J Mol Evol (1999) 48:577–585

JOURNAL OF MOLECULAR EVOLUTION

© Springer-Verlag New York Inc. 1999

A Novel Tripartite Structure Comprising a *Mariner*-Like Element and Two Additional Retrotransposons Found in the *Bombyx mori* Genome

Yumiko Nakajima,[1,2] Kazuo Hashido,[2] Kozo Tsuchida,[2] Naoko Takada,[2] Teiichiro Shiino,[3] Hideaki Maekawa[2]

[1] Gene Research Center, University of the Ryukyus, Chihara 1, Nishihara-cho, Okinawa 903-0213, Japan
[2] Division of Radiological Protection and Biology, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-0052, Japan
[3] Laboratory of Epidemiology, AIDS Research Center and Committee in Research Information Laboratory, National Institute of Infectious Disease, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-0052, Japan

Received: 12 January 1998 / Accepted: 3 November 1998

Abstract. A clone of a DNA-mediated mobile element (transposon) corresponding to a *mariner*-like element (MLE) was obtained by carrying out the polymerase chain reaction with genomic DNA of *Bombyx mori* using a *Hyalophora cecropia* MLE sequence as a primer. This clone had a size of about 4.2 kb and, after sequencing, was found to contain an RNA-mediated, shorter retrotransposon named L1Bm, which was in turn integrated with a much longer retrotransposon named BMC1. Thus, the mobile elements made a novel tripartite structure. The BMC1 and L1Bm moieties of the composite structure each contained a 63-bp conserved sequence which was subsequently found to be highly conserved in all BMC1 and L1Bm elements registered so far. We propose that the 63-bp stretch may be a recognition site for a retrotransposition mechanism conducted by a reverse transcriptase and an endonuclease complex. On the basis of this inference, we propose a model that predicts how different types of BMC1 and L1Bm elements are dispersed in the genome. In addition, a phylogenetic tree made from the current and extant BMC1 and L1Bm sequences indicated that these elements can be classified into Subfamilies I and II.

Key words: Silkworm — *Bombyx mori* — Transposon — *Mariner*-like element — Retrotransposon — BMC1 — L1Bm — LINE — Retrotransposition — SINE

Correspondence to: H. Maekawa; *e-mail:* hidemae@nih.go.jp

Introduction

Retrotransposons are integrated via RNA intermediates into genomes by the function of a reverse transcriptase (RT) and an endonuclease (EN) complex (RT-EN). This mechanism involves specific recognition of the 3′ end of the RNA by RT-EN (Luan et al. 1993; Luan and Eickbush 1995; Zimmerly et al. 1995a, b). The element, named L1, which belongs to the long interspersed elements (LINEs), is a retrotransposon of the class without long terminal repeats (LTR) and is located at variable sites of mammalian genomes (Hutchinson et al. 1989; Moran et al. 1996). The R1 and R2 elements are also non-LTR-type retrotransposons but are integrated at specific sites in the genomes of insects including *Bombyx mori* (Luan et al. 1993; Luan and Eickbush 1995). The group called short interspersed elements (SINEs) are also integrated into genomes by an RT-EN mechanism similar to that known in LINEs, since both groups have a common sequence at the 3′ end (Ohshima et al. 1996; Okada and Hamada 1997; Okada et al. 1997). Elements of SINEs and LINEs constitute a major part of many genomes and hence analyses of their integration processes as well as their relative locations are of impor-

tance for the understanding of the structure and evolution of chromosomes.

A sequence called BMC1 (full-length unit of 5.1 kb) has been detected in highly abundant copy numbers of about 3500 per haploid genome of *B. mori* (Ogura et al. 1994). Having an open reading frame (ORF) encoding RT, this element can be classified into the non-LTR class of retrotransposons like R1, R2, Jockey, and mouse L1 (Maekawa unpublished; Priimagi et al. 1988; Xiong and Eickbush 1990; Burke et al. 1993). The integration sites and flanking regions of different BMC1 elements that have been found to date are highly variable in different geographical strains of *B. mori.* If a conserved copy of BMC1 could be found in most strains of *B. mori*, it could be considered an ancient form of this retrotransposon. We report here that we found such a BMC1 element in the *B. mori* genome during the study of a DNA-mediated mobile element (transposon) named *mariner* and its relatives, called *mariner*-like elements (MLEs). The latter are known to be widespread from nematodes to human and have been characterized in several lepidopteran species, including *Hyalophora cecropia* (Lidholm et al. 1991; Robertson 1993), *B. mori* (Tomita et al. 1997), and *Attacus atlas* (Nakajima et al. 1998). The present BMC1 element was found to be integrated into an another retrotransposon, L1Bm (previously called Bm2 as a SINE-like retroposon), which was recently characterized by Ichimura et al. (1997), and the L1Bm was inserted into an MLE at the first, thus making a unit with a novel, tripartite structure. The BMC1 and L1Bm regions each contained a 63-bp conserved sequence, which we subsequently confirmed in all previously registered BMC1 and L1Bm elements. We discuss the conserved features of these elements and suggest a potential mechanism by which different types of BMC1 and L1Bm elements arise in the genome.

Materials and Methods

Silkworm Strains

The *B. mori* strain N4 was supplied by the Faculty of Agriculture, Nagoya University, Nagoya, Japan, and reared at the Division of Radiological Protection and Biology, National Institute of Infectious Diseases, Tokyo, Japan. Gun-Po × Shu-Gyoku was supplied by Kanebo Silk Elegance Co. Ltd., Japan; p22 and p50, by the Institute of Genetic Resources, Kyushu University, Fukuoka, Japan; and Seihaku, Oolong, Aojuku, and w30, by the National Institute of Sericultural and Entomological Science, Tsukuba, Japan. All larvae were reared on mulberry leaves except for N4 and Gun-Po × Shu-Gyoku, which were reared on an artificial diet.

Amplification of an MLE Sequence by PCR and Cloning of a Product

High molecular weight, genomic DNA was isolated from posterior or middle larval silk glands (10 males and 10 females) on days 3 to 4 of the fifth instar (Maekawa et al. 1988) and subjected to PCR with *EX Taq* DNA polymerase (Takara) using a commercial buffer developed for long-range reactions. The primer used (5′-TTAATATTAGGTCCTTACATAT GAAATTAGC GTTTTGT-3′) was designed from the inverted repeat (IR) of the *H. cecropia* MLE (Lidholm et al. 1991; Robertson 1993). After denaturation for 2 min at 95°C, amplification was conducted in 30 rounds, each for 1 min at 90°C, 20 s at 98°C, and 10 min at 68°C, followed by a final round of extension for 5 min at 75°C. The products were separated by electrophoresis on a 1% agarose gel. The fragment in the 4.2-kb band was extracted from the gel using a Qiagen kit, ligated with the vector pBluescript-II-SK(+) after digestion with *Sma*1, and used as a probe for Southern blot hybridization or subjected to DNA sequencing.

Genomic Southern Blot Analysis

Genomic DNA was incubated extensively with the restriction enzymes indicated and the digests were separated by agarose gel electrophoresis as described above. The gels were blotted onto Hybond-N nylon membranes (Amersham) and hybridized with a ^{32}P-labeled probe at 65°C overnight (Sambrook et al. 1989). The filters were washed twice in 0.1× SSC/0.1% SDS at room temperature, twice in 2× SSC/0.1% SDS at 65°C, and, finally, twice in 0.1× SSC/0.1% SDS at 65°C, dried, and visualized by a BAS 1000 Bio Image analyzer system (Fuji Film Co. Ltd.) or by autoradiography.

Preparation of Deletion Mutants and Sequencing of the 4.2-kb Fragment

Deletion mutants of the 4.2-kb fragment were prepared using a Kilo-sequence deletion mutant kit (Takara) and transformed into high-efficiency competent cells of *E. coli* XL-1 Blue (Takara). Appropriate clones were chosen on the basis of size, blotted onto Hybond-N nylon membranes, and probed with a BMC1 sequence [a clone, pBmT 6435 (Ogura et al. 1994)] for preliminary mapping, after which the clones were labeled with [α-^{32}P]dCTP using a random oligonucleotide priming kit (Pharmacia), and sequenced by the dideoxy method (Sanger et al. 1977).

Sequence Comparison and the Construction of a Phylogenetic Tree

Sequence homology was analyzed using a BLAST database search system. Multiple alignments were made with the GCG pileup program (Wisconsin Package, Version 9.0, Genome Computer Group, Inc., Madison, WI). A phylogenetic tree was constructed by PHYLO_WIN, a "graphic interface" for molecular phylogeny written by Nicholas Galtier of the University of Lyon (galtier@biomserv.univ-lyon1.fr); up to 400 bp of sequences including the 63-bp conserved regions (see below) in the registered sequences was utilized, although sequences shorter than 63 bp were used without trimming.

Results

Isolation of an MLE Clone Containing BMC1

Amplification of the genomic DNA from eight *B. mori* strains was carried out by PCR and the products were separated by electrophoresis (Fig. 1A). The IR primer designed from the *H. cecropia* MLE sequence worked well [a somewhat similar IR primer has also been reported by Tomita et al. (1997)]. A band with a size of 1.3

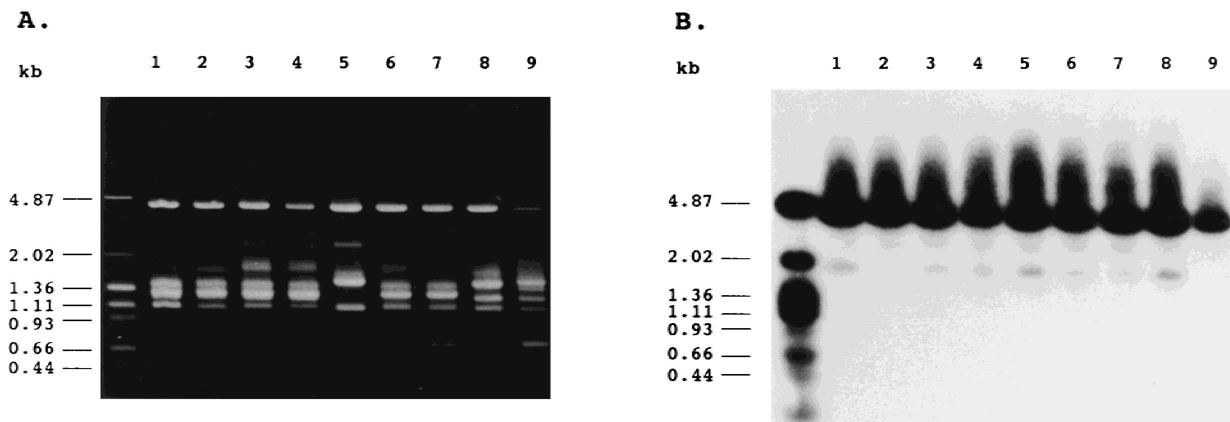


Fig. 1. PCR amplification of MLE sequences using *B. mori* genomic DNAs. The inverted repeat (IR) sequence of the *H. cecropia* MLE was used as a primer under long-range PCR conditions. DNA was isolated from the posterior silk glands unless otherwise indicated. The strains of *B. mori* used are N4 (Lane 1), Seihaku (Lane 2), Oolong (Lane 3), Aojuku (Lane 4), p50 (Lane 5), Gun-Po × Shu-Gyoku (Lane 6 = middle silk gland DNA), Gun-Po × Shu-Gyoku (Lane 7), p22 (Lane 8), and w30 (Lane 9). Leftmost lane, pHY markers (Takara), whose sizes are drawn along the margin. All strains contained 1.3- and 4.2-kb bands, except p50 (lane 5), in which the 1.3-kb band probably shifted to 1.4 kb inserted by other sequence. No complete ORFs of *mariner*-like sequences were found in *B. mori* and hence the presence of a 1.3-kb fragment does not necessarily mean that it is intact. Therefore, although the 1.3-kb band was confirmed to contain an MLE (details omitted), the 4.2-kb band was used for analysis in subsequent experiments. **A** Ethidium bromide stain. **B** Southern blot hybridization with a BMC1 sequence [pBmT6435 (Ogura et al. 1988)] as a probe (see text).

kb which contained a predicted complete MLE was observed in most strains, together with several bands with presumed insertions or deletions. It is noteworthy that all strains also exhibited a 4.2-kb band. This large size suggested that there was a large insertion into an MLE or a highly repetitious sequence with homology to the MLE primers. We carried out the subsequent analyses only with this band, which was extracted from the gel, radiolabeled, and used for Southern blot hybridization to *B. mori* genomic DNA after digestion with different restriction enzymes (Fig. 2). The hybridization pattern was reminiscent of a previous one probed with the retrotransposon BMC1 (Ogura et al. 1994), suggesting that the 4.2-kb band contained a BMC1 element. This inference was highly likely to be the case since a marked signal was obtained at the 4.2-kb band when the gel shown in Fig. 1A was hybridized with a BMC1 fragment (Fig. 1B; note that the faint 4.2-kb band in Lane 9 was also hybridizable). Hence the 4.2-kb fragment was subjected to cloning. Eventually four clones were obtained, all giving the same restriction map (not shown). One of them, named pBmTNML1, was used in the following experiments. Repeated trials including PCR amplification with MLE and L1Bm (see below) sequences as primers, as well as Southern blot hybridization using the same sequences as probes, suggested that the unit corresponding to pBmTNML1 might be present in a single copy per haploid genome (unpublished observations).

The Sequence of pBmTNML1 Reveals a Tripartite Structure

The complete base sequence of pBmTNML1 (the 4.2-kb fragment) was determined and subjected to a homology search using available DNA databases; results are illustrated in Figs. 3A and B. As predicted, this clone contained both an MLE and a BMC1 sequence. Surprisingly, another unit was also detected, corresponding to a retrotransposon, Bm2, recently redesignated L1Bm (Ichimura et al. 1997). Thus the overall structure made a tripartite combination: An L1Bm element is inserted into an MLE, and then a BMC1 element is integrated into the L1Bm. Both of the BMC1 and L1Bm regions had putative target site duplications (see the hatched boxes in Fig. 3A, labeled as target sequences A and B for L1Bm and BMC1, respectively; see also the double underlines in Fig. 3B). Interestingly, a common stretch about 63 bp in length could be discriminated within both the L1Bm and the BMC1 regions (see the horizontally striped box labeled as the conserved sequence in Fig. 3A and the outlined letters in Fig. 3B); the one in the L1Bm region was interrupted by the BMC1 insertion. This 63-bp homologous sequence is located approximately 150 bp upstream of the 3′ poly(A) tail of each element. A similar 63-bp region was present in the previously reported complete sequence of an L1Bm element (Ichimura et al. 1997). Thus we searched for this region by BLAST analysis in all BMC1 and L1Bm (Bm2) elements registered to date. An alignment of the sequences found (boxed in Fig. 4) indicated that this homologous region was conserved at a 3′ subterminal location in all of the copies reported (note that the BMC1 and L1Bm in the tripartite structure are abbreviated BMC1-ml and Bm2-ml, respectively, in Figs. 4, 6, and 7).

Relationship Between BMC1 and L1Bm

The 63-bp region was highly conserved, showing approximately 90% sequence identity among the 28 ele-

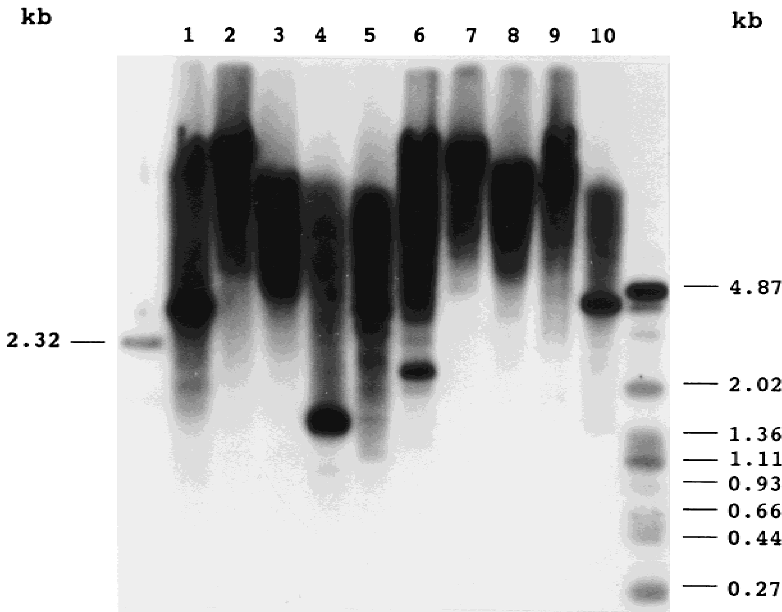


Fig. 2. Southern blot hybridization to genomic DNA from *B. mori* (N4) using the 4.2-kb sequence (clone pBmTNML1) as a probe. DNA was isolated from the posterior silk glands. The digests with restriction enzymes *Ava*1 (Lane 1), *Bam*H1 (Lane 2), *Eco*R1 (Lane 3), *Eco*R5 (Lane 4), *Hin*d3 (Lane 5), *Hpa*1 (Lane 6), *Not*1 (Lane 7), *Pst*1 (Lane 8), *Pvu*1 (Lane 9), and *Sal*1 (Lane 10) were electrophoresed through a 1% agarose gel. Rightmost lane, size markers (see Fig. 1); leftmost lane, *Hin*d3 digest of λ phage DNA as another size marker.

ments compared; further, no specific difference was found between BMC1 and L1Bm elements (Fig. 4). For example, the BMC1 element in the tripartite structure was identical to Bm2-8 (arbitrarily selected from registered L1Bm sequences) except for four bases.

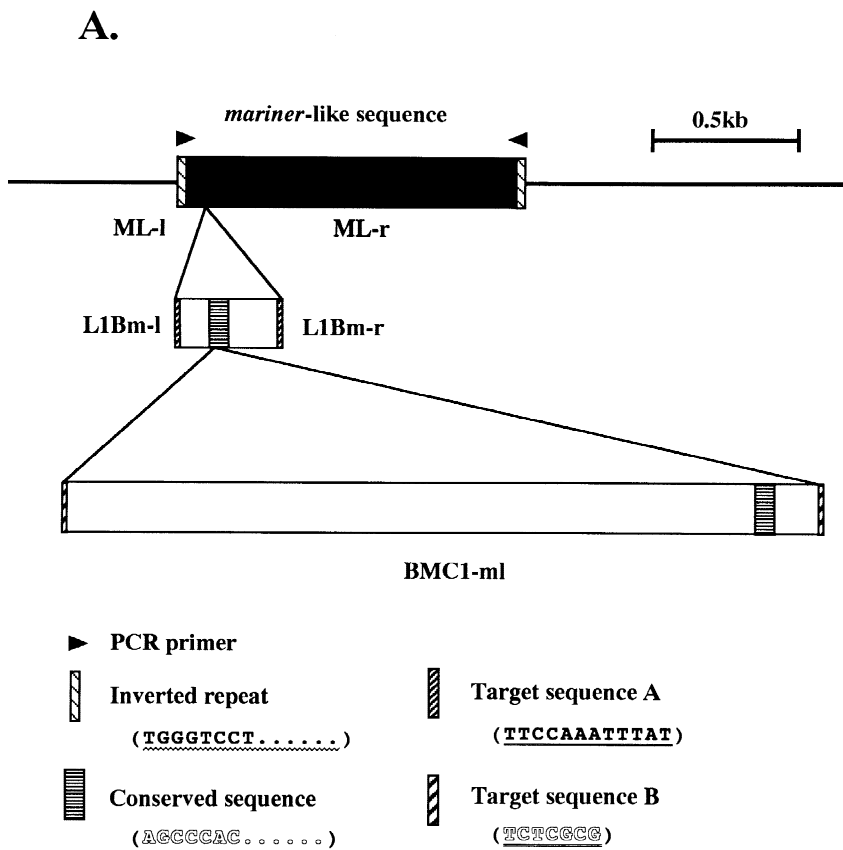
We analyzed all available sequences by PHYLO_WIN in order to construct a phylogenetic tree (Fig. 5). The analysis showed that the elements could be classified into two major groups, Subfamilies I and II. The classification was based on the sequence differences in the upstream and downstream regions flanking the conserved 63-bp stretch. Figure 6 illustrates the degree of similarity for the conserved and flanking regions both within and between the two subfamilies, using representative elements of each.

In order to clarify the mechanism of transposition for these elements, we also assessed, by BLAST search, whether or not the putative target site duplications contained in the upstream and downstream boundary regions in different BMC1 and L1Bm elements were identical. The sequence alignment shown in Fig. 7 indicated that the duplications in question (underlined) were variable from element to element. Further, no duplications similar to those found here in the BMC1 and L1Bm moieties of the tripartite unit occurred in other elements (for target sequences in the tripartite structure, see also Figs. 3A and B).

Discussion

The tripartite structure of the *B. mori* 4.2-kb genomic element, analyzed using the clone pBmTNML1, was unique in that a DNA-mediated transposon (MLE) was integrated with an RNA-mediated retrotransposon (L1Bm) and this L1Bm element itself was further integrated with another retrotransposon (BMC1). The presence of the well-conserved 63-bp stretch, which was detected in each of the BMC1 and L1Bm domains and was confirmed to be present in all previously registered BMC1 and L1Bm elements, suggested that it could serve as a recognition signal for the site-specific insertion events conducted by RT-EN. A preliminary examination of the possible secondary structure of an RNA molecule encoded by this element using the program DNASIS indicated that the conserved 63-bp sequences are able to assume a very common stem-loop structure, which may extrude a tertiary structure recognizable by the enzyme complex (unpublished observations). Previous results using the target DNA-primed RT assay of a LINE-like insertion found in *B. mori* [R2Bm (Luan et al. 1993; Luan and Eickbush 1995)] have also indicated the importance of the 3′ untranslated region of the RNA product as a recognition site. The inference that the highly conserved sequence of the 63-bp region is critical as the starting point of insertion is further supported by the present finding that the sequences of putative target site duplications residing in the flanking regions are variable in different BMC1 and R1Bm elements; probably the RT-EN complexes are less specific with regard to the target site duplications than to the conserved sequence.

The fact that the same 4.2-kb unit was detected in all strains of *B. mori* analyzed by PCR followed by Southern blot hybridization with a BMC1 probe, as well as in all but one strain reported by other investigators (Tomita et al. 1997), implies that the structure with BMC1 arose earlier than the time when geographic strains of *B. mori* were separated, which is believed to have occurred some 5000 years ago. Nevertheless, we were unable to detect the 4.2-kb unit in *B. mandarina* (Japanese type), a wild



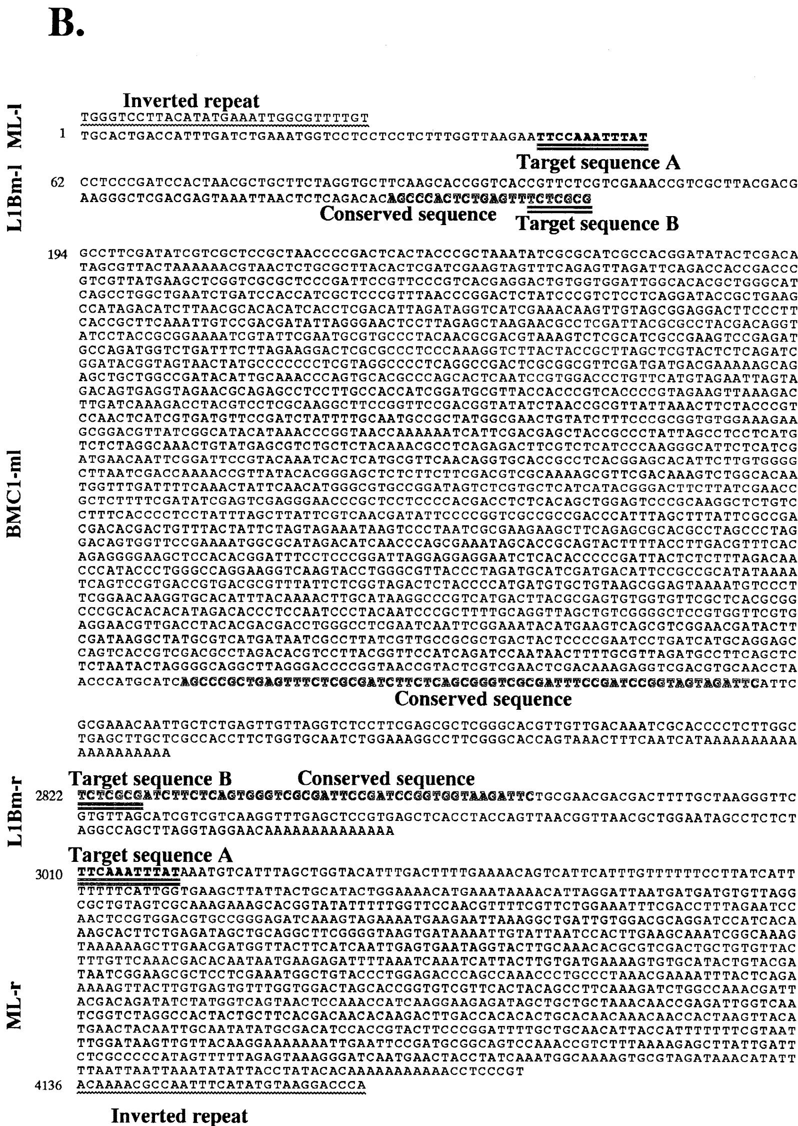


Fig. 3. The primary structure of the 4.2-kb unit clone (pBmTNML1). **A** The schematically illustrated structure. An MLE is inserted with L1Bm, which in turn contains BMC1 (l and r represent left and right arms, respectively). BMC1-ml denotes a BMC1 integrated into an MLE. The average 63-bp conserved sequence residing in both the L1Bm and the BMC1 moieties is represented by *horizontally hatched boxes.* Putative target duplications for L1Bm and BMC1 (tentatively named A and B, respectively) are shown by *boxes with oblique lines.* **B** The complete sequence. The *outlined letters* show the two copies of the 63-bp conserved sequence. *Double underlined regions* stand for the putative target sequences A and B. The *letters with wavy underlines* are an IR pair derived from the PCR primer (the relevant regions are also seen in A).

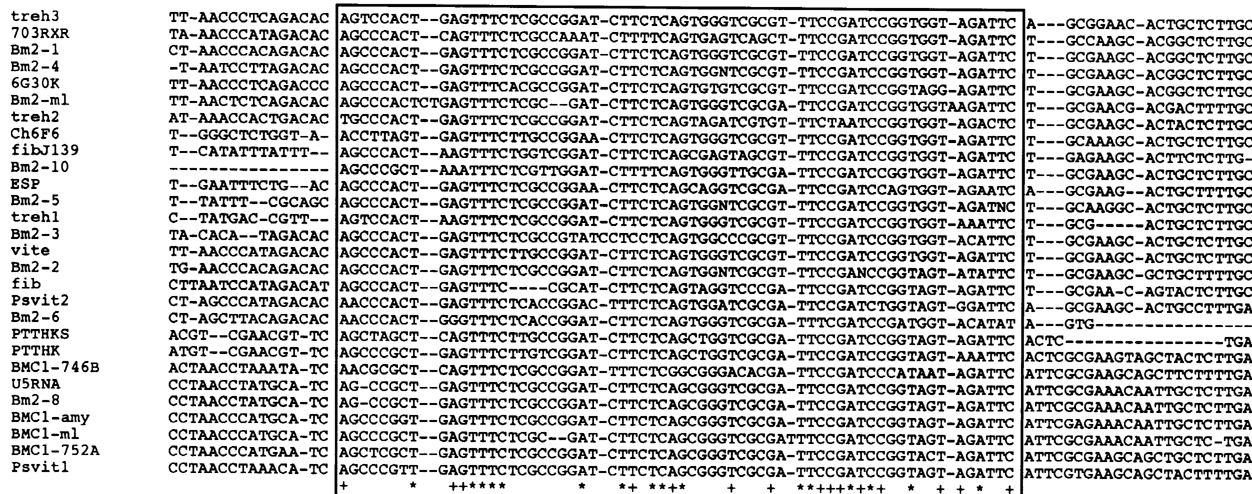


Fig. 4. Multialignment of the presumed recognition sequence in previously registered BMC1 and L1Bm (Bm2) families in comparison with the present ones (abbreviated as BMC1-ml and Bm2-ml) from the clone pBmTNML1. All elements exhibited the conserved sequences as indicated in the *box*. An *asterisk* indicates that the residue is the same in all members, whereas a *plus* means that it is different in only one of the members. The accession numbers are treh3 (D86212), 703RXR (006073), Bm2-1 (X70928), Bm2-4 (X90931), 6G30K (X54735), Bm2-ml (AB008779), treh2 (D86212), Ch6F6 (X66164), fibJ139 (Z26887), Bm2-10 (X70937), ESP (D12521), Bm2-5 (X70932), treh1 (D86212), Bm2-3 (X70930), vite (D30732), Bm2-2 (X70924), fib (M76430), Psvit2 (D16233), Bm2-6 (X70933), PTTHKS (X75942), PTTHK (X75941), BMC1-746B (D26008), U5RNA (S42609), Bm2-8 (X70935), BMC1-amy (U07847), BMC1-ml (AB008780), BMC1-752A (D26009), and Psvit1 (D16233). These include BMC1 and Bm2 elements registered as such and those picked up spontaneously from the sequences registered as other genes, e.g., the PTTH-K gene. The BMC1 sequence found in the amylase gene intron and tentatively named BMC1-amy in the present study corresponds to L1BmcAmy of Ichimura et al. (1997).

silkworm very close to *B. mori* (Murakami and Imai 1974). We therefore conclude that it is likely that BMC1 integrated into an MLE later than the time when these species were separated, which is believed to have occurred some 3 million years ago (Maekawa et al. 1988), although the alternative possibility that this unit was lost from *B. mandarina* during this long period of time cannot be excluded.

The similarity of BMC1 elements and L1Bm (Bm2) elements has been reported previously (Ogura et al. 1994; Ichimura et al. 1997), with emphasis on their shared common 3′ sequences. However, the more extensive sequence comparisons made here, in which we used 400 bp encompassing the 63-bp conserved region in the expectation that this would include most regions of L1Bm elements and the 3′ half of BMC1 elements as far as possible, revealed that the elements can be divided into Subgroups I and II. The elements named BMC1 belong to the former and the elements named Bm2 (L1Bm) belong to the latter subgroups. This nomenclature may be applicable to the differently named members, e.g., PTTKHS etc.

Non-LTR retrotransposons frequently suffer from 5′ truncations, probably occurring during the insertion into multiple locations in genomes, and it is difficult to find a complete copy. In the current tripartite unit, the L1Bm moiety is probably inactive in retrotransposition since it is highly 5′ truncated, and the much longer BMC1 moiety may also be nonfunctional since its ORF is incomplete. However, a conserved BMC1 element of 3.7 kb with a complete ORF for a putatively active RT-EN (Hinkel, unpublished data) has been found in an intron of the *B. mori* amylase gene, although this element seems to lack a promoter region. This observation, plus the discovery of the tripartite element and other observations reported here, suggests how many 5′ truncated L1Bm elements may have been dispersed in the *B. mori* genome (Ichimura et al. 1997) and how the presently found composite structure was constructed. Thus, we imagine that there is a prototype BMC1 element which is equipped with a promoter and a complete ORF and propose that a prototype BMC1 may coexist, with its defective descendants including highly 5′ truncated ones, i.e., L1Bm elements, as illustrated schematically in Fig. 8. This model includes a prototype BMC1 which is inserted next to a promoter site. The positioning of the promoter (and of a *gag*-like domain) upstream of the EN and RT domains permits transcription and translation of an enzyme complex, which is able to recognize its own RNA using the 63-bp recognition sequence near the 3′ end. The recognition may be assisted by different types of target site duplications, e.g., target sites A, B, and C drawn in the model. As shown at target site A, a copy without a promoter but with a complete ORF (thus a 5′ truncated active BMC1) could be produced. This is the type of element found in the amylase intron of *B. mori* (Hinkel unpublished data). We infer that this sequence was integrated at a rather late period of evolution and thus its ORF has remained intact. As shown at the target site B, even when an element which is more or less defective

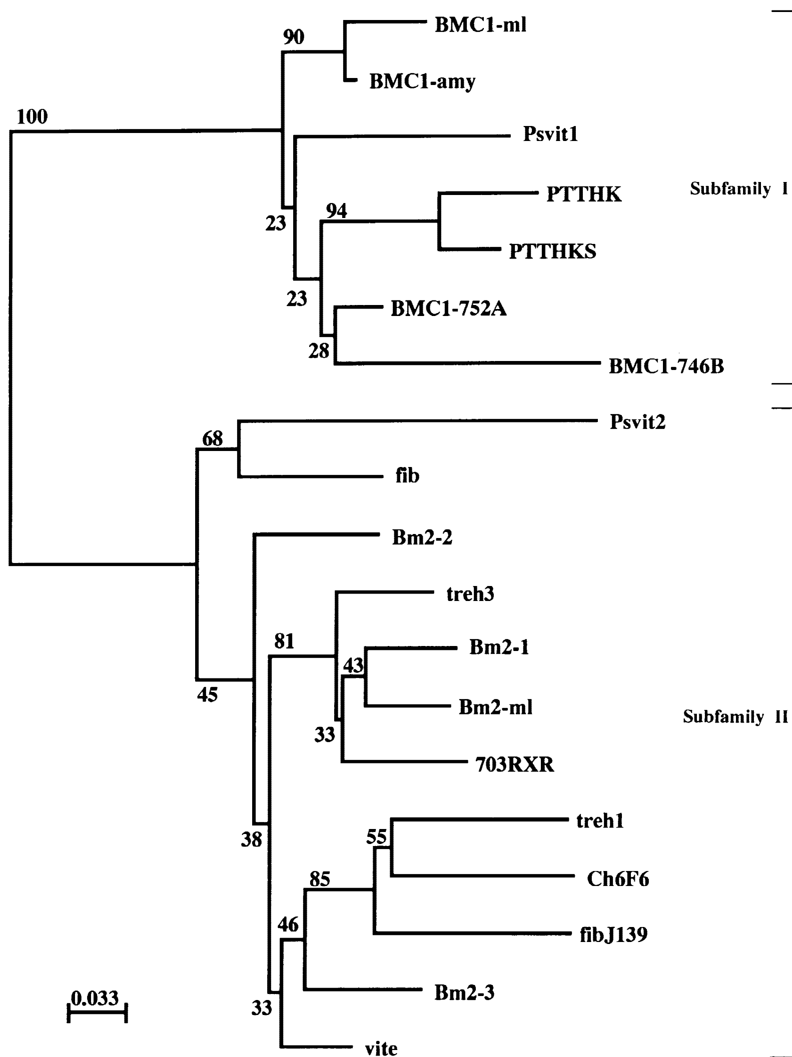


Fig. 5. Phylogenetic tree for the BMC1 and L1Bm elements. The present ones are abbreviated BMC1-ml and Bm2-ml, respectively. The pileup program for GCG was applied for up to 400 bp of each sequence. All are classified into the major two Subfamilies I and II. For accession numbers see the legend to Fig. 4.

	upstream region	conserved region	downstream region
Subfamily I	85.0%	92.0%	86.2%
	59%	88%	52%
Subfamily II	67.7%	89.6%	79.6%

Fig. 6. Sequence homology among BMC1 and L1Bm elements in the upstream, conserved, and downstream regions analyzed separately. For Subfamilies I and II, see Fig. 5. *Numerals in the boxes:* Comparisons within the same subfamily. Six arbitrary members of Subfamily I were compared with BMC1-amy [which contains an intact ORF (Hinkel, unpublished data)] as a representative. Similarly, seven arbitrary members of Subfamily II were compared with Bm2-2 chosen arbitrarily as a representative. *Numerals between the boxes:* Comparison between Subfamilies I and II made between BMC1-amy and Bm2-2 as representatives, respectively. All results are expressed as average percentages.

occurs, it can propagate if it still contains a recognition sequence and enzyme domains located downstream of a promoter. On the other hand, extensive 5′ truncation will produce a series of L1Bm elements (target site C) which exist in multiple, differing copies and which are by no means active without promoters. However, we suppose that an L1Bm unit may also acquire retropositional activity by itself if it resides next to a promoter, based on analogy to the tRNA-derived SINEs, wherein a sequence homologous to tRNA serves as a primer like Bm1 (Eickbush 1995). Okada et al. proposed that a strong-stop RNA transcript from a retrovirus might be integrated into the upstream region of a LINE sequence. However, it is also possible that a LINE could be integrated into the downstream region of a tRNA homologous sequence derived from an integrated strong-stop RNA (Ohshima et al. 1993, 1996; Okada and Hamada 1997; Okada et al. 1997). RNA polymerase III products could then be recognized by an endogenous reverse transcriptase and hence give rise to a SINE.

Further study is required to substantiate this model and to understand more fully the mode of evolution of *B. mori* and *B. mandarina* in terms of chromosome structures. Detailed analyses of the recognition mechanisms including the binding sites of a BMC1 unit containing a complete ORF will be useful and a study along these lines is in progress in our laboratories.

```
Bm2-1        CCTAATATCACATTTTTATTAACACCGTCA        AAAAACATTTTATTCTTTTTTTTG
Bm2-2        AAGAACCCACTTGTCGTGGCCGCTG             AAAAATTGTCTATGATTGCC
Bm2-3        CCGTATAGGACTCCCGACC                   TATCAGTATAGGTAGGAAAAAT
Bm2-4        GATATCCTTACAATTCTCATAATCCTTAG         AAAAACAATTCTCAATTTGGTTTA
Bm2-5        AGTGGCATATCTTGGAAACTTTACCT            CGCAATTGGATTCTATTATT
Bm2-6        AACTGGTCTTGTAGATTCCTCCGGTGC           TAGTGGTAGTTAGACTTTCAA
Bm2-8        TCGACGTGCAACCTAACCTATGCATC            AAAAAAACCTAATGTGGGTAAT
Bm2-9-1      TCANATCTNAGGCTCTNGCTAGTGCCA           AAAGAGGCTATCCTAACCACAC
Bm2-9-2      AAGAATAGTTAGGGGAAAAAAGAGGC            GTTTNACGGGGACGGTGCGTAT
Bm2-10       TTCACTATGATATTGGTAAAAGCCCGCTAA        AAAAATTATGGTAAAAGGCAATATA
Bm2-7        AGCCCATAGACATGTAAAACGGCCCCCTCCT       AAAAACATGTAAAACGTAAATGC
Ch6F6        GGCTCTGGTAACCTTAGTGAGTTT              AAAAAAACCTGTAACTATTG
BMC1-amy     ACGGCAGAAATAGGTAAGGTCACGCCACGG        AAAAATAGGTAAGGTGGTGGTACCT
Bm2-ml       TGGTTAAGAATTCCAAATTTATCCTCCCGATC      AAAAATTCAAATTTATAAATGTCATTT
BMC1-ml      ACTCTGAGTTTCTCGCGGCCTTCGATA           AAAAATCTCGCGATCTTCTCAG
gly          GGGTGCTTTTAGATACCTCAAGAACC            AAAAAGTATACTATATA
PTTHKS       CTTGGAAACGGATAGCATAACTGGG             AAAAAGATAGTAGCTCCCAT
PTTHK        CTTGGAAACGGATAGCATAACTGGG             AAAAAGATAGTACTCTCC
ErA4         TTCCGATCCGATAGTAGATTCAT               AAAAAGATACATATATT
fib          CGTCATCCTCAAAGCCAATTTCAATTT           AAAAGAAAGCCATGTGAGTTTA
Psvit2       ATCCCCAACGTAAATGCGCCACCCACC           AACAATAAATGCCTCTCATACCT
703RXR       GATCCACTAACGGTGCTTTTAGGTA             AAAAGGGTGAAATCCAAAA
treh2        TTACCATCAGCTGGACCGTACGCCCG            AAAAAGGACAAAGTAATAA
treh3        GAATGGAAAGTAGTATATG                   AAATAGAAAT
U5RNA        TCGACGTGCAACCTAACCTATGCATC            AAAAAAACCTAATGTGGGTAAT
Psvit1       GTGATCTGAGGAGCCCATGTGCTAGA            ATAAAGACCGCCTCTAT
ErB2         CTGCTCATCATGCGATATCATCGACGCCTTC       AAAAATGCGATATCATAGCTCATTGT
BMC1-752A    CAAACCATACCCTGGGCCCCAATTCCGA          AAAAAACCCTGGGCCAGGAAGGTCA
BMC1-746A    ATAGTACATTCTCTTTCCCATTCCGAT           AAAAACTCTTTCGAATCATTTT
BMC1-746B    ATCAACAGTTCACGTTTCACACGTCGT           AAAAACACATTTTTCCAAAGCA
```

Fig. 7. Multialignment of the sequences in the flanking regions including putative target duplications for the insertion of BMC1 and L1Bm elements (the present ones are abbreviated BMC1-ml and Bm2-ml, respectively). The duplicated putative target sequences are *underlined.* The accession numbers are gly (L08106), ErA4 (X58447), ErB2 (S58449), and BMC1-746A (D26007); for others see the legend to Fig. 4.

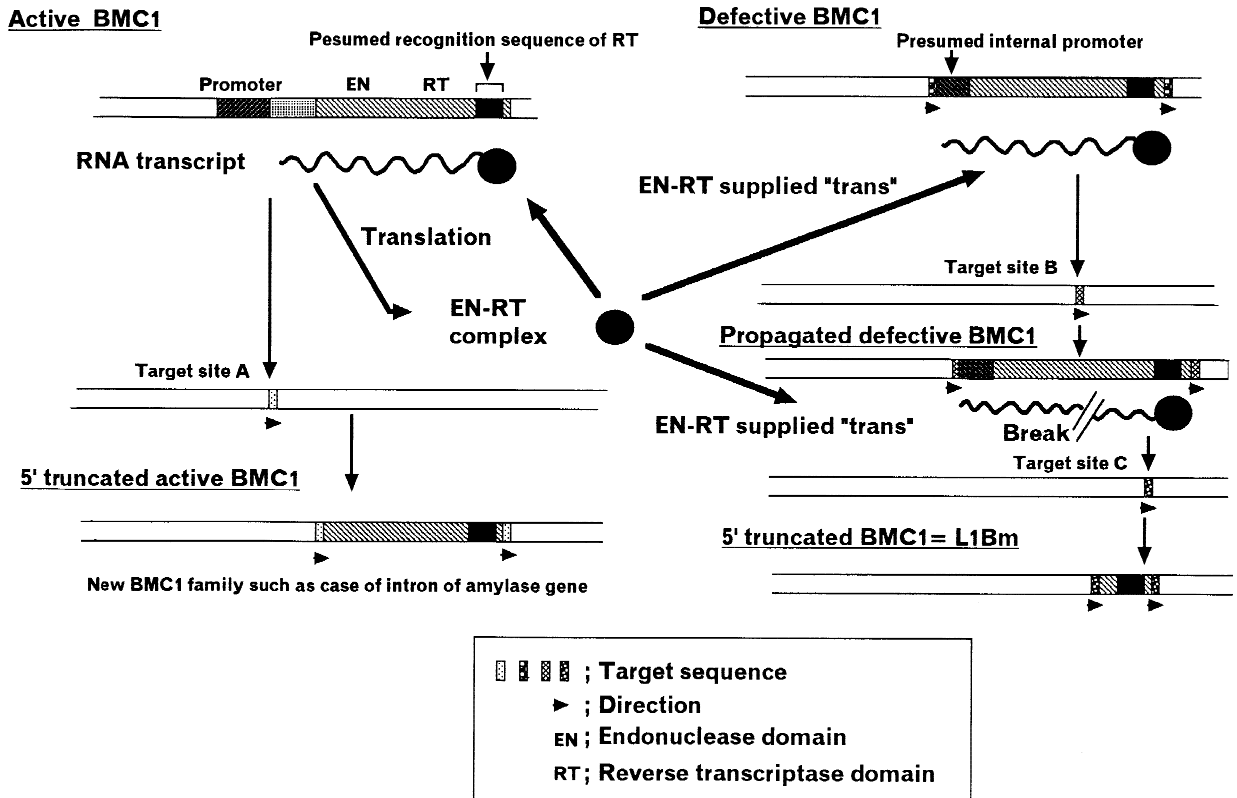


Fig. 8. A model for evolution of various types of BMC1 and L1Bm elements. The presence of a prototype, active BMC1 is assumed. The 63-bp conserved sequence is assigned to the presumed recognition sequence for the enzyme complex EN-RT. If the complex is supplied in "trans" (*thick arrows*), it makes cDNA which is inserted at a target site (*thin arrows*). *Arrowheads* indicate the orientation of target sites.

Acknowledgments. We sincerely thank Dr. T. Eickbush, Dr. M.R. Goldsmith, and Dr. K. Koga for their critical reading and thoughtful advice. This work was partly supported by grants from the Ministry of Education, Science and Culture of Japan and Science and Technology Agency of Japan.

References

Burke WD, Eickbush DG, Xiong Y, Jakubczak JL, Eickbush TH (1993) Sequence relationship of retrotransposable elements R1 and R2 within and between divergent insect species. Mol Biol Evol 10:163–185

Eickbush T (1995) Mobile elements of lepidopteran genomes. In: Goldsmith MR, Wilkins AS (eds) Molecular model systems in the lepidoptera. Cambridge University Press, New York, pp 77–105

Hutchinson CA 3d, Hardies SSC, Loeb DD, Shehee WR, Edgell MH (1989) LINEs and related retroposons: long interspersed repeated sequences in the eukaryotic genome. In: Berg D, Howe MM (eds) Mobile DNA. Am Soc Microbiol, Washington, DC, pp 593–617

Ichimura S, Mita K, Sugaya K (1997) A major non-LTR retrotransposon of *Bombyx mori,* L1Bm. J Mol Evol 45:253–264

Lidholm DA, Gudmundsson GH, Boman HG (1991) A highly repetitive, *mariner*-like element in the genome of Hyalophora cecropia. J Biol Chem 266:11518–11521

Luan DD, Eickbush TH (1995) RNA template requirements for target DNA-primed reverse transcription by the R2 retrotransposable element. Mol Cell Biol 15:3882–3891

Luan DD, Korman MH, Jakubczak JL, Eickbush TH (1993) Reverse transcription of R2Bm RNA is primed by nick at the chromosomal target site: A mechanism for non-LTR retroposition. Cell 72:595–605

Maekawa H, Takada N, Mikitani K, et al. (1988) Nucleotide organizers in the wild silkworm *Bombyx mandarina* and the domesticated silkworm *B. mori.* Chromosoma 96:263–269

Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD, Kazazian HH Jr (1996) High frequency retrotransposition in cultured mammalian cells. Cell 87:917–927

Murakami A, Imai H (1974) Cytological evidence for holocentric chromosomes of the silkworms, *Bombyx mori* and *B. mandarina* (Bombycidae, Lepidoptera). Chromosoma 47:167–178

Nakajima Y, Hashido K, Shiino T, Tsuchida K, Nagamine M, Maekawa H (1998) Isolation of a *mariner*-like sequence containing a complete open reading frame for transposase from *Attacus atlas* and its phylogenetic relationships within the Ditrysia of Lepidoptera. J Seric Sci Jpn 67:271–278

Ogura T, Okano K, Tsuchida K, et al. (1994) A defective non-LTR retrotransposon is dispersed throughout the genome of the silkworm, *Bombyx mori.* Chromosoma 103:311–323

Ohshima K, Koishi R, Matsuo M, Okada N (1993) Several short interspersed repetitive elements (SINEs) in distant species may have originated from a common ancestral retrovirus: Characterization of a squid SINE and a possible mechanism for generation of tRNA-derived retroposons. Proc Natl Acad Sci USA 90:6260–6264

Ohshima K, Hamada M, Terai Y, Okada N (1996) The 3′ ends of tRNA-derived short interspersed repetitive elements are derived from the 3′ ends of long interspersed repetitive elements. Mol Cell Biol 16:3756–3764

Okada N, Hamada M (1997) The 3′ ends of tRNA-derived SINEs oriented from the 3′ ends of LINEs: A new example from the bovine genome. J Mol Evol 44:S52–S56

Okada N, Hamada M, Ogiwara I, Ohshima K (1997) SINEs and LINEs share common 3′ sequences: A review. Gene 205:229–243

Priimagi AF, Mizrokhi LJ, Ilyin YV (1988) The Drosophila mobile element jockey belongs to LINEs and contains coding sequences homologous to some retroviral proteins. Gene 70:253–262

Robertson HM (1993) The *mariner* transposable element is widespread in insects. Nature 362:241–245

Sambrook J, Fritsch FF, Maniatis T (1989) Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Tomita S, Sohn B-H, Tamura T (1997) Cloning and characterization of the *mariner*-like element in the silkworm, *Bombyx mori.* Genes Genet Syst 72:219–228

Xiong Y, Eickbush TH (1990) Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J 9:3353–3362

Zimmerly S, Guo H, Eskes R, Yang J, Perlman PS, Lambowitz AM (1995a) A group II intron RNA is a catalytic component of a DNA endonuclease involved in intron mobility. Cell 83:529–538

Zimmerly S, Guo H, Perlman PS, Lambowitz AM (1995b) Group II intron mobility by target DNA-primed reverse transcription. Cell 82:545–554