Transgene integration, organization, and expression in cereals

P. Christou, R.M. Twyman, Xiangdong Fu, E. Wegel, A. Kohli, and E. Stoger

As in other direct DNA transfer systems, transgenes delivered by particle bombardment integrate into the cereal genome by illegitimate recombination assisted by short regions of homology. Exogenous plasmid DNA tends to concatemerize prior to integration, resulting in tandemly arranged contiguous copies. However, recent research has shown that such concatemers may positively influence the integration of further transgenes nearby, perhaps by recruiting DNA repair complexes to the site of the original break. Transgenic loci therefore tend to comprise clusters of contiguous copies interspersed with short regions of genomic DNA. The analysis of transgenic wheat chromosomes by fluorescence *in situ* hybridization (FISH) indicates a further level of organization, where several transgene clusters integrate in the same region of the chromosome arm, but individual clusters produce separable FISH signals at metaphase, suggesting they are interspersed by large segments of genomic DNA. A segregating transgene locus therefore has a hierarchical structure, showing up to three levels of organization.

The analysis of transgene integration shows that recombination plays a strong role in determining integrated transgene structure. Particularly, certain sites in the transformation plasmid can provide hotspots for recombination, leading to particular types of transgene rearrangement. Since the failure of transgene expression often reflects such rearrangements, the modification of transformation vectors to eliminate troublesome sequences should improve the efficiency of transformation. The "clean DNA" system takes this strategy to its logical extreme by removing all unnecessary backbone elements and using just minimal linear cassettes (promoter, coding region, terminator) for transformation. Such experiments consistently generate transgenic plants with simple integration patterns, lower transgene copy numbers, fewer rearrangements, and stable transgene expression. The clean DNA system also allows direct transformation with multiple genes, but the integration patterns remain simple and silencing occurs very rarely.

Silencing often correlates with DNA methylation, and it is of particular interest to identify sequences in plant transgenes that induce methylation and result in unstable expression. There is evidence that methylation can be induced by interactions between homologous transgene copies (or between a transgene and homologous endogene) or may reflect genomic position ef-

fects. Recent data suggest, however, that more complex underlying processes may be involved. In a line of plants containing a single-copy three-gene transgenic locus, different forms of silencing could occur *de novo* and affect individual transgenes, even though the three were separated by only a few nucleotides. Remarkably, two completely different modes of silencing, associated with different methylation patterns, could be established in adjacent transgenes. Transposon mutagenesis is one of the most powerful tools available to gain an understanding of factors underlying transgene expression and stability. We have a population of more than 500 independent indica rice transformants containing the maize *Ac* transposon, representing a total of more than 1,500 plants. We have observed active excision and re-insertion of the *Ac* element in this population, which can now be scaled up to look for genes affecting genomic methylation and transgene expression.

The genetic manipulation of cereals is a key objective of both fundamental and applied research. This reflects the importance of cereals as a major food crop as well as the emergence of rice in particular as a model in functional genomics. It is therefore essential to gain a fundamental understanding of the mechanisms of transgene integration and how transgene structure and organization affect the stability of expression. Recent research carried out in our laboratory has shown that the mechanism of transgene integration at the molecular level is similar in cereals and all other plants (Kohli et al 1999). However, the use of direct DNA transfer methods such as particle bombardment to produce transgenic cereals appears to generate a unique higher-order transgene organization, which may have specific effects on transgene expression. The purpose of this review is to summarize our current knowledge of transgene integration and organization in cereals, particularly rice and wheat, and how this information has been used to improve the efficiency with which useful transgenic plants are generated.

DNA transfer to cereals

Since most monocots, including all cereals, were thought to be outside the natural host range of *Agrobacterium tumefaciens*, early attempts to transform cereals focused on direct DNA transfer methods. Originally, this involved the transformation of protoplasts, a technically demanding procedure requiring long periods of tissue culture (Krens et al 1982, Negrutiu et al 1987, Fromm et al 1987). The advent of particle bombardment, which is simple both conceptually and in practice, radically streamlined the production of transgenic cereals (Klein et al 1987, 1988, Christou et al 1988). Although advances in *Agrobacterium* methodology now permit the reliable transformation of cereals (Komari et al 1998), particle bombardment has emerged as the most robust method, since it is in principle unrestricted by either species or genotype barriers, relying on physical parameters rather than genetic ones. DNA can be introduced into a range of organized and easily regenerable tissues, and transgenic plants pro-

duced by organogenesis, somatic embryogenesis, or the germination of transformed seeds (Twyman et al, n.d.).

Another important advantage of particle bombardment is that it allows the stable and heritable introduction of many different genes at once using different plasmids, as these tend to cluster at a single locus. Conversely, transformation with more than two genes using conventional *Agrobacterium* systems currently requires the cointegration of all the genes in the same T-DNA. Chen et al (1998) reported the biolistic cotransformation of rice with 14 separate plasmids containing various genes and showed data confirming the integration of at least 13 of the plasmids in one plant. Cotransformation has also been used to introduce up to four agronomically important genes into rice, producing plants showing resistance to a spectrum of insect pests and plants with resistance pyramided against individual pests (Bano-Maqbool and Christou 1999, Tang et al 1999). Cointegration prevents different transgenes from segregating at meiosis. This is very important in breeding programs in which plants carry two or more transgenes required to generate a single protein, for example, in plants expressing recombinant human antibodies.

Transgene organization and expression

The stable expression of transgenes over multiple generations is essential if transgenic plants are to be of long-term value to producers and consumers. Transgene expression is influenced by several factors, many of which in some way reflect the structure of the transgenic locus. Such factors include the integrity of individual transgenes, the number of transgene copies, and their organization within the transgene locus. By investigating the relationship between transgene structure/organization and expression, it may be possible to determine the factors that are most important for stable transgene expression and to improve the design of transformation constructs to eliminate nonproductive transgene structures. This has been demonstrated, for example, by the success of the clean DNA transformation system discussed later in this chapter.

Three-level hierarchy of transgene organization in cereals

Transgene structure/organization can be investigated at different levels of resolution and this has provided evidence for a hierarchical organization. Polymerase chain reaction (PCR) amplification and sequencing provide the finest resolution and show that transgenic loci in rice often contain contiguous copies of the transforming plasmid (Kohli et al 1998, 1999). These may be arranged either head-to-tail or as inverted repeats and are thought to form prior to integration through the concatemerization (end-to-end joining) of exogenous DNA fragments.

Southern blot analysis is often used to characterize transgenic plants, but, as well as confirming transgene integration, it also demonstrates a second level of transgene organization (Kohli et al 1998, 1999, Fu et al 2000b; Fig. 1). Restriction enzymes that cut once in the transforming plasmid should generate different hybridizing bands for each transgene copy and thus provide an estimate of copy number. Restriction enzymes that cut either side of the transgene should release a cassette of diagnostic size

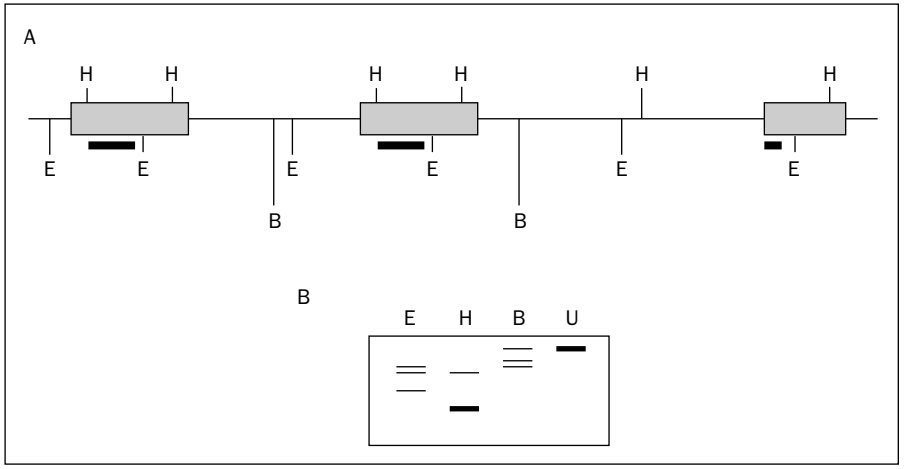

Fig. 1. How Southern blots can be used to characterize transgene organization. (A) A hypothetical transgenic locus comprising two intact transgenes and one truncated transgene, each separated by a short stretch of genomic DNA. Each intact transgene carries one site for the restriction enzyme *Eco*R1 (E), two sites for the enzyme *Hin*dIII (H), and no sites for the enzyme *Bam*H1 (B). (B) Typical Southern blot results. The *Eco*R1 digest reveals three bands. This indicates that there are three copies of the transgene since each hybridizing fragment is delineated by one *Eco*R1 site within the transgene and one randomly placed in the surrounding genomic DNA. The *Hin*dIII digest should generate a specific diagnostic hybridizing fragment of defined length, delineated by the two sites within the transgene. Additional fragments shorter or longer than expected indicate transgene rearrangements, in this case a truncation. The aberrant fragment is longer than expected because the 5′ *Hin*dIII site has been deleted in the truncation, so the fragment is formed by the next available *Hin*dIII site in the genomic DNA. The *Bam*H1 digest should generate a single high-molecular-weight fragment if there are no *Bam*H1 sites between the transgenes, often similar to the fragment seen in U (undigested DNA). If such sites do exist, individual fragments would be generated as shown. Experiments with several such "noncutter" enzymes reveal the existence of genomic DNA between transgene copies.

that can be used to confirm transgene integrity at the gross level (i.e., point mutations will not be identified). Restriction enzymes that do not cut in the transformation plasmid should release the transgenic locus as a single high-molecular-weight fragment regardless of the number of copies, as dictated by the positions of restriction sites in the surrounding genomic DNA. However, such digests often produce two or more hybridizing bands, providing evidence for the presence of variable-length stretches of genomic DNA between transgenes or contiguous transgene arrays. Such genomic stretches can be isolated by long PCR, indicating that they are tens rather than hundreds of kilobase pairs in length. This indicates that individual transgene copies and contiguous arrays are interspersed with genomic DNA and organized into transgene clusters.

In wheat, we began an investigation of the distribution of transgene loci to determine whether there was any bias to integration sites in terms of specific chromosomes or chromosome regions. This series of experiments was carried out using fluorescence *in situ* hybridization (FISH) and provided some surprising findings concern-

ing transgene organization. FISH analysis of wheat lines shown to have a single segregating transgenic locus revealed that each locus could comprise several separable signals, suggesting that individual transgenes and transgene clusters could be interspersed with sections of genomic DNA visible at the cytogenetic level, that is, in the order of megabase pairs (Abranches et al, n.d.). Analysis of rice and wheat transgenic loci has therefore demonstrated a three-tier organization of transgenes (Fig. 2), starting with individual copies and contiguous arrays, organized at a second level in clusters containing short regions of genomic DNA, which are in turn distributed in a local region of the chromosome and interspersed with longer regions of DNA, so that they generate discrete FISH signals at the cytogenetic level. Each of these levels of organization provides clues as to the mechanisms involved in transgene integration by direct DNA transfer, as discussed below.

Transgene expression

Transgene organization reflects three properties of a transgenic locus: the transgene copy number, the spatial arrangement of transgenes within the locus, and their structural integrity. In animal systems, evidence has accumulated that all three of these properties can profoundly affect transgene expression. For example, where two transgenes are arranged as an inverted repeat, the copies can form ectopic pairs or

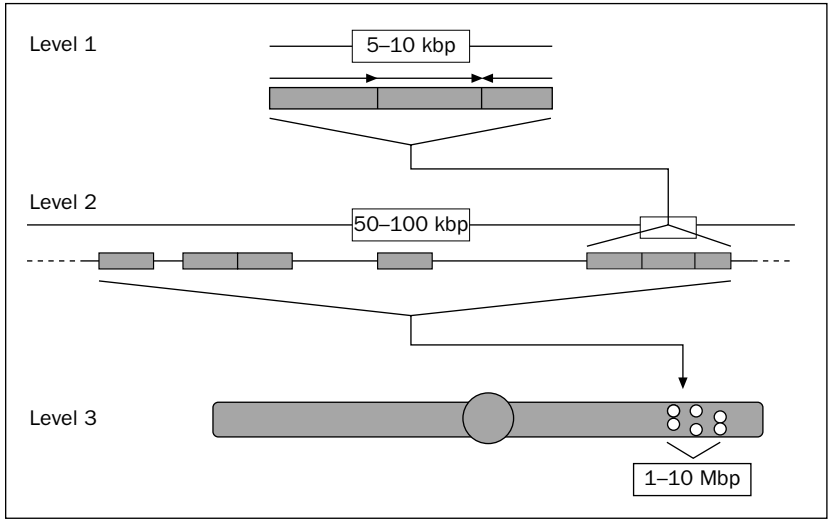

Fig. 2. Three-tier hierarchical transgene organization in cereals. Level 1 comprises individual transgenes and contiguous transgene copies (either tandem or inverted repeats, and containing intact and/or truncated copies), which concatemerize prior to integration. Level 2 comprises transgene clusters, i.e., groups of individual transgenes and contiguous transgene copies, interspersed by short regions of genomic DNA that can be isolated by polymerase chain reaction. Level 3 comprises groups of separable fluorescence *in situ* hybridization (FISH) signals, each corresponding to a transgene cluster, separated by megabases of genomic DNA. The groups of FISH signals behave as a single locus in genetic segregation analysis.

transcription through one transgene and into the next can generate antisense RNA (Twyman and Whitelaw 2000). Both these effects can lead to transgene silencing. In plants, the influence of the spatial arrangement of transgenes has not received much attention. A great deal of evidence suggests that multiple transgene copies provoke silencing, either by ectopic pairing in *cis* or in *trans* or by posttranscriptional mechanisms involving aberrant RNA species (reviewed by Gallie 1998, Grant 1999). All such studies, however, have been carried out in model dicot plants such as petunia, tobacco, and *Arabidopsis*. Work carried out in our laboratory has provided no evidence for a link between transgene copy number and expression level. We have generated hundreds of cereal plants ranging from those with single-copy transgenes to those carrying tens of copies, and there appears to be no correlation between copy number and silencing. Therefore, it appears that transgene integrity and spatial arrangement are most important when it comes to the effect of transgene organization on expression, and that even when intact single-copy transgenes are analyzed, there can still be pronounced and inexplicable *de novo* silencing effects at both the transcriptional and posttranscriptional levels even in plants with intact, single-copy transgenes (Fu et al 2000a).

Mechanisms of transgene integration

Level 1: integration at the molecular level

Studies in dicots have identified a possible link between double-strand break repair (DSBR) and transgene integration, both requiring illegitimate recombination. This form of recombination is a ubiquitous repair mechanism in eukaryotes, involving the covalent joining of sequences with either microhomology or no homology (Tsukamoto and Ikeda 1998). Illegitimate recombination junctions are often characterized by one or more of the following properties: (1) deletion of nucleotides at one or both of the recombining ends; (2) microhomology between the recombining ends, involving up to nine nucleotides; and (3) the presence of short stretches of additional DNA, known as filler DNA, at the junction. Equivalent studies in cereals have been carried out only recently. Takano et al (1997) used calcium phosphate-mediated transformation to study plasmid-genomic junctions in rice protoplasts, whereas, in our laboratory, we have carried out a series of studies involving plasmid-genomic and plasmid-plasmid junctions in transgenic rice and maize plants generated by particle bombardment. The sequences of novel recombination junctions suggest that the transforming plasmid had undergone rearrangements involving illegitimate recombination in a fashion similar to that of monocot and dicot species transformed by other methods. Therefore, at the molecular level, transgene integration following direct DNA or *Agrobacterium*-mediated T-DNA transfer is likely to involve the same components of the DNA-repair machinery.

Illegitimate recombination is the predominant mechanism of DSBR in higher eukaryotes, probably because the large genome size prevents effective homology searching and because the higher-order chromatin structure would be expected to hold broken DNA ends in close proximity. However, where unorthodox substrates are involved, illegitimate recombination can lead to large-scale genome rearrange-

ments and the integration of exogenous DNA (Salomon and Puchta 1998). In both *Agrobacterium*-mediated DNA delivery and the direct physical DNA transfer into plant cells, a wound response is elicited, which involves the activation of nucleases and DNA-repair enzymes to maintain the integrity of the host genome (Hunold et al 1994). Any exogenous DNA entering the cell is therefore simultaneously exposed to breakdown and repair enzymes. As a consequence, much of the exogenous DNA may be degraded but some may be used as a substrate for DNA repair, resulting in its rearrangement and/or incorporation into the genome (Matsumoto et al 1990, Gheysen et al 1991, Mayerhofer et al 1991, Takano et al 1997, Gorbunova and Levy 1997). Intact plasmids are relatively poor substrates for illegitimate recombination, so DNA ends may be provided by the activity of nucleases, which create single-strand nicks or double-strand breaks in the DNA.

The plasmid-plasmid recombination junctions we characterized in rice and maize plants generated by particle bombardment were similar to those found in tobacco following polyethylene glycol-mediated (Gorbunova and Levy 1997) or *Agrobacterium*-mediated transformation (De Groot et al 1994, Salomon and Puchta 1998). The analysis revealed typical features of illegitimate recombination, including (1) short patches of 4–8 homologous nucleotides at the junctions, (2) filler DNA, and (3) deletions at one of the recombining ends. Junctions were predominantly formed by microhomology-dependent illegitimate recombination. The favored mechanism for this type of reaction involves interaction between DNA ends with short single-stranded tails, the annealing of complementary tails, and repair synthesis over the remaining gaps. In addition, weak homology was often seen surrounding the junction site—this may have stabilized the initial interaction between the recombination substrates.

Illegitimate recombination junctions studied in animals suggest that homology with more than four nucleotides is rare, whereas the majority of the junctions we characterized in rice involved homology with more than four nucleotides. We found that junction sites preferentially comprised purine-rich tracts. In only two cases did junction formation appear not to involve the above mechanism, that is, there was no homology between the recombining partners and recombination may have occurred simply by blunt-end ligation of two DNA molecules. Gorbunova and Levy (1997) found only three out of 19 tobacco protoplast-derived clones in which the blunt or cohesive ends of their restriction-digested substrate joined without any degradation. Salomon and Puchta (1998) found two blunt-end ligation events among the 28 junctions they characterized in tobacco.

In one line, the junction contained an insertion of a 23-bp fragment of filler DNA. The presence of direct repeats (5′ TCCGG 3′) flanking this insert suggested one of two possible mechanisms. Illegitimate recombination may have occurred between the two ends and may have involved the synthesis of untemplated nucleotides. The direct repeats may represent short tails of imperfect complementarity responsible for the initiation of recombination. Alternatively, the insertion may represent a transposition event. The presence of staggered breaks in a target DNA molecule (in this case represented by two plasmids) may have acted as a substrate for the transposase or integrase encoded by an endogenous plant transposable element, resulting in the in-

sertion of an adventitious DNA fragment. Such events are thought to be responsible for the generation of processed pseudogenes, which make up a significant component of many higher eukaryotic genomes (Berg and Howe 1989). The direct repeats would in this case represent repair synthesis over the staggered break, generating the "target site duplication" characteristic of transposition events in all species, including the integration of T-DNA into the plant genome (Tinland 1996). Gorbunova and Levy (1997) found insertions ranging from 2 bp to 1.2 kbp in nearly 30% of the plasmid junctions they studied. This so-called "filler DNA" was sometimes genomic in origin, sometimes it appeared to derive from the transforming plasmid, and in other cases the origin was uncertain. The similarities between recombination junctions generated by particle bombardment in our study and those generated by alternative transformation methods in both monocots and dicots in other studies strongly suggest that the underlying mechanisms controlling plasmid rearrangement and transgene integration in plants are likely to be the same. The evidence we have provided suggests that several categories of illegitimate recombination are likely to be involved in this process, although microhomology-mediated recombination predominates (Kohli et al 1999; Fig. 3).

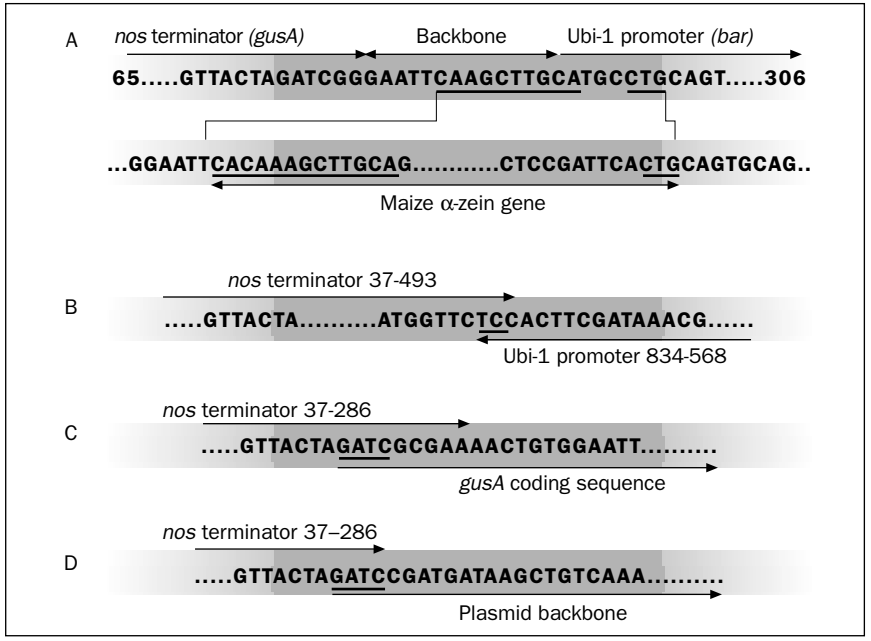

Fig. 3. Microhomology-mediated recombination in transgenic maize. The figure shows transgene junction sequences from transgenic loci in four transgenic maize plants. In each case, short regions of homology (3–10 nt) are involved in the recombination process and these regions of common sequence between the two recombining partners are shown as underlined. This can result in (A) the insertion of genomic DNA into transgene (B) "head-to-head" junctions or (C, D) "head-to-tail" junctions.

Level 2: activation of local repair complexes

Most of the transgenic plants we have analyzed have only a single transgenic locus as defined by genetic segregation analysis. However, as discussed above, individual transgenes and contiguous transgene arrays have been shown to be arranged in a local cluster, also containing significant stretches of genomic DNA. This is not the filler DNA described above, which is usually only a few nucleotides in length.

The presence of significant stretches of genomic DNA within a transgene cluster suggests that there is a pronounced tendency for fragments of exogenous DNA to integrate at one site. Since transgenic loci appear to arise randomly, probably reflecting naturally occurring DNA breaks, we proposed that a primary integration event promoted secondary integration events nearby (Kohli et al 1998). This two-phase mechanism could reflect the recruitment of DNA-repair complexes to the original integration site and the consequent introduction of additional double-strand DNA breaks in the local region. This mechanism would result in a concentrated core of DNA repair and transgene integration, which would likely cause a significant amount of transgene rearrangement as well as the loss of stretches of genomic DNA. Indeed, the deletion of genomic DNA around transgene integration sites is a well-known phenomenon in both transgenic plants and animals.

Level 3: integration at the chromatin level

The third level of transgene organization involves the dispersion of transgene copies and clusters such that individual signals can be visualized by FISH on metaphase chromosomes. This demonstrates that transgene clusters may be separated by megabases of DNA. How is this different from the level 2 organization discussed above? Could it be that repair complexes are being activated over a greater area to generate transgene sites separated by larger intervening regions of genomic DNA? Two pieces of evidence suggest not. First, the intervening regions at level 2 organization can be isolated by long PCR, a reasonable distance to be spanned by locally recruited repair complexes. However, repair complexes would unlikely be recruited to sites megabases away from the initial integration site. Second, the analysis of the same wheat nuclei at interphase shows the remarkable phenomenon that the FISH signals that are separated at metaphase are brought back together at interphase (Abranches et al, n.d.). Since this phenomenon occurs in multiple nuclei from somatic tissues of a given transgenic plant, and in progeny thereof, it confirms that the interphase chromatin is highly and reproducibly organized in the nucleus. We have put forward three models to explain this (Abranches et al, n.d.). First, it is possible that the homologous DNA sequences are associating in *trans*. Second, copies of the same promoter may be recruited to a common transcription factory in the nucleus. Our third model, which provokes the most thought with respect to the mechanisms of transgene integration, is that the FISH signals permit visualization of the three-dimensional configuration of the nucleus at the moment of transformation. It has been shown that cells successfully transformed by particle bombardment usually have the metal particle lodged in the nucleus. Therefore, the metal particle may cause localized damage to DNA in a particular region of the nucleus. The way that chroma-

tin is folded into loops may result in damage to several discrete sites that are separated in *cis* by hundreds of kilobase pairs of DNA but are close together in *trans*.

Transgene rearrangements and the role of recombination "hotspots"

The importance of transgene rearrangements

We have found that transgene rearrangements have profound effects on transgene expression. Transgene rearrangement is a pitfall of all direct DNA transfer methods, but is perhaps more acute in particle bombardment because the forces involved may cause more DNA fragmentation than other methods and because bombarded plant cells may be induced to produce DNA degradation and repair enzymes in response to their injury (Hunold et al 1994). Two important recent discoveries in this area are that transgene rearrangements can be subtle, yet still contribute to loss of activity, and that rearrangements often involve so-called recombination hotspots in the transformation construct.

Detection of subtle rearrangements

We generated several transgenic maize plants carrying three transgenes: *hpt*, *bar*, and *gusA*. We examined all plants by Southern blot analysis and found that even those with apparently intact transgene cassettes could be inactive. To investigate whether transgene inactivity reflected mutation or epigenetic phenomena, we used long PCR to characterize the transgene loci (Mehlo et al, n.d.). We found that, in most cases where intact transgenes gave no expression, there were subtle rearrangements that could not be detected by Southern blot analysis. For example, some lines showed small internal deletions, point mutations, or external erosions. These had the effect of removing PCR primer sites so that amplification of the entire expression cassette failed, whereas typical PCR testing of a small region within the transgene showed no abnormalities. From these experiments, we concluded that copy number/expression correlations should be interpreted with caution since undetected subtle mutations could account for many instances of transgene inactivity.

Recombination hotspots

Our analysis of rice plants concentrated on the structure of plasmid-plasmid junctions. As discussed above, this showed that the formation of contiguous plasmid arrays and the integration of exogenous DNA both involved predominantly microhomology-mediated recombination. Remarkably, however, more than 50% of the junctions we analyzed involved the same region of the plasmid, a palindromic element from within the CaMV 35S promoter (Kohli et al 1999). The same region in the full-length CaMV RNA is known to promote recombination between the genomes of different strains of the virus in plants (Gal et al 1992, Swoboda et al 1993). It has been suggested that virus-specific *cis*-acting elements and virus-encoded enzymes might play a role in determining the nature of recombination events (Schoelz et al 1993). Even in the absence of these factors, however, the CaMV 35S DNA fragment underwent recombination at a high frequency, an observation that

strongly suggested that plant factors can direct recombination events by recognizing and using the highly recombinogenic viral sequences. We established that, at the 3′ end of the CaMV 35S promoter, an imperfect palindrome of 19 bp, in conjunction with specific flanking sequences derived from the transforming plasmid, acts as a hotspot for recombination.

Clean DNA transformation

In all transformation methods, exogenous DNA tends to undergo rearrangement and recombination events leading to the integration of multiple fragmented, chimeric, and rearranged transgene copies. Such clusters are prone to internal recombination and may therefore be unstable as well as promote transgene silencing.

A major breakthrough has been achieved in the last year in the production of transgenic plants without integrated vector backbone sequences. Since the backbone is a major source of recombinogenic elements, such as the origin of replication, a logical approach is to remove these sequences from the transforming vector. The "clean DNA" system (Fu et al 2000b) takes this strategy to its logical extreme by removing all unnecessary backbone elements and using just minimal linear cassettes for transformation: the promoter, open reading frame, and terminator. Transgenic rice plants regenerated from callus bombarded with such cassettes showed very simple integration patterns, with one or a few hybridizing bands on Southern blots. Comparative transformation experiments using whole-plasmid DNA resulted in much more complex integration patterns, with multiple bands of different sizes (Fig. 4). Furthermore, release of a diagnostic fragment from the transgene showed that fewer rearrangement events had occurred in the plants carrying the linear cassettes. The progeny of these plants were examined for transgene expression and there were no instances of silencing among the transgenes studied. Conversely, whole-plasmid transformation by particle bombardment can lead to silencing in up to 20% of the resulting transgenic plants. The simplicity of clean DNA transformation is probably achieved by reducing the overall recombinogenicity (propensity to undergo recombination) of the construct and thereby simplifying the structural hierarchy of the locus. This work showed that minimal linear cassettes promoted the same efficiency of transformation as normal plasmids. Similarly, multiple linear fragments showed the same efficiency of cotransformation as multiple plasmids, so the clean DNA strategy is a promising technique for the generation of transgenic plants carrying multiple transgenes.

Transgene silencing and DNA methylation

Silencing often correlates with DNA methylation, and it is of particular interest to find sequences in plant transgenes that induce methylation and result in unstable expression. There is evidence that methylation can be induced by interactions between homologous transgene copies (or between a transgene and homologous endogene) or may reflect genomic position effects (Gallie 1998, Grant 1999). Recent data suggest, however, that more complex underlying processes may be involved. We addressed

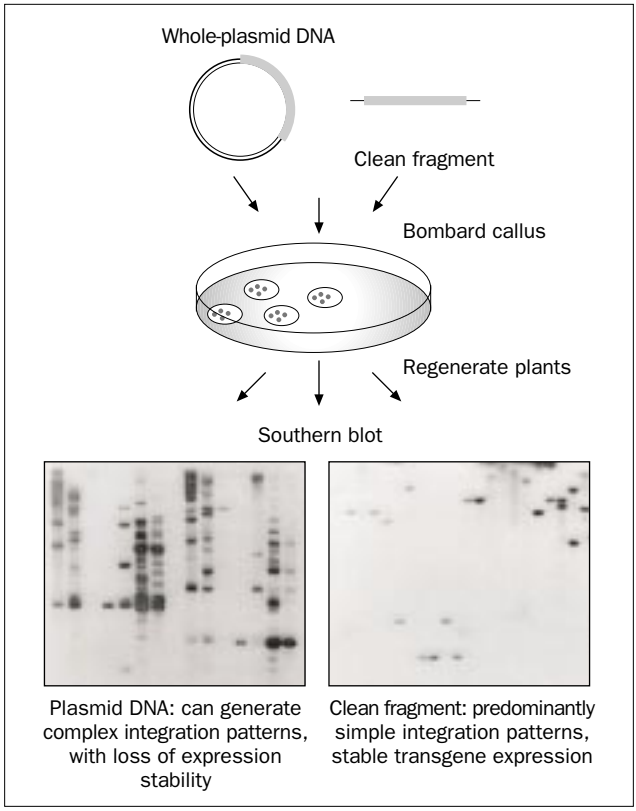

Fig. 4. Clean DNA transformation. Clean DNA is the linear minimal transgene expression cassette, cleaved from the plasmid using restriction enzymes or generated by polymerase chain reaction. In comparative experiments, clean DNA generates simpler integration patterns and a greater proportion of plants with stable transgene expression, even though overall transformation frequencies are similar using the two methods.

this issue by studying a single transgenic rice line containing a single copy of a construct comprising three heterologous transgenes: *hpt*, *gusA*, and *bar* (Fu et al 2000a). By concentrating on this line, we eliminated variation caused by different position effects and different transgene copy numbers, since the transgene locus was identical in structure in each plant we analyzed. Furthermore, because there was only a single copy of the locus and because none of the transgenes were homologous to endogenous rice genes, we eliminated the possibility that ectopic pairing could influence transgene expression.

Despite these precautions, we found that silencing still occurred *de novo* in several R_1, R_2, and R_3 plants. In all cases, silencing was associated with DNA methylation of the silenced locus, but in some cases this was heritable through meiosis and in

other cases it was not. We observed a novel form of silencing in the *hpt* transgene, involving the methylation of one DNA strand, with the result that the silencing phenotype was inherited by 50% of the progeny in each generation. We also observed developmentally regulated silencing of *gusA*, which was active at 3 weeks postgermination but inactive at 6 weeks. We found that silencing could affect the *hpt* gene without influencing the adjacent *gusA* gene, and vice versa, even though the two transgenes were separated by less that 50 bp of intervening DNA. Most remarkably of all, we found that the *hpt* and *gusA* genes could be differentially silenced in the same plant, with each silencing mode associated with an entirely distinct and nonspreading type of DNA methylation (Fig. 5). For all these variations to occur in sibling plants with the same transgenic locus suggests that there is still much to learn about the factors responsible for DNA methylation and silencing.

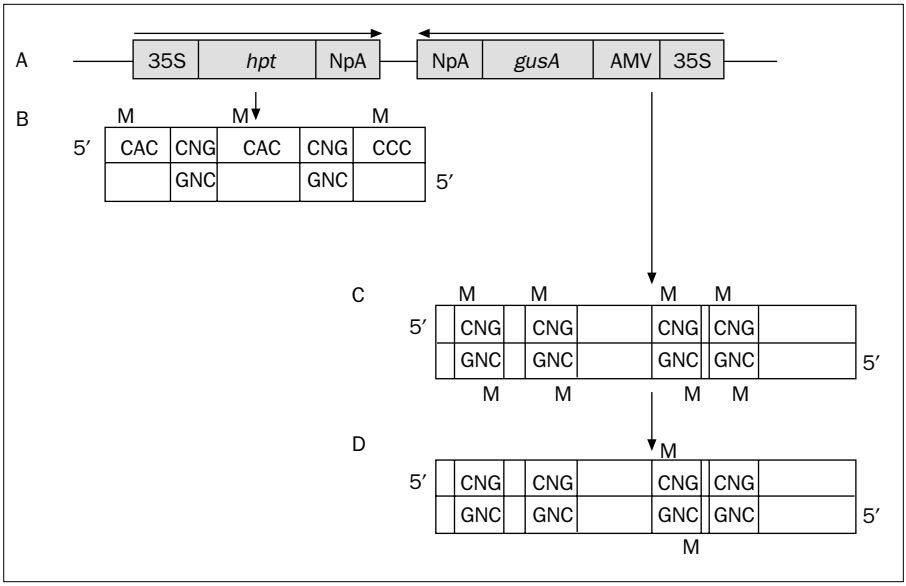

Fig. 5. Methylation and transgene silencing in a single rice plant. (A) Transgenic rice line C549 carried a single-copy transgenic locus comprising three heterologous transgenes: *bar* (not shown), *hpt*, and *gusA*. Each transgene was separated by less than 100 bp of DNA. In one transgenic R_1 plant, the *hpt* and *gusA* genes simultaneously underwent different modes of silencing, associated with distinct, nonspreading methylation patterns. (B) The *hpt* gene underwent transcriptional silencing, associated with hemimethylation of the coding region at nonconventional (non-CNG) sites. The silencing was transmitted to 50% of progeny plants, presumably those inheriting the methylated DNA strand. (C) The *gusA* gene concurrently underwent posttranscriptional silencing, associated with methylation at conventional CNG sites in the coding region. However, while these CNG sites were methylated at 3 weeks postgermination (C), most were demethylated at 6 weeks postgermination (D), corresponding to the onset of *gusA* gene expression.

Concluding comments

Experiments involving cotransformation with multiple transgenes, in combination with FISH analysis, have given us a unique perspective of how hierarchical organization affects the level and stability of transgene expression in cereals. Undoubtedly, further studies need to be carried out to completely unravel the factors that permit stable and predictable transgene expression, and these studies may offer a glimpse of the complex interacting mechanisms involved. In the future, it may be possible to design transformation constructs that permit full predictable control over transgene integration and therefore provide guaranteed routes to expression. Until this time, it is important to continue investigating the mechanisms that control transgene integration, organization, and expression in the cereals we all rely on so heavily.

References

Abranches R, Santos AP, Williams S, Wegel E, Castilho A, Christou P, Shaw P, Stoger E. n.d. Widely-separated multiple transgene integration sites in wheat chromosomes are brought together at interphase. Plant J. (In press.)

Bano-Maqbool S, Christou P. 1999. Multiple traits of agronomic importance in transgenic indica rice plants: analysis of transgene integration patterns, expression levels and stability. Mol. Breed. 5:471-480.

Berg DE, Howe M. 1989. Mobile DNA. Washington, D.C. (USA): American Society of Microbiology.

Chen LL, Marmey P, Taylor NJ, Brizard JP, Espinoza C, de Cruz P, Huet H, Zhang SP, de Kochko A, Beachy RN, Fauquet CM. 1998. Expression and inheritance of multiple transgenes in rice plants. Nature Biotechnol. 16:1060-1064.

Christou P, McCabe D, Swain WF. 1988. Stable transformation of soybean callus by DNA-coated gold particles. Plant Physiol. 87:671-674.

De Groot MJ, Offringa R, Groet J, Does MJ, Hooykaas PJ, van dan Elzen PJ. 1994. Non-recombinant background in gene targeting: illegitimate recombination between *hpt* gene and defective 5′ deleted *nptII* gene can restore a Km^r phenotype in tobacco. Plant Mol. Biol. 25:721-733.

Fromm M, Callis J, Taylor LP, Walbot V. 1987. Electroporation of DNA and RNA into plant protoplasts. Methods Enzymol. 153:351-366.

Fu X, Kohli A, Twyman RM, Christou P. 2000a. Alternative silencing effects involve distinct types of non-spreading cytosine methylation at a three-gene single-copy transgenic locus in rice. Mol. Gen. Genet. 263:106-118.

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

Gal S, Pisan B, Hohn T, Grimsley N, Hohn B. 1992. Agroinfection of transgenic plants leads to viable cauliflower mosaic virus by intermolecular recombination. Virology 187:525-533.

Gallie DR. 1998. Controlling gene expression in transgenics. Curr. Opin. Plant Biol. 1:166-172.

Gheysen G, Villarroel R, van Montagu M. 1991. Illegitimate recombination in plants: a model for T-DNA integration. Genes Dev. 5:287-297.

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.
Grant SR. 1999. Dissecting the mechanisms of post-transcriptional gene silencing: divide and conquer. Cell 96:303-306.
Hunold R, Bronnel R, Hahne G. 1994. Early events in microprojectile bombardment—cell viability and particle location. Plant J. 5:593-604.
Klein TM, Wolf ED, Wu R, Stanford JC. 1987. High-velocity microprojectiles for delivering nucleic acids into living cells. Nature 327:70-73.
Klein TM, Fromm M, Weissinger A, Tomes D, Schaaf S, Sletten M, Stanford JC. 1988. Transfer of foreign genes into intact maize cells with high-velocity microprojectiles. Proc. Natl. Acad. Sci. USA 85:4305-4309.
Kohli A, Leech MJ, 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, 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.
Komari T, Hiei Y, Ishida Y, Kumashiro T, Kubo T. 1998. Advances in cereal gene transfer. Curr. Opin. Plant Biol. 1:161-165.
Krens FA, Molendijk L, Wullems GJ, Schilperoort RA. 1982. In vitro transformation of plant protoplasts with Ti-plasmid DNA. Nature 296:72-74.
Matsumoto S, Ito Y, Hosoi T, Takahashi Y, Machida Y. 1990. Integration of T-DNA into tobacco chromosome: possible involvement of DNA homology between T-DNA and plant DNA. Mol. Gen. Genet. 224:309-316.
Mayerhofer R, Koncz-Kalman Z, Nawrath C, Bakkeren G, Crameri A, Angelis K, Redei GP, Schell J, Hohn B. 1991. T-DNA integration: a mode of illegitimate recombination in plants. EMBO J. 10:697-704.
Mehlo L, Mazithulela G, Twyman RM, Boulton MI, Davies JF, Christou P. n.d. Structural analysis of transgene rearrangements and effects on expression in transgenic maize plants generated by particle bombardment. Maydica. (In press.)
Negrutiu I, Shillito R, Potrykus I, Biasini G, Sala F. 1987. Hybrid genes in the analysis of transformation conditions. I. Setting up a simple method for direct gene transfer in plant protoplasts. Plant Mol. Biol. 8:363-373.
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.
Schoelz JE, Wintermantel WM. 1993. Expansion of viral host range through complementation and recombination in transgenic plants. Plant Cell 5:1681-1692.
Swoboda P, Hohn B, Gal S. 1993. Somatic homologous recombination *in planta*: the recombination frequency is dependent on the allelic state of the recombining sequences and may be influenced by genomic position effects. Mol. Gen. Genet. 273:33-40.
Takano M, Egawa H, Ikeda JE, Wakasa K. 1997. The structures of integration sites in transgenic rice. Plant J. 11:353-361.
Tang K, Tinjuangjun P, Xu Y, Sun X, Gatehouse JA, Ronald PC, Qi H, Lu X, Christou P, Kohli A. 1999. Particle bombardment-mediated co-transformation of elite Chinese rice cultivars with genes conferring resistance to bacterial blight and sap-sucking insect pests. Planta 208:552-563.

Tinland B. 1996. The integration of T-DNA into plant genomes. Trends Plant Sci. 1:178-184.
Tsukamoto Y, Ikeda H. 1998. Double-strand break repair mediated by DNA end-joining. Genes Cells 3:135-144.
Twyman RM, Whitelaw CAB. 2000. Animal cell genetic engineering. In: Spier RE, editor. Encyclopedia of cell technology. New York (USA): John Wiley & Sons. p 737-819.
Twyman RM, Christou P, Stoger E. n.d. Genetic transformation of plants and their cells. In: Oksman-Caldentey KM, Barz W, editors. Plant biotechnology and transgenic plants. New York (USA): Marcel & Dekker. (In press.)

Notes

Authors' address: Molecular Biotechnology Unit, John Innes Center, Colney Lane, Norwich NR4 7UH, U.K.
Citation: Khush GS, Brar DS, Hardy B, editors. 2001. Rice genetics IV. Proceedings of the Fourth International Rice Genetics Symposium, 22-27 October 2000, Los Baños, Philippines. New Delhi (India): Science Publishers, Inc., and Los Baños (Philippines): International Rice Research Institute. 488 p.