Plant Molecular Biology **42:** 251–269, 2000.

Transposable element contributions to plant gene and genome evolution

Jeffrey L. Bennetzen
Department of Biological Sciences, Purdue University, West Lafayette, IN 47907-1392, USA (e-mail: maize@bilbo.bio.purdue.edu)

Key words: genome structure, mutation, repetitive DNA, retroelements, transposition

Abstract

Transposable elements were first discovered in plants because they can have tremendous effects on genome structure and gene function. Although only a few or no elements may be active within a genome at any time in any individual, the genomic alterations they cause can have major outcomes for a species. All major element types appear to be present in all plant species, but their quantitative and qualitative contributions are enormously variable even between closely related lineages. In some large-genome plants, mobile DNAs make up the majority of the nuclear genome. They can rearrange genomes and alter individual gene structure and regulation through any of the activities they promote: transposition, insertion, excision, chromosome breakage, and ectopic recombination. Many genes may have been assembled or amplified through the action of transposable elements, and it is likely that most plant genes contain legacies of multiple transposable element insertions into promoters. Because chromosomal rearrangements can lead to speciating infertility in heterozygous progeny, transposable elements may be responsible for the rate at which such incompatibility is generated in separated populations. For these reasons, understanding plant gene and genome evolution is only possible if we comprehend the contributions of transposable elements.

Introduction

Before they were observed to transpose [82], transposable elements were perceived as exceptionally mutagenic agents that acted on individual genes [34] or overall genome structure [81]. After their original discovery and characterization in maize, transposable elements of one type or another have been found in all organisms, including all plant species that have been investigated. In many plants with large and complex genomes, transposable elements make up over 50% of the nuclear DNA [104], yet they still lack any proven positive role in the fitness of an individual member of a species. In every plant species, transposable elements are the major identified type of non-genic DNA, so they do provide raw material that can be used to assemble or otherwise modify genetic function [65, 121, 122]. Moreover, most of the activities of a transposable element give rise to changes in gene and/or genome structure, often with accompanying alterations in gene activity. It was for this reason that McClintock originally named these entities 'controlling elements' [83] and that was why she proposed that one of their major roles in evolution was to serve as a source of hypermutagenicity that could create surviving individuals from a population that was stressed to the point of annihilation [86]. For all of these reasons, it would be impossible to have any meaningful conception of plant genome structure and evolution without understanding the contributions of transposable elements.

Transposable element types

All transposable elements share two basic properties. The first is the ability to move from place to place in the genome – hence their designation as mobile DNAs or transposable elements. The second is their ability to amplify their copy number within the genome via this transposition, thereby providing a selectable function that can make them selfish or parasitic DNAs [28, 89].

How these two activities are accomplished and regulated is quite different, however, particularly between the two major classes of transposable elements, the DNA transposable elements and the retroelements.

DNA transposable elements were the first identified in plants, primarily because they gave rise to altered gene or genome phenotypes at very high frequencies in both germinal and somatic tissues. These elements include the *Ac/Ds*, *Spm/dspm* (*En/I*) and *Mutator* systems of maize and the *Tam* elements of snapdragon [39, 65]. Most of these elements range in size from a few hundred bases to about 10 kb. An abundant class of small transposable elements in plants, the miniature inverted-repeat transposable elements (MITEs), also have a structure indicating that they are likely to be DNA transposable elements [18, 121]. DNA transposable elements are found in all organisms, and are the major class of transposable DNAs in all prokaryotes characterized. These elements all have terminal inverted repeats (TIRs), ranging in size from 11 bp (*Ac/Ds*) to a few hundred bases (*Mutator*), although imperfect repeats within the first few hundred bases of the termini may play an important role in the activity of all of these elements (Figure 1).

A family of DNA transposable elements is defined by the fact that they share the same TIR sequences. Hence, all *Ac* and all *Ds* transposable elements have approximately the same 11 bp TIR, while the 13 bp inverted repeat termini shared by all *Spm/dspm* elements are completely different from the *Ac/Ds* TIR. Within a family, one or more members will encode an enzyme, called a transposase, that recognizes the family's TIRs. For instance, *Spm* encodes a transposase that will interact with the termini shared by *Spm* and *dspm* elements, and lead to transposition of those elements. The *Spm* transposase does not recognize or bind the TIRs of *Ac/Ds* elements or any other transposable element family in maize. Hence, *Spm* is called the autonomous member of the *Spm/dspm* family because it encodes the potential for its own transposition as well as the ability to transactivate transposition of other (i.e. *dspm*) elements in the family. *Ac* is the autonomous member of the *Ac/Ds* family. In most cases, the non-autonomous members of a transposable element family (like *Ds* in the *Ac/Ds* family) contain deletions or are otherwise defective derivatives of the autonomous element that have lost the ability to encode a functioning transposase [39, 65].

The MITEs are unusual in having no identified autonomous elements [121]. The existence of undiscovered autonomous elements that encode MITE-specific transposases is likely, although it is also possible that these tiny elements utilize a *trans*-acting transposition function that is not itself encoded on a mobile DNA. Such an activity might be specified by a standard 'host' gene involved in some other cellular process (e.g. DNA replication, recombination or repair), although there are no obvious candidates at this time.

In a successful transposition event, the transposase encoded by a DNA transposable element recognizes the TIRs of its family, causing the excision of the DNA between the TIRs and its reinsertion elsewhere in the genome. The gap left by the excision of the element is then repaired, in some cases by a simple ligation across the gap (leading to a net excision) and in some cases by recombinational gene conversion across the gap using either the other homologue or the sister chromatid as template (and thus leading to no net excision).

Retroelements, or RNA transposable elements, are particularly abundant in eukaryotes. In most or all plant species, they comprise the greatest mass of transposable elements. Retroelements make up over 70% of the nuclear DNA in maize [104] and are equally or even more numerous in other plant species with large complex genomes. All retroelements transpose through reverse-transcription of an RNA intermediate. That is, the DNA version of a retroelement encodes an RNA that is reverse-transcribed into DNA that then integrates. Hence, these elements do not excise when they transpose. Instead, they make a copy that inserts elsewhere. There are five types of retroelements, and plants contain representatives of at least four types [8, 45]. The presumably most ancient [126] class of retroelements is the long interspersed nuclear elements (LINEs). These elements have the structures of an integrated DNA version of an mRNA (Figure 1). A fully intact LINE will encode both gag proteins (involved in intracellular packaging of the RNA transcript) and a polymerase (*pol*) function that includes the enzyme reverse transcriptase. The *pol* functions have the ability to reverse-transcribe a LINE RNA into DNA, while an endonuclease (EN) also encoded by the element is probably associated with integration into the genome [23]. LINE insertions are flanked by short direct duplications of target DNA, like those seen for all other mobile DNAs, usually created by the action of a transposase or integrase.

The most numerous class of large retroelements in plants are the retrotransposons that contain direct long terminal repeats (LTRs) (Figure 1) [8, 45].

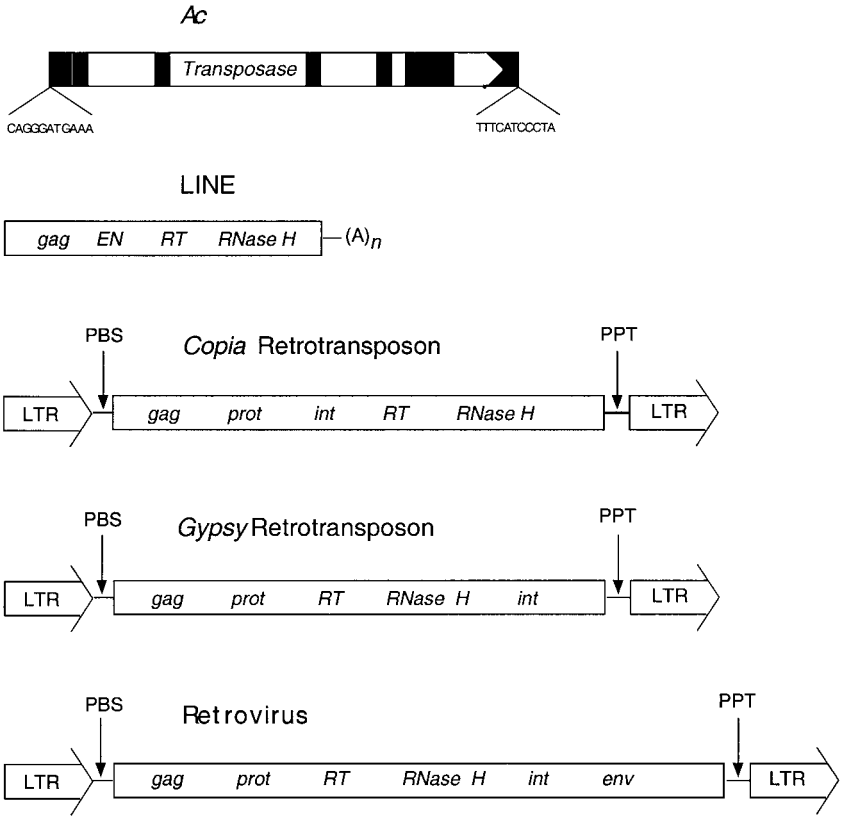

Figure 1. Mobile DNA structures. The *Ac* figure shows transposase encoded by a transcript that is provided by the open boxes, while the shaded regions are either not transcribed or are introns. The letters below the *Ac* figure are the near-identical 11 bp inverted repeats. The open boxes all show various retroelements, with the proteins that they specify designated within the boxes (see text). PBS is the primer binding site and PPT is the polypurine tract, key sequence regions needed for replication/transposition of the LTR-containing elements. Horizontal arrows in the *Ac* transcript and LTRs show the predicted directions of transcription. All of the figures are drawn to approximate scale with the 4.8 kb *Ac* element, although different retroelement families can be very different in size.

LTR-retrotransposons vary in size from several hundred bases to over 10 kb, with LTRs that are usually a few hundred bases to several thousand bases in length. These elements presumably evolved from a LINE that acquired LTRs [126], perhaps as an outcome of a tandem insertion preference like that seen for the *Drosophila* Het-A and TART elements [26, 92]. Regardless of their mechanism of origin, LTR-retrotransposons are found in all eukaryotes either because they originated early in the eukaryotic lineage or because of their potential horizontal transfer. LTR-retrotransposons encode an integrase (IN) function that allows them to incorporate the circular product of reverse transcription into the chromosome. The two major subclasses of LTR-retrotransposons (named after their first representatives observed in *Drosophila*) are called *gypsy* and *copia* elements, and they differ in the position of integrase within the encoded polyprotein (Figure 1). In animals, a *gypsy* retrotransposon apparently acquired an envelope (*env*) gene that allowed it to be packaged in a membranous envelope, leading to intercellular (and interorganismal) infectivity [126]. These infectious retroelements are called retroviruses, and are believed to be found only in animals (Figure 1).

The last class of retroelements is represented by the small interspersed nuclear elements (SINEs). In some animals, for instance with the *Alu* sequences of man [107], SINEs are highly abundant. In plants, SINEs are relatively rare in most genomes that have been investigated [8, 45]. SINEs are usually only 100 to 300 bp in size, and appear to be derived from reverse transcription of RNA polymerase III products. They encode no known peptides, and must use *trans*-acting polymerase and integrase functions in order to transpose. The SINEs are usually derivatives of tRNA

or snRNA genes that have mutated to a structure that can be reverse-transcribed and integrated. Because these RNA polymerase III-transcribed genes carry a promoter specified within the RNA itself, a newly inserted element can usually be transcribed in any active part of the genome, thus creating a high potential for amplification [71].

Genes transcribed by RNA polymerase II also can sometimes be reverse-transcribed and integrated into the genome by the action of *trans*-acting polymerase and integrase functions. These integrated RNA copies are seen as intronless pseudogenes. They lack introns because they are usually derived from mature mRNA, they often have an integrated poly(A) tail at their 3′ end (as do most LINEs), and they usually lack a promoter (unless they happen to insert near one). Relative to man, for instance, plants have relatively few intronless pseudogenes [31, 70]. Eickbush has argued that intronless pseudogenes (and, one might add, SINEs as well) are likely to utilize *trans*-acting functions encoded or induced by LINE elements [33]. The correlated deficiency of LINEs, SINEs and intronless pseudogenes in plants compared to mammals agrees with this model.

Unlike DNA transposable elements, the definition of a family of retroelements is not functionally unambiguous. Retroelement polymerase and/or integrase functions may show some degree of preference for action on elements related to those that encoded them, but that preference is not likely to be absolute. Otherwise, it is not clear how any of the steps in the transposition/creation of a SINE or an intronless pseudogenes could occur. Moreover, the first mobile retroelement isolated in plants, *Bs1* of maize, was identified as having transposed into the *adh1* gene in a genetic background that contained no *Bs1*-related element that encoded a reverse transcriptase [57, 59]. In the absence of a functional definition of a retroelement family, families have been defined by their degree of sequence or structural similarity. In some cases, for instance in distinguishing between LINE elements and LTR-retrotransposons or between *gypsy* and *copia* LTR-retrotransposons, the distinction is solidly based on the presence/absence or location of major motifs. In other cases, it is based on the degree of sequence divergence. For example, the Bennetzen lab arbitrarily has chosen 50% sequence identity between the LTRs of a retrotransposon as definitional of a family in maize. As more elements are sequenced, and with the potential for chimeric elements as an outcome of ectopic recombination, the differences between members of different families can begin to shade into a gray zone. Moreover, different laboratories will not necessarily set the same definitional standards. However, at least for the moment, this sequence-relatedness criterion provides a usable tool until informative distinctions can be made on the basis of function.

Transposable element origins

The ubiquity of transposable elements in all living organisms suggests an early origin of these mobile DNAs. However, their mobility makes transposable elements particularly likely candidates for horizontal transmission. In bacteria, for instance, they are often found on plasmids and are activated by the process of mating and concurrent DNA replication [62]. The retroviruses can travel both within members of a species and among species very efficiently. Hence, it is not known when these elements arose, nor are the specific mechanisms of this origin clear. It seems likely, however, that DNA transposable elements and retroelements are derivatives of independent evolutionary creations. The concept of selfish or parasitic DNA [28, 89] suggests that the ability to amplify within a genome would be selected for any sequence and, as long as this did not significantly decrease the fitness of the host, would give rise to such elements, perhaps through multiple independent origins. Analysis of different SINEs indicates just such a series of independent origins, from different RNA polymerase III products [71].

At a different level, one can also ask about the origins of a particular transposable element family. The *Ac/Ds* family of maize, for instance, has closely related elements (with very similar TIRs and encoded transposase) in many other plant species, as does the *Spm/dspm* (*En/I*) family [17, 39, 50, 65]. Hence, these elements may have been present in the primordial angiosperm. Moreover, given that these elements tend to evolve more rapidly than the genic DNA within a genome (see below), it is entirely possible that they share ancestry and vertical evolutionary descent with similar transposable elements in animals.

It is clear, though, that some elements are more abundant in some genomes than in others, and that some families of elements are found primarily in one species or another. Many MITEs, for instance, are primarily found in monocotyledonous plants [18, 121]. Certain subfamilies of MITEs appear to be distinctive to particular lineages. Similarly, one subclass of *gypsy*

retrotransposons has been found only in plants [125]. In these cases, it seems likely that new transposable element families are derived from previous transposable element families at a relatively high frequency. For defective elements that do not need to encode their own transposition processes, this may occur very rapidly via extensive rearrangement of internal sequences or by acquisition of sequences from other parts of the genome [13, 58, 110]. MITEs might be created *de novo*, much like SINEs, through a small number of minor mutations that convert a small AT-rich region of the genome into a sequence composition that is recognized by MITE-specific *trans*-acting factors. However, creation of a new transposase with a new recognition specificity, and appropriate ends, is likely to be a much rarer event.

Within the maize genome, a very large number of LTR-retrotransposons have been found that appear to have a very recent origin [105, 106]. These several different element families, which combine to make up over 70% of the maize nuclear genome [104], mostly appear to have arisen within the last 2 to 6 million years [105]. It is possible that low copy numbers of these elements existed in the maize genome long before this time, and that their amplification was a recent event. Alternatively, they may have arisen via horizontal transfer within this short time frame. One possible horizontal source for these elements would be a wide cross that might transfer only mobile DNAs, because one of the participating sets of chromosomes would be progressively lost [1]. Another possibility is a horizontal transfer of element DNA or RNA, either as a naked nucleic acid or within a packaged virus. These types of transfers might occur quite commonly into damaged tissues (e.g. insect feeding sites), given the propensity of eukaryotic cells to take up and incorporate exogenous DNAs into their genomes. However, only rarely would such a transfer be likely to occur in tissues that would give rise to gametes and thus be transmitted to the next generation. This model suggests that a greater number of transposable elements and a resultant tendency towards larger genomes might be observed in plant species that often reproduce vegetatively.

Interestingly, the LTR-retrotransposons of maize have several properties in common with retroviruses, including the ability to acquire (and perhaps transduce?) sequences from other genes [19, 58, 90] and extra sequence information in the part of the element where *env* is usually encoded [55, 58, 66, 125]. Could it be that some of the LTR-retrotransposons that are so abundant in plants are defective retroviruses [8]? It is possible that they could be derived from animal retroviruses, perhaps imparted by insects that fed on gametophytic tissues. Although an *env*-packaged retrovirus would not be likely to be infective in plants, given the presence of a cell wall, an insect retrovirus might be able to replicate intracellularly, thus becoming an LTR-retrotransposon. Moreover, selection against *env* function would be expected in such an element because packaging and extracellular export would be counterproductive, perhaps accounting for the current highly defective appearance of the putative *env*-derived regions of many current plant LTR-retrotransposons [55, 66, 125].

Transposable element specificities

Plant transposable elements were first isolated from alleles of genes that they had inserted into or near. These alleles often exhibited unusual behaviors, such as high somatic and germinal instability. This instability was commonly perceived as phenotypic reversion, and usually required the action of an unlinked autonomous transposable element. Hence, their interesting behaviors made the cloning of transposable elements an early goal of plant molecular genetics. The approach was to use traditional cDNA isolation of a regular coding gene as the route to isolating alleles of the gene that also carried an element [14, 37]. These elements were first cloned because they had inserted into genes, but it was not clear whether such genic associations were a common location for transposable elements.

However, studies of sequences flanking DNA transposable elements have routinely indicated preferences for insertion and/or maintenance in active regions of the genome that contain genes. The first studies indicated that *Mu1* (a defective member of the *Mutator* family) and *Spm* both preferentially insert into the minority of the maize genome that is unmethylated in adult tissues [10, 22]. Subsequent studies indicated that the DNA around *Ac/Ds*, *Mu1* and other DNA transposable elements was primarily of a low-copy-number type [21, 25]. In a recent study, Tikhonov *et al.* [111] found 33 MITEs in a 225 kb region of the maize genome flanking *adh1*. None of the MITEs in this region were found within the 166 kb (74%) of DNA that was occupied by LTR-retrotransposons. Hence, MITEs show a strong insertion preference for genic DNA, and especially for the regions 5′ and 3′ to a gene where matrix attachment regions (MARs) are found [3, 111] (see below). Hence,

with no exceptions to date, all DNA transposable elements exhibit preferential insertion and/or retention within unmethylated, genetically active, presumably euchromatic regions of plant genomes.

Some LTR-retrotransposons have been associated with insertional mutations [8, 45, 59, 99], and with preferential insertion into low-copy-number DNA [53]. In this regard, these elements behave like the DNA transposable elements in plants. The copy numbers of these families of gene-preferring elements are usually much less than a few hundred per genome. However, the most abundant LTR-retrotransposons of maize are primarily associated with methylated, presumably heterochromatic DNA [12]. In the *adh1* region of maize, these LTR-retrotransposons are commonly found as nested clusters of elements inserted within each other [105, 106]. These 'intergene LTR-retrotransposons' (IRPs) exhibit an approximate five-fold preference for insertion into the LTRs of retrotransposons, relative to their insertion in other parts of the retroelements. IRPs come in all varieties, including *gypsy* and *copia* types but, despite comprising over 70% of the maize genome [104], IRPs have not been associated within any mutated gene in maize. Hence, we believe that IRPs have evolved a specificity for avoidance of genes [106]. We do not currently have any criterion that allows us to recognize an IRP from raw DNA sequence, but we have observed a general correlation that all high-copy-number LTR-retrotransposons in maize are IRPs [106].

It appears that IRPs have evolved a lifestyle different from that of large DNA transposable elements (which usually have a copy number of a few to a few dozen) and some low-copy-number LTR-retrotransposons [106]. Any low-copy-number element can gain an advantage from inserting into a transcriptionally active part of the genome, thereby providing an opportunity for further transcription and transposition. MITEs, despite copy numbers in the thousands, may also be able to gain this advantage due to their small size (and, hence, relatively low potential to make severe mutations) and a bias for insertion near matrix attachment regions rather than the coding parts of genes [3]. The large IRPs, though, with copy numbers in the thousands, would cause thousands of mutations if they exhibited a gene-specific insertion preference. Such a mutational load would probably be fatal to the 'host' genome. If an element can amplify to copy numbers as high as some of these IRPs, and is interspersed throughout the genome, then passage into the next generation is essentially guaranteed by random segregation and transmission. Such an element would exhibit selection for insertion into inactive regions of the genome, where it would create the least genetic load.

Targeting mechanisms for insertion into active or inactive portions of plant genomes have not been identified. However, insertion into genic/euchromatic DNA might be simply preferred because of its more open nature. In *Saccharomyces cerevisiae*, the Ty5 LTR-retrotransposon preferentially inserts into (the relatively few) inactive portions of the yeast genome. This targeting is determined by a specific interaction between chromatin proteins that keeps these regions silent and the Ty5 integrase [129]. Similarly, IRPs may associate with a heterochromatin-specific protein or proteins in the plant nucleus. By the same token, MITEs may commonly interact with a protein that is found at MARs while other elements that show a gene-specific preference might interact with specific euchromatin proteins. In this regard, many transposable elements in eukaryotes show preferences for insertion into promoter regions of genes, areas that are both very open and associated with a large number of expression-specific proteins.

In situ hybridization analyses have shown that many highly repetitive DNAs show extensive bias for one part of the genome versus another [32, 87, 94, 96, 101, 108]. Some of these repetitive DNAs have been shown to be LTR-retrotransposons. In the sugar beet, Heslop-Harrison and coworkers have found that some highly repetitive LTR-retrotransposons are scattered throughout the genome, but preferentially associate with heterochromatin near centromeres [108]. Others have seen elements that are highly biased toward centromeres or other heterochromatic regions in other species [87, 94, 96, 101]. Edwards and coworkers demonstrated that the highly repetitive IRPs found in the maize *adh1* region were interspersed throughout the chromosome arms, but some were notably deficient in centromeric heterochromatin, knobs and the nucleolar organizer [32]. Hence, different interspersed highly repetitive elements show different biases for insertion and/or accumulation, perhaps caused by differences in the types of heterochromatin (e.g. protein composition) in those different locations.

Elegant genetic studies with the *Ac/Ds* system indicated a preferential transposition of this element family to linked sites in the genome, providing insertions at unlinked sites in only about one half of the transposition events [48]. Transposable elements in all organisms tend to show some biases for par-

ticular regions within any given gene (promoters are commonly hot spots for insertion). Some also show some sequence bias, like the preferred TAA insertion site target of *Tourist* [121], although the biases may be subtle [9].

Transposable element arrangements in plant genomes

With the numerous different levels and degrees of specificity/bias for insertion or accumulation of different transposable elements, a plant genome can arrive at many different arrangements of elements. Moreover, different regions of a genome are likely to have very different arrangements. Centromeric heterochromatin, for instance, has few if any genes but many classes of repetitive DNA. We expect that most of these repetitive DNAs will be mobile (and far outnumber the centromeric repeats [56]). Moreover, the exceptional instability of transposable elements (see below) indicates that they will rapidly rearrange, so that the arrangement of elements in a genome may be very different in a lineage with recent transposable activity compared to one where most of the activity occurred millions of years ago.

Even now, our most comprehensive understanding of the organization of repetitive sequences comes from DNA renaturation studies that were initiated about 30 years ago. These experiments indicated that large plant genomes were largely composed of repetitive DNA, and that most of this repetitive DNA was interspersed with genes [40]. Only recently has it been shown that most of these interspersed repetitive DNAs are mobile DNAs, mainly LTR-retrotransposons [8, 45, 104, 106, 111]. In maize, one study now indicates that these elements are mainly arrayed as blocks of nested LTR-retrotransposons, intermixed with genic blocks of one to a few genes each [106]. However, only two other fairly large segments ($>$50 kb) of complex plant genomes have been sequenced, and these regions were notable for exceptionally high gene densities. In these two cases, a 22 kDa zein gene cluster of maize [69] and the region around the *mlo* locus of barley [91], only a few LTR-retrotransposons were found, arranged as intact or highly rearranged singlets. In all of these large regions of genomic sequence, LTR-retrotransposons made up the greatest quantity of interspersed repetitive DNA, although MITEs were more numerous.

The most extensive characterization to date of plant genome organization comes from the ongoing sequencing of the *Arabidopsis* genome. In the genic parts of the Arabidopsis genome, the standard patterns observed are rare interspersed repetitive DNAs, the largest being retroelements found about once every 200 kb [15]. Because solo LTRs from LTR-retrotransposons have few distinguishing features, they would usually be missed by sequence analyses unless an intact element or other solo LTR of the same family had been sequenced elsewhere. Solo LTRs are a common outcome of unequal recombination between the two directly repeated LTRs of a single element. In yeast, which has a very high recombination rate per kb of DNA, solo LTRs far outnumber intact elements. Hence, it is possible that a great deal of the DNA between genes in *Arabidopsis* is made up of unidentified solo LTRs. In general, though, it is clear that most of the *Arabidopsis* genome is free of large repetitive DNAs. In fact, the anomalously small genome of *Arabidopsis* served as part of the justification for its choice as a model plant species. Hence, we probably should not look to *Arabidopsis* as a model for the more typical (and more complex) plant genomes.

Given that we know so little about the linear arrangement of sequences within any complex plant genome, it is not surprising that we know even less about the three-dimensional organization of repetitive and genic DNAs. From gel blot hybridization analyses conducted mostly in the 1980s, it is clear that most highly repetitive DNAs in plants are cytosine methylated in most or all tissues, at the sequences 5′-CG-3′ and 5′-CNG-3′ [12]. Because DNA methylation of this type usually correlates with genetic inactivity, it is likely that these repeats (which we now know are mostly LTR-retrotransposons in large-genome monocots) are in condensed and heterochromatic structures.

The nuclear matrix is a proteinaceous structure that is believed to be responsible for the regulated folding of chromatin in the interphase nucleus of all eukaryotes. Particular eukaryotic DNA sequences, called matrix attachment regions (MARs), exhibit specific high-affinity binding to the matrix. It is thought that two MARs flanking a chromosomal region will define a physically isolated loop, providing insulation from the chromatin structure and genetic functions in any adjacent loop [3, 44]. Analysis of MARs in the *adh1* region of maize indicates that they usually flank each gene, often separating a gene from an adjacent LTR-retrotransposon block [2]. The MARs may serve as insulator elements that prevent a spreading of the inac-

tivated/methylated state of the retroelement block into adjacent genes.

In summary, we can conclude that different plant genomes have different compositions and arrangements of their repetitive and genic DNAs. Small plant genomes have fairly few repetitive DNAs, and most of these are found in large blocks (e.g. satellites [95], centromeres [56], telomeres [103], and centromere-associated regions). Larger plant genomes have mostly interspersed repetitive DNAs [40]. Some of these repetitive DNAs are interspersed with genes, but much of it is intermixed primarily with other repeats in centromere-associated regions [108]. Most of this interspersed repetitive DNA (at least in monocots) is composed of LTR-retrotransposons. Some of the intact and/or fragmented LTR-retrotransposons are associated with genes, but most are in intergenic blocks that are methylated and presumably heterochromatic. In grasses, large DNA transposable elements are fairly rare and will often be found in or near genes, as will the more numerous MITEs. The structure of the intergenic LTR-retrotransposons (IRPs) blocks may be a nested one like that seen in maize [106], or could be a more complex arrangement of tandemly amplified or otherwise rearranged elements. Although the same types of elements will be found in different plant species, the specific families will largely be distinctive to a genus [8, 39, 45, 65, 76, 115] and some types may predominate in some species, while others are more abundant in another (e.g. LTR-retrotransposons in yeast versus LINEs in man). However, there is still a lot to be discovered in this area. In fact, we can conclude that we still do not really know the rules of arrangement for any plant genome (perhaps other than *Arabidopsis*), and we have little idea of what the nature of most of the exceptional regions will be.

Transposable element activities

Plant transposable elements have a range of possible activities, all of them associated with possible alterations in genome and/or gene structure and function. Chromosome breakage, chromosomal rearrangement, insertional mutation, altered gene regulation and sequence amplification are all identified outcomes of the transpositional and/or recombinational potential of all of the retrotransposon types.

Transposable elements were originally perceived because of their ability to rearrange genes and genomes [33, 81, 82]. McClintock first characterized *Ds* as an element that could break chromosomes, thereby serving to 'Dissociate' the acentric fragment from the rest of the chromatid [81]. Chromosome-breaking *Ds* elements have a structure indicative of an element that has inserted into itself, thereby creating a complex element with three or more ends [30, 36]. It is believed that the inability of the transposase to properly recognize, mobilize and/or repair an attempted excision/transposition of such a 'double *Ds*' element leads to the breakage [36, 118].

Transposable elements now are best known for their transposition, wherein an element moves from one place in the genome to another in a homology-independent manner. All transposable elements manage to increase their copy number via this transposition process. *Ac* elements excise directly after or during the time that a DNA replication fork goes through an element, and then the excised element preferentially transposes to a site that has not yet replicated in that S phase [49]. Hence, this creates three copies of *Ac* (one on the replicated non-participating donor chromatid, two on the replicated target chromatids) at the end of S phase, where replication alone would have only led to two of the replicated copies of the original chromosome. Many other DNA transposable elements, like *Mutator* of maize, transpose much more commonly than they appear to excise [9]. This is probably because the excised element donor site is repaired using the other chromosome or chromatid that has not undergone the excision as a template, as has been shown for bacterial elements, Tc1 of nematodes and P elements of *Drosophila* [4, 35, 98]. Retroelements do not excise at all during transposition, instead using their transcripts as a template to make additional integration-competent DNA copies.

Transposition of an element into a gene will often lead to inactivation of that locus. Because DNA transposable elements can sometimes excise during somatic or gametophytic development, they are often associated with alleles that are highly unstable or 'mutable' in plants. The simplest form of mutability is exact reversion to wild type. However, reversion events often leave behind sequences, sometimes including the flanking direct target repeats that were generated upon insertion, a small segment of the element, and/or a few bases of sequences generated either by illegitimate conversion or other forms of repair of the excision site [30, 93]. Other excision events may be associated with small deletions of DNA (usually from one end) of the target site [67]. Hence, mutations caused by transposable elements in individual genes

include both insertions of the elements, and subsequent sequence changes associated with later rounds of excision.

Transposable elements also appear to change their own structure much faster than do genic sequences within the same genome. Deletions and other internal rearrangements are common, perhaps as an outcome of failed transposition events [85]. IRPs primarily target insertions into the LTRs of other IRPs, thereby presumably inactivating the target IRP [106]. Beyond these self-mutagenic activities, one might imagine that genomes have evolved ways to degrade or otherwise remove these potentially hypermutagenic agents. In the *adh1* region of maize, it has been observed that the elements have an approximately threefold higher than normal ratio of transition to transversion mutations [105]. This could be due to the extensive 5-methylation of cytosines in these elements. Chromosomal 5-methylcytosine has been shown to be a hot spot for transition (C-to-T) mutations, probably due both to an increased rate of deamination in 5-methylcytosine and to the fact that deaminated 5-methylcytosine is identical to thymidine [24]. Hence, cytosine methylation may assist both in transcriptional silencing and in sequence decay of the IRPs.

Transposable elements carry with them regulatory sequences that can alter the expression of adjacent loci. At the simplest level, an insertion of such an element into a promoter of a gene can bring that gene's regulation under the control of the transposable element [74, 75]. Some DNA elements have terminal sequences that allow them to act as introns under some circumstances [61, 119, 120], and the binding of transposase to such elements can lead to a transposase-dependent suppression of gene activity [75].

Finally, some transposable elements can amplify DNA sequences from other parts of the genome. The action of the reverse transcriptase complex from retroelements can potentially turn any RNA (with a fortuitous primer source) into a DNA that can be integrated into the genome. Hence, *trans*-acting retroelement functions can convert a tRNA into a SINE or an mRNA into an intronless pseudogene [31, 70, 71]. Other elements, notably *Mutator* [13, 110] and *Bs1* [19, 58, 90] of maize, have taken up portions of other sequences (e.g. genes) within the elements themselves. Transposition then amplifies these acquired segments along with the rest of the element, thereby leading to a more complex genome.

Transposable element regulation in plants

Like all other expressed sequences in eukaryotes, transposable elements are differentially active in different tissues, at different times in development, and/or under different induction regimes. For instance, some LTR-retrotransposons are most active during male gametophytic development [113] or in root tissues [100], and many are induced by abiotic and/or biotic stresses [52, 81, 86, 97, 99, 116]. Various reviews exhaustively discuss this subject [46, 63, 65]. The degree to which any transposable element can or will rearrange a gene or genome will depend on its level of activity and whether these activities occur in tissues that contribute to the next generation. However, from an evolutionary perspective, transposable elements show two types of control that have not been associated with the regulation of most plant genes. The first of these is a possible tendency toward self-inactivation and the second is a host-determined (or, at least, host-assisted) process of epigenetic silencing.

Whether viewed as parasitic/selfish DNAs or as mobile elements with some possible beneficial role, it is clear that a very high level of transposable element activity can be deleterious to individuals, presumably due to both genic and chromosomal mutations [60]. It is not surprising, then, that most transposable elements are usually inactive in any given individual or population. Part of this inactivity is due to the fact that most transposable elements within a genome are defective. For instance, most maize lines have zero or one active (or potentially active) *Ac* element, but a few hundred *Ds* elements [30, 64]. Many LTR-retrotransposons in a plant appear to be defective as well, existing as solo LTRs or with internal deletions, rearrangements, and/or replacements [55, 58]. This predominance of defective elements is partly due to the self-mutagenic properties of the DNA elements [64, 85], but is also likely to be associated with an intrinsic higher mutation rate of cytosine-methylated DNA [105].

Even in plant genomes that have structurally complete (e.g. autonomous) transposable elements, activity is often lacking. This deficiency is associated with 5-methylation of cytosines within the elements, in the sequences 5′-CG-3′ and 5′-CNG-3′ [7, 9, 12, 27, 117]. Methylation-associated inactivation of DNA transposable elements has been extensively studied with the *Mutator* system of maize, where the methylated nucleotides are largely delimited to the elements and the inactivation appears to be induced by high transpos-

able element copy numbers and/or activity [9]. This epigenetic modification is associated with inactivation, as confirmed by the loss of the 5-methylation of cytosines in elements that have been reactivated [116]. By analogy with large methylated and heterochromatic blocks of plant genomic DNA, it is likely that transposable element methylation is associated with an inactive state of the chromatin, although it is not clear whether the methylation or chromatin alteration occurs first.

The precise mechanism(s) of this epigenetic regulation of transposable elements remain(s) unclear, but the phenomenon does have similarities with the homology-based silencing that has been observed with plant transgenes [77, 79]. In fact, it is highly likely that the transgene silencing process is a secondary outcome of an evolved plant mechanism for the inactivation of plant viruses and transposable elements.

Wide crosses can reactivate silenced transposable elements in *Drosophila* [60], and a large amplification of genome size associated with a wide cross has also been observed in wallabies [88]. Apparently, a subtle regulatory incompatibility of these otherwise balanced genomes leads to a transient loss of the inactivational status quo. Chromosome breakage has also been observed to activate quiescent transposable elements [81, 86, 116], possibly by causing extensive DNA repair that overwhelms the capacities of maintenance DNA methylases and/or heterochromatin assembly factors. LTR-retrotransposons in plants are known to be activated by pathogen infection or wounding of tissues (e.g. insect feeding) and by the release of cell wall fragments during the generation of protoplasts [8, 45, 52, 99]. *Ac* elements can also be activated by passage through tissue culture [97]. All of these reactivations can be lumped under the general context of stress activation.

McClintock viewed stress activation as a clue to the central role for plant transposable elements [86]. McClintock felt that induction of these elements under severe stress could lead to a very large number of new mutations, with a slight chance that one multiply-mutated individual would then survive a stress so severe that all unmutated individuals in the population or species would perish. However, it is hard to see how such a massive mutation capability would be retained over many generations unless there was fairly common selection for its use. Moreover, the occasional success of this process would presumably give rise to species that, although closely related, differed tremendously in genome organization and possibly functioning gene content. Comparative map analysis in plants, particularly in the grasses, suggests a great deal of conservation of gene content and order [41]. In addition, using a traditional analogy for mutation, it is reasonable to conceive that driving a nail through a functioning watch could very rarely create a superior watch, but it seems impossible that driving hundreds of nails through a watch would ever make it better, unless one had trillions of watches.

A simpler model for transposable element activation posits a possible selfish or parasitic origin [28, 89]. It is not unusual for known viral parasites to exist within a prokaryotic or eukaryotic genome in a relatively benign state, managing to survive by replicating passively with the host DNA. However, these viruses can be activated by a stress (e.g. starvation) that indicates their host's survival is in doubt. Then the integrated virus becomes active, making new copies that can find and infect a new host. Plant transposable elements share this stress activation potential, perhaps because they can occasionally undergo horizontal transfer. A wide cross could create an opportunity for transfer to another species, in some cases where the chromosomes that carried the element are not maintained [1]. In particular, the activation of some LTR-retrotransposons by insect feeding or pathogen infection makes very good sense from a retroviral perspective, as this type of tissue stress suggests that a vector for horizontal transfer to another individual or species is present.

Transposable elements and the evolution of genome structure/function

From a quantitative perspective, it is easy to see that transposable elements are the most significant factors in determining the structure of a complex plant genome. In many cases, they make up the majority of such a genome [40, 104]. Equally important, every aspect of their life cycle has the potential for alteration of genome structure and adjacent gene function.

Although the chromosome-breaking *Ds* elements are rare, their potential to rearrange the genome is exceptional. Any time a chromosome breaks, this leads to an abnormal ('sticky') chromosome end that either must be repaired by telomerase or will fuse with any other broken chromosome end. Usually this leads to a breakage-fusion-bridge cycle, if the fusion is either to the sister chromatid or any other broken chromosome in the same nucleus [80]. However, if the broken

end is repaired by fusion with an acentric fragment from another chromosome, then this gives rise to a stable translocation. Other breakage and fusion events could generate inversions, deletions and duplications, all events seen to originate at a *Ds* site [84]. These types of rearrangements will give rise to unbalanced gametes in heterozygotes, thereby leading to a loss of fertility. Individuals that are heterozygous for a handful of large rearrangements (e.g. full-arm inversions and translocations) will be essentially sterile, yielding two parents that can now found separate species. In general, any activity that frequently breaks chromosomes will lead to large DNA rearrangements, and transposable elements can be a substantial source of such breakage.

In yeast, the various Ty retroelements can act as agents of genome rearrangement primarily because they serve as sources of homology for ectopic (or unequal) recombination. Unequal recombination between directly repeated elements at adjacent sites will give rise to reciprocal duplications and deletions of the DNA between the two elements, while unequal exchange between elements in opposite orientations will yield an inversion of the DNA between the elements. Similar ectopic exchange between elements on different chromosomes can give rise to reciprocal translocation. All of these rearrangements, and more complex events requiring more than one ectopic recombination event, have been observed in yeast [124]. In *Drosophila*, such unequal recombination events have also been observed, both between the two LTRs of an LTR-retrotransposon to give a solo LTR and between two distant transposable elements [68]. Such an unequal recombination was the source of the first gene duplication event ever reported, generating the Bar eye phenotype in *Drosophila* [109, 112]. In plants, as in other eukaryotes, most recombination is limited to genes [29, 127], so this should limit ectopic recombination between most transposable elements. The IRPs, in particular, usually appear to be locked away in methylated and heterochromatic blocks that may undergo very little equal or unequal recombination. However, these blocks must have some recombination activity, as they can have a few solo LTRs [106]. These can only have been created by unequal recombination. However, only two out of over twenty elements have generated solo LTRs in the last several million years in the *adh1* region of maize [105], suggesting that recombination of any type is very rare within these blocks.

An obvious outcome of transposable element activity is the amplification of genome size. Current evidence suggests that these elements will make up the majority of complex plant nuclear genomes [8, 11, 45, 104]. In at least some cases, possible mechanisms that could reduce these elements' quantitative contribution to complex plant genomes may be missing or unable to seriously compete with frequent amplifications [11]. Depending on the specificities of these elements, they could lead to larger genes (for elements that often insert into introns or near 3′ or 5′ ends of genes) or to large blocks of heterochromatin (for IRPs and other elements that avoid genes). The placements of these heterochromatic blocks in a given species might depend on the biases of the elements themselves. For instance, some elements may preferentially insert into centromeric heterochromatin [108] and avoid intergenic regions, while another species might predominantly have IRPs that prefer intergenic regions but avoid centromeric heterochromatin [32]. Somewhat surprisingly, wide variations in genome size are not correlated with any catastrophic changes in the biology or fitness of a host. However, to whatever degree overall genome size and heterochromatin content/placement affects gene or genome function [5, 16, 102], amplified transposable elements are likely to be a common underlying cause.

Partly because of their ability to acquire segments of the genome (including genes) [13, 58, 110] and move them to new locations, transposable elements can increase gene copy numbers. Usually these genes would be highly fragmented, but they could provide segments that might be assembled into new composite genes. This model for the assembly of new genes from individual domains of different genes has a long history, including as a possible explanation for the existence of introns as a way of resolving the initially sloppy linkages at the borders of the assembled domains [42]. The ability of some plant transposable elements to act as fairly good introns [61, 119, 120] provides reasonable support for this argument. However, only one known plant intron in a 'wild type' allele has the obvious legacies of a transposable element origin [43], so gene creations of this type must have mostly occurred in the distant past, if they have occurred at all. A second route to increasing gene number would be by two adjacent transposable elements acquiring and transposing the DNA between them. This is the likely origin of many bacterial transposons, from the acquisition of a conditionally useful gene (e.g. heavy-metal or antibiotic resistance) by two

insertion sequences (IS elements) of the same family. Although this has not been seen in plants yet, it would be theoretically possible either via a coordinated transposition of two adjacent elements or by their movement of the intervening gene(s) by ectopic recombination [124].

The centromeric regions of all species of grasses that have been examined contain conserved retrotransposons [87, 101], a surprising result when one considers that these elements usually are not conserved enough outside of their genus to be identified by cross-hybridization [8]. The simplest interpretation of this exceptional conservation of retrotransposons in centromeric regions is that they are performing a function important to chromosome segregation [87, 101]. Although this may seem a surprising possibility, it has long been known that large regions of heterochromatic DNA can serve as partly functional centromeres in plants and other species [128]. Once a retrotransposon arrives in a centromeric region, there is no obvious reason why natural selection would not act on such a sequence if random mutations within it somehow allowed it to assist, or replace, the ancestral centromeric sequences. In this same way, particular *Drosophila* retrotransposons have apparently replaced the requirements for a telomerase gene and the standard type of eukaryotic telomeres [92].

The Avramova lab has recently observed that some MITEs can act as MARs, at least as determined by the *in vitro* matrix binding assay [2, 3]. Although these MARs differ in their matrix-binding activity, and some may be conditional/regulated *in vivo*, it is somewhat disconcerting to think that a property as basic as the folding of a genome could be determined by mobile DNAs. However, the preferential insertion of MITEs near MARs [3] would minimize any disruptive effect.

In summary, transposable elements make up much of the DNA of many genomes [8, 32, 104], and blocks of these repeats comprise major components of cytogenetic features such as centromere-associated heterochromatin. Whether through their transposition, unequal recombination or associated chromosome breakage, these elements can and do rearrange plant genomes (Table 1). Although most of a plant's transposable elements are inactive at any given time, due to both epigenetic regulation and their propensity to acquire defective forms, occasional activity can have enormous effects. Although the breakpoints of large chromosomal rearrangements such as inversions and translocations have not been characterized yet in plants, it is likely that some of them will be associated with transposable elements. Gene numbers, structures, and patterns of interspersion with other sequences will all be determined (at least in part) by transposable elements. The degree to which transposable elements are responsible for the various aspects of chromosome structure and evolution will depend on the number and predominant types of elements that are present, and these two characteristics can differ significantly between even closely related plant species [38, 63, 76, 114, 115].

Transposable elements and gene evolution

There does not appear to be any limitation to the ways in which transposable elements can affect the structure and evolution of individual genes. As described above, they have the potential to increase gene numbers and create new genes by serving as modules for transposition and unequal exchange [31, 42, 58, 70, 112, 124]. Some plant transposable elements can serve as variably functional introns [43, 119, 120], although the mostly conserved positions of introns in distantly related eukaryotes (e.g. plants and animals) suggests that such generations of retained new introns are not common. The insertion of a transposable element into a gene obviously provides the raw material for possible use of those sequences for a new protein-coding potential [65, 72, 121]. In maize, over 50% of genes have a segment of a transposable element (often a MITE or an LTR-retrotransposon fragment) within their promoters and/or transcribed regions, and this number will grow as genomic sequence databases become more complete [111, 121]. We have little or no evidence that an improved gene function has evolved in this way for any plant gene, though, perhaps because of the large size of the these elements and an initial selection against a likely inactivated allele. However, when a DNA transposable element excises, it often leaves behind a 'footprint' of sequences that may include the flanking target direct repeat [30, 64, 93]. These kinds of adjacent short repeats are commonly found within plant genes, including in the coding portions, suggesting that the composition of a gene can be slowly built up by cycles of insertion and excision [30]. However, unequal recombination and replication slippage could also give rise to this kind of sequence pattern within a gene. Deletions or acquisition of apparently unrelated sequences at the donor site associated with transposable element excision would also give rise to an altered sequence within a gene [67, 93].

Table 1. Genome alterations that can be caused by transposable elements.

Element activity	Outcome	Likelihood
Transposition	Increased element copy number	Almost always
Transposition	Increased genome size	Can be common and substantial with LTR-retrotransposons
Transposition	Large cytogenetic structures	Highly amplified elements are a major component of some heterochromatic blocks
Transposition	Create intronless pseudogenes	Rare in plants
Transposition	Move genes or segments to new sites, amplifying region	Frequency unknown in plants
Transposition	Create new chromosome folding patterns	Possible with MITEs acting as MARs
Transposition	Horizontal gene transfer	Not proven in plants
Chromosome breakage	Chromosomal rearrangements (inversions, duplications, translocations, deletions)	Rare class of element, but all of these events can occur at insertion sites of DNA elements
Ectopic recombination	Chromosomal rearrangements, as above, but more likely with nearby elements, hence a bias for smallish duplication/deletion	May be relatively frequent, but the nature of such rearrangements has not been well studied in plants

Perhaps the most likely outcome of a transposable element interaction with a gene, other than insertional inactivation, is the acquisition of new regulatory potential by that gene. Many transposable elements prefer to insert into or near genes, while MITEs and other elements appear to have an even stronger bias for the regulatory regions that are often 5′ to a gene [3, 111]. Because many transposable elements are transcribed themselves, they often carry their own promoter elements, with their own regulatory regimes. Insertions in plant gene regulatory elements can give rise to new tissue specificities for the affected locus [20, 47] and/or place the gene under the control of epigenetic regulation that is directed at the element [47, 74]. The paramutational phenomena associated with some maize genes [54] often may have its epigenetic nature associated with regulatory elements acquired from epigenetically inactivated transposable elements [73, 78].

When a transposable element or any other sequence is inserted into a promoter region, natural selection will act on how the gene does or does not utilize those sequences. Investigations of plant gene sequences indicates that many promoters have fragments of transposable elements in them, often at sites that have been found to be bound by proteins that regulate the gene's expression [65, 121, 122]. It is entirely possible that all plant genes have promoters that are descended from transposable element contributions.

In these numerous ways, transposable elements can influence the evolution of structure and regulation in any gene (Table 2). As with alterations in overall genome structure, the types and frequencies of change observed in genes are a function of the predominant types, abundances, and levels of activity of any particular element within a plant genome. DNA transposable elements and some low-copy-number retroelements have insertion specificities suggesting that they will be the most frequent contributors to genic evolution. Given the propensity of IRPs to reside in apparently inactive portions of the genome, it is likely that these elements contribute relatively little to gene evolution or the evolution of new/altered biological capabilities. Elements that prefer to insert near genes, particularly near their promoters, should be the most likely to have major biological effects by altering gene regulation.

Viewing the genome as an ecosystem

Plant genomes, even exceptionally simple versions like that of *Arabidopsis* [15, 51], are complex structures. The ability to replicate is essential to the survival of any sequence within a genome. The ability of all transposable elements to increase their copy number within a genome by transposition should be a selectable attribute. Those elements that make the most new copies will have the most similar progeny elements, which can then make additional copies. In addition, some studies in bacteria and yeast suggest that populations with functioning transposable elements will generally out-compete populations lacking

Table 2. Alterations in genes that can be caused by transposable elements.

Element activity	Outcome	Likelihood
Transposition	Gene inactivation via insertion	Common with low-copy-number elements
Transposition	Creation of new sequences within a gene that can serve as raw material for the evolution of new gene functions	Common with low-copy-number elements
Transposition	Altered gene regulation	Exceedingly common, with all elements that mainly insert in or near genes
Transposition	Creation of a new intron in a gene	Theoretically possible, but no evidence for common occurrence in current genes
Transposition	Assembly of a new gene from components carried by an element or elements	Possible origin for many (most?) genes, but no proof, although some elements do carry other gene domains
Excision	Small changes in local sequence, often legacies of direct target duplication or other small insertions/deletions	Common with some active DNA elements

these active factors [123]. Running counter to this trend are the possible negative effects that huge copy numbers of these elements within a genome might create, like possible disadvantages of a large genome size [11, 16] or high levels of mutation [60]. All of these criteria could describe the action of a parasitic or selfish DNA [28, 89]. Like most coevolved parasite-host interactions, minimization of the negative aspects of the interaction are of value both to the host and the parasite. Hence, one can view the plant nuclear genome as an ecosystem, where the mobile DNAs are commonly the most abundant feature [104].

High-copy-number transposable elements have apparently managed to avoid too much of a detrimental effect on their host genomes by integrating into primarily inactive parts of the genome. Their transmission to subsequent generations is assured by their high copy number and by their dispersal onto several chromosomes. IRPs may preferentially target other IRP LTRs as a way of inactivating the target IRP (hence, decreasing its ability to compete with them). In addition, an IRP inserted into the LTR of another IRP will acquire access to the enhancers and other promoter elements of the targeted element.

Low-copy-number transposable elements are less likely to be transmitted to the next generation by chance. In fact, if they are mutagenic in nature, then progeny segregating for loci containing or not containing a transposable element insertion will often exhibit preferential survival of individuals without the element. For a low-copy-number element, insertion into a genetically active part of the genome would provide the opportunity to retain activity, including the potential to transpose again and thereby make new copies. Elements of this type might be most effective if they target active regions without actually inserting into the relatively non-malleable components of genes (like the regions that encode the protein). MITEs, for instance, appear to have a bias toward insertion near MARs, thereby making it less likely that they would fully inactivate a gene. Having a structure similar to that of a MAR, and a high potential for evolution into a MAR, may make MITEs less detrimental even if they do insert in a way that could inactivate MAR function. Similarly, preferential insertion into a promoter by a transposable element that carries promoter functions will often lead to a gene with a largely retained function, although with some possible change in regulation.

In general, it seems appropriate to view the effects of transposable elements both from the perspective of the host and of the element. Their ability to amplify via transposition guarantees that mobile DNAs will be selected for that activity, and they should compete with other transposable element individuals, families and types for presumably finite genomic space and resources. Different plant species have different populations of these transposable elements, different exposures to activating stresses, possibly different sets (or efficiencies) of processes to inactivate and/or remove these element, and conceivably different constraints on how much the effects of these elements can be tolerated. Hence, one expects different outcomes regarding the abundance, arrangement and genetic contributions of these elements. This great variability is exactly what is observed, even in closely related

plant species [6, 11, 111]. Like any other DNA within a heritable genome, plant transposable elements can provide a substrate for selection of superior host fitness. Hence, a transposable element that consistently increases telomere length can remove the requirement for telomerase in *Drosophila* [92], segments of transposable elements can become the regulatory promoter elements of a gene [65, 121, 122] or possibly a conserved component of centromere function [87, 101]. Use and/or removal of transposable elements by a host may be a relatively slow process, however, compared to their selfish/parasitic amplification and dispersal.

Unanswered questions

In some genomes with a recent history of active transposable elements, most variation in genome structure will be due to transposable elements. In other species, transposable elements may be few, almost exclusively defective, or efficiently down-regulated. In these cases, variation in genome structure may be mainly generated by other extrinsic or intrinsic activities, such as radiation-induced chromosome breakage or the natural properties of DNA replication and repair enzymes. In complex plant genomes, if for no other reason than the abundance of these elements, it seems likely that the latter scenario will be rare. However, much further experimentation is needed to see both what genetic changes have occurred in plants, and what are the responsible factors. To date, we have investigated relatively few plant species, and most investigations have been at the level of individual gene structures. Analyses of genome organization across contiguous multigenic segments will be more valuable [15, 69, 91, 106], particularly if they are comparable across species [3, 111]. Analyses of the junction points of chromosomal rearrangements would be very informative, especially if the frequent types and locations of such rearrangements differed in closely related species. Study of the breakpoint(s) of a recent chromosomal rearrangement would be most likely to yield information on the cause(s) of the rearrangement before they are obscured by subsequent events.

A second unanswered question is the nature(s) of transposable element origins within a species. Are the very abundant elements in some plants derived from low-copy-number elements that have been within the species for a very long time, but amplified during one stage of the evolutionary history of the organism? If so, then why were only some elements amplified and not others? Or were they transmitted horizontally, perhaps as viruses or during a wide cross? An important related question is why are transposable elements much more abundant in one species than in another? Do some species have particularly good mechanisms for shutting down and/or removing some or all transposable elements? Perhaps Arabidopsis is exceptionally effective in homology-based silencing [77, 79]. Or maybe *Arabidopsis* has such high homologous recombination rates per kb of DNA that it has turned most of its LTR-retrotransposons into now unrecognized solo LTRs? Alternatively, perhaps the difference is in exposure to transposable element amplifying conditions. If many of these transposable elements came via a horizontal route, then maybe *Arabidopsis* has not lived under circumstances where wide crosses, germline insect feeding, etc. have provided elements that could amplify greatly. Although all plant species, including *Arabidopsis*, have many different families of transposable elements, only a few of these transposable element families appear to have the competence to amplify into the tens or hundreds of thousands. In general, the plants with the largest genomes tend to have LTR-retrotransposons with the highest copy numbers per genome [8, 45]. Maize, for instance, has thousands of families of LTR-retrotransposons, yet only a half dozen of these have amplified up to tens of thousands of copies per nucleus [104]. Perhaps *Arabidopsis* has not been exposed to the thousand or so families of LTR-retrotransposons that would be required to find one that can amplify to such high copy numbers and thereby create a large genome. Plants also have differed in the degree to which they are exposed to environmental stresses, and this could create different frequencies and levels of reactivation of quiescent elements into an amplifying state.

A third important question is the nature of the insertion and amplification specificities that we observe for different transposable elements. Although some low-copy-number LTR-retrotransposons can insert into genes, perhaps preferentially, the high-copy-number IRPs appear to avoid genic insertions. Yet, at a DNA sequence/structure level, we cannot distinguish between IRPs and other LTR-retrotransposons. Although it is likely that the insertion specificities are associated with recognition of particular chromatin proteins and/or structures, we can only guess what these targeting features may be. Moreover, the primary sequence of an LTR-retrotransposon does not currently indicate to the investigator whether it can

amplify to tens of thousands of copies per nucleus or not. It is not clear, in fact, whether transposable elements do differ in amplification potential, or if chance alone determines whether a particular element family will amplify to make up a large portion of a genome.

A fourth important question concerns how often and for how long these elements are active. It is likely that this question will have very different answers for different transposable elements and in different species, but we currently have no precise answer to this question in any plant species or for any element.

Because some of these elements are active even when transferred to other species by transformation, we can begin to investigate many of these questions in detail. Equally important, we need to characterize the actual ground state of current plant genomes; what do they contain, how is it arranged, and how has it changed? Only in *Arabidopsis* do we have a serious beginning to a significant understanding of the nature of a plant genome, and the unusually small size of the *Arabidopsis* genome guarantees that its structure will be at least somewhat anomalous.

Conclusions

Transposable elements constitute large portions of many plant genomes, and are potentially hyperactive in changing genes and genomes in all plants. Overall genome structure can be changed by transposable element action, including such changes as large inversions and translocations that can contribute to reproductive isolation and subsequent speciation. Individual genes are also impacted by transposable elements, particularly as a source of potential regulatory elements. Transposable elements appear to be the major determinant of genome size in at least some species. However, for all of the other possible contributions of transposable elements to genome and gene structure and evolution, we lack any comprehensive understanding of the frequency and primary types of these contributions in any plant. Additional experimentation is needed to determine exactly how transposable elements actually have contributed to the evolution of particular genomes, and plant genomes in general. Comprehensive analyses of genome structure and transgenic studies of identified elements will provide the tools for these investigations. If recent history is any indication, the results will be both surprising and tremendously informative.

Acknowledgements

I thank Alex Tikhonov and Sue Wessler for their many helpful comments and Sheri Frank for preparing Figure 1. The preparation of this paper was supported by a grant from the USDA CSREES (97-35300-4594).

References

1. Ananiev EV, Riera-Lizarazu O, Rines HW, Phillips RL: Oat-maize chromosome addition lines: a new system for mapping the maize genome. Proc Natl Acad Sci USA 94: 3524–3528 (1997).
2. Avramova Z, SanMiguel P, Georgieva E, Bennetzen JL: Matrix attachment regions and transcribed sequences within a long chromosomal continuum containing maize *adh1*. Plant Cell 7: 1667–1680 (1995).
3. Avramova Z, Tikhonov A, Chen M, Bennetzen JL: Matrix attachment regions and structural collinearity in the genomes of two grass species. Nucl Acids Res 26: 761–767 (1998).
4. Benjamin HW, Kleckner N: Intramolecular transposition by Tn*10*. Cell 59: 373–383 (1989).
5. Bennett MD: Variation in genome form in plants and its ecological implications. New Phytol 106 (Suppl): 177–200 (1987).
6. Bennett MD, Leitch IJ: Nuclear DNA amounts in angiosperms: 583 new estimates. Ann Bot 80: 169–196 (1997).
7. Bennetzen JL: The regulation of *Mutator* function and *Mu1* transposition. In: Freeling M (Ed.) Plant Genetics, pp. 343–353. Alan R. Liss, New York (1985).
8. Bennetzen JL: The contributions of retroelements to plant genome organization, function and evolution. Trends Microbiol 4: 347–353 (1993).
9. Bennetzen JL: The *Mutator* transposable element system of maize. Curr Top Microbiol Immunol 204: 195–229 (1996).
10. Bennetzen JL, Brown WE, Springer PS: The state of DNA modification within and flanking maize transposable elements. In: Nelson OE (ed), Plant Transposable Elements, pp. 237–250. Plenum, New York (1988).
11. Bennetzen JL, Kellogg EA: Do plants have a one-way ticket to genomic obesity? Plant Cell 1509–1514 (1997).
12. Bennetzen JL, Schrick K, Springer PS, Brown WE, SanMiguel P: Active maize genes are unmodified and flanked by diverse classes of modified, highly repetitive DNA. Genome 37: 565–576 (1994).
13. Bennetzen JL, Springer PS: The generation of *Mutator* transposable element subfamilies in maize. Theor Appl Genet 87: 657–667 (1994).
14. Bennetzen JL, Swanson J, Taylor WC, Freeling M: An insertion in the first intron of maize *Adh1* affects transcript levels: cloning of progenitor and mutant alleles. Proc Natl Acad Sci USA 81: 4125–4128 (1984).
15. Bevan M, Bancroft I, Bent E, Love K, Goodman H, Dean C, 62 others: Analysis of 1.9 Mb of contiguous DNA sequence from chromosome 4 of *Arabidopsis thaliana.* Nature 391: 485–488 (1998).
16. Biradar DP, Rayburn AL: Heterosis and nuclear DNA content in maize. Heredity 71: 300–304 (1993).
17. Bonas U, Sommer H, Saedler H: The 17-kb *Tam1* element of *Antirrhinum majus* induces a 3-bp duplication upon integra-

tion into the chalcon synthase gene. EMBO J 3: 1015–1019 (1984).
18. Bureau TE, Wessler SE: Mobile inverted-repeat elements of the *Tourist* family are associated with the genes of many plant genomes. Proc Natl Acad Sci USA 91: 1411–1415 (1994).
19. Bureau TE, White SE, Wessler SR: Transduction of a cellular gene by a plant retroelement. Cell 77: 479–480 (1994).
20. Chen C-H, Oishi KK, Kloeckener-Gruissem B, Freeling M: Organ-specific expression of maize *Adh1* is altered after a *Mu* transposon insertion. Genetics 117: 109–116 (1987).
21. Chen J, Greenblatt IM, Dellaporta SL: Transposition of *Ac* from the *P* locus of maize into unreplicated chromosomal sites. Genetics 117: 109–116 (1987)
22. Cone KC, Schmidt RJ, Burr B, Burr FA: Advantages and limitations of using *Spm* as a transposon tag. In: Nelson OE (ed), Plant Transposable Elements, pp. 149–159. Plenum, New York (1988).
23. Cost GJ, Boeke JD: Targeting of human retrotransposon integration is directed by the specificity of the L1 endonuclease for regions of unusual DNA structure. Biochemistry 22: 18081–18093 (1998).
24. Coulondre C, Miller JH, Farabaugh PJ, Gilbert W: Molecular basis of base substitution hotspots in *Escherichia coli.* Nature 274: 775–780 (1978).
25. Cresse AD, Hulbert SH, Brown WE, Lucas JR, Bennetzen JL: *Mu1*-related transposable elements of maize preferentially insert into low copy number DNA. Genetics 140: 315–324 (1995).
26. Danilevskaya ON, Arkhipova JR, Traverse KL, Pardue ML: Promoting in tandem: the promoter for telomere transposon *HeT*-A and implications for the evolution of retroviral LTRs. Cell 88: 647–655 (1997).
27. Dellaporta SL, Chomet PS: The action of maize controlling elements. In: Hohn B, Dennis ES (eds), Plant Gene Research: Genetic Flux in Plants, pp. 169–216, Springer-Verlag, Berlin (1985).
28. Doolittle WF, Sapienza C: Selfish genes, the phenotype paradigm and genome evolution. Nature 284: 601–603 (1980).
29. Dooner HK, Martinez-Ferez IM: Recombination occurs uniformly within the *bronze* gene, a meiotic recombination hotspot in the maize genome. Plant Cell 9: 1633–1645 (1997).
30. Doring J-P, Starlinger P: Molecular genetics of transposable elements in plants. Annu Rev Genet 20: 175–200 (1986).
31. Drouin G, Dover GA: A plant processed pseudogene. Nature 328: 557–558 (1987).
32. Edwards KJ, Veuskens J, Rawles H, Daly A, Bennetzen JL: Characterization of four dispersed repetitive DNA sequences in *Zea mays* and their use in constructing contiguous DNA fragments using YAC clones. Genome 39: 811–817 (1996).
33. Eickbush TH: Transposing without ends: the non-LTR retrotransposable elements. New Biol 4: 430–440 (1992).
34. Emerson RA: The inheritance of a recurring somatic variation in variegated ears of maize. Am Nat 48: 87–115 (1914).
35. Engels WR, Johnson-Schilz DM, Eggleston WB, Sved J: High-frequency P element loss in Drosophila is homolog dependent. Cell 62: 515–525 (1990).
36. English J, Harrison K, Jones JDG: A genetic analysis of DNA sequence requirements for *Dissociation* state-I activity in tobacco. Plant Cell 5: 501–514 (1993).
37. Fedoroff N, Wessler S, Shure M: Isolation of the transposable controlling elements *Ac* and *Ds.* Cell 35: 243–251 (1983).
38. Flavell AJ, Dunbar E, Anderson R, Pearce SR, Hartley R, Kumar A: Ty1-*copia* group retrotransposons are ubiquitous and heterogeneous in higher plants. Nucl Acids Res 20: 3639–3644 (1992).
39. Flavell AJ, Pearce SR, Kumar A: Plant transposable elements and the genome. Curr Opin Genet Dev 4: 838–844 (1994).
40. Flavell RB, Bennett MD, Smith JB, Smith DB: Genome size and proportion of repeated nucleotide sequence DNA in plants. Biochem Genet 12: 257–269 (1974).
41. Gale MD, Devos KM: Plant comparative genetics after 10 years. Science 282: 656–659 (1998).
42. Gilbert W: Introns and exons: Playgrounds of evolution. In: Axel R, Maniatis T, Fox CF (eds), Eucaryotic Gene Regulation, pp. 1–12. Academic Press, New York (1979).
43. Giroux MJ, Clancy M, Baier J, Ingham L, McCarty D, Hannah LC: De novo synthesis of an intron by the maize transposable element *Dissociation.* Proc Natl Acad Sci USA 91: 12150–12154 (1994).
44. Gross D, Garrard W: Poising chromatin for transcription. Trends Biochem Sci 12: 293–297 (1987).
45. Grandbastien M-A: Retroelements in higher plants. Trends Genet 8: 103–108 (1992).
46. Grandbastien M-A: Activations of plant retrotransposons under stress conditions. Trends Plant Sci 3: 181–187 (1998).
47. Green B, Walko R, Hake S: *Mutator* insertions in an intron of the maize *knotted-1* gene result in dominant suppressible mutations. Genetics 138: 1275–1285 (1994).
48. Greenblatt IM: A chromosome replication pattern deduced from pericarp phenotypes resulting from movements of the transposable element, Modulator, in maize. Genetics 108: 471–485 (1984).
49. Greenblatt IM, Brink RA: Twin mutations in medium variegated pericarp maize. Genetics 47: 489–501 (1962).
50. Hehl R, Nacken W, Krause A, Saedler H, Sommer H: Structural analysis of *Tam*3, a transposable element from *Antirrhinum majus*, reveals homologies to the *Ac* element from maize. Plant Mol Biol 16: 369–371 (1991).
51. Heslop-Harrison JS, Murata M, Ogura Y, Schwarzacher T, Motoyoshi F: Polymorphisms and genomic organization of repetitive DNA from centromeric regions of Arabidopsis chromosomes. Plant Cell 11: 31–42 (1999).
52. Hirochika H: Activation of tobacco retrotransposons during tissue culture. EMBO J 12: 2521–2528 (1993).
53. Hirochika H, Sugimoto K, Otsuki Y, Tsugawa H, Kanda M: Retrotransposons of rice involved in mutations induced by tissue culture. Proc Natl Acad Sci USA 93: 7783–7788 (1996).
54. Hollick JB, Dorweiler JE, Chandler VL: Paramutation and related allelic interactions. Trends Genet 13: 302–308 (1997).
55. Hu W, Das OP, Messing J: *Zeon-l*, a member of a new maize retrotransposon family. Mol Gen Genet 248: 471–480 (1995)
56. Jiang J, Nasuda S, Dong F, Scherrer CW, Woo S-S, Wing RA, Gill BS, Ward DC: A conserved repetitive DNA element located in the centromeres of cereal chromosomes. Proc Natl Acad Sci USA 93: 14210–14213 (1996).
57. Jin Y-K, Bennetzen JL: Structure and coding properties of *Bs1*, a maize retrovirus-like transposon. Proc Natl Acad Sci USA 86: 6235–6239 (1989).
58. Jin Y-K, Bennetzen JL: Integration and nonrandom mutation of a plasma membrane proton ATPase gene fragment within the *Bs1* retroelement of maize. Plant Cell 6: 1177–1186 (1994).

59. Johns MA, Mottinger J, Freeling M: A low copy number, *Copia*-like transposon in maize. EMBO J 4: 1093–1102 (1985).
60. Kidwell MG, Kidwell JF, Sved JA: Hybrid dysgenesis in *Drosophila melanogaster*: syndrome of aberrant traits including mutation, sterility, and male recombination. Genetics 86: 813–833 (1977).
61. Kim H-Y, Schiefelbein JW, Raboy V, Furtek DB, Nelson OE Jr: RNA splicing permits expression of a maize gene with a defective *Suppressor-mutator* transposable element insertion in an exon. Proc Natl Acad Sci USA 84: 5863–5867 (1987).
62. Kleckner N: Regulation of transposition in bacteria. Annu Rev Cell Biol 6: 297–327 (1990).
63. Kumar A, Bennetzen JL: Plant retrotransposons. Annu Rev Genet 33, in press.
64. Kunze R: The maize transposable element *Activator* (*Ac*). In: Saedler H, Gierl A (eds), Transposable Elements, pp. 161–194. Springer-Verlag, Berlin (1996).
65. Kunze R, Saedler H, Lonnig WE: Plant transposable elements. Adv Bot Res 27: 331–470 (1997).
66. Laten H, Majumdar A and Gaucher EA: *SIRE-1*, a *copia/Ty1*-like retroelement from soybean, encodes a retroviral envelope-like protein. Proc Natl Acad Sci USA 95: 6897–6902 (1998).
67. Levy AA, Walbot V: Molecular analysis of the loss of somatic instability in the *bz2*:*mu1* allele of maize. Mol Gen Genet 229: 147–151 (1991).
68. Lim JK, Simmons JM: Gross chromosome rearrangements mediated by transposable elements in *Drosophila melanogaster.* BioEssays 16: 269–273 (1994).
69. Llaca V, Messing J: Amplicons of maize zein genes are conserved within genic but expanded and constricted in intergenic regions. Plant J 15: 211–20 (1998).
70. Loguercio LL, Wilkins TA: Structural analysis of a hmg-coA reductase pseudogene: insights into evolutionary processes affecting the hmgr gene family in allotetraploid cotton (*Gossypium hirsutum* L.). Curr Genet 34: 241–249 (1998).
71. Maraia R: The Impact of Short Interspersed Elements (SINEs) on the Host Genome. Springer-Verlag, New York (1995).
72. Marillonnet S, Wessler SR: Retrotransposon insertion into the maize *waxy* gene results in tissue-specific RNA processing. Plant Cell 9: 967–978 (1997).
73. Martienssen RA: Epigenetic phenomena: paramutation and gene silencing in plants. Curr Biol 6: 810–813 (1996).
74. Martienssen RA, Barkan A, Taylor WC, Freeling M: Somatically heritable switches in the DNA modification of *Mu* transposable elements monitored with a suppressible mutant in maize. Genes Dev 4: 331–343 (1989).
75. Masson P, Surovsky R, Kingsbury J, Fedoroff NV: Genetic and molecular analysis of the *Spm*-dependent *a-m2* alleles of the maize *a* locus. Genetics 177: 117–137 (1987).
76. Matsuoka Y, Tsunewaki K: Evolutionary dynamics of Ty1-*copia* group retrotransposons in grasses shown by reverse transcriptase domain analysis. Mol Biol Evol 16: 208–217 (1999).
77. Matzke MA, Matzke AJ: Epigenetic silencing of plant transgenes as a consequence of diverse cellular defence responses. Cell Mol Life Sci 54: 94–103 (1998).
78. Matzke MA, Matzke AJ, Eggleston W: Transgene silencing and paramutation: a common response to invasive DNA? Trends Plant Sci 1: 382–388 (1996).
79. Matzke MA, Primig M, Trnovsky J, Matzke AJ: Reversible methylation and inactivation of marker genes in sequentially transformed tobacco plants. EMBO J 8: 643–649 (1989).
80. McClintock B: The fusion of broken ends of chromosomes following nuclear fusion. Proc Natl Acad Sci USA 28: 458–463 (1942).
81. McClintock B: Maize genetics. Carnegie Inst Washington Year Book 45: 176–186 (1946).
82. McClintock B: Mutable loci in maize. Carnegie Inst Washington Year Book 47: 155–169 (1948)
83. McClintock B: Mutable loci in maize. Carnegie Inst Washington Year Book 48: 142–154 (1949).
84. McClintock B: Mutable loci in maize. Carnegie Inst Washington Year Book 49: 157–167 (1950).
85. McClintock B: Topographical relations between elements of control systems in maize. Carnegie Inst Washington Year Book 61: 448–461 (1962).
86. McClintock B: The significance of responses of the genome to challenge. Science 226: 792–801 (1984).
87. Miller JT, Dong F, Jackson SA, Song J, Jiang J: Retrotransposon-related DNA sequences in the centromeres of grass chromosomes. Genetics 150: 1615–1623 (1998).
88. O'Neill RJW, O'Neill MJ, Graves JAM: Undermethylation associated with retroelement activation and chromosome remodeling in an interspecific mammalian hybrid. Nature 393: 68–72 (1998).
89. Orgel LE, Crick FHC: Selfish DNA: the ultimate parasite. Nature 284: 604–607 (1980).
90. Palmgren MG: Capturing of host DNA by a plant retroelement: *Bs1* encodes plasma membrane H^+-ATPase domains. Plant Mol Biol 25: 137–140 (1994).
91. Panstruga R, Buschges R, Piffanelli P, Schulze-Lefert P: A contiguous 60 kb genomic stretch from barley reveals molecular evidence for gene islands in a monocot genome. Nucl Acids Res 26: 1056–1062 (1998).
92. Pardue ML, Danilevskaya ON, Traverse KL, Lowenhaupt K: Evolutionary links between telomeres and transposable elements. Genetica 100: 73–84 (1997).
93. Peacock WJ, Dennis ES, Gerlach WL, Sachs MM, Schwartz D: Insertion and excision of *Ds* controlling elements in maize. Cold Spring Harbor Symp Quant Biol 49: 347–354 (1984).
94. Pearce SR, Pich U, Harrison G, Flavell AJ, Heslop-Harrison JS, Schubert I, Kumar A: The Ty1-*copia* group retrotransposons of *Allium cepa* are distributed throughout the chromosomes but are enriched in the terminal heterochromatin. Chrom Res 4: 357–364 (1996).
95. Pelissier T, Tutois S, Deragon JM, Tourmente S, Genestier S, Picard G: *Athila*, a new retroelement from *Arabidopsis thaliana.* Plant Mol Biol 29: 441–452 (1995).
96. Pelissier T, Tutois S, Tourmente S, Deragon JM, Picard G: DNA regions flanking the major *Arabidopsis thaliana* satellite are principally enriched in *Athila* retroelement sequences. Genetica 97: 141–151 (1996).
97. Peschke VM, Phillips RL, Gengenbach BG: Discovery of transposable element activity among progeny of tissue culture-derived maize plants. Science 238: 804–807 (1987).
98. Plasterk RHA, Groenen TM: Targeted alterations of the *Caenorhabditis elegans* genome by transgene instructed DNA double-strand break repair following Tc1 excision. EMBO J 11: 287–290 ((1992).
99. Pouteau S, Grandbastien M-A, Boccara M: Microbial elicitors of plant defence responses activate transcription of a retrotransposon. Plant J 5: 535–542 (1994).

100. Pouteau S, Spielmann A, Meyer C, Grandbastien M-A, Caboche M: Effects of *Tnt1* tobacco retrotransposon insertion on target gene transcription. Mol Gen Genet 228: 233–239 (1991).
101. Presting GG, Malysheva L, Fuchs J, Schubert I: A *TY3/GYPSY* retrotransposon-like sequence localises to the centromeric region of cereal chromosomes. Plant J 16: 721–728 (1998).
102. Price HJ: Nuclear DNA content variation within angiosperm species. Evol Trends Plants 2: 53–60 (1988).
103. Richards EJ, Ausubel FM: Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. Cell 53: 127–136 (1988).
104. SanMiguel P, Bennetzen JL: Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. Ann Bot 82: 37–44 (1998).
105. SanMiguel P, Gaut BS, Tikhonov A, Nakajima Y, Bennetzen JL: The paleontology of intergene retrotransposons of maize. Nature Genet 20: 43–45 (1998).
106. SanMiguel P, Tikhonov A, Jin YK, Motchoulskaia N, Zakharov D, Melake-Berhan A, Springer PS, Edwards KJ, Lee M, Avramova Z, Bennetzen, JL: Nested retrotransposons in the intergenic regions of the maize genome. Science 274: 765–768 (1996).
107. Schmid CW: Alu: Structure, origin, evolution, significance and function of one tenth of human DNA. Prog Nucl Acid Res Mol Biol 53: 283–319 (1996).
108. Schmidt T, Kubis S, Heslop-Harrison JS: Analysis and chromosomal location of retrotransposons in sugar beet (*Beta vulgaris*): LINEs and Ty1-*copia*-like elements as major components of the genome. Chrom Res 3: 335–345 (1995).
109. Sturtevant AH, Morgan TH: Reverse mutation of the bar gene correlated with crossing over. Science 57: 746–747 (1923).
110. Talbert LE, Chandler, VL: Characterization of a highly conserved sequence related to *Mutator* transposable elements in maize. Mol Biol Evol 5: 519–529 (1988).
111. Tikhonov AP, SanMiguel PJ, Nakajima Y, Gorenstein ND, Bennetzen JL, Avramova, Z: Collinearity and its exceptions in orthologous *adh* regions of maize and sorghum. Proc Natl Acad Sci. USA 96: 7409–7414 (1999).
112. Tsubota SI, Rosenberg D, Szostak H, Rubin D Schedl P: The cloning of the *Bar* region and the *B* breakpoint in *Drosophila melanogaster*: evidence for a transposon-induced rearrangement. Genetics 122: 881–890 (1989).
113. Turcich MP, Bokharri-Riza A, Hamilton DA, He CP, Messier W, Stewart CB, Mascarenhas JP: PREM-2, a *copia*-type retroelement in maize is expressed preferentially in early microspores. Sexual Plant Reprod 9: 65–74 (1996).
114. Vicient CM, Suoniemi A, Anamthawat-Jonsson K, Tanskanen J, Beharav A, Nevo E, Schulman AH: Retrotransposon *BARE*-1 and its role in genome evolution in *Hordeum*. Plant Cell, in press (1999).
115. Voytas DF, Cummings MP, Konieczny A, Ausubel FM, Rodermel SR: *Copia*-like retrotransposons are ubiquitous among plants. Proc Natl Acad Sci USA 89: 7124–7128 (1992).
116. Walbot V: Reactivation of the *Mutator* transposable element system following gamma irradiation of seed. Mol Gen Genet 212: 259–264 (1988).
117. Walbot V, Chandler V, Taylor L: Alterations in the *Mutator* transposable element family of *Zea mays*. In: Freeling M (ed), Plant Genetics, pp. 333–342. Alan R. Liss, New York (1985).
118. Weil CF, Wessler SR: Molecular evidence that chromosome breakage by *Ds* elements is caused by aberrant transposition. Plant Cell 5: 512–522 (1993).
119. Wessler SR: The splicing of transposable elements and its role in intron evolution. Genetica 86: 295–305 (1992).
120. Wessler SR, Baran G, Varagona M: The maize transposable element *Ds* is spliced from RNA. Science 237: 916–918 (1987).
121. Wessler SR, Bureau TE, White SE: LTR-retrotransposons and MITEs: important players in the evolution of plant genomes. Curr Opin Genet Dev 5: 814–821 (1995).
122. White SE, Habera LF, Wessler SR: Retrotransposons in the flanking regions of normal plant genes: a role of *copia*-like elements in the evolution of gene structure and expression. Proc Natl Acad Sci USA 91: 11792–11796 (1994).
123. Wilke CM, Adams J: Fitness effects of Ty transposition in *Saccharomyces cerevisiae*. Genetics 131: 31–42 (1992).
124. Williamson VM: Transposable elements in yeast. Int Rev Cytol 83: 1–25 (1983).
125. Wright DA, Voytas DF: Potential retroviruses in plants: Tat*1* is related to a group of *Arabidopsis thaliana* Ty3/*gypsy* retrotransposons that encode envelope-like proteins. Genetics 149: 603–715 (1998).
126. Xiong Y, Eickbush TH: Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J 9: 3353–3362 (1990).
127. Xu X, Hsia A-P, Zhang L, Nikolau BJ, Schnable PS: Meiotic recombination break points resolve at high rates at the 5′ end of a maize coding sequence. Plant Cell 7: 2151–2161 (1995).
128. Yu H-G, Hiatt EN, Chan A, Sweeney M, Dawe RK: Neocentromere-mediated chromosome movement in maize. J Cell Biol 139: 831–840 (1997).
129. Zou S, Voytas DF: Silent chromatin determines target preference of the *Saccharomyces* retrotransposon Ty5. Proc Natl Acad Sci USA 94: 7412–7416 (1997).