MOLECULAR AND CELLULAR BIOLOGY, Jan. 2000, p. 54–60 Vol. 20, No. 1
0270-7306/00/$04.00+0
Copyright © 2000, American Society for Microbiology. All Rights Reserved.

A Double-Strand Break in a Chromosomal LINE Element Can Be Repaired by Gene Conversion with Various Endogenous LINE Elements in Mouse Cells

ANNIE TREMBLAY,[1] MARIA JASIN,[2] AND PIERRE CHARTRAND[1,3]*

Molecular Biology Program[1] and Department of Pathology and Cellular Biology,[3] University of Montreal, Montreal, Quebec, Canada, and Cell Biology and Genetics Program, Memorial Sloan-Kettering Cancer Center and Cornell University Graduate School of Medical Sciences, New York, New York[2]

Received 21 July 1999/Returned for modification 10 September 1999/Accepted 27 September 1999

A double-strand break (DSB) in the mammalian genome has been shown to be a very potent signal for the cell to activate repair processes. Two different types of repair have been identified in mammalian cells. Broken ends can be rejoined with or without loss or addition of DNA or, alternatively, a homologous template can be used to repair the break. For most genomic sequences the latter event would involve allelic sequences present on the sister chromatid or homologous chromosome. However, since more than 30% of our genome consists of repetitive sequences, these would have the option of using nonallelic sequences for homologous repair. This could have an impact on the evolution of these sequences and of the genome itself. We have designed an assay to look at the repair of DSBs in LINE-1 (L1) elements which number 10^5 copies distributed throughout the genome of all mammals. We introduced into the genome of mouse epithelial cells an L1 element with an I-*Sce*I endonuclease site. We induced DSBs at the I-*Sce*I site and determined their mechanism of repair. We found that in over 95% of cases, the DSBs were repaired by an end-joining process. However, in almost 1% of cases, we found strong evidence for repair involving gene conversion with various endogenous L1 elements, with some being used preferentially. In particular, the T_F family and the L1Md-A2 subfamily, which are the most active in retrotransposition, appeared to be contributing the most in this process. The degree of homology did not seem to be a determining factor in the selection of the endogenous elements used for repair but may be based instead on accessibility. Considering their abundance and dispersion, gene conversion between repetitive elements may be occurring frequently enough to be playing a role in their evolution.

DNA double-strand breaks (DSBs) represent a threat to genomic integrity that have to be repaired efficiently or they will lead to serious consequences, such as programmed cell death (45). In mammalian cells, a single DSB can cause a cell cycle arrest (17). Mammalian DSB repair is a complex process that involves multiple proteins that vary depending on the cell cycle phase (16). Two main repair pathways are utilized (22). The first one involves the rejoining of the broken ends with or without the loss and/or addition of DNA. The second one involves homologous recombination with an intact copy of the broken DNA segment acting as a template. In most cases the homologous template could be the allelic copy present on the sister chromatid or homologous chromosome. However, since 30% of the genome is made of highly repetitive sequences, this raises the possibility that nonallelic homologous templates could be used for the repair of a DSB. Indeed, several groups have shown that a DSB occurring in a chromosomal construct containing tandemly repeated sequences could be repaired by homologous recombination at a frequency of as high as 30% (7, 12, 21, 22, 39, 44). Also, it was shown that homologous sequences located at a nonallelic position on a distinct chromosome could be used for DSB repair, albeit at a much lower frequency (34).

We wanted to determine how a DSB in highly repetitive sequences would be repaired. In particular, we were interested in the repair of a DSB occurring in LINE-1 (L1) elements. L1 elements are present in all mammals at around 10^5 copies dispersed throughout the genome and are more conserved within a given species than between closely related species (18). This concerted evolution of L1 elements could be due in part to homologous recombination between nonallelic L1 elements. Indeed, Saxton and Martin (40) suggested that recombination between particular L1 subtypes in mouse could have created the T_F family. There is some evidence that neighboring L1 elements can be involved in homologous recombination leading to chromosomal deletions (9, 43). Also, recombination between L1 elements would be partly responsible for the present hominid Y chromosome structure (42). However, it could not be determined if DSB repair was implicated in these processes.

In order to study DSB repair of L1 elements, we devised an assay based on the I-*Sce*I system (19, 36). We integrated at several distinct sites in the genome of a mouse epithelial cell line an L1 element containing an I-*Sce*I site and then introduced a DSB at that site by the transfection in these cells of a vector expressing the I-*Sce*I endonuclease. The results were that, over 95% of the time, the DSBs in the L1 sequences were repaired by an end-joining process. However, in almost 1% of the cases we found strong evidence for homologous recombination repair involving gene conversion with endogenous L1 elements. Several distinct endogenous L1 elements were used for homologous recombination repair, some preferentially. In particular, the T_F family and L1Md-A2 subfamily, which are currently the most active in retrotransposition, appeared to be contributing most of the elements in these gene conversion events.

* Corresponding author. Mailing address: Centre Hospitalier de l'Université de Montréal (CHUM), Hôpital Notre-Dame, 1560 rue Sherbrooke Est, Montréal, Quebec, H2L 4M1, Canada. Phone: (514) 281-6000, ext. 6939. Fax: (514) 896-4689. E-mail: chartrap@magellan.umontreal.ca.

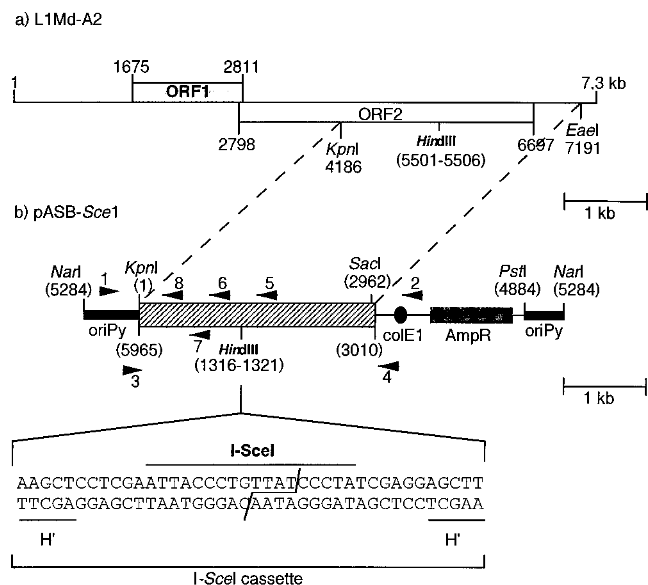

FIG. 1. Structure of L1Md-A2 (23) (a) and the pASB-*Sce*I vector used in our assay (b), which contains the 3-kb *Kpn*I-*Eae*I fragment of L1Md-A2. A linker containing the I-*Sce*I restriction site was introduced at the *Hin*dIII site: H′ designates the partial *Hin*dIII sites flanking the I-*Sce*I cassette, and the I-*Sce*I cut site is shown. oriPy, polyomavirus origin that was present in the initial pASB-*Hin*dIII vector (4). Position of PCR primers (1 to 4) and sequencing primers (5 to 8) in pASB-*Sce*I are as follows: 1 (5850 to 5872), 2 (3090 to 3070), 3 (5887 to 5916), 4 (3043 to 3014), 5 (1500 to 1481), 6 (1107 to 1087), 7 (872 to 850), and 8 (562 to 542).

MATERIALS AND METHODS

Cells and growth conditions. Nod-2 are nonciliated epithelial cells of the bronchioles derived from a pulmonary adenocarcinoma (nodule) of an FVB transgenic mouse that harbors a polyomavirus large-T-antigen (T-Ag) transgene (20). Nod-2 cells are probably derived from precursor cells within the basal bronchiolar epithelium. Cells were maintained in Dulbecco modified Eagle medium (DMEM) medium supplemented with 10% fetal bovine serum plus 50 μg of gentamicin per ml and kept in a 5% CO_2 atmosphere at 37°C. When used, hygromycin was added at a final concentration of 150 μg/ml.

Vector construction. pASB-*Sce*I is derived from pASB-*Hin*dIII which contains the 3-kb *Kpn*I-*Eae*I fragment of the BALB/c L1Md-A2 (4). An I-*Sce*I restriction site was introduced at the unique *Hin*dIII site by using an oligodimer (see Fig. 1b).

Transfection. A total of 2×10^6 nod-2 cells were transfected in DMEM by electroporation with 10 μg of *Nar*I-digested pASB-*Sce*I and 1 μg of *Bam*HI-linearized hygromycin vector, p3′SS (Stratagene). The electroporation conditions used were as follows: 0.4-cm gap chamber, 960 μF, and 270 V. Hygromycin was added 48 h after the electroporation, and Hyg^r colonies were picked 3 weeks later.

Transient I-*Sce*I expression. pCBASce (34) is an expression vector that contains the coding sequences of the I-*Sce*I endonuclease from *Saccharomyces cerevisiae*, under the control of the β-actin gene promoter. A total of 4×10^6 cells in DMEM of the different pASB-*Sce*I containing clones were electroporated (see conditions in the previous section) with 100 μg of uncut pCBASce vector or without vector as a control. Viable cells were harvested 72 h later for genomic DNA extraction and analysis.

DNA analysis. Southern analysis was performed by using 13 μg of digested genomic DNA according to Ausubel et al. (3). The *Nar*I-*Kpn*I polyomavirus origin fragment of pASB-*Sce*I was used as a probe. All of the restriction enzymes used were purchased at New England Biolabs except Meganuclease I-*Sce*I (Boehringer Mannheim). For all PCR amplifications, the Expand High-Fidelity PCR System (Boehringer Mannheim) was used. The primers used for PCR were as follows: 1, 5′-CAGAGGAGGTGTATGGGTTTGTC-3′; 2, 5′-CGAGTCAG TGAGCGAGGAAGC-3′; 3, 5′-GTTTGTAAGTCGAACAGCGGGGGCTAT ATG-3′; and 4, 5′-TCTCCCCGCGCGTTGGCCGATTCATTAATG-3′. PCR1 (primers 1 and 2) amplification was performed on 150 ng of digested genomic DNA as follows: 94°C for 2 min; 10 cycles of 94°C for 1 min, 62°C for 1 min, and 68°C for 2 min; and 20 cycles of 94°C for 1 min, 62°C for 1 min, and 68°C for 2 min (plus a 20-s/cycle). For PCR2 (primers 3 and 4), the same procedure was followed except that the annealing temperature was changed from 62 to 70°C. For DNA sequencing of PCR products, PCR fragments were cloned (*Kpn*I-*Sac*I) or subcloned (*Kpn*I-*Hin*dIII) into pBluescript SK(+) (Stratagene). Sequencing was done by the dideoxy chain termination reaction (38) by using Sequenase (Amersham). Sequence data were compiled and analyzed by using Genetic Computer Group software.

RESULTS

Experimental design. The strategy that we used was to introduce in the genome of a mouse cell line, an L1 fragment in which we had replaced a highly conserved restriction site by an I-*Sce*I site. We then created a DSB at the I-*Sce*I site in vivo and scored for repair events by the loss of the I-*Sce*I. To enrich for homologous recombination repair events, we selected for the reacquisition of the highly conserved restriction site.

The vector we used has been described previously (4). It contains a 3-kb fragment from the 3′ end of the well-characterized L1Md-A2 element (Fig. 1a). We chose the 3′-end fragment since most L1 elements of the mouse genome are truncated at their 5′ end (47). Certain restriction sites are highly conserved in mouse L1 elements, and this is the case for the *Hin*dIII site present in the L1 fragment we used, which is present in more than 50% of the mouse L1 elements (4, 8) (GenBank-EMBL L1 sequences). We introduced in the *Hin*dIII site an I-*Sce*I cassette to create the vector pASB-*Sce*I (Fig. 1b).

pASB-*Sce*I was transfected in a mouse epithelial cell line (nod-2 [20]), along with the hygromycin selection vector p3′SS. Three independent clones were selected (clones 19, 22, and 25); each had integrated into their genome a single complete copy of the *Nar*I-digested pASB-*Sce*I vector, as indicated by the presence of a single complete insert seen by Southern analysis (Fig. 2). We then created a DSB in the integrated L1 sequence in clones 19, 22, and 25 by transfecting an I-*Sce*I expression vector (pCBASce [34]).

To analyze the repair events, the genomic DNA was first digested with *Pst*I to break down the genome and I-*Sce*I to cleave the integrated L1 sequences which had not lost the I-*Sce*I site. Integrated L1 sequences which had lost the I-*Sce*I site were then amplified by PCR with primers 1 and 2 (Fig. 1b). The resulting PCR products were digested with *Kpn*I and *Sac*I, cloned, and sequenced by using primer 5. To enrich for homologous recombination events involving the reacquisition of the *Hin*dIII site, after the first PCR the procedure was repeated (I-*Sce*I digestion and PCR amplification) with internal primers 3 and 4. The resulting products were digested with

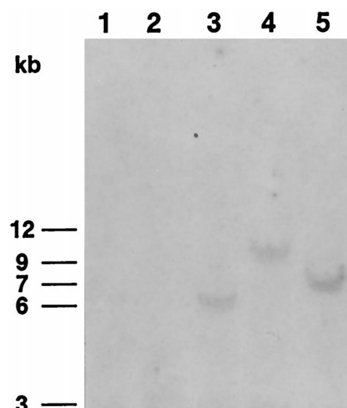

FIG. 2. Southern blot analysis of the *Pst*I-digested genomic DNA isolated from clone 19 (lane 3), clone 22 (lane 4), and clone 25 (lane 5) derived from nod-2 cells transfected with pASB-*Sce*I and p3′SS, and from parental nod-2 cells (lane 1) and nod-2 transfected with the selection vector p3′SS alone (lane 2). The *Nar*I-*Kpn*I polyomavirus origin fragment of pASB-*Sce*I was used as a probe (Fig. 1b).

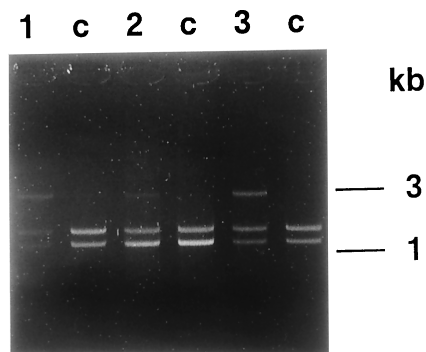

FIG. 3. I-*Sce*I digestion of PCR1 products amplified from *Pst*I-digested genomic DNA isolated from the pCBASce-transfected clones 19 (lane 1), 22 (lane 2), and 25 (lane 3). Lanes c correspond to the same clones subjected to electroporation but in the absence of pCBASce. This control indicates that the genomic DNA from each clone could be digested by I-*Sce*I. The I-*Sce*I digestion of the PCR1 product of 3 kb should yield bands of 1.7 and 1.3 kb. In clones 19 (lane 1) and 25 (lane 3), ca. 35% of the PCR product is resistant to I-*Sce*I digestion. In the case of clone 22 (lane 2), only 20% of the PCR product is resistant to I-*Sce*I digestion. The numbers at the side of the figure indicate the molecular weight.

*Kpn*I-*Hin*dIII, cloned, and sequenced from the *Hin*dIII site, with primers flanking the cloning site and primers 6, 7, and 8 if necessary.

Nature of the repair events. The I-*Sce*I cleavage efficiency was determined for each of the six independent pCBASce transfection assays performed and was found to vary depending on the clone used. After each transfection with the pCBASce vector, the genomic DNA was extracted and digested with *Pst*I, amplified by PCR with primers 1 and 2 (Fig. 1b), digested with I-*Sce*I, and analyzed on a gel. The relative proportion of digested to nondigested DNA was quantified by densitometry. For clones 19 (three assays) and 25 (one assay), ca. 35% of the PCR products were resistant to I-*Sce*I digestion, while for clone 22 (two assays), only ca. 20% were resistant to I-*Sce*I digestion (Fig. 3). Thus, for clones 19 and 25, 35% of the integrated L1 sequences were involved in a repair event, while it appears that for clone 22 only 20% were involved in a repair event. It should be noted that it is not possible with this assay to determine if repair events occurred that would have regenerated the I-*Sce*I site.

The largest number of repair events were analyzed from clone 19. Of 111 repair events defined by the loss of the I-*Sce*I site, 4 had deletions larger than 1 kb and 81 had small deletions (on average, 10 bp), including 6 with additions of small DNA fragments. Another 20 repair events had additions of small DNA fragments without deletion. Thus, at least 105 of the 111 repair events (95%) were the result of end joining. The remaining six repair events involved the acquisition of an *Hin*dIII site replacing the I-*Sce*I inserted sequences. To determine if this was reproducible, PCR products from other assays with clones 19, 22, and 25 were tested for the acquisition of the *Hin*dIII site. As in the initial experiment, *Hin*dIII-acquiring events were also found to represent fewer than 5% of the events in each case, as evaluated by restriction enzyme analysis. The *Hin*dIII-acquiring events were due to DSB repair, since we were not able to isolate *Hin*dIII-acquiring events in controls not transfected with pCBASce.

To enrich for homologous recombination events, 304 *Hin*dIII-acquiring events were cloned and sequenced from six independent pCBASce electroporation assays. From these, 20 had acquired an *Hin*dIII site as a result of a fortuitous deletion of vector sequences around the I-*Sce*I site that generated a de novo *Hin*dIII site or as a result of the acquisition at the I-*Sce*I site of a pCBASce fragment that contained an *Hin*dIII site. Of the remaining 284 *Hin*dIII-acquiring events, the vast majority (252) had sequences identical to that of the L1 sequences present in our vector (or with 1 or 2 random base changes per 300 bp; see below). Since, on average, endogenous mouse L1 elements differ in their sequences from one another by 4% (24) and since a number of polymorphic bases are located near the *Hin*dIII site (Table 1), a gene conversion event might be expected to generally introduce other base changes beside the acquisition of the *Hin*dIII site. Instead, these were most probably generated by end-to-end joining at the two half *Hin*dIII sites bordering the I-*Sce*I cassette (see H′; figure 1b) rather than gene conversion events. In this case, end-joining would occur at the 4-bp H′ repeat, perhaps by a single-strand annealing type of mechanism that uses such microhomologies. Indeed, in our analysis of the 111 PCR products described above, we found single-strand annealing-type events at repeats of 2 to 4 bases at about the same frequency (8 of 111) as the *Hin*dIII-acquiring events (6 of 111). Previous studies had already shown that repair events involving strand annealing of a few bases are a common end-joining repair process (11, 28, 29, 31, 35), including those occurring at short repeats flanking an I-*Sce*I site (37).

The remaining 32 *Hin*dIII-acquiring events appeared to be bona fide gene conversion events with endogenous L1 elements, based on at least two of the following criteria: (i) they had at least one base change beside the acquisition of the *Hin*dIII site; (ii) specific base changes were repeated in different *Hin*dIII-acquiring events originating from different PCR amplifications; (iii) the base changes could be regrouped in conversion tracts of different lengths; (iv) the base changes were skewered toward the *Hin*dIII site; (v) the base changes corresponded to diagnostic bases of various L1 subfamilies.

Since base changes could be produced as an artifact of PCR, we did a control experiment to evaluate the number of base changes that could be attributed to PCR. We spiked nod-2 genomic DNA with 1 copy per genome of the pASB-*Sce*I plasmid and followed the experimental procedure described above. Of 5,225 bp sequenced from 20 different amplified pASB-*Sce*I molecules, we found altogether seven base changes located at random positions. Thus, we excluded from the conversion events infrequent *Hin*dIII-acquiring events that had only one or two base changes that did not repeat themselves in an independent PCR amplification.

Taken together, these results indicate that 5% (6 of 111) of the repair events were *Hin*dIII-acquiring events. From these, 11% (32 of 284) could be identified as bona fide homologous recombination repair events, giving an overall homologous recombination frequency of almost 0.6%. It is possible that we are underestimating the frequency of gene conversion events, since those occurring with endogenous L1 elements without an *Hin*dIII site would not be part of the *Hin*dIII-acquiring events. However, these are apparently infrequent, since none were seen in the 111 events analyzed. Also, events with extremely short conversion tracts, i.e., those that did not extend past the *Hin*dIII site to the first mismatched base between the endogenous L1 element and the pASB-*Sce*I L1 sequence, would not have been scored as a conversion event. Short conversion tracts have been observed in other studies involving the repair of chromosomal DSBs (13, 44). Nevertheless, the proximity and abundance of polymorphic bases near the *Hin*dIII site should have permitted the detection of most events (Table 2).

Gene conversion events with endogenous L1 elements. The conversion tracts of the 32 gene conversion events that were identified in this study are presented in Table 1. The extent of conversion from the *Hin*dIII site that could be scored by the presence of base changes was as short as 13 bp or as long as 780

TABLE 1. List of the 32 gene conversion events of clones 19, 22, and 25[a]

Gene conversion event(s)	Sequence change from pASB-*Sce*I at position (base):																	No. of base changes with pASB-*Sce*I	Minimum conversion tract length
	1064 (A)	1196 (T)	1210 (G)	1231 (T)	1232 (G)	1233 (T)	1252 (T)	1262 (A)	1266 (G)	1283 (A)	1288 (T)	1292 (G)	1294 (C)	1295 (G)	1301 (A)	1305 (T)	1309 (A)		
19-A.01, 19-3.02	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	T	1	13
19-5.09	—	—	—	—	—	—	—	—	—	—	—	—	T	—	—	—	T	2	28
19-9.01	—	—	—	—	—	—	—	—	—	—	C	—	T	—	—	—	T	4	34
19-1.02, 19-2.08, 19-4.06, 19-8.01	—	—	—	—	—	—	—	—	C	—	C	—	T	—	—	—	T	4	56
19-7.08	—	—	—	—	—	—	G	—	C	—	C	—	T	—	—	—	T	5	70
19-0.01	—	—	A	C	A	—	G	—	C	—	C	—	T	—	—	—	T	24	783
19-α.18	—	—	—	—	—	—	—	—	—	—	—	0	—	—	G	—	—	3	37
19-6.05	—	—	—	—	—	—	—	G	—	—	—	—	—	—	—	—	—	3	122
19-6.06	—	—	—	—	—	—	—	G	—	G	—	—	—	—	—	—	—	2	60
19-2.21	—	—	—	—	—	—	—	—	—	G	—	—	—	—	—	—	—	2	140
19-1.03	—	—	—	—	—	—	—	—	—	G	—	—	—	—	—	—	—	1	39
19-6.01	—	C	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2	154
19-2.17	—	C	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	1	126
19-E.05, 19-B.07	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	C	—	1	17
19-A.04, 19-α.06	—	—	—	—	—	—	—	—	—	—	—	—	—	A	—	—	—	1	27
22-1.04	—	—	—	—	—	—	—	—	—	—	—	0	—	—	G	—	—	3	57
22-4.03	G	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2	258
22-9.06	G	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2	258
22-2.02	—	—	—	C	—	—	—	—	—	—	—	—	—	—	—	—	—	1	91
22-H.03	—	—	A	C	A	—	—	—	C	—	C	—	T	—	—	—	T	9	118
25-1.07, 25-5.03	—	—	—	—	—	C	—	—	—	—	—	—	—	—	—	—	—	1	89
25-4.08	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	3	246
25-4.05	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	T	2	133
25-2.19	—	—	—	—	—	—	—	—	—	—	C	—	T	—	—	—	T	4	87
25-3.01	—	—	—	—	—	—	—	—	C	—	C	—	T	—	—	—	T	5	135

[a] The base changes with the pASB-*Sce*I LINE (L1Md-A2) sequence observed in at least two distinct gene conversion events are indicated. The positioning of the base changes refers to Fig. 1b. The "0" indicates that there is no base at the corresponding position. For each gene conversion event, the first number identifies the clone from which it was isolated; the second number (first pCBASce electroporation) or letter (second electroporation) or symbol (third electroporation experiment) identifies the individual independent PCR, and the third number identifies the specific event. The length of the conversion tract is delineated by the farthest base change from the *Hin*dIII site (positions 1316 to 1321). —, No change.

bp, with the majority being shorter than 100 bp. The conversion tracts occurred between L1 sequences that were as much as 10% heterologous, with most being less than 4%.

For each clone, several different patterns of gene conversion were seen that involved distinct endogenous L1 partners (11 of 21 for clone 19, 5 of 5 for clone 22, and 5 of 6 for clone 25). In at least one instance, it appeared that several independent gene conversion events involved the same endogenous L1 partner, resulting in a set of nested conversion tracts (the first 10 gene conversion events for clone 19). Interestingly, the longest conversion tract of this cohort (783 bp, event 19-0.01), which has incorporated 24 base changes compared to pASB-*Sce*I, is 100% identical to the sequence of a L1 element that has recently been reported as actively being involved in retrotransposition events ($L1_{spa}$, AF016099 GenBank/EMBL accession number) (27).

Because their sequences are identical and they come from the same pCBASce electroporation assay (but not the same PCR amplification), gene conversion events 19-1.02, 19-2.08, 19.4.06, and 19-8.01 could represent the same clonal event, as could be the case for events 19-E.05 and 19-B.07 or events 25-1.07 and 25-5.03.

Comparative analysis of mouse L1 elements indicate that they can be divided into at least three groups or families of more-related sequences that have been designated the F, A, and T_F families (10). The F family would be the oldest and most diverged one, while the A and T_F families would be more recent, with a higher degree of homology among their members (97 and 99.8%, respectively). The T_F family would possibly be the most active in retrotransposition (27). The L1 sequences present in pASB-*Sce*I are derived from L1Md-A2 (23), which is part of an A subfamily in which members share 99.5% of homology and are also active in retrotransposition (41). The T_F family and the L1Md-A2 subfamily have been estimated to contain the same number of full-length members, i.e., 2,000 to 3,000 per diploid genome (40). The L1 segments that were sequenced in this study include a number of diagnostic bases that are identified with the L1Md-A2 subfamily or T_F family. These are indicated in Table 2. Analysis of the gene conversion tracts we observed indicated that the majority of events can be classified as possibly having involved a L1Md-A2 subfamily partner (17 events) or a T_F partner (10 events). Five events have a mixed T_F and A subfamily conversion tract that could have arisen from either mismatch repair or recombination with a mosaic T_F/A L1 partner (40).

Gene conversion event 22-H.03 has an interesting structure (Fig. 4a) that suggests a one-sided invasion homologous recombination event. This event has an 118-bp L1 duplication containing 9-bp changes predominantly of the T_F pattern, and a deletion of 12 bases from the I-*Sce*I cassette. Figure 4b illustrates the different steps that have possibly been involved in the genesis of event 22-H.03, based on the one-sided homologous recombination model (4, 5). The steps are invasion of the L1 donor sequence by only one side of the double-strand break, followed by DNA synthesis and then release of the invading end, and end-to-end joining with the other side of the DSB. This had also been observed in earlier studies on homologous recombination between extrachromosomal and endogenous chromosomal L1 elements in mouse cells (4, 6).

DISCUSSION

In the present study, we have shown that a DSB in LINE elements could be repaired by homologous recombination with

TABLE 2. Categorization of the 32 gene conversion events with respect to the predominant L1 polymorphic bases from the *Hin*dIII site[a]

Family (subfamily) pattern and event(s)	Sequence change at position:											
	1048	1051	1119	1180	1210	1231	1232	1252	1266	1288	1294	1309
L1Md-A2 subfamily	A	T	T	C	G	T	G	T	G	T	C	A
22-9.06, 22-4.03, 25-4.08	—	—	T	C	G	T	G	T	G	T	C	A
19-6.01	—	—	—	C	G	T	G	T	G	T	C	A
19-6.05, 19-2.21, 19-2.17	—	—	—	—	G	T	G	T	G	T	C	A
25-1.07, 25-5.03	—	—	—	—	—	—	—	T	G	T	C	A
19-6.06, 22-1.04	—	—	—	—	—	—	—	—	G	T	C	A
19-α.18, 19-1.03	—	—	—	—	—	—	—	—	—	T	C	A
19-E.05, 19-B.07, 19-A.04, 19-α.06	—	—	—	—	—	—	—	—	—	—	—	A
T_F family	g	c	c	t	a	c	a	g	c	c	t	t
19-0.01	g	c	c	t	a	c	a	g	c	c	t	t
19-7.08	—	—	—	—	—	—	—	g	c	c	t	t
19-1.02, 19-2.08, 19-4.06, 19-8.01	—	—	—	—	—	—	—	—	c	c	t	t
19-9.01	—	—	—	—	—	—	—	—	—	c	t	t
19-5.09	—	—	—	—	—	—	—	—	—	—	t	t
19-A.01, 19-3.02	—	—	—	—	—	—	—	—	—	—	—	t
Mixed L1Md-A2/T_F												
22-0.02	—	—	—	—	—	c	G	T	G	T	C	A
25-4.05	—	—	—	—	G	T	G	T	G	T	C	t
25-2.19	—	—	—	—	—	—	—	T	G	c	t	t
25-3.01	—	—	—	—	G	T	G	T	c	c	t	t
22-H.03	—	—	—	—	a	c	a	T	c	c	t	t

[a] For the 12 most polymorphic L1 bases, the bases are specified for the L1Md-A2 subfamily (in capital letters) and T_F family (in lowercase letters) and the 32 gene conversion events for their respective conversion tracts (Table 1). Highly polymorphic L1 bases were determined from an alignment of 20 GenBank-EMBL *Hin*dIII-containing mouse L1 elements (accession numbers AC002315, AC002406, AC003019, AE000663, AF016099, AF021335, AF037352, D84391, K00582, K00587, K00588, M13002, M11515, M23236, M29324, M29325, M68842, MMMHC29N7, U15647, and X14061). —, No change.

nonallelic endogenous L1 elements. To our knowledge, this is the first demonstration of homologous repair of a chromosomal DSB by endogenous genomic repeat sequences in mammal cells. However, repair of a DSB in Ty elements by nonallelic Ty chromosomal sequences has been reported in yeast cells (30). We found that the dominant repair mechanism leading to I-*Sce*I site loss was by far end joining rather than gene conversion with an endogenous L1 element. This contrasts with yeast, in which nonallelic repetitive elements are readily used (30). Nevertheless, considering the large number of repetitive elements, a frequency of almost 1% DSB repair by homologous recombination with nonallelic L1 elements could be enough to have played a role in the evolution of repetitive elements. In this regard, it is interesting that the selection of partners was not based primarily on the degree of homology, since partners with higher (L1Md-A2 subfamily) and lower (T_F family) homology appeared to have participated equally in the process. A lack of preference for highly homologous partners had also been seen for homologous recombination between extrachromosomal and chromosomal L1 elements (6, 33). This could explain the genesis of mosaic lineages in L1 elements (40).

The choice of partners for DSB repair by homologous recombination could be more a question of accessibility than homology. We do not know the relative location of the endogenous L1 partners relative to our pASB-*Sce*I L1 element. It could have been intrachromosomal proximal or distal, or interchromosomal, or even extrachromosomal. This last possibility seems less likely since what appears to be the same partner was in some cases involved repeatedly in the process, resulting in a set of encompassing conversion tracts (see clone 19, Table 1). In any case, several partners were clearly accessible for DSB repair.

It would seem that the most likely endogenous L1 partners were members of the L1-MdA2 subfamily or the T_F family. These represent only a fraction of the total L1 elements of the mouse genome. The full-length members in these two groups have been estimated at 2,000 to 3,000 for each (40). However, they are by far the most active in retrotransposition (10, 27, 40, 41). It is tempting to speculate that this could somehow influence their accessibility. In particular, we observed for one en-

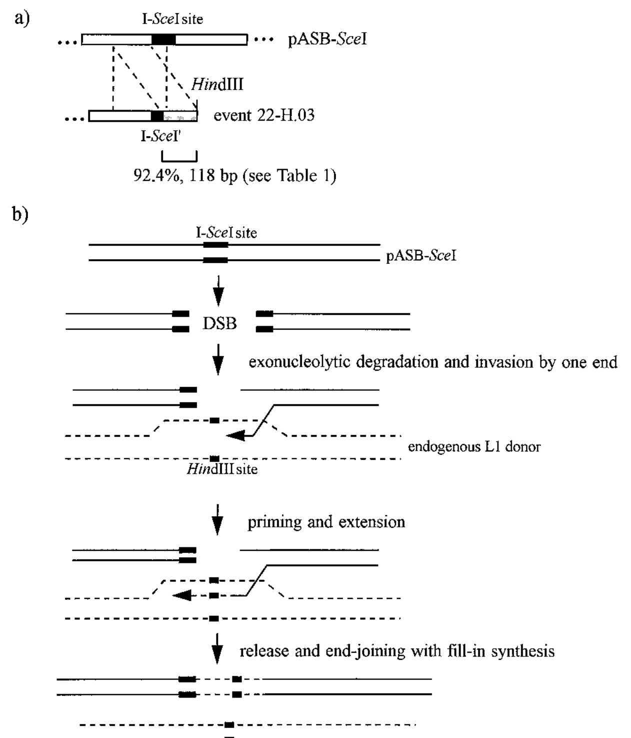

FIG. 4. Proposed mechanism (b) for the production of gene conversion event 22-H.03 (a).

dogenous L1 partner a 100% homology over 800 bp with a T_F element, $L1_{spa}$, that has been shown to be currently involved in retrotransposition (27).

In one case, we observed that the DSB repair involved a gene conversion event with a structure compatible with a one-sided invasion (OSI) mechanism, as we had described previously (4). OSI appears to be a universal homologous recombination mechanism seen in mammals (1, 4, 14, 26, 34), yeasts (15), plants (32), insects (46), and birds (25). It can lead to the transfer of several kilobases of nonhomologous sequences between nonallelic homologous sites (2, 4, 15). Thus, OSI between L1 elements could have had an impact on the evolution of the mammalian genome, permitting the transfer of unique sequences from one location to another via repetitive elements that would provide numerous entry points. In this regard, it is interesting to note that we have readily detected homologous recombination between nonallelic L1 elements in germ cells of transgenic mice (A. Tremblay et al., unpublished data).

ACKNOWLEDGMENTS

We gratefully acknowledge Dominic Corbeil, Hélène D'anjou, Julie Guay, Nadine Gusew, Benoit Houle, Roxane Lussier, and Hugo Würtele for their technical assistance. We also thank Anne-Marie Mes-Masson for her interest in this work and for providing the nod-2 cells.

This study has been supported by a grant from the Medical Research Council (MRC) of Canada. A.T. is a recipient of an MRC fellowship.

REFERENCES

1. **Adair, G. M., R. S. Nairn, J. H. Wilson, M. M. Seidman, K. A. Brotherman, C. Mackinnon, and J. B. Scheerer.** 1989. Targeted homologous recombination at the adenine phosphoribosyltransferase locus in Chinese hamster cells. Proc. Natl Acad. Sci. USA **86:**4574–4578.
2. **Aratani, Y., R. Okazaki, and H. Koyama.** 1992. End extension repair of introduced targeting vectors mediated by homologous recombination in mammalian cells. Nucleic Acids Res. **20:**4795–4801.
3. **Ausubel, F. M., K. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1992. Current protocols in molecular biology. John Wiley and Sons, New York, N.Y.
4. **Belmaaza, A., J. C. Wallenburg, S. Brouillette, N. Gusew, and P. Chartrand.** 1990. Genetic exchange between endogenous and exogenous LINE-1 repetitive elements in mouse cells. Nucleic Acids Res. **18:**6385–6391.
5. **Belmaaza, A., and P. Chartrand.** 1994. One-sided invasion events in homologous recombination at double-strand breaks. Mutat. Res. **314:**199–208.
6. **Belmaaza, A., E. Milot, J.-F. Villemure, and P. Chartrand.** 1994. Interference of DNA sequence divergence with precise recombinational DNA repair in mammalian cells. EMBO J. **13:**5355–5360.
7. **Brenneman, M., F. S. Gimble, and J. H. Wilson.** 1996. Stimulation of intrachromosomal homologous recombination in human cells by electroporation with site-specific endonucleases. Proc. Natl Acad. Sci. USA **93:**3608–3612.
8. **Brown, S. D. M., and G. Dover.** 1981. Organization and evolutionary progress of a dispersed repetitive family of sequences in widely separated rodent genomes. J. Mol. Biol. **150:**441–466.
9. **Burwinkel, B., and M. W. Kilimann.** 1998. Unequal homologous recombination between LINE-1 elements as a mutational mechanism in human. J. Mol. Biol. **277:**513–517.
10. **DeBerardinis, R. J., J. L. Goodier, E. M. Ostertag, and H. H. Kazazian, Jr.** 1998. Rapid amplification of a retrotransposon subfamily is evolving in the mouse genome. Nat. Genet. **20:**288–290.
11. **Derbyshire, M. K., L. H. Epstein, C. S. H. Young, P. L. Munz, and R. Fishel.** 1994. Nonhomologous recombination in human cells. Mol. Cell. Biol. **14:**156–169.
12. **Donoho, G., M. Jasin, and P. Berg.** 1998. Analysis of gene targeting and intrachromosomal homologous recombination stimulated by genomic double-strand breaks in mouse embryonic stem cells. Mol. Cell. Biol. **18:**4070–4078.
13. **Elliott, B., C. Richardson, J. Winderbaum, J. A. Nickoloff, and M. Jasin.** 1998. Gene conversion tracts from double-strand break repair in mammalian cells. Mol. Cell. Biol. **18:**93–101.
14. **Ellis, J., and A. Bernstein.** 1989. Gene targeting with retroviral vectors, recombination by gene conversion into regions of nonhomology. Mol. Cell. Biol. **9:**1621–1627.
15. **Haber, J. E.** 1992. Mating-type gene switching in *Saccharomyces cerevisiae*. Trends Genet. **8:**446–452.
16. **Hendrickson, E. A.** 1997. Insights from model systems: cell-cycle regulation of mammalian DNA double-strand-break repair. Am. J. Hum. Genet. **61:**795–800.
17. **Huang, L.-C., K. C. Clarkin, and G. M. Wahl.** 1996. Sensitivity and selectivity of the DNA damage sensor responsible for activating p53-dependent G_1 arrest. Proc. Natl Acad. Sci. USA **93:**4827–4832.
18. **Hutchison, C. A., III, S. C. Hardies, D. D. Loeb, W. R. Shehee, and M. H. Edgell.** 1989. LINEs and related retroposons: long interspersed repeated sequences in the eucaryotic genome, p. 593–617. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
19. **Jasin, M.** 1996. Genetic manipulation of genomes with rare-cutting endonucleases. Trends Genet. **12:**224–228.
20. **Lebel, M., M. Webster, W. J. Muller, A. Royal, J. Gauthier, and A.-M. Mes-Masson.** 1995. Transgenic mice bearing the polyomavirus large T antigen directed by 2,1 kb of the keratin 19 promoter develop bronchiolar papillary tumors with progression to lung adenocarcinomas. Cell Growth Differ. **6:**1591–1600.
21. **Liang, F., P. J. Romanienko, D. T. Weaver, P. A. Jeggo, and M. Jasin.** 1996. Chromosomal double-strand break repair in Ku80-deficient cells. Proc. Natl Acad. Sci. USA **93:**8929–8933.
22. **Liang, F., M. Han, P. J. Romanienko, and M. Jasin.** 1998. Homology-directed repair is a major double-strand break repair pathway in mammalian cells. Proc. Natl Acad. Sci. USA **95:**5172–5177.
23. **Loeb, D. D., R. W. Padgett, S. C. Hardies, W. R. Shehee, M. B. Comer, M. H. Edgell, and C. A. Hutchison III.** 1986. The sequence of a large L1md element reveals a tandemly repeated 5′ end and several features found in retrotransposons. Mol. Cell. Biol. **6:**168–182.
24. **Martin, S. L., C. F. Voliva, S. C. Hardies, M. H. Edgell, and C. A. Hutchison III.** 1985. Tempo and mode of concerted evolution in the L1 repeat family of mice. Mol. Biol. Evol. **2:**127–140.
25. **McCormack, W. T., and C. B. Thompson.** 1990. Chicken Ig_L variable region gene conversions display pseudogene donor preference and 5′ to 3′ polarity. Genes Dev. **4:**548–558.
26. **Mudgett, J. S., and W. D. Taylor.** 1990. Recombination between irradiated shuttle vector DNA and chromosomal DNA in African green monkey kidney cells. Mol. Cell. Biol. **10:**37–46.
27. **Nass, T. P., R. J. Deberardinis, J. V. Moran, E. M. Ostertag, S. F. Kingsmore, M. F. Seldin, Y. Hayashizaki, S. L. Martin, and H. H. Kazazian.** 1998. An actively retrotransposing, novel subfamily of mouse L1 elements. EMBO J. **17:**590–597.
28. **Nicolas, A. P., and C. S. H. Young.** 1994. Characterization of DNA end joining in a mammalian cell nuclear extract: junction formation is accompanied by nucleotide loss, which is limited and uniform but not site specific. Mol. Cell. Biol. **14:**170–180.
29. **North, P., A. Ganesh, and J. Thacker.** 1990. The rejoining of double-strand breaks in DNA by human cell extracts. Nucleic Acids Res. **18:**6205–6210.
30. **Parket, A., O. Inbar, and M. Kupiec.** 1995. Recombination of Ty elements in yeast can be induced by a double-strand break. Genetics **140:**67–77.
31. **Pfeiffer, P., and W. Vielmetter.** 1988. Joining of nonhomologous DNA double strand breaks in vitro. Nucleic Acis Res. **16:**907–924.
32. **Puchta, H., B. Dujon, and B. Hohn.** 1996. Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. Proc. Natl Acad. Sci. USA **93:**5055–5060.
33. **Richard, M., A. Belmaaza, N. Gusew, J. C. Wallenburg, and P. Chartrand.** 1994. Integration of a vector containing a repetitive LINE-1 element in the human genome. Mol. Cell. Biol. **14:**6689–6695.
34. **Richardson, C., M. E. Moynahan, and M. Jasin.** 1998. Double-strand break repair by interchromosomal recombination: suppression of chromosomal translocations. Genes Dev. **12:**3831–3842.
35. **Roth, D. B., and J. H. Wilson.** 1986. Nonhomologous recombination in mammalian cells: role for short sequence homologies in the joining reaction. Mol. Cell. Biol. **6:**4295–4304.
36. **Rouet, P., F. Smith, and M. Jasin.** 1994. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. Proc. Natl Acad. Sci. USA **91:**6064–6068.
37. **Rouet, P., F. Smith, and M. Jasin.** 1994. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. Mol. Cell. Biol. **14:**8096–8106.
38. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl Acad. Sci. USA **74:**5463–5467.
39. **Sargent, R. G., M. A. Brenneman, and J. H. Wilson.** 1997. Repair of site-specific double-strand breaks in a mammalian chromosome by homologous and illegitimate recombination. Mol. Cell. Biol. **17:**267–277.
40. **Saxton, J. A., and S. L. Martin.** 1998. Recombination between subtypes creates a mosaic lineage of LINE-1 that is expressed and actively retrotransposing in the mouse genome. J. Mol. Biol. **280:**611–622.
41. **Schichman, S. A., D. M. Severynse, M. H. Edgell, and C. A. Hutchinson III.** 1992. Strand-specific LINE-1 transcription in mouse F9 cells originates from the youngest phylogenetic subgroup of LINE-1 elements. J. Mol. Biol. **224:**559–574.
42. **Schwartz, A., D. C. Chan, L. G. Brown, R. Alagappan, D. Pettay, C. Disteche, B. McGillivray, A. de la Chapelle, and D. C. Page.** 1998. Reconstructing hominid Y evolution: X-homologous block, created by X-Y transposition,

was disrupted by Yp inversion through LINE-LINE recombination. Hum. Mol. Genet. **7:**1–11.

43. **Segal, Y., B. Peissel, A. Renieri, M. de Marchi, A. Ballabio, Y. Pei, and J. Zhou.** 1999. LINE-1 elements at the sites of molecular rearrangements in Alport syndrome-diffuse leiomyomatosis. Am. J. Hum. Genet. **64:**62–69.
44. **Taghian, D. G., and J. A. Nickoloff.** 1997. Chromosomal double-strand breaks induce gene conversion at high frequency in mammalian cells. Mol. Cell. Biol. **17:**6386–6393.
45. **Vamvakas, S., E. H. Vock, and W. K. Lutz.** 1997. On the role of DNA double-strand breaks in toxicity and carcinogenesis. Crit. Rev. Toxicol. **27:** 155–174.
46. **Van der Ploeg, L. H. T., K. Gottesdien, and M. G.-S. Lee.** 1992. Antigenic variation in the African trypanosome. Trends Genet. **8:**452–456.
47. **Voliva, C. F., C. L. Jahn, M. B. Comer, C. A. Hutchison III, and M. H. Edgell.** 1983. The L1Md long interspersed repeat family in the mouse: almost all examples are truncated at one end. Nucleic Acids Res. **11:**8847–8859.