



## Recombinase-directed plant transformation for the post-genomic era

David W. Ow

Plant Gene Expression Center, Agricultural Research Service, United States Department of Agriculture, Albany, CA 94710, USA (\*author for correspondence; e-mail ow@pgec.ars.usda.gov) and Plant & Microbial Biology, University of California, Berkeley, CA 94720, USA

**Key words:** Cre-*lox*, DNA integration, FLP-*FRT*, phiC31, site-specific recombination

### Abstract

Plant genomics promises to accelerate genetic discoveries for plant improvements. Machine-driven technologies are ushering in gene structural and expressional data at an unprecedented rate. Potential bottlenecks in this crop improvement process are steps involving plant transformation. With few exceptions, genetic transformation is an obligatory final step by which useful traits are engineered into plants. In addition, transgenesis is most often needed to confirm gene function, after deductions made through comparative genomics, expression profiles, and mutation analysis. This article reviews the use of recombinase systems to deliver DNA more efficiently into the plant genome.

### Introduction

Since the pioneering transformation advances of the early 1980s, much of the research efforts have been directed, and rightly so, to a horizontal spread of the technology. As a result of this emphasis, it is now possible to transform a wide variety of plant species. The trade-off, however, has been less attention devoted to advancing the efficiency of the transformation process itself. Compared to many microbial systems, plant transformation appears somewhat antiquated. Whereas millions of independent transformants are routinely obtained with many microbial systems, in plants the numbers are generally in the single- to double-digit range. Hence a shotgun transformation approach to gene discovery is an option that has not been seriously entertained.

Gene transfer in many microbial systems also produces highly consistent phenotypes. Relatively few representative clones are needed for every construct analyzed. In plants, independent transformants show highly variable levels and patterns of expression. So, for a typical DNA construct, 20–50 independent primary transformants are needed. For the commercial development of a new trait, hundreds of independent

transformants are screened for the few with suitable transgene structure and expression.

The underlying reasons for the high variability in transgene expression are not completely understood, but at least four factors are involved in this phenomenon.

1. *Tissue culture*. Somaclonal variation has long been associated with tissue culture-regenerated plants. Changes in chromosome structure and ploidy, DNA sequence, DNA modification, and transposon activity have all been reported in somaclonal variants (Peschke and Phillips, 1992; Kaeppler *et al.*, 2000).
2. *Integration site*. Chromosomal structures such as telomeres or heterochromatin are known to affect the expression of nearby genes (Stavenhagen and Zakian, 1994; Howe *et al.*, 1995; Wallrath and Elgin, 1995). As a transgene integrates at random locations, chromosomal influences on transgene expression can be expected to differ among independent transformants (Meyer, 2000).
3. *Transgene redundancy*. Transformed plants often contain variable numbers of transgenes. Rarely is there a positive correlation between gene expression and copy number. On the contrary, many cases have linked extra full or partial transgene copies

to post-transcriptional and transcriptional gene silencing (Muskins *et al.*, 2000; Matzke *et al.*, 2000).

4. *Genetic mutations.* As expected for any genetic manipulation there is always the possibility of acquiring point mutations, deletions or rearrangements (Battacharyya *et al.*, 1994).

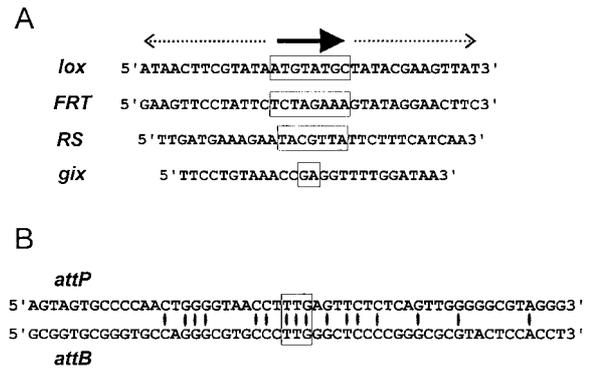
This article reviews the use of site-specific recombination in plant transformation, beginning with a brief overview of the recombination systems, followed by recombinase-based methods that seek to address some of the current limitations in plant transformation.

### Site-specific recombination systems

Many site-specific recombination systems have been described in prokaryotic and lower eukaryotic cells. The simplest recombination systems comprise two identical or nearly identical recombination sites and a single-polypeptide recombinase protein that catalyzes the recombination event (Sadowski, 1993). More complex recombination systems may have recombination sites that are dissimilar in sequence, different recombinases for the forward or backward reaction, and/or accessory protein(s) for recombination.

Site-specific recombination in plants has been reported for several of the simplest recombination systems (Ow and Medberry, 1995). These are the Cre-*lox* system from bacteriophage P1 (Dale and Ow, 1990, 1991; Odell *et al.*, 1990, 1994; Bayley *et al.*, 1992; Russell *et al.*, 1992), the FLP-*FRT* system of *Saccharomyces cerevisiae* (Lyznik *et al.*, 1993; Lloyd and Davis, 1994; Sonti *et al.*, 1995; Kilby *et al.*, 1995; Bar *et al.*, 1996), the R-*RS* system of *Zygosaccharomyces rouxii* (Onouchi *et al.*, 1991, 1995), and a modified Gin-*gix* system from bacteriophage Mu (Maeser and Kahmann, 1991). Cre, FLP, R, and Gin are the recombinases, and *lox*, *FRT*, *RS* and *gix* the respective recombination sites. In the case of the Gin-*gix* system, the stable transformation of the *gin* recombinase gene has not been reported as constitutive production of the Gin protein adversely affects plant regeneration.

The recombination systems above have in common the property that a single polypeptide recombinase catalyzes the recombination between two sites of identical or nearly identical sequences. Each recombination site consists of a short, 2 to 8 bp, asymmetric spacer sequence where strand exchange takes place, flanked by an inverted repeat of some 12 to 13 bp, where recombinases bind (Figure 1A). The asymme-



**Figure 1.** DNA sequence of recombination sites that operate in plants. A. The *lox*, *FRT*, *RS* and *gix* sites are structurally similar in that an asymmetric spacer sequence (boxed) is flanked by inverted repeats. B. The  $\phi$ C31 *attP* and *attB* sites share only 16 bp of identity in a 53 bp stretch centered at TTG (boxed), the recombination crossover point.

try of the spacer sequence gives an orientation to the recombination site, and dictates the outcome of a recombination reaction. Recombination between directly or indirectly oriented sites *in cis* excises or inverts the intervening DNA. Recombination between sites *in trans* causes a reciprocal translocation of two linear DNA molecules, or cointegration if at least one of the two molecules is circular. Since the product sites generated by recombination are themselves substrates for subsequent recombination, the reaction is freely reversible. In practice, however, excision is essentially irreversible because the probability of an intramolecular interaction, where the two recombination sites are closely linked, is much higher than an intermolecular interaction between unlinked sites. The corollary is that the DNA molecule inserted into a genomic recombination site will readily excise out.

### Resolution-based single-copy transgenic lines

Current methods in plant gene transfer often produce a complex integration structure at the insertion locus. Typically, multiple full and/or partial copies of the introduced molecule are arranged as direct and/or indirect repeats. These complex patterns are not necessarily an impediment to research, but they are not desirable for commercial use. Structural documentation is a prerequisite for regulatory approval and a simple integration pattern is easier to characterize. Repetitive DNA also tends to be structurally and functionally unstable. Repeat sequences can participate in intra- and inter-chromosomal recombination. Even if a

complex integration locus yields a suitable phenotype, it may be difficult to maintain the original structure, along with its defined expression pattern, through the numerous crosses involved in breeding and seed production programs. Multiple gene copies, particularly if some are arranged as indirect repeats, are frequently associated with homology-dependent gene silencing (Iyer *et al.*, 2000; Muskens *et al.*, 2000).

To obtain transgenic plants that harbor a single intact copy of the introduced DNA, the conventional approach has been to screen molecularly for the rare individuals among a large pool of transformants. With fewer trait genes at hand, conducting this exercise may not seem onerous. With orders of magnitude more provided by plant genomics data, such screening efforts would be much less manageable. Additionally, the screening for single-copy plants can be feasible only for crop varieties where transformants are relatively easy to obtain. For varieties that are difficult to transform, accumulating a sizable pool of independent transformants can be resource-intensive. The class of plants with low transformation efficiency includes not only certain species, but also many of the commercial cultivars of species where their laboratory counterparts are considered transformation-proficient.

Genetic transformation of many agronomically important crop plants is technically challenging. In the biolistic transformation of wheat, for example, thousands of dissected embryos are bombarded with DNA and cultured into embryogenic calluses, but only a few plants are obtained that yield transgenic progeny. Since single-copy insertions occur at a frequency of 5% or less, some 400 000 dissected embryos are needed to obtain 20 independent single-copy transformants. This amount of effort, through tissue culture, greenhouse propagation and molecular analyses, may be practical only when a few genes are being engineered. When hundreds of genes are made available through plant genomics research, even a ten-fold scale-up of resources will be inadequate.

Recombinase-mediated excision of redundant copies has provided a solution that obviates the need to generate and screen a large pool of primary transformants. Site-specific recombination can convert a complex multiple-copy integration pattern into a single-copy state. The conversion step saves human and capital resources that would otherwise be employed in the generation and characterization of an additional 5- to 20-fold greater number of plants.

#### *Subsequent resolution of complex integration patterns*

The initial version of a resolution-based strategy relies on the introduction of a recombinase gene by a genetic cross (Srivastava *et al.*, 1999). The transformation vector is designed to reduce, by site-specific excision, multiple copies of a gene construct to a single unit. At the same time, it can remove the selectable marker gene. Inversely oriented recombination sites flank the gene construct, while another set of directly oriented recombination sites flank the marker gene. For this work, the Cre-*lox* system was used, although the same principle applies with other recombination systems such as FLP-*FRT* and R-*RS*.

The transformation vector pVS11 contains a *bar* gene encoding resistance to bialaphos and which is flanked by two wild-type *lox* sites (*loxP*) in same orientation (Figure 2). The *bar* gene is expressed under the control of the rice actin gene promoter (designated as *P1*). The *bar* gene is followed by a maize ubiquitin gene promoter (designated as *P2*). The entire fragment is flanked by two oppositely oriented *lox511* sites. The *lox511* allele is a mutant variant of the wild-type *loxP* sequence with a base change within the 8 bp spacer (Hoess *et al.*, 1986). The two heterospecific alleles do not recombine with one another efficiently. Only a 1% efficiency of recombination was observed for a *loxP* × *lox511* reaction *in vitro* (Hoess *et al.*, 1986). This makes the *loxP* × *lox511* reaction unlikely unless *cre* expression is sufficiently strong (see section on heterospecific sites). Thus, regardless of the number of copies integrated between the outermost transgenes, and the relative orientation of the outermost pVS11 fragments, recombination between the outermost *lox511* sites should resolve multiple units into a single copy, and the recombination between the *loxP* alleles should delete the *bar* marker.

From the bombardment of several-thousand dissected embryos of wheat, four transgenic lines were obtained. Each of the four transgenic wheat lines initially contained multiple copies of the introduced DNA, but at a single integration locus. After *cre* was introduced through a genetic cross, F<sub>1</sub> progeny was obtained that had resolved the multiple copies to a single-copy state. As designed, the *bar* transformation marker was also removed to produce the *lox511-P1-loxP-P2-lox511* configuration. The *cre* locus was subsequently segregated away in the F<sub>2</sub> population.

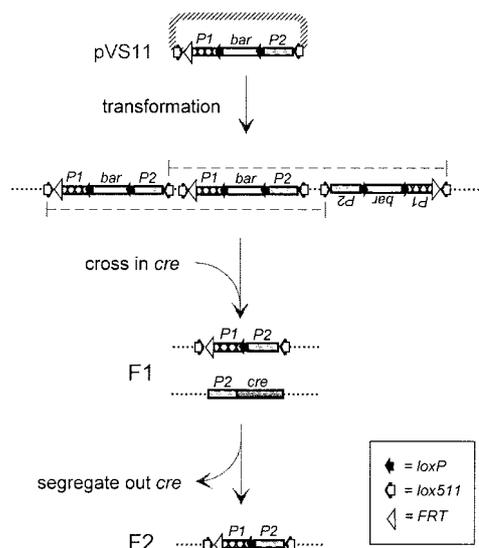


Figure 2. Resolution of complex integration pattern. A relevant segment of transformation molecule is flanked by *lox511* sites of opposite orientation. The *cre* locus is introduced through a genetic cross. Recombination between outermost sites of the same orientation (indicated by dash lines) deletes interior copies. The *bar* marker is excised by flanking *loxP* sites of the same orientation. The *cre* locus is segregated away by the F<sub>2</sub> generation. For simplicity, the terminator element following each coding region is not shown in this and other figures. Abbreviations: *P1*, rice actin promoter; *P2*, maize ubiquitin promoter; *bar*, coding region for bialaphos (basta) resistance.

Immediate resolution via co-introduction of recombinase

The strategy above is generally applicable for most plant transformation systems. In some instances, however, an alternative strategy may be necessary. For instance, the need to cross in and out the recombination gene may not be compatible with breeding programs that primarily use vegetative propagation methods. Additionally, crossing in and segregating out the recombinase gene requires additional generation time. Hence, an alternative version of the resolution strategy has been described (Srivastava and Ow, 2001a).

This variation of the resolution-based strategy relies on the co-introduction of the recombinase gene. The transient synthesis of recombinase is used to catalyze the recombination between the outermost recombination sites. Whether this happens before and/or after DNA integration, it should enrich the number of single-copy insertions in primary transformants. The fate of the co-introduced recombinase gene is an important consideration. Since it is not selected for, some

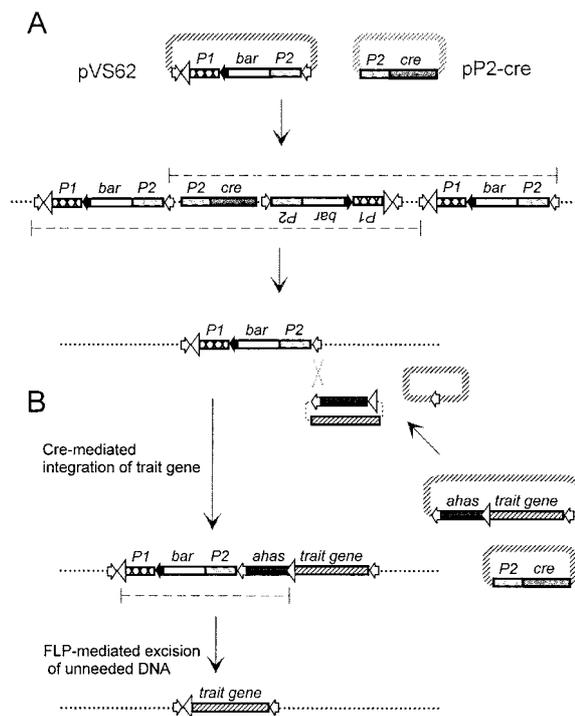


Figure 3. Single-copy target for site-specific integration and subsequent removal of extraneous DNA. A. Co-introduced *cre* transiently provides recombinase to resolve complex integration patterns, including *cre* DNA that lies within the outermost *lox511* sites. Recombination between outermost sites indicated by dash lines. B. Hypothetical strategy to target a trait gene into the chromosomal target through a *lox511* × *lox511* reaction, followed by *FRT* × *FRT* recombination to remove unneeded DNA. Abbreviations not in Figure 2: *ahas*, acetohydroxyacid synthase (imazethapyr resistance) coding region.

plants may not retain this DNA. In practice, however, when two different DNA molecules are co-introduced through direct DNA transformation methods, both molecule types integrate more often than not into the same locus. Given this expectation, the recombinase gene is likely to be among the cluster of introduced DNA. However, whether the recombinase gene will be retained in the genome will depend on whether one or more copies lie outside the outermost recombination sites (Figure 3A). Copies of DNA that lie within the outermost recombination sites will be deleted.

This transient recombinase expression approach was tested in maize (Srivastava and Ow, 2001a). Since the presence of a selectable marker is needed to detect the transformation event, the *loxP* site downstream of *bar* was removed from pVS11 to generate pVS62. This modification permits maintenance of the *bar* gene. If the desired final product were a

marker-free transgenic plant, additional downstream manipulations would be needed (see section on marker removal).

Maize embryogenic calluses were bombarded with a combination of pVS62 and pP2-cre DNA. Primary transformants regenerated from bialaphos-resistant calluses were analyzed by Southern hybridization. Out of 13 plants, 5 were single-copy lines. Re-probing of the blots showed that 2 of these 5 plants lacked *cre* DNA. In these two plants, the *cre* DNA either had not integrated at all into the genome, or had integrated only within the outermost *lox511* sites, and was subsequently deleted.

There are two considerations worth noting with a resolution-based strategy. The first concerns the frequency of gene copies interspersed with host DNA, as opposed to being in a contiguous tandem array. Two recent articles have described non-transgenic DNA interspersed among transgene copies in plants (Kohli *et al.*, 1998; Pawlowski and Somers, 1998), but the origin of the nontransgenic DNA is not known. Ectopic gene targeting may be a possibility. This is the situation where the introduced DNA strand invades the homologous locus and replicates a segment of DNA from that locus. The transgene linked with the copied host DNA then integrates at a different site (Offringa *et al.*, 1993; Risseeuw *et al.*, 1995; Dellaire *et al.*, 1997). If there were interspersed DNA, and if that DNA were derived from elsewhere in the genome, its removal along with the extra transgene copies would most likely be inconsequential. If the nontransgenic DNA were native to the locus, however, the resolution process would generate a chromosomal deletion. In the light of the latter possibility, for practical implementation of this strategy, it would be prudent to obtain several independent single-copy lines.

The second consideration is that since the DNA is flanked by a set of inverted recombination sites, the final transgene configuration could be in either orientation within the integration locus. Hence, two types of transgenic lines, differing in orientation and possibly also in expression, could be produced at a given integration locus.

It is worth noting that a recent article described a two-copy transgenic locus, where the copies are arranged as an inverted repeat (De Buck *et al.*, 2001). This locus was poorly expressed, and was able to trans-silence homologous genes. *Cre-lox* site-specific recombination was used to remove one of the two copies. The resulting single-copy locus became highly expressed and lost the ability to cause gene silencing.

### Site-specific integration

The recombinase-based resolution of complex integration structures bypasses the need for extensive screening for single-copy transgenic lines. However, the method neither controls the site of insertion nor insures that the single intact copy transgene is completely free of repetitious DNA, as small amounts of transgenic DNA may lie outside of the outermost recombination sites. To deliver a DNA molecule more precisely would be through site-specific integration. This approach requires the prior construction of a genomic target with the appropriate recombination sites in place.

For the moment, whether it is necessary or desirable to deliver a transgene to a known location is not at all clear. There is insufficient knowledge on what chromosome locations are best for the expression of a DNA construct. Moreover, in the absence of effective homologous integration (Vergunst and Hooykaas, 1999; Puchta, this issue), it is not possible to create chromosome targets through design. Instead, targets can only be generated through prior analysis of a collection of randomly placed genomic recombination sites. For immediate applications, engineering a trait gene through site-specific integration would be more laborious than screening for a collection of transgene insertions for effective expression. Given that the recombinase-based resolution strategy could reduce 90% of the effort needed to obtain single-copy plants, this alone could make the transformation process much more manageable.

On the other hand, there are some advantages to being able to direct the DNA into chosen locations. There is evidence supporting a position effect on transgene expression. Once a genomic target is shown to confer suitable expression, subsequent deliveries can be made to that site. Additionally, site-specific integration is a precise process with respect to transgene structural fidelity. From a scientific standpoint, precise integration may yield a higher percentage of stable expression. From a public relations point of view, greater precision in genetic transformation may be more comforting to end-users. Finally, in conjunction with promoter-trap strategies, it is possible to generate a collection of randomly placed target sites that lie downstream of endogenous plant promoters. As such, the subsequent delivery of new DNA into the chromosome target will place the new trait gene under the control of host elements. Although this hypothesis

remains to be tested, one expectation may be a more faithful expression pattern for the introduced trait.

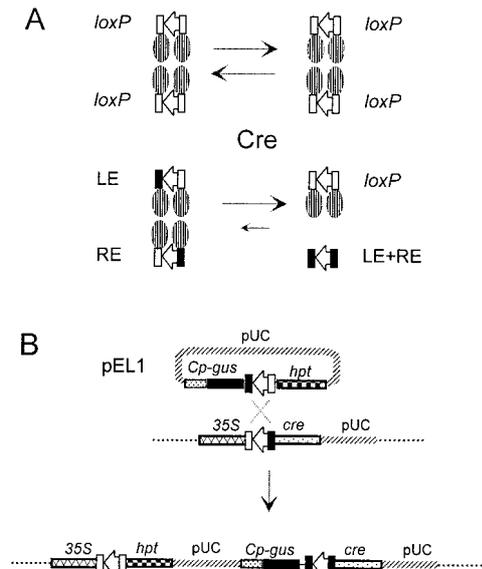
#### Targeting by cointegration

Site-specific integration can proceed through a cointegration or a DNA exchange reaction. The product of a cointegration reaction contains two recombination sites of the same orientation and subsequent recombination between these two sites can resolve the cointegrate. Since Cre-*lox* and similar systems catalyze freely reversible reactions, intramolecular excision is kinetically favored over intermolecular cointegration. For the integration of a circular DNA substrate into a linear chromosome, this would mean that recovery of insertion events would be inherently difficult. Hence, Albert *et al.* (1995) used two approaches to help stabilize DNA integration. The first relies on generating new recombination product sites that are inefficiently recognized by the recombinase. The second relies on abolishing recombinase activity after DNA integration.

#### Less recombinogenic product site

Recombination sites can be designed to favor insertion over excision. Cre and FLP occupy their respective sites as a dimer, with each monomer situated at each of the 13 bp inverted repeats that flank the asymmetric 8 bp spacer. These binding elements can be mutated to reduce recombinase recognition. Since the two monomers bind cooperatively, a mutation in only one element has less of an effect on dimer occupation of the site. Hence, recombination can occur between a left element (LE) mutant and a right element (RE) mutant (Figure 4A). The reaction regenerates a wild-type site, and a new left and right element (LE+RE) mutant site. Mutations on both elements can substantially reduce recombinase recognition. Therefore, as one of the product sites is no longer efficiently recognized by the recombinase, the reaction is much less reversible. With the FLP-*FRT* system, Huang *et al.* (1991) tested a pair of LE and RE sites in *Escherichia coli*. This mutant pair's LE  $\times$  RE forward reaction is at least 5-fold less efficient than that of wild-type sites. Consequently, despite a more stable integration event, the less efficient forward reaction produced 100-fold fewer integrants than did the use of wild-type sites.

An analogous strategy was developed for the Cre-*lox* system (Albert *et al.*, 1995). An *in vitro* analysis of three mutant pairs, *lox66*  $\times$  *lox71*, *lox76*  $\times$  *lox75*, and *lox43*  $\times$  *lox44*, showed that the greatest difference



**Figure 4.** Site-specific integration using *lox* sites mutated in recombinase binding elements. **A.** The *lox* site is depicted by rectangles (13 bp inverted repeats) flanking an open arrowhead (8 bp asymmetric spacer). Filled rectangles represent mutant sequences with reduced recognition by Cre. Unlike *loxP*  $\times$  *loxP* reactions, a LE  $\times$  RE reaction yields a LE+RE double mutant site that is poorly recognized by the recombinase (shaded ovals). **B.** DNA integration in tobacco using mutant *lox* sites in conjunction with disruption of *cre* expression. Recombination indicated by dotted cross. Abbreviations: 35S, CaMV 35S RNA promoter; *hpt*, hygromycin resistance coding region; *Cp-gus*, commelina yellow mottle virus promoter fused to  $\beta$ -glucuronidase coding region.

in the forward (LE  $\times$  RE mutants) over the reverse reaction (*loxP*  $\times$  LE+RE mutant) occurs at low Cre concentration. At high Cre concentration, a difference between the forward and reverse reaction was not seen. This indicates that a site with reduced affinity for Cre recognition can be compensated by a high enzyme concentration (see section on heterospecific sites). Compared to *loxP*  $\times$  *loxP*, the three mutant pairs are also less efficient in the forward reaction. A titration of Cre enzyme required for detectable recombination was observed at 0.001, 0.002, 0.004 and 0.008 units, respectively, for the reactions *loxP*  $\times$  *loxP*, *lox66*  $\times$  *lox71*, *lox76*  $\times$  *lox75*, and *lox43*  $\times$  *lox44*.

#### Eliminating post-integration recombinase activity

The transient provision of the recombinase is one solution that was first tested in mammalian systems (Sauer and Henderson, 1990; O'Gorman *et al.*, 1991; Baubonis and Sauer, 1993). The recombinase DNA, mRNA, or protein can be introduced for a brief period during

the transformation process. As the cell depletes the supply of recombinase, the rare molecules that remain inserted at the target site would be stable. This was one strategy adopted by Albert *et al.* (1995). Plant lines were generated with a *35S-lox-luc* target, where *35S* is the CaMV 35S RNA promoter and *luc* is the firefly luciferase gene. Through co-transformation with a *35S-cre* plasmid, the production of Cre catalyzed the insertion of a *lox-hpt* plasmid, where *hpt* is a promoterless hygromycin phosphotransferase coding region. The newly fused *35S-lox-hpt* linkage conferred hygromycin resistance.

A second strategy was the disruption of recombinase gene expression upon site-specific integration into the genomic target. Cre was produced from a *35S-lox-cre* construct placed into the plant genome. Upon introduction of a circular *lox-hpt* substrate, site-specific integration displaced the *cre*-coding region from its promoter to terminate *cre* transcription (Figure 4B).

#### *Polyethylene glycol-mediated targeting in tobacco*

Albert *et al.* (1995) tested the insertion of circular DNA at 8 different target sites in the tobacco genome. Wild-type and mutant *lox* sites were compared in conjunction with the two strategies to eliminate post-integration recombinase activity. Polyethylene glycol-treated protoplasts containing a genomic *35S-lox* target were transformed with a test *lox-hpt* plasmid. Since random integration of the *lox-hpt* plasmid would not be detected unless it integrated behind a genomic promoter, the frequency of random integration was deduced from a parallel transformation with a *35S-hpt* construct. To confirm that the *lox-hpt* plasmid integrated at the *35S-lox* target, molecular analysis was conducted on representative calluses and regenerated plants. Only the *lox66 × lox71* and the *lox76 × lox75* pairs were tested, since the *in vitro* data on *lox43 × lox44* suggested a forward reaction that might be too inefficient to yield a high integration rate.

Considerable variation in transformation efficiencies was found. In particular, there was a tendency for a lack of site-specific integration with protoplast preparations that showed low transformation competency. For example, among 6 attempts on two *35S-lox66-cre* lines, two of them failed to produce site-specific events. In these two experiments, the transformation efficiency of the *35S-hpt* plasmid averaged only  $3.5 \times 10^{-5}$ , an order of magnitude below expectation. With the other 4 experiments, the *35S-hpt*

plasmid had an average transformation efficiency of  $3.6 \times 10^{-4}$ . In these same experiments, site-specific recombination by the *lox71-hpt* plasmid averaged  $2.5 \times 10^{-4}$ . Likewise, for 4 out of 6 experiments from two different *35S-lox76-cre* lines, the transformation efficiencies of the *35S-hpt* plasmid and the *lox75-hpt* plasmid averaged  $5.1 \times 10^{-4}$  and  $3.8 \times 10^{-4}$ , respectively. In contrast, from a total of 8 attempts, the transformation of two *35S-loxP-cre* lines by a *loxP-hpt* plasmid failed to produce a single integrant, whereas the control *35S-hpt* plasmid averaged  $1.3 \times 10^{-4}$ . Thus, for this integration strategy, there was a clear advantage in mutant over wild-type sites.

In the transient expression approach, however, the mutant sites were only slightly better. Counting only experiments where both the *lox75-hpt* plasmid and the *35S-hpt* yielded transformants (3 of 7 experiments), the average transformation efficiency obtained from a single *35S-lox76-luc* line by either the *lox75-hpt* plasmid or the *35S-hpt* plasmid was  $1.6 \times 10^{-4}$  or  $1.4 \times 10^{-4}$ , respectively. The average transformation efficiency, from 3 out of 7 experiments, on a single *35S-loxP-luc* line by the *loxP-hpt* plasmid or the *35S-hpt* plasmid was  $7.9 \times 10^{-5}$  or  $2.9 \times 10^{-4}$ , respectively.

The difference in *loxP × loxP* integration with the two experimental approaches can be attributed to the relative abundance of Cre. In the promoter displacement strategy, Cre is readily available within each cell. As a *loxP × loxP* excision occurs more readily than a *loxP × LE+RE* excision, this could account for a less stable *loxP × loxP* integration. With the transient expression approach, each cell within a transformed population experiences a different level of Cre concentration depending on the amount of *cre* DNA taken up. Those with a rapid depletion of the recombinase will more likely trap-integrated molecules.

In mouse embryonic stem cells, Araki *et al.* (1997) tested one of the mutant pairs for targeted integration. The test plasmid was designed to distinguish between site-specific and random integration. The mean value from 7 data points from *loxP × loxP* site-specific integration was 0.2% of random integration. In contrast, the mean value from 16 data points for *lox61 × lox71* site-specific integration was 4.7% of random integration.

#### *Biostic-mediated targeting in rice*

A drawback with protoplast transformation is that it is a cumbersome procedure for many crop plants. A

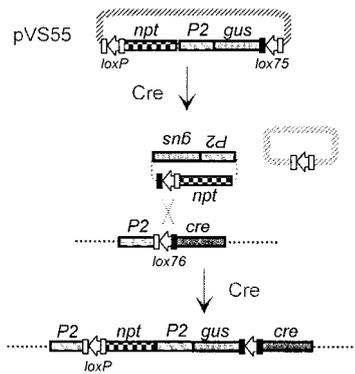


Figure 5. Site-specific integration in rice. The transformation vector is resolved by Cre into two molecules, one of which then integrates into the rice target to generate a *P2-npt* transcription unit conferring resistance to geneticin.

more convenient method of DNA delivery is through microparticle bombardment. A recent study described site-specific integration in rice (Srivastava and Ow, in press). The most effective strategy found in the earlier report was used (Albert *et al.*, 1995), combining both the disruption of *cre* transcription along with a pair of *lox75* and *lox76* alleles. Two target cell lines were generated harboring a *P2-lox76-cre* construct (Figure 5). Calluses were biolistically transformed with pVS55, a plasmid with a promoterless neomycin phosphotransferase coding region (*npt*) and a  $\beta$ -glucuronidase coding region (*gus*) under the control of *P2*. This *npt-P2-gus* fragment is flanked by a *loxP* upstream of *npt* and a mutant *lox75* allele downstream of *gus*. Cre-mediated intra-molecular recombination between *loxP* and *lox75* separates pVS55 into two molecules: a circular plasmid backbone with a *loxP* site, and a circular fragment consisting of *lox75-npt-P2-gus*. Insertion of the latter molecule into the genomic *lox76* target forms a *P2-loxP-npt* linkage to confer geneticin resistance. As before, *cre* transcription terminates through displacement of *cre* from its promoter, resulting in a *gus-lox75/76-cre* junction.

From the PCR analysis of 36 geneticin resistant calluses collected from the 2 target cell lines, the selected *P2-npt* junction was found in every line. The unselected *gus-cre* junction was also found in all but three clones. These three failed to express *gus*, possibly due to the DNA acquiring a mutation, or was silenced. However, a PCR product corresponding to the *P2-cre* target junction was detected in every callus line, indicating that the geneticin-resistant clones were chimeric, with target cells growing along integrant cells. Due to the chimeric calluses, further

analysis was confined to regenerated plants that express the *gus* gene. Only three GUS<sup>+</sup> integrant plants regenerated, and hybridization analysis indicated site-specific integration in each case. However, only one of the three plants showed the site-specifically integrated molecule as the sole copy in the genome. The other two plants harbored additional molecules that are most likely the result of random integration elsewhere in the genome. Moreover, one of the two plants also retained the plasmid backbone that was resolved from the intramolecular recombination of pVS55.

#### Agrobacterium-mediated targeting in Arabidopsis

Although direct DNA delivery methods are effective in a wide range of crop species, many plants are most easily transformed via *Agrobacterium*. Recombinase-catalyzed reactions require a double-stranded DNA substrate. Therefore, the T-DNA must first acquire second-strand synthesis before it can be a substrate for recombination. It is not clear whether the T-DNA must first integrate into the plant genome before a double-stranded form emerges. After all, the initial burst of high-level expression by T-DNA encoded transgenes is thought to arise from nonintegrated, presumably duplex T-DNA templates.

*Agrobacterium*-mediated delivery was explored with the Cre-*lox* system in *Arabidopsis*. In one attempt, the *Arabidopsis* genome was transgenic for a promoterless *lox-npt* construct (Vergunst and Hooykaas, 1998). Two insertion vectors were used, one with a *35S-lox* fragment, and a second with a *lox-35S-lox* fragment. The Cre recombinase was provided through *cre* transient expression by a co-transforming vector. Site-specific insertion into the chromosomal target would confer kanamycin resistance from formation of a *35S-lox-npt* linkage. One precise integration event was obtained by the *lox-35S-lox* construct, none by the *35S-lox* construct. This suggests that a circular *35S-lox* molecule, excised from the *lox-35S-lox* construct, was the integration substrate. In another attempt, a *35S-lox-cre* construct was placed into the *Arabidopsis* genome and was targeted by a *lox-npt-lox* construct (Vergunst *et al.*, 1998). A *lox-npt* circular molecule would be produced by site-specific recombination prior to insertion into the genomic target. With this strategy, kanamycin-resistant clones appeared at a frequency of 1–2% of the frequency of random integration.

### Site-specific insertion of large DNA molecules

Large DNA libraries are useful for mapping and isolating novel genes. Methods that facilitate transformation with large DNA should help with complementation analysis. It may even help clone quantitative trait genes, if they map within a particular YAC or BAC clone. Choi *et al.* (2000) tested the delivery of BAC clones into the tobacco genome via Cre-*lox* site-specific integration. A *lox-hpt* fragment was placed into a BAC vector for cointegration at a genomic *35S-lox-cre* target. Three BAC clones, containing 50, 100 or 150 kb of cotton DNA, were biolistically transformed into a *35S-lox-cre* tobacco line. Between 2 and 14 independently derived hygromycin-resistant plants were obtained from each BAC clone. PCR and Southern analyses showed that these plants have the correct recombination junctions. Unfortunately, it proved difficult to determine whether the large inserts were intact. The cotton DNA cross-hybridized extensively to tobacco DNA, making such determination impractical.

### Targeting via DNA exchange

Targeted integration with a pair heterospecific recombination sites has been reported in mammalian cells (Bethke and Sauer, 1997; Seibler and Bode, 1997; Feng *et al.*, 1999). The experiment of Bethke and Sauer (1997) is a typical example of this strategy. A mouse cell line was made transgenic for a promoter-*loxP-lacZ-lox511* construct (Figure 6). This cell line was then co-transfected by a *cre*-expressing plasmid and a second plasmid containing a *loxP-neo-lox511* fragment. In both constructs, *loxP* and *lox511* were in an inverted orientation, which insures that the intervening DNA would not be deleted through the possible but unlikely *loxP* × *lox511* reaction (see section on heterospecific sites). Cre-mediated crossovers, *loxP* × *loxP* and *lox511* × *lox511*, exchanged the genomic *lacZ* with *neo* DNA. This replacement strategy yielded integration events 2- to 3-fold above the rate of random integration, and about 15- to 20-fold above the rate for a comparable *loxP* × *loxP* cointegration strategy. This proficiency was attributed to a reduction of the reverse reaction.

In plants, this approach has not been reported in the refereed literature, but has been described in a patent (Baszczynski *et al.*, 2001b). Like the Cre-*lox* system, nucleotide changes within the 8 bp spacer can create new recognition sites for FLP (Seibler and Bode, 1997). In this patent, maize cell lines with a

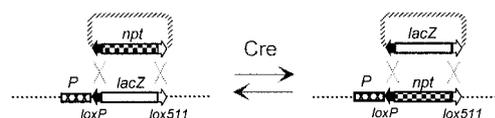


Figure 6. DNA replacement in mouse cells. Genomic *lacZ* ( $\beta$ -galactosidase) fragment is replaced by plasmid *npt* fragment in a dual-recombination reaction. The *P-npt* linkage confers neomycin resistance.

genomic target set of heterospecific *FRT* sites were re-transformed by a promoterless *gus* construct flanked by the same set of heterospecific *FRT* sites. *GUS*<sup>+</sup> calluses were recovered for 18 target lines, at very high frequencies that range from 2% to 67% of the total transformation frequency, as scored by another marker present on the construct. This high efficiency might be attributed to an improved FLP recombinase, as a separate patent described codon changes in the *FLP* gene that enhanced its recombination activity in maize (Baszczynski *et al.*, 2001a). On the other hand, since the patent lacks a description of the *GUS*<sup>+</sup> frequency obtained in the absence of a co-transformed *FLP* construct, the rate of FLP-independent random promoter-trap events is not known. Also missing are data on the downstream junction, without which it cannot be determined what relative proportion of the events is replacement and not cointegration.

### Recombination between heterospecific sites

A word of caution should be noted on the use of heterospecific recombination alleles. An example is the *loxP* × *lox511* reaction. Although it has been described to have an efficiency of 1% *in vitro* (Hoess *et al.*, 1986), this does not necessarily mean that the reaction could not be more efficient with a higher recombinase concentration. In targeting via DNA exchange, Feng *et al.* (1999) observed *loxP* × *lox511* recombination. It is possible that the burst of *cre* expression from a co-transfected construct was sufficiently high, at least in a sizable portion of the transfected cells. In *Arabidopsis*, Vergunst and Hooykaas have recently detected *loxP* × *lox511* recombination in a construct that also expresses *cre* (personal communication).

On the other hand, other reports have not observed recombination between heterospecific sites. Seiber and Bode (1997) and Baszczynski *et al.* (2001b) used pairs of directly oriented heterospecific *FRT* sites for their DNA exchange constructs. Directly oriented *loxP* and *lox511* sites were used in the DNA resolution experiments in wheat and maize (Srivastava *et al.*, 1999; Srivastava and Ow, 2001). In wheat, Southern blotting

data did not detect *loxP* × *lox511* recombination during the F<sub>1</sub> generation that *cre* was introduced through a cross. With maize, a *loxP* × *lox511* reaction would not have been detected, as the *bar* selectable marker would have been lost (Figure 3).

### Site-specific transgene expression

Day *et al.* (2000) sought to address whether expression of a transgene at a target site would be similar among independently transformed lines, and whether expression would differ among different chromosome locations. The targeting strategy was analogous to that described earlier (Albert *et al.*, 1995) except that the revised *lox-hpt* plasmid, pEL1, carried a *Cp-gus* fragment, where *Cp* is the commelina yellow mottle virus promoter with activity in vascular and reproductive tissue (Figure 4B). Selection for active expression was imposed on *35S-lox-hpt* but not on the linked *Cp-gus* reporter. Several independently derived integrant plants from each of the 5 sites, a total of 22 lines, were shown to contain the correct single-copy integration structure.

#### *Gene imprinting from site-specific integration*

When examined for *Cp-gus* expression, approximately half of the integrant lines in each of 4 target sites showed the expected pattern of expression, namely expression throughout vascular tissues. The remaining integrant lines, however, showed various degrees of gene silencing. The GUS staining pattern of the F<sub>1</sub> plants showed sharp boundaries that define blue sectors, suggesting a clonal origin. The expression patterns were maintained through plant development, and were transmitted to subsequent generations. Thus, this implicated an imprinting phenomenon that was stable and heritable. Both PCR and Southern analysis concluded that the integrated DNA remained in place. Hence, inactivity of the *Cp-gus* transgene was not due to excision of pEL1 from the target site.

Methylation-sensitive and -insensitive enzymes were used to probe the extent of DNA methylation in active and silenced *Cp-gus* transgenes at the different loci. From this analysis, a difference in degree of DNA methylation was found. For the lines that showed full expression in vascular tissues, there were fewer sites methylated. For the silenced lines, however, hypermethylation was observed. Most surprising was the finding that the DNA outside of the newly integrated

transgene did not show a change in the methylation pattern. It was as though the ‘old’ DNA was fixed into a certain state of methylation, whereas different degrees of imprinting were imposed on each ‘new’ DNA.

The mechanism for this imprinting is not known. However, amidst the current models of gene silencing, the RNA-directed DNA methylation model (Wassenegger *et al.*, 2000) seems most plausible for this phenomenon. During DNA uptake by competent cells, aberrant or duplex RNAs may be produced by transient expression of the transformed DNA, especially from rearranged and/or concatameric extrachromosomal molecules. Although, in each case, a single molecule integrated site-specifically, the amount of initial DNA uptake can vary considerably among individual recipient cells. Those with the highest initial copies may therefore be more prone to an RNA-directed DNA methylation effect. In this model, the imprinting may occur prior to or soon after DNA integration. Current efforts are aimed at minimizing this transformation-associated imprinting effect through inhibiting cellular transcription during the initial period following DNA uptake.

The biolistic targeting in rice followed a similar pattern (Srivastava and Ow, in press). Among the 36 gentamicin-resistant calluses collected from the 2 target cell lines, 17 of them tested negative for *gus* expression. Only three of these lacked a correct downstream junction to suggest imprecise integration. The rest appeared to be correct. In light of the results of Day *et al.* (2000), hypermethylation-associated gene silencing seems to be a probable cause.

#### *Position effect on gene expression*

The second finding from the Day *et al.* (2000) study is derived from a comparison of the full-expression integrant lines. These plants all showed the expected vascular-specific expression pattern. The intensity of expression was similar among independent lines of the same site, but dissimilar among different integration sites. For example, the level of GUS activity was ca.10-fold higher at one integration site than at another, although both sites could give rise to the same expression pattern. This supports the hypothesis that the chromosome location can exert an effect on the expression of the integrated transgene.

### Dedicated integration system

The recombination systems described above catalyze freely reversible reactions. This is not ideal for integrating DNA as the inserted molecule will be more prone to excise or exchange out of the genome than to remain in place. More complex recombination systems can control the reversibility of the reactions. One such system is from *Streptomyces* phage  $\phi$ C31 (Thorpe and Smith, 1998). A 68 kDa single-polypeptide integrase, encoded by the *int* gene, catalyzes the recombination between the bacterial and phage attachment sites, known as *attB* and *attP* respectively. The  $\phi$ C31 integration reaction, *attB*  $\times$  *attP*, is not reversible unless an additional phage excisionase protein is provided. The phage and bacterial *att* sites share only three base pairs of homology at the point of cross-over, and a total of 16 bp of identity within 50 bp (Figure 1).

The  $\phi$ C31 system has been tested in fission yeast, mammalian cells and transgenic plants. In the fission yeast *Schizosaccharomyces pombe*, Thomason *et al.* (2001) placed a  $\phi$ C31 *attP* (84 bp) site into the *leu1* locus of chromosome II (Figure 7A). This strain was transformed with a non-replicating plasmid pLT45, which harbors *ura4<sup>+</sup>* and an *attB* (280 bp) sequence. When pLT45 was introduced by itself, Ura<sup>+</sup> colonies appeared from random integration of the DNA at a frequency of  $6.3 \times 10^{-6}$ . When introduced along with pLT43 that produced  $\phi$ C31 integrase, the number of Ura<sup>+</sup> transformants increased up to 22-fold at optimal integrase concentration, and this increase was due mainly (94%) to site-specific recombination. Southern analysis showed that 88% of the transformants were perfect integration events, 6% were correct integrants, but with additional random integrations, and only 6% showed an integration pattern that could not be attributed to site-specific integration. For two site-specific integration events, sequence analysis showed precise *attP*  $\times$  *attB* recombination. Relative to the transformation of a replicating plasmid, site-specific insertion into the genome was ca. 20% as efficient at optimal integrase concentration.

An exchange reaction was also examined in a strain that harbors two *attP* sites in direct orientation (Figure 7B). The replacement substrate was either pLT50, or a linear *attB-ura4<sup>+</sup>-attB* fragment purified or synthesized by PCR from pLT50. The recombination between *attP* and *attB* sites replaced the *attP*-flanked genomic DNA with the *attB*-flanked *ura4<sup>+</sup>* marker. This exchange reaction yielded a much greater num-

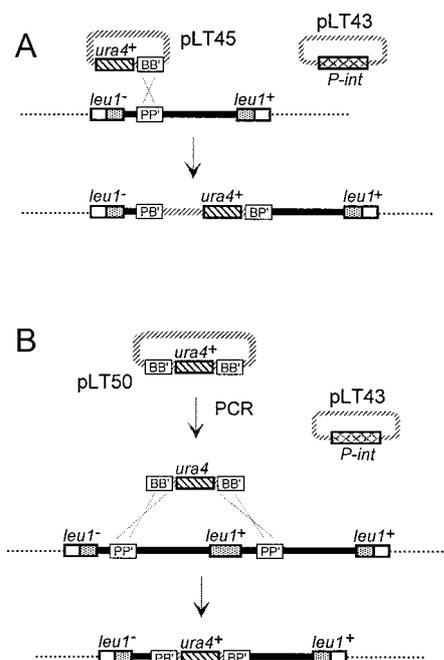


Figure 7. Insertion and replacement reactions in fission yeast. A. Integrase-promoted recombination between *attB* and *attP* (depicted as *BB'* and *PP'*, respectively) inserts the *ura4<sup>+</sup>* pLT45 into the chromosomal *leu1* locus. B. Linear DNA derived from pLT50, or pLT50 itself was co-transformed with pLT43 into a strain with two direct copies of *attP*. The *ura4<sup>+</sup>* fragment is placed site-specifically into the chromosome in an exchange reaction.

ber of Ura<sup>+</sup> transformants than did the insertion of circular substrates. At optimal integrase concentration, the enhancement over background integration was to up to 150-fold. This is close to the frequency obtained by an autonomously replicating plasmid. Since the  $\phi$ C31 reaction is not reversible, the difference between an exchange reaction and an insertion reaction cannot be attributed to the rate of reaction reversal. Rather, this difference may be due to substrate availability. Direct DNA delivery methods are known to yield the concatamerization of extrachromosomal molecules, which will reduce the number of the single copy circular substrates for the cointegration reaction. For an exchange reaction, concatamers may still be as effective, as all that is required in a substrate is two directly oriented recombination sites.

In a human cell line, Groth *et al.* (2000) conducted intra and intermolecular recombination on episomic vectors. For the intramolecular reaction, a plasmid containing a marker gene flanked by *attP* and *attB* sites was transfected along with a  $\phi$ C31 integrase-expressing plasmid. For the intermolecu-

lar reaction, stable cell lines were used that harbor an Epstein Bar viral vector with an *attB* target. A second *attP*-containing plasmid was subsequently co-transfected into these cells, along with a  $\phi$ C31 integrase-expressing plasmid. To assay for recombination, plasmid DNA was recovered from the transfected cells for introduction into *E. coli*. The expected deletion, or cointegration of two plasmids, was detected and confirmed by molecular analyses. Episomal excision and cointegration efficiencies, using full-length *att* sites, were highest at 54% and 7.5%, respectively. Minimal sites of 34 bp for *attB* and 39 bp for *attP* were still functional but less efficient.

More recently, the same group reported integration into human and mouse chromosomes (Thyagarajan *et al.*, 2001). An *attP* target was engineered into the genome. Upon transfecting the target line with an *attB* plasmid along with an integrase-expressing construct, the number of stably transfected colonies increased 5- to 17-fold over background random integration (without co-introduced integrase construct). Surprisingly, however, only 15% of the transfected colonies were integrations at the *attP* target, which means that site-specific integration was only 0.75- to 2.6-fold above background. A majority of the background events were insertions into pseudo-*attP* sequences that share only limited primary sequence identity with  $\phi$ C31 *attP*. The authors estimated ca. 100 different pseudo-*attP* sites in these mammalian genomes. This finding suggests that similar pseudo-*attP* sites may also be present in plant DNA.

As a follow-up of the fission yeast experiments, this laboratory has also examined the  $\phi$ C31 integrase system in a mammalian cell line (Ow *et al.*, 2001). Chinese hamster ovary cell lines were generated with stably integrated *attB* target constructs pFY12, pFY14 or pFY15 (Figure 8). These plasmids harbor an *attB* site of various lengths located between the human cytomegalovirus promoter and the *lacZ* coding region. Plasmids pFY12, pFY14 and pFY15 contain, respectively, 90, 50 and 30 bp of the *attB* sequence. Four lines of each construct were used for integration experiments to average out the possibility of chromosome position effects. Each of the 16 lines was transfected with pFY6, a  $\phi$ C31 integrase expression plasmid, along with an integration vector, pFY17, pFY19, or pFY20. The plasmids pFY17, pFY19 and pFY20 harbor an *attP* sequence of 90, 50 and 32 bp, respectively. The *attP* sequence is situated upstream of the *hpt* open reading frame. Recombination between *attP* and *attB* should place the target site promoter upstream of *hpt* to

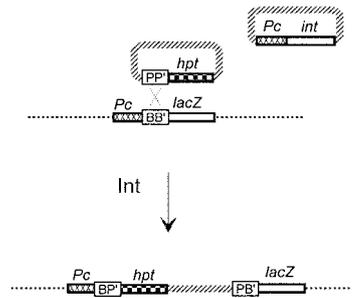
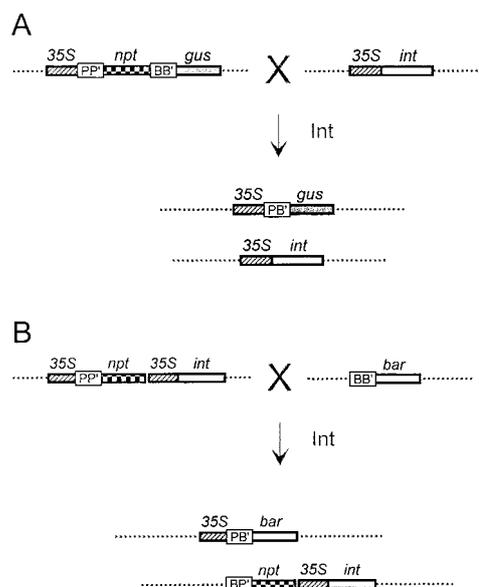


Figure 8. Gene targeting in Chinese hamster ovary cells. The *attP-hpt* plasmid recombines with a chromosomal *Pc-attB-lacZ* construct to form a *Pc-hpt* linkage and a hygromycin-resistant phenotype. *Pc*, human cytomegalovirus promoter. Integrase is provided by a co-transfected plasmid.

confer hygromycin resistance. For control, pBSK-*hpt*, isogenic to the test plasmids but lacking an *attP* sequence, was used to monitor the frequency of random promoter fusion to *hpt*.

The control plasmid pBSK-*hpt* yielded at most 2 hygromycin-resistant colonies per million cells ( $\leq 2 \times 10^{-6}$ ). In contrast, pFY17, pFY19 and pFY20 yielded up to a thousand-fold higher number of colonies, depending on the particular integration plasmid and the particular cell line. Although different target lines could exhibit intrinsic differences resulting from the site of integration, an overall trend was seen. Most efficient combination appear to be 90 bp *attB*  $\times$  50 bp *attP* ( $9 \times 10^{-4}$ ) and 90 bp *attB*  $\times$  90 bp *attP* ( $6 \times 10^{-4}$ ). A number of other combinations, 50 bp *attB*  $\times$  50 bp *attP*, 90 bp *attB*  $\times$  32 bp *attP*, and 30 bp *attB*  $\times$  50 bp *attP* were slightly less efficient (ca.  $2 \times 10^{-4}$ ), and inefficient combinations were 50 bp *attB*  $\times$  90 bp *attP* and 50 bp *attB*  $\times$  32 bp *attP* ( $\leq 6 \times 10^{-5}$ ). The shorter sites were still functional, but only in combination with certain longer sites. PCR detected the expected recombination junction from representative clones. Unlike the experiments of Thyagarajan *et al.* (2001), these integration constructs produce a selectable phenotype only when fused behind a genomic promoter. Hence, integration at random or at pseudo-*attB* sites would not be detected.

In plants, intra- and intermolecular recombination have been observed respectively in tobacco and *Arabidopsis*. The CHO results suggested that 50 bp sites would suffice. As synthetic sites of this length are more convenient and faithful to produce, 50 bp sites were chosen for the plant experiments. In tobacco, intramolecular recombination was tested with two types of transgenic lines (Ow *et al.*, 2001). The first type is



**Figure 9.** Intra- and interchromosomal recombination in plants. A. Tobacco plant with a test construct is crossed with an integrase-expressing plant. Excision of *npt* results from *attB* × *attP* intramolecular recombination. B. *Arabidopsis* plant with an *attP* site and an integrase-expressing gene is crossed with a plant harboring an *attB*-*bar* construct. Intermolecular *attB* × *attP* recombination links 35S to *bar* to yield basta resistance.

transgenic for 35S-*int*, and the second type for a 35S-*attB*-*npt*-*attP*-*gus* construct (Figure 9A). Integrase-promoted site-specific recombination between *attB* and *attP* should delete the *npt* marker and fuse 35S to *gus*, separated by a hybrid *attP*/*attB* sequence. This linkage should produce GUS active plants. Indeed, when the hemizygous parents were crossed, GUS active progeny were obtained in a quarter of the progeny, and PCR analysis detected the recombination junction. Sequencing of the PCR product confirmed a precise recombination event.

In *Arabidopsis*, site-specific recombination between nonhomologous chromosomes was detected in the progeny of crosses between plants containing a 35S-*attP* construct, and plants with an *attB*-*bar* construct (Figure 9B, this laboratory, unpublished data). The 35S-*attP* plants also harbor a 35S-*int* fragment. Some of the progeny from these crosses exhibited resistance to the herbicide basta. Southern analysis confirmed the presence of the recombination junction. As Cre-*lox*-mediated recombination between nonhomologous chromosomes is a rare event (Qin *et al.*, 1994; Koshinsky *et al.*, 2000; Vergunst and Hooykaas, 2000a), the  $\phi$ C31 *attB* × *attP* reaction may offer im-

provement for chromosome engineering within and between plant species (Ow, 1996).

### Site-specific integration followed by marker removal

The co-incorporation of selection genes in commercial products has received considerable attention over the past decade. Many of the most effective marker genes confer resistance to antibiotics. As only a small fraction of the bacterial kingdom can be cultured for analysis, the possibility, probability and predicted consequences of transmission to unintended hosts will likely remain an open question. For markers that are not relevant to the intended traits to be introduced, a diffusion of this controversy is to do away with them.

The removal of a marker gene after successful gene transfer has a more important virtue. It permits appropriate re-use in subsequent transformations. For plants that are propagated by vegetative cuttings, introducing a second trait to an already transformed plant would require the use of a new marker gene. Yet, for any particular plant, only a few markers are available. To develop a new marker gene for each new trait appended to the existing stock is not practical, not to mention the costs involved for safety evaluations of each new gene. For sexually propagated plants, different traits can be introduced with the same marker into separately transformed lines, followed by combining the traits through genetic crosses. However, the progeny with the multiple traits would end up with multiple copies of the same marker DNA, and homologous sequences can lead to gene silencing. If only the marker genes were silenced, it would be inconsequential. But should gene silencing spread to the closely linked trait gene, the whole exercise of engineering the value-added trait could be undermined. Given that future commercial products will surely be engineered through the stacking of multiple traits, recycling the use of a marker gene, which necessitates its removal after each use, would be a most appropriate solution.

Several approaches for removing marker DNA are available, including homologous recombination, transposition and co-transformation (Ow, 2000; Hohn *et al.*, 2001). An early description is based on site-specific recombination (Dale and Ow, 1991; Russell *et al.*, 1992). This method consists of flanking the marker gene with directly oriented recombination sites such that subsequent recombination excises the marker from the genome. This strategy has also been

extended to the removal of marker genes from plant chloroplasts (Corneille *et al.*, 2001; Hajdukiewicz *et al.*, 2001). The recombinase for the reaction can be provided by a genetic cross to a recombinase-producing plant or by a second round of gene transfer with the recombinase gene. The recombinase locus is subsequently removed by genetic segregation. For plants that are bred through vegetative means, the recombinase will have to be provided by the subsequent transient expression of a recombinase gene (Lyznik *et al.*, 1996; Gleaves *et al.*, 1999), or by the induced expression of a co-transformed recombinase gene (Sugita *et al.*, 2000; Zuo *et al.*, 2001). With the latter strategy, the recombinase gene is transcriptionally repressed when co-introduced with the marker gene, and is subsequently activated to remove itself as well as other unneeded DNA. This strategy can be more effective, convenient, and applicable to a wide spectrum of sexually and asexually propagated species (Ow, 2001).

The use of a recombination system to direct DNA integration can be combined with a second recombination system for marker removal. An illustration of this concept is shown in Figure 3B. (Note that the single-copy target site in Figure 3A, generated by a resolution strategy, can then be followed by site-specific integration of a trait gene, and finally by another resolution step to remove unneeded DNA.)

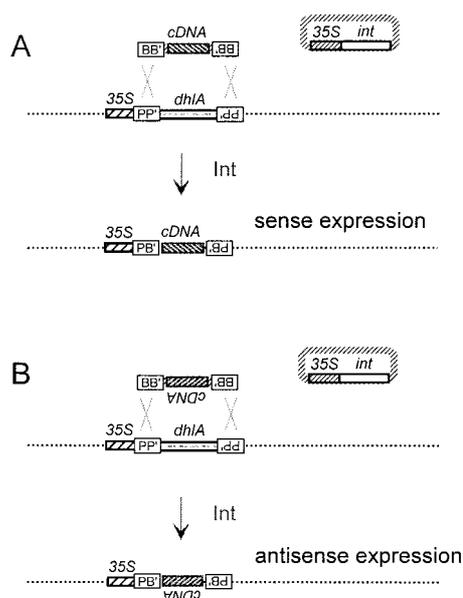
### Controlling recombinase activity

The future deployment of recombination systems will likely incorporate a greater control of recombinase activity. A novel method for delivering the Cre protein to the target site has recently been described (Vergunst *et al.*, 2000b). Attached to the NH<sub>2</sub>-terminus of the VirE2 or VirF protein, the Cre fusion protein was delivered by the *Agrobacterium* VirB/D4 transport system for site-specific recombination in the plant nucleus. This method bypasses the need to incorporate a recombinase gene into the host genome. On another front is the development of inducible recombinase systems (Lyznik *et al.*, 1995; Sieburth *et al.*, 1998; Kilby *et al.*, 2000; Sugita *et al.*, 2000; Hoff *et al.*, 2001; Zuo *et al.*, 2001). Such systems permit the co-transformation of a silent recombinase gene along with the trait genes, followed by subsequent activation of the recombinase gene to rearrange the DNA. Limiting recombinase activity can also help minimize possible deleterious effects from constitutive

recombinase activity (Ow, 2001). As recombinases are DNA-binding proteins, it is possible that their hyper-accumulation may interfere with normal chromosome activities, such as transcription, replication and segregation.

As mentioned earlier, a stable transformant expressing *gin* was not obtained, presumably due to deleterious effects exerted by the Gin recombinase (Maeser and Kahmann, 1991). With the Cre-*lox* system, the presence of a highly expressed *cre* construct in petunia and tomato has been associated with crinkled leaves and/or reduced fertility (Que *et al.*, 1998; Mark van Haaren, personal communication). A similar, though less drastic, effect can be seen in some tobacco and *Arabidopsis* lines (Mark van Haaren, personal communication). Fortunately, the abnormal phenotype co-segregates away with *cre* DNA, so the recombinase is not likely to have caused permanent genetic changes. With the  $\phi$ C31 system, crinkled leaves have also been observed in some of the *Arabidopsis* lines transgenic for a *35S-int* construct in conjunction with an *attP* site as depicted in Figure 9B (unpublished observation, this laboratory).

Unexpected genetic changes, however, have been documented in recent reports on Cre-*lox*-mediated marker removal in tobacco plastids (Corneille *et al.*, 2001; Hajdukiewicz *et al.*, 2001). When the *cre* gene was introduced by a second round of *Agrobacterium* transformation, both articles described unexpected chloroplast DNA deletions in a small percentage of the plants. However, when *cre* was introduced through pollen, unexpected deletions were not found (Corneille *et al.*, 2001). It is possible that the combined steps of re-transformation and regeneration had contributed to the genetic changes. *Agrobacterium*-mediated transformation can yield an initial high rate of transgene expression (in this case, *cre* expression) and plant regeneration from tissue culture is known to yield somaclonal variants. In animals, one report described infertility resulting from high *cre* expression in the spermatids of transgenic mice (Schmidt *et al.*, 2000). Although the article cited genetic changes as a likely cause, the data are also compatible with Cre interference on normal DNA activity. Nonetheless, the above observations suggest that it might be best to limit the expression of a recombinase gene to the designated recombination event, such as in using tightly regulated recombinase expression systems, or by providing the recombinase transiently.



**Figure 10.** A strategy for direct cDNA transformation. PCR-derived linear cDNA with *attB* ends is directed into a genomic target by a double crossover event. Replacement event is detected through loss of the negative selectable marker *dhIA*. Orientation of insertion yields sense or antisense expression.

## Future prospects

### Additional recombination systems

Besides the recombination systems described above, several others have recently been reported in the mammalian literature. They are the  $\beta$ -recombinase-*six* system from a *Bacillus subtilis* plasmid, the  $\gamma\delta$  resolvase-*res* system from the bacterial transposon *Tn1000*, and the integrase-*att* system from bacteriophage  $\lambda$ . Most probably, these systems can also be adapted for use in plants.

The  $\beta$  recombinase-*six* system catalyzes intra- but not intermolecular recombination. A chromatin-associated protein, such as prokaryotic HU or eukaryotic HMG1, is required for recombinase binding (Alonso *et al.*, 1995). As HMG1 is present in mammalian cells, Diaz *et al.* (1999) showed that expression of the single-polypeptide recombinase was sufficient for intramolecular recombination.

The  $\gamma\delta$  resolvase recombinase *res* sites to resolve co-integrate structures generated through transposition. This reaction requires negatively supercoiled DNA. In transient expression studies on mammalian cells, the wild-type resolvase failed to recombine the episomal substrates, implying that episomal DNA is not negatively supercoiled. In contrast, mutant  $\gamma\delta$

resolvases that recognize topologically relaxed substrates were able to excise DNA from episomal molecules (Schwikardi and Dorge, 2000).

With the phage  $\lambda$  system, integrase-mediated *attB*  $\times$  *attP* or *attL*  $\times$  *attR*, respectively, directs the integration into or the excision out of the bacterial genome. These reactions require DNA supercoiling of the target sites, and accessory proteins IHF and FIS, and for excision, also XIS. However, mutant integrase proteins can perform intramolecular reactions without these requirements. Using these mutant  $\lambda$  integrases, Lorbach *et al.* (2000) demonstrated DNA inversions in recombination targets introduced into the human genome.

### Targeting for gene discovery

The development of recombinase-directed DNA integration for gene discovery has been initiated by attempts to integrate large BAC molecules (Choi *et al.*, 2000). Introducing large DNA molecules would certainly expedite genetic complementation efforts. Although site-specific integration was achieved, it remains to be demonstrated whether large inserts can integrate with minimal rearrangement.

A second possibility lies in the direct transformation of cDNA molecules. The experiments of Thomson *et al.* (2001) showed that linear DNA flanked by  $\phi$ C31 *att* sites is efficiently integrated into the *S. pombe* chromosome, approaching the efficiency of plasmid transformation by an autonomously replicating vector. Should this be achieved in plants, even with lower efficiency, this can be an effective approach to help elucidate gene function. One scenario would begin with the target plant engineered to harbor a promoter-*attB*-counterselection-gene-*attB* construct, with the *attB* sites as inverted repeats (Figure 10). The haloalkane dehalogenase gene *dhIA* is listed as one example of a counterselectable marker (Naested *et al.*, 1999). An mRNA implicated by comparative genomic or transcript-profiling analysis can be selectively amplified by PCR using primers with *attP* ends. In the example shown, the *attP* ends would be in opposite orientation such that the cDNA can insert into the target in either orientation. The dual recombination reaction would fuse the cDNA behind the target promoter for sense or antisense expression, with the intention that it would lead to hyper- or hypoproduction of the gene product. Direct transformation of a cDNA molecule without passing it through *E. coli*

not only saves an extra cloning step, but also bypasses the selective effects exerted by bacteria.

A third possibility lies in developing high-frequency integration. If high integration rates can be achieved, direct library transformation may become an option for gene discovery, much as functional complementation has been for microbial systems. A major difference between the transformation of, for example, *E. coli* and higher-plant cells is that the bacterium uses stably replicating vectors. The efficiency of transformation is therefore mainly determined by the rate of DNA uptake. In plants, it is the combination of DNA uptake, DNA integration, and plant cell regeneration. With respect to DNA integration, at least, there is the prospect of enhancing its efficiency with integrase systems. Many phage chromosomes lyzogenize their host genomes via site-specific integration at near 100% efficiency. Towards that goal, the  $\phi$ C31 system has been the most promising to date. Whereas other recombination systems evolved for excision or inversion, the  $\phi$ C31 system is naturally used for DNA integration. In conjunction with a linear DNA transformation system, it may be possible to shotgun-transform small collections of cDNAs, such as pre-selected collections identified through mRNA profiling analyses.

### Acknowledgements

I am grateful to L. Gilbertson, P. Maliga, M. van Haaren, A. Vergunst, and members of this laboratory (F.-Y. Bih, C. Day, V. Srivastava, B. Stuart and L. Xiong) for communication of information not yet available in the refereed literature.

### References

- Albert, H., Dale, E.C., Lee, E. and Ow, D.W. 1995. Site-specific integration of DNA into wild-type and mutant *lox* sites placed in the plant genome. *Plant J.* 7: 649–659.
- Alonso, J.C., Weise, F. and Rojo, F. 1995. The *Bacillus subtilis* histone-like protein Hbsu is required for DNA resolution and DNA insertion mediated by the  $\beta$  recombinase of plasmid pSM19035. *J. Biol. Chem.* 270: 2938–2945.
- Araki, K., Araki, M. and Yamamura, K. 1997. Targeted integration of DNA using mutant *lox* sites in embryonic stem cells. *Nucl. Acids Res.* 25: 8968–8972.
- Battacharyya, M.K., Stermer, B.A. and Dixon, R.A. 1994. Reduced variation in transgene expression from a binary vector with selectable markers at the right and left T-DNA borders. *Plant J.* 6: 957–968.
- Baubonis, W. and Sauer, B. 1993. Genomic targeting with purified Cre recombinase. *Nucl. Acids Res.* 21: 2025–2029.
- Bayley, C.C., Morgan, M., Dale, E.C. and Ow, D.W. 1992. Exchange of gene activity in transgenic plants catalyzed by the Cre-*lox* site-specific recombination system. *Plant Mol. Biol.* 18: 353–361.
- Bethke, B. and Sauer, B. 1997. Segmental genomic replacement by Cre-mediated recombination: genotoxic stress activation of the p53 promoter in single-copy transformants. *Nucl. Acids Res.* 25: 2828–2834.
- Baszczynski, C.L., Bowen, B.A., Drummond, B., Gordon-Kamm, W.J., Peterson, D.J., Sandahl, G.A., Tagliani, L.A., Zhao, Z.-Y. 2001a. Novel nucleic acid sequence encoding FLP recombinase. US patent 6,175,058 B1.
- Baszczynski, C.L., Bowen, B.A., Peterson, D.J., Tagliani, L.A. 2001b. Composition and methods for genetic modification of plants. US patent 6,187,994 B1.
- Choi, S., Begum, D., Koshinsky, H., Ow, D.W. and Wing, R.A. 2000. A new approach for the identification and cloning of genes: the pBACwch system using Cre/*lox* site-specific recombination. *Nucl. Acids Res.* 28: e19, i–vii.
- Cornelle, S., Lutz, K., Svab, Z. and Maliga, P. 2001. Efficient elimination of selectable marker genes from the plastid genome by the Cre-*lox* site-specific recombination system. *Plant J.* 27: 171–178.
- Dale, E.C. and Ow, D.W. 1990. Intra- and intermolecular site-specific recombination in plant cells mediated by bacteriophage P1 recombinase. *Gene* 91: 79–85.
- Dale, E.C. and Ow, D.W. 1991. Gene transfer with the subsequent removal of the selection gene from the host genome. *Proc. Natl. Acad. Sci. USA* 88: 10558–10562.
- Day, C.D., Lee, E., Kobayashi, J., Holappa, L.D., Albert, H. and Ow, D.W. 2000. Transgene integration into the same chromosomal location can produce alleles that express at a predictable level, or alleles that are differentially silenced. *Genes Dev.* 14: 2869–2880.
- De Buck, S., Van Montagu, M. and Depicker, A. 2001. Transgene silencing of invertedly repeated transgenes is released upon deletion of one of the transgenes involved. *Plant Mol. Biol.* 46: 433–445.
- Dellaire, G., Lemieux, N., Belmaaza, A. and Chartrand, P. 1997. Ectopic gene targeting exhibits a bimodal distribution of integration in murine cells, indicating that both intra- and interchromosomal sites are accessible to the targeting vector. *Mol. Cell. Biol.* 17: 5571–5580.
- Diaz, V., Rojo, F., Martinez, A.C., Alonso, J.C. and Bernad, A. 1999. The prokaryotic  $\beta$ -recombinase catalyzes site-specific recombination in mammalian cells. *J. Biol. Chem.* 274: 6634–6640.
- Feng, Y-Q, Seibler, J., Alami, R., Eisen, A., Westerman, K.A., Leboulch, P., Fiering, S. and Bouhassira, E.E. 1999. Site-specific chromosomal integration in mammalian cells: highly efficient Cre recombinase-mediated cassette exchange. *J. Mol. Biol.* 292: 779–785.
- Gleave, A.P., Mitra, D.S., Mudge, S.R. and Morris, B.A.M. 1999. Selectable marker-free transgenic plants without sexual crossing: transient expression of *cre* recombinase and use of a conditional lethal dominant gene. *Plant Mol. Biol.* 40: 223–235.
- Groth, A.C., Olivares, E.C., Thyagarajan, B. and Calos, M.P. 2000. A phage integrase directs efficient site-specific integration in human cells. *Proc. Natl. Acad. Sci. USA* 97: 5995–6000.
- Hajdukiewicz, P.T., Gilbertson, L.A. and Staub, J.M. 2001. Multiple pathways for Cre/*lox*-mediated recombination in plastids. *Plant J.* 27: 161–170.

- Hoess, R.H., Wierzbicki, A. and Abremski, K. 1986. The role of the *loxP* spacer region in P1 site-specific recombination. *Nucl. Acids Res.* 14: 2287–2300.
- Hoff, T., Schnorr, K.M. and Mundy, J. 2001. A recombinase-mediated transcriptional induction system in transgenic plants. *Plant Mol. Biol.* 45: 41–49.
- Hohn, B., Levy, A.A., Puchta, H. 2001. Elimination of selection markers from transgenic plants. *Curr. Opin. Biotechnol.* 12: 139–143.
- Howe, M., Dimitri, P. Berloco M. and Wakimoto, B.T. 1995. *Cis*-effects of heterochromatin on heterochromatic and euchromatic gene activity in *Drosophila melanogaster*. *Genetics* 140: 1033–1045.
- Huang, L.C., Wood E.A. and Cox, M.M. 1991. A bacterial model system for chromosomal targeting. *Nucl. Acids Res.* 19: 443–448.
- Iyer, L.M., Kumpatla, S.P., Chandrasekharan, M.B. and Hall, T.C. 2000. Transgene silencing in monocots. *Plant Mol. Biol.* 43: 323–346.
- Kaeppler, S.M., Kaeppler, H.F. and Rhee, Y. 2000. Epigenetic aspects of somaclonal variation in plants. *Plant Mol. Biol.* 43: 179–188.
- Kilby, N.J., Davies, G.J., Snaith, M.R. and Murray, J.A.H. 1995. FLP recombinase in transgenic plants: constitutive activity in stably transformed tobacco and generation of marked cell clones in *Arabidopsis*. *Plant J.* 8: 637–652.
- Kilby, N.J., Fyvie, M.J., Sessions, R.A., Davies, G.J. and Murray, J.A.H. 2000. Controlled induction of GUS marked clonal sectors in *Arabidopsis*. *J. Exp. Bot.* 51: 853–863.
- Kohli, A., Leech, M., Vain, P., Laurie, D.A. and Christou, P. 1998. Transgene organization in rice engineered through direct DNA transfer supports a two-phase integration mechanism mediated by the establishment of integration hot spots. *Proc. Natl. Acad. Sci. USA* 95: 7203–7208.
- Koshinsky, H.A., Lee, E. and Ow, D.W. Cre-*lox* site-specific recombination between *Arabidopsis* and tobacco chromosomes. *Plant J.* 23: 715–722.
- Lloyd, A.M. and Davis, R.W. 1994. Functional expression of the yeast FLP/*FRT* site-specific recombination system in *Nicotiana tabacum*. *Mol. Gen. Genet.* 242: 653–657.
- Lorbach, E., Christ, N., Schwikardi, M. and Droge, P. 2000. Site-specific recombination in human cells catalyzed by phage  $\lambda$  integrase mutants. *J. Mol. Biol.* 296: 1175–1181.
- Lyznik, L.A., Mitchell, J.C., Hirayama, L. and Hodges, T.K. 1993. Activity of yeast FLP recombinase in maize and rice protoplasts. *Nucl. Acids Res.* 21: 969–975.
- Lyznik, L. A., Hirayama, L., Rao, K.V., Abad, A. and Hodges, T.K. 1995. Heat-inducible expression of *FLP* gene in maize cells. *Plant J.* 8: 177–186.
- Lyznik, L.A., Rao, K.V. and Hodges, T.K. 1996. FLP-mediated recombination of *FRT* sites in the maize genome. *Nucl. Acids Res.* 24: 3784–3789.
- Maeser, S. and Kahmann, R. 1991. The Gin recombinase of phage Mu can catalyze site-specific recombination in plant protoplasts. *Mol. Gen. Genet.* 230: 170–176.
- Matzke, M.A., Mette, M.F. and Matzke, A.J.M. 2000. Transgene silencing by the host genome defense: implications for the evolution of epigenetic control mechanisms in plants and vertebrates. *Plant Mol. Biol.* 43: 401–415.
- Meyer, P. 2000. Transcriptional transgene silencing and chromatin component. *Plant Mol. Biol.* 43: 221–234.
- Muskens, M.W.M., Vissers, A.P.A., Mol, J.N.M. and Kooter, J.M. 2000. Role of inverted DNA repeats in transcriptional and post-transcriptional gene silencing. *Plant Mol. Biol.* 43: 243–260.
- Naested, H., Fennema, M., Hao, L., Andersen, M., Janssen, D.B. and Mundy, J. 1999. A bacterial haloalkane dehalogenase gene as a negative selectable marker in *Arabidopsis*. *Plant J.* 18: 571–576.
- Odell, J., Caimi, P., Sauer, B. and Russell, S. 1990. Site-directed recombination in the genome of transgenic tobacco. *Mol. Gen. Genet.* 223: 369–378.
- Odell, J.T., Hoopes, J.L. and Vermerris, W. 1994. Seed-specific gene activation mediated by Cre-*lox* site-specific recombination system. *Plant Physiol.* 106: 447–458.
- 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.
- O’Gorman, S., Fox, D.T. and Wahl, G.M. 1991. Recombinase-mediated gene activation and site-specific integration in mammalian cells. *Science* 251: 1351–1355.
- Onouchi, H., Yokoi, K., Machida, C., Matzuzaki, H., Oshima, Y., Matsuoka, K., Nakamura, K. and Machida, Y. 1991. Operation of an efficient site-specific recombination system of *Zygosaccharomyces rouxii* in tobacco cells. *Nucl. Acids Res.* 19: 6373–6378.
- Onouchi, H., Nishihama, R., Kudo, M., Machida, Y. and Machida, C. 1995. Visualization of site-specific recombination catalyzed by a recombinase from *Zygosaccharomyces rouxii* in *Arabidopsis thaliana*. *Mol. Gen. Genet.* 247: 653–660.
- Ow, D.W. and Medberry, S.L. 1995. Genome manipulation through site-specific recombination. *Crit. Rev. Plant Sci.* 14: 239–261.
- Ow, D.W. 1996. Recombinase-directed chromosome engineering in plants. *Curr. Opin. Biotechnol.* 7: 181–186.
- Ow, D. 2000. Marker Genes. Joint FAO/WHO Expert Consultation on Foods Derived from Biotechnology (May 29–June 2, 2000, Geneva, Switzerland); [http://www.who.int/fsf/GMfood/Consultation\\_May2000/Biotech\\_00\\_14.pdf](http://www.who.int/fsf/GMfood/Consultation_May2000/Biotech_00_14.pdf)
- Ow, D.W. 2001. The right chemistry for marker gene removal? *Nature Biotechnol.* 19: 115–116.
- Ow, D.W., Calendar, R. and Thomason, L. 2001. DNA recombination in eukaryotic cells by the bacteriophage phiC31 recombination system. International patent filing, WO 01/07572.
- Pawlowski, W.P. and Somers, D.A. 1998. Transgenic DNA integrated into the oat genome is frequently interspersed by host DNA. *Proc. Natl. Acad. Sci. USA* 95: 12106–12110.
- Peschke, V.M. and Phillips, R.L. 1992. Genetic implications of somaclonal variation in plants. *Adv. Genet.* 30: 41–75.
- Puchta, H. 2001. Gene replacement by homologous replacement in plants. *Plant Mol. Biol.*, this issue.
- Qin, M., Bayley, C., Stockton, T. and Ow, D.W. 1994. Cre recombinase mediated site-specific recombination between plant chromosomes. *Proc. Natl. Acad. Sci. USA* 91: 1706–1710.
- Que, Q.; Wang, H.-Y. and Jorgensen, R. 1998. Distinct patterns of pigment suppression are produced by allelic sense and anti-sense chalcone synthase transgenes in petunia flowers. *Plant J.* 13: 401–409.
- Risseuw, E., Offringa, R., Franke-van Dijk, M.E.I. and Hooykaas, P.J.J. 1995. Targeted recombination in plants using *Agrobacterium* coincides with additional rearrangements at the target locus. *Plant J.* 7: 109–119.
- Russell, S.H., Hoopes, J.L. and Odell, J.T. 1992. Directed excision of a transgene from the plant genome. *Mol. Gen. Genet.* 234: 49–59.
- Sadowski, P.D. 1993. Site-specific genetic recombination: hops, flips, and flops. *FASEB J.* 7: 760–767.

- Sauer, B. and Henderson, N. 1990. Targeted insertion of exogenous DNA into the eukaryotic genome by the Cre recombinase. *New Biol.* 2: 441–449.
- Schmidt, E.E., Taylor, D.S., Prigge, J.R., Barnett, S. and Capecchi, M.R. 2000. Illegitimate Cre-dependent chromosome rearrangements in transgenic mouse spermatids. *Proc. Natl. Acad. Sci. USA* 97: 13702–13707.
- Schwikardi, M. and Droge P. 2000. Site-specific recombination in mammalian cells catalyzed by  $\gamma\delta$  resolvase mutants: implications for the topology of episomal DNA. *FEBS Lett.* 471: 147–150.
- Seibler, J. and Bode, J. 1997. Double-reciprocal crossover mediated by FLP-recombinase: a concept and an assay. *Biochemistry* 36: 1740–1747.
- Sieburth, L.E., Drews, G.N. and Meyerowitz, E.M. 1998. Non-autonomy of AGAMOUS function in flower development: use of a Cre/lox method for mosaic analysis in *Arabidopsis*. *Development* 125: 4303–4312.
- Sonti, R.V., Tissier, A.F., Wong, D., Viret, J.-F. and Signer, E.R. 1995. Activity of the yeast FLP recombinase in *Arabidopsis*. *Plant Mol. Biol.* 28: 1127–1132.
- Srivastava, V., Anderson, O.A. and Ow, D.W. 1999. Single-copy transgenic wheat generated through the resolution of complex integration patterns. *Proc. Natl. Acad. Sci. USA* 96: 11117–11121.
- Srivastava, V. and Ow D.W. 2001a. Single copy primary transformants of maize obtained through the co-introduction of a recombinase-expressing construct. *Plant Mol. Biol.* 46: 561–566.
- Srivastava, V. and Ow D.W. 2001b. Biolistic mediated site-specific integration in rice. *Mol. Breed.*, in press.
- Stavenhagen, J.B. and Zakian, V.A. 1994. Internal tracts of telomeric DNA act as silencers in *Saccharomyces cerevisiae*. *Genes Dev.* 8: 1411–1422.
- Sugita, K., Kasahara, T., Matsunaga, E. and Ebinuma, H. 2000. A transformation vector for the production of marker-free transgenic plants containing single copy transgene at high frequency. *Plant J.* 22: 461–469.
- Thomason, L.C., Calendar, R. and Ow, D.W. 2001. Gene insertion and replacement in *Schizosaccharomyces pombe* mediated by the *Streptomyces* bacteriophage  $\phi$ C31 site-specific recombination system. *Mol. Genet. Genomics* 265: 1031–1038.
- Thorpe, H.M. and Smith M.C. 1998. *In vitro* site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. *Proc. Natl. Acad. Sci. USA* 95: 5505–5510.
- Thyagarajan, B., Olivares, E.C., Hollis, R.P., Ginsburg, D.S. and Calos, M.P. 2001. Site-specific genomic integration in mammalian cells mediated by phage  $\phi$ C31 integrase. *Mol. Cell. Biol.* 21: 3926–3934.
- Vergunst, A.C. and Hooykaas, P.J.J. 1998. Cre/lox-mediated site-specific integration of *Agrobacterium* T-DNA in *Arabidopsis thaliana* by transient expression of cre. *Plant Mol. Biol.* 38: 393–406.
- 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.
- Vergunst, A.C., Jansen, L.E.T. and Hooykaas, P.J.J. 1998. Site-specific integration of *Agrobacterium* T-DNA in *Arabidopsis thaliana* mediated by Cre recombinase. *Nucl. Acids Res.* 26: 2729–2734.
- Vergunst, A.C., Jansen, L.E.T., Fransz, P.F., de Jong, J.H. and Hooykaas, P. J. J. 2000a. Cre/lox-mediated recombination in *Arabidopsis*: evidence for a transmission of a translocation and a deletion event. *Chromosoma* 109: 287–297.
- Vergunst, A.C., Schrammeijer, B., den Dulk-Ras, A., de Vlaam, C.M.T., Regensburg-Tuink, R. and Hooykaas, P.J.J. 2000b. VirB/D4 dependent protein translocation from *Agrobacterium* into plant cells. *Science* 290: 979–982.
- Wallrath, L.L. and Elgin, S.C.R. 1995. Position effect variegation in *Drosophila* is associated with an altered chromatin structure. *Genes Dev.* 9: 1263–1277.
- Wassenegger, M. 2000. RNA-directed DNA methylation. *Plant Mol. Biol.* 43: 203–220.
- Zuo, J., Niu, Q.-W., Moller, S.G. and Chua, N.-H. 2001. Chemical-regulated, site-specific DNA excision in transgenic plants. *Nature Biotechnol.* 19: 157–161.