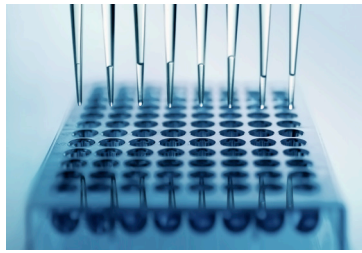
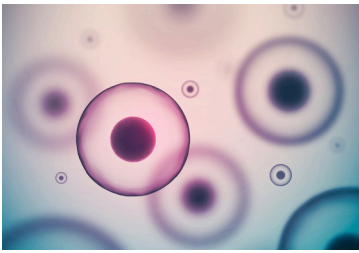


Current status of emerging technologies for plant breeding:

Biosafety and knowledge gaps of site directed nucleases and
oligonucleotide-directed mutagenesis



Biosafety report 02/2015

Biosafety Report 2015/02
GenØk - Centre for Biosafety, Tromsø, Norway
January 2015

Current status of emerging technologies for plant breeding:

Biosafety and knowledge gaps of site directed nucleases and
oligonucleotide-directed mutagenesis

Sarah Z. Agapito-Tenfen
Odd-Gunnar Wikmark

GenØk – Centre for Biosafety, Forskningsparken, PB. 6418, 9294 Tromsø, Norway
Tel.: +47 77 64 66 20 - post@genok.no – <http://www.genok.com>

*Cover photos: lonely #74687963/angellodeco #42501384/alice_photo #75620250/Claude #73273098
(all from dollarphotoclub.com)*

This report was commissioned by the Norwegian Environment Agency and Norad.

GenØk – Centre for Biosafety (www.genok.no) is an independent research institute founded in 1998 and located in Tromsø, Norway. GenØk is engaged in the field of biosafety and gene ecology research on modern biotechnology, nanotechnology, synthetic biology and other technologies emerging from these. This institution also works on capacity building and advisory activities related to biosafety. GenØk focuses on a precautionary, holistic and interdisciplinary approach to biosafety. In 2007, GenØk was appointed national competence center on biosafety by Norwegian authorities.

Table of Contents

Executive Summary	4
Objectives and scope.....	5
1. Background information	7
2. Nuclease-based techniques	8
2.1. What are site-direct nucleases?	8
2.2. Types of site-direct nucleases and modes-of-action.....	10
2.2.1. Meganucleases.....	10
2.2.2. Zinc-finger nucleases (ZFNs).....	12
2.2.3. Transcription activator-like effector nucleases (TALENs).....	13
2.2.4. Clustered regularly interspersed short palindromic repeats (CRISPR) and associated proteins (Cas)..	14
2.3. Intended modifications by site-direct nucleases (SDN-1, SDN-2 and SDN-3)	16
2.4. Identification of unintended effects and consideration of uncertainty of SDN-based crops	17
2.4.1. Specificity of target cleavage	18
2.4.2. Specificity of double-stranded break repair mechanisms.....	21
3. Oligonucleotide-directed mutagenesis-based techniques	23
3.1. What are oligonucleotide-directed mutagenesis?	23
3.2. Types of oligonucleotide-directed mutagenesis	26
3.2.1. Single-stranded oligo-deoxynucleotides (SSOs or ssODMs)	26
3.2.2. Chimeric RNA-DNA oligonucleotide molecules (RDOs)	27
3.2.3. Small Fragment Homologous Replacement (SFHR).....	28
3.2.4. Triple helix-forming oligonucleotides (TFOs)	29
3.3. Identification of unintended effects and consideration of uncertainty of ODM-based crops	29
4. Conclusions	32
5. Recommendations for follow up and additional research needs	34
6. References	36

Executive Summary

The new plant breeding techniques provide the emergence of novel products that challenges our current regulations and our management practices of what we traditionally have viewed as a genetically modified organism (GMO). International regulations, such as the Cartagena Protocol on Biosafety, operate with definitions of GMOs that may not be applicable to products arising from some of these new techniques. The question then arises on how society and regulatory bodies should view and regulate the products. This report does not approach that problem *per se*, but as a crucial step in management, we sum up the current scientific understanding of two new plant breeding techniques, site directed nucleases (SDN) and oligonucleotide directed mutagenesis (ODM). The underlying mode of action of both of these techniques are the plants natural repair systems and how this can be utilized to achieve genomic modifications. Herein also lies the main challenge for risk assessment - our limited knowledge about the function of these systems, factors involved and potential off-target effects.

This report aims at providing an overview of the current status of scientific knowledge concerning SDN and ODM. We have reviewed up to date peer reviewed scientific publications on the mechanisms and natural functions that are utilized by SDN and ODM techniques in an effort to understand potential risks such as unintentional changes in the genome of plants. Finally, recommendations for action are outlined.

SDNs

Site-direct nucleases are enzyme complexes that recognize specific DNA sequences in the genome and cleave them. The cleaved DNA is subsequently repaired by the organisms natural DNA repair systems. Currently, these enzymes are divided into four categories: meganucleases, zinc-finger, TALEN and CRISPR. The end product of the use of SDN techniques will largely depend on the design of the nuclease DNA recognition protein and the available template for repair. This report outlines potential unintended effects derived from the use of SDN that are mainly related to uncertainties regarding DNA target sequence specificity and DNA double-stranded break repair mechanism. There are several factors that influence both DNA binding and DNA repair, unfortunately they are to a large extent not fully understood. The lack of mechanistic understanding is a severe limitation for identifying potential hazards from SDNs and more research in this field is greatly recommended. Identifying unintentional effects in a system which is not fully understood becomes very difficult. However, using comprehensive untargeted profiling methods (such as omics) in order to detect and identify unintentional mutations with methodologies that are available today could minimize the potential hazards from SDN products.

ODM

Oligonucleotide-directed mutagenesis uses oligonucleotides to induce sequence specific mutations of native genomic sequences (i.e. genome editing). The introduced DNA is complementary to the genomic target sequence with the exception of a modification that usually is a deletion, insertion or a mismatch between the introduced synthesized DNA and the genomic DNA. There is scientific dispute on the DNA repair pathways and mechanisms that are induced, but briefly the cells DNA repair mechanisms detects the mismatch and repairs the genomic DNA using the introduced

modified DNA as a template. There are some reports on collateral and unspecific background mutations as side effects, but several research groups highlighted the uncertainties regarding those studies. Unintended effects that are described as induced by ODMs are related to cellular death, unpredicted mutations, mutation inheritance discrepancies, and others.

In conclusion, the two techniques reviewed in this report are not fully scientifically understood and thus poses many uncertainties connected to mode of action as well as potential unintentional effects. The safety assessment of such products should take into account risks associated with other existing user practices and habits, and the sources and nature of uncertainty that could not be addressed during the preceding steps of the risk assessment.

Finally, several recommendations have been proposed for the development of future research to contribute to the better understanding of the modes-of-action of such techniques and also to seek fulfilling the biosafety knowledge gaps. Overall, the results of this report indicate that biosafety considerations regarding new plant breeding technologies could, in principle, be addressed by the general approach developed for conducting the risk assessment of genetically modified crops. However, it is important to understand that the limitation that resides in the current understanding of these techniques regarding their potential adverse effects. Therefore, according to the requirements of a scientifically based risk assessment and the application of the precautionary principle, further biosafety research is highly recommended.

Objectives and scope

The aim of this report is to provide an assessment and identification of scientific knowledge gaps and uncertainties that are related to two new emerging technologies for plant breeding. In this context, new emerging technologies for plant breeding are mainly about changing the process of creating a new crop rather than changes in the traits carried by these organisms.

Within the site-directed nuclease-based (SDN) group of methods we have focused on the zinc finger technology as well as giving an overview of the other SDNs and oligonucleotide-directed mutagenesis (ODM) techniques. We have chosen not to discuss possible implementation of regulation concerning products resulting from the utilization of these plant-breeding methods.

Because commercialized products from ODM and SDN are not yet available in the market, it is not possible to address realistic risk scenarios, as well as case examples, that are connected to the choice of method used to generate modifications to the host plant genome. In addition, SDN and ODM are methods that promote the integration of foreign DNA or the modification of existing plant DNA. In any given plant modified by SDN or ODM techniques, the connected risk assessment must also take into account the transformation method and possibly the plant regeneration method as well as the inserted or modified DNA. This report does not address those tightly connected methods due to the lack of products available on the market today resulting from SDNs or ODMs.

As mentioned above, there is a lot of knowledge concerning the mechanism or mode of action of the nucleases and repair mechanisms activated that needs to be uncovered. The purpose of this report is to show in what area these knowledge gaps reside. Therefore, we have reviewed scientific evidence including latest findings concerning unintentional changes in the genome of plants induced by these

techniques. Finally, the report provides insight into strategies for identification and management of these risks.

In addition, answers to the questions on the regulation of products from these new techniques, whether it falls within GMO legislation or not, were not within the scope of this report. However, this issue should be carefully taken into account not only to comply with national and international regulations but also to evaluate to what extent the precautionary principle should be invoked on products that might fall outside of existing safety regulations. The characteristics of these new crops and products may have adverse impacts in the environment, as well as socio-economic implications that have not yet been fully investigated.

The utilization of ODMs and SDNs is only one step in the generation of the final plant that will be grown in crop fields. Still there is the question of delivery of DNA/proteins to the cell and the generation of a plant from single cells. In addition crossing of the laboratory parental strain into elite crops raises the same issues as when generating elite crops from traditional genetic engineering, albeit with a much more difficult task of monitoring and detection.

1. Background information

Plant science has made considerable progress in the past years towards the development of new plant breeding techniques to alter genetic and epigenetic factors more efficiently than current modern biotechnological breeding techniques (e.g. transgenesis). Recent advances in the field of genome engineering have successfully described site-directed modifications to the genomes of crop species, including targeted mutations, gene insertions, and gene replacements (Lusser et al., 2012). The development of such new technologies has been also recently used to create new and valuable agronomic traits (Lusser et al., 2012). As a key step to enable site-directed genome modifications are sequence-specific nucleases or oligonucleotide molecules that trigger cell's native DNA repair pathways within sequences of genes of interest.

New plant breeding techniques are under development and already being used in breeding programs by agribusiness companies (Camacho et al., 2014). A few crops that have been developed by these techniques have already reached the last commercial development phase, which comprises regulatory submission (if applicable), seed bulk-up and pre-marketing. A full list of these products can be found in Lusser et al. (2012) and a list of regulatory inquiries can be found in Camacho et al. (2014), the latter revises US cases only. Oilseed rape and maize developed thru ODM for herbicide tolerance were pointed to reach phase 2 and 3, respectively. And tomato, oilseed rape and maize developed thru ZFN were by phase 3, but traits were not disclosed. According to Lusser et al. (2012), phase 2 is related to trait development, pre regulatory data and large-scale transformation. Phase 3 is trait integration, field-testing and regulatory data generation (if applicable).

A literature search showed an unbalanced number of studies that apply these genome-editing techniques to plants species. While there are dozens of papers with ZFN application, there are very few that apply ODM or other SDNs to plant species. Targeting native agricultural crops and not just model plants has been widely seen in SDN applications. But this report revealed a lack of published scientific literature for ODM techniques applied to plant species. This poses extra challenges for the identification of potential unintended effects and thus raises knowledge gaps. It is clear that some of the effects observed in animal models cannot be assumed to occur in plant models due to metabolic differences between them.

New biotechnology-based plant breeding techniques and products thereof raise many regulatory challenges, as they do not necessarily fit into known product definitions, regulatory frameworks and risk assessment approaches. Therefore, this report examines the history of the application of the two new plant-breeding techniques (SDN and ODM), with a focus on unintentional effects and knowledge gaps involving these techniques. The report also shows similarities in the functioning of such technologies, which are mainly related to their efficiency in targeting a specific region in the host genome. As new information becomes available, these knowledge gaps will no doubt extenuate. The reason for this analysis is to both create a historical record of the emergence of these technologies and for their potential application to serve as another case study in how 'early warnings' may be incorporated into risk assessments at the cutting edge of technology.

2. Nuclease-based techniques

Summary

- Site-direct nucleases are enzyme complexes that recognize specific DNA sequences in the genome and cleave them. They combine DNA binding domains to non-specific nuclease.
- There are currently 4 types: meganucleases, zinc-finger, TALEN and CRISPR.
- The outcome product of the use of such nuclease-based techniques will depend on the nuclease design, consequently, the intended genome modification.
- Safety analysis of such products categorizes them into three major classes: SDN-1, SDN-2 and SDN-3.
- SDN-1 produces random mutations at target site, SDN-2 introduces predefined mutations at target site, SDN-3 adds large DNA sequences (transgenes) at target site.
- New unintended effects derive from the use of such nucleases are mainly related to uncertainties regarding target specificity and double-stranded break repair.
- Specificity of target cleavage depends on: DNA-binding domain (e.g. DNA-binding energy, tolerance of mutant sequences), DNA-cleavage domain (e.g. nuclease dimerization), Inter-domain linker, and others (e.g. target chromatin structure, cell type, nuclease concentration (levels), environmental stimuli that have an effect in nuclease expression).
- Double-stranded break repair cannot be predicted a priori, neither on-target or off-target sites.
- Therefore, comprehensive untargeted profiling methods (such as *omics*) should be applied in order to detect and identify unintentional mutations in the entire host genome.

2.1. What are site-direct nucleases?

Strategies for target and specific modification of the genome, or the genetic information, have been referred to as *genome editing*. Most genome editing techniques direct the repair of DNA double-strand-breaks (DSBs) achieved in specific genome sequences by sequence-specific nucleases, or site-specific nucleases, also called site-direct nucleases (SDNs). Such enzymes recognize and cleave genomic regions that shares sequence homology to their DNA binding enzymatic domains. The repair of DSBs can be directed to create a variety of targeted DNA sequence modifications, ranging from DNA deletions to the insertion of large arrays of transgenes. There are currently four major classes of SDNs: meganucleases, zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and clustered regularly interspersed short palindromic repeats (CRISPR)/Cas9 reagents (Camacho et al., 2014; Voytas and Gao, 2014).

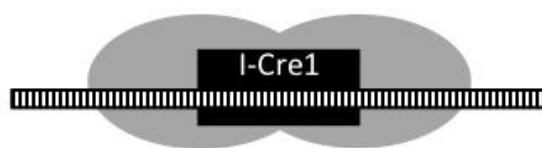
These kinds of site-directed nuclease share two main features: (i) the capacity of binding to DNA at a specific sequence and (ii) to cleave those. The double stranded DNA break is then restored by the cell innate DNA repair machinery. DSBs stimulate the homologous recombination (HR) or error-prone non-homologous end-joining (NHEJ) (Box 1) DNA repair machinery, which can then lead to site-specific mutagenesis or the integration of foreign DNA.

Site-directed nucleases are composed of two main parts: (i) the DNA-binding domain, which determines the DNA (site) specificity of the endonuclease; and the (ii) the DNA cleavage domain, which is responsible for the actual cleaving of the endogenous DNA (non-specific nucleases). Table 1

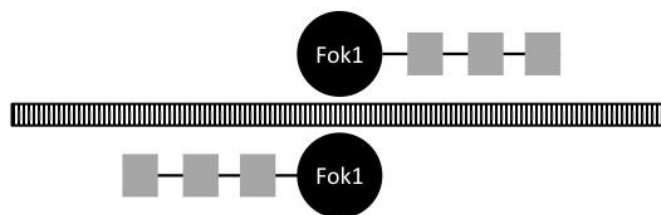
summarizes the main components and modes-of-action of the four types of site-directed nucleases that will be further discussed later. And Figure 1 is a schematic drawing of each nuclease in order to show basic structural differences between them.

SDN type	Components	Mode-of-action
Meganucleases	Homing endonuclease with a large DNA recognition site (12–40 bp)	The catalytic domain also determines the DNA sequence specificity. Induces DSB in target DNA sequence.
Zinc-finger nucleases (ZFNs)	Multiple zinc finger peptides (each targeting 3 bp of genomic sequence) fused to restriction nuclease (usually <i>FokI</i>)	ZFN targets are bound by two zinc-finger DNA binding domains (dimers) separated by a 5–7-bp spacer sequence. Induces DSB in target DNA sequence and the <i>FokI</i> cleavage occurs within the spacer.
Transcription activator-like effector nucleases (TALENs)	Arrays of the TAL effector motifs (each targeting 1 bp of genomic sequence) fused to restriction nuclease (usually <i>FokI</i>)	The two TALEN targets sites (dimers) are typically separated by a 15-20-bp spacer sequence. Induces DSB in target DNA sequence and the <i>FokI</i> cleavage occurs within the spacer.
Clustered regularly interspersed short palindromic repeats (CRISPR)/Cas9	20-nt crRNA fused to tracrRNA and Cas9 endonuclease that has two nuclease domains and each cleaves one DNA strands.	The CRISPR-based guide RNA (gRNA) recognizes target DNA sequences by base pairing. Cas9 creates DSB at DNA target sites with complementarity to the 5'-end of a gRNA.

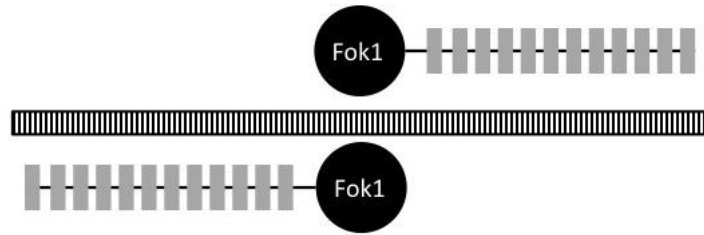
Table 1. Overview of site-directed nuclease (SDN) types, their main components and modes of action. Note that 'bp' stands for base pairs (for review see Curtin et al., 2012 and Voytas and Gao, 2014).



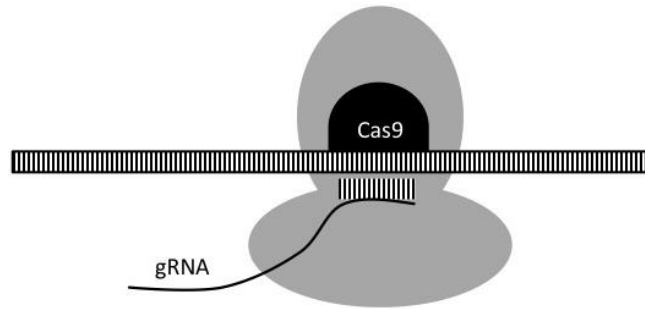
a. Meganucleases, specific I-Cre1 residues are responsible for target recognition and cleavage.



b. Zinc-finger nucleases, Fok1 is responsible for non-specific cleavage and target recognition is performed by the binding of two zinc-finger DNA binding domains (dimers) separated by a spacer sequence.



c. TALEN nucleases, two dimers also bind to DNA target sequence and these are fused to Fok1.



d. CRISPR/Cas9 system, DNA recognition is performed by a guide RNA strand and cleavage occurs by two nuclease domains and each cleave one strand of target DNA.

Figure 1. Schematic representation of the four main types of site-direct nucleases. Gray drawings represent DNA binding domain and black drawings represent cleavage domains. The shape of these drawings are merely illustrative and do not represent the true shape of these enzymes. These drawings were based on Voytas and Gao, 2014.

2.2. Types of site-direct nucleases and modes-of-action

2.2.1. Meganucleases

Meganucleases were initially recognized as microbial endonucleases that catalyze genomic modification, rearrangement, protection, and repair in a gene-specific matter. The term *meganuclease* reflects the name “Omega” that was coined for the first known homing system as well as for designating sequence specific endonucleases proteins that could recognize large targets (from 12 to 40 bp) (Thierry and Dujon, 1992). Homing is a process in which microbial self-splicing intervening sequences are specifically inherited, thus duplicated, into recipient alleles of their host gene that initially lacks the sequence (Stoddard, 2011). The term has progressively become synonymous with naturally occurring enzymes that bind and cleave large DNA sequence targets, and should no longer include artificial endonucleases such as zinc finger nucleases. This protein family has grown to include several hundred of identified members, found in eukaryotes, bacteria and archae, with the notable exception of metazoans (Chevalier and Stoddard, 2001) and many could be shown to initiate homing events.

Over the past 25 years, a number of studies have applied meganucleases in animal and mammalian models for gene targeting recombination, thus showing the applicability of this technology. Although several hundreds of other meganucleases had been identified in between, the collection of cleavable sequences was too limited to address the complexity of the genomes (Pâques and Duchateau, 2007). Engineering the sequence specificity of meganucleases was the next step to overcome this issue. But the lack of knowledge of protein structure and the relation between structure and function has not only delayed the development of new meganucleases sequences but also illustrated clearly the complexity of designing novel proteins.

Box1: Double-stranded break (DSB) repair mechanisms. DSB is a form of DNA damage that occurs when both DNA strands are cleaved. Site-direct nucleases induce targeted DSBs, which then stimulate the cellular native DNA repair machinery, including homologous recombination (HR) or error-prone non-homologous end joining (NHEJ) (for review see Wyman and Kanaar, 2006).

Double-stranded DNA break repair

DSB are DNA lesions that can be both detrimental and beneficial to organisms. A deleterious effect example is the potential genome rearrangements that can initiate carcinogenesis or apoptosis. By contrast, DSBs can be also beneficial when they occur in a controlled manner in the context of specialized events that demand genome sequences to be rearranged, such as during development of the immune system and generation of genetic diversity in meiosis (Wyman and Kanaar, 2006). In order to avoid DSB-induced genetic loss of information and disruption of vital processes, such as replication and transcription, cells possess robust mechanisms to repair DSBs and each mechanism is triggered by the structure of the broken DNA (Podevin et al., 2013). There are two major types of DSB repair mechanisms (Paquês and Haber, 1999), and they are briefly described below:

Homologous recombination

Homologous recombination is the exchange of base-paired partners between two homologous DNA molecules. The process requires extensive regions of DNA homology and repairs DSBs accurately by using information on the undamaged sister chromatid. When one of the two sister chromatids has suffered a DSB, processing results in single-stranded tails at the break with 3'-hydroxyl ends. The tails are a substrate for nucleoprotein filament formation with a recombinase. The nucleoprotein filament directed homology recognition and DNA strand exchange lead to joint molecule formation between the broken DNA and the intact sister chromatid, after several DNA synthesis steps, this will result in non-crossover recombinants.

Non-homologous end joining

Non-homologous DNA end joining uses extremely limited or no sequence homology (DNA repair template) to rejoin juxtaposed ends. Therefore, DNA ends may be processed to expose or create ends, able to ligate (i.e. come together) a 3'-hydroxyl and a 5'-phosphate. In such cases, sequence information can be lost upon rejoining, making this process error prone.

The first homing endonuclease structure to be characterized was the I-Cre1 protein (Heath et al., 1997; Jurica et al., 1998). Mutations of specific residues of I-Cre1 have shown substantial cleavage of new targets and loss of affinity to its natural site (Sussmann et al., 2004). Millions of combinations were screened resulting in several hundreds of mutants with locally altered specificity (Pâques and Duchateau, 2007), but of course the randomization of the nine residues of the I-Cre1 meganuclease

family (LAGLIDADG) would not result in enough variations to engineer the target site. In addition, most homing endonucleases are not clearly separated from the catalytic domains, thus making protein engineering procedures more complex and often compromising nuclease activity (Taylor et al., 2012). This has hindered their widespread use (Voytas and Gao, 2014).

To the best of our knowledge, there are only one publicly interest on the commercialization of a crop developed by the use of meganucleases. This was seen in a publicly available letter of inquiry to US Department of Agriculture (USDA) Animal and Plant Health Inspection Service (APHIS) made on the regulatory status of a product that had used meganucleases (Camacho et al., 2014). Although little information was available, the inquiry in 2011 from a company called “Cellestis”, required information regarding the regulatory status of a potential crop genome modification product derived from the use of I-Cre1 meganuclease. The I-Cre1 endonucleases are derived from the chloroplasts of *Chlamydomonas reinhardtii* (Thompson et al., 1992) and experimental selection and computational design methods have been further developed to re-engineer the cleavage specificity of I-Cre1 and its homologs (Chevalier et al., 2002; Ashworth et al., 2006; Gao et al., 2010).

Numerous methods have been developed to assess mutated meganucleases for activity and altered target specificities (Seligman et al., 2002; Rosen et al., 2006; Gao et al., 2010; Stoddard, 2011). The first report of an reengineered meganuclease to target a chromosomal locus of choice in a crop species, applied an reengineered I-Cre1 to target a sequence neighboring the maize liguleless locus and effectively generated heritable monoallelic and biallelic mutations at the target site (Gao et al., 2010), thereby knocking out the gene transcription.

These and other studies have shown that the modification of a homing endonuclease’s cleavage specificity would be required in order to target and modify endogenous target sites in various biological genomes (Stoddard et al., 2007). Thus, a variety of biotechnology applications can potentially make use of the properties of a homing endonuclease that could lead to mutation, knockout, modification, or insertion of exogenous coding DNA into the endogenous gene target.

2.2.2. Zinc-finger nucleases (ZFNs)

Zinc fingers are small protein domains that are fold stabilized by a zinc metal molecule. These enzymes are structurally diverse and present among proteins that perform a broad range of functions in various cellular processes, such as replication and repair, transcription and translation, metabolism and signaling, cell proliferation and apoptosis (Krishna et al., 2003). They work as interaction modules and bind to several compounds, ranging from nucleic acids to proteins and other small molecules.

Zinc fingers nucleases have emerged in the first attempt to stimulate gene targeting by applying novel chimeric restriction enzymes, fusing the catalytic domain of a well-known restriction enzyme, *Fok1* from *Flavobacterium okeanoicoites*, to another DNA binding domain to alter the DNA specificity of such enzymes. These researchers have enhanced gene targeting with designed zinc finger nucleases (Bibikova et al., 2003; Porteus and Baltimore, 2003).

The ZFN architecture links the DNA-binding domain of a versatile class of eukaryotic transcription factors, so called the zinc finger proteins, with the nuclease domain of the *Fok1* restriction enzyme. Originally, the number and order of cysteine and histidine amino acid residues was used to classify

different types of zinc fingers (e.g., Cys₂His₂, Cys₄, and Cys₆). More recently, a systematic method classifies zinc finger proteins based on the overall shape of the protein backbone in the folded domain (Krishna et al., 2003) and the Cys₂–His₂ zinc-finger domain is among the most common types of DNA-binding motifs found in eukaryotes.

The flexibility of zinc finger proteins makes it a robust platform for the design of novel DNA binding domains (Cathomen and Joung, 2008). These domains contain a tandem array of Cys₂–His₂ fingers, each recognizing approximately 3 bp of DNA (Miller et al., 1985). In primary studies, three fingers were applied to bind a 9-bp target, which enabled ZFN dimers to specify 18 bp of DNA per cleavage site. But recent studies have added more fingers (up to six per ZFN) to specify longer and rarer cleavage targets (Urnov et al., 2010). ZFN works as dimers because *Fok1* must dimerize in order to cleave DNA. Thus, cleavage by *Fok1* as part of a ZFN requires two adjacent and independent binding events, which must occur in both the correct orientation and with appropriate spacing to permit dimer formation (Wah et al., 1998; Bitinaite et al., 1998).

The initial successes in plant genome engineering with ZFNs were reported in *Arabidopsis thaliana* and tobacco (*Nicotiana tabacum*) (Lloyd et al., 2005; Wright et al., 2005). Further developments included a small number of crop plants, such as maize and soybean, and introduced important agronomic traits. Herbicide-resistance mutations were introduced into tobacco endogenous acetolactate synthase genes (ALS SuRA and SuRB) loci by ZFN-mediated gene targeting at frequencies exceeding 2% of transformed cells for mutations as far as 1.3 kilobases from the ZFN cleavage site (Townsend et al., 2009). In addition, gene disruption and repair by the NHEJ pathway was observed in the tobacco protoplasts with more than 40% of recombinant plants having modifications in multiple SuR alleles. These authors have also tested for HR-based gene replacement by electroporation of ZFN plasmids into tobacco protoplasts along with SurB donor templates modified at positions known to confer herbicide resistance. Successful gene replacement was observed at frequencies ranging from 0.2 to 4% (Townsend et al., 2009). Another example of gene editing using ZFNs was published by Shukla et al. (2009) in which the authors have generated ZFNs that targeted the maize inositol-1,3,4,5,6-pentakisphosphate enzyme gene (IPK1). The disruption of IPK1 gene reduces the level of phytate, a known anti-nutritional compound found in seeds of many cereals. The IPK1 locus was disrupted by ZFN-mediated DSBs and repaired by HR using donor templates with an herbicide resistance gene and short, locus-specific homology arms (Shukla et al., 2009).

2.2.3. Transcription activator-like effector nucleases (TALENs)

TALENs are similar to ZFNs in that they have a DNA-binding domain derived from TALE proteins fused to *Fok1* cleavage domain. TALE proteins are transcription factors from the plant bacterial pathogen *Xanthomonas*. The bacteria produce such proteins to target and control the expression of specific plant gene promoters in order to mimic host transcription factors and thus achieve successful pathogen infection (Bogdanove et al., 2010).

DNA recognition by TALENs is achieved through arrays of TAL effector motifs. TALEs typically consists of a 16 to 20 single repeat monomers, with each monomer 34 amino acid residues in length. Although highly conserved, the exception is the hypervariable amino acid residues at positions 12 and 13, called repeat-variable di-residues (RVDs). The DNA-binding specificity of TALE repeats is apparently solely based on the nature of the RVDs in a typical 34-aa repeat (Boch and Bonas, 2010).

thus suggesting that the amino acids at position 12 and 13 probably interact directly with the DNA bases.

The consecutive array of repeats binds to a consecutive DNA sequence and rearranging repeat units generates novel custom-designed DNA-binding specificities with high potential for biotechnology (Curtin et al., 2012). Computational and molecular biological analyses made it possible to decipher the TALE code for DNA recognition (Boch et al., 2009; Moscou and Bogdanove, 2009). Likelywise, the number of repeats and the sequence of the RVD region will determine the length and sequence of the target DNA. DNA-binding domains of individual TALENs typically contain between 15 and 30 RVDs and can then target 15 up to 30 nucleotides in the host genome accordingly (Cermak et al., 2011). However, in contrast to zinc-finger proteins, the re-engineering of the linkage between repeats is not necessary to construct long arrays of TALEs (Christian et al., 2010) because of the lack of interaction between every 3 nucleotides.

To date TALENs have been used to generate targeted modifications in a variety of plant species, such as *Arabidopsis thaliana* (Cermak et al., 2011), tobacco (Mahfouz et al., 2011; Mahfouz and Li, 2011) and rice (Li et al., 2012). Li et al. (2012) have presented one of the first successful demonstrations of a nuclease-mediated modification of agronomic importance. These authors have exploited TALEN technology to edit a specific susceptibility gene *Os11N3* (also called *OsSWEET14*) in rice to thwart the virulence strategy of *Xanthomonas oryzae* and thereby engineer heritable genome modifications for resistance to bacterial blight, a devastating disease for rice. During infection the bacteria secretes effector proteins that target DNA sequences in the promoter region of the rice sucrose-efflux transporter gene (*OsSWEET14*) (Curtin et al., 2012). It is interesting though, is that *OsSWEET14* also plays a crucial role in the development of the plant, thus obtaining a knockout mutant to circumvent the effects of the pathogen was not feasible. The authors successfully reduced the pathogen's virulence by disrupting the promoter sequence bound by the pathogen effector.

2.2.4. Clustered regularly interspersed short palindromic repeats (CRISPR) and associated proteins (Cas)

Over the past two years, a more versatile genome editing tool has been developed based upon the bacterial clustered regularly interspaced short palindromic repeats (CRISPR) associated protein (Cas) adaptive immune system and tested in different organisms including plants (Kumar and Jain, 2014). The first description of a CRISPR array was made in 1987 by researchers who found repeat sequences interspersed by short non-repeating spacer sequences in *Escherichia coli*, and in subsequent years, similar CRISPR arrays were found in *Mycobacterium tuberculosis* and other bacteria and archaea (for review see Sorek et al., 2008).

Bacterial CRISPR/Cas systems process short segments of foreign DNA into small elements that are then inserted into the CRISPR locus in the bacterial genome. RNAs derived from CRISPR loci are constitutively transcribed and processed into small RNAs (crRNA) of exogenously derived and incorporated DNA. After processing, these small RNAs guide other Cas proteins to mediate sequence-specific degradation of the foreign DNA thereby recognizing sequences encountered before in a immune-system like manner (Kumar and Jain, 2014). The key components of the CRISPR/Cas system are Cas proteins (cleavage site) and CRISPR RNAs (recognition site). These two components are the

basic constituents of this genome editing system and they occur under three major CRISPR/Cas systems (Makarova et al., 2011).

Type I and III systems share some overarching features. In both, Cas endonucleases process the precursor CRISPR RNA (pre-crRNA) to crRNA that then assembles into a large multi-Cas protein complex capable of recognizing and cleaving nucleic acids with complementarity to the crRNA. In contrast, type II systems require a trans-activating crRNA (tracrRNA), with complementary sequences, to trigger pre-crRNAs processing by the double-stranded RNA-specific ribonuclease RNase III in the presence of the Cas9 protein. Cas9 is thought to be the sole protein responsible for crRNA-guided silencing of foreign DNA (Jinek et al., 2012).

The type II system are considered simpler than other CRISPR systems (Pennisi, 2013), and its endonuclease family can be programmed with single RNA molecules to cleave specific DNA sites, thereby raising the exciting possibility of developing a simple and versatile RNA-directed system to generate DNA DSB for genome targeting and editing (Jinek et al., 2012).

That is the major reason why the Cas9 system (CRISPR/Cas9) is the most explored Cas protein to date. It has been demonstrated that the two acting RNA molecules (crRNA and the tracrRNA) can be fused artificially to form a chimaeric RNA molecule termed single guide RNA (sgRNA) (Mali et al., 2013; Li et al., 2013). The site-specific catalytic action of this sgRNA–Cas9 complex is defined by a sequence of only ~20 consecutive nucleotides. It is reasonable that designing and producing a synthetic 20 nt sgRNA long is much simpler compared with the production of custom ZFNs and TALENs enzymes. A brief comparison of meganuclease, ZFN, TALEN, and CRISPR/Cas9 technologies is presented in Table 1 (adapted from Kumar and Jain 2014).

By modifying these basic constituents for the use in other organisms, the CRISPR/Cas9 system has been shown to be a useful tool for gene editing and silencing. CRISPR sequences can be engineered which give rise to crRNA directed against specific endogenous genes in various organisms. Significant advances for the use of this technique have been promoted by the generation of a “single-guide RNA” (sgRNA) that combines the function of the tracrRNA and crRNA in a chimeric molecule (Jinek et al., 2012; Jinek et al., 2014).

Kumar and Jain (2014) have recently reviewed the literature on the CRISPR/Cas system advances and opportunities for plant genome editing. These authors report that the usage of CRISPR/Cas9 system-mediated plant genome engineering are limited so far and it was only in 2013, that three different research groups simultaneously reported targeted genome modifications in plants using CRISPR/Cas9 (Li et al., 2013; Nekrasov et al., 2013; Shan et al., 2013b). Later, a number of other reports demonstrated the efficacy/potential of the CRISPR/Cas9 technology as an efficient genome editing system in model and crop plants (Belhaj et al., 2013; Feng et al., 2013; Jiang et al., 2013; Mao et al., 2013; Miao et al., 2013; Sugano et al., 2014; Jia and Wang, 2014; Wang et al., 2014).

An example of genome editing of crop plants was performed by Shan et al. (2013) who designed several sgRNAs and codon-optimized Cas9, which target different DNA strands of four rice endogenous genes and one wheat gene. These authors have demonstrated different levels of mutagenesis (frequency of insertions and deletions) and the possibility of homology-directed repair by co-transformation of Cas9, sgRNA and single-stranded DNA oligos into plant cells. Another report demonstrated the use of a multiplex CRISPR/Cas9 technique for gene simultaneously edition of two sites (CHLOROPHYLL A OXYGENASE1 and LAZY1) in the *Arabidopsis* genome by applying a single CRISPR/Cas9 construct harboring two sgRNA expression cassettes (Mao et al., 2013). In addition, it

has been demonstrated that the CRISPR/Cas9 system can provide homozygous gene editing within one reproductive generation (Zhang et al., 2014).

2.3. Intended modifications by site-direct nucleases (SDN-1, SDN-2 and SDN-3)

The outcome product of genome editing depends on mechanism used for the DSB repair and on the presence of a DNA template (donor DNA) that may be added artificially. Therefore, depending on the type of SDN and the presence of a template or not, the plant cell might trigger different repair mechanisms (Figure 2).

In line with the EFSA Panel on Genetically Modified Organisms Scientific opinion addressing the safety assessment of plants developed using Zinc Finger Nuclease 3 and other Site-Directed Nucleases with similar function (EFSA, 2012) and the Joint Research Centre Report on New Plant Breeding Techniques: State of the art and prospects for commercial development (Lusser et al., 2011), the classification used for addressing such techniques was based on the types of modifications made by nucleases. In this regard, the intentional modifications should be categorized into three major classes:

- SDN-1: SDN-1 techniques use SDNs to generate site-specific random mutations at target sequences harboring the nuclease recognition sequence. SDN-1 nucleases are delivered into plant cells without the presence of a donor DNA repair template. In most cases, single DSBs are repaired mainly by NHEJ. There are two kinds of outcome products of such class of SDN techniques. The first kind is resulted from unfaithful repair of NHEJ mechanisms that create small nucleotide deletions and/or insertions (i.e. indels). The second kind of outcome is the deletion of entire regions (e.g. to delete regulatory regions, exons/introns, whole genes or even parts of chromosome) between the two target sites when two DSBs are introduced (Podevin et al., 2013). There is also evidence of plant gene targeting using SDN-1 techniques that have caused genomic inversions or translocations between the two induced DSBs (Pacher et al., 2007).
- SDN-2: SDN-2 techniques deliver SDN complexes along with an externally supplied DNA template (donor DNA). Donor DNA sequences contain the targeted desired mutations (i.e. few nucleotide substitutions or short insertion/deletions) and sequences with high homology to the endogenous gene target in order to trigger homologous recombination (HR). A successful HR event substitutes the endogenous sequence by the donor DNA introducing the desired mutations. The introduced modifications can be target to repair undesirable spontaneous mutations (i.e. targeted gene correction) or to introduce targeted mutations and new phenotypes (i.e. gene/allele replacement) (Podevin et al., 2013).
- SDN-3: Similarly to SDN-2, SDN-3 techniques apply a SDN complex together with donor DNA between flanking DNA sequences showing homology to the endogenous target locus. However, in the case of SDN-3, a large stretch of donor DNA is used (i.e. plasmid DNA with promoter, gene of interest, terminator, gene reporter sequences, etc). Usually, homologous

sequences will trigger HR, thus facilitating the target integration of exogenous sequences, which otherwise would integrate randomly in the genome (Podevin et al., 2013). Therefore, the SDN-3 approach can be used for targeted addition of genes of interest, gene replacement and trait stacking (Lusser et al., 2012).

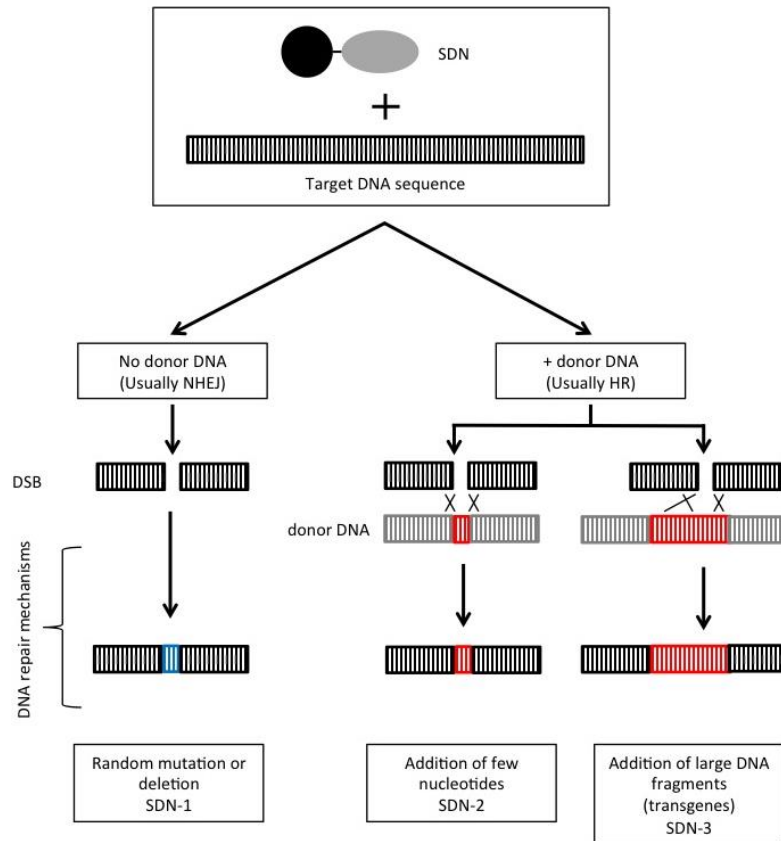


Figure 2. Schematic representation of intended modification by site-direct nuclease (SDN) types. Each SDN technique type (SDN-1, SDN-2 and SDN-3) is defined by the product outcome and will depend on the SDN design. This graphic shows a SDN in association with the target sequence. Double-strand break (DSB) repair can occur via non-homologous end-joining (NHEJ) or homologous recombination (HR) when a donor DNA is present. SDN-1 can result in site-specific random mutations or deletion. SDN-2 can result in the addition of a few specific nucleotides. SDN-3 can result in the integration of large DNA fragments (transgenes). This graphic is based on the work of Podevin et al., 2013.

2.4. Identification of unintended effects and consideration of uncertainty of SDN-based crops

Site-direct nuclease-based techniques are subject to a number of possible unintended effects as any other technology applied to living organisms. In addition, due to the recent discovery of these new plant breeding techniques and the current lack of knowledge on the details of the mechanisms, significant uncertainties might be associated to the assessment of unintended effects.

We propose a step-by-step approach in order to contextualize and identify unintended effects and associated uncertainties of nuclease techniques on each stage of a SDN-based crop development (Box 2). In order for a crop to be developed, other techniques than the genome editing technique itself are necessary for the delivery of the SDN and the regeneration and/or multiplication of the plant. Although we focus on the safety issues related to the use of SDN only, other biotechnology techniques (such as transgenesis, plant cell tissue culture, etc.) may be applied and also need to be considered.

2.4.1. Specificity of target cleavage

Genome stability is normally maintained in the face of potentially dangerous DSBs by the combined activity of non-homologous DNA end-joining and homologous recombination (Wymann and Kanaar, 2006). In gene targeting approaches, the location of inducible genomic DSBs will depend on the DNA-binding specificity of the nuclease. The specificity will then determine the activity and toxicity of nucleases.

Off-target cleavage has been most studied in ZFNs. Although at different levels, the imperfect specificity of engineered zinc-fingers domains has been linked to cellular toxicity (Cornu et al., 2008).

Off-target cleavage of ZFNs has been investigated by Pattanayak and coworkers (2011) utilizing an *in vitro* selection method. The authors find that information about the specificity of ZFNs has been based on unproven assumptions that (i) dimeric ZFNs cleave DNA with the same sequence specificity with which isolated monomeric ZF domains bind DNA, and (ii) the binding of one ZF domain does not influence the binding of the other ZF domain in the functional dimeric protein. As a result, DNA-binding specificities of monomeric ZFN domains have been used to predict potential off-target cleavage sites of dimeric ZFNs. The methodology proposed by Pattanayak et al. (2011) identified several sites in the human genome that can be cleaved *in vitro* by two different ZFNs, they have then further examined these sites for evidence of ZFN-induced mutagenesis in cultured human cells expressing both ZFNs. The authors have confirmed that cleavage was likely to be dependent on cell type and ZFN concentration. Doyon et al. (2011) have also reported that transient hypothermia increased ZFN-mediated gene disruption in several cell lines between 1.5–15-fold. Although it is not clear yet whether this observation is generally applicable to any system, the authors showed that the environment (i.e. temperature shift), can influence ZFN protein levels in the cells and thus have an impact on nuclease activity (Doyon, et al., 2010). Similar results were also obtained by Pruett-Miller et al. (2009) who showed that by regulating protein levels, they could maintain high rates of ZFN-mediated gene targeting was maintained while ZFN toxicity was reduced.

Box2: The outcome product of nuclease-based techniques is preceded by several stages of the crop development. Each of these stages has potential unintended effects and uncertainties thereof. The following step-by-step approach is intended to contextualize and identify safety issues related to the use of SDN at all stages of a SDN-based crop development.

Step 1: What SDN design?

Points to consider:

- SDN type
- SDN type + Donor DNA

Questions to be addressed:

1. How to predict and detect off-target cleavage?
2. How to predict and detect mutation/insertion at target and off-target sites?
3. How to predict and detect positional effects?

Step 2: What delivery method?

Points to consider:

- Transgenesis
- Transient
- Direct SDN delivery

Questions to be addressed:

1. Is there foreign DNA introgression into host genome? Where and how?
2. Does the expression of SDN alters/affects host cell metabolism?

Step 3: What plant regeneration method?

Points to consider:

- Plant cell tissue culture
- No plant regeneration method applied

Question to be addressed:

1. Do these techniques apply mutagenic reagents?
2. Do these techniques alter host cell metabolism?

In regards to cell-type off-target cleavage dependency, other authors have also observed that ZFNs that can cleave at sites in one cell line may not necessarily function in a different cell line most likely because of local differences in chromatin structure (Maeder et al., 2008). In addition, their results indicate that excess DNA-binding energy results in increased off-target ZFN cleavage activity and suggest that ZFN specificity can be enhanced by (i) designing ZFNs with decreased binding affinity, (ii) by lowering ZFN expression and (iii) by choosing target sites that differ by at least three base pairs from their closest sequence relatives in the genome. Furthermore, the authors show that a four-finger ZFN showed a more diffuse range of cleavage positions with relaxed specificity and a higher tolerance of mutant sequences than a three-finger ZFN. Nonetheless, half or more of all potential substrates with one or two site mutations could be cleaved by ZFNs, suggesting that binding affinity between ZFN and DNA substrate is sufficiently high for cleavage to occur even with suboptimal molecular interactions at mutant positions (Pattanayak et al., 2011). That same study also reveals that *in vitro* spacer preferences do not necessarily reflect spacer preferences in cells. Their results suggest that the dimeric *FokI* cleavage domain can influence ZFN target-site recognition (Pattanayak et al., 2011).

Consistent with these results, Gabriel et al. (2011) have investigated the specificity of ZFN by a different method that was based on detecting off-target DSBs that have captured integrase-defective lentiviral vectors by NHEJ. These authors have analyzed genome-wide integration sites mapping *in vivo* cleavage activity of four ZFN pairs. Their results showed that the ranking of *in vivo* cleavage sites could not be predicted *in silico*. Thus, indicating the need for comprehensive maps of ZFN activity in *in vivo* models.

Other groups have also put their efforts in alternative strategies to decrease ZFN off-target cleavage (Miller et al., 2007; Doyon et al., 2011; Sander et al., 2013). Doyon et al. (2011) for example, have identified critical residues involved in off-target homo-dimerization activity of ZFN and they have successfully designed ZFN suppressing this homo-dimerization. An extensive review about the parameters (i.e. DNA-binding domain, DNA-cleavage domain, Inter-domain linker and others) that improve ZFN activity and specificity is provided by Händel and Cathomen (2011). These authors highlight the need for a proper assessment of each type of ZFN toxicity in order to be able to distinguish between cytotoxicity, genotoxicity, immunotoxicity or teratogenicity.

Undesired off-target cleavage has been also recognized for TALENs. According to several research groups, this remains an important issue and may cause severe side effects to the model organism under investigation (Hockemeyer et al., 2011; Mussolino et al., 2011; Osborn et al., 2013; Tesson et al., 2011).

Similar results to ZFNs were obtained when assessing TALEN specificity. Guilinger et al. (2014) have profiled 30 unique TALENs with varying target sites. Computational analysis predicted 76 off-target substrates in the human genome, 16 of which were accessible and modified by TALENs in human cells. In addition, their results suggested that (i) TALE repeats bind DNA relatively independently; (ii) longer TALENs are more tolerant of mismatches, yet are more specific in a genomic context; and (iii) excessive DNA-binding energy can lead to reduced TALEN specificity in cells (Guilinger et al., 2014). TALEN protein concentration has been also observed as a factor contributing to off-target cleavage. Tesson et al. (2011) have observed results suggesting dose-dependent mutation rates. TALE arrays are reported with on-target DNA binding affinities sufficient enough to theoretically saturate target sites even when expressed at modest, mid-nM concentrations in the cell (Meckler et al., 2013).

Fine et al. (2014) have recently highlighted that experimental identification of ZFN and TALEN off-target sites is a daunting task because of the availability and size of host genomes and the large number of potential cleavage sites to assay. More importantly, these authors also present literature from previous attempts to identify new off-target sites that were based entirely on bioinformatics search methods and that have all failed to locate any off-target cleavage sites, which has led to the belief that identifying off-target activity based on sequence homology alone would not be fruitful. Several time-consuming experimental approaches were then developed, such as SELEX, bacterial one-hybrid, *in vitro* cleavage or IDLV trapping (Fine et al., 2014). But the overall conclusion of the recent literature review is that although several experimental and theoretical methods have improved ZFN and TALEN activity, genome-wide specificity of ZFNs is still under scrutiny.

The historical record about the determinants of the specificity of CRISPR/Cas9-based techniques seems even less substantial than those for ZFNs and TALENs. Hallmark papers on comprehensive analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases date to a year ago or even later (Cradick et al. 2013; Fu et al., 2013; Cho et al. 2014; Fu et al., 2013; Pattanayak et al., 2013). Very recently, three groups independently showed that RNA-guided

nucleases (RGNs) indeed induce off-target mutations, even at sites that differ by 5 nt from target sites in human cells (Fu et al. 2013; Hsu et al. 2013; Pattanayak et al. 2013). Fu et al. (2013) have applied a human cell-based reporter assay to characterize off-target cleavage of CRISPR-associated Cas9-based RGNs. These authors found that off-target sites were mutagenized with frequencies comparable to (or higher than) those observed at the intended on-target site. In addition, their work demonstrates that RGNs can be highly active even with imperfectly matched RNA-DNA interfaces in human cells, a finding that might confound their use in research and other therapeutic applications.

Behaj et al. (2013) have reviewed targeted mutagenesis in model and crop plants that have used CRISPR/Cas system. These authors show the outcomes of four reports that have addressed CRISPR system to off-target activity when applied in plants. While Nekrasov et al. (2013) did not detect off-target activity in *N. benthamiana* for 18 off-sites with sequence similarity to the target; two of the four reports detected experimental evidence of off-target activity in rice (Shan et al., 2013; Xie et al., 2013).

These reports have shown that the 3' end of the guide sgRNA sequence is the one responsible for target specificity of the CRISPR/Cas system. In addition, researchers have also identified a protospacer adjacent motif (PAM) 20 bp right after the target sequence as essential for Cas9 cleavage activity. Mismatches located within the last 8-10 bp of the 20 bp target sequence often block target recognition by Cas9, while mismatches towards the 5' end of the target are better tolerated. Although this might work as a general rule, Hsu et al. (2013) reported that a noncanonical variant sequence of the PAM retains some activity and that the number and position of acceptable mismatches is target-dependent and should not be generalized to the reported rates. Most importantly, Behaj et al. (2013) highlight that the overall number of tested off-sites in all CRISPR/Cas studies was relatively small and general conclusions would be premature. These authors strongly suggested comprehensive studies based on whole genome sequencing of mutant plants in order to fully address the off-target cleavage issue *in planta*.

In conclusion, the recent literature demonstrates off-target cleavage activity of all nuclease types under different various experimental conditions. In addition, it is clear that there are recognized uncertainties regarding the probabilities and the characteristics by which these off-targets are generated. The difficulty in predicting these effects leads to the necessity of prior *in vivo* testing of these nucleases activities, as well as comprehensive untargeted profiling approaches (such as *omics* techniques) in order to detect these off-target cleavage sites in the host genome. The inheritance of such unintentional genomic changes is even less studied and cannot be assumed to vanish from the host genome organism just by crossing.

2.4.2. Specificity of double-stranded break repair mechanisms

Considering that DSBs can occur at target and off-target sites, it is necessary to verify intentional and unintentional mutations that might have resulted from the cell innate DNA repair mechanism at all DSB sites. As presented in Box 2, in principle, DSBs can be repaired via two main pathways: homologous recombination (HR) and non-homologous end-joining (NHEJ).

HR requires regions of DNA homology and repairs DSBs by using information on the undamaged sister chromatid or, in the case of genome editing techniques, the donor DNA. Whereas sequences for HR are linked via regions that are identical to one another, sequence information does not play a

major role in the rejoining of the two DSB ends for NHEJ. Therefore, NHEJ uses extremely limited or no sequence homology to rejoin juxtaposed ends in a manner that is not required to be error free (Wyman and Kanaar, 2006).

The question whether the application of a SDN into the plant cell will trigger HR or NHEJ will depend on many factors, and yet most of them are still unknown. This is because all breaks are not equally created and the particular repair mechanism to be used will depend largely on what is possible and needed based on the structure of the broken DNA (Wyman and Kanaar, 2006). In addition, it is important to mention that the frequency of SDN-mediated HR has been also shown to be dependent on the architecture of the donor DNA. Critical parameters range from the length of the homology arms and non-homologous sequences to the linear or circular DNA donor topology (Cornu et al., 2008; Orlando et al., 2010). The question to what extent the homology arm lengths as well as the topology and the configuration of the donor DNA influence the ratio between HR and illegitimate NHEJ has not been conclusively answered yet (Handel et al., 2011).

Consistent with these evidences, Ramirez et al. (2012) highlights that although HR-mediated alterations can be efficiently introduced using engineered nucleases, alleles can also acquire NHEJ-mediated mutations; thus, unwanted alterations at other off-target genomic sites can also be introduced by NHEJ-mediated repair. And as an alternative to the potential undesirable consequences of introducing DSBs in living cells, these authors propose a single-stranded break model (SSB or nicks) to genome editing. They have converted engineered ZFNs into zinc finger nickases by inactivating the catalytic activity of one monomer in a ZFN dimer. The concept applying nicks for stimulation of HR has been previously suggested in the context of theoretical models and recombination induced by RAG proteins (Lee et al., 2004).

A very recent comparative analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases has shown improvements in nuclease specificity (Cho et al., 2014). These authors have measured mutation frequencies of paired nickases and nucleases at off-target sites using deep sequencing. Cas9 nucleases induced off-target mutations at six sites that differ by 1 or 2 nt from their corresponding on-target sites with frequencies that ranged from 0.5% to 10%. In contrast, paired Cas9 nickases did not produce indels above the detection limit of 0.1% at any of the six off-target sites (Cho et al., 2014).

These mutations can have a variety of effect on endogenous gene expression (Pattanayak et al., 2011). Small deletions or insertions can cause gene knockout, or be completely ineffective from normal expression rate perspective among other impacts. Some of these mutations might not represent a phenotypic change and thus be detected by an unsupervised eye, but still trigger safety considerations. Furthermore, it cannot be assumed that they will not be inherited by daughter cells (Meng et al., 2008). It is clear at this stage that knowledge about the DSB DNA repair mechanisms applied by host cells that have been exposed to nucleases is still in its infancy. Most interestingly, the large majority of the studies are still based on mammalian models and not plant models. The off-target cleavage of such nucleases seems to be reaching a consensus in the scientific community but the outcomes of such breaks as well as the repair mechanisms are far from being elucidated.

3. Oligonucleotide-directed mutagenesis-based techniques

Summary

- Oligonucleotide-directed mutagenesis uses oligonucleotide to induce sequence specific modifications or mutations of native genomic sequences.
- The oligonucleotides are homologous to the target sequence with exception of few (1 to 5 bp/nt).
- The potential genetic changes are: introduction of mutation (replacement of few base pairs) and short deletions.
- ODM comprises a two-step mechanism: (1) DNA strand pairing and (2) DNA repair. DNA strand pairing is catalyzed by RAD54 protein family. DNA repair pathways are controversial. Proposed models to explain DNA repair pathways in ODM are: DNA strand pairing and formation of a D-loop region that triggers DNA repair pathways. Model 1: direct integration of base mismatches into the helix or corrected by DNA repair enzymes, probably mismatch-repair family. Or yet via nucleotide excision repair. Model 2: via transcriptional-coupled DNA repair enzymes, or nucleotide-excision repair systems or mismatch repair mediated by recombinases.
- There are four types of ODM approaches: SSO, RDOs, SFHR and TFOs. Each is composed by a different strandedness (single or double), size (from 10 to 2000-nt) and features (chemical modifications).
- There has been evidence about collateral and unspecific background mutations but several research groups highlighted the uncertainties regarding those. Unintended effects can be described as those related to cellular death, unpredicted mutations, mutation inheritance discrepancies, etc.

3.1. What are oligonucleotide-directed mutagenesis?

Oligonucleotide-directed mutagenesis (ODM) is an approach that uses oligonucleotide (25-200 nucleotides or base pairs (bp) in length) or polynucleotide (>200 nucleotides or bp in length) sequences to induce sequence-specific modifications (i.e. mutations) of endogenous genomic target sequences. Oligo/polynucleotide generally carry a limited number (between 1 and 5) of modifying bases within a segment of DNA that is otherwise homologous to the target sequences (Sargent et al., 2011). Most importantly, this strategy does not use viruses as delivery vehicles; neither does it aim to add more genetic material to the target gene (Liu et al., 2003) unless, otherwise combine to other genome editing or transgenesis techniques.

Moerschell, et al. (1988) were the first to show that single-stranded DNA oligonucleotides could direct *in vivo* base changes in *Saccharomyces cerevisiae*. To date, numerous publications report that synthetic oligonucleotides can induce small sequence changes in nuclear DNA of eukaryotic cells. The genetic changes that can be obtained using oligonucleotide-directed mutagenesis include the introduction of a new mutation (replacement of one or a few base pairs), the reversal of an existing mutation (gene repair) or the induction of short deletions in native genomic sequences (Lusser et al., 2011).

The general mechanism involved in oligonucleotide-directed mutagenesis-based gene repair is believed to encompass a two-step process, namely (i) DNA strand pairing and (ii) DNA repair (also called gene conversion). However, the exact mechanisms by which each of these pathways occur seems

controversial in the literature. There are two main models that explain the results obtained by several research groups. This report does not aim at deciphering the details of each proposed model, but to show discrepancy on understanding the modes-of-action by different research groups. This highlights knowledge gaps in the basic functioning of such technique and also shows that most assays are error-prone trials with large variation in the efficiency. In addition, most of these mechanistic studies are based on animal evidences and thus cannot be fully applied to plants (Yoshiyama et al., 2013).

The strand pairing is the first step for ODM. Strand pairing involves the synthetic molecule that is brought into homologous region at the site in the chromosome that is designated for change, and the event is catalyzed by members of the RAD54 epistasis group in plants (Osakabe et al., 2006). The invasion and binding of oligonucleotides with the target sequence resulting in the formation of a loop structure often referred to as D-loop (Laible et al., 2006). Whether this step is a rate-limiting step in the process of gene correction by oligonucleotides is yet controversial (Liu et al., 2003; Igoucheva et al., 2004a; Igoucheva et al., 2004b).

In the second step, after alignment, one or few base mismatches are present between the oligonucleotide molecule and the target strands. The following steps are controversial in the literature. In addition, it has been also presented evidence that the different types of oligonucleotide molecules might trigger distinct cell responses. Here it is presented a summary of the two most accepted mechanistic models for the DNA repair pathways. Figure 3 is a schematic representation and a summary of the main steps provided by each model.

- In one model, Kmiec's group believes that several pathways can be triggered by the oligonucleotide molecule once it is paired. The molecule can be simply detached from the genome and no modification is observed or, alternatively, the oligonucleotide assimilates into the target site and replaces one of the parental strands, perhaps integrating directly into the helix, and the synthetic oligonucleotide strand is further removed from the parental target strand by nuclease activity. In a third scenario, the mismatched base pair (formed between the vector and the target) is corrected or reversed by DNA repair enzymes, which are probably members of the DNA mismatch-repair family. There is, however, a fourth explanation from other data that implicate nucleotide excision repair as an active pathway in gene repair (for review see Liu et al., 2003).
- Yet another research group led by Igoucheva propose an alternative model for gene repair mechanism. In this model, an oligonucleotide is assimilated to the target DNA during active transcription, leading to formation of a transient D-loop structure. The trafficking of RNA polymerase is interrupted by the D-loop, and the stalled RNA polymerase complex may signal for recruitment of DNA repair proteins, including transcription-coupled DNA repair and nucleotide-excision repair (Igoucheva et al., 2004b).

But not only the mechanisms by which oligonucleotide-directed mutagenesis occur but also the nomenclature of such agents and approaches are not consensus, further complicating the literature search. Lusser et al. (2011), for instance, report countless descriptive and operational names to designate ODM, these are: oligonucleotide-mediated gene modification, targeted gene correction, targeted gene repair, RNA-mediated DNA modification, RNA-templated DNA repair, induced targeted mutagenesis, targeted nucleotide exchange, chimeroplasty, genoplasty, oligonucleotide-mediated gene editing, chimeric oligonucleotide-dependent mismatch repair, oligonucleotide-mediated gene

repair, triplex-forming oligonucleotides induced recombination, oligodeoxynucleotide-directed gene modification, therapeutic nucleic acid repair approach.

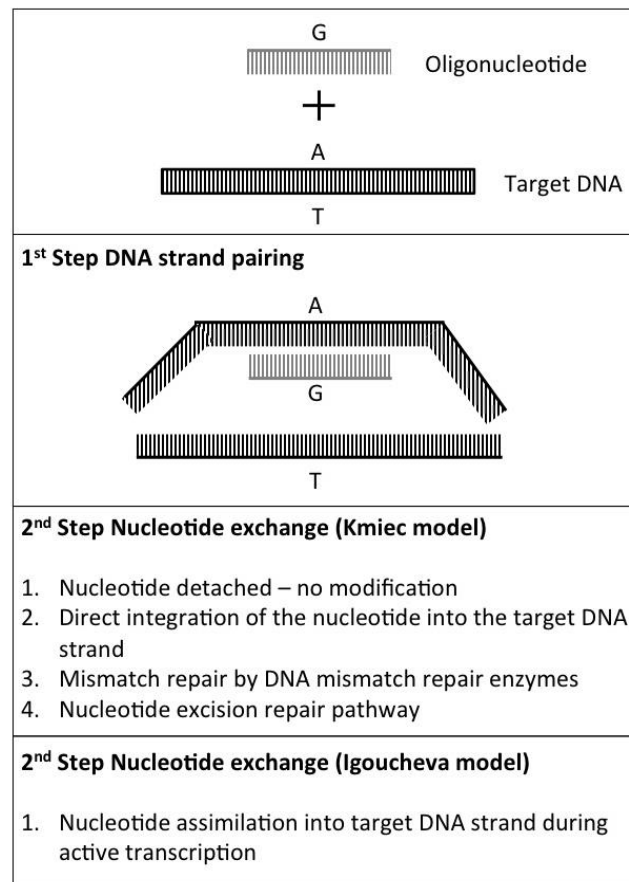


Figure 3. Schematic representation and summary points of the main theoretical models for oligonucleotide-directed mutagenesis (ODM) mode-of-action. The oligonucleotide is delivered into cells (first square), and it targets DNA strand, the process is known as the DNA pairing step (1st Step). The second step (DNA repair step) is explained by two main models. In Kmiec's model (Liu et al., 2003) four main possibilities take place in the cell and are based on the activity of mismatch repair enzymes or the direct exchange (integration or excision) of mismatched nucleotides. In Igoucheva's model (Igoucheva et al., 2004b), mismatched nucleotides are exchanged during transcription activity in a transient D-loop DNA structure.

At present, the potential applicability of gene repair by ODM seems to be limited by the lack of knowledge on how these molecules are functioning and, consequently, by low efficiency of correction. For most ODM specialized researchers, mechanistic studies to identify critical proteins involved in the gene correction process will undoubtedly play an important role for designing rational approaches toward increasing the targeting efficiency, efficacy and safety.

3.2. Types of oligonucleotide-directed mutagenesis

Different types of oligonucleotide-directed mutagenesis methods shows that different types of oligonucleotide molecules have been applied to various cell species with distinct cell-type versatility and apparent various efficiencies of gene modification.

In order to enhance the understanding about modes-of-action and implications of each described oligonucleotide type, we will follow the categorization provided by Igoucheva et al. (2004a), Laible et al. (2006) and Sargent et al. (2011). A summary of the main characteristics of each oligonucleotide type is provided in Table 2.

Oligonucleotide type	Components	Mode-of-action*
Single-stranded oligo-deoxynucleotides (SSOs or ssODMs)	ssDNA (<200-nt oligonucleotide)	DNA strand pairing and formation of a D-loop region that triggers DNA repair pathways.
Chimeric RNA-DNA oligonucleotide molecules (RDOs)	ssRNA-DNA or dsRNA-DNA oligonucleotide	Model 1: direct integration of base mismatches into the helix or corrected by DNA repair enzymes, probably mismatch-repair family. Or yet via nucleotide excision repair
Small Fragment Homologous Replacement (SFHR)	dsDNA ou ssDNA (200–2000-nt polynucleotide)	Model 2: via transcriptional-coupled DNA repair enzymes, or nucleotide-excision repair systems or mismatch repair mediated by recombinases.
Triple helix-forming oligonucleotides (TFOs)	ssDNA (10–40-nt oligonucleotide) that target poly-purine or polypyrimidine regions of DNA	

Table 2. Overview of oligonucleotide-directed nuclease (ODM) types, their main components and modes of action.

Note that 'bp' stands for base pairs. For review see Liu et al. (2003) and Sargent et al. (2011). Note: * The mode-of-action for each oligonucleotide type has not reached consensus in the literature. Therefore, this Table does not specify DNA pairing and DNA repair pathways triggered by each oligonucleotide. Table 2 summarizes major DNA repair models proposed by leading research groups.

3.2.1. Single-stranded oligo-deoxynucleotides (SSOs or ssODMs)

SSOs are single stranded oligonucleotides molecules of about <200 bp in length and comprised of a single mismatch to the target sequence that is generally in the middle of the molecule (Sargent et al., 2011). Although SSOs have been synthesized with phosphorothioate backbones, 2'-O-methyl uracil or with 50 or 30 thymidine clamps to inhibit degradation, they have also been used without modification to facilitate homologous exchange (Igoucheva et al., 2004a).

The first studies using homologous SSOs of 20-70 nucleotides containing a single mismatch were shown to trigger defined point mutations in mammalian cells and yeast (Simon and Moore, 1987; Moerschell et al., 1988; Campbell et al., 1989). The frequency of correction ranged from 10^{-7} from 10^{-3} (one in 10^{th} million cells to one in a thousand cells) in mammalian cells and yeast, respectively,

depending on the amount, length, and polarity of SSO, as well as the genetic background of the recipient yeast.

The activity of such small molecules seems to be influenced by several biological factors, thus leading to low efficiency rates (Laible et al., 2006). However, high correction rates of approximately 0.05% for nuclear extracts, 1% for episome, and 0.1% for chromosome of mammalian cells were observed by Igoucheva et al. (2001) when applying relatively short single-stranded oligodeoxynucleotides, 25–61 bases homologous to the target sequence except for a single mismatch to the targeted base in the mutant *b-galactosidase* gene. Igoucheva et al. (2004a) have made an extensive review of studies that utilize SSOs and concluded that a large variation in frequencies of gene alteration has been observed by researchers.

Accordingly, several studies on SSO do not report coherent or consistent results. While some studies suggest that transcription may be a factor in the SSO correction process due to a higher efficiency of correction with the anti-sense non-transcribed SSO target (Igoucheva et al., 2003; Pierce et al., 2003); others attribute this result to the context of chromatin accessibility due to the integration of ODM transgenes in different regions of the genome. According to Sargent et al. (2011) these would represent chromosomal environment differences in terms of gene expression. On the other hand, Dekker et al. (2003) also suggested that there are no strand-associated differences in SSO-mediated targeting.

In conclusion, the methodology seems to be currently limited by its low frequency of repair events, variability amongst different experimental settings, and low viability of mutated cells. But whether contradictory findings reflect differences in the cells used in the targeting studies requires further analysis (Sargent et al., 2011).

3.2.2. Chimeric RNA-DNA oligonucleotide molecules (RDOs)

Chimeric RNA-DNA oligonucleotides are composed of a contiguous stretch of RNA and DNA residues has been developed to facilitate correction of single-base mutations in mammalian cells. The original design of the chimeric RNA-DNA oligonucleotide consisted of a double-hairpin capped duplex comprising a 25-nucleotide-long DNA stretch (DNA-strand) containing the mismatch, paired to a fully complementary RNA stretch (RNA-strand) with a pentameric DNA interruption in the middle (Igoucheva et al., 2004a).

While the first constructs were double-stranded oligonucleotides that consisted of a one DNA strand and another strand with a central DNA stretch but terminal regions of RNA, the second generation RDOs single-stranded DNA oligonucleotides with terminal RNA stretches or simply uracil residues (Liu et al., 2003). The idea of replacing the complementary DNA strand to a RNA strand emerged from the observation of previous studies that showed increased efficiency in RNA-DNA chimeric molecules.

Chimeric double-stranded oligonucleotides are designed to be self-complementary and produce double-hairpin structures. Within the duplex conformation, two complementary strands can be distinguished: a chimeric strand consisting of an interposed DNA fragment (5 bp) and two stretches of RNA (10 bp) flanking this region, and another strand that is composed only of DNA contains a mutation relative to the homologous target sequence (Kochevenko and Willmitzer, 2003).

The first experimental evidence of successful gene correction by RDO was demonstrated in 1996 by correction of a point mutation in alkaline phosphatase cDNA in Chinese hamster ovary (CHO) cells (Yoon et al., 1996). The RNA-DNA sequence was designed to align with the sequence of the mutant alkaline phosphatase locus and to contain the desired nucleotide change. The chimeric molecule was introduced into human cells containing the mutant gene on an extra-chromosomal plasmid, correction of the point mutation was accomplished with a frequency approaching 30% (Yoon et al., 1996). The design of this RNA-DNA hybrid molecule makes them highly active in homologous pairing reactions *in vitro* and the observations suggested that recombination of these exogenous RNA-DNA sequences is protected from exonucleolytic degradation by capping both ends (Cole-Strauss et al., 1995).

Some experiments have extended the utility of RDO molecules to plants. Apparently, this seems to be the preferred oligonucleotide types used in plant studies. For example, Kmiec et al. (2001) used a plant cell-free extracts system to study DNA-repair mutants and the role of plant proteins in the DNA repair process. In this assay, cell-free extracts were mixed with a plasmid containing a specific mutation (point or frameshift mutations). The authors also used a genetic readout system in bacteria and chimeric or modified DNA oligonucleotides designed to direct the conversion of mutations in antibiotic resistance genes in a spinach chloroplast lysate system. They reported the genetic repair of a point frameshift mutation directed by RDOs.

Interesting to note, large variations among RDO experiments were observed and these have been attributed to the use of different species, cell types, genes, methods of delivery and assays, which have hindered efficiency comparisons (Igoucheva et al., 2004a).

3.2.3. Small Fragment Homologous Replacement (SFHR)

A widely used approach to generate transgenic animals was based on the delivery of homologous DNA and selectable markers to embryonic stem cells. In principle, a portion of a gene can be introduced into the nucleus allowing pairing of a DNA fragment with its homologous loci by means of proteins involved in HR, thus modifying genomic DNA in a sequence-specific manner (Igoucheva et al., 2004a). This approach has been further refined to the use of SFHRs for relatively high level of targeted correction (1%) and achieve therapeutic level.

SFHRs are distinct from SSO because they are larger molecules (generally between 200 and 2000 bp/nt in length) and can be individual ssDNA, complementary ssDNA, or double-stranded DNA (Sargent et al., 2011). In addition, SFHR relies on two main features (i) a SFHR carrying a single or multiple base alterations, finding its sequence homolog, and (ii) the cellular enzymatic pathways that facilitate homologous exchange between target sequences and the SFHR. Although there has been speculation about the mechanism(s) that underlie SFHR, there is still only limited knowledge of the cellular factors that influence SFHR (Sargent et al., 2011). To the best of our knowledge, there have been only a limited number of studies that have explored the mechanisms that underlie SFHR-mediated modification. Even fewer apply SFHR in plants.

3.2.4. Triple helix-forming oligonucleotides (TFOs)

TFOs are single-stranded oligonucleotides, generally 10–40 nt in length, that bind to specific regions in duplex DNA as a third strand to form a triple helix at poly-purine or poly-pyrimidine regions of DNA bound via Hoogsteen hydrogen bonds (Sargent et al., 2011).

The triple-helix-forming oligonucleotide recognizes the sequence surrounding a targeted base and the coupled reactive group or DNA elicits DNA repair and/or recombination process, resulting in sequence alteration. Such oligonucleotides have been used previously to change targeted DNA sequences by approximately 1% but are limited by the restrictions of the target sequence, which must consist of homopurine or homopyrimidine stretches for triplex formation (Igoucheva et al., 2001).

TFOs have a bipartite structure that contains sequences that are favorable for triple helix formation and donor DNA sequences homologous to the chromosomal target sequence to be modified. Although covalent linkage of the TFO portion of the oligonucleotide was initially used to orient the contiguous DNA sequence to make available for repair of the DNA, most studies now use unlinked TFOs and donor oligo/ polynucleotides to stimulate HR and/or other DNA repair pathways (Sargent et al., 2011). In addition, TFOs coupled to functional chemical groups or double-stranded DNA have been also used to alter, specifically or non-specifically, genomic DNA in cultured cells and *in vivo* (Pierce et al., 2003).

Studies evaluating the enzymatic pathways that may underlie the TFO-mediated homologous exchange indicate that nucleotide excision repair is involved (Christensen et al., 2008). We have not found plant examples utilizing TFOs in the recent literature. Many examples can be found in mammalian and other animal models (Igoucheva et al., 2004b).

TFOs can also be utilized to create permanent heritable changes in the genome by both site-specific mutagenesis and site-specific recombination. However, mechanistic questions remain unanswered. The exact mechanisms behind cellular processes that play a role in mediating these genetic changes are unknown. There is a need to determine which factors can influence the availability of genomic target to DNA binding molecules, transcriptions levels, cell cycle phase, nucleosome binding and histone state that may all play important roles (Knauert and Glazer, 2001; Sargent et al., 2011).

3.3. Identification of unintended effects and consideration of uncertainty of ODM-based crops

The identification of similarities and differences in the mechanisms underlying the various ODM gene-targeting approaches has been hindered by various experimental factors, such as the variability in the exogenous DNA, the genetic target, the cell systems employed, and the approaches used to analyze the efficiency of homologous exchange.

Similar to SDN techniques, ODM-mediated homologous exchange process relies on the native cellular machinery to perform the exchange between the exogenous nucleotide sequences and the endogenous target sequences. And since there are multiple DNA repair and replication pathways that depend not only on the type of DNA modification, but also on the cell type and the stage of the

cell, it is difficult to generalize about the involvement of specific mechanisms in the homologous exchange process for an individual strategy (Sargent et al., 2011).

The situation becomes even more complicated when most of the ODM examples were performed in animal models, with a very few applications in plants. With that in mind, it is relevant to state that the discussion on unintended effects below are restricted to animal examples just because this has never been studied in plant species.

It is clear that the events of the second phase where the 'correction' of the targeted mutant gene actually occurs are poorly understood. The current challenge centers on the fate of the oligonucleotide paired at the target site and its role in the correction event. From Parekh-Olmedo et al. (2005) perspective, the answers may lead to an understanding of not only how it works but also which cellular factors are involved in its regulation.

According to these same authors, as more and more systems are developed, the central questions about this technology are changing from "does it work?" and "how does it work?" to "what happens if it works?".

Olsen et al. (2003) have tested a direct *in vivo* gene modification approach using a CHO cell assay containing a stable integrated mutated enhanced green fluorescent protein (EGFP) reporter. These authors observed specific correction of mutated EGFP to a functional wild-type sequence after oligonucleotide administration *in vivo*. Most interestingly, they have observed that the repair of site-specific DS break by homologous recombination is severely reduced in daughter cells compared to the paternal V79-4 cell line. The authors further mention the fact that different chromosomal integration of the EGFP transgene in daughter and parental cell lines might have influenced these results. They also observed that both predicted and unpredicted alterations have occurred in the mEGFP2 template. In addition, 9.3% of the amplified sequences have a corrected stop codon but also show additional flanking mutations that are not contained in the oligonucleotide. The interesting work of Olsen et al. (2003) have also shown high apoptosis rates that has been suggested to be coupled to the targeted sequence alteration rather than to general toxicity. But whether the apoptosis observed in the experiments is a general problem associated with the presence of large amounts of oligonucleotides in cells, whether it is a more specific process related to cells that have undergone sequence conversion and whether any of the above mechanisms account for the observed selective apoptosis should be further investigated.

Liu et al. (2009) have also observed that during SSO-mediated gene repair reaction, a significant population of corrected cells failed to divide, and were much more prone to undergo apoptosis. In addition, their results showed that while a defective mismatch repair mutant greatly enhanced the efficiency of gene correction, compromising the mismatch repair system did not yield any viable corrected clone, indicating that the mismatch repair machinery, although plays a critical role in determining SSO-directed repair, was not involved in the observed cellular genotoxic responses (Liu et al., 2009).

On the other hand, Parekh-Olmedo et al. (2005) highlights that such cellular apoptotic responses should not be surprising but rather predicted due to the common high oligonucleotide dosage delivery into cells. These authors also recall the fact that it is widely recognized that free ends induce a DNA damage response, which leads to cell cycle arrest, replication fork stalling and the activation of repair pathways, including HR. And that the combination of these activities determines the fate of

the cell; whether the damage will be repaired or the cells will be programmed for death (Parekh-Olmedo et al., 2005).

Not only the amount of oligonucleotide delivered into cells, but also length and polarity have shown to highly influence chromosomal gene corrections. The strand orientation, sense or anti-sense, showed a 1000-fold difference in frequency, indicating a possible influence of transcription on gene correction (Igoucheva et al., 2001).

The large variability in the results obtained by these experiments shows the numerous factors that influence the functioning of ODM and the huge knowledge gaps surrounding them. In addition, some of these gaps have been followed by the application of wrong experimental methods. Igoucheva et al. (2004a) highlights that most studies have detected the gene conversion event by PCR amplification of genomic DNA, without direct confirmation at the protein level. Consequently, concerns have been raised about a potential artifact during PCR amplification or by cross-contamination between cells.

Yoon et al. (2004) suggest that the delivery and quality of an oligonucleotide, two favorite candidates for the reported variation among different systems, are not sufficient to contribute the large variation seen in experiments. These authors revealed that mechanistic studies from their group and others showed that the DNA-strand of oligonucleotide, for examples, might have strong influence in the gene repair activity. They further suggest that, based on these results, the active domain (DNA-strand) needs to be exposed at the right time and at the right locale to the target DNA for successful gene correction. But since the processing of an oligonucleotide by nucleases or helicases to expose the active domain is random, the subsequent strand-pairing reaction to the target DNA is also random and hard to control (Yoon et al., 2004).

There seems to be increasing evidence that links cellular toxicity to unintended effects from oligonucleotide delivery into cells. But a specific focus on efficacy of target correction. This might be explained by the uncertainties on the modes-of-action of such molecules. It seems that, unless more experimental evidence elucidates the mechanisms by which these molecules function, robust investigation about possible unintended effects will remain marginal.

4. Conclusions

New plant breeding techniques have emerged as a group that is different to both conventional breeding approaches and standard genetically modification or genetic engineering developments. Regardless of their regulatory status, which will depend on national legislation, this reports shows that the application and adaptation of current existing GMO safety regulations might be of great importance to identify the potential risks of such new products.

However, a part from the criteria present in both EFSA Scientific opinion on guidance for risk assessment of food and feed from genetically modified plants (EFSA, 2011) and EFSA Guidance on the environmental risk assessment of genetically modified plants (EFSA, 2010), other criteria should be incorporated into the risk assessment that might be of help to address potential hazards associated with the application of SDN and ODM techniques. In addition, a number of cross cutting issues have been already discussed in previous sections, which are also important for the evaluation of such techniques.

Specific aspects relevant for the risk assessment of SDN and ODM crops can be summarized as follows: (i) the kind of DNA/genome or RNA/transcriptome modifications introduced into the host genome, (ii) the kind of traits generated by the application of such techniques, (iii) the presence of exogenous DNA or RNA sequences, and (iv) alteration in gene expression. Although some of these issues have been already addressed by EFSA Scientific opinion addressing the safety assessment of plants developed using Zinc Finger Nuclease-3 and other Site-Directed Nucleases with similar function (EFSA, 2012), a more comprehensive analysis has been made by the Austrian Environmental Agency this year (Eckerstorfer et al., 2014).

Potential risks regarding the use of both SDN and ODM techniques are related to unintentional effects and uncertainties thereof. It is important to keep in mind that, in order for a crop to be commercially developed, not only these techniques are applied. Other techniques, such as transfection or transformation protocols, as well as plant regeneration processes are commonly needed. It is then relevant to state that this report only focuses on the risk issues related to the new technologies *per se*. But when performing a risk assessment for decision-making processing of approval or not of a crop product, the organism should be taken as a whole, on a case-by-case basis.

Some of the products derived from new emerging plant breeding techniques might be regulated differently from country to country according to their national legislation (Podevin et al., 2012; Camacho et al., 2014; Voytas and Gao, 2014; Jones, 2015). In addition, some of these new techniques might also be within the scope and methods referred as 'synthetic biology' (e.g. SDNs) .In fact, an international effort is being made on the discussion on whether or not these techniques should be addressed along with GM international regulation and guidelines. For instance, the latest report from the Ad Hoc Technical Expert Group on Risk Assessment and Risk Management under the Cartagena Protocol on Biosafety (AHTEG, 2014) has raised concerns regarding the need for further guidance on products from synthetic biology. Simultaneously, The Subsidiary Body on Scientific, Technical and Technological Advice under the Convention on Biological Diversity (CBD) has drafted two reports on the potential impacts and possible gaps and overlaps with the applicable provisions of the Convention, its protocols and other relevant agreements related to components, organisms and products resulting from synthetic biology techniques (SBSTTA, 2014a; SBSTTA, 2014b). These draft reports highlights the disagreement among synthetic biologists, ecologists, industry and civil society,

on how well the potential dangers related to synthetic biology are known and if there is duty to undertake an environmental impact assessment. But on the other hand, these reports reference scientific opinions about the duty to carry out an environmental impact assessment for industrial activities that may have a significant adverse impact in a transboundary context as an important development that might require clarification as to its precise implications. In accordance with Article 14 of the Convention, governments should take a precautionary approach, when addressing threats of significant reduction or loss of biological diversity posed by organisms, components and products resulting from synthetic biology, in accordance with domestic legislation and other relevant international obligations.

The two techniques reviewed in this report presented robust evidence of unintentional effects and uncertainties thereof and those were mainly related to the potential off-target mutation and the unpredictable nature of the cell native repair mechanisms. The latter is related to the unknown mutation outcome of each target and off-target site.

The recommendation on the acceptability of the potential risks should take into account any available scientific analysis of potential benefits for the environment, biodiversity, and human health and should also take into account risks associated with other existing user practices and habits (CBD, 2012). Further, the sources and nature of uncertainty that could not be addressed during the preceding steps of the risk assessment should be described in relation to how they could affect the conclusions of the risk assessment. For assessments where uncertainties could not be addressed, difficulties encountered during the risk assessment should be made transparent to the decision makers (CBD, 2012).

Several comparative risk analyses have been proposed and accepted by many regulatory bodies for the risk assessment of GMOs. These include a comprehensive molecular characterization, environmental and health risk assessment. To the best of our knowledge, the living products from the used of such techniques should be assessed for risks.

Further care should be taken regarding the molecular characteristics of such products, as some of them might not now be distinguishable from the products of other kinds of techniques. In addition, these techniques are too new to make strong claims that all outcomes are predictable and known.

5. Recommendations for follow up and additional research needs

This report aims at providing the state-of-art of the research into the potential applications and knowledge gaps of two emerging new breeding techniques, site-directed mutagenesis and oligonucleotide directed mutagenesis. After collecting available data and carrying out specific evaluation of the current literature, we concluded by focusing on the identification of knowledge gaps and uncertainties that additional research is needed to answer biosafety questions that have not yet been properly addressed. While most attention has been given in the past years to issues concerning the regulatory status of new emerging breeding techniques, biosafety aspects have received much less attention.

In order to contribute fulfilling open biosafety questions, it is our recommendation that a risk assessment should consider that:

- The current lack of knowledge and uncertainties regarding genome location of intended and unintended modification should be taken into account; further research should be developed to investigate enzyme binding sites specificity focusing in the detection of small nucleotide insertions and deletions across the genome by the use of high throughput genome sequencing technologies. A similar approach should be also developed to detect minimal sequence complementarity of oligonucleotide mutagenesis. The results of such analysis should be also compared to standard used bioinformatics tools in order to verify the prediction parameters and algorithms of such software.
- Considering that the genetic modification resulting from the application of such enzymes will also depend on the way the cell machinery responds to DNA lesions and that insufficient knowledge is available; further research should be developed to investigate what are the relevant factors affecting DNA repair mechanisms. Several different DNA lesions types should be facilitated and each repair mechanism should be followed up by sequencing analysis in order to be able to correlate the type of DNA lesion to a specific DNA repair mechanism. External factors such as genetic background, cell environmental conditions as well as enzyme types should be also addressed under highly controlled conditions.
- Acknowledging that plant-to-plant natural variability and the complexities of how plant genomes interact might mask the biological relevance of unintentional genetic changes; further research should be performed in order to define biological and genetic controls and materials to serve as proper comparators for future risk assessments. A comparison between a conventional near isogenic line and a genetically modified line created through the use of any of these new breeding technologies by the use of “*omics*” approaches (i.e. transcriptomics, proteomics, mirnomics, etc) will contribute to building baseline knowledge on natural variability of plant varieties, the parameters that might influence that variability and relevant molecular changes that should be further investigated for their safety.
- Considering currently existing labeling regimes; further research should be also developed to investigate the practical and technical constrains of detecting products of new breeding technologies. Since some literature has indicated the full lack of detectability of such

products, it has been also indicated that traceability can only be ensured by continuous documentation. In that sense, not only laboratory detection and identification techniques should be investigated but also the regulatory framework has to be discussed.

Overall, the results of this report indicate that biosafety considerations regarding new plant breeding technologies could, in principle, be addressed by the general approach developed for conducting the risk assessment of genetically modified crops. However, it is important to inform that insufficient knowledge is available about both of the technologies here under analysis regarding their potential adverse effects or unknown level of risk. Therefore, according to the requirements of a scientifically based risk assessment and the application of the precautionary principle, further biosafety research needs to be performed *a priori* to commercial release.

6. References

- AHTEG. (2014, June 6) Report of the Ad Hoc Technical Expert Group on risk assessment and risk management under the Cartagena Protocol on Biosafety. <http://www.cbd.int/doc/meetings/bs/bsrarm-05/official/bsrarm-05-06-en.pdf>: *United Nations Environment Programme Convention for Biodiversity* UNEP/CBD/BS/AHTEG-RA&RM/5/6, 1-14.
- Ashworth, J. *et al.* Computational redesign of endonuclease DNA binding and cleavage specificity. *Nature* **441**, 656–659 (2006).
- Belhaj, K., Chaparro-Garcia, A., Kamoun, S. & Nekrasov, V. Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant Methods* **9**, 39 (2013).
- Bibikova, M., Beumer, K., Trautman, J. K. & Carroll, D. Enhancing Gene Targeting with Designed Zinc Finger Nucleases. *Science* **300**, 764 (2003).
- Bitinaite, J., Wah, D. A., Aggarwal, A. K. & Schildkraut, I. FokI dimerization is required for DNA cleavage. *Proc. Natl Acad. Sci. USA* **95**, 10570–10575 (1998).
- Boch, J. & Bonas, U. Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annu. Rev. Phytopathol.* **48**, 419–36 (2010).
- Boch, J. *et al.* Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509–12 (2009).
- Bogdanove, A. J., Schornack, S. & Lahaye, T. TAL effectors: finding plant genes for disease and defense. *Curr. Opin. Plant Biol.* **13**, 394–401 (2010).
- Camacho, A., Van Deynze, A., Chi-Ham, C. & Bennett, A. B. Genetically engineered crops that fly under the US regulatory radar. *Nat. Biotechnol.* **32**, 1087–91 (2014).
- Campbell, C. R. *et al.* Homologous recombination involving small single-stranded oligonucleotides in human cells. *New Biol* **1**, 223–227 (1989).
- Cathomen, T. & Joung, J. K. Zinc-finger nucleases: the next generation emerges. *Mol. Ther.* **16**, 1200–7 (2008).
- CBD. (2012, July). Guidance on Risk Assessment of Living Modified Organisms. <http://www.cbd.int/doc/meetings/bs/mop-06/official/mop-06-13-add1-en.pdf>: *United Nations Environment Programme Convention for Biodiversity*, UNEP/CBD/BS/COP-MOP/6/13/Add.1, 1-64.
- Cermak, T. *et al.* Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* **39**, e82 (2011).
- Chevalier, B.S. & Stoddard, B.L. Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Res.* **29**, 3757–3774 (2001).
- Chevalier, B.S. *et al.* Design, activity and structure of a highly specific artificial endonuclease. *Mol. Cell* **10**, 895–905 (2002).

- Cho, S. W. *et al.* Analysis of off-target effects of CRISPR / Cas-derived RNA-guided endonucleases and nickases. 132–141 (2014). doi:10.1101/gr.162339.113.Freely
- Christensen, L. A., Wang, H., Van Houten, B. & Vasquez, K. M. Efficient processing of TFO-directed psoralen DNA interstrand crosslinks by the UvrABC nuclease. *Nucleic Acids Res.* **36**, 7136–7145 (2008).
- Christian, M. *et al.* Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* **186**, 757–761 (2010).
- Cole-strauss, A. *et al.* Correction of the mutation responsible for sickle cell anemia by an RNA-DNA oligonucleotide. *Science* **273**, 1386-9 (1996).
- Cornu, T. I. *et al.* DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. *Mol. Ther.* **16**, 352–8 (2008).
- Cradick, T. J., Fine, E. J., Antico, C. J. & Bao, G. CRISPR/Cas9 systems targeting β -globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res.* **41**, 9584–92 (2013).
- Curtin, S. J., Voytas, D. F. & Stupar, R. M. Genome engineering of crops with designer nucleases. *Plant Genome J.* **5**, 42 (2012).
- Dekker, M. Targeted gene modification in mismatch-repair-deficient embryonic stem cells by single-stranded DNA oligonucleotides. *Nucleic Acids Res.* **31**, e27 (2003).
- Doyon, Y. *et al.* Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat. Methods* **8**, 74–9 (2011).
- Eckerstorfer, M., Miklau, M. & Gaugitsch, H. New plant breeding techniques and risks associated with their application. *Umweltbundesamt Report REP-0477*, 1-92 (2014). http://www.umweltbundesamt.at/aktuell/publikationen/publikationssuche/publikationsdetail/?pub_id=2054
- EFSA. EFSA Panel on Genetically Modified Organisms (GMO): Guidance on the environmental risk assessment of genetically modified plants. *EFSA Journal* **8**, 1– 111 (2010).
- EFSA. EFSA Panel on Genetically Modified Organisms (GMO): Scientific opinion on guidance for risk assessment of food and feed from genetically modified plants. *EFSA Journal* **9**, 2150 (2011).
- EFSA. EFSA Panel on Genetically Modified Organisms (GMO): Scientific opinion addressing the safety assessment of plants developed using Zinc Finger Nuclease 3 and other Site-Directed Nucleases with similar function. *EFSA Journal* **10**, 2943 (2012).
- Feng, Z. *et al.* Efficient genome editing in plants using a CRISPR/Cas system. *Cell Res.* **23**, 1229-1232 (2013).
- Fine, E. J., Cradick, T. J., Zhao, C. L., Lin, Y. & Bao, G. An online bioinformatics tool predicts zinc finger and TALE nuclease off-target cleavage. *Nucleic Acids Res.* **42**, e42 (2014).
- Fu, Y. *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* **31**, 822–6 (2013).

- Gabriel, R. *et al.* An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat. Biotechnol.* **29**, 816–23 (2011).
- Gao, H. *et al.* Heritable targeted mutagenesis in maize using a designed endonuclease. *Plant J.* **61**, 176–87 (2010).
- Guilinger, J. P. *et al.* Broad specificity profiling of TALENs results in engineered nucleases with improved DNA-cleavage specificity. *Nat. Methods* **11**, 429–35 (2014).
- Händel, E. & Cathomen, T. Zinc-finger nuclease based genome surgery : It ' s all about specificity. *Curr Gene Ther.* **11**, 28–37 (2011).
- Heath, P. J., Stephens, K. M., Monnat, R. J., Jr. & Stoddard, B.L. The structure of I-Crel, a group I intron-encoded homing endonuclease. *Nat. Struct. Biol.*, **4**, 468-476 (1997).
- Hockemeyer, D. *et al.* Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. Biotechnol.* **29**, 731–734 (2011).
- Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–32 (2013).
- Igoucheva, O. Transcription affects formation and processing of intermediates in oligonucleotide-mediated gene alteration. *Nucleic Acids Res.* **31**, 2659–2670 (2003).
- Igoucheva, O., Alexeev, V. & Yoon, K. Oligonucleotide-directed mutagenesis correction : a mechanistic point of view targeted. *Curr Mol Med.* **4**, 445–463 (2004a).
- Igoucheva, O., Alexeev, V. & Yoon, K. Review mechanism of gene repair open for discussion. *Oligonucleotides* **14**, 311-321(2004b).
- Igoucheva, O., Alexeev, V. & Yoon, K. Targeted gene correction by small single-stranded oligonucleotides in mammalian cells. *Gene Ther.* **8**, 391–399 (2001).
- Jia, H. & Wang, N. Targeted genome editing of sweet orange using Cas9/sgRNA. *Plos One* **9**, e93806 (2014).
- Jiang, W. *et al.* Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Res.* **41**, e188 (2013).
- Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–21 (2012).
- Jinek, M. *et al.* Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* **343**, 1247997 (2014).
- Jones, H. D. Regulatory uncertainty over genome editing. *Nature Plants* **1**, 14011 (2015).
- Jurica, M. S., Monnat, R. J., Jr. & Stoddard, B.L. DNA recognition and cleavage by the LAGLIDADG homing endonuclease I-Crel. *Mol. Cell* **2**, 469-476 (1998).
- Kmiec, E. B., Johnson, C. & May, G. D. Chloroplast lysates support directed mutagenesis via modified DNA and chimeric RNA / DNA oligonucleotides. *Plant J.* **27**, 267-74 (2001).

- Knauert, M. P. & Glazer, P. M. Triplex forming oligonucleotides : sequence-specific tools for gene targeting. *Hum Mol Genet.* **10**, 2243–2252 (2001).
- Kochevenko, A. & Willmitzer, L. Modification of the tobacco acetolactate synthase gene. **132**, 174–184 (2003).
- Krishna, S. S. Structural classification of zinc fingers: Survey and Summary. *Nucleic Acids Res.* **31**, 532–550 (2003).
- Kumar, V. & Jain, M. The CRISPR-Cas system for plant genome editing: advances and opportunities. *J. Exp. Bot.* (2014).
- Laible, G., Wagner, S. & Alderson, J. Oligonucleotide-mediated gene modification and its promise for animal agriculture. *Gene* **366**, 17–26 (2006).
- Lee, G. S., Neiditch, M. B., Salus, S. S. & Roth, D. B. RAG proteins shepherd double-strand breaks to a specific pathway, suppressing error-prone repair, but RAG nicking initiates homologous recombination. *Cell* **117**, 171–184 (2004).
- Li, J.-F. *et al.* Multiplex and homologous recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. *Nat. Biotechnol.* **31**, 688–91 (2013a).
- Li, T., Liu, B., Spalding, M. H., Weeks, D. P. & Yang, B. High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nat. Biotechnol.* **30**, 390–2 (2012).
- Liu, C. *et al.* Cell death caused by single-stranded oligodeoxynucleotide-mediated targeted genomic sequence modification. *Oligonucleotides* **19**, 281–6 (2009).
- Liu, L. Parekh-Olmedo, H. & Kmiec, E.B. The development and regulation of gene repair. *Nat Rev Genet.* **4**, 679–89 (2003).
- Lloyd, A., Plaisier, C. L., Carroll, D. & Drews, G. N. Targeted mutagenesis using zinc-finger nucleases in Arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 2232–2237 (2005).
- Lusser, M., Parisi, C., Plan, D. & Rodríguez-Cerezo, E. (2011, May) New plant breeding techniques: State-of-the-art and prospects for commercial development. *JRC Scientific and Technical Reports* EUR 24760 EN, 1–219.
- Lusser, M., Parisi, C., Plan, D. & Rodríguez-Cerezo, E. Deployment of new biotechnologies in plant breeding. *Nature Biotechnology* **30**, 231–239 (2012).
- Maeder, M. L. *et al.* Rapid “open-source” engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol Cell.* **31**, 294–301 (2008).
- Mahfouz, M. M. & Li, L. TALE nucleases and next generation GM crops. *GM Crops* **2**, 99–103 (2011).
- Mahfouz, M. M. *et al.* De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 2623–8 (2011).

- Makarova, K. S. *et al.* Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* **9**, 467–77 (2011).
- Mali, P. *et al.* CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature Biotechnology* **31**, 833–838 (2013).
- Mao, Y. *et al.* Application of the CRISPR-Cas system for efficient genome engineering in plants. *Mol. Plant* **6**, 2008–2011 (2013).
- Meckler, J. F. *et al.* Quantitative analysis of TALE-DNA interactions suggests polarity effects. *Nucleic Acids Res.* **41**, 4118–28 (2013).
- Meng, X. *et al.* Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. *Nat. Biotechnol.* **26**, 695–701 (2008).
- Miao, J. *et al.* Targeted mutagenesis in rice using CRISPR-Cas system. *Cell Res.* **23**, 1233–1236 (2013).
- Miller, J. C. *et al.* An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat. Biotechnol.* **25**, 778–85 (2007).
- Miller, J. C., McLachlan, A. D. & Klug, A. Repetitive zinc-binding domains in the protein transcription factor IIIA from xenopus oocytes. *EMBO J.* **4**, 1609–1614 (1985).
- Moerschell, R. P., Tsunasawa, S. & Sherman, F. Transformation of yeast with synthetic oligonucleotides. *Proc. Natl. Acad. Sci. USA* **85**, 524–528 (1988).
- Moscou, M. J. & Bogdanove, A.J. A simple cipher governs DNA recognition by TAL effectors. *Science* **326**, 1501 (2009).
- Mussolino, C. *et al.* A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.* **39**, 9283–93 (2011).
- Nekrasov, V. *et al.* Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. *Nat Biotechnol.* **31**, 691–3 (2013).
- Olsen, P. A, McKeen, C. & Krauss, S. Branched oligonucleotides induce in vivo gene conversion of a mutated EGFP reporter. *Gene Ther.* **10**, 1830–40 (2003).
- Orlando, S. J. *et al.* Zinc-finger nuclease- driven targeted integration into mammalian genomes using donors with limited chromosomal homology. *Nucleic Acids Res.* **38**, e152 (2010).
- Osakabe, K. *et al.* Isolation and characterization of the RAD54 gene from *Arabidopsis thaliana*. *Plant J.* **48**, 827–42 (2006).
- Osborn, M.J. *et al.* TALEN-based gene correction for epidermolysis bullosa. *Molecular Therapy.* **21**, 1151–9 (2013).
- Pâques, F. & Duchateau, P. Meganucleases and DNA double-strand break-induced recombination: Perspectives for gene therapy. *Curr Gene Ther* **7**, 49–66 (2007).

- Pâques, F. & Haber, J.E. Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* **63**, 349-404 (1999).
- Parekh-Olmedo, H., Ferrara, L., Brachman, E. & Kmiec, E. B. Gene therapy progress and prospects: targeted gene repair. *Gene Ther.* **12**, 639–46 (2005).
- Pattanayak, V. *et al.* High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nature Biotechnology* **31**, 839–843 (2013).
- Pattanayak, V., Ramirez, C. L., Joung, J. K. & Liu, D. R. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nature Methods* **8**, (2011).
- Pennisi, E. The CRISPR craze. *Science* **341**, 833-6 (2013).
- Pierce, E. *et al.* Oligonucleotide-directed single-base DNA alterations in mouse embryonic stem cells. *Gene Ther.* **10**, 24–33 (2003).
- Podevin, N., Davies, H. V., Hartung, F., Nogu , F. & Casacuberta, J. M. Site-directed nucleases: a paradigm shift in predictable, knowledge-based plant breeding. *Trends Biotechnol.* **31**, 375–83 (2013).
- Podevin, N., Devos, Y., Davies, H. V. & Nielsen, K. M. Transgenic or not? No simple answer! New biotechnology-based plant breeding techniques and the regulatory landscape. *EMBO Rep.* **13**, 1057–61 (2012).
- Porteus, M. H. & Baltimore, D. Chimeric nucleases stimulate gene targeting in human cells. *Science* **300**, 763 (2003).
- Pruett-Miller, S. M., Reading, D. W., Porter, S. N. & Porteus, M. H. Attenuation of zinc finger nuclease toxicity by small-molecule regulation of protein levels. *PLoS Genet.* **5**, e1000376 (2009).
- Ramirez, C. L. *et al.* Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects. *Nucleic Acids Res.* **40**, 5560–8 (2012).
- Rosen, L. E. *et al.* Homing endonuclease I-CreI derivatives with novel DNA target specificities. *Nucleic Acids Res.* **34**, 4791-4800 (2006).
- Sander, J. D. *et al.* In silico abstraction of zinc finger nuclease cleavage profiles reveals an expanded landscape of off-target sites. *Nucleic Acids Res.* **41**, e181 (2013).
- Sargent, R. G., Kim, S. & Gruenert, D. C. Oligo/polynucleotide-based gene modification: strategies and therapeutic potential. *Oligonucleotides* **21**, 55–75 (2011).
- SBSTTA. (2014a, April) Draft Report of the Subsidiary Body on Scientific, Technical and Technological Advice under the Convention on Biological Diversity: New and emerging issues relating to the conservation and sustainable use of biodiversity - possible gaps and overlaps with the applicable provisions of the Convention, its protocols and other relevant agreements related to components, organisms and products resulting from synthetic biology techniques. <http://www.cbd.int/emerging/>: *United Nations Environment Programme Convention for Biodiversity*, 1-60.
- SBSTTA. (2014b, April) Draft Report of the Subsidiary Body on Scientific, Technical and Technological Advice under the Convention on Biological Diversity: New and emerging issues relating to the conservation

- and sustainable use of biodiversity - potential positive and negative impacts of components, organisms and products resulting from synthetic biology techniques on the conservation and sustainable use of biodiversity. <http://www.cbd.int/emerging/>: *United Nations Environment Programme Convention for Biodiversity*, 1-60.
- Seligman, L. M. *et al.* Mutations altering the cleavage specificity of a homing endonuclease. *Nucleic Acids Res.* **30**, 3870-3879 (2002).
- Shan, Q. *et al.* Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 686–8 (2013).
- Shukla, V. K. *et al.* Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. *Nature* **459**, 437–41 (2009).
- Simon, J.R. & Moore, P.D. Homologous recombination between single-stranded DNA and chromosomal genes in *Saccharomyces cerevisiae*. *Mol Cell Biol* **7**, 2329–2334 (1987).
- Sorek, R., Kunin, V. & Hugenholtz, P. CRISPR - a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat Rev Microbiol.* **6**, 181-6 (2008).
- Stoddard, B. L. Homing endonucleases: from microbial genetic invaders to reagents for targeted DNA modification. *Structure* **19**, 7–15 (2011).
- Stoddard, B. L., Monnat, R. J., Jr., & Scharenberg, A.M. (2007). Advances in engineering homing endonucleases for gene targeting: ten years after structures. In *Progress in Gene Therapy: Autologous and cancer stem cell gene therapy*, R. Bertolotti and K. Ozawa, eds. (Singapore: World Scientific eBooks), pp. 135–167.
- Sugano, S. S. *et al.* CRISPR/Cas9-mediated targeted mutagenesis in the liverwort *Marchantia polymorpha* L. *Plant Cell Physiol.* **55**, 475–481 (2014).
- Sussman, D. *et al.* Isolation and characterization of new homing endonuclease specificities at individual target site positions. *J. Mol. Biol.* **342**, 31-41 (2004).
- Taylor, G.K. *et al.* LAHEDES: The LAGLIDADG homing endonuclease database and engineering server. *Nucleic Acids Res.* **40**, W110–W116 (2012).
- Tesson, L. *et al.* Knockout rats generated by embryo microinjection of TALENs. *Nat. Biotechnol.* **29**, 695–6 (2011).
- Thierry, A. & Dujon, B. Nested chromosomal fragmentation in yeast using the meganuclease I-Sce I: a new method for physical mapping of eukaryotic genomes. *Nucleic Acids Res.*, **20**, 5625-5631 (1992).
- Thompson, A. J., Yuan, X., Kudlicki, W. & Herrin, D. L. Cleavage and recognition pattern of a double-strand-specific endonuclease (I-Crel) encoded by the chloroplast 23S rRNA intron of *Chlamydomonas reinhardtii*. *Gene* **119**, 247-251 (1992).
- Townsend, J. *et al.* High-frequency modification of plant genes using engineered zinc-finger nucleases. *Nature* **459**, 442–5 (2009).

- Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S. & Gregory, P. D. Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* **11**, 636–46 (2010).
- Voytas, D. F. & Gao, C. Precision genome engineering and agriculture: opportunities and regulatory challenges. *PLoS Biol.* **12**, e1001877 (2014).
- Wah, D. A., Bitinaite, J., Schildkraut, I., Aggarwal, A.K. Structure of FokI has implications for DNA cleavage. *Proc Natl Acad Sci USA.* **95**, 10564–9 (1998).
- Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E.S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80–84 (2014).
- Wright, D. A. *et al.* High- frequency homologous recombination in plants mediated by zinc-finger nucleases. *Plant J.* **44**, 693–705 (2005).
- Wyman, C. & Kanaar, R. DNA double-strand break repair: all's well that ends well. *Annu. Rev. Genet.* **40**, 363–83 (2006).
- Xie, K. & Yang, Y. RNA-guided genome editing in plants using a CRISPR-Cas system. *Mol. Plant* (2013). doi: 10.1093/mp/sst119
- Yoon, K. Montagna Symposium on Epidermal Stem Cells Oligonucleotide-Directed Gene Correction in Epidermis. *J Investig Dermatol Symp Proc.* **9**, 276–83 (2004).
- Yoon, K., Cole-strauss, A. & Kmiec, E. B. Targeted gene correction of episomal DNA in mammalian cells mediated by a chimeric RNA.DNA oligonucleotide. *Proc Natl Acad Sci USA* **93**, 2071–2076 (1996).
- Yoshiyama, K. O., Sakaguchi, K. & Kimura, S. DNA damage response in plants: conserved and variable response compared to animals. *Biology (Basel)* **2**, 1338–56 (2013).
- Zhang, H. *et al.* The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. *Plant Biotechnol J.* **12**, 797–807 (2014).