Chapter from the book *Genetic Engineering - An Insight into the Strategies and Applications*

Downloaded from: http://www.intechopen.com/books/genetic-engineering-an-insight-into-the-strategies-and-applications

Interested in publishing with InTechOpen?
Contact us at book.department@intechopen.com
Abstract

Recently developed methods for genome editing, representing a major breakthrough in the field of genetic engineering, will enable researchers to produce transgenic plants in a more convenient and safer way. Double-strand breaks (DSBs) are triggered by synthetic nucleases that later induce DNA repair mechanisms known as nonhomologous-end joining (NHEJ) or homology-directed repair (HDR) in the presence of a donor DNA. Gene targeting (GT) was earlier demonstrated in rice and maize genomes by exploiting several genes (Acetohydroxyacid synthase, waxy, ALS, OS11N3 etc.), while zinc finger nucleases (ZFNs) were used to modify IPK1 gene in maize. Clustered regularly interspaced short palindromic repeats (CRISPR-CAS) system has been shown to be efficient for targeted mutagenesis in wheat that has a hexaploid complex genome, rice, maize, and recently in barley. The CRISPR system is considered as advantageous over previous approaches due to its easy use and efficiency, however, needs to be improved for high off-target effects.

Keywords: genome editing, TALENs, ZFNs, CRISPR-CAS, rice, maize, wheat, barley

1. Introduction

Genome editing refers to the ability to perform controlled changes in the genome using specific nucleases. The ability of a recombination initiation by inducing double-strand breaks (DSBs) is a breakthrough in efficient genome editing and engineering of plants. Site-directed mutagenesis and gene replacement have been possible by these mechanisms that will lead to crop improvement and progress in functional genomics studies. Cereals, on the other hand, represent an important group in agriculture as those directly supply main carbohydrate sources for human food and animal feeding, e.g., rice for Asia, wheat for the whole world, and maize for the Americas. Grass family (known as Poaceae) consists of agronomically important plants such as wheat, rice, maize, sorghum, oat, and barley whose grains have high nutritional value having a rich source of fibers, vitamins, and minerals. Substantial amount
of research has been conducted during the past two decades for the improvement of cereal varieties through conventional or molecular breeding, or a combination of both. Conventional methods such as hybridization, selection, and hybrid breeding have been applied and a large number of genes and quantitative trait loci (QTLs) for various traits have been tagged with molecular markers to apply marker-assisted selection (MAS) for trait improvement [1]. It is now possible to measure gene expression, and with the recent methods, obtaining knockout plants for different genes, which lead to an understanding of the roles and functions of genes and their effects under changing environments.

Genome modification studies have been launched in plants two decades ago with low-targeted integration frequencies [2]. By the discovery of nucleases inducing DSBs in specific loci, GT frequencies dramatically enhanced. In maize, acetohydroxyacid synthase genes (AHAS108 and AHAS109) were modified using chimeric RNA/DNA oligonucleotides (ONDs) with a frequency of $10^{-4}$, which was higher than spontaneous mutations and GT

<table>
<thead>
<tr>
<th>Plant</th>
<th>Explant</th>
<th>Transformation method</th>
<th>Genome-editing approach</th>
<th>Gene/locus</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maize</td>
<td>Embryogenic cell cultures</td>
<td>Whisker-mediated</td>
<td>ZFNs</td>
<td>$\text{Inositol pentakisphosphate 2-kinase (IPK1)}$</td>
<td>Shukla et al. [8]</td>
</tr>
<tr>
<td>Rice</td>
<td>Embryogenic cells</td>
<td>Agrobacterium tumefaciens</td>
<td>TALENs</td>
<td>OsITN3</td>
<td>Li et al. [9]</td>
</tr>
<tr>
<td>Wheat</td>
<td>Cell suspensions</td>
<td>Agrobacterium tumefaciens</td>
<td>CRISPR</td>
<td>$\text{Inositol oxygenase (INOX)}$ and $\text{phytoene desaturase (PDS)}$</td>
<td>Upadhyay et al. [10]</td>
</tr>
<tr>
<td>Rice</td>
<td>Callus</td>
<td>Particle bombardment</td>
<td>CRISPR</td>
<td>Chlorophyll A oxygenase 1 (CAO1) and Lazy I</td>
<td>Miao et al. [11]</td>
</tr>
<tr>
<td>Maize</td>
<td>Callus</td>
<td>Agrobacterium tumefaciens</td>
<td>GT by HR</td>
<td>Bar, gfp, npt II</td>
<td>Ayar et al. [12]</td>
</tr>
<tr>
<td>Barley</td>
<td>Immature embryos</td>
<td>Agrobacterium tumefaciens</td>
<td>CRISPR</td>
<td>HvPM19</td>
<td>Lawrenson et al. [13]</td>
</tr>
<tr>
<td>Wheat</td>
<td>Protoplasts, immature embryos</td>
<td>Polyethylene glycol, particle bombardment</td>
<td>CRISPR</td>
<td>TaGASR7</td>
<td>Zhang et al. [14]</td>
</tr>
<tr>
<td>Rice</td>
<td>Callus</td>
<td>Agrobacterium tumefaciens</td>
<td>CRISPR</td>
<td>OsERF922</td>
<td>Wang et al. [15]</td>
</tr>
<tr>
<td>Maize</td>
<td>Protoplast callus</td>
<td>Polyethylene glycol, Agrobacterium tumefaciens</td>
<td>CRISPR</td>
<td>Zmzb7</td>
<td>Feng et al. [16]</td>
</tr>
</tbody>
</table>

ZFNs: zinc finger nucleases; TALENs: transcription activator-like effector nucleases; CRISPR: clustered regularly interspaced short palindromic repeats (CRISPR/CAS system); GT: gene targeting; HR: homologous recombination.

Table 1. Genome-editing applications in cereals.
by homologous recombination (HR) [3]. GT experiments were conducted and improved by various studies in which different genes were targeted in maize, rice, wheat, and barley. A negative/positive selection approach was demonstrated for targeting the waxy gene of rice [4]. Gao et al. [5] used a sequence-specific meganuclease I-CreI for NHEJ-mediated targeted mutagenesis in liguleless1 locus of maize. HR-mediated targeting studies in agronomical traits such as herbicide tolerance have been the main object in model cereals [3, 6]. Besides conferring herbicide tolerance, genes that are difficult to mutate by conventional mutagenesis have been successfully targeted and analyzed for their putative functions, e.g., ROS1 of rice, which is associated with cytosine DNA demethylation and thus epigenetic modifications in plants [7].

By the advancement of zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and clustered regularly interspaced short palindromic repeats (CRISPR-CAS) system, these technologies were performed in Arabidopsis and tobacco, as models, and also cereals, as summarized in Table 1, to modify specific gene/locus. In this chapter, the major technologies for genome editing are described with an emphasis on the applications of cereal model species.

2. Principles of genome editing

The basis of genome editing relies on the formation of DSBs at specific loci and triggering DNA repair mechanisms. DSBs can be formed in eukaryotic cells by chemical and physical factors (reactive oxygen species, ionized-radiation, etc.) or by natural events like meiotic recombination. During the last decades, it was demonstrated that DSBs can be induced by synthetic nucleases and lately by the bacterial defence system CRISPR as well. A common feature of these synthetic nucleases is the combination of the bacterial type IIS restriction endonuclease FokI nuclease subunit with a synthetic DNA-binding domain. This combination results in a specific DNA-binding domain for target and a nonspecific DNA cleavage domain. Zinc finger nucleases (ZFNs), TALENs, and dCas9-Fok (hybrid of FokI nuclease subunit with deactivated Cas9) are all based on this principle. It should be noted that mutations generated by FokI-based nucleases show small deletions or small deletions with insertions (so-called “indels”).

The strategies for genome editing are based on the endogenous cellular processes related to DNA repair and recombination. It is well known that recombination occurs naturally during meiosis and in many cases, involves chromatin exchange between homologous sequences. Such a recombination is designated as homologous recombination (HR) and is the governing recombination type and DNA repair mechanism in lower organisms such as bacteria, yeasts, and moss [17]. Homologous recombination frequency in lower organisms such as yeast and the moss Physcomitrella patens can reach to over 10% or even 90% of transformants, respectively [17, 18].

DNA repair through homologous recombination is designated as homology directed repair (HDR). This pathway is initiated by a DSB in DNA, which is a result of DNA damage or an
endonuclease activity. In the presence of a donor DNA template and specific endonucleases, this pathway enables the replacement of a specific sequence. Therefore, one can use oligonucleotides such as triplex forming oligonucleotides (TFOs), short ssDNA or dsDNA donors such as T-DNA to induce HR. It should be noted that while T-DNA of Agrobacterium is infiltrated in the plant cell as a ssDNA coated by VirE2, it turns into dsDNA shortly after and probably integrates into the plant genome as dsDNA [19, 20]. While unassisted HR levels are very low, and it is highly induced by specific genomic DSBs [21]. Therefore, to induce gene replacement, one needs a nuclease/nickase to induce single specific DSB/nick and a donor DNA with homology arms to the genomic target sequence. The designed donor can then be used from a single known mutation of one base to a complete gene replacement and the integration of a new sequence into the genome.

When looking at DNA repair mechanisms, a nonlegitimate recombination or nonhomologous-end joining (NHEJ) is the dominant repair pathway in higher organisms such as flowering plants and humans. In NHEJ pathway, two broken ends of DNA were ligated together without the need of a homologous template (Figure 1). This pathway can be looked on as an “SOS” pathway, where the cell is “panicked” and quickly repair the damaged DNA with putative errors in the process. The NHEJ pathway is usually recognized with many errors and, therefore, is an excellent choice for gene disruption. NHEJ can achieve all editing objectives, i.e., mutations including small deletions or insertions [22], as well as gene insertion and gene replacement [19, 23]. However, while getting a mutation is certain, the mutations are completely random, and unlike the homologous directed recombination (HDR), there is no way to predict which mutation will occur and what will be the final result.

Gene insertion is a combination of single DSB with a supplied donor DNA. Here, we mimic the T-DNA integration by Agrobacterium, which is known to integrate randomly into an existing genomic DSBs [24]. Integration of supplied donor DNA either as T-DNA or simple

Figure 1. Two DNA repair mechanisms underlying genome-editing studies.
dsDNA (such as PCR product) into the desired location are increased by causing a specific DSB [24, 25]. However, the donor will incorporate also into many other genomic locations, and massive screening should be taken to exclude undesired integrations.

Gene replacement can be achieved by generating DSBs flanking the gene of interest and supplying a donor DNA. This may result in deletion and targeted insertion leads to a gene replacement [19].

In both NHEJ and HDR, the main challenge is screening and recognizing the relevant HDR event. When designing genome editing, these two DNA repair pathways (NHEJ and HDR) would be the main guidelines that should be considered. In general, sequence replacement can be achieved by HDR while mutations, deletions, and insertions can be achieved utilizing NHEJ (Figure 1).

3. Zinc finger nucleases (ZFNs)

Zinc finger nucleases (ZFNs) are synthetic endonucleases combining zinc finger DNA-binding domain with a nuclease subunit (typically FokI endonuclease). ZFNs have evolved from transcription factors harboring zinc-finger domain in their DNA-binding domain. ZFNs are built from C2H2 zinc-finger domains where each finger recognizes three nucleotides. Therefore, a 3-finger nuclease will bind to a DNA sequence of nine nucleotides. The first ZFN known as Zif268 is a combination between transcription factor Zif268 DNA-binding domain and FokI nuclease subunit [26]. FokI forms a dimer to cleave DNA thus two ZFNs’ different monomers should be designed and used to cleave a single target sequence.

Major problems in using ZFNs are low specificity resulting in genotoxicity [27–30] and high complexity leading to low success rates of the designed enzymes [31].

DSBs by ZFNs have been applied mainly as a proof of concept and for research [32–34] in model plants and thus, all genome-editing strategies were explored by this pioneering system. In maize, inositol pentakisphosphate 2-kinase (IPK1) gene was targeted by generating a panel of 66 ZFNs against five intragenic positions [8] (Table 1). IPK1 gene was chosen for its importance in phytate reduction as an agronomic and ecological trait. Sequencing of genomic PCR products confirmed that addition of PAT gene conferring the herbicide tolerance into IPK1 had occurred precisely in a homology-directed manner. A recent study was conducted to explore noncoding genomic regions suitable for site-specific integrations to ensure stability and high gene expression in rice, using ZFNs. As a result, 28 genomic regions including only one noncoding have been discovered for safe integration of ZFN constructs carrying a β-glucuronidase gene [35].

4. TAL effector proteins (TALENs)

Transcription activator-like effector nucleases (TALEN) are synthetic nucleases combining FokI nuclease subunit with DNA-binding domain composed of repeats. Repeat number may
vary and is typically between 16 and 30 forming a protein encoded by about 3.7 Kb open reading frame (ORF). Each repeat binds to a single nucleotide and is composed of 33–34 amino acids. Amino acids 12 and 13 are variable and known as repeat variable diresidue (RVDs). These variations enable the binding of different nucleotides, whereas NI for adenosine, HD for cytosin, NN for guanine, and NG for thymidine [36].

TALENs where evolved from the *Xanthomonas* AvrBs3 superfamily of type III effectors acting as transcription factors *in planta* [36, 37]. The different proteins in this family contain different number of DNA-binding repeats that govern the pathogen host range. Analysis of repeats/targets resulted with the discovery of a new set of DNA-binding domains that are simpler and more specific than the zinc finger sets used for ZFNs. The new repeat combination enables TALENs to have high target specificity and high DNA affinity, which results in both low genotoxicity and high genome-editing rates. TALENs are probably the most accurate systems for genome editing with high success levels but the system suffers from several drawbacks.

Similar to ZFNs, TALENs use FokI nuclease subunit working as a dimer and, therefore, two monomers should be designed for each genomic target. The resulted ORF size is huge and, therefore, cannot be used in viral vectors. Furthermore, size and the need to synthesize a new pair of enzyme for each genomic target may hinder the ability to edit several genomic targets.

TALEN-directed mutations were generated in *Os11N3* gene in rice, which is normally activated by TAL-effectors (named AvrXa7 or PthXo3) of a rice pathogen causing bacterial blight disease [9]. That study showed that TALENs can be successfully employed for the modification of a S gene promoter to prevent its induction by bacterial effectors. The authors discussed the possibility of editing multiple susceptibility genes to confer resistance to other forms of bacterial blight.

5. CRISPR-CAS system

CRISPR or clustered regularly interspaced short palindromic repeats is a bacterial defence mechanism against bacteriophages. Although usually this system is composed of a cascade of many different proteins, in *Streptococcus pyogenes*, most cascade proteins are provided as a single self-operating protein designated as Cas9.

The Cas9 has several functional characteristics including enabling binding of sgRNA, searching for complementary sequence, and cleaving the target sequence (HNH domain that cleaves the complementary DNA strand and RuvC domain cleaves the noncomplementary DNA strand). The most important is the PAM recognition domain that distinguishes the bacterial encoding RNA from the bacteriophage target sequence. In Cas9, this sequence requires NGG downstream to the targeted sequence. The Cas9 first binds the PAM sequence and then opens the DNA, allowing RNA/DNA hybridization or R-loop formation and then cleaves both DNA/RNA and ssDNA strands [38–40].

In 2013, several articles have been published reporting the plant genome engineering, using CRISPR system that five of them resulted in the generation of mutant plants with specific
targeted loci [41]. Besides the applicability of CRISPR, the use of protoplast cultures for transient expression assays and agroinfiltration of leaf tissues have been preferred due to their advantages. Upadhyay et al. [10] targeted inositol oxygenase (INOX) and phytoene desaturase (PDS) genes in the suspension cultures of wheat, which has a complex hexaploid genome (17 Gb). The authors reported that CRISPR system was simpler than ZFNs and TALENs with high efficiency even in large genomes at each of the multiple targeted locations. In the same year, targeted mutagenesis by CRISPR-CAS in Chlorophyll A oxygenase 1 (CAO1) and Lazy 1 genes were demonstrated in rice [11]. These genes were selected for their easily detectable phenotypes for screening, e.g., pale green leaves for CAO1 and tiller-spreading appearance for Lazy 1 gene, respectively. Finally, induction of heritable mutations (transmitted stably to T2 plants) by CRISPR has been shown in barley which is an important model with its diploid nature [13]. In another study, high activities of sgRNA were optimized using protoplast cultures of wheat and both protoplasts and immature embryos were tested as a broadly applicable system for genome editing [14]. Recently, insertion and deletion mutations were successfully introduced to a rice gene OsERF922, coding an ERF family transcription factor and resistance to rice blast was enhanced in the resulted plants [15].

6. Conclusions and future perspectives

There have been substantial efforts to develop efficient technologies for GT in plants including cereals. For this aim, synthetic nucleases, including the mitochondrial I-SceI from yeast and chloroplast I-CreI from Chlamydomonas reinhardtii, have been used for higher GT frequencies [5, 12, 42]. However, it was recently shown that CRISPR-CAS system can greatly facilitate the modification of targeted locus in rice, maize, wheat, and barley tissue cultures [13–16]. The primary application of genome-editing tools is obtaining knockout plants, and in time, the other applications that extend to crop improvement are expected, e.g., abiotic stress tolerance will be important for near future to resolve stress response and adaptation pathways [43]. For example, barley has been used for a long time in genomics studies as a highly adaptive and tolerant model for environmental stresses [44, 45].

Frequency of HDR in plants (Arabidopsis and tobacco) is typically $10^{-4}$–$10^{-5}$ [46], whereas gene replacement by HR in plants may be increased to $10^{-2}$ through transient expression of meganucleases, which induces double-strand breaks (DSBs) [2]. NHEJ levels as shown in the form of T-DNA integration are 3–15 times higher when compared to HDR events in plants and transgenes integrated into the correct site in about 1% of the transformants [23].

There are several considerations and limitations for the major genome-editing technologies. TALENs are considered as the most precise genome-editing system for today. This suggests not only hitting the correct genomic location but more importantly less cytotoxicity from off-targeting effect. The high precision enables targeting multiple targets with confidence. High efficiency in genome editing is translated to the amount of screened plants in order to reach the desired modified plant. ZFNs are considered to be less efficient than TALENs and Cas9 shows higher efficiency. Both ZFNs and TALENs have to be redesigned for each target, while CRISPR-based methods require redesign of RNA molecules. Therefore, CRISPR methods are
easier to employ and are more suitable for large genomic screenings or multiplex gene targeting. However, more evidence has raised Cas9 low specificity effect [47] and low homologous recombination ratio [48], whereas several approaches were taken to overcome these limitations of which dCas9-Fok might be the most promising [49–53], none addresses the size constraints presented by Cas9.

The huge size constrains of the current genome-editing tools prevent applying plant viral vector as genome-editing tool, thus the researcher should use meganucleases for these applications. In general, Cas9 and its derivative technologies would be sufficient for research of Agrobacterium transformable and regenerative plants. Nevertheless, the need for a more precise and smaller system exists, and we can expect that future technologies will answer these restrictions.

**Author details**

Dan M. Weinthal¹ and Filiz Gürel²*

*Address all correspondence to: filiz@istanbul.edu.tr

1 TargetGene Biotechnologies LTD, Rehovot, Israel

2 Department of Molecular Biology and Genetics, Vezneciler, Faculty of Science, Istanbul University, Istanbul, Turkey

**References**


[14] Zhang Y. et al. Efficient and transgene-free genome editing in wheat through transient expression of CRISPR/Cas9 DNA or RNA. Nat Commun. 2016; 7: 12617. DOI: 10.1038/ncomms12617


