# 12. Inactivation of Repeated Genes – DNA-DNA Interaction?

MARJORI MATZKE, ANTONIUS J.M. MATZKE and
ORTRUN MITTELSTEN SCHEID

**Contents**



## Introduction

Homology-dependent gene silencing, in which the presence of two or more totally or partially homologous copies of a nuclear gene can lead to inactivation of at least one gene copy, has provided a fresh perspective for understanding fundamental aspects of gene expression in plants. Several



*J. Paszkowski (ed.), Homologous Recombination and Gene Silencing in Plants*, 271–307.

reviews have appeared recently on this topic (Matzke & Matzke 1993a, 1993b; Jorgensen 1992b, 1993), and various possible mechanisms have been proposed. Our aim is not to duplicate these previous reviews, but to consider one specific hypothetical mechanism that involves the recognition and pairing of homologous DNA sequences in somatic cells, followed by the transmission of epigenetic information between these regions. As should be immediately evident, the initiating steps of this mechanism – homology sensing and pairing – can be common to gene inactivation and homologous recombination. Therefore, we regard homology-dependent gene silencing as an integral part of a discussion on homologous recombination in plants.

The rationale for considering DNA-DNA pairing is that at least some cases of gene silencing in plants, particularly those involving transgene-transgene interactions, are strikingly reminiscent of known pairing-dependent phenomena that inactivate gene expression and produce epigenetic modifications in other organisms. It is important to emphasize, however, that a single mechanism may not suffice to explain all cases of homology-dependent gene silencing in plants. Alternative mechanisms involving post-transcriptional controls or autoregulation are also conceivable. As discussed by other authors in this volume (see Mol et al. chapter 00 and Meins & Kuntze chapter 00), these mechanisms might be particularly suited to explain cases of co-suppression involving endogenous genes and homologous transgenes. For the purpose of this review, however, we will adopt the point of view that many, if not all, of the gene silencing cases in plants are due to DNA-DNA pairing.

We begin with a survey on homology-dependent gene inactivation in plants, including features which might promote or otherwise influence pairing between homologous regions. We will then discuss recent studies on somatic homologous recombination in plants and compare these findings to general characteristics of homology-dependent gene silencing. To support the proposal that homology-dependent gene silencing in plants results from DNA-DNA pairing, we will present examples from *Drosophila* and fungi which illustrate that pairing-dependent phenomena in these organisms can both alter gene expression and produce epigenetic modifications. We will also discuss the seeming absence of such effects in transgenic mammals, and possible reasons for this. Ways to test the pairing hypothesis will also be considered. Finally, we will present our own views on possible evolutionary and developmental implications of homology-dependent gene silencing.

## Survey on Homology-Dependent Gene Silencing in Plants

Homology-dependent gene silencing can occur between multiple gene copies that are closely linked or arranged in tandem arrays, and between homologous genes at allelic or non-allelic (ectopic) locations. There have been several variations of homology-dependent gene silencing reported, and this has given rise to individualistic terminology. Figure 1 shows the relationships among

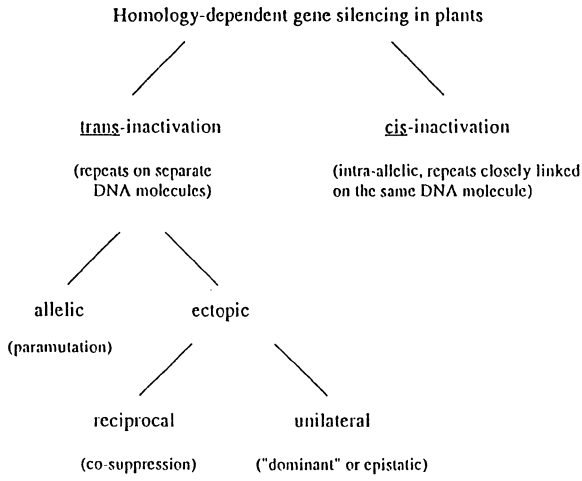


*Figure 1.* Variants of homology-dependent gene silencing in transgenic plants. Inactivation of homologous genes (also termed repeat-induced gene silencing (RIGS), Assaad et al. 1993) can occur if they are present as closely linked repeats on the same DNA molecule (intra-allelic or *cis*-inactivation), or if the repeats are present on separate DNA molecules (*trans*-inactivation). *Trans*-inactivation can be further subdivided into those interactions which involve alleles on homologous chromosomes, and those which occur between homologous genes present at unlinked (ectopic) loci. An allelic interaction that results in a heritable epigenetic change of one allele in the presence of the other is termed paramutation in plants. Ectopic interactions include cases in which both copies of the homologous gene pair are inactivated (co-suppression), and cases in which only one genetic locus is changed by the interaction ("dominant" or epistatic).

these different variants of the phenomenon. In this chapter, we will use the term *cis*-inactivation to refer to the suppression of multiple, closely linked copies; allelic *trans*-inactivation and ectopic *trans*-inactivation will be used to refer to inactivation of multiple copies at either allelic or non-allelic sites, respectively.

Homology-dependent gene inactivation has become a topic of broader interest only recently as a consequence of gene transfer techniques. To our knowledge, all studies on homology-dependent silencing of transgenes were initiated by involuntary observations, when experimental results differed from experimenters' predictions. In particular, numerous cases of transgene inactivation were revealed by the following types of experiments: 1) attempts to introduce multiple genes by sequential transformation; 2) unintended integration of more than one transgene copy; 3) trials to achieve over-expression of plant genes; and 4) use of "sense" controls in studies on the effect of antisense constructs. Since the first publications reflecting a clear perception of the phenomenon date back only five years, most of the data must be extracted from initial studies in which experiments were not designed to approach the underlying mechanism. In Table 1, we have compiled the

*Table 1.* Survey on homology dependent transgene inactivation in plants.

| plant | gene | promoter | copy number | minimal length of homology | description of effect | pattern of effect | frequency of affected transformants | reference |
|---|---|---|---|---|---|---|---|---|
| Nicotiana tabacum | NPTII/HPT NOS/OCS | NOSpro | 2 unlinked transgene loci | 0.3 kb (promoter) | loss of kanR loss of NOS production | complete, different degree after segregation | 15 % | Matzke et al. 1989 Matzke and Matzke 1990, 1991, 1993 a,b Matzke et al. 1993, and unpublished results |
| Petunia hybrida | CHS | CaMV 35S CHSpro | transgene and endogenous copy | 1.17 kb (coding region) | loss of CHS expression, reduced anthocyanin synthesis | complete or partial (variegated), symmetric or asymmetric | 32-100 % | Napoli et al. 1990, Jorgensen 1990, 1992a,b, Taylor and Jorgensen 1992 |
| Petunia hybrida | CHS DFR | CaMV 35S CHSpro | transgene and endogenous copy, no correlation with transgene copy number | 1.2-8 kb 1.3 kb (coding region) | loss of CHS/DFR expression, reduced anthocyanin biosynthesis | complete or partial (variegated), symmetric or asymmetric | 5-25 % | van der Krol et al. 1990 |
| Lycopersicon esculentum | PG | CaMV 35S | transgene and endogenous copy | 0.73 kb (coding region) | reduction of PG activity | | 93 % | Smith et al. 1990 |
| Nicotiana tabacum | GUS | CaMV 35S | 2 transgene copies and more | 2.2 kb (coding region) | reduction of GUS expression | either high or low expression | 30 % | Hobbs et al. 1990 Hobbs et al. 1993 |
| Nicotiana tabacum | PAL | CaMV 35S-PALpro fusion | transgene and several endogenous genes | | reduction of the endogenous PAL expression, phenotypic symptoms | | 30 % | Elkind et al. 1990 |
| Petunia hybrida | A1 | CaMV 35S | 1 - <17 | ca. 4.3 kb (promoter and coding region | loss of flower pigmentation | complete or partial (variegated) | 65 % | Linn et al. 1990, Meyer et al. 1992, Meyer et al. 1993 |
| Nicotiana tabacum | NOS | NOSpro CaMV 35S | 1 complete and one truncated transgene copy | 0.86 kb (coding region) | loss of NOS production | complete, stronger effect than antisense construct | 100 % | Goring et al. 1991 |
| Arabidopsis thaliana | HPT | CaMV 35S | ca. 10 transgene copies in single loci | 1.8 kb (promoter and coding region) | loss of hygR | complete | 50 % | Mittelsten Scheid et al. 1991, and unpublished data |
| Nicotiana sylvestris | CHN | CaMV 35S | 1-2 heterologous transgenes and endogenous copy | ca. 1 kb (coding region) | loss of chitinase expression | either high or low expression | 100 % | Neuhaus et al. 1991, Hart et al. 1992 |
| Arabidopsis thaliana | NPTII | NOSpro | line E and G: 2 and 5 transgenes in single loci | ca. 5 kb (promoter and coding region) | loss of kanR | complete | ca. 50 % in total, 29 % in later generations | Kilby et al. 1992 |
| Nicotiana tabacum | PK | CaMV 35S | 2 unlinked transgene loci | ca. 4 kb (promoter and coding region) | reduction of PK activity | | | Gottlob-McHugh et al. 1992 |
| Nicotiana tabacum | GUS | CaMV 35S | | | loss of GUS expression, no loss of kanR | | 3 % | Bochardt et al. 1992 |
| Medicago sativa | PAT | CaMV 35S | single transgene copy | | loss of PPT resistance | complete | (all cultures derived from one transformant) | Walter et al. 1992 |
| Nicotiana tabacum | GN1 | CaMV 35S | only heterologous transgene | ca. 2 kb (coding region without introns) | loss of glucanase expression from transgene | | 20 % | de Carvalho et al. 1992 |
| Arabidopsis thaliana | HPT NPTII | CaMV 35S | most pronounced with 2 and more transgene copies | ca. 4 kb (promoter and coding region) | loss of hygR and kanR | complete or partial | (all lines derived from one transformant) | Assaad et al. 1993 |
| Nicotiana tabacum | NiR, GUS, NPTII, HPT, ALS, NR, rolA, rolC, Aux2 | CaMV 35S, CaMV 19S, NOSpro, 1'pro, NRpro | 2 transgene copies or transgene and endogenous gene | 0.09 kb (promoter) | loss of NiR activity loss of marker gene expression | complete or partial | | H. Vaucheret, 1993 |

*Abbreviations*

| | | | |
|---|---|---|---|
| ALS | acetolactate synthase | CHS | chalcone synthase |
| Aux2 | gene from Agrobacterium rhizogenes | CHSpro | promoter of chalcone synthase |
| A1 | dihydroflavonol 4-reductase of Zea mays | DFR | dihydroflavonol reductase |
| | | GN1 | β-1,3-glucanase 1 |
| CaMV 35S | promoter of cauliflower mosaic virus 35S transcript | GUS | β-glucuronidase |
| | | HPT | hygromycin phosphotransferase |
| CaMV 19S | promoter of cauliflower mosaic virus 19S transcript | hygR | hygromycin resistance |
| | | kanR | kanamycin resistance |
| CHN | chitinase | NiR | nitrite reductase |

*Table 1. (Continued).*

| reversibility | inheritance | maternal/ paternal bias | environ-mental influence | type of interaction | methylation, hetero-chromatin | modifiers | references |
|---|---|---|---|---|---|---|---|
| reversible on segregation and outcrossing | stable, some persistence even after segregation possible | yes | | trans, ectopic, after transformation and crossing, only one locus inactivated, epistatic | specific methylation | different degree of suppression and susceptibility, difference homo-/hemizygotes | Matzke et al. 1989 Matzke and Matzke 1990, 1991, 1993 a,b Matzke et al.1993, and unpublished results |
| spontaneous within a plant and after segregation | stable, no persistence after segregation | | yes | trans, ectopic, co-suppression | | different degree among transformants | Napoli et al. 1990, Jorgensen 1990,1992a,b, Taylor and Jorgenson 1992 |
| spontaneous within a plant | | | yes | trans, ectopic, co-suppression | | different degree among transformants | van der Krol et al. 1990 |
| | stable, no persistence after segregation | | | trans, ectopic, co-suppression, linked to expression | | | Smith et al. 1990 |
| reversible after segregation | stable, no persistence | no | no | cis, intra allelic and trans, ectopic, co-suppression on sequential trans-formation, additive for active copies, epistatic | specific methylation | | Hobbs et al. 1990, Hobbs et al. 1993 |
| reversible on segregation | | | yes | trans, ectopic, only endogenous gene inactivated | | difference homo-/hemizygotes | Elkind et al. 1990 |
| weak reversion on outcrossing | stable persistence (paramutation) | yes | yes | (single copy), trans, allelic, dominant or semi-dominant | specific methylation, but no absolute correlation | different degree among transformants, developmental factors | Linn et al. 1990, Meyer et al. 1992, Meyer et al. 1993 |
| | stable, some persistence after segregation possible | | | trans, ectopic, co-suppression | no | | Goring et al. 1991 |
| reversible on outcrossing and by tissue culture | stable | no | | cis, intra-allelic and trans, ectopic, all copies inactivated, epistatic | some methylation, but no specific correlation | different degree of suppression and susceptibility, difference homo-/hemizygotes | Mittelsten Scheid et al. 1991, and unpublished data |
| stable in one plant, reversible in next generation | no | | yes | trans, ectopic, co-suppression | no | developmental factors, difference homo-/hemizygotes | Neuhaus et al. 1991, Hart et al. 1992 |
| | yes | | | cis, intra-allelic or trans, allelic? | specific methy-lation, but no absolute corre-lation, reacti-vation by 5aza | progressive in subsequent generations | Kilby et al. 1992 |
| | yes | | | trans, ectopic, co-suppression | | tissue-specific suppression | Gotlieb McHugh et al. 1992 |
| reversed in young progeny | yes | | | | reactivation by 5aza | progressive with plant maturation | Bochardt et al. 1992 |
| not observed | | | yes (tempe-rature) | | | type of culture | Walter et al. 1992 |
| reversible after segregation and outcrossing | stable | | | no allelic interaction, gene dosage effect, endogenous gene not affected, posttranscriptional | | developmental factors, difference homo-/hemizygotes | de Carvalho et al. 1992 |
| spontaneous within single plants | stable | | yes (light) | trans, allelic or cis, intra-allelic | specific methylation | different degree in different genotypic combinations | Assaad et al. 1993 |
| | | | | trans, ectopic, after crossing, co-suppression | | | Vaucheret, 1993 |

| | | | |
|---|---|---|---|
| NOS | nopaline synthase | PAT | phosphinotricin acetyltransferase |
| NOSpro | promoter of nopaline synthase | PG | polygalacturonase |
| NR | nitrate reductase | PK | pyruvate kinase |
| NRpro | promoter of nitrate reductase | RolA | gene from Agrobacterium rhizogenes |
| NPTII | neomycin phosphotransferase II | RolB | gene from Agrobacterium rhizogenes |
| OCS | octopine synthase | 1'pro | mannopine synthase 1' promoter |
| PAL | phenylalanine ammonia-lyase | 5aza | 5-azacytidine |
| PALpro | promoter of phenylalanine ammonia-lyase | | |

available information according to criteria that might be relevant for discussing a probable involvement of homology recognition by DNA-DNA interaction. We hope that the reduction of information in Table 1 is balanced by the opportunity for synoptic comparison. Although these examples all involve transgenes, it is our view that homology-dependent gene inactivation can also occur between repeated endogenous genes. Transgenes have allowed a more precise analysis of the phenomenon, and can serve as reporters of basic processes affecting all nuclear genes.

*Plants*

The predominance of *Solanaceae* among the plant species listed is probably not relevant for the incidence of homology-dependent gene silencing, but rather reflects the ease of regeneration and transformation techniques for this family. In fact, there are indications that instability and variability of transgene expression occur also in monocotyledonous plants (Klein et al. 1990). Therefore, it can be assumed that transgene interactions leading to loss of activity are a widespread phenomenon among higher plants.

Because methylation has been shown to play a role in a number of cases of homology-dependent gene silencing (see below), it might be interesting to consider the variations in GC content and methylcytosine (mC) among different species. The genomes of the monocots maize and rye have GC contents of ca. 50% and ca. 7% mC. The genomes of the dicots tobacco and *Arabidopsis* both have GC contents of ca. 40%, but whereas the mC content of tobacco is ca. 7%, *Arabidopsis* has only ca. 1% mC (Matassi et al. 1992). Whether this low mC content in *Arabidopsis* is a reflection of the relative lack of repetitive sequences, which could become methylated, or other factors is not known.

*Genes*

The genes that are affected by homology-dependent inactivation are composed of two groups, the widely used selectable or screenable marker genes of mostly bacterial origin, and a variety of plant derived genes. Whether these genes as a group share any particular features that make them susceptible to inactivation is not known. Bacterial genes may differ from plant genes in their overall base composition or in the density of potential methylation sites, perhaps making them more sensitive to inactivation. Plant genes, however, especially those that are homologous to the residing genes in the transformed species, cannot be expected to carry any such label in their sequence. Their susceptibility to undergo inactivation could nevertheless still be influenced by the length of homology uninterrupted by introns as well as the copy number of the residing gene families (see below). The number of endogenous plant genes analysed until now is far too small to allow any comparisons of sequence requirements for being either sensitive or resistant to silencing.

*Promoters*

Most transgenes under investigation are controlled by either the nopaline synthase promoter (NOSpro) derived from *Agrobacterium tumefaciens* or the promoter of the 35S transcript of cauliflower mosaic virus (CaMV 35S). Although both promoters are presumably of non-plant origin, they have evolved to give high levels of transcription in plant cells. They nevertheless represent genetic material from invading pathogens, which conceivably could be more easily controlled than endogenous sequences by inactivation through methylation (Bestor 1990; Doerfler 1991). It should be emphasized, however, that regulation by one of these two foreign promoters is not a requirement for obtaining homology-dependent gene silencing, because inactivation has also been seen with an introduced plant gene regulated by its own promoter (van der Krol et al. 1990). Furthermore, in cases of ectopic *trans*-inactivation involving the NOSpro, the suppression and methylation observed are clearly not simply a response to the foreign promoter, since the inactivation only occurs after the introduction of a second construct containing additional copies of the NOSpro (Matzke et al. 1989, 1993).

*Copy Number*

As a general rule, gene inactivation is observed more frequently as soon as two or more sequences with a certain degree of homology are present in the same genome. However, actually counting the copy numbers of the genes involved is not as easy as it might appear at first sight. Obviously, in all cases of attempted over-expression of a homologous, endogenous gene there are usually three copies present, two allelic copies coming from the diploid genome and one ectopic transgene copy in the case of a hemizygote transformant. This number can be higher if transformants are bred homozygous for the transgene or if multiple copies are integrated simultaneously (these can either be linked or unlinked). It is also difficult to "count" the number of gene copies when an introduced plant gene is derived from a different ecotype, line or species, because the degree of sequence similarity to the corresponding endogenous gene, which has usually not been isolated in these cases, cannot be determined.

For non-plant transgenes, a correlation between the number of integrated copies and the frequency of inactivation is well documented in many publications. This correlation is particularly solid for copies arranged in *cis*. A reduction of copy number inside a locus was shown to increase gene expression (Mittelsten Scheid et al. 1994) or decrease the suppressing effect (Matzke et al. 1994). For unlinked copies, the situation is more complicated, since it is clear in a number of cases that multiple unlinked copies can be expressed without interference in some transformants. The multiplicity does not need to include the whole gene: duplication of either the promoter or addition of a truncated transgene coding region can cause decreased expression.

*Minimal Length of Homology*

The values listed in Table 1 for the length and degree of homology that are required to stimulate gene inactivation cannot be regarded as conclusive, and a systematic evaluation of this aspect is required. As mentioned above, in many cases, plant genes from a different species were chosen for transformation since the endogenous genes have not yet been characterized. In such cases, the degree and distribution of homology between the interacting genes is not known. This is further complicated when residing genomic copies contain relatively long introns that are lacking in the transgenic copy if this is derived from a cDNA template. In fact, the division of coding sequences into fairly small exons might serve to prevent pairing with a pseudogene or cDNA copy (Kricker et al. 1992).

In the case of interactions between repeated copies of the same transgene, there is 100% homology along the length of the transgene. In addition, there may be flanking sequences derived from the vector that provide additional regions of complete homology. Inactivation has been seen with a truncated transgene of 730 bp (Smith et al. 1990), and in the case of homology only within the NOSpro with a length of ca. 300 bp (Matzke et al. 1989). Even 90 bp overlap may be sufficient to affect gene expression (Vaucheret 1993).

*Description of Effect*

All phenomena listed here were described as gene inactivation, silencing or suppression characterized by a total loss or severe reduction of gene expression. Quantitative measurements are not very conclusive in those cases where expression is at or near the limits of detection. Usually such measurements have been made often on the activity of gene products. These measurements reflect different levels of regulation including stability and degradation of proteins and might distort the image of gene activity. In a few individual cases, transcription initiation from silent genes has been determined by nuclear run-off experiments (Mol et al. 1991; de Carvalho et al. 1992). More widespread use of this technique should resolve whether other silent genes are still transcribed.

The focus on gene inactivation should not overlook a report about ectopic activation of an endogenous gene by introduction of a homologous transgene (Hirel et al. 1992). Although only one such case has been described so far, the phenomenon of transvection in *Drosophila* supplies a precedent for "*trans*-activation" (see below, Pairing-dependent phenomena that affect gene expression in *Drosophila* and fungi). Therefore, one should not disregard the possibility that gene interaction in plants may also result in coordinated activation as well as inactivation of homologous genes.

*Pattern of Effect*

The ability to describe the pattern of gene inactivation, that is the distribution of cells containing the silenced gene within a plant, depends on the resolution of the assayed phenotype. Inactivation of marker genes with a cell autonomous change of phenotype allows the detection of gene inactivation patterns on the cellular level. This is clearly seen in the case of the flower pigmentation genes in petunia. The frequent occurrence of variegated flowers demonstrates that gene inactivation can be partial with regard to organs or tissue where the gene is expected to be expressed. The patterns observed are either symmetric or asymmetric. However, histological investigation has shown that, for a single cell, even in such a variegated tissue the gene is either "on" or "off". The activity status does not necessarily correlate with the clonal origin of the individual cells since borders of pigmentation are not identical with tissues derived from different layers of the meristem (Jorgensen 1992a).

Activity patterns of non-cell autonomous genes are difficult to evaluate and require the analysis of samples containing thousands of cells. In spite of the inherent uncertainties in investigating mixtures of cells with active or inactive genes respectively, there are examples where tissue samples can be clearly divided into two categories of either high or low expression (Hobbs et al. 1990; Hart et al. 1992).

Intermediate levels of gene activity seem to be observed frequently for the antibiotic resistance markers, and this is most evident after genetic transmission and segregation. Whether this is due to an intermediate level of gene expression in each individual cell or is a result of epichimeric tissues having a different ratio between cells with completely active or inactive genes, respectively, can only be answered partially by resistance assays with whole seedlings or explants. Green and white patches in seedlings on selective medium indicate cellular mosaicism in the expression of antibiotic resistance markers (Matzke & Matzke 1991; Assaad et al. 1993). Methylation patterns (see below, Gene expression affected by pairing of duplicated sequences in filamentous fungi) after clonal propagation of protoplast derived single cells originating from epichimeric leaves indicate that the epichimeric state is re-established in all cases, although with a varying ratio between methylated and unmethylated copies (M Matzke et al. unpublished observation). A similar reconstitution of mosaic methylation patterns has been observed in cloned lines of animal cells (Silva et al. 1993).

*Frequency of Affected Transformants*

In assessing the frequency of affected transformants, several factors must be considered. First, the cases in Table 1 must be distinguished on the basis of whether the foreign DNA has been integrated by direct gene transfer (Linn et al. 1990; Mittelsten Scheid et al. 1991) or T-DNA-based vectors (all others). In the case of the former, genetically stable single-copy lines could be

established (Potrykus et al. 1985), but multiple copies in *cis* are more often obtained, and one might therefore expect that many of the individual transformants show inactivation of some or all copies. In the case of T-DNA integration, ca. 30–50% of the inserts are single copies, and initial genetic studies indicated that these were reliably recovered as active inserts (Heberle-Bors et al. 1988; Deroles & Gardner 1988). The remaining T-DNA transformants can contain multiple inserts that can be present at one locus (in which case they are often in inverse orientations or complex arrangements) or at multiple segregating loci. The former in particular were associated with weaker-than-expected activity in early experiments (Jones et al. 1987).

It is also necessary to consider separately cases involving interactions between multiple transgene copies that have been integrated in a single transformation step and the cases of ectopic *trans*-inactivation, in which a sequence homologous to an endogenous gene is introduced. In these latter cases, the locus that is suppressed after introduction of a homologous sequence was previously expressed. Therefore, the frequency of affected transformants means those that show inactivation of a previously active gene, not those that show only inactivation of the introduced sequences. In cases of ectopic *trans*-inactivation, the percentage of transformants showing inactivation of one or both homologous genes has ranged from 3–100% in different experiments. This large spread in values could reflect both the condition of the plant material that was transformed and the susceptibility of the endogenous gene to homology-dependent gene inactivation. However, an average value of ca. 50% would be consistent with the ca. 50% of weakly expressing or silent T-DNA inserts. Although some reports have claimed no correlation between transgene copy number and ability to suppress endogenous genes (e.g. van der Krol et al. 1990), it is probable that multicopy transgene inserts are more effective suppressors (Matzke et al. 1993; Matzke et al. 1994).

While inactivation associated with transgenes is not a rare phenomenon, it is also not the norm. Stable transgene expression is often obtained, particularly with single transgene copies lacking repetitive elements. The numerous examples of genetically engineered plants entering into applications also illustrate that transgenes can be stably expressed for several generations (Leemans 1993).

*Reversibility*

Reversibility is an important criterion to all the cases of gene inactivation considered here as epigenetic phenomena, since this excludes gene inactivation resulting from the loss of the functional gene sequence by deletion, recombination or mutation. In fact, reversion from an inactive to an active state was observed so repeatedly that back mutation can be dismissed as an explanation for the gene silencing phenomena in plants. The transient nature of the inactive state was confirmed for nearly all experimental systems.

Reversion can either happen spontaneously within a plant, or be stimulated by tissue culture, and it can also occur after genetic segregation or outcrossing. The reversed state is eventually further propagated to progeny cells or progeny plants. However, reversion may be incomplete or even unstable, with a high tendency for repeated inactivation (Mittelsten Scheid et al. 1991).

*Inheritance*

Plants with inactivated genes often give rise to progeny with the same phenotype. Therefore, the maintenance of the inactive state in successive generations has been described as stable inheritance. In some cases, this transmission was coupled to the presence of the transgene(s), being reversed directly after segregation from a hemizygous parental plant or by backcrosses with the wild-type. Other cases in which an effect persisted after segregation can be regarded as an imprinting of heritable information on the locus or a paramutation-like effect. This, together with the reversibility of the phenomenon, classifies it as epigenetic behavior. Accordingly, the transmission does not necessarily follow a Mendelian segregation and can give rise to unexpected ratios of segregation (Mittelsten Scheid et al. 1991; Assaad et al. 1993). Judgment of transmission is problematic if the phenotype can be scored only in late developmental stages, e.g. flower pigmentation. Since inactivation is known to occur frequently during somatic growth, it cannot be determined whether genes are "on" or "off" in the cells of young plantlets. Screening at the stage of seedlings, as in the case of resistance assays, is closer to the level of germinal inheritance but is also not precise enough to provide a satisfying answer. There is one exception where silenced genes are reset to active forms during transmission (Hart et al. 1992). In the same study, plants with an inactive gene produce progeny seeds that often undergo silencing in later developmental stages. Inheritance here could imply the transmission of a program to undergo inactivation with a certain probability, an alternative that would not exclude the examples with "obvious" sexual transmission.

*Maternal/Paternal Bias*

It is well known for certain animal genes and transgenes that the inheritance of gene expression states can be influenced by the transmission of a gene through either the male or female parent (Surani 1991). A few examples of parental imprinting of endogenous plant genes and chromosomal regions are also known (Matzke & Matzke 1993b). Reciprocal crosses to reveal these differences for plant transgenes have not been done extensively, and most of them did not show a significant difference. A parental bias has been described in two cases (Matzke et al. 1993; Meyer et al. 1993).

*Environmental Influence*

Among the factors that can influence the frequency of gene inactivation, environmental conditions were repeatedly observed to have striking effects. Again, these features were discovered unintentionally due to variations in plant growth conditions. Increased light intensity, increased temperature or other, as yet undefined parameters inherent in a field trial (Meyer et al. 1992) or growth in axenic cultures versus greenhouse conditions (Hart et al. 1992) are able to stimulate gene inactivation, but there is no quantitative comparison in strictly controlled environments. Although obtained on the basis of cell culture, a report about stimulation of gene inactivation after a week of growth at elevated temperature under otherwise constant conditions (Walter et al. 1992) indicates that single exogenous factors might exhibit drastic effects on transgene expression. On the other hand, a test of sibling plants derived from 10 different transformants under three different growth conditions (Hobbs et al. 1990) indicates that gene expression levels in plants with active as well as with inactive genes can be also independent from environmental factors.

*Type of Interaction*

The data listed in Table 1 demonstrate that there is no influence of the relative location of interacting genes in the genome. It is easy to imagine that multiple copies integrated at one site could frequently interact, and the data on homology recognition leading to intrachromosomal recombination in plants confirm this possibility (see below). Accordingly, inactivation has also been observed in correlation with *cis*-interaction inside one locus. The higher frequency of inactivation in *Arabidopsis* transformants obtained with direct gene transfer (which results often in multiple copy integration) versus *Agrobacterium* transformants (that integrates single copies more often) might be evidence for a frequent *cis*-inactivation within repeats.

Interaction in *trans* between alleles also occurs, and is suggested by those cases where interaction and suppression are found more often in homozygotic versus hemizygotic genotypes (see below). In addition, the well-documented phenomenon of paramutation is the best evidence for allelic interaction of genes in plants. Paramutation at endogenous loci has been recognized for some time (Brink et al. 1968; Hagemann 1969; Brink 1973), and a similar type of dominant silencing involving a transgene has been described recently (Meyer et al. 1993). It should be emphasized that these allelic interactions occur in somatic cells and are not limited to pairing of homologous chromosomes in meiosis.

Most surprising for our present understanding of genome organization are the numerous cases of ectopic interactions between loci. Although it is still a matter of discussion whether transgenes integrate randomly into the genome, the interactions between homologous transgenes clearly involve unlinked loci that can be separated by segregation (Matzke et al. 1989, 1993; Hobbs et al.

1993). Most transgenes and their endogenous counterparts have not been genetically mapped, but it can be assumed that the interacting loci in the cases of co-suppression are unlinked and on different chromosomes. It is not yet known when and how an interaction takes place between ectopic genes in somatic cells. The growing insight into the nuclear structure and the influence of scaffold attachment regions on neighboring transgenes may eventually reveal whether ectopic genes have a chance for spatial alignment and transfer of epigenetic information.

An interaction can be either reciprocal or nonreciprocal. Reciprocal inactivation, resulting in silencing of all genes involved, is more prevalent so far. However, it is unclear whether this mutual inactivation is really a coordinate process or whether it affects one partner initially and then spreads to the other gene(s). For cases of nonreciprocal interaction, either the residing transgene or an endogenous gene is unaffected by the silencing of the second transgene, or the endogenous gene is switched off while the transgene is still expressed (Elkind et al. 1990).

The mode by which the genes are combined in the same genome does not seem to play an important role: interaction can occur after transformation as well as after sexual crossing.

*Involvement of Methylation*

Besides changes of the nucleotide sequence, methylation is a chemical modification of DNA that is well documented and can serve as a basis for epigenetic mechanisms. Nuclear DNA in plants is frequently methylated at cytosine residues in the sequences CG or CNG, with these palindromes offering the possibility for methylation maintenance during replication. Cytosines not present in such a di- or trinucleotide may also be methylated, affecting the activity of restriction enzymes (e.g. *Hin*dIII) that are not normally considered to be methylation-sensitive (Renckens et al. 1992). Residing silent genes are often found to be hypermethylated, and this is true for silent transgenes as well. In some cases of homology-dependent gene silencing, the expression status of a transgene can be reliably correlated with the methylation of a specific recognition site within the construct. There is also some evidence that methyltransferase inhibitors like 5-azacytidine enhance the reversion to an active state. In other cases, the transgene locus can be methylated to a variable degree, but there is no unequivocal correlation with gene expression. There are also examples in which no methylation has been found in any of the interacting genes. The current method to study site-specific methylation using methylation-sensitive restriction endonucleases cannot render negative results conclusive since these enzymes only score a few of many putative methylation sites in the constructs and might miss other CG or CNG sequences of importance for gene activity. Mutant plants with a reduced methyltransferase activity and the use of transgenes with a reduced number of methylation sites offer another way to study the role of methylation in the process of gene

inactivation. Gene inactivation can also be the consequence of any structural change that renders the regulatory sequences inaccessible for the transcription machinery. Such a change from eu- to heterochromatin structure is another possibility to "label" a gene as inactive. How such a secondary structure is stably maintained, and how it can be conferred in a sequence-specific manner from one gene to another is not yet known.

*Modifying Factors*

The variable degree and frequency of gene inactivation among independent transformants of the same plant species with the same transgene indicate that there are additional factors that influence the probability of silencing. It has been observed that independent transgenic plants harboring comparable copy numbers of the same construct are not always affected by silencing (e.g. Linn et al. 1990). It is probable that the integration site of the transgene can make it more or less susceptible to silencing, possibly via neighboring sequences, extent of heterochromatinization, fixation to nuclear scaffold elements or spatial accessibility for proteins. The influence of other genes interacting in *trans* with such loci may simply enhance an intrinsic tendency towards stability or instability, respectively (Matzke et al. 1993; Neuhuber et al. 1994).

Different grades of silencing have also been observed within individual plants or isogenic populations. This has been explained by developmental dependency. Whereas young plants were shown to exhibit no silencing in the first few weeks after germination, the same plants were affected in later stages of development (F Meins personal communication; D Inzé personal communication; Neuhuber et al. 1994; O Mittelsten Scheid unpublished observation). Factors related to development could also explain differences in silencing due to the type of culture conditions (Hart et al. 1992).

Another factor modifying the degree and stability of inactivation is the number of genes, not calculated as copy number per haploid genome but rather in relation to the whole genome: a more drastic and less reversible inactivation in plants homozygous for the gene versus the hemizygote genotype is common in several systems. Although there is no evidence for a long term association of homologous chromosomes in nuclei of somatic cells (Heslop-Harrison & Bennett 1990), the doubling of the copy number at allelic positions may increase the chance for homology recognition and a subsequent exchange or transfer of the epigenetic state during any transient or partial pairing.

*Avoiding Gene Silencing in Transgenic Plants*

To obtain stable expression and inheritance of transgenes in genetically modified plant lines, the following criteria should be considered.

- Gene silencing is frequently observed upon integration of complex inserts, rearranged copies or duplications. Therefore, plants with a single insert of

the transferred gene, without duplications in the form of tandem or inverted repeats and consisting of unique elements may have a higher chance of stable expression.

- Irrespective of the exact mechanism, homology on DNA or RNA levels seems to stimulate silencing events. If homology (e.g. to residing DNA) cannot be avoided, length and degree of homology should be controlled by interrupting perfect homology with either mismatches or intron sequences (Kricker et al. 1992).
- In addition to the structure of the integrated DNA itself, the environment of the insert may influence the stability of gene expression. Integration into single-copy, unmethylated sequences of the plant genome may increase the probability for continuous stable expression.
- Based on the expectation that DNA methylation may spread along chromosomes, transgene expression possibly can be stabilized by "insulation" with sequences that minimize the expansion of methylation and/or heterochromatin. The AT-rich nature of scaffold attachment sites (SARs), which contain proportionally fewer methylatable cytosines, might exhibit such an effect, although there is no evidence supporting this speculation.
- Gene silencing occasionally becomes evident only after transmission of the gene to the next generation (Kilby et al. 1992). In some cases, transgene expression was decreased progressively over subsequent generations. Therefore, continuous monitoring of expression levels in progeny – even of well-established transgenic lines – may be a precaution against unexpected epigenetic modifications.

## Homologous recombination in somatic plant cells

The topic is covered extensively by other contributors to this volume, and our purpose here is to discuss briefly a possible relation between homologous recombination and homology-dependent gene silencing, since the initial process of homology recognition in both cases could be related. We restrict this discussion to those recombination approaches where homologous repeats were integrated in the genome (see Chapter 00), since interaction of transgenic sequences leading to silencing is assumed to occur once they are integrated in genomic DNA, although an earlier interaction with endogenous during the transformation process cannot be ruled out (Hobbs et al. 1993).

Intrachromosomal recombination is believed to be initiated by recognition of homology between the repeats, followed by single or double strand breaks, heteroduplex formation, repair synthesis and religation. Which of these subsequent steps limits the recombination frequency is not known and may vary between different systems. In plants, the frequency of somatic recombination between homologous sequences in close proximity ranges from $10^{-4}$ to $10^{-6}$ (Peterhans et al. 1990; Tovar & Lichtenstein, 1992; Assaad and Signer

1992; Swoboda et al. 1993). Although the frequency of homology-dependent gene inactivation has not been determined in relation to the number of single cells, it can be assumed to be relatively high, since silencing has been observed without selection or screening.

As mentioned previously, homology-dependent gene silencing can involve sequences in intra-allelic, inter-allelic or ectopic configurations. Whether all three of these variations also apply to homologous recombination is difficult to assess, since there is, to our knowledge, no evidence for interchromosomal homologous recombination events between plant genes. Similar frequencies for mitotic recombination between homologous sequences in ectopic and allelic positions, respectively, in yeast (Lichten & Haber 1989) make it nevertheless probable that plants have a similar capability for recognizing homologous sequences in somatic cells. In addition, chromosome translocations are cytogenetic evidence for interchromosomal recombination, although it is not known whether homologous DNA is involved in their formation.

The terms homology recognition, DNA alignment and pairing allow one to visualize the molecular structures involved in recombination, which presumably requires intimate contact between the recombining molecules, allowing the formation of heteroduplexes or three- or four stranded structures over a certain region (Stasiak 1992). Whether such a close encounter is also required for gene inactivation is a matter of speculation as long as the molecular nature of the interaction and epigenetic change has not been clarified. It is possible that the initial comparison of chromosomal regions containing homologous sequences can be achieved by recognition of similarities in chromatin structure or DNA-protein binding patterns. It is not yet known to what extent DNA modifications can interfere with recombination or inactivation. In the case of inactivated and methylated genes in *Ascobolus immersus*, it has been demonstrated that methylation does not impair a second round of pairing and inactivation of a third gene copy (Faugeron et al. 1990). In plants, the influence of DNA methylation on recombination has been studied so far only in an experiment designed to detect extrachromosomal recombination; methylation did not interfere with recombination rates (Puchta et al. 1992).

Neither homology-dependent inactivation nor recombination occur inevitably but seem to happen with a certain probability, which can be influenced by external factors (environment) as well as internal factors (flanking sequences, location of genes). Moreover, neither interaction is restricted to a specific functional part of the gene (coding or intron sequence, promoter). How often repeated sequences present in a genome sense each others' homology and interact remains an open question, but it probably occurs more often than the observed frequency of actual recombinational events. The interactions that align homologous DNA sequences but do not lead to structural alterations may nevertheless produce phenotypic variations resulting from epigenetic changes.

## Pairing-Dependent Phenomena that Affect Gene Expression in *Drosophila* and Fungi

The various cases of homology-dependent gene silencing in plants share several features with known pairing-dependent phenomena that modify gene expression in *Drosophila* and filamentous fungi. Our description of these phenomena will be followed by a discussion of their similarities to homology-dependent gene silencing in plants. There are other examples of gene silencing which are not yet known to rely on pairing. These include the silent yeast mating type loci (Haber 1992), X-chromosome inactivation in mammals (Ballabio & Willard 1992), silent variant cell surface glycoprotein genes in trypanosomes (van der Ploeg et al. 1992) and telomeric position effects in yeast (Wright & Shay 1992; Sandell & Zakian 1992). Some of these have been discussed in a recent review (Rivier & Rine 1992).

### *Gene Expression Affected by Pairing of Alleles in* Drosophila

In *Drosophila*, it is generally assumed that homologous chromosomes are paired in diploid somatic cells during interphase, placing alleles in close proximity. As has been recognized for decades, this arrangement influences the activity of a number of genes. We will discuss three general cases in which gene expression is affected as the result of pairing of alleles. (There are, to our knowledge, no recognized cases of *Drosophila* gene expression being affected as a result of the ectopic pairing of homologous genes or regions, although such pairing can and does occur (Hiraoka et al. 1993).) In accordance with the distinctions made by Henikoff and Dreesen (1989), we will consider separately transvection, the *zeste-white* interaction and dominant position effect variegation (PEV). These cases can be distinguished on the basis of whether they constitute an activation or suppression of the interacting alleles, and whether one allele is dominant to the other.

#### *Transvection*

In a recent review, Pirrotta (1990) has defined transvection as "a pairing-dependent complementation [in which] regulatory elements on one chromosome [can] control the expression of the corresponding gene on the paired homologous chromosome" (Figure 2). This complementation generally occurs between particular regulatory mutations or between regulatory and structural mutations. It can be detected as the pairing-dependent ability of an intact "distant regulatory element" of one allele to regulate the promoter of the other allele (Pirrotta 1990). First characterized by Lewis (1954), transvection has been described for several different loci in *Drosophila*, the best characterized being *white* (*w*), *Ultrabithorax* (*Ubx*), *decapentaplegic* (*dpp*) and *yellow* (*y*) (Judd 1988; Pirrotta 1990). As has been noted (Henikoff & Dreesen 1989), transvection requires pairing to activate, not inactivate, a gene, and does not involve heterochromatin. Several (but not all) loci which undergo trans-

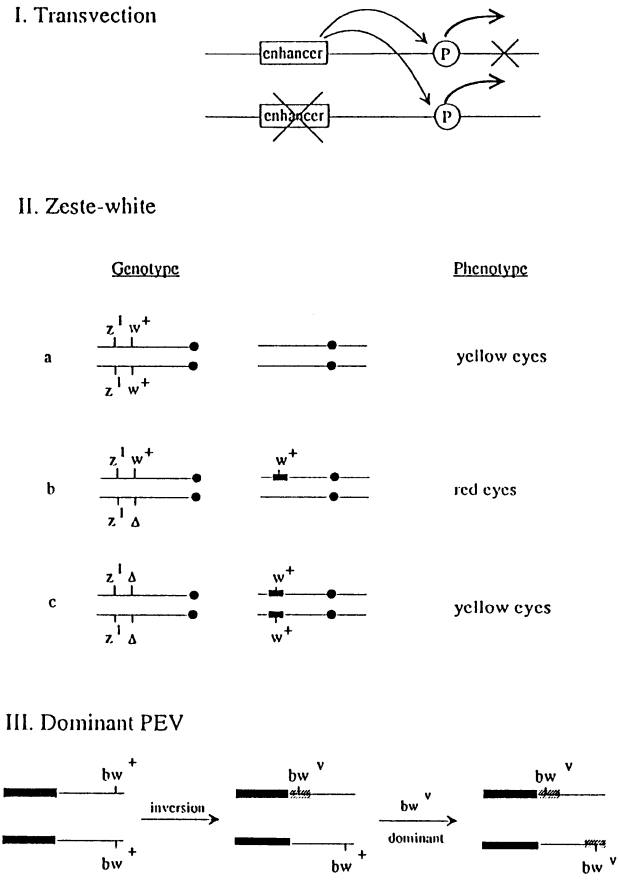


*Figure 2.* Pairing-dependent phenomena that affect gene expression in Drosophila. (*I*) *Transvection* is defined as the pairing-dependent ability of an intact upstream regulatory element (enhancer) on one allele to activate the promoter (P) of the other allele. It usually involves a structural mutation in one gene copy and a regulatory mutation in the other (X's) (redrawn from Pirrotta 1990). (*II*) The *zeste-white* interaction is the pairing-dependent repression of both *white* alleles in the presence of a mutant zeste protein. In this drawing, X-chromosomes are to the left, a pair of autosomes is to the right. Paired *white* ($w^+$) alleles on the X-chromosome are repressed in the presence of $z^1$, resulting in mutant yellow eyes. If the second copy of *white* is present on an autosome such that pairing between *white* alleles is prevented, eyes are wild-type red. Placing both copies of *white* at allelic sites on an autosome again allows pairing and both alleles are again repressed, resulting in mutant yellow eyes (redrawn from Wu & Goldberg 1989). (*III*) *Dominant PEV* occurs at the normally euchromatic *brown* locus when one brown allele ($bw^+$) undergoes a rearrangement which places it adjacent to heterochromatin (black bar). Presumably because of variable spreading of heterochromatin into the rearranged allele (hatched region), variegated expression results ($bw^v$). In this case, the $bw^v$ allele is dominant, and can induce in *trans* the heterochromatinization and variegated expression of the unrearranged wild-type allele when both are paired.

vection require the activity of the *zeste* (*z*) gene. The product of *zeste* is a DNA binding protein that can bind to multiple sites on polytene chromosomes (Biggin et al. 1988). Its role is possibly to draw two alleles together thus allowing *trans*-activation of a promoter by a regulatory region on the other gene copy (Pirrotta 1990). As an example, the dominant *Contrabithorax* (*Cbx*) mutation causes the ectopic expression of the *Ubx* gene. This occurs when *Cbx* is paired with the wild-type *Ubx* gene, and the mutated *Cbx* regulatory element activates *Ubx* in cells where it is otherwise not expressed (Pirrotta 1990).

*Zeste-White Interaction*

This is a phenomenon, first described by Gans (1953), in which female flies homozygous for the recessive X-chromosome mutation $z^1/z^1$ have yellow eyes, in contrast to $z^1/Y$ males which have wild-type red eyes. This sexual dimorphism is due to the dosage of the *white* gene, also present on the X chromosome. The $z^1$ mutation causes abnormal eye color only in homozygote *w/w* female flies, which have two X chromosomes. Only *white* alleles which are paired or in close proximity exhibit such reduced activity in the presence of certain mutant *zeste* alleles, as demonstrated by the fact that ectopic copies of *white* do not inactivate in the presence of *zeste* (Judd 1988; Wu & Goldberg 1989) (Figure 2). Even though the *zeste/white* interaction requires pairing, this does not necessarily have to be normal homologous pairing. For example, heterozygous inversions that move the *white* gene around on the X chromosome decrease but do not eliminate the *zeste* effect unless the breakpoint is very close to the *white* gene. The close association or pairing of *white* alleles necessary for the effect recalls transvection, although the *zeste-white* interaction involves a coordinate inactivation of paired *white* alleles.

*Dominant PEV*

PEV was first discovered by Muller (1930) in studies on the *white* locus (reviewed by Henikoff 1990, 1992; Spradling & Karpen 1990). In general, it can arise when a normally euchromatic gene is placed by a rearrangement next to heterochromatin. The heterochromatin apparently spreads variably into the euchromatic gene, resulting in different degrees of inactivation in individual cells. The inactivated state can be inherited clonally, producing mosaic patterns of gene expression. Most cases of PEV are recessive; in some, however, the rearranged allele is dominant and able to inactivate in *trans* the wild-type allele on the homologous chromosome. Thus, the heterochromatic state of the rearranged allele is imposed on the normally euchromatic homolog (Figure 2). The classic example of dominant PEV is the *brown* gene. (Although up to 100 variegating rearrangements of *white* are known, none are dominant (Henikoff & Dreesen 1989).) The mechanism by which this transmission of the inactive state from one allele to the other occurs is not known. Henikoff and Dreesen have proposed a model invoking pairing of the mutant and wild-type alleles, because the effect was sensitive to chromosome configurations

that disturbed pairing (Henikoff & Dreesen 1989). Furthermore, copies of *brown* at ectopic locations were not *trans*-inactivated by rearrangements affecting the endogenous gene (Dreesen et al. 1991). The dominance of brown is best explained by contact between heterochromatin proteins associated with the *cis*-inactivated copy of the *brown* gene and a positive regulator present on the *trans* wild-type copy rather than by spreading of heterochromatin along the *trans* copy (Henikoff 1992).

Although it is relatively straight-forward to imagine PEV as a result of heterochromatin infiltrating a euchromatic gene translocated to its vicinity, there are exceptions to this simple scheme and possible alternatives to the heterochromatin spreading model (Henikoff 1990, 1992; Spradling & Karpen 1990). For example, some normally heterochromatic genes variegate when placed by rearrangements next to euchromatin. There are a number of unlinked modifiers of PEV, which enhance or suppress the variegated phenotype. These often work in a reciprocal dosage-dependent manner, suppressing when present in one copy and enhancing when present in three copies, or vice versa (Tartof et al. 1989; Henikoff 1992). Some of the modifier loci seem to encode proteins involved in the formation of heterochromatin (Reuter & Spierer 1992; Henikoff 1992). Others, e.g. heterochromatin on the Y chromosome, may serve to titrate out factors needed for heterochromatin formation (Henikoff 1990, 1992; Spradling & Karpen 1990).

### *Gene Expression Affected by Pairing of Duplicated Sequences in Filamentous Fungi*

Working with the filamentous fungus *Neurospora crassa*, Selker & co-workers reported in 1987 that sequence duplications could trigger the inactivation and methylation of both copies in haploid nuclei during the sexual cycle (reviewed by Selker 1990). The methylation and inactivation were associated with extensive conversion of methylated cytosines to thymines. The acronym RIP was coined to refer to this "repeat-induced point mutation". Both closely linked and unlinked duplications can be affected by this process, although the latter are "ripped" less frequently. In *Neurospora*, methylation is irreversible, even when two unlinked gene copies are segregated to separate cells (Selker & Garrett 1988; Cambareri et al. 1989). Selker has discussed the evolutionary implications of RIP, which can be viewed as a means to accelerate the rate by which gene duplications diverge in nucleotide sequence, and so provide raw material for evolution (Selker 1990).

The action of RIP was originally confined to haploid cells in the sexual phase of filamentous fungi (whose genomes do not contain substantial amounts of repetitive DNA), and this was accompanied by C to T conversions in the time span of a single generation. However, other work has demonstrated that RIP is just one process acting on duplicated sequences. For example, Faugeron and co-workers reported a similar phenomenon in which duplicated genes were also methylated and inactivated during the sexual cycle of another filamen-

tous fungus, *Ascobolus immersus*. In *Ascobolus*, however, methylation and inactivation of duplications were reversible and apparently not associated with C to T conversions (Goyon & Faugeron 1989; Faugeron et al. 1990; Rhounim et al. 1992). Therefore, mutation is not an obligatory consequence of repeat-induced methylation. Rhounim and co-workers have suggested that the process acting in *Ascobolus* is qualitatively different from RIP in *Neurospora* and have introduced the acronym MIP (methylation-induced pre-meiotically) to distinguish the two phenomena (Rhounim et al. 1992).

The discussion on the action of RIP and MIP has limited it so far to the period of sexual reproduction in haploid fungal cells. However, another way in which the original definition of RIP in fungi could be expanded is to include the possibility that also during the asexual cycle (i.e. in diploid somatic cells) duplications can be reversibly methylated and inactivated. Two recent reports demonstrate that this apparently can occur (Pandit & Russo 1992; Roman & Macino 1992), although it is not yet clear whether these cases have any connection with RIP or MIP.

A final way to amend the original definition of RIP is to allow for its operation in higher organisms. RIP was initially believed to be advantageous for fungi, which maintain small genomes with little repetitive DNA, since it could rapidly produce DNA sequence divergence between duplications, thus limiting their ability to recombine and create deleterious rearrangements of the genome. However, it is not necessary to restrict RIP to small, non-repetitive genomes: it can also apparently act as a germ line process in large genomes with abundant repetitive DNA. This has been demonstrated in a recent analysis of CpG depletion in repeated versus unique sequences in mammals (Kricker et al. 1992).

Thus, the original conception of RIP in fungi is proving to be only a single variation of a more widespread and potentially less disruptive process. A general term describing simply repeat-induced methylation and gene silencing would encompass the main features shared by all the phenomena discussed above. Additional special modifications would include whether the process is germinal or somatic, and whether it is reversible or mutagenic. In any case, releasing the definition of "RIP" from the constraints of frequent mutation occurring sexually in haploid fungal cells allows valid comparisons to cases of *cis*- and *trans*-inactivation of repeated genes in somatic plant cells.

What is the evidence that RIP or MIP requires the pairing of duplicated sequences? A pairwise interaction is suggested by the fact that in *Neurospora*, RIP mutates either both copies of the duplicated sequence or neither copy, whereas the presence of three or more copies leads to a combination of altered and unaltered copies (Selker 1990). A study in *Ascobolus* has also supported the role of pairing of repeated DNA sequences (Faugeron et al. 1990). In this case, however, up to three gene copies could become methylated and inactivated, suggesting both that one copy could undergo successive cycles of pairing, and that pairing could occur between methylated and unmethylated copies. Pairing is also the best explanation for the observation that

methylation in *Ascobolus* is exactly coextensive with the region of homology (Barry et al. 1993). Studies with a *Neurospora* mutant defective in chromosome pairing have shown that the pairing of homologous DNA sequences is not dependent on pairing of homologous chromosomes (Foss & Selker 1991).

### *Comparisons between Pairing-Dependent Phenomena in* Drosophila *and Fungi to Cases of Homology-Dependent* Cis- *and* Trans-*Inactivation of Gene Expression in Plants*

How might the various cases of pairing-dependent phenomena described above serve as models for understanding homology-dependent *cis-* and *trans-*inactivation of allelic or ectopic copies of plant genes?

To our knowledge, transvection, defined as an interallelic complementation, has not been identified in plants. A possible exception, in which the expression of a glutamine synthetase transgene was associated with the ectopic activation of the homologous endogenous gene (Hirel et al. 1992), was mentioned in Part II. Since all cases in plant systems involve *inactivation* of the interacting genes, then the best parallels are probably provided by dominant PEV and RIP/MIP.

#### *Similarities to Dominant PEV*

Dominant PEV shares with the plant phenomena the inactivation, as well as the dominance that has been observed in at least several cases in plants. These include examples of ectopic *trans*-inactivation described in tobacco (Matzke et al. 1989, 1993; Hobbs et al. 1990, 1993) and *Arabidopsis* (O Mittelsten Scheid unpublished observation) as well as the allelic *trans-*inactivation (paramutation) of *A1* alleles in petunia (Meyer et al. 1993). As with dominant PEV, paramutation as described by Meyer and co-workers includes the imposition of an inactive (in this case, hypermethylated) state of one allele on its active, hypomethylated partner. Similarly, several dominant transgene loci are also hypermethylated (Hobbs et al. 1993; Matzke et al. 1994) and can presumably impose this state on unlinked, partially homologous, and normally hypomethylated transgene loci. Since dominant PEV has been shown to require pairing or close proximity of alleles, it is reasonable to assume that the cases from plants which seem to involve the transfer of epigenetic information, even between unlinked copies, also are pairing- or proximity-dependent. One limitation of dominant PEV as a model for the plant phenomena is that it occurs in *Drosophila* only between alleles on homologous chromosomes; ectopic copies are not affected. This is not a concern in the case of paramutation in plants, which is an allelic interaction. However, the examples of ectopic gene silencing would require that pairing as well as methylation or heterochromatin transfer occur between unlinked copies. The likelihood of this being possible is discussed below in connection with RIP/MIP.

In contrast to the examples of dominant *trans*-inactivation described above,

cases of co-suppression in plants, which involve the coordinate (simultaneous?) inactivation of a transgene and its unlinked endogenous counterpart, are more difficult to fit into the dominant PEV model. It is conceivable that the transgene becomes inactivated first by methylation or heterochromatinization and then rapidly imposes this on the endogenous gene, such that the inactivation appears coordinate. However, another problem is that both the transgene and endogenous gene are not only inactivated coordinately, but also reactivated coordinately, i.e. *both* are changed in the same way by the interaction. In contrast, the allele inducing the change in dominant PEV or paramutation (in the case of ectopic effects, the inducing locus) remains substantially unchanged by the interaction. If pairing is involved in cases of reversible co-suppression, then these might be more similar to the example of the *zeste-white* interaction, in which both *white* alleles are inactivated when paired, or RIP/MIP (see below), which also affects both copies of the duplicated gene. Since other transcription factors in *Drosophila* besides zeste might be able to promote such effects (Pirrotta 1990), it is possible that transcription factors in plants can do so as well in the absence of any homolog to the *Drosophila* zeste protein.

Assuming that dominant PEV is the best model for some cases of allelic and ectopic *trans*-inactivation in plants, what steps are required for this to occur in the plant systems studied to date? To begin with, dominant PEV in *Drosophila* requires a rearrangement to heterochromatinize and inactivate one allele. There are several possible ways to hypermethylate or heterochromatinize plant transgenes. First, a single copy of a transgene might integrate into, or close to, methylated plant DNA, and itself become more or less methylated and inactivated. As in PEV, one could imagine that methylation (or condensed chromatin) in the adjacent plant DNA could spread variably into the transgene locus in different cells. It might be possible to recover distinct "epialleles" comprising the two possible extremes of hypermethylation and hypomethylation, as has been accomplished with an *A1* transgene locus in petunia (Meyer et al. 1993). These *A1* epialleles could be maintained rather stably as homozygotes. When combined in the heterozygote, however, the hypermethylated allele could partially impose this state on its unmethylated partner (Meyer et al. 1993), a scenario similar to dominant PEV.

The second way in which transgenes could become methylated is by integrating multiple copies, either in tandem arrays or closely linked, at a single locus. Such closely linked transgene repeats can inactivate in *cis* spontaneously (Mittelsten Scheid et al. 1991; Assaad et al. 1993), apparently by *de novo* methylation (Doerfler 1991). Such a locus could then become a suppressor in *trans* for either an unmethylated allelic form of the transgene (Assaad et al. 1993), or unlinked homologous copies (Hobbs et al. 1993; O Mittelsten Scheid unpublished observation; Matzke et al. 1994).

A third means by which transgenes or endogenous genes could become inactivated by methylation or heterochromatin formation is if they are subject to parental imprinting. In a few cases, paternal transmission of transgenes

has resulted in weaker expression in seedlings (Matzke et al. 1993; M Matzke & A Matzke unpublished observation) or a stronger suppressing effect (Meyer et al. 1993). In selfed progeny, an interaction between an active, hypomethylated maternal epiallele and a weakened, hypermethylated paternal epiallele could diminish the activity of the former by imposition of the hypermethylated state. This recalls the paramutation of *A1* alleles (Meyer et al. 1993). In fact, earlier studies on parental imprinting of the *R* locus in maize demonstrated that *R* alleles that were imprinted were also subject to paramutation, leading Brink to suggest that the parental effect is an aspect of paramutation (Brink 1973).

In summary, the two features of dominance (or semi-dominance) and DNA alteration (in the case of plants, methylation has been identified so far; in *Drosophila*, which lacks significant cytosine methylation, heterochromatinization) seem to be common to at least some cases of *trans*-inactivation in plants and dominant PEV in *Drosophila*. Although ectopic copies are not affected in *Drosophila*, this may be due to the stable somatic pairing of homologous chromosomes which would tend to promote primarily allelic interactions. In plants, it does not appear that homologs undergo extensive somatic pairing (discussed in Matzke & Matzke 1993b), and this arrangement might permit ectopic interactions between homologous sequences.

*Similarities to RIP/MIP*

As discussed above, the phenomenon originally defined as RIP in *Neurospora* could probably be extended to include reversible, somatic phenomena. This is essential for drawing close comparisons with homology-dependent gene silencing in plants, because all cases occur in somatic cells, and most are reversible. Further considerations follow:

*Homology Searching.* An aspect of fungal RIP/MIP that might be particularly relevant for understanding ectopic *trans*-inactivation in plants is that duplications in fungi can be methylated and inactivated at ectopic positions, albeit at a lower frequency than duplications within a genetic locus (in *Neurospora*, practically all closely linked duplications are altered whereas only 10–50% of ectopic duplications are affected) (Foss & Selker 1991). RIP/MIP, therefore, sets a precedent for ectopic interactions, unlike the examples from *Drosophila* which are limited to alleles. This demonstrates that a homologous sequence can scan the genome for a partner, resulting in DNA modifications and inactivation. Therefore, the unexpected findings that unlinked homologous sequences could interact in plants could be explained by invoking such a genome-wide homology-searching mechanism that can operate in diploid or multiploid somatic cells. The recent results, demonstrating that methylation and inactivation of duplications occurs not only pre-meiotically but also somatically in *Neurospora* (Pandit & Russo 1992; Romano & Macino 1992), provide support for the existence of such a mechanism in somatic cells.

*Coordinate Inactivation.* In RIP/MIP, both copies of a duplicated gene are methylated and inactivated. In the case of MIP in *Ascobolus* and somatic "RIP" in *Neurospora*, this is a reversible process, whereas for pre-meiotic RIP in *Neurospora*, it is irreversible and associated with C to T conversions. The coordinate inactivation of both gene copies in RIP/MIP is similar to the *cis*-inactivation of transgene repeats and cases of co-suppression observed between transgenes and homologous endogenous genes in plants (Table 1).

*Mutagenicity Vs. Reversibility.* In cases of plant allelic or ectopic *trans*-inactivation where methylation has been observed, it is unlikely that there are extensive C to T conversions. In one report, a site for a methylation-sensitive restriction enzyme was preserved through several sexual generations despite being methylated (Matzke et al. 1993). Sequencing data from recloned silenced genes in *Arabidopsis* gave no evidence for a RIP-like process (Mittelsten Scheid et al. 1994). Therefore, although further studies are necessary, it is probable that extensive RIP does not result from repeat-induced methylation in plants. In this sense, the plant phenomena resemble MIP in *Ascobolus* or somatic "RIP" in *Neurospora.*

The less frequent alteration of ectopic duplications versus closely linked copies by RIP/MIP in fungi is possibly related to observations in plant systems. Many transgenic plants are obtained that contain unlinked gene duplications that do not exhibit reduced expression of the duplicated gene, whereas closely linked repeats seem to be more readily inactivated. In addition to homology, several other factors might influence whether two unlinked genes interact in the plant genome. First, the chromosomal locations of two potentially interacting (i.e. homologous) genes may determine whether they are able to find each other and pair. This favorable juxtapositioning could be either a transient (perhaps developmentally regulated) or permanent condition, and obviously one extreme is that the two gene copies never "see" each other at all and are unable to interact.

Second, some genes are probably more susceptible to ectopic *trans*-inactivation than others. For example, the chalcone isomerase gene in petunia (J Mol personal communication) or the peroxidase gene family in tobacco (Lagrimini et al. 1990; S Rothstein personal communication) appear to be more refractory to inactivation after the introduction of homologous transgenes. What could make a gene more or less susceptible? In the case of paramutation at the *R* locus in maize, a sensitive (paramutable) allele was found to be inherently unstable, even when hemizygous (Brink 1973). A paramutagenic allele served only to enhance this basic instability and push the paramutable allele further in the direction of reduced activity. Similarly, a susceptible transgene locus was also slightly unstable, and could occasionally become partially methylated in the absence of a suppressor locus (Matzke et al. 1993). Recent data comparing the sensitivity to *trans*-inactivation of the same transgene construct integrated at different locations in the tobacco genome has suggested

that partial methylation of the transgene (presumably imposed by the methylation status of flanking plant DNA) tends to make it unstable and sensitive to allelic and ectopic *trans*-effects (Neuhuber et al. 1994). Susceptible endogenous genes might also be inherently unstable due to partial and variable methylation patterns among individual cells.

A third factor to consider is the methylation state of a potentially interacting gene pair. If both are unmethylated, then an interaction along with the activity of a *de novo* methylase would be required for methylation and inactivation. However, if one member is more methylated than the other, then this differential state could conceivably initiate the transfer of methylation to the unmethylated partner via a maintenance methyltransferase. This could occur if pairing between homologous regions aligned methylated and unmethylated DNA strands, producing a hemimethylated intermediate. This has been shown to be the preferred substrate for a mammalian DNA methyltransferase (Gruenbaum et al. 1982). If this model is valid, then an additional factor influencing some cases of homology-dependent *trans*-inactivation in plants would be that one allele or locus must have a different state of methylation than its homologous partner. This has been shown to be true in several cases that clearly involved methylation (Meyer et al. 1993; Hobbs et al. 1993; Matzke et al. 1994). This requirement for an intrinsic epigenetic difference would be consistent with the fact that ectopic *trans*-inactivation in plants often involves transgene loci comprising multiple copies with complex arrangements (Hobbs et al. 1993; Matzke et al. 1993; Matzke et al. 1994). Because of their repetitive nature, these may have a high frequency of spontaneous de novo methylation, which could endow them with suppressing capabilities. Conversely, single copies of transgenes do not generally show suppressing effects on unlinked homologous (trans)genes (Matzke et al. 1993; Hobbs et al. 1993). Plants in which methylation has not been correlated with gene silencing may label their inactive genes by other means such as chromatin condensation, allowing the exchange of epigenetic information analogous to methylation (Jorgensen 1990).

In summary, the principle features shared between RIP/MIP and the plant phenomena seem to be that homologous genes, even in ectopic locations, can be detected, reversibly inactivated and, in some cases, methylated. Closely linked repeats in plants probably trigger their own *de novo* methylation leading to *cis*-inactivation, much like the situation in *Ascobolus*. However, most cases of *trans*-inactivation associated with methylation in plants are probably best explained by a mechanism involving a pre-existing epigenetic difference between homologous regions and a maintenance methylase activity. This is slightly different from the apparent *de novo* methylation of ectopic repeats in fungi. Although it cannot be ruled out that ectopic unmethylated repeats in plants can become methylated *de novo* there are up until now no reported cases of this phenomenon. It might, however, occur with low frequency.

## Related Cases in Transgenic Animals or Animal Cell Lines

Homology-dependent *trans*-inactivation involving transgenes has not yet been reported in transgenic mammalian systems. However, an awareness of the cases in plants, where such phenomena might be particularly easy to detect, could eventually lead animal biologists to similar findings. Two reports have demonstrated that transgenes can be methylated and inactivated in animal cell lines (Mehtali et al. 1990; Palmer et al. 1991). In one of these studies, a chimeric reporter gene under the control of the promoter of a housekeeping gene (*HMGCR*, encoding 3-hydroxy 3-methylglutaryl CoA reductase) became hypermethylated and inactive only in cell lines containing many copies (up to 260) of the transgene arranged in tandem. The expression of the endogenous copy of the *HMGCR* gene remained unaffected, however, despite the 100-fold increase in the copy number of the *HMGCR* promoter contributed by the transgene copies. Although this result seems to contradict the plant work showing that methylated tandemly arranged transgenes can *trans*-inactivate homologous genes, the results must be considered inconclusive until additional cases are examined. The possibility that methylation of transgenes in animals is an "immune" response to foreign DNA has been discussed (Bestor 1990; Doerfler 1991). This seems an unlikely explanation for all of the results obtained with plant transgenes, particularly those in which hypermethylation of a transgene was only observed after introducing an unlinked, partially homologous transgene locus (Matzke et al. 1993; Hobbs et al. 1993). Cases in which *cis*-inactivation and methylation do not spread entirely through a multicopy transgene locus (Matzke et al. 1994) also argue against a foreign DNA response.

Cumulative studies on the expression and epigenetic modification of transgenes in mice have revealed several striking similarities to transgene methylation and PEV in *Drosophila* (Allen et al. 1990; Sapienza 1991; Engler et al. 1991). These similarities include control by modifying loci, parental effects and cellular mosaicism (Engler et al. 1991). As described above, some cases of homology-dependent gene silencing in plants also share features with PEV. Since it is possible that all of these phenomena are fundamentally related, it is premature to conclude that homology-dependent gene silencing (or variations thereof) does not occur at all in animals. Recall that in plants, paramutation is perhaps the consequence of parental imprinting, which could set up an epigenetic difference between maternal and paternal alleles, leading to a weakening of the hypomethylated allele in the heterozygote. Since the action of parental imprinting is now well established in mammals, paramutation-like phenomena might also be detectable in these organisms.

Increasing attention is being paid to the existence of pairing-dependent allelic interactions that affect gene expression in animals (Wu 1993). Tsai and Silver (1991) have suggested that somatic transfer of genomic imprinting between paired homologs could explain the usual ability of a duplicated, paternally-transmitted *T-associated maternal effect* (*Tme*) locus to rescue a mutant

phenotype in mice. Laird has proposed that dominant PEV in *Drosophila* could be used as a model for Huntington's disease in humans (Laird 1990), and has suggested that the associated *trans*-inactivation of the $HD^+$ allele could persist through meiosis, thus explaining several unusual features of the inheritance of Huntington's disease (Sabl & Laird 1992). *Trans*-inactivation that is retained through meiosis has been termed "epigene conversion", which is postulated to "non-reciprocally change the functional [epigenetic] state of one allele to that of its homologous allele, without a change in nucleotide sequence" (Sabl & Laird 1992). Since this concept is strikingly similar to cases of *trans*-inactivation that involve the apparent transfer of methylation between repeated transgenes in plants (see above, Pairing-dependent phenomena that affect gene expression in *Drosophila* and fungi), "epigene conversion" might be a suitable generic term for all cases in which an epigenetic differential is equalized between allelic or ectopic homologous genes.

In examining differences between plants and mammals, and the different degrees to which they might exhibit homology-dependent gene silencing involving transgenes, one must consider the processes that plant and animal biologists use to produce transgenic organisms. Transgenic mice are obtained by injecting DNA either into paternal pronuclei of fertilized eggs or, for site-specific integration, into nuclei of embryonic stem cells. Since progeny mice resulting from these two procedures are mosaic or chimeric, respectively, a sexual generation must ensue before mice are obtained that are composed entirely of cells containing the transgene. In contrast to this process, in which cells at very early developmental stages are transformed, the procedure often used by plant biologists relies on transformation of, and regeneration from, differentiated somatic cells. Either cells of "leaf disks" are transformed using *Agrobacterium*, or foreign DNA is introduced into leaf protoplasts by techniques of direct gene transfer. Unfortunately, there is no guarantee that these cells are epigenetically or even genetically identical. In *Arabidopsis*, for example, the ploidy level of individual somatic cells can contain up to 16 times the 1C DNA amount (Heslop-Harrison 1992). The phenomenon of somaclonal variation demonstrates that a variety of different genetic changes are recovered in regenerated plants (Larkin et al. 1989). Furthermore, different epigenetic states of a given gene in individual leaf cells can also apparently be stably inherited in plants regenerated from single cells (Neuhuber et al. 1994). Therefore, the state of methylation or chromatin condensation of any given endogenous gene could follow a pattern of cellular mosaicism throughout a leaf (or tissue to be transformed), and these non-equivalent epigenetic states might determine the subsequent response of the endogenous gene to introduced homologous transgenes.

Another consideration is the different mode of development of higher plants and mammals. Foremost among these differences are the continuous growth and development of plant vegetative (and eventually reproductive) structures from meristems, and the absence of a sequestered germ line in plants. Because of these features, plants might be particularly well-suited for observing the

epigenetic changes which could arise from homology-dependent gene interactions. Perhaps a directed search for the possibility of such interactions might reveal similar processes in transgenic animals during organ development when cells are still dividing.

### *Effects of Heterochromatic Regions on Euchromatic Genes in* Drosophila

We have argued that homology-dependent *trans*-inactivation in plants requires a homology-searching mechanism which operates in diploid somatic cells, and that it often involves an intrinsic epigenetic difference between two potentially interacting alleles or homologous loci, so that epigenetic information (methylation or heterochromatin structure) on one allele (or locus) can be imposed on the other. Given the latter putative requirement, it might be worthwhile to mention cases of interactions between heterochromatic regions and euchromatic genes (for reviews, see Pimpinelli et al. 1986; Hilliker & Sharp 1988; Pardue & Hennig 1990; Gatti & Pimpinelli 1992). Although the highly repetitive, non-coding DNA sequences that generally make up constitutive heterochromatin are considered superfluous and biologically uninteresting by some (Miklos & Cotsell 1990), there are several known cases from *Drosophila* which argue against the genetic inertness of at least some regions of heterochromatin. Specifically, several genes located in euchromatin are affected in a defined manner by discrete regions of heterochromatin. These euchromatic genes include *Segregation Distorter* (*Sd*) and the *abnormal oocyte* (*abo*) gene. The exact nature of the heterochromatic effects are still under investigation in both cases.

Many of these effects of heterochromatic regions on genes in euchromatin are quantitative (Pardue & Hennig 1990). The sliding scale of phenotypes and degree of methylation observed in some cases of homology-dependent gene inactivation in plants (Matzke & Matzke 1991) demonstrates that quantitative effects are also encountered in these systems. Moreover, transgene inserts with suppressing effects often comprise complex arrays of multiple transgene copies which might spontaneously methylate or form heterochromatin.

## Test of the DNA-DNA Pairing Hypothesis

The pairing of interacting or non-interacting transgene alleles (or homologous loci) can be studied by determining their respective locations in interphase nuclei by in situ hybridization. If paramutating alleles at a particular transgene locus are paired, then one hybridization signal should be observed in interphase nuclei of both homozygotes and hemizygotes, although it would be stronger in the former (Hiraoka et al. 1993). If two separate and distinct signals are observed in the homozygote but only one in the hemizygote, then pairing would be unlikely. Since pairing could be transient, different devel-

opmental stages or tissue types should be examined. Does the technology exist to do such studies? Recently, a transgene suppressor locus (containing 4–5 closely linked copies, each ca. 10 kb, of a transgene complex) has been localized in interphase nuclei as well as mitotic tobacco cells (Matzke et al. 1994). Therefore, if the transgene inserts are large enough (probably 20 kb is the lower limit (Ambros et al. 1986)), this approach can be used to study somatic pairing of alleles or unlinked loci during interphase.

In *Drosophila*, requirements for pairing have been determined by studying genotypes in which pairing is disrupted or prevented by a translocation or insertion of a gene into ectopic sites in the genome. Similar experiments could also be done with transgene constructs in plants. Transgene constructions that disrupted pairing with a homologous endogenous gene or a second transgene could be made by either adding regions of non-homologous DNA to one construct, or by making partial deletions in homologous regions. The Cre-*lox* system for directed excision of transgenic sequences from the plant genome (Bayley et al. 1992; Russell et al. 1992) could also be exploited to remove or alter selectively homologous regions that might be involved in pairing.

A third way to test the requirement for DNA-DNA pairing is to couple a known pairing-dependent process, such as transvection, to homology-dependent *trans*-inactivation. Different vectors could be constructed such that complementation of a promoter-less gene on one transgene complex is dependent on *trans*-activation by an enhancer on a second unlinked transgene complex. If genes containing regions of homology were also present on these two vectors, then it might be possible to obtain *trans*-activation with one pair of genes and have this coupled to *trans*-inactivation of the adjacent homologous gene pair.

## Developmental and Evolutionary Implications

Homology-dependent gene silencing is often regarded in a negative sense, as something undesirable and to be avoided, especially in transgene applications. It is possible, however, that it plays an essential role in both higher plants and animals by gradually restricting the activity of endogenous genes during differentiation. An excellent case has been made for the progressive repression of genes in the control of development (Caplan & Ordahl 1978). In differentiated plant and animal cells, a large proportion of the genome is not expressed (Kamalay & Goldberg 1984; Kolata 1985). The embryonic lethal phenotype of mice homozygous for a targeted disruption of the DNA methyltransferase gene (Li et al. 1992) confirms that methylation is essential for normal development in these organisms. The totipotency of many plant cells demonstrates that any process restricting the activity of genes must also be potentially reversible, a criterion fulfilled by methylation. Might homology-dependent gene silencing accompanied by methylation and/or chromatin

condensation be a means to passively, but reversibly, inactivate large portions of the genome whose inappropriate expression would be deleterious? Most repetitive sequence families in eukaryotic genomes are methylated, often to varying extents at different stages of development (Chapman et al. 1984; Ponzetto-Zimmerman & Wolgemuth 1984; Sanford et al. 1984; Howlett & Reik 1991). These hypermethylated repetitive DNA sequences might suppress structural genes that are adjacent to an interspersed member of the repeat family. Interactions between repeat family members could be developmentally regulated in a manner that depends on the controlled movement of chromatin (Borden & Manuelidis 1988; De Boni & Mintz 1986), such that homologous sequences only encounter each other at specific stages.

Studies with transgenes show only what is possible in the plant nucleus, and not how frequently these events actually occur with endogenous plant genes. However, one can imagine that during evolution, situations similar to those generated artificially by the random insertion of transgenes have been encountered by endogenous genes as well. For example, translocations could place duplicated sequences in different arrangements relative to each other, with new possibilities for homology-dependent *trans*-interactions. Duplications or amplifications of a gene in *cis* may produce compound loci that can contain both an expressed copy of the gene and silenced copies that can suppress other homologous loci or alleles. The amplification of satellite sequences and translocations involving heterochromatic regions could increase opportunities for homology-dependent *trans*-interactions. Homologous transposable elements moving around a genome could also create new possibilities for such interactions. Transposable elements in fact may be evolutionary links, their frequent inactivation by methylation being a remnant of a "defense" reaction against an accumulation of foreign DNA (Bestor 1990; Doerfler 1991), combined with a spontaneous reactivation and contribution to mutation rates and genetic variability.

In considering the generality of homology-dependent gene silencing, one might legitimately ask how essential genes can be expressed reliably in a nuclear environment that is rife with multigene families, satellite sequences and other dispersed repeats. The spatial arrangement of repeated sequences and their relationship to each other could be variable, so that homologous sequences only interact at specific periods of development (Swoboda et al. 1993). Repeated genes could also be distributed between more stable and less stable domains of the nucleus, the latter becoming particularly useful when an immobile plant encounters environmental stress. In their studies on environmentally induced genotype changes in flax, Cullis and co-workers have shown that the 5S ribosomal RNA gene family comprises two subsets, one of which readily undergoes rearrangements under stress while the other is considerably more stable (Cullis 1986; Schneeberger & Cullis 1991).

Finally, epigenetic variability induced by homology-dependent gene interactions in plants might have survival value if it permits a plant to vary the genes expressed in different cells of meristems (Matzke & Matzke 1990) or

in different meristems of the same plant. The inheritance of epigenetic variants occurs quite frequently in plants and this might play a more important role in evolution than is currently appreciated (Jablonka et al. 1992).

## Acknowledgements

We want to thank our colleagues V. Chandler, D. Inzé, R. Jorgensen, F. Meins, P. Meyer, J.-L. Rossignol, E. Signer, and H. Vaucheret for preprints and communicating unpublished results. We also want to thank R. Jorgensen and V. Pirrotta for helpful comments on the manuscript. M. Matzke and A. Matzke acknowledge the financial support of the Austrian Fonds zur Förderung der wissenschaftlichen Forschung (grant no. S6006-BIO) and the Bundesministerium für Wissenschaft und Forschung. O. Mittelsten Scheid thanks the Swiss National Science Foundation for financial support by grant 3100-030118/90.

## References

Allen ND, Norris ML & Surani MA (1990) Epigenetic control of transgene expression and imprinting by genotype-specific modifiers. Cell 61: 853–861.

Ambros PF, Matzke AJM & Matzke MA (1986) Localization of *Agrobacterium rhizogenes* T-DNA in plant chromosomes by *in situ* hybridization. EMBO J 5: 2073–2077.

Assaad FF & Signer ER (1992) Somatic and germinal recombination of a direct repeat in *Arabidopsis*. Genetics 132: 553–566.

Assaad FF, Tucker KL & Signer ER (1993) Epigenetic repeat-induced gene silencing (RIGS) in *Arabidopsis*. Plant Mol Biol, in press.

Ballabio A & Willard HF (1992) Mammalian X-chromosome inactivation and *XIST* gene. Curr Op Genet Devel 2: 439–447.

Barry C, Faugeron G & Rossignol JL (1993) Methylation induced premeiotically in *Ascobolus*: coextension with DNA repeat lengths and effect on transcript elongation. Proc Natl Acad Sci USA 90: 4557–4561.

Bayley C, Morgan M, Dale EC & Ow DW (1992) Exchange of gene activity in transgenic plants catalyzed by the Cre-*lox* site-specific recombination system. Plant Mol Biol 18: 353–361.

Bestor TH (1990) DNA methylation: evolution of a bacterial immune function into a regulator of gene expression and genome structure in higher eukaryotes. Phil Trans R Soc (Lond) B 326: 179–187.

Biggin MD, Bickel S, Benson M, Pirrotta V & Tjian R (1988) *Zeste* encodes a sequence-specific transcription factor that activates the *Ultrabithorax* promoter *in vitro*. Cell 53: 713–722.

Bochardt A, Hodal L, Palmgren G, Mattsson O & Okkels FT (1992) DNA methylation is involved in maintenance of an unusual expression pattern of an introduced gene. Plant Physiol 99: 409–414.

Borden J & Manuelidis L (1988) Movement and the X chromosome in epilepsy. Science 242: 1687–1691.

Brink RA (1973) Paramutation. Annu Rev Genet 7: 129–152.

Brink RA, Styles ED & Axtell JD (1968) Paramutation: directed genetic change. Science 159: 161–170.

Cambareri EB, Jensen BC, Schabtach E & Selker EU (1989) Repeat-induced G-C to A-T mutations in *Neurospora*. Science 244: 1571–1575.

Caplan AI & Ordahl CP (1978) Irreversible gene repression model for control of development. Science 201: 120–130.

Chapman V, Forrester L, Sanford J, Hastie N & Rossant J (1984) Cell lineage-specific undermethylation of mouse repetitive DNA. Nature 307: 284–286.

Cullis CA (1986) Phenotypic consequences of environmentally induced changes in plant DNA. Trends Genet 2: 307–309.

De Boni U & Mintz AH (1986) Curvilinear, three-dimensional motion of chromatin domains and nucleoli in neuronal interphase nuclei. Science 234: 863–866.

de Carvalho F, Gheysen G, Kushnir S, Van Montagu M, Inzé D & Castresana C (1992) Suppression of β-1,3-glucanase transgene expression in homozygous plants. EMBO J 11: 2595–2602.

Deroles SC & Gardner RC (1988) Expression and inheritance of kanamycin resistance in a large number of transgenic petunias generated by *Agrobacterium*-mediated transformation. Plant Mol Biol 11: 355–364.

Doerfler W (1991) Patterns of DNA methylation – evolutionary vestiges for foreign DNA inactivation as a host defense mechanism. Biol Chem Hoppe-Seyler 372: 557–564.

Dreesen TD, Henikoff S & Loughney K (1991) A pairing-sensitive element that mediates *trans*-inactivation is associated with the *Drosophila brown* gene. Genes Develop 5: 331–340.

Elkind Y, Edwards R, Mavandad M, Hedrick SA, Ribak O, Dixon RA & Lamb CJ (1990) Abnormal plant development and down-regulation of phenylpropanoid biosynthesis in transgenic tobacco containing a heterologous phenylalanine ammonia-lyase gene. Proc Natl Acad Sci USA 87: 9057–9061.

Engler P, Haasch D, Pinkert CA, Doglio L, Glymour M, Brinster R & Storb U (1991) A strain-specific modifier on mouse chromosome 4 controls the methylation of independent transgene loci. Cell 65: 939–947.

Faugeron G, Rhounim L & Rossignol J-L (1990) How does the cell count the number of ectopic copies of a gene in the premeiotic inactivation process acting in *Ascobolus immersus*? Genetics 124: 585–591.

Foss HM & Selker EU (1991) Efficient DNA pairing in a *Neurospora* mutant defective in chromosome pairing. Mol Gen Genet 231: 49–52.

Gans M (1953) Etude génétique et physiologique du mutant *z* de *Drosophila melanogaster*. Bull Biol France Belg (Suppl) 38U: 1–19.

Gatti M & Pimpinelli S (1992) Functional elements in *Drosophila melanogaster* heterochromatin. Annu Rev Genet 26: 239–275.

Goring DR, Thomson L & Rothstein SJ (1991) Transformation of a partial nopaline synthase gene into tobacco suppresses the expression of a resident wild-type gene. Proc Natl Acad Sci USA 88: 1770–1774.

Gottlob-McHugh SG, Sangwan RS, Blakeley SD, Vanlerberghe GC, Ko K, Turoin DH, Plaxton WC, Miki BL & Dennis DT (1992) Normal growth of transgenic tobacco plants in the absence of cytosolic pyruvate kinase. Plant Physiol 100: 820–825.

Goyon C & Faugeron G (1989) Targeted transformation of *Ascobolus immersus* and *de novo* methylation of the resulting duplicated DNA sequences. Mol Cell Biol 9: 2818–2827.

Gruenbaum Y, Cedar H & Razin A (1982) Substrate and sequence specificity of a eukaryotic DNA methylase. Nature 295: 620–622.

Haber JE (1992) Mating-type gene switching in *Saccharomyces cerevisiae*. Trends Genet 8: 446–452.

Hagemann R (1969) Somatic conversion (paramutation) at the sulfurea locus of *Lycopersicon esculentum* Mill. III Studies with trisomics. Can J Genet Cytol 11: 346–358.

Hart CM, Fischer B, Neuhaus JM & Meins F (1992) Regulated inactivation of homologous gene expression in transgenic *Nicotiana sylvestris* plants containing a defense-related tobacco chitinase gene. Mol Gen Genet 235: 179–188.

Heberle-Bors E, Charvat B, Thompson D, Schernthaner JP, Barta A, Matzke AJM & Matzke MA (1988) Genetic analysis of T-DNA insertions into the tobacco genome. Plant Cell Rep 7: 571–574.

Henikoff S (1990) Position-effect variegation after 60 years. Trends Genet 6: 422–426.

Henikoff S (1992) Position effect and related phenomena. Curr Op Genet Devel 2: 907–912.

Henikoff S & Dreesen TD (1989) Trans-inactivation of the *Drosophila brown* gene: evidence for transcriptional repression and somatic pairing dependence. Proc Natl Acad Sci USA 86: 6704–6708.

Heslop-Harrison JS (1992) Nuclear architecture in plants. Curr Op Genet Devel 2: 913–917.

Heslop-Harrison JS & Bennett MD (1990) Nuclear architecture in plants. Trends Genet 6: 401–405.

Hilliker AJ & Sharp CB (1988) New perspectives on the genetics and molecular biology of constitutive heterochromatin. In: Gustafson JP & Appels R (eds) Proceedings of the 18th Stadler Genetics Symposium: Chromosome Structure and Function-Impact of New Concepts (pp 91–116).

Hiraoka Y, Dernburg AF, Parmelee SJ, Rykowski MC, Agard DA & Sedat JW (1993) The onset of homologous chromosome pairing during *Drosophila melanogaster* embryogenesis. J Cell Biol 120: 591–600.

Hirel B, Marsolier MC, Hoarau A, Hoarau J, Brangeon J, Schafer R & Verma DPS (1992) Forcing expression of a soybean root glutamine synthetase gene in tobacco leaves induces a native gene encoding cytosolic enzyme. Plant Mol Biol 20: 207–218.

Hobbs SLA, Kpodar P & DeLong CMO (1990) The effect of T-DNA copy number, position and methylation on reporter gene expression in tobacco transformants. Plant Mol Biol 15: 851–864.

Hobbs SLA, Warkentin TD & DeLong CMO (1993) Transgene copy number can be positively or negatively associated with transgene expression. Plant Mol Biol 21: 17–26.

Howlett SK & Reik W (1991) Methylation levels of maternal and paternal genomes during preimplantation development. Development 113: 119–127.

Jablonka E, Lachmann M & Lamb MJ (1992) Evidence, mechanisms and models for the inheritance of acquired characters. J Theor Biol 158: 245–268.

Jones JDG, Gilbert DE, Grady KL & Jorgensen RA (1987) T-DNA structure and gene expression in petunia plants transformed by *Agrobacterium tumefaciens* C58 derivatives. Mol Gen Genet 207: 478–485.

Jorgensen RA (1990) Altered gene expression in plants due to trans interactions between homologous genes. Trends Biotechnol 8: 340–344.

Jorgensen RA (1992a) Elicitation of organized pigmentation patterns by a chalcone synthase transgene. In: Amasino R (ed) Cellular Communications in Plants (pp 87–92). Plenum Press, New York.

Jorgensen RA (1992b) Silencing of plant genes by homologous transgenes. AgBiotech News and Info 4: 265N–273N.

Jorgensen RA (1993) The germinal inheritance of epigenetic information in plants. Phil Trans R Soc Lond B 339: 173–181.

Judd BH (1988) Transvection: allelic cross talk. Cell 53: 841–843.

Kamalay JC & Goldberg RB (1984) Organ-specific nuclear RNAs in tobacco. Proc Natl Acad Sci USA 81: 2801–2805.

Kilby NJ, Leyer HMO & Furner IJ (1992) Promoter methylation and progressive transgene inactivation in *Arabidopsis*. Plant Mol Biol 20: 103–112.

Klein TM, Kornstein L & Fromm ME (1990) Genetic transformation of maize cells by particle bombardment and the influence of methylation on foreign gene expression. In: Gustafson JP (ed) Gene Manipulation in Plant Improvement II (pp 265–288). Plenum Press, New York.

Kolata G (1985) Fitting methylation into development. Science 228: 1183–1184.
Kricker MC, Drake J & Radman M (1992) Duplication-targeted DNA methylation and mutagenesis in the evolution of eukaryotic chromosomes. Proc Natl Acad Sci USA 89: 1075–1079.
Lagrimini LM, Bradford S & Rothstein S (1990) Peroxidase-induced wilting in transgenic tobacco plants. Plant Cell 2: 7–18.
Laird CD (1990) Proposed genetic basis of Huntington's disease. Trends Genet 6: 242–247.
Larkin PJ, Banks PM, Bhati R, Brettell RIS, Davies PA, Scowcroft WR, Spindler LH & Tanner GJ (1989) From somatic variation to variant plants: mechanisms and applications. Genome 31: 705–711.
Leemans J (1993) Ti to tomato, tomato to market. Bio/Technol 11: S22–S26.
Lewis EB (1954) The theory and application of a new method of detecting chromosome rearrangements in *Drosophila melanogaster*. Am Nat 88: 225–239.
Li E, Bestor TH & Jaenisch R (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell 69: 915–926.
Lichten M & Haber JE (1989) Position effects in ectopic and allelic mitotic recombination in *Saccharomyces cerevisiae*. Genetics 123: 261–268.
Linn F, Heidmann I, Saedler H & Meyer P (1990) Epigenetic changes in the expression of the maize A1 gene in *Petunia hybrida*: role of numbers of integrated gene copies and state of methylation. Mol Gen Genet 222: 329–336.
Matassi G, Melis R, Kuo KC, Macaya G, Gehrke CW & Bernardi G (1992) Large scale methylation patterns in the nuclear genomes of plants. Gene 122: 239–245.
Matzke AJM & Matzke MA (1993a) Epigenetic variation as a consequence of homology dependent gene interactions in transgenic plants. Seminars in Developmental Biology 4: 83–89.
Matzke AJM, Neuhuber F, Park Y-D, Ambros PF & Matzke MA (1994) Homology-dependent gene silencing in transgenic plants: epistatic silencing loci contain multiple copies of methylated transgenes. Mol. Gen. Genet., in press.
Matzke MA & Matzke AJM (1990) Gene interactions and epigenetic variation in transgenic plants. Develop 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.
Matzke M & Matzke AJM (1993b) Genomic imprinting in plants: parental effects and *trans*-inactivation phenomena. Annu Rev Plant Physiol Plant Mol Biol 44: 53–76.
Matzke MA, Primig M, Trnovsky J & Matzke AJM (1989) Reversible methylation and inactivation of marker genes in sequentially transformed tobacco plants. EMBO J 8: 643–649.
Matzke MA, Neuhuber F & Matzke AJM (1993) A variety of epistatic interactions can occur between partially homologous transgene loci brought together by sexual crossing. Mol Gen Genet 236: 379–386.
Mehtali M, LeMeur M & Lathe R (1990) The methylation-free status of a housekeeping transgene is lost at high copy number. Gene 91: 179–184.
Meyer P, Linn F, Heidmann I, Meyer ZAH, 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.
Meyer P, Heidmann I & Niedenhof I (1993) Differences in DNA methylation are associated with a paramutation phenomenon in transgenic petunia. Plant J, in press.
Miklos GLG & Cotsell JN (1990) Chromosome structure at interfaces between major chromatin types: *alpha*- and *beta*- heterochromatin. BioEssays 12: 1–6.
Mittelsten Scheid O, Paszkowski J & Potrykus I (1991) Reversible inactivation of a transgene in *Arabidopsis thaliana*. Mol Gen Genet 228: 104–112.
Mittelsten Scheid O, Afsar K & Paszkowski J (1994) Gene inactivation in *Arabidopsis thaliana* is not accompanied by an accumulation of repeat-induced point mutations. Mol Gen Genet, in press.
Mol J, van Blokland R & Kooter J (1991) More about co-suppression. Trends Biotechnol 9: 124.

Muller HJ (1930) Types of visible variations induced by X-rays in *Drosophila*. J Genet 22: 299–334.

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.

Neuhaus JM, Ahl-Goy P, Hinz U, Flores S & Meins F (1991) High-level expression of a tobacco chitinase gene in *Nicotiana sylvestris*. Susceptibiltiy of transgenic plants to *Cercospora nicotianae* infection. Plant Mol Biol 16: 141–151.

Neuhuber F, Park Y-D, Matzke AJM & Matzke MA (1994) Susceptibility of transgene loci to homology-dependent gene silencing. Mol. Gen. Genet., in press.

Palmer TD, Rosman GJ, Osborne WRA & Miller AD (1991) Genetically modified skin fibroblasts persist long after transplantation but gradually inactivate introduced genes. Proc Natl Acad Sci USA 88: 1330–1334.

Pandit NN & Russo VEA (1992) Reversible inactivation of a foreign gene, *hph*, during the asexual cycle in *Neurospora crassa* transformants. Mol Gen Genet 234: 412–422.

Pardue ML & Hennig W (1990) Heterochromatin: junk or collectors item? Chromosoma 100: 3–7.

Peterhans A, Schlüpmann H, Basse C & Paszkowski J (1990) Intrachromosomal recombination in plants. EMBO J 9: 3437–3445.

Pimpinelli S, Bonaccorsi S, Gatti M & Sandler L (1986) The peculiar genetic organization of *Drosophila* heterochromatin. Trends Genet 2: 17–20.

Pirrotta V (1990) Transvection and long-distance regulation. BioEssays 12: 409–414.

Ponzetto-Zimmerman C & Wolgemuth DJ (1984) Methylation of satellite sequences in mouse spermatogenic and somatic DNAs. Nucl Acids Res 12: 2807–2822.

Potrykus I, Paszkowski J, Saul MW, Petruska I & Shillito RD (1985) Molecular and general genetics of a hybrid foreign gene introduced into tobacco by direct gene transfer. Mol Gen Genet 199: 169–177.

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.

Renckens S, De Greve H, Van Montagu M & Hernalsteens J-P (1992) *Petunia* plants escape from negative selection against a transgene by silencing the foreign DNA via methylation. Mol Gen Genet 233: 53–64.

Reuter G & Spierer P (1992) Position effect variegation and chromatin proteins. BioEssays 14: 605–612.

Rhounim L, Rossignol J-L & Faugeron G (1992) Epimutation of repeated genes in *Ascobolus immersus*. EMBO J 11: 4451–4457.

Rivier DH & Rine J (1992) Silencing: the establishment and inheritance of stable, repressed transcription states. Curr Op Genet Devel 2: 286–292.

Romano N & Macino G (1992) Quelling: transient inactivation of gene expression in *Neurospora crassa* by transformation with homologous sequences. Molec Microbiol 6: 3343–3353.

Russell SH, Hoopes JL & Odell JT (1992) Directed excision of a transgene from the plant genome. Mol Gen Genet 234: 49–59.

Sabl JF & Laird CD (1992) Epigene conversion: a proposal with implications for gene mapping in humans. Am J Hum Genet 50: 1171–1177.

Sandell LL & Zakian VA (1992) Telomeric position effect in yeast. Trends Cell Biol 2: 10–14.

Sanford J, Forrester L & Chapman V (1984) Methylation patterns of repetitive DNA sequences in germ cells of *Mus musculus*. Nucl Acids Res 12: 2823–2836.

Sapienza C (1991) Genome imprinting and carcinogenesis. Biochim Biophys Acta 1072: 51–61.

Schneeberger RG & Cullis CA (1991) Specific DNA alterations associated with the environmental induction of heritable changes in flax. Genetics 128: 619–630.

Selker EU (1990) Premeiotic instability of repeated sequences in *Neurospora crassa*. Annu Rev Genet 24: 579–613.

Selker EU & Garrett PW (1988) DNA sequence duplications trigger gene inactivation in *Neurospora crassa*. Proc Natl Acad Sci USA 85: 6870–6874.

Silva AJ, Ward K & White R (1993) Mosaic methylation in clonal tissue. Dev Biol 156: 391–398.
Smith CJS, Watson CF, Bird CR, Ray J, Schuch W & Grierson D (1990) Expression of a truncated tomato polygalacturonase gene inhibits expression of the endogenous gene in transgenic plants. Mol Gen Genet 224: 477–481.
Spradling AC & Karpen GH (1990) Sixty years of mystery. Genetics 126: 779–784.
Stasiak A (1992) Three-stranded DNA structure; is this the secret of DNA homologous recognition? Molec Microbiol 6: 3267–3276.
Surani MA (1991) Genomic imprinting: developmental significance and molecular mechanisms. Curr Op Genet Develop 1: 241–246.
Swoboda P, Hohn B & Gal S (1993) Somatic homologous recombination *in planta*: The recombination frequency is dependent on the allelic state of recombining sequences and may be influenced by genomic position effects. Mol Gen Genet 237: 33–40.
Tartof KD, Bishop C, Jones M, Hobbs CA & Locke J (1989) Towards an understanding of position effect variegation. Develop Genet 10: 162–176.
Taylor LP & Jorgensen RA (1992) Conditional male fertility in chalcone synthase-deficient *Petunia*. J Heredity 83: 11–17.
Tovar J & Lichtenstein C (1992) Somatic and meiotic chromosomal recombination between inverted duplications in transgenic tobacco plants. Plant Cell 4: 319–332.
Tsai J-Y & Silver LM (1991) Escape from genomic imprinting at the mouse *T-associated maternal* (*Tme*) locus. Genetics 129: 1159–1166.
van der Krol AR, Mur LA, Beld M, Mol JNM & Stuitje AR (1990) Flavonoid genes in *Petunia*: addition of a limited number of gene copies may lead to a suppression of gene expression. Plant Cell 2: 291–299.
van der Ploeg LHT, Gottesdiender K & Lee MG-S (1992) Antigenic variation in African trypanosomes. Trends Genet 8: 452–457.
Vaucheret H (1993) Identificaton of a general silencer for 19S and 35S promoters in a transgenic tobacco plant: 90 bp of homology in the promoter sequence are sufficient for trans-inactivation. C R Acad Sci Paris 316: 1471–1483.
Walter C, Broer I, Hillemann D & Pühler A (1992) High frequency, heat treatment-induced inactivation of the phosphinotricin resistance gene in transgenic single cell suspension cultures of *Medicago sativa*. Mol Gen Genet 235: 189–196.
Wright WE & Shay JW (1992) Telomere positional effects and the regulation of cellular senescence. Trends Genet 8: 193–197.
Wu C-T (1993) Transvection, nuclear structure, and chromatin proteins. J Cell Biol 120: 587–590.
Wu C-T & Goldberg M (1989) The *Drosophila zeste* gene and transvection. Trends Genet 5: 189–194.