloning 179
4 CRISPR-Cas12 specific cloning requirements 182
5 Expanding CRISPR/Cas9 activities with modular cloning 182
6 The polycistronic strategy for gRNA multiplexing 184
7 Case study: multiple gene targeting in tobacco with GoldenBraid cloning 186
8 References 191

8 Strategies for CRISPR/Cas9-mediated genome editing: from delivery to production of modified plants 195
William Gordon-Kamm, Pierluigi Barone, Sergei Svitashev, Jeffry D. Sander, Sandeep Kumar and Todd Jones, Corteva Agriscience, USA

1 Introduction 195
2 Delivery of genome editing components into plant cells 196
3 Delivery methods for genome editing reagents: delivery into single cells 200
4 Delivery methods for genome editing reagents: delivery into intact tissues 201
5 Alternatives to DNA delivery 207

6 Morphogenic genes increase transformation efficiency and extend genotype range 208
7 Morphogenic genes permit transformation in nontraditional explants 211
8 Future trends 212
9 Where to look for further information 214
10 References 215

9 Advances in screening plants for edits and off-targets 231
Chun Wang and Kejian Wang, China National Rice Research Institute, China

1 Introduction 231
2 PCR/RE assay 232
3 Mismatch cleavage assay 232
4 PCR/RNP 233
5 High-resolution melting (HRM) analysis 234
6 ACT-PCR 234
7 Analysis of Sanger sequencing chromatograms 235
8 Next generation sequencing (NGS) 238
9 Detection of off-target effects 240
10 Conclusion and future trends 241
11 Where to look for further information 243
12 References 243

10 Targeted modification of promoters 247
Andika Gunadi and Ning Zhang, Boyce Thompson Institute, USA; and John J. Finer, The Ohio State University, USA

1 Introduction 247
2 Plant promoters, *cis*-acting regulatory elements and leading introns 249
3 Principles of targeted promoter modification 259
4 Approaches for targeted promoter modification 261
5 Conclusion and future trends 266
6 Where to look for further information 267
7 References 267

11 The regulation of genome-edited crops 279
Gregory Jaffe, Center for Science in the Public Interest, USA

1 Introduction 279
2 Background on genome editing and its regulation 280
3 Whether the Cartagena Biosafety Protocol applies to genome-edited crops 286
4 Country case studies demonstrating regulation of genome-edited crops 288
5 Analysis of regulations in different countries 308
6 Conclusion and future trends 315
7 References 317

Part 2 Applications

12 Genome editing of barley 325
Martin Becker, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Germany; and Goetz Hensel, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Germany, and Palacký University, Czech Republic

1 Introduction 325
2 Expression systems used for targeted mutagenesis 328
3 Enhancing plant development and performance 330
4 Molecular understanding of resistance against plant pathogens 333
5 Targeted mutagenesis for molecular farming purposes 333
6 Conclusion 334
7 Where to look for further information 335
8 References 336

13 Genome editing of maize 341
Jacob D. Zobrist, Morgan McCaw, Minjeong Kang, Alan L. Eggenberger, Keunsub Lee, and Kan Wang, Iowa State University, USA

1 Introduction 341
2 Site-specific engineering in maize: an overview 343
3 Current challenges in maize genome editing 355
4 Case study: gene editing of a recalcitrant line by co-bombardment of morphogenic genes (*Babyboom/Wuschel*) and CRISPR/Cas9 reagents 360
5 Conclusion and future trends 364
6 Acknowledgements 365
7 Where to look for further information 365
8 References 367

14 Genome editing in sorghum 377
Aixia Li, Shandong University, China; and David R. Holding, University of Nebraska-Lincoln, USA

1 Introduction 377
2 Advances of gene editing in sorghum 379
3 Application of CRISPR/Cas9 gene editing for sorghum protein improvement 382
4 Future trends 396
5 Conclusion 397
6 References 398

15 CRISPR/Cas9-mediated genome editing in *Brassica* 403
Cheng Dai, Xia Tian and Chaozhi Ma, Huazhong Agricultural University, China

1 Introduction 403
2 Gene disruption in *Brassica* species 405
3 Homologous recombination-based editing in *Brassica* species 408
4 Case study: generating mutants by CRISPR/Cas9 toolbox 408
5 Conclusion and future trends 415
6 Where to look for further information 415
7 References 416

16 Genome editing of tomatoes and other Solanaceae 421
Joyce Van Eck, The Boyce Thompson Institute and Cornell University, USA

1 Introduction 421
2 Gene editing proof-of-concept investigations 423
3 DNA sequence insertion 425
4 Application of CRISPR/Cas ribonucleoprotein complexes 426
5 Base editing 427
6 Modification of gene function for improved disease tolerance 428
7 Editing to affect fruit and tuber quality 429
8 Plant growth and development modifications 432
9 Case study: application of gene editing for de novo domestication 432
10 Conclusion and future trends 434
11 Where to look for further information 434
12 References 436

17 Genome editing of woody perennial trees 441
Chung-Jui Tsai, University of Georgia, USA

1 Introduction 441
2 The heterozygous genome challenge in gene editing 442
3 Genome resource limitation of woody perennial crops 444
4 Gene-editing reagents 445
5 Stability 446
6 CRISPR/Cas9-empowered characterization of duplicated genes 447
7 Conclusion and future trends 450
8 Where to look for further information 451
9 Acknowledgements 452
10 References 452

Index 459

Introduction

Genome editing is rapidly transforming plant research. The technique offers unparalleled precision in breeding without the need to introduce foreign DNA into plants. CRISPR/Cas systems have established themselves as the leading technique in genome editing. This collection of chapters takes stock of the wealth of research on these techniques and their potential in crop breeding in improving traits such as yield, disease resistance, drought tolerance and nutrient use efficiency.

Part 1 of this volume reviews advances in techniques such as TALENS and zinc finger nucleases, double-strand break repair techniques, insertion-based genome edits, base editing, guide RNAs and gRNA/Cas9 constructs. Chapters in this part of the book also discuss advances in screening plants as well as regulatory issues. Part 2 examines applications of gene editing in key cereal crops, specifically focusing on barley, maize and sorghum. Other chapters cover crops such as brassicas, tomatoes and perennials.

Part 1 Genome editing techniques

Part 1 begins with a chapter that examines using transcription activator-like effector nucleases (TALENs) for genome editing in plants. TALENs are powerful tools for precise and efficient locus-specific editing. Among the various genome editing tools that were developed prior to the CRISPR/Cas system, TALENs have been the most widely used, and they remain worthy of consideration due to their unique targeting properties and the current intellectual property landscape, which give them distinct advantages over CRISPR/Cas for some applications. Chapter 1 briefly reviews the use of TALENs in plants including their important distinguishing features, as well as design principles and tools, methods for construct assembly, and other available resources to assist researchers using this technology.

Chapter 2 reviews double strand break (DSB) repair pathways in plants and their application in genome engineering. In genome engineering, after targeted induction of DSBs researchers take advantage of the organisms' own repair mechanisms to induce different kinds of sequence changes into the genome. Therefore, understanding of the underlying mechanisms is essential. The chapter describes in detail the two main pathways of DSB repair in plant cells, non-homologous end joining (NHEJ) and homologous recombination (HR) and sums up what we have learned over the last few decades about them. The chapter also summarises the different models that have been proposed and set these into relation with the molecular outcomes of different classes of

DSB repair. Moreover, it describes the factors that have been identified to be involved in these pathways. Applying this knowledge of DSB repair should help us to improve the efficiency of different types of genome engineering in plants.

The next chapter assesses advances in the generation of insertion-based genome edits in plants. Tremendous progress has been achieved in the field of gene editing in plants, such as with the use of zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and clustered regularly interspaced short palindromic repeats (CRISPR). Because of the potential advantages associated with mutant creation and crop germplasm innovation, genome editing technology has been rapidly developed and widely used in crop improvement in recent years. Chapter 3 documents some of the important recent developments and applications of genome-editing tools, especially with respect to gene knock-ins. It introduces the mechanism underlying knock-ins and different outcomes of insertion. The chapter also discusses genome editing tools and methods developed to improve insertion efficiencies. Additionally, the chapter reviews the recent trends in genetic editing biotechnologies; several strategies are being developed to further improve the efficiency of plant gene knock-ins. Undoubtedly, CRISPR/Cas technology will boost the development of new plant breeding techniques tremendously.

The subject of Chapter 4 is using viruses as vectors for the delivery of gene editing reagents. A significant challenge for plant gene editing is the delivery of editing reagents to germline or regenerable cells to recover heritable genetic modifications. Reagent delivery using biolistics or Agrobacterium is only possible with a limited range of species and genotypes, and inefficient editing or lengthy tissue culture steps further limit throughput. Viruses are natural vectors for nucleic acids, and both DNA and RNA plant viruses have been engineered to extend or replace conventional vectors for delivery of gene editing reagents. The chapter reviews aspects of viral biology essential for engineering vectors, highlights landmark studies using viruses to overcome traditional limitations in gene editing, and outlines important considerations for the use of viral vectors in new systems or for new targets. Motivated by fundamental differences in both their infection modes and utility as vectors, DNA and RNA viruses are discussed separately.

Moving on from Chapter 4, Chapter 5 addresses the progress that has been made in precise and predictable genome editing in plants with base editing. Base editors are gene editing tools that allow targeted nucleic acid conversions, most commonly C>T and A>G, through pairing of deamination domains with impaired nucleases. Multiple deaminase domains and architectures have been demonstrated in planta across a wide array of species, with both cytosine and adenine base editing frequencies being observed at over 80%. The chapter first focuses on progress that has been made in mammalian systems, then moves on to discuss cytosine and adenine base editing in plants. A section

on protospacer adjacent motif (PAM) accessibility is also included, followed by a discussion on sequence diversification in crops, off-target base editing and current applications of base editors in crops.

The next chapter analyses the advances in guide RNA design for editing plant genomes using CRISPR-Cas systems. To date, many bioinformatic tools have been developed to facilitate the design of sgRNAs for CRISPR-based genome editing in eukaryotes. Chapter 6 focuses specifically on the general rules for selecting target sites for genome editing using CRISPR-Cas technology and summarises the bioinformatic tools that can be used to design sgRNA sequences. It also discusses predicting the outcome of CRISPR-mediated genome editing, before concluding with an overview of why designing a sgRNA is a crucial step for CRISPR-based genome editing.

Chapter 7 focuses on the advances in assembling gRNA/Cas9 constructs in genome editing of plants. It reviews the principles of CRISPR cloning in binary vectors and the different methods and elements employed, including the nucleases alternative to Cas9. It pays special attention to modular cloning strategies and multiplexing tools as well as the engineering of expanded Cas activities. Finally, the chapter includes a case study of the cloning of a nine-gRNA multiplex construct and the analysis of its transformants in tobacco plants.

The subject of Chapter 8 is strategies for CRISPR/Cas9-mediated genome editing: from delivery to production of modified plants. Despite the conspicuous and rapid development of genome editing tools, implementing this technology in plants often remains constrained by our inability to regenerate fertile genome-modified plants. Finding a viable combination of delivery method, genome editing reagents, and plant regeneration system has often been both crop-specific and crop-limited. Recent progress using morphogenic genes such as Wuschel2 (Wus2) and Babyboom (Bbm) has alleviated some of these crop-specific challenges, and along with other improvements, continues to expand the range of plants that can be edited. The chapter focuses specifically on delivery of genome editing components into plant cells, then moves on to examine the delivery methods for genome editing reagents in terms of delivery into single cells and intact tissues. A section on alternatives to DNA delivery is also provided. This is followed by discussions on morphogenic genes and how they can increase transformation efficiency, extend genotype range and permit transformation in non-traditional explants.

Chapter 9 considers the advances in screening plants for edits and off targets. Genome editing technology greatly accelerates the development of plant molecular research and crop breeding. However, accurately detecting mutations induced by sequence-specific nucleases between large-scale plants produced from transformation become a challenge. The chapter introduces some frequently used molecular methods of identify genome editing induced

mutation or off-targets, which developed in recent years. It also discusses the advantages and disadvantages of these methods and the scope of application, providing a useful reference for researchers who will identify genome-edited plants.

The next chapter assesses targeted modification of promoters. Although most genome editing efforts focus on modifications to gene coding regions, this chapter emphasizes genome editing of the upstream regulatory regions. Thoughtful editing of the promoter region will ultimately lead to improved plants, modified for more precise control of the intensity and specificity of native gene expression. Chapter 10 presents an overview of the promoter or upstream regulatory region of a gene and describes how this sequence is defined and studied. It then describes how the composition and arrangements of cis-regulatory elements within the promoter and the leading intron associated with the promoter region have been studied using classical transgenic approaches to reveal what regulatory components might be suitable for genome editing approaches. Finally, the chapter offers some suggestions for pursuit of promoter editing and gene expression modulation, which will eventually lead to modified plants with an altered regulation of native gene expression.

The final chapter of Part 1 reviews the regulation of genome edited crops. To explore the regulation of genome edited crops, Chapter 11 describes the regulatory systems adopted by different countries as well as how the Cartagena Biosafety Protocol might apply. It finds there is a continuum of oversight, with some countries applying their GMO biosafety laws to all genome edited crops and other countries exempting certain genome edited techniques and/or products from any special oversight. For countries with what on the surface seem like similar regulatory policies, when their regulations are applied to identical products, they often produce different results, which will make international harmonization difficult. In addition, how much oversight to impose on different crops tends not to be based on scientific evidence. Finally, the regulatory treatment for multi-edited products is unclear. It will take time and additional the development of additional genome edited crops before there is sufficient clarity on how all types of genome edited crops will be regulated.

Part 2 Applications

Part 2 opens with a discussion of genome editing of barley. Although barley is of great importance for the brewing and animal feed industries and is regarded as a model for small grain cereals, only a few results on targeted gene modification using CRISPR/Cas endonuclease technology have been published to date. Chapter 12 examines the frontiers and achievements of the

currently used techniques in barley genome modification. It focuses specifically on the expression systems used for targeted mutagenesis, enhancing plant development and performance and also examines the molecular understanding of resistance against plant pathogens. A section on targeted mutagenesis for molecular farming purposes is also included. The chapter concludes by emphasising the importance of developing genome editing techniques in barley.

Moving on from Chapter 12, Chapter 13 focuses on genome editing of maize. Developed over thousands of years largely through human intervention, the modern maize genome can now be precisely modified for agricultural improvement and scientific research. The chapter focuses on the recent progress that has been made in utilizing site-specific nuclease (SSN) technologies in maize genome engineering. Many SSNs, such as meganucleases, zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated proteins (Cas) have been used in maize for both functional analysis and trait improvement. The chapter summarizes the recent innovations related to maize genome editing using SSN technologies, the type of approaches, target genes and traits, and reagent delivery methods. It also discusses the current challenges as well as potential improvements for maize genome engineering protocols.

Chapter 14 addresses genome editing in sorghum. Genetic engineering plays a key role in plant functional research and genetic improvement. A novel and powerful gene editing technique, CRISPR/Cas9, which was developed from a type II bacterial immune system, opened up a new era in precision genetic engineering in plants. This technique is based on a non-permanent transgene system and is starting to be adopted for precise gene editing in major cereal crops. It offers tremendous potential to accelerate crop improvement in a way that potentially reduces or eliminates the cumbersome and expensive regulatory processes associated with traditional transgenic crops. The chapter describes the advance of gene editing applied to sorghum, a drought tolerant C4 crop, and a successful strategy of CRISPR/Cas9 mediated gene family editing to improve sorghum digestibility and protein quality. It also discusses future prospects of CRISPR/Cas9 gene editing for sorghum genetic improvement.

The next chapter examines CRISPR/Cas9-mediated genome editing in Brassica. CRISPR/Cas9 is a valuable tool for both basic and applied research that has been widely applied to different plant species. Chapter 15 reviews the application of the CRISPR/Cas9 genome editing toolkit in various *Brassica* crops. The chapter also provides a case study in *Brassica napus*, specifically describing how mutants are generated using the CRISPR/Cas9 toolbox. Collectively, the results demonstrate that CRISPR/Cas9 is an efficient tool for

creating targeted genome modifications at multiple loci in B. napus. These findings open many doors for biotechnological applications in oilseed crops.

Chapter 16 focuses on genome editing of tomatoes and other Solanaceae. Application and effectiveness of different editing approaches based on CRISPR/Cas of solanaceous species are outlined in this chapter. The chapter begins by discussing gene editing proof-of-concept investigations, which is then followed by a review of DNA sequence insertion. The chapter moves on to examine base editing and modification of gene function for improved disease tolerance, specifically focusing on Solanaceae species. Gene editing to affect fruit and tuber quality is also discussed and a section on plant growth and development modifications is provided. The chapter also provides a case study on application of gene editing for de novo domestication.

The final chapter of the book examines genome editing of woody perennial trees. Chapter 17 discusses the challenges facing woody perennial trees, namely the heterozygosity of outcrossing species and limited genomic resources. Gene knockouts (KO) represent the predominant applications of the CRISPR technology in woody perennial crops to date and will be the focus of this chapter. The issues considered for gene KO are generally applicable to other CRISPR applications targeting regulatory sequences or non-coding genes. Case studies from the Populus 4-coumarate:CoA ligase (4CL) family are presented to demonstrate the power of CRISPR in elucidating functional redundancy as well as specificity of duplicated genes in a perennial woody species.

Chapter 1

Using TALENs for genome editing in plants

Hilal Betul Kaya, Manisa Celal Bayar University, Turkey; Rhitu Rai, ICAR-National Institute for Plant Biotechnology, India; and Adam J. Bogdanove, Cornell University, USA

1 Introduction
2 How TALENs work
3 How have TALENs been used?
4 How to use TALENs
5 Conclusion
6 Acknowledgements
7 References

1 Introduction

1.1 Why consider TALENs?

Given the ease of use, proven efficacy, and focused attention on CRISPR/Cas systems ('CRISPR') for genome editing, it is reasonable to ask why anyone should consider using transcription activator-like (TAL) effector nucleases (TALENs). Of the four main site-directed nucleases used to date for genome editing—zinc finger nucleases, meganucleases, TALENs, and CRISPR (reviewed in Carroll, 2014; Bogdanove et al., 2018)—CRISPR, is by far the most widely used. In contrast to the other reagents, targeting CRISPR to a sequence of choice is a simple matter of generating an RNA construct that hybridizes to that sequence. The specificity of the other nucleases is determined by their direct interactions with the DNA and thus requires protein engineering to customize. This is most challenging for meganucleases, which recognize long sequences (typically 12–40 bases and up to 60 bases (Stoddard, 2014)) through DNA contacts that are structurally complex. It is less so but still arduous for zinc finger nucleases, in which individual structural units ('fingers') specify triplet nucleotides in the bound DNA sequence but are influenced by the neighboring fingers. TAL effectors emerged as the first truly modular DNA-binding proteins, their specificity being determined by polymorphic structural repeats that individually and independently specify single nucleotides in the

http://dx.doi.org/10.19103/AS.2020.0082.02

bound DNA sequence, in a contiguous fashion. TALENs are customized simply by assembling an array of repeats in the TAL effector domains that specify the desired sequence. Several efficient protocols using publicly available libraries of repeat modules have been developed for low-cost assembly in as little as three days using standard molecular biology techniques. Along with the ease of engineering TALENs, there are three important features of TALENs distinct from CRISPR that make these reagents worth considering, described below.

1.2 TALENs target long sequences

TALENs are described in detail in the next section of this chapter. Relevant here, they function as two monomers that bind distinct sequences on the DNA separated by a short spacer and on opposite strands. Generally, TAL effector-binding sites are no shorter than 12 bases and can measure as long as 30 or more bases (Boch and Bonas, 2010). Thus, the two monomers together specify a binding site length rivaling that of meganucleases, likely to be unique even within a large genome. This contrasts with the 20 base recognition sequence common for CRISPR. Theoretically, TALEN monomers could be designed with super-long repeat arrays. However, extending a repeat array past about 18 repeats has been observed to decrease overall specificity by increasing the non-specific affinity for DNA, while leaving affinity for the targeted DNA sequence relatively unchanged (Rinaldi et al., 2017).

1.3 TALENS offer the potential for specificity tuning

Different repeat types in TAL effectors differ not only in their base preference but in the stringency of that preference. Indeed, there are highly specific repeat types for each of the four bases, there are repeat types with dual specificity, and there are 'wildcard' repeat types that are able to accommodate any of the four bases (Miller et al., 2015; Yang et al., 2014). Thus, a repeat array could be assembled to judiciously relax specificity at particular positions, enabling targeting of multiple, moderately conserved sequences that match a consensus. With CRISPR, in which target recognition is governed by Watson and Crick base pairing, this can be achieved by introducing multiple gRNAs. But where the number of gRNAs needed becomes a limiting factor, TALENs could be a useful alternative, taking advantage of different repeat types to tune the specificity to capture the range of sequences desired (discussed further in Bogdanove et al., 2018).

1.4 The TALEN intellectual property landscape is settled

In contrast to CRISPR, for which (at the time of this writing) an intellectual property (IP) dispute continues, ownership of the dominant IP for TALENs is

uncontested. The TAL effector-DNA binding 'code' was patented by members of one of the groups that co-discovered it (Boch et al., 2009), and rights to that intellectual property are managed by the 2Blades Foundation, a non-profit organization dedicated to advancing technologies to improve crop disease resistance. For uses outside of plants, 2Blades licensed the code exclusively to ThermoFisher Scientific. For uses in plants, however, 2Blades retains the rights to grant additional licenses. For the TALEN technology itself, the University of Minnesota and Iowa State University own the rights but granted an exclusive license to the biopharmaceutical company Cellectis. Cellectis established the subsidiary Calyxt, an independent company now that uses TALENs to develop healthier food crops. Commercial use of TALENs requires licenses both from 2Blades and from Calyxt. Depending on the specific architecture and delivery methods used, additional licenses from these or other parties may be required.

The remainder of this chapter describes how TALENs work, discusses how they have been used in plant genome editing, and provides a brief overview of the workflow for a TALEN-mediated genome editing experiment.

2 How TALENs work

TALENs create targeted double-strand breaks in DNA for genome editing via non-homologous end-joining (NHEJ) or homology-dependent repair (HDR) (Carroll, 2014). Each TALEN monomer is a fusion of a customized TAL-effector DNA recognition and binding domain to the N-terminus of the non-specific catalytic domain of the Type IIS restriction endonuclease *Fok*I. *Fok*I functions as a dimer. Thus, TALEN monomers function in pairs, binding the DNA on opposite strands and on opposite sides of a short-spacer sequence, in a tail-to-tail orientation, so that the *Fok*I domains come together across that spacer to make the break (Fig. 1a). TAL effectors in their native context are produced by members of the plant pathogenic bacterial genus *Xanthomonas*, which, upon interacting with the host plant, secrete them into the host nucleus, where the proteins mimic host transcription factors by binding to and activating specific sequences in promoters of host genes that may contribute to disease development, or trigger defense (Christian et al., 2010). A TAL effector is structurally composed of an N-terminal secretion signal, a central repeat region (CRR) with nearly identical direct repeats of 33–35 amino acid residues followed by a truncated half repeat with 20 amino acids, and a C-terminal region with nuclear localization signals and an acidic activation domain. TAL effectors, including the CRR, are highly conserved, except at positions 12 and 13 in each repeat, which are together known as the repeat-variable-diresidue (RVD). The RVD of each repeat determines its base specificity, such that the sequence of RVDs across the CRR determines the overall DNA sequence specificity of the protein (Boch et al., 2009; Moscou and Bogdanove, 2009). Importantly,

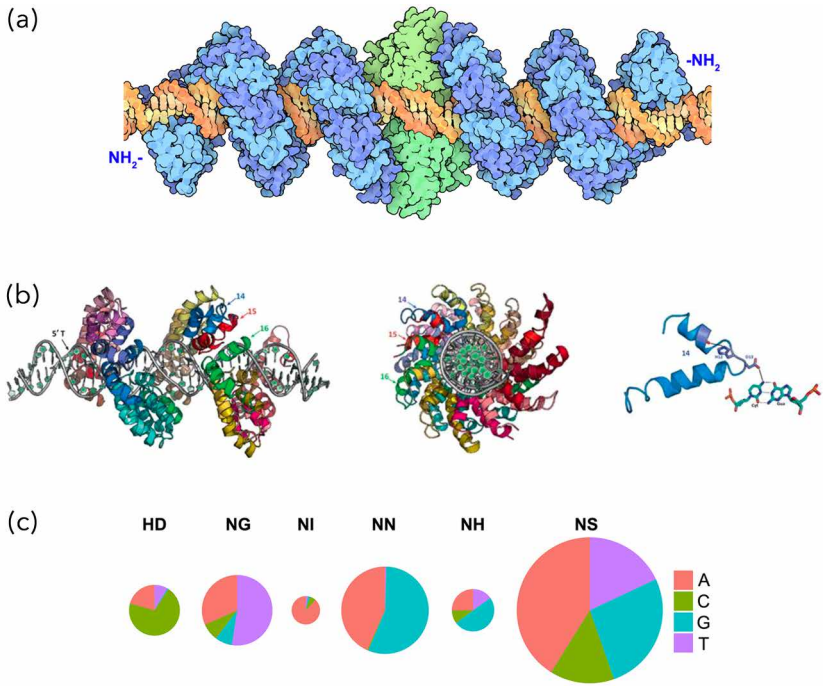

Figure 1 TALEN structure and mechanism of DNA recognition: (a) A TALEN drawn by David Goodsell from PDB files 1FOK (nuclease/DNA) and 3UGM (TALE/DNA); CC license 3.0. The paired TAL effector-*Fok*I fusion proteins bind the DNA in tail-to-tail orientation. The TAL effector domains are blue, the *Fok*I domains are green, and the DNA is orange; (b) Ribbon cartoon views of TAL effector PthXo1 bound to DNA. Left to right: the full complex viewed along the long axis of the DNA, an end view of the complex, and a zoomed view of a single repeat, with the RVD 'HD' interacting with cytosine in the plus strand of the DNA. Images reprinted with permission from Bogdanove et al. (2018); (c) Base specificities and relative binding contributions of commonly used RVDs, determined by Miller et al. (2011). Data were generated by measuring relative binding in an ELISA of a test TAL effector in which one repeat in the middle of the CRR was substituted with a repeat harboring the indicated RVD, on DNA targets with an A, G, C, or T at the position corresponding to that repeat. The slices of each pie represent the ELISA signal for the TAL effector with that RVD on the target with that base, and the overall size of the pie represents the sum of those relative binding affinities.

TAL effectors bind efficiently only when the CRR-defined DNA sequence is preceded by a 'T'. A 'C' is sometimes tolerated, but is thought to reduce binding affinity (Yu et al., 2011; Miller et al., 2011; Sun et al., 2012). Immediately N-terminal to the CRR, a 'cryptic' repeat that does not contain RVDs appears to contribute to the specificity for this initial base, in a structurally complex way (Doyle et al., 2013). Each repeat in a TAL effector comprises two alpha helices, with the RVD at the interhelical loop, and the overall protein wraps the DNA as a super helix, tracking the major groove, with the RVDs of each repeat oriented

inward for base-specific contacts (Mak et al., 2012; Deng et al., 2012) (Fig. 1b). The most common RVDs in nature, HD, NI, NG, NN, NS, and N* (missing the second amino acid of the RVD), respectively specify C, A, T, G/A, A/T/G/C and C/T (Moscou and Bogdanove, 2009; Boch et al., 2009), and the less common RVD NH specifies G (Streubel et al., 2012). For engineering, this small set thus provides RVDs specific for each of the four bases, two dual specificity RVDs, and a wildcard (Fig. 1c). As further detailed in this chapter later, generating a TALEN with the desired DNA sequence specificity is a simple matter of assembling a sequence of repeats with the necessary RVDs selected from among this set into the CRRs of the respective monomer constructs. Specificity and affinity contributions for all 400 possible RVDs have been characterized (Miller et al., 2015; Yang et al., 2014) and offer even greater potential for specificity tuning.

3 How have TALENs been used?

The precision, efficiency, and ease of use of TALENs offer great potential for a wide range of applications to accelerate crop improvement. TALEN-based genome editing is capable of yielding different types of mutations in target sequences, including targeted gene knockout, replacement, insertion, and large deletions and chromosomal rearrangements (Gurushidze et al., 2014; Sawai et al., 2014; Forner et al., 2015; Kelliher et al., 2017; Li et al., 2016b; Butler et al., 2016; Budhagatapalli et al., 2015; Wendt et al., 2013; Christian et al., 2013), and since the first two proof of principle reports of TALEN-based editing in *Arabidopsis* protoplasts and *Nicotiana benthamiana* leaves in 2011 (Cermak et al., 2011; Mahfouz et al., 2011), TALENs have been used in a wide variety of plant species including *Arabidopsis*, barley, *Brachypodium*, Brassica, maize, *N. benthamiana*, potato, rice, sugarcane, soybean, tobacco, tomato, and wheat (Table 1). TALENs have been applied in plants both to investigate gene function and to create new varieties with desirable traits such as enhanced nutritional content and biotic or abiotic stress tolerance (Table 1; reviewed also in (Zhang et al., 2017; Malzahn et al., 2017; Zhang et al., 2018)).

Most TALEN-mediated genome editing studies have been NHEJ-based (see Table 1). However, precise, HDR-based genome editing using TALENs has been successfully implemented in *N. benthamiana* and tobacco (Zhang et al., 2013; Mahfouz et al., 2011), barley (Budhagatapalli et al., 2015), potato (Forsyth et al., 2016; Butler et al., 2016), rice (Li et al., 2016b), and tomato (Cermak et al., 2015). To date, approximately 60 genes including protein-coding and non-protein-coding genes have been targeted for editing using TALENs in plants. Gene-regulatory regions have also been targeted by TALENs in some studies (Li et al., 2012; Holme et al., 2017). As shown in Table 1, a large portion of published studies is proof of concept and method development studies, which helped to optimize TALENs in plants (e.g. Zhang et al., 2013).

Index

2Blades Foundation 5
4CL. *see* 4-coumarate:CoA ligase (4CL)
4-CoAS. *see* Arabidopsis 4-*coumaryl synthase* (4-CoAS) intron
4-coumarate:CoA ligase (4CL) 442, 448–450

AAV. *see* Adeno-associated viral (AAV) plasmids
ABE. *see* Adenine base editor (ABE)
Acetolactate Synthase (*ALS*) gene 136, 137, 208, 334, 352, 423, 425, 428
Acetyl-coenzyme A carboxylase gene 137
ACT-PCR. *see* Annealing at critical temperature (ACT)-polymerase chain reaction (PCR)
Adcy3. *see Adenylyl cyclase 3* (*Adcy3*)
Addgene repository 177
Adenine base editor (ABE) 69–70, 125, 132, 151, 241
Adeno-associated viral (AAV) plasmids 76
Adenylyl cyclase 3 (*Adcy3*) 74
ADP ribosylation 39
AGEseq 238
Agrobacterium-mediated transformation 201, 203–205, 213, 214, 385, 386, 408
Alcobaca (*ALC*) gene 431
AlleleAnalyzer 444
Allene Oxide Cyclase (*AOC*) gene 426
ALS. *see Acetolactate Synthase* (*ALS*) gene
Alternative non-homologous end joining (a-NHEJ) 30–33, 37–38, 40
American Seed Trade Association 285
Amplicon sequencing 238
a-NHEJ. *see* Alternative non-homologous end joining (a-NHEJ)
Annealing at critical temperature (ACT)-polymerase chain reaction (PCR) 234–235
Annealing temperature 235
ANT1. *see Anthocyanin Mutant 1* (*ANT1*)
Anthocyanin 2 gene 426
Anthocyanin Mutant 1 (*ANT1*) 105, 107, 430
AOC. *see Allene Oxide Cyclase* (*AOC*) gene
Application programming interface (API) 162
Arabidopsis 7, 31–33, 36–38, 43, 44, 107
Arabidopsis 4-*coumaryl synthase* (4-CoAS) intron 65
Ascorbate peroxidase/4-coumarate 3-hydroxylase 450
Automated image collection 255

BADH2 12
Barley 329, 335–336
  overview 325–328
  plant development and performance 330–333
  resistance against plant pathogens 333
  targeted mutagenesis 326
    expression systems used for 328
    for molecular farming purposes 333–334
Barley Stripe Mosaic Virus 206
Barley Yellow Striate Mosaic Virus (BYSMV) 113
Base editing in plants 126, 128–129, 138–139
  adenine 132
  base editors application 136–137
  cytosine 127, 130–132
  mammalian systems 125, 127
  off-target 135–136
  overview 123–125
  protospacer adjacent motif (PAM) accessibility 132–134
  sequence diversification 134–135
Basic helix-loop-helix (bHLH) genes 405

BBM protein 208–211, 360–361, 363, 451
Bean Dwarf Mosaic Virus (BDMV) 102
Bean yellow dwarf virus (BeYDV) 77, 99, 102, 107, 109, 206
Beet necrotic yellow vein virus (BNYVV) 77, 113, 205
Beet severe curly top virus (BSCTV) 78
Beta carotene hydroxylase 2 (*CrtR-b2*) 429
Beta-galactanase (*TBG4*) 431
BeYDV. *see* Bean yellow dwarf virus (BeYDV)
bHLH. *see* Basic helix-loop-helix (bHLH) genes
Bio-Innovation Report 285
Biolistics. *see* Particle bombardment
BioNano 213
Biosafety Clearinghouse mechanism 288
Biosafety laws 284, 286, 289
Blasticidin deaminase (*BSD*) marker 70
BnTT8s gene 405
BNYVV. *see* Beet necrotic yellow vein virus (BNYVV)
BoMS1 gene 405
BoSRK3 gene mutation 405, 408
Brassica plants 406–407, 411, 413–414, 416
  gene disruption 405, 408
  generating mutants by CRISPR/Cas9 toolbox 408–410, 412, 415
  homologous recombination-based editing 408
  overview 403–405
BsaI enzyme 186, 188
BSCTV. *see* Beet severe curly top virus (BSCTV)
BSD. *see* Blasticidin deaminase (*BSD*) marker
BsmBI enzyme 186–188
BYSMV. *see* Barley Yellow Striate Mosaic Virus (BYSMV)

Cabbage Leaf Curl Virus (CaLCuV) 102, 206
Caffeic acid 450
Caffeoyl shikimate esterase 450
CaLCuV. *see* Cabbage Leaf Curl Virus (CaLCuV)
Calyxt 5
CaMV35S. *see* Cauliflower Mosaic Virus 35S (CaMV35S) promoter
Capsid/coat protein (CP) 101, 110
Carbon nanodots 335
Carotenoids 429
Cartagena Act (2004) 295, 297
Cartagena Biosafety Protocol (CBP) 279, 286–288, 295, 299, 315
Cas9-gRNA RNP complex 80, 202, 233, 353
Cas9/sgRNA complex 150, 151, 241
Cas-Analyzer 238
Cas-OFFinder 160–162, 240
CAST. *see* CRISPR-associated transposase (CAST)
Cauliflower Mosaic Virus 35S (CaMV35S) promoter 249, 256–257, 265, 430
CBD. *see* Convention on Biological Diversity (CBD, 1992)
CBE. *see* Cytosine base editor (CBE)
CBP. *see* Cartagena Biosafety Protocol (CBP)
CCT motif family (CMF) 332
Cellectis 5
CENH3 protein 355
Central repeat domain (CRD) 65
Central repeat region (CRR) 5–7, 12
Chemical mutagens and mutagenesis 65, 134, 282, 284, 294
ChIP-chip 254
ChIP-exo 254
ChIP-seq 254
Chloroplasts 332
Chlorsulfuron 208
Chromatin immunoprecipitation 254
Chromosomal rearrangements (CRs) 36–38
CIRCLE-seq 354
CIYVV. *see* Clover yellow vein virus (CIYVV)
CKX. *see* Cytokinin oxidase/dehydrogenase (CKX)
Classic/canonical non-homologous end joining (c-NHEJ) 30–33, 37–39, 66
CLEAVE-seq 354
Clover yellow vein virus (CIYVV) 137
Clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9 (CRISPR/Cas9)-mediated genome editing 3, 28, 48, 65–69, 75, 81, 106–107, 109, 112, 124, 147, 175, 198–199, 213–215, 280, 332, 343, 364, 396–398, 404, 422, 441
  alternatives to DNA delivery 207–208
  co-bombardment of *Bbm/Wus* genes and reagents 360–364
  components delivery into plant cells 196–197
  delivery into intact tissues 201
    agrobacterium 203–204
    electroporation 202–203
    nanoparticles 206
    particle bombardment 202
    viral-based replicons 205–206

delivery into single cells 200–201
empowered characterization of duplicated genes 447–450
in maize 351–355
overview 195–196
planta gene targeting in tomato 106–107
ribonucleoprotein complexes application 426–427
for sorghum protein improvement 382–396
transformation efficiency and extending genotype range 208–211
transformation in nontraditional explants 211–212
virus-mediated genome editing 112–113
CLV3/embryo-surrounding region (*SlCLE*) 432
CMF. *see* CCT motif family (CMF)
CMS-C. *see* Type C cytoplasmic male sterility (CMS-C)
c-NHEJ. *see* Classic/canonical non-homologous end joining (c-NHEJ)
CO. *see* Crossover (CO) products
Compact TALEN (cTALEN) 13
Comparative genomic sequence analysis 254
COMT 12
CONABIA. *see* National Advisory Commission on Agricultural Biotechnology (CONABIA)
Condensed tannins 448
Constitutive promoter 249
Conventional breeding techniques 64, 282, 283, 285, 311
Convention on Biological Diversity (CBD, 1992) 286, 287
'Coordinated Framework on Biotechnology' (1986) 300
Corteva Agriscience 209
CP. *see* Capsid/coat protein (CP)
CRD. *see* Central repeat domain (CRD)
CRE recombinase protein 202, 204
CRISPR activation (CRISPRa) 262, 263
CRISPR-associated transposase (CAST) 81
CRISPR/Cas9 system. *see* Clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9 (CRISPR/Cas9)-mediated genome editing
CRISPR/Cas-RNP complexes 426
CRISPResso 238
CRISPR-GA 238
CRISPR interference (CRISPRi) 262, 263
CRISPRitz tool 240
CRISPR RNA (crRNA) 46, 74, 75, 148, 241
Crossover (CO) products 43
CRR. *see* Central repeat region (CRR)
crRNA. *see* CRISPR RNA (crRNA)
CRs. *see* Chromosomal rearrangements (CRs)
CrtR-b2. *see Beta carotene hydroxylase 2* (*CrtR-b2*)
CsLOBP promoter 264
Csy4 processing signals 184
cTALEN. *see* Compact TALEN (cTALEN)
Cytokinin oxidase/dehydrogenase (CKX) 331
Cytosine base editor (CBE) 69, 125, 127, 130, 138, 139, 151, 241, 354, 355

dCas9. *see* Deactivated Cas9 (dCas9)
Deactivated Cas9 (dCas9) 151, 262, 263
Degenerate sequence decoding (DSD) 237
dHJ. *see* Double Holliday junction (dHJ)
Direct DNA-free genome editing 207
Direct transfer 175
Displacement loop (D-loop) 42
DNA methyltransferases 263
DNA microarray 254
DNA-PK 78, 79
DNASTAR 383
DNA structural features 250, 251
Domestication syndrome 433
Donor repair template (DRT) 74–76
Double Holliday junction (dHJ) 42
Double strand break (DSB) repair pathways 65, 69, 71, 81, 103, 150, 167, 232, 265, 326, 327, 343, 404
homologous recombination (HR) application in gene targeting 45–48 mechanisms 40, 42–45
non-homologous end joining (NHEJ) 29 genomic changes 32–40 mechanisms 30–32
overview 27–29
perspectives 48
Double-stranded DNA (dsDNA) 71, 72, 98, 195, 327, 359
DRT. *see* Donor repair template (DRT)
DSB. *see* Double strand break (DSB) repair pathways
DSB site-specific (DS) primer 235
DSD. *see* Degenerate sequence decoding (DSD)

DSDecode 237, 410
dsDNA. *see* Double-stranded DNA (dsDNA)
DsRED2. *see* Red fluorescence protein (*DsRED2*) gene
DuPont Pioneer 381

ECJ. *see* European Court of Justice (ECJ)
EFSA. *see* European Food Safety Agency (EFSA)
Electroporation 200–203
Elongation factor-1α (*SLEF1 α*) gene 425
EMS. *see* Ethyl methanesulfonate (EMS)
Endo-N-Acetyl-b-D-Glucosaminidase (*ENGase*) 333
Enhanced sgRNA (esgRNA) 69
Environmental Protection Agency (EPA) 283, 299, 300, 302
EPA. *see* Environmental Protection Agency (EPA)
esgRNA. *see* Enhanced sgRNA (esgRNA)
ESTs. *see* Expressed sequence tags (ESTs)
Ethametsulfuron 211
Ethyl methanesulfonate (EMS) 27, 396, 403
EU Council of Ministers 295
EU Directive 2001/18 294
European Commission 294, 295
European Court of Justice (ECJ) 294, 295, 416
European Food Safety Agency (EFSA) 294
Exon-intron junctions 444
Expressed sequence tags (ESTs) 445

FAD. *see* Fatty acid desaturase (*FAD*)
Fast-Flowering Mini-Maize A (FFMM-A) 360, 361
Fast Ligation-based Automatable Solid-phase High-throughput (FLASH) assembly 14, 15
Fatty acid desaturase (*FAD*) 12
FDA. *see* Food and Drug Administration (FDA)
Federal Food, Drug and Cosmetic Act (FFDCA) 302, 303
Federal Fungicide, Insecticide and Rodenticide Act (FIFRA) 302
FFDCA. *see* Federal Food, Drug and Cosmetic Act (FFDCA)
FFMM-A. *see* Fast-Flowering Mini-Maize A (FFMM-A)
FIFRA. *see* Federal Fungicide, Insecticide and Rodenticide Act (FIFRA)
FLASH. *see* Fast Ligation-based Automatable Solid-phase High-throughput (FLASH) assembly
FlashFry 240
FLP-mediated site-specific recombination 204, 209
*Fok*I domains 5, 13
FoMV. *see* Foxtail Mosaic Virus (FoMV)
Food and Drug Administration (FDA) 300, 302–303
FORECasT 168
Foxtail Mosaic Virus (FoMV) 113
Friends of the Earth report 285
FRIGIDA 12

Gateway cloning. *see* Type IIS REs
GB. *see* GoldenBraid (GB) modular cloning system
G-box element 251
GB software 186, 188
GBSS. *see Granule Bound Starch Synthase* (*GBSS*) gene
Geminiviral replicons (GVRs) 98–103
  based editing strategies 108–109
  deconstruction 101–102
  delivery and expression 102–103
  DNA replicons for plant genome engineering 104–105
  geminiviruses 98–99
  gene targeting in cereals with Wheat Dwarf Virus (WDV) 107–108
  high-frequency, precision modification of tomato genome 105–106
  mastrevirus genome structure, function, and replication 99, 101
  planta gene targeting in tomato 106–107
Gene functional analysis 249, 259
Gene targeting (GT) 45–48, 65, 67, 103, 105–108, 186–189, 191, 325, 326, 408
Gene technology 292
Gene Technology Act (2000) 292, 309
Genetically modified (GM) crops 64
Genetically modified organisms (GMOs) 279, 281, 284–286, 289–303, 308, 316, 317, 378
Genetic engineering 285, 301
Genome-edited mutants screening 242–243
  annealing at critical temperature (ACT)-polymerase chain reaction (PCR) 234–235
  high-resolution melting (HRM) analysis 234
  mismatch cleavage assay 232–233
  next generation sequencing (NGS) 238–240

off-target effects detection 240–241
overview 231–232
polymerase chain reaction (PCR)/restriction enzyme (RE) assay 232
polymerase chain reaction (PCR)/ribonucleoprotein (RNP) 233–234
Sanger sequencing chromatograms 235–237
Genomic rearrangements (GR) 68–69
GFP. *see* Green fluorescent protein (*GFP*)
Gibson assembly 177
Gif1. *see Grf-Interacting Factor1* (*Gif1*)
GL2 gene 12
GM crops. *see* Genetically modified (GM) crops
GMO Directive 2001/18/EC 416
GMOs. *see* Genetically modified organisms (GMOs)
GoldenBraid (GB) modular cloning system 177–178, 184, 186–189, 191
Golden Gate cloning technology 14, 15, 177
Gold-functionalized mesoporous silica 206
GR. *see* Genomic rearrangements (GR)
Granule Bound Starch Synthase (*GBSS*) gene 332, 427
Green biotechnology 28
Green fluorescent protein (*GFP*) 70, 107, 108, 202, 255, 335
Green Revolution 64
Grf4. *see Growth-Stimulating Factor4* (*Grf4*)
Grf-Interacting Factor1 (*Gif1*) 209
gRNA. *see* Guide RNA (gRNA)
gRNA/Cas9. *see* Guide RNA (gRNA)/Cas9 constructs
Growth-Stimulating Factor4 (*Grf4*) 209
GT. *see* Gene targeting (GT)
GT-Scan 160
Guide RNA (gRNA) 4, 66, 73, 76–77, 135, 326, 378, 442, 444
Guide RNA (gRNA)/Cas9 constructs 178, 190
CRISPR/Cas9 cloning basics 179–181
expanding activities 182–184
modular cloning tools 177, 179
requirements 182
multiple gene targeting in tobacco 186–189, 191
overview 175–177
polycistronic strategy 184–186
GUS gene 43, 112, 203
GVRs. *see* Geminiviral replicons (GVRs)
Hazardous Substances and New Organism Act (1996) 298
HDHL. *see* High digestibility and high lysine (HDHL)
HDR. *see* Homology-directed repair (HDR)
HEK293T cells 74
Herbicide tolerance 136–137
High Court of New Zealand 299
High digestibility and high lysine (HDHL) 392, 396
High-resolution melting (HRM) analysis 234
High-throughput sequencing. *see* Next generation sequencing (NGS)
High-throughput tracking of mutations (Hi-TOM) 238–241
HiScribe T7 In Vitro Transcription Kit 233
Histone modifiers 263
HITI. *see* Homology-independent targeted integration (HITI)
Hi-TOM. *see* High-throughput tracking of mutations (Hi-TOM)
HMEJ. *see* Homology-mediated end joining (HMEJ)
Homologous recombination (HR) 29, 40, 42–45, 47, 48, 68, 72, 408, 425
application in gene targeting 45–48
mechanisms 40, 42–45
Homology-directed repair (HDR) 7, 65, 66, 68–72, 75, 76, 78, 80, 97, 98, 103–104, 107, 108, 124, 196, 264, 326, 327, 335, 343, 350, 358–359, 404, 425
Homology-independent targeted integration (HITI) 75–76
Homology-mediated end joining (HMEJ) 72
HPT. *see Hygromycin phosphotransferase* (*HPT*) gene
HR. *see* Homologous recombination (HR)
HRM. *see* High-resolution melting (HRM) analysis
Hygromycin phosphotransferase (*HPT*) gene 68

Illumina HiSeq platform 239
IME. *see* Intron-mediated enhancement (IME)
inDelphi model 168, 169
Indirect encoding 175–176
Inducible promoter 249
Innovative Biotechnology 289
In planta methods 47, 197
Insertion-based genome edits 81–85
DNA fragments integration

donor repair templates with only the left homology arm 74–75
donor repair templates with two long homology arms 70–72
donor repair templates with two short homology arms 72–74
homology-independent targeted integration (HITI) 75–76
precise integration into the target chromosome (PITCh) 75
RNA donor repair templates 76
single-stranded donor repair templates 76
efficacy and accuracy of knock-ins
Cas9 variants 79–80
gRNA structure 76–77
NHEJ-mediated pathway 78–79
virus-based vectors 77–78
knock-ins outcomes 66
gene insertion and substitution 67–68
gene stacking 70
genomic rearrangements (GR) 68–69
targeted nucleotide exchange 69–70
overview 63–66
Insertion of transposable elements by gRNA-assisted targeting (INTEGRATE) 81
INTEGRATE. *see* Insertion of transposable elements by gRNA-assisted targeting (INTEGRATE)
Integrated Chip 14, 15
Intellectual property (IP) 4–5
Intron-mediated enhancement (IME) 258, 259
In vitro assembled RNP complexes 207
In vitro nucleic acid techniques 287
In vitro transcribed RNA molecules (IVTs) 207
Iowa State University 5
IP. *see* Intellectual property (IP)
IPK1 gene 349, 350
I-*Sce*I 28, 37, 43–45
Iterative Capped Assembly 14, 15
I-TevI DNA cleavage domain 13
IVTs. *see In vitro* transcribed RNA molecules (IVTs)

Japan Biosafety Clearinghouse 296

KU70/KU80 heterodimer 30, 31, 38

LbCpf1 endonuclease 74, 75
Lig4 78–79
Ligation-Independent Cloning Assembly 14, 15
Lignin biosynthesis 448, 449
Living modified organisms (LMO) 286–289, 295
lncRNAs. *see* Long non-coding RNAs (lncRNAs)
Long non-coding RNAs (lncRNAs) 430
Loop-out SD-MMEJ 33
LOX3 gene 12
Lycopene 430

MAFF. *see* Ministry of Agriculture, Forestry and Fisheries (MAFF)
Maize 344–348, 365–367
challenges in genome editing 355
chromosome rearrangements in biolistic transformation 358
DNA-free editing-RNPs 359–360
fidelity and efficiency 357–358
increasing HDR frequency for gene modification/insertion 358–359
regulatory landscape 360
tissue culture and transformation bottleneck 356–357
co-bombardment of *Bbm/Wus* genes and CRISPR/Cas9 reagents 360–364
overview 341–343
site-specific engineering
CRISPR/Cas 351–355
meganucleases 343, 349
transcription activator-like effector nucleases (TALENs) 350–351
zinc finger nucleases (ZFNs) 349–350
Maize Streak Virus (MSV) 99, 102
MAR1 gene 334
Marker-assisted selection breeding (MAS) 64, 396
MAS. *see* Marker-assisted selection breeding (MAS)
MegAlign 383
Meganucleases 28, 65, 326, 343, 349
MHLW. *see* Ministry of Health, Labor and Welfare (MHLW)
MHs. *see* Microhomologies (MHs)
Microchidia (MORC) proteins 333
Microhomologies (MHs) 30, 31, 34, 35
Microhomology-mediated end joining (MMEJ) 30, 31, 33, 39, 66, 72, 75, 150, 151, 167
MicroTom 427
Mildew resistance locus (MLO) 12

Ministry of Agriculture, Forestry and Fisheries (MAFF) 295, 296
Ministry of Environment (MOE) 295–297
Ministry of Health, Labor and Welfare (MHLW) 295, 297, 298
Mirus Bio 359
MLO. *see* Mildew resistance locus (MLO)
MMEJ. *see* Microhomology-mediated end joining (MMEJ)
Modern biotechnology 287, 289
MORC. *see* Microchidia (MORC) proteins
Movement protein (MP) 101, 109
MP. *see* Movement protein (MP)
MSV. *see* Maize Streak Virus (MSV)
MTL gene 12
Multiplexing 176
Mungbean yellow mosaic virus (MYMIV) 104

National Advisory Commission on Agricultural Biotechnology (CONABIA) 290
National Biosafety Management Agency (NBMA) 299, 300, 308
National Biosafety Management Agency Act (2015) 299
National Center for Biotechnology Information (NCBI) 385
NBMA. *see* National Biosafety Management Agency (NBMA)
NBMA Act (2019) 299
NBT. *see* New Breeding Techniques (NBT)
nCas9. *see* Nickase Cas9 (nCas9)
NCBI. *see* National Center for Biotechnology Information (NCBI)
NCO. *see* Non-crossover (NCO) product
NER. *see* Nucleotide excision repair (NER)
New Breeding Techniques (NBT) 289, 290
New England Biolabs 335
New Techniques Working Group (NTWG) 294
Next generation sequencing (NGS) 64, 213, 238–240, 249, 254
NHEJ. *see* Non-homologous end-joining (NHEJ)
Nickase Cas9 (nCas9) 69
Nicotiana benthamiana 109, 112, 113, 115, 212, 257
Nitrate reductase (*NR*) gene 427
Nitrogen-use efficiency (NUE) 64
NLS. *see* Nuclear localization signal (NLS)
NMD. *see* Nonsense-mediated mRNA decay (NMD)
Non-crossover (NCO) product 42, 43
Non-GMO Project 285
Non-homologous end-joining (NHEJ) 5, 7, 29, 36, 37, 46–48, 65, 68, 71, 72, 75, 149, 167, 261, 264, 327, 335, 343, 358, 404, 425
  genomic changes 32–40
  mechanisms 30–32
  mediated pathway 78–79
Nonsense-mediated mRNA decay (NMD) 450
NR. *see Nitrate reductase* (*NR*) gene
NtPDR6 transporter 423
NTWG. *see* New Techniques Working Group (NTWG)
Nuclear localization signal (NLS) 130, 179
Nucleotide excision repair (NER) 35
Nud1 gene 331
NUE. *see* Nitrogen-use efficiency (NUE)

Off-Spotter 240
OpenMTA initiative 177
Open-reading frames (ORFs) 99, 102
Optimal temperature 235
ORFs. *see* Open-reading frames (ORFs)

PA. *see* Phytic acid (PA)
PAGE. *see* Polyacrylamide gel electrophoresis (PAGE)
PAM. *see* Protospacer adjacent motif (PAM) accessibility
PAPhy_a gene 331
PAR. *see* Poly(ADP-ribose) (PAR)
PARP1. *see* Poly(ADP-ribose) polymerase 1 (PARP1)
Particle bombardment 200–202, 205, 207, 208, 213, 255, 385
PAT. *see Phosphinothricin acetyl transferase* (*PAT*)
PCBs. *see* Polychlorinated biphenyls (PCBs)
pcoCas9. *see* Plant codon-optimized SpCas9 (pcoCas9)
PCR/RE assay. *see* Polymerase chain reaction (PCR)/restriction enzyme (RE) assay
PCR/RNP. *see* Polymerase chain reaction (PCR)/ribonucleoprotein (RNP)
PDS. *see* Phytoene desaturase gene (*PDS*)
Pea early browning virus (PEBV) 77, 111, 113, 205
Pectate Lyase (*PL*) 431
PEG. *see* Polyethylene glycol (PEG)

pegRNA. *see* Prime editing guide RNA (pegRNA)
PG2a. *see Polygalacturonase 2a* (*PG2a*)
Phosphinothricin acetyl transferase (*PAT*) 65, 349, 350
Phytic acid (PA) 351
'Phytobricks syntax' 177
Phytoene desaturase gene (*PDS*) 113, 212, 351, 423
Phytoene synthase 1 (*PSY1*) 429
Phytosensing system 257
Phytozome 385
PIPs. *see* Plant-incorporated protectants (PIPs)
PITCh. *see* Precise integration into target chromosome (PITCh) system
PL. *see Pectate Lyase* (*PL*)
PLACE database 253
PlantCARE 253
Plant codon-optimized SpCas9 (pcoCas9) 71
Plant-incorporated protectants (PIPs) 302
Plant Protection Act 300
Pleiotropic genes 208
Pol III. *see* Polymerase III (Pol III)
Polyacrylamide gel electrophoresis (PAGE) 387
Poly(ADP-ribose) (PAR) 39
Poly(ADP-ribose) polymerase 1 (PARP1) 31, 39
Polychlorinated biphenyls (PCBs) 283
Polyethylene glycol (PEG) 200, 201, 351, 354, 426
Polygalacturonase 2a (*PG2a*) 431
Polymerase chain reaction (PCR)/restriction enzyme (RE) assay 232
Polymerase chain reaction (PCR)/ribonucleoprotein (RNP) 233–234
Polymerase III (Pol III) 445, 446
Polymerase theta (Pol θ) 32
Pol θ. *see* Polymerase theta (Pol θ)
'Position effect' 342
Post-transcriptional RNA silencing (PTGS) mechanism 111
Precise integration into target chromosome (PITCh) system 66, 75
Prime editing guide RNA (pegRNA) 265
PROGNOS 14
Promoter dissection 250, 251
PromPredict 250
Protein Targeting to Starch 1 (PTST1) 332
Protoplasts 200, 201, 386
Protospacer adjacent motif (PAM) accessibility 46, 47, 71, 117, 125, 132–134, 148, 149, 153, 168, 182, 265, 266, 326, 351, 357, 378, 385
PSY1. *see Phytoene synthase 1* (*PSY1*)
PTGS. *see* Post-transcriptional RNA silencing (PTGS) mechanism
PTST1. *see* Protein Targeting to Starch 1 (PTST1)

Quantitative trait loci (QTLs) 64

Random mutagenesis 342, 404
RdRP. *see* RNA-dependent RNA polymerase (RdRP)
REAL. *see* Restriction Enzyme and Ligation (REAL) assembly
REAL-Fast Assembly 14
Recalcitrant species 196
Red fluorescence protein (*DsRED2*) gene 379
Regulation of genome-edited crops 304–307, 316–317
  analysis in countries
    addressing organisms with multiple/complex genomic changes 314–315
    criteria for case-by-case review 309–311
    harmonization 312–313
    provisions are not science-based 311–312
    specific regulation for genome-edited products 308–309
    transparency is critical 313–314
  Cartagena Biosafety Protocol 286–288
  case studies 288
    Argentina 289–292
    Australia 292–293
    European Union 293–295
    Japan 295–298
    New Zealand 298–299
    Nigeria 299–300
    United States 300–303, 308
  genome editing techniques and focus on 280–284
  international debate 284–285
  overview 279
Regulatory Status Review (RSR) 302, 313
Repeat-variable-diresidue (RVD) 5–7, 13–14
REs. *see* Restriction enzymes (REs)
Restriction Enzyme and Ligation (REAL) assembly 14

Restriction enzymes (REs) 177
Reverse Transcription Quantitative PCR (RTqPCR) 431
RIN gene 430
RNA-aptamer strategies 184
RNA-dependent RNA polymerase (RdRP) 109
RNA interference (RNAi) 330, 382, 447
RNA-seq 445, 449, 450
Rolling circle replication 206
Roundup-Ready 2 Yield soybean 258
RSR. *see* Regulatory Status Review (RSR)
RTqPCR. *see* Reverse Transcription Quantitative PCR (RTqPCR)
RVD. *see* Repeat-variable-diresidue (RVD)

Saccharomyces cerevisiae 28
Sb-CENH3. see Sorghum *centromere-specific histone H3* (*Sb-CENH3*)
SbFT gene 382
SbGA2ox5 gene 382
SD-MMEJ. *see* Synthesisdependent MMEJ (SD-MMEJ)
SDNs. *see* Site-directed nucleases (SDNs)
SDSA. *see* Synthesis-dependent strand annealing (SDSA)
Second end capture 42
SECURE. *see* Sustainable, Ecological, Consistent, Uniform, Responsible, Efficient, (SECURE) rule
Self-pruning (*SP*) gene 432
Sequence analysis 237, 250, 251
Sequence-specific endonucleases 65, 325, 326
Sequence-specific nucleases (SSNs) 65, 103, 105, 111, 259, 266, 343, 364
sgRNAs. *see* Single-guide RNAs (sgRNAs)
Shootmeristemless (*STM*) 212
Short hairpin RNA (shRNA) 79
shRNA. *see* Short hairpin RNA (shRNA)
Single-guide RNAs (sgRNAs) 71, 73, 113, 115, 149–152, 154, 156–157, 159, 326, 351, 409, 410
  bioinformatics programs
    Breaking-Cas 160
    Cas-designer, Cas-OFFinder, Cas-Database, and Cpf1-Database 161–162
    CGAT 164–165
    CHOPCHOP 162–163
    CRISPOR 162
    CRISPRdirect 165
    CRISPR multitargeter 160
    CRISPRseek 165
    CRISRP-GE 158
    CRISRP-P and CRISRP-local 155, 158
    CRISRP-PLANT 158
    E-CRISP 163–164
  designing
    editing window for base editing 151
    guidelines 153, 155
    off-target prediction 149–150
    on-target efficiency prediction 150
    target activation and repression 151
    targeted deletion of DNA fragment 150–151
  overview 147–149
  predicting CRISPR-mediated genome editing 166–169
Single nucleotide polymorphisms (SNPs) 232, 427, 442, 445, 447
Single-nucleotide variants (SNVs) 135, 241
Single-strand annealing (SSA) 31, 37, 40, 43–45, 47–48
Single-strand breaks (SSB) 29, 33–35
Single-stranded DNA (ssDNA) viruses 98, 125, 327, 359
Single-stranded oligodeoxynucleotide (ssODN) 72–74, 80, 408
Single-walled carbon nanotubes 335
Site-directed nucleases (SDNs) 123, 124, 138, 292, 293
Site-specific nucleases (SSNs) 98, 103
Site-specific recombination 202
SlCLE. see CLV3/embryo-surrounding region (*SlCLE*)
SlClV3 gene 431, 432
SLEF1 a. see Elongation factor-1a (*SLEF1 a*) gene
SlMYB12 transcription factor 430
Snap-back SD-MMEJ 33
SNAP-RNP-repair template complex 335
SNGD-mediated gene editing 81
SNPs. *see* Single nucleotide polymorphisms (SNPs)
SNVs. *see* Single-nucleotide variants (SNVs)
Somatic embryogenesis 211
Sonchus Yellow Net Nucleorhabdovirus (SYNV) 116, 117
Sorghum 380, 384, 389–391, 394–395, 397–398
  advances of gene editing 379, 381–382
  CRISPR/Cas9 gene editing for protein improvement 382

Cas9/gRNA system delivery and regeneration of transgenic sorghum plants 385–386
digestibility and protein quality assessment 392
high lysine/high digestibility line selection and breeding strategies 392–393, 396
screening of stably edited, transgene-free lines 388, 392
sgRNA design of k1C gene family and Cas9/gRNA vector construction 383, 385
T7E1 assay and Illumina MiSeq analysis 386–388
overview 377–379
Sorghum *centromere-specific histone H3* (*Sb-CENH3*) 381
SP. *see Self-pruning* (*SP*) gene
Spatiotemporal promoter 249
SpCas9 (Cas9 from *Streptococcus pyogenes*) 46, 70, 79, 148, 149, 153, 179, 328
SPL. *see* Squamosa-promoter binding protein-like (SPL) family
Splicing process 258, 259
SPROUT 168
Squamosa-promoter binding protein-like (SPL) family 186
SSA. *see* Single-strand annealing (SSA)
SSB. *see* Single-strand breaks (SSB)
ssDNA. *see* Single-stranded DNA (ssDNA) viruses
SSNs. *see* Sequence-specific nucleases (SSNs); Site-specific nucleases (SSNs)
ssODN. *see* Single-stranded oligodeoxynucleotide (ssODN)
StALS1. *see Acetolactate synthase 1* (*StALS1*) gene
Stiff-Stalk 211
STM. *see Shootmeristemless* (*STM*)
Sulfonylurea (SU) 408
SunTag peptides 183, 263
SuRB gene 426
Surveyor nuclease 233
Sustainability Council of New Zealand Trust 299
Sustainable, Ecological, Consistent, Uniform, Responsible, Efficient, (SECURE) rule 301, 310, 312, 314, 315
SWEET14 sucrose transporter gene 12
Synthesisdependent MMEJ (SD-MMEJ) 33
Synthesis-dependent strand annealing (SDSA) 33, 40, 42, 43, 45, 47, 73
SYNV. *see* Sonchus Yellow Net Nucleorhabdovirus (SYNV)

T7 endonuclease 1 (T7E1) 233, 386–388
TALENoffer 14
TALENs. *see* Transcription activator-like (TAL) effector nucleases (TALENs)
Targeted promoter modification 266–267
approaches 261
base editing 264
chromatin modification and targeted methylation 263
prime editing 265
targeted integration 264–265
targeted mutagenesis and deletion 263–264
targeted transcriptional modification 262–263
leading introns and their roles 258–259
overview 247–249
plant promoters and *cis*-acting regulatory elements
identification 253–254
measuring activity using transgenic approaches 254–256
promoter region 249–251
roles 251–253
synthetic promoters 256–258
principles 259–261
TATA box 248, 250
TBG4. *see Beta-galactanase* (*TBG4*)
TCV. *see* Turnip crinkle virus (TCV)
TEB. *see* TEBICHI (TEB)
TEBICHI (TEB) 32
Template-directed mutations 237
Ternary pVir system 203, 204
TFBSs. *see* Transcription factor-binding sites (TFBSs)
TGMV. *see* Tomato Golden Mosaic Virus (TGMV)
TGS. *see* Third generation sequencing (TGS)
ThermoFisher Scientific 5
Third generation sequencing (TGS) 64
TIDE. *see* Tracking of indels by decomposition (TIDE)
TIDER. *see* Tracking of Insertions, Deletions and Recombination events (TIDER)
Tissue culture strategy 196
TKC. *see* Transgene killer CRISPR (TKC) system

TLPs. *see* Trait landing pads (TLPs)
Tobacco Mosaic Virus (TMV) 113, 205
Tobacco rattle virus (TRV) 77, 111, 113, 115, 205
ToLCV. *see* Tomato Leaf Curl Virus (ToLCV)
Tomatoes and Solanaceae family 435–436
  base editing 427–428
  CRISPR/Cas ribonucleoprotein complexes application 426–427
  de novo domestication 432–434
  DNA sequence insertion 425–426
  editing to affect fruit and tuber quality 429–432
  gene editing proof-of-concept investigations 423–425
  gene function modification for improved disease tolerance 428–429
  overview 421–423
  plant growth and development modifications 432
Tomato Genome Consortium (2012) 429
Tomato Golden Mosaic Virus (TGMV) 102
Tomato Leaf Curl Virus (ToLCV) 102
Tomato yellow leaf curl virus (TYLCV) 429
Tracer RNA (trRNA) 184, 187, 326
Tracking of indels by decomposition (TIDE) 237
Tracking of Insertions, Deletions and Recombination events (TIDER) 237
tracrRNA. *see* Transactivating crRNA (tracrRNA)
Trait landing pads (TLPs) 350
Transactivating crRNA (tracrRNA) 148
Transcription activator-like (TAL) effector nucleases (TALENs) 8–11, 16–18, 28, 65–68, 75, 103, 206, 280, 293, 299, 326, 331, 343, 350–351, 364, 404, 422
  consideration 3–4
  intellectual property landscape 4–5
  specificity tuning 4
  target long sequences 4
  use 7, 12
    architectures 12–13
    assembly principles and reagents 14–15
    design software 13–14
    workflow 19–20
  working 5–7
Transcriptional activator 184
Transcriptional units (TUs) 177, 179, 184, 187
Transcription factor-binding sites (TFBSs) 253, 254, 263
Transcription start site (TSS) 248, 250
Transcript-templated (TT)-HDR 76
TRANSFAC 253
Transgene killer CRISPR (TKC) system 388, 392
Transgenic plants 255, 256
Transient expression 255
TransIT®-2020 359
Translational fusion 184
Transposon tagging 342
trRNA. *see* Tracer RNA (trRNA)
TRV. *see* Tobacco rattle virus (TRV)
TSS. *see* Transcription start site (TSS)
TT-HDR. *see* Transcript-templated (TT)-HDR
Turnip crinkle virus (TCV) 333
TUs. *see* Transcriptional units (TUs)
TYLCV. *see* Tomato yellow leaf curl virus (TYLCV)
Type C cytoplasmic male sterility (CMS-C) 353
Type IIS REs 177

Ubiquitin 1 promoter 328
UGI. *see* Uracil glycosylase inhibitor (UGI)
ULtiMATE. *see* USER-Based Ligation Mediated Assembly of TAL Effector (ULtiMATE)
Ultra-performance liquid chromatographytandem mass spectrometry (UPLC-MS/MS) system 392
Uniform Mu program 342
Unit Assembly 14
United Nations 81, 286
University of Minnesota 5
Untranslated regions (UTRs) 444
UPLC-MS/MS. *see* Ultra-performance liquid chromatographytandem mass spectrometry (UPLC-MS/MS) system
Uracil glycosylase inhibitor (UGI) 69, 124, 125
U.S. Department of Agriculture (USDA) 300–302, 313–315
USER-Based Ligation Mediated Assembly of TAL Effector (ULtiMATE) 14
US National Academy of Sciences 280
U.S. Surgeon General 283
UTRs. *see* Untranslated regions (UTRs)

VIGS. *see* Viral-induced gene silencing (VIGS)
Vlnv gene 12
Viral-induced gene silencing (VIGS) 111, 422
Viruses for gene-editing reagents delivery
  DNA viruses 98

geminiviral replicons (GVRs)
based editing strategies 108–109
deconstruction 101–102
delivery and expression 102–103
DNA replicons for plant genome engineering 104–105
geminiviruses 98–99
gene targeting in cereals with Wheat Dwarf Virus (WDV) 107–108
high-frequency, precision modification of tomato genome 105–106
mastrevirus genome structure, function, and replication 99, 101
planta gene targeting in tomato 106–107
oveview 97
replicon vectors for homology-directed repair (HDR) 98, 103–104
RNA viruses 109
CRISPR/Cas9 system 112–113
expanding to other species 113, 115
modifications into viral vectors 110–111
multiplexed heritable gene editing and mobile single guide RNAs 115
nontransgenic genome modification 111–112
viral vectors 116–117
viral vectors for virus-induced gene silencing 111
VR-ACS1 gene 256
Vrs1 gene 330, 331

W568L mutation 426
WDV. *see* Wheat Dwarf Virus (WDV)
WGS. *see* Whole-genome sequencing (WGS)
Wheat Dwarf Virus (WDV) 99, 102, 107–109, 206
WHISKERS™ 349, 350
Whole-genome sequencing (WGS) 240
Woody perennial trees 451
CRISPR/Cas9 technology 447–450
gene-editing reagents 445–446
genome resource limitation 444–445
heterozygous genome challenge 442–444
overview 441–442
stability 446–447
World Bank 81
World Economic Forum 282, 285
WUS protein 208–212, 360–361, 363, 451

Xanthomonas 5
XRCC4-XLF complex 38
XRCC6 78, 79

Yeast-one hybrid screening system 254

ZFNs. *see* Zinc finger nucleases (ZFNs)
ZFPs. *see* Zinc finger proteins (ZFPs)
Zinc finger nucleases (ZFNs) 28, 65, 67, 68, 70, 71, 105, 111, 112, 123, 206, 280, 293, 294, 299, 326, 343, 349–350, 364, 404
Zinc finger proteins (ZFPs) 65, 254
ZmIPK 12
ZmTMS5 gene 352