Vertical Transmission of the Retrotransposable Elements R1 and R2 During the Evolution of the Drosophila melanogaster Species Subgroup

Danna G. Eickbush and Thomas H. Eickbush

Department of Biology, University of Rochester, Rochester, New York 14627

Manuscript received April 15, 1994
Accepted for publication October 15, 1994

ABSTRACT

R1 and R2 are non-long-terminal repeat retrotransposable elements that insert into specific sequences of insect 28S ribosomal RNA genes. These elements have been extensively described in Drosophila melanogaster. To determine whether these elements have been horizontally or vertically transmitted, we characterized R1 and R2 elements from the seven other members of the melanogaster species subgroup by genomic blotting and nucleotide sequencing. Each species was found to have homogenous families of R1 and R2 elements with the exception of erecta and arena, which have no R2 elements. The DNA sequences of multiple R1 and R2 copies from each species indicated nucleotide divergence within each species averaged only 0.48% for R1 and 0.35% for R2, well below the level of divergence among the species. Most copies of R1 and R2 (40 of 47) sequenced from the seven species were potentially functional, as indicated by the absence of premature termination codons or translational frameshifts that would destroy the open reading frame of the element. The sequence relationships of both the R1 and R2 elements from the various members of the melanogaster subgroup closely followed that of the species phylogeny, suggesting that R1 and R2 have been stably maintained by vertical transmission since the origin of this species subgroup 17–20 million years ago. The remarkable stability of R1 and R2, compared to what has been suggested for transposable elements that insert at multiple locations in these same species, may be due to their unique specificity for sites in the rRNA gene locus. Under low copy number conditions, when it is essential for any mobile element to transpose, the insertion specificities of R1 and R2 ensure uniform developmentally regulated target sites that can be occupied with little or no detrimental effect on the host.

Transposable elements are prevalent in all eukaryotes, frequently comprising 10–20% of the total genome (see reviews in Berg and Howe 1989). Many features of present day genomes and much of the diversity increasingly apparent in genome organization may have evolved under selective pressure to accommodate or reject a variety of mobile DNAs. Understanding the impact these elements have had on the evolution of genomes will involve knowing how the elements themselves have evolved. One fundamental question that must be addressed is whether mobile elements can remain stable active components of a genome over long evolutionary periods.

Many reports have appeared that raise the possibility of the transfer of mobile elements between species (reviewed in Kidwell 1993). This movement of elements between species is usually referred to as "horizontal" or "lateral" transfer. Many studies of horizontal transfer have been conducted within the genus Drosophila because of the large number of well-characterized species and transposable elements and the low copy number of these elements. Examples of potential horizontal transfer include P, which appears to have been introduced into Drosophila melanogaster from Drosophila willistoni (Daniels et al. 1990); jockey, which may have been introduced into Drosophila funebris from D. melanogaster (Mizukaki and Mato 1990); I and hobo, which appear to have been lost from D. melanogaster and then reacquired from Drosophila simulans (Bucheton et al. 1986; Pascaud and Periquet 1991; Simmons 1992); and mariner, which may have entered the melanogaster species group from the genus Zaprionus (Maruyama and Hartl 1991). The horizontal transfer of mobile elements has also been reported in plants based on the presence of highly divergent elements in the same species (Konieczny et al. 1991; Flavell et al. 1992). Finally, the most convincing evidence for the horizontal transfer of mobile elements has been studies of the mariner element that appears to have transferred between major groups of animals (Robertson 1993; Robertson and Macleod 1993).

In contrast to these numerous reports of horizontal transfer, only a few reports have appeared suggesting that specific classes of mobile elements have remained stable components of a species phylogeny (Hardies et al. 1986; Springer et al. 1991; Pascale et al. 1993; Vanderwiel et al. 1993). It is interesting to note that
A

rDNA repeat

18S 5.8S 28S

R1 3' UT

D. melanogaster rDNA unit (LONG and DAWID 1979; Kidd and Glover 1981; Jamrich and Miller 1984). Type I and type II elements were subsequently identified in the silkmoth, Bombyx mori (Eickbush and Robins 1985). The fraction of the 28S rRNA genes that are inserted with these elements can vary significantly in different strains of D. melanogaster and B. mori (Xiong et al. 1988; Lyckegaard and Clark 1991; Jakubczak et al. 1992). Nucleotide sequence analysis of these elements (renamed R1 and R2) from both D. melanogaster and B. mori revealed that they contain open reading frames (ORFs) typical of retrotransposable elements of the non-long-terminal repeat (non-LTR) type (Burke et al. 1987; Xiong and Eickbush 1988; Jakubczak et al. 1990). R1 and R2 are no more related to each other than to any other non-LTR retrotransposable element; thus they appear to represent distinct elements that have independently become specialized for insertion into the 28S gene (Eickbush 1992).

A few copies of R1 and R2 have also been detected outside the rRNA gene loci in B. mori and a large number of R1 elements have been detected in the centromeric heterochromatin of D. melanogaster (Kidd and Glover 1980; Peacock et al. 1981; Roha et al. 1981; Xiong et al. 1988). Most of these non-rDNA copies are located in fragments of the 28S gene or inserted into sequences with identity to the 28S gene insertion sites, suggesting that each element encodes a sequence specific endonuclease that recognizes the insertion site. Direct analysis of the enzymatic activities encoded by the ORF of R2 from B. mori revealed that this element does indeed encode an endonuclease capable of cleaving the 28S target site (Xiong and Eickbush 1988b). A 3' hydroxyl group exposed by this DNA cleavage serves as primer for reverse transcription of the RNA template (Luan et al. 1993).

The specificity of R1 and R2 elements to sites in the 28S gene has made it possible to detect these elements in most insect species (Jakubczak et al. 1991). Studies
to date indicate that some insect species contain single families of R1 and R2, whereas other species contain multiple distinct families of either R1 or R2 (BURKE et al. 1993). We report here the first systematic attempt to trace the evolution of R1 and R2 elements in a species phylogeny. The eight members of the melanogaster species subgroup of Drosophila were chosen for this study because they represent one of the best-characterized insect species phylogenies (reviewed in LACHAISE et al. 1988). Our results suggest that R1 and R2 have been vertically transmitted during the evolution of this species subgroup.

**MATERIALS AND METHODS**

**Strains and DNA isolation:** Stocks of Drosophila sechellia and Drosophila mauritiana were obtained from C.-I. WU, Drosophila arena from J. TRUE and C. C. LAURIE, Drosophila yakuba from M. ASHburnER and Drosophila teissierii from F. LEMEUIN. Additional strains were obtained from the Drosophila stock centers at Bowling Green University and Indiana University. DNA was isolated from ~100 adult individuals of each species by the method of CUTH et al. (1985), except that after RNase A treatment, the aqueous layer was extracted with an equal volume of phenol-chloroform, an equal volume of chloroform alone and ethanol precipitated. Genomic blots based on restriction site polymorphisms between species in the rDNA genes or the R1 and R2 elements were conducted from at least two stocks of each species to confirm the species identity and purity of each stock. The genomic blot analysis and all nucleotide sequences presented in this report were of DNA obtained from one stock of each species: D. simulans (0251.3, Bowling Green), D. mauritiana (0241.0, Bowling Green), D. sechellia (3578, Bowling Green), D. teissierii (1015, Indiana), D. yakuba (0261.0, Bowling Green), Drosophila erecta (1013, Indiana) and D. arena (C. C. LAURIE).

**Genomic DNA blots:** All genomic blots were performed as previously described (JAKUBCZAK et al. 1992). For each species, ~1 μg of total adult DNA was digested with restriction enzymes, electrohoresed on a 1% agarose gel, transferred to nitrocellulose, baked and hybridized with [32P]-labeled restriction fragments. All restriction fragments used as probes were excised from agarose gels, purified (GeneClean, Bio101 Inc.) and labeled by random priming. The 280-bp 28S gene probe representing sequences immediately 3′ of the R1 site (Figure 1B) was obtained from D. melanogaster clone a56 (JAKUBCZAK et al. 1991). The 853-bp BamHI fragment from near the 3′ end of R1 and the 345-bp EcoRI/SpI fragment from near the 3′ end of R2 were obtained from the D. melanogaster clones a56 and p305, respectively (JAKUBCZAK et al. 1990).

**PCR amplifications:** The 3′ ends of R1 and R2 elements were obtained by PCR amplification of genomic DNA using the primer positions shown in Figure 1B. R1 sequences were amplified using the degenerate primer, 5′-GGTGGCNCN-CARGGNTC-3′ (N, any nucleotide; Y, T or C; R, A or G), termed the R1 primer. This R1 primer was complementary to the sequence encoding the amino acids GCPQGS, found in motif four of the reverse transcriptase domain of all R1 elements (XIONG and ECKBUSH 1990; BURKE et al. 1995). R1 sequences were also amplified using a degenerate primer, 5′-GCNTWGWCGNAYGA-3′ (W, A or T), termed the R1/R2 primer. The R1/R2 primer was complementary to the sequence encoding the amino acids A(Y/F)ADD found in motif five of the reverse transcriptase domain of both R1 and R2 elements. Both primers were used in combination with the 28S gene primer, 5′-AACAGGGCCACATCGAAAGGATC-3′, complementary to 28S gene sequences 670 bp downstream of the R1 insertion site. Two degenerate primers were also used in combination with a 28S gene primer to amplify R2 elements. One primer 5′-CARGGNGAYCCNYTNTC-3′, termed the R2 primer, was complementary to the sequence encoding the amino acids QGDPLS, found in motif four of the reverse transcriptase domain of all R2 elements. The second primer was the same R1/R2 primer used for the R1 amplifications. Both the R2 primer and the R1/R2 primer were used in combination with the primer 5′-GCYGCGATGATGAT-YAACAG-3′, complementary to the 28S gene 20 bp downstream of the R2 insertion site. AmpliTaq DNA polymerase (Perkin-Elmer Cetus) was used under conditions specified by the supplier. Approximately 0.2 μg genomic DNA was amplified in 30 cycles of 94° for 1 min, 60° for 1 min and 72° for 3 min.

**Nucleotide sequence determinations:** The PCR amplification products were purified from primer sequences (GeneClean, Bio101 Inc.) and cloned into the sequencing vector mp18 digested at the SmaI site and dT-tailed with terminal transferase (HOLT and GRAHAM 1991). Multiple mp18 clones from each species were subjected to sequence analysis by the single-stranded dyeoxy chain termination method (SANGER et al. 1977). The 1.7–1.8 kb of R1 or R2 sequence on those clones containing the anti-sense strand of the elements were determined using three separate sequencing reactions. Starting at the primer sequence within the reverse transcriptase domain (referred to as the 5′ end of the PCR product), the first 650 bp was obtained using a Universal sequencing primer (U.S. Biochemicals). The middle third of the amplified product was obtained using a sequencing primer complementary to the R1 or R2 element ~600 bp from the 5′ end of the PCR product (5′-GACTGCTGACGTTCTC-3′ for R1, 5′-GTTGGCTCAGATTCGATC-3′ for R2). The sequence of the final third of the amplified product of each element was obtained using a sequencing primer complementary to the R1 or R2 element ~1,050 bp from the 5′ end of the PCR product (5′-GACTTGGAGTGGAC-GC-3′ for R1, 5′-GTCGGATCCTCGG-3′ for R2). Because a 329 bp duplication at the 3′ end of certain teissier I elements prevented the completion of the sequence in these three steps, a 2.3 kb HindIII fragment containing the R1 sequences was excised from these clones and reinserted in the opposite orientation in mp18. The duplication was sequenced using the primer 5′-TCGCGATTAGTGTCAGAT-3′ complementary to the antisense strand of the 28S gene 65 bp downstream of the R1 site. This 28S downstream primer and the Universal primer were used to sequence additional 3′ R1 and R2 junctions, respectively, in those original mp18 clones that contained the sense strand of the elements. Determination of the 28S gene sequence adjacent to the R1 elements was conducted using the primer 5′-GACTTGGACTCTAATCTGGC-3′, which was complementary to the sense strand of the 28S gene 65 bp downstream of the R1 site. R1 and R2 sequences have been submitted to GenBank under the following accession numbers: D. simulans, R1sim1-U13028; R2sim1-U13032; D. mauritiana, R1mau3-U13030; D. sechellia, R1seu4-U13027; R2seu1-U13031; D. teissierii, R1te2-U13035; R2tei1-U13033; D. yakuba, R1yak2-U13029, R2yak1-U13034; D. erecta, R1er4-U130024; D. arena, R1aen5-U13026. All within species nucleotide variations have been annotated on the single R1 and R2 sequence submitted for each species. Copies of the sequence alignments are available from the authors upon request.

**Estimates of the PCR and sequence determination error**
rates: Two sources of error are associated with our sequence determinations: the misincorporation of nucleotides during the PCR amplifications and the error rate associated with the sequence determination of the cloned products. To estimate the misincorporation rates associated with our PCR reactions, a 1.36-kb fragment of the rDNA repeat of D. melanogaster was amplified under the PCR reactions described above using primers, 5'-GAAGCTGAGGAGGAGGATC-3' and 5'-GCCTCCGTCATAATGGCC-3', which are complimentary to the rDNA repeat positions 1968 and 3392 (TAUTZ et al. 1988). An individual m13 clone, ITSmel9 (0.901 µg), was used in a second amplification with these primers. The PCR product was inserted into the sequencing vector mp18 as described above and 20 clones containing the sense strand of the rDNA repeat were sequenced for 450 bp using the Universal primer. PCR errors were scored by running the sequencing gel with each of the four bases from the different clones group together, similar to that described by NACHMAN et al. (1994) for the determination of polymorphisms within a species. Of the 9.0 kb of DNA determined by this method, 20 substitutions were detected, resulting in an error rate of 0.22%. This rate is similar to the 0.25% determined by SAIKU et al. (1988) using similar PCR conditions.

To minimize the potential errors associated with the determination of sequence on only one strand of each clone, the sequence of all copies of R1 or R2 from the same species where first compared. All nucleotide positions that differed (typically 0.5% of the nucleotide positions) were rechecked on the original X-ray films to eliminate most of the human error involved with the determination of nucleotide sequences. To estimate the remaining error rate that could be removed by sequencing the second stand, the following experiment was conducted. The inserts from five R2 clones (two each from D. sechellia and D. teissieri and one from D. yakuba) and seven R1 clones (two each from D. simulans, D. teissieri, D. erecta and one from D. yakuba) were excised by digest with EcoRI and HindIII and reinserted in the opposite orientation in mp19. The primers 5'-CAGTCTGTTCTCC-TCC-3' and 5'-GGCGGACGGCATAAC-3', which are complementary to sequences ~675 bp downstream of the 5' end of the PCR-amplified segments of R1 and R2, respectively, were used to sequence a 645-bp region of the elements on the second strand. In the 7.7 kb of DNA that was sequenced on the second strand, only five errors were obtained. This error rate (0.06%) is well below the error rate of 0.22% resulting from the PCR amplification reactions, indicating little improvement in our estimation of the sequence variation within or between the species would be obtained by sequencing both strands of each DNA clone.

Sequence analysis: DNA and protein sequence analyses were analyzed by the computer program package MacVector (International Biotechnologies, Inc.) on a Macintosh computer and with the Wisconsin Genetics Computer Group programs on a VAX computer. The phylogenetic relationship of the R1 and R2 sequences was determined with maximum parsimony algorithms using the PAUP package of programs (SWOFFORD 1991).

RESULTS

Selection of the 3' end of R1 and R2 for analysis: All copies of the R1 and R2 elements of D. melanogaster terminate at the same positions at their 3' ends but many are truncated at their 5' ends (DAMID and REBBERT 1981; ROHHA et al. 1981). Truncations of the 5' end at a variable distance from a conserved 3' end are common in most non-LTR retrotransposable elements (EICKBUSH 1992) and are probably generated during the retrotransposition event itself (LUAN et al. 1993). Because of this 5' length variation, our assays for the presence of R1 and R2 in a species, as well as our sequence analyses to determine their phylogenetic relationship, have focused on the 3' half of each element (JAKUBCZAK et al. 1991; BURKE et al. 1993).

Genomic blot assays for R1 and R2 in the eight species of the melanogaster subgroup: Genomic blot assays have previously suggested that R1 and R2 elements, which can cross-hybridize with those from D. melanogaster, are present in most species of the melanogaster subgroup (GOEN et al. 1982; ROHHA et al. 1983). We have devised genomic blotting procedures that are highly sensitive in detecting the presence of these elements and in estimating the fraction of the 28S genes inserted (JAKUBCZAK et al. 1992). An example of the genomic blot assay is shown in Figure 2A. In D. melanogaster, HindIII cleaves the 28S gene 280 bp downstream of the R1 insertion site, whereas PsiI cleaves upstream of this site within the 5.8S gene (Figure 1A). When a genomic blot of PsiI + HindIII digested DNA is hybridized with the 28S gene sequences immediately downstream of the R1 insertion site (Figure 1B), a 3.5-kb hybridizing band representing the uninserted 28S genes within D. melanogaster can be observed (Figure 2A). This 3.5-kb band is present in all members of the melanogaster species subgroup, indicating that these HindIII and PsiI restriction sites within the rDNA unit are conserved. In D. melanogaster, PsiI cleaves the R1 and R2 elements 1.1 and 0.6 kb from their 3' ends (Figure 1B); thus 28S genes containing R1 and R2 insertions are represented in Figure 2A as 1.4- and 0.9-kb hybridizing bands, respectively. Because the 28S gene probe is located downstream of the R1 insertion site, 28S genes containing both R1 and R2 insertions are scored on the blot as only containing an R1 insertion. Thus the intensity of the 3.5- and 1.4-kb bands represent an accurate ratio of the uninserted and R1 inserted genes, whereas the intensity of the 0.9-kb band underestimates the level of R2 insertions.

Hybridizing bands ranging from 0.9 to 1.4 kb in length are also apparent in each of the seven other species (Figure 2A). To determine which of these bands corresponded to R1 and R2 insertions, blots of PsiI + HindIII digested genomic DNA from each species were probed with a fragment from either the 3' end of the D. melanogaster R1 element (Figure 2B) or the 3' end of the D. melanogaster R2 element (Figure 2C). R1 insertions in the 28S genes are visible as hybridizing bands present in Figure 2, A and B, and R2 insertions within the 28S genes are visible as bands present in Figure 2, A and C.

Comparison of Figure 2, A and B, indicates that all species contained R1 elements with the possible excep-
FIGURE 2.—Genomic blots showing the abundance of R1 and R2 elements in the eight members of the melanogaster species subgroup. The species from which the DNA was isolated is shown at the top of each lane. (A) Total adult DNA was digested with PstI and HindIII and hybridized with a 32P-labeled BamHI-HindIII probe shown in Figure 1B. Sizes at left are in kilobases. The intense band at 3.5 kb corresponds to uninserted rDNA units, the bands at 1.0–1.4 kb correspond to rDNA units containing an R1 element and the bands at 0.9 kb correspond to rDNA units containing an R2 insertion. (B) Similar blot to that in A but probed with the 0.8-kb BamHI fragment from the D. melanogaster R1 element (Figure 1B). (C) Similar blot to that in A but probed with the 0.3-kb EcoRI-SphI fragment from the D. melanogaster R2 element (Figure 1B).

Comparison of the hybridization of the 28S gene probe (Figure 2A) with that of the R2 probe (Figure 2C) indicates that the predominant 0.9-kb band, present in all species except D. erecta and D. orena, represents R2 insertions within the 28S genes. In addition to the 0.9-kb band, a series of fainter hybridizing bands is also observed upon longer exposure of the blot probed with R2 sequences. Most of these faint bands do not correspond to R2 insertions outside of the 28S genes. Rather, these bands are the result of R2 insertions in 28S genes that also contain an R1 insertion. Because one end of the hybridizing restriction fragments in all three panels in Figure 2 is dependent on cleavage at the HindIII site downstream of the R1 insertion site, those R2 elements inserted in a 28S gene also containing an R1 element give rise to hybridizing restriction fragments that are not 0.9 kb in length. The 5’ truncations associated with the R1 elements result in hybridizing fragments that vary in size within each species (JAKUBCZAK et al. 1992).

PCR amplification of R1 and R2 from each species: To determine the sequence relationship of the R1 and R2 elements within and between members of the melanogaster species subgroup, multiple copies of the 3’ end of R1 and R2 from each species were PCR amplified and sequenced. For each amplification, one PCR primer was complementary to a highly conserved region encoding part of the reverse transcriptase domain of the element and the second primer was complementary to the 28S gene downstream of the insertion site (Figure 1B and MATERIALS AND METHODS). The primers to the reverse transcriptase domain were designed to be sufficiently degenerate to amplify R1 and R2 elements from species of Diptera, Lepidoptera, Hymenoptera and Coleoptera (JAKUBCZAK et al. 1992; BURKE et al. 1993). The use of such highly degenerate primers
was complementary to a elements in most species are located outside the rDNA d contained three termination codons or changes in frame, and a large deletion that removes the last six codons of the OW, repeats (Figure 2). This ensured that all families of R1 or R2 elements would be amplified independent of their sequence relationship to the R1 and R2 elements of D. melanogaster.

The 28S gene primer used for amplification of the R1 elements was located 670 bp downstream of the R1 insertion site because a significant fraction of the R1 elements in most species are located outside the rDNA repeats (Figure 2B). Most of these non-rDNA copies do not hybridize with the 280 bp 28S gene probe downstream of the R1 site (Figure 2A) and therefore should not be PCR amplified by a 28S gene primer located even further downstream of the insertion site. The 28S gene primer used in the amplification of R2 elements was complementary to a 28S gene sequence between the R2 and R1 insertion sites. This ensured that all R2 elements would be amplified independent of whether they were in a 28S gene also containing an R1 insertion.

In the case of the R1 amplifications, the expected 2.3–2.4 kb products (see Figure 1B) were detected with genomic DNA from all eight species of the melanogaster subgroup. Only the PCR product from D. teissieri appeared unusual in that it migrated on agarose gels as a diffuse doublet 2.3–2.7 kb in length. This doublet suggested that R1 copies in D. teissieri were polymorphic in size. In the case of the R2 amplifications, the expected 1.7–1.8 kb products were observed in all species of the melanogaster subgroup except D. orena and D. erecta, where no amplified product was detected. Two additional R2-specific PCR primers were used to attempt the amplification of R2 elements from D. orena and D. erecta. One of these primers was complementary to a region 550 bp closer to the 3′ end of the R2 element and was used in combination with 28S gene sequences downstream of the insertion site. The second R2 primer was complementary to the sequences encoding the highly conserved cysteine-histidine motif located at the 5′ end of R2 elements (JAKUBCZAK et al. 1990) and oriented to be used in combination with a 28S gene primer 5′ of the insertion site. Once again these two additional R2-specific PCR primers gave the expected amplified products from the other species of the melanogaster subgroup and no product in D. orena and D. erecta (data not shown). Thus all four pairs of PCR primers confirmed the genomic blots shown in Figure 2, indicating that D. orena and D. erecta have no R2 elements inserted within their 28S genes.

The PCR-amplified R1 and R2 products from each species were cloned in mp18, and multiple copies of the antisense strand of each element were sequenced (see MATERIALS AND METHODS and Table 1). Additional copies of the 3′ junctions of R1 and R2 with the 28S gene were obtained using those mp18 clones containing the sense strand of the element.

### 3′ junctions with the 28S gene: Non-LTR retrotransposable elements contain poly(A) tails or short A-rich nucleotide repeats at their 3′ ends (EICKBUSH 1992). We have recently shown that the reverse transcriptase encoded by the R2 element of B. mori has the ability to add nontemplated residues, predominantly T's, on the
cDNA strand during the initiation of reverse transcription (D. D. Luan and T. H. Eckbush, unpublished data). These nontemplated additions result in A-rich repeats at the 3' end of the element. Consistent with a model in which most of the sequence variation detected within the R1 and R2 elements of each species is generated by retrotransposition, the highest level of this sequence variation between elements from the same species or from different species was detected at their 3' junction with the 28S gene (Figure 3).* 

The 3' junctions of all sequenced R2 elements ended in a poly(A) tail from 10 to 25 nucleotides in length (Figure 3B). In four elements this poly(A) tail was interrupted by a G or T nucleotide. R1 elements from the eight Drosophila species also terminated in a repeated sequence, but these repeats were more complex than those of the R2 elements. As shown in Figure 3A, the 3' end of the R1 elements appears to be composed of two sequence features. The conserved R1 sequences (stippled region) end in a short (1-6 nucleotide) A tract. This A tract is then followed by short tandem repeats (arrows) that vary in sequence even between closely related species: two to five AC (or AT) repeats in D. melanogaster, D. teissieri, D. yakuba, D. erecta and D. orena; two to four GTC repeats in D. mauritiana and one AC and one GTC repeat in D. simulans and D. sechellia. This high level of sequence variation found at the 3' junctions of R1 and R2 elements, which is similar to that found in other non-LTR retrotransposable elements, suggests that these elements have undergone retrotransposition events and are not simply being maintained in the rDNA locus by recombinational mechanisms alone.

Sequence comparisons: The PCR-amplified regions of the R1 and R2 elements available for sequence analysis included the 3' untranslated region of each element (~550 bp for R1 and ~250 bp for R2) and the region encoding the carboxyl terminal portion of the ORF (387 codons for R1 and 476 for R2). This region of the ORF contains part of the reverse transcriptase domain and a carboxyl domain containing a highly conserved cysteine-histidine motif (Burke et al. 1993). Over 85% of the R1 and R2 copies sequenced from the seven species (40 of 47) were potentially functional insofar as they contained neither premature termination codons nor insertions or deletions leading to a change in translational reading frame in the region amplified by PCR (Table 1). In our previous report on D. melanogaster, only 1 of 18 copies of R1 and R2 contained a disruption of the ORF based on a 540-bp sequenced region (Jakubczak et al. 1992). It should be emphasized that there is no data to confirm that such a high percentage of the R1 and R2 elements are actually capable of retrotransposition. Therefore we will use the term "intact" elements in this section to describe those elements that based on our available sequences encode undisrupted ORFs and "defective" elements to describe those copies with disrupted ORFs.

The nucleotide sequence divergences of the R1 and R2 elements within each member of the melanogaster species subgroup are presented in Table 1. In Figure 3.-Variation in the nucleotide sequences at the 3' junction of R1 and R2 elements. The precise junction between the R1 or R2 element with the 28S gene is indicated by a solid vertical line. A greater number of 3' junctions are presented in this figure than are indicated as sequenced elements. The precise junction between the R1 and R2 elements appears to be composed of two sequence features. The conserved R1 sequences (stippled region) end in a short (1-6 nucleotide) A tract. This A tract is then followed by short tandem repeats (arrows) that vary in sequence even between closely related species: two to five AC (or AT) repeats in D. melanogaster, D. teissieri, D. yakuba, D. erecta and D. orena; two to four GTC repeats in D. mauritiana and one AC and one GTC repeat in D. simulans and D. sechellia. This high level of sequence variation found at the 3' junctions of R1 and R2 elements, which is similar to that found in other non-LTR retrotransposable elements, suggests that these elements have undergone retrotransposition events and are not simply being maintained in the rDNA locus by recombinational mechanisms alone.

Sequence comparisons: The PCR-amplified regions of the R1 and R2 elements available for sequence analysis included the 3' untranslated region of each element (~550 bp for R1 and ~250 bp for R2) and the region encoding the carboxyl terminal portion of the ORF (387 codons for R1 and 476 for R2). This region of the ORF contains part of the reverse transcriptase domain and a carboxyl domain containing a highly conserved cysteine-histidine motif (Burke et al. 1993). Over 85% of the R1 and R2 copies sequenced from the seven species (40 of 47) were potentially functional insofar as they contained neither premature termination codons nor insertions or deletions leading to a change in translational reading frame in the region amplified by PCR (Table 1). In our previous report on D. melanogaster, only 1 of 18 copies of R1 and R2 contained a disruption of the ORF based on a 540-bp sequenced region (Jakubczak et al. 1992). It should be emphasized that there is no data to confirm that such a high percentage of the R1 and R2 elements are actually capable of retrotransposition. Therefore we will use the term "intact" elements in this section to describe those elements that based on our available sequences encode undisrupted ORFs and "defective" elements to describe those copies with disrupted ORFs.

The nucleotide sequence divergences of the R1 and R2 elements within each member of the melanogaster species subgroup are presented in Table 1. In Figure 3.-Variation in the nucleotide sequences at the 3' junction of R1 and R2 elements. The precise junction between the R1 or R2 element with the 28S gene is indicated by a solid vertical line. A greater number of 3' junctions are presented in this figure than are indicated as sequenced elements. The precise junction between the R1 and R2 elements appears to be composed of two sequence features. The conserved R1 sequences (stippled region) end in a short (1-6 nucleotide) A tract. This A tract is then followed by short tandem repeats (arrows) that vary in sequence even between closely related species: two to five AC (or AT) repeats in D. melanogaster, D. teissieri, D. yakuba, D. erecta and D. orena; two to four GTC repeats in D. mauritiana and one AC and one GTC repeat in D. simulans and D. sechellia. This high level of sequence variation found at the 3' junctions of R1 and R2 elements, which is similar to that found in other non-LTR retrotransposable elements, suggests that these elements have undergone retrotransposition events and are not simply being maintained in the rDNA locus by recombinational mechanisms alone. 

Sequence comparisons: The PCR-amplified regions of the R1 and R2 elements available for sequence analysis included the 3' untranslated region of each element (~550 bp for R1 and ~250 bp for R2) and the region encoding the carboxyl terminal portion of the ORF (387 codons for R1 and 476 for R2). This region of the ORF contains part of the reverse transcriptase domain and a carboxyl domain containing a highly conserved cysteine-histidine motif (Burke et al. 1993). Over 85% of the R1 and R2 copies sequenced from the seven species (40 of 47) were potentially functional insofar as they contained neither premature termination codons nor insertions or deletions leading to a change in translational reading frame in the region amplified by PCR (Table 1). In our previous report on D. melanogaster, only 1 of 18 copies of R1 and R2 contained a disruption of the ORF based on a 540-bp sequenced region (Jakubczak et al. 1992). It should be emphasized that there is no data to confirm that such a high percentage of the R1 and R2 elements are actually capable of retrotransposition. Therefore we will use the term "intact" elements in this section to describe those elements that based on our available sequences encode undisrupted ORFs and "defective" elements to describe those copies with disrupted ORFs.

The nucleotide sequence divergences of the R1 and R2 elements within each member of the melanogaster species subgroup are presented in Table 1. In Figure 3.-Variation in the nucleotide sequences at the 3' junction of R1 and R2 elements. The precise junction between the R1 or R2 element with the 28S gene is indicated by a solid vertical line. A greater number of 3' junctions are presented in this figure than are indicated as sequenced elements. The precise junction between the R1 and R2 elements appears to be composed of two sequence features. The conserved R1 sequences (stippled region) end in a short (1-6 nucleotide) A tract. This A tract is then followed by short tandem repeats (arrows) that vary in sequence even between closely related species: two to five AC (or AT) repeats in D. melanogaster, D. teissieri, D. yakuba, D. erecta and D. orena; two to four GTC repeats in D. mauritiana and one AC and one GTC repeat in D. simulans and D. sechellia. This high level of sequence variation found at the 3' junctions of R1 and R2 elements, which is similar to that found in other non-LTR retrotransposable elements, suggests that these elements have undergone retrotransposition events and are not simply being maintained in the rDNA locus by recombinational mechanisms alone. 

Sequence comparisons: The PCR-amplified regions of the R1 and R2 elements available for sequence analysis included the 3' untranslated region of each element (~550 bp for R1 and ~250 bp for R2) and the region encoding the carboxyl terminal portion of the ORF (387 codons for R1 and 476 for R2). This region of the ORF contains part of the reverse transcriptase domain and a carboxyl domain containing a highly conserved cysteine-histidine motif (Burke et al. 1993). Over 85% of the R1 and R2 copies sequenced from the seven species (40 of 47) were potentially functional insofar as they contained neither premature termination codons nor insertions or deletions leading to a change in translational reading frame in the region amplified by PCR (Table 1). In our previous report on D. melanogaster, only 1 of 18 copies of R1 and R2 contained a disruption of the ORF based on a 540-bp sequenced region (Jakubczak et al. 1992). It should be emphasized that there is no data to confirm that such a high percentage of the R1 and R2 elements are actually capable of retrotransposition. Therefore we will use the term "intact" elements in this section to describe those elements that based on our available sequences encode undisrupted ORFs and "defective" elements to describe those copies with disrupted ORFs.

The nucleotide sequence divergences of the R1 and R2 elements within each member of the melanogaster species subgroup are presented in Table 1.
4, the within species and between species sequence divergences are summarized as a possible phylogenetic relationship of the R1 and R2 elements based on maximum parsimony algorithms.

Only low levels of within-species nucleotide divergence were found for those R1 and R2 elements that appeared intact (Table 1). The apparent level of within-species sequence divergence averaged 0.48% for R1 and 0.35% for R2. Because the error rate associated with the PCR amplification of each element copy was 0.22% (see MATERIALS AND METHODS), these values are mostly due to the misincorporation of nucleotides during the PCR amplifications. Thus the actual level of sequence variation within each species is extremely low and cannot be quantified by the PCR approach used in this report. Previous studies of widely distributed geographical strains of D. melanogaster revealed that divergence of the R1 and R2 elements was not significantly greater than that found within each strain (JAKUBCZAK et al. 1992). Assuming a similar situation in the other species, the data suggest that each member of the melanogaster species subgroup contained a single highly homogeneous family of R1 elements and all but D. orena and D. erecta contained a homogeneous family of R2 elements.

Table 1 also summarizes the sequence divergence of the cloned R1 and R2 copies that appeared defective. These defective copies are indicated by “d” following the clone number in Figure 4. The sequence divergence of these defective elements with the intact elements present in the same species varied from 2.0 to 3.0% for the R1 elements and 0.4 to 1.9% for the R2 elements. In an attempt to determine whether the four defective R1 elements were located in a 28S gene segment within or outside the rDNA loci, the nucleotide sequence of the 670 bp of the 28S gene amplified along with each R1 element was also determined (see Figure 1B). In the case of the defective D. orena R1 element (ore7-d), the 28S gene segment adjacent to this element differed by 1.4% from the 28S sequences adjacent to the five intact R1 copies from that species. Thus the 28S gene in which ore7-d resided was no longer undergoing the process of concerted evolution with the other 28S genes of the species and was therefore probably not located in an rDNA locus.

In the case of the D. teissieri R1 elements, a clear separation of intact R1 elements inside the rDNA loci and defective copies outside the rDNA loci was not found. The three defective elements contained flanking 28S gene sequences that averaged 1.1% nucleotide divergence from the DNA flanking the intact elements. Two of these clones also contained large deletions that eliminated the entire 3' untranslated region and 451 bp of the 28S gene. These three defective elements were probably outside the rDNA loci of D. teissieri. Surprisingly, two of the apparently intact R1 elements

![Figure 4](image_url)
(tet10 and tet40) contained a 29-bp deletion of the 28S gene sequence, suggesting that they were also located outside the rDNA loci. It is not clear why a greater number of defective R1 elements was recovered from D. teissieri or why the intact copies from this species had twice the level of sequence divergence detected in any other species (Table 1). The Southern blots in Figure 2B clearly indicated that the D. teissieri genome did not contain a large number of R1 elements outside the rDNA loci. We suggest that the low number of R1 elements present in D. teissieri (Figure 2A) resulted in a higher probability of selecting copies of R1 outside the rDNA locus, even if the absolute number of these non-rDNA copies was not high.

Determination of whether the defective R2 elements listed in Table 1 were located within or outside the rRNA gene locus could not be made because only a short segment of the 28S gene was amplified along with the rDNA units. The genomic blot in Figure 2C suggested that few R2 elements are outside the rDNA loci. Only one defective R2 element (maul8-d) had >1% divergence with the functional copies of the species and was thus likely to be outside the rDNA loci. The nucleotide sequence of maul8-d indicates that it was equally divergent from the intact elements of D. mauritiana and D. sechellia (Figure 4B) and thus may have existed outside the rDNA locus since before the speciation event separating D. mauritiana and D. sechellia.

Many D. teissieri R1 elements contain unusual 3′ untranslated regions: A total of nine R1 elements from D. teissieri were sequenced because the PCR-amplified products migrated as two diffuse bands of 2.4 and 2.7 kb rather than the distinct 2.4-kb band produced by all other species. Figure 5 summarizes the nature of the length variations found in seven of the D. teissieri R1 clones. (Two of the teissieri clones scored as defective in Table 1 contained a deletion eliminating their 3′ untranslated regions.) Only three of the sequenced D. teissieri R1 elements contained a 3′ untranslated region similar in length to the R1 elements from the other members of the melanogaster subgroup. The remaining D. teissieri R1 copies contained major deletions or duplications in their 3′ untranslated region. Two clones contained a 115-bp deletion of the 3′ untranslated region starting at residue −296 bp, and three clones contained a duplication of the 329 bp at the 3′ end of the untranslated region. The 115-bp deletion was present in copies with or without the 329-bp duplication, thus explaining why the original PCR product appeared as two diffuse bands.

A 14-bp segment of the 28S gene was located between the 329-bp duplicated sequences in clones tet4, tet19-d and tet22. This 28S sequence corresponded to the 14-bp target site duplication generated upon the insertion of R1 elements into the 28S gene of most insects. The presence of this 14-bp 28S gene sequence suggests that the apparent DNA duplication in these clones was actually generated by two R1 integration events in the same 28S gene. One integration involved a highly truncated R1 element (presumably 329 bp in length) and the second involved a R1 element at least 1.7 kb in length.

**Vertical transmission of R1 and R2 elements:** Based on a variety of morphological and molecular data sets, including allozyme data, mitochondrial DNA sequences, total single copy DNA hybridizations, the nucleotide sequence of individual nuclear genes and chromosomal inversion patterns, LACHAISE et al. (1988) have derived the consensus phylogenetic relationship of the eight members of the melanogaster species subgroup shown in Figure 6A. The age of this subgroup has been estimated to be 17–20 million years (reviewed in LACHAISE et al. 1988). Three distinct species complexes are recognized within the melanogaster subgroup: the melanogaster complex, which includes D. melanogaster, D. simulans, D. mauritiana and D. sechellia; the yakuba complex, which includes D. yakuba and D. teissieri, and the erecta complex, which includes D. erecta and D. aerea. There are two equivocal topologies to this phylogeny. The first deals with the relationship of the yakuba complex with the other two complexes. Most of the data are consistent with the phylogeny in Figure 6A, in which the erecta complex forms the sister group of the other species. However, some data suggest that the melanogaster complex is the sister group to the erecta and yakuba
complexes. The second uncertainty is the relationship of the sibling species D. simulans, D. mauritiana and D. sechellia. Even though the relationship of these three species has been the subject of numerous reports, no clear consensus has been reached.

The phylogeny of the R1 and R2 sequences found within the melanogaster species subgroup is shown in Figure 6, B and C. Because the level of sequence variation detected between the copies of intact R1 and R2 elements from the same species was not much greater than that expected of the PCR misincorporation rate, consensus R1 and R2 sequences have been used for each species. The phylogeny was determined by maximum parsimony using PAUP (Swofford 1991). The branch and bound method was conducted using furthest addition of sequences. The bootstrap data (numbers included within circles) were derived from 1,000 branch and bound replications.

The phylogenetic relationships of the R1 elements (Figure 6B) as well as the R2 elements (Figure 6C) were in agreement with the consensus species phylogeny (Figure 6A). R1 and R2 elements from each of the three species complexes are closer to each other than to members of the other species complexes. R1 and R2 of D. melanogaster are clearly sister sequences to those from the sibling species D. simulans, D. mauritiana and D. sechellia. Only the relationship of the elements within these three sibling species is not clear. In the case of R1, D. sechellia elements form a sister group to those of D. simulans and D. mauritiana (however, the bootstrap value supporting this arrangement is only 73%), whereas the R2 elements of D. simulans form a sister group to D. mauritiana and D. sechellia. As described above, there is at present no consensus relationship between these three species (LACHAISE et al. 1988).

Thus, with the single exception of the loss of R2 elements from the D. erecta and D. orena lineage, R1 and R2 appear to have been stably maintained by vertical transmission since the origin of the melanogaster species subgroup. These data do not rule out the possible transfer of R1 and R2 between species of the same species complex. Indeed, the short branch lengths of the R1 and R2 elements in D. teissieri and D. yakuba imply either introgression events or a slowing of the rate of nucleotide evolution. This issue is addressed in the following report in which the rates of nucleotide sequence evolution of R1 and R2 are compared with that of various nuclear genes (EICKBUSH et al. 1994).

DISCUSSION

There are two nonexclusive explanations for the presence of R1 and R2 elements in over 90% of the insect species tested (JAKUBCZAK et al. 1991). The elements may be highly stable in most species and maintained in each lineage by vertical transmission through the germline. Alternatively, the elements may be unstable and occasionally lost from any species, but such species may be continually reinfected with R1 and R2 elements from other species. The results in this report have given us the first indication of the relative contribution of each of these models in explaining the distribution of R1 and R2. The sequence phylogeny of the R1 and R2 elements from members of the melanogaster species subgroup follows that of the species phylogeny, suggesting that the elements have been stably associated with the species of this group during the estimated 17–
20 million years of its existence (Lachaise et al. 1988). The only exception to this stability is the absence of R2 elements from the *arena*/erecta species complex. We have recently isolated R2 elements from the *takahashii* and *ananassae* species subgroups as well as from the *obscura* species group (Eickbush et al. 1995; W. Lathe, unpublished data). The sequence divergence of these R2 elements compared with those from the *melanogaster* subgroup are generally consistent with their species phylogeny. Therefore we believe R2 elements were lost from the *arena*/erecta lineage rather than gained in the lineages leading to the other two species complexes.

Few reports have appeared detailing the vertical transmission of retrotransposable elements within a species phylogeny. The most extensively studied is that of the *LI* family of elements in rodents. The phylogeny of these elements follows that of the rodent species phylogeny, suggesting vertical transmission within these species (Hardies et al. 1986; Pascale et al. 1993). Other attempts to trace the vertical transmission of retrotransposable elements in a lineage include the SURL element in four sea urchin species (Springer et al. 1991) and the copia-like elements in three species of wheat (Vanderwiel et al. 1993). In each of these examples, analysis of the sequence relationships of the element in different species is complicated by the fact that each species contains thousands of copies of each element, and the level of sequence divergence of the elements within each species can be as high as the level of divergence between the species. This sequence diversity is a result of nearly all newly transposed elements being defective and rapidly accumulating mutations at the neutral rate. Therefore extensive characterization of large numbers of these elements in each species is required to trace the phylogeny of the active elements within any species history.

R1 and R2, which are only distantly related non-LTR retrotransposable elements, represent the two clearest examples of how small families of transposable elements can remain stably associated with their host's genome over extended evolutionary periods. The unambiguous nature of this conclusion is a result of the remarkable uniformity of the R1 and R2 sequences in each species of the lineage. The observed level of within-species sequence divergence for the R1 and R2 elements was similar to that of the PCR error rate used to obtain these sequences. Thus the level of sequence variation was below our level of sensitivity. The likely explanation for the sequence uniformity of R1 and R2 is the highly specialized niche each element has found within the host genome. The rDNA genes of eukaryotes undergo the process of concerted evolution by recombination mechanisms involving unequal crossovers and gene conversions (Petes 1980; Szostak and Wu 1980; Coen et al. 1982; Seperack et al. 1988). Recombination in the rDNA units of *D. melanogaster* is sufficiently high to generate a fivefold range in the number of repeats in flies obtained from the same population (Lyckegaard and Clark 1989, 1991). R1 and R2 elements are also subject to significant changes in copy number. Based on a survey of 27 geographical strains of *D. melanogaster*, we have shown (Jarubczak et al. 1992) that the number of R1 elements varied from 14 to 150 and could occupy from 17 to 67% of the total rDNA units. The number of R2 elements varied from 3 to 60 and could occupy from 2 to 28% of the rDNA units. A similar range of R1 and R2 levels has been detected even within a single population of *D. melanogaster* (Lyckegaard and Clark 1991). This dramatic variation in numbers of R1 and R2 elements suggests that they also turnover rapidly by the unequal crossovers within the rRNA gene loci.

That the uniformity in sequence of the R1 and R2 copies within a species is dependent on their being located within the rDNA loci is shown by an analysis of copies that are outside the rRNA gene loci. In *B. mori*, copies of R1 and R2 located outside the rDNA units showed significant sequence variation (Xiong et al. 1988). In *D. melanogaster*, copies of R1 elements inserted in fragments of the 28S gene located in the centromeric heterochromatin have 8% nucleotide divergence with those copies inserted in the rDNA gene loci (Kidd and Glover 1980; Peacock et al. 1981). In this report, the only examples of either R1 or R2 elements exhibiting significant within-species sequence variation were the R1 elements of *D. tussieri*. Intact (i.e., no premature termination codons or frameshifts) R1 elements of *D. teissieri* had twice the level of sequence variation found in any other species. Here too the sequence variation could be explained by the inadvertent cloning of R1 elements from outside the rRNA gene locus. Sequence analysis of the 670 bp of the 28S gene flanking the *D. teissieri* R1 elements revealed that at least two of the R1 elements scored as intact were located within 28S genes that were not undergoing the process of concerted evolution with the majority of the 28S genes of the species and thus are probably located outside the rRNA gene loci.

The ability of R1 and R2 to maintain themselves in the rDNA locus implies that they are capable of increasing their numbers by retrotransposing when their level in the genome is reduced by recombination. Similar models based on the ability of the element to transpose when their copy number is reduced have been suggested for both prokaryotic and eukaryotic transposable elements (Kleckner 1990; Rio 1990). Some of these transposable elements have been shown to encode a repressor molecule that inhibits the transposition process until the gradual elimination or inactivation of elements reduces the level of this repressor. If such repressor molecules are produced by R1 and R2, they may be involved in transcription of the rRNA genes themselves.
because \( R1 \) or \( R2 \) insertions inactivate transcription of the entire rDNA unit (LONG and DAWID 1979; KIDD and GLOVER 1981; JAMRICH and MILLER 1984).

**Why are \( R1 \) and \( R2 \) elements stable?** The distribution and sequence relationship of many RNA- and DNA-mediated transposable elements has been studied in the *melanogaster* species subgroup (MARTIN et al. 1983; BROOKFIELD et al. 1984; STACEY et al. 1986; SILBER et al. 1989; MONTCHAMP-MOREAU et al. 1993). However, the stabilities of these mobile elements in the subgroup are not well established. In most cases not all eight members of the subgroup were screened, and in many of the early studies using genomic blots, controls were not conducted for the level of hybridization that would be expected in each species. In the case of \( I \) and \( hobo \), which are believed to be present in all eight species, detailed analysis in *melanogaster* has suggested that all active copies were at one time lost from the genome and only more recently reentered the species (BUCHETON et al. 1986; PASCUAL and PERIQUET 1991; SIMMONS 1992). These findings, along with the evidence for the horizontal transfer of \( P \) and \( mariner \) elements from outside the species group (DANIELS et al. 1990; MARUYAMA and HARTL 1991), have strengthened the arguments that horizontal transfers between related species may be necessary for some transposable elements to remain active over extended evolutionary periods in Drosophila. It has been suggested that the low equilibrium copy number per genome maintained by most Drosophila mobile elements and the tendency of these elements to accumulate defective copies that are still able to transpose should lead to the stochastic loss of all functional elements from the genome (KAPLAN et al. 1985).

How have \( R1 \) and \( R2 \) overcome the apparent instability seen in many other elements? The unique location of \( R1 \) and \( R2 \) in the rDNA locus does not eliminate the problem of accumulating defective copies. Unequal crossovers in the rDNA locus are as likely to spread defective copies in the locus as they are functional copies. One explanation for the stability of \( R1 \) and \( R2 \) in Drosophila is that they supply a beneficial function to the host. The elements may supply endonuclease activities that promote recombination within the rDNA locus (HAWLEY and MARCUS 1989; RASOOLY and ROBBINS 1991) or they may provide a mechanism to slow the growth rate of the organism during development (TEMPLETON et al. 1993). These suggestions are intriguing but do not explain why both \( R1 \) and \( R2 \) elements are stable within a lineage or why some species contain multiple families of \( R1 \) or \( R2 \) elements (BURKE et al. 1993). Thus, even for the relatively simple situation in which mobile elements insert into a unique region of the genome, it remains unresolved as to whether the elements have a beneficial effect on the host or whether they can be considered purely selfish genetic elements (see discussions in SYVANEN 1984; CHARLESWORTH and LANGLEY 1989). Whether \( R1 \) and \( R2 \) are eventually found to have a positive or entirely negative fitness effect on the host, we suggest that certain characteristics of \( R1 \) and \( R2 \) give them greater stability within the host genome compared with transposable elements that insert in many locations throughout the genome.

1. \( R1 \) and \( R2 \) are assured of a population of uniform target sites for the insertion of new copies. Newly formed \( R1 \) and \( R2 \) copies will insert into sites identical to those occupied by the donor elements; thus the new elements can be developmentally regulated in the same manner as the donor elements. Transposable elements that insert at numerous locations throughout the genome run the risk of inserting newly made copies into regions of the genome where the elements can not be expressed or where their expression can not be appropriately regulated.

2. Under the conditions where the retrotransposition of \( R1 \) and \( R2 \) elements is most likely to occur, that is, at low copy numbers in the genome, the insertion of new elements has little effect on the host. The large range seen in the number of rDNA units per genome suggests that most species have more than sufficient rDNA units for their survival (LONG and DAWID 1980). Thus a new \( R1 \) or \( R2 \) insertion is likely to have minimal effect on fitness. By contrast, each time a nonspecific transposable element jumps or makes a new copy for insertion, there is the danger that it will affect the expression of a normal host gene and thus be selected against.

3. \( R1 \) and \( R2 \) do not induce ectopic exchange. Recombination between \( R1 \) elements or \( R2 \) elements in the rDNA locus would be no different from recombinations involving the rDNA unit itself. Nonspecific transposable elements inserted in euchromatin have been postulated to give rise to chromosomal rearrangements by ectopic exchange between elements (CHARLESWORTH and LANGLEY 1989; CHARLESWORTH et al. 1992a,b). Thus there is likely to be selective pressure against transposable elements in euchromatic regions, even when they do not directly affect the expression of a host gene.

We suggest that by eliminating the randomness associated with the insertion of new copies and the problems this can cause in their own expression as well as the expression of host genes, \( R1 \) and \( R2 \) can maintain their numbers in the genome of their hosts over considerable evolutionary periods. \( R1 \) and \( R2 \) would therefore depend less on horizontal transfers between species to maintain themselves in a lineage than do nonspecific transposable elements. However, the results of this report should not be interpreted as indicating that these elements never undergo horizontal transfer. Horizontal transfers may still be necessary to explain the existence
in some insect species of multiple highly divergent families of R1 and R2 (Burke et al. 1993).

We thank W. Burke and W. Lathe for critical comments on the manuscript. We thank S. Connors for assistance in the genomic blot analysis, and W. Lathe for assistance in the utilization of the PAUP programs. This research was supported by National Science Foundation grant MCB-9219123.

LITERATURE CITED


Jaruzczak, J. L., Y. Xiong and T. H. Eickbush, 1990 Type I (R1) and Type II (R2) ribosomal DNA insertions of Drosophila melanogaster are retrotransposable elements closely related to those of Bombus mori. J. Mol. Biol. 212: 37–52.


Jamrich, M., and O. L. Miller, 1984 The rare transcripts of interrupted rDNA genes in Drosophila melanogaster are processed or degraded during synthesis. EMBO J. 3: 1541–1545.


Peacock, W. J., R. Appels, S. Endow and D. Glover, 1981 Chromo-

Rex are repeated neomorphic loci in the Drosophila melanogaster ribosomal DNA. Genetics 129: 119-132.


Xiong, Y., W. D. Burke, J. L. Jakuschak and T. H. Eickbush, 1988 Ribosomal DNA insertion elements R1Bm and R2Bm can transpose in a sequence specific manner to locations outside the 28S genes. Nucleic Acids Res. 16: 10561-10573.