EVOLUTION

INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

PUBLISHED BY

THE SOCIETY FOR THE STUDY OF EVOLUTION

Vol. 55 **January 2001** **No. 1**

Evolution, 55(1), 2001, pp. 1–24

PERSPECTIVE: TRANSPOSABLE ELEMENTS, PARASITIC DNA, AND GENOME EVOLUTION

MARGARET G. KIDWELL[1] AND DAMON R. LISCH[2]

[1]*Department of Ecology and Evolutionary Biology, The University of Arizona, Tucson, Arizona 85721*
E-mail: kidwell@azstarnet.com

[2]*Department of Plant and Microbial Biology, University of California, Berkeley, California 94720*

Abstract.—The nature of the role played by mobile elements in host genome evolution is reassessed considering numerous recent developments in many areas of biology. It is argued that easy popular appellations such as ''selfish DNA'' and ''junk DNA'' may be either inaccurate or misleading and that a more enlightened view of the transposable element-host relationship encompasses a continuum from extreme parasitism to mutualism. Transposable elements are potent, broad spectrum, endogenous mutators that are subject to the influence of chance as well as selection at several levels of biological organization. Of particular interest are transposable element traits that early evolve neutrally at the host level but at a later stage of evolution are co-opted for new host functions.

Key words.—Gene regulation, genome evolution, parasitic DNA, retrotransposon, transposon.

Received July 20, 1999. Accepted July 12, 2000.

During the last thirty years, numerous families of transposable elements (TEs) have been discovered and characterized in all species that have been adequately examined. Their ubiquity raises a number of questions concerning the nature of the relationships between TEs and their hosts and the significance and long-term consequences of these elements on the evolution of host genomes. The existence of these units of mobile DNA was largely unforeseen by classical geneticists during the first half of the 20th century. Divergent opinions about the relevance of TEs for the evolution of their hosts have been expressed over the years since their discovery in the mid-1940s. For example, a number of authors have stressed the actual or potential benefits of TEs to their hosts, but have often not balanced this with a description of their other properties. In contrast, the selfish and junk DNA concepts have often been accepted blindly and rigidly to the exclusion of other host-element relationships. Although the concept of TEs as molecular parasites has been useful for understanding the nature of host-element relationships, this approach needs to be expanded to consider a possible continuum of interactions between elements and their hosts, ranging from parasitism at one extreme, to mutualism or symbiosis at the other (Brosius 1999a; Kidwell and Lisch 2000). We will argue that host-element relationships are so dynamic and intimate that any rigid distinction between the two is meaningless. The purpose of this article is to integrate a number of points of view to develop a balanced perspective of the evolution of eukaryotic TEs with that of their hosts. New experimental evidence bearing on these issues will also be described and discussed.

HISTORICAL PERSPECTIVE

The relevance of TEs to the biology of their hosts was of considerable interest to Barbara McClintock, who discovered these elements in maize in the 1940s. Indeed she chose the name ''controlling elements'' because of the ability of these elements ''to regulate gene expression in precise ways'' (McClintock 1956, 1984). This perspective was consistent with the prevailing assumption of the time that all, or most, characteristics of organisms were adaptive and thus were subject to positive natural selection. Later there was a growing awareness that a large fraction of many genomes was made up of noncoding sequences consisting largely of highly repetitive and middle repetitive DNA. TEs were shown to constitute a physically significant component of the middle repetitive DNA fraction in many eukaryotic genomes. The characteristics of these sequences made it difficult to determine an obvious function for the organism. The publication of Dawkins' *The Selfish Gene* (Dawkins 1974) and the development of the neutral theory of molecular evolution (Kimura 1983) provided a backdrop for serious questioning of the

assumptions underlying the idea that mobile elements play a largely beneficial role in the evolution of their hosts.

In their classic papers, Doolittle and Sapienza (1980) and Orgel and Crick (1980) pointed out that the ubiquity of repetitive sequences could be best explained by mechanisms independent of a direct selectable benefit to the host. The replicative advantage of TEs could explain their increase in frequency and maintenance in natural populations (natural selection at the level of DNA) without invoking a positive selective advantage at the level of the individual organism. Following the lead of Dawkins, the term ''selfish DNA'' was often used to describe the ability of mobile elements to amplify in host genomes—a critical stage in their life style. The word ''selfish'' is defined by the *Oxford English Dictionary* to mean ''Devotion to, or concern for, one's own advantage, or welfare, to the exclusion of regard for others.'' This clearly anthropomorphic term had value in capturing a property essential for the propagation of TEs. Unfortunately, the connotation of exclusivity of interest that is implied in the definition, and emphasized by the original proponents, has often lead to disregard of other characteristics of TEs that may provide short and long-term benefits to hosts.

The earlier, mistaken, notion that selfish DNA has ''little or no phenotypic effects'' (Dawkins 1974; Orgel and Crick 1980) has still not been totally abandoned, despite overwhelming evidence to the contrary. Similar to most other mutator mechanisms, TEs produce phenotypic effects that are deleterious to the fitness of the host organism and most of these will be removed from the host population over time by purifying natural selection. This idea was incorporated by Hickey (1982) who proposed a quantitative population genetics model for the evolution of transposable genetic elements. This model demonstrated that ''selfish'' DNA sequences do not necessarily have to be selectively neutral at the organismal level; indeed, a significant point was that such DNA can produce major deleterious effects in the host organism and still spread through the population. Charlesworth et al. (1994) stressed the ability of ''selfish DNA'' to replicate within the genome, but also pointed out that the mutations induced can cause genetic diseases and confer significant fitness losses on the organism. The extent of fitness loss has been estimated quantitatively for *P* elements in *Drosophila*, (e.g., Fitzpatrick and Sved 1986; Mackay 1986, 1989; Mackay et al. 1992; Currie et al. 1998) and found to be significant.

The demonstration that TEs can spread in host genomes despite a significant fitness loss has been explained by several authors as an example of natural selection acting in opposite directions at different levels of the biological hierarchy (e.g., Gould 1983; Kiyasu and Kidwell 1984; Good et al. 1989). At the most fundamental level of natural selection, that of DNA, the replicative advantage of TE sequences over nonmobile sequences allows their proliferation in genomes. This proliferation may be offset by negative selection, at the level of the individual host organism, in so far as organismal fitness is reduced by new, deleterious, TE-induced mutations. However, selection for reduced transposition, at the level of the host, is a weak force compared to selection for increased transposition, at the level of the transposon (Charlesworth and Langley 1986).

Besides producing neutral and deleterious effects, however, TEs can also produce mutations and genomic changes that make significant contributions to the evolution of their hosts. In fact, there is increasing evidence that the activity of TEs can influence host phenotypes in many ways: some of their effects are immediate, direct, and obvious; others are very subtle and may take long periods of evolutionary time to fully manifest. A dramatic example of the former kind is a TE-induced mutation affecting a fitness trait that was recently reported in *D. melanogaster* (Lin et al. 1998). The life span of a mutant line of flies was extended approximately 35% over nonmutant lines. The mutant line also showed increased resistance to starvation, high temperature and other stress factors. This phenotype resulted from a *P*-element insertion in the third intron of the *methuselah* (*mth*) gene. A more detailed examination of the evidence for more subtle TE-induced phenotypes is given below, following a brief introduction to the biology of TEs.

Nature and Distribution of TEs

TEs are DNA sequences that have the capacity to change genomic locations. This capacity may, or may not, be intrinsic to the element itself. Those elements that encode enzymes required for their own movement, or transposition, are described as autonomous. In contrast, nonautonomous elements are dependent on autonomous elements within the same or a different gene family for their movement. The ''life cycle'' of a TE in its host organism can be visualized as having three main phases (Fig. 1). Initially there is a rapid invasion phase in which amplification of copy number occurs accompanied by mutation which renders some elements inactive, followed by a stage of maturity of variable length, in which amplification and loss of copies are roughly in balance. Finally, a stage of senescence extends for possibly millions of years. At this stage, all autonomous elements have been lost, no amplification occurs, and nonautonomous sequences are either lost, become deleted, or otherwise diverge. Autonomous elements are represented by continuous lines and nonautonomous elements by dashed lines. Like other molecular parasites, many TEs are subject to high mutation rates (Nee and Maynard Smith 1990), and the frequency of nonautonomous elements is often high. Additional properties shared by most TEs include (1) variability in copy number; (2) insertion site polymorphism; and (3) the duplication of a few base pairs of host DNA at each end of every new element insertion site.

TEs are divided into two main classes according to their structural organization and mechanism of transposition (Finnegan 1989; Capy et al. 1997a). As members of the larger family of retroid agents that also includes retroviruses (McClure 1999), Class I elements use an RNA-mediated mode of transposition and encode a reverse transcriptase (RT). There are three distinct subclasses of eukaryotic RT-encoding TEs, the retrotransposons, the retroposons, and the retrointrons (McClure 1999). The retroposons include the short interspersed nuclear elements (SINES) and the long interspersed nuclear elements (LINES). The Class II elements, the transposons (*sensu strictu*) use a DNA-mediated mode of transposition. In contrast to exogenous agents such as retroviruses that are transmitted both vertically and horizontally,

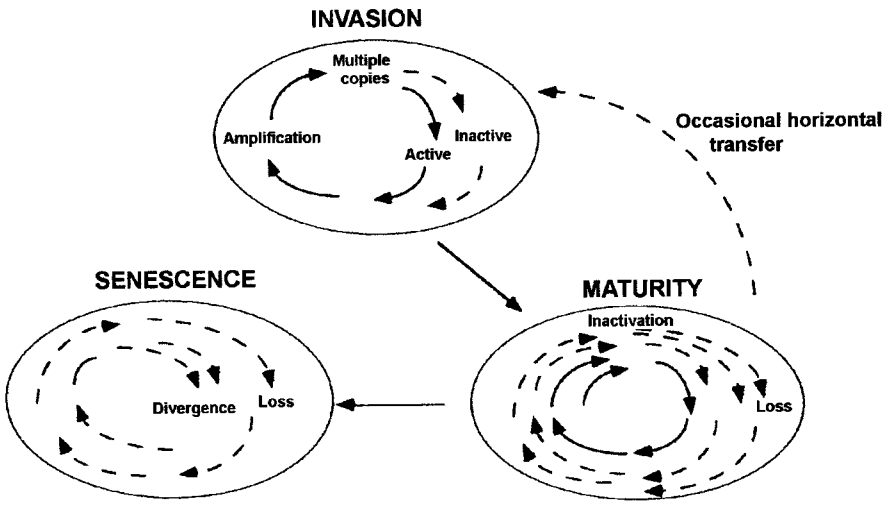

FIG. 1. General features of the life cycle of a Class II transposable element. Three main stages are illustrated: a rapid invasion phase in which amplification of copy number occurs accompanied by mutation which renders some elements inactive; a stage of maturity of variable length, when amplification and loss are roughly in balance; and a stage of senescence, possibly extending for millions of years, when all autonomous elements have been lost, no amplification occurs and nonautonomous sequences are lost or diverge. Autonomous elements are represented by continuous lines and nonautonomous elements by dashed lines.

endogenous TEs are mainly transmitted vertically with occasional exceptions (e.g., Robertson and Lampe 1995).

TEs are present in copy numbers ranging from just a few elements to tens, or hundreds, of thousands per genome. In the latter case they can represent a major fraction of the genome, especially in some plants. For example, TEs have recently been estimated to make up more than 50% of the maize genome (SanMiguel et al. 1996). The genome of *Homo sapiens* contains more than a million copies of the *Alu* (SINE) family of elements (Smit and Riggs 1996), and more than 100,000 copies of the *L1* (LINE) element (Kazazian and Moran 1998). Identifiable TEs make up approximately 35% of the genome of our own species (Bestor 1998), and possibly as much as 50%, if diverged and degraded TEs are included. It is becoming increasingly recognized that the large fraction of human and other genomes consisting of TEs may be a manifestation of the evolutionary benefits of genomic flexibility (Henikoff et al. 1997).

Because of their abundance, and the irreversible and independent nature of their insertions, some TEs, such as SINES, provide very useful markers for phylogeny reconstruction (Shedlock and Okada 2000). For example, distinct families of *Alu* sequences have amplified so recently in human evolution that they display significant amounts of interpopulational differences (Batzer et al. 1996). Polymorphic *Alu* insertions can therefore be very useful for tracing the evolutionary history of individual populations.

Origin and Evolutionary History

The appearance of a new TE family within a species can be explained in one of two different ways. TEs can originate spontaneously from non-transposable sequences as shown with human *Alu* elements (Quentin 1992). Alternatively, TEs can be transferred horizontally between species (Daniels et al. 1990; Robertson and Lampe 1995). Although the frequency of horizontal transfer is low, it seems to be higher for TEs than for conventional genes (Kidwell 1993; Capy et al. 1994). This makes sense because the intrinsic properties of TEs that enable them to move within genomes also occasionally allow them to transfer between individuals of different species. The specific mechanisms responsible for horizontal transfer of TEs remain elusive.

Although the de novo appearance of TEs is possible, by far the majority of individuals inherit their TEs by vertical transmission from their parents. At least some domains of most TE lineages appear to have ancient common origins (Capy et al. 1997b). Their molecular phylogenies can be reconstructed from sequence data in a manner similar to those of conventional genes. However, the phylogenies of TEs are often incongruent with phylogenies based on nonmobile genes from the same host species. Lack of congruence is most often due to horizontal transfer, species hybridization, or ancient polymorphisms. These explanations are not necessarily mutually exclusive. In addition, phylogenies based on different regions, or domains, of the same elements may produce incongruent phylogenies due to occasional xenologous recombination (the replacement of a resident gene by a homologous foreign gene). For example, phylogenies based on different domains found within some retrotransposons are sometimes not congruent with one another (McClure 1991). Thus the evolution of TEs is most accurately traced though the phylogenies of individual domains rather than through the phylogenies of whole elements. For example, the RT domain of Class I elements may have originated from group II introns, after migrating from organelles to the nucleus of early eukaryotes (Cavalier-Smith 1991; Capy et al. 1997b).

Although it is premature to understand the ultimate origins of TEs with a strong degree of confidence, it appears that their evolution has occurred predominantly by the serial addition of domains. However, the reverse process, deletion or loss of domains, also appears to have occurred in some lineages (Capy et al. 1997a). Both integrase/transposase and RT

domains seem likely to have evolved from bacteria, and have been assembled in eukaryotes, leading to retrotransposons and then to retroviruses by the acquisition of an envelope gene (Capy et al. 1997b). Therefore it seems likely that some, but not all eukaryotic TEs or parts of TEs, have an ancient bacterial origin. However, TEs of prokaryotes and eukaryotes have a very different evolutionary biology (Nee and Maynard Smith 1990), and this essay will mainly be restricted to consideration of TEs in eukaryotes.

Sources of New Genetic Variation

Major Roles in Early Evolution?

Questions concerning the origin and early evolution of TEs may never be unambiguously resolved. It is possible, for example, that Retroid agents are the descendants of the agents responsible for conversion of RNA to DNA in the ancient world (McClure 1999). Various authors have also speculated that TEs have played important roles during the evolution of life as we know it today, as illustrated in the following examples:

(1) At an early stage of biological evolution, it has been proposed that small modules, or duplication units, that constituted the first exons may have been the ultimate building blocks for protein assembly and that these modules were highly mobile (Dwyer 1998). Transposition and splicing of modules was further postulated to lead to increasingly complex proteins, as well as multigene families of proteins.

(2) It has been conjectured that parasitic elements, such as transposons and introns, might have provided the genes responsible for the establishment of the sexual cycle in eukaryotic organisms (Hickey 1982; Hickey and Rose 1988; Bell 1993).

(3) TEs may have played a critical role in two major macroevolutionary events: the origin of eukaryotes and the origin of vertebrates (McDonald 1998). The rationale for this proposition is that TEs were the driving force for the evolution of the two major epigenetic mechanisms, chromatin formation and methylation, that were prerequisites for the quantum expansions in gene number that accompanied the emergence of the eukaryotes and vertebrates, respectively (Bird 1995).

Broad-Spectrum Mutator Mechanisms

It is becoming apparent that TEs likely play an especially useful role as mutators in evolution because of the broad spectrum of mutations produced by their activity. The genetic changes induced by TE activity range from modifications in the size and arrangement of whole genomes to substitutions, deletions, and insertions of a single nucleotide. TEs, like other mutators, can produce major effects on phenotypic traits or small silent changes that are detectable only at the DNA sequence level. Furthermore, the spectrum of TE-induced mutations is arguably broader than that produced by any other mutator mechanism. It is important to note that TEs produce their mutagenic effects not simply on initial insertion into host DNA. They may also produce mutations when they excise imprecisely, leaving either no identifying sequence, or only small ''footprints'' of their previous presence. In addition, some TE-induced mutations that are of evolutionary significance to their hosts, such as mutations in regulatory sequences (Britten 1996a), may take long periods of time to evolve these new beneficial functions. Consequently, they may have lost their original identification as TEs. For these reasons, the properties of TE sequences in the genomes of contemporary species of animals and plants may not adequately reflect the properties of elements that have had a long association with their host species.

It has recently become apparent that a hierarchy of mutational events is required for the rapid generation of protein diversity in evolution (Bogarad and Deem 1999). Although simple DNA base substitutions are well suited for the generation, diversification, and optimization of local protein space (Maeshiro and Kimura 1998), a hierarchy of natural mutational events is required for the rapid generation of protein diversity. Experimental evolution (Moore et al. 1997; Zhang et al. 1997; Crameri et al. 1998) and computer simulation experiments (Bogarad and Deem 1999) showed that sequence shuffling has the potential to improve protein function significantly better than does point mutation alone. Because they are uniquely competent to reshuffle DNA sequences, TEs and viruses are important generators of the more complex types of mutations in the mutational hierarchy.

Another interesting aspect of the mutator role of TEs in plant evolution is their relative silence during normal development and their activation by stresses, including wounding, pathogen attack, and cell culture (Wessler 1996). Quiescence during normal development can be seen as a mechanism to ensure against massive and potentially catastrophic expression that would lower the fitness of both the host and the element. An increase in transposition rate following stress could have evolved as a way to guarantee that at least one offspring of the stressed individual would carry a transpositionally competent TE, but the effect of the consequent spike in the mutation rate of the host could be advantageous to the host's offspring by providing new variation at a time when it would be likely to be needed.

Additionally, the evolution of conditionally expressed TEs may have other benefits to the host as well. The tissue-specific expression of *B1* and *B2* SINE RNAs of mice in vivo, together with evidence for regulated changes in their expression by physiological stresses is highly suggestive that these elements perform a function important to their hosts (Li et al. 1999).

The Role of TEs in Restructuring Genomes

The ability of TEs to induce chromosomal mutations has been appreciated for some time. Indeed, McClintock saw genome restructuring mediated by TE activity as an essential component of the hosts' response to stress (McClintock 1978). In addition to movement and insertion of sequences of more than a few nucleotides, either spontaneously or by transduction, transposon activity can also result in large scale changes in host genome. A number of ways that TEs influence genome structure are now briefly described.

Insertion and excision.—The insertion of TEs can disrupt or alter gene function in a number of ways. TEs that insert within the exons of genes are most likely to result in null mutations because of the sensitivity of these regions to frame shift mutations and the lack of tolerance of highly conserved

regions to mutations of any kind. However, those mutations that survive the selection process can provide interesting and sometimes spectacular phenotypic variability (Kidwell and Lisch 1997). A large number of TE insertions are fixed within some mammalian species, leaving open the possibility that a subset of insertions has been subject to positive selection (Deininger 1989).

In addition to their role in insertional mutagenesis, TEs may generate evolutionarily significant genetic variation through the excision process (reviewed by Wessler 1988). This is because excision of a TE from a host insertion site is often not precise. The excision process may result in either the addition of new sequences or deletion of host sequences flanking the insertion site (Schiefelbein et al. 1988). A particularly interesting recent example of sequence addition with potentially important evolutionary implications is provided by the *Ds*-like transposon *Ascot-1* in the fungus *Ascobolus immersus* (Colot et al. 1998). A study of excision products of this element in a spore-color gene revealed small palindromic nucleotide additions in many of them. A number of parallels are suggested between the excision reaction of these and other *hAT* transposons, on the one hand, and V(D)J recombination, on the other. However, whereas the latter is a somatic process with little direct evolutionary consequence, the addition of small palindromic additions in the excision of *hAT* transposons can occur during meiosis with the potential to contribute significantly to germline variability and the evolution of phenotypic traits.

Ectopic recombination.—Multigene families are susceptible to nonhomologous or ectopic recombination in which genetic exchange takes place, either intrachromosomally or interchromosomally, between two members of the family located in different chromosomal regions. Such events can lead to duplications, deficiencies, and new linkage relationships, often with consequent fitness reduction. This can potentially lead to selection against increases in copy number and is considered by some (e.g., Charlesworth et al. 1997) to be a more important factor in containing TE copy number than selection against insertion mutations (Biemont et al. 1997).

However, some products of nonhomologous exchange do survive and have been documented to affect genome organization. For example, evidence was provided by Kim et al. (1998) that *Ty* retrotransposons are important in the formation of chromosomal rearrangements. They conducted a genome-wide survey of *Saccharomyces cerevisiae* retrotransposons and identified a total of 331 insertions of members of five TE families. There was evidence for recombination near many of these elements, and 5′- and 3′-flanking sequences were often duplicated and rearranged among multiple chromosomes. This suggests that recombination between retrotransposons has occurred relatively frequently and has had an important influence on both genome organization (Kim et al. 1998) and the evolution of the *Ty* family of elements itself (Jordan and McDonald 1998).

The multiplicity of SINES and LINES in the human genome, described above, provides a good substrate for unequal recombination. Thirty three cases of germline genetic diseases and 16 cases of cancer were reported to be caused by unequal homologous exchange between *Alu* elements, representing a significant fraction of human diseases (0.3%) (Deininger and Batzer 1999). Less frequent, but still important is unequal recombination between LINE elements in humans. For example a 7.5 kb deletion of genomic DNA results from unequal exchange between two neighboring LINE-1 sequences, giving rise to glycogen storage disease through phosphorylase kinase deficiency (Burwinkel and Kilimann 1998). Unequal exchange among LINE elements also gives rise to Alport syndrome and associated diffuse leiomyomatosis, a syndrome of disseminated smooth-muscle tumors involving the esophagus, large airways, and female reproductive tract (Segal et al. 1999).

Translocations and inversions.—Besides insertions and deletions, the activity of TEs is known to be associated with the production of a number of chromosomal structural changes, such as translocations and inversions. For example, many break points in hybrid dysgenesis-induced chromosomal rearrangements were shown to occur at, or very near to, the sites of *P* and *hobo* element insertions (Engels and Preston 1984; Lim 1988). In a similar way, a significant fraction of break points of newly induced chromosomal rearrangement induced by hybrid dysgenesis in *D. virilis* were found to contain *Penelope* and *Ulysses* TEs (Evgen'ev et al. unpubl. ms.). Furthermore, a TE, named *Odysseus*, was recently discovered at the distal breakpoint of a polymorphic inversion in *Anopheles arabiensis* (Mathiopoulos et al. 1998). There are several possible mechanisms by which a chromosomal inversion might be formed, including ectopic recombination between two inverted TE copies, or breakage at or near the site of insertion of two elements in the same orientation, with subsequent inversion of the intervening sequences. TEs are more likely to be found associated with endemic inversions than cosmopolitan ones, because of the tendency of TEs to be lost over time from inversion break points (Engels and Preston 1984; Mathiopoulos et al. 1998).

Pseudogene formation and genome duplication.—In addition to the mutagenic effects of insertion and excision of TEs themselves, these elements may play a significant role in genome reorganization by promoting the insertion of other gene sequences and the formation of pseudogenes (Brosius 1999). For example, Esnault et al. (2000) have recently demonstrated the ability of human LINEs to mobilize transcribed DNA not associated with a LINE sequence. The formation of processed pseudogenes by LINE-mediated retroposition may thus be an important mechanism for genome duplication and genome remodeling. There is evidence that among the retroelements, pseudogene formation may be specific to LINEs (Esnault et al. 2000).

Transposon transduction and exon shuffling.—Some TEs can provide a vehicle for the mobilization of flanking sequences accompanying aberrant transposition events. Gene sequences, such as exons or promoters, can be transduced and inserted into other existing genes and may provide a general mechanism for the evolution of new genes. An elegant demonstration of ''exon retroshuffling,'' that is, exon shuffling mediated by the human *L1* element, has been reported in tissue culture cells (Moran et al. 1999). They showed that *L1* can efficiently comobilize a 3′ flanking sequence of non-*L1* DNA to new genomic locations. The transduction of downstream sequences by endogenous *L1* elements were also earlier documented in humans, (e.g., Holmes

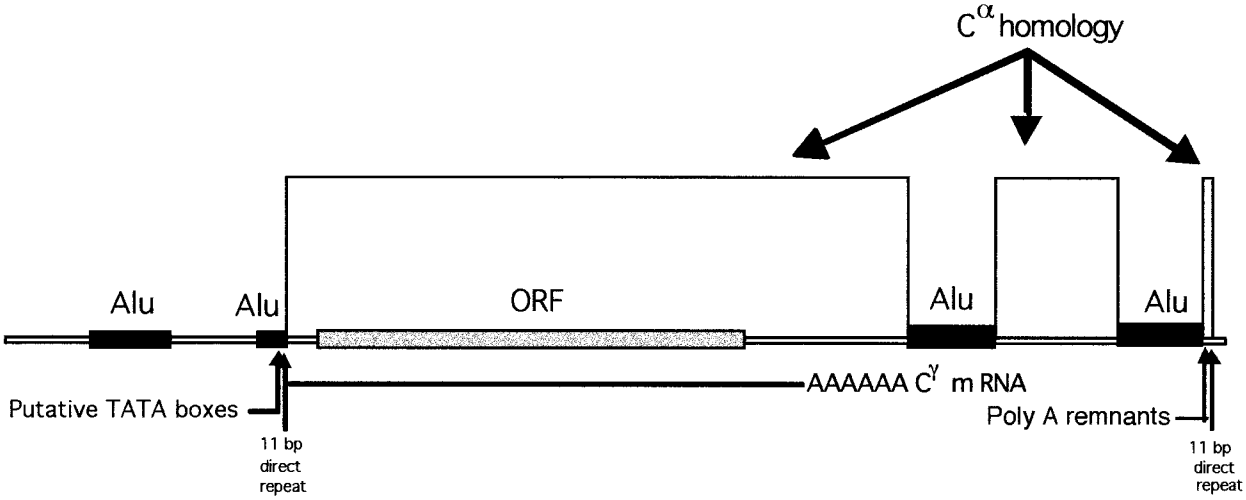

FIG. 2. Acquisition of a new function by a retrotransposed pseudogene. Schematic representation of the retrotransposed gene encoding the C gamma catalytic subunit of cAMP-dependent protein kinase (*C*γ). Open reading frame of *C*γ gene, *Alu* element insertions, Poly A remnants, 11 bp direct repeats, and putative TATA box are all as indicated. The new hybrid gene contains components of both transposon and host genes. Importantly, the regulation of expression of this gene is likely to be mediated by flanking transposon sequences (from Reinton et al. 1998).

et al. 1994; McNaughton et al. 1997). The potential significance of transposon-mediated transduction for host evolution has recently been highlighted, using a bioinformatics approach (Pickeral et al. 2000). Computational analysis of human genome sequences indicated that at least 10% of 129 full-length active *L1* elements have an associated putative 3′-transduced segment. Extrapolating this result to the whole genome, at least 1% of the human genome may represent sequences transduced by *L1* elements (Pickeral et al. 2000).

A specialized transducing transposon *Tpn1*, has been recognized in the Japanese morning glory (Takahashi et al. 1999). *Tpn1* is a Class II element that is found in the second intron of the *Dfr-B* gene for flower pigmentation, in the mutable line ''flecked'' that shows variegation for flower color. In this location the element has apparently captured a genomic DNA segment containing at least four host exon sequences. Among the *Mu* elements of maize, the nonautonomous element *Mu1* is homologous to a host gene, *Mrs-A* (*Mu*-related sequence) (Talbert and Chandler 1988). More recent database searches, using other nonautonomous *Mu* elements, reveal that all of them appear to contain transduced host sequences (D. R. Lisch, unpubl. results).

The presence of large amounts of RT has also facilitated the production of large numbers of intronless pseudogenes in many organisms. In some cases, these retrotransposed genes have acquired new functions, either as a direct consequence of mutation during retrotransposition, and/or because of the acquisition of new regulatory functions at their new position. A recent example is provided by *Cg*, the gene encoding the *C*γ catalytic subunit of cAMP-dependent protein kinase in humans (Reinton et al. 1998). This gene, which is expressed specifically in the testis, is an apparently functional retrotransposed portion of the homologous *C*α gene. The *C*γ gene lacks introns, contains a remnant of a poly(A) tail, and is flanked by direct repeats. Given that this gene can be translated in vitro (Foss et al. 1994), is expressed in a highly regulated manner (Beebe et al. 1992), and is conserved in monkeys and humans, it is very likely to produce a functional protein. As the start of transcription of this gene is likely to be located outside the retrotransposed portion of the gene (in an *Alu* element), it is likely that its tissue specificity arose because of borrowed signals from its new position (See Fig. 2).

Genomic interspersion patterns.—It is possible that TEs may play an important role in the determination of genomic interspersion patterns. The kinds and numbers of transposable elements can vary radically in different species. This becomes quite apparent when comparing TE compliments in *D. melanogaster* and the yellow fever mosquito, *Aedes aegypti*. More than thirty families of transposable elements have recently been found in *A. aegypti* (Z. Tu, pers. comm.). These include a number of families of highly repetitive miniature inverted repeat transposons (MITEs) (Tu 1997), non-LTR retrotransposons (Tu et al. 1998), and a family of SINEs (Tu 1999). The three families of LTR retrotransposons in *A. aegypti* are low in copy number (Warren et al. 1997; Z. Tu, unpubl. data). Many of these TEs tend to be associated with other families of TEs, which is one of the reasons why so many families of TEs have been discovered in a relatively short period of time. In addition, some of the TEs, especially the MITEs, tend to be associated with genes in *A. aegypti*. The genome of *A. aegypti* is approximately 800 Mb, five times larger than that of its dipteran relative, *D. melanogaster*. In addition, the *A. aegypti* genome is organized in a ''short period interspersion pattern'' while the *D. melanogaster* genome is organized in a ''long period interspersion pattern''. It is possible that the presence of highly repetitive MITEs, SINEs, and non-LTR retrotransposons may have contributed to the pattern of short period interspersion in *A. aegypti*.

Genome size and the C value paradox.—Although basal genome sizes (C values) of Archea and Eubacteria vary only by about an order of magnitude (Krawiec and Riley 1990),

those of eukaryotes vary >80,000 fold (Li 1997). In turn, some taxonomic groups of eukaryotes, such as birds, mammals, and teleost fishes show little variation in basal genome size (Gregory and Hebert 1999). Others, such as many plants and invertebrates, often show patterns of quantum shifts of C-values. The mechanisms responsible for quantum size jumps are only poorly understood, but it seems clear that, along with other mechanisms such as polyploidization, the amplification of TEs is a major contributing mechanism. There is now increasing evidence that both shifts in intron lengths, and in the amounts of other noncoding DNA, are driven by genome-wide changes in deletion and insertion frequencies (Petrov et al. 1996; Gregory and Hebert 1999). In turn, C-values are generally positively correlated with cell and nuclear volume (Cavalier-Smith 1982). Changes in genome size, through their influence on cell size can be subject to strong selection pressures acting on the organismal phenotype under certain conditions (Gregory and Hebert 1999).

Chromosome inactivation and speciation.—Whereas the mobilization of TEs has been shown to be associated with high rates of karyotypic change in *Drosophila* (Lim 1988) and yeast (Codon et al. 1997), a direct link between TEs and speciation has not been demonstrated. Although it was earlier thought that hybrid dysgenesis could, under some conditions, lead to speciation (e.g., Ginzburg et al. 1984), no clear experimental evidence supports this idea. Some computer simulations suggest that the necessary combination of conditions would be only infrequently present (J. M. C. Ribeiro, pers. comm.). An interesting alternative model was used by McFadden and Knowles (1997) to investigate the effect of transposon-mediated mutations in populations of asexual digital organisms as they compete for resources in an adaptive landscape. In contrast to nontransposon mutations, deleterious transposon-induced mutations allowed populations to escape evolutionary stasis and promoted speciation events.

The enriched occurrence of the TRAM element in the evolving neo-Y chromosome of *Drosophila miranda* provides strong evidence in favor of the view that Y chromosome degeneration is driven by the accumulation of transposable elements (Steinemann and Steinemann 1997). Also the massive amplification of retroelements in a single generation in wallaby species hybrids (O'Neill et al. 1998) provides an impressive example of the capacity of TEs to change the structural architecture of host genomes without disrupting the normal functioning of somatic cells. In this study, genome-wide undermethylation of retroelements sequences was demonstrated in a species hybrid, in contrast to the normal methylation of nonhybrid parents. Similar observations of deficient methylation and spontaneous chromosome changes were also reported in other mammalian hybrids. Such changes have the potential to facilitate rapid karyotypic evolution, but, again, the potential for these hybrids to produce new species remains speculative. This area of research appears to be ripe for further investigation.

TE Survival Strategies and the Ecology of the Genome

The Parasitic-Mutualistic Continuum

In our introduction, we asserted that the concept of selfishness was valuable in bringing out a property of TEs that is essential for understanding the ability of these elements to invade and spread in genomes, but the use of this term as the sole descriptor tends to be misleading as it excludes the possibility of the evolution of significant beneficial properties of TEs in a species genome. Indeed, description of TEs in terms of host-parasite relationships, rather than selfish DNA has been preferred by a number of writers. Although the expected dynamics of molecular parasites have not yet been rigorously addressed (see Nee and Maynard Smith 1990), it may be useful to consider whether some of the theory applicable to nonmolecular host-parasite associations may be applicable to molecular parasites.

For example, the old conventional wisdom that a well adapted parasite is invariably a benign parasite can no longer be defended (e.g., Ewald 1995). Modern theory predicts that the degree of virulence of parasites depends on a number of factors, such as a parasite's mode of transmission, the degree of host mobility, and the evolutionary response of the host to the parasite. A strong host response to TE invasion may lead to an ''arms race'' between the host and the parasite. Indeed the relationship between hosts and TEs has been characterized by some authors in just this way (e.g., McDonald 1998). Again, this appears to be just one of several possible aspects of the association that may develop in specific cases, rather than representing a universal characteristic of the TE-host relationship.

One possibly useful application of host-parasite theory to TEs relates to modes of transmission. Modern theory predicts that vertical transmission should favor evolution towards parasite benignness (Ewald 1995). Experimental work on plasmids and viruses has provided evidence that restriction of horizontal transfer between bacteria leads to a loss of virulence (Bouma and Lenski 1988; Bull and Rice 1991). In the case of TEs, the predominant mode of transmission is vertical inheritance from one generation to the next. Phylogenetic analysis of some retrotransposon families has indicated that horizontal transfer between species rarely, if ever, occurs. In contrast, Class II elements, notably the *mariner* and *P* families, tend to have a significantly higher frequency of horizontal transfer than Class I elements. Is there evidence that the relationship of Class I elements with their hosts is more benign than the more frequently horizontally transferred Class II elements? Perhaps the very low rates of transposition of some Class I elements (Nuzhdin and Mackay 1994) may be interpreted in this way.

Rather than labeling TE-host associations as either selfish or parasitic, we prefer the idea of a continuum, ranging from aggressive parasitism at one extreme, through a neutral middle ground, to mutualism at the other extreme (Ewald 1987). It seems likely that the association of a specific element with its host might vary along this continuum, perhaps according to the length of association with the host. In particular, we might expect a recently invaded element family to have strong parasitic tendencies. Later, if strong host defenses evolve, then an arm race may occur. However, without strong host defenses, the documented high frequency of vertical transfer might be expected to promote long TE-host associations with coevolution towards the mutualistic end of the continuum. Therefore, while the characteristics of modern invasions may be best described as parasitic, this term is inappropriate for

the established lineages of TEs that have coevolved with their hosts and are now integral to host function (McClure 1999).

Contrary to a strict parasite-host relationship, the association of TEs with their hosts over evolutionary time can, in a number of circumstances, lead to a coevolutionary process that drives the evolution of three types of mechanisms or changes: (1) The coevolution of TE-derived mechanisms to minimize negative effects on the host (and thus, indirectly on the TE). In this category are included transposon self-regulation, tissue specificity, targeting and genome partitioning. (2) The evolution of host defense mechanisms. In this category are included the evolution of epigenetic effects and host suppressors. (3) The evolution of new and altered host functions and large scale genomic changes; these new host functions are frequently achieved by the recruitment of TE functions.

The Coevolution of Mechanisms to Minimize Host Negative Effects

Transposon Self-Regulation

Self-regulation refers to a property of TEs themselves that allows them to regulate their own rate of transposition. It has been shown theoretically that self-regulation of TEs does not evolve if increased copy number alone is responsible for deleterious effects on host fitness (Charlesworth and Langley 1986). However, if the deleterious effects are immediate and occur as a direct consequence of transposition itself, then there may be a selective advantage to elements with reduced transposition rates that allow them to spread in the genome (reviewed by Brookfield 1995).

The transposition of *P* elements in *D. melanogaster* has been shown to result in a significant loss of fitness (Fitzpatrick and Sved 1986; Mackay 1989) and the *P* element system shows clear evidence for self-regulation. *P* activity in *D. melanogaster* is regulated by two types of element-encoded repressor products. A 66-kD, *P* element-encoded, repressor of transposition and excision is responsible for a cellular condition known as *P* cytotype (Robertson and Engels 1989). The genomic location of repressor elements determines the maternal versus zygotic inheritance of *P* cytotype (Misra and Rio 1990). A second type of *P* element-encoded repressor of *P* element transposition has large internal deletions, is sensitive to genomic location, but shows no maternal inheritance (Engels 1996). A similar example in maize is provided by the product of one of the genes encoded by this element, *tnpA*, which can protect *Spm* from methylation, but may also act as a repressor of *Spm* (Fedoroff et al. 1995; Schlappi et al. 1994). Additionally, as with *P* elements, some deleted *Spm* elements can repress full length *Spm* elements in *trans* (Cuypers et al. 1988).

Tissue Specificity of Transposition

The restriction of transposition to the germline appears to be a highly efficient strategy for TEs to ensure transmission to the next generation yet minimize damage to the host. Such a strategy appears to have been adopted by a broad range of elements. For example, expression of active LINE-1 sequences appears to be favored in cells of germ line origin in the human (Singer et al. 1993) and mouse (Trelogan and Martin 1995) genomes. Also, the expression of intracisternal A-particles (IAPs) in the mouse genome was found to be essentially restricted to the male germ line (Dupressoir and Heidmann 1996). The *P* and *I* elements in *D. melanogaster* also provide good examples of the restriction of transposition to the germ line. In the case of the *P* element, the 2–3 intron is spliced only in germ cells. The absence of transposase in somatic cells results from the production of a truncated 66-kD protein due to a premature stop codon following a frameshift. Two different proteins encoded by host genes have been implicated in the somatic inhibition (Siebel et al. 1992, 1995). Among plants, one of the two genes encoded by the *MuDR* transposon in maize is specifically expressed only in those cells that are potentially transmitted to the next generation. Expression of this gene is down-regulated in terminally differentiated tissue and markedly up-regulated in floral tissues (Donlin et al. 1995).

Nonrandom Distributions and Genome Partitioning

Although TEs have been found in nearly every portion of host genomes examined, their distributions are decidedly nonrandom. For example, in the compact genome of yeast, the *Ty1*, *Ty2*, *Ty3* and *Ty4* retroelement families, target the upstream region of the RNA polymerase III-transcribed genes. In contrast, the *Ty5* family inserts into silenced regions of the genome such as the telomeres and mating type loci (Kim et al. 1998). In addition to *Het-A* and *TART* that exclusively target telomere ends of *Drosophila* (Biessmann and Mason 1997; Levis et al. 1993; Pardue et al. 1996), *copia* retrotransposons found in *Allium cepea* (Pearce et al. 1996) and two retroposons, *Zepp* from *Chlorella* (Higashiyama et al. 1997) and *SART1* from *Bombyx mori* are preferentially inserted into telomeric repeats (Takahashi et al. 1997).

Repetitive elements are found in heterochromatic clusters in many less compact genomes, such as *Drosophila*. In theory, this distribution can be partially explained by increased survival in heterochromatic, compared with euchromatic, genomic locations, as a direct consequence of the differing density of genes in the two types of chromatin. However, experiments in *D. melanogaster* have provided no support for the hypothesis that the primary reason for accumulation of TEs in heterochromatin is selection against TEs inserted into euchromatin, either directly at insertion sites or indirectly through ectopic recombination (Dimitri and Junakovic 1999). In contrast, there is evidence for specific targeting of some heterochromatic regions. For example, the lethal mutation frequency induced by the *I* element in 13 loci located in the proximal heterochromatin of *D. melanogaster* chromosome 2 was about an order of magnitude higher than that of euchromatic genes located on the same chromosome (Dimitri et al. 1997). These data provide evidence that *I* elements transpose with high frequency into pericentric regions of chromosome 2 that are enriched with heterochromatin. Further, even within heterochromatin, TEs are not evenly distributed. In one study, all of 11 elements examined in *D. melanogaster* exhibited preferential and discrete clustering into specific heterochromatic regions, suggesting that different elements may have distinct insertional preferences (Pim-

pinelli et al. 1995). This clustering was also observed for *P* and *hobo* element insertions, despite their relatively recent arrival into the *D. melanogaster* genome (Kidwell 1994; Simmons 1992).

An extreme example of local clustering is found in maize in which nearly all of the 240 kb between two host genes was composed of retrotransposon sequences (SanMiguel et al. 1996). Many of these elements are inserted into other elements, suggesting that they target their own kind for insertion. The region between maize genes is much less recombinogenic than are the genes themselves, suggesting that recombination between these elements (which is likely to be disastrous for the host) is suppressed (Civardi et al. 1994). Thus, it appears that the intergenic regions of maize represent a specific domain of the maize genome specialized for habitation by retrotransposons.

Not all TEs target heterochromatin, however. Like the retroelements, miniature inverted repeat transposons (MITEs) are present in high copy number in maize (Wessler et al. 1995). Despite their ubiquity, MITEs are found almost exclusively in or near host genes, suggesting their exclusion from the specialized domain inhabited by retrotransposons (Zhang et al. 2000). Similarly, *Mu* elements, highly mutagenic transposons in maize, show strong preference for single copy sequences (Cresse et al. 1995), suggesting that they too are excluded from the intergenic concentrations of retroelements.

The nonrandom distribution of clusters of TEs, in a variety of organisms, suggests that genomes can be partitioned into a variety of ''ecological niches,'' which can be exploited in different ways by different transposons. The formation of these niches, as well as their exploitation by the transposons, is most likely a result of a long-term interaction between host and parasite. Heterochromatin proteins can recognize and silence transposable elements (Fanti et al. 1998), some of which are known to target heterochromatin for insertion (Terrinoni et al. 1997). Thus, the evolution of heterochromatin could have lead to a self-perpetuating expansion of domains rich in transposable elements. Transposons that avoid heterochromatin and specifically target gene sequences, like the *Mu* elements of maize, may be taking advantage of the open chromatin configuration and local concentration of transcription factors around host genes, but at the cost of increased selection against new insertions. Not surprisingly, the copy number of this class of elements in maize is orders of magnitude lower than that of the retroelements. MITE elements, on the other hand, can be present in extremely high numbers (Wessler et al. 1995). However, they are very small, and are found almost exclusively in untranslated portions of host genes, either because of selection, or because their insertion profile represents yet another ecological niche.

The evolution of large scale domains within genomes can have a significant effect on host evolution. Through investigation of Position Effect Variegation in *D. melanogaster*, it is known that heterochromatin can affect the expression of adjacent genes (reviewed by Henikoff 1996) and there is strong evidence that local chromatin configuration has a basic role in the unfolding of the developmental program in species as diverse as *Drosophila* (reviewed by Pirrotta 1996) and mice (van der Lugt et al. 1994). Although it is not known whether transposons were the factor that caused heterochromatin to evolve in the first place, because of the intimate relationship between TEs and various aspects of chromatin organization it seems likely that their coevolution has affected the means by which the genome is structured and thus some of the ways that host genes are expressed.

Mobile Introns

If, as has been suggested (Orgel and Crick 1980), nuclear introns evolved from a form of self-splicing transposon, these introns are among the most ancient and well-adapted mobile elements. Although they may have evolved the capacity to be spliced from mRNA to minimize their effect on host gene expression, they have become intimately tied to the biology of the host. Not only can differential splicing affect protein expression (e.g., Larsson et al. 1997), but introns can carry enhancers (Gillies et al. 1983; Muller et al. 1999), or repressors (Wight and Dobretsova 1997) of transcription. Because of their position within gene sequences they have become an integral mechanism for the modulation of gene expression.

One school of thought holds that spliceosomal nuclear introns evolved from an ancestor similar to two extant classes of self-splicing introns found in the genomes of organelles such as the mitochondria. These exceptional classes of introns provide clues as to the structure of the ancient transposon-like progenitors of the predominant class of nuclear introns found in most eukaryotic genes.

Group II introns present in some yeast mitochondrial genes are mobile elements that encode a protein that has maturase, reverse transcriptase and endonuclease activities. Group II introns of plastids can also encode reverse transcriptase (Natvig et al. 1984). These introns are self-splicing and can move to both allelic and nonallelic targets. The ability to move to intronless allelic sites is called retrohoming (Curcio and Belfort 1996).

Group I introns are mobile self-splicing elements found mainly in nuclear rRNA genes and the genomes of organelles. An intron found in the mitochondrial *cox1* gene has a widespread but patchy distribution in angiosperms. Four lines of evidence lead Cho et al. (1998) to conclude that the 48 angiosperm genera found to contain this cox1 intron acquired it by 32 separate horizontal transfer events. This explosive invasion appears to have been entirely of recent occurrence and illustrates the similarities of mobile introns and TEs.

Group I and II introns both encode homing endonucleases, an enzyme involved in site-specific double-stranded DNA cleavage. Due to their capacity to rapidly spread throughout a given population, evidence for frequent horizontal transmission, and widespread but sometimes patchy distribution, homing endonucleases are generally considered to by highly effective parasites whose frequent association with introns and inteins ensures relative invisibility to selective pressures. There is evidence, however, that some homing endonucleases have been coopted for use by hosts. For instance, it has been suggested that HO endonuclease in yeast, which is required for mating-type conversion, may have evolved from an originally selfish-intron-encoded homing endonuclease (Gimble 2000). In another remarkable example, an intronic *E. coli*

homing endonuclease functions to kill off competing strains of *E. coli* by cleaving their DNA (James et al. 1996).

Evolution of Host Defenses

Epigenetic Effects

A fundamental problem facing all complex multicellular organisms is that of selective gene expression. For proper differentiation to occur, a given cell must only express a subset of its genes and that pattern of expression must be heritable in a given differentiated lineage. Heritable changes in gene expression, in the absence of changes in the DNA, are referred to as epigenetic changes. Along with differentiation in multicellular organisms, epigenetic alterations in gene expression are associated with phenomena as diverse as X chromosome inactivation in mammals, transgene silencing in plants, and selective gene expression in some bacteria (reviewed by Torreblanca and Casadesus 1996).

One of the more prevalent and intriguing epigenetic mechanisms involves cytosine methylation of DNA. We should note here that although we have chosen to focus on it here because of its frequent association with changes in TE activity, methylation is just one of a variety of interrelated epigenetic mechanisms that can inactivate TEs, including those such as chromatin remodeling and double-stranded RNA mediated silencing.

In hosts, cytosine methylation of promoters is known to inhibit transcription (Meehan et al. 1992; Muiznieks and Doerfler 1994), and is heritable (Wigler 1981). Because it can repress transcriptional activity and can be replicated via mechanisms which recognize hemimethylated DNA, methylation is an appealing explanation for heritable alterations in gene expression. Indeed, genetic imprinting, which involves the specific inactivation of an allele from one parent, is associated with methylation of the inactive allele (Li et al. 1993), and failures in imprinting are associated with a number of diseases (Reik and Maher 1997). Further, mice that are homozygous for a mutation (*Dnmt*) in a mammalian methyl transferase die soon after gastrulation, implying that methylation function is essential for normal mammalian development (Li et al. 1992). In a similar way, *Arabidopsis* plants homozygous for an antisense version of a plant methyltransferase cDNA (METI) exhibit a number of developmental abnormalities (Finnegan et al. 1996). Inactive X chromosomes in mammals are known to be extensively methylated (reviewed by Singer-Sam and Riggs 1993), suggesting that methylation plays an important role in the normal repression of one copy of the female X chromosome.

Because of these kinds of correlations, methylation has long been viewed as an essential component of the epigenetic regulation of mammalian genes that evolved to make differentiation in complex organisms possible (Holliday and Pugh 1975). However, that hypothesis has recently been challenged by the view that methylation evolved as a genomic defense against parasitic DNA (Barlow 1993; Yoder et al. 1997). These authors argue that, what has been interpreted as regulation of host genes is in fact a system that evolved to repress transposon activity. Similarly, it has been argued that transgene transcriptional silencing in plants, which is often associated with DNA methylation, is clearly related to the defense of the plant genome against repetitive or foreign DNA (reviewed by Matzke and Matzke 1998). In mammals, TEs are interspersed within the host genes (Yoder et al. 1997), and a majority of those elements are methylated soon after implantation (Monk et al. 1987; Sanford et al. 1987). Thus, it has been suggested that what had been believed to be a carefully regulated change in the methylation status in the host genes represents instead efforts of the host to downregulate the activity of these interspersed TE sequences.

Interestingly, one of the first cases of epigenetic alteration of gene expression described was Barbara McClintock's examination of the *Spm* transposon in maize (McClintock 1958). It was her belief that the programmable, heritable changes in *Spm* expression could serve as a paradigm for gene regulation. It was subsequently discovered that maize elements undergo reversible methylation associated with changes in their activity (Banks et al. 1988; Dennis and Brettell 1990). Indeed, methylation of transposons (in hosts that methylate their DNA) is the most reliable single predictor of that activity.

There is an increasing body of evidence that regulation of transposable elements by DNA methylation and the regulation of host genes are mechanistically linked. For instance, changes in methylation of TEs in maize are clearly nonrandom. McClintock noted that *Spm* could cycle between active and inactive phases in heritable and programmed patterns during development (reviewed by Federoff 1989). Similarly, Martienssen et al. (1990) have observed that the *Mutator* system of transposons in maize is progressively methylated as the plant meristem ages. These observations lend credence to the idea that shifts in the degree of silencing of these transposable element are correlated with host-encoded shifts in global epigenetic regulatory pathways. More recently, it has been discovered that mutations that prevent paramutation (an epigenetically regulated process of gene inactivation) in maize, also reverse *Mu* element silencing, suggesting that the regulation of both host and parasite can be mediated by a similar suite of genes (D. R. Lisch et al., unpubl. ms.).

It is impossible to say at this point, how epigenetic mechanisms such as methylation arose. Certainly they can play different roles in different species. In *Ciona intestinalis*, for instance, it is active genes that are predominantly methylated and transposons that are hypomethylated (Simmen et al. 1999), and some organisms, such as *Drosophila melanogaster* have little if any methylation (Urieli-Shoval et al. 1982). However, given its prevalence and frequent association with gene silencing, and thus its potential utility, we suggest that DNA methylation almost certainly plays an important role in epigenetic regulation of host genes as well as transposons. The two perspectives are not, after all, mutually exclusive. DNA methylation in many species may have evolved as a response to repetitive DNA. However, the presence of such a flexible mechanism for reversible gene inactivation presents the host with an invaluable tool for nuanced gene expression. Further, regardless of how it arose, the massive methylation of the large fraction of the genome of organisms, such as maize and humans, almost certainly has affected larger structural characteristics of these genomes. In a remarkable example of this, O'Neill et al. (1998) have shown in wallaby hybrids that large scale changes in methylation are accom-

panied by gross alterations of entire chromosomes due to massive retrotransposon amplification.

Interestingly, transposon systems in maize are frequently associated with suppressible mutations, which effectively link systems of epigenetic regulation with host gene expression (reviewed by Girard and Freeling 1999). In these cases, expression of the mutant phenotype becomes dependent on the expression of the transposase source, and variations in transposon activity due to epigenetic regulation of the transposon can and do result in changes in host gene expression. As we have pointed out earlier, regulatory regions of host genes have in some cases incorporated transposon sequences, and transposon coding regions have become host genes. Thus, the blurring of the line between host gene and parasitic element represents instances in which mechanisms that may have arisen to control repetitive DNA may have had a profound impact on the regulation of host development and evolution.

Host Recruitment of Transposon Functions

As we have shown, transposons can introduce a distinct spectrum of variability through insertion mutagenesis, suggesting that TEs can introduce selectively advantageous variation otherwise not easily available to the host. In these cases it is the products of TE activity that are selectively advantageous. However, TEs also carry with them enzymatic machinery that is itself potentially useful. Any host function that involves the breakage and rejoining of DNA, for instance, could potentially benefit from the presence of a diverse pool of rapidly evolving enzymes specializing in that process. This category is often not easy to identify because recruitment may occur long after the elements initially entered the host genome; these sequences may even have lost some of the features that distinguish them as of TE origin. Four subcategories of transposon recruitment are described and illustrated below: those providing (1) regulatory functions; (2) structural functions; (3) enzymatic functions; and (4) new coding sequences.

Host Recruitment of Regulatory Functions

The potential importance for evolution of mutations affecting regulation of gene expression began to be appreciated in the 1960s and 1970s (Britten and Davidson 1969, 1971). This preceded the growing understanding of the genetic basis of development in multicellular eukaryotes as a process in which genes acting early in development serve as regulators of subsequent, more regionally specialized, pathways that progressively direct cells into more highly specialized functions (McDonald 1990). Class I TEs, in particular, have the properties required of any mechanism postulated to be at the foundation of major developmental shifts that have occurred over evolutionary time. These properties are their ability to be completely or partial suppressed and their ability to increase in frequency under conditions of environmental stress (McDonald 1990). More recently there is growing appreciation of the dominant evolutionary role played by reshuffling of different combinations of ready-to-use regulatory elements between different genes (Frigerio et al. 1986; Xue and Noll 1996), and the alteration of expression induced by mutation of cis-regulatory elements (Xue and Noll 1996).

The classical model for the evolution of genes that have undergone gene duplication implies that the only mechanism for the permanent preservation of duplicate genes is the fixation of rare beneficial mutations. However, a new model, the duplication degeneration complementation (DDC) model (Force et al. 1999), posits that deleterious or degenerative mutations in gene regulatory elements can increase, rather than decrease, the probability of duplicate gene preservation. The latter model incorporates the recent view that genes often have several functions, or subfunctions, that are controlled by different DNA regulatory elements (see Fig. 3). Thus the DDC model allows for the evolution of new subfunctions in the absence of new beneficial mutations in the duplicated genes themselves. As illustrated in the following examples, this changed perspective is likely to be important for fully understanding the significance and role of TEs in regulatory evolution. The reason is that TEs, like mutators in general, produce a greater proportion of mutations that are deleterious rather than beneficial to their hosts. The number of cases in which TEs have been shown to affect host gene regulation is steadily growing.

MITEs in gene regulation in plants and other species.—MITEs are a common feature of introns and the 5′ and 3′ regions flanking the coding sequences of plant genes. These elements are short (100–400 bp in length) and most do not have any coding potential. MITEs are present in high copy number (3,000–10,000) per genome and in plants have DNA target site preferences for TA or TAA. They are exemplified by the *Tourist* element in maize (Bureau and Wessler 1992) and the *Stowaway* element in Sorghum (Bureau and Wessler 1994). Similar elements have recently been found in species as diverse as *Xenopus* (Unsal and Morgan 1995), Zebra fish (Izsvak et al. 1999), the yellow-fever mosquito, *Aedes aegypti* (Tu 1997), *Arabidopsis thaliana* (Casacuberta et al. 1998), and *Homo sapiens* (Smit and Riggs 1996). The propensity of these elements to insert into (and, in some cases, even make up) the regulatory regions of genes in all of these species suggests an important role in host regulatory evolution. Additionally, analysis of two grass species (Avramova et al. 1998) suggested that some MITEs preferentially insert near Matrix Attachment Regions (MARs) and/or that they can serve as MARs suggesting that these elements, like the retroelements in maize, may have a major role in partitioning the genome of their hosts.

Mammalian retroposons in gene regulation.—Tomilin (1999) has documented the growing number of well-established examples of mammalian retroposons that are involved in the modulation of expression of protein-coding genes transcribed by RNA polymerase II (Pol II). Retroposons, such as SINES and LINES encode conserved binding sites for some important Pol II transcription factors and sequences involved in regulation of mRNA stability. One example of these functional Pol II Transcription-Modulating Elements (TMEs) (Tomilin 1999) in retroposons is the tRNA-derived rat ID retroposons (Weiner et al. 1986; Kim et al. 1994). These retroposons act as enhancers of RNA polymerase II gene transcription in cell lines that express these RNAs. Based on increasingly solid evidence, Tomilin (1999) con-

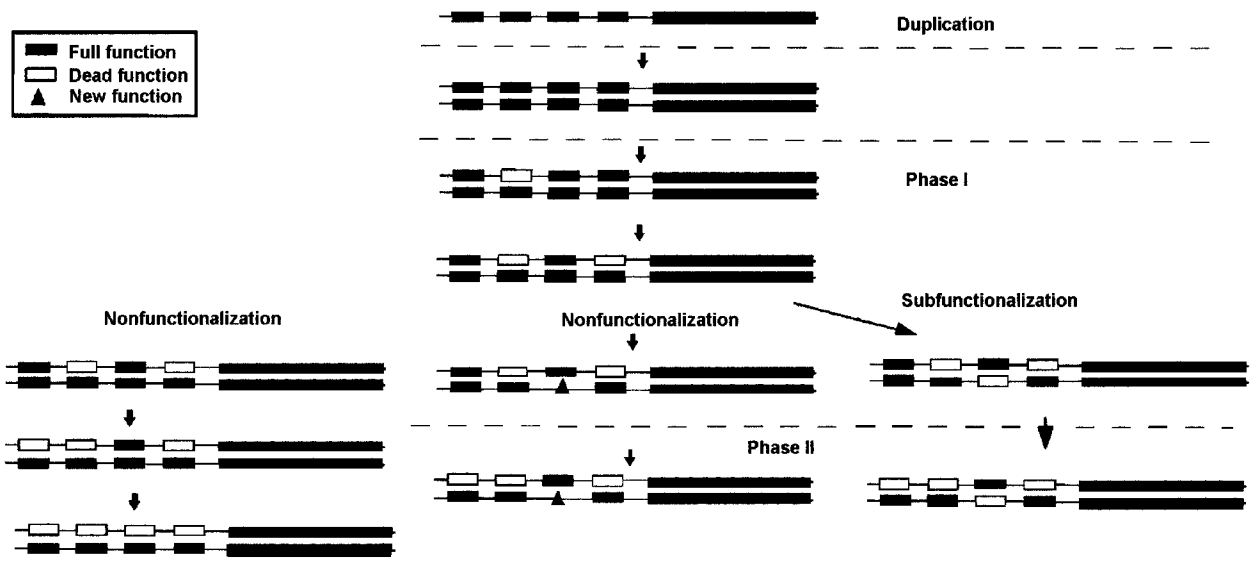

FIG. 3. Illustration of the duplication degeneration complementation (DDC) model for the evolution of duplicate genes. Three potential fates of duplicate gene pairs with multiple regulatory regions are shown according to this model. The small boxes denote regulatory elements with unique functions, and the large boxes denote transcribed regions. Solid boxes denote intact regions of a gene, whereas open boxes denote null mutations, and triangles denote the evolution of a new function. Because the model focuses on mutations that are fixed in populations, the diagram shows the state of a single gamete. In the first two steps, one of the copies acquires null mutations in each of two regulatory regions. On the left, the next fixed mutation results in the absence of a functional protein product from the upper copy. Because this gene is now a nonfunctional pseudogene, the remaining regulatory regions associated with this copy eventually accumulate degenerative mutations. On the right, the lower copy acquires a null mutation in a regulatory region that is intact in the upper copy. Because both copies are now essential for complete gene expression, this third mutational event permanently preserves both members of the gene pair from future nonfunctionalization. The fourth regulatory region, however, may still eventually acquire a null mutation in one copy or the other. In the center, a regulatory region acquires a new function that preserves that copy. If the beneficial mutation occurs at the expense of an otherwise essential function, then the duplicate copy is preserved because it retains the original function (after Force et al. 1999).

cluded that ''mammalian retroposons play a significant role in the evolution of regulatory elements controlling expression of protein-coding genes crucial for the functioning of organisms.''

Other TE-associated regulatory functions.—Britten has used stringent criteria for the identification of strong cases of the involvement of TEs in the actual evolution of gene regulation (Britten 1996a,b, 1997). He maintains that a long term perspective is necessary in identifying and understanding mutations important for gene regulation. The number of cases he has identified is small, but growing. In addition to the plant MITE examples discussed above, he includes cases involving the association of a retrovirus-related element with androgen regulation of the sex-limited protein (*Slp*) gene in mice (Stavenhagen and Robins 1988) and inverted repeats in the *CyIIIa* actin gene of sea urchins (Anderson et al. 1994). In plants, computer-assisted database searches using plant copia-like retroelements as query sequences revealed that ancient, degenerate retrotransposon insertions are found in close proximity to 21 previously sequenced plant genes. The data suggest that these elements may be involved in gene duplication and the regulation of gene expression (White et al. 1994).

A number of additional examples of the effects of TEs on gene regulation are reviewed by Kidwell and Lisch (1997). Of particular interest is an example involving the *Tam3* element in *Antirrhinum* (Lister et al. 1993). A series of TE-associated mutations exemplifies the potential for TE-mediated ''rewiring'' of regulatory networks, in this case by bringing new regulatory sequences in proximity to exonic sequences via an inversion, followed by an imprecise excision event. A recent example of a TE becoming an integral part of the regulation of a host gene is provided by the *apolipoprotein* (*a*) gene in humans (Yang et al. 1998). An enhancer of this gene, located 20 kilobases upstream of the start of transcription, is located entirely within a LINE element.

Differential developmental and tissue-specific expression.—Insertion mutations appear to provide a ready mechanism for the fast evolution of complex patterns of developmental regulation and tissue specific expression. A clear example of TE-induced differential expression of a gene product is provided by the arylalkylamine N-acetyltransferase (aaNAT1) locus in *Drosophila* (Brodbeck et al. 1998). In insects, aaNAT1 has been implicated in sclerotinization, inactivation of certain neurotransmitters and catalysis of a rate-limiting step in melatonin formation. In *D. melanogaster*, two alleles of this locus have been studied in detail. The allele that encodes the more abundant isoform has two TEs, *MDG412* and *blastopia*, that together constitute 12 kb, inserted downstream of the first exon. This isomer is first expressed in embryogenesis and continues to be detectable throughout development and in the adult brain. In contrast, the less abundant isoform, encoded by the allele lacking the TE insert, appears first during the late pupal stage of development.

Host Recruitment of Structural Functions

Telomeres.—Another fundamental problem facing all eukaryotes is that presented by the replication of linear chromosomes. Because DNA polymerase is insufficient to replicate the very ends of these chromosomes, additional specialized enzymatic machinery is necessary to ensure that entire linear chromosomes are replicated successfully. In most eukaryotes the key enzyme involved is telomerase. This enzyme uses short RNA molecules and a reverse transcriptase function to add residues that would otherwise be lost during replication.

As it is a fundamentally important survival mechanism having to do with an ancient problem faced by eukaryotes, one would expect the use of some form of telomerase to be extremely well conserved, as it is in most species examined. However, insects of the order Diptera, which includes *Drosophila melanogaster*, turned out to have a remarkable surprise in store. Rather than telomerase, *Drosophila* uses two families of retrotransposons, *Het-A* and *TART*, to solve the same basic problem (Levis et al. 1993; Sheen and Levis 1994). At the conclusion of each round of DNA replication, one or more of these elements transposes from source sites a distance from the telomere to the termini of each replicated chromosome. Both these elements are retrotransposons of the non-LTR type, an ancient and widely distributed class of mobile elements (Flavell 1995). At the telomeres there are a series of elements that have transposed to the end of the chromosome and then subsequently been ''eaten away'' before the addition of the next transposon. The overall length of chromosomes is kept relatively constant by the continual addition of new transposon sequences.

Interestingly, another insect, the silkworm *Bombyx mori*, has a non-LTR retroelement that also specifically targets telomeres (Takahashi et al. 1997). In this species, however, rather than replacing telomerase function, the element transposes into a more typical telomeric sequence, suggesting that in this case the TE is simply targeting a ''safe haven'' for insertion, rather than acting as a surrogate telomerase. Similarly, in the green algae *Chlorella vulgaris*, a non-LTR retrotransposon, *Zepp*, accumulates in nested clusters immediately behind the telomeric repeats (Higashiyama et al. 1997). These observations suggest that some non-LTR retroelements in *Diptera* may have evolved from parasitic exploitation of transcriptionally silent regions in the telomeres to providing a vital role in host telomere maintenance.

As it turns out, the situation is even more complex. Telomerase itself may have evolved from an originally selfish genetic element. The RT domain of telomerase is functionally and structurally similar to that found in a wide range of mobile genetic elements (Eickbush 1994; Nakamura et al. 1997). Phylogenetic analysis has shown that telomerase represents a deep branch in the evolution of reverse transcriptases, strongly suggesting a common origin with TEs (Eickbush 1997; Nakamura et al. 1997). Because of the wide divergences between sequences and the uncertainty of any molecular clock based on hypermutable mobile elements, the exact phylogeny of RT sequences is difficult to determine (Nakamura and Cech 1998). However, based on such an analysis, Eickbush (1997) has proposed that the RT function of telomerase was recruited from an originally parasitic element. An alternative explanation is that telomerase RT is phylogenetically more basal, and that non-LTR elements arose as parasitic ''escaped'' telomerase sequences (Eickbush 1994, 1997). In either event, it is clear that telomerase and reverse transcriptase-based mobile elements represent two ends of a dynamic spectrum of possible relationships between host and parasite. The essential role of *TART* and *Het-A* elements in *Drosophila* and the more parasitic relationship between telomere-targeting retroposons and other hosts suggests that a single class of element (non-LTR elements) can occupy a range of positions within that spectrum.

In each of the cases outlined above, transposable elements may have had profound effects on the mode and perhaps even the direction of host evolution. Certainly, vivipary and V(D)J recombination represent fundamental turning points in vertebrate evolution, and telomere maintenance is a problem that all eukaryotes face. In this case, and in most others perhaps, alternative pathways were available. Yeast, for instance, can use a recombination based pathway to maintain its chromosome ends in the absence of telomerase (Le et al. 1999; Lundblad and Blackburn 1993). However, it is clear that in many cases the best available mechanism was encoded by previously selfish genetic elements.

Genome reorganization in ciliated protozoa.—The programmed somatic excision of interstitial segments of DNA that occurs as part of the regular life cycle of ciliated protozoa provides a fascinating example of the cooption of TEs for a host function. Each organism contains both a micronucleus and a macronucleus. The micronucleus consists of conventional eukaryotic chromosomes and can be viewed as the ''germline'' nucleus (Klobutcher and Herrick 1995). It is transcriptionally inactive during vegetative propagation, but is active during sexual reproduction. The genome of the macronucleus consists of a subset of the micronucleus DNA which is extensively rearranged by chromosome fragmentation, de novo telomere synthesis and DNA amplification. During the reorganization process, large numbers of interstitial segments of DNA (internal eliminated sequences; IES) are exactly excised from the somatic nucleus through DNA breaking, rejoining, and splicing. There is evidence that TEs are a major component of the highly variable structures of IESs (Klobutcher and Herrick 1995; Seegmiller et al. 1996; Seegmiller and Herrick 1998). It has been proposed (Klobutcher and Herrick 1997) that transposon invasions have generated the ciliate DNA excision phenomena. Further, it has been suggested that short IESs are ancient transposon IESs that have shrunk by loss of internal sequences unnecessary for host developmental excision. A series of evolutionary steps is envisaged (Klobutcher and Herrick 1997). This includes four phases: invasion, bloom, abdicate, and fade (IBAF) (see Fig. 4). Of particular interest in the present context is the change occurring during the ''abdicate'' phase in which a transposon-encoded excisase comes under the control of a host promoter. Note also the similarities of the IBAF phases with those of a TE life cycle (Fig. 1).

Centromeres and heterochromatin.—Centromeres are the regions of chromosomes devoted to segregation of sister and homologous chromosomes during cell division. They generally contain few transcribed genes and are often hetero-

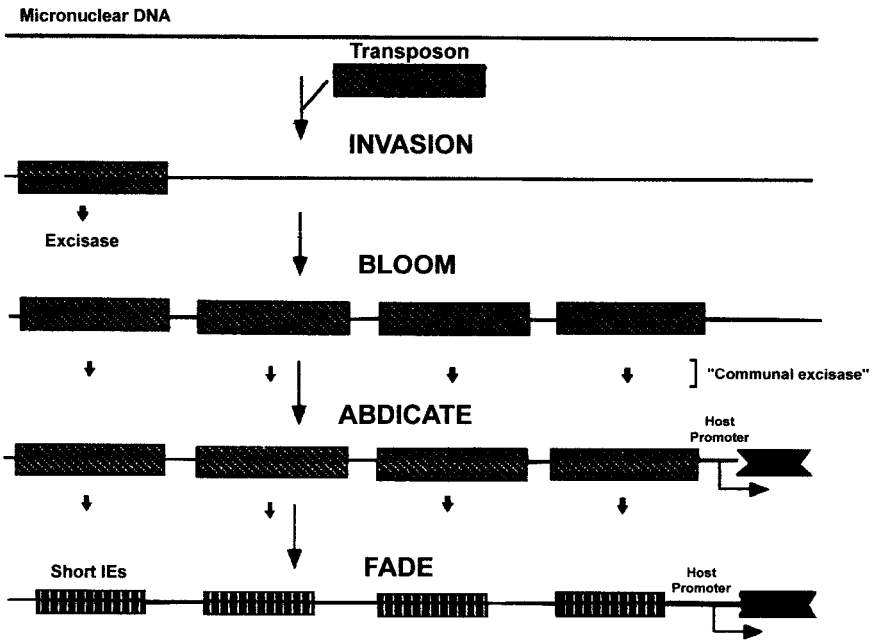

FIG. 4. The IBAF (invasion, bloom, abdicate, and fade) cycle in ciliated protozoa. Micronuclear DNA is shown as a line, and transposons and short IESs as black rectangles. In the abdicate step, a part of a transposon encoding excisase functions is represented as coming under the control of a strong host promoter (from Klobutcher and Herrick 1997).

chromatic in character. They are essentially defined by their late replication in S phase and their function in storing specific DNA-binding proteins in cell division (Csink and Henikoff 1998). In a number of species they are made up largely of transposons and simple repeated sequences. In the yeast *Neurospora crassa*, for example, sequenced portions of the centromere are composed entirely of degenerate transposons, one of which, *Tcen*, is located exclusively in the centromere (Cambareri et al. 1998). Similarly, a centromere of *D. melanogaster* was observed to be primarily made up of satellite sequences and single intact transposable elements (Sun et al. 1997). These sequences were neither unique to the centromere, nor were they present in all centromeres of *D. melanogaster*, supporting the view that it is the general structural characteristics of these centromeres that are important, rather than specific sequences.

In maize, sequenced centromeric regions were also rich in both simple repeat and TE sequences, some of which, like *Tcen*, were localized exclusively to centromeres (Ananiev et al. 1998). Interestingly, heterochromatic chromosomal knobs in maize, which are also composed of simple repeats interspersed with transposable elements, can, in some circumstances, act as neocentromeres (Yu et al. 1997). These neocentromeres exhibit meiotic drive, meaning they are over represented in the progeny (Dawe and Cande 1996), a suspiciously ''selfish'' behavior. Any TE contribution to centromeric function can, in some instances, result in a change in the frequency of TEs in a population.

The SGM family of TEs in *Drosophila guanche* provides an interesting example of the cooption of a formerly active TE to form a major satellite DNA (Miller et al. 2000). The SGM family of TEs has structural and functional similarities to MITES. SGM elements were active in the common ancestor of the closely-related species *Drosophila subobscura*, *Drosophila madeirensis* and *Drosophila guanche* in the *Drosophila obscura* species group. They gave rise to the A-type promoter of the *P*-element neogene cluster (Miller et al. 2000, see also section below). SGM elements were active in *D. subobscura* and *D. madeirensis* only 1–2 MYA. However, in *D. guanche* they gave rise to a major satellite DNA that now occupies approximately 10% of the genome and is found mainly in the centric heterochromatin. Other examples of the cooption of TEs to form centromeric satellite DNA are provide by mammals generally (Kipling and Warburton 1997), Cetaceans (Kapitonov et al. 1998), and other Drosophilids (Zelentsova et al. 1986; Heikkinen et al. 1995).

Recent work in both yeast and mammals suggest that transposons may play an even more active role in centromeric function than was previously thought. In mammals, centromeric repeat binding proteins bind repetitive motifs within centromeric regions. One of these proteins, CENP-B, is clearly related to the *pogo* TE of *Drosophila*, and the *Tigger* transposon in humans (Earnshaw et al. 1987). Homologues of CENP-B are found in species as diverse as fission yeast (Lee et al. 1997) and *Arabidopsis thaliana* (Lin et al. 1999). Although sequences similar to those bound by CENP-B in humans are found in centromeric regions of a number of organisms, they are apparently much less widespread than CENP-B itself (Canapa et al. 2000; Haaf et al. 1995; Heslop-Harrison et al. 1999; Nonomura and Kurata 1999). Interestingly, the *Tigger* terminal inverted repeat, which is presumably interacting with the transposase, contains a close match to the region bound by CENP-B. Thus, it appears that at least one function of the *Tigger* transposon, i.e., the binding of a specific DNA sequence, has been retained by CENP-B.

In some genetic backgrounds, mice that are mutant for the

Cenp-b locus show significant reduction in fertility, perhaps due to lesions in uterine remodeling, a process which requires extensive mitotic division (Fowler et al. 2000). The observed dependence on genetic background in mice may be related to redundancy of function. In support of this hypothesis, in fission yeast, when two redundant homologues of CENP-B were mutated the resulting strains were severely compromised in growth rate and morphology, and they exhibited marked chromosome missegregation (Baum and Clarke 2000). These data suggest that CENP-B, which is clearly related to an ancient and widespread family of TEs, can play an important role in host chromosome segregation.

Interestingly, at least two other genes that are clearly related to the *Pogo* superfamily of transposases appear to have functional roles in their hosts. The *jerkey* gene causes pronounced neurological problems when mutant in mice (Toth et al. 1995), and another relative of *Pogo*, the *rag3* gene in yeast, is involved in Pyruvate decarboxylase synthesis (Prior et al. 1996). This phylogenetic intermixing, in which a series of apparently functional host genes are related to a series of transposons, suggests that this class of element may be particularly likely to be ''domesticated'' (Miller et al. 1997) for a variety of roles.

Genomic interspersion patterns.—The kinds and numbers of transposable elements can vary radically in different species. This becomes quite apparent when comparing TE compliments in *D. melanogaster* and the yellow fever mosquito, *Aedes aegypti*. More than thirty families of transposable elements have recently been found in *A. aegypti* (Z. Tu, pers. comm.). These include a number of families of highly repetitive MITEs (Tu 1997), non-LTR retrotransposons (Tu et al. 1998), and a family of SINEs (Tu 1999). The three families of LTR-retrotransposons in *A. aegypti* are low in copy number (Warren et al. 1997; Z. Tu, unpubl. data). Many of these TEs tend to be associated with other families of TEs, which is one of the reasons why so many families of TEs have been discovered in a relatively short period of time. In addition, some of the TEs, especially the MITEs, tend to be associated with genes in *A. aegypti*. The genome of *A. aegypti* is approximately 800 Mb, five times larger than that of its dipteran relative, *D. melanogaster*. In addition, the *A. aegypti* genome is organized in a ''short period interspersion pattern'' while the *D. melanogaster* genome is organized in a ''long period interspersion pattern''. It is possible that the presence of highly repetitive MITEs, SINEs, and non-LTR retrotransposons may have contributed to the pattern of short period interspersion in *A. aegypti*.

Host Recruitment of Enzymatic Functions

Repair functions.—In addition to the healing of chromosome ends in *Drosophila*, described above, retrotransposon-encoded RT can be responsible for the repair of double-strand chromosome breaks (DSBs) in *Saccharomyces cerevisiae* (Moore and Haber 1996; Teng et al. 1996). Note that in this case, as in others described here, coadaptation may have grown out of parasitic element behavior; DSBs may simply represent an especially good target for efficient TE insertion and, in turn, these insertions may, in some instances, also be the most efficient repair pathway available to the host. The net effect, rapid insertional repair of breaks, is expected to benefit both host and TE. The ability to heal breaks, either in the middle, or at the ends of chromosomes may represent an evolutionary conserved role for retrotransposons (Labrador and Corces 1997). The widespread distribution of retroelements in centromeric heterochromatin may also be relevant in this regard. Such elements may play a role in healing the breaks that this type of chromatin is prone to because of its late replicating nature (Labrador and Corces 1997).

Perhaps the most remarkable example of host recruitment involves the vertebrate immune system. Pathogens, such as viruses, are capable of rapid evolutionary change (Fitch 1996). This capacity for extreme plasticity over short periods of time presents large, slowly evolving organisms with a profound problem. They must match the flexibility of their pathogens with a mechanism capable of producing the same degree of variability within the somatic cells of a single individual. Recent work on the V(D)J recombination system in vertebrates provides a clue as to how such a system may have evolved.

Diversity in the immune system of vertebrates is generated in lymphocytes. In many vertebrates, antibodies and T cell receptors (immunoglobins that bind directly to antigens) are produced in somatic cells via a process of chromosome breakage and rejoining (Tonegawa 1983). This process joins together several segments of DNA that together make up sequences that encode immunoglobins. Because the joining is imprecise and may involve any one of several possible sets of sequences, this process results in a remarkably diverse set of distinct immunoglobin molecules. Because it is repeated in the somatic lymphocytes of each individual, it is possible for each individual to carry with it a mechanism for generating the same degree of variability as is present among all potential pathogens. As it turns out, transposons appear to have provided vertebrates with the enzymatic functions necessary to achieve this degree of somatic hypermutability.

V(D)J recombination requires the activity of two enzymes, RAG1 and RAG2 (Oettinger et al. 1990; McBlane et al. 1995). These enzymes are encoded by two genes, *rag1* and *rag2*, that are tightly linked to each other (Lundblad and Blackburn 1993; Oettinger et al. 1990). During the course of investigating the activity of RAG1 and RAG2 it became apparent that the mechanism by which V(D)J recombination is achieved is strikingly similar to that by which transposons move (reviewed by Lewis and Wu 1997). As the mechanism was so similar, it was proposed that the two may have a common evolutionary origin (Spanopoulou et al. 1996). More recent work (Agrawal et al. 1998; Hiom et al. 1998), has demonstrated that RAG1 and RAG2 can, under certain circumstances, catalyze both inter- and intramolecular transposition in vitro, lending credence to the theory that the original function of these two enzymes was to catalyze transposition. Together, these data provide strong evidence that the system for V(D)J recombination arose from an ancestral transposition pathway that was modified to eliminate reinsertion of intervening transposon sequences (Fig. 5).

The generation of immunological diversity through V(D)J recombination is a key achievement of the vertebrate lineage. It is difficult to imagine the success of this lineage in the absence of a means by which pathogens could be successfully

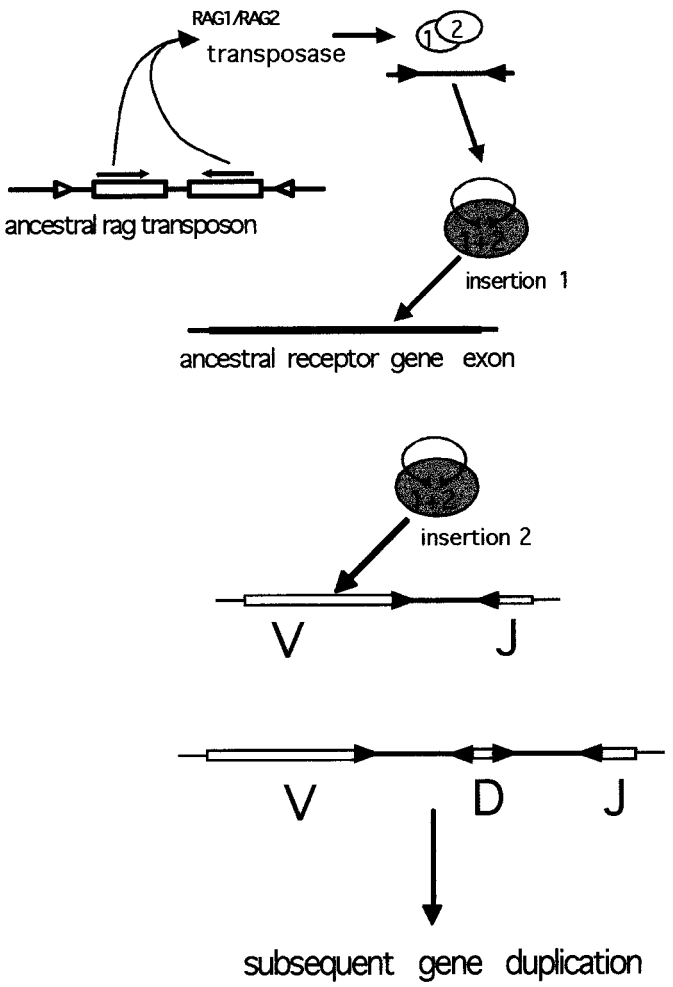

FIG. 5. Hypothetical depiction of the evolution of the V(D)J recombination system in vertebrates. Arrows above the ancestral ''*rag*'' transposon represent the two transcripts rag1 and rag2. Triangles represent recombination signals. The figure depicts the sequential addition of recombination signals into a receptor gene progenitor via transposition of the ''*rag*'' transposon or a nonautonomous derivative. The result of this subdivision of the ancestral host gene is hypothesized to result in the V, D, and J gene segments. Restoration of gene function through somatic excision of the transposons would be selectively favored, as would be the continued presence of the transposase source. Inaccuracies in the excision products (e.g., footprints) would contribute to the variability of the host gene. Subsequent gene duplication would contribute additional variation to the system (from Agrawal et al. 1998).

fought off. As it turns out, transposons almost certainly provided one of the mechanisms by which this could be achieved. The collaboration between host and parasite in this case is so intimate that they no longer can be viewed as separate entities. The capacity to break and recombine DNA, a key feature of transposon biology, appears to have been successfully coopted by the host, and the result has had a profound impact on the evolution of vertebrates.

Viral resistance.—There is good evidence that the ''domestication'' of some components of endogenous RT-based retroviruses has played a role in the evolution of resistance to viral infections. Reverse transcriptase-based viruses are ubiquitous, diverse, and occasionally, highly pathogenic. Thus, mechanisms to reduce their virulence are selectively advantageous. One of the more interesting of these mechanisms involves the recruitment of viral protein products by the host as a defense against exogenous viral attack.

The *Fv1* locus in mice was identified originally as a viral resistance gene (Rowe and Sato 1973). Mice carrying specific alleles of this gene confer resistance to infection by certain strains of murine leukemia virus. When cloned, the gene was identified as a homolog of the *Gag*-encoding gene of a newly described class of HERV-L-related mouse endogenous retroviruses (Benit et al. 1997; Best et al. 1996). The *Fv1*-gene product acts on the murine leukemia virus at a stage after entry into the cell, but before integration into the host genome (Pryciak and Varmus 1992). Genetic evidence suggests that it interacts with a component of the viral preintegration complex (Bowerman et al. 1989).

The host *Fv1* gene appears to have been recruited from an endogenous or exogenous retrovirus. Thus, what had been a purely selfish sequence of DNA was modified for the purpose of defense against infection. Interestingly, part of that modification probably involved the insertion of additional TEs, as the polyadenylation signal in all *Fv-1* alleles examined is within insertion elements (Best et al. 1996). Other viral resistance genes, such as *Fv-4* (Kozak et al. 1984; Ikeda and Sugimura 1989; Nihrane et al. 1997), which encodes a homolog of a viral *env* gene, have also proved to be portions of viral genomes, suggesting that viral recruitment may be a common way to augment host defenses.

Vivipary.—Viral recruitment may have played an important role in the evolution of vivipary as well. Viviparous organisms with active immune systems face a fundamental problem: the developing embryo represents a ''foreign'' organism within its mother. If unprotected, it would be subject to immunological attack by the maternal immune system. Although it is as yet unknown exactly how this is achieved, there are strong indications that endogenous retroviral sequences are involved. In all mammals examined to date, there are at least a few endogenous retroelements that are complete and active (Villarreal 1997). These elements are often well conserved and widespread within a given host species (Jurka et al. 1995; Singer et al. 1993), and are expressed at high levels in specific tissues. A highly conserved portion of the *env* gene of both endogenous and exogenous retroelements has been shown to be immunosuppressive (Denner 1998). Further, the tissues in which it is expressed, such as the placenta, are precisely those that one would expect if the purpose of the expression were to protect the embryo from attack (Huang and Calarco 1981; Larsson et al. 1994; Venables et al. 1995). Thus, it has been proposed that some endogenous retroelements in mammals have been recruited in order to make vivipary possible (Venables et al. 1995; Villarreal 1997). If true, this would be an excellent example of the host making use of a function, immunosuppression, that originally evolved in viruses for supremely selfish reasons. Certainly, immunosuppression in mammals could have evolved from existing host genes, but the presence of a diverse and efficient system, that was preexisting in the viral genome, apparently made that unnecessary.

An interesting analog for this putative immunosuppressive function in mammals exists in some endoparasatoid wasp

species (reviewed by Beckage 1998). In this case, wasp eggs are deposited in the larval host species along with large amounts of wasp-encoded DNA polydnavirus. The polydnavirus is an endogenous component of the wasp genome that is inactive in the wasp, but transcriptionally active in the larval host. The viral products act to prevent the normal immunological response of the larvae, thus allowing development of the parasitoid. Thus, the polydnavirus can be seen as an essential component of the wasp reproductive system.

Recombination template.—The human major histocompatibility (MHC) region provides an excellent example of the range of TE-induced changes that have contributed to host evolution over a long period of time. In particular, this region provides a good example of the effects of retroelements on genomic plasticity of the host in an area in which recombination is particularly valuable for generating diversity. In addition to more than 200 genes, densely clustered areas of retroelements have been identified in the MHC region (Andersson et al. 1998). The predominant classes are SINEs, LINEs, and HERVs (human endogenous retroviruses). The central region of the MHC contains two duplicated segments, peri-B (53 kb) and peri-C (48 kb) that are the result of a gene duplication that occurred prior to the divergence of the New World primate/human lineage split (Kulski et al. 1997). The identification of numerous fragmented and intact retroelements (*L1*, *Alu*, LTR, and *THE* sequences) flanking two genes located in a region flanked by (HLA)-B and tumor necrosis factor (TNF) suggested that these retroelements are involved in the duplication process (Gaudieri et al. 1997). Analysis of the retroelements contained within the HLA-B and HLA-C genes that reside within the duplicated segments has shown the utility of retroelements, such as *Alu* repeats in the dating of the duplication to 44–81 MYA (Kulski et al. 1997). This analysis has also pointed out the major role played by the mobility of these elements in the mutation and diversification of the HLA-B and HLA-C genes following duplication.

Molecular domestication.—The stationary *P* element-related gene clusters of *Drosophila guanche*, *D. madeirensis*, and *D. subobscura* provide an interesting example of molecular domestication (Miller et al. 1992, 1997). Each cluster unit consists of a cis-regulating section composed of different insertion sequences followed by the first three exons of a *P* element that encodes a 66 kDa ''repressor-like'' protein. In contrast to this normal repressor function, these stationary *P* element repeats appear to have evolved the function of transcription factors (Miller et al. 1995). Remarkably, the *D. guanche* P-protein produces an enhancer-like effect, rather than repressing canonical *P* element activity in transgenic *D. melanogaster* (W. Miller, pers. comm.). The insertion sequence which gave rise to the de novo A-type promoter of this *P*-gene cluster has recently been identified (Miller et al. 2000). This insertion belongs to a new MITE-like TE family, designated SGM, that is related to poorly characterized *IS* elements of other *obscura* group species.

Evidence that molecular domestication of a TE family may recur in a host lineage is provided by a truncated immobile *P* sequence cloned from *Drosophila tscasi*, a species of the *Drosophila montium* subgroup (Nouaud and Anxolabéhère 1997). This truncated element produces a polyadenylated RNA that also has a coding capacity for a 66-kDa ''repressor-like'' protein. This domestication event has led to the evolution of a new promoter and a new intron (Nouaud et al. 1999). Although the immobile *obscura* and *montium P* sequences were derived from the same ancestral mobile *P* element family, the structures of the flanking regions of these sequences indicates that they were produced by two completely independent evolutionary events.

It has been argued (Miller et al. 1999) that these examples of domestication represent more than just isolated examples of an infrequent process. Rather, the adaptive integration of a short piece of autonomous DNA into a complex regulatory network that occurs in molecular domestication might be considered to represent an evolutionary process that has been common in the evolution of increasingly complex forms of life.

Discussion and Conclusions

Much has been made about the distinction between ''selfish'' and ''altruistic'' DNA. Transposable element evolution, in all of its rich complexity, reveals the artificiality of this distinction. Instead of maintaining an unchanging relationship with hosts, transposons occupy a range of positions within a spectrum of relationships, ranging from extreme parasitism (e.g., the world-wide invasion of *P* elements in *D. melanogaster*) to unalloyed mutualism (e.g., the cooption of RAG1/RAG2 in the evolution of V(D)J recombination in vertebrates).

Transposable element evolution is primarily selfish (especially during the early stages of population invasion) in the sense that it is largely driven by selection in favor of competence to increase copy number. It is constrained, however, by selection against the negative effects of transposition, such as disruption of gene function, which has lead to a complex array of mechanisms to modulate the effects of TE activity. Further, the rapid ebb and flow of transposable elements within host genomes and the rapid pace of evolution that is possible within large populations of transposable elements in the absence of selection at the level of the host has resulted in the evolution of traits that, in the end, can prove quite beneficial to the host.

The position at which any given mobile element finds itself in the spectrum between pure parasite and essential component of the host genome depends both on the selective forces acting upon it at a given time in a given species and on a variety of chance factors. As selection can operate both at the level of the host and at the level of the gene simultaneously and the balance between these selective forces can vary with time and situation, characterization of mobile elements as selfish or beneficial in any global sense is not very meaningful.

It is not that TEs act as agents of evolutionary change in spite of their selfishness, nor that they have necessarily evolved in any proximal way to serve their hosts by increasing their ''evolvability.'' Rather, selection on TEs at the level of the gene has resulted in the evolution of a remarkably facile ''machine'' for the generation of a sophisticated range of molecular tools, some of which turn out to be selectively advantageous to the host. Selection at the level of the gene can be particularly beneficial to the host because evolution

at that level is not constrained by demands at every step for a trait to be selectively advantageous to the host. In this sense, TEs function as agents of evolutionary change, not in spite of their selfish characteristics, but because of them. Selection at the level of the gene on a diverse array of TEs provides a vast pool of rapidly evolving enzymatic and regulatory functions, all of which are potentially available to the host.

This process is illustrated quite clearly in the evolution of V(D)J recombination in vertebrates. In this case, presumably, the capability to splice and reshuffle DNA efficiently was developed by a putative transposon as a means of increasing its copy number. Long after that capability had been developed and refined it took on a role in the generation of immunological diversity. The competence to generate that diversity arose only because many of the early stages in its evolution were dominated by selection at the level of the gene. In a similar way, the *Fv-1* locus in mice evolved as a viral coat protein gene, but many of the properties that made it useful to viruses also made it effective in preventing viral infection. The host did not need to evolve a response de novo because it already had endogenous sequences available that had previously evolved the potential to provide that response.

Who are we and who are they, who is host, and who is parasite can be seen as a function of how selection is operating at any given time, and, because selection can operate at multiple levels simultaneously, these distinctions can become meaningless in specific cases. A host gene that functions in the repair of double-stranded gaps produced by a *P* element excision is a component of a selectively advantageous system of DNA repair, but it is also a component of a system to amplify *P* element copy number. A deletion derivative of a *P* element that suppresses *P* element activity is a selfish molecular parasite of autonomous *P* elements, but it is also a component of a system that limits *P* element amplification. Methyl transferase activity can act to modulate TE and viral activity, but it may also act to regulate host gene transcription during development. In each of these cases the ''function'' of the mechanism involved depends on the nature of the selective forces operating on a particular lineage.

From a broad perspective, TEs are increasingly being viewed as important players in the ability of genomes to enhance their own evolution and as a major source of tools for generating the diversity to respond successfully to changes in environment. The emergence of such tools has been termed second order selection (Arber 1991, 2000). The very properties that lead some to label TEs as ''junk DNA'' may have in fact provided genomes with a plasticity otherwise unavailable to them. Indeed, rather than resulting from selection, many of these properties may be the outcome of constructive neutral evolution (Stoltzfus 1999).

Given the opportunistic nature of evolution and recent evidence that morphological change is primarily a result of changes in regulatory networks, it would be surprising indeed if a ready-made tool-kit for the generation of diversity was not periodically exploited by the host species. We argue that the dynamic interplay of selection on transposable elements at various levels provides genomes with a broad and flexible capacity to evolve.

Acknowledgments

This work was supported by National Science Foundation (NSF) grant DEB 9815754. We thank A. Holyoake and J. Silva for comments on the manuscript.

Literature Cited

Agrawal, A., Q. M. Eastman, and D. G. Schatz. 1998. Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. Nature 394:744–751.

Ananiev, E. V., R. L. Phillips, and H. W. Rines. 1998. Chromosome-specific molecular organization of maize (*Zea mays* L.) centromeric regions. Proc. Natl. Acad. Sci. 95:13073–13078.

Anderson, R., R. J. Britten, and E. H. Davidson. 1994. Repeated sequence target sites for maternal DNA-binding proteins in genes activated in early sea urchin development. Dev. Biol. 163: 11–18.

Andersson, G., A. C. Svensson, N. Setterblad, and L. Rask. 1998. Retroelements in the human MHC class II region. Trends Genet. 14:109–114.

Arber, W. 1991. Elements in microbial evolution. J. Mol. Evol. 33: 4–12.

———. 2000. Genetic variation: molecular mechanisms and impact on microbial evolution. FEMS Microbiol Rev. 24:1–7.

Avramova, Z., A. Tikhonov, M. Chen, and J. L. Bennetzen. 1998. Matrix attachment regions and structural colinearity in the genomes of two grass species. Nucleic Acids Res. 26:761–767.

Banks, J. A., P. Masson, and N. Fedoroff. 1988. Molecular mechanisms in the developmental regulation of the maize *Suppressor-mutator* transposable element. Genes Dev. 2:1364–1380.

Barlow, D. P. 1993. Methylation and imprinting: from host defense to gene regulation? Science 260:309–310.

Batzer, M. A., S. S. Arcot, J. W. Phinney, M. Alegria-Hartman, D. H. Kass, S. M. Milligan, C. Kimpton, P. Gill, M. Hochmeister, P. A. Ioannou, R. J. Herrera, D. A. Boudreau, W. D. Scheer, B. J. Keats, P. L. Deininger, and M. Stoneking. 1996. Genetic variation of recent *Alu* insertions in human populations. J. Mol. Evol. 42:22–29.

Baum, M., and L. Clarke. 2000. Fission yeast homologs of human CENP-B have redundant functions affecting cell growth and chromosome segregation. Mol. Cell. Biol. 20:2852–2864.

Beckage, N. E. 1998. Modulation of immune responses to parasitoids by polydnaviruses. Parasitology 116(S):57–64.

Beebe, S. J., P. Salomonsky, T. Jahnsen, and Y. Li. 1992. The C gamma subunit is a unique isozyme of the cAMP-dependent protein kinase. J. Biol. Chem. 267:25505–25512.

Bell, G. 1993. The sexual nature of the eukaryote genome. J. Hered. 84:351–359.

Benit, L., N. De Parseval, J. F. Casella, I. Callebaut, A. Cordonnier, and T. Heidmann. 1997. Cloning of a new murine endogenous retrovirus, MuERV-L, with strong similarity to the human HERV-L element and with a *gag* coding sequence closely related to the *Fv1* restriction gene. J. Virol. 71:5652–5657.

Best, S., P. Le Tissier, G. Towers, and J. P. Stoye. 1996. Positional cloning of the mouse retrovirus restriction gene *Fv1*. Nature 382: 826–829.

Bestor, T. H. 1998. The host defense function of genomic methylation patterns. Novartis Found Symp 214:187–195.

Biemont, C., A. Tsitrone, C. Vieira, and C. Hoogland. 1997. Transposable element distribution in *Drosophila*. Genetics 147: 1997–1999.

Biessmann, H., and J. M. Mason. 1997. Telomere maintenance without telomerase. Chromosoma 106:63–69.

Bird, A. P. 1995. Gene number, noise reduction and biological complexity. Trends Genet. 11:94–100.

Bogarad, L. D., and M. W. Deem. 1999. A hierarchical approach to protein molecular evolution. Proc. Natl. Acad. Sci. USA 96: 2591–2595.

Bouma, J. E., and R. E. Lenski. 1988. Evolution of a bacteria/plasmid association. Nature 335:351–352.

Bowerman, B., P. O. Brown, J. M. Bishop, and H. E. Varmus. 1989.

A nucleoprotein complex mediates the integration of retroviral DNA. Genes Dev 3:469–478.
Britten, R. J. 1996a. DNA sequence insertion and evolutionary variation in gene regulation. Proc. Natl. Acad. Sci. USA 93: 9374–9377.
———. 1996b. Cases of ancient mobile element DNA insertions that now affect gene regulation. Mol. Phylo. Genet. Evol. 5: 13–17.
———. 1997. Mobile elements inserted in the distant past have taken on important functions. Gene 205:177–182.
Britten, R. J., and E. H. Davidson. 1969. Gene regulation for higher cells: a theory. Science 165:349–358.
———. 1971. Repetitive and non-repetitive DNA sequences and a speculation on the origin of evolutionary novelty. Q. Rev. Biol. 46:111–138.
Brodbeck, D., R. Amherd, P. Callaerts, E. Hintermann, U. A. Meyer, and M. Affolter. 1998. Molecular and biochemical characterization of the aaNAT1 (Dat) locus in *Drosophila melanogaster*: differential expression of two gene products. DNA Cell Biol 17: 621–633.
Brookfield, J. F. Y. 1995. Transposable elements as selfish DNA. Pp. 130–153 *in* D. J. Sherratt, eds. Mobile genetic elements. IRL Press, Oxford, U.K.
Brosius, J. 1999a. Vertebrate genomes were forged by massive bombardments with retroelements and retrosequences. Genetica 107: 209–238.
———. 1999b. RNAs from all categories generate retrosequences that may be exapted as novel genes or regulatory elements. Gene 238:115–134.
Bull, J. J., and W. R. Rice. 1991. Distinguishing mechanisms for the evolution of co-operation. J. Theor. Biol. 149:63–74.
Bureau, T. E., and S. R. Wessler. 1992. *Tourist*: a large family of small inverted repeat elements frequently associated with maize genes. Plant Cell 4:1283–1294.
———. 1994. *Stowaway*: a new family of inverted repeat elements associated with the genes of both monocotyledonous and dicotyledonous plants. Plant Cell 6:907–916.
Burwinkel, B., and M. W. Kilimann. 1998. Unequal homologous recombination between LINE-1 elements as a mutational mechanism in human genetic disease. J. Mol. Biol. 277:513–517.
Cambareri, E. B., R. Aisner, and J. Carbon. 1998. Structure of the chromosome. VII. Centromere region in *Neurospora crassa*: degenerate transposons and simple repeats. Mol. Cell. Biol. 18: 5465–5477.
Canapa, A., M. Barucca, P. N. Cerioni, and E. Olmo. 2000. A satellite DNA containing CENP-B box-like motifs is present in the Antarctic scallop *Adamussium colbecki*. Gene 247:175–180.
Capy, P., D. Anxolabéhère, and T. Langin. 1994. The strange phylogenies of transposable elements: are horizontal transfers the only explanation? Trends Genet. 10:7–12.
Capy, P., C. Bazin, D. Higuet, and T. Langin. 1997a. Dynamics and evolution of transposable elements. Landes Bioscience, Austin, TX.
Capy, P., T. Langin, D. Higuet, P. Maurer, and C. Bazin. 1997b. Do the integrases of LTR-retrotransposons and class II element transposases have a common ancestor? Genetica 100:63–72.
Casacuberta, E., J. M. Casacuberta, P. Puigdomenech, and A. Monfort. 1998. Presence of miniature inverted-repeat transposable elements (MITEs) in the genome of *Arabidopsis thaliana*: characterisation of the *Emigrant* family of elements. Plant. J. 16: 79–85.
Cavalier-Smith, T. 1982. Skeletal DNA and the evolution of genome size. Annu. Rev. Biophys. Bioeng. 11:273–302.
———. 1991. Intron phylogeny: a new hypothesis. Trends Genet 7:145–148.
Charlesworth, B., and C. H. Langley. 1986. The evolution of self-regulated transposition of transposable elements. Genetics 112: 359–383.
Charlesworth, B., P. Sniegowski, and W. Stephan. 1994. The evolutionary dynamics of repetitive DNA in eukaryotes. Nature 371: 215–220.
Charlesworth, B., C. H. Langley, and P. D. Sniegowski. 1997. Transposable element distributions in *Drosophila*. Genetics 147: 1993–1995.
Cho, Y., Y. L. Qiu, P. Kuhlman, and J. D. Palmer. 1998. Explosive invasion of plant mitochondria by a group I intron. Proc. Natl. Acad. Sci. USA 95:14244–14249.
Civardi, L., Y. Xia, K. J. Edwards, P. S. Schnable, and B. J. Nikolau. 1994. The relationship between genetic and physical distances in the cloned a1-sh2 interval of the *Zea mays* L. genome. Proc. Natl. Acad. Sci. USA 91:8268–8272.
Codon, A. C., T. Benitez, and M. Korhola. 1997. Chromosomal reorganization during meiosis of *Saccharomyces cerevisiae* baker's yeasts. Curr Genet 32:247–259.
Colot, V., V. Haedens, and J. L. Rossignol. 1998. Extensive, non-random diversity of excision footprints generated by *Ds*-like transposon *Ascot-1* suggests new parallels with V(D)J recombination. Mol. Cell. Biol. 18:4337–4346.
Crameri, A., S. A. Raillard, E. Bermudez, and W. P. Stemmer. 1998. DNA shuffling of a family of genes from diverse species accelerates directed evolution. Nature 391:288–291.
Cresse, A. D., S. H. Hulbert, W. E. Brown, J. R. Lucas, and J. L. Bennetzen. 1995. Mu1-related transposable elements of maize preferentially insert into low copy number DNA. Genetics 140: 315–324.
Csink, A. K., and S. Henikoff. 1998. Something from nothing: the evolution and utility of satellite repeats [see comments]. Trends Genet. 14:200–204.
Curcio, M. J., and M. Belfort. 1996. Retrohoming: cDNA-mediated mobility of group II introns requires a catalytic RNA. Cell 84: 9–12.
Currie, D. B., T. F. Mackay, and L. Partridge. 1998. Pervasive effects of *P* element mutagenesis on body size in *Drosophila melanogaster*. Genet. Res. 72:19–24.
Cuypers, H., S. Dash, P. A. Peterson, H. Saedler, and A. Gierl. 1988. The defective *En-I102* element encodes a product reducing the mutability of the *En-Spm* system of *Zea mays*. Embo J. 7: 2953–2960.
Daniels, S. B., K. R. Peterson, L. D. Strausbaugh, M. G. Kidwell, and A. Chovnick. 1990. Evidence for horizontal transmission of the *P* transposable element between *Drosophila* species. Genetics 124:339–355.
Dawe, R. K., and W. Z. Cande. 1996. Induction of centromeric activity in maize by suppressor of meiotic drive 1. Proc. Natl. Acad. Sci. USA 93:8512–8517.
Dawkins, R. 1974. The selfish gene. Oxford Univ. Press, Oxford, U.K.
Deininger, P. L. 1989. SINES: Short interspersed repeated DNA elements in higher eukaryotes. Pp. 619–636 *in* D. Berg and M. Howe, eds. Mobile DNA. American Society of Microbiology, Washington, DC.
Deininger, P. L., and M. A. Batzer. 1999. Alu repeats and human disease. Mol Genet. Metab. 67:183–193.
Denner, J. 1998. Immunosuppression by retroviruses: implications for xenotransplantation. Ann. N.Y. Acad. Sci. 862:75–86.
Dennis, E. S., and R. I. Brettell. 1990. DNA methylation of maize transposable elements is correlated with activity. Philos. Trans. R. Soc. Lond. B Biol. Sci. 326:217–229.
Dimitri, P., and N. Junakovic. 1999. Revising the selfish DNA hypothesis: new evidence on accumulation of transposable elements in heterochromatin. Trends Genet. 15:123–124.
Dimitri, P., B. Arca, L. Berghella, and E. Mei. 1997. High genetic instability of heterochromatin after transposition of the LINE-like I factor in Drosophila melanogaster. Proc. Natl. Acad. Sci. USA 94:8052–8057.
Donlin, M. J., D. Lisch, and M. Freeling. 1995. Tissue-specific accumulation of MURB, a protein encoded by *MuDR*, the autonomous regulator of the *Mutator* transposable element family. Plant Cell 7:1989–2000.
Doolittle, W. F., and C. Sapienza. 1980. Selfish genes, the phenotype paradigm and genome evolution. Nature 284:601–603.
Dupressoir, A., and T. Heidmann. 1996. Germ line-specific expression of intracisternal A-particle retrotransposons in transgenic mice. Mol. Cell. Biol. 16:4495–4503.
Dwyer, D. S. 1998. Assembly of exons from unitary transposable

genetic elements: implications for the evolution of protein-protein interactions. J. Theor. Biol. 194:11–27.
Earnshaw, W. C., K. F. Sullivan, P. S. Machlin, C. A. Cooke, D. A. Kaiser, T. D. Pollard, N. F. Rothfield, and D. W. Cleveland. 1987. Molecular cloning of cDNA for CENP-B, the major human centromere autoantigen. J. Cell. Biol. 104:817–829.
Eickbush, T. H. 1994. Origin and evolutionary relationship of retroelements. Pp. 121–157 *in* S. S. Morse, eds. Evolutionary biology of viruses. Raven Press.
———. 1997. Telomerase and retrotransposons: which came first? Science 277:911–912.
Engels, W. R. 1996. *P* elements in *Drosophila*. Pp. 103–123 *in* H. Saedler and A. Gierl, eds. Transposable elements. Springer-Verlag, Berlin.
Engels, W. R., and C. R. Preston. 1984. Formation of chromosome rearrangements by *P* factors in *Drosophila*. Genetics 107: 657–678.
Esnault, C., J. Maestre, and T. Heidmann. 2000. Human LINE retrotransposons generate processed pseudogenes. Nat. Genet. 24: 363–367.
Evgen'ev, M. B., H. Zelentsova, H. Poluectova, G. T. Lyozin, V. Veleikodvorskaja, K. I. Pyatkov, L. A. Zhivotovsky, and M. G. Kidwell. 2000. Mobile elements and chromosomal evolution in the *virilis* group of *Drosophila*. Proc. Natl. Acad. Sci. USA. 97: 11337–11342.
Ewald, P. W. 1987. Transmission modes and evolution of the parasitism-mutualism continuum. Ann. N.Y. Acad. Sci. 503: 295–306.
———. 1995. The evolution of virulence: a unifying link between parasitology and ecology. J. Parasitol. 81:659–669.
Fanti, L., D. R. Dorer, M. Berloco, S. Henikoff, and S. Pimpinelli. 1998. Heterochromatin protein 1 binds transgene arrays. Chromosoma 107:286–292.
Fedoroff, N. 1989. Maize Transposable Elements. Pp. 375–411 *in* D. Berg and M. Howe, eds. Mobile elements. American Society of Microbiology, Washington, DC.
Fedoroff, N., M. Schlappi, and R. Raina. 1995. Epigenetic regulation of the maize *Spm* transposon. Bioessays 17:291–297.
Finnegan, D. J. 1989. Eukaryotic transposable elements and genome evolution. Trends Genet. 5:103–107.
Finnegan, E. J., W. J. Peacock, and E. S. Dennis. 1996. Reduced DNA methylation in *Arabidopsis thaliana* results in abnormal plant development. Proc. Natl. Acad. Sci. USA 93:8449–8454.
Fitch, W. M. 1996. The variety of human virus evolution. Mol. Phylogenet. Evol. 5:247–258.
Fitzpatrick, B., and J. A. Sved. 1986. High levels of fitness modifiers induced by hybrid dysgenesis in *Drosophila melanogaster*. Genet. Res. 48:89–94.
Flavell, A. J. 1995. Retroelements, reverse transcriptase and evolution. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 110: 3–15.
Force, A., M. Lynch, F. B. Pickett, A. Amores, Y. Yan, and J. Postlethwait. 1999. Preservation of duplicate genes by complementary, degenerative mutations. Genetics 151:1531–1545.
Foss, K., B. Landmark, B. Skalhegg, K. Tashken, E. Jellum, V. Hansson, and T. Jahnsen. 1994. Characterization of in-vitro-translated human regulatory and catalytic subunits of cAMP-dependent protein kinases. Eur. J. Biochem. 220:217–223.
Fowler, K. J., D. F. Hudson, L. A. Salamonsen, S. R. Edmondson, E. Earle, M. C. Sibson, and K. H. Choo. 2000. Uterine dysfunction and genetic modifiers in centromere protein B-deficient mice. Genome Res. 10:30–41.
Frigerio, G., M. Burri, D. Bopp, S. Baumgartner, and M. Noll. 1986. Structure of the segmentation gene paired and the *Drosophila* PRD gene set as part of a gene network. Cell 47:735–746.
Gaudieri, S., C. Leelayuwat, D. C. Townend, J. K. Kulski, and R. L. Dawkins. 1997. Genomic characterization of the region between HLA-B and TNF: implications for the evolution of multicopy gene families. J. Mol. Evol. 44:S147–154.
Gillies, S. D., S. L. Morrison, V. T. Oi, and S. Tonegawa. 1983. A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell 33:717–728.
Gimble, F. S. 2000. Invasion of a multitude of genetic niches by mobile endonuclease genes. FEMS Microbiol. Lett. 185:99–107.
Ginzburg, L. R., P. M. Bingham, and S. Yoo. 1984. On the theory of speciation induced by transposable elements. Genetics 107: 331–341.
Girard, L., and M. Freeling. 1999. Regulatory changes as a consequence of transposon insertion. Dev. Genet. 25:291–296.
Good, A. G., G. Meister, H. Brock, T. A. Grigliatti, and D. Hickey. 1989. Rapid spread of transposable *P* elements in experimental populations of *Drosophila melanogaster*. Genetics 122:387–396.
Gould, S. J. 1983. Hen's teeth and horse's toes. W. W. Norton, New York.
Gregory, T. R., and P. D. Hebert. 1999. The modulation of DNA content: proximate causes and ultimate consequences. Genome Res. 9:317–324.
Haaf, T., A. G. Mater, J. Wienberg, and D. C. Ward. 1995. Presence and abundance of CENP-B box sequences in great ape subsets of primate-specific alpha-satellite DNA. J. Mol. Evol. 41: 487–491.
Heikkinen, E., V. Launonen, E. Muller, and L. Bachmann. 1995. The *pvB370 Bam*III satellite DNA family of the *Drosophila virilis* group and its evolutionary relation to mobile dispersed genetic *pDv* elements. J. Mol. Evol. 41:604–614.
Henikoff, S. 1996. Position-effect variegation in *Drosophila*: recent progress. Pp. 319–334 *in* V. E. A. Russo, R. A. Martienssen, and A. D. Riggs, eds. Epigenetic mechanisms of gene regulation. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Henikoff, S., E. A. Greene, S. Pietrokovski, P. Bork, T. K. Attwood, and L. Hood. 1997. Gene families: the taxonomy of protein paralogs and chimeras. Science 278:609–614.
Heslop-Harrison, J. S., M. Murata, Y. Ogura, T. Schwarzacher, and F. Motoyoshi. 1999. Polymorphisms and genomic organization of repetitive DNA from centromeric regions of Arabidopsis chromosomes. Plant Cell 11:31–42.
Hickey, D. A. 1982. Selfish DNA: a sexually-transmitted nuclear parasite. Genetics 101:519–531.
Hickey, D. A., and M. R. Rose. 1988. The role of gene transfer in the evolution of eukaryotic sex. Pp. 161–175 *in* R. Michod and B. Levin, eds. The evolution of sex: an examination of current ideas. Sinauer, Sunderland, MA.
Higashiyama, T., Y. Noutoshi, M. Fujie, and T. Yamada. 1997. Zepp, a LINE-like retrotransposon accumulated in the Chlorella telomeric region. Embol. J. 16:3715–3723.
Hiom, K., M. Melek, and M. Gellert. 1998. DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations. Cell 94:463–470.
Holliday, R., and J. E. Pugh. 1975. DNA modification mechanisms and gene activity during development. Science 187:226–232.
Holmes, S. E., B. A. Dombroski, C. M. Krebs, C. D. Boehm, and H. H. Kazazian, Jr. 1994. A new retrotransposable human L1 element from the LRE2 locus on chromosome 1q produces a chimaeric insertion. Nat. Genet. 7:143–148.
Huang, T. T., Jr., and P. G. Calarco. 1981. Immunoprecipitation of intracisternal A-particle-associated antigens from preimplantation mouse embryos. J. Natl. Cancer Inst. 67:1129–1134.
Ikeda, H., and H. Sugimura. 1989. *Fv-4* resistance gene: a truncated endogenous murine leukemia virus with ecotropic interference properties. J. Virol. 63:5405–5412.
Izsvak, Z., Z. Ivics, N. Shimoda, D. Mohn, H. Okamoto, and P. B. Hackett. 1999. Short inverted-repeat transposable elements in teleost fish and implications for a mechanism of their amplification. J. Mol. Evol. 48:13–21.
James, R., C. Kleanthous, and G. R. Moore. 1996. The biology of E colicins: paradigms and paradoxes. Microbiology 142: 1569–1580.
Jordan, I. K., and J. F. McDonald. 1998. Evidence for the role of recombination in the regulatory evolution of *Saccharomyces cerevisiae* Ty elements. J. Mol. Evol. 47:14–20.
Jurka, J., E. Zietkiewicz, and D. Labuda. 1995. Ubiquitous mammalian-wide interspersed repeats (MIRs) are molecular fossils from the mesozoic era. Nucleic Acids Res. 23:170–175.
Kapitonov, V. V., G. P. Holmquist, and J. Jurka. 1998. L1 repeat

is a basic unit of heterochromatin satellites in Cetaceans. Mol. Biol. Evol. 15:611–612.
Kazazian, H. H., Jr., and J. V. Moran. 1998. The impact of L1 retrotransposons on the human genome. Nat. Genet. 19:19–24.
Kidwell, M. G. 1993. Lateral transfer in natural populations of eukaryotes. Ann. Rev. Genet. 27:235–256.
———. 1994. The Wilhelmine E. Key 1991 Invitational Lecture. The evolutionary history of the *P* family of transposable elements. J. Hered. 85:339–346.
Kidwell, M. G., and D. Lisch. 1997. Transposable elements as sources of variation in animals and plants. Proc. Natl. Acad. Sci. USA 94:7704–7711.
———. 2000. Transposable elements and host genome evolution. Trends Ecol. Evol. 15:95–99.
Kim, J., J. A. Martignetti, M. R. Shen, J. Brosius, and P. Deininger. 1994. Rodent BC1 RNA gene as a master gene for ID element amplification. Proc. Natl. Acad. Sci. USA 91:3607–3611.
Kim, J. M., S. Vanguri, J. D. Boeke, A. Gabriel, and D. F. Voytas. 1998. Transposable elements and genome organization: A comprehensive survey of retrotransposons revealed by the complete *Saccharomyces cerevisiae* genome sequence. Genome Res. 8: 464–478.
Kimura, M. 1983. The neutral theory of molecular evolution. Cambridge Univ. Press, Cambridge, U.K.
Kipling, D., and P. E. Warburton. 1997. Centromeres, CENP-B and *Tigger* too. Trends Genet. 13:141–145.
Kiyasu, P. K., and M. G. Kidwell. 1984. Hybrid dysgenesis in *Drosophila melanogaster*: the evolution of mixed P and M populations maintained at high temperature. Genet. Res. 44: 251–259.
Klobutcher, L. A., and G. Herrick. 1995. Consensus inverted terminal repeat sequence of Paramecium IESs: resemblance to termini of *Tc1*-related and *Euplotes Tec* transposons. Nucleic Acids Res. 23:2006–2013.
———. 1997. Developmental genome reorganization in ciliated protozoa: the transposon link. Prog. Nucleic Acid Res. Mol. Biol. 56:1–62.
Kozak, C. A., N. J. Gromet, H. Ikeda, and C. E. Buckler. 1984. A unique sequence related to the ecotropic murine leukemia virus is associated with the *Fv-4* resistance gene. Proc. Natl. Acad. Sci. USA 81:834–837.
Krawiec, S., and M. Riley. 1990. Organization of the bacterial chromosome. Microbiol. Rev. 54:502–539.
Kulski, J. K., S. Gaudieri, M. Bellgard, L. Balmer, K. Giles, H. Inoko, and R. L. Dawkins. 1997. The evolution of MHC diversity by segmental duplication and transposition of retroelements. J. Mol. Evol. 45:599–609.
Labrador, M., and V. G. Corces. 1997. Transposable element-host interactions: regulation of insertion and excision. Annu. Rev. Genet. 31:381–404.
Larsson, E., A. C. Andersson, and B. O. Nilsson. 1994. Expression of an endogenous retrovirus (ERV3 HERV-R) in human reproductive and embryonic tissues—evidence for a function for envelope gene products. Ups. J. Med. Sci. 99:113–120.
Larsson, E., P. Venables, A. C. Andersson, W. Fan, S. Rigby, J. Botling, F. Oberg, M. Cohen, and K. Nilsson. 1997. Tissue and differentiation specific expression on the endogenous retrovirus ERV3 (HERV-R) in normal human tissues and during induced monocytic differentiation in the U-937 cell line. Leukemia 3: 142–144.
Le, S., J. K. Moore, J. E. Haber, and C. W. Greider. 1999. RAD50 and RAD51 define two pathways that collaborate to maintain telomeres in the absence of telomerase. Genetics 152:143–152.
Lee, J. K., J. A. Huberman, and J. Hurwitz. 1997. Purification and characterization of a CENP-B homologue protein that binds to the centromeric K-type repeat DNA of *Schizosaccharomyces pombe*. Proc. Natl. Acad. Sci. USA 94:8427–8432.
Levis, R. W., R. Ganesan, K. Houtchens, L. A. Tolar, and F. M. Sheen. 1993. Transposons in place of telomeric repeats at a *Drosophila* telomere. Cell 75:1083–1093.
Lewis, S. M., and G. E. Wu. 1997. The origins of V(D)J recombination. Cell 88:159–162.
Li, E., T. H. Bestor, and R. Jaenisch. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell 69:915–926.
Li, E., C. Beard, and R. Jaenisch. 1993. Role for DNA methylation in genomic imprinting. Nature 366:362–365.
Li, T., J. Spearow, C. M. Rubin, and C. W. Schmid. 1999. Physiological stresses increase mouse short interspersed element (SINE) RNA expression in vivo. Gene 239:367–372.
Li, W.-H. 1997. Molecular evolution. Sinauer Associates, Sunderland, MA.
Lim, J. K. 1988. Intrachromosomal rearrangements mediated by *hobo* transposons in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA 85:9153–9157.
Lin, Y. J., L. Seroude, and S. Benzer. 1998. Extended life-span and stress resistance in the *Drosophila* mutant *methuselah*. Science 282:943–946.
Lin, X., S. Kaul, S. Rounsley, T. P. Shea, M. I. Benito, C. D. Town, C. Y. Fujii, T. Mason, C. L. Bowman, M. Barnstead, T. V. Feldblyum, C. R. Buell, K. A. Ketchum, J. Lee, C. M. Ronning, H. L. Koo, K. S. Moffat, L. A. Cronin, M. Shen, G. Pai, S. Van Aken, L. Umayam, L. J. Tallon, J. E. Gill, M. D. Adams, A. J. Carrera, T. H. Creasy, H. M. Goodman, C. R. Somerville, G. P. Copenhaver, D. Preuss, W. C. Nierman, O. White, J. A. Eisen, S. L. Salzberg, C. M. Fraser, and J. C. Venter. 1999. Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. Nature 402:761–768.
Lister, C., D. Jackson, and C. Martin. 1993. Transposon-induced inversion in *Antirrhinum* modifies *nivea* gene expression to give a novel flower color pattern under the control of cycloidearadialis. Plant Cell 5:1541–1553.
Lundblad, V., and E. H. Blackburn. 1993. An alternative pathway for yeast telomere maintenance rescues est1-senescence. Cell 73:347–360.
Mackay, T. F. 1986. Transposable element-induced fitness mutations in *Drosophila melanogaster*. Genet. Res. 48:77–87.
Mackay, T. F. 1989. Transposable elements and fitness in *Drosophila melanogaster*. Genome 31:284–295.
Mackay, T. F., R. F. Lyman, and M. S. Jackson. 1992. Effects of *P* element insertions on quantitative traits in *Drosophila melanogaster*. Genetics 130:315–332.
Maeshiro, T., and M. Kimura. 1998. The role of robustness and changeability on the origin and evolution of genetic codes. Proc. Natl. Acad. Sci. USA 95:5088–5093.
Martienssen, R., A. Barkan, W. C. Taylor, and M. Freeling. 1990. Somatically heritable switches in the DNA modification of *Mu* transposable elements monitored with a suppressible mutant in maize. Genes Dev. 4:331–343.
Mathiopoulos, K. D., A. della Torre, V. Predazzi, V. Petrarca, and M. Coluzzi. 1998. Cloning of inversion breakpoints in the *Anopheles gambiae* complex traces a transposable element at the inversion junction. Proc. Natl. Acad. Sci. USA 95:12444–12449.
Matzke, A. J., and M. A. Matzke. 1998. Position effects and epigenetic silencing of plant transgenes. Curr. Opin. Plant. Biol. 1: 142–148.
McBlane, J. F., D. C. van Gent, D. A. Ramsden, C. Romeo, C. A. Cuomo, M. Gellert, and M. A. Oettinger. 1995. Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. Cell 83:387–395.
McClintock, B. 1956. Controlling elements and the gene. Cold Spring Harbor Symp. Quant. Biol. 21:197–216.
———. 1958. The *Suppressor-mutator* system of control of gene action in maize. Carnegie Inst. Wash. Yearbook 57:415–429.
———. 1978. Mechanisms that rapidly reorganize the genome. Stadler Symp. 10:25–47.
———. 1984. The significance of responses of the genome to challenge. Science 226:792–801.
McClure, M. A. 1991. Evolution of retroposons by acquisition or deletion of retrovirus-like genes. Mol. Biol. Evol. 8:835–856.
———. 1999. The retroid agents: disease, function, and evolution. Pp. 163–195 *in* E. Domingo, R. Webster, and J. Holland, eds. Origin and evolution of viruses. Academic Press, London.
McDonald, J. F. 1990. Macroevolution and retroviral elements. BioScience 40:183–191.

———. 1998. Transposable elements, gene silencing and macroevolution. Trends Ecol. Evol. 13:94–95.
McFadden, J., and G. Knowles. 1997. Escape from evolutionary stasis by transposon-mediated deleterious mutations. J. Theor. Biol. 186:441–447.
McNaughton, J. C., G. Hughes, W. A. Jones, P. A. Stockwell, H. J. Klamut, and G. B. Petersen. 1997. The evolution of an intron: analysis of a long, deletion-prone intron in the human dystrophin gene. Genomics 40:294–304.
Meehan, R., J. Lewis, S. Cross, X. Nan, P. Jeppesen, and A. Bird. 1992. Transcriptional repression by methylation of CpG. J. Cell. Sci. Suppl. 16:9–14.
Miller, W. J., S. Hagemann, E. Reiter, and W. Pinsker. 1992. *P*-element homologous sequences are tandemly repeated in the genome of *Drosophila guanche*. Proc. Natl. Acad. Sci. USA 89: 4018–4022.
Miller, W. J., N. Paricio, S. Hagemann, M. J. Martinez-Sebastian, W. Pinsker, and R. de Frutos. 1995. Structure and expression of clustered *P* element homologues in *Drosophila subobscura* and *Drosophila guanche*. Gene 156:167–174.
Miller, W. J., J. F. McDonald, and W. Pinsker. 1997. Molecular domestication of mobile elements. Genetica 100:261–270.
Miller, W. J., J. F. McDonald, D. Nouaud, and D. Anxolabéhère. 1999. Molecular domestication—more than a sporadic episode in evolution? Genetica 107:197–207.
Miller, W. J., A. Nagel, J. Bachmann, and L. Bachmann. 2000. Evolutionary dynamics of the SGM transposon family in the *Drosophila obscura* species group. Mol. Biol. Evol. 17: 1597–1609.
Misra, S., and D. C. Rio. 1990. Cytotype control of *Drosophila P* element transposition: the 66 kd protein is a repressor of transposase activity. Cell 62:269–284.
Monk, M., M. Boubelik, and S. Lehnert. 1987. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. Development 99:371–382.
Moore, J. C., H. M. Jin, O. Kuchner, and F. H. Arnold. 1997. Strategies for the in vitro evolution of protein function: enzyme evolution by random recombination of improved sequences. J. Mol. Biol. 272:336–347.
Moore, J. K., and J. E. Haber. 1996. Capture of retrotransposon DNA at the sites of chromosomal double-strand breaks. Nature 383:644–646.
Moran, J. V., R. J. DeBerardinis, and H. H. Kazazian, Jr. 1999. Exon shuffling by *L1* retrotransposition. Science 283: 1530–1534.
Muiznieks, I., and W. Doerfler. 1994. The topology of the promoter of RNA polymerase II- and III-transcribed genes is modified by the methylation of 5′-CG-3′ dinucleotides. Nucleic Acids Res. 22:2568–2575.
Muller, F., B. Chang, S. Albert, N. Fischer, L. Tora, and U. Strahle. 1999. Intronic enhancers control expression of zebrafish sonic hedgehog in floor plate and notochord. Development 126: 2103–2116.
Nakamura, T. M., and T. R. Cech. 1998. Reversing time: origin of telomerase. Cell 92:587–590.
Nakamura, T. M., G. B. Morin, K. B. Chapman, S. L. Weinrich, W. H. Andrews, J. Lingner, C. B. Harley, and T. R. Cech. 1997. Telomerase catalytic subunit homologs from fission yeast and human. Science 277:955–959.
Natvig, D. O., G. May, and J. W. Taylor. 1984. Distribution and evolutionary significance of mitochondrial plasmids in *Neurospora* spp. J. Bacteriol. 159:288–293.
Nee, S., and J. Maynard. 1990. The evolutionary biology of molecular parasites. Parasitology 100(S):5–18.
Nihrane, A., I. Lebedeva, M. S. Lyu, K. Fujita, and J. Silver. 1997. Secretion of a murine retroviral *Env* associated with resistance to infection. J. Gen. Virol. 78:785–793.
Nonomura, K. I., and N. Kurata. 1999. Organization of the 1.9-kb repeat unit RCE1 in the centromeric region of rice chromosomes. Mol. Gen. Genet. 261:1–10.
Nouaud, D., and D. Anxolabéhère. 1997. P element domestication: a stationary truncated *P* element may encode a 66-kDa repressor-like protein in the *Drosophila montium* species subgroup. Mol. Biol. Evol. 14:1132–1344.
Nouaud, D., B. Boeda, L. Levy, and D. Anxolabéhère. 1999. A *P* element has induced intron formation in *Drosophila*. Mol. Biol. Evol. 16:1503–1510.
Nuzhdin, S. V., and T. F. Mackay. 1994. Direct determination of retrotransposon transposition rates in *Drosophila melanogaster*. Genet. Res. 63:139–144.
O'Neill, R. J., M. J. O'Neill, and J. A. Graves. 1998. Undermethylation associated with retroelement activation and chromosome remodeling in an interspecific mammalian hybrid. Nature 393:68–72.
Oettinger, M. A., D. G. Schatz, C. Gorka, and D. Baltimore. 1990. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. Science 248:1517–1523.
Orgel, L. E., and F. H. Crick. 1980. Selfish DNA: the ultimate parasite. Nature 284:604–607.
Pardue, M. L., O. N. Danilevskaya, K. Lowenhaupt, F. Slot, and K. L. Traverse. 1996. *Drosophila* telomeres: new views on chromosome evolution. Trends Genet. 12:48–52.
Pearce, S. R., U. Pich, G. Harrison, A. J. Flavell, J. S. Heslop-Harrison, I. Schubert, and A. Kumar. 1996. The *Ty1-copia* group retrotransposons of *Allium cepa* are distributed throughout the chromosomes but are enriched in the terminal heterochromatin. Chromosome Res. 4:357–364.
Petrov, D. A., E. R. Lozovskaya, and D. L. Hartl. 1996. High intrinsic rate of DNA loss in *Drosophila*. Nature 384:346–349.
Pickeral, O. K., W. Makaowski, M. S. Boguski, and J. D. Boeke. 2000. Frequent human genomic DNA transduction driven by LINE-1 retrotransposition. Genome Res. 10:411–415.
Pimpinelli, S., M. Berloco, L. Fanti, P. Dimitri, S. Bonaccorsi, E. Marchetti, R. Caizzi, C. Caggese, and M. Gatti. 1995. Transposable elements are stable structural components of *Drosophila melanogaster* heterochromatin. Proc. Natl. Acad. Sci. USA 92: 3804–3808.
Pirrotta, V. 1996. Stable chromatin states regulating homeotic genes in Drosophila. Pp. 489–505 *in* V. E. A. Russo, R. A. Martienssen, and A. D. Riggs, eds. Epigentic Mechanisms of Gene Regulation. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Prior, C., L. Tizzani, H. Fukuhara, and M. Wesolowski-Louvel. 1996. *RAG3* gene and transcriptional regulation of the *pyruvate decarboxylase* gene in *Kluyveromyces lactis*. Mol. Microbiol. 20: 765–772.
Pryciak, P. M., and H. E. Varmus. 1992. *Fv-1* restriction and its effects on murine leukemia virus integration in vivo and in vitro. J. Virol. 66:5959–5966.
Quentin, Y. 1992. Origin of the *Alu* family: a family of *Alu*-like monomers gave birth to the left and the right arms of the *Alu* elements. Nucleic Acids Res. 20:3397–3401.
Reik, W., and E. R. Maher. 1997. Imprinting in clusters: lessons from Beckwith-Wiedemann syndrome. Trends Genet. 13: 330–334.
Reinton, N., T. B. Haugen, S. Orstavik, B. S. Skalhegg, V. Hansson, T. Jahnsen, and K. Tasken. 1998. The gene encoding the C gamma catalytic subunit of cAMP-dependent protein kinase is a transcribed retroposon. Genomics 49:290–297.
Robertson, H. M., and W. R. Engels. 1989. Modified *P* elements that mimic the P cytotype in *Drosophila melanogaster*. Genetics 123:815–824.
Robertson, H. M., and D. J. Lampe. 1995. Recent horizontal transfer of a *mariner* transposable element among and between Diptera and Neuroptera. Mol. Biol. Evol. 12:850–862.
Rowe, W. P., and H. Sato. 1973. Genetic mapping of the *Fv-1* locus of the mouse. Science 180:640–641.
Sanford, J. P., H. J. Clark, V. M. Chapman, and J. Rossant. 1987. Differences in DNA methylation during oogenesis and spermatogenesis and their persistence during early embryogenesis in the mouse. Genes Dev. 1:1039–1046.
SanMiguel, P., A. Tikhonov, Y. K. Jin, N. Motchoulskaia, D. Zakharov, A. Melake-Berhan, P. S. Springer, K. J. Edwards, M. Lee, Z. Avramova, and J. L. Bennetzen. 1996. Nested retrotransposons in the intergenic regions of the maize genome. Science 274: 765–768.

Schiefelbein, J. W., D. B. Furtek, H. K. Dooner, and O. E. Nelson, Jr. 1988. Two mutations in a maize *bronze-1* allele caused by transposable elements of the *Ac-Ds* family alter the quantity and quality of the gene product. Genetics 120:767–777.

Schlappi, M., R. Raina, and N. Fedoroff. 1994. Epigenetic regulation of the maize *Spm* transposable element: novel activation of a methylated promoter by TnpA. Cell 77:427–437.

Seegmiller, A., and G. Herrick. 1998. A short internal eliminated sequence with central conserved sequences interrupting the *LA-MSC* gene of the 81 locus in the hypotrichous ciliates *Oxytricha fallax* and *O. trifallax*. J. Eukaryot. Microbiol. 45:55–58.

Seegmiller, A., K. R. Williams, R. L. Hammersmith, T. G. Doak, D. Witherspoon, T. Messick, L. L. Storjohann, and G. Herrick. 1996. Internal eliminated sequences interrupting the Oxytricha 81 locus: allelic divergence, conservation, conversions, and possible transposon origins. Mol. Biol. Evol. 13:1351–1362.

Segal, Y., B. Peissel, A. Renieri, M. de Marchi, A. Ballabio, Y. Pei, and J. Zhou. 1999. *LINE-1* elements at the sites of molecular rearrangements in Alport syndrome-diffuse leiomyomatosis. Am. J. Hum. Genet. 64:62–69.

Shedlock, A. M., and N. Okada. 2000. SINE insertions: powerful tools for molecular systematics. Bioessays 22:148–160.

Sheen, F. M., and R. W. Levis. 1994. Transposition of the LINE-like retrotransposon TART to *Drosophila* chromosome termini. Proc. Natl. Acad. Sci. USA 91:12510–12514.

Siebel, C. W., L. D. Fresco, and D. C. Rio. 1992. The mechanism of somatic inhibition of *Drosophila P*-element pre-mRNA splicing: multiprotein complexes at an exon pseudo-5′ splice site control U1 snRNP binding. Genes Dev. 6:1386–1401.

Siebel, C. W., A. Admon, and D. C. Rio. 1995. Soma-specific expression and cloning of PSI, a negative regulator of *P* element pre-mRNA splicing. Genes Dev. 9:269–283.

Simmen, M. W., S. Leitgeb, J. Charlton, S. J. Jones, B. R. Harris, V. H. Clark, and A. Bird. 1999. Nonmethylated transposable elements and methylated genes in a chordate genome. Science 283:1164–1167.

Simmons, G. M. 1992. Horizontal transfer of hobo transposable elements within the *Drosophila melanogaster* species complex: evidence from DNA sequencing. Mol. Biol. Evol. 9:1050–1060.

Singer, M. F., V. Krek, J. P. McMillan, G. D. Swergold, and R. E. Thayer. 1993. LINE-1: a human transposable element. Gene 135: 183–188.

Singer-Sam, J., and A. D. Riggs. 1993. X chromosome inactivation and DNA methylation. Exs 64:358–384.

Smit, A. F., and A. D. Riggs. 1996. *Tiggers* and DNA transposon fossils in the human genome. Proc. Natl. Acad. Sci. USA 93: 1443–1448.

Spanopoulou, E., F. Zaitseva, F. H. Wang, S. Santagata, D. Baltimore, and G. Panayotou. 1996. The homeodomain region of Rag-1 reveals the parallel mechanisms of bacterial and V(D)J recombination. Cell 87:263–276.

Stavenhagen, J. B., and D. M. Robins. 1988. Evolution in the genus *Drosophila*. Cell 55:247–254.

Steinemann, M., and S. Steinemann. 1997. The enigma of *Y* chromosome degeneration: *TRAM*, a novel retrotransposon is preferentially located on the *Neo-Y* chromosome of *Drosophila miranda*. Genetics 145:261–266.

Stoltzfus, A. 1999. On the possibility of constructive neutral evolution. J. Mol. Evol. 49:169–181.

Sun, X., J. Wahlstrom, and G. Karpen. 1997. Molecular structure of a functional *Drosophila* centromere. Cell 91:1007–1019.

Takahashi, H., S. Okazaki, and H. Fujiwara. 1997. A new family of site-specific retrotransposons, SART1, is inserted into telomeric repeats of the silkworm, Bombyx mori. Nucleic Acids Res. 25:1578–1584.

Takahashi, S., Y. Inagaki, H. Satoh, A. Hoshino, and S. Iida. 1999. Capture of a genomic HMG domain sequence by the En/Spm-related transposable element *Tpn1* in the Japanese morning glory. Mol. Gen. Genet. 261:447–451.

Talbert, L. E., and V. L. Chandler. 1988. Characterization of a highly conserved sequence related to mutator transposable elements in maize. Mol. Biol. Evol. 5:519–529.

Teng, S. C., B. Kim, and A. Gabriel. 1996. Retrotransposon reverse-transcriptase-mediated repair of chromosomal breaks. Nature 383:641–644.

Terrinoni, A., C. D. Franco, P. Dimitri, and N. Junakovic. 1997. Intragenomic distribution and stability of transposable elements in euchromatin and heterochromatin of *Drosophila melanogaster*: non-LTR retrotransposon. J. Mol. Evol. 45:145–153.

Tomilin, N. V. 1999. Control of genes by mammalian retroposons. Int. Rev. Cytol. 186:1–48.

Tonegawa, S. 1983. Somatic generation of antibody diversity. Nature 302:575–581.

Torreblanca, J., and J. Casadesus. 1996. DNA adenine methylase mutants of *Salmonella typhimurium* and a novel dam-regulated locus. Genetics 144:15–26.

Toth, M., J. Grimsby, G. Buzsaki, and G. P. Donovan. 1995. Epileptic seizures caused by inactivation of a novel gene, *jerky*, related to centromere binding protein-B in transgenic mice. Nat. Genet. 11:71–75.

Trelogan, S. A., and S. L. Martin. 1995. Tightly regulated, developmentally specific expression of the first open reading frame from LINE-1 during mouse embryogenesis. Proc. Natl. Acad. Sci. USA 92:1520–1524.

Tu, Z. 1997. Three novel families of miniature inverted-repeat transposable elements are associated with genes of the yellow fever mosquito, *Aedes aegypti*. Proc. Natl. Acad. Sci. USA 94: 7475–7480.

———. 1999. Genomic and evolutionary analysis of *Feilai*, a diverse family of highly reiterated SINEs in the yellow fever mosquito, *Aedes aegypti*. Mol. Biol. Evol. 16:760–772.

Tu, Z., J. Isoe, and J. A. Guzova. 1998. Structural, genomic, and phylogenetic analysis of *Lian*, a novel family of non-LTR retrotransposons in the yellow fever mosquito, *Aedes aegypti*. Mol. Biol. Evol. 15:837–853.

Unsal, K., and G. T. Morgan. 1995. A novel group of short interspersed repetitive elements (SINES) in *Xenopus*: Evidence of a specific target site for DNA mediated transposition of inverted-repeat SINES. J. Mol. Biol. 248:812–823.

Urieli-Shoval, S., Y. Gruenbaum, J. Sedat, and A. Razin. 1982. The absence of detectable methylated bases in *Drosophila melanogaster* DNA. FEBS Lett. 146:148–152.

van der Lugt, N. M., J. Domen, K. Linders, M. van Roon, E. Robanus-Maandag, H. te Riele, M. van der Valk, J. Deschamps, M. Sofroniew, and M. van Lohuizen. 1994. Posterior transformation, neurological abnormalities, and severe hematopoietic defects in mice with a targeted deletion of the *bmi-1* proto-oncogene. Genes Dev. 8:757–769.

Venables, P. J., S. M. Brookes, D. Griffiths, R. A. Weiss, and M. T. Boyd. 1995. Abundance of an endogenous retroviral envelope protein in placental trophoblasts suggests a biological function. Virology 211:589–592.

Villarreal, L. P. 1997. On viruses, sex, and motherhood. J. Virol. 71:859–865.

Warren, A. M., M. A. Hughes, and J. M. Crampton. 1997. Zebedee: a novel *copia-Ty1* family of transposable elements in the genome of the medically important mosquito *Aedes aegypti*. Mol. Gen. Genet. 254:505–513.

Weiner, A. M., P. L. Deininger, and A. Efstratiadis. 1986. Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. Annu. Rev. Biochem. 55:631–661.

Wessler, S. R. 1988. Phenotypic diversity mediated by the maize transposable elements *Ac* and *Spm*. Science 242:399–405.

———. 1996. Turned on by stress: plant retrotransposons. Curr. Biol. 6:959–961.

Wessler, S. R., T. E. Bureau, and S. E. White. 1995. LTR-retrotransposons and MITEs: important players in the evolution of plant genomes. Curr. Opin. Genet. Dev. 5:814–821.

White, S. E., L. F. Habera, and S. R. Wessler. 1994. Retrotransposons in the flanking regions of normal plant genes: a role for copia-like elements in the evolution of gene structure and expression. Proc. Natl. Acad. Sci. USA 91:11792–11796.

Wight, P. A., and A. Dobretsova. 1997. The first intron of the myelin proteolipid protein gene confers cell type-specific expression by

a transcriptional repression mechanism in non-expressing cell types. Gene 201:111–117.
Wigler, M. H. 1981. The inheritance of methylation patterns in vertebrates. Cell 24:285–286.
Xue, L., and M. Noll. 1996. The functional conservation of proteins in evolutionary alleles and the dominant role of enhancers in evolution. Embo J. 15:3722–3731.
Yang, Z., D. Boffelli, N. Boonmark, K. Schwartz, and R. Lawn. 1998. Apolipoprotein(a) gene enhancer resides within a LINE element. J. Biol. Chem. 273:891–897.
Yoder, J. A., C. P. Walsh, and T. H. Bestor. 1997. Cytosine methylation and the ecology of intragenomic parasites. Trends Genet. 13:335–340.
Yu, H., E. Hiatt, A. Chan, M. Sweeney, and R. Dawe. 1997. Neocentromere-mediated chromosome movement in maize. J. Cell Biol. 139:831–840.
Zelentsova, E. S., R. P. Vashakidze, A. S. Kraev, and M. B. Evgen'ev. 1986. Dispersed repeats in *Drosophila virilis*: elements mobilized by interspecific hybridization. Chromosoma 93: 469–476.
Zhang, J. H., G. Dawes, and W. P. Stemmer. 1997. Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening. Proc. Natl. Acad. Sci. USA 94:4504–4509.
Zhang, Q., J. Arbuckle, and S. R. Wessler. 2000. Recent, extensive, and preferential insertion of members of the miniature inverted-repeat transposable element family *Heartbreaker* into genic regions of maize. Proc. Natl. Acad. Sci. USA 97:1160–1165.

Corresponding Editor: T. Markow