View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

Cell, Vol. 97, 157–160, April 16, 1999, Copyright ©1999 by Cell Press

Gene Silencing: Repeats that Count

Minireview

Eric U. Selker
University of Oregon
Institute of Molecular Biology
Eugene, Oregon 97403-1229

Repeated DNA and cytosine methylation are common features of most eukaryotic genomes and are frequently found associated with one another. Much repeated DNA, such as the highly repeated satellite sequences, is of dubious value, and some, such as transposable elements, may be purely parasitic. Methylation of such sequences may reflect the action of a genome defense system that can silence repetitive DNA. Indeed, silencing mechanisms directed specifically at repeated DNA, and involving methylation, have been identified in fungi. Furthermore, DNA methylation is frequently associated with silencing of multicopy sequences in plants and animals. The correlation between repetitive DNA and methylation is far from perfect, however, indicating that repeats, per se, do not always trigger methylation and silencing. Recent findings in plants, including those reported by Bender and colleagues in the April issue of *Molecular Cell* (Luff et al., 1999), suggest that certain repeats count more than others: inverted repeats seem to be especially potent *trans*-acting inducers of silencing and methylation. Interestingly, inverted repeats have also been implicated in methylation-independent silencing in *Drosophila* (Dorer and Henikoff, 1994). Inverted repeats are common products of natural chromosomal rearrangements and are frequently found in the complex arrays of transgenes that typically result from DNA-mediated transformations of plants, animals, and fungi. Thus, inverted repeats might be responsible for many cases of silencing. How might inverted repeats cause silencing? Mechanisms based on DNA:DNA, RNA:RNA, and RNA:DNA interactions are all in the running. It is worth considering the precedents for these possibilities.

Silencing Driven by Repeats

About a dozen years ago, we learned that repeats can cause gene silencing, at least in some fungi (see Selker, 1997 and references therein). It was noticed that transforming DNA is subject to mutation and DNA methylation in the sexual phase of the *Neurospora crassa* life cycle when more than one copy is introduced or when the transforming DNA is homologous to an endogenous sequence. Any sequence above a minimum length of about 500 bp is susceptible to inactivation when duplicated, regardless of its transcriptional status or arrangement in the genome. Inactivation is normally irreversible because it involves multiple G:C to A:T transition mutations. Several additional observations support the idea that this silencing process, named RIP (repeat-induced point mutation), results from direct DNA:DNA interactions. First, RIP occurs in a pairwise manner (RIP never inactivates just one copy of a duplication but can inactivate just two copies of a triplication). Second, closely linked duplications (direct or inverted) are discovered by RIP more readily than unlinked duplications. Third, RIP is nucleus limited. We know this because the specialized cells that carry out RIP are heterokaryotic, with nuclei from both parents. When a sequence is duplicated in one nucleus but not the other, the lone copy is never silenced.

The mechanism of RIP is unknown but probably involves DNA:DNA pairing (Figure 1) followed by methylation and deamination of some cytosines in the paired region. Products of RIP are typically found methylated, even when they segregate away from all homologous sequences. In some cases, the observed methylation simply results from "maintenance methylation" (defined as methylation triggered by preexisting methylation), but in many cases the sequences mutated by RIP serve as triggers for de novo methylation.

A close, but interestingly different, relative of RIP, named MIP (methylation-induced premeiotically), occurs in the fungus *Ascobolus immersus*. MIP and RIP are virtually identical except that silencing by MIP does not involve mutations. Methylation resulting from both RIP and MIP causes transcriptional arrest. An *Ascobolus* gene with features expected of a DNA methyltransferase gene is required for efficient MIP but not for maintenance of methylation in vegetative cells (Malagnac et al., 1997). Methylation resulting from MIP, like that associated with genes inactivated by RIP and that associated with gene inactivation in plants, is not limited to symmetrical sites such as CpGs and CpNpGs, which are thought to be required for maintenance methylation. As discussed below, there are clues from plants that methylation at nonsymmetrical sites may depend on methylation of nearby symmetrical sites.

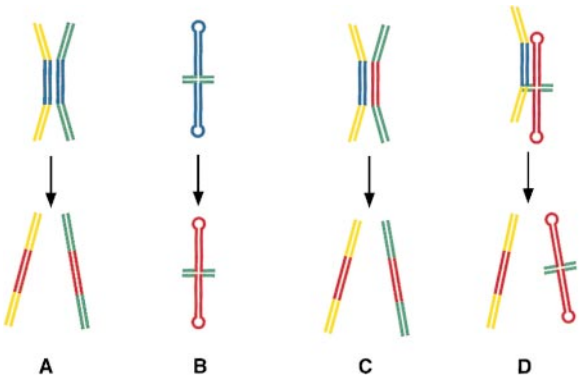

Figure 1. DNA Self-Control

Gene silencing can result from repeat-induced DNA methylation.
(A) In certain circumstances (e.g., in the premeiotic cells of *Neurospora* and *Ascobolus*), pairing of unlinked (or linked) unmethylated homologous sequences (blue segments) triggers de novo methylation (red segments).
(B) In other situations (e.g., in plants), pairing of unlinked repeats may not trigger methylation, but inverted repeats (symbolized by cruciform structure) may be still subject to de novo methylation.
(C and D) Once methylated, dispersed repeats (C) or inverted repeats (D) may induce methylation of homologous sequences even under conditions that for some reason are not conducive to de novo methylation.

Repeats along for the Ride

To date, RIP and MIP are the only silencing mechanisms that are known to "count" repeats. They are not the only mechanisms that silence repeated genes, however. Even *Neurospora* has at least one additional silencing mechanism, named "quelling," that is sometimes triggered by introduced DNA (see Selker, 1997 and references therein). All indications are that quelling is equivalent to "cosuppression," or posttranscriptional gene silencing (PTGS), discovered first in plants (see Vaucheret, 1998 and references therein) and subsequently in animals (see Sharp, 1999 and references therein). Although the detailed mechanism of PTGS is not yet known in any organism, it is clear that this family of repeat-associated silencing phenomena is fundamentally different from that of RIP and MIP. There are several distinctive features of PTGS. (1) Single-copy sequences can experience PTGS even in haploid cells; thus, while commonly repeat associated, PTGS is not repeat induced. (2) PTGS induced by transgenic DNA in one nucleus can spread to untransformed nuclei. Indeed, in plants, sequence-specific silencing signals can travel long distances systemically. (3) Not all sequences appear capable of triggering PTGS. (4) Not every transformant bearing a PTGS-susceptible sequence shows silencing, even when multiple copies of the DNA are present. (5) PTGS commonly appears dependent on transcription. Strong transgene promoters promote silencing, as does robust expression from the corresponding endogenous gene. (6) DNA methylation is frequently associated with PTGS, but silencing can occur in the absence of methylation, for example in a *N. crassa* mutant lacking DNA methylation and in *C. elegans* and *D. melanogaster*, which are devoid of DNA methylation. (7) PTGS is frequently, if not always, associated with abnormal RNA molecules related to the silenced genes and, at least in some cases, can be triggered by RNA in the absence of DNA. Double-stranded RNA seems to be the most potent inducer of PTGS (see Waterhouse et al., 1998; Sharp, 1999 and references therein). PTGS may result from a chain reaction initiated directly by double-stranded RNA or by single-stranded "aberrant RNA" that pairs with normal and/or abnormal mRNAs (Figure 2). Degradation of such complexes may generate more aberrant RNAs, which could trigger additional degradation (Metzlaff et al., 1997). Moreover, any double-stranded RNA may serve as a template for transcription by an RNA-dependent RNA polymerase, which could produce the actual silencing molecules. Initiation and maintenance of PTGS can be mechanistically uncoupled and DNA:DNA or DNA:RNA interactions may initiate the process in some cases (see Voinnet et al., 1998).

PTGS may be the most common form of repeat-associated gene silencing in plants, but it is not the only one. Transcriptional gene silencing (TGS) associated with repeated DNA has been observed in both transformed and untransformed plants. Although it is too early to conclude that TGS and PTGS phenomena are mechanistically unrelated, they show distinctive features in addition to the level at which silencing occurs (see Vaucheret, 1998). Genes silenced by TGS, but not PTGS, can remain silent one or more generations after the silencer allele segregates away, and in some cases they can transmit their silent state and their silencing ability to naive alleles. Such behavior was first observed with some alleles of genes in untransformed plants and is referred to as "paramutation" (see Hollick et al., 1997). DNA methylation is commonly found in the promoter regions of genes silenced at the transcriptional level, and this methylation can be "contagious," consistent with the possibility that it is involved in silencing (Figure 1).

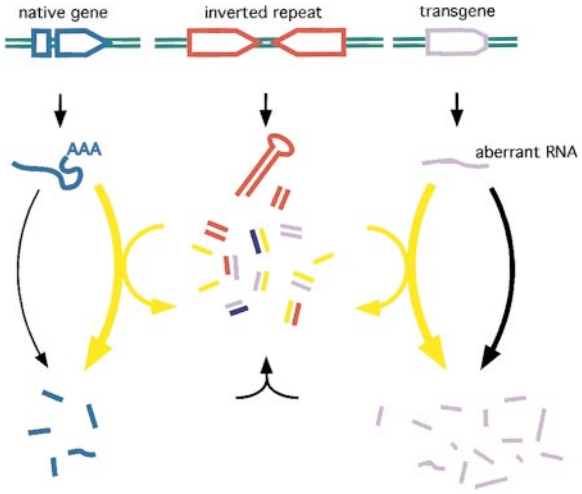

Figure 2. RNA Self-Control

Posttranscriptional silencing can result from massive degradation (large yellow arrows) of mRNA triggered by double-stranded RNA (smaller yellow arrows). Double-stranded RNA may result directly by transcription of an inverted repeat (shown in red; transcript represented as hairpin structure), by annealing of complementary segments of RNAs (center), or perhaps by enzymatic synthesis of cRNA (short yellow segments) by an RNA-dependent RNA polymerase. Transcription of double-stranded RNA templates by such a polymerase may produce the actual silencing molecules. Aberrant RNAs (at right), which are regarded as particularly prone to degradation (heavy black arrow), could result from abnormal RNA processing or template abnormalities including DNA methylation (not shown). Methylation may be triggered by DNA:DNA or RNA:DNA interactions. The latter could result in a positive feedback loop in addition to that illustrated involving double-stranded RNA.

Inverted Repeats: Special Powers?

A number of findings suggest that inverted repeats can be particularly potent silencers of gene expression. A good example involves a gene family in *Arabidopsis* specifying the enzyme phosphoribosylanthranilate isomerase (PAI). The standard strain of *Arabidopsis*, Columbia (Col), has three unlinked *PAI* genes (*PAI1*, *PAI2*, and *PAI3*), presumably accounting for failure to isolate *PAI* mutants in this strain. *PAI* mutants were readily isolated in another ecotype, however, and paradoxically, the fruitful strain, Wassilewskija (WS), actually contains one additional *PAI* gene compared to Col (Bender and Fink, 1995). Bender and Fink showed that the extra copy resulted from a nearly perfect, tail-to-tail duplication of the *PAI1* gene, and that the presence of this inverted repeat was correlated with DNA methylation of all *PAI* genes of WS. All *PAI* genes are unmethylated in Col, and two are expressed, *PAI1* and *PAI2*. *PAI2* is silenced in WS, while *PAI3* appears nonfunctional in both ecotypes. Curiously, the inverted repeat in WS (*PAI1-PAI4*) produces abundant functional mRNA despite its heavy

methylation. A spontaneous deletion of *PAI1-PAI4*, mediated by direct repeats flanking the pair of genes, resulted in the *PAI*-deficient plants. This deletion was associated with an immediate loss of some methylation from the remaining unlinked *PAI* sequences (*PAI2* and *PAI3*), and the residual methylation was unstable, consistent with the notion that the inverted repeat was somehow responsible for the methylation of all *PAI* sequences of WS. Further reduction in methylation, whether spontaneous, induced by treatment with the drug 5-azacytidine, or caused by a mutation in a gene required for efficient maintenance of methylation, was associated with derepression of *PAI2* (Bender and Fink, 1995; Jeddeloh et al., 1998).

Recent results from the Bender laboratory provide additional strong evidence that methylation of all *PAI* sequences in WS is caused by the *PAI1-PAI4* inverted repeat (Luff et al., 1999). Introduction of the inverted repeat into the Col background, by crossing, induced methylation of Col *PAI* sequences. Interestingly, several generations were required to reach full methylation, and de novo methylation of the *PAI3* pseudogene was particularly slow. The methylated *PAI2* and *PAI3* genes from WS were unable to trigger methylation of Col *PAI* sequences, but, as expected from the previous study, substantial methylation of *PAI2* and *PAI3* from WS was maintained in the absence of the inverted repeat. However, detailed analysis of methylation by genomic sequencing demonstrated that methylation at asymmetric (non CpG or CpNpG) sites was largely lost in the absence of the inverted repeat. This supports prior suggestions that methylation at asymmetric sites may reflect de novo methylation, while maintenance methylation may depend on methylation at symmetric sites, perhaps as originally proposed in the 1970s (see Dieguez et al., 1998). Bender and colleagues suggest that the inverted repeat is inherently prone to de novo methylation and is capable of triggering methylation of homologous sequences through DNA:DNA interactions. In keeping with this idea, methylation of *PAI* sequences covered the regions of homology including promoter and intron sequences and terminated almost exactly at the start of the nonhomologous sequences. Although RIP and MIP provide precedent for the proposed DNA:DNA interactions, other possibilities cannot yet be ruled out. For example, in light of recent evidence that DNA methylation can be triggered by RNA (Wassenegger et al., 1994; Jones et al., 1998), it seems possible that methylation and silencing of *PAI2* are directed by RNA transcribed from the inverted repeat as discussed below.

The observed specificity of methylation of *PAI* sequences in Col/WS hybrids supports the idea that the *PAI* inverted repeat triggers de novo methylation of homologous sequences. It remained possible, however, that sequences linked to the inverted repeat are partially or fully responsible for the observed effects. As a first step to investigate this possibility, Bender and colleagues generated transgenic plants with a promoterless *pai1-pai4* inverted repeat construct (Luff et al., 1999). Partial methylation of the inverted repeat was observed after two generations, and heavier methylation was observed two generations later. The authors conclude that the inverted repeat, per se, can induce de novo methylation. Caution is advised, however, until other transgenic experiments are carried out, for example to test (1) a nonduplicated *PAI* sequence, (2) a *PAI* direct repeat, and (3) the possible role of sequences flanking the *PAI1-PAI4* inverted repeat. It is also noteworthy that the promoterless *pai1-pai4* construct, when present as a single copy, failed to trigger methylation of *PAI2* or *PAI3* sequences. It will be interesting to learn whether constructs including the *PAI* promoters can trigger methylation in *trans*, as expected from the behavior of the natural *PAI1-PAI4* inverted repeat.

Two additional cases of silencing in untransformed plants support the idea that inverted repeats have special power(s) in gene silencing. Chalcone synthase (CHS), an enzyme in the anthocyanin pigment pathway, is encoded by a multigene family in plants that have been examined. Curiously, in both snapdragon and soybean, dominant or semidominant alleles are known that cause marked reduction in steady-state mRNA from the *CHS* gene family. The dominant-negative alleles have inverted repeats not present in wild-type alleles. Phenotypic reversion is associated with loss of an inverted repeat, and new dominant-negative alleles show gain of an inverted repeat (Bollmann et al., 1991; Todd and Vodkin, 1996).

Analyses of transformed plants have also implicated inverted repeats in both transcriptional and posttranscriptional gene silencing. Either type of silencing can occur in the absence of inverted repeats, but detailed analyses of integrated transgenic DNA have revealed that the presence of inverted repeats is strongly correlated with silencing (Hobbs et al., 1993; Jorgensen et al., 1996; Stam et al., 1997). In one study, Stam and colleagues showed in petunia that all of the rare instances of PTGS caused by a promoterless *CHS* transgene were associated with both an inverted repeat arrangement of the transgenic DNA and DNA methylation (Stam et al., 1998). Methylation and silencing were maximal near the central junction of the inverted repeats. Analyses of genetic crosses showed that while the inverted repeats could silence both the endogenous *CHS* gene and single-copy transgenes, methylation was only triggered in transgenes. Silencing occurred in the absence of detectable transcription from the inverted repeat, but it appeared to depend on transcription of the endogenous gene.

Although there are other suggestions that inverted repeats can be potent silencers in the absence of transcription, several recent investigations implicate transcription of the inverted repeat (Hamilton et al., 1998; Waterhouse et al., 1998; Mette et al., 1999). A study to test whether the nopoline synthase promoter (pNOS) could be silenced in *trans* by RNA homologous to the promoter itself provides an interesting example. In general, TGS was not induced by pNOS RNA, but an exceptional plant with an inverted repeat including pNOS showed silencing and de novo methylation of an unlinked gene driven by this promoter. When the promoter responsible for transcription through pNOS was itself silenced, silencing and methylation of pNOS were eliminated, suggesting that RNA, or the act of transcription, was involved (Mette et al., 1999). In light of recent evidence that double-stranded RNA is a potent silencer of genes in worms and flies (see Sharp, 1999), it is attractive to suppose that transcription through inverted repeats

creates double-stranded RNA that then triggers silencing and, in some cases, DNA methylation (Figure 2). Consistent with this possibility, Waterhouse and colleagues recently demonstrated in tobacco that either simultaneous expression of sense and antisense constructs or transcription through an inverted repeat caused silencing of homologous sequences (Waterhouse et al., 1998). It is noteworthy in this context that the *PAI1-PAI4* inverted repeat of WS produces an RNA species that may result from readthrough across the inverted repeat (Bender and Fink, 1995). Thus, it may be premature to conclude that the special silencing power of this inverted repeat reflects DNA:DNA interactions. It would be interesting to learn whether silencing by the *PAI1-PAI4* inverted repeat depends on its tail-to-tail orientation and on the absence of a transcriptional terminator between the genes. It would also be instructive to learn whether the special powers of the *PAI1-PAI4* allele are nucleus limited.

Future Prospects

With complete genomes being churned out as previously unimaginable, one might predict that geneticists will soon need to be content dotting their i's and crossing their t's. But new genetic mechanisms continue to be uncovered, even in the most intensively studied model organisms, as exemplified by the repeat-associated gene silencing phenomena discussed briefly above. (I regret that imposed limits preclude direct citation of many key studies.) Besides the cases discussed, repeated DNA is associated with two prominent silencing processes in mammals, genomic imprinting and X inactivation, as pointed out by Barlow and colleagues and Lyon, respectively. Silencing can result from mutation, DNA methylation, repressive chromatin structure, RNA degradation, or a combination of factors. Unfortunately, confusion often follows discovery of novel biological phenomena, and in the area of gene silencing, the confusion is still formidable. No repeat-associated silencing mechanism is yet understood in detail, and it is not even clear how many distinct mechanisms are responsible for the various examples of gene silencing that have come to light. It is virtually certain that some silencing phenomena result from DNA:DNA interactions and that some others result from RNA:RNA interactions, but the basis of silencing remains uncertain in most cases. We are also largely in the dark about the role of RNA:DNA interactions (e.g., as a trigger of DNA methylation). An important aim at this stage is to distinguish between superficial and fundamental similarities among silencing phenomena. In the case of inverted repeats, for example, we must determine whether the prominence of inverted repeats in gene silencing reflects both their tendency to produce double-stranded RNA and their potential to form cruciform structures that may promote DNA:DNA interactions and be prone to de novo DNA methylation.

Selected Reading

Bender, J., and Fink, G.R. (1995). Cell *83*, 725–734.

Bollmann, J., Carpenter, R., and Coen, E.S. (1991). Plant Cell *3*, 1327–1336.

Dieguez, M.J., Vaucheret, H., Paszkowski, J., and Mittelsten Scheid, O. (1998). Mol. Gen. Genet. *259*, 207–215.

Dorer, D.R., and Henikoff, S. (1994). Cell *77*, 993–1002.

Hamilton, A.J., Brown, S., Yuanhai, H., Ishizuka, M., Lowe, A., and Solis, A.-G.A. (1998). Plant J. *15*, 737–746.

Hobbs, S.L.A., Warkentin, T.D., and DeLong, C.M.O. (1993). Plant Mol. Biol. *21*, 17–26.

Hollick, J.B., Dorweiler, J.E., and Chandler, V.L. (1997). Trends Genet. *13*, 302–308.

Jeddeloh, J.A., Bender, J., and Richards, E.J. (1998). Genes Dev. *12*, 1714–1725.

Jones, A.L., Thomas, C.L., and Maule, A.J. (1998). EMBO J. *17*, 6385–6393.

Jorgensen, R.A., Cluster, P.D., English, J., Que, Q., and Napoli, C.A. (1996). Plant Mol. Biol. *31*, 957–973.

Luff, B., Pawlowski, L., and Bender, J. (1999). Mol. Cell *3*, 505–511.

Malagnac, F., Wendel, B., Goyon, C., Faugeron, G., Zickler, D., Rossignol, J.L., Noyer-Weidner, M., Vollmayr, P., Trautner, T.A., and Walter, J. (1997). Cell *91*, 281–290.

Mette, M.F., van der Winden, J., Matzke, M.A., and Matzke, A.J. (1999). EMBO J. *18*, 241–248.

Metzlaff, M., O'Dell, M., Cluster, P.D., and Flavell, R.B. (1997). Cell *88*, 845–854.

Selker, E.U. (1997). Trends Genet. *13*, 296–301.

Sharp, P.A. (1999). Genes Dev. *13*, 139–141.

Stam, M., de Bruin, R., Kenter, S., van der Hoorn, R.A.L., van Blokland, R., Mol, J.N.M., and Kooter, J.M. (1997). Plant J. *12*, 63–82.

Stam, M., Viterbo, A., Mol, J.N.M., and Kooter, J.M. (1998). Mol. Cell Biol. *18*, 6165–6177.

Todd, J.J., and Vodkin, L.O. (1996). Plant Cell *8*, 687–699.

Vaucheret, H. (1998). Plant J. *16*, 651–659.

Voinnet, O., Vain, P., Angell, S., and Baulcombe, D.C. (1998). Cell *95*, 177–187.

Wassenegger, M., Heimes, S., Riedel, L., and Sanger, H.L. (1994). Cell *76*, 567–576.

Waterhouse, P.M., Graham, M.W., and Wang, M.B. (1998). Proc. Natl. Acad. Sci. USA *95*, 13959–13964.