# Chapter 17

# DNA Break Repair in Plants and Its Application for Genome Engineering

Carla Schmidt, Michael Pacher, and Holger Puchta

## Abstract

Genome engineering is a biotechnological approach to precisely modify the genetic code of a given organism in order to change the context of an existing sequence or to create new genetic resources, e.g., for obtaining improved traits or performance. Efficient targeted genome alterations are mainly based on the induction of DNA double-strand breaks (DSBs) or adjacent single-strand breaks (SSBs). Naturally, all organisms continuously have to deal with DNA-damaging factors challenging the genetic integrity, and therefore a wide range of DNA repair mechanisms have evolved. A profound understanding of the different repair pathways is a prerequisite to control and enhance targeted gene modifications. DSB repair can take place by nonhomologous end joining (NHEJ) or homology-dependent repair (HDR). As the main outcome of NHEJ-mediated repair is accompanied by small insertions and deletions, it is applicable to specifically knock out genes or to rearrange linkage groups or whole chromosomes. The basic requirement for HDR is the presence of a homologous template; thus this process can be exploited for targeted integration of ectopic sequences into the plant genome. The development of different types of artificial site-specific nucleases allows for targeted DSB induction in the plant genome. Such synthetic nucleases have been used for both qualitatively studying DSB repair in vivo with respect to mechanistic differences and quantitatively in order to determine the role of key factors for NHEJ and HR, respectively. The conclusions drawn from these studies allow for a better understanding of genome evolution and help identifying synergistic or antagonistic genetic interactions while supporting biotechnological applications for transiently modifying the plant DNA repair machinery in favor of targeted genome engineering.

**Key words** DSB repair, HDR, NHEJ, Programmable nucleases, Gene targeting, Genome engineering

## 1 Randomly Induced DNA Damage

Different endogenous and exogenous factors continuously damage genetic information, causing various kinds of DNA lesions (Fig. 1). As photoautotrophic organisms, plants use light energy for photosynthesis and thus are also exposed to ultraviolet radiation (UV). This radiation ranges from 100 to 400 nm in wavelength. Three types of UV radiation are described, namely UV-A, UV-B, and UV-C, with UV-C being directly absorbed by the DNA and thus

Sandeep Kumar et al. (eds.), *Transgenic Plants: Methods and Protocols*, Methods in Molecular Biology, vol. 1864,
https://doi.org/10.1007/978-1-4939-8778-8_17, 

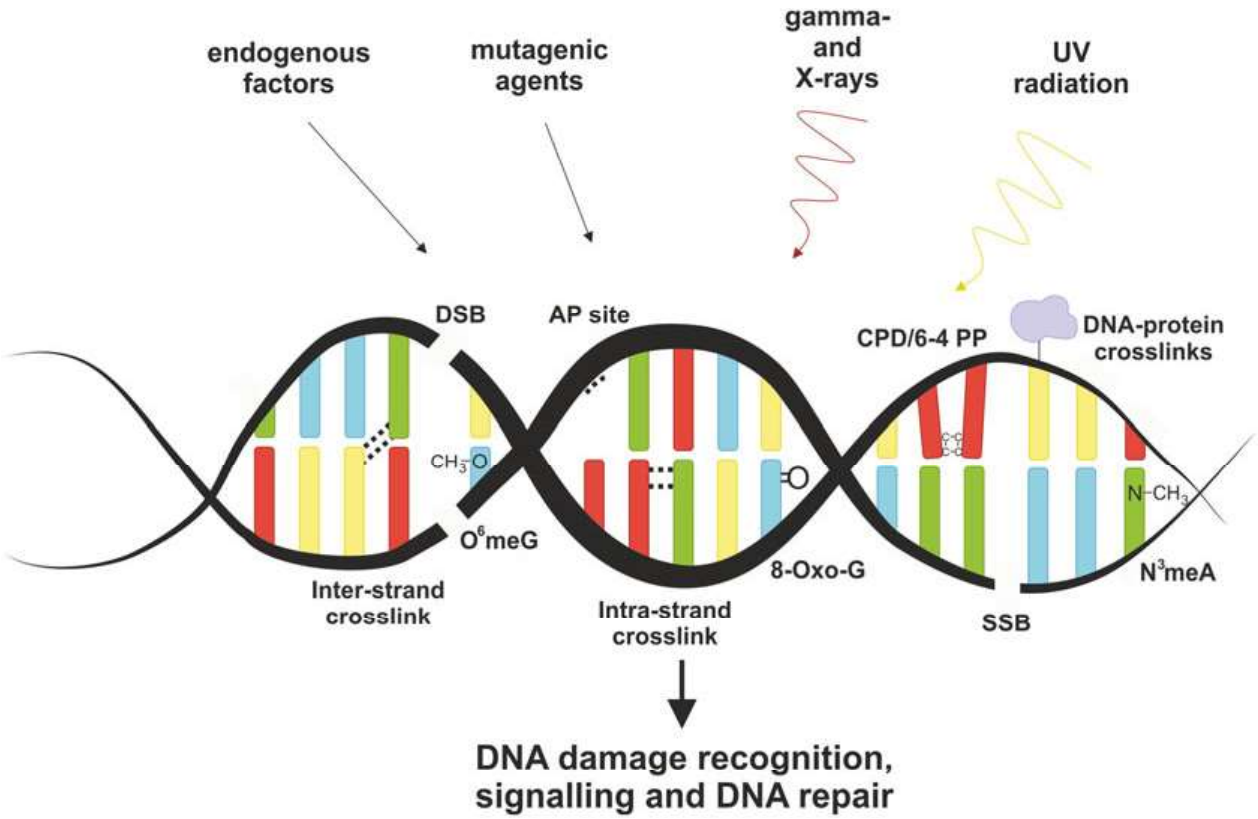


**Fig. 1** Possible DNA lesions induced by various factors. The DNA double helix is exposed to different endogenous and exogenous factors, resulting in different DNA damages, followed by damage recognition and DNA repair. CPDs and 6–4 PPs are mainly produced via UV radiation, while gamma- and X-rays directly ionize the DNA, resulting in DSBs and SSBs. Both UV and ionizing radiations are also involved in the formation of ROS. This highly reactive species can cause several kinds of base damages like 8-oxo-guanine, $O^6$-meG, $N^3$meA, and AP sites. ROS are formed endogenously due to the presence of free radicals. Mutagenic agents can cause cross-links between adjacent and paired bases, as well as DNA-protein cross-links

being most harmful for genetic material. UV radiation mainly causes the formation of cyclobutane pyrimidine dimers (CPDs) and 6–4 photoproducts (6–4 PP) by covalent binding of $C_5$ and $C_6$ carbon atoms of the deoxyribose backbone. Additionally, oxidative DNA damage can be caused by UV irradiation, as the high-energy rays produce reactive oxygen species (ROS) by interactions with molecules within a living cell [1]. Free radicals can also appear in plant cells resulting in ROS formation, due to an inevitable leakage of electrons onto $O_2$ molecules during electron transport in chloroplasts, mitochondria, and plasma membranes, or as a by-product of different metabolic pathways [2]. ROS can cause alkylation or oxidation of bases (e.g., 8-oxo-guanine), as well as hydrolysis of the N-glycosidic bond of bases, both resulting in an apurinic/apyrimidinic site (AP site). Besides these endogenous factors, agents like methyl-methanesulfonate (MMS) and ethyl-methanesulfonate (EMS) can alkylate bases at their O/N-positions, resulting, for example, in $O^6$-methylguanine ($O^6$-meG) or $N^7$-methylguanine/$N^3$-methyladenine ($N^7$meG/$N^3$meA). During

replication, the $O^6$-meG mispairs with thymine instead of cytosine, leading to a point mutation. The mutagenic effect of EMS is widely used for random mutagenesis to perform crop improvements and execute forward genetic studies [3]. The most noted procedure is called targeting-induced local lesions in genomes (TILLING), in which traditional chemical mutagenesis is combined with a high-throughput genome-wide screening for point mutations [4, 5]. Aside from methylation, DNA bases can covalently bind to an adjacent base on the same strand, leading to intra-strand cross-links, or mispairing with bases of the complementary strand can result in inter-strand cross-links. Agents like mitomycin C and cisplatin, typically used as cytostatic agents, create inter- and intra-strand cross-links and are also applied for mutagenic treatments of plants. Cross-links can occur not only between two DNA strands, but also between the DNA backbone and proteins. Protein-DNA interactions are required for nearly all essential DNA transactions and need to be highly dynamic; however, such transient interactions can also result in covalent binding of proteins to the DNA. Such DNA-protein cross-links are highly toxic, as they interfere with transcription and replication processes [6]. These DNA damages can occur naturally or can be induced by camptothecin, leading to covalent binding of topoisomerase I to DNA. Besides base damages and cross-linking, the induction of DSBs into the DNA backbone is highly mutagenic due to the absence of a complementary strand, usually serving as repair template. DSBs can arise endogenously during processes involved in transmission, expression, and maintenance of the genetic information, as well as by exogenous factors like ionizing radiation or radiomimetic agents, e.g., bleomycin. Ionizing radiation like X- and gamma rays was discovered in the late nineteenth century. Its highly mutagenic potential is caused by its ability to directly ionize DNA molecules or to produce highly reactive radicals and free electrons. Ionizing radiation can lead to DNA fragmentation and especially to modifications at the break sites, like the formation of 3′-phosphate, 3′-phosphoglycolate, and 3′-phosphoglycaldehyde ends, which need to be processed before ligation [7]. AP sites, base oxidations, and SSBs can also occur via secondary DNA ionization by highly reactive radicals or free electrons. The first mutagenesis studies applying X-rays in eukaryotes were made in maize, barley, wheat and *Drosophila* [8, 9]. Since then, ionizing radiation has been applied as a useful tool to create mutants and novel traits and to study DNA repair mechanisms. Large-scale mutagenesis of plants using gamma rays was performed on the gamma field of the Japanese Institute of Radiation Breeding [10]. The number of mutations arising in plant cells can differ depending on the exposure time and the distance to the radiation source. Subsequently, laborious screening of plants carrying mutations associated to desired phenotypes or to the production of greater yields is indispensable.

Mutagens like bleomycin are also able to induce DSBs by mimicking the effects of ionizing radiation. Due to its cytostatic effect, this family of glycopeptides are commonly known from cancer treatments. The DSB-inducing property of bleomycin is exploited for studying DNA repair in plants, while it only rarely plays a role in mutagenesis [11, 12].

Plants continuously face the challenge to repair all of these different base lesions, cross-links, SSBs, and DSBs (reviewed in [13]). Therefore, a complex network of various damage surveillance and repair mechanisms is in place, acting through highly specific proteins and signaling cascades. To get a deeper insight into how these mechanisms are regulated and which protein functions are required for adequate repair, plant mutant lines lacking certain protein functions are required. For the model plant *Arabidopsis thaliana*, many mutant lines are available for such studies that have been produced via Agrobacterium-mediated T-DNA transformation [14]. Based on the elucidation of repair mechanisms in plants, different DNA repair models have been developed, which now provide the basis for controlled sequence modifications, mediated by targeted DNA damage. The discovery of synthetic nucleases with specific or adaptable target sequences led to a revolution of plant genome engineering approaches [15].

## 2 Site-Specific Endonucleases for Targeted Genome Manipulations

The first experiments in plants to induce site-specific breaks in the genome were performed with the homing endonuclease I-*Sce*I, which was discovered in *Saccharomyces cerevisiae* [16, 17]. Due to its non-palindromic 18mer recognition sequence [18, 19], I-*Sce*I is also referred to as megaendonuclease. The probability of the I-*Sce*I target site occurring randomly in the genome of a plant species is very low and correlated to the respective genome size. Obviously, this discovery and the characterization of this unique enzyme provided an unprecedented tool for targeted DSB induction in vivo, by I-*Sce*I expression in plants harboring transgenes with the 18 nt recognition site. It was demonstrated by Puchta et al. that I-*Sce*I can induce DSBs in plant cells and can therefore be used for precise genome alterations and to study basic mechanisms of DSB repair [20]. A disadvantage of I-*Sce*I, however, is its static intrinsic 18 nt recognition motif, as it is desirable to target any sequence in the genome. For that reason, different types of artificial nucleases with adaptive recognition sites have been investigated. As the characteristics of these nucleases and their applications are discussed in detail in the following chapters, we will provide only a short overview here.

First attempts to create site-specific nucleases were based on the combination of a DNA-binding domain with a nuclease domain.

This was initially realized by a combination of a highly variable class of zinc-finger transcription factors and a cleavage domain of the restriction enzyme *Fok*I [21]. Different combinations of these zinc fingers, with one zinc finger binding three adjacent bases, implemented sequence-specific recognition of the target sequence. However, sometimes interactions between zinc fingers and proximal nucleotides led to unpredictable binding; thus each nuclease required thorough testing. As cloning and testing of these zinc-finger nucleases (ZFNs) are quite time consuming, a similar approach was investigated using transcription activator-like effectors (TALEs), which were discovered in the plant pathogen species *Xanthomonas*. The DNA-binding domain of these TALEs consists of tandem repeats, varying only in amino acids 12 and 13, referred to as repeat-variable diresidue. The variable region of the tandem repeat was found to provide the specificity for all four DNA bases and thus a combination of the TALEs with a *Fok*I nuclease domain, called a TALE nuclease (TALEN), can be applied for targeting a DSB. Both ZFNs and TALENs were successfully used in different plant species for genome engineering approaches (reviewed in [22]).

The programmable nuclease discovered most recently was originally described as an adaptive immune system in bacteria and archaea, consisting of a clustered regularly interspaced short palindromic repeat (CRISPR) locus and Cas proteins [23]. This CRISPR/Cas system can be used for site-specific break induction by combining the Cas9 nuclease with a sgRNA complementary to the target sequence [24]. Besides the 20 nt spanning target sequence, a protospacer adjacent motif (PAM) is required at the target locus for break induction via Cas9. The adaption of the sgRNA to the target site and cloning is simple and can be conducted very quickly in any molecular biology laboratory. This technique is currently used in many eukaryotic systems and has proved successful as a tool for genome engineering in plants (reviewed in [25–27]).

The availability of these programmable site-specific nucleases provides the basis of targeted genome modifications and tremendously broadens the possibilities of genome engineering. However, these new biotechnological opportunities are limited by the properties of the intrinsic DNA repair machinery. Thus, a thorough understanding of the plant DSB repair mechanisms and its regulation is a mandatory prerequisite for knowledgeable optimization of genome editing approaches.

## 3 DNA Repair Mechanisms in Plants and Their Application for Genome Engineering

Plants have to protect their genome against several exogenous and endogenous factors. Mostly, only one DNA strand is affected, but

the more severe damages are DSBs, due to the absence of a complementary strand, normally serving as a repair template. Therefore, eukaryotic organisms evolved different pathways to repair these lesions and to preserve genomic integrity. Some of these pathways are error prone and thus cause mutations; however, there are also repair pathways resulting in an accurate restoration of the genome. Most applications of genome engineering in living organisms are based on the exploitation of repair pathways, e.g., mutagenesis, induced chromosomal rearrangements, and gene targeting (GT). In particular, NHEJ as well as HDR are the bases for most applications of genome engineering [28]. While base editing and introduction of epigenetic modifications are indirect ways of either exchanging single nucleotides or modulating gene expression activity, any targeted genetic rearrangement requires the transient opening of the genetic code itself with a site-specific cut, followed by changes therein catalyzed by the intrinsic DNA repair machinery. Thus, numerous studies have been conducted over the last 25 years, in order to understand how DNA double-strand breaks (DSBs) are repaired in plants.

### 3.1 HDR

In somatic plant cells, HDR is a minor pathway occurring mainly during S and $G_2$ phases of the cell cycle [29]. In somatic cells, two different HDR mechanisms have been described, the single-strand annealing (SSA) and the synthesis-dependent strand annealing (SDSA), which are the mechanisms employed for the majority of HDR repair outcomes, while the double-strand break repair (DSBR) model [30] and the dissolution pathway play only a minor role (for overview, *see* Fig. 2). DSBR and dissolution occur most frequently during meiosis, resulting in crossover (CO) and non-crossover (NCO) products, and thus DSBR is the most prominent mechanism for explaining the observed meiotic recombination events [31].

All HR pathways have one common step, mediated by the MRN (Mre11, Rad50, Nbs1) complex, leading to the formation of 3′ single-stranded overhangs. To protect the produced DNA single strands from exonuclease-mediated degradation, a heterodimeric protein complex is formed by the three RPA subunits RPA1, RPA2, and RPA3 in plants [32].

In the presence of homologies on both ends of the break site, the nonconservative SSA pathway can be utilized for repair (Fig. 2a). Following end resection by the MRN complex, the protein RAD52 is recruited to the RPA-bound single-stranded DNA nucleoprotein complex in yeast [33]. In *Arabidopsis*, homologs of RAD52 have been identified; however, they display only a moderate effect on homologous recombination [34]. Instead, it was demonstrated that the *Arabidopsis* RAD51 paralogues XRCC2, RAD51B, and RAD51D are involved in SSA [35]. In contrast to the MMEJ pathway, it was shown in yeast that the homology length

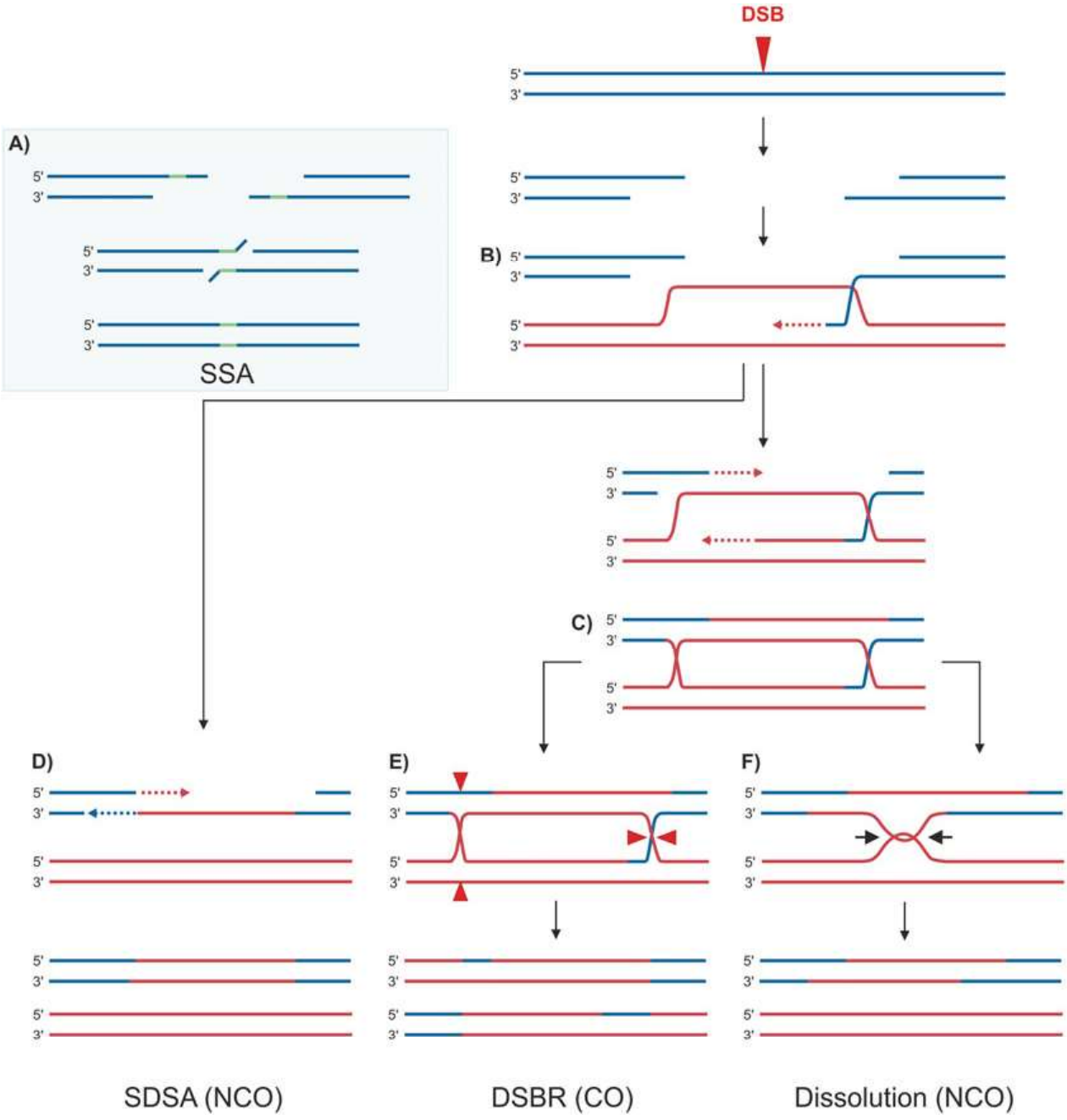


**Fig. 2** Overview of HR-related DSB repair mechanisms. After DSB induction, single-stranded 3′ overhangs are produced. (**a**) In the case of homologies (green) flanking the break site, the repair reaction occurs via SSA. Subsequent to annealing of homologous sequences, the 3′ overhanging flaps are removed by an endonuclease and the gaps are filled by a polymerase. (**b**) When a homologous template is accessible, the 3′ end is able to invade into the donor molecule (red), forming a D-loop. Subsequently, elongation of the invaded single strand can take place, using the complementary donor strand as a template. If the displaced strand realigns with the other end of the break and DNA synthesis occurs, a double Holliday junction (dHJ) is formed (**c**). In the case of SDSA, the invading strand is released after elongation and realigns with the original break site, resulting in a NCO event, while the information is only copied to one strand (**d**). The dHJ can be resolved via HJ-processing endonucleases, inducing orthogonal orientated cuts (red triangles), resulting in a CO event (**e**). Alternatively, the dHJ can be dissolved by the formation of a hemicatenane involving a helicase followed by strand separation by a topoisomerase, resulting in a NCO event (**f**)

should add up to a minimum size of between 63 and 89 bp, but in the absence of competition SSA can utilize repeats as small as 29 bp [36]. During subsequent strand annealing, overhanging flaps are removed by an endonuclease, resulting in a deletion between the homologies. In plants, the RAD1/RAD10 heterodimer is involved in trimming these flaps before ligation [37]. It is obvious that the homologous template for this kind of repair reaction has to be located in close proximity to the break site and that DSBs in genomic regions containing such homologies are preferentially repaired via SSA. Besides the factors discussed so far, there are indications hinting to a minor involvement of the helicase FANCM and the nuclease MUS81 during SSA [38]. Given the current limited knowledge regarding the factors involved in SSA, it is likely that additional helicases and nucleases can substitute each other, hampering the detection of SSA key factors.

In the case of a DSB occurring in a genomic region lacking repetitive homologies, a homologous double-stranded template elsewhere can serve as the repair template. Following the initial MRN-mediated formation of 3′ single-stranded ends, these overhangs are able to invade into the donor molecule, forming a structure called a D-loop (Fig. 2b). The formation of such a D-loop can lead to either repair via SDSA or processing of the double Holliday junction (dHJ) as an intermediate, resulting in repair via DSBR or dissolution. A dHJ is formed when the displaced strand aligns with the original break site, so that DNA synthesis can also take place at the other broken end (Fig. 2c). This process results in a structure in which information is copied from both strands of the matrix and separation of the homologous molecules can only be achieved by resolution or dissolution of the dHJ. For resolution, specific HJ-processing endonucleases are essential which can bind and introduce symmetric cuts with opposite orientation at the junction (Fig. 2e). Resolution is mainly observed during meiosis and results in CO events leading to an exchange of chromosome arms and is required for mixing of parental genomes. Dissolution occurs via the RTR complex, composed of a RECQ helicase, type 1 topoisomerase, and a structural protein. The helicase forms the hemicatenane in which the topoisomerase cleaves one strand, enabling strand separation (Fig. 2f).

During SDSA, sequence information of the donor strand is replicated to close the gap, with the invading strand being released to realign with the original break site. Thus, the repair reaction always results in a NCO, while the information of the matrix is only copied to one strand (Fig. 2d). Strand invasion of the 3′ ends is mediated by RAD51 and its paralogs [39, 40]. Subsequently, the repair reaction is performed utilizing the complementary donor strand as a template. The binding of the RAD51-ssDNA complex to the homologous strand is stabilized by the SWI2/SNF2 class chromatin remodeler RAD54, which harbors a DNA-binding

domain, mediating homology search [41, 42]. It was shown before that RAD54, as well as RAD51 and its paralogs RAD51C and XRCC3, is essential for SDSA, but not for SSA, while the nuclease MUS81 is involved in both pathways [38, 42].

The mechanistic HR models discussed above have been sustained qualitatively and quantitatively by analyzing the repair of DSBs induced in the genome of model plants. Early studies in tobacco showed that in the absence of DSBs, the occurrence of HR in somatic tissue is very low [43, 44]. First experiments to study DNA repair *in planta*, using transposons and their intrinsic property of cut-and-paste transposition, showed that these DSBs are mainly repaired via NHEJ [45]. An enhancement of HR-mediated repair was observed, however, when homologous sequences were localized close to the break site [46]. Using I-*Sce*I in tobacco protoplasts allowed Puchta et al. to conclude that homologous recombination can be significantly enhanced by targeted DSB induction in vivo [20]. Since then, recombination traps harboring one or two I-*Sce*I recognition sites have been used to study DSB repair mechanisms *in planta*.

When homologies are in close vicinity to the induced DSB, the SSA-mediated repair is dependent on the flanking homologies, resulting in the loss of the sequence between the homologous repeats. Quantification of SSA frequency is possible using a stably integrated *ß-glucuronidase* (*gus*) reporter, in which the *gus* gene harbors a segmentally duplicated region with an I-*Sce*I recognition site and the functional *gus* gene can be restored via SSA repair. This way, it could be demonstrated that one out of three DSBs is processed via SSA in regions with sequence homologies [47]. In a similar approach, Orel et al. established that SSA-mediated DSBR is approximately five times more efficient than interchromosomal SDSA in the presence of flanking homologies [48]. Very recently, Vu et al. analyzed the mutation spectra at three different endogenous loci in *Arabidopsis* by deep sequencing whereby it was found that the distance between the repeats has an influence on the absolute frequency of SSA, with the efficiency of SSA being inversely correlated to the distance between homologies [49]. Thus, SSA is a very efficient way of repairing DSBs in genomic regions with tandem repeats, depending on their distance. Naturally, tandem repeats and sequence duplications are not uniformly distributed over the genome, implying that repair mechanisms can differ due to the genomic context. Moreover, HDR is dependent on the accessibility of homologous sequences serving as a template. While dissolution and DSBR occur mainly during meiosis, with the dHJ and the synaptonemal complex formation being prerequisites for correct pairing of homologous chromosomes, it was shown by Puchta that the SDSA model is best suitable to describe DSBR in somatic plant cells [50]. In contrast to SSA, the frequency of SDSA-dependent repair is difficult to measure as several types of

homologous donor sequences can be a substrate for SDSA. During the $G_2$ and S phases of the cell cycle, the homology of the sister chromatid can be utilized for repair via SDSA. The interaction of sister chromatids is mediated by the "structural maintenance of chromosome" genes SMC6A and SMC6B, factors that are required for DNA repair and HDR in *Arabidopsis* [42, 51]. Unfortunately, the frequency of such an occurrence cannot be quantified, as the accurately restored sequence is indistinguishable from the original sequence. Apart from the sister chromatid, there are different genomic sites, which can possibly be used as a repair template during HDR (Fig. 3). Allelic sequences on the homologous chromosome can act as a template as well as homologous sequences

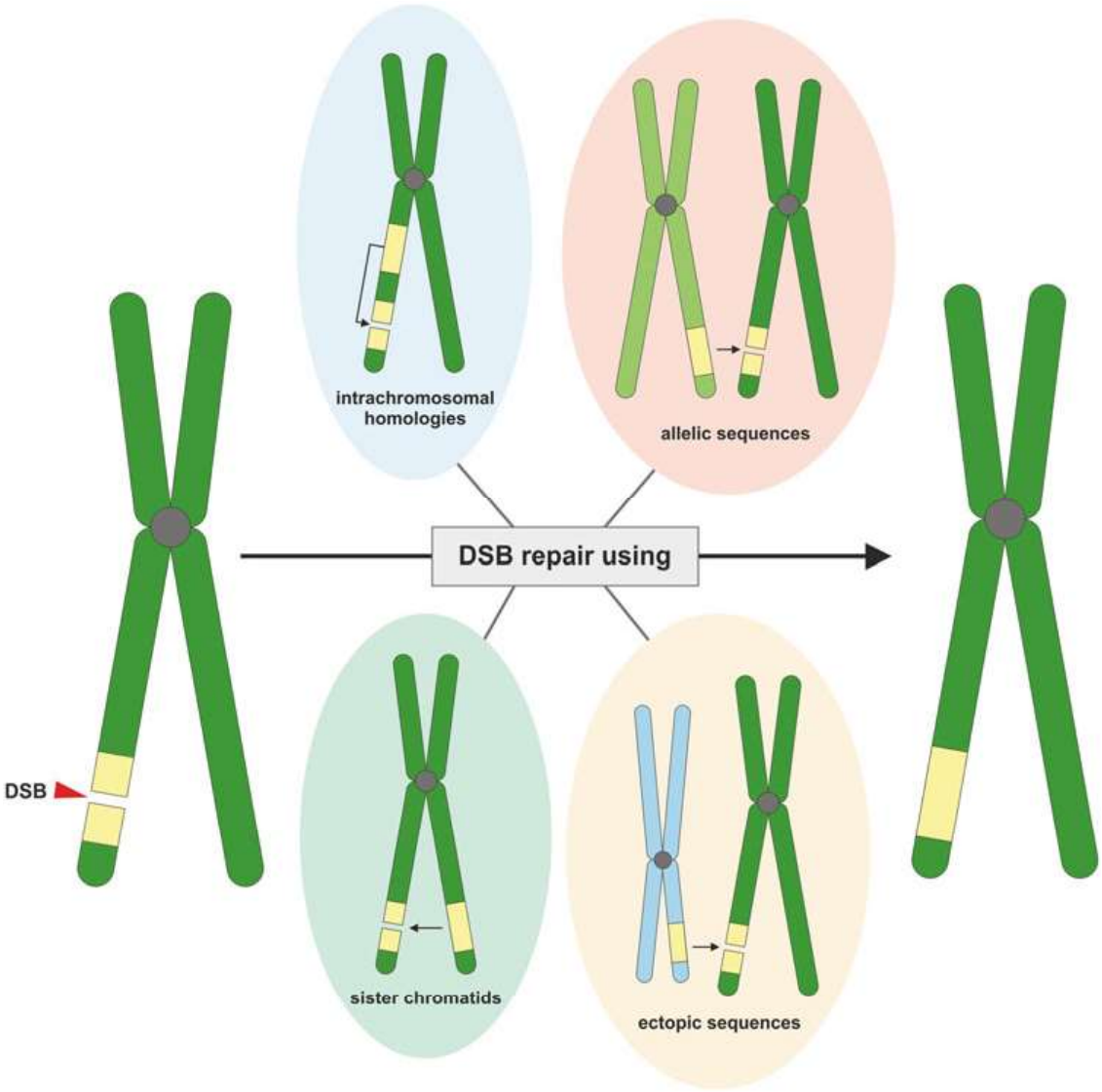


**Fig. 3** Possible donor sequences for HR. Induced DSBs can be repaired using several templates for correct repair via SDSA. During G2 and S phases of the cell cycle, the sister chromatid can be used, leading to an accurate restoration of the interrupted sequence. Furthermore, intrachromosomal homologies located on the same chromosome can serve as a template for HDR, as well as allelic sequences from homologous chromosomes, primarily occurring during meiosis. Ectopic sequences located on other chromosomes or incoming ectopic DNA are also possible templates for HDR

located on the same chromosome, e.g., in the case of gene duplications or transposable elements. Additionally, ectopic sequences on different chromosomes or from extrinsic DNA are also possible templates for SDSA (for review, *see* [29]).

As discussed before, the discrimination of HDR involving the sister chromatid is not possible per se and thus relatively difficult to measure. To get an insight into how frequently the sister chromatid is used as a template during $G_2$ and S phases, Vu et al. displayed the sister chromatid exchange via ethynyldeoxyuridine-based staining [52]. Interestingly, it could be shown this way that most of the I-*Sce*I-induced DSBs were repaired using the sister chromatid as a template, leading to the conclusion that the frequency of this kind of repair is underestimated. Although these findings are highly interesting for fundamental research regarding DNA repair, the applications for genome engineering are negligible due to the fact that it is impossible to induce sequence changes via sister chromatid exchange.

The second most common template for DSB repair via SDSA is intrachromosomal homology. Frequencies can be measured using a nonfunctional GUS reporter with homologies on the same chromosome within close proximity, serving as a template for the restoration of the GUS marker gene. Applying this kind of reporter system, it was shown that I-*Sce*I-induced breaks can be repaired via SDSA using an intrachromosomal homology and about 1 out of 15 DSBs is repaired via this kind of HDR [48]. Besides this analysis, many studies have been published, indicating that homologous sequences on the same chromosome are an appropriate matrix for repair in plants [42, 47, 53].

During meiosis, homologies in allelic positions serve as a template for HR and it was shown that such a template can also be exploited in somatic plant cells [54]. In this study, two different transgenic lines were combined, one with an intact I-*Sce*I recognition site and the other with a disrupted I-*Sce*I site. After crossing these two lines, induction of a DSB via I-*Sce*I led to recombination events involving the allelic sequence. Besides loss-of-function mutations due to NHEJ-related repair representing the majority of events, Gisler et al. detected allelic HR frequencies of $10^{-4}$ [54]. Unfortunately, the transgenic lines were lacking homologies directly at the I-*Sce*I break site in this setup. Taken together, this led to the conclusion that intrachromosomal somatic HR is 100–1000-fold more efficient than gene conversion with donor sequences in ectopic or allelic positions. Very recently, a study in tomato revealed that Cas9-induced targeted DSBs are more often repaired via somatic HR, with the homologous chromosome acting as a template. In this study, an allele-specific assay was set up in *S. lycopersicum* with DSB induction occurring in a region that allows discrimination between repair footprints, originating from the homologous allele or from nonhomologous repair. In this

experimental setup, 14% of all detectable DSB repair events were allele dependent and three cases could be identified in which HR-dependent repair events were transmitted through the germline [55]. In contrast to the findings of Gisler et al., the homologies were directly present at the break site, with the marker for allele-specific recombination being only a single-nucleotide polymorphism [54]. These results are promising for genome engineering approaches relying on HDR, such as the integration of an extrachromosomal homologous template into a break site, using homologous chromosome arms.

The integration of an ectopic sequence into a break site is another possibility of repair by the use of homologous genomic sequences, especially relevant for the large plant genome with its many repetitive and widely spread genetic elements. Efficiencies for ectopic DSB-induced HR were tested in tobacco using an acceptor and a donor locus on different chromosomes. The endonuclease I-*Sce*I was transiently expressed via *Agrobacterium* in those plants, inducing a DSB in the acceptor locus. Following successful recombination involving the donor locus on a different chromosome, a kanamycin resistance cassette could be restored. Puchta established that ectopic sequences are used extremely rarely as a HR template, with frequencies in the range of $10^{-5}$ [56]. In a similar approach by Shalev and Levy, frequencies of $10^{-3}$–$10^{-5}$ were reported using transposons for DSB induction in *Arabidopsis* [57]. The low occurrence of such events is thought to be a prerequisite for stabilizing genetic information in somatic tissue [56] and might be the major challenge for the integration of DNA during GT. Taken together, the chromosomal site of the HR template and the occurrence of homologies close to the break site strongly influence the repair pathway choice. Furthermore, recent studies found that the genome size of barley and *Arabidopsis* correlates with the abundance of different repair outcomes, leading to the conclusion that frequencies of the entirety of repair pathways differ between plant species [52, 58]. These species-specific differences were described earlier for *Arabidopsis* and tobacco, demonstrating that there is a positive correlation between genome size and DSB processing [59, 60]. It is evident from the studies discussed here that DSB repair pathway choice and the natural prevalence of NHEJ-mediated repair in somatic plant cells pose bottlenecks for biotechnological approaches, focused on targeted genome modifications.

### 3.2 Gene Targeting in Plants

The GT technique is based on HDR and can be applied to integrate new genes at a specific locus or to modify the sequence of a respective gene (reviewed in [61]). GT was initially tested in tobacco protoplasts without the induction of a site-specific DSB. Targeted events could be selected by GT-mediated restoration of a nonfunctional kanamycin resistance gene with plasmids [62], as well as with incoming T-DNA or chimeric RNA/DNA

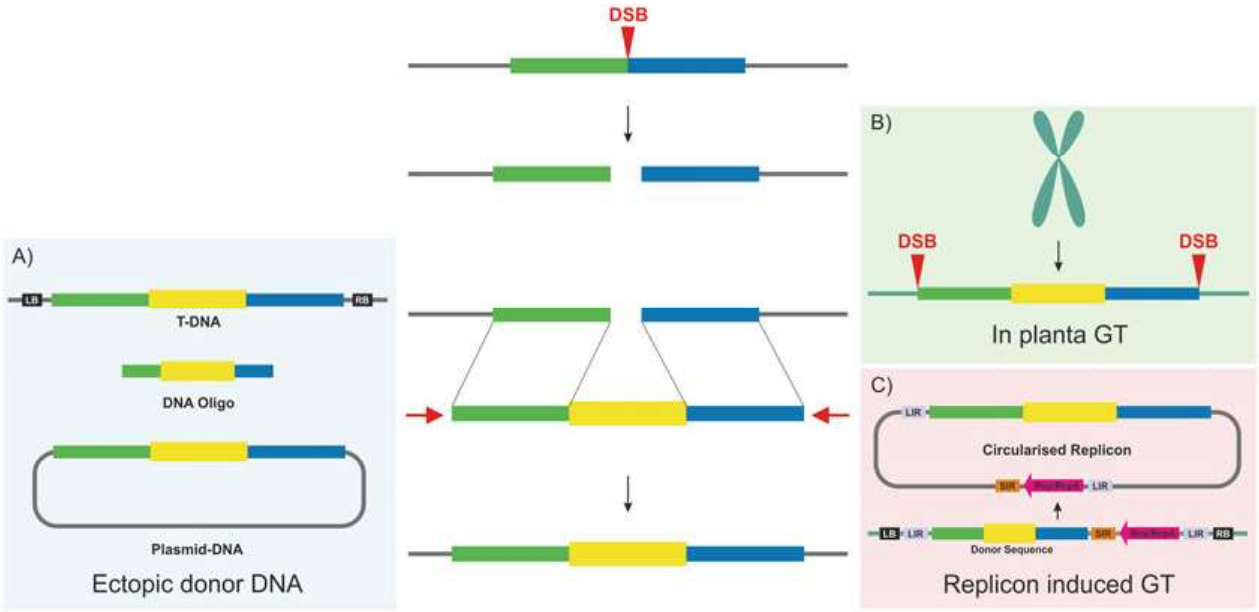


**Fig. 4** Possible donor molecules for GT in plants. Following DSB induction, different donor sequences can be used for HDR. Homologies (green/blue) to the area flanking the break site need to be present for correct incorporation of the donor sequence (yellow). (**a**) Possible ectopic templates for GT are T-DNAs transferred via Agrobacterium, as well as plasmid DNA and short DNA oligos. (**b**) If the donor sequence is stably integrated into the genome, three simultaneous breaks are required to release the target sequence and simultaneously activate the donor sequence. (**c**) Enhancement of GT frequencies can be achieved by the use of geminivirus-based replicons: the transferred T-DNA consists of the donor sequence combined with a site-specific nuclease, a donor, and minimal parts necessary for geminivirus rolling-circle replication, leading to the formation of circularized replicons

oligonucleotides [63, 64] (Fig. 4a). I-*Sce*I was the first nuclease to enhance GT frequencies in tobacco by targeted DSB induction in an artificial target site integrated within the genome. The transiently expressed I-*Sce*I cassette led to the induction of a DSB and the break being repaired using a T-DNA with flanking homology stretches, leading to an increase of GT frequency by up to two orders of magnitude [65].

Besides I-*Sce*I, different synthetic nucleases have been applied to increase HR frequencies in plants, thus becoming independent of one particular recognition sequence. ZFNs were used to restore a disrupted marker gene via GT in tobacco protoplasts, reaching GT frequencies up to 10% in selected lines [66], with applications in maize, *Arabidopsis*, and tobacco cells in the following years [67–69]. Subsequently, GT experiments were performed using TALENs in tobacco protoplasts, altering the ALS gene with frequencies up to 4% [70]. A similar gene alteration frequency was successfully obtained using the Cas9-sgRNA system, combined with different donor DNAs (vector DNA, single-stranded oligos), targeting the ALS2 gene in maize protoplasts [71]. Moreover, Cas9-mediated GT approaches were also successfully applied in *Linum usitatissimum* and soybean [72, 73].

Another GT approach, referred to as "*in planta* GT," is independent of transfected donor DNA, as the donor sequence is stably integrated into the genome. For this purpose, three simultaneous cuts are required, one within the target site and two cuts to release the linear donor sequence from the genome (Fig. 4b). Applying such an *in planta* GT technique based on I-*Sce*I-mediated DSBs in *Arabidopsis*, Fauser et al. were able to reach GT efficiencies of up to 1% [74]. More recently, *in planta* GT using the Cas9-sgRNA system with an adaptable recognition site allowed targeting of an endogenous locus, without the requirement of integrating a transgene upfront [75]. An advantage of this *in planta* GT technique is a minimization of unwanted random integrations of the donor, as with the case of a hemizygous line only one donor can be excised and integrated per genome. Furthermore, *in planta* GT is independent of transformation efficiencies, with this making it especially attractive for crop plants that are more difficult to transform.

To increase the number of donor molecules, *in planta* GT can also be combined with geminivirus-based replicons (Fig. 4c). For this application, the transformed T-DNA consists of a site-specific nuclease, donor template, and the minimal parts necessary for geminivirus replication. Recently, the successful application of this technique was demonstrated using ZFNs, TALENs and Cas9, respectively [76, 77].

As HDR is a minor pathway in higher organisms, GT frequencies are relatively low. Therefore, various attempts were tested to enhance HDR by overexpression of HDR-related genes or by repression of NHEJ-related genes. It could be demonstrated in *Arabidopsis* that overexpression, as well as egg cell-specific expression of RAD54 from *S. cerevisiae*, can enhance GT efficiency [78, 79]. Furthermore, it was shown that a knockout of the SMC6B gene leads to impaired sister chromatid exchange and therefore GT efficiencies could be increased in *Arabidopsis* [80]. Besides knockout or overexpression of HDR-related proteins, different NHEJ factors are also able to influence GT. Endo et al. successfully obtained enhanced GT efficiencies by using *lig4*-deficient rice calli cells, lacking an important factor for NHEJ [81]. This indicates that besides HDR pathways, DSB repair via NHEJ plays an important role for genome engineering and deserves our attention.

### 3.3 NHEJ

DSBs in somatic plant cells are most frequently repaired via NHEJ, either by rejoining the broken DNA ends without or with small sequence changes. In contrast to HDR pathways, there is no need for homologous sequences during the repair process. The mechanism ensures that in the absence of homologies, a DSB can be repaired easily without significant sequence loss. The NHEJ pathway described initially is characterized by the KU70/KU80 heterodimer and is commonly referred to as canonical-NHEJ (c-NHEJ).

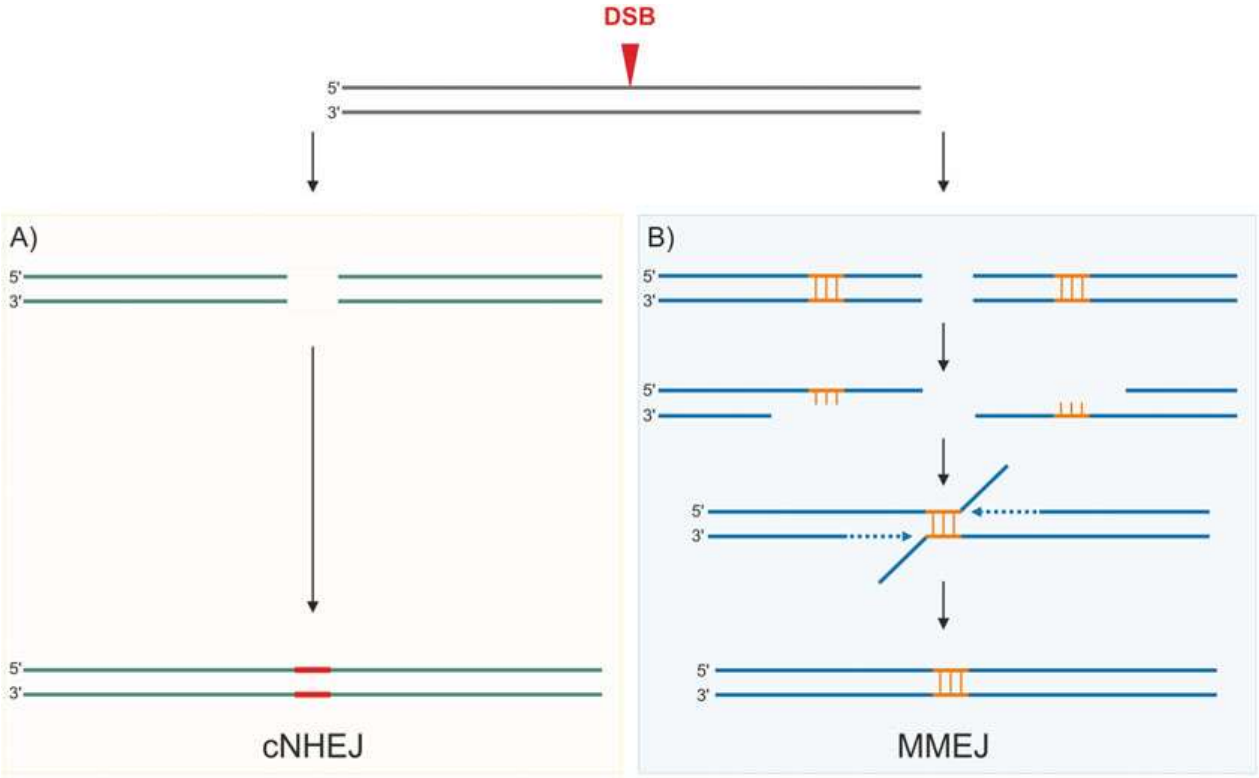


**Fig. 5** DSB repair via NHEJ mechanisms. Repair of a DSB can occur via two major NHEJ pathways. (**a**) If the KU heterodimer is present, the most frequently used pathway is referred to as cNHEJ. During cNHEJ, a direct ligation of the broken ends can be performed, leading to INDELS (red) or accurate repair of the break site. (**b**) In the absence of the KU heterodimer, MMEJ takes over for DSB repair. The MMEJ pathway is dependent on microhomologies (orange) close to the break site. The first step of the MMEJ pathway is end resection, leading to 3′ overhangs. Subsequently, annealing of the microhomology region is performed, leading to 3′ flaps. The removal of heterologous flaps is followed by fill-in synthesis and finally the occurrence of a ligation reaction. MMEJ is highly error-prone as it mainly leads to deletions between microhomologies

cNHEJ results either in small insertions and/or deletions (InDels) or in an error-free ligation of the break site (Fig. 5a). The KU70 and KU80 proteins protect the free DNA ends from a significant loss of nucleotides during end resection [82]. Following binding of KU70/KU80, the cNHEJ pathway in vertebrates involves the recruitment of DNA-PKcs and Artemis, as well as X-family polymerases, followed by XLF, XRCC4, and LIG4 [83]. In plants, homologs for KU70, KU80, XRCC4, LIG4 have been described [84–86], however, no homologs for DNA-PKcs, ARTEMIS, and XLF were found. During cNHEJ, only little or no genetic information is lost; however, if only one or two nucleotides are deleted or inserted, the mutation can result in a frameshift and cause a complete knockout of gene function. The KU70/KU80 heterodimer is the key factor in cNHEJ and is supposed to repress HR in plants [87], which may explain shortening of telomeric sequences in *Arabidopsis* [88] and barley [89].

Loss-of-function mutations in *A. thaliana Ku70/80, LIG4.* and *XRCC4* genes do not cause any growth defects in the

respective plants [90]. Additionally, in the absence of the KU-heterodimer, various alternatively named NHEJ pathways known as backup-NHEJ (b-NHEJ), alternative-NHEJ (a-NHEJ), and microhomology-mediated end joining (MMEJ) can take over. However, it remains unclear as to if all these designations are synonyms for each other, or if there are indeed several alternative NHEJ pathways. There are indications that PARP1/PARP2 and XRCC1 are involved in one of these alternative pathways [91, 92]. Studies in *Arabidopsis ku80* mutants suggest that in the absence of the KU heterodimer, somatic plant cells shift to a more error-prone end-joining mechanism, although no significant difference in repair outcomes was observed when triple mutants of *ku80, parp1*, and *parp2* were tested [93].

All of these backup pathways are characterized by the presence of small microhomologies, with a length of 1–20 bp, close to the break site. The MMEJ pathway consists of five steps: resection of the DSB ends, annealing of a microhomology region, removal of heterologous flaps, fill-in synthesis, and ligation [94] (Fig. 5b). The first step of end resection is driven by the MRN complex, the same complex which initiates end resection during HDR [95]. The exact annealing mechanism for sequences with microhomologies and protein factors involved in this process are not yet fully understood, but it is assumed that in mammals DNA polymerase Θ (POLΘ) plays a major role in this step [96, 97]. Subsequently, the overhanging 3′-tails are removed by a nuclease. In mammals, there are indications that MRE11 and XPF-ERCC1 are candidates for this nuclease activity [98]; however, until now there are no reports regarding the involvement of their homologs RAD1/RAD10 in plants. The following fill-in synthesis is performed by PolΘ, which seems to be the key player of MMEJ. There are several studies in higher eukaryotes supporting the prominent role of POLΘ in MMEJ [99, 100], as well as involvement in the appearance of chromosomal translocations in mice [101]. Additionally, POLΘ exhibits a RAD51-binding motif and this binding capacity may help to direct DSB repair to MMEJ and to repress the HDR pathways [102]. The plant homolog of POLΘ, TEBICHI (TEB), is involved in DNA replication, recombination, and gene expression, as well as in DNA repair [103]. *Arabidopsis teb* mutants are sensitive to MMC and MMS, indicating that TEB is involved in the repair of inter- and intra-strand cross-links [104]. Furthermore, the *Arabidopsis teb* mutant is not accessible for T-DNA integration; thus TEB is a prerequisite for T-DNA integration into the plant genome [105].

It is possible that a knockout of cNHEJ factors, like KU70 or LIG4 in plants, can lead to a shift of the repair pattern to MMEJ and therefore result in more and larger deletions. The repair patterns of different NHEJ repair mutants were analyzed in *Drosophila* and *Arabidopsis*, indicating that knockouts of KU70, LIG4, or POLΘ, for example, can influence the occurrence of deletions and insertions following DSB induction [93, 106].

### 3.4 Formation of Insertions

The entirety of described repair mechanisms discussed so far all result in INDELS of various sizes or an accurate restoration of the damaged DNA region. However, an explanation for larger insertions at the break site is not evident. Gorbunova and Levy obtained relatively large insertions after DSB induction in tobacco cells, with these inserted sequences originating from either internal regions of transfected plasmids or tobacco genomic DNA [107]. In 1998, Salomon and Puchta were able to demonstrate that DNA insertions can occur during DSB repair in higher eukaryotes, while repair was associated with a broad spectrum of unique and repetitive genomic sequences [108].

Based on their experiments, both groups concluded that an annealing and priming reaction seems to be involved in the formation of genomic insertions. Furthermore, it was assumed that the DSB is processed via an illegitimate recombination pathway, based on the SDSA model of recombination. During this SDSA-like mechanism, the 3′ single-stranded overhang is able to invade into a double-stranded donor and alignment takes place with only a short patch of homology. The linkage with the original break site can occur via ligation to a blunt end, with an annealing step mediated by microhomologies (Fig. 6).

NHEJ-mediated gene insertion has been already achieved with Cas9. Geisinger et al. [109] showed in human and mouse cells that precise sequence insertion is possible after DSB induction. Furthermore, Li et al. reported site-specific insertions via NHEJ after DSB induction, using the Cas9-sgRNA system in rice [110]. Targeted gene insertions were obtained at a frequency of 2.2% in this study. This recently developed approach opens up new possibilities that are promising for the field of genome engineering.

Furthermore, in 2010, Yu et al. investigated a model explaining the occurrence of direct or inverted repeats on the basis of the MMEJ pathway [106]. This synthesis-dependent MMEJ (SD-MMEJ) model assumed that POLΘ is involved in the generation of template insertions, when multiple microhomologies flank the DSB. The conclusion was drawn that direct repeats are produced by a mechanism called loop-out SD-MMEJ (Fig. 7a) and that in contrast inverted repeats can be explained by snap-back SD-MMEJ (Fig. 7b). Both processes result in the formation of a repeated motif, arising from flanking sequences on both sites. Substantial numbers of complex insertions, composed of multiple overlapping copies of nearby sequences, were also obtained in this experimental setup, leading to the assumption that multiple rounds of synthesis and dissociation can occur.

Recent findings revealed that besides DSBs, adjacent SSBs can also lead to the formation of tandem sequence duplications in plants. Schiml et al. induced SSBs on the same and opposite DNA strand, in different intervals and in different chromatic regions [111]. The data indicate that opposing SSBs are highly mutagenic

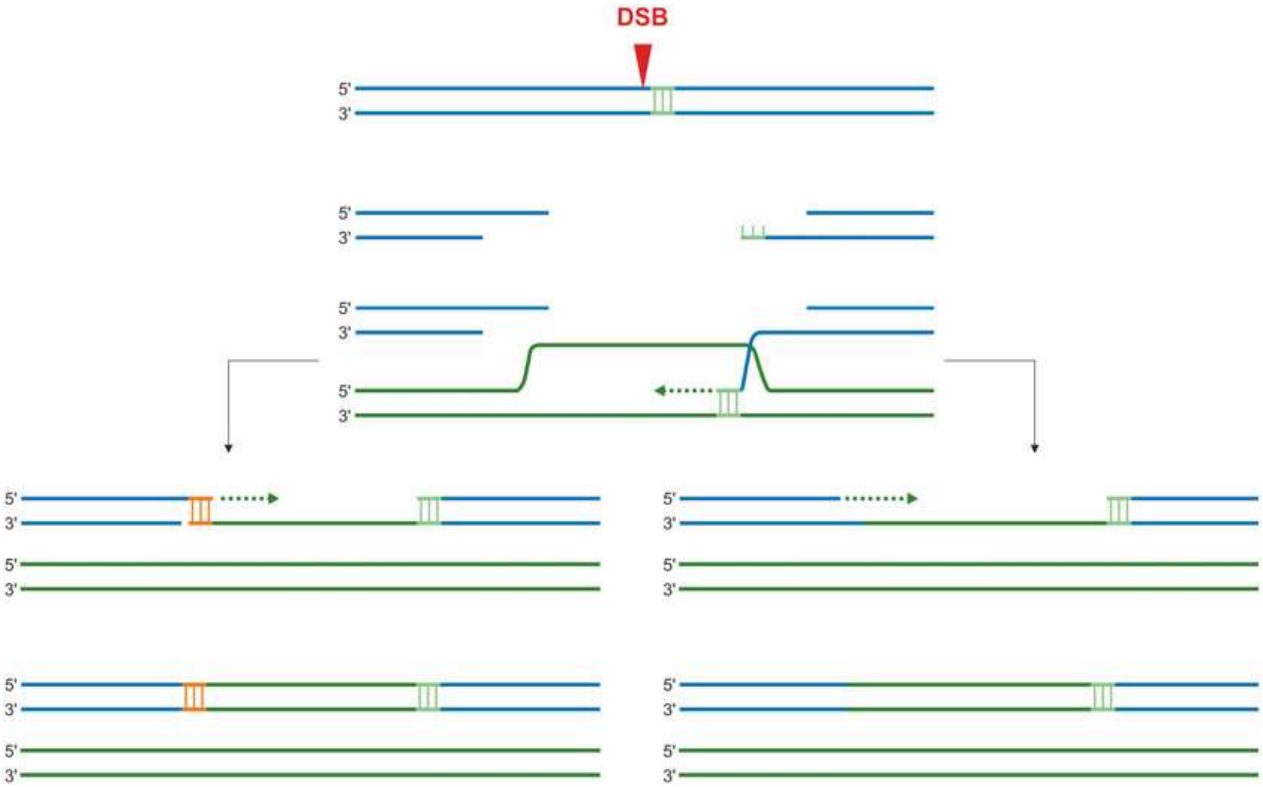


**Fig. 6** SDSA-like mechanism leading to insertions. After end resection at the break site, 3′ overhangs are produced and the single strand is able to invade into the donor molecule (green). Microhomologies (light green) are a prerequisite for alignment with the donor sequence and a D-loop is formed, analogous to the SDSA mechanism. Subsequently, strand displacement can occur via a second microhomology (orange) or via blunt-end ligation. This SDSA-like repair mechanism is leading to insertions at the break site, with any genomic sequence or ectopic DNA being used as a template

and are especially suitable for the production of mutant plants. Furthermore, the tested intervals between the break sites prove that a distance of 50 bps, leading to the formation of 5′ overhangs, is suitable to generate heritable mutations. In this approach, both deletions and insertions were obtained with frequencies up to 77% and it was shown that deletion sizes correlate with SSB distance. Besides repair outcomes related to cNHEJ, deletions with flanking microhomologies were obtained, indicating that these mutations were formed via MMEJ-mediated repair.

The observed insertions originated directly from adjacent sequences or from areas in close proximity, and might be explained by a patch-mediated model for tandem duplications, based on simultaneous base excision repair [112]. During this kind of repair, single-stranded 5′ ends can hybridize via microhomologies and the internal gaps are repaired by fill-in synthesis, while 5′ flaps are removed (Fig. 8a). Schiml et al. assumed that in the absence of microhomologies at the break site, synthesis from the 3′ end and simultaneous degradation of the 5′ end, followed by ligation via cNHEJ, can lead to duplicated sequences (Fig. 8b) [111].

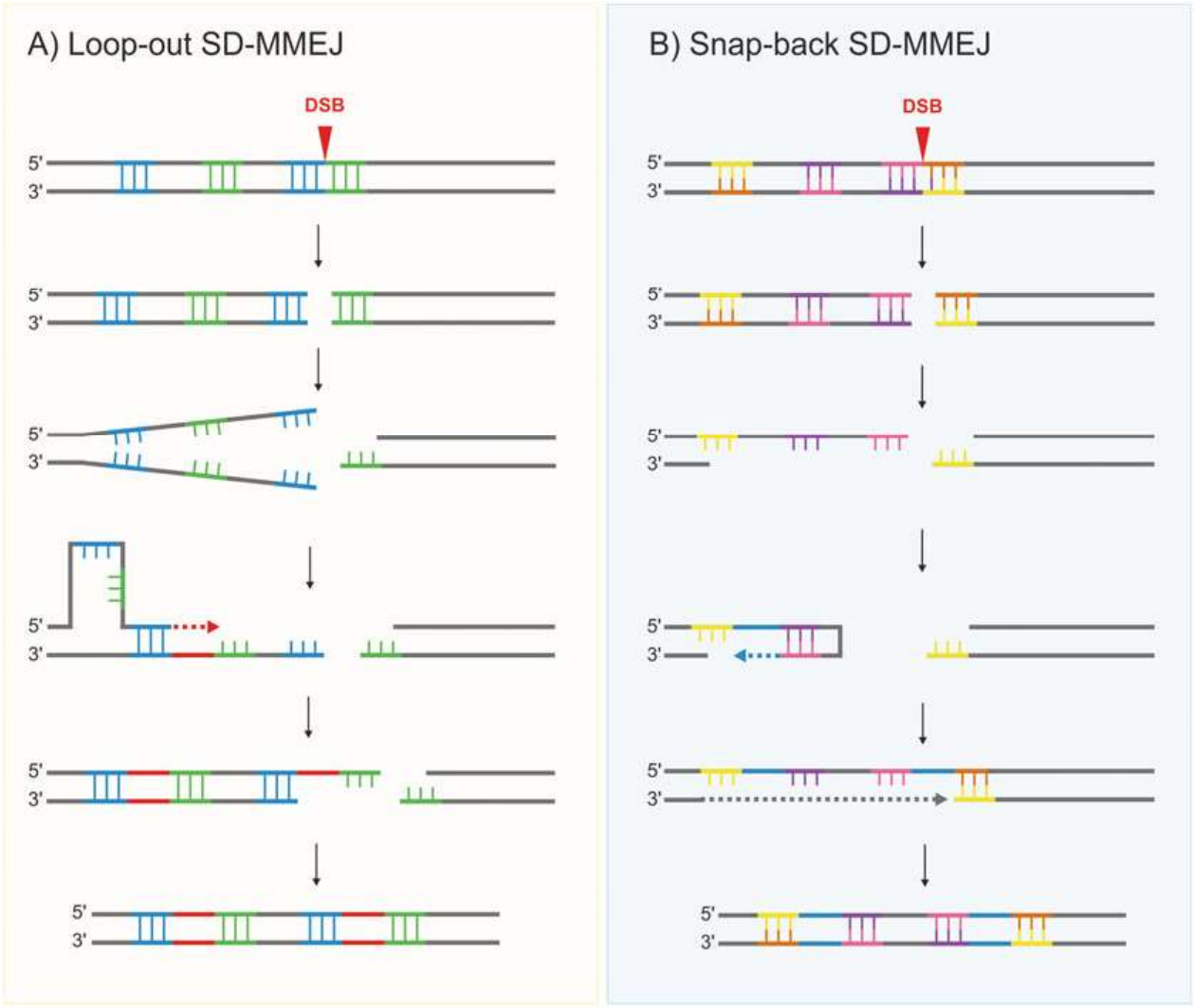


**Fig. 7** Mechanisms leading to insertions via SD-MMEJ. SD MMEJ is leading to direct or inverted repeats at the site of a DSB. (**a**) For an insertion via loop-out SD-MMEJ, two pairs of short direct repeats are required, with the first pair being located on one break side (blue) and the second pair spanning the DSB (green). Initially, the two complementary strands are separated and the blue homology close to the break site aligns with the second one distal to the DSB, forming a loop-out structure. Next, a template-dependent DNA polymerase is required for elongation and the newly synthesized part forms the insertion. The loop-out structure is then dissolved and the replicated microhomology (green) aligns with the second end of the DSB. (**b**) If an insertion is created by snap-back SD-MMEJ, the DSB should be flanked by two pairs of short inverted repeats. The first pair of microhomologies spans the DSB (yellow/orange), while the second pair is oriented in parallel (pink/purple). The strand separation occurs via end resection or unwinding of the DNA, and hereafter the second pair of homologies form a hairpin. Analogous to loop-out SD-MMEJ DNA, elongation is performed using a template-dependent DNA polymerase and the newly synthesized part forms the insertion. The replicated microhomology (orange) realigns with the other break site, resulting in an inverted repeat

In contrast, the induction of two paired SSBs on one DNA strand led to relatively low mutation frequencies below 1%. Nevertheless, deletions were obtained correlating with the SSB position and additionally insertions originating directly from adjacent sequences or from areas in close proximity (20–50 bp). The formation of deletions and inversions can possibly be associated with the

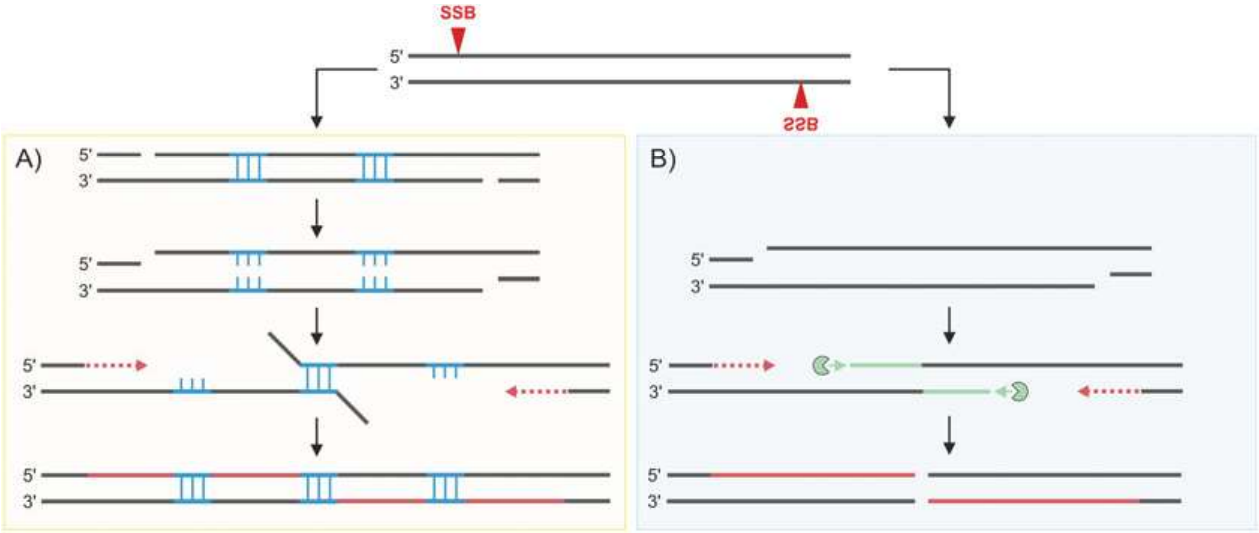


**Fig. 8** Adjacent SSBs on opposite DNA strands, leading to sequence duplications. The induction of two oppositely oriented SSBs results in a DSB with staggered ends. (**a**) When microhomologies (blue) are present in between the two SSBs, unwinding of the complementary strands is followed by realignment of the distal homology, located on both single strands. Template-dependent elongation (red) fills up the gaps, resulting in sequence duplications. (**b**) Without microhomologies, unwinding of the single-stranded break ends allows for parallel degradation of the 5′ ends (green) and polymerase-dependent synthesis from the 3′ ends (red). The resulting blunt ends are subsequently ligated via NHEJ, resulting in sequence duplications that are independent of microhomologies

release of the single-stranded DNA fragment in between the two SSBs. Breakage of this single-stranded gap can end up in a DSB, with overhanging ends that need to be processed, resulting in a larger deletion. Furthermore, the formation of insertions may be supported by replication slippage. It can be assumed that adjacent SSBs on opposing strands are highly mutagenic, while adjacent SSBs on the same strand are only slightly mutagenic, and both might be relevant for plant genome evolution, when occurring naturally.

Plant genome sizes vary between different species, but the related molecular mechanisms of genome size evolution are still under discussion. It is obvious that polyploidization and transposable element proliferation can lead to increased genome size [113, 114]. Genome shrinkage can be due to a dysploid reduction of chromosome number, a loss of entire chromosomes, as well as transposon-mediated excision, replication slippage, and illegitimate recombination [115, 116]. Besides these factors, recent data suggest that inaccurate DSB repair may have an underestimated role in genome size evolution. Initially, it was shown that expansion through retrotransposon proliferation can be compensated by an increase in DNA removal during DSB repair, in small-sized genomes [117]. Thereafter, species-specific differences between *Arabidopsis thaliana* and *Nicotiana tabacum* were observed for deletions occurring during the classical NHEJ pathway, and Kirik

et al. demonstrated a positive correlation between genome size and DSB processing via NHEJ [59]. In the much larger tobacco genome, deletions were about 30% smaller than in *Arabidopsis* and were accompanied, in 50% of the cases, by the insertion of filler sequences. The larger deletions in the latter species rarely coincided with insertions of only a few base pairs. Such differences are believed to originate from different exonucleolytic degradation rates of free DSB ends before downstream processing, thus shaping plant genomes during evolutionary timescales [60, 118]. In addition, the data of Vu et al. showed significant differences in DSB repair outcomes between *Arabidopsis* and barley [58]. Larger deletions were obtained on average for *Arabidopsis* and the proportion of mutated reads, containing a large deletion, was significantly higher in *Arabidopsis* than in barley. Furthermore, Vu et al. found strong distinctions between the proportion of insertions [58]. In *Arabidopsis*, only 5.3% of imprecisely repaired sequences showed insertions, as compared to 22.9% insertions in barley. Therefore, inaccurate DSB repair and repair of adjacent SSBs can be a significant source of genome variation and is involved in eukaryotic genome shaping.

## 4 Genomic Modifications

Combining the knowledge about DNA repair and the specific induction of DSBs using programmable site-specific nucleases provides a powerful tool for successful genome engineering. Synthetic nucleases are able to induce one or more DSBs at defined loci, predominantly repaired via NHEJ. As NHEJ is mostly error-prone, this allows for performing highly specific gene knockouts. Unfortunately, the outcome of NHEJ mediated repair can hardly be predicted. A study in human cell culture analyzing DNA repair profiles of hundreds of target sites of Cas9 indicates that repair outcomes are rather determined by the protospacer sequence than the genomic context [119]. In contrast, Vu et al. analyzed the mutation spectra after break induction via Cas9, at three selected endogenous loci, and concluded that the sequence context in close vicinity to the break has an impact on the used repair pathway and thus on repair outcomes [49]. Therefore, it remains to be elucidated as to how DSBs are repaired in a specific sequence context around the break site and how these different repair outcomes can be explained.

To bring matters to a head in the field of genome engineering, synthetic nucleases, and especially the Cas9/Cpf1 endonucleases, allow for the induction of more than one break, simultaneously. Therefore, a combination of the respective Cas9/Cpf1 nucleases with multiple spacer sequences and their suitable (s)gRNA sequences is needed for multiplexing [120–124]. Usually, the

repair of more than one DSB at a time is catalyzed by mechanisms that keep broken ends in close proximity, to reduce the risk of chromosomal aberrations [125]. As these mechanisms are not always accurate, it is possible that two originally not-associated DNA strands align, resulting in deletions, inversions, or translocations. Whole-genome alignments of the two *Arabidopsis* ecotypes, *Columbia* (Col-0) and *Landsberg erecta* (Ler-1), showed that there are several inversions and inter- and intrachromosomal transpositions, which occurred naturally in those *Brassicaceae* genomes [126]. For this reason, it is possible that a low number of break sites results in the alignment of unequal strands when several breaks are induced within the genome. When two DSBs are induced within close proximity, in plants, the part between the two breaks can be deleted from the genome [47, 127–129] (Fig. 9a). In particular, deletions smaller than 100 bp can be induced easily; however it was also shown that deletions up to 245 kb can occur at a low frequency in *Arabidopsis* and rice [130, 131]. This allows for the deletion of whole genes as opposed to mutating only single bases or functional domains. To achieve a controlled deletion, different repair reactions need to be considered. The direct way is to join the broken ends via cNHEJ, leading to a more or less precise elimination of the part in between. Alternatively, when (micro-) homologies are present, the alignment of the break sites is mediated by MMEJ or SSA repair factors; thus it is possible to obtain additional deletions at the junction site. Besides the targeted deletion, two break sites can provide a basis for a flip of the flanked fragment and the integration in an inverted orientation (Fig. 9b). In *Arabidopsis*, there is a 2.1 Mbp heterochromatic Knob (hk4S) region and

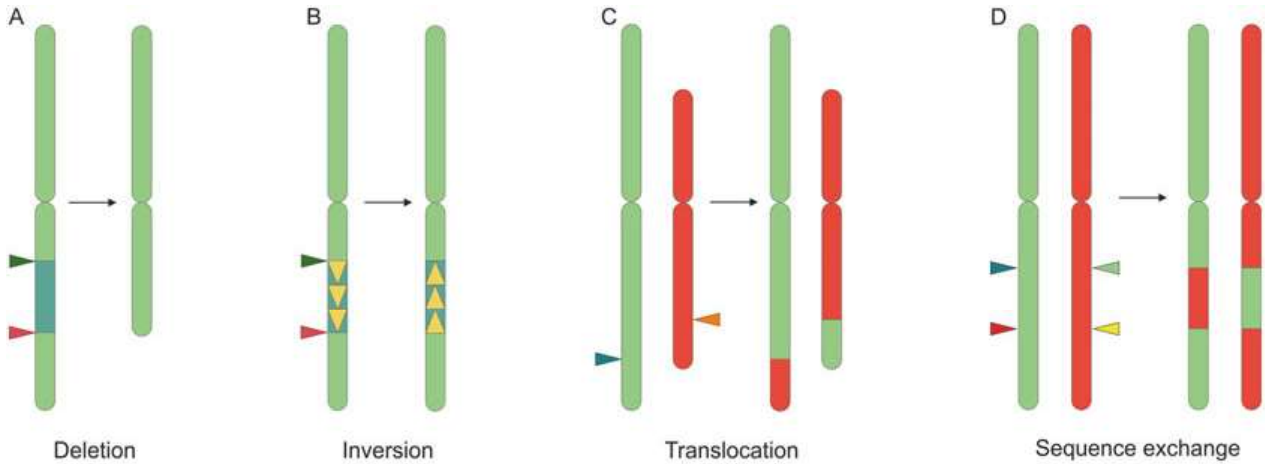


**Fig. 9** Induction of genomic rearrangements using synthetic nucleases. Applications of genome engineering by the induction of several breaks at a time can lead to the formation of chromosomal rearrangements. When two breaks are induced in relative close proximity on the same chromosome arm, a deletion (**a**) or inversion (**b**) of the internal part can occur. (**c**) Furthermore, break induction can take place on different chromosomes, leading to an exchange of chromosome parts. (**d**) In the case of a simultaneous induction of four DSBs (two on each chromosome), an exchange of sequences between different chromosomes is feasible

it is postulated that this originates from an inversion event that moved DNA sequences from the pericentromere to a more distal position in the euchromatin [132]. Targeted inversions induced by ZFNs and Cas9 were shown in mammalian cells, as well as in *Arabidopsis* [133–135]. Furthermore, besides break induction on the same chromosome arm, it is possible to induce several breaks on different chromosome arms, to exchange sequences within a plant genome (Fig. 9c). Formation of a DSB-induced translocation has been demonstrated in tobacco, using the homing endonuclease I-*Sce*I, in which two previously unlinked broken ends were rejoined via NHEJ or SSA [136]. Interestingly, there is a study in mammalian cell lines showing an increase in translocation frequency when cells are lacking KU70, and being even higher if KU70 and POLΘ are mutated [96]. The induction of at least four breaks allows for the exchange of chromosomal segments. For this purpose, in each case, two DSBs should be induced on both ends containing the target sequence to be exchanged (Fig. 9d). Gene replacements, involving an incoming T-DNA, were performed using ZFNs in plant cells [137]. There are several studies investigating gene replacements with the aid of homology arms [138–140]; however, this technology is even more attractive if this can occur by NHEJ and thus without the use of existing homologies as shown for rice [110].

Taken together, the current understanding of DSB repair in plants, combined with the tremendous possibilities provided by the increasingly available site-specific and adaptable nucleases, allows for unprecedented progress in plant genome engineering and thus provides the tool for a new green revolution, with significant progress in plant breeding technologies. The following book chapters will provide a more in-depth insight into the molecular tools to be applied in state-of-the-art genome engineering approaches.

## References


1. Watanabe K, Yamada N, Takeuchi Y (2006) Oxidative DNA damage in cucumber cotyledons irradiated with ultraviolet light. J Plant Res 119(3):239–246. https://doi.org/10.1007/s10265-006-0266-2
2. Sharma P, Jha AB, Dubey RS et al (2012) Reactive oxygen species, oxidative damage, and antioxidative defense mechanism in plants under stressful conditions. J Bot 2012 (1):1–26. https://doi.org/10.1155/2012/217037
3. Sikora P, Chawade A, Larsson M et al (2011) Mutagenesis as a tool in plant genetics, functional genomics, and breeding. Int J Plant Genomics 2011:314829. https://doi.org/10.1155/2011/314829
4. Chen L, Hao L, Parry MAJ et al (2014) Progress in TILLING as a tool for functional genomics and improvement of crops. J Integr Plant Biol 56(5):425–443. https://doi.org/10.1111/jipb.12192
5. Till BJ, Reynolds SH, Greene EA et al (2003) Large-scale discovery of induced point mutations with high-throughput TILLING. Genome Res 13(3):524–530. https://doi.org/10.1101/gr.977903
6. Stingele J, Jentsch S (2015) DNA-protein crosslink repair. Nat Rev Mol Cell Biol 16 (8):455–460. https://doi.org/10.1038/nrm4015
7. Schärer OD (2003) Chemistry and biology of DNA repair. Angew Chem Int Ed Engl 42

(26):2946–2974. https://doi.org/10.1002/anie.200200523
8. Stadler LJ (1928) Genetic effects of X-rays in maize. Proc Natl Acad Sci U S A 14(1):69–75
9. Muller HJ (1927) Artificial transmutation of the gene. Science 66(1699):84–87. https://doi.org/10.1126/science.66.1699.84
10. Nakagawa H (2009) Induced mutations in plant breeding and biological researches in Japan. Rome: 51
11. West CE, Waterworth WM, Story GW et al (2002) Disruption of the Arabidopsis AtKu80 gene demonstrates an essential role for AtKu80 protein in efficient repair of DNA double-strand breaks in vivo. Plant J 31 (4):517–528
12. Manova V, Gecheff K, Stoilov L (2006) Efficient repair of bleomycin-induced double-strand breaks in barley ribosomal genes. Mutat Res 601(1–2):179–190. https://doi.org/10.1016/j.mrfmmm.2006.07.004
13. Mannuss A, Trapp O, Puchta H (2012) Gene regulation in response to DNA damage. Biochim Biophys Acta 1819(2):154–165. https://doi.org/10.1016/j.bbagrm.2011.08.003
14. Alonso JM, Stepanova AN, Leisse TJ et al (2003) Genome-wide insertional mutagenesis of Arabidopsis thaliana. Science 301 (5633):653–657. https://doi.org/10.1126/science.1086391
15. Schiml S, Puchta H (2016) Revolutionizing plant biology: multiple ways of genome engineering by CRISPR/Cas. Plant Methods 12:8. https://doi.org/10.1186/s13007-016-0103-0
16. Jacquier A, Dujon B (1985) An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. Cell 41(2):383–394
17. Perrin A, Buckle M, Dujon B (1993) Asymmetrical recognition and activity of the I-SceI endonuclease on its site and on intron-exon junctions. EMBO J 12(7):2939–2947
18. Monteilhet C, Perrin A, Thierry A et al (1990) Purification and characterization of the in vitro activity of I-Sce I, a novel and highly specific endonuclease encoded by a group I intron. Nucleic Acids Res 18 (6):1407–1413
19. Moure CM, Gimble FS, Quiocho FA (2003) The crystal structure of the gene targeting homing endonuclease I-SceI reveals the origins of its target site specificity. J Mol Biol 334 (4):685–695
20. Puchta H, Dujon B, Hohn B (1993) Homologous recombination in plant cells is enhanced by in vivo induction of double strand breaks into DNA by a site-specific endonuclease. Nucleic Acids Res 21 (22):5034–5040
21. Kim YG, Cha J, Chandrasegaran S (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci U S A 93(3):1156–1160
22. Voytas DF (2013) Plant genome engineering with sequence-specific nucleases. Annu Rev Plant Biol 64:327–350. https://doi.org/10.1146/annurev-arplant-042811-105552
23. Wiedenheft B, Sternberg SH, Doudna JA (2012) RNA-guided genetic silencing systems in bacteria and archaea. Nature 482 (7385):331–338. https://doi.org/10.1038/nature10886
24. Jinek M, Chylinski K, Fonfara I et al (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337(6096):816–821. https://doi.org/10.1126/science.1225829
25. Pacher M, Puchta H (2017) From classical mutagenesis to nuclease-based breeding - directing natural DNA repair for a natural end-product. Plant J 90(4):819–833. https://doi.org/10.1111/tpj.13469
26. Zhang K, Raboanatahiry N, Zhu B et al (2017) Progress in genome editing technology and its application in plants. Front Plant Sci 8:177. https://doi.org/10.3389/fpls.2017.00177
27. Puchta H (2017) Applying CRISPR/Cas for genome engineering in plants: the best is yet to come. Curr Opin Plant Biol 36:1–8. https://doi.org/10.1016/j.pbi.2016.11.011
28. Gorbunova V, Levy AA (1999) How plants make ends meet: DNA double-strand break repair. Trends Plant Sci 4(7):263–269
29. Puchta H (2005) The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. J Exp Bot 56 (409):1–14. https://doi.org/10.1093/jxb/eri025
30. Szostak JW, Orr-Weaver TL, Rothstein RJ et al (1983) The double-strand-break repair model for recombination. Cell 33(1):25–35
31. Osman K, Higgins JD, Sanchez-Moran E et al (2011) Pathways to meiotic recombination in Arabidopsis thaliana. New Phytol 190 (3):523–544. https://doi.org/10.1111/j.1469-8137.2011.03665.x
32. Eschbach V, Kobbe D (2014) Different replication protein A complexes of Arabidopsis thaliana have different DNA-binding properties as a function of heterotrimer composition.

Plant Cell Physiol 55(8):1460–1472. https://doi.org/10.1093/pcp/pcu076

33. van Dyck E, Stasiak AZ, Stasiak A et al (2001) Visualization of recombination intermediates produced by RAD52-mediated single-strand annealing. EMBO Rep 2(10):905–909. https://doi.org/10.1093/embo-reports/kve201
34. Samach A, Melamed-Bessudo C, Avivi-Ragolski N et al (2011) Identification of plant RAD52 homologs and characterization of the Arabidopsis thaliana RAD52-like genes. Plant Cell 23(12):4266–4279. https://doi.org/10.1105/tpc.111.091744
35. Serra H, Da Ines O, Degroote F et al (2013) Roles of XRCC2, RAD51B and RAD51D in RAD51-independent SSA recombination. PLoS Genet 9(11):e1003971. https://doi.org/10.1371/journal.pgen.1003971
36. Sugawara N, Ira G, Haber JE (2000) DNA length dependence of the single-strand annealing pathway and the role of Saccharomyces cerevisiae RAD59 in double-strand break repair. Mol Cell Biol 20 (14):5300–5309
37. Dubest S, Gallego ME, White CI (2002) Role of the AtRad1p endonuclease in homologous recombination in plants. EMBO Rep 3 (11):1049–1054. https://doi.org/10.1093/embo-reports/kvf211
38. Mannuss A, Dukowic-Schulze S, Suer S et al (2010) RAD5A, RECQ4A, and MUS81 have specific functions in homologous recombination and define different pathways of DNA repair in Arabidopsis thaliana. Plant Cell 22 (10):3318–3330. https://doi.org/10.1105/tpc.110.078568
39. Bleuyard J-Y, Gallego ME, Savigny F et al (2005) Differing requirements for the Arabidopsis Rad51 paralogs in meiosis and DNA repair. Plant J 41(4):533–545. https://doi.org/10.1111/j.1365-313X.2004.02318.x
40. Charlot F, Chelysheva L, Kamisugi Y et al (2014) RAD51B plays an essential role during somatic and meiotic recombination in Physcomitrella. Nucleic Acids Res 42 (19):11965–11978. https://doi.org/10.1093/nar/gku890
41. Osakabe K, Abe K, Yoshioka T et al (2006) Isolation and characterization of the RAD54 gene from Arabidopsis thaliana. Plant J 48 (6):827–842. https://doi.org/10.1111/j.1365-313X.2006.02927.x
42. Roth N, Klimesch J, Dukowic-Schulze S et al (2012) The requirement for recombination factors differs considerably between different pathways of homologous double-strand break repair in somatic plant cells. Plant J 72 (5):781–790. https://doi.org/10.1111/j.1365-313X.2012.05119.x
43. Dulieu HL (1975) Somatic variations on a yellow mutant in Nicotianatabacum L. (a1+/a1a2+/a2) II. Reciprocal genetic events occurring in leaf cells. Mutat Res Fundam Mol Mech Mutagen 28(1):69–77. https://doi.org/10.1016/0027-5107(75)90316-4
44. Carlson PS (1974) Mitotic crossing-over in a higher plant. Genet Res 24(1):109–112. https://doi.org/10.1017/S0016672300015123
45. Rinehart TA, Dean C, Weil CF (1997) Comparative analysis of non-random DNA repair following Ac transposon excision in maize and Arabidopsis. Plant J 12(6):1419–1427
46. Athma P, Peterson T (1991) Ac induces homologous recombination at the maize P locus. Genetics 128(1):163–173
47. Siebert R, Puchta H (2002) Efficient repair of genomic double-strand breaks by homologous recombination between directly repeated sequences in the plant genome. Plant Cell 14(5):1121–1131
48. Orel N, Kyryk A, Puchta H (2003) Different pathways of homologous recombination are used for the repair of double-strand breaks within tandemly arranged sequences in the plant genome. Plant J 35(5):604–612
49. Vu GTH, Cao HX, Fauser F et al (2017) Endogenous sequence patterns predispose the repair modes of CRISPR/Cas9-induced DNA double strand breaks in Arabidopsis thaliana. Plant J 92(1):57–67. https://doi.org/10.1111/tpj.13634
50. Puchta H (1998) Repair of genomic double-strand breaks in somatic plant cells by one-sided invasion of homologous sequences. Plant J 13(3):331–339. https://doi.org/10.1046/j.1365-313X.1998.00035.x
51. Watanabe K, Pacher M, Dukowic S et al (2009) The structural maintenance of chromosomes 5/6 complex promotes sister chromatid alignment and homologous recombination after DNA damage in Arabidopsis thaliana. Plant Cell 21(9):2688–2699. https://doi.org/10.1105/tpc.108.060525
52. Vu GTH, Cao HX, Watanabe K et al (2014) Repair of site-specific DNA double-strand breaks in barley occurs via diverse pathways primarily involving the sister chromatid. Plant Cell 26(5):2156–2167. https://doi.org/10.1105/tpc.114.126607
53. Chiurazzi M, Ray A, Viret JF et al (1996) Enhancement of somatic intrachromosomal homologous recombination in Arabidopsis

by the HO endonuclease. Plant Cell 8 (11):2057–2066. https://doi.org/10.1105/tpc.8.11.2057

54. Gisler B, Salomon S, Puchta H (2002) The role of double-strand break-induced allelic homologous recombination in somatic plant cells. Plant J 32(3):277–284
55. Filler Hayut S, Melamed Bessudo C, Levy AA (2017) Targeted recombination between homologous chromosomes for precise breeding in tomato. Nat Commun 8:15605. https://doi.org/10.1038/ncomms15605
56. Puchta H (1999) DSB-induced recombination between ectopic homologous sequences in somatic plant cells. Genetics 152:1173–1181
57. Shalev G, Levy AA (1997) The maize transposable element Ac induces recombination between the donor site and an homologous ectopic sequence. Genetics 146 (3):1143–1151
58. Vu GTH, Cao HX, Reiss B et al (2017) Deletion-bias in DNA double-strand break repair differentially contributes to plant genome shrinkage. New Phytol 214 (4):1712–1721. https://doi.org/10.1111/nph.14490
59. Kirik A, Salomon S, Puchta H (2000) Species-specific double-strand break repair and genome evolution in plants. EMBO J 19 (20):5562–5566. https://doi.org/10.1093/emboj/19.20.5562
60. Orel N, Puchta H (2003) Differences in the processing of DNA ends in Arabidopsis thaliana and tobacco: possible implications for genome evolution. Plant Mol Biol 51 (4):523–531
61. Steinert J, Schiml S, Puchta H (2016) Homology-based double-strand break-induced genome engineering in plants. Plant Cell Rep 35(7):1429–1438. https://doi.org/10.1007/s00299-016-1981-3
62. Paszkowski J, Baur M, Bogucki A et al (1988) Gene targeting in plants. EMBO J 7 (13):4021–4026
63. Offringa R, de Groot MJ, Haagsman HJ et al (1990) Extrachromosomal homologous recombination and gene targeting in plant cells after Agrobacterium mediated transformation. EMBO J 9(10):3077–3084
64. Beetham PR, Kipp PB, Sawycky XL et al (1999) A tool for functional plant genomics: chimeric RNA/DNA oligonucleotides cause in vivo gene-specific mutations. Proc Natl Acad Sci U S A 96(15):8774–8778
65. Puchta H, Dujon B, Hohn B (1996) Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. Proc Natl Acad Sci U S A 93(10):5055–5060
66. Wright DA, Townsend JA, Winfrey RJ et al (2005) High-frequency homologous recombination in plants mediated by zinc-finger nucleases. Plant J 44(4):693–705. https://doi.org/10.1111/j.1365-313X.2005.02551.x
67. Shukla VK, Doyon Y, Miller JC et al (2009) Precise genome modification in the crop species Zea mays using zinc-finger nucleases. Nature 459(7245):437–441. https://doi.org/10.1038/nature07992
68. Townsend JA, Wright DA, Winfrey RJ et al (2009) High-frequency modification of plant genes using engineered zinc-finger nucleases. Nature 459(7245):442–445. https://doi.org/10.1038/nature07845
69. de Pater S, Pinas JE, Hooykaas PJJ et al (2013) ZFN-mediated gene targeting of the Arabidopsis protoporphyrinogen oxidase gene through Agrobacterium-mediated floral dip transformation. Plant Biotechnol J 11 (4):510–515. https://doi.org/10.1111/pbi.12040
70. Zhang Y, Zhang F, Li X et al (2013) Transcription activator-like effector nucleases enable efficient plant genome engineering. Plant Physiol 161(1):20–27. https://doi.org/10.1104/pp.112.205179
71. Svitashev S, Young JK, Schwartz C et al (2015) Targeted mutagenesis, precise gene editing, and site-specific gene insertion in maize using Cas9 and guide RNA. Plant Physiol 169(2):931–945. https://doi.org/10.1104/pp.15.00793
72. Sauer NJ, Narváez-Vásquez J, Mozoruk J et al (2016) Oligonucleotide-mediated genome editing provides precision and function to engineered nucleases and antibiotics in plants. Plant Physiol 170(4):1917–1928. https://doi.org/10.1104/pp.15.01696
73. Li Z, Liu Z-B, Xing A et al (2015) Cas9-guide RNA directed genome editing in soybean. Plant Physiol 169(2):960–970. https://doi.org/10.1104/pp.15.00783
74. Fauser F, Roth N, Pacher M et al (2012) In planta gene targeting. Proc Natl Acad Sci U S A 109(19):7535–7540. https://doi.org/10.1073/pnas.1202191109
75. Schiml S, Fauser F, Puchta H (2014) The CRISPR/Cas system can be used as nuclease for in planta gene targeting and as paired nickases for directed mutagenesis in Arabidopsis resulting in heritable progeny. Plant J

80(6):1139–1150. https://doi.org/10.1111/tpj.12704

76. Baltes NJ, Gil-Humanes J, Cermak T et al (2014) DNA replicons for plant genome engineering. Plant Cell 26(1):151–163. https://doi.org/10.1105/tpc.113.119792
77. Čermák T, Baltes NJ, Čegan R et al (2015) High-frequency, precise modification of the tomato genome. Genome Biol 16:232. https://doi.org/10.1186/s13059-015-0796-9
78. Even-Faitelson L, Samach A, Melamed-Bessudo C et al (2011) Localized egg-cell expression of effector proteins for targeted modification of the Arabidopsis genome. Plant J 68(5):929–937. https://doi.org/10.1111/j.1365-313X.2011.04741.x
79. Shaked H, Melamed-Bessudo C, Levy AA (2005) High-frequency gene targeting in Arabidopsis plants expressing the yeast RAD54 gene. Proc Natl Acad Sci U S A 102 (34):12265–12269. https://doi.org/10.1073/pnas.0502601102
80. Qi Y, Zhang Y, Zhang F et al (2013) Increasing frequencies of site-specific mutagenesis and gene targeting in Arabidopsis by manipulating DNA repair pathways. Genome Res 23(3):547–554. https://doi.org/10.1101/gr.145557.112
81. Endo M, Mikami M, Toki S (2016) Biallelic gene targeting in rice. Plant Physiol 170 (2):667–677. https://doi.org/10.1104/pp.15.01663
82. Mimitou EP, Symington LS (2010) Ku prevents Exo1 and Sgs1-dependent resection of DNA ends in the absence of a functional MRX complex or Sae2. EMBO J 29 (19):3358–3369. https://doi.org/10.1038/emboj.2010.193
83. Lieber MR (2010) The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu Rev Biochem 79:181–211. https://doi.org/10.1146/annurev.biochem.052308.093131
84. Tamura K, Adachi Y, Chiba K et al (2002) Identification of Ku70 and Ku80 homologues in Arabidopsis thaliana: evidence for a role in the repair of DNA double-strand breaks. Plant J 29(6):771–781
85. West CE, Waterworth WM, Jiang Q et al (2000) Arabidopsis DNA ligase IV is induced by gamma-irradiation and interacts with an Arabidopsis homologue of the double strand break repair protein XRCC4. Plant J 24 (1):67–78
86. Friesner J, Britt AB (2003) Ku80- and DNA ligase IV-deficient plants are sensitive to ionizing radiation and defective in T-DNA integration. Plant J 34(4):427–440
87. Nishizawa-Yokoi A, Nonaka S, Saika H et al (2012) Suppression of Ku70/80 or Lig4 leads to decreased stable transformation and enhanced homologous recombination in rice. New Phytol 196(4):1048–1059. https://doi.org/10.1111/j.1469-8137.2012.04350.x
88. Gallego ME, Jalut N, White CI (2003) Telomerase dependence of telomere lengthening in Ku80 mutant Arabidopsis. Plant Cell 15 (3):782–789
89. Stolarek M, Gruszka D, Braszewska-Zalewska A et al (2015) Functional analysis of the new barley gene HvKu80 indicates that it plays a key role in double-strand DNA break repair and telomere length regulation. Mutagenesis 30(6):785–797. https://doi.org/10.1093/mutage/gev033
90. Bleuyard J-Y, Gallego ME, White CI (2006) Recent advances in understanding of the DNA double-strand break repair machinery of plants. DNA Repair (Amst) 5(1):1–12. https://doi.org/10.1016/j.dnarep.2005.08.017
91. Charbonnel C, Gallego ME, White CI (2010) Xrcc1-dependent and Ku-dependent DNA double-strand break repair kinetics in Arabidopsis plants. Plant J 64(2):280–290. https://doi.org/10.1111/j.1365-313X.2010.04331.x
92. Jia Q, den Dulk-Ras A, Shen H et al (2013) Poly(ADP-ribose)polymerases are involved in microhomology mediated back-up non-homologous end joining in Arabidopsis thaliana. Plant Mol Biol 82(4–5):339–351. https://doi.org/10.1007/s11103-013-0065-9
93. Shen H, Strunks GD, Klemann BJPM et al (2017) CRISPR/Cas9-induced double-strand break repair in arabidopsis nonhomologous end-joining mutants. G3 (Bethesda) 7 (1):193–202. https://doi.org/10.1534/g3.116.035204
94. Wang H, Xu X (2017) Microhomology-mediated end joining: new players join the team. Cell Biosci 7:6. https://doi.org/10.1186/s13578-017-0136-8
95. Truong LN, Li Y, Shi LZ et al (2013) Microhomology-mediated end joining and homologous recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. Proc Natl Acad Sci U S A 110(19):7720–7725. https://doi.org/10.1073/pnas.1213431110
96. Wyatt DW, Feng W, Conlin MP et al (2016) Essential roles for polymerase θ-mediated end joining in the repair of chromosome breaks.

Mol Cell 63(4):662–673. https://doi.org/10.1016/j.molcel.2016.06.020

97. Zahn KE, Averill AM, Aller P et al (2015) Human DNA polymerase θ grasps the primer terminus to mediate DNA repair. Nat Struct Mol Biol 22(4):304–311. https://doi.org/10.1038/nsmb.2993
98. Bennardo N, Cheng A, Huang N et al (2008) Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. PLoS Genet 4(6):e1000110. https://doi.org/10.1371/journal.pgen.1000110
99. Chan SH, Yu AM, McVey M (2010) Dual roles for DNA polymerase theta in alternative end-joining repair of double-strand breaks in Drosophila. PLoS Genet 6(7):e1001005. https://doi.org/10.1371/journal.pgen.1001005
100. Kent T, Chandramouly G, McDevitt SM et al (2015) Mechanism of microhomology-mediated end-joining promoted by human DNA polymerase θ. Nat Struct Mol Biol 22(3):230–237. https://doi.org/10.1038/nsmb.2961
101. Mateos-Gomez PA, Gong F, Nair N et al (2015) Mammalian polymerase [thgr] promotes alternative NHEJ and suppresses recombination. Nature 518(7538):254–257
102. Ceccaldi R, Liu JC, Amunugama R et al (2015) Homologous-recombination-deficient tumours are dependent on Pol-θ-mediated repair. Nature 518(7538):258–262. https://doi.org/10.1038/nature14184
103. Inagaki S, Nakamura K, Morikami A (2009) A link among DNA replication, recombination, and gene expression revealed by genetic and genomic analysis of TEBICHI gene of Arabidopsis thaliana. PLoS Genet 5(8):e1000613. https://doi.org/10.1371/journal.pgen.1000613
104. Klemm T, Mannuß A, Kobbe D et al (2017) The DNA translocase RAD5A acts independently of the other main DNA repair pathways and requires both its ATPase and RING domain for activity in Arabidopsis thaliana. Plant J 91(4):725–740. https://doi.org/10.1111/tpj.13602
105. van Kregten M, de Pater S, Romeijn R et al (2016) T-DNA integration in plants results from polymerase-θ-mediated DNA repair. Nat Plants 2(11):16164. https://doi.org/10.1038/nplants.2016.164
106. Yu AM, McVey M (2010) Synthesis-dependent microhomology-mediated end joining accounts for multiple types of repair junctions. Nucleic Acids Res 38(17):5706–5717. https://doi.org/10.1093/nar/gkq379
107. Gorbunova V, Levy AA (1997) Non-homologous DNA end joining in plant cells is associated with deletions and filler DNA insertions. Nucleic Acids Res 25(22):4650–4657
108. Salomon S, Puchta H (1998) Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. EMBO J 17(20):6086–6095. https://doi.org/10.1093/emboj/17.20.6086
109. Geisinger JM, Turan S, Hernandez S et al (2016) In vivo blunt-end cloning through CRISPR/Cas9-facilitated non-homologous end-joining. Nucleic Acids Res 44(8):e76. https://doi.org/10.1093/nar/gkv1542
110. Li J, Meng X, Zong Y et al (2016) Gene replacements and insertions in rice by intron targeting using CRISPR-Cas9. Nat Plants 2:16139. https://doi.org/10.1038/nplants.2016.139
111. Schiml S, Fauser F, Puchta H (2016) Repair of adjacent single-strand breaks is often accompanied by the formation of tandem sequence duplications in plant genomes. Proc Natl Acad Sci U S A 113(26):7266–7271. https://doi.org/10.1073/pnas.1603823113
112. Vaughn JN, Bennetzen JL (2014) Natural insertions in rice commonly form tandem duplications indicative of patch-mediated double-strand break induction and repair. Proc Natl Acad Sci U S A 111(18):6684–6689. https://doi.org/10.1073/pnas.1321854111
113. Adams KL, Wendel JF (2005) Polyploidy and genome evolution in plants. Curr Opin Plant Biol 8(2):135–141. https://doi.org/10.1016/j.pbi.2005.01.001
114. Bennetzen JL, Wang H (2014) The contributions of transposable elements to the structure, function, and evolution of plant genomes. Annu Rev Plant Biol 65:505–530. https://doi.org/10.1146/annurev-arplant-050213-035811
115. Devos KM, Brown JKM, Bennetzen JL (2002) Genome size reduction through illegitimate recombination counteracts genome expansion in Arabidopsis. Genome Res 12(7):1075–1079. https://doi.org/10.1101/gr.132102
116. Schubert I, Vu GTH (2016) Genome Stability and Evolution: Attempting a Holistic View. Trends Plant Sci 21(9):749–757. https://doi.org/10.1016/j.tplants.2016.06.003

117. Hawkins JS, Proulx SR, Rapp RA et al (2009) Rapid DNA loss as a counterbalance to genome expansion through retrotransposon proliferation in plants. Proc Natl Acad Sci U S A 106(42):17811–17816. https://doi.org/10.1073/pnas.0904339106
118. Cavalier-Smith T (2005) Economy, speed and size matter: evolutionary forces driving nuclear genome miniaturization and expansion. Ann Bot 95(1):147–175. https://doi.org/10.1093/aob/mci010
119. van Overbeek M, Capurso D, Carter MM et al (2016) DNA repair profiling reveals nonrandom outcomes at Cas9-mediated breaks. Mol Cell 63(4):633–646. https://doi.org/10.1016/j.molcel.2016.06.037
120. Cong L, Ran FA, Cox D et al (2013) Multiplex genome engineering using CRISPR/Cas systems. Science 339(6121):819–823. https://doi.org/10.1126/science.1231143
121. Li J-F, Norville JE, Aach J et al (2013) Multiplex and homologous recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. Nat Biotechnol 31(8):688–691. https://doi.org/10.1038/nbt.2654
122. Wang H, Yang H, Shivalila CS et al (2013) One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell 153(4):910–918. https://doi.org/10.1016/j.cell.2013.04.025
123. Wang M, Mao Y, Lu Y et al (2017) Multiplex gene editing in rice using the CRISPR-Cpf1 system. Mol Plant 10(7):1011–1013. https://doi.org/10.1016/j.molp.2017.03.001
124. Zetsche B, Heidenreich M, Mohanraju P et al (2017) Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. Nat Biotechnol 35(1):31–34. https://doi.org/10.1038/nbt.3737
125. Williams GJ, Lees-Miller SP, Tainer JA (2010) Mre11-Rad50-Nbs1 conformations and the control of sensing, signaling, and effector responses at DNA double-strand breaks. DNA Repair (Amst) 9(12):1299–1306. https://doi.org/10.1016/j.dnarep.2010.10.001
126. Zapata L, Ding J, Willing E-M et al (2016) Chromosome-level assembly of Arabidopsis thaliana Ler reveals the extent of translocation and inversion polymorphisms. Proc Natl Acad Sci U S A 113(28):60. https://doi.org/10.1073/pnas.1607532113
127. Petolino JF, Worden A, Curlee K et al (2010) Zinc finger nuclease-mediated transgene deletion. Plant Mol Biol 73(6):617–628. https://doi.org/10.1007/s11103-010-9641-4
128. Kapusi E, Corcuera-Gómez M, Melnik S et al (2017) Heritable genomic fragment deletions and small indels in the putative ENGase gene induced by CRISPR/Cas9 in barley. Front Plant Sci 8:540. https://doi.org/10.3389/fpls.2017.00540
129. Qi Y, Li X, Zhang Y et al (2013) Targeted deletion and inversion of tandemly arrayed genes in Arabidopsis thaliana using zinc finger nucleases. G3 (Bethesda) 3(10):1707–1715. https://doi.org/10.1534/g3.113.006270
130. Ordon J, Gantner J, Kemna J et al (2017) Generation of chromosomal deletions in dicotyledonous plants employing a user-friendly genome editing toolkit. Plant J 89(1):155–168. https://doi.org/10.1111/tpj.13319
131. Zhou H, Liu B, Weeks DP et al (2014) Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. Nucleic Acids Res 42(17):10903–10914. https://doi.org/10.1093/nar/gku806
132. Fransz PF, Armstrong S, de Jong JH et al (2000) Integrated cytogenetic map of chromosome arm 4S of A. thaliana: structural organization of heterochromatic knob and centromere region. Cell 100(3):367–376
133. Lee HJ, Kweon J, Kim E et al (2012) Targeted chromosomal duplications and inversions in the human genome using zinc finger nucleases. Genome Res 22(3):539–548. https://doi.org/10.1101/gr.129635.111
134. Li Y, Park AI, Mou H et al (2015) A versatile reporter system for CRISPR-mediated chromosomal rearrangements. Genome Biol 16:111. https://doi.org/10.1186/s13059-015-0680-7
135. Zhang C, Liu C, Weng J et al (2017) Creation of targeted inversion mutations in plants using an RNA-guided endonuclease. Crop J 5(1):83–88. https://doi.org/10.1016/j.cj.2016.08.001
136. Pacher M, Schmidt-Puchta W, Puchta H (2007) Two unlinked double-strand breaks can induce reciprocal exchanges in plant genomes via homologous recombination and nonhomologous end joining. Genetics 175(1):21–29. https://doi.org/10.1534/genetics.106.065185
137. Weinthal DM, Taylor RA, Tzfira T (2013) Nonhomologous end joining-mediated gene replacement in plant cells. Plant Physiol 162

(1):390–400. https://doi.org/10.1104/pp.112.212910

138. Zheng Q, Cai X, Tan MH et al (2014) Precise gene deletion and replacement using the CRISPR/Cas9 system in human cells. BioTechniques 57(3):115–124. https://doi.org/10.2144/000114196
139. Byrne SM, Ortiz L, Mali P et al (2015) Multi-kilobase homozygous targeted gene replacement in human induced pluripotent stem cells. Nucleic Acids Res 43(3):e21. https://doi.org/10.1093/nar/gku1246
140. Arazoe T, Miyoshi K, Yamato T et al (2015) Tailor-made CRISPR/Cas system for highly efficient targeted gene replacement in the rice blast fungus. Biotechnol Bioeng 112 (12):2543–2549. https://doi.org/10.1002/bit.25662