Chapter 7
Transgene Integration, Expression and Stability in Plants: Strategies for Improvements

Ajay Kohli, Berta Miro, and Richard M. Twyman

7.1 Introduction

The transfer of DNA into plants has been common practice for over 20 years, and transgenic plants are now a burgeoning industry. In 2007, over 114 million ha (282.4 million acres) of transgenic crops were grown commercially in 23 countries, the most prevalent traits being herbicide tolerance, pest resistance, or both traits stacked together (James 2007). In the laboratory, one encounters a vastly greater diversity of traits, including disease resistance, stress tolerance, nutritional improvement, modified development, and the use of plants to produce specific, high-value molecules, such as secondary metabolites, chemical precursors, antibodies, vaccine subunits, and industrial enzymes. It is notable that in the majority of cases, the purpose of gene transfer into plants is to achieve a specific, desirable phenotype. Plants that fail to live up to expectations are routinely discarded so that the best performers can be nurtured.

Despite the focus on phenotype, over the last decade there has been an increasing interest in creating transgenic plants to study the process of gene transfer itself (Kohli et al. 2003). On the academic side, it has been appreciated for many years that the structure of a transgene locus can have a major influence on the level and stability of transgene expression; thus, researchers have studied DNA integration mechanisms, particularly with regard to how transgenes interact with the plant's DNA repair and genome defense systems. On the applied side, the global adoption of transgenic crops and the development of transgenic plants producing pharmaceuticals and other important molecules have attracted the interest of regulatory authorities (Ramessar et al. 2008). The demand for robust risk assessment practices

A. Kohli (✉)
Plant Molecular Biology Laboratory, Plant Breeding Genetics and Biotechnology, International Rice Research Institute, DAPO-7777, Metro Manila, The Philippines
e-mail: a.kohli@cgiar.org

C. Kole et al. (eds.), *Transgenic Crop Plants*,
DOI 10.1007/978-3-642-04809-8_7,

means that transgenic plants have to be characterized in great detail, including information on the sequence, structure, organization, and genomic position of the transgenic locus. Recently, this has culminated in the first report of the genome sequence of a transgenic plant, including the analysis of the transgenic locus (Ming et al. 2008). The principles and practices of transgenic technology have come under scrutiny, leading to research focusing on the use and elimination of marker genes, the role of vector sequences that integrate along with the transgene, and the random nature of transgene integration events with regard to copy number, transgene orientation, and transgene rearrangements. Researchers, therefore, have practical as well as academic reasons for studying transgene integration and expression, and have developed new ways to analyze transgenic loci. Current research focuses on ways to better control the way DNA integrates into the plant genome.

In this chapter, we describe the methods used to study transgene locus structure and discuss evidence supporting current models of transgene integration for both *Agrobacterium*-mediated transformation and direct transfer methods. We discuss how transgene loci are organized and how this affects the level and stability of transgene expression from generation to generation. Finally, we look to the future by describing how recent research has advanced the state of the art in gene transfer technology.

7.2 Methods for the Analysis of Transgenic Loci

Most gene transfer experiments are phenotype driven, by which we mean that successfully transformed plants tend to be identified on the basis of the phenotype conferred by the transgene rather than the structure of the transgene itself. This is pertinent because the appearance of the desired phenotype is prima facie evidence that the transgene has integrated into the genome and is intact, thus allowing expression of the encoded protein. Since most transgenic plants are regenerated under selection for the product of a selectable marker gene, the fact that a transgenic plant exists at all indicates that at least one intact copy of the marker gene is present in the genome. Similarly, the phenotypes conferred by any other transgenes can be used as evidence to support successful integration and expression. This information is of limited value, however, because it divides all plants into just two categories – (a) plants transformed with at least one intact transgene and (b) plants not transformed at all or transformed with a nonfunctional transgene. It provides no quantitative information, yet every gene transfer experiment produces a population of plants with a range of phenotypes reflecting the level of transgene expression. Since the same input DNA is used in each case, the only explanation for quantitative differences in phenotype is differences in the structure and activity of the integrated transgenes.

The technique used most commonly for a definitive analysis of transgenic loci is the Southern blot, in which genomic DNA is digested with one or more restriction

enzymes, fractionated by agarose gel electrophoresis, denatured, transferred to a membrane, and hybridized to a labeled probe. Many different types of information can be obtained from Southern blots depending on the restriction enzymes and probes used. One of the most common strategies is to use an enzyme that cuts once within the transgene in combination with a probe that hybridizes to the body of the transgene. This generates DNA fragments whose size depends on the distance between a fixed point in the transgene (the restriction site) and the adjacent restriction site in the genomic DNA, which of course varies according to the site of insertion. Where multiple copies of a transgene have integrated, a single cutter enzyme tends to generate a unique pattern of bands that serves as a genetic fingerprint of that plant and all its descendants, thereby helping to identify clonal relatives of the original transformant and allowing transgene segregation to be followed through generations. Because it is unlikely that any of the transgene copies will generate identical-sized bands (unless they are perfect concatemers, in which case the band size will correspond exactly to the size of the transgene), this method also provides an estimate of transgene copy number. A variant of the technique is to use a probe that hybridizes to the vector backbone instead of the transgene body, which helps to identify inserts of vector DNA.

Another handy method is to use an enzyme that cuts twice in the transgene and liberates a specific DNA cassette in combination with a probe that hybridizes to that cassette. If all copies of the transgene are intact, there should be only one hybridizing band, corresponding to the size of the cassette, and the intensity of the hybridization signal will be proportional to the number of transgene copies (since one cassette should be released from each integrated transgene copy). Copy number determination is best achieved by "spiking" genomic DNA from an untransformed plant with a known amount of transforming plasmid DNA, and then digesting this and the genomic DNA from genuine transgenic plants. It may be useful to set up a series of control DNA samples, containing for example one, five, and ten copies of the transforming plasmid per genome equivalent of DNA. In this way, a calibration curve of signal intensities relative to copy number can be generated, onto which any transgenic plant can be mapped. Band sizes greater or smaller than the diagnostic fragment indicate truncations or rearrangements. It is also useful to digest genomic DNA with an enzyme known not to cut within the transgene. If all copies of the transgene have integrated at a single locus as a concatemer, digestion with such enzymes should liberate the locus as a single, high-molecular-weight fragment. Thus, the presence of two or more bands suggests either the presence of two or more independent transgenic loci (this can be confirmed by segregation analysis as discussed below), or the presence of interspersed genomic DNA between transgene copies at a single locus (this can be confirmed, if required, by fiber-FISH as discussed below).

The polymerase chain reaction (PCR) is a rapid technique that can be used to confirm transgene integration through the use of primer combinations that generate a transgene-specific product. Although quicker than Southern blots and more

amenable to multiplexing, false positives can occur through the amplification of episomal plasmid DNA so the PCR should only be used indicatively, with Southern blots used for definitive confirmation of DNA integration. Long PCR is a variation that allows larger products to be amplified and is potentially useful for analyzing larger transgenic loci (Mehlo et al. 2000). Another PCR variant, real-time quantitative PCR, is now used for the rapid estimation of transgene copy number by comparison with a control sample in which a single-copy endogenous gene is amplified. The relative signal intensities of the control and transgenic samples reveal the transgene copy numbers in the transgenic plants (Li et al. 2004; Yang et al. 2005). Although the PCR can show the presence or absence of a transgene and provide a dependable copy number estimate, it provides little in the way of information about the structure of a transgenic locus unless the genomic flanking sequences are already known. DNA sequencing is the highest-resolution transgene analysis method, and permits the precise definition of structural organization and rearrangement. It also allows the nature of transgene-genomic and transgene-transgene junctions to be investigated at the nucleotide level, and the integration site in genomic DNA to be identified. Prior to sequencing, the transgenic locus must be isolated, which can be achieved by standard molecular cloning or through a variety of methods such as inverse PCR, thermal asymmetric interlaced (TAIL)-PCR, or plasmid rescue (Twyman and Kohli 2003). The complete sequencing of a transgenic papaya variety, SUNUP, was recently reported, allowing intricate structural analysis of the transgenic locus (Ming et al. 2008).

All these methods involve the identification of discrete DNA fragments of precise length, which is useful for fine structural analysis but not for the characterization of transgenic loci at the chromosome level. In this context, fluorescence in situ hybridization (FISH) can be useful as it allows target sequences to be identified in isolated DNA fibers, interphase chromatin, and even metaphase chromosomes. FISH involves the use of fluorescently labeled nucleic acid probes to identify particular target sequences, revealing higher-order transgene organization and the distribution of integration sites. FISH to metaphase chromosomes allows the insertion sites to be mapped cytogenetically and simultaneous analysis in interphase allows the nuclear territory of transgenes to be determined (Abranches et al. 2000). Fiber-FISH on extended chromatin gives an overview of locus structure, revealing the presence of single-copy inserts, transgene concatemers, and interspersed genomic DNA (Jackson et al. 2001). The resolution of FISH lies somewhere between that of genetic segregation and Southern blot hybridization and can provide important correlative data for both techniques.

Segregation analysis involves studying the transmission of particular DNA sequences or the phenotypes thus conferred over several generations of transgenic plants. Closely linked transgene copies are unlikely to be separated by recombination, while widely separated loci are likely to segregate at meiosis in some plants. This allows the number of transgenic loci to be determined. Problems with phenotype analysis include the misleading results caused by epigenetic gene silencing, but analysis of DNA sequence segregation by Southern blot hybridization can be highly informative.

7.3 Locus Structure in Plants Transformed by *Agrobacterium tumefaciens*

7.3.1 Principles of Gene Transfer

Agrobacterium tumefaciens is a soil pathogen that colonizes wounded plant cells and induces the formation of a tumor (or crown gall) that produces special amino acid derivatives called opines, which the bacteria are able to use as a carbon and nitrogen source. The ability of virulent *Agrobacterium* strains to induce tumor growth and opine synthesis, and the capacity to utilize opines, is conferred by a resident tumor-inducing plasmid (Ti-plasmid). During the colonization process, a segment of DNA from this plasmid called the transferred DNA (T-DNA) is transferred to the plant nuclear genome. The T-DNA encodes enzymes that synthesize auxins and cytokinins, resulting in unregulated cell proliferation, and enzymes that synthesize opines from standard amino acids (reviewed by Gelvin 2003).

The Ti plasmid is a naturally occurring vector for plant transformation, but wild-type Ti-plasmids are not suitable vectors for genetic engineering in plants because they are too big to manipulate, and the oncogenes contained in the T-DNA cause uncontrolled proliferation of transformed plant cells and prevent efficient regeneration. The T-DNA must therefore be moved to a smaller, more convenient vector, and disarmed by deleting the oncogenes. A marker gene must also be included to allow transformed cells to be propagated. T-DNA transfer is controlled by about 30 genes located in a separate virulence (*vir*) region of the Ti plasmid, and these must be supplied in trans using a binary vector system if the T-DNA is placed on a smaller plasmid. Modern binary vectors contain multiple unique cloning sites within the T-DNA, a *lac*Z marker gene for blue-white selection of recombinants, and a choice of selectable markers to identify transformed plant cells (Hellens et al. 2000).

The transfer mechanism is pertinent to the resulting locus structure. The T-DNA is flanked by 25-bp imperfect direct repeats known as border sequences, which are not transferred to the plant genome intact, but they are required for the transfer process. T-DNA transfer is mediated by the *vir*A and *vir*G gene products, which transduce external signals and activate other *vir* genes resulting in the construction of a pilus for DNA transfer, and the release of the T-DNA by an endonuclease comprising the products of the *vir*D1 and *vir*D2 genes. This introduces either single-strand nicks or a double-strand break at the 25-bp borders of the T-DNA. It is thought that the intermediate formed (a double stranded T-DNA or a single T-strand) may depend on the virulence functions particular to the *Agrobacterium* strain (Steck 1997). Either the left or right border sequence can initiate T-DNA transfer, although it is more usual for initiation to occur at the right border due to the presence of an adjacent overdrive sequence, which is recognized by the VirC1 and VirC2 proteins and acts as a transfer enhancer (Shaw et al. 1984). For this reason, deletion of the right T-DNA border severely reduces the efficiency of transfer, whereas deletion of the left border has little effect (Jen and Chilton 1986).

The VirD2 protein remains covalently attached to the 5′ end of the processed T-DNA strand and has been proposed to protect the T-DNA against nucleases, to target the DNA to the plant cell nucleus, and to help integrate it into the plant genome (Tzfira et al. 2000).

7.3.2 T-DNA Locus Structure

Most investigations of T-DNA transfer have suggested that there may be preferential integration into transcription units, with up to 90% of events occurring in genes (e.g., Lindsey et al. 1993). In petunia, FISH analysis showed that T-DNA inserts were found preferentially at the gene-dense distal chromosome sites (Wang et al. 1995; Ten Hoopen et al. 1996), and a comparative analysis in *Arabidopsis* and rice showed that T-DNA inserted randomly in the *Arabidopsis* genome (which is globally gene-rich, with little repetitive DNA) but homed in on the 10–20% of the rice genome known to be gene-dense while avoiding the more widespread heterochromatic regions (Barakat et al. 2000). It has also been suggested that T-DNA integration occurs preferentially in regions showing microhomology to the T-DNA borders (Matsumoto et al. 1990), which may also be enriched in the transcribed part of the genome. The bias in T-DNA insertion is valuable for gene-tagging experiments, ensuring a high gene hit rate in genomes with large tracts of gene-poor heterochromatin while showing little bias among different genes (Jeon et al. 2000; Weigel et al. 2000; Hsing et al. 2007; Wan et al. 2009).

The structure and complexity of transgenic loci generated by *Agrobacterium* depends on the strain, plant species, and explant type, but generally gives rise to lower transgene copy numbers than direct transformation methods. An informative experiment was performed by Cheng et al. (1997) by transforming wheat using both *Agrobacterium* and particle bombardment. Of 26 *Agrobacterium*-mediated transformants, more than one-third contained a single T-DNA insert, half contained 2–3 copies, and the remainder (about 15%) contained 4–5 copies. There were no transformants containing more than five T-DNAs. In contrast, from the population of 77 bombarded transformants, only 13 (17%) contained a single copy of the transgene. The maximum number of transgene copies in this population was not reported. Hu et al. (2003) also observed more complex transgene insertions from particle bombardment than from *Agrobacterium*-mediated techniques. More recently, similar experiments in barley showed that all the *Agrobacterium*-derived lines contained 1–3 copies of the transgene, while 60% of the transgenic lines derived by particle bombardment contained more than eight copies (Travella et al. 2005). Dai et al. (2001) found in rice that the average transgene copy numbers were 1.8 for *Agrobacterium*-derived lines and 2.7 for plants obtained by particle bombardment. However, Khanna and Raina (2002) observed multiple transgene insertions in rice transformants generated through both techniques together with the transfer of partial T-DNA fragments.

The organization of integrated T-DNA sequences differs among *Agrobacterium* strains, but a common feature of nopaline-type derivatives such as C58 is the preferential integration of T-DNA as dimers with an inverted repeat configuration, linked either at the left or right borders (Jones et al. 1987; Jorgensen et al. 1987). Where cotransformation is carried out with two T-DNAs containing different markers, the different T-DNAs were often present as heterodimer inverted repeats, preferentially around the right border (De Block and Debrouwer 1991). Similarly, cotransformation of rice with the vectors pGreen and pSoup (each containing different selectable and visible markers) resulted in 56% of plants with the two T-DNAs cointegrated, although there was also a high proportion of plants containing separate integration events (Afolabi et al. 2004). In contrast, Spielmann and Simpson (1986) carried out transformation using the octopine *Agrobacterium* strain LBA4404. They found only two integration events among the 22 characterized transformants that resulted in dimer formation, while most of the rest were single-copy integrations. When cotransformation experiments were carried out with this strain (McKnight et al. 1987), three double transformants were obtained and in all cases the two T-DNAs were genetically unlinked. These results suggest that the virulence functions carried by a particular *Agrobacterium* strain strongly influence the structure of the transgene locus.

Another important aspect of locus structure is the amount and types of transgene rearrangement. Occasionally, it has been reported that T-DNA has undergone spontaneous rearrangement prior to or during integration (e.g., Offringa et al. 1990; Puchta et al. 1992), and this has been demonstrated directly by fiber-FISH in potato (Wolters et al. 1998). In some cases, rearrangements may be induced by specific recombinogenic sequences such as the CaMV 35S promoter (Kohli et al. 1999), which may have been responsible for T-DNA rearrangements in some transgenic potato lines (Porsch et al. 1998). In many cases, however, rearrangements may reflect "collateral damage" occurring spontaneously during the transfer process. Afolabi et al. (2004) found that nonintact T-DNAs were present in >70% of transgenic rice lines, in most cases reflecting loss of the mid to right border portion of the T-DNA. Similarly, Rai et al. (2007) found that about 50% of rice plants transformed with a T-DNA containing the phytoene synthase (*psy*) and phytoene desaturase (*crt*I) genes showed evidence of T-DNA rearrangements, and in the majority of cases the rearrangements occurred in the *crt*I expression cassette, which was adjacent to the right T-DNA border.

7.3.3 T-DNA Integration Mechanism

A number of groups have investigated the structure of genomic/T-DNA and T-DNA/T-DNA junctions in plants and have concluded that integration occurs by illegitimate recombination (see Salomon and Puchta 1998; Somers and Makarevistch 2004). A strand invasion mechanism of integration has been proposed (reviewed by Tinland 1996), in which the 3′ end of the T-strand initiates the

integration process by hybridizing to a short region of homology in the plant genome, the second strand being completed by primer extension of the plant DNA. Other models suggest conversion of the T-strand into a double-stranded intermediate, which integrates at the site of naturally occurring chromosome breaks via double-strand DNA break repair. This is supported by experiments that show transformation efficiency increases following UV irradiation, which generates nicks and breaks in genomic DNA. However, since T-DNA integration occurs normally, if less frequently, in DNA repair mutants, it is possible both mechanisms occur simultaneously.

DNA repair models argue that proteins encoded by the host plant have a much more important role in T-DNA integration than *Agrobacterium* proteins, such as VirD2, which are imported into the plant with the T-DNA. However, since VirD2 protein remains covalently attached to the 5′ end of the T-strand during transfer it is also likely to influence integration (Ward and Barnes 1988). In an in vitro assay, VirD2 can ligate together a cleaved T-DNA border sequence but cannot ligate T-DNA to other genomic targets unless plant cell extracts are also present (Pansegrau et al. 1993; Ziemienowicz et al. 2000).

Plant proteins are certainly required for integration, as a number of *Arabidopsis* mutants have been identified that are deficient for T-DNA insertion. The role of DNA strand break repair in T-DNA integration was supported by the discovery of *Arabidopsis* mutants *uvh1* and *rad5*, which are hypersensitive to UV and gamma irradiation, respectively, and show a low frequency of stable transformation by *Agrobacterium*. Since these mutants showed normal levels of transient expression, it was suggested that they caused deficiencies in the repair of radiation-induced breaks and that break repair is essential for T-DNA integration (Sonti et al. 1995). However, Nam et al. (1998) showed that *uvh1* is no less transformation proficient than wild-type plants and that *rad5* is deficient for both transient and stable transformation, indicating that the dysfunction affects a process occurring much earlier than T-DNA integration. Other mutants resistant to *Agrobacterium* transformation (*rat* mutants) have been identified, and five are thought to be blocked at the point of T-DNA integration (Nam et al. 1999). One of the corresponding genes, *rat5*, encodes a histone protein, suggesting that efficient T-DNA integration is dependent on chromatin structure at the integration site.

Much can be learned about the T-DNA integration mechanism by the inspection of borders, especially the borders between adjacent T-DNA sequences in multicopy insertions. The formation of heterodimers during cotransformation argues in favor of T-DNA concatemerization prior to integration. Although inverted repeats around the right border are often precise, those around the left border and those separating direct T-DNA repeats are often characterized by the insertion of variable-sized regions of filler DNA, which may be derived from the T-DNA sequence or from plant genomic DNA (De Buck et al. 1999; Kumar and Fladung 2000, 2002). This suggests either the simultaneous integration of multiple T-DNAs at a single locus, or a two-phase mechanism, in which a primary T-DNA integration event stimulates further secondary integrations in the same area, similar to those proposed for particle bombardment (see Sect. 7.4.3). Zhu et al. (2006) carried out a

comprehensive study of T-DNA border characteristics in a population of transgenic rice plants including 156 T-DNA/genomic DNA junctions, 69 T-DNA/T-DNA junctions, and 11 T-DNA/vector backbone junctions, which included 171 left borders and 134 right borders. Conserved cleavage was observed in 6% of left and 43% of right borders, microhomology was observed in 58% of T-DNA/genomic DNA, 43% of T-DNA/T-DNA, and 82% of T-DNA/vector junctions, mostly at left borders, and about one-third of the T-DNA/genomic DNA and T-DNA/T-DNA junctions showed evidence of filler DNA (up to 344 bp). This was derived mainly from the T-DNA region adjacent to the breakpoint and/or from the rice genomic DNA flanking the T-DNA integration site, with T-DNA/T-DNA filler DNA showing the greatest complexity. Interestingly, when two T-DNAs were integrated in the inverted repeat configuration, significant truncation was always observed in one of the two T-DNAs, whereas with direct repeat configuration, large truncations were rare. These data suggested no single integration mechanism could account for all observations, but the presence of filler DNA at many of the junctions argued that a template-driven DNA synthesis mechanism must be involved, probably reflecting abortive gap repair through a synthesis-dependent strand annealing (SDSA) process. For example, a 16-bp filler DNA that was identical to a reversed T-DNA fragment close to the right border was observed at a left/right border junction. This was most likely produced by invasion of the 3′ end of a T-DNA into another T-DNA near the right border in reverse orientation during recombination or interaction of these two T-DNAs. When the right border is not protected by VirD2, it is subjected to 5′ exonuclease degradation that creates a free 3′ end in its complementary strand. This 3′ end is able to invade another template to produce filler DNA at the right border end. Multiple template switches can be used to explain the origin of complex filler DNA structures, and longer regions of homologous DNA might reflect a single-strand annealing process in addition to SDSA.

7.3.4 Cotransfer of Vector Backbone Sequences

Agrobacterium was initially thought to be a clean transformation method because the T-DNA is more or less precisely defined (cleavage occurs at a precise position within the right border repeat and the cleavage site at the left border varies by about 100 bp). However, it is now evident that T-DNA transfer is much less precise than originally envisaged, and 25–30% of transformants may commonly contain vector sequences linked to the T-DNA insert, indicating that the cleavage reaction during T-DNA transfer can be rather inefficient (Martineau et al. 1994; Rai et al. 2007). Other studies have shown that in some systems, the frequency of vector backbone transfer can reach as high as 66% (Afolabi et al. 2004). Ramanathan and Veluthambi (1995) constructed binary vectors, in which the selectable marker was located outside the left T-DNA border. In accordance with the T-DNA transfer mechanisms discussed above, it was considered likely that this strategy would catch those transfer events in which transfer, initiated at the right border, overran the left

border and terminated somewhere along the plasmid backbone. Surprisingly, these investigators found that none of the transformants contained any T-DNA sequences, indicating that, in these cases, transfer had initiated at the left border and had proceeded around the plasmid away from the T-DNA, presumably breaking off before completing the circuit and reaching the right border sequence. Further investigations have shown that vector sequence transfer is probably a very common event, occasionally involving the entire plasmid backbone with or without the T-DNA. Concatemers of the entire binary vector have also been seen, indicating that transfer does not necessarily terminate at the T-DNA border even after one or more complete circuits of the vector (Wenck et al. 1997). The exact structure of the insert and the presence or absence of T-DNA in recovered transgenic plants depend of course on the position of the selectable marker. In the strategy of Ramanathan and Veluthambi (1995), the external position of the marker allowed non-T-DNA transformants to be recovered. The experiments carried out by Kononov et al. (1997) are particularly informative because this group constructed binary vectors, in which a selectable marker was present within the T-DNA and a screenable marker gene was present outside either the left or the right borders of the T-DNA. Over 200 transformants were obtained under selection and 75% were shown to carry the external screenable marker gene *gus*A. Interestingly, both vectors appeared to transfer *gus*A to the plant genome with equal efficiency, suggesting that T-DNA transfer could be initiated nonselectively at either the left or right borders. It is also notable that Kononov and colleagues used three alternative *Agrobacterium* strains: LBA4404, GV3101, and EHA105, representing octopine, nopaline, and agropine-type virulence functions, respectively. There were no significant differences among the strains in terms of the frequency of vector sequence transfer. Finally, these investigators reported that they could also detect independent integration events involving plasmid backbone sequences alone. Since the selectable marker in these experiments was located within the T-DNA, such vector-only integrations must have occurred in addition to the T-DNA-linked integration events. This indicates that in the natural course of transformation, many vector-only integration events may occur, but will not be recovered under selection. It is also likely that vector-only integration events occur, undetected, in many plant transformation experiments.

7.4 Locus Structure in Plants Transformed by Direct DNA Transfer

7.4.1 Principles of Gene Transfer

A number of direct DNA transfer methods have been developed to transform plants recalcitrant to *Agrobacterium*-mediated transformation (reviewed by Twyman et al. 2002). Among these methods, particle bombardment has become the most successful

because it is based on purely mechanical principles and is therefore not dependent on the biological factors that restrict the *Agrobacterium* "host range". Particle bombardment works with any plant species, variety, and explant, leaving the regeneration of fertile plants rather than the DNA transfer process itself as the only significant bottleneck (Altpeter et al. 2005). Particle bombardment involves the acceleration of small DNA-coated metal particles (either gold or tungsten) into plant tissue with sufficient force to break through the cell wall and membrane. Some of the particles reach the nucleus, where the DNA is released, probably by a simple diffusion mechanism (Altpeter et al. 2005). Notably, the foreign DNA entering a bombarded cell is naked, double-stranded, and competent for both transient episomal expression and integration into the genome. Transient expression also occurs in the process of *Agrobacterium*-mediated transformation, but the T-strand must first be converted into a double-stranded intermediate (Narasimhulu et al. 1996). Other direct DNA transfer methods are gentler, using chemicals (e.g., PEG, calcium phosphate) or physical methods (e.g., electroporation) to persuade plant protoplasts to take up DNA from the surroundings. However, this DNA must ultimately find its way to the nucleus, and integration occurs in the same way as described below for particle bombardment.

While *Agrobacterium*-mediated transformation involves a number of virulence gene products that must be supplied either on the same plasmid as the T-DNA or on a separate binary vector, particle bombardment has no such requirements because the introduction of DNA is governed entirely by external physical factors (Sanford et al. 1993). For convenience, therefore, vectors used for direct transfer are generally based on bacterial cloning plasmids, and incorporate a selectable marker and origin of replication functional in bacteria. In *Agrobacterium*-mediated transformation, the T-DNA is meant to be excised from the vector during the transformation process, and any vector backbone transfer results from inefficient processing. In contrast, there is no such processing in particle bombardment, although this can be achieved before transformation by excising the linear cassette, purifying it, and using just this cassette as the substrate for coating the metal particles (Fu et al. 2000). This practice has the interesting side effect of reducing the complexity of transgene loci as discussed in Sect. 7.4.3.

7.4.2 Transgenic Locus Structure

There have been few studies, in which integration sites generated by particle bombardment have been carefully mapped, so whether there is a preference for inserting in transcription units is not so clear as in the case of T-DNA integration. The variable nature of the input DNA linear cassette sequences should remove any sequence-dependent bias (compared to the preserved ends of the T-DNA), but as discussed in Sect. 7.3.4, the T-DNA cutting process can overshoot the left and/or right border, so it is likely that the substrates for integration are equally variable in T-DNA transfer. Chen et al. (1998) noted that in rice plants cotransformed with up

to 13 plasmids, there was no preference for the integration of particular transgenes, indicating that the insertion mechanisms operated independent of input gene sequence. Svitashev et al. (2000) showed by FISH analysis of transgenic oat that integration occurred randomly with respect to the A/D and C genomes, and there was no preference for chromosomes from a particular genome. However, the majority of integration events occurred at telomeric and subtelomeric regions, which are typically gene-rich. It is also possible that this preferential integration may reflect some aspect of the nuclear architecture in oat rather than the distribution of genes, since FISH analysis of a limited number of transgenic wheat plants generated by particle bombardment showed no preferential integration in terms of the chromosome region. In the commercial papaya variety SUNUP, five of the six sequences flanking the three indentified transgene integration sites were genomic copies of plastid genes (Ming et al. 2008). Since the plastid genome is more AT-rich than typical genomic DNA, this both supports the possibility of preferential insertion in or near genes and matches the observation of AT-rich sequences at the insertion sites in other transgenic lines generated by *Agrobacterium* and direct DNA transfer.

Unlike the situation with *Agrobacterium*-mediated transformation, a vast literature has accumulated on the structure and complexity of transgenic loci generated by direct DNA transfer, particularly particle bombardment. As discussed in Sect. 7.3.2, T-DNA integration usually occurs with a low copy number, rarely exceeding five copies, and the T-DNA is generally intact. In contrast, direct DNA transfer often generates much larger transgenic loci. Typically, these contain from 1 to 20 transgene copies (e.g., Klein et al. 1987; Register et al. 1994; Cooley et al. 1995; Dai et al. 2001; Travella et al. 2005). The structure of such loci is highly variable, comprising single copies, tandem or inverted repeats, concatemers, intact transgenes, truncated and rearranged sequences, and interspersed genomic DNA. The analysis of transgenic cereal plants by FISH to extended DNA fibers, metaphase chromosomes, and interphase chromatin has revealed a higher-order level of organization where discrete integration events are interspersed by large fragments of genomic DNA, up to several hundred kilobase pairs in length. This organization, which generates immense (megabase) transgenic loci, appears unique to particle bombardment, and could thus reflect the nature of the transformation process itself (see Sect. 7.4.3).

A useful overview of transgene organization in wheat has been reported by Jackson et al. (2001) using the technique of fiber-FISH. This study showed that transgene loci in bombarded wheat plants can be organized in three ways. The simplest arrangement, described as a type III locus, is characterized by a single discrete fiber-FISH signal corresponding approximately to the length of the transforming plasmid. This represents an intact, single copy transgene. Type III loci may be present uniquely in a given plant, or there may be two or more unlinked inserts representing multiple genetic loci. These two possibilities can be distinguished by FISH to metaphase chromosomes and genetic segregation analysis. Other loci, described as type I loci, are longer than the single plasmid copy yet still generate a continuous signal along the extended chromatin fiber. For example, Jackson and colleagues reported a type I transgenic locus with a continuous signal of 77 kb,

representing 11 contiguous plasmid copies. Such loci represent concatemers of the transforming plasmid and are characterized by the absence of intervening genomic DNA. The presence of concatemers can also be confirmed by Southern blot analysis and sequencing across plasmid/plasmid junctions. Loci thus characterized have been described by Kohli et al. (2003) as "transgene arrays" (Fig. 7.1). Until the late 1990s, both head-to-head and head-to-tail concatemers had been sporadically reported in the literature, but it was unusual for the structure of a transgenic locus to be examined in such detail. Concatemerization is probably quite a common phenomenon. Extensive concatemerization, for example, has been reported by Hadi et al. (1996) in transgenic soybean simultaneously transformed with 12 different plasmid vectors. The remaining class of locus (type II) is the most complex. It is characterized by fiber-FISH signals that extend for a significant distance (>100 kb) over the chromosome, but which are punctuated regularly by intervening segments

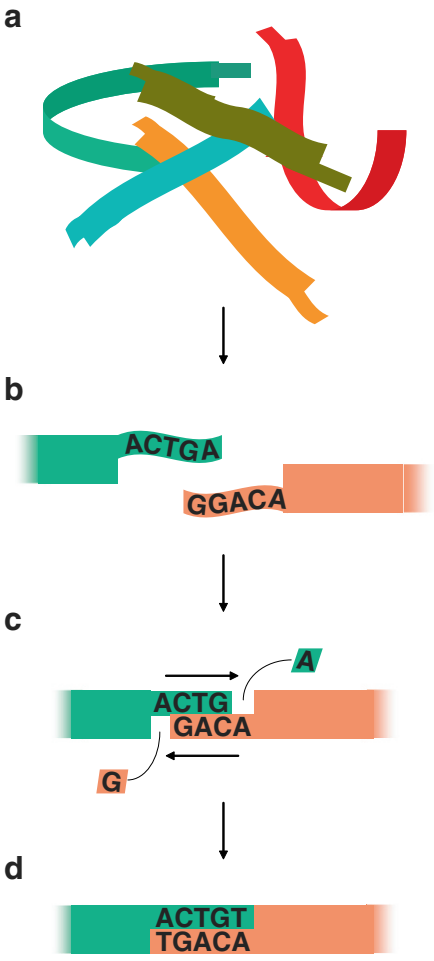

Fig. 7.1 Mechanism for transgene integration at regions of microhomology. A mixture of DNA fragments with ragged ends (**a**) interacts with a double-stranded DNA break with partially complementary ragged ends (**b**). Repair synthesis across the gap (**c**) generates a recombination junction (**d**) which may be completely conserved if the homology is precise, or may involve either the loss of terminal sequences or the insertion of filler DNA if the homology is partial

of genomic DNA (no signals). Such loci have also been identified in transgenic oat, rice, barley, and maize. In barley, for example, some transgene integration sites showed simple structures represented by one single FISH signal, whereas in others it was possible to identify up to six spots organized in a linked cluster and separated by barley DNA, making the locus several megabase pairs long (Travella et al. 2005). Kohli et al. (2003) defined such loci as "transgene clusters." Type II loci contain genomic interspersions ranging from a few tens of base pairs to approximately 10 kb (Fig. 7.2). Although dispersed over a distance of up to 100 kbp, such

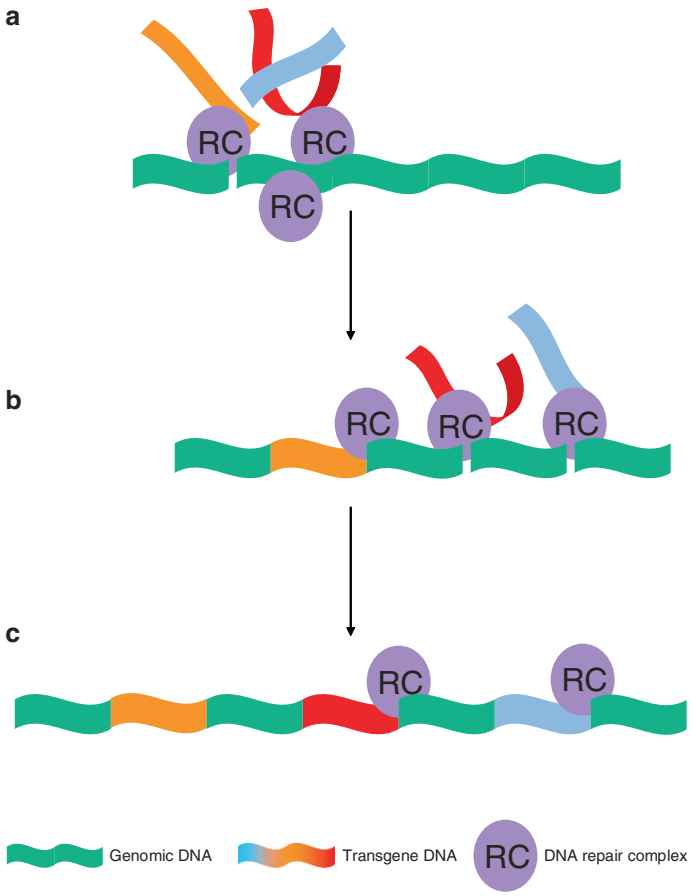

Fig. 7.2 Explanation for the formation of transgene arrays and transgene clusters interspersed with genomic DNA. A mixture of DNA fragments interacts with a double-stranded DNA break where a repair complex has already assembled (**a**). The repair complex may stitch together DNA fragments to form concatemers prior to integration, or may integrate single copies. The first integration event stimulates further repair complex activity nearby, resulting in additional nicks and breaks in the genomic DNA that act as further integration sites (**b**). This results in a cluster of transgenes (single copies and concatemers) interspersed with short regions of genomic DNA (**c**)

loci would still be expected to generate a single discrete signal if FISH analysis was applied to metaphase chromosomes due to the low resolution of this technique. However, the analysis of metaphase wheat chromosomes by FISH has revealed an unexpected third level of organization, involving the dispersion of transgene arrays and/or clusters over a larger area comprising megabase pairs of DNA (Abranches et al. 2000). Instead of discrete spots for each transgenic locus, two or more separable FISH signals were often observed, restricted to a particular chromosome region (Fig. 7.3). To be separable at the cytogenetic level, each signal must be interspersed by hundreds of kilobase pairs of genomic DNA. Similarly large genomic interspersions have been seen in transgenic oat (Svitashev and Somers 2001).

Interestingly, FISH analysis of interphase chromatin and metaphase chromosomes in the same transgenic wheat plants showed that the dispersed metaphase FISH signals could come together at interphase (Abranches et al. 2000). Occasionally, the signals clustered at a specific region of the nucleus but remained discrete.

7.4.3 Mechanisms of Transgene Integration

The analysis of plasmid/plasmid and plasmid/genomic junctions in transgenic plants generated by particle bombardment reveals features characteristic of illegitimate recombination similar to those seen for T-DNA junctions, suggesting that the same overall integration mechanisms may be involved (Svitashev et al. 2002). For example, such junctions are characterized by regions of microhomology, filler DNA, trimming of the DNA ends so sequences are lost and AT-rich elements surrounding the junction site, with similarity to topoisomerase I binding/cleavage sites (Fig. 7.1). In the analysis of multiple plasmid/plasmid junctions in 12 transgenic rice lines, Kohli et al. (1998) observed ten plants with microhomology at the junctions and two plants where junctions appeared to be generated by blunt ligation, with no overlap. A similar ratio of conserved end-joining to microhomology-mediated recombination was observed by Gorbunova and Levy (1997) and Salomon and Puchta (1998). Topoisomerase I sites were also observed adjacent to 10 out of 12 junctions characterized in transgenic *Arabidopsis* plants generated by particle bombardment (Sawasaki et al. 1998) and in four of the six junctions in the commercial SUNUP variety of papaya (Ming et al. 2008). Illegitimate recombination, therefore, appears to be responsible both for the integration of foreign DNA into the plant genome and the linking of multiple plasmid copies, which is similar to the mechanism proposed for T-DNA integration (Sect. 7.3.3).

Any model for transgene integration following particle bombardment must take into account the three-tier organization revealed in transgenic cereals: contiguous arrays, interspersed clusters, and widely dispersed FISH signals. Two-phase transgene integration mechanisms have been proposed to explain the first two levels of organization, and in such models concatemerization is proposed to occur prior to integration, while interspersion occurs during the integration process (Kohli et al.

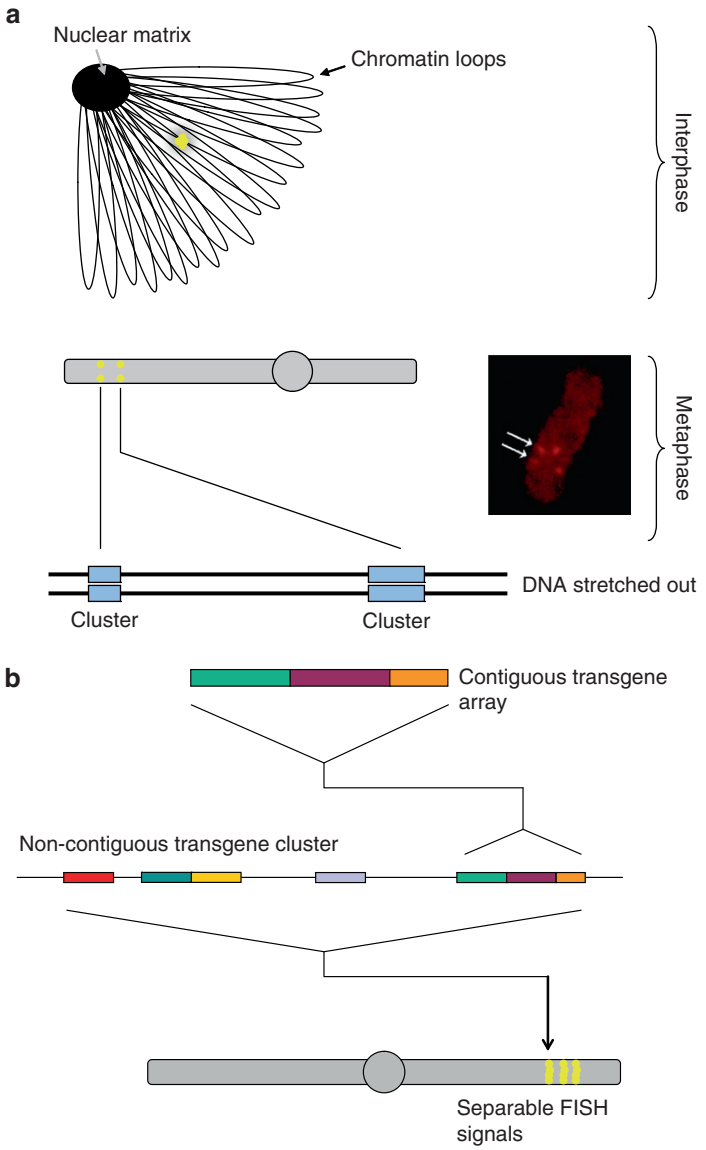

Fig. 7.3 Higher order transgene locus organization in cereals transformed by particle bombardment. Transformation occurs during interphase, when the chromatin is distributed into specific nuclear zones and territories. If a metal particle causes localized damage, DNA repair complexes will form at these sites and initiate transgene integration (**a**). During metaphase, when FISH analysis is generally carried out, loci that are brought together in interphase may be separated, resulting in multiple signals from the same transformation event (**b**). If the DNA were stretched out, this would reveal large (megabase) interspersed sequences, which have also been observed in fiber-FISH experiments

1998; Pawlowski and Somers 1998; Svitashev et al. 2002) (Fig. 7.2). In each model, penetration of the cell is proposed to elicit a wound response, which would include the induction of DNA repair enzymes, such as nucleases and ligases. The presence of these enzymes and an excess of foreign DNA would result in the linking together of several copies to form concatemers, which would be the substrates for integration. This might be stimulated by homology between individual copies of transforming plasmids, and "backbone" homology might also result in the concatemerization of plasmids carrying different transgenes in cotransformation experiments. However, cotransformation and cointegration were also shown to occur when two nonhomologous minimal cassettes were used for transformation, so homology might not be as important as the presence of free DNA ends (Fu et al. 2000). Kohli et al. (1998) suggested that transgene clusters arise in a second phase where a primary integration event occurring by illegitimate recombination at a chromosome break generates a "hot-spot" for further integration events in the same area. This might be due, for example, to the presence of local repair complexes that slide along the DNA and introduce nicks which can be exploited by more foreign DNA (Gelvin 1998). Pawlowski and Somers (1998) suggested an alternative second phase where a number of discrete transgene concatemers integrate simultaneously at a site containing multiple replication forks. Although there is no direct evidence for either mechanism, it is interesting to note that DNA integration is stimulated in rapidly dividing cells and is blocked in *Arabidopsis* mutants lacking essential components of the DNA recombination machinery.

The higher order organization of transgenic loci observed by metaphase FISH is thus far unique to particle bombardment and demands a model which takes into account the three-dimensional structure of the nucleus. In one scenario, it is possible that the transformation event affects a local region of the interphase nucleus. For example, it is possible that the metal particle causes damage to a particular area of chromatin, which is arranged in loops attached to the nuclear matrix. If the particle "skims" several loops, there will be regions of DNA damage close together in trans, but widely separated in the cis configuration were the DNA to be stretched out (Fig. 7.3). Each of these sites could act as a nucleation point where foreign DNA diffusing from the metal particle is used to patch up double-strand breaks, generating widely separated arrays and/or clusters (Abranches et al. 2000; Kohli et al. 2003). In support of this induced break and repair model, Svitashev et al. (2000) have shown that in six of 25 transgenic oat plants generated by particle bombardment, transgene integration sites were associated with rearranged chromosomes. This suggests that DNA breaks caused by incoming particles are repaired with foreign DNA and may also result in deletions, inversions, and translocations involving genomic DNA. Chromosomal rearrangements have also occasionally been seen associated with T-DNA integration (Nacry et al. 1998; Laufs et al. 1999).

The model above suggests that dispersed metaphase signals come together at interphase due to the physical position of the transgenic loci at the moment of transformation. In another scenario, the bringing together of transgene sites at interphase could represent recruitment, for example to a common transcription

factory in the nucleus (Cook 1999). A further scenario involves transgenes that are brought together by virtue of their homology, perhaps as a consequence of their initial placing in the same region of the nucleus. This is an exciting prospect because the coincidence of FISH signals observed in wheat nuclei could represent a physical basis of the postulated DNA-DNA interactions that precede transcriptional transgene silencing in plants (see below).

Transgene rearrangements following particle bombardment have been widely reported in the literature, and many publications repeat the "lore" that direct DNA transfer is more likely than T-DNA transfer to generate complex rearranged loci. The number of rearrangements that can be detected depends entirely on the resolution of the method being used. Thus, careful analysis of locus structure by Southern blot hybridization, PCR, and DNA sequencing has recently shown that rearrangements may be more widespread than first envisaged in both transformation methods. The analysis of transgenic oat loci by Somers and colleagues has shown that transgene rearrangements can be extensive and extremely complex, with multiple small insertions, inversions, and deletions within any transgene, plus the presence of filler DNA (Svitashev et al. 2000). In maize, Mehlo et al. (2000) noted that every single plant among the population they analyzed showed some form of rearrangement, and they speculated that undetected "minor" rearrangements could be responsible for many instances of transgene silencing otherwise attributed to epigenetic effects (see Sect. 7.5). In particular, certain transgene rearrangements were not detectable by Southern blot hybridization because they were too subtle, but they could be picked up by PCR and sequencing. Since in most cases, Southern blot hybridization is used to determine whether a given locus is intact or rearranged, this suggests caution should be used in relying on such results, since only "major" rearrangements can be detected in this manner.

Few researchers have characterized transgene rearrangements in detail, but work by Kohli et al. (1999) has shown that rearrangements may involve palindromic sequences in the transforming plasmid, which tend for form three-dimensional structures such as hairpins and cruciforms. These investigators characterized 12 transgenic rice lines, transformed by particle bombardment, which had been shown to contain rearranged transgenes. Interestingly, they found that an imperfect palindrome in the CaMV 35S promoter was involved in one-third of all rearrangements, i.e., the sequence of this palindrome was adjacent to the rearrangement junction. Similar phenomena have been noted in T-DNA transformants containing the same promoter (Sect. 7.3.2). This sequence has the ability to adopt a cruciform secondary structure, which may stimulate recombination events. Many other promoters contain palindromic sequences of variable length within 100 bp of the transcription start site. The DNA secondary structures formed at these sites enable DNA-protein interactions for transcription under normal circumstances, but may also participate in aberrant recombination events. The fully sequenced papaya genome (Ming et al. 2008) also revealed a number of previously unidentified transgene rearrangements, i.e., a 1,533-bp fragment composed of a truncated, nonfunctional *tet*A gene and flanking vector backbone sequence, and a 290-bp nonfunctional fragment of the *npt*II gene, in addition to the intact, primary transgene conferring virus resistance.

7.5 Locus Structure and Transgene Stability

One of the most profound insights to come from the detailed analysis of transgene loci over the last decade is that many integrated transgenes contain minor rearrangements. As discussed above, these are difficult to pinpoint using low-resolution detection methods such as Southern blot hybridization and FISH, but high-resolution methods such as sequencing are rarely used as a routine analysis tool. Therefore, the impact of physical rearrangements on transgene expression is likely to be vastly underestimated, since unstable loci are often blamed on epigenetic phenomena with no further analysis to draw confirmatory evidence.

There are many factors that influence transgene stability, and these lead to highly variable expression within populations of plants generated in the same gene transfer experiment. One of the most important factors is the position effect, which reflects the influence of genomic DNA surrounding the site of transgene integration (Wilson et al. 1990). Another is the structure of the locus, including the number of transgene copies, their intactness, and their relative arrangement, which influences the likelihood of physical interactions and further recombination within the locus (physical instability) and the induction of silencing through DNA methylation and/or the production of aberrant RNA species from the locus (Heinrichs 2008).

7.5.1 Position Effects

Specific position effects result from the influence of local regulatory elements on the transgene. For example, an integrated transgene may come under the influence of a nearby enhancer, such that its expression profile is modified. The effect is transgene-specific because the enhancer interacts with regulatory elements in the transformation construct to control transcription; hence, the final expression pattern reflects the combined influence of both regulatory elements. Such effects are clearly revealed by entrapment constructs, which contain minimal control sequences linked to a visible marker gene and therefore "report" the activity of local regulatory elements (e.g., Goldsbrough and Bevan 1991).

As well as specific position effects governed by local regulatory elements, nonspecific position effects can also be generated by the surrounding chromatin architecture. Where the local environment is favorable for transgene expression, i.e., a positive position effect, it is generally taken for granted. However, nonspecific and repressive position effects reflect the integration of the transgene into a chromosomal region containing repressed chromatin (heterochromatin). The molecular features of heterochromatin, including its characteristic nucleosome structure, deacetylated histones, and hypermethylated DNA, spread into the transgene causing it to be inactivated (Pikaart et al. 1998). Analysis of the genomic context of silenced transgenes suggests that integration in the vicinity of certain repetitive DNA sequences, such as microsatellites and retrotransposon remnants, may predispose the transgene to silencing (Tanako et al. 1997). The chromosomal

location is important, since in many plants, the genes are restricted to a small portion of the genome known as gene space, and the majority of the DNA is taken up by repetitive sequences. Thus, stable transgene expression has been associated with gene-rich telomeric and subtelomeric integration sites, whereas mosaic expression and silencing occurs at predominantly heterochromatic centromeric loci.

A third type of position effect reflects the tolerance of the surrounding DNA for "invasion" by foreign DNA. In this case, the effect is not automatic (as above) but is set off by the presence of the transgene. It appears that certain sequences can trigger de novo methylation, perhaps because the GC-content or sequence architecture is recognized as abnormal (reviewed by Kumpatla et al. 1998). Prokaryotic DNA may be recognized in this manner, since silencing is often associated with the presence of prokaryotic vector backbone DNA, particular binary vector sequences joining T-DNA to genomic DNA (Iglesias et al. 1997).

7.5.2 Locus Structure Effects

At least three aspects of locus structure influence transgene stability and expression: copy number, intactness, and arrangement. It is natural to assume that increasing the number of copies of a particular transgene will lead to an increase in the level of its product. However, even from the earliest plant transformation experiments, it was appreciated that multiple transgene copies could induce transgene silencing and that the phenomenon was associated with DNA methylation at the transgenic locus (e.g., Gelvin et al. 1983; Hepburn et al. 1983). A strikingly visual demonstration of this effect was provided by introducing the maize *AI* gene into mutant petunia plants with white flowers. Expression of the transgene resulted in pelargonidin production, generating a red pigment. However, it was shown that red flowers generally appeared on plants with single copy transgenes, while plants with multiple transgene copies had white or variegated flowers. Where transgene silencing had occurred, increased methylation of the transgene DNA was observed (Meyer et al. 1992). Similarly, it was thought that the amount of pigment in wild-type petunia flowers could be increased by introducing extra copies of the chalcone synthase (*chs*) gene (Napoli et al. 1990). Chalcone synthase converts coumaryol-CoA and 3-malonyl-CoA into chalcone, a precursor of anthocyanin pigments. The presence of multiple transgene copies was expected to increase the level of enzyme and hence cause stronger flower pigmentation. However, in about 50% of the plants recovered from the experiment, exactly the opposite effect was observed. The flowers were either pure white, or variegated with purple and white sectors. It appeared that integration of multiple copies of the transgene led not only to the suppression of transgene expression, but also to the cosuppression of the homologous endogenous gene.

Rooke et al. (2003) looked at the integration, inheritance, and expression of transgenes in six transgenic wheat lines generated by particle bombardment with two plasmids containing genes encoding a glutelin subunit and a selectable marker, respectively. Transgene insertion number ranged from 1 to 15, with most lines carrying multiple copies consistent with previous reports (Becker et al. 1994;

Blechl and Anderson 1996; Srivastava et al. 1996; Stoger et al. 1998; Cannell et al. 1999). Four of the transgenic loci were clusters interspersed with genomic DNA, in some cases enough to allow independent segregation which contrasts with previous reports in which cointegration and cosegregation were the norm, and independent segregation rare (Stoger et al. 1998). There was no evidence for a direct correlation between transgene copy number and expression level, and no evidence for cosuppression of endogenous glutelin genes even in multicopy lines. The presence of multiple transgene copies has been implicated in transgene silencing, but other studies in cereals have shown that multiple copies do not necessarily lead to silencing and can even enhance expression levels in proportion to copy number (Stoger et al. 1998; Gahakwa et al. 2000). In contrast, Spencer et al. (1992) failed to recover progeny expressing the marker transgenes from maize lines containing more than five or six copies of the integrated plasmid, while Cannell et al. (1999) observed silencing or a gradual reduction in marker gene expression over three generations of transgenic wheat lines. It has been suggested that the production of lines with single transgenes or low copy numbers is desirable as such lines may be more stable and less likely to exhibit transgene silencing (Finnegan and McElroy 1994), but with several studies reporting contrary results, this may indicate that chance plays a role in the impact of copy number, perhaps reflecting insertion events near to boundary elements (Sect. 7.5.3).

Variation in transgene expression levels can also result from uncontrolled differences in experimental protocols reflecting gene-environment interactions, which means that proper comparisons between transgenic lines should take place in a standardized environment. To study sources of spurious variation, transgene expression levels were quantified over five homozygous generations in two independent transgenic rice lines created by particle bombardment (James et al. 2004a, b). Both lines contained the same *gus*A expression unit which was stably inherited, and all plants were cultured and sampled using previously developed standardized protocols. Plants representative of each generation (T_2–T_6) were grown either all together or across several different growth periods. Where the plants were grown and characterized independently, the amount of extraneous variation in transgene expression levels was up to three-fold higher than in plants grown and analyzed together. This study therefore provided important evidence that the growth and analysis of all plants from all generations together, using standard operating procedures (SOP), can reduce extraneous variation associated with transgene expression and is the key to improving the reproducibility of transgenic studies conducted over multiple generations (James et al. 2004a, b).

7.5.3 Overcoming Position and Locus Structure Effects by Buffering the Transgene

As discussed earlier, analysis of the genomic context of transgene integration sites has shown that silenced transgenes are often surrounded by repetitive elements, which are sequestered into repressed chromatin. The same studies have also shown

that stably expressed transgenes are often associated with matrix attachment regions (MARs) (Iglesias et al. 1997). MARS are AT-rich elements that attach chromatin to the nuclear matrix and organize it into topologically isolated loops (Holmes-Davis and Comai 1998). A number of highly expressed endogenous plant genes have also been shown to be flanked by matrix attachment regions (e.g., Chinn and Comai 1996). One strategy that has been proposed to overcome position effects is therefore to protect or buffer the transgene by flanking it with MARs prior to transformation. In this way, it is hoped that the transgene will form a discrete chromatin loop which will be isolated from surrounding chromatin.

Several experiments have been carried out in which a reporter gene such as *gus*A has been flanked by MARs. Such constructs have been introduced into transgenic plants and compared to populations containing the same reporter gene without MARs (e.g., Mlynarova et al. 1994, 1995, 1996; Van Leeuwen et al. 2001; Mlynárová et al. 2002). Generally MARs do have a positive effect on transgene expression and can significantly reduce position effects, but they cannot rescue all lines and restore full expression. It is acceptable to say that they generally reduce expression variability within a population (e.g., Breyne et al. 1992). Expression may increase as much as five-fold, but some remarkable exceptions include a 25-fold enhancement using a yeast MAR and a 140-fold enhancement using a tobacco MAR in tobacco callus (Allen et al. 1993, 1996).

7.5.4 *Overcoming Position and Locus Structure Effects by Homologous Recombination*

As discussed in Sects 7.3.3 and 7.4.3, transgene integration in higher plants occurs almost universally by illegitimate recombination, which may involve microhomology but is not dependent upon it. Since there is only minimal sequence relationship between the transgene and the genomic region into which it integrates, the experimenter has little control over the integration site. In other systems, notably yeast, homologous recombination is favored over illegitimate recombination if the vector carries a homology region that matches the yeast genome, allowing endogenous genes to be altered by gene targeting (Schiestl and Petes 1991). In the context of controlling transgene integration, it also allows transgenes to be inserted at specific loci, a strategy that should allow favorable sites for transgene integration to be chosen, theoretically abolishing position effects and reducing the complexity of locus structure.

Although widely used in microbial systems, homologous recombination occurs with a very low efficiency in plants (illegitimate recombination occurs about 10^5 times more frequently than homologous recombination, making genuine targeting events difficult to isolate). Only one plant species has been shown to undergo efficient nuclear homologous recombination (the moss *Physcomitrella patens*) and the results in higher plants have been much less impressive, with targeting efficiencies as low as 10^{-6} (Lee et al. 1990; Offringa et al. 1990; Miao and Lam 1995; Risseeuw et al. 1995, 1997; Kempin et al. 1997; Reiss et al. 2000; Hanin et al. 2001).

A transgene has also been repaired by homologous recombination in tobacco (Paszkowski et al. 1988). More recently, promising results have been achieved using a T-DNA-mediated gene-targeting strategy involving a long homology region in combination with a strong counterselectable marker in rice (Terada et al. 2002). Targeting frequencies of up to 1% have been achieved using this system (reviewed by Ida and Terada 2004 and Cotsaftis and Guiderdoni 2005). Gene targeting has also been reported recently in maize (D'Halluin et al. 2008).

There has also been interest in the use of zinc-finger endonucleases to make targeted double-strand breaks in the plant genome, so that homologous recombination is favored at such sites (Kumar et al. 2005). The modular nature of zinc-finger transcription factors means that recombinant DNA technology can be used to "mix and match" these DNA-binding domains to create recombinant proteins with unique sequence specificities. Zinc fingers are motifs approximately 30 amino acids in length which coordinate a Zn^{2+} ion and bind to DNA sequences three base pairs long. Combining different zinc fingers in series allows proteins to be tailor made to bind longer DNA sequences. When a nonspecific DNA endonuclease is incorporated into such a protein, it becomes a targeted DNA cutting tool (Lloyd et al. 2005; Wright et al. 2005; Zeevi et al. 2008; Cai et al. 2009).

7.5.5 Overcoming Position and Locus Structure Effects by Organelle Transformation

The plant cell contains not only a nuclear genome, but also organellar genomes in the chloroplasts and mitochondria. The chloroplast is a useful target for gene transfer because tens of thousands of chloroplasts may be present in a single plant cell, and each chloroplast may contain multiple copies of its chromosome. Genetic engineering of the plastid genome offers several advantages over nuclear transformation including that integration occurs by homologous recombination, the high copy number of transgenes in a homoplasmic cell, and the absence of gene silencing phenomena due to the lack of position and locus structure effects, and the absence of DNA methylation in the plastid genome (Daniell et al. 2005; Daniell 2006; Bock 2007). The recombination machinery is very active in chloroplasts and can induce rearrangements, as observed in some of the first tobacco transformants generated with the *aad*A selectable marker (Svab and Maliga 1993). The stability of a plastid transgene has been evaluated in soybean transformants over six generations. These transformants had integrated the *aad*A selection cassette in the intergenic region between the *rps*12/7 and *trn*V genes. Three independent homoplasmic T_0 transformation events were selected and ten plants from each event propagated to generation T_5 in the absence of selection pressure. Neither transgene rearrangement nor wild-type plastids were detected in generation T_5 by Southern blot analysis. All tested progenies were uniformly resistant to spectinomycin. Therefore, soybean transformants of generations T_0 and T_5 appear to be genetically and phenotypically identical (Dufourmantel et al. 2006).

7.5.6 *Overcoming Position and Locus Structure Effects by Site-specific Recombination*

Site-specific recombination is a form of recombination that occurs at short, specific recognition sites rather than DNA sequences with long regions of homology but no particular sequence specificity, as is the case for homologous recombination. Site-specific recombination is not ubiquitous – indeed different organisms encode their own very specific systems that include the cis-acting recombinogenic sites and the enzymes that recognize them and carry out the recombination event. Therefore, the target sites for site-specific recombination can be introduced easily and unobtrusively into transgenes, but recombination will only occur in a heterologous cell if a source of the specific recombinase enzyme is also supplied. As with the homologous recombination strategy discussed above, position and locus structure effects can be eliminated by introducing foreign DNA at a specific, favorable locus. A number of different site-specific recombination systems have been identified and several have been studied in detail (Sadowski 1993). The most extensively used are Cre recombinase from bacteriophage P1 (Lewandoski and Martin 1997) and FLP recombinase from the yeast *Saccharomyces cerevisiae* (Buchholz et al. 1998). These have been shown to function in many heterologous eukaryotic systems including transgenic plants (Metzger and Feil 1999). Both recombinases recognize 34-bp sites (*lox*P and *FRP*, respectively) comprising a pair of 13-bp inverted repeats surrounding an 8-bp central sequence. *FRP* possesses an additional copy of the 13-bp repeat sequence, although this is nonessential for recombination.

The Cre-*lox*P system has been used most widely in plants, often for controlled transgene excision (particularly selectable marker genes after transformation) but more recently for controlled transgene insertion (Gilbertson 2003; Lyznik et al. 2003; Puchta 2003; Marjanac et al. 2008). Marker genes are usually excised in the T_1 generation once transgene expression is verified, allowing the separately introduced *cre* gene to segregate in T_2 plants. This method has been used in many crops including wheat (Srivastava and Ow 2001, 2003), maize (Kerbach et al. 2005; Djukanovic et al. 2006; Hu et al. 2006; Vega et al. 2007, 2008), rice (Chen et al. 2004; Srivastava et al. 2004; Chawla et al. 2006; Moore and Srivastava 2006; Vega et al. 2008), potato (Kopertekh et al. 2004a,b), and tomato (Gidoni et al. 2003; Coppoolse et al. 2005). Controlled integration has been studied in transgenic plants already engineered to contain recipient *lox*P sites (Srivastava et al. 2004). In this study, three different recipient wheat lines were generated by bombarding plants with the *lox*P sequence, and these were subsequently bombarded with a *gus*A construct also containing flanking *lox*P sequences and a *cre* gene. Following transformation, about 80% of lines contained *gus*A at the recipient site, many with single-copy transgenes and others with concatemers. Both types of locus were stably inherited. There was much less variation in expression among the single copy lines (Srivastava et al. 2004).

Chawla et al. (2006) generated 18 different transgenic rice lines containing a precise single copy of *gus*A at a designated site. In seven of these lines, additional

copies of the transgene integrated at random sites by illegitimate recombination, while 11 showed "clean" integration by site-specific recombination only. The single-copy lines were stable over at least four generations and showed consistent levels of expression, which doubled in homozygous plants. In contrast, the multi-copy lines showed variable expression and some fell victim to transgene silencing. Interestingly, where the site-specific and illegitimate integration loci segregated in later generations, transgene expression was reactivated in the plants carrying the site-specific integration site alone, whereas close linkage between the site-specific and random integration prevented segregation in other lines and the silencing persisted.

An exciting recent development is the GENE DELETOR system, which is a hybrid of the Cre-*lox*P and FLP-*FRT* systems. The GENE DELETOR is based on a fusion recognition site (*lox*P-*FRT*), which is inefficient when both recombinases are expressed but highly efficient when either one of the recombinases is expressed alone, giving up to 100% efficiency in populations of up to 25,000 T_1 transgenic plants (Luo et al. 2007).

Another use for Cre-*lox*P is the simplification of locus structure by resolving multicopy loci to a single transgene copy (Srivastava et al. 1999). A strategy was developed in which the transformation vector contained a transgene flanked by *lox*P sites in an inverted orientation. Regardless of the number of copies integrated between the outermost transgenes, recombination between the outermost sites resolved the integrated molecules into a single copy. The principle was proven by resolving four multicopy loci successfully into single-copy transgenes.

7.5.7 Overcoming Position and Locus Structure Effects Using Minichromosomes

In bacteria, plasmid vectors are maintained as episomal replicons to make cloning and isolating recombinant DNA a simple procedure. When it comes to expressing heterologous genes in eukaryotic cells, episomal vectors are widely used to avoid position effects, hence the development of yeast episomal vectors, yeast artificial chromosomes, mammalian plasmid vectors carrying virus origins of replication (e.g., SV40-based vectors, herpesvirus-based vectors), and plant expression vectors based on plant viruses (all of which replicate episomally). The yeast artificial chromosome system is the most relevant in this context because it allows genes of any size to be introduced into the yeast genome as an independent replicating unit that is treated by the cell as an additional chromosome. YACs comprise a yeast centromere and telomeres, the origin of replication (autonomous replicating sequence) and selectable markers. More recently, analogous systems have been developed to maintain genes as episomal minichromosomes in plants. These have many advantages for plant genetic engineering including the ability to express large transgenes or groups of transgenes, and the ability to rapidly introduce new linkage groups into diverse germplasm.

Carlson et al. (2007) created plant minichromosomes by combining the *DsRed* and *npt*II marker genes with 7–190 kb of maize genomic DNA fragments containing satellites, retroelements, and other repeat sequences commonly found in centromeres. The circular constructs were introduced into embryogenic maize tissue by particle bombardment and transformed cells were regenerated and propagated for several generations without selection. The minichromosomes were maintained as extrachromosomal replicons through mitosis and meiosis, and showed roughly Mendelian segregation ratios (93% transmission as a disome with 100% expected, 39% transmission as a monosome crossed to wild type with 50% expected, and 59% transmission in self-crosses with 75% expected). The *DsRed* reporter gene was expressed over four generations, and Southern blot analysis indicated the genes were intact.

7.6 Epigenetic Silencing Phenomena Resulting From Complex Locus Structures and High-Level Expression

As stated earlier, the earliest plant transformation experiments showed that multiple transgene copies could induce transgene silencing, in some cases associated with the cosuppression of homologous endogenes. Transgene silencing can occur through two overlapping pathways, one acting at the transcriptional level (characterized by the reduction or abolition of transcription from one or more copies of the transgene) and one acting post-transcriptionally (transcription from the silenced locus is required to initiate silencing) (Hammond et al. 2001). Transcriptional silencing is often correlated with increased methylation in the promoter regions of affected loci, and both the methylation and the silencing tend to be heritable through meiosis. Post-transcriptional silencing requires homology in the transcribed regions, which may become methylated, and the silencing effect can be reset at meiosis. Post-transcriptional silencing is also known as RNA silencing.

Transcriptional silencing occurs when transgene repeats somehow act as a trigger for de novo DNA methylation. It has been shown that inverted repeats can form secondary structures that are favored substrates for methylation, and thus it is likely that cis DNA-DNA pairing may be involved in such processes. However, transgene silencing can also occur in trans, i.e., silencing interactions may occur between unlinked loci. This has been shown, for example, in sequential transformation events with homologous transgenes, or where two plant lines carrying homologous transgenes have been crossed (Matzke and Matzke 1990, 1991). In this situation, it is likely that physical interactions between transgenes may occur to mediate silencing, and that DNA methylation may somehow be transferred from one site to another. As discussed in Sect. 7.4.3, FISH studies in transgenic wheat provide tantalizing evidence for such interactions in the interphase nucleus (Abranches et al. 2000). Since the CaMV 35S promoter is frequently used for transgene expression and can form cruciform structures that induce transgene rearrangements (Kohli et al. 1999), it may also play a role in transcriptional

silencing under certain circumstances. Supporting evidence for this has been provided in studies of activation tag lines in which the CaMV 35S enhancer is used as a random insertional mutagen to hyperactivate adjacent genes and generate gain-of-function phenotypes. It has been noted that such screens using T-DNA cassettes containing the enhancer elements from the CaMV 35S promoter return a low frequency of morphological mutants (Chalfun et al. 2003). Detailed analysis revealed a correlation between the number of T-DNA insertion sites, the methylation status of the enhancer sequence and enhancer activity. All plants containing more than a single T-DNA insert were methylated on the enhancer and its activity was reduced, with the amount of methylation and the reduction of enhancer activity correlating with the number of T-DNA copies, particularly those with right border inverted repeats (Chalfun et al. 2003). Even so, methylation was still detected at a lower frequency in plants without right border inverted repeats suggesting other triggers were active in these lines.

A recurring theme in post-transcriptional silencing is the presence of double-stranded RNA. Double-stranded RNA introduced into the plant cell can trigger the catalytic degradation of homologous RNA molecules and the methylation of homologous DNA sequences in the genome (e.g., Tavernarakis et al. 2000). When carried out deliberately through the expression of hairpin RNA constructs, this process (RNA interference, RNAi) is a potent method for silencing individual genes, generating phenocopies of mutant phenotypes (e.g., see McGinnis et al. 2005; Gordon and Waterhouse 2007). It has been suggested that complex multicopy transgenic loci could also generate hairpin dsRNA, e.g., if two transgenes are present as inverted repeats, or if truncation and/or rearrangements (some perhaps undetectable by standard screening methods) generated small, aberrant dsRNA species (Jorgensen et al. 1996; Que et al. 1997; Muskens et al. 2000). Experiments designed to test this hypothesis specifically have shown that inverted repeat T-DNA configurations and arrangements of tandem repeated transgenes may not be sufficient in all cases to trigger transgene silencing (Lechtenberg et al. 2003), whereas many reports show post-transcriptional silencing in plants with intact transgenes. In such cases it has been suggested that the level of transgene expression may be an important trigger, with "runaway expression" resulting in the most potent silencing effects (Lindbo et al. 1993; Vaucheret et al. 1998; Schubert et al. 2004). Experiments comparing the frequency and potency of cosuppression by sense chalcone synthase transgenes driven by different promoters have shown that a strong promoter is required for high-frequency cosuppression of chalcone synthase genes and for the production of the full range of cosuppression phenotypes (Que et al. 1997). Indeed the correlation between transgene copy number and silencing in some systems may reflect the higher expression level in multicopy loci triggering silencing (Schubert et al. 2004) suggesting that transgenic lines escaping this effect may fall below the threshold for triggering silencing (e.g., Stoger et al. 1998).

The expression threshold model accounts for RNA silencing in intact transgenic loci but it is also possible that such loci are prone to silencing because their high expression promotes the formation of more aberrant RNA products than a poorly expressed transgene. If true, the trigger would still be aberrant dsRNA, the same as

produced by complex, rearranged loci, and it should be possible to mitigate the effects and generate plants with extremely high expression levels. Several studies have shown that RNA silencing in transgenic plants is accompanied by the accumulation of incorrectly processed mRNAs (often lacking polyadenylate tails) (e.g., van Eldik et al. 1998; Metzlaff et al. 2000; Wang and Waterhouse 2000; Han and Grierson 2002) and in at least one case it has been shown specifically that tandem repeats can generate small interfering RNAs (Ma and Mitra 2002). Since dsRNA is unlikely to be generated directly from tandem repeats (as opposed to the situation with inverted repeats), the process must involve an RNA-dependent RNA polymerase. In agreement with this, Luo and Chen (2007) found that RNA silencing in transgenic *Arabidopsis* could be induced by three direct repeats of the *gus*A open reading frame, and this was dependent on the RNA-dependent RNA polymerase encoded by *RDR6*. Normal plants transformed with either three tandem copies of *gus*A or a single copy lacking a polyadenylation site were able to silence a normal *gus*A transgene cassette in trans, but there was no silencing in *rdr6* mutants, which also accumulated long RNA molecules corresponding to *gus*A read-through transcripts of various lengths. Therefore, it appears that the read-through of termination sites leading to the production of long RNA products triggers RNA silencing in an RDR6-dependent manner. A further transgenic line containing a *gus*A transgene with two polyadenylation sites produced fewer read-through transcripts, less siRNA, and therefore showed higher levels of GUS activity. Transgene silencing in tandem repeat transgenes may therefore be triggered by a defense mechanism that evolved to reduce errors caused by read-through transcription (Luo and Chen 2007).

7.7 Conclusions

Transgene integration following *Agrobacterium*-mediated transformation and direct DNA transfer occur by very similar mechanisms, involving illegitimate recombination between genomic DNA and invading transgene DNA strands, and the repair of double-stranded breaks in the host genome. There is often microhomology between the recombining partners, although direct blunt end ligation also occurs. Both transformation methods induce a wound response, resulting in the activation of nucleases, ligases, and recombinases in the host cell. The foreign DNA is simultaneously degraded and concatemerized resulting in transgene arrays containing intact and/or truncated and rearranged copies. Several integration events may occur simultaneously at a cluster of replication forks, or a primary integration event may stimulate further integrations in the local area. Regardless of the mechanism, the result is a transgene cluster interspersed with genomic DNA. In the case of particle bombardment, clusters and arrays may be widely dispersed, generating very large transgenic loci. The position of transgene integration is essentially random within the “gene space” of the plant species. The transgene is thus subject to position effects which may influence its expression, resulting in some cases in transcriptional silencing as the new DNA is sequestered into the

surrounding chromatin. The structure of a transgenic locus may also induce silencing via a number of mechanisms. These include de novo DNA methylation in response to DNA-DNA interactions, the expression of aberrant RNA species (particularly small hairpin RNAs) from truncated and rearranged transgenes or partial transgenes, and the expression of aberrant RNA products from inefficiently terminated transcripts. Position effects can be reduced by buffering the transgene with matrix attachment regions or controlling the site of integration through homologous recombination or site-specific recombination. Alternatively, it may be possible to introduce the transgene into the plastid genome, which does not suffer from position effects. More recent developments such as minichromosomes may provide a method to introduce entire linkage groups and maintain them stably and episomally. Site-specific recombination can also be used to simplify locus structure, by reducing the number of repeats, which may help to reduce the likelihood of RNA silencing. Even so, many reports show that high-level transgene expression is possible in plants with multiple transgene copies, suggesting that the overall level of expression may be relevant, i.e., there may be a trigger level at which silencing is induced. This may involve the detection of high levels of transgene mRNA or may simply reflect the greater likelihood of aberrant RNA products being generated as collateral damage. The recent publication of the full draft sequence of the transgenic SUNUP papaya genome shows that the detailed characterization of the transgene sequence and its flanking regions is not an insurmountable obstacle. Perhaps such intensive analysis will, in the future, allow the accurate prediction of transgene behavior and stability in transgenic plants.

References

Abranches R, Santos AP, Wegel E, Williams S, Castilho A, Christou P, Shaw P, Stoger E (2000) Widely separated multiple transgene integration sites in wheat chromosomes are brought together at interphase. Plant J 24:713–723

Afolabi AS, Worland B, Snape JW, Vain P (2004) A large-scale study of rice plants transformed with different T-DNAs provides new insights into locus composition and T-DNA linkage configurations. Theor Appl Genet 109:815–826

Allen GC, Hall G Jr, Michalowski S, Newman W, Spiker S, Weissinger AK, Thompson WF (1996) High-level transgene expression in plant cells: effects of a strong scaffold attachment region from tobacco. Plant Cell 8:899–913

Allen GC, Hall GE Jr, Childs LC, Weissinger AK, Spiker S, Thompson WF (1993) Scaffold attachment regions increase reporter gene expression in stably transformed plant cells. Plant Cell 5:603–613

Altpeter F, Baisakh N, Beachy R, Bock R, Capell T, Christou P, Daniell H, Datta K, Datta S, Dix PJ, Fauquet C, Huang N, Kohli A, Mooibroek H, Nicholson L, Nguyen TT, Nugent G, Raemakers K, Romano A, Somers DA, Stoger E, Taylor N, Visser R (2005) Particle bombardment and the genetic enhancement of crops: myths and realities. Mol Breed 15:305–327

Barakat A, Gallois P, Raynal M, Mestre Ortega D, Sallaud C, Guiderdoni E, Delseny M, Bernardi G (2000) The distribution of T-DNA in the genomes of transgenic Arabidopsis and rice. FEBS Lett 471:161–164

Becker D, Brettschneider R, Lörz H (1994) Fertile transgenic wheat from microprojectile bombardment of scutellar tissue. Plant J 5:299–307

Blechl AE, Anderson OD (1996) Expression of a novel high-molecular-weight glutenin subunit gene in transgenic wheat. Nat Biotechnol 14:875–879

Bock R (2007) Plastid biotechnology: prospects for herbicide and insect resistance, metabolic engineering and molecular farming'. Curr Opin Biotechnol 18:100–106

Breyne P, van Montagu M, Depicker N, Gheysen G (1992) Characterization of a plant scaffold attachment region in a DNA fragment that normalizes transgene expression in tobacco. Plant Cell 4:463–471

Buchholz F, Angrand PO, Stewart AF (1998) Improved properties of FLP recombinase evolved by cycling mutagenesis. Nat Biotechnol 16:657–662

Cai CQ, Doyon Y, Ainley WM, Miller JC, Dekelver RC, Moehle EA, Rock JM, Lee YL, Garrison R, Schulenberg L, Blue R, Worden A, Baker L, Faraji F, Zhang L, Holmes MC, Rebar EJ, Collingwood TN, Rubin-Wilson B, Gregory PD, Urnov FD, Petolino JF (2009) Targeted transgene integration in plant cells using designed zinc finger nucleases. Plant Mol Biol 69(6):699–709

Cannell ME, Doherty A, Lazzeri PA, Barcelo P (1999) A population of wheat and tritordeum transformants showing a high degree of marker gene stability and heritability. Theor Appl Genet 99:772–784

Carlson SR, Rudgers GW, Zieler H, Mach JM, Luo S, Grunden E, Krol C, Copenhaver GP, Preuss D (2007) Meiotic transmission of an in-vitro-assembled autonomous maize minichromosomes. PLOS Genet 3:1965–1974

Chalfun A, Mes JJ, Mlynarova L, Aarts MGM, Angenent GC (2003) Low frequency of T-DNA based activation tagging in Arabidopsis is correlated with methylation of CaMV 35S enhancer sequences. FEBS Lett 555:459–463

Chawla R, Ariza-Nieto M, Wilson AJ, Moore SK, Srivastava V (2006) Transgene expression produced by biolistic-mediated, site-specific gene integration is consistently inherited by the subsequent generations. Plant Biotechnol J 4:209–218

Chen L, Marmey P, Taylor NJ, Brizard JP, Espinoza C, D'Cruz P, Huet H, Zhang S, de Kochko A, Beachy RN, Fauquet CM (1998) Expression and inheritance of multiple transgenes in rice plants. Nat Biotechnol 16:1060–1064

Chen SB, Liu X, Peng HY, Gong WK, Wang R, Wang F, Zhu Z (2004) Cre/lox-mediated marker gene excision in elite indica rice plants transformed with genes conferring resistance to lepidopteran insects. Acta Bot Sin 46:1416–1423

Cheng M, Fry JE, Pang S, Zhou H, Hironaka CM, Duncan DR, Conner TW, Wan Y (1997) Genetic transformation of wheat mediated by *Agrobacterium tumefaciens*. Plant Physiol 115:971–980

Chinn AM, Comai L (1996) The heat shock cognate 80 gene of tomato is flanked by matrix attachment regions. Plant Mol Biol 32:959–968

Cook PR (1999) The organization of replication and transcription. Science 284:1790–1795

Cooley J, Ford T, Christou P (1995) Molecular and genetic characterization of elite transgenic plants produced by electric-discharge particle acceleration. Theor Appl Genet 90:97–104

Coppoolse ER, de Vroomen MJ, van Gennip F, Hersmus BJM, van Haaren MJJ (2005) Size does matter: Cre-mediated somatic deletion efficiency depends on the distance between the target lox-sites. Plant Mol Biol 58:687–698

Cotsaftis O, Guiderdoni E (2005) Enhancing gene targeting efficiency in higher plants: rice is on the move. Transgenic Res 14:1–14

Dai S, Zheng P, Marmey P, Zhang S, Tian W, Chen S, Beachy RN, Fauquet C (2001) Comparative analysis of transgenic rice plants obtained by *Agrobacterium*-mediated transformation and particle bombardment. Mol Breed 7:25–33

Daniell H (2006) Production of biopharmaceuticals and vaccines in plants via the chloroplast genome. Biotechnol J 1:1071–1079

Daniell H, Kumar S, Dufourmantel N (2005) Breakthrough in chloroplast genetic engineering of agronomically important crops. Trends Biotechnol 23:238–245

De Block M, Debrouwer D (1991) Two T-DNAs co-transformed into *Brassica napus* by a double *Agrobacterium tumefaciens* infection are mainly integrated at the same locus. Theor Appl Genet 82:257–263

De Buck S, Jacobs A, Van Montagu M, Depicker A (1999) The DNA sequences of T-DNA junctions suggest that complex T-DNA loci are formed by a recombination process resembling T-DNA integration. Plant J 20:295–304

D'Halluin K, Vanderstraeten C, Stals E, Cornelissen M, Ruiter R (2008) Homologous recombination: a basis for targeted genome optimization in crop species such as maize. Plant Biotechnol J 6:93–102

Djukanovic V, Orczyk W, Gao HR, Sun XF, Garrett N, Zhen SF, Gordon-Kamm W, Barton J, Lyznik LA (2006) Gene conversion in transgenic maize plants expressing FLP/FRT and Cre/loxP site-specific recombination systems. Plant Biotechnol J 4:345–357

Dufourmantel N, Tissot G, Garcon F, Pelissier B, Dubald M (2006) Stability of soybean recombinant plastome over six generations. Transgenic Res 15:305–311

Finnegan J, McElroy D (1994) Transgene inactivation: plants fight back! Biotechnology 12:883–888

Fu X, Duc LT, Fontana S, Bong BB, Tinjuangjun P, Sudhakar D, Twyman RM, Christou P, Kohli A (2000) Linear transgene constructs lacking vector backbone sequences generate low-copy-number transgenic plants with simple integration patterns. Transgenic Res 9:11–19

Gahakwa D, Maqbool SB, Fu X, Sudhakar D, Christou P, Kohli A (2000) Transgenic rice as a system to study the stability of transgene expression: multiple heterologous transgenes show similar behaviour in diverse genetic backgrounds. Theor Appl Genet 101:388–399

Gelvin SB (1998) Multigene plant transformation: more is better! Nat Biotechnol 16:1009–1010

Gelvin SB (2003) Agrobacterium-mediated plant transformation: the biology behind the gene jockeying tool. Microbiol Mol Biol Rev 67:16–37

Gelvin SB, Karcher SJ, DiRita VJ (1983) Methylation of the T-DNA in *Agrobacterium tumefaciens* and in several crown gall tumors. Nucleic Acids Res 11:159–174

Gidoni D, Fuss E, Burbidge A, Speckmann GJ, James S, Nijkamp D, Mett A, Feiler J, Smoker M, de Vroomen MJ, Leader D, Liharska T, Groenendijk J, Coppoolse E, Smit JJ, Levin I, de Both M, Schuch W, Jones JD, Taylor IB, Theres K, van Haaren MJ (2003) Multi-functional T-DNA/Ds tomato lines designed for gene cloning and molecular and physical dissection of the tomato genome. Plant Mol Biol 51:83–98

Gilbertson L (2003) Cre-lox recombination: Cre-ative tools for plant biotechnology. Trends Biotechnol 21:550–555

Goldsbrough A, Bevan M (1991) New patterns of gene activity in plants detected using an Agrobacterium vector. Plant Mol Biol 16:263–269

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

Gordon KH, Waterhouse PM (2007) RNAi for insect-proof plants. Nat Biotechnol 25:1231–1232

Hadi MZ, McMullen MD, Finer JJ (1996) Transformation of 12 different plasmids into soybean via particle bombardment. Plant Cell Rep 15:500–505

Hammond SM, Caudy AA, Hannon GJ (2001) Post-transcriptional gene silencing by double-stranded RNA. Nat Rev Genet 2:110–119

Han Y, Grierson D (2002) Relationship between small antisense RNAs and aberrant RNAs associated with sense transgene mediated gene silencing in tomato. Plant J 29:509–519

Hanin M, Volrath S, Bogucki A, Briker M, Ward E, Paszkowski J (2001) Gene targeting in Arabidopsis. Plant J 28:671–677

Heinrichs A (2008) Small RNAs: united in silence. Nat Rev Mol Cell Biol 9:496

Hellens R, Mullineaux P, Klee H (2000) Technical focus: a guide to Agrobacterium binary Ti vectors. Trends Plant Sci 5:446–451

Hepburn AG, Clarke LE, Pearson L, White J (1983) The role of cytosine methylation in the control of nopaline synthase gene expression in a plant tumor. J Mol Appl Genet 2:315–329

Holmes-Davis R, Comai L (1998) Nuclear matrix attachment regions and plant gene expression. Trends Plant Sci 3:91–97

Hsing YI, Chern CG, Fan MJ, Lu PC, Chen KT, Lo SF, Sun PK, Ho SL, Lee KW, Wang YC, Huang WL, Ko SS, Chen S, Chen JL, Chung CI, Lin YC, Hour AL, Wang YW, Chang YC, Tsai MW, Lin YS, Chen YC, Yen HM, Li CP, Wey CK, Tseng CS, Lai MH, Huang SC,

Chen LJ, Yu SM (2007) A rice gene activation/knockout mutant resource for high throughput functional genomics. Plant Mol Biol 63:351–364
Hu Q, Nelson K, Luo H (2006) FLP-mediated site-specific recombination for genome modification in turfgrass. Biotechnol Lett 28:1793–1804
Hu T, Metz S, Chay C, Zhou HP, Biest N, Chen G, Cheng M, Feng X, Radionenko M, Lu F, Fry J (2003) Agrobacterium-mediated large-scale transformation of wheat (*Triticum aestivum* L.) using glyphosate selection. Plant Cell Rep 21:1010–1019
Ida S, Terada R (2004) A tale of two integrations, transgene and T-DNA: gene targeting by homologous recombination in rice. Curr Opin Biotechnol 15:132–138
Iglesias VA, Moscone EA, Papp I, Neuhuber F, Michalowski S, Phelan T, Spiker S, Matzke M, Matzke AJM (1997) Molecular and cytogenetic analyses of stably and unstably expressed transgene loci in tobacco. Plant Cell 9:1251–1264
Jackson SA, Zhang P, Chen WP, Phillips RL, Friebe B, Muthukrishnan S, Gill BS (2001) High resolution structural analysis of biolistic transgene integration into the genome of wheat. Theor Appl Genet 103:56–62
James C (2007) Global status of commercialized biotech/GM crops: 2007. ISAAA Brief No 37. ISAAA, Ithaca, NY
James VA, Worland B, Snape JW, Vain P (2004a) Development of a standard operating procedure (SOP) for the precise quantification of transgene expression levels in rice plants. Physiol Plant 120:650–656
James VA, Worland B, Snape JW, Vain P (2004b) Strategies for precise quantification of transgene expression levels over several generations in rice. J Exp Bot 55:1307–1313
Jen GC, Chilton MD (1986) The right border region of pTiT37 T-DNA is intrinsically more active than the left border region in promoting T-DNA transformation. Proc Natl Acad Sci USA 83:3895–3899
Jeon JS, Lee S, Jung KH, Jun SH, Jeong DH, Lee J, Kim C, Jang S, Yang K, Nam J, An K, Han MJ, Sung RJ, Choi HS, Yu JH, Choi JH, Cho SY, Cha SS, Kim SI, An G (2000) T-DNA insertional mutagenesis for functional genomics in rice. Plant J 22:561–570
Jones JDG, Gilbert DE, Grady KL, Jorgensen RA (1987) T-DNA structure and gene expression in petunia plants transformed by *Agrobacterium tumefaciens* C58 derivates. Mol Gen Genet 207:478–485
Jorgensen RA, Cluster PD, English J, Que Q, Napoli CA (1996) Chalcone synthase cosuppression phenotypes in petunia flowers: Comparison of sense vs. antisense constructs and single-copy vs. complex T-DNA sequences. Plant Mol Biol 31:957–973
Jorgensen RA, Snyder C, Jones JDG (1987) T-DNA is organized predominantly in inverted repeat structures in plants transformed with *Agrobacterium tumefaciens* C58 derivatives. Mol Gen Genet 207:471–477
Kempin SA, Liljegren SJ, Block LM, Rounsley SD, Yanofsky MF, Lam E (1997) Targeted disruption in Arabidopsis. Nature 389:802–803
Kerbach S, Lorz H, Becker D (2005) Site-specific recombination in Zea mays. Theor Appl Genet 111:1608–1616
Khanna HK, Raina SK (2002) Elite indica transgenic rice plants expressing modified Cry1Ac endotoxin of *Bacillus thuringiensis* show enhanced resistance to yellow stem borer (*Scirpophaga incertulas*). Transgenic Res 11:411–423
Klein TM, Wolf ED, Wu R, Sanford JC (1987) High-velocity microprojectiles for delivering nucleic acids into living cells. Nature 327:70–73
Kohli A, Griffiths S, Palacios N, Twyman RM, Vain P, Laurie DA, Christou P (1999) Molecular characterization of transforming plasmid rearrangements in transgenic rice reveals a recombination hotspot in the CaMV 35S promoter and confirms the predominance of microhomology-mediated recombination. Plant J 17:591–601
Kohli A, Leech M, Vain P, Laurie DA, 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

Kohli A, Twyman RM, Abranches A, Wegel E, Shaw P, Christou P, Stoger E (2003) Transgene integration, organization and interaction in plants. Plant Mol Biol 52:247–258

Kononov ME, Bassuner B, Gelvin SB (1997) Integration of T-DNA binary vector "backbone" sequences into the tobacco genome: Evidence for multiple complex patterns of integration. Plant J 11:945–957

Kopertekh L, Juttner G, Schiemann J (2004a) PVX-Cre-mediated marker gene elimination from transgenic plants. Plant Mol Biol 55:491–500

Kopertekh L, Juttner G, Schiemann J (2004b) Site-specific recombination induced in transgenic plants by PVX virus vector expressing bacteriophage P1 recombinase. Plant Sci 166:485–492

Kumar S, Fladung M (2000) Transgene repeats in aspen: molecular characterisation suggests simultaneous integration of independent T-DNAs into receptive hotspots in the host genome. Mol Gen Genet 264:20–28

Kumar S, Fladung M (2002) Transgene integration in aspen: structures of integration sites and mechanism of T-DNA integration. Plant J 31:543–551

Kumar S, Allen GC, Thompson WF (2005) Gene targeting in plants: fingers on the move. Trends Plant Sci 11:159–161

Kumpatla SP, Chandrasekharan MB, Iyer LM, Li G, Hall TC (1998) Genome intruder scanning and modulation systems and transgene silencing. Trends Plant Sci 3:97–104

Laufs P, Autran D, Traas J (1999) A chromosomal paracentric inversion associated with T-DNA integration in Arabidopsis. Plant J 18:131–139

Lechtenberg B, Schubert D, Forsbach A, Gils M, Schmidt R (2003) Neither inverted repeat T-DNA configurations nor arrangements of tandemly repeated transgenes are sufficient to trigger transgene silencing. Plant J 34:507–517

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

Lewandoski M, Martin GR (1997) Cre-mediated chromosome loss in mice. Nat Genet 17:223–225

Li Z, Hansen JL, Liu Y, Zemetra RS, Berger PH (2004) Using real-time PCR to determine transgene copy number in wheat. Plant Mol Biol Rep 22:179–188

Lindbo JA, Silva-Rosales L, Proebsting WM, Dougherty WG (1993) Induction of a highly specific antiviral state in transgenic plants: implications for regulation of gene expression and virus resistance. Plant Cell 5:1749–1759

Lindsey K, Wei W, Clarke MC, McArdle HF, Rooke LM, Topping JF (1993) Tagging genomic sequences that direct transgene expression by activation of a promoter trap in plants. Transgenic Res 2:33–47

Lloyd A, Plaisier CL, Carroll D, Drews GN (2005) Targeted mutagenesis using zinc-finger nucleases in *Arabidopsis*. Proc Natl Acad Sci USA 102:2232–2237

Luo K, Duan H, Zhao D, Zheng X, Deng W, Chen Y, Stewart CN Jr, McAvoy R, Jiang X, Wu Y, He A, Pei Y, Li Y (2007) 'GM-gene-deletor': fused loxP-FRT recognition sequences dramatically improve the efficiency of FLP or CRE recombinase on transgene excision from pollen and seed of tobacco plants. Plant Biotechnol J 5:263–274

Luo Z, Chen Z (2007) Improperly terminated, unpolyadenylated mRNA of sense transgenes is targeted by RDR6-mediated RNA silencing in Arabidopsis. Plant Cell 19:943–958

Lyznik LA, Gordon-Kamm WJ, Tao Y (2003) Site-specific recombination for genetic engineering in plants. Plant Cell Rep 21:925–932

Ma C, Mitra A (2002) Intrinsic direct repeats generate consistent post-transcriptional gene silencing in tobacco. Plant J 31:37–49

Marjanac G, De Paepe A, Peck I, Jacobs A, De Buck S, Depicker A (2008) Evaluation of Cre-mediated excision approaches in *Arabidopsis thaliana*. Transgenic Res 17:239–250

Martineau B, Voelker TA, Sanders RA (1994) On defining T-DNA. Plant Cell 6:1032–1033

Matsumoto S, Ito Y, Hosoi T, Takahashi Y, Machida Y (1990) Integration of Agrobacterium T-DNA into a tobacco chromosome: possible involvement of DNA homology between T-DNA and plant DNA. Mol Gen Genet 224:309–316

Matzke MA, Matzke AJM (1990) Gene interactions and epigenetic variation in transgenic. plants. Dev Genet 11:214–223
Matzke MA, Matzke AJM (1991) Differential inactivation and methylation of a transgene in plants by two suppressor loci containing homologous sequences. Plant Mol Biol 16:821–830
McGinnis K, Chandler V, Cone K, Kaeppler H, Kaeppler S, Kerschen A, Pikaard C, Richards E, Sidorenko L, Smith T, Springer N, Wulan T (2005) Transgene-induced RNA interference as a tool for plant functional genomics. RNA Interf 392:1–24
McKnight TD, Lillis MT, Simpson RB (1987) Segregation of genes transferred to one plant cell from two separate Agrobacterium strains. Plant Mol Biol 8:439–445
Mehlo L, Mazithulela G, Twyman RM, Boulton MI, Davies JW, Christou P (2000) Structural analysis of transgene rearrangements and effects on expression in transgenic maize plants generated by particle bombardment. Maydica 45:277–287
Metzger D, Feil R (1999) Engineering the mouse genome by site-specific recombination. Curr Opin Biotechnol 10:470–476
Metzlaff M, O'Dell M, Hellens R, Flavell RB (2000) Developmentally and transgene regulated nuclear processing of primary transcripts of chalcone synthase A in petunia. Plant J 23:63–72
Meyer P, Linn F, Heidmann I, Meyer H, Niedenhof I, Saedler H (1992) Endogenous and environmental factors influence 35S promoter methylation of a maize A1 gene construct in transgenic petunia and its colour phenotype. Mol Gen Genet 231:345–352
Miao ZH, Lam E (1995) Targeted disruption of the *TGA3* locus in *Arabidopsis thaliana*. Plant J 7:359–365
Ming R, Hou S, Feng Y, Yu Q, Dionne-Laporte A, Saw JH, Senin P, Wang W, Ly BV, Lewis KL, Salzberg SL, Feng L, Jones MR, Skelton RL, Murray JE, Chen C, Qian W, Shen J, Du P, Eustice M, Tong E, Tang H, Lyons E, Paull RE, Michael TP, Wall K, Rice DW, Albert H, Wang ML, Zhu YJ, Schatz M, Nagarajan N, Acob RA, Guan P, Blas A, Wai CM, Ackerman CM, Ren Y, Liu C, Wang J, Wang J, Na JK, Shakirov EV, Haas B, Thimmapuram J, Nelson D, Wang X, Bowers JE, Gschwend AR, Delcher AL, Singh R, Suzuki JY, Tripathi S, Neupane K, Wei H, Irikura B, Paidi M, Jiang N, Zhang W, Presting G, Windsor A, Navajas-Pérez R, Torres MJ, Feltus FA, Porter B, Li Y, Burroughs AM, Luo MC, Liu L, Christopher DA, Mount SM, Moore PH, Sugimura T, Jiang J, Schuler MA, Friedman V, Mitchell-Olds T, Shippen DE, dePamphilis CW, Palmer JD, Freeling M, Paterson AH, Gonsalves D, Wang L, Alam M (2008) The draft genome of the transgenic tropical fruit tree papaya (Carica papaya Linnaeus). Nature 452:991–997
Mlynarova L, Jansen RC, Conner AJ, Stiekema WJ, Nap JP (1995) The MAR-mediated reduction in position effect can be uncoupled from copy number-dependent expression in transgenic plants. Plant Cell 7:599–609
Mlynarova L, Keizer L, Stiekema WJ, Nap JP (1996) Approaching the lower limits of transgene variability. Plant Cell 8:1589–1599
Mlynarova L, Loonen A, Heldens J, Jansen RC, Keizer P, Stiekema WJ, Nap JP (1994) Reduced position effect in mature transgenic plants conferred by the chicken lysozyme matrix-associated region. Plant Cell 6:417–426
Mlynárová L, Loonen A, Mietkiewska E, Jansen RC, Nap JP (2002) Assembly of two transgenes in an artificial chromatin domain gives highly coordinated expression in tobacco. Genetics 160:727–740
Moore SK, Srivastava V (2006) Efficient deletion of transgenic DNA from complex integration locus of rice mediated by Cre/lox recombination system. Crop Sci 46:700–705
Muskens MW, Vissers AP, Mol JN, Kooter JM (2000) Role of inverted DNA repeats in transcriptional and post-transcriptional gene silencing. Plant Mol Biol 43:243–260
Nacry P, Camilleri C, Courtial B, Caboche M, Bouchez D (1998) Major chromosomal rearrangements induced by T-DNA transformation in Arabidopsis. Genetics 149:641–650
Nam J, Mysore KS, Gelvin SB (1998) *Agrobacterium tumefaciens* transformation of the radiation hypersensitive *Arabidopsis thaliana* mutants *uvh1* and *rad5*. Mol Plant Microbe Interact 11:1136–1141

Nam J, Mysore KS, Zheng C, Knue MK, Matthysse AG, Gelvin SB (1999) Identification of T-DNA tagged Arabidopsis mutants that are resistant to transformation by Agrobacterium. Mol Gen Genet 261:429–438

Napoli C, Lemieux C, Jorgensen R (1990) Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in trans. Plant Cell 2:279–289

Narasimhulu SB, Deng XB, Sarria R, Gelvin SB (1996) Early transcription of Agrobacterium T-DNA genes in tobacco and maize. Plant Cell 8:873–886

Offringa R, de Groot MJ, Haagsman HJ, Does MP, van den Elzen PJ, Hooykaas PJ (1990) Extrachromosomal homologous recombination and gene targeting in plant cells after Agrobacterium mediated transformation. EMBO J 9:3077–3084

Pansegrau W, Schoumacher F, Hohn B, Lanka E (1993) Site-specific cleavage and joining of single-stranded DNA by VirD2 protein of *Agrobacterium tumefaciens* Ti plasmids: analogy to bacterial conjugation. Proc Natl Acad Sci USA 90:11538–11542

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

Pawlowski WP, Somers DA (1998) Transgenic DNA integrated into the oat genome is frequently interspersed by host DNA. Proc Natl Acad Sci USA 95:12106–12110

Pikaart M, Feng J, Villeponteau B (1998) The polyomavirus enhancer activates chromatin accessibility on integration into the HPRT gene. Mol Cell Biol 12:5785–5792

Porsch P, Jahnke A, During K (1998) A plant transformation vector with a minimal T-DNA. II. Irregular integration patterns of the T-DNA in the plant genome. Plant Mol Biol 37:581–585

Puchta H (2003) Marker-free transgenic plants. Plant Cell Tissue Organ Cult 74:123–134

Puchta H, Kocher S, Hohn B (1992) Extrachromosomal homologous DNA recombination in plant cells is fast and is not affected by CpG methylation. Mol Cell Biol 12:3372–3379

Que Q, Wang HY, English JJ, Jorgensen RA (1997) The frequency and degree of cosuppression by sense chalcone synthase transgenes are dependent on transgene promoter strength and are reduced by premature nonsense codons in the transgene coding sequence. Plant Cell 9:1357–1368

Rai M, Datta K, Parkhi V, Tan J, Oliva N, Chawla HS, Datta SK (2007) Variable T-DNA linkage configuration affects inheritance of carotenogenic transgenes and carotenoid accumulation in transgenic indica rice. Plant Cell Rep 26:1221–1231

Ramanathan V, Veluthambi K (1995) Transfer of non-T-DNA portions of the *Agrobacterium tumefaciens* Ti plasmid pTiA6 from the left terminus of TL-DNA. Plant Mol Biol 28:1149–1154

Ramessar K, Capell T, Twyman RM, Quemada H, Christou P (2008) Trace and traceability – a call for regulatory harmony. Nat Biotechnol 26:975–978

Register JC 3rd, Peterson DJ, Bell PJ, Bullock WP, Evans IJ, Frame B, Greenland AJ, Higgs NS, Jepson I, Jiao S, Lewnau CJ, Sillick JM, Wilson HM (1994) Structure and function of selectable and non-selectable transgenes in maize after introduction by particle bombardment. Plant Mol Biol 25:951–961

Reiss B, Schubert I, Kopchen K, Wendeler E, Schell J, Puchta H (2000) RecA stimulates sister chromatid exchange and the fidelity of double-strand break repair, but not gene targeting, in plants transformed by Agrobacterium. Proc Nat Acad Sci USA 97:3358–3363

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

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

Rooke L, Steele SH, Barcelo P, Shewry PR, Lazzeri PA (2003) Transgene inheritance, segregation and expression in bread wheat. Euphytica 129:301–309

Sadowski PD (1993) Site-specific genetic recombination: hops, flips, and flops. FASEB J 7:760–767

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

Sanford JC, Smith FD, Russell JA (1993) Optimizing the biolistic process for different biological applications. Methods Enzymol 217:483–509

Sawasaki T, Takahashi M, Goshima N, Morikawa H (1998) Structures of transgene loci in transgenic Arabidopsis plants obtained by particle bombardment: junction regions can bind to nuclear matrices. Gene 218:27–35

Schiestl RH, Petes TD (1991) Integration of DNA fragments by illegitimate recombination in *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 88:7585–7589

Schubert D, Lechtenberg B, Forsbach A, Gils M, Bahadur S, Schmidt R (2004) Silencing in Arabidopsis T-DNA transformants: the predominant role of a gene-specific RNA sensing mechanism versus position effects. Plant Cell 16:2561–2572

Shaw CH, Watson MD, Carter GH, Shaw CH (1984) The right hand copy of the nopaline Ti-plasmid 25 bp repeat is required for tumour formation. Nucleic Acids Res 12:6031–6041

Somers DA, Makarevistch I (2004) Transgene integration in plants: poking or patching holes in promiscuous genomes? Curr Opin Biotechnol 15:126–131

Sonti RV, Chiurazzi M, Wong D, Davies CS, Harlow GR, Mount DW, Signer ER (1995) Arabidopsis mutants deficient in T-DNA integration. Proc Natl Acad Sci USA 92:11786–11790

Spencer TM, O'Brien JV, Start WG, Adams TR, Gordon-Kamm WJ, Lemaux PG (1992) Segregation of transgenes in maize. Plant Mol Biol 18:201–210

Spielmann A, Simpson RB (1986) T-DNA structure in transgenic tobacco plants with multiple independent integration sites. Mol Gen Genet 205:34–41

Srivastava V, Anderson OD, Ow DW (1999) Single-copy transgenic wheat generated through the resolution of complex integration patterns. Proc Natl Acad Sci USA 96:11117–11121

Srivastava V, Ariza-Nieto M, Wilson AJ (2004) Cre-mediated site-specific gene integration for consistent transgene expression in rice. Plant Biotechnol J 2:169–179

Srivastava V, Ow DW (2001) Single-copy primary transformants of maize obtained through the co-introduction of a recombinase-expressing construct. Plant Mol Biol 46:561–566

Srivastava V, Ow DW (2003) Rare instances of Cre-mediated deletion product maintained in transgenic wheat. Plant Mol Biol 52:661–668

Srivastava V, Vasil V, Vasil IK (1996) Molecular characterization of the fate of transgenes in transformed wheat. Theor Appl Genet 92:1031–1037

Steck TR (1997) Ti plasmid type affects T-DNA processing in Agrobacterium tumefaciens. FEMS Microbiol Lett 147:121–125

Stoger E, Williams S, Keen D, Christou P (1998) Molecular characteristics of transgenic wheat and the effect on transgene expression. Transgenic Res 7:463–471

Svab Z, Maliga P (1993) High-frequency plastid transformation in tobacco by selection for a chimeric *aadA* gene. Proc Natl Acad Sci USA 90:913–917

Svitashev SK, Ananiev E, Pawlonski WP, Somers DA (2000) Association of transgene integration sites with chromosome rearrangements in hexaploid oat. Theor Appl Genet 100:872–880

Svitashev SK, Pawlowski WF, Makarevitch I, Plank DW, Somers DA (2002) Complex transgene locus structures implicate multiple mechanisms for plant transgene rearrangement. Plant J 32:433–445

Svitashev SK, Somers DA (2001) Genomic interspersions determine the size and complexity of transgene loci in transgenic plants produced by microprojectile bombardment. Genome 44:691–697

Tanako M, Egawa H, Ikeda J, Wakasa K (1997) The structures of integration sites in transgenic rice. Plant J 11:353–361

Tavernarakis N, Wang SL, Dorovkov M, Ryazanov A, Driscoll M (2000) Heritable and inducible genetic interference by double-stranded RNA encoded by transgenes. Nat Genet 24:180–183

Ten Hoopen R, Robbins TP, Fransz PF, Montijn BM, Oud O, Gerats A, Nanninga N (1996) Localization of T-DNA insertions in petunia by fluorescence in situ hybridization: physical evidence for suppression of recombination. Plant Cell 8:823–830

Terada R, Urawa H, Inagaki V, Tsugane K, Lida S (2002) Efficient gene targeting by homologous recombination in rice. Nat Biotechnol 20:1030–1034

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

Travella S, Ross SM, Harden J, Everett C, Snape JW, Harwood WA (2005) A comparison of transgenic barley lines produced by particle bombardment and Agrobacterium-mediated techniques. Plant Cell Rep 23:780–789

Twyman RM, Christou P, Stoger E (2002) Genetic transformation of plants and their cells. In: Oksman-Caldentey KM, Barz W (eds) Plant biotechnology and transgenic plants. Marcel, New York, pp 111–141

Twyman RM, Kohli A (2003) Genetic modification: Insertional and transposon mutagenesis. In: Thomas B, Murphy DJ, Murray B (eds) Encyclopedia of applied plant sciences. Elsevier Science, London, UK, pp 369–377

Tzfira T, Rhee Y, Chen MH, Kunik T, Citovsky V (2000) Nucleic acid transport in plant-microbe interactions: the molecules that walk through the walls. Annu Rev Microbiol 54:187–219

van Eldik GJ, Litiere K, Jacobs JJ, Van Montagu M, Cornelissen M (1998) Silencing of beta-1, 3-glucanase genes in tobacco correlates with an increased abundance of RNA degradation intermediates. Nucleic Acids Res 26:5176–5181

Van Leeuwen W, Mlynárová L, Nap JP, van der Plas LH, van der Krol AR (2001) The effect of MAR elements on variation in spatial and temporal regulation of transgene expression. Plant Mol Biol 47:543–554

Vaucheret H, Beclin C, Elmayan T, Feuerbach F, Godon C, Morel JB, Mourrain P, Palauqui JC, Vernhettes S (1998) Transgene-induced gene silencing in plants. Plant J 16:651–659

Vega JM, Yu WC, Han FP, Kato A, Peters EM, Birchler JA (2007) Transfer of a Cre/lox recombination system to maize for gene targeting and chromosome engineering. Chromosome Res 15:1–11

Vega JM, Yu WC, Han FP, Kato A, Peters EM, Zhang ZJ, Birchler JA (2008) Agrobacterium-mediated transformation of maize (*Zea mays*) with Cre-lox site specific recombination cassettes in BIBAC vectors. Plant Mol Biol 66:587–598

Wan S, Wu J, Zhang Z, Sun X, Lv Y, Gao C, Ning Y, Ma J, Guo Y, Zhang Q, Zheng X, Zhang C, Ma Z, Lu T (2009) Activation tagging, an efficient tool for functional analysis of the rice genome. Plant Mol Biol 69:69–80

Wang J, Lewis ME, Whallon JH, Sink KC (1995) Chromosome. mapping of T-DNA inserts in transgene. Petunia. by in situ hybridization. Transgenic Res 4:241–246

Wang MB, Waterhouse PM (2000) High-efficiency silencing of a beta-glucuronidase gene in rice is correlated with repetitive transgene structure but is independent of DNA methylation. Plant Mol Biol 43:67–68

Ward ER, Barnes WM (1988) VirD2 protein of *Agrobacterium tumefaciens* very tightly linked to 5′ end of T-strand DNA. Science 242:927–930

Weigel D, Ahn JH, Blázquez MA, Borevitz JO, Christensen SK, Fankhauser C, Ferrándiz C, Kardailsky I, Malancharuvil EJ, Neff MM, Nguyen JT, Sato S, Wang ZY, Xia Y, Dixon RA, Harrison MJ, Lamb CJ, Yanofsky MF, Chory J (2000) Activation tagging in Arabidopsis. Plant Physiol 122:1003–1013

Wenck A, Czakó M, Kanevski I, Márton L (1997) Frequent collinear long transfer of DNA inclusive of the whole binary vector during Agrobacterium-mediated transformation. Plant Mol Biol 34:913–922

Wilson C, Bellen HJ, Gehring WJ (1990) Position effects on eukaryotic gene expression. Ann Rev Cell Biol 6:679–714

Wolters AMA, Trindade LM, Jacobsen E, Visser RGF (1998) Fluorescence in situ hybridization on extended DNA fibres as a tool to analyse complex T-DNA loci in potato. Plant J 13:837–847

Wright DA, Townsend JA, Winfrey RJ Jr, Irwin PA, Rajagopal J, Lonosky PM, Hall BD, Jondle MD, Voytas DF (2005) High-frequency homologous recombination in plants mediated by zinc-finger nucleases. Plant J 44:693–705

Yang L, Ding J, Zhang C, Jia J, Weng H, Liu W, Zhang D (2005) Estimating the copy number of transgenes in transformed rice by real-time quantitative PCR. Plant Cell Rep 23:759–763

Zhu QH, Ramm K, Eamens AL, Dennis ES, Upadhyaya NM (2006) Transgene structures suggest that multiple mechanisms are involved in T-DNA integration in plants. Plant Sci 171:308–322

Zeevi V, Tovkach A, Tzfira T (2008) Increasing cloning possibilities using artificial zinc finger nucleases. Proc Natl Acad Sci USA 105:12785–12790

Ziemienowicz A, Tinland B, Bryant J, Gloeckler V, Hohn B (2000) Plant enzymes but not Agrobacterium VirD2 mediate T-DNA ligation in vitro. Mol Cell Biol 20:6317–6322