Chapter 3

Gene Targeting and Genetic Transformation of Plants

Richard Mundembe

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/56335

1. Introduction

A broad definition of gene targeting includes any method that can lead to permanent site-specific modification of the genome [1], preferably with predetermined outcomes. More specifically, gene targeting is the alteration of a specific DNA sequence in an endogenous gene at its original locus in the genome, and often refers to the conversion of the endogenous gene into a designed sequence [2]. Rapid developments in the field of gene targeting, and the potential of the technology to revolutionalise genomics and plant biotechnology in particular has led to the adoption of this broad definition, over earlier definitions such as that by [3] and [4] that restricted gene targeting to homologous recombination mechanisms.

While gene targeting does not necessarily lead to marker-free, vector backbone-free transformation, gene targeting certainly brings these desired outcomes of plant transformation research closer. Such marker-free, vector backbone-free plants will be truly and precisely engineered plants, and might actually be non-transgenic, depending on the source of the sequences used. Gene targeting in *Drosophila*, mice and yeast is now more or less routine [5]. Transgenic organisms for use in research are 'made-to-order' via gene targeting and are sold by commercial companies. Gene targeting in animals is accomplished via homologous recombination (HR). However, the same cannot be said of plants. Approaches adapted from gene targeting in yeast, insect and animal models have failed to give comparable results in plants mainly because the predominant mechanism of recombination in somatic cells of plants is not HR, but is non-homologous end joining, NHEJ, also known as illegitimate recombination [6].

Double-stranded breaks in plant genomic DNA are repaired either via HR or NHEJ [7]. Homologous recombination mechanisms involve linkage of DNA fragments to regions of identical sequence, such as the other member of the homologous partner, as template for accurate repair of the double stranded break. This mechanism is therefore only functional in the S/G2 phase of the cell cycle. Non-homologous end-joining mechanisms of recombination however are functional in all phases of the life cycle and do not require significant homology

to join two fragments of a broken DNA molecule. While HR has been very successful in insects and animals [8], it has remained unavailable for the manipulation of plant transformation; it is NHEJ that is useful for plant transformation.

Occurrence of single-stranded breaks on a DNA molecule does not normally pose a challenge to the plant genome because these can be repaired by ligation without change to the nucleotide sequence. Faithful strand replacement or nick translation may take place starting at the single-strand break, again with no changes to the nucleotide sequence.

Double-stranded breaks, however, have dire consequences if not repaired, or if repaired incorrectly. A double-stranded break effectively results in two fragments of the chromosome, and only one of the fragments might have a centromere to enable separation after cell division; the other fragment might be 'lost'. Also, if unprotected, the double-stranded breaks are exposed to the exonucleases of the cell and may be misconstrued as foreign and will therefore be degraded.

Living cells therefore need efficient mechanisms for detecting chromosomal double-stranded breaks and initiation of appropriate repair mechanisms for replication to be successful. The repair of double-stranded breaks takes place by one of two main pathways for double-stranded break repair: the HR pathway or the NHEJ pathway, or both. Coincidentally, these are the two mechanisms by which exogenous DNA may also integrate into a host genome [7].

2. Homologous and non-homologous recombination

Recombination evolved in nature to repair DNA damage that may occur during the cell cycle, and to generate diversity through meiotic recombination of genetic material which in turn has enabled sexually-reproducing eukaryotes to become extremely adaptive to their ever-changing environment and is partly responsible for their success on earth.

2.1. Homologous recombination

In homologous recombination (HR) a long and extended region of homology such as that found between sister chromatids is required for the two DNA molecules to line up adjacent to each other. There are many variations to this pathway, but the basics of two popular models are illustrated in Figure 1. A cellular protein, Spo I, may induce double-stranded breaks in the chromosome. These double-stranded breaks are repaired exclusively by HR using one of several possible homologous matrices: copied from elsewhere in the genome (ectopic HR), copied from the homologue (allelic HR), or copied from the same chromosome (intra-chromosomal HR) [6, 9].

Ectopic HR is a minor pathway, and was reported to be responsible for the repair of only one in 10 000 double-stranded breaks. In some of the cases, both homologous and non-homologous end-joining mechanisms were involved in repairing different ends of the same double-stranded break [6, 10]. Of the possible ectopic recombination models, the synthesis-dependent

http://dx.doi.org/10.5772/56335

strand annealing (SDSA) model is the one that is conservative and is consistent with these observations. Figure 1(a) below illustrates this model.

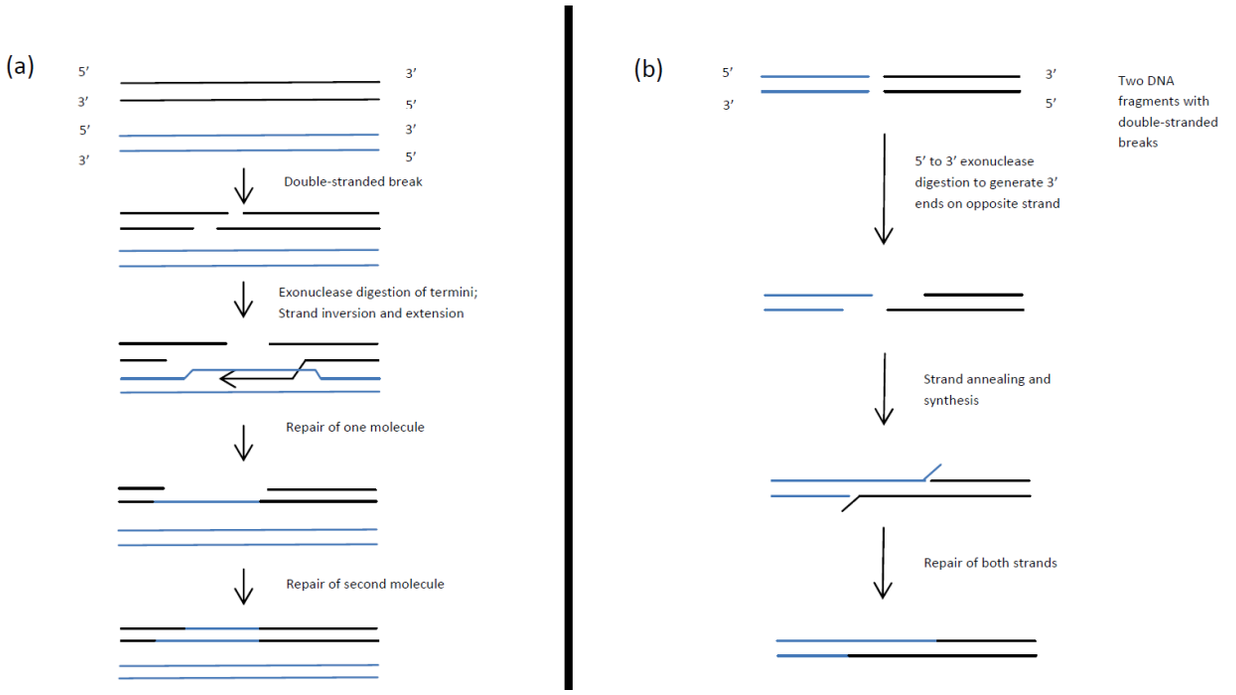

Figure 1. Models for double-stranded break repair mechanisms. (a) The synthesis-dependent strand-annealing (SDSA) for ectopic recombination. (b) The single-strand annealing (SSA) model. Polarity of the DNA molecules is shown on the first set of molecules only for simplicity. Modified from [7].

This model predicts that double-stranded break repair is not accompanied by crossing-over. Also, both perfect integrations into target sites by homologous recombination and imperfect integrations by HR are possible on one end of the target site. Integration by NHEJ is possible on the other end of the same double-stranded break of transgene, as well as ectopic integrations elsewhere in the genome, after copying of transgene sequences.

Allelic HR occurs during meiosis, to repair double-stranded breaks using sequences of the homologues in a process that involves formation of Holliday junctions to resolve into the crossover or gene conversion products. Allelic HR is not significant in somatic cells but is the classic HR that occurs in meiotic cells. In nature this essential process takes place during meiosis I to result in recombination for sexually reproducing species. Extensive lengths of homology (several hundreds or thousands of nucleotides) are required for this process, and ensures that recombination takes place between sister chromatids.

Intrachromosomal HR utilizes sequences close to the double-stranded break, on the same chromosome or on the sister chromatid (in G_2 stage only) as a matrix for repair. This can result in deletion as predicted by single-strand annealing (SSA) model (Figure 1b) or gene conversion as predicted by the conservative SDSA model depending on the structure of the chromosomal locus [11]. The SSA pathway was shown to be five time more efficient than the SDSA pathway [12]. SSA-like pathways have also been described for NHEJ.

2.2. Non-homologous end-joining recombination

The second pathway is non-homologous end-joining (NHEJ) pathway, also known as illegitimate recombination. It also requires some homology, albeit much reduced. This limited homology is required at the ends of the DNA strands on which the double-stranded breaks occur. The double-stranded breaks can be sticky ends or blunt ends. The homology present within the sticky ends may be sufficient for this mechanism, and the properly aligned ends will be ligated together. For blunt ends, binding of a specific protein complex, such as the Ku complex in mammalian cells, to the broken ends of the DNA limits nucleolytic degradation, and unlike HR repair, prevents exposure of single-stranded regions [8]. The bound protein may also function directly or indirectly to bring the DNA ends together for processing and ligation. Alignment of the termini by complementary micro-homologies of 1 – 4 nucleotides is usually required. The process might also require either limited unwinding or limited exonucleolytic digestion to expose the ends for alignment, and DNA polymerase to fill-in gaps. Single-stranded deletion of short segments at the 5'-end may expose single-stranded regions that will be used to search for homologies in the other DNA fragment, which will then form the basis of the alignment and repair [8]. The process of NHEJ is illustrated in Figure 2. The arrangement of chromosomal DNA into loops attached to a matrix that restricts the mobility of DNA promotes the re-joining of previously linked DNA ends [8].

NHEJ is the predominant pathway for double-stranded break repair in somatic cells of higher eukaryote, including plants. Simple ligation will result in junctions with no homology. Short stretches of homology may be a result of SSA-like mechanisms [13], while longer stretches might be from an SDSA copying of ectopic chromosomal DNA into the break [14].

NHEJ is also the mechanism by which transgene integration occurs following either *Agrobacterium*-mediated or direct transformation of plant cells. The integration sites are generally random, but transcriptionally active sites seem to be preferred.

When we consider the evolution of gene targeting research, HR pathways were initially considered the only route with potential to achieve this because of high levels of fidelity observed in HR during meiosis. The levels of homology involved in meiotic recombination are large and would make this approach unworkable for routine plant genetic engineering. The extent of homology required is extensive, and may elongate the transgenes required in plant transformation to impractical levels. Induction of double-stranded breaks on the DNA by exposure to X-rays or by transposon activity was shown to increase HR [15, 16]. Site-specific recombination systems therefore became a potential route to achieving gene targeting by HR, since they can introduce double-stranded breaks in DNA, and repair these in via an HR mechanism that utilizes shorter homologies.

The objective of many plant transformation research groups is to study genomics and generate improved crops. While transgenic plants produced for genomics study have little regulatory requirements since they are for contained use, transgenic plants for general release have to comply with governmental regulations and must also meet consumer acceptance. Gene targeting will make it easier for genetically modified plants to meet these requirements. A strategy for gene targeting that has been explored extensively by researchers is that of site-specific recombination.

http://dx.doi.org/10.5772/56335

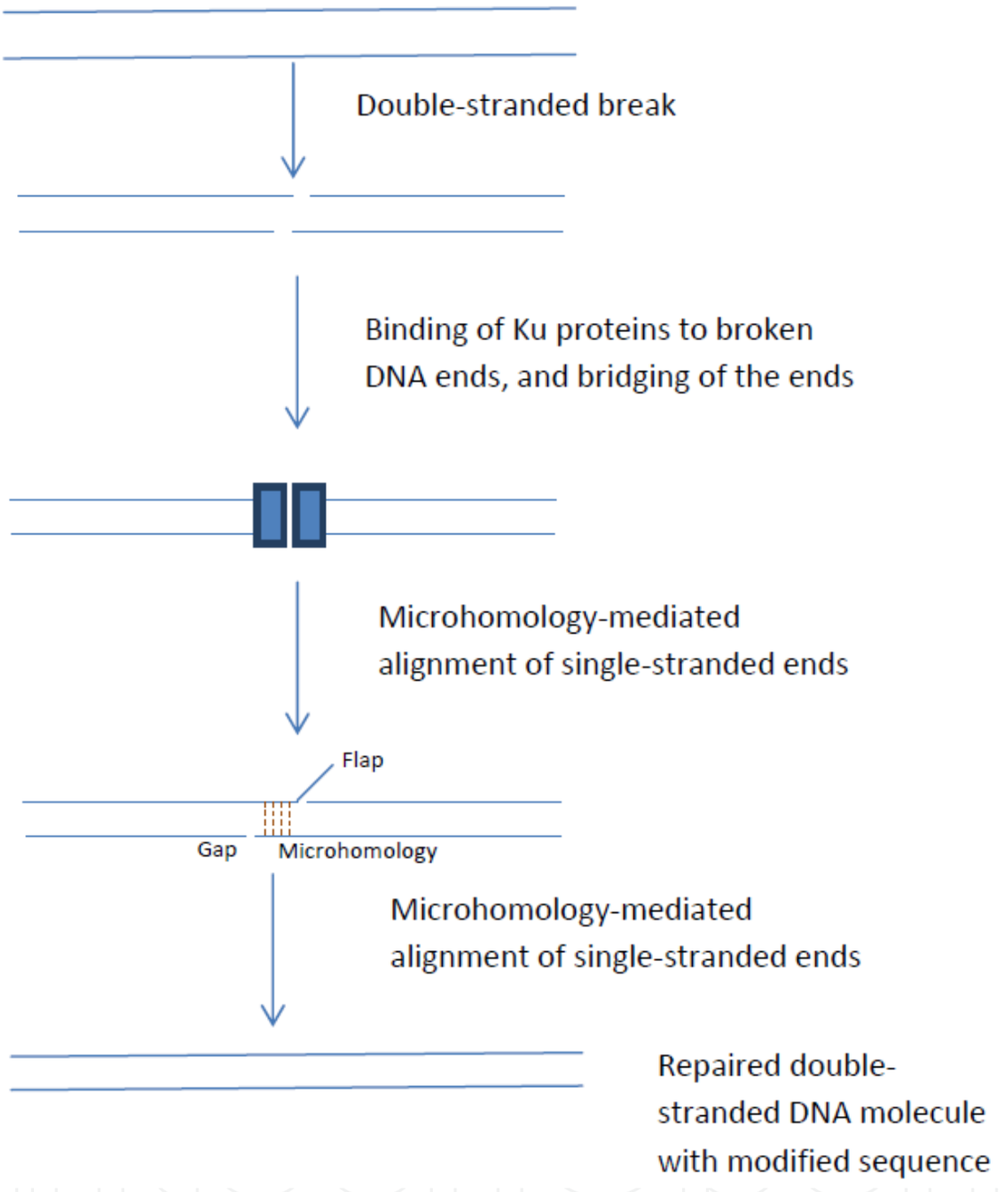

Figure 2. Model for repair of double-stranded breaks by non-homologous end-joining. Adapted from [8].

Site-specific recombination systems consist of a recombinase and donor sites. The recombinase is a protein that mediates a recombination reaction between a target site characterized by particular target sequence for that protein, and the donor site, also with a characteristic nucleotide sequence. In general, the results of the ensuing recombination reaction are excision, integration or inversion.

The site-specific recombination systems that can be utilized for gene targeting include the tyrosine family recombination system, the serine family recombination system and the newly

developed hybrid system consisting of zinc finger DNA sequence recognition motifs in combination with a rare-cutting restriction endonuclease. Each of these systems will now be considered in turn, and the potential to contribute to gene targeting discussed.

3. Tyrosine family recombination systems

The tyrosine family recombination systems include the Cre/*lox P*, FLP/*FRP*, λ integrase and variations thereof [17]. Also known as integrases, they use the hydroxyl group of the catalytic tyrosine for a nucleophilic attack on the phosphodiester bond of the target DNA, and function through a Holliday junction intermediate [17, 18]. Their function depends on the cofactors supplied.

Cre and FLP recombinases are the most popular members of the integrase family because they are simple and unrestrictive, requiring no auxillary factors other than their recombinase monomers and their cognate targets. Cre recombinase recombines 34 bp *lox P* sites in the absence of accessory proteins or auxillary DNA sequences [17, 19]. The FLP target site has been trimmed from the original 599 bp in the 2 μm yeast plasmid to 30 bp in *FRT* sites [17]. The wild-type *FRT* and *lox P* sites are unchanged by the recombination reaction, making the reaction reversible; and there are many different possible recombination intermediates in each case. This has however made it difficult to utilize the tyrosine family recombinase systems more extensively in vector construction, while the irreversible λ integrase is more popularly used in vectors [17].

4. Serine family recombinase systems

The Serine family recombinase systems such as ϕC31, Hin and Gin [17] are also known as the resolvases or invertases. They have a conserved serine residue that is used to create the covalent link between the recombinase and the DNA target site [18]. Serine family recombinases initiate strand-exchange by making double-stranded breaks at two sites in the DNA molecules. Each site of the double-stranded break is associated with a dimer of the recombinase, and the two dimers will come together bringing the two broken ends together and forming an active tetramer in a process that is elaborately controlled [20].

The general scheme for using site-specific recombination systems in gene targeting involves, first, the genetic engineering of the recombination target sites into the particular genomic location of the plant to be transformed. This can be achieved by standard transformation procedures followed by screening to identify transformation events in 'acceptable' locations. Transposon tagging has also been used with the recombination target sites incorporated within the transposon.

The second requirement is that the incoming transgene should have unique DNA sequences that constitute the donor sites. Finally, there should be a mechanism for expression or intro-

duction of the recombinase, to mediate the recombination reaction between donor and target sites. In this scheme, a second transformation experiment targets the genes into which the recombinase target sequences were integrated by the first transformation experiment.

Both *Agrobacterium*-mediated and direct gene transfer (bombardment, electroporation and PEG-mediated transformation) have been used for the initial transformation to introduce target and donor sites. The recombinase may be expressed constitutively, transiently or may be induced. Recombinase expression as well as stability of the transgene may vary [21].

These approaches were based on the need to improve homologous recombination at the target site. In these approaches, homology is limited to target and donor site compatibility for the particular recombinase being considered. With elegant engineering, site-specific recombination systems can be used to remove marker genes from transgenic plants before their commercialization. But the process is far from routine. Also, the footprint that remains on the chromosome is associated with genetic instability. The search for a better system continues, and that is why zinc finger nucleases are being considered.

5. Zinc finger nuclease and gene targeting

Zinc finger nucleases (ZFN) are artificial restriction endonucleases composed of a fusion between an artificial Cys_2His_2 zinc finger protein DNA binding domain and the cleavage domain of the *Fok* I endonuclease. The sequence-specific DNA binding domain of zinc fingers could be engineered to recognize a variety of specific DNA nucleotide sequences of the researcher's choice [22]. While the *Fok* I endonuclease activity is non-specific, the enzyme only functions when it forms a dimer, whose assembly will be guided by proper alignment of the two zinc finger monomers at the target site. Assembly of the ZFN therefore enables site-specific cleavage [3, 23]. The two zinc finger monomers are usually designed to flank a 5 – 6 bp long target sequence, within which *Fok* I cleavage will occur. The zinc finger domain itself is composed of 3 – 4 individual fingers, each of which recognizes 3 bp sequences [1]. Overall, a unique sequence of about 24 bp is specifically recognized, and this is large enough to be unique in most genomes.

The most common forms of the ZFN recognition sites are $(NNY)_3N_6(RNN)_3$, of which $(NNC)_3N_6(GNN)_3$ has been extensively studied [23, 24]. The double-stranded breaks will significantly increase integration of DNA into the target site by HR by up to 100 times in plants [1]. But even then, double-stranded breaks induced by restrictions endonucleases or transposons have been shown to be predominantly repaired by NHEJ, often accompanied by some level of mutagenesis [1, 7]. A high proportion of the double-stranded breaks will therefore be repaired by NHEJ, since it is the predominant repair mechanism in plants.

Once the double-stranded break is made, early approaches were to try and increase the chances of their being repaired by HR, over the more predominant NHEJ. The approach has not been very successful. Research efforts should rather focus on ensuring that the repair by NHEJ does not mutate the nucleotide sequence of the target gene in an undesirable manner.

There have been attempts to increase the chances of HR after inducing double-stranded breaks with ZFN. For example, use of ZFN in combination with recombinases and chromatin remodeling proteins, this system increases both targeting precision and transformation efficiencies by HR. Further development of the system should optimise removal or exclusion of marker and reporter genes as well as vector backbone sequences.

Gene targeting using ZFN was first demonstrated for the yellow locus in *Drosophila* [25]. The approach has since become standard for many animal species having been demonstrated even in humans [26]. In higher plants, the technology lagged behind, but has shown a lot of promise with for instance the use of a novel TRV-based vector to achieve non-transgenic genome modification in plant cells [27]. ZFN may also be used for gene deletion [28], and removal of marker genes. Genetically engineered plants released into the environment should not have unnecessary transgene sequences.

Failure to increase HR in plants does not mean that all is lost. In fact, maybe plant transformation efforts may well benefit from NHEJ, which is the predominant mechanism of recombination in plants cells anyway. If one considers for instance a scenario where one needs to disrupt an endogenous gene whose phenotype is easily assayed for. Transient expression of a ZFN that targets the gene should introduce double-stranded breaks in the gene, and most of the breaks will be repaired by NHEJ. Errors introduced during the repair process should inactivate the gene. The precisely engineered plant you get is non-transgenic because there are no foreign sequences integrated into its genome, but the genome would have been elegantly edited! Sequencing of the edited gene should be used to confirm and characterize the mutation. Marton and coworkers reported on a successful experiment with this approach using a disarmed *Tobacco rattle virus* (TRV) vector to deliver the ZFN to cells of intact tobacco and petunia plants [27]. The mutations that were induced were stable and heritable.

There are many other possibilities for gene targeting in plants. For instance, the efficacy of oligonucleotide-directed plant gene targeting has been demonstrated, again with the possibility of the plants being considered non-transgenic [29].

6. Prospects for further development

The efficacy of gene targeting in plants has now been demonstrated, and genetically engineered plants using this technology are being developed. These plants are expected to be low copy number, reflecting on the target gene, and in genomic locations that correspond to the natural locations of the targets. Gene targeting approaches utilize the vast amount of genomic data that is now readily available in databases and can be correlated with the stability of the modification introduced at particular genomic sites. This would be ideal to enhance agricultural attributes of crop, for instance by increasing the expression of a desired product or shutting down a competing or undesirable pathway. With the levels of precision and true engineering that comes with gene targeting, the dependency on reporter genes and even marker genes is reduced. New and elegant ways of delivering DNA to plant cells, such as oligonucleotides and minimal cassettes will enable plant transformation without the use of

plant transformation vectors whose backbones are notorious for integrating into the plant genome. Marker-free, vector backbone-free precisely engineered agricultural crops are what farmers, consumers and the environment need.

7. Conclusions

Gene targeting technology in plants has come a long way, and several alternative approaches to gene targeting have been evaluated. It is now possible and desirable for new plant transformation experiments to give some consideration as to which region of the genome they would want to target, and also give special consideration to reporter genes, marker genes and vector backbone sequences that might be associated with the experiment. It is hoped that the dream of reporter-free, marker-free, vector backbone-free truly and precisely engineered plants will soon be a reality.

Author details

Richard Mundembe*

Address all correspondence to: rmunde01@yahoo.com mundember@cput.ac.za

The Biotechnology Programme, Department of Agricultural and Food Sciences, Cape Peninsula University of Technology, Cape Town, South Africa

References

[1] Weinthal, D, Tovkach, A, Zeevi, V, & Tzfira, T. (2010). Genome editing in plant cells by zinc finger nucleases. Trends in Plant Science , 15, 308-321.

[2] Iida, S, & Terada, R. (2005). Modification of endogenous natural genes by gene targeting in rice and other higher plants. Plant Molecular Biology , 59, 205-219.

[3] Durai, S, Mani, M, Kandavelou, K, Wu, J, Porteus, M. H, & Chandrasegaran, S. (2005). Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. Nucleic Acids Research , 33(18), 5978-5990.

[4] Terada, R, Johzuka-hisatomi, Y, Saitoh, M, Asao, H, & Iida, S. (2007). Gene targeting by homologous recombination as a biotechnological tool for rice functional genomics. Plant Physiology , 144, 846-856.

[5] Evans, M. J, Smithies, O, & Capecchi, M. R. (2001). Mouse gene targeting. Nat Med , 7, 1081-1090.

[6] Puchta, H. (1999). DSB-induced recombination between ectopic homologous sequences in somatic plant cells. Genetics , 152, 1173-1181.

[7] Putcha, H. (2005). The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. Journal of Experimental Botany , 56(409), 1-14.

[8] Hefferin, M. L, & Tomkinson, A. E. (2005). Mechanism of DNA double-strand break repair by non-homologous end joining. DNA Repair , 4, 639-648.

[9] Keeney, S. (2001). Mechanism and control of meiotic recombination initiation. Current Topics in Developmental Biology , 52, 1-53.

[10] Shaley, G, & Levy, A. A. (1997). The maize transposable genetic element Ac induces recombination between donor site and a homologous ectopic sequence. Genetics , 146, 1143-1151.

[11] Fishman-lobell, J, Rudin, N, & Haber, J. E. (1992). Two alternative pathways of double-strand break repair that are kinetically separable and independently modulated. Molecular and Cellular Biology , 12, 1292-1303.

[12] Orel, N, Kirik, A, & Puchta, H. (2003). Different pathways of homologous recombination are used for repair of double-stranded breaks within tandemly arranged sequences in the plant genome. The Plant Journal , 35, 604-712.

[13] Nicolas, A. L, Munz, P. L, & Young, C. S. (1995). A modified single-strand annealing model best explains the joining of DNA double-strand breaks in mammalian cells and cell extracts. Nucleic Acids Research , 23, 1036-1043.

[14] Salomon, S, & Puchta, H. (1998). Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. EMBO Journal , 17, 6086-6095.

[15] Tovar, J, & Lichtenstein, C. (1992). Somatic and meiotic chromosomal recombination between inverted duplications in transgenic tobacco plants. The Plant Cell , 4, 319-332.

[16] Xiao, Y. L, & Peterson, T. (2000). Interchromosomal homologous recombination in Arabidopsis induced by a maize transposon. Molecular and General Genetics , 263, 22-29.

[17] Lyznik, L. A, Gordon-kamm, W, Gao, H, & Scelonge, C. (2007). Application of site-specific recombination systems for targeted modification of plant genomes. Transgenic Plant Research , 1(1), 1-9.

[18] Grindley, N. D, Whiteson, K. L, & Rice, P. A. (2006). Mechanisms of site-specific recombination. Annual Reviews in Biochemistry , 75, 567-605.

[19] Gopaul, D. N, Guo, F, & Van Duyne, G. D. (1998). Structure of the Holliday junction intermediate in Cre-loxP site-specific recombination. The EMBO Journal , 17(14), 4175-4187.

[20] Marshall, S. W, Boocock, M. R, Olorunniji, F. J, & Rowland, S. J. Intermediates in serine recombinase-mediated site-specific recombination. Biochemical Society Transections , 39(2), 617-622.

[21] Kumar, S, Franco, M, & Allen, G. C. (2006). Gene targeting: Development of novel systems for genome engineering in plants. Floriculture, Ornamental and Plant Biotechnology Volume IV, Chapter 8, Global Science Books., 84-98.

[22] Fauser, F, Roth, N, Pacher, M, & Ilg, G. Sanchez-Fernandez, Biesgen, C. and Puchta, H. ((2012). In plant gene targeting. PNAS , 109(19), 7535-7540.

[23] Carroll, D, Morton, J. J, Beumer, K. J, & Segal, D. J. (2006). Design, construction and in vitro testing of zinc finger nucleases. Nature Protocols , 1, 1329-1341.

[24] Liu, Q, Xia, Z. Q, Zhong, X, & Case, C. C. (2002). Validated zinc finger protein designs for all 16 GNN DNA triplet targets. Journal of Biological Chemistry , 277, 3850-3856.

[25] Bibikova, M, Carroll, D, Segal, D. J, Trautman, J. K, Smith, J, Kim, Y-G, & Chandrasegaran, S. (2001). Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. Molecular and Cell Biology , 21, 289-297.

[26] Porteus, M. H, & Carroll, D. (2005). Gene targeting using zinc finger nucleases. Nature Biotechnology , 23, 967-973.

[27] Marton, I, Zuker, A, Shklarman, E, Zeevi, V, Tovkach, A, Roffe, S, Ovadis, M, Tzfira, T, & Vainstein, A. (2010). Non-transgenic genome modification in plant cells. Plant Physiology , 154, 1079-1087.

[28] Petolino, J. F, Worden, A, Curlee, K, Connell, J, Moynahan, T. L. S, Larsen, C, & Russell, S. (2010). Plant Molecular Biology , 73, 617-628.

[29] Oh, T. J, & May, G. D. (2001). Oligonucleotide-directed gene targeting. Current Opinion in Biotechnology12: 169 172.

IntechOpen