Genetic exchange between endogenous and exogenous LINE-1 repetitive elements in mouse cells

Abdellah Belmaaza[1], John C.Wallenburg, Suzanne Brouillette, Nadine Gusew and Pierre Chartrand[1]*
Canadian Red Cross Society, Research and Development, Montreal Centre, 3131 Sherbrooke St. E, Montreal, Quebec H1W 1B2 and [1]Department of Microbiology and Immunology, Université de Montreal, Montreal, Canada

Received June 13, 1990; Revised and Accepted September 11, 1990

ABSTRACT

The repetitive LINE (L1) elements of the mouse, which are present at about 10^5 copies per genome and share over 80% of sequence homology, were examined for their ability to undergo genetic exchange with exogenous L1 sequences. The exogenous L1 sequences, carried by a shuttle vector, consisted of an internal fragment from L1Md-A2, a previously described member of the L1 family of the mouse. Using an assay that does not require the reconstitution of a selectable marker we found that this vector, in either circular or linear form, acquired DNA sequences from endogenous L1 elements at a frequency of 10^{-3} to 10^{-4} per rescued vector. Physical analysis of the acquired L1 sequences revealed that distinct endogenous L1 elements acted as donors and that different subfamilies participated. These results demonstrate that L1 elements are readily capable of genetic exchange. Apart from gene conversion events, the acquisition of L1 sequences outside the region of homology suggested that a second mechanism was also involved in the genetic exchange. A model which accounts for this mechanism is presented and its potential implication on the rearrangement of L1 elements is discussed.

INTRODUCTION

The LINE-1 (L1) family of long interspersed repetitive elements has been found in all mammalian species investigated (reviewed in references 1–3). L1 elements are present at about 10^5 copies per genome and, within species, share over 80% of sequence homology. Full-length elements are 6–7 kbp long but approximately 95% of them are heterogenously truncated from their 5'end and many of them are rearranged. Restriction mapping (4, 5) and DNA sequence analyses (6) have indicated that these elements are evolving in concert such that they are more similar to each other within than between species. Gene conversion and recent dispersion have been invoked as potential mechanisms to explain this concerted evolution (6–8). Analyses of L1 sequences have also revealed features indicating that L1 elements are retroposons inserted into the genome as reverse transcripts of RNA (reviewed in references 9, 10). Consensus L1 sequences feature two open reading frames (11–19) one of which shows patchy homology to retroviral reverse transcriptases (11, 17, 18, 20). This suggests that L1 elements, unlike other mammalian retroposons, may be responsible for their own dispersion. The fact that their presence or absence causes allelic variations at a number of genetic loci (21–27) and that their *de novo* insertion causes mutations (28) suggest that L1 elements are still mobile. Although full-length L1 polyadenylated transcripts have been found (12, 29) no L1 protein products have been identified.

Despite our current knowledge on the structure and genomic organization of the L1 elements there is still much to be learned about them, such as the number of active L1 elements, the products of their open reading frames, the regulation of their transcription, and the mechanisms of their dispersion, rearrangement and truncation. The answers to many of these questions lies in the identification and analysis of functional L1 elements. However, the high copy number of nonfunctional L1 elements renders this very difficult. The ability to 'label' individual L1 elements (by genetic exchange with an exogenous L1 element) would greatly facilitate their analysis. As a first step we needed to determine whether or not L1 elements were capable of genetic exchanges since their heterogeneity both in terms of sequence and structure could interfere with this process. Current assays for genetic exchange that rely on the reconstitution of a selectable marker could not be used for this purpose. Thus we devised a nonselective assay based solely on the acquisition of endogenous L1 sequences by an exogenous L1 element. We found that genetic exchange occurred readily, and that many distinct endogenous L1 elements could act as donors.

MATERIALS AND METHODS

Cell lines and culture

MOP-8 cells (ATCC No.: CRL 1709) are NIH3T3 mouse cells carrying an integrated copy of the early region of the polyoma (Py) virus genome wich allows the replication of exogenous DNA

* To whom correspondence should be addressed

molecules with Py origins of replication (30). Cells were maintained in DMEM medium supplemented with 10% calf serum and in 5% CO_2 atmosphere at 37°C.

Vector construction

The 3.8 kbp *Eco*RV(3410)-*Eae*I(7191) fragment from L1Md-A2 (11) was isolated, the recessed end of *Eae*I was filled-in using the Klenow polymerase (Amersham), and the fragment was cloned into the unique PvuII site of pSBL19, (pSBL19 was generated by modifying the unique *Ssp*I site of pUC19 to a *Pst*I site), to generate pASB-E2. pASB-Hind was generated by replacing the *Kpn*I-*Pst*I fragment of pASB-E2 with the *Kpn*I(4693)-*Pst*I(486) origin-containing fragment of polyomavirus. pASB-Xho was generated by digesting pASB-Hind with *Hin*dIII, filling-in the recessed ends with the Klenow polymerase and ligating in an XhoI linker (Pharmacia). pASB-Eco was generated by deleting the 1.4 kbp *Eco*RI fragment of pASB-Xho, recircularizing the plasmid and reisolating it in bacteria.

Transfection and plasmid rescue

1ug of plasmid DNA was introduced into MOP-8 mouse cells by the DEAE-DEXTRAN procedure (30). Transfected cells (2×10^6) were incubated at 37°C for 48 hours. Low molecular weight plasmid DNA was extracted by the Hirt method (31) and resuspended in 20 ul of TE-RNase (10 mM Tris-HCl pH 8.0, 1.0 mM EDTA, RNase 50 ug/ml). The Hirt extract was used to transform *E.coli* RecA$^-$ DH5 competent cells (Bethesda Research Laboratories). Two procedures were used to identify recombinants. In the first one the transformed bacteria were grown overnight at 37°C in the presence of ampicillin (50 ug/ml). Plasmid DNA was extracted from this library, cleaved with *Hin*dIII and ligated to a *Hin*dIII kanamycin resistance gene cassette derived from *Hin*dIII cleavage of the plasmid pUC4KIXX (Pharmacia Inc.). The ligation mixture was used to transform DH5 bacterial cells which were then plated onto kanamycin-containing agar plates. In the second procedure, the bacteria transformed with the Hirt preparation were immediately plated on agar plates containing ampicillin (50 ug/ml), and the resulting colonies were transfered to nylon membranes and hybridized *in situ* with ^{32}P-labelled probes (32).

Southern analyses

DNA from MOP-8 cells was extracted and purified as previously described (33). Enzyme-restricted DNA was electrophoresed in agarose gels and transfered to nylon membranes (Hybond N, Amersham) by the technique of Reed and Mann (34), with the modifications suggested by the manufacturer. Hybridization of these membranes were done as previously described (35).

DNA sequencing

Sequencing was done by the dideoxy chain termination reaction (36) using *Taq* DNA polymerase (TaqTrack kit, Promega).

RFLP analyses

A 2 kbp fragment between bases 4450 and 6420 was amplified in a Perkin-Elmer Cetus thermal cycler using a GeneAmp kit (Perkin-Elmer Cetus), following the instructions of the manufacturer. The amplified fragments were restricted with *Hin*fI, *Taq*I and *Mbo*II, and seperated on 2% Nuseive (FMC), 1% NA (Pharmacia) agarose gels.

RESULTS

Experimental design

The endogenous L1 elements of the mouse were analysed for their capacity to interact with exogenous L1 sequences. The exogenous L1 sequences consisted of an internal 3.0 kbp fragment from L1Md-A2, a previously described L1 element of the mouse (11). This fragment was cloned in a shuttle vector to generate pASB-Hind (Fig. 1). The cloned L1 fragment contains two *Eco*RI sites that delineate a 1.4 kbp fragment, which includes a unique *Hin*dIII site that subdivides it into fragments of 760 and 620 bp. These three sites are conserved in most of the mouse endogenous L1 elements (4). The *Hin*dIII restriction site was changed to an *Xho*I site using a synthetic *Xho*I linker and the resulting pASB-Xho plasmid was introduced into mouse cells by the DEAE-DEXTRAN procedure (30). Genetic exchange between this plasmid and most of the endogenous L1 elements was expected to regenerate the initial *Hin*dIII site. However, this would not lead to a selectable phenotype. Thus to score for this event, plasmids were rescued from mouse cells by the Hirt procedure (31) and used to transform a RecA$^-$ strain of *E.coli*. The resulting library was allowed to amplify in ampicillin containing medium and total plasmid DNA was extracted and digested with *Hin*dIII. Plasmid DNA molecules that had regained the *Hin*dIII site were recovered by inserting a *Hin*dIII Kanamycin resistance gene cassette into this site. Kanamycin resistant (Kan^r) colonies were selected after bacterial transformation. To determine whether the Kan^r colonies contained recombinants (molecules where exogenous L1 sequences had been replaced by endogenous L1 sequences), plasmid DNA was extracted and analysed by restriction enzymes.

Frequency of genetic exchange

Since the kanamycin cloning procedure is not quantitative, the frequency of genetic exchange between exogenous and

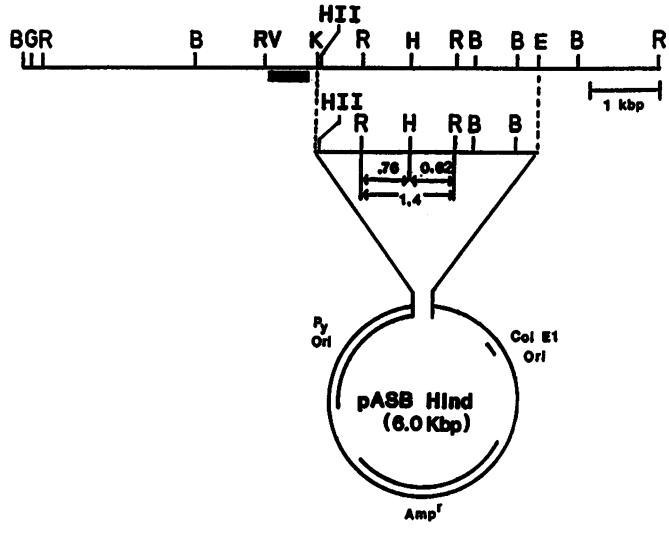

Figure 1. Maps of pASB-Hind and pASB-Xho. Top line: a partial restriction enzyme map of L1Md-A2 (11). Middle line: a 3.0 kbp internal fragment [from *Kpn*I (4186) to *Eae*I (7191)] of L1Md-A2 was cloned into a pUC19 derivative plasmid (pSBL19, Materials and Methods) containing the polyoma A2 origin of replication (Py Ori) to generate pASB-Hind (bottom circle). In pASB-Xho, the *Hin*dIII site has been replaced with an *Xho*I site. The filled rectangle represents the *Eco*RV-*Kpn*I probe. Restriction sites: B, *Bam*HI; G, *Bgl*I; R, *Eco*RI; H, *Hin*dIII; HII, *Hin*cII; K, *Kpn*I; RV, *Eco*RV. Col E1: col EI origin of replication. Amp^r: ampicillin resistance gene from pUC19. Fragment lengths are shown in kilobase pairs (kbp).

endogenous L1 sequences was determined as follows: the Hirt preparation of each experiment was fractionated to determine the minimum number of rescued plasmids from which we could still recover a recombinant molecule. At least ten independant libraries were generated from each Hirt preparation. Each library was tested for the presence of recombinants. The frequency of recombinants per rescued plasmid was calculated to be equal to the number of positive libraries divided by the total number of plasmids used to generate the libraries. When mouse cells were transfected with circular pASB-Xho, the average frequency from four experiments was 0.5×10^{-4}. Linearization of pASB-Xho by either *Xho*I or *Eco*RI (*Eco*RI cleavage creates a gap of 1.4 kbp in the L1Md-A2 sequences (Fig. 1)) before transfection resulted in frequencies of 0.4×10^{-3} in both cases (averages of four experiments).

To confirm that the recombinants were generated in mouse cells prior to bacterial transformation the following control experiments were done in parallel. Circular or linear pASB-XhoI DNA, either alone, or mixed with a Hirt preparation derived from nontransfected mouse cells, was introduced in bacteria and amplified. Plasmid DNA was extracted from these libraries, digested with *Hin*dIII and ligated to the kanamycin cassette. Ten times more transformants were screened than for the experiments and no Kan^r colonies were found. Thus we conclude that no *Hin*dIII site had been generated by passage in bacteria.

Since cleavage of pASB-Xho by *Eco*RI deletes a fragment of 1.4 kbp of L1Md-A2 sequences, the frequency of recombinants per rescued plasmid could be determined directly by colony hybridization with the deleted fragment. To prevent any contamination with the deleted fragment pASB-XhoI was digested with *Eco*RI, recircularized and purified by cloning. The resulting plasmid, pASB-Eco was linearized at its unique *Eco*RI site and introduced into mouse cells. Plasmids which had recovered the deleted L1 sequences were identified by *in situ* hybridization. The average frequency from six experiments was 0.5×10^{-3} and was comparable to that obtained for this substrate by the kanamycin cloning procedure described above, thus confirming the accuracy of the latter procedure.

Restriction enzyme analysis of the recombinants

Extensive restriction enzyme analysis allowed the recombinants to be classified into three groups as follows:
1. recombinants which were identical in structure to the input plasmid, 2. recombinants with rearrangements in their L1 sequences, and 3. recombinants in which the genetic exchange involved only one end of the vector.

The recombinants belonging to the first group had the same structure as pASB-Hind and, presumably, had been generated by gene conversion between the vector and endogenous L1 sequences. In some cases restriction sites had been gained or lost. This group of recombinants represented about 50% of all recombinants. The *Eco*RI/*Hin*dIII restriction patterns of representatives of this group are illustrated in Fig. 2A. When digested with *Eco*RI, pASB-Hind generates a fragment of 1380 bp (lane 2), that is cut by *Hin*dIII into two fragments of 760 bp and 620 bp (lane 3). Recombinants with this restriction pattern were recovered from each of the four different types of assays used in this study (lanes 4, 5, 6 and 7). This restriction pattern is the one most frequently found in endogenous L1 elements of the mouse (4). Recombination was not restricted however to L1 elements with this pattern. Some recombinants had regained the *Hin*dIII site, but were missing one or both of the *Eco*RI sites, generating, for example, only the 620 bp fragment and not the 760 bp fragment (lane 8). Others had the *Hin*dIII site positioned elsewhere inside (lane 9, 1080 bp and 300 bp) or outside the *Eco*RI fragment (lane 10, 1380 bp and 180 bp) thus generating new fragments. In the assays using pASB-Eco, where the presence of a *Hin*dIII site was not selected for, we obtained molecules that had regained the *Eco*RI fragment devoid of *Hin*dIII sites (data not shown). In all cases the origin of the fragments was confirmed, by hybridization, to be derived from the L1 region containing the *Eco*RI fragment (Fig. 2B).

In a previous restriction enzyme analysis of mouse genomic

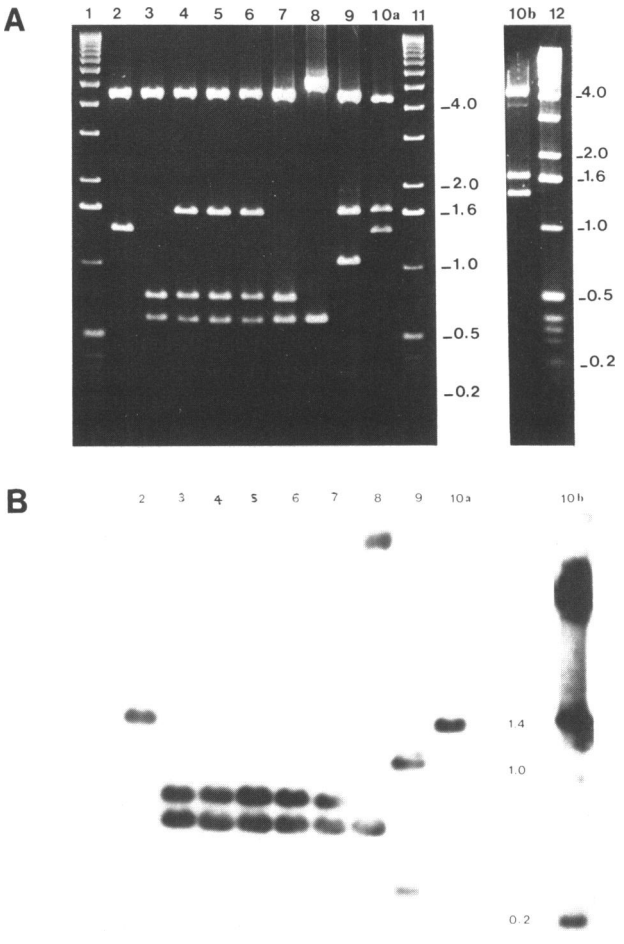

Figure 2. *Eco*RI/*Hin*dIII restriction analysis of group 1 recombinants. A. Digestion of pASB-*Hin*d with *Eco*RI (lane 2) generates two fragments: the 1380 bp L1 *Eco*RI fragment and a 4.5 kbp fragment that represents the rest of the vector. Further digestion with *Hin*dIII cleaves the *Eco*RI fragment into 760 and 620 bp fragments (lane 3). Lanes 4–10: *Eco*RI/*Hin*dIII digestions of recombinants. The recombinants were generated from the following assays: circular pASB-Xho (lane 4); *Xho*I-linearized pASB-Xho (lane 5); *Eco*RI-gapped pASB-Xho (lanes 6, 9, and 10); *Eco*RI-linearized pASB-Eco (lanes 7 and 8). The 1.6 kbp fragment in lanes 4, 5, 6, 9, and 10 is the *Hin*dIII kanamycin gene cassette. In Lane 8, the 4.5 kbp fragment is longer by 760 bp due to the absence of one of the *Eco*RI sites. Lane 10b is a repeat of lane 10a but run on a different gel to be hybridized with a second probe. Lanes 1, 11 and 12 are 1 kbp ladders (BRL). Numbers refer to length in kbp. B. Hybridization of lanes 2–10a in Figure 2A with the 1.4 kbp *Eco*RI L1 fragment illuminates the 1380 bp *Eco*RI fragment and the 760 and 620 bp *Eco*RI/*Hin*dIII fragments. The 5.3 kbp vector fragment in lane 8 hybridizes because of the presence of the 760 bp fragment. Lane 9: the *Hin*dIII digestion of the *Eco*RI fragment generates fragments of 1080 and 300 bp. Lane 10b was hybridized with a probe consisting of all of the L1 sequences present in pASB-Hind. This illuminates the vector fragment, the 1380 *Eco*RI fragment, and an additional 180 bp fragment which does not hybridize with the 1380 *Eco*RI fragment.

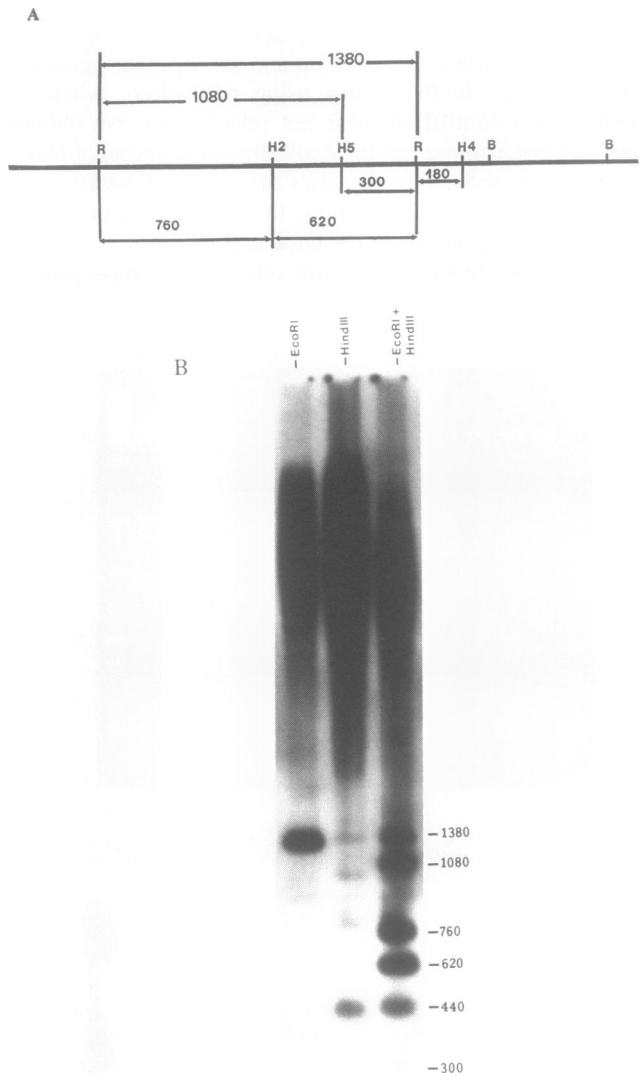

Figure 3. A. Position of the *Hin*dIII sites found in various recombinants. The positions of *Hin*dIII sites were determined by co-digestion with *Eco*RI, *Bam*HI, *Cla*I, *Kpn*I and *Pst*I. The nomenclature of the *Hin*dIII (H2 and H4) sites is according to Brown and Dover (4). H5 is described in the text. Lengths are given in base pairs. B, *Bam*HI; R, *Eco*RI. B. Autoradiogram of restricted mouse DNA after migration in agarose gel, transfer to nylon membrane and hybridization to 1.4 kbp ^{32}P-labelled *Eco*RI fragment of pASB-Hind. *Eco*RI digestion produces a major band of 1.4 kbp and a high molecular weight smear indicating L1 elements heterogenous for the *Eco*RI sites. *Hin*dIII digestion produces fragments of different sizes resulting from a variety of combinations of *Hin*dIII sites, including a predominant 440 bp fragment. *Hin*dIII/*Eco*RI co-digestion produces 5 fragments (1380 bp, 1080 bp, 760 bp, 620 bp, and 300 bp) found in the recombinants, as well as the 440 bp *Hin*dIII fragment.

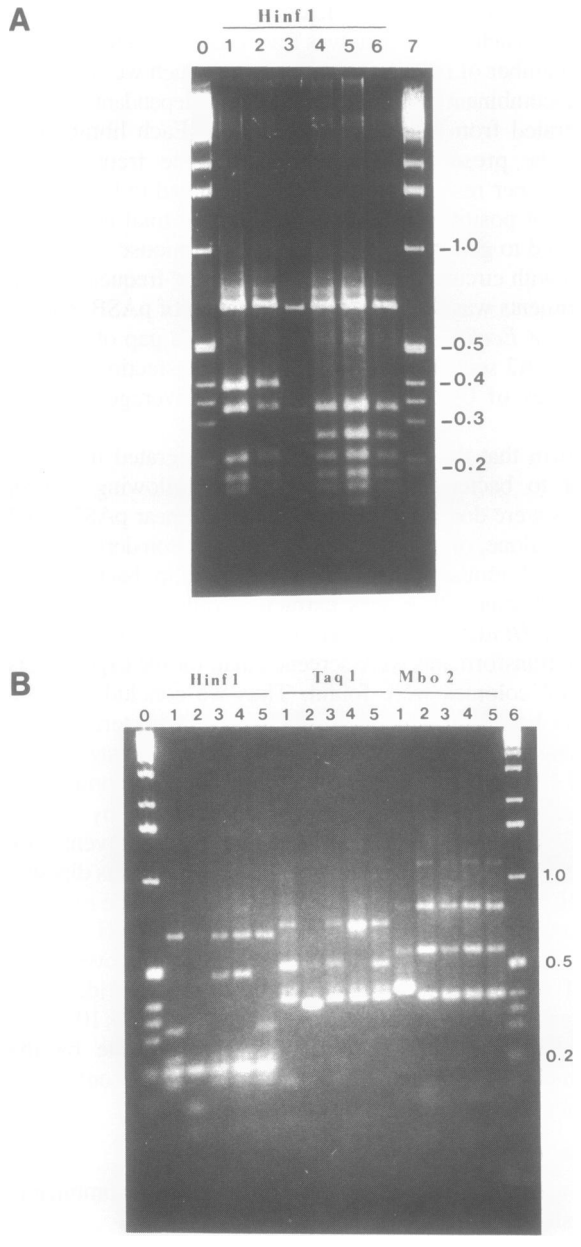

Figure 4. RFLP analysis of recombinants having the same *Eco*RI/*Hin*dIII profile as pASB-Hind. A region containing the *Eco*RI fragment of L1 sequences of each recombinant and pASB-Hind was amplified by PCR and digested with *Hin*fI (A) or *Hin*fI, *Taq*I, and *Mbo*II (B). A. Restriction patterns of recombinants derived from distinct experiments with circular pASB-Xho (lane 1, 2, 3, 4 and 5). The kanamycin gene cassette was removed from these recombinants before amplification. Note that the recombinants in lanes 1 and 2 differ from pASB-Hind (lane 6). Lanes 0 and 7: 1 kbp molecular weight ladder(BRL). Fragment lengths are given in kbp. B. Restriction patterns of pASB-Hind (lanes 5) and of recombinants derived from distinct experiments with linearized pASB-Eco (lanes 1, 2, 3 and 4). Note that *Hin*fI digestion serves to distinguish lanes 2, 3, and 4 from pASB-Hind (lane 5) and lane 2 from lanes 3 and 4. The *Taq*I digestion further distinguishes between lanes 3 and 4. The *Mbo*II (*Mbo*2) digestion serves to distinguish between lanes 1 and pASB-Hind (lane 5). Lanes 0 and 6: 1 kbp molecular weight ladder(BRL). Fragment lengths are given in kbp.

DNA, Brown and Dover (4) identified four conserved *Hin*dIII sites (H1-H4) in and around the 1380 bp *Eco*RI fragment and used these sites to define subfamilies of L1 elements. Interestingly, we found recombinants that had acquired H2 or H4 but not H1 or H3. We also recovered an additional *Hin*dIII (H5) site situated 1080 bp and 300 bp from the left and right *Eco*RI sites, respectively (Fig. 3A). Hybridization of *Hin*dIII/*Eco*RI restricted mouse genomic DNA with the *Eco*RI fragment reveals the high incidence in endogenous L1 elements of the fragments observed in our recombinants including the 1080 and 300 bp fragments (Fig. 3B). It also shows the presence of a 440 bp fragment not observed in our recombinants. Thus genetic exchange was not restricted to a given subset of L1 elements although some were underrepresented in our recombinants.

The recombinants that had the same *Eco*RI/*Hin*dIII restriction patterns could have acquired their L1 sequences from the same L1 element or alternatively, from distinct L1 elements. Since the endogenous L1 elements are not 100% homologous, distinct

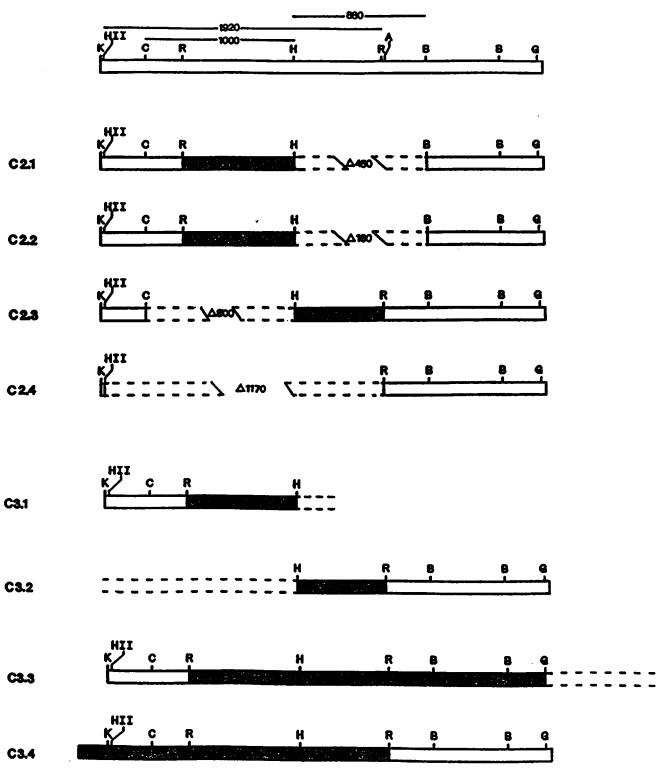

Figure 5. Examples of the L1 sequences present in group 2 and 3 recombinants. The top line represents restriction map of the L1Md-A2 fragment present in pASB-Hind; the numbers above the open box refer to fragment lengths in base pairs. C2.1 to C2.4: group 2 recombinants. C3.1 to C3.4: group 3 recombinants. Stippled boxes: L1 sequences necessarily recovered from endogenous LINEs. Broken lines: acquired L1 sequences the exact nature of which have not been determined. Open triangles indicate deletions in the acquired L1 sequences compared to L1Md-A2, the deletion lengths are indicated in base pairs. K, *Kpn*I; HII, *Hinc*II; C, *Cla*I; R, *Eco*RI; H, *Hin*dIII; A, *Ava*I; B, *Bam*HI; G, *Bgl*I.

recombinants were likely to have L1 sequences with some variations in their restriction patterns. A 2 kbp region including the *Eco*RI fragment from nine recombinants with identical *Eco*RI/*Hin*dIII patterns, was amplified by the polymerase chain reaction. When digested with *Hin*fI the amplified fragment of 2 out of 5 recombinants derived from the assays with circular pASB-Xho were found to be different from pASB-Hind (Fig. 4A). Four recombinants derived from the assays with pASB-Eco and analysed with three restriction enzymes were found to be different (Fig. 4B). From this we conclude that most recombinants analysed had acquired sequences from distinct L1 elements.

The second group of recombinants had rearranged L1 sequences. They all originated from the assays with *Eco*RI linearized pASB-Eco and were identified by hybridization to the *Eco*RI fragment. Representatives are illustrated in Fig. 5 (C2.1-C2.4). The rearrangemants could be most easily explained by deletions in the *Eco*RI fragment (that ranged in size from 100 bp to 1100 bp) and that could extend beyond. One explanation for the generation of these recombinants is that the rearrangements were already present in the donor L1 sequences and were transfered as such to the recipient molecule by gene conversion as in the case of the first group of recombinants. Alternatively they may have resulted from a distinct mechanism of genetic transfer (see discussion).

In the third group of recombinants, only one end of the vector

```
                                                                    KpnI
L1Md GCTATACTCAACAAAACTGGAAAACCTGGACGAAATGGACAAATTTCTGGACAGATACCAGGTACC
C3-4 ............................................................

L1Md AAAGT-TGAATCAGGATCAAGTTGACCATCTAAACAGTCCCATATCACCTAAAGAAATA
C3-4 .....T.....................................................A.....
```

Figure 6. Sequence of recombinant C3.4. The top line gives the sequence of L1Md-A2 from base 4126–4249 (11). The bottom line is the sequence of C3.4 towards the 3′ end. Sequences are shown 5′ to 3′, and the 5′-end of the LINE sequence is to the left. Only bases dissimilar to the L1Md-A2 sequence are shown in the sequence of C3.4, and a 1-base insertion in C3.4 is shown as a space in the L1Md-A2 sequence. The L1 sequences in C3.4 before the *Kpn*I site (base 4186) do not exist in the vector and were necessarily acquired from an endogenous LINE.

could have participated in the genetic exchange. This was due to the fact that the molecules that were involved in the genetic exchange were all constituted of head to head dimers of the vector sequences such that both ends of the molecule were identical, though in opposite orientation. Head to head ligation of transfected DNA molecules is not an uncommon event (37). These recombinants had acquired various lengths of L1 sequences continuous with one of the vector's ends (examples illustrated in Fig. 5, C3.1-C3.4). The other end had formed a nonhomologous junction with the acquired sequences. In the case of recombinant C3.4, 2 kbp of L1 sequences had been acquired including the entire *Eco*RI fragment, the sequences between the *Eco*RI and *Kpn*I restriction sites, and an additional 64 bp upstream of the *Kpn*I site. The *Kpn*I site represents the 5′end of the L1Md-A2 fragment cloned in the vector. Hybridization analysis using the *Eco*RV-*Kpn*I fragment from the L1Md-A2 sequences as a probe (Fig. 1) confirmed that the additional 64 bp originated from endogenous L1 sequences (data not shown). This region was sequenced and found to be continuous with the sequences originally present in the vector and to have a 1-base difference and a 1-base addition when compared to the L1Md-A2 sequence (Fig. 6).

In order to confirm that sequences outside the region of homology (external to the L1 sequences present in the vector) could be acquired by a recombination mechanism involving only one homologous end, we did the following assay. pASB-XhoI was cut by *Kpn*I to generate a molecule with only one end homologous to endogenous L1 elements (Fig. 1). Mouse cells were transfected, and vector DNA rescued as before. Recombinant plasmids that had acquired endogenous L1 sequences upstream of the 5′ *Kpn*I site were identified by *in situ* hybridization using the *Eco*RV-*Kpn*I fragment as a probe (Fig. 1). The average frequency from four such experiments was one positive signal per 10,000 rescued plasmid molecules. Restriction enzyme analysis of four recombinants showed that they had acquired several kbp of endogenous sequences (data not shown). These results demonstrate that homologus recombination proceeding from a single homologous end can acquire L1 sequences from outside the region of homology.

DISCUSSION

We have introduced in mouse cells a vector containing a segment of L1 sequences and have rescued, at a frequency of 10^{-3} to 10^{-4}, molecules in which the L1 sequences had been modified. It is obvious that the molecules that had acquired the missing 1.4 Kbp L1 fragment must have been generated by recombination

with endogenous L1 elements. However, in the assays involving the reconstitution of the *Hin*dIII site it is conceivable that this event could also have resulted from mutagenesis. Several authors have shown that transfected vectors are mutagenized at a high frequency during their sojourn in mammalian cells (38–43). We feel this is an unlikely explanation for the following reasons. Although it has been shown (38, 40, 41) that inactivation of genes present on a shuttle vector can occur at a frequency of 10^{-2}, Miller *et al* (43) have demonstrated that specific mutations occurred only at a frequency of 10^{-5} or less. This is much lower than the frequency we observed for the reconstitution of the *Hin*dIII site. Secondly, a number of the molecules we analysed had sustained multiple modifications, which Would imply multiple independent mutagenic events per molecule. Finally the modifications observed were specific to certain sites and corresponded to polymorphism already present in the endogenous L1 sequences. Mutagenesis of transfected DNA has been shown to be non-specific (39, 41, 43).

Another point to be considered is the possibility that recombination of the vector we introduced occurred primarily with L1 sequences carried by the extrachromosomal free circles present in mammalian cells (44, 45). We and others have shown that the primary product of extrachromosomal recombination is non-homologous recombinants (46–50). If recombination had occured frequently between the vector and the extrachromosomal DNA containing L1 sequences we would have expected to find mainly non-homologous recombinants. Actually, all the examined molecules that had acquired L1 sequences had done so by homologous recombination.

Our results indicate that genetic exchange between exogenous and endogenous L1 elements can be readily detected in mouse cells. Physical analyses of the acquired L1 sequences revealed that numerous distinct donor L1 elements had been involved. Since the acquired L1 sequences could be assigned to several previously described subfamilies (4) genetic exchange was not restricted to a given subfamily, however certain subfamilies were not found. The transfer of genetic information between L1 elements has been hypothesized as one of the mechanisms to explain their homogeneity (7, 8). Comparison of their sequences reveals a patchwork pattern of homology suggestive of gene conversion (51, 52). Our results indicate that L1 elements are capable of such a process.

The frequency we obtained of transfer of genetic information from endogenous L1 sequences to an extrachromosomal molecule is comparable to the upper limit of what has been reported for the acquisition by a vector of sequences present only once in the genome (53, 54). One might have expected this frequency to be significantly higher because of the high copy number of L1 elements in the genome. However there are differences between individual L1 elements and it has been shown that heterogeneity between homologous sequences negatively affects the frequency of recombination (55). Additionally it has been reported that copy number is not a limiting factor, at least for homologous integration (56).

For the majority of the recombinants analysed, the acquisition of endogenous L1 sequences can be explained by gene conversion. Current models of gene conversion require that homologous sequences from both sides of the converted region be involved in the process (reviewed in reference 57). However, these models cannot explain those recombinants that had acquired endogenous L1 sequences by a recombination mechanism involving only one homologous end. The single-strand annealing model of nonconservative homologous recombination (58–61) was proposed to explain extrachromosomal nonconservative homologous recombinants with one homologous and one nonhomologous junction. However this model invokes double-strand breaks in both substrates, which would appear to be an unlikely event for chromosomal DNA. Furthermore, it would not lead to the acquisition of chromosomal sequences by the exogenous molecule, but rather to its integration (62). Therefore to account for these recombinants, we propose a model based on priming and extension.

In this model (Fig. 7) a 3′ single strand end of the recipient molecule (A) invades homologous sequences of the donor thus creating a D-loop (B). DNA synthesis primed on the invading 3′ end extends the D-loop (C). The newly synthesized strand could be released either by unwinding (D) or by single-strand breaks of the template strand (E), then ligated back to the recipient molecule. DNA synthesis completes the process. Although the initiation step is dependent on homology between the substrates, the length and nature of the sequences acquired by the recipient are only dependent on the extent of polymerization, and independent of homology. In this way if the polymerization extends beyond the homologous sequences shared with the vector and continues into the flanking sequences, nonhomologous sequences would be acquired by the vector as was the case for recombinant C3.4. The recombinants with partial gap repair could also be explained by this model if the polymerization does not extend across the entire region of homology. Both these kinds of recombinants have been described previously (63–65). In one report, partial gap repair accounted for 20% of the recombinants in yeast (63), and in another, more than 60% of the recombinants

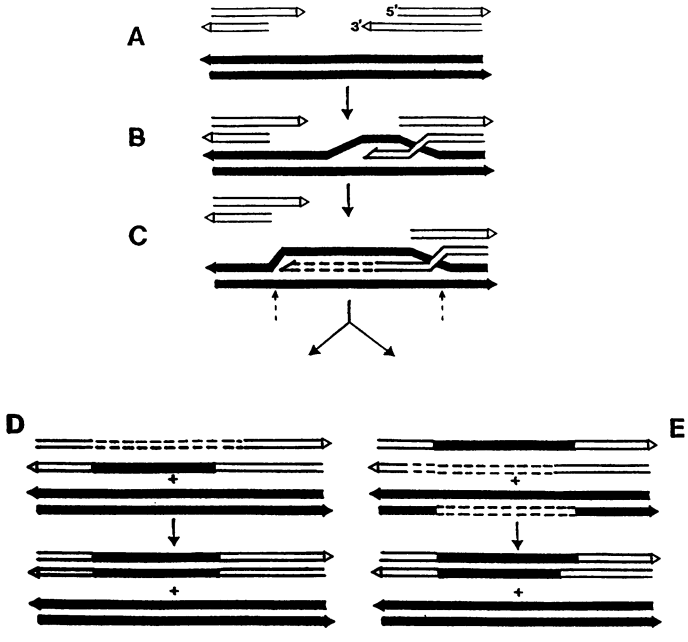

Figure 7. Primer-extension model of genetic exchange. Open lines: recipient molecule. Solid lines: donor molecule. Broken lines: newly synthesized DNA. Arrow heads identify the 3′-ends of DNA molecules. In this model, recombination would be initiated by a double-strand break (A). This would be followed by single-strand invasion from one end (B). DNA synthesis primed at the invading strand would create a D-loop (C). Resolution could occur by (D) unwinding followed by ligation of the newly synthesized strand to the non-invading end of the lower strand of the recipient molecule. Alternatively, (E) resolution could occur as a result of single-strand nicks at the extremities of the D-loop (see broken arrows, step C), followed by ligation of the bottom donor strand to the top recipient strand. Repair synthesis would complete the process in both cases and give the final products shown at the bottom.

obtained in mammalian cells had acquired flanking nonhomologous sequences from the chromosome (64). Thus, primer-extension could represent a common mechanism that may often go undetected in systems using selection. Precedents of this model have been proposed to explain the replication of yeast telomeres (66) and the mobility of mitochondrial group 1 introns (67).

If two chromosomal sequences were involved, an interesting possibility is that this mechanism, amongst others, could account for the duplication and scrambling which are frequently observed in L1 elements (9). If primer-extension involved the (A-rich) tails of two retroposons this would lead to insertion of one retroposon into the tail of the other. This has been frequently observed (9). If an A-rich region devoid of an L1 element invaded the A-rich tail of an existing L1 element, priming and extension would lead to the acquisition, by this region, of the invaded L1 element or a 5'-truncated portion of it. Thus priming and extension could participate in L1 elements dispersal.

In conclusion, we have demonstrated that endogenous L1 elements can exchange genetic information with exogenous L1 sequences. This can be used to tag individual endogenous L1 elements with a selectable marker. The study of L1 elements is rendered difficult by their high copy number and homogeneity, tagging will permit to study individual copies with regard to their transcription and transposition. This work is currently in progress.

ACKNOWLEDGEMENTS

We are grateful to Clyde A. Hutchison III and Daniel B. Loeb for the generous gift of L1Md-A2, Roxanne Lussier for technical assistance, Pierre Trudel and Gilles Roy for helpful discussion and Nicole Langlois for typing of the manuscript. This work was supported by a grant from the Medical Research Council of Canada.

REFERENCES

1. Hutchison III,C.A., Hardies,S.C., Loeb,D.D., Shehee,W.R. and Edgell,M.H. (1989) In Berg,D.E., and Howe,M.M. (ed.), Mobile DNA. ASM, Washington, DC, pp. 593–617.
2. Burton,F.H., Loeb,D.D., Voliva,C.F., Martin,S.L., Edgell,M.H. and Hutchison III,C.A. (1986) *J. Mol. Biol.*, **187**, 291–304.
3. Fanning,T.G. and Singer,M.F. (1987) *Biochim. Biophys. Acta* **910**, 203–212.
4. Brown,S.D.M. and Dover,G (1981) *J. Mol. Biol.*, **150**, 441–466.
5. Grimaldi,G. and Singer,M.F. (1983) *Nucleic Acids Res.*, **11**, 321–338.
6. Martin,S.L., Voliva,C.F., Hardies,S.C., Edgell,M.H. and Hutchison III,C.A. (1985) *Mol. Biol. Evol.*, **2**, 127–140.
7. Dover,G. (1982) *Nature* **299**, 111–117.
8. Hardies,S.C., Martin,S.L., Voliva,C.F., Hutchison III, C.A. and Edgell,M.H. (1986) *Mol. Biol. Evol.*, **3**, 109–125.
9. Rogers,J. (1985) *Int. Rev. Cytol.*, **93**, 187–279.
10. Singer,M.F. and Skowronski,J. (1985) *Trends Biochem. Sci.*, **10**, 119–122.
11. Loeb,D.D., Padgett,P.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.
12. Skowronski,J., Fanning,T.G. and Singer,M.F. (1988) *Mol. Cell. Biol.*, **8**, 1385–1397.
13. Shehee,W.R., Chao,S.F., Loeb,D.D., Comer,M.B., Hutchison III,C.A. and Edgell,M.H. (1987) *J. Mol. Biol.*, **196**, 757–767.
14. D'Ambrosio,E., Waitzkin,S.D,. Witney,F.R., Salemme,A. and Furano,A.V. (1986) *Mol. Cell. Biol.*, **6**, 411–424.
15. Demers,G.W., Matunis,M.J. and Hardison,R.C. (1989). *J. Mol. Biol. Evol.* **29**, 3–19.
16. Hattori,M., Hidaka,S. and Sakaki,Y. (1985) *Nucleic Acids Res.*, **13**, 7813–7827.
17. Hattori,M., Kuhara,S., Takenaka,O. and Sakaki,Y. (1986) *Nature* **32**, 625–628.
18. Sakaki,Y., Hattori,M., Fujita,A., Yoshioka,K., Kuhara,S. and Takenaka,O. (1986) Cold Spring Harbor Symp. Quant. Biol., **51**, 465–469.
19. Scott,A.F., Schmeckpeper,B.J., Abdelrazik,M., Thiesen Comey,C., O'Hara,B., Pratt Rossiter,J., Cooley,T., Heath, P., Smith,K.D. and Margolet,L. (1987) *Genomics* **1**, 113–125.
20. Fanning,T. and Singer,M.F. (1987) *Nucleic Acids Res.*, **15**, 2251–2260.
21. Burton,F.H., Loeb,D.D. Chao,S.F., Hutchison III,C.A. and Edgell,M.H. (1985) *Nucleic Acids Res.*, **13**, 5071–5084.
22. Casavant,N.C., Hardies,S.C., Funk,F.D., Comer,M.B., Edgell,M.H. and Hutchison III,C.A. (1988) *Mol. Cell. Biol.*, **8**, 4669–4674.
23. Shyman,S. and Weaver,S. (1985) *Nucleic Acids Res.*, **13**, 5085–5093.
24. Economou-Pachnis,A., Lohse,M.A., Furano,A.V. and Tsichlis,P.N. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 2857–2861.
25. Lakshmikumaran,M.S., D'Ambrosio,E., Laimans,L,A., Lin,D.T. and Furano,A.V. (1985) *Mol. Cell. Biol..*, **5**, 2197–2203.
26. Katzir,N., Rechavi,G., Cohen,B.J., Unger,T., Simono,F., Segal,S., Cohen,D. and Givol,D. (1985). *Proc. Natl. Acad. Sci. USA* **82**, 1054–1058.
27. Woods-Samuels,P., Wong,C., Mathias,S.L., Scott,A.F., Kazazian,Jr.H.H. and Antonarakis,S.E. (1989) *Genomics* **4**, 20–296.
28. Kazazian,H.H., Wong,Jr,C., Youssoufian,H., Scott,A.F., Phillips,D.G. and Antonarakis,S.E. (1988) *Nature* **332**, 164–166.
29. Skowronski,J. and Singer,M.F. (1986) Cold Spring Harbor Symp. Quant. Biol., **51**, 457–464.
30. Muller,W.J., Naujokas,M.A. and Hassel,J.A. (1984) *Mol. Cell. Biol.*. **4**, 2406–2412.
31. Hirt,B.J. (1967) *J. Mol. Biol.*, **26**, 365–368.
32. Grunstein,M. and Hogness,D. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3961.
33. Bowtell,D. (1987) *Anal. Biochem.*, **162**, 453–465.
34. Reed,K.C. and Mann,D.A. (1985) *Nucleic Acids Res.*, **13**, 7207–7221.
35. Singh,L. and Jones,K.W. (1984) *Nucleic Acids Res.*, **12**, 5627–5638.
36. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467
37. Bishop,J.O. and Smith,P. (1989) *Mol. Biol. Med.* **6**, 283–294.
38. Razzaque,A., Mizusawa, H. and Seidman,M.M. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 3010–3014.
39. Razzaque,A., Chakrabarti,S., Joffee,S. and Seidman,M. (1984) *Mol. Cell. Biol.***4**, 435–441.
40. Calos,M.P., Lebkowski,J.S. and Botchan,M.R. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 3015–3019.
41. Ashman, C.R. and Davidson, R.C. (1984) *Mol. Cell. Biol.* **4**, 2266–2272.
42. Lebkowski, J.S., DuBridge, R.B., Antell, E.A., Greisen, K.S. and Calos, M.P. (1984) *Mol. Cell. Biol.* **4**, 1951–1960.
43. Miller, J.H., Lebkowski, J.S., Greisen, K.s. and Calos, M.P. (1984) *EMBO J.* **3**, 3117–3121.
44. Krolewski, J.J. and Rush, M.G. (1983) *J. Mol. Biol.* **174**, 31–40.
45. Kunisada, T. and Yamagishi, H. (1984) *Gene* **31**, 213–223.
46. Brouillette, S. and Chartrand, P. (1987) *Mol. Cell. Biol.* **7**, 2248–2255.
47. Chakrabarti, S., Joffee, S. and Seidman, M. (1985) *Mol. Cell. Biol.* **5**, 2265–2271.
48. Roth, D.B. and Wilson, J.H. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 3355–3359.
49. Stringer, J.R. (1985) *Mol. Cell. Biol.* **5**, 1247–1259.
50. Subramani, S. and Berg, P. (1983) *Mol. Cell. Biol.* **3**, 1040–1052.
51. Rodgers, J.H. (1985) *Biochem. Biophys. Acta* **824**, 113–120.
52. Jubier-Maurin, V., Dod, B.J., Bellis,M., Piechaczyk,M. and Roizes, G. (1985) *J. Mol. Biol.*, **184**, 547–564.
53. Subramani, S. (1986) *Mol. Cell. Biol.* **6**, 1320–1325.
54. Jasin, M., de Villiers, J., Weber, F. and Schaffner, W. (1985) *Cell* **43**, 695–703.
55. Waldman, A.S. and Liskay, R.M. (1988) *Mol. Cell. Biol.* **8**, 5350–5357.
56. Zheng, H. and Wilson, J.H. (1990) *Nature* **344**, 170–173.
57. Orr-Weaver, T.L. and Szostak, J.W. (1985) *Microbiol. Rev.* **49**, 33–58.
58. Anderson, R.A. and Eliason, S.L. (1986) *Mol. Cell. Biol.* **6**, 3246–3252.
59. Chakrabarti, S. and Seidman, M.M. (1986) *Mol. Cell. Biol.* **6**, 2520–2526.
60. Lin, F.-L, Sperle, K. and Sternberg, N. (1984) *Mol. Cell. Biol.* **4**, 1020–1034.
61. Wake, C.T., Vernaleone, F. and Wilson, J.H. (1985) *Mol. Cell. Biol.* **5**, 2080–2089.
62. Lin, F.-L., Sperle, K. and Sternberg, N. (1990) *Mol. Cell. Biol.* **10**, 103–112.
63. Orr-Weaver, T.L. and Szoztak, J.W. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 4417–4421.
64. Mudgett, J.S. and Taylor, W.D. (1990) *Mol. Cell. Biol.* **10**, 37–46.
65. Ellis, J. and Bernstein, A. (1989) *Mol. Cell. Biol.* **9**, 1621–1627.
66. Wang, S.S. and Zakian, V.A. (1990) *Nature* **345**, 456–458.
67. Colleaux, L., D'Auriol, L., Galibert, F. and Dujon,B. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 6022–6026.