

---

**Ribosomal DNA insertion elements R1Bm and R2Bm can transpose in a sequence specific manner to locations outside the 28S genes**

---

Yue Xiong, William D.Burke, John L.Jakubczak and Thomas H.Eickbush

---

Department of Biology, University of Rochester, Rochester, NY 14627, USA

---

Received August 12, 1988; Accepted October 19, 1988

---

**ABSTRACT**

A fraction of the ribosomal 28S genes in some insects are interrupted at specific sites by insertion elements R1 and R2 (also called Type I and II). These elements contain long open-reading frames with homology to reverse transcriptase. We have identified in the silkworm, *Bombyx mori*, copies of these elements which have inserted into sites outside the ribosomal DNA (rDNA) units. The 3' ends of all "non-rDNA" elements are identical to the elements within the 28S genes; however their 5' ends are often truncated. Each non-rDNA copy has inserted into sequences that exhibit similarity to their target sites in the 28S gene. We also demonstrate by genomic blot analysis of different strains of *B. mori* that insertions of R1 and R2 outside the rDNA units have been infrequent, while considerable turnover of elements has occurred within the rDNA locus. One race of *B. mori* has lost all copies of R1 from its rDNA units, while retaining normal levels of R2. The level of both R1 and R2 have significantly increased in a tissue culture line. These findings add considerable support to the model that R1 and R2 are retrotransposable elements that utilize sequence specific endonucleases in their integration into the genome.

**INTRODUCTION**

Transposable elements and retroviruses of eukaryotes are capable of inserting into a large number of sites within the genome of the organism they inhabit (1). This flexibility in the selection of target sites occasionally results in their insertion into or flanking active genes, which can have striking detrimental effects on their host. If one views these mobile elements as "selfish" or "parasitic" (2,3) then it might be more advantageous for the element to become specialized for insertion into sites within the genome that could be tolerated by the host.

A fraction of the ribosomal genes in a number of insect species have been found to contain specific DNA insertions (reviewed in 4). The best characterized of these elements are the type I and type II inserts first found in *Drosophila melanogaster* (5-7). Type I elements have also been found in *D. virilis* (8), the blow fly *Calliphora erythrocephala* (9) and the silkworm *Bombyx mori* (10,11). Type II elements have only been found in sibling species of *D. melanogaster* (12) and in *B. mori*. (10,11). We have recently suggested (13) that these elements be renamed R1 and R2 to avoid confusion with other insertion elements, particularly the fungal mitochondrial intron sequences to which they are related (14,15).

Complete nucleotide sequence determination of the R1 and R2 elements from *B. mori* (13,16) has revealed that each element contains an over 1,000 amino acid open-reading frame. A

central portion of this open-reading frame contains sequences similar to reverse transcriptases found in retroviruses and retrotransposable elements. The R1 element contains a second open-reading frame which has several properties of retroviral *gag* genes. Sequence similarities in the reverse transcriptase region are highest to a series of mobile elements which like R1 and R2 do not end in long terminal repeats (15,16).

Based upon their structural similarities to retrotransposons, we have suggested that R1 and R2 are, or at least were at one time, transposable elements (13,16). If these elements are mobile, then it is necessary that they be remarkably sequence specific in their insertion since they are found in the identical nucleotide position in multiple 28S genes in each species. An alternative explanation for the sequence specific location of the R1 and R2 elements is that the recombination mechanisms which maintain very high levels of sequence homology between the ribosomal DNA (rDNA) repeats, have simply expanded and maintained the number of R1 and R2 elements within the rDNA loci of different species. In this latter model it is not necessary that these elements were originally sequence specific, nor is it necessary that they remain active. In this report we present data which indicates that R1 and R2 can on rare occasions insert in a sequence specific manner to sites outside the rDNA locus. These findings support the model that R1 and R2 are highly specialized for transposition into particular sequences in the eukaryotic genome.

## MATERIALS AND METHODS

### Isolation and sequence determination of non-rDNA R1 and R2

All clones were isolated from a lambda phage library of partial Sau3AI digested genomic DNA of strain 703 cloned into Charon 35 (gift of B. Hibner). Clones containing non-rDNA R1 and R2 elements were isolated from this library by hybridizing one set of plaque lifts with the 1.7 kb BamHI - XbaI fragment corresponding to the 3' half of the 28S gene in clone B108 (11), and a duplicate set of filters with either the 2.5 kb KpnI fragment isolated from the R1 element in clone B78 (16) or the 1.4 kb PstI - SstI fragment from clone B131 (13). Hybridizations were conducted at 68°C in 0.6 M saline (0.6 M NaCl, 0.12 M Tris hydrochloride [pH 8], 4 mM EDTA) as previously described (17). After preliminary restriction mapping and Southern analysis, the appropriate regions of the lambda clones were subcloned into pUC18. Specific restriction fragments containing the junction sequences were cloned into M13mp18 and M13mp19 vectors (18) for sequence determination (19). The direction and extent of sequence determination for each element are shown in Figures 1 and 3.

### Genomic blot analysis

Genomic DNA was isolated from five sibling female moths from each of the following strains: 703, a homozygous wildtype line derived from European strain 703 (20); C108, a standard Chinese strain used as a tester strain in genetic studies of the chorion locus (21); Akajuku, a Japanese race; Shokei, a Chinese race; Old European 16 (OE16), a European race; and

Cambodge, a tropical multivoltine race from Cambodia. The last four are standard geographical races approved by the Committee on Silkmoth Genetics of the Japanese Society of Sericultural Science, and were obtained from the Sericultural Experimental Station, Ministry of Agriculture, Fishery and Forestry (Tsukuba, Japan). DNA was also isolated from the tissue culture line BM5 (22). The DNA probes for R1, R2 and the rDNA unit, as well as the conditions of the hybridization, were the same as those used in the genomic library screen described above.

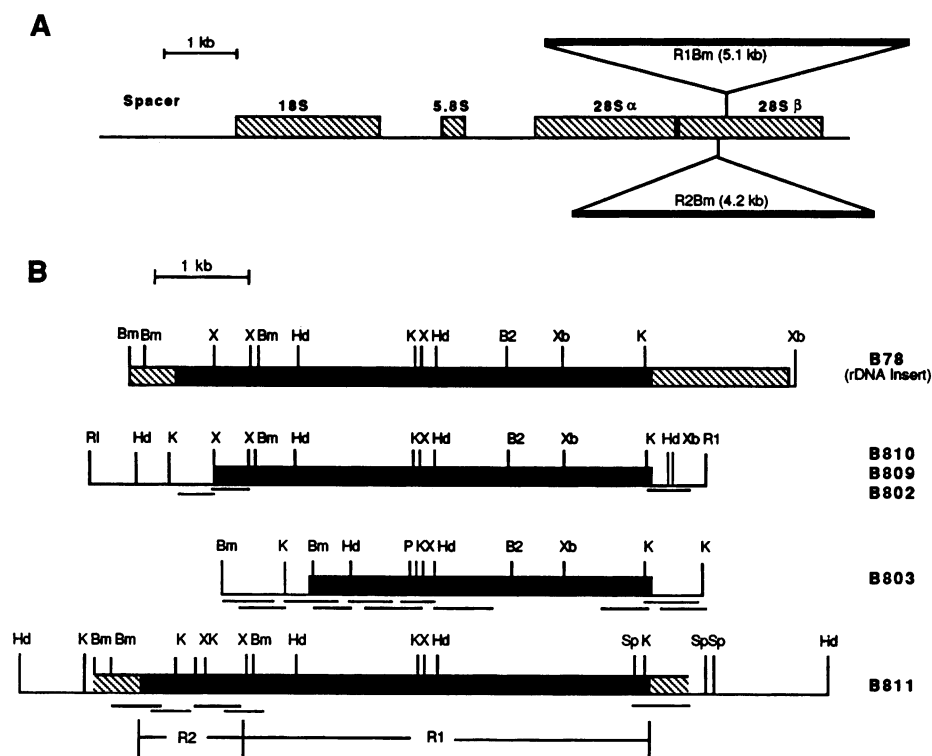
## RESULTS

Based upon genomic blot analyses, *B. mori* strain 703 contains approximately 20 copies of R1 and 20 copies of R2 located within the 28S genes of its estimated 240 rDNA repeats (13, 16,23). Only a limited number of R1 and R2 elements are located outside the rDNA repeats (13,16). To isolate these non-rDNA copies, 60,000 clones from a Sau3A genomic library of *B. mori* DNA were double-screened: one set of filters was probed with the 3' half of the 28S gene (the region containing the R1 and R2 insertion sites, Figure 1A), a duplicate set of filters was probed with internal fragments from either R1 or R2 (see MATERIALS and METHODS). All lambda phage clones which hybridized to the R1 and R2 probes but at reduced levels with the 28S gene probe were isolated, and will be referred to as "non-rDNA" copies of R1 and R2.

### Non-rDNA R1 Elements

Four lambda clones hybridizing to R1 but not to the 28S gene probe were obtained from the genomic library screen. Three of these clones B802, B809 and B810, while different in their end points, contained the same element (Figure 1B). The R1 element in these clones contained all the restriction endonuclease cleavage sites found in R1 copies within the rDNA units (represented by clone B78). The fourth clone isolated from the genomic library, B803 (Figure 1B), corresponded to a second non-rDNA R1 element. This R1 copy did not contain the two XhoI sites at the 5' end of the R1 element, and the distance between the KpnI and HindIII sites within the element was shorter than that of the R1 element in either B78 or B810. Thus the B803 R1 element appeared to be truncated at its 5' end and contained an approximately 0.6 kb internal deletion.

A fifth non-rDNA R1 clone isolated from the genomic library, B811, hybridized to R1 and R2 probes, and at a reduced level to the 28S gene probe. Restriction endonuclease (Figure 1B) and nucleotide sequence analysis (data not presented) indicated that this clone contained a typical 5' end of an R2 element inserted into the 28S gene site. The R2 sequence extended to nucleotide 1324 and then continued with the R1 sequence, starting at nucleotide 665 (nucleotide positions in R1 and R2 are numbered as in ref. 13, 16). The R1 element ended at a typical 3' junction with the 28S gene. We assume that the origin of this clone was an rDNA repeat with both R1 and R2 insertions. A subsequent deletion event between these two elements resulted in the loss of the 3' end of the R2 element, the 5' end of the R1 element as well as the 74 bp of intervening 28S sequence. The rDNA unit itself has also undergone considerable DNA rearrange-



**Figure 1. Structure of the non-rDNA R1 elements.** Diagonally shaded bars, rRNA genes; solid bars, R1 and R2 elements. (A) Diagram of an rDNA unit showing the location of the R1 and R2 insertions. A 30 bp hidden break is found near the center of the 28S gene (24). (B) Restriction maps of the non-rDNA R1 copies compared to a typical R1 insertion within a 28S gene (clone B78). Clone B811 contains a fused R2 and R1 element and is flanked by a short segment of the 28S gene (see text). Abbreviations for the restriction sites: Bm, BamHI; X, XhoI; Hd, HindIII; K, KpnI; B2, BglII; Xb, XbaI; Sp, SphI; P, PstI; R, EcoRI; and S, SstI. Shown below each restriction map are horizontal lines indicating the extent of the sequence determinations.

ment, since homology to the 28S gene extends for less than 500 bp at either end of the fused R2-R1 element.

Restriction analysis of the R1 elements in clones B803, B810 and B811 revealed that they corresponded to the three aberrant R1 copies detected in genomic blots of *B. mori* DNA (16). More than three aberrant bands are frequently seen on genomic blots (see Figure 5A), due to a low level of restriction polymorphism of R1 elements located within the rDNA units (16).

The 5' and 3' borders of the R1 elements present in clones B810 and B803 were sequenced and are compared in Figure 2A with the 28S gene junction sequences (represented by clone B78). Both elements contained the identical 3' end found in all sequenced R1 elements located within the 28S genes. The 5' ends of the non-rDNA R1 elements, on the other hand,

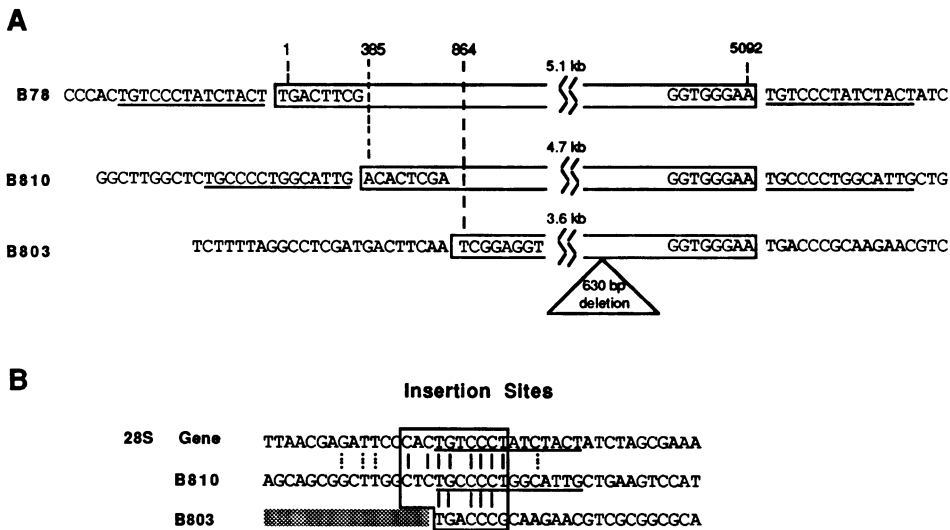


Figure 2. Nucleotide sequences at the borders of non-rDNA R1 elements. 28S gene sequences flanking a rDNA R1 element (represented by clone B78) are also shown. (A) R1 element sequences are boxed. The 14 bp flanking duplication in B78 and B810 are underlined. Sequence identity between the R1 element in B78 and the element in B810 begins at nucleotide 385, while sequence identity with the element in B803 begins at nucleotide 864. Nucleotides are numbered as in ref. 16, Figure 2. Clone B803 has an additional internal deletion of 630 bp. (B) Comparison of the 28S insertion site with the putative non-rDNA insertion sites. Flanking 14 bp duplications are underlined. Sequences at the 5' end of the element in B803 did not contain similarity to the 28S gene and have not been included since this end of the element has probably undergone a deletion (see text).

were missing 384 bp relative to a complete R1 element in the case of B810, and 863 bp in the case of B802. Neither element was flanked by 28S gene sequences. A 14 bp duplication was found flanking the R1 element in B810. Since 14 bp duplications of the 28S gene target site are found flanking all R1 elements within the rDNA units of *B. mori* (10,11) as well as R1 elements in three other insect species (6-9), the missing 5' end of the B810 R1 element was probably not the result of a deletion subsequent to insertion. In the case of the R1 element in B803, a deletion of the 5' end subsequent to the insertion event appears likely since it is not flanked by a DNA duplication. The B803 R1 element also contained a 630 bp internal deletion (nucleotide 1496 to 2125). The DNA sequences immediately flanking the 3' end of the R1 elements in B803 and B810 belong to the large repetitive DNA fraction of the *B. mori* genome. These repetitive sequences are highly heterogeneous and thus we have not been successful in cloning examples of the unoccupied insertion sites.

We have compared the R1 insertion site of the 28S genes with the presumed insertion sites of the two non-rDNA copies of R1 to determine if they have inserted in a sequence specific manner. As shown in Figure 2B the insertion site of B810 contained considerable sequence

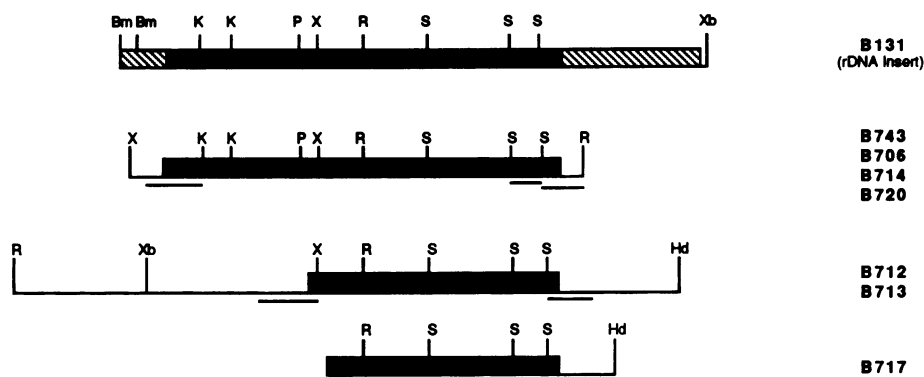


Figure 3. Structure of the non-rDNA R2 elements. Diagonally shaded bars, rRNA genes; solid bars, R2 elements. Restriction maps of the non-rDNA copies are compared to a typical R2 element within a 28S gene (clone B131). The R2 element in clone B717 has been truncated by cloning at a *Sau*3A site. Restriction site abbreviations are as in Figure 1. Shown below each restriction map are horizontal lines indicating the extent of the sequence determinations.

similarity to the 28S site, centered at the 5' end of the 14 bp target site duplication (boxed region). The sequence similarity is 8 of 10 nucleotides, which can be extended to 11 of 16 nucleotides by expanding the comparison another 6 bases in the 5' direction. Also presented in Figure 2B is the 3' junction of the R1 element in B803. The 5' junction of this element was not included in the figure, since it probably has undergone a deletion, and no similarity of the current B803 5' junction with the 28S gene target site was detected. It is significant that the sequence 3' of the B803 R1 element contained the same 5 of 6 bp match with the 28S sequence (TGNCCC) as was found in B810. We conclude that both copies of the R1 element which have inserted outside the rDNA repeat, have done so in a specific manner into sequences that are similar to the 28S gene target site. Furthermore, their effect on the target sequence, at least in the case of B810, is a target site duplication identical in length to that found in the rDNA locus.

#### Non-rDNA R2 Elements

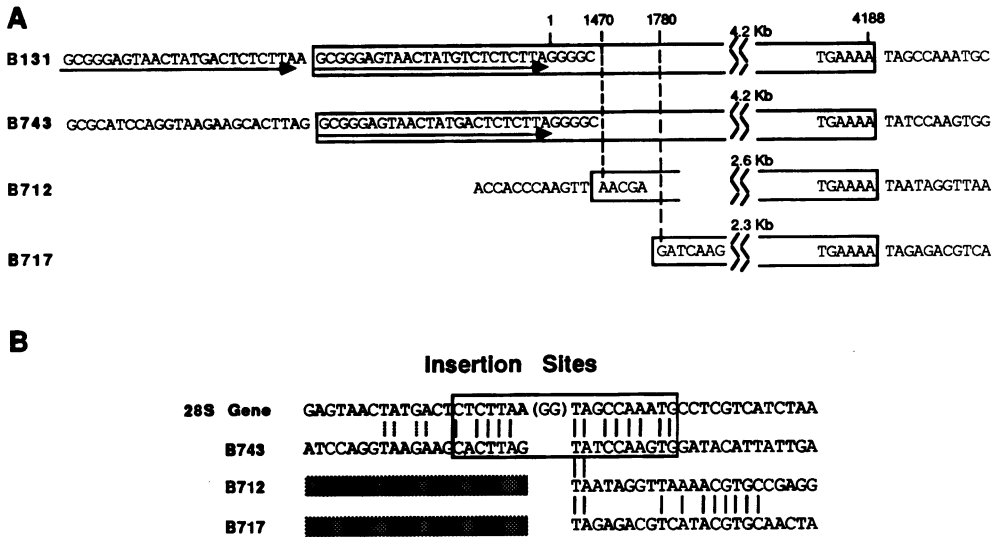
Screening of the *B. mori* genomic library resulted in eight clones that hybridized to the R2 element but did not hybridize to the 28S gene. Restriction maps of these elements are compared in Figure 3 to the map of a typical R2 element inserted into a 28S gene, clone B131 (13). The nucleotide sequences of the boundaries of these elements with the flanking chromosomal sequences are presented in Figure 4A. Four clones B706, B714, B720 and B743, while different in their end points all contained the same non-rDNA R2 element. The restriction map of this element was identical to that of R2 elements in the 28S genes, and sequence analysis indicated that the 5' and 3' ends of the non-rDNA element were identical to that in B131. This element also contained a 24 bp sequence from the 28S gene identical to the 24 bp tandem 5' duplication found in two of six copies of R2 located in the rDNA repeat (13). We believe this 28S gene

sequence was part of the transposition unit that inserted outside the rDNA repeats, rather than a short fragment of a 28S gene into which an R2 element inserted. No DNA duplication was found flanking the R2 element in B743. This finding is consistent with R2 insertions in the rDNA unit where target site duplications have not been observed (11,13). Indeed when R2 inserts into a 28S gene it appears that two nucleotides are deleted from the target sequence. This can be seen in Figure 4 by comparing the DNA flanking the R2 insert in B131 (part A) with the sequence of the target site (part B). The adenine nucleotide immediately upstream of the two deleted guanines is also deleted in two of the six rDNA insertions analyzed (clones B98 and B703 in Figure 3 of ref. 13).

The presumed insertion site of the R2 element in B743 has considerable similarity (12 of 15 bp) to the 28S gene insertion site (Figure 4B, boxed region). In order to obtain this level of similarity it was necessary to assume that as with R2 insertions in the 28S genes a 2 bp deletion had occurred in the B743 insertion site. Thus the mechanism by which the genomic DNA was cleaved and an R2 element inserted outside the rDNA locus appears similar to that in the rDNA locus. Cloning of the unoccupied B743 insertion site to confirm that 2 bp were deleted upon R2 insertion was not possible. Southern blots using flanking DNA sequences as probe have revealed that the B743 insertion site is occupied by an R2 element in every strain of *B. mori* we have tested (data not included).

A second non-rDNA R2 element obtained from the library screen was found in clones B712 and B713 (Figures 3 and 4). This element contained a complete 3' end and a 5' end truncated at nucleotide 1470. Due to the absence of target site duplications associated with R2 insertions we are unable to distinguish whether the B712 R2 element was truncated before insertion, or deleted after insertion. The third non-rDNA R2 element obtained from the library screen, found in clone B717, also contained a complete 3' end. Its 5' end was not available since the phage clone ends at a *Sau3A* site within the element (position 1780). The identity of the 3' ends of the B712 and B717 non-rDNA elements to that found in R2 elements inserted into the 28S gene, suggests that they have undergone specific integration events, rather than fragments of the rDNA unit which have inserted at random outside the nucleolus. Only the first two (TA) or three (TAG) nucleotides of the 3' flanking DNA of these non-rDNA elements are identical to the 28S target sequence. Thus these two target sequences do not have as much similarity to the 28S gene as the target sequence in clone B743. However, starting 9 bp from the 3' end of the R2 element there are 8 of 10 identical nucleotides in B712 and B717. This region contains the sequence (CGT) also found at the same position in the 28S sequence. Apparently all three non-rDNA elements have inserted in a sequence specific manner, but the sequences involved in recognition in the B712 and B717 insertions are somewhat different from that in the B743 insertion.

A fourth non-rDNA R2 element isolated from the genomic library (clone B710) has not been included in either Figures 3 or 4, because it is a highly diverged copy of R2. The restriction



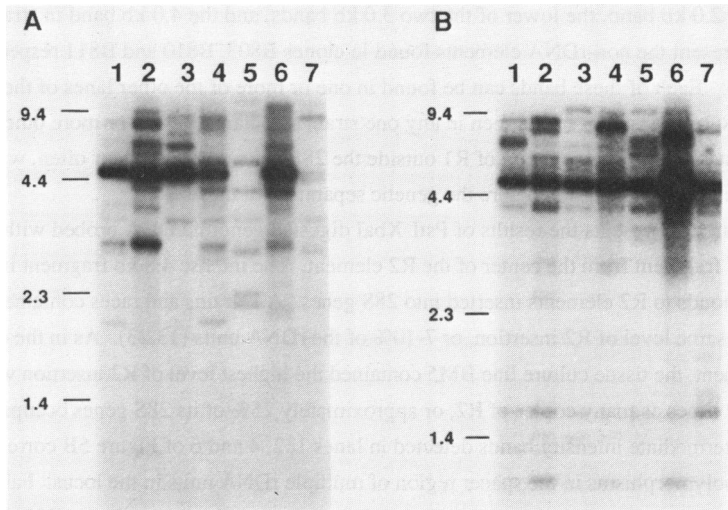
**Figure 4. Nucleotide sequences at the borders of non-rDNA R2 elements.** 28S gene sequences flanking a rDNA R2 element (represented by clone B131) are also shown. (A) R2 sequences are boxed. The 23 bp duplication at the 5' end of the element in B131 that is also present in the B743 R2 element is underlined. Nucleotides are numbered as in ref. 13 Figure 3, except that nucleotide number one is the nucleotide following the 23 bp duplication. The element in B717 has been truncated by cloning at a *Sau*3A site. (B) Comparison of the 28S insertion site with the putative non-rDNA insertion sites. The two nucleotides deleted from the 28S gene upon R2 insertion are included in parenthesis. It is not known whether a two base pair deletion also occurs in the B743 site. Sequences at the 5' end of the element in B712 did not contain similarity to the 28S gene and have not been included since this end of the element may have undergone a deletion.

map of the B710 R2 element was entirely different from the consensus R2 map, and Southern analysis indicated that at least one large (1.5 kb) insertion was present within the element. Nucleotide sequence determination of an internal segment revealed that sequence divergence between the B710-R2 element and the consensus R2 element was nearly 13% (88/680 bp). Attempts to sequence the borders of this element revealed several smaller insertions and deletions had occurred in the element, such that the exact 5' and 3' boundaries could not be unambiguously determined.

**Distribution of R1 and R2 in Different Geographical Races of *B. mori***

In an effort to monitor the extent to which expansions or contractions in the number of R1 and R2 elements have occurred within the rDNA loci, and the frequency with which they move outside the rDNA units, we have conducted genomic blots of DNA from several standard geographical races of *B. mori* and the tissue culture line, BM5. Figure 5A presents the results of *Hind*III digested DNA probed with the 1.6 kb *Bgl*II-*Kpn*I fragment from the 3' end of the R1





**Figure 5. Genomic blot hybridizations of *B. mori* DNA from different sources probed with R1 or R2 sequences.** Lane 1, strain 703; lane 2, race Shokei; lane 3, race OE16; lane 4, strain C108; lane 5, race Cambodge; lane 6, tissue culture line BM5; and lane 7, race Akajuku. (A) For each lane 2 micrograms of genomic DNA was digested with HindIII and probed with the 1.6 kb BglII - KpnI fragment from the 3' end of R1. (B) Each lane contained 2 micrograms of genomic DNA digested with PstI- XbaI and probed with the 1.1 kb PstI -SstI fragment from R2. Numbers at left indicate length (in kilobases) of DNA standards.

element. The intense 5.1 kb band in five of the lanes represents R1 elements inserted within the 28S genes. The less intense bands in each lane correspond to either restriction polymorphisms of the R1 elements within the rDNA units or to R1 elements located outside typical rDNA repeats. The very weakly hybridizing bands seen in some lanes probably represent highly divergent copies of the R1 element .

Each DNA source in Figure 5 contained approximately the same number of rDNA repeats per genome (data not shown). Since R1 elements occupy approximately 7-9% of the 240 rDNA units in strain 703 (16,23), the fraction of the 28S genes occupied by R1 elements in the different races of *B. mori* can be directly estimated from this Figure. C108 (lane 4) contains approximately the same level of occupation as 703 (lane 1). Shokei (lane 2) and OE16 (lane 3) have from 12-15% of their rDNA units occupied, while the tissue culture line (lane 6) had an even higher level of insertion (20%). Perhaps the most interesting strains were Akajuku (lane 7) and Cambodge (lane 5). Based upon the blot in Figure 5A, as well as additional genomic blots not shown, Cambodge has approximately six copies of R1, however none of these copies appear to be inserted into rDNA repeats. Akajuku also has approximately six copies of R1 of which two appear to be inserted within the 28S genes.

The 2.0 kb band, the lower of the two 3.0 kb bands, and the 4.0 kb band in strain 703 (lane 1) represent the non-rDNA elements found in clones B803, B810 and B811 respectively (Figure 1B). Each of these bands can be found in one or more of the other lanes of the blot. Indeed, most of the variant bands seen in any one strain are shared by one or more other strains. This suggests that the transposition of R1 outside the 28S gene does not occur often, with many of the current copies generated before the genetic separation of these races.

Figure 5B presents the results of PstI-XbaI digested genomic DNA probed with the 1.1 kb PstI-SstI fragment from the center of the R2 element. The intense 4.8 kb fragment in each lane corresponds to R2 elements inserted into 28S genes. All strains and races contained approximately the same level of R2 insertion, or 7-10% of the rDNA units (13,23). As in the case of the R1 element, the tissue culture line BM5 contained the highest level of R2 insertion with nearly three times as many copies of R2, or approximately 25% of its 28S genes occupied by R2. The intermediate intensity bands detected in lanes 1, 2, 4 and 6 of Figure 5B correspond to restriction polymorphisms in the spacer region of multiple rDNA units in the locus. Fainter bands correspond to non-rDNA copies, or restriction polymorphisms in a single rDNA unit. The non-rDNA R2 elements in clones B743 and B712 correspond to two of the fragments slightly larger than the the major 4.8 kb band (the location of the B717 element in this digest is not known). As in the case of R1, most of the fainter bands are found in a number of the different strains. This suggests that many of the R2 elements inserted outside the rDNA locus before the genetic separation of the different races.

## DISCUSSION

R1 and R2 elements appear to suppress the transcription of the rDNA units they occupy in *D. melanogaster* (25-27). As a result one would predict that negative selective pressure and the recombinational mechanisms that give rise to the concerted evolution of the ribosomal genes should eventually eliminate these elements. It has been suggested that the maintenance of these elements in the rDNA units could be explained if they are capable of repeated insertions (28,29). Clearly the structure of these elements as revealed in *B. mori* are consistent with their origins as retrotransposable elements (13,16). The identification of R1 and R2 elements that have moved outside the rDNA repeats to sites with similarity to the 28S insertion sites is evidence that these elements are both mobile and sequence specific.

R1 and R2 elements lack long terminal repeats (LTRs) or inverted repeats, which probably explains why they were not immediately recognized as transposable elements. It has recently been shown that there is a distinct class of such non-LTR retrotransposable elements distributed throughout the plant and animal kingdoms (13,16,30-35). Based on the homology of their reverse transcriptase encoded sequences, these non-LTR elements are more closely related to mitochondrial class II intron sequences of fungi, than to retroviral or copia-like transposable elements (15).

If the maintenance of R1 and R2 in the rDNA locus is to be explained by their ability to transpose, they must be remarkably specific in their insertion not only for the rDNA locus but also for a particular sequence of the rDNA repeat. This degree of specificity has not been observed in any other retrotransposable element or retrovirus. While some retrotransposable elements have been shown to exhibit some sequence preference (for example, refs. 36-38), this preference is not sufficient to limit their insertion into a distinct set of sequences within the genome. One possible explanation of the remarkable insertion specificity of R1 and R2 is that they utilize the unique rRNA processing, ribosomal assembly or DNA recombination mechanisms that occur within the nucleolus. For example, they could insert at the RNA level followed by reverse transcription and DNA recombination back into the rDNA locus. However, several features of the non-rDNA copies of both R1 and R2 indicate that this is not the case. First, the R1 and R2 elements that have inserted outside rRNA genes contain the same 3' end found in all rDNA copies. This suggests that the priming for reverse transcription and/or the processing of this transcript is identical to that used to generate the rDNA inserted copies. Second, the effect the non-rDNA copies of R1 and R2 have had on the target sequence is the same as in the rDNA locus. The R1 element in B803 is flanked by a 14 bp duplication, as are all sequenced R1 elements within the 28S genes. The target site in the B743 R2 element appears to have undergone a 2 bp deletion, as have all sequenced R2 copies within the 28S genes. Third, the non-rDNA copies of R1 and R2 have inserted into DNA sequences that exhibit similarity to their target site in the 28S gene. Clearly the mechanism for the insertion of R1 and R2 into the genome is not dependent upon the unusual features of the 28S gene repeating unit or its transcription. The simplest model to explain the data is to assume that R1 and R2 utilize sequence specific integrases. As direct confirmation of this model we have recently expressed the open reading-frame of R2 in *E. coli* and have found that it does encode an endonuclease activity with high specificity for the 28S gene insertion site (39). The cleavage made by this endonuclease is consistent with it serving as an intermediate for integration.

It is interesting to compare the distribution of R1 and R2 in the silkworm with those in *D. melanogaster*, the only other species where both elements are known to exist. In *D. melanogaster*, R1 elements occupy 60% of the rDNA units on the X chromosome, while R2 elements occupy 15% of the rDNA units on both the X and Y chromosomes (5,40). Most if not all rRNA genes in *B. mori* reside on one autosome (41). R1 and R2 each occupy 7-15% of these units, however some races have lost all or almost all of the R1 elements. Many 5' truncated versions of R1 and R2 are present in the rDNA genes of *D. melanogaster* (5,40), indicating that these elements do not need a complete 5' end for sequence specific insertion. In the silkworm, no truncated copies have been found in the rRNA repeats, suggesting perhaps that the reverse transcription process is more efficient in the silkworm. Finally, a truncated R1 element, corresponding to the 0.8 kb 3' end of a complete element, has been characterized in *D. melanogaster* (42). This truncated element generated a 10 bp target site duplication. The left end of this duplication

contains sequence similarities to the 28S gene target site. However, most if not all of the remaining non-rDNA copies of R1 and R2 in *D. melanogaster* are located as tandem arrays in the pericentric heterochromatin (7,43,44). In *Drosophila* defective copies of all transposable elements are rapidly eliminated from the genome except those in the repetitive DNA of the chromacenter (45). Silkmoths, on the other hand, contain a dispersed pattern of repeated sequence with a high content of repeated sequences including an Alu-like retroposon (46). Thus the greater numbers of non-rDNA copies of R1 and R2 we have found in the silkmoth may be explained by the greater stability of such defective copies in the silkmoth genome.

In conclusion, we have previously suggested (13) that the ribosomal locus may be a favorable niche in the eukaryotic genome for the invasion of selfish mobile elements. R1 and R2 are not the only elements that appear to occupy this niche. In the nematode, *Ascaris lumbricoides*, a 4.5 kb element occupies approximately 5% of the rDNA units. The elements insert with accompanying target site duplications at a unique location midway between the R1 and R2 sites (47). Only truncated copies have been found outside the rDNA loci (48). In *D. melanogaster*, the G element has recently been shown to contain an ORF with homology to reverse transcriptase, and is a member of the non-LTR class of retrotransposons (35). This element appears to insert into a specific sequence in the rDNA spacer which contains remarkable sequence similarity to the region within the 28S gene between the R2 and R1 insertion sites (44). Finally, in several *Tetrahymena* species, a self-splicing 0.4 kb insert occupies all copies of the large rRNA genes (49,50). The location of the *Tetrahymena* insertions corresponds to a site only 3 bp from R2 insertion. The presence of these rDNA insertion elements in different phyla suggests that many additional species will be found to contain these or other similar elements.

### ACKNOWLEDGEMENTS

We thank B. Sakaguchi for supplying the geographical races of *B. mori*, M. Goldsmith, for strain C108 and K. Iatrou for BM5 cells. Thanks to N. Johnson for assistance in the analysis of clone B710. This work was supported by National Institutes of Health Grant R01 GM31867 and was aided by Grant No. IN-18 from the American Cancer Society.

### REFERENCES

1. Shapiro, J.A. (ed.) (1983) Mobile Genetic Elements. Academic Press, Inc., Orlando, Fla.
2. Doolittle, W.F. and Sapienza, C. (1980) *Nature* **284**, 601-604.
3. Orgel, L.E. and Crick, F.H.C. (1980) *Nature* **284**, 604-608.
4. Beckingham, K. (1982) In Busch, H. and Rothblum, L. (eds) *The Cell Nucleus*. Academic Press, New York, Vol. X, pp. 205-269.
5. Dawid, I.B., Wellauer, P.K. and Long, E.O. (1978) *J. Mol. Biol.* **126**, 749-768.
6. Dawid, I.B. and Rebert, M.L. (1981) *Nucleic Acids Res.* **9**, 5011-5020.
7. Roiha, H., Miller, J.R., Woods, L.C. and Glover, D.M. (1981) *Nature* **290**, 749-753.
8. Rae, P.M.M., Kohorn, B.D. and Wade, R.P. (1980) *Nucleic Acids Res.* **8**, 3491-3505.
9. Smith, V.L. and Beckingham, K. (1984) *Nucleic Acids Res.* **12**, 1707-1724.
10. Fujiwara, H., Ogura, T., Takada, N., Miyajima, N., Ishikawa, H. and Maekawa, H. (1984) *Nucleic Acids Res.* **12**, 6861-6869.
11. Eickbush, T.H. and Robins, B. (1985) *EMBO J.* **4**, 2281-2285.

12. Roiha, H., Read, C.A., Browne, M.J. and Glover, D.M. (1983) *EMBO J.* **2**, 721-726.
13. Burke, W.D., Calalang, C.C. and Eickbush, T.H. (1987) *Mol. Cell. Biol.* **7**, 2221-2230.
14. Michel, F. and Dujon, B. (1983) *EMBO J.* **2**, 33-38.
15. Xiong, Y. and Eickbush, T.H. (1988) *Mol. Biol. Evol.* **5**, (6) in press.
16. Xiong, Y. and Eickbush, T.H. (1988) *Mol. Cell. Biol.* **8**, 114-123.
17. Eickbush, T.H. and Kafatos, F.C. (1982) *Cell* **29**, 633-643.
18. Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) *Gene* **33**, 103-119.
19. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.
20. Iatrou, K., Tsitilou, S.C., Goldsmith, M.R. and Kafatos, F.C. (1980) *Cell* **20**, 659-669.
21. Goldsmith, M. R. and Clermont-Rattner, E. (1979) *Genetics* **92**, 1173-1185.
22. Grace, T.D.C. (1967) *Nature* **216**, 613-614.
23. Lecanidou, R., Eickbush, T.H. and Kafatos, F.C. (1984) *Nucleic Acids Res.* **11**, 4703-4713.
24. Fujiwara, H. and Ishikawa, H. (1986) *Nucleic Acids Res.* **14**, 6393-6401.
25. Long, E.O. and Dawid, I. B. (1979) *Cell* **18**, 1185-1196.
26. Kidd, S.J. and Glover, D.M. (1980) *Cell* **19**, 103-119.
27. Jamrich, M. and Miller, O.L. (1984) *EMBO J.* **3**, 1541-1545.
28. Dover, G. and Coen, E. (1981) *Nature* **290**, 731-732.
29. Roiha, H. and Glover, D.M. (1981) *Nucleic Acids Res.* **9**, 5521-5532.
30. Loeb, D.D., Padgett, R.W., Hardies, S.C., Shehee, W.R., Comer, M.B., Edgell, M.H. and Hutchison III, C.A. (1986) *Mol. Cell. Biol.* **6**, 168-182.
31. Fawcett, D.H., Lister, C.K., Kellet, E. and Finnegan, D.J. (1986) *Cell* **47**, 1007-1015.
32. Kimmel, B.E., ole-Moiyoi, O.K. and Young, J.R. (1987) *Mol. Cell. Biol.* **7**, 1465-1475.
33. Di Nocera, P.P. and Casari, G. (1987) *Proc. Natl. Acad. Sci. USA* **87**, 5843-5847.
34. Schwarz-Sommer, Z., Leclercq, L., Gobel, E. and Saedler, H. (1987) *EMBO J.* **6**, 3873-3880.
35. Di Nocera, P.P. (1988) *Nucleic Acids Res.* **16**, 4041-4052.
36. Ikenaga, H. and Saigo, K. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 4143-4147.
37. Inouye, S., Yuki, S. and Saigo, K. (1984) *Nature* **310**, 332-333.
38. Freund, R. and Meselson, M. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 4462-4464.
39. Xiong, Y. and Eickbush, T.H. (1988) *Cell* **55**, in press.
40. Wellauer, P.K., Dawid, I.B. and Tartof, K.D. (1978) *Cell* **14**, 269-278.
41. Maekawa, H., Takada, N., Mikitani, K., Ogura, T., Miyajima, N., Fujiwara, H., Kobayashi, M. and Ninaki, O. (1988) *Chromosoma* **96**, 263-269.
42. Browne, M.J., Read, C.A., Roiha, H. and Glover, D.M. (1984) *Nucleic Acids Res.* **12**, 9111-9122.
43. Peacock, W.J., Appels, R., Endow, S. and Glover, D. (1981) *Genet. Res. Camb.* **37**, 209-214.
44. Di Nocera, P.P., Graziani, F. and Lavorgna, G. (1986) *Nucleic Acids Res.* **14**, 675-691.
45. Finnegan, D.J. and Fawcett, D.H. (1986) In *Oxford Surveys on Eukaryotic Genes*, Oxford University Press, Oxford, Vol. 3, pp. 1-62.
46. Adams, D.S., Eickbush, T.H., Herrera, R.J. and Lizardi, P.M. (1986) *J. Mol. Biol.* **187**, 465-478.
47. Back, E., Van Meir, E., Muller, F., Schaller, D., Neuhaus, H., Aeby, P. and Tobler, H. (1984) *EMBO J.* **3**, 2523-2529.
48. Neuhaus, H., Muller, F., Etter, A. and Tobler, H. (1987) *Nucleic Acids Res.* **15**, 7689-7707.
49. Wild, M.A. and Sommer, M. (1980) *Nature* **283**, 693-694.
50. Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E. and Cech, T.R. (1982) *Cell* **31**, 147-157.