Journal of
Experimental
Botany

REVIEW ARTICLE

The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution

Holger Puchta*

Botany II, University of Karlsruhe, D-76128 Karlsruhe, Germany

Abstract

The efficient repair of double-strand breaks (DSBs) in genomic DNA is important for the survival of all organisms. In recent years, basic mechanisms of DSB repair in somatic plant cells have been elucidated. DSBs are mainly repaired by non-homologous end-joining (NHEJ). The repair can be associated with deletions, but also insertions due to copying genomic sequences from elsewhere into the break. Species-specific differences of NHEJ have been reported and an inverse correlation of deletion size to genome size has been postulated, indicating that NHEJ might contribute significantly to evolution of genome size. DSB repair by homologous recombination (HR) might also influence genome organization. Whereas homology present in an allelic or an ectopic position is hardly used for repair, the use of homologous sequences in close proximity to the break is frequent. A 'single-strand annealing' mechanism that leads to sequence deletions between direct repeats is particularly efficient. This might explain the accumulation of single long terminal repeats of retroelements in cereal genomes. The conservative 'synthesis-dependent strand annealing' mechanism, resulting in conversions without crossovers is also prominent and seems to be significant for the evolution of tandemly arranged gene families such as resistance genes. Induction of DSBs could be used as a means for the controlled manipulation of plant genomes in an analogous way for the use of marker gene excision and site-specific integration.

Key words: *Arabidopsis*, deletion formation, gene conversion, homologous recombination, I-*Sce*I, non-homologous end-joining, single-strand annealing, synthesis-dependent strand-annealing, tobacco.

Introduction

Double-strand breaks (DSBs) have to be eliminated before genomes can be replicated. Therefore, the repair of DSBs is critical for the survival of all organisms. Generally, DSBs can be repaired via two different pathways, either via homologous recombination (HR) or via non-homologous end-joining (NHEJ; also known as illegitimate recombination). Whereas in the former pathway the sequences are linked in regions that are identical to each other, in the latter the sequence information does not play a major role in the rejoining of the two double strands. An important question of recent research has been to find out under what conditions each pathway is used (for a general review see Paques and Haber, 1999). Many studies have been performed using bacteria and, in the last decade, the basic mechanisms of DSB repair has been determined in yeast. In recent years the focus of research has shifted to higher eukaryotes, particularly as rare cutting restriction endonucleases like I-*Sce*I (Perrin *et al.*, 1993) have become available for the induction of unique site-specific genomic DSBs (Jasin, 1996). The first report on the use of the endonuclease I-*Sce*I in a higher eukaryote, tobacco, dates back to 1993 (Puchta *et al.*, 1993). Since then, in the mammalian field, more than 100 studies on DSB repair using restriction endonucleases have appeared in the literature, and on plants many studies, not only based on restriction enzymes but also on transposable elements as means to induce breaks, have been published. Based on these data a broader view of the pathways involved in DSB repair in plants is now possible. After a short introduction of the means to induce breaks, the main part of this review will focus on the role of the two basic pathways of DSB repair, HR and NHEJ. In addition, besides presenting a detailed picture of the ways that breaks are repaired, the possible evolutionary consequences for plant genomes will be discussed. Being able to predict the outcome of DSB

* Fax: +49 721 6084874. E-mail: holger.puchta@bio.uka.de

repair reactions, the induction of DSB can also be used for the controlled induction of genomic changes in plants, a point that will be described in the final section of this review. As many recent publications have described the current understanding of the enzyme machinery involved in DNA repair in plants and the problems of gene targeting, these topics will not be addressed (Britt and May, 2003; Hanin and Paszkowski, 2003; Hartung and Puchta, 2004; Hays, 2002; Puchta, 2002, 2003*a*; Ray and Langer, 2002; Reiss, 2003).

Means to make ends

The prerequisite for the study of DSB repair is the production of DNA breaks in the plant nucleus at a given time point. In principle, DNA that has already been broken can be transferred into the nucleus or a break can be induced *in vivo* (e.g. by radiation, application of chemicals, during the replication of transposons, or by the use of restriction endonucleases). Plasmid molecules linearized by restriction enzymes *in vitro* ('broken' DNA) are used as a substrate for the transient transformation of plant cells. Transfer can be mediated by PEG transfection, electroporation or particle bombardment (for transfer techniques see Potrykus and Spangenberg, 1995). Indeed, by using extrachromosomal DNA, many basic features of HR and NHEJ can be characterized in plant cells. For NHEJ the production of filler DNAs has been demonstrated (Gorbunova and Levy, 1997) and for HR the occurrence of the non-conservative single-strand annealing mechanism has been proven (Puchta and Hohn, 1991; De Groot *et al.*, 1992; Bilang *et al.*, 1992). Although studies on extrachromosomal recombination have been productive in the determination of specific repair pathways it has to be kept in mind that the respective reactions occur early after transformation (Puchta *et al.*, 1992), when the non-replicating plasmid DNA, in contrast to chromosomal DNA, is most probably not associated with chromatin. Moreover, hundreds or even thousands of DNA molecules enter a single cell simultaneously. Thus, the repair enzyme machinery might be outnumbered by the number of breaks that have to be processed. As a consequence, the repair process itself might differ, at least for specific pathways, from a reaction in which only one or a few DSBs are being processed in a single cell (for a detailed discussion see also Puchta and Meyer, 1994). Most of the more recent studies that are discussed in this review concentrate on the induction of genomic DSBs *in vivo*.

A classical way to induce DSBs is by the application of X-rays to the organism under study. Indeed irradiation has been used to induce mutations in plants for decades (e.g. in *Arabidopsis*: Redei and Li, 1969; Shirley *et al.*, 1992). Using transgenic tobacco plants that carry intrachromosomal recombination substrates, it was demonstrated that the application of X-rays strongly enhances homologous recombination frequencies of the marker (Tovar and Lichtenstein, 1992; Lebel *et al.*, 1993), a phenomenon that has been explained by the induction of DSBs into the marker DNA. However, one disadvantage of the use of X-rays is that, beside DNA breaks, other kinds of damage may be induced in the cell. In addition, although the number of breaks can be influenced by the application of different doses, it is not possible to influence where in the genome a break is induced. So in only some of the irradiated cells may a break occur at a specific locus. Thus, for a defined site the frequencies of recombination obtained by X-rays are many fold lower than with a site-specific break introduced uniquely at a precise location. For the induction of a site-specific DSB, transposable elements and rare cutting restriction endonucleases have been used in plants. DNA transposons are excised by the transposase out of the donor locus before integration somewhere else in the genome takes place. During this process the transient DSB is mainly repaired by NHEJ (Rinehart *et al.*, 1997). If homologous sequences are close to a transposon a strong increase in HR has been detected (Athma and Peterson, 1991; Lowe *et al.*, 1992; Xiao and Peterson, 2000). When a restriction site of a rare cutting nuclease like HO or I-*Sce*I was included within the marker gene rather than a transposon, a similar enhancement of homologous recombination could be observed after expression of the enzyme, indicating that indeed the DSB is the inducing principal (Chiurazzi *et al.*, 1996; Siebert and Puchta, 2002; Orel *et al.*, 2003). By regulation of expression of the respective transposase (in the case of a non-autonomous element), or endonuclease, the time point of break formation can be controlled. Expression can be achieved by transient transformation (Puchta, 1999*a*) or by the control of a constitutive, organ-specific or inducible promoter after stable integration of the respective open reading frame into the plant genome (Schmidt-Puchta *et al.*, 2004).

DSB repair by homologous recombination

The majority of studies on DSB repair in plants concentrate on homologous recombination. In part, this is because there is a general interest in HR for its application to genome manipulation ('gene targeting'), but also the consequence of DSBs having been demonstrated to be intermediates in meiotic recombination in yeast. In meiosis, DSBs are induced by the Spo11 protein, which is thought to originate from the subunit A of the archaebacterial topoisomerase VI. These breaks are repaired exclusively by HR using the sequence of the allelic chromosome as a matrix (for review see Keeney, 2001). As the focus of this is on DSB repair in somatic cells, readers are advised to consult recent reviews on meiotic recombination in plants for a deeper insight into this rapidly evolving field (Caryl *et al.*, 2002; Jones *et al.*, 2003). In somatic plant cells HR, in general, is a minor pathway, the efficiency of which is dramatically influenced by the availability of a homologous matrix. This not only

influences the frequency of repair via HR versus NHEJ, but also may depend on the site of the homologous sequence in the genome as different kinds of HR pathways might be used. In principle, homologous information for repair can be copied from elsewhere in the genome ('ectopic'), from the homologue as in meiosis ('allelic') or from the same chromosome ('intrachromosomal') (Fig. 1). Moreover, after replication the information could also be taken from the sister chromatid.

Ectopic recombination

In two independent studies, one based on DSB induction by the transposase of *Ac*, and the other by the endonuclease I-*Sce*I, the question of how often an ectopic sequence can be used as the matrix for the repair of a DSB in plants was analysed (Shalev and Levy, 1997; Puchta, 1999*b*). In both systems the transgenic donor locus and an unlinked acceptor locus at which the DSB was induced carried parts of non-functional marker genes, namely a β-glucuronidase and a kanamycin resistance gene, respectively. Restoration of the marker was only possible via copying of information from the ectopic donor to the acceptor locus. Indeed, in both systems, restoration of the marker by homologous DSB repair could be detected. It was estimated that around one out of 10 000 breaks is repaired by HR. The detailed molecular analysis of several of the events indicated that the repair was not necessarily linked to homologous recombination of both ends of the break. DSB repair by a combination of HR and NHEJ was also found (Puchta, 1999*b*).

This result was reminiscent of studies in which an I-*Sce*I-induced genomic break was repaired by the use of homologous sequences on an incoming T-DNA from *Agrobacterium tumefaciens* (Puchta *et al.*, 1996). Integration of the T-DNA into the acceptor locus was mediated either by HR alone or by a combination of HR and NHEJ. As the same acceptor locus and the same donor sequences were used in both studies (T-DNA: Puchta *et al.*, 1996; ectopic DNA: Puchta, 1999*b*), a comparison is possible and indicates that T-DNA is a better template for homologous DSB repair than an ectopic sequence. This is most probably due to the fact that the incoming T-DNA is more easily accessible for the copying processes than a chromosomal site. The original single-stranded T-DNA seems to be transformed to a double-stranded molecule in the nucleus (Tinland *et al.*, 1994; Tzfira *et al.*, 2003) and thus might be activated by the presence of factors involved in DNA repair or replication. Moreover, in contrast to a chromatin-packed ectopic locus, there might be fewer constraints of a steric nature for an interaction between the acceptor sequences and a 'free' T-DNA.

The fact that a break can be repaired by a combination of HR and NHEJ clearly demonstrates that, in somatic cells, the classical double-strand break repair (DSBR) model of recombination (Szostak *et al.*, 1983) is not appropriate, as it predicts that both ends of the break would be repaired by HR in all cases. By contrast, other models such as the one-sided invasion (OSI) model (Belmaaza and Chartrand, 1994) or the synthesis-dependent strand-annealing (SDSA) model of recombination (Formosa and Alberts, 1986; Nassif *et al.*, 1994; Keeler and Gloor, 1997) could explore the combination of HR and NHEJ on the repair of a DSB. However, whereas the OSI model could account for only the combination of HR and NHEJ, the SDSA model has the capacity to explain both kinds of events, repair by HR alone as well as by HR and NHEJ. To discriminate further between the models, experiments were performed with a T-DNA that carried homology to only one end of the break. The frequencies were compared with experiments in which a T-DNA was used harbouring homology to both ends of the break in the target locus (Fig. 2). According to the DSBR model the reaction should only proceed when homology to both ends of a break is present. The recombination frequencies obtained with the repair construct harbouring homology at both ends were only one-third higher than with the one-ended construct (Puchta, 1998). Even this small difference could be explained by one-sided invasion of the unique second homologous end of the respective T-DNA. Hence, homology to only one end of the DSB is sufficient for an efficient homologous recombination reaction and DSB repair reactions in plant cells are initiated by one-sided initiation events. Using Ockham's razor (postulating the least number of entities required for
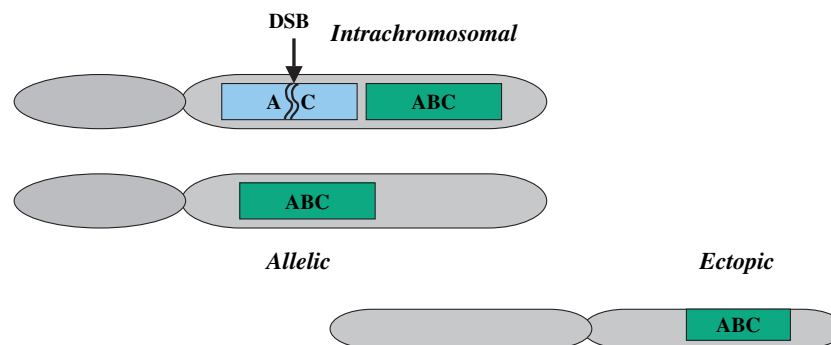

Fig. 1. Homologous sequences at different positions in the genome can be used as matrix for the repair of a DSB.
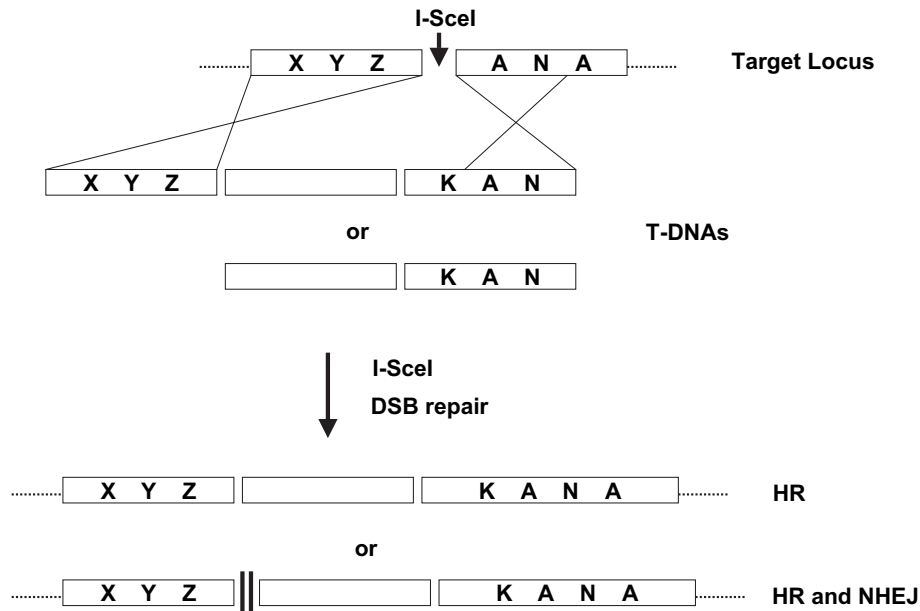
Fig. 2. Schematic diagram of the experimental set-up that was used to demonstrate that a DSB could be repaired by a combination of homologous recombination (HR) and non-homologous end-joining (NHEJ). By expression of I-*Sce*I a DSB is induced in the target locus. The break can be repaired by restoration of the kanamycin gene (KANA) by homologous recombination with an incoming T-DNA. Two different T-DNAs were used: one with homologies to both ends of the break and the other with homology to only one end of the break. Both kinds of constructs could be targeted to the genomic locus. As only the SDSA model of recombination is able to predict both kinds of events, this mechanism may be the most appropriate to account for gene conversion events in plants.

the explanation of a phenomenon) the best suited model for repair of genomic DSB in somatic plant cells is the SDSA model (Fig. 3) as it is based on one-sided initiation and it is able to describe recombination events due to HR as well as due to a combination of HR and NHEJ events. The application of the SDSA model seems to be even broader (see below): during repair of breaks by NHEJ, in some cases sequences from elsewhere in the genome are copied into the break (Rubin and Levy, 1997; Salomon and Puchta, 1998; Gorbunova and Levy, 1999). The respective junctions between broken DNA and inserted sequence are characterized by some base pairs (bp) of common homology, indicating a template switch, which is most easily explained by a SDSA mechanism (Fig. 4).

Using the SDSA model as the central paradigm of somatic DSB repair, two other important predictions can be made: somatic DSB repair events between ectopic sequences are not accompanied by crossovers and, in gene targeting experiments, not only perfect integration events at the target locus by homologous recombination will be found. Indeed, since the early days of gene targeting in plants, recombinants were characterized in which only one end of the targeting vector integrated by homology (Risseeuw *et al.*, 1995) or after copying sequences from the transgene locus the vector can be found integrated elsewhere in the genome ('ectopic targeting', Offringa *et al.*, 1993). Such cases were found repeatedly and represent a major class of targeting events (Reiss *et al.*, 2000; Hanin *et al.*, 2001; Hohn and Puchta, 2003). Recently, it was demonstrated that ectopic recombin-

ation could be overcome by the integration of counter selection markers into the targeting vector outside the homologous regions. Using an efficient selection procedure an enrichment of 'perfect' targeted recombinants has been achieved for rice (Terada *et al.*, 2002).

The SDSA model describes double-strand break repair without formation of double Holliday junctions during the repair process. Consequently, this reaction will lead either to the non-homologous repair of one end of the break or in the case of perfect homologous recombination to the exclusive occurrence of gene conversion events not accompanied by crossovers (Fig. 3). Thus, due to their biological role, pathways of gene conversion at least partly differ in somatic and meiotic cells in higher eukaryotes. Whereas in meiosis double Holliday junctions occur (Schwacha and Kleckner, 1996) and crossovers are a prerequisite for alterations in genetic coupling arrangements, in somatic cells the prevalence of gene conversion events avoids unpredictable deleterious genomic change by crossovers between ectopic sequences (Richardson *et al.*, 1998). This might be particularly important for preserving the stability of the plant genome, as many species contain ample amounts of repetitive DNA and somatic change can be transferred to the germ line (Walbot, 1996).

Allelic recombination

During meiosis, DSBs are induced and repaired by the use of sequences from the homologue, resulting in either crossover or gene conversion. It was therefore important
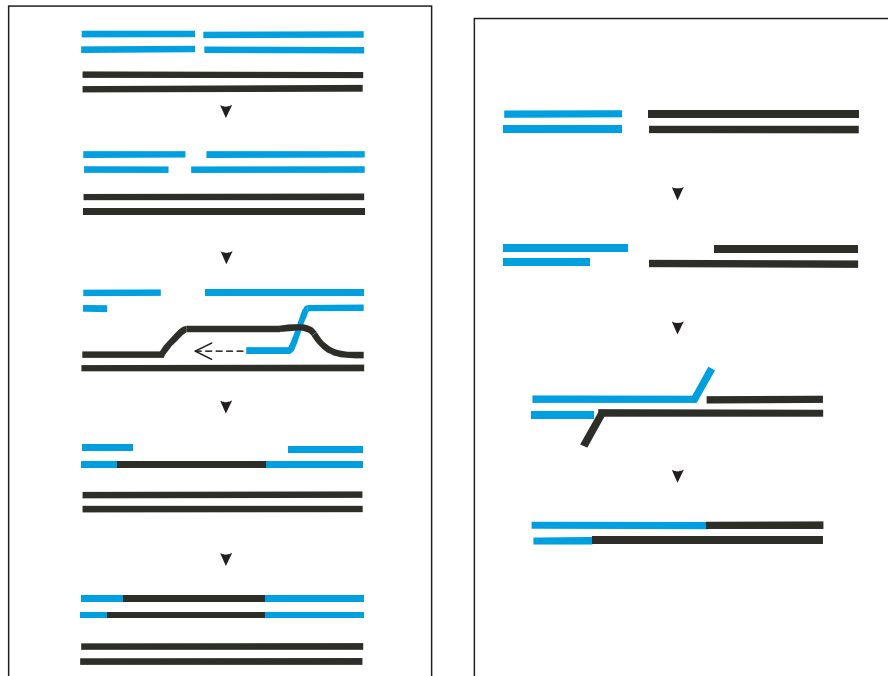
Fig. 3. Models describing the pathways of DSBs repair in somatic plant cells. On the left the synthesis-dependent strand annealing-mechanism (SDSA) and on the right the single-strand annealing (SSA) mechanism is depicted. SDSA model: After induction of a double-strand break in the acceptor molecule a 3′-single-strand overhang is produced via exonuclease-catalysed digestion. By forming a D-loop a free 3′ end invades the double-stranded donor and repair synthesis occurs. For the further processing of the intermediate two possible outcomes can be envisaged: if the 3′ of the acceptor molecule is elongated up to the homology of the second 3′ end of the DSB, the single strands anneal and the molecule is repaired. If, on the other hand, the 3′-end of the elongated acceptor molecules is not complementary to the 3′ end of the DSB, the break is closed via NHEJ, which might be initiated by the annealing of a few nucleotides common to the newly synthesized strand and the other broken end. SSA: according to this mechanism a double-strand break is induced between repeated sequences. After exonuclease-catalysed digestion 3′, single-strand overhangs are produced. If the overhangs carry complementary sequences they can anneal and a chimeric DNA molecule is formed. If the chimeric molecule has overhanging ends, as shown, they are resected. Otherwise putative single-stranded regions are filled in by DNA synthesis. Thus all information between the repeated sequences is lost in the reaction.
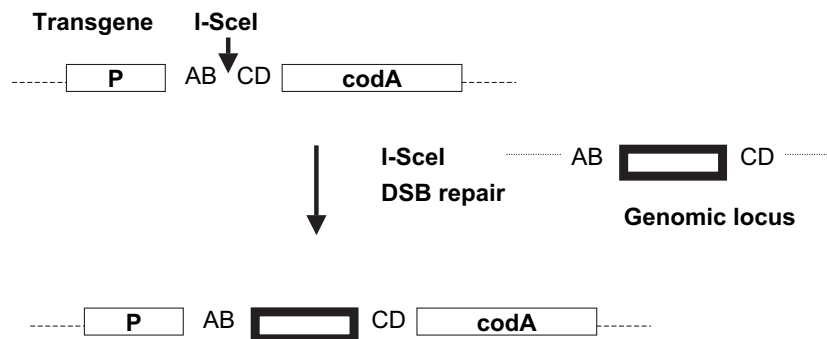

Fig. 4. Insertion of genomic sequences during DSB repair: Schematic diagram of an assay system used to detect insertions. A DSB is induced between promoter (P) and open reading frame of a negative selectable marker (*cod*A, cytosine deaminase) by I-*Sce*I expression. In some cases in which the marker became non-functional the repair was correlated with insertion of DNA from elsewhere in the genome. Interestingly short stretches of common nucleotides (AB, CD) were found at the junction between I-*Sce*I site and genomic sequence, indicating that the DNA was inserted by a copying process as described by the SDSA model (Fig. 3).

to test whether such a pathway is also operating in somatic cells. As gene targeting is not feasible for flowering plants it is, in contrast to mammalian (Moynahan and Jasin, 1997) and yeast cells (Hiraoka *et al.*, 2000), not possible to manipulate directly allelic positions in plant genomes. To overcome this limitation a strategy was developed in which the transgene locus carrying the negative selectable marker cytosine deaminase (*cod*A) with an I-*Sce*I originally produced for the analysis of NHEJ (Salomon and Puchta, 1998; see below) was rendered non-functional by the induction of a DSB repair reaction. Plants homozygous for the modified sequence were then crossed with plants

homozygous for the original transgene. Thus, plant lines could be obtained with two different alleles at the transgene locus on homologous chromosomes. One allele contained an I-*Sce*I site within the *cod*A gene and the other either a deletion or an insertion within the region of the *cod*A expression cassette that rendered both the I-*Sce*I site and the *cod*A gene non-functional. This approach was similar to the one of Dooner and Martinez-Ferrez (1997) in which naturally occurring alleles of the *bz* gene of maize were combined. One of the alleles carried an active transposable element. However, the germinal excision of the element did not stimulate meiotic recombination, a fact that might be due to the programmed induction of DSBs at a specific time point during meiosis. After DSB induction by I-*Sce*I expression selection was performed in heteroallelic tobacco plants for cells resistant to 5-fluorocytosine (5-FC) due to a non-functional *cod*A gene. The occurrence of DSB-induced allelic recombination could be demonstrated (Gisler *et al.*, 2002). However, the vast majority of events resulted in a loss of function of the marker gene due to NHEJ. The frequency of allelic recombination was estimated to be about 10^{-4}, similar to that of ectopic recombination obtained after DSB repair in tobacco by two different studies (see above; Shalev and Levy, 1997; Puchta, 1999*b*). Allelic recombination is induced by DSB at a specific site by at least three orders of magnitude. Thus, compared with meiosis, allelic HR is not a significant DSB repair pathway in somatic cells. As in these cells DSB repair occurs mainly by NHEJ, a process which might lead to major genomic alterations (see below), the inefficiency of allelic HR seems to be a prerequisite for the fixation of major changes in the genome. As somatic changes can be transferred to the germ line, this might be an important factor responsible for the enhanced plasticity postulated for plant genomes.

Intrachromosomal recombination

By contrast with the use of allelic and ectopic positions, sequences close to the break on the same chromosome or on the sister chromatid are more efficient matrixes for repair. Sister chromatids can be used as a template in the G_2 stage of the cell cycle, and this pathway works efficiently in mammals (Johnson and Jasin, 2000), however, due to experimental limitations it has not been possible to address this question in plant cells in connection with DSB induction until now (Molinier *et al.*, 2004). By contrast, a number of studies have been published on intrachromosomal DSB repair in plants (Athma and Peterson, 1991; Chiruazzi *et al.*, 1996; Orel *et al.*, 2003; Siebert and Puchta, 2002; Xiao and Peterson, 2000; Xiao *et al.*, 2000) and a more detailed picture of the frequencies of repair as well as on the pathways has been determined. Depending on the structure of the respective chromosomal locus at least two different kinds of pathways may be used to repair the break. Whereas one pathway results in the formation of a deletion

between direct repeats, the product of the other is a gene conversion event (Fishman-Lobell *et al.*, 1992). The deletion pathway is described by the single-strand annealing (SSA) model, whereas gene conversion can be explained by the synthesis-dependent strand-annealing (SDSA) model. The SSA model (Lin *et al.*, 1984, 1990) describes a non-conservative reaction, resulting in a loss of information that is positioned between the annealing repeats (Fig. 3). SSA-like models have also been postulated to explain the loss of information during NHEJ (Nicolas *et al.*, 1995). As discussed previously the SDSA model is conservative, it describes transfer of information from a homologous donor sequence to the break site without loss of sequences at the donor locus (Fig. 3). The efficiency of the pathways can be addressed by using a different configuration of marker genes. According to the SDSA model the orientation of the donor sequence to the break has no influence on gene conversion and the break should lie within the homologous region. By contrast, constructs measuring the efficiency of the SSA mechanism should have a break site within two direct repeats. Using the respective markers in a recent study (Orel *et al.*, 2003) the frequencies of both pathways could be directly compared, indicating that the SSA pathway seems to be about five times more efficient than the gene conversion pathway.

However, of equal importance is to obtain an estimate of the number of breaks that are repaired in this situation by HR compared with NHEJ. In an attempt to determine this an assay was set up in tobacco based on interrupted overlapping halves of a β-glucuronidase gene. Between the overlaps a negative selectable marker gene flanked by two I-*Sce*I sites was inserted into the transgene (Fig. 5). As a consequence of I-*Sce*I expression, after the excision of a negative selectable marker gene, a single genomic DSB had to be repaired. As the marker gene is lost, all resulting repair events (independent of their nature) could be isolated. It was demonstrated that the DSB was repaired in about one in three cases by HR. Thus, SSA is a highly efficient mechanism for DSB repair. SSA-like models can also be used to describe a class of NHEJ events (see below). Thus, SSA and SSA-like mechanisms might be the most prominent mode for the rejoining of broken DNA molecules in higher eukaryotes. If there are homologous sequences available in close proximity to the break they are regularly used for the annealing reaction, if they are not available, short patches of homology, which are present abundantly, might be used.

The demonstration that, in plants, DSBs can be repaired efficiently by the use of nearby homologies seems also to be relevant for plant genome evolution: In the retroelement *BARE*-1 only a 7–42-fold excess of single long terminal repeats (LTRs) over its own internal regions has been found (Vicient *et al.*, 1999; Kalendar *et al.*, 2000; Shirasu *et al.*, 2000). It seems surprising that such a high incidence of single LTRs was found, as this situation could only have

arisen by a homology-dependent deletion of the internal region of the retrotransposon. *BARE*-1 is active in barley (Kalendar *et al.*, 2000; Vicient *et al.*, 2001) and during their spread retroelements integrate regularly into one another (San Miguel *et al.*, 1996). If one assumes that transient breaks within the internal part of *BARE*-1 are induced either due to the activity of transposons (Xiao *et al.*, 2000) or due to aberrant integration events of retroelements, a DSB-induced HR between the two flanking LTRs may yield solo LTRs (Fig. 6). Thus, it is tempting to speculate that accumulation of multiple solo LTRs in plant genomes is linked to the presence of active mobile genetic elements that are able to induce breaks during their life cycle.

If all experimental data are taken into account, a detailed picture can be obtained of how DSBs are repaired in somatic plant cells. The most efficient way to repair a break is by NHEJ (see below). If homologous sequences are available close to the break, the repair can take place in up to one-third of the cases by a single-strand annealing pathway. In about one out of 15 breaks, repair could also proceed via the SDSA pathway. This is two to three orders of magnitude more efficient than a gene conversion event that uses homology from an allelic or an ectopic site. DSB-induced gene conversion between members of tandemly arranged gene families might well play an important role in plant genome evolution. Various resistance genes are organized in tandem arrays in plant genomes and gene conversion events between the various members have been

Fig. 5. Schematic diagram of an assay system used to compare the frequencies of HR and NHEJ for DSB repair in case homologous sequences are available in close proximity to the break. Between two I-*Sce*I sites, a negative selectable marker gene is positioned. After expression of the endonuclease the gene is cut out of the genome and the remaining single break can be repaired either by homologous recombination (HR) leading to the restoration of a functional marker gene (GUS, β-glucuronidase) or non-homologous end-joining (NHEJ). In about one-third of the cases the break was repaired by HR. This set-up can also be used for the excision of transgene sequences (e.g. selection markers) out of the plant genome.

demonstrated by sequence analysis (Parniske *et al.*, 1997; Parniske and Jones, 1999). Such events might not necessarily be due to meiotic recombination but may, at least partially, also occur during vegetative growth. Indeed, pathogen attack is inducing intrachromosomal recombination between tandemly repeated sequences (Lucht *et al.*, 2002). Moreover, certain sites in plant genomes might be less stable for repeated sequences than others. This is documented by findings that certain transgene sequences are deleted from the genome at high rates (Zubko *et al.*, 2000), a phenomenon that has been postulated to be due to transient DSB-induction (Puchta, 2000). It seems possible that various gene conversion events, detected in tandemly arranged resistance gene clusters, might be correlated with higher rates of transient breaks in these regions of the plant genome (Ramakrishna *et al.*, 2002).

DSB repair by non-homologous end joining

In higher eukaryotes, NHEJ seems to be the main mode of DSB repair (Sargent *et al.*, 1997). NHEJ might lead, in a certain fraction of cases, to genomic changes such as deletions, insertions or various kinds of genomic rearrangements (Pipiras *et al.*, 1998). As genomic alterations in meristematic cells can be transferred to the offspring, it is especially important to understand this pathway in plants. Different classes of NHEJ repair events were characterized in eukaryotes: on one side the repair of the break was accompanied by incorporation of filler DNAs, or the break ends were rejoined with or without deletions. The resulting junctions of the latter pathway contained either no or several nucleotides of sequence identity between the break ends. Junctions without homologies were explained by simple ligation, whereas small patches of homologous nucleotides were considered to be a prerequisite for the operation of a single-strand annealing-like mechanism (Nicolas *et al.*, 1995; Mason *et al.*, 1996; Lehman *et al.*, 1994). In plants, all different kinds of reaction products have been detected with transfected plasmid molecules as well as with an induced DSB in the genome (Gorbunova and Levy, 1997; Salomon and Puchta, 1998). The repair of a genomic DSB was studied using tobacco harbouring a single I-*Sce*I site between promoter and open reading frame of the cytosine deaminase gene (Fig. 4). After the selection of 5-FC resistant clones besides deletions (see below) a broad spectrum of unique and abundant genomic sequences was found inserted into chromosomal breaks of tobacco (Salomon and Puchta, 1998), a phenomenon drastically different from yeast, where only cDNA of retrotransposons or mitochondrial DNA was found to be inserted into breaks (Moore and Haber, 1996; Teng *et al.*, 1996; Ricchetti *et al.*, 1999). The presence of stretches of several nucleotides homologous between break site and copied sequences strongly argues for the occurrence of a synthesis-dependent strand annealing mediated copying of ectopic chromosomal DNA into the break. In most cases
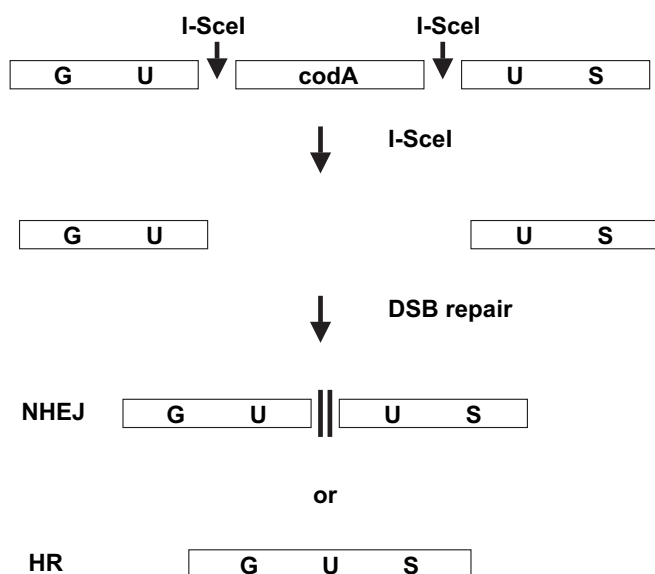
apparently only one 3′ end was actively invading a genomic donor sequence (Salomon and Puchta, 1998), however, there were also indications for a combination of sequences of different origin due to independent invasion events of both 3′ ends of the break (Gorbunova and Levy, 1997). Older models explain the amplification of genomic sequences by unequal homologous recombination between repetitive sequences in tandem or by replicative transposition of retroelements. However, these models do not explain how transposition-inactive sequences can be transferred into new positions within the plant genome. Using DSB repair, nearly any kind of genomic sequence that is available for a copying process can be inserted into new genomic positions. This phenomenon might be of general importance for higher plants and their genome evolution. DSB repair is also a prominent source of deletions. The explanation for the large differences in the nuclear DNA content of eukaryotes, particularly plants, known as the 'C value paradox', has been a matter of debate for some time (Cavalier-Smith, 1985; Dove and Flavell, 1988). Even closely related species with a similar phenotype may differ significantly in their diploid genome size. One mechanism responsible for these size differences could be related to a species-specific increase/reduction of repetitive sequences. In principle, genomes may become larger via duplications and insertions or smaller via deletions. Species-specific spread of retrotransposons was postulated as a main route for the enlargement of plant genomes (SanMiguel *et al.*, 1996, 1998; Bennetzen and Kellog, 1997). Alternatively, deletions might reduce genome size and counterbalance enlargements (Petrov *et al.*, 1996, 2000; Petrov, 2001). A comparison of deletion formation in somatic cells of tobacco and *Arabidopsis*, two dicotyledonous plant species with an over 20-fold difference in genome size (Bennett and Leitch, 1997), was performed using the same construct as described above (Kirik *et al.*, 2000). Indeed, surprisingly large differences in DSB repair were found: whereas the mechanism of junction formation itself was not different between the two species. In tobacco, almost every second deletion event associated with the loss of function of the marker was accompanied by the insertion of filler sequences. By contrast, no insertions were isolated from *Arabidopsis*. The overall length of the deletions was about one-third shorter in tobacco than in *Arabidopsis*. Thus, as suggested theoretically before for insects (Petrov *et al.*, 2000), an inverse correlation between genome size and the medium length of deletions could be detected. Although just two species were compared, and a final conclusion on the matter is, of course, not yet justified, it is nevertheless tempting to speculate that species-specific differences in DSB repair pathways could significantly contribute to the evolution of eukaryotic genome size. As *Arabidopsis* has undergone a number of segmental duplications or possibly a complete genome duplication during evolution (Arabidopsis Genome Initiative, 2000) deletion formation must have played a prominent role

resulting in the small size of the present day genome of this plant species. Indeed, evidence that *Arabidopsis* has lost DNA over evolutionary time periods is supported by a recently completed elegant computer study based on the analysis of deletions within LTRs of retroelements (Devos *et al.*, 2002).

But what are the molecular causes for the differences in DSB repair between *Arabidopsis* and tobacco? During DSB repair the size of a deletion depends on the processing of DNA ends. If broken ends are not religated directly the processing of such ends might result in the loss of DNA at the break site. Depending on the efficiency of DNA degradation more or less information will be lost. To address this question, particle bombardment was used to test the stability of linearized plasmid DNA in tobacco and *Arabidopsis* cells (Orel and Puchta, 2003). Indeed indications were found that plasmid DNA is degraded to a lesser extent in tobacco than *Arabidopsis*. If a break was induced between promoter and open reading frame of a marker gene the expression level has reduced by a factor of two in *Arabidopsis* compared with tobacco. This observation was in contrast to the similar expression levels obtained for both species when the break was introduced a long distance from the marker gene. Whereas in the first case exonucleolytic digestion of few nucleotides would result in the destruction of the expression cassette, in the latter case the removal of 2 kb would be a prerequisite for the destruction of the ORF. Because endonucleolytic degradation would destroy all the information on a plasmid at the same rate, independent of a pre-existing break site, no differences should be found in the presence of such a nuclease activity. However, as the restoration of the marker ORF required a ligation step, the data of the expression assay alone do not suffice to discriminate between a more efficient exonucleolytic destruction of the marker gene or a less efficient ligation reaction in *Arabidopsis* cells. As a Southern blot analysis revealed differences in the degradation of linearized plasmid DNA between *Arabidopsis* and tobacco, it seems that most of the effect is due to a more efficient exonucleolytic degradation of DNA in *Arabidopsis*. However one question remains: What is the reason for the enhanced DNA degradation in *Arabidopsis*? This could be either due to less active DNA exonucleases or due to the better protection of DNA ends from degradation, for example, by DNA end binding factors (Liang *et al.*, 1996).

Biotechnological applications: marker gene excision and site-specific integration

After the elucidation of the main mechanisms of DSB repair in plants it is of course an obvious step to use the induction of breaks at specific sites for the controlled manipulation of plant genomes. Two major applications have been developed: The excision of sequences from the genome and the integration of transgenes into specific sites of the genome.

Marker gene excision

Site-specific recombinases have been used efficiently for the excision of selectable marker genes (for a review see Hohn *et al.*, 2001; Hare and Chua, 2002; Puchta 2003*a*, *b*). However, using this approach, recognition sites of the recombinases are left behind in the genome. Thus, for the performance of multiple genomic changes, a combination of different site-specific recombination systems has to be applied. The development of alternative approaches for site-specific alterations of genomes is of great interest in biotechnology. A very promising approach would be the combined elimination of the transgene sequences and the respective recognition sequence. Therefore, the use of highly specific restriction endonucleases to remove transgenic sequences from the plant genome is a useful alternative to the established site-specific techniques. The removal of the recognition site of the enzyme makes the process irreversible. It has been demonstrated that a marker gene can be efficiently eliminated from the plant genome either via HR or via NHEJ (including religation) by inducing DSBs by a rare cutting restriction enzyme (Fig. 5; Siebert and Puchta, 2002). Thus, any sequence flanked by restriction sites could, in principle, be eliminated from the plant genome. *In praxis*, the open reading frame of the endonuclease, under an inducible promoter, could be included in the 'I-*Sce*I cassette' of the transgene to eliminate all superfluous sequence from the genome in one step (Puchta, 2003*b*). The use of several other rare cutting endonucleases might also be useful for this purpose.

DSB-induced site-specific integration

Using co-transformation it could be demonstrated that different T-DNAs integrate into a single genomic locus (De Block and Debrouwer, 1991; De Neve *et al.*, 1997). This indicates that a surplus of extrachromosomal DNA ends is recognized by the cell as breaks to be ligated, yielding a tandem arrangement of transgenes, which can then be linked to free genomic ends. It seems that the number of free genomic ends might be rate-limiting for the integration step of transgenes and indeed X-rays, which induce DSBs in the genome, can enhance transgene integration in plant cells (Köhler *et al.*, 1989). T-DNA integration into a DSB was demonstrated directly in a study on NHEJ using the *cod*A gene harbouring an I-*Sce*I site (see above; Salomon and Puchta, 1998). In three out of four cases almost the entire sequence of the 2.4 kb long T-DNA was integrated. More detailed analysis on the frequencies of such events was performed recently using different site-specific recombinases. In approximately 1% of the transformants the T-DNA was found inserted into the break (Chilton and Que, 2003; Tzfira *et al.*, 2003). Where I-*Sce*I sites flanked the transformation marker on the T-DNA, integration of the DNA into the genomic I-*Sce*I site by simple ligation could be demonstrated. This indicates that the T-DNA became double-stranded and was then cut, prior to integration, by the enzyme. These studies clearly demonstrate that DNA can integrate during DSB repair into the plant genome. Nevertheless, taking the frequencies, as well as the integration patterns, into account, it seems that DSB
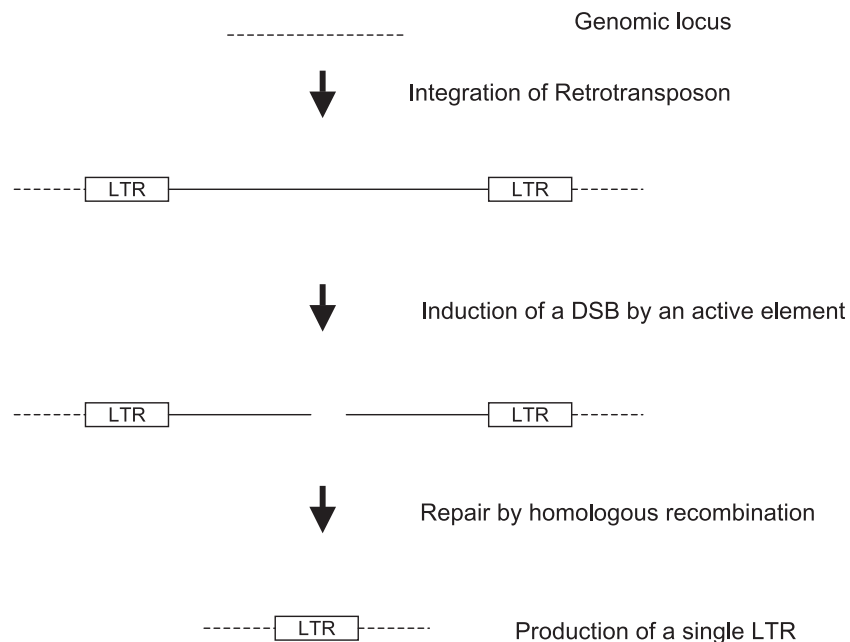

Fig. 6. An hypothesis proposed to explain the accumulation of single long terminal repeats (LTRs) in cereal genomes. Active retroelements spread in the cereal genome to new genomic locations, resulting in an integrated DNA element that is flanked by two identical LTRs. Due to further activities of DNA transposons or retrotransposons a DSB might be induced into the central part of the element by an aberrant integration reaction. In such a situation in at least one-third of the cases the break will be repaired by homologous recombination leaving behind a single LTR in the cereal genome.

repair is not the only and not even the main mechanism of T-DNA integration. Various studies on T-DNA integration indicate that there are, in principle, two different classes of integration patterns (Fig. 7; Brunaud *et al.*, 2002). In most cases, the right border is conserved in the recombinants and contains no or very little microhomology to the target locus whereas longer homologies exist between the left border (combined with truncations) and the preinsertion site. In these events the two ends behave differently, and it has been postulated that integration occurs via single-strand annealing followed by the right border being linked to genomic sequences directly or indirectly by the VirD2 protein that is covalently bound to the 5′ end of the T-strand (Tinland, 1996). In a second class of events, not only the left but also the right border is partly truncated in the recombinants, and common nucleotides can be found in most cases at the junction site also between the right border and the pre-insertion. This pattern of integration might be indicative for DSB repair. T-DNA molecules are transferred as single strands to the nucleus, where some of them become double-stranded before integration (Tinland *et al.*, 1994; Tzfira *et al.*, 2003). By interaction with break ends, a single- or double-stranded T-DNA could be copied into the genome via synthesis-dependent strand annealing. Alternatively, a double-stranded T-DNA could also be integrated at the break by two consecutive single-strand annealing reactions.

Homologous integration into a genomic DSB was also achieved by the use of I-*Sce*I. In up to 1% of the transformation events the target vector integrated into the homologous target locus (Puchta *et al.*, 1996). Besides events in which both ends of the target vector were integrated by HR, combinations of HR and NHEJ were also found and, indeed, targeting with a vector with homology only to one end of the break resulted in similar frequencies (Puchta, 1998; see also Fig. 2). However, this strategy is not applicable for targeting genes at will, as such

a break can only be induced at the specific recognition sites of the rare cutting restriction enzyme. The technique can only be used for the integration of transgenes at a specific predetermined site in the genome and is, as such, an alternative to site-specific integration mediated by sequence specific recombinases (for a review see Ow, 2002). However, an alternative would be to construct restriction endonucleases that recognize specific sequences in the gene of interest. Indeed, chimeric endonucleases (Kim *et al.*, 1996; Chevalier *et al.*, 2002; Epinat *et al.*, 2003) and zinc finger nucleases (for a review see Carroll, 2004) are being developed to fulfil this task. The application of zinc finger nucleases for gene targeting in vertebrates has already been demonstrated (Bibikova *et al.*, 2003; Porteus and Baltimore, 2003) and it may be only a matter of time before the first reports of the use of the technique will appear for plants.

Another gene targeting approach using DSB induction for the production of a recombination intermediate has been developed recently for *Drosophila melanogaster* by the group of Kent Golic (Rong and Golic, 2000; Rong *et al.*, 2002). In this method the construct for targeting is integrated in the host genome flanked by two recognitions sites of a site-specific recombinase and includes a site for a rare cutting restriction endonuclease. By induced expression of the site-specific recombinase a DNA circle is excised from the genome. This circle is then linearized after the restriction enzyme (in this case I-*Sce*I) has been expressed resulting in an 'activated' DNA molecule with both ends homologous to the target sequence (for a simplified scheme see Fig. 8). In the female germline of *Drosophila*, gene targeting occurred in about one out of 500 cases. As gene targeting is a technique urgently needed in plant biotechnology there is considerable interest in determining whether the strategy is also applicable to plants (for discussion see also Kumar and Fladung, 2001). On the one hand the set-up of the system seems to be very complex, as besides construction of a donor sequence
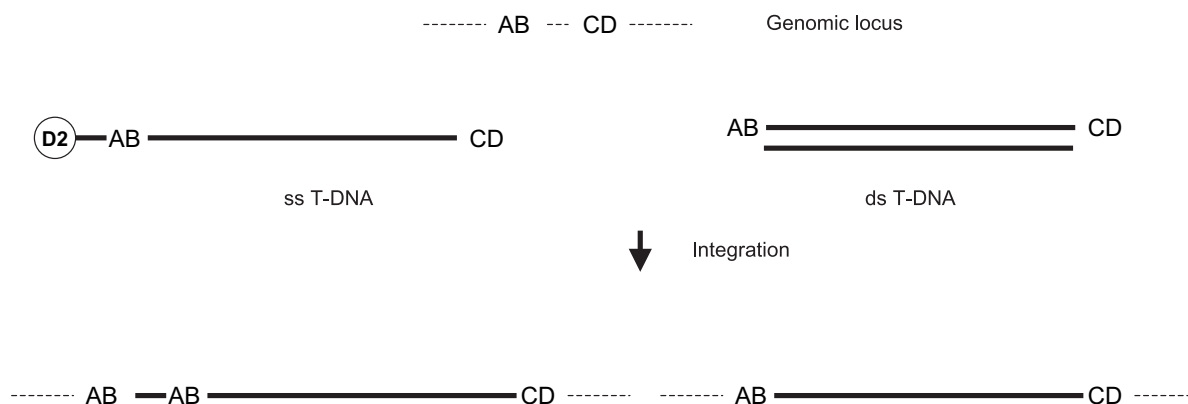

Fig. 7. A simplified model to describe different ways of T-DNA integration into the plant genome. In general, two classes of integrated T-DNA molecules have been found that can be classified by the nature of the T-DNA–plant DNA junctions. In one class (left) the 3′ end of the T-DNA and the genomic locus are joined by a few homologous nucleotides (CD), whereas in another minor class small homologies are found at both junctions between the integrated T-DNA and the genomic locus. Whereas in the first case it can be speculated that a single-stranded T-DNA intermediate with a 5′ end protected by the VirD2 protein was involved in integration, in the case of the second class the pattern of the repair event is identical to patterns found during integration of double-stranded DNA molecules (ss: single-stranded, ds: double-stranded).
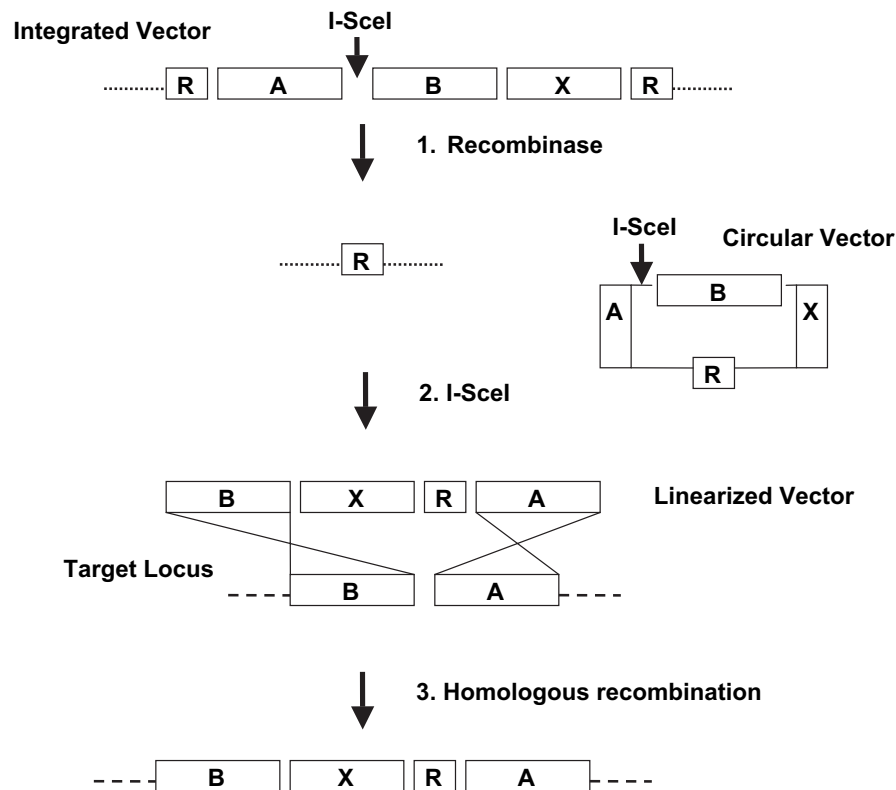
Fig. 8. Simplified scheme for a targeting approach as performed in *Drosophila*. A targeting vector containing sites for a recombinase (R) and I-*Sce*I is integrated in the genome. After expression of the recombinase the vector is excised from the genome as a circular DNA molecule and, in a next step, linearized by I-*Sce*I. The linearized vector is then able to integrate at the target locus via homologous recombination (X: gene of interest to be integrated at the target locus).

with sites for a recombinase and a restriction enzyme, expression cassettes of both enzymes have to be included into the transgene construct or supplied *in trans*. On the other hand, if the reaction occurs in an efficient way *in planta*, every single seedling should represent an excision event. Thus, by the use of respective marker genes, large numbers of plants can be produced and easily screened.

One has to see if and when the two new approaches, both based on the induction of DSBs, will help to solve the long-lasting problem of gene targeting in plants.

Outlook

Over the last ten years a substantial body of information has accumulated relating to DSB repair in plants. The characterization of different pathways involved in DSB repair and the estimation of the respective frequencies makes it now possible for us to draw a quite detailed picture of the changes that occur during DSB repair in the plant genome. As sequenced genomes of an increasing number of plant species become available and, as a consequence, the evolutionary changes can be analysed at the whole genome level, current knowledge of DSB repair pathways will help to explain why genomes evolved in size and organization, resulting in the complexity that is found today. However, although major

insights have been made, this picture of DSB repair in plant cells is far from complete, and neither questions on DSB-induced sister chromatid exchange nor on DSB-induced translocation have been addressed until now. A major challenge will be to refine techniques of genome manipulation further and this holds true particularly for the set-up of a reliable targeting technique. The induction of meiotic recombination by site-specific DSBs may be an important step along this path over the coming years.

Acknowledgements

References

Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408,** 796–815.

Athma P, Peterson TA. 1991. Ac induces homologous recombination at the maize P locus. *Genetics* **128,** 163–173.

Belmaaza A, Chartrand P. 1994. One-sided invasion events in homologous recombination at double-strand breaks. *Mutation Research* **314,** 199–208.

Bennett MD, Leitch IJ. 1997. Nuclear DNA amounts in angiosperms—583 new estimates. *Annals of Botany* **80,** 169–196.

Bennetzen JL, Kellog EA. 1997. Do plants have a one-way ticket to the genomic obesity? *The Plant Cell* **9,** 1509–1514.

Bibikova M, Beumer K, Trautman JK, Carroll D. 2003. Enhancing gene targeting with designed zinc finger nucleases. *Science* **300,** 764.

Bilang R, Peterhans A, Bogucki A, Paszkowski J. 1992. Single-stranded DNA as recombination substrate in plants assessed by stable and transient expression. *Molecular and Cellular Biology* **12,** 329–336.

Britt AB, May GD. 2003. Re-engineering plant gene targeting. *Trends in Plant Science* **8,** 90–95.

Brunaud V, Balzergue S, Dubreucq B, *et al.* 2002. T-DNA integration into the *Arabidopsis* genome depends on sequences of pre-insertion sites. *EMBO Reports* **3,** 1152–1157.

Carroll D. 2004. Using nucleases to stimulate homologous recombination. *Methods in Molecular Biology* **262,** 195–207.

Caryl AP, Jones GH, Franklin FC. 2002. Dissecting plant meiosis using *Arabidopsis thaliana* mutants. *Journal of Experimental Botany* **54,** 25–38.

Cavalier-Smith T. 1985. *The evolution of genome size*. Chichester: John Wiley and Sons.

Chevalier BS, Kortemme T, Chadsey MS, Baker D, Monnat RJ, Stoddard BL. 2002. Design, activity, and structure of a highly specific artificial endonuclease. *Molecular Cell* **10,** 895–905.

Chilton MD, Que Q. 2003. Targeted integration of T-DNA into the tobacco genome at double-stranded breaks: new insights on the mechanism of T-DNA integration. *Plant Physiology* **133,** 956–965.

Chiurazzi M, Ray A, Viret J-F, Perera R, Wang X-H, Lloyd A, Signer ER. 1996. Enhancement of somatic intrachromosomal homologous recombination in *Arabidopsis* by HO-endonuclease. *The Plant Cell* **8,** 2057–2066.

De Block M, Debrouwer D. 1991. Two T-DNA's co-transformed into *Brassica napus* by a double *Agrobacterium tumefaciens* infection are mainly integrated at the same locus. *Theoretical and Applied Genetics* **82,** 477–484.

De Groot MJA, Offringa R, Does MP, Hooykaas PJJ, van den Elzen PJM. 1992. Mechanisms of intermolecular homologous recombination in plants as studied with single- and double-stranded DNA molecules. *Nucleic Acids Research* **20,** 2785–2794.

De Neve M, De Buck S, Jacobs A, Van Montagu M, Depicker A. 1997. T-DNA integration patterns in co-transformed plant cells suggests that T-DNA repeats originate from co-integration of separate T-DNAs. *The Plant Journal* **11,** 15–29.

Devos KM, Brown JK, Bennetzen JL. 2002. Genome size reduction through illegitimate recombination counteracts genome expansion in *Arabidopsis*. *Genome Research* **12,** 1075–1079.

Dooner KH, Martinez-Ferrez IM. 1997. Germinal excision of the maize transposon *activator* do not stimulate meiotic recombination or homology-dependent repair at the *bz* locus. *Genetics* **147,** 1923–1932.

Dove GA, Flavell RB. 1988. *Genome evolution*. London: Academic Press.

Epinat JC, Arnould S, Chames P, Rochaix P, Desfontaines D, Puzin C, Patin A, Zanghellini A, Paques F, Lacroix E. 2003. A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. *Nucleic Acids Research* **31,** 2952–2962.

Fishman-Lobell J, Rudin N, Haber JE. 1992. Two alternative pathways of double-strand break repair that are kinetically separable and independently modulated. *Molecular and Cellular Biology* **12,** 1292–1303.

Formosa T, Alberts BM. 1986. DNA synthesis dependent on genetic recombination: characterization of a reaction catalyzed by purified bacteriophage T4 proteins. *Cell* **47,** 793–806.

Gisler B, Salomon S, Puchta H. 2002. The role of double-strand break-induced allelic homologous recombination in somatic plant cells. *The Plant Journal* **32,** 277–284.

Gorbunova V, Levy AA. 1997. Non-homologous DNA end-joining in plant cells is associated with deletions and filler DNA insertions. *Nucleic Acids Research* **25,** 4650–4657.

Gorbunova V, Levy AA. 1999. How plants make ends meet: DNA double-strand break repair. *Trends in Plant Science* **4,** 263–269.

Hanin M, Volrath S, Bogucki A, Briker M, Ward E, Paszkowski J. 2001. Gene targeting in *Arabidopsis*. *The Plant Journal* **28,** 671–677.

Hanin M, Paszkowski J. 2003. Plant genome modification by homologous recombination. *Current Opinion in Plant Biology* **6,** 157–162.

Hare PD, Chua NH. 2002. Excision of selectable marker genes from transgenic plants. *Nature Biotechnology* **20,** 575–586.

Hartung F, Puchta H. 2004. What comparative genomics tells us about the evolution of the eukaryotic recombination machinery. *Current Genomics* **5,** 109–121.

Hays JB. 2002. *Arabidopsis thaliana*, a versatile model system for study of eukaryotic genome-maintenance functions. *DNA Repair* **1,** 579–600.

Hiraoka M, Watanabe K, Umezu K, Maki H. 2000. Spontaneous loss of heterozygosity in diploid *Saccharomyces cerevisiae* cells. *Genetics* **156,** 1531–1548.

Hohn B, Levy A, Puchta H. 2001. Elimination of selection markers from transgenic plants. *Current Opinion in Biotechnology* **12,** 139–143.

Hohn B, Puchta H. 2003. Some like it sticky: gene targeting in rice. *Trends in Plant Science* **8,** 51–53.

Jasin M. 1996. Genetic manipulation of genomes with rare-cutting endonucleases. *Trends in Genetics* **12,** 224–228.

Jones GH, Armstrong SJ, Caryl AP, Franklin FC. 2003. Meiotic chromosome synapsis and recombination in *Arabidopsis thaliana*; an integration of cytological and molecular approaches. *Chromosome Research* **11,** 205–215.

Johnson RD, Jasin M. 2000. Sister chromatid gene conversion is a prominent double-strand break repair pathway in mammalian cells. *EMBO Journal* **19,** 3398–3407.

Kalendar R, Tanskanen J, Immonen S, Nevo E, Schulman AH. 2000. Genome evolution of wild barley (*Hordeum spontaneum*) by BARE-1 retrotransposon dynamics in response to sharp microclimatic divergence. *Proceedings of the National Academy of Sciences, USA* **97,** 6603–6607.

Keeler KJ, Gloor G-B. 1997. Efficient gap repair in *Drosophila melanogaster* requires a maximum of 31 nucleotides of homologous sequence at the searching ends. *Molecular and Cellular Biology* **17,** 627–634.

Keeney S. 2001. Mechanism and control of meiotic recombination initiation. *Current Topics in Developmental Biology* **52,** 1–53.

Kim YG, Cha J, Chandrasegaran S. 1996. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences, USA* **93,** 1156–1160.

Kirik A, Salomon S, Puchta H. 2000. Species-specific double-strand break repair and genome evolution in plants. *EMBO Journal* **19,** 5562–5566.

Köhler F, Cardon G, Pöhlman M, Gill R, Schieder O. 1989. Enhancement of transformation rates in higher plants by low-dose irradiation: are DNA repair systems involved in the incorporation of exogenous DNA into the plant genome? *Plant and Molecular Biology* **12,** 189–199.

Kumar S, Fladung M. 2001. Controlling transgene integration in plants. *Trends in Plant Science* **6,** 155–159.

Lebel EG, Masson J, Bogucki A, Paszkowski J. 1993. Stress-induced intrachromosomal recombination in plant somatic cells. *Proceedings of the National Academy of Sciences, USA* **90,** 422–426.

Lehman CW, Trautman JK, Carroll D. 1994. Illegitimate recombination in *Xenopus*: characterization of end-joined junctions. *Nucleic Acids Research* **22,** 434–442.

Liang F, Romanienko PJ, Weaver DT, Jeggo PA, Jasin M. 1996. Chromosomal double-strand break repair in Ku80-deficient cells. *Proceedings of the National Academy of Sciences, USA* **93,** 8929–8933.

Lin F-L, Sperle, K, Sternberg N. 1984. Model for homologous recombination during transfer of DNA into mouse L cells: role for DNA ends in the recombination process. *Molecular and Cellular Biology* **4,** 1020–1034.

Lin F-L, Sperle K, Sternberg N. 1990. Intermolecular recombination between DNAs introduced into mouse L cells is mediated by a nonconservative pathway that leads to crossover products. *Molecular and Cellular Biology* **10,** 103–112.

Lowe B, Mathern J, Hake S. 1992. Active Mutator elements suppress the knotted phenotype and increase recombination at the Kn1-O tandem duplication. *Genetics* **132,** 813–822.

Lucht JM, Mauch-Mani B, Steiner HY, Metraux, JP, Ryals J, Hohn B. 2002. Pathogen stress increases somatic recombination frequency in *Arabidopsis*. *Nature Genetics* **30,** 311–314.

Mason RM, Thacker J, Fairman MP. 1996. The joining of non-complementary DNA double-strand breaks by mammalian extracts. *Nucleic Acids Research* **24,** 4946–4953.

Molinier J, Ries G, Bonhoeffer S, Hohn B. 2004. Interchromatid and interhomolog recombination in *Arabidopsis thaliana*. *The Plant Cell* **16,** 342–352.

Moore JK, Haber JE. 1996. Capture of retrotransposon DNA at the sites of chromosomal double-strand breaks. *Nature* **383,** 644–646.

Moynahan ME, Jasin M. 1997. Loss of heterozygosity induced by a chromosomal double-strand break. *Proceedings of the National Academy of Sciences, USA* **94,** 8988–8993.

Nassif N, Penny J, Pal S, Engels WR, Gloor GB. 1994. Efficient copying of nonhomologous sequences from ectopic sites via P-element induced gap repair. *Molecular and Cellular Biology* **14,** 1613–1625.

Nicolas AL, Munz PL, Young CS. 1995. A modified single-strand annealing model best explains the joining of DNA double-strand breaks in mammalian cells and cell extracts. *Nucleic Acids Research* **23,** 1036–1043.

Offringa R, Franke-van Dijk ME, de Groot MJA, van den Elzen PJM, Hooykaas PJJ. 1993. Nonreciprocal homologous recombination between *Agrobacterium* transferred DNA and a plant chromosomal locus. *Proceedings of the National Academy of Sciences, USA* **90,** 7346–7350.

Orel N, Kirik 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. *The Plant Journal* **35,** 604–612.

Orel N, Puchta H. 2003. Differences in the processing of DNA ends in *Arabidopsis* and tobacco and its implication for genome evolution. *Plant Molecular Biology* **51,** 523–531.

Ow DW. 2002. Recombinase-directed plant transformation for the post-genomic era. *Plant Molecular Biology* **48,** 183–200.

Paques F, Haber JE. 1999. Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiology and Molecular Biology Reviews* **63,** 349–404.

Parniske M, Hammond-Kosack KE, Golstein C, Thomas CM, Jones DA, Harrison K, Wulff BB, Jones JD. 1997. Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the Cf-4/9 locus of tomato. *Cell* **91,** 821–832.

Parniske M, Jones JD. 1999. Recombination between diverged clusters of the tomato Cf-9 plant disease resistance gene family. *Proceedings of the National Academy of Sciences, USA* **96,** 5850–5855.

Perrin A, Buckle M, Dujon B. 1993. Asymmetrical recognition and activity of the I-*Sce*I endonuclease on its site and on intron-exon junctions. *EMBO Journal* **12,** 2939–2947.

Petrov DA. 2001. Evolution of genome size: new approaches to an old problem. *Trends in Genetics* **17,** 23–28.

Petrov DA, Lozovskaya ER, Hartl DL. 1996. High intrinsic rate of DNA loss in *Drosophila*. *Nature* **384,** 346–349.

Petrov DA, Sangster TA, Johnston JS, Hartl DL, Shaw KL. 2000. Evidence for DNA loss as a determinant of genome size. *Science* **287,** 1060–1062.

Pipiras E, Coquelle A, Bieth A, Debatisse M. 1998. Interstitial deletions and intrachromosomal amplification initiated from a double-strand break targeted to a mammalian chromosome. *EMBO Journal* **17,** 325–333.

Porteus MH, Baltimore D. 2003. Chimeric nucleases stimulate gene targeting in human cells. *Science* **300,** 763.

Potrykus I, Spangenberg G. 1995. *Gene transfer to plants*. Berlin: Springer.

Puchta H. 1998. Repair of genomic double-strand breaks in somatic plant cells by one-sided invasion of homologous sequences. *The Plant Journal* **13,** 331–339.

Puchta H. 1999*a*. Use of I-*Sce*I to induce double-strand breaks in *Nicotiana*. *Methods in Molecular Biology* **113,** 447–451.

Puchta H. 1999*b*. DSB-induced recombination between ectopic homologous sequences in somatic plant cells. *Genetics* **152,** 1173–1181.

Puchta H. 2000. Removing selectable marker genes: taking the shortcut. *Trends in Plant Science* **5,** 273–274.

Puchta H. 2002. Gene replacement by homologous recombination in plants. *Plant Molecular Biology* **48,** 173–182.

Puchta H. 2003*a*. Towards the ideal GMP: homologous recombination and marker gene excision. *Journal of Plant Physiology* **160,** 743–754.

Puchta H. 2003*b*. Marker-free transgenic plants. *Plant Cell Tissue and Organ Culture* **74,** 123–134.

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 Research* **21,** 5034–5040.

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. *Proceedings of the National Academy of Sciences, USA* **93,** 5055–5060.

Puchta H, Hohn B. 1991. The mechanism of extrachromosomal homologous DNA recombination in plant cells. *Molecular and General Genetics* **230,** 1–7.

Puchta H, Kocher S, Hohn B. 1992. Extrachromosomal homologous DNA recombination in plant cells is fast and is not affected by CpG methylation. *Molecular and Cellular Biology* **12,** 3372–3379.

Puchta H, Meyer P. 1994. Substrate specificity of plant recombinases determined in extrachromosomal recombination systems. In: Paszkowski J, ed. *Homologous recombination in plants*. Dordrecht, The Netherlands: Kluwer, 123–155.

Ramakrishna W, Emberton J, SanMiguel P, Ogden M, Llaca V, Messing J, Bennetzen JL. 2002. Comparative sequence analysis of the sorghum rph region and the maize rp1 resistance gene complex. *Plant Physiology* **130,** 1728–1738.

Ray A, Langer M. 2002. Homologous recombination: ends as the means. *Trends in Plant Science* **7,** 435–440.

Redei GP, Li SL. 1969. Effects of 'X' rays and ethyl methanesulfonate on the chlorophyll B locus in the soma and on the thiamine loci in the germline of *Arabidopsis*. *Genetics* **61,** 453–459.

Reiss B. 2003. Homologous recombination and gene targeting in plant cells. *International Review of Cytology* **228,** 85–139.

Reiss B, Schubert I, Köpchen K, Wendeler E, Schell J, Puchta H. 2000. RecA stimulates sister chromatid exchange and the fidelity of double-strand break repair, but not gene targeting, in plants transformed by *Agrobacterium*. *Proceedings of the National Academy of Sciences, USA* **97,** 3358–3363.

Ricchetti M, Fairhead C, Dujon B. 1999. Mitochondrial DNA repairs double-strand breaks in yeast chromosomes. *Nature* **402,** 96–100.

Richardson C, Moynahan ME, Jasin M. 1998. Double-strand break repair by interchromosomal recombination: suppression of chromosomal translocations. *Genes and Development* **12,** 3831–3842.

Rinehart TA, Dean C, Weil CF. 1997. Comparative analysis of non-random DNA repair following Ac transposon excision in maize and *Arabidopsis*. *The Plant Journal* **12,** 1419–1427.

Risseeuw E, Offringa R, Franke-van Dijk ME, Hooykaas PJ. 1995. Targeted recombination in plants using *Agrobacterium* coincides with additional rearrangements at the target locus. *The Plant Journal* **7,** 109–119.

Rong YS, Golic KG. 2000. Gene targeting by homologous recombination in *Drosophila*. *Science* **288,** 2013–2018.

Rong YS, Titen SW, Xie HB, Golic MM, Bastiani M, Bandyopadhyay P, Olivera BM, Brodsky M, Rubin GM, Golic KG. 2002. Targeted mutagenesis by homologous recombination in *D.melanogaster*. *Genes and Development* **16,** 1568–1581.

Rubin E, Levy AA. 1997. Abortive gap repair: underlying mechanism for Ds element formation. *Molecular and Cellular Biology* **17,** 6294–6302.

Salomon S, Puchta H. 1998. Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. *EMBO Journal* **17,** 6086–6095.

SanMiguel P, Gaut BS, Tikhonov A, Nakajima Y, Bennetzen JL. 1998. The paleontology of intergene retrotransposons of maize. *Nature Genetics* **20,** 43–45.

SanMiguel P, Tikhonov A, Jin YK, *et al*. 1996. Nested retrotransposons in the intergenic regions of the maize genome. *Science* **274,** 765–768.

Sargent RG, Brenneman MA, Wilson JH. 1997. Repair of site-specific double-strand breaks in a mammalian chromosome by homologous and illegitimate recombination. *Molecular and Cellular Biology* **17,** 267–277.

Schmidt-Puchta W, Orel N, Kirik A, Puchta H. 2004. Intrachromosomal homologous recombination in *Arabidopsis thaliana*. *Methods in Molecular Biology* **262,** 25–35.

Schwacha A, Kleckner N. 1996. Identification of double Holliday junctions as intermediates in meiotic recombination. *Cell* **83,** 783–791.

Shalev G, Levy AA. 1997. The maize transposable element *Ac* induces recombination between the donor site and an homologous ectopic sequence. *Genetics* **146,** 1143–1151.

Shirasu K, Schulman AH, Lahaye T, Schulze-Lefert P. 2000. A contiguous 66-kb barley DNA sequence provides evidence for reversible genome expansion. *Genome Research* **10,** 908–915.

Shirley BW, Hanley S, Goodman HM. 1992. Effects of ionizing radiation on a plant genome: analysis of two *Arabidopsis* transparent testa mutations. *The Plant Cell* **4,** 333–347.

Siebert R, Puchta H. 2002. Efficient repair of genomic double-strand breaks via homologous recombination between directly repeated sequences in the plant genome. *The Plant Cell* **14,** 1121–1131.

Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW. 1983. The double-strand break repair model of recombination. *Cell* **33,** 25–35.

Teng SC, Kim B, Gabriel A. 1996. Retrotransposon reverse-transcriptase-mediated repair of chromosomal breaks. *Nature* **383,** 641–644.

Terada R, Urawa H, Inagaki Y, Tsugane K, Iida S. 2002. Efficient gene targeting by homologous recombination in rice. *Nature Biotechnology* **20,** 1030–1034.

Tinland B. 1996. The integration of T-DNA into plant genomes. *Trends in Plant Science* **1,** 178–184.

Tinland B, Hohn B, Puchta H. 1994. *Agrobacterium tumefaciens* transfers single-stranded T-DNA into the plant cell nucleus. *Proceedings of the National Academy of Sciences, USA* **91,** 8000–8004.

Tovar J, Lichtenstein C. 1992. Somatic and meiotic chromosomal recombination between inverted duplications in transgenic tobacco plants. *The Plant Cell* **4,** 319–332.

Tzfira T, Frankman LR, Vaidya M, Citovsky V. 2003. Site-specific integration of *Agrobacterium tumefaciens* T-DNA via double-stranded intermediates. *Plant Physiology* **133,** 1011–1023.

Vicient CM, Suoniemi A, Anamthawat-Jonsson K, Tanskanen J, Beharav A, Nevo E, Schulman AH. 1999. Retrotransposon BARE-1 and its role in genome evolution in the genus *Hordeum*. *The Plant Cell* **11,** 1769–1784.

Vicient CM, Jaaskelainen MJ, Kalendar R, Schulman AH. 2001. Active retrotransposons are a common feature of grass genomes. *Plant Physiology* **125,** 1283–1292.

Walbot V. 1996. Sources and consequences of phenotypic and genotypic plasticity in flowering plants. *Trends in Plant Science* **1,** 27–33.

Xiao YL, Li X, Peterson T. 2000. Ac insertion site affects the frequency of transposon-induced homologous recombination at the maize p1 locus. *Genetics* **156,** 2007–2017.

Xiao YL, Peterson T. 2000. Intrachromosomal homologous recombination in *Arabidopsis* induced by a maize transposon. *Molecular and General Genetics* **263,** 22–29.

Zubko E, Scutt C, Meyer P. 2000. Intrachromosomal recombination between attP regions as a tool to remove selectable marker genes from tobacco transgenes. *Nature Biotechnology* **18,** 442–445.