Gene replacement by homologous recombination in plants

Holger Puchta
*Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK), Corrensstrasse 3, 06466 Gatersleben, Germany
(e-mail puchta@ipk-gatersleben.de)*

Abstract

After the elucidation of the sequence of the yeast genome a major effort was started to elucidate the biological function of all open reading frames of this organisms by targeted gene replacement via homologous recombination. The establishment of the complete sequence of the genome of *Arabidopsis thaliana* would principally allow a similar approach. However, over the past dozen years all attempts to establish an efficient gene targeting technique in flowering plants were in the end not successful. In contrast, in *Physcomitrella patens* an efficient gene targeting procedure has been set up, making the moss a valuable model system for plant molecular biologists. But also for flowering plants recently several new approaches – some of them based on the availability of the genomic sequence of *Arabidopsis* – were initiated that might finally result on the set up of a general applicable technique. Beside the production of hyper-recombinogenic plants either via expression or suppression of specific gene functions or via undirected mutagenesis, the application of chimeric oligonucleotides might result in major progress.

Introduction

Two different ways for alteration of covalent linkage between DNA molecules is possible: either a region of sequence identity between the partners is needed (homologous recombination) or no sequence-specific requirements have to be fulfilled (illegitimate recombination also referred to as non-homologous end joining). For the controlled manipulation of genomes homologous recombination is of special interest as the product of the reaction is predictable. In transformation experiments in prokaryotes and lower eukaryotes such as yeast homologous recombination is predominating. In yeast, for instance, an external piece of DNA harbouring homologous sequences to the genome will almost exclusively integrate into the corresponding genomic position, a phenomenon called 'gene targeting'. Transformation with sequences that carry no homology to the yeast genome is hardly possible. At present reverse genetics is used to elucidate the role of specific proteins in an organism. Using information obtained from various genome sequencing programs the gene coding for the protein of interest can be modified to render it non-functional. The transformed organism carrying a loss-of-function mutation can then be analysed for its phenotype. Gene targeting is thus a powerful tool for directed 'knockout' of genes. Furthermore, using this technique it is possible to perform subtle changes such as the modification of single nucleotides in a gene *in situ*, an achievement that cannot be obtained by random insertion mutagenesis. In higher eukaryotes including plants DNA integrates mainly via illegitimate recombination in an undirected, sequence-independent manner in the genome. Tremendous efforts have been undertaken to improve gene targeting in animals and plants and the development of a feasible technique for mouse embryonic stem cells has revolutionized basic as well as applied research in this area. Unfortunately, similar improvements have not been achieved for higher plants. However, a number of approaches were developed recently that might indeed lead to a feasible gene targeting technique in the near future. The main emphasis of this review will lie on these

new developments; mainly gene targeting experiments performed in plants will be discussed. Results from other organisms or basic aspects of recombination in plants will only be discussed were it is necessary for understanding. To obtain a more detailed picture of specific aspects of homologous recombination and gene targeting in plants the reader is recommended to consult a row of recently published reviews (Puchta, 1998a; Gorbunova and Levy, 1999; Hohn and Puchta, 1999; Mengiste and Paszkowski, 1999; Vergunst and Hooykaas, 1999; Oh and May, 2001; Schaefer, 2001).

Attempts to improve gene targeting via the targeting vector

Mammals as guiding light

In mammals gene targeting has become a routine technique for knocking out gene functions. Since the pioneering work of the groups of Smithies (Doetschman *et al.*, 1987) and Capecci (Thomas and Capecci, 1987) thousands of loss-of-function mutations of mouse have been generated by gene targeting in embryonic stem (ES) cells. Numerous experiments indicate gene targeting frequencies of 10^{-2} (one homologous integration event per hundred random integration events) or higher in ES cells (for review see Jasin *et al.*, 1996). Unfortunately, similar frequencies have not been obtained in gene targeting experiments in plants. Since the pioneering first report on the targeting of a transgene locus in plants (Paszkowski *et al.*, 1988) no convincing enhancement of the targeting frequency could be achieved within the next dozen of years. In principle, there are two ways to produce transgenic plants. Either the DNA is directly transferred into plants cells by PEG transformation or electroporation of protoplasts or by the bombardment of various plant tissues (for details of plant transformation see Potrykus and Spangenberg, 1996). Alternatively, *Agrobacterium tumefaciens* is used as a vector for plant transformation (for recent revieuws see Rossie *et al.*, 1996; Tinland, 1996; Hansen and Chilton, 1999; Zupan *et al.*, 2000). Both direct gene transfer (Paszkowski *et al.*, 1988; Halfter *et al.*, 1992) and *Agrobacterium*-mediated T-DNA transformation (Offringa *et al.*, 1990; Lee *et al.*, 1990) were applied in the initial studies in which the restoration of selectable marker genes (confering kanamycin or hygromycin resitance) was used to estimate the frequency of gene targeting in plants. In all cases the observed gene targeting frequencies were low, 10^{-4} to 10^{-5}, independent of the plant species (tobacco or *Arabidopsis*) and the transformation method. Considering the factors that improved gene targeting in mouse, various attempts have been undertaken in the following years to improve the ratio of homologous to illegitimate recombination. However, neither extending the length of homology in the transferred DNA to up to 22 kb (Thykjaer *et al.*, 1997) nor including negative selectable markers outside of the homology of the targeting vector to be able to select against random integration (Risseeuw *et al.*, 1997; Gallego *et al.*, 1999; for a detailed discussion on negative selectable markers, see Vergunst and Hooykaas, 1999) did result in a significantly higher frequency of gene targeting in plants.

Nevertheless, it is well possible to enhance homologous recombination in plant cells by 'activating' the target site in the genome by the induction of a double-strand break (DSB). Recombination is inducible by DSBs and all current recombination models are based on the repair of such breaks (for review see Pasques and Haber, 1999). Via transient expression of the restriction enzyme I-*Sce*I a DSB was introduced at a transgenic restriction site within the tobacco genome *in vivo* resulting in a homologous integration frequency of up to 10^{-2} at this site (Puchta *et al.*, 1996). However, this strategy is not applicable for targeting genes at all, as such a break can only be induced at the specific recognation sites of the rare-cutting restriction enzyme. Therefore this 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 as described in the accompanying review by David Ow (Ow, 2001).

The targeting of 'natural' genes in Arabidopsis thaliana

Two cases have been reported over the years in which natural genes (the loss of function cannot be selected for) were knocked out in *Arabidopsis thaliana* by gene targeting after *Agrobacterium*-mediated transformation. Eric Lam's group was able to isolate one *Arabidopsis* callus out of 2580 tissue culture transformants in which targeting of the *TGA3* locus had occurred. However, the respective callus was a chimera for the targeted locus and could not be regenerated (Mia and Lam, 1995). Two years later Martin Yanofsky's group in collaboration with Eric Lam was finally

able to obtain a mutant *Arabidopsis* plant by knocking out the *AGL5* MADS-box gene. Using vacuum infiltration of bolting plants, they were able to recover one targeted event in 750 transformations (Kempin *et al.*, 1997). This relatively high frequency induced hopes that the application of vacuum infiltration might have helped to transform especially active in homologous recombination, reminiscent of mouse ES cells (Essers *et al.*, 2000) and that indeed gene targeting would become feasible for *Arabidopsis*. However, no statistically sound conclusion as to the frequencies of targeting could be drawn from the reported single event and the results were discussed controversially (Puchta 1998a; Liljegren and Yanofsky, 1998). Unfortunately, for the following years no further successful gene targeting experiments with vacuum infiltration were reported, so that the initial hopes were not fulfilled.

Ectopic recombination: a not wanted by-product of gene targeting

Beside the low frequencies of gene targeting another cause of disillusionment for several researchers was the fact that although the respective marker gene was restored via homologous recombination, in some cases the target locus was not changed in the expected way (Offringa *et al.*, 1993; Risseeuw *et al.*, 1995; Reiss *et al.*, 2000). The classical double-strand break repair (DSBR) model of recombination (Szostak *et al.*, 1983) predicts that integration of both ends of the extrachromosomal piece of DNA into the target locus should occur via homologous recombination. However gene targeting experiments in tobacco revealed two further classes of homologous recombinants. Firstly, integration of the extrachromosomal DNA in the target locus by a combination of homologous and illegitimate recombination was detected. Secondly, sequence homology can also be copied from the target locus to extrachromosomal DNA. This modified sequence may then integrate elsewhere in the genome ('ectopic targeting'). Recent studies in somatic plant cells indicate that these findings are not suprising because the main mechanism of recombination is described best by a synthesis-dependent strand annealing (SDSA)-like model of recombination (Rubin *et al.*, 1997; Puchta 1998b; for review see Gorbunova and Levy, 1999) rather then by the DSBR model. It could be demonstrated that DNA ends can react independently and one end might be repaired by homologous and the other by illegitimate recombination. Therefore, gene targeting experiments in higher plants have to include a thorough Southern analysis of recombinants to identify the desired knock-out, in which both ends of the vector sequence have been integrated via homology in the target locus. But why is the SDSA mechanism operating in somatic plant cells? It seems to be a prerequisite for the stability of the highly repetitive plant genome. Homologous interactions between ectopic sequences can, according to the DSBR model, result in chromosomal rearrangements as translocations whereas according to the SDSA model translocations are avoided (for detailed discussion on DSB repair, see Puchta, 1999; Gorbunova and Levy, 1999).

Production of a targeting vector in vivo*: a solution to the problem?*

An apparently highly efficient variant of gene targeting has been developed recently for *Drosophila melanogaster* by the group of Kent Golic (Rong an Golic, 2000, 2001). In this method the contruct 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. In the female germ line of *Drosophila* gene targeting occurred in about one out of 500 cases. Although questions were raised about the general applicability of the technique (Engels, 2000) and certain product classes obtained are hard to explain by a simple mechanism, it seems worthwhile to test such a strategy in plants, too (for discussion see also Kumar and Fladung, 2001). On the one hand, the set-up of the system seems to be very complex, as beside constructions of a donor sequence with sites for recombinase and restriction enzyme, expression cassettes 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.

Productions op plants with a hyper-recombination phenotype

Over-expression of heterologous genes

If homologous recombination is not working efficiently in plants why not try to transfer the enzyme machinery from an organism in which gene targeting is working well to plants? This idea is followed since a row of years in the field, and pioneering work was done by the group of Bernd Reiss. Transgenic tobacco plants were produced which expressed the key actor of homologous recombination in *Escherichia coli*, RecA. As expected, the rate of intrachromosomal homologous recombination was enhanced by one order or magnitude (Reiss *et al.*, 1996). Hopes were flying high that also gene targeting frequencies would be enhanced by RecA, however in the respective experiments for two different transgene loci no significant of the gene targeting frequency could be achieved with *Agrobacterium*-mediated transformation (Reiss *et al.*, 2000). Although one can speculate that the T-DNA complexed with bacterial proteins might have hindered homologous recombination, experiments in yeast demonstrate that in this organism T-DNA is perfectly integrated via homology (Bundock *et al.*, 1996). In a different approach the group of Avi Levy over-expressed the RuvC gene in tobacco. RuvC is responsible for the resolution of homologous recombination intermediates (Holliday junctions) in *E. coli.* They were able to demonstrate that frequencies of extrachromosomal, intrachromosomal as well as interchromosomal recombination were enhanced up to two orders of magnitude, although no conclusive explanation for this phenomenon could be supplied (Shalev *et al.*, 1999). It will be interesting to see whether a similar effect can also be obtained with gene targeting. A lesson learnt from these experiments is, however, that it is very hard to predict the effects of the expression of a heterologous protein in plants on recombination. Due to complex interactions with different host proteins it might be possible to influence some or even different but not all of the reactions in which the respective protein is participating in its natural 'host'.

Modulating the function of plant genes

Basic features of the enzyme machinery involved in recombination are conserved between yeast mammals and plants. Thus, factors of putatively similar function can be identified by sequence homologies (Vergunst and Hooykaas, 1999; Bhatt *et al.*, 2001). With the complete sequence of the *Arabidopsis* genome at hand, multiple open reading frames with similarities to known genes of other organisms involted in repair and recombination can be identified (see the compilation of homologous to known repair and recombination genes of other organisms; Arabidopsis Genome Initiative, 2000). Moreover, it has been reported for vertebrates that mutations in certain genes lead to enhanced homologous recombination frequencies. For example, the knockout of the *Mre 11* gene of chicken leads to increased gene targeting frequencies (Yamaguchi-Iwai *et al.*, 1999). *Mre 11* together with *RAD50* is in yeast part of a multifunctional protein complex that can act as nuclease and is involved in certain reactions of homologous as well as in illegitimate recombination (for review see Haber, 1998). Homologous factors exist in *Aribidopsis* (Hartung and Puchta, 1999; Gallego *et al.*, 2001) and the group of Charles White could indeed demonstrate that a RAD50 T-DNA insertion mutant showed an increase of intrachromosomal homologous recombination by one order of magnitude (Gherbi *et al.*, 2001). It will be interesting to see whether such a similar increase can also be detected for gene targeting.

There are other genes the mutation or suppression of which might lead to higher targeting frequencies. In chicken enhanced target integration of DNA was found for a mutant of the gene responsible for Bloom's syndrome (BLM) (Wang *et al.*, 2000). BLM is a member of the RecQ helicase family involved in pro- and eukaryotes in DNA replication and recombination (for review see Karow *et al.*, 2000). Recently several RecQ homologues have been characterized in *Arabidopsis* (Hartung *et al.*, 2000).

Blocking illegitimate recombination should increase the relation of homologous to random integration events. Two components of this pathway, DNA Ligase IV and XRCC4, have already been characterized from *Arabidopsis* (West *et al.*, 2000) and more will follow. Without question, in the next few years a series of mutants with insertions within different genes involved in recombination processes will be produced and characterized. Analysis of their effects on recombination will lead to identification of plants with enhanced gene targeting frequencies. Alternatively, also over-expression of *Arabidopsis* genes involved in homologous recombination or its regulation might help to increase gene targeting efficiency. For the MIM gene (Hannin *et al.*, see below) it has already been recently demonstrated that over-expression is

correlated with increased homologous recombination frequencies.

Undirected mutagenesis can result in the production of hyper-recombinogenic plants

The classical genetic method to use undirected mutagenesis to isolate a phenotype of interest can provide important new insights because, in contrast to gene isolation by homology search, no assumptions need to be made about which genes might be involved in the phenomenon of interest. This strategy was followed by the group of Jurek Paszkowski. As defects in recombination should be correlated with an impaired ability to repair DSBs resulting in an increased sensitivity against X-rays, screening of *Arabidopsis* populations, mutagenized by T-DNA insertions, for radiation-sensitive mutants should lead to the identification of genes involved in this process. Indeed, an insertion mutant into a gene (MIM) closely related to the SMC (structural maintenance of chromosomes) gene family of other eukaryotes was correlated with an X-ray-sensitive phenotype. The mutant showed a decreased frequency of intrachromosomal homologous recombination (Mengiste *et al.*, 1999). Overexpression of the MIM gene in wild-type *Arabidopsis* plants, on the other hand, did indeed increase the frequency of homologous intrachromosomal recombination, although the effect was too small to expect a decisive improvement of gene targeting frequencies (Hannin *et al.*, 2000).

A more direct strategy would be to screen for hyper-recombination mutants. For this purpose the group of Avi Levy used a 'sulfur' mutation of tabacco. In this mutant somatic cross-overs between homologues can be detected *in planta* as yellow/green twin sectors, although the molecular nature of the process is not characterized in detail (Carlson, 1974). Using a transposon as insertion mutagen, they were able to obtain a plant in which the frequency of interchromosomal recombination was enhanced by three orders of magnitude. Moreover, the mutant proved to be resistant to elevated doses of radiation. Interestingly, although the frequency of homologous recombination between extrachromosomal substrates was also increased, no enhancement could be detected with an introchromosomal substrate (Gorbunova *et al.*, 2001). It will be very interesting to test this mutant for gene targeting, especially as an enhancement of several orders of magnitude would indeed be a major breakthrough. Even if this will not be the case for this mutant, the approach taken seems to be promising. There is no question that further mutant screens, for example also with intrachromosomal recombination substrates and also with *Arabidopsis* (e.g. Swoboda *et al.*, 1994), should lead to the identification of more hyperrecombining plant mutants making this strategy a very promising path that indeed might lead to the set-up of a feasible gene targeting technique in plants.

Chimeraplasty

A technique for targeted genome mutations

Based on findings that transcription enhances RecA-mediated homologous pairing of DNA *in vitro*, Eric Kmiec and his group came to the conclusion that pairing can be enhanced in the presence of complementary RNA. Therefore they constructed self-complementary chimeric oligonucleotides that consist of a combination of DNA and RNA sequences (for reviews see Ye *et al.*, 1998; Rice *et al.*, 2001). Chimeric oligonucleotides consist of a DNA 'mutator' region of 5 or 6 nucleotides complementary to the target including a mutation which should be introduced into the genome, framed by two $2'$ *O*-methyl RNA bridges of 8–12 nucleotides that are also complementary to the target locus (*O*-methylation increases the stability of the RNA against degradation within cells). These regions in turn are flanked by hairpin loops of 3 or 4 T nucleotides each. The hairpins are connected to a DNA region of some 30 nucleotides, fully complementary to the 'upper' RNA-DNA-RNA strand. The 'lower' strand contains a strand break that seems to be necessary for the topological interwinding of the chimera into the target DNA. That chimeric oligonucleotides can be used to induce mutations in a mammalian genomic target sequence *in vivo* was first demonstrated by using chimeric oligonucleotides with a single-base mismatch corresponding to the wild-type sequence within the β-globin gene. This construct was used for repairing the 'sickle' allele, resulting in a chromosomal reversion to the wild-type β-globulin. An astonishingly high reversion frequency of 20% was reported (Cole-Strauss *et al.*, 1996). Many reports have been published since then about the successful use of chimeric oligonucleotides in mammalian cells, demonstrating the usefulnesss of the technique. However, the mutation frequencies reported differed dramatically within and between experiments (e.g. van der Steege *et al.*, 2001).

Targeted mutations in plants

Facing the gene targeting dilemma in plants it was therefore obvious to use the technique for the targeted mutation of plant genomes (for review see Oh and May, 2001). In 1999, in parallel studies in tobacco (Beetham *et al.*, 1999) and maize (Zhu *et al.*, 1999), an endogenous gene and a transgene were mutated by chimeric oligonucleotides. In both studies the first enzyme of the biosynthetic pathway of branched chain amino acids was chosen as endogenous gene. In tobacco, the gene is referred to as acetolacetate synthase (ALS) and in maize as acetohydroxy acid synthase (AHAS). The mutation of certain amino acids in the protein results in a resistant phenotype, which can be selected for by application of imidazoline and sulfonylurea herbicides. In contrast to animals cells, where chimeric oligonucleotides are normally delivered via cationic liposomes, the chimeric oligonucleotides were delivered into the plant cells by biolistics. The strategy chosen in the experiments was to obtain a selectable, dominant mutation in one gene, as both plant species harbour gene families of ALS or AHAS. Two different chimeric oligonucleotides were applied in both studies and mutation frequencies were assessed to be arround 10^{-4} (Zhu *et al.*, 1999; Hohn and Puchta, 1999). Surprisingly, the specificity of the reaction was not as precise as expected. In tobacco none of the isolated resistant plant cell lines proved to harbour the expected target mutation. In all cases one nucleotide $5'$ to the mismatched nucleotide was changed, thus, strictly speaking, no targeted but only 'semi-targeted' events were obtained. As controls indicate that chimeric oligoncleotides do not enhance mutation rates all over the genome, the chimeric oligonucleotides – at least in plants – seems to induce mutation not only at the specific mismatch but within the DNA stretch complementary to the target locus. In maize the predicted target mutations were indeed recovered, although 'semi-targeted' mutations were also found (Zhu *et al.*, 1999, 2000). In both studies transgenic model systems using the green fluorescent protein (GFP) as marker gene were set up. Plants containing an inactive GFP gene with either a frame shift mutation (Beetham *et al.*, 1999) or an internal termination codon (Zhu *et al.*, 1999) were bombarded with the respective COs to reverse the mutations. Putative revertants that are able to express a functional GFP can easily be detected by fluorescence microscopy. In tobacco as well as in maize GFP-positive cells were found at frequencies at least as

high as reported for chimeric oligonucleotide-induced mutations in the ALS or AHAS genes. The molecular analysis of one mutant line in maize indicated that the induced mutation occurred indeed at the mismatch position of the chimeric oligonucleotide, but instead of the intended T a C was detected. The mutated maize line was regenerated and it could be domonstrated that the mutation was transferred faithfully to the progeny (Zhu *et al.*, 1999). Similarly stable propagation of the desired targeted mutations in the AHAS gene could be demonstrated recently for maize (Zhu *et al.*, 2000). Taking advantage of the ease of the GFP assay system the establishment of the mutated lines in tobacco and in maize will definitely help to further optimize the efficiency of chimeric oligonucleotide induced mutations in plants.

In vitro *assays and a bunch of open questions*

Experiments were also performed using chimeric oligonucleotides *in vitro* with extracts of mammalian and, more recently, plant cells. It could be demonstrated with human cell extracts that the repair reaction requires the presence of the mismatch repair protein Msh2. In extracts of a cell line devoid of Msh2 and wild-type cell extracts depleted of Msh2 by a specific antibody, a reduced efficiency of repair was measured (Cole-Strauss *et al.*, 1999). Thus, whereas gene targeting relies exclusively on factors involved in homologous recombination, chimeric oligonucleotide-directed repair seems to require factors involved in mismatch reapair. In line with this argument is the finding that via the use of chimeric oligonucleotides only base substitutions and, with less efficiency, deletions or insertions of one or two bases on the target locus can be achieved (Ye *et al.*, 1998), whereas via homologous recombination genomic sequences of tens of kilobases can be changed, inserted or deleted. *In vitro* experiments were also performed with nuclear extracts of different plant species, leading to the discovery that in extracts of tobacco the repair was in some cases less precise then in the extracts of other plant species (Rice *et al.*, 2000). This finding is in accordance with the data obtained *in vivo* (Beetham *et al.*, 1999). However, recent result of Eric Kmiec's group indicate that with single-stranded DNA oligonucleotides with varying numbers of phosophorothiote linkages at the $3'$ and $5'$ ends (to block exonucleolitic degradation) higher repair efficiencies can be achieved than with chimeric oligonucleotides in animal as well as in plant

extracts (Gamper *et al.*, 2000). Moreover, the involvement of Msh2 in the repair process could not be sustained, when the phosphorothioate-containing oligonucleotides were used. Thus, at present several questions remain to be answered, the most important of which are whether phosphorothioate-containing oligonucleotides can also be used to mutate genes *in vivo* and whether the low frequencies obtained for genomic repair in plants can be drastically increased in future. The situation is somehow similar to that after the first report on gene targeting in plants (Paszkowski *et al.*, 1988). It has been unambiguously demonstrated that a basic principle works in plants, and hopes are flying high that a feasible technique will be available soon. Hopefully, the oligionucleotide-mediated DNA repair will leads us in the next years up the hills of glory and not into a similar valley of tears as gene targeting, when we where trying to optimize the technique.

Physcomitrella patens : an alternative to be taken seriously

Gene targeting in moss

In contrast to flowering plants, the moss *Physcomitrella patens* is able to integrate DNA efficiently by homologous recombination (for a recent review see Schaefer, 2001). This finding, first reported by Didier Schaefer and Jean-Pierre Zryd (Schaefer an Zryd, 1997), marked a major breakthrough for plant molecular biology. Soon afterwards it could be demonstrated by Ralf Reski's group that the technique can indeed be used efficiently for the knock-out of gene functions (Strepp *et al.*, 1998). The moss has turned out to be a very valuable system for investigation of basic processes of plant biology. Mosses share with other plants the same kind of hormones, high similarities in light and stress signalling response systems, as well as features of cell differentation and nucleus organelle interactions. This is also demonstrated by the fact that most of the ESTs of *Physecomitrella patens* show strong homology to genes of other plants (e.g. Reski *et al.*, 1998). Thus, it seems likely that in many cases where the biological function of a specific gene is under investigation in a higher plant a homologue might be found in *Physcomitrella*. A gene targeting experiment in *Physcomitrella* might therefore help to elucidate its general function. Although this strategy was under debate in the beginning (Puchta, 1998; Reski, 1998a;

Liljegren and Yanofsky, 1998), the reports of successful knockouts of moss are starting to accumulate (e.g. Hofmann *et al.*, 1999; Girod *et al.*, 1999) and the technique is in the process of becoming a popular tool in plant biology. Also from a practical point of view experiments with *Physcomitrella* are attractive. The life cycle takes 2–3 months per generation and consists mainly of the gametophyte stage. Thus, in contrast to diploid organisms, recessive phenotypes can be scored immediately after a single gene targeting reaction. After fertilization multiple spores can be harvested easily for analysis (for recent reviews on the biology of *Physcomitrella patens* see Cove *et al.*, 1997; Reski, 1998b, 1999).

Why is gene targeting so efficient in moss?

An extremely important question that remains to be answered – especially as it might help to establish gene targeting in higher plants – is why homologous integration takes place so efficiently in moss? Many possible explanations can be brought forward. Is the enzyme machinery different or the regulation of its expression? Apparently PEG-mediated transformation is especially efficient for obtaining gene targeting in moss (Schaefer, 2001) but not in higher plants (Mengiste and Paszkowski, 1999). Transformation apparently takes place during a specific phase of the cell cycle, at the G_2/M transition, and it can be speculated that homologous recombination is especially efficient at this time (Reski, 1998a). Indeed, indirect evidence exists that in mammals homologous recombination is more efficient during S and G_2 phase than within G_1 (Dronkert *et al.*, 2000). However, it remains to be shown whether a putative moss-specific cell cycle block is indeed responsible for the dramatic enhancement in homologous recombination. It was postulated that the haploid state *per se* is a prerequisite for the efficiency of gene targeting (Schaeffer and Zryd, 1998); however, in higher plants transformation experiments with haploid tissue did not result in higher gene targeting frequencies (Mengiste and Paszkowski, 1999). An intriguing finding is the fact that in many cases several copies of the plasmid DNA were integrated per target locus (Schaeffer and Zryd, 1997) and that extrachromosomal DNA can persist in *Physcomitrella* over longer periods (Ashton *et al.*, 2000). This hints to the possibility that incomming plasmid DNA might become more 'activated' for homologous interactions. Only further studies about the recombination mechanisms in *Physcomitrella patens* and the factors involved will give us a deeper un-

derstanding of the recombination behaviour of this organism.

Conclusions

A major goal of plant biotechnology is to develop new and improve existing elite cultivars. To meet this goal, it will be necessary to both improve existing and develop novel strategies for plant genome manipulation. The ability to modify a resident gene *in situ* or to integrate a transgene at a specific genomic position in a controlled way is a central issue in this context. Gene targeting will permit validation of gene function in plants through knockout of target genes. This reverse genetics approach is expected to become increasingly important in the post-genomic era in order to demonstrate gene function thus permitting protein engineering and overcoming problems of gene silencing and position effects so often encountered in transgenic plants. Time will tell when and by which approach gene targeting in higher plants will be achieved in a feasible way. That it will be achieved in the long run seems to me to be out of question, but is is hard to guess when it will be the case. As we have various approaches at hand, we should not be too pessimistic about time frames. In general, two different strategies have been developed to improve gene targeting frequencies in plants. The first one aims at achieving gene targeting in a wild-type genetic background and includes the production of a targeting vector *in vivo*, chimeraplasty and the moss *Physcomitrella patens* in which gene targeting is naturally efficient. The second strategy aims at improving gene targeting efficiencies in a mutated genetic background and includes the over-expression of heterologous recombination genes, the modulation of the expression of endogenous plant genes involved in recombination and the identification of hyper-recombinogenic mutants. Wheres the first strategy will result in mutants with a stable genetic background, further offorts have to be taken to avoid unwanted secondary effects due to a mutated genetic background obtained with the second strategy. All these approaches are followed up in an active community by a growing number of researchers and we have to wait and see what their results will be.

Acknowledgements

References

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

Ashton, N.W., Champagne, C.E.M., Weiler, T. and Verkoczy, L.K. 2000. The bryophyte *Physcomitrella patens* replicates extrachromosomal transgenic elements. New Phytol. 146: 391–402.

Beetham, P.R., Kipp, P.B., Sawycky, X.L., Arntzen, C.J. and May, G.D. 1999. A tool for functional plant genomics: chimeric RNA/DNA oligonucleotides cause *in vivo* genespecific mutations. Proc. Natl. Acad. Sci. USA 96: 8774–8778.

Bhatt, A.M., Canales, C. and Dickinson, H.G. 2001. Plant meiosis: the means to 1N. Trends Plant Sci. 6: 114–121.

Bundock, P., den Dulk-Ras, A., Beijersbergen, A. and Hooykaas, P.J. 1995. Trans-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae.* EMBO J. 14: 3206–3214.

Carlson, P.S. 1974. Mitotic crossing over in a higher plant. Genet. Res. Camb. 24: 109–112.

Cole-Strauss, A., Gamper, H., Holloman, W.K., Munoz, M., Cheng, N. and Kmiec, E.B. 1999. Targeted gene repair directed by the chimeric RNA/DNA oligonucleotide in a mammalian cell-free extract. Nucl. Acids Res. 27: 1323–1330.

Cole-Strauss, A., Yoon, K., Xiang, Y., Byrne, B.C., Rice, M.C., Gryn, J., Holloman, W.K. and Kmiec, E.B. 1996. Correction of the mutation responsible for sickle cell anemia by an RNA-DNA oligonucleotide. Science 273: 1386–1389.

Cove, D.J., Knight, C.D. and Lamparter, T. 1997. Mosses as model systems. Trends Plant Sci. 2: 99–105.

Doetschman, T., Gregg, R.G., Maeda, N., Hooper, M.L., Melton, D.W., Thompson, S. and Smithies, O. 1987. Targeted correction of a mutant HPRT gene in mouse embryonic stem cells. Nature 330: 576–578.

Dronkert, M.L., Beverloo, H.B., Johnson, R.D., Hoeijmakers, J.H., Jasin, M. and Kanaar, R. 2000. Mouse RAD54 affects DNA double-strand break repair and sister chromatid exchange. Mol. Cell Biol. 20: 3147–3156.

Engels, W.R. 2000. Reversal of fortune for *Drosophila* geneticists? Science 288: 1973–1975.

Essers, J., van Steeg, H., de Wit, J., Swagemakers, S.M., Vermeij, M., Hoeijmakers. J.H. and Kanaar, R. 2000. Homologous and non-homologous recombination differentially affect DNA damage repair in mice. EMBO J. 19: 1703–1710.

Gallego, M.E., Jeanneau, M., Granier, F., Bouchez, D., Bechtold, N. and White I. 2001. Disruption of the *Arabidopsis RAD50* gene leads to plant sterility and MMS sensitivity. Plant J. 25: 31–41.

Gallego, M.E., Sirand-Pugnet, P. and White, C.I. 1999. Positive-negative selection and T-DNA stability in *Arabidopsis* transformation. Plant Mol. Biol. 39: 83–93.

Gamper, H.B., Parekh, H., Rice, M.C., Bruner, M., Youkey, H. and kmiec, EB. 2000. The DNA strand of chimeric RNA/DNA oigonucleotides can direct gene repair/conversion activity in

mammalian and plant cell-free extracts. Nucl. Acids Res. 28: 4332–4339.

Gherbi, H., Gallego, M.E., Jalut, N., Lucht, J.M., Hohn, B. and White, C.I. 2001. Homologous recombination in planta is stimulated in the absence of Rad50. EMBO Rep. 2: 287–291.

Girke, T., Schmidt, H., Zahringer, U., Reski, R. and Heinz, E. 1998. Identification of a novel △6-acyl-group desaturase by targeted gene disruption in *Physcomitrella patens*. Plant J. 15: 39–48.

Girod, P.A., Fu, H., Zryd, J.P. and Vierstra, R.D. 1999. Multiubiquitin chain binding subinit MCB1 (RPN10) of the 26S proteasome is essential for developmental progression in *Physcomitrella patens*. Plant Cell 11: 1457–1472.

Gorbunova, V., Avivi-Ragolski, N., Shalev, G., Kovalchuk, I., Abbo, S., Hohn, B. and Levy, A.A. 2000. A new hyperrecombinogenic mutant of *Nicotiana tabacum*. Plant J. 24: 601–611.

Gorbunova, V. and Levy, A.A. 1999. How plants make ends meet: DNA double-strand break repair. Trends Plant Sci. 4: 263–269.

Haber, J.E. 1998. The many interfaces of Mre11. Cell 95: 583–586.

Halfter, U., Morris, P.C., and Willmitzer, L. 1992. Gene targeting in *Arabidopsis thaliana*. Mol. Genet. 231: 186–193.

Hanin, M., Mengiste, T., Bogucki, A. and Paszkowski, J. 2000. Elevated levels of intrachromosomal homologous recombination in *Arabidopsis* overexpressing the MIM gene. Plant J. 24: 183–189.

Hansen, G. and Chilton, M.D. 1999. Lessons in gene transfer to plants by a gifted microbe. Curr. Top. Microbiol. Immunol. 240: 21–57.

Hartung, F., Plchova, H. and Puchta, H. 2000. Molecular characterization of RecQ homologues in *Arabidopsis thaliana*. Nucl. Acids Res. 28: 4275–4282.

Hartung, F. and Puchta, H. 1999. Isolation of the complete cDNA of the Mre11 homologue of *Arabidopsis* (Accession No. AJ243822) indicates conservation of DNA recombination mechanisms between plants and other eucaryotes (PGR 99–132). Plant Physiol. 121: 311.

Hofman, A.H., Codon, A.C., Ivascu, C., Russo, V.E., Knight, C., Cove, D., Schaefer, D.G., Chakhparoian, M. and Zryd, J.P. 1999. A specific member of the Cab multigene family can be efficiently targeted and disrupted in the moss *Physcomitrella patens*. Mol. Gen. Genet. 261: 92–99.

Hohn, B. and Puchta, H. 1999. Gene therapy in plants. Proc. Natl. Acad. Sci. USA 96: 8321–8323.

Jasin, M., Moynaham, M.E. and Richardson, C. 1996. Targeted transgenesis. Proc. Natl. Acad. Sci. USA 93: 8804–8808.

Karow, J.K., Wu, L. and Hickson, I.D. 2000. RecQ family helicases: roles in cancer and ageing. Curr. Opin. Genet. Dev. 10: 32–38.

Kempin, S.A., Liljegren, S.J., Block, L.M., Rounsley, S.D., Yanofsky, M.F. and Lam E. 1997. Targeted disruption in *Arabidopsis*. Nature 389: 802–803.

Kumar, S. and Fladung, M. 2001. Controlling trangene integration in plants. Trends Plant Sci. 6: 155–159.

Lee, K.Y., Lund, P., Lowe, K. and Dunsmuir, P. 1990. Homologous recombination in plant cells after *Agrobacterium*-mediated transformation. Plant Cell 2: 415–425.

Liljegren, S.J and Yanofsky, M.F. 1998. Towards targeted transformation in plants. Response: targeting *Arabidopsis*. Trends Plant Sci. 3: 79–80.

Mengiste, T. and Paszkowski, J. 1999. Prospects for the precise engineering of plant genomes by homologous recombination. Biol. Chem. 380: 749–758.

Mengiste, T., Revenkova, E., Bechtold, N. and Paszkowski, J. 1999. An SMC-like protein is required for efficient homologous recombination in *Arabidopsis*. EMBO J. 18: 4505–4512.

Miao, Z.-H. and Lam, E. 1995. Targeted disruption of the TGA3 locus in *Arabidopsis thaliana*. Plant J. 7: 359–365.

Offringa, R., de Groot, M.J.A., Haagsman, H.J., Does, M.P., van den Elzen, P.J.M. and Hooykaas, P.J.J. 1990. Extrachromosomal homologous recombination and gene targeting in plant cells after *Agrobacterium*-mediated transformation. EMBO J. 9: 3077–3084.

Offringa, R., Franke-van Dijk, M.E.I., de Groot, M.J.A., van den Elzen, P.J.M. and Hooykaas, P.J.J. 1993. Nonreciprocal homologous recombination between *Agrobacterium* transferred DNA and a plant chromosomal locus. Proc. Natl. Acad. Sci. USA 90: 7346–7350.

Oh, T.J. and May, G.D. 2001. Oligonucleotide-directed plant gene targeting. Curr. Opin. Biotechnol. 12: 169–172.

Ow, D. 2001. Applications of site-specific recombination. Plant Mol. Biol., this issue.

Paques, F. and Haber, J.E. 1999. Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. Microbiol. Mol. Biol. Rev. 63: 349–404.

Paszkowski, J., Baur, M., Bogucki, A. and Potrykus I. 1988. Gene targeting in plants. EMBO J. 7: 4021–4026.

Potrykus, I. and Spangenberg, G. 1995. Gene Transfer to Plants. Springer-Verlag, Berlin.

Puchta, H. 1998a. Towards targeted transformation in plants. Trends Plant Sci. 3: 77–78.

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

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

Puchta, H., Dujon, B. and 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. USA 93: 5055–5060.

Reiss, B., Klemm, M., Kosak, H. and Schell, J. 1996. RecA protein stimulates homologous recombination in plants. Proc. Natl. Acad. Sci. USA 93: 3094–3098.

Reiss, B. Schubert, I., Köpchen, K., Wendeler, E., Schell, J. and 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*. Proc. Natl. Acad. Sci. USA 97: 3358–3363.

Reski, R. 1998a. *Physcomitrella* and *Arabidopsis*: the David and Goliath of reverse genetics. Trends Plant Sci. 3: 209–210.

Reski, R. 1998b. Development, genetics and molecular biology of mosses. Bot. Acta 111: 1–15.

Reski, R. 1999. Molecular genetics of *Physcomitrella*. Planta 208: 301–309.

Reski, R., Reynolds, S., Wehe, M., Kleber-Janke, T. and Kruse, S. 1998. Moss (*Physcomitrella patens*) expressed sequence tags include several sequences which are novel for plants. Bot. Acta 111: 145–151.

Rice, M.C., May, G.D., Kipp, P.B., Parekh, H. and Kmiec, E.B. 2000. Genetic repair of mutations in plant cell-free extracts directed by specific chimeric oligonucleotides. Plant Physiol. 123: 427–438.

Rice, M.C., Czymmek, K. and Kmiec, E.B. 2001. The potential of nucleic acid repair in functional genomics. Nature Biotechnol. 19: 321–326.

Risseeuw, E., Offringa, R., Franke-van Dijk, M. E. and Hooykaas, P.J. 1995. Targeted recombination in plants using *Agrobacterium* coincides with additional rearrangements at the target locus. Plant J. 7: 109–119.

182

Risseeuw, E., Franke-van Dijk, M.E. and Hooykaas, P.J. 1997. Gene targeting and instability of *Agrobacterium* T-DNA loci in the plant genome. Plant J. 11: 717–728.

Rong, Y.S. and Golic, K.G. 2000. Gene targeting by homologous recombination in *Drosophila*. Science 288: 2013–2018.

Rong, Y.S. and Golic, K.G. 2001. A targeted gene knockout in *Drosophila*. Genetics 157: 1307–1312.

Rossi, L., Tinland, B. and Hohn, B. 1996. Role of the virulence protein of *Agrobacterium tumefaciens* in the plant cell. In: H. Spaink, P. Hooykaas and A. Kondorosi (Eds.) The Rhizobiaceae. Kluwer Academic Publishers, Dordrecht, Netherlands, pp. 303–320.

Rubin, E. and Levy, A.A. 1997. Abortive gap repair: the underlying mechanism for Ds elements formation. Mol. Cell Biol. 17: 6294–6302.

Schaefer, D.G. 2001. Gene targeting in *Physcomitrella patens.* Curr. Opin. Plant Biol. 4: 143–150.

Schaefer, D.G. and Zryd, J.P. 1997. Efficient gene targeting in the moss *Physcomitrella patens.* Plant J. 11: 1195–1206.

Shalev, G., Sitrit, Y., Avivi-Ragolski, N., Lichenstein, C. and Levy A.A. 1999. Stimulation of homologous recombination in plants by expression of the bacterial resolvase RuvC. Proc. Natl. Acad. Sci. USA 96: 7398–7402.

Strepp, R., Scholz., S., Kruse, S., Speth, V. and Reski R. 1998. Plant nuclear gene knockout reveals a role in plastid division for the homolog of the bacterial cell division protein FtsZ, an ancestral tubulin. Proc. Natl. Acad. Sci. USA 95: 4368–4373.

Swoboda, P., Gal, S., Hohn, B. and Puchta, H. 1994. Intrachromosomal homologous recombination in whole plants. EMBO J. 13: 484–489.

Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J. and Stahl, F.W. 1983. The double-strand break repair model of recombination. Cell 33: 25–35.

Thomas, K.R. and Capecchi, M.R. 1987. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell 51: 503–512.

Thykjaer, T., Finnemann, J., Schausser, L., Christensen, L., Poulsen, C. and Stougaard, J. 1997. Gene targeting approaches using positive-negative selection and large flanking regions. Plant Mol. Biol. 35: 523–530.

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

van der Steege, G., Schuilenga-Hut, P.H., Buys, C.H., Scheffer, H., Pas, H.H. and Jonkman M.F. 2001. Persistent failures in gene repair. Nature Biotechnol. 19: 305–306.

Vergunst, A.C. and Hooykaas, P.J.J. 1999. Recombination in the plant genome and its application in biotechnology. Crit. Rev. Plant. Sci. 18: 1–31.

Wang, W., Seki, M., Narita, Y., Sonoda. E., Takeda, S., Yamada, K., Masuko, T., Katada. T. and Enomoto, T. 2000. Possible association of BLM in decreasing DNA double strand breaks during DNA replication. EMBO J. 19: 3428–3435.

West, C.E., Waterworth, W.M., Jiang, Q. and Bray C.M. 2000. *Arabidopsis* DNA ligase IV is induced by gamma-irradiation and interacts with an *Arabidopsis* homologue of the double stand break repair protein XRCC4. Plant J. 24: 67–78.

Yamaguchi-Iwai, Y., Sonoda, E., Sasaki, M.S., Morrison, C., Haraguchi, T., Hiraoka, Y., Yamashita, Y.M., Yagi, T., Takata, M., Price, C., Kakazu, N. and Takeda, S. 1999. Mre11 is essential for the maintenance of chromosomal DNA in vertebrate cells. EMBO J. 18: 6619–6629.

Ye, S., Cole-Strauss, A.C., Frank, B. and Kmiec, E.B. 1998. Targeted gene correction: a new strategy for molecular medecine. Mol. Med. Today 4: 431–437.

Zhu, T., Peterson, D.J., Tagliani, L., St. Clair, G., Baszczynski, C. and Bowen B. 1999. Targeted manipulation of maize genes *in vivo* using chimeric RNA/DNA oligonucleotides. Proc. Natl. Acad. Sci. USA 96: 8768–8773.

Zhu, T., Mettenburg, K., Peterson, D.J., Tagliana, L. and Baszczynski, C.L. 2000. Engineering herbicide-resistant maize using chimeric RNA/DNA oligonucleotides. Nature Biotechnol. 18: 555–558.

Zupan, J., Muth, T.R., Draper, O. and Zambryski, P. 2000. The transfer of DNA from *Agrobacterium tumefaciens* into plants: a feast of fudamental insights. Plant J. 23: 11–28.