

# Circular retrotransposition products generated by a LINE retrotransposon

Jeffrey S. Han\* and Shirley Shao

Department of Embryology, Carnegie Institution for Science, Baltimore, MD 21218, USA

Received July 24, 2012; Revised August 20, 2012; Accepted August 21, 2012

## ABSTRACT

**Non-long terminal repeat (non-LTR) retrotransposons are highly abundant elements that are present in chromosomes throughout the eukaryotic domain of life. The long interspersed nuclear element (LINE-1) (L1) clade of non-LTR retrotransposons has been particularly successful in mammals, accounting for 30–40% of human genome sequence. The current model of LINE retrotransposition, target-primed reverse transcription, culminates in a chromosomally integrated end product. Using a budding yeast model of non-LTR retrotransposition, we show that in addition to producing these ‘classical’, chromosomally integrated products, a fungal L1 clade member (Zorro3) can generate abundant, RNA-derived episomal products. Genetic evidence suggests that these products are likely to be formed via a variation of target-primed reverse transcription. These episomal products are a previously unseen alternative fate of LINE retrotransposition, and may represent an unexpected source for *de novo* retrotransposition.**

## INTRODUCTION

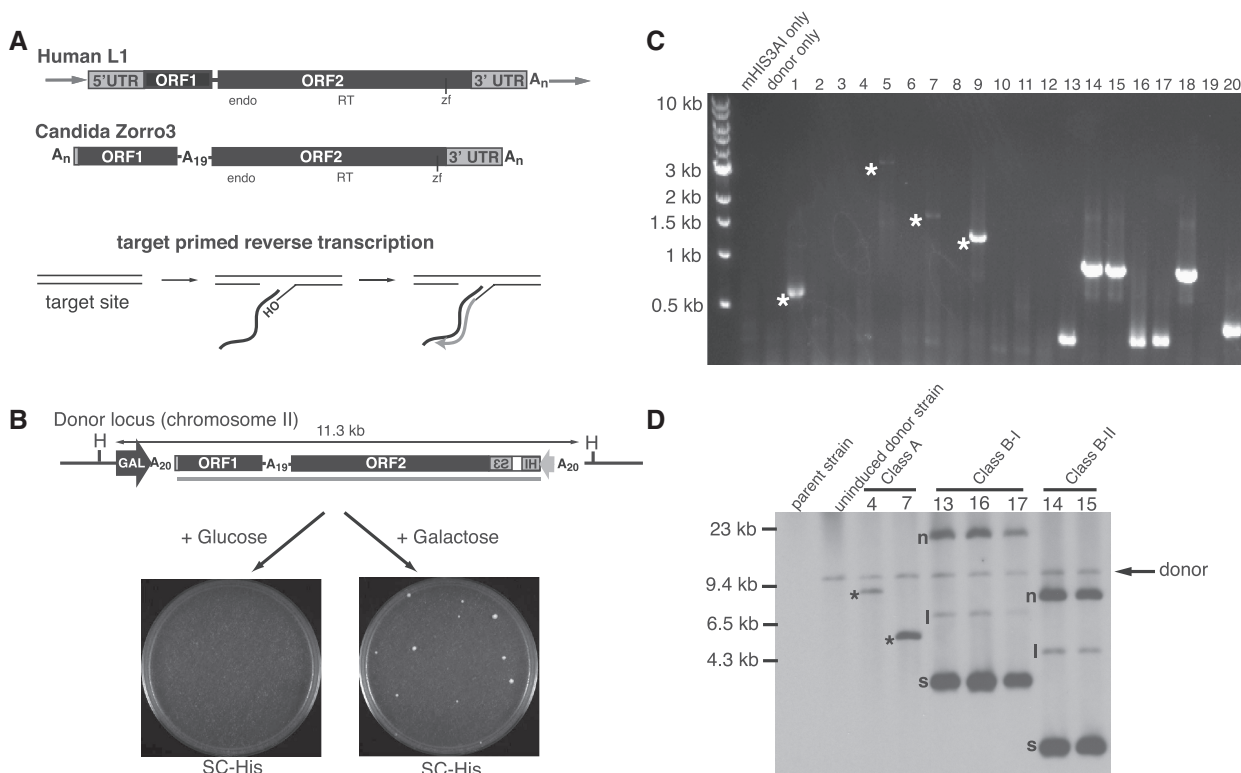
Non-long terminal repeat (non-LTR) retrotransposons are highly abundant elements that are present in chromosomes throughout the eukaryotic domain of life. Members of the L1 clade of non-LTR retrotransposons are of particular interest, as they have amplified to high copy number in mammalian genomes. Sequencing projects have shown that 15–30% of mammalian genomes comprise long interspersed nuclear element (LINE) sequence (1–8). An additional 8–20% of mammalian sequence is derived from non-LINE sequences (e.g. Alu elements, SVA elements and processed pseudogenes) that parasitize LINE machinery to retrotranspose (9). The process of L1 retrotransposition is still occurring in humans today and is

a source of DNA damage (10,11), *de novo* mutations (12,13) and genetic variability (14–20).

Most available information on the biology of L1 clade members is derived from experiments with the mammalian L1 elements. An L1 element produces a bicistronic transcript encoding two proteins, ORF1 and ORF2 (Figure 1A). ORF1 protein binds to single-stranded nucleic acids with high affinity and possesses nucleic acid chaperone activity (21–25). ORF2 encodes endonuclease (26) and reverse transcriptase activities (27–30). All of these activities are important for retrotransposition (26,31–33). Complete mechanistic details of how LINE elements retrotranspose are unclear; however, a model of the initial steps of reverse transcription, called target-primed reverse transcription (TPRT), has emerged from biochemical experiments with the related R2 element of *Bombyx mori* (34). Later experiments with human L1 support this TPRT model (35), in which the L1 ribonucleoprotein particle (RNP), consisting of L1 RNA, ORF1 and ORF2, enters the nucleus and nicks a chromosomal target site. Reverse transcription ensues using L1 RNA as the template and the free 3' end at the chromosomal nick as the primer (Figure 1A). Thus, L1 cDNA synthesis and chromosomal attachment is concurrent (in contrast to retroviral and LTR-retrotransposon cDNA synthesis). Completion of the insertion involves a second nick on the opposite DNA strand of the target site, which serves as the primer for plus strand DNA synthesis. Consistent with this mechanism, all characterized products of authentic L1 retrotransposition to date have been chromosomally integrated (26,33,36–40).

The *Candida albicans* genome contains an active yeast non-LTR retrotransposon called Zorro3 (41,42). Phylogenetic analysis of the Zorro3 reverse transcriptase domain places Zorro3 in the L1 clade (41). Genomic Zorro3 elements tend to be flanked by polyadenosine [poly(A)] tracts (Figure 1A). The poly(A) flanks are likely derived from three sources—the target site [Zorro3 tends to integrate into poly(A) tracts], a possible target site duplication [if both endonuclease nicks are within the target site poly(A) tract], and the mRNA poly(A) tail.

\*To whom correspondence should be addressed. Tel: +1 410 246 3055; Fax: 410 243 6311; Email: han@ciwemb.edu



**Figure 1.** A novel class of LINE retrotransposon products. **(A)** Structure of L1 clade members and proposed mechanism of non-LTR retrotransposon replication. Top, comparison of the structures of human L1 and *Candida albicans* Zorro3. Both elements contain two ORFs and endonuclease (endo), reverse transcriptase (RT) and zinc finger (zf) domains. Genomic L1s are typically flanked by target site duplications (flanking arrows). Genomic Zorro3s are flanked by poly(A) tracts, which may include poly(A) target site duplications. Bottom, general mechanism for the initiation of non-LTR retrotransposon replication. **(B)** Assay to induce Zorro3 retrotransposition. A donor Zorro3 marked with a *HIS3AI* (47) retrotransposition cassette (Zorro3mHIS3AI) is present at the *LYS2* locus of chromosome II. To induce retrotransposition, cells are grown on galactose and then plated on SC-His to select retrotransposition events. H = *HaeII*. Bar below Zorro3 represents the probe used for Southern blots in all subsequent figures. **(C)** Example of LM-PCR products. Each lane represents the LM-PCR product from an independent Zorro3 retrotransposition event. Owing to variable *EcoRI* distances from the Zorro3 3' end, Class A LM-PCR products are not always amplified. 'mHIS3AI only' is an LM-PCR reaction from a strain containing only the mHIS3AI cassette. 'Donor only' is an LM-PCR reaction from a strain containing Zorro3mHIS3AI, but not a retrotransposition event. Class A products are noted with an asterisk. **(D)** Southern analysis of class A and class B events. Genomic DNA from the indicated clones was digested with *HaeII* and probed for Zorro3 sequence. Arrow indicates the band representing the original Zorro3mHIS3AI donor. \* = class A insertions. n = nicked. l = linear. s = supercoiled.

As these sequences are identical homopolymers, we cannot unambiguously determine the source for all of the nucleosides in the poly(A) tracts. Another distinguishing feature between Zorro3 and mammalian L1 elements is that the former contains a type I ORF1, whereas the latter contains a type II ORF1 (43). Type I ORF1s contain CCHC zinc knuckle motifs instead of the conserved mammalian C-terminal domain, and are found in at least 5 different non-LTR retrotransposon clades (including non-mammalian members of the L1 clade), indicating their ancient origin (43). We previously redesigned and synthesized an *Saccharomyces cerevisiae*-compatible version of Zorro3 (40). We found that Zorro3 is able to retrotranspose in *S. cerevisiae* and maintains the requirement for ORF1 and ORF2 motifs. In addition, chromosomally integrated Zorro3 retrotransposition events revealed structural similarities to L1 integrations in human and mouse, suggesting that Zorro3 retrotransposition in *S. cerevisiae* is mechanistically similar to mammalian L1 retrotransposition. Here, we describe a previously unknown second class (class B) of

retrotransposition products from Zorro3. Rather than being chromosomally integrated, class B products are RNA-derived DNA circles that represent the majority of retrotransposition products formed by Zorro3. This is the first time LINE elements have been shown to produce extrachromosomal products from the retrotransposition reaction. These episomal products may be a previously unrecognized side product or an intermediary source of active LINE retrotransposition.

## MATERIALS AND METHODS

### Yeast strains

The yeast strain used in Figures 1 through 3 (JHY 339) has been already described (40). The Zorro3pA<sup>-</sup> strain, JHY343, was generated by converting wt Zorro3 to Zorro3pA<sup>-</sup> by two-step gene replacement using pRS406-derived plasmids, as previously described (40). To test nucleic acid enzyme mutants for effects on class B frequency, we used the strain JHY869 to knock out

complete open reading frames of desired genes by PCR-mediated disruption (44). JHY869 contains the Zorro3mHIS3AI and *His3Δ200* alleles backcrossed four times to YAD373 (45), to generate a retrotransposition strain with an s288c background with high sporulation frequency. Our laboratory has since switched to this strain for all Zorro3 retrotransposition studies because it is isogenic with the sequenced yeast genome and the yeast knockout collection, and it contains three alleles [RME1(ins-308 A), TAO3(E1493Q) and MKT1(D30G)] that allow it to sporulate with high efficiency for easier genetic manipulation. All yeast strains and genotypes are listed in Supplementary Table S1.

### Primers

All primers in this study are listed in Supplementary Table S2.

### Southern analysis

Genomic DNA was prepared as described previously (46). One microgram of genomic DNA was digested with the appropriate restriction enzyme, run on a 0.8% agarose gel, and blotted in 0.4 M NaOH to positively-charged nylon membrane (Millipore). Hybridizations were performed in Ultrahyb (Ambion) at 42°C. Riboprobes were biotin-16-UTP labelled T7 *in vitro* transcription products of JH328/JH329 PCR products (Zorro3mHIS3 probe) or CD15/CD16 PCR products (tubulin probe). Detection was performed with the Phototope-star detection kit (New England Biolabs).

### Class B product stability assays

A His<sup>+</sup> colony from a SC-His plate was patched onto a yeast extract peptone dextrose (YPD) plate and incubated at 30°C for 24 hours. Cells from this YPD plate were used to inoculate a 2 mL YPD culture and incubated at 30°C for 24 h. Serial dilutions of this culture were plated on YPD plates, and colonies were allowed to form for 2–3 days at 30°C. Plates with ~100–300 colonies were replica plated to both SC-His plates and new YPD plates. Percentage stability is defined as the number of colonies that grow on the SC-His plate divided by the number of colonies that grow on the YPD plate.

### Zorro3 retrotransposition assays and colony PCR

Quantitative Zorro3 retrotransposition assays were performed essentially as described previously (40). When isolating retrotransposition events to assay for class B products, induction of retrotransposition was performed by patching strains on SC+galactose plates (instead of cultures) to ensure that each His<sup>+</sup> colony was an independent event. The induction plates were replica plated to SC-His to select for His<sup>+</sup> events. Colony PCR (cPCR) was performed by inoculating 2.5 μL of H<sub>2</sub>O. This mixture was subjected to two freeze-thaw cycles, then used in a 20 μL PCR reaction containing 500 nM each primer, 200 nM dNTPs and 0.25 units ExTaq (Takara). JH774/JH775 amplifies class B-I, and JH774/JH776 amplifies class B-II. Conditions were 94°C for 4 min, 42 cycles of

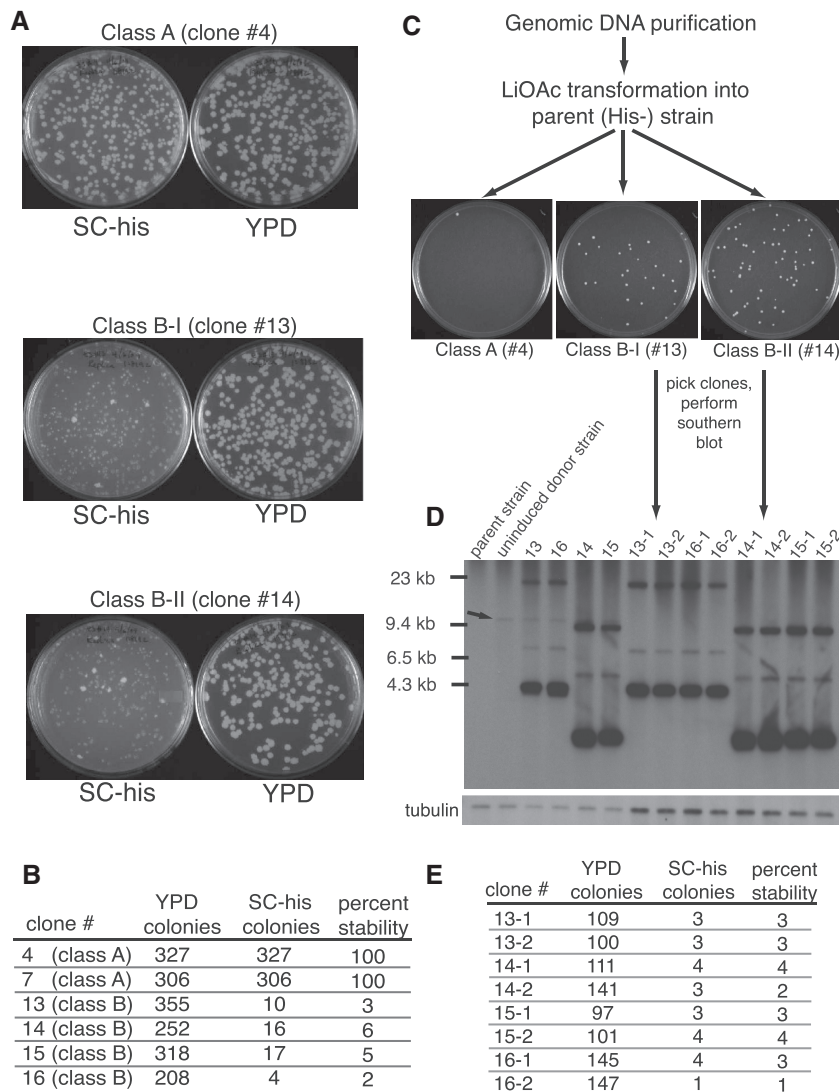
94°C for 15 s, 56°C for 15 s, 72°C for 1 min, followed by a 2-min 72°C extension. Class B events were cloned into a standard TA vector for sequencing.

## RESULTS

### An unusual class of Zorro3 retrotransposition products

We used a previously described assay (Figure 1B) to detect Zorro3 retrotransposition events in budding yeast as His<sup>+</sup> cells (40). This assay depends on the splicing of a synthetic intron out of a reverse reporter (47) placed in the Zorro3 3' UTR. A functional HIS3 reporter is made only after Zorro3 transcription, splicing and reverse transcription/integration. Zorro3 induction produced cells that formed colonies of heterogeneous sizes on SC-His plates, with a large proportion of small colonies. Ligation-mediated PCR (LM-PCR) of 3' flanking sequence from independently derived His<sup>+</sup> clones revealed two main classes of products (schematic in Supplementary Figure S1). The first class was typically present in the biggest colonies and generated LM-PCR products of various sizes (e.g. Figure 1C, lanes 1, 5, 7, 9). This class represents chromosomally integrated retrotransposition events, which are described in detail elsewhere (40). We will henceforth refer to these classic, chromosomally integrated products as 'class A' products. The second class of retrotransposition events was present in some of the large colonies and almost all of the small colonies, and gave rise to LM-PCR products with characteristic band sizes of either ~400 base pairs or ~1000 base pairs (Figure 1C, lanes 13–18, 20). Sequencing revealed that the 3' flanking sequence in these PCR products was either Zorro3 5'UTR/ORF1 or Zorro3 interORF/ORF2, respectively (Supplementary Figure S1). We classified these retrotransposition events as 'class B', and further distinguished them as 'class B-I' (3' flanked by ORF1) or 'class B-II' (3' flanked by ORF2).

Zorro3 has a preference to integrate into chromosomal poly(A) tracts (40,42). As the donor Zorro3 on chromosome II contains poly(A) tracts at the 5' and interORF regions (Figure 1B), we hypothesized that class B products were simply integrants at these locations. To test this hypothesis, we performed Southern analysis on HaeII-digested genomic DNA from strains containing class A, B-I and B-II products (Figure 1D), using Zorro3mHIS3 sequence as a probe. As expected, class A products (clones 4 and 7) were integrated at a new site, generating a band distinct from the original donor band. We predicted that if class B products were integrated into the donor Zorro3, the original donor band would increase in size. However, the donor band was unaltered in class B strains, indicating that class B products are not integrated into the donor locus. Instead, class B products appeared as three new bands in addition to the donor band (Figure 1D, clones 13–17). Within each subclass (B-I or B-II), the banding pattern was virtually identical. Because the His<sup>+</sup> clones used for Southern analysis were produced independently, this suggests that class B-I and B-II products are specific structures that are reproducibly generated on induction of Zorro3 retrotransposition.



**Figure 2.** Instability and transfer of class B products between strains. (A) Representative examples of class A and class B product stability. (B) Quantitation of the experiment described in panel A. (C) Transfer of class B product between strains. Total genomic DNA purified from strains containing a class A or class B insertion was used to transform His<sup>-</sup> yeast cells without pre-existing Zorro3 sequence. Transformations were plated on SC-His. His<sup>+</sup> colonies represent successful maintenance of the retrotransposition product under selection. (D) Class B transformants harbour the class B episome. Genomic DNA from His<sup>+</sup> transformants described in panel C was digested with HaeII and subjected to Southern analysis. The donor (arrow) and class B products are both present in the original class B strains (13–16), but only the class B product is transferred to the transformants (13-1, 13-2, etc.). (E) Class B products transferred to new strains exhibit instability. His<sup>+</sup> stability testing described in panels A and B was repeated with class B transformants generated in panel C.

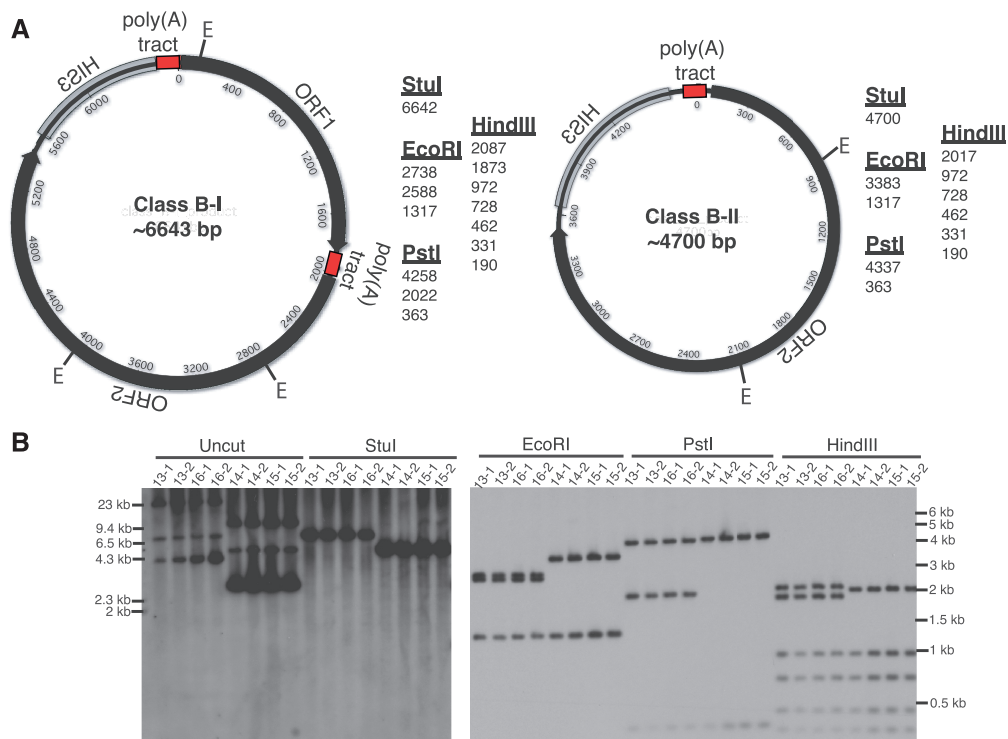
The three bands representing class B products bear a striking resemblance to the three forms of DNA (nicked, linear and supercoiled) often seen in plasmid preparations, and the intensity of the bands relative to the donor band suggests they are present in multiple copies. Thus, we next sought to investigate whether class B products were DNA circles.

### Class B products are DNA circles

As class A products are integrated into a host chromosome, they are expected to be faithfully replicated and partitioned to daughter cells along with the rest of the chromosome. Episomes, on the other hand, may

eventually be lost from a dividing population, unless they contain a centromere or other partitioning system. To test whether class B products are stably transmitted to daughter cells, we grew class B, His<sup>+</sup> clones under non-selective conditions for 2 days. The resultant population quickly became predominantly His<sup>-</sup> (Figure 2A and B). In contrast, class A clones remained 100% His<sup>+</sup> after non-selective growth. These results are consistent with hypothesis that class B products are episomal circles.

We next tested the transformation ability of class B products. We transformed total genomic DNA preparations from class A or class B clones into a Zorro3<sup>-</sup>, His<sup>-</sup> strain. We reasoned that for class A genomic DNA to convert this parental strain to His<sup>+</sup>, a fragment of



**Figure 3.** Structural characterization of class B products. (A) Predicted structures and restriction fragment sizes of class B-I and class B-II circles. (B) Verification of predicted class B structures. Genomic DNA from donorless strains containing a class B circle were digested with the indicated enzyme and subjected to Southern analysis with the Zorro3mHIS3 probe shown in Figure 1B.

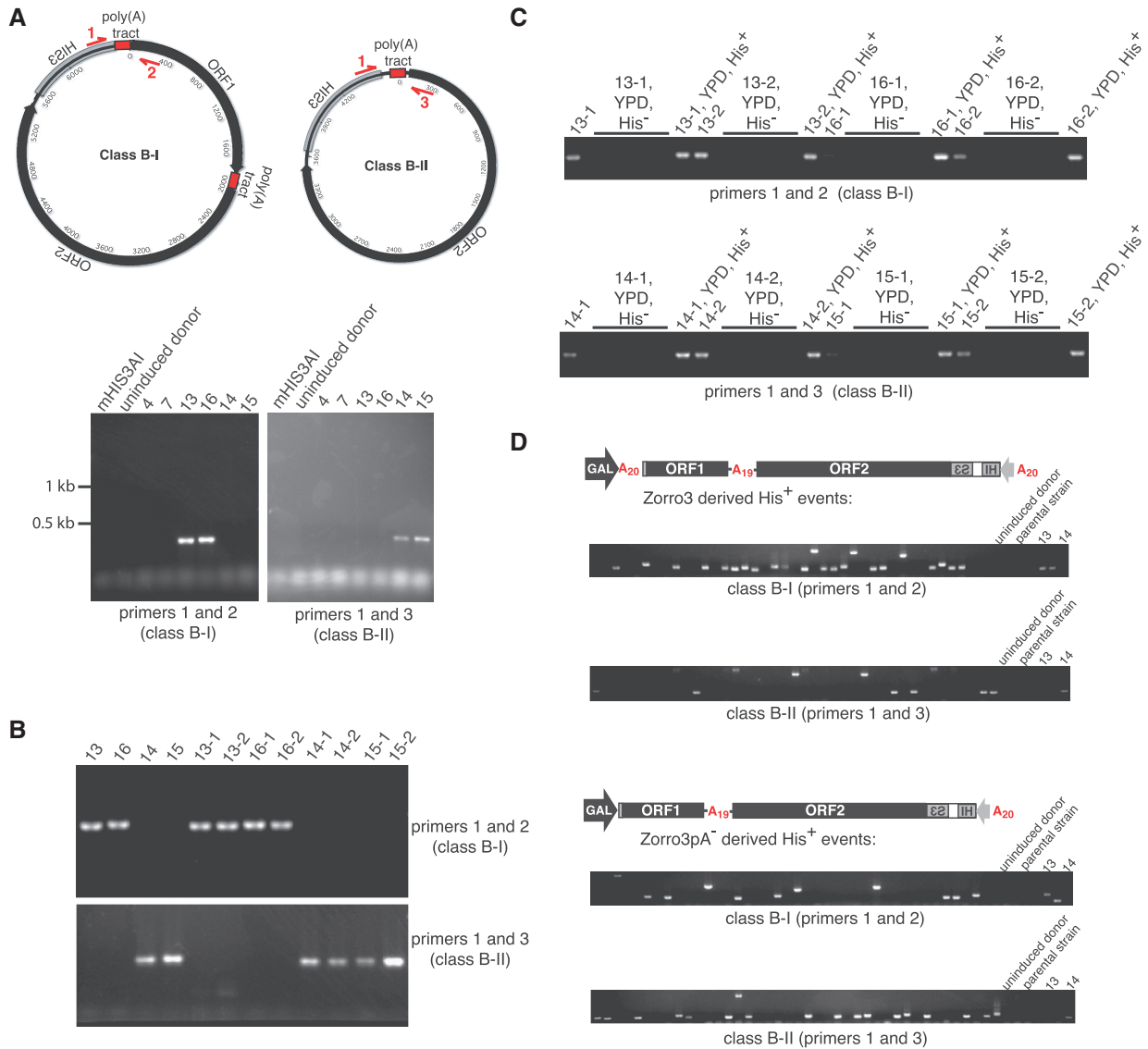
transformed DNA containing the spliced *HIS3* gene would need to integrate into the host chromosome via homologous recombination or non-homologous end joining. Given that DNA fragments containing the *HIS3* gene are expected to be a small fraction of the total nucleic acid in a genomic DNA prep, transformation to His<sup>+</sup> by class A DNA is a rare event (Figure 2C, left). However, transformation with the same amount of class B DNA led to comparatively high number of His<sup>+</sup> colonies (Figure 2C, middle and right). This suggests that class B products are in the more easily transformable circular form. To confirm this, we isolated total genomic DNA from 'recipient' strains (e.g. 13-1, 13-2, etc.) newly transformed with class B DNA and probed for Zorro3 sequence (Figure 2D). The respective class B products were present in the recipient strains, but the chromosomally integrated donor from the original strains was not transferred (arrow in Figure 2D). The His<sup>+</sup> phenotype exhibited the same instability shown in original class B strains (Figure 2E). These data are consistent with class B products being plasmids.

As Zorro3 has three polyadenosine [poly(A)] tracts corresponding to the locations where class B circles are 'joined', we inferred the general structure of class B-I and class B-II forms (Figure 3A). We then confirmed these structures by restriction digests of genomic DNA from strains containing a class B-I or class B-II product (but not the original integrated donor) and probing for

Zorro3 (Figure 3B). In all cases, we obtained the predicted restriction pattern, demonstrating unambiguously that non-LTR retrotransposition can, in some cases, result in circular products. The circular structure of these products also explains the size of the bands obtained from LM-PCR (shown in Supplementary Figure S2 schematic).

### Class B products can form with high frequency

To simplify analysis of large numbers of retrotransposition events, we used a colony PCR (cPCR assay). Primers flanking the class B-I or class B-II junctions (primers 1 and 2 or primers 1 and 3 in Figure 4A) specifically detected the presence or absence of class B circles (Figure 4A–C). We used the cPCR assay to measure class B product frequencies. We induced 44 independent retrotransposition events. Remarkably, 24/44 clones contained a class B-I product, and 8/44 clones contained a class B-II product (Figure 4D, top). Thus class B retrotransposition can be a frequent event. As the formation of circles is possibly augmented by a template switch between the poly(A) tracts of the Zorro3 mRNA and the A-rich top strand of the Zorro3 target site, we investigated whether deleting the 5' poly(A) tract of Zorro3 (Zorro3pA<sup>-</sup>) eliminated class B-I products. We found that 12/44 retrotransposition clones derived from Zorro3pA<sup>-</sup> contained a class B-I event, suggesting that the long stretch of homology [in this case, poly(A)] is not required for circularization (Figure 4D, bottom).



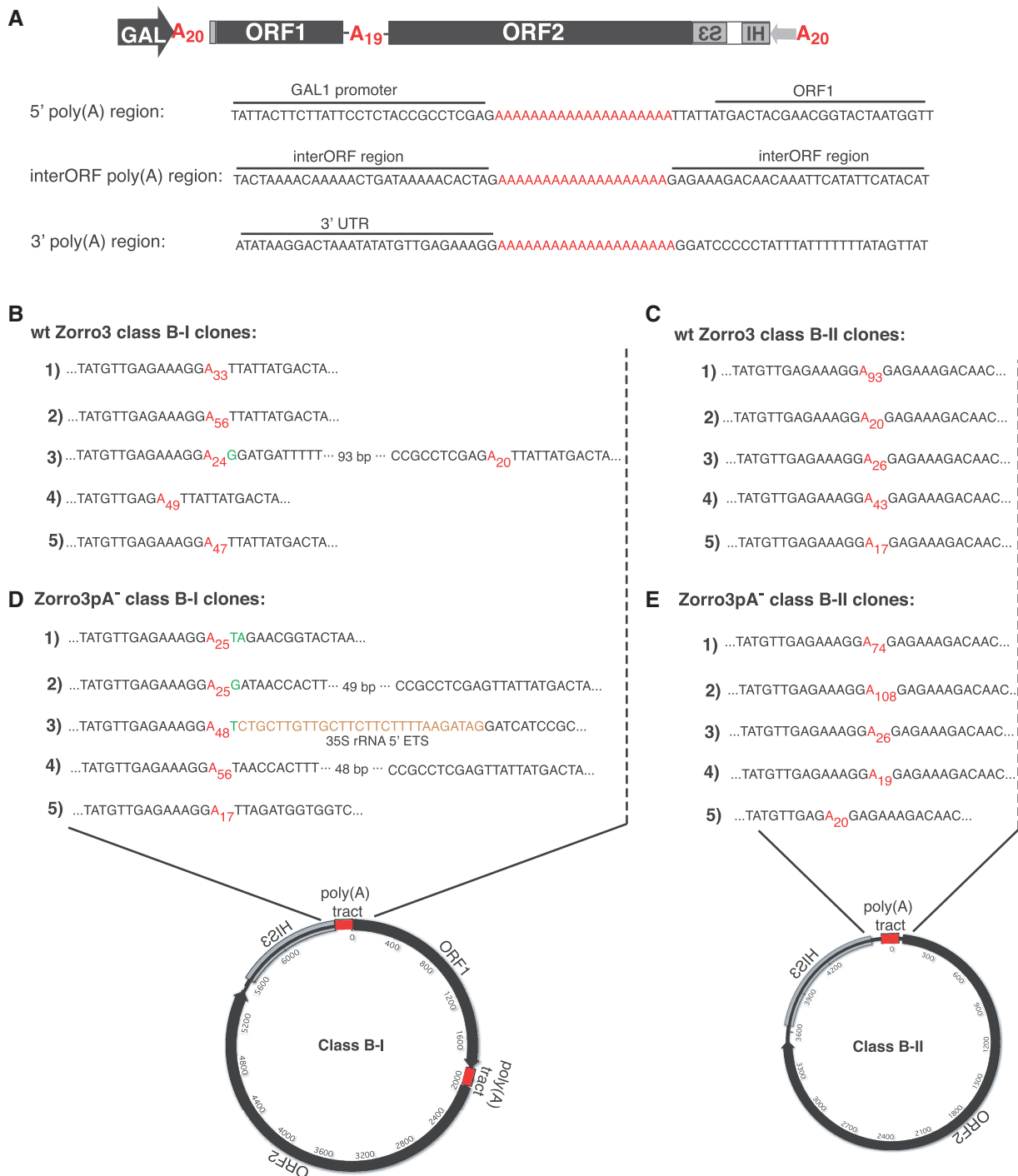
**Figure 4.** Class B products form with high frequency. **(A)** A cPCR assay to detect class B products. The indicated primers (in red) were used for colony PCR reactions to specifically amplify products in strains containing class B-I or class B-II circles. **(B)** The cPCR assay is sensitive and specific for class B products. Original class B strains and strains converted to His<sup>+</sup> by class B genomic DNA transformation (Figure 2C and D) were subjected to the cPCR assay. **(C)** His<sup>+</sup> phenotype tracks with the class B circle. A His<sup>+</sup> stability experiment (described in Figure 2E) was performed on parent strains transformed with a class B product. Colonies that lost the His<sup>+</sup> phenotype (YPD, His<sup>-</sup>) and rare colonies that retained the phenotype (YPD, His<sup>+</sup>) were tested by the cPCR assay for the presence of a class B episome. **(D)** Class B products constitute the majority of Zorro3 retrotransposition events. For two strains (containing Zorro3mHIS3AI or Zorro3pA<sup>-</sup>mHIS3AI), 44 independently derived retrotransposition events (all unlabelled lanes) were tested by cPCR for the presence of class B-I or class B-II products. 13 = an original class B-I clone (control). 14 = an original class B-II clone (control).

### Pre-existing homology is not required for class B product formation

To learn more about the nature of the junctions of class B circles (junction defined as the region corresponding to where the 5' and 3' ends of the retrotransposed element are joined to form the circle), we cloned and sequenced the junctions of five independent class B-I and class B-II clones derived from either wild-type Zorro3 or Zorro3pA<sup>-</sup> (Figure 5A–E). Sequencing revealed that 3/5 wild-type class B-I clones (#1, 2, 5 from Figure 5B) comprised full-length Zorro3 sequence with the 3' and 5' ends joined by a poly(A) tract. The only difference between these clones

was the length of the joining poly(A) tract. A fourth clone (#4) was nearly identical, but missing two Gs at the 3' end before the poly(A) junction. This could be due to an alternative cleavage and polyadenylation site, or internal priming of Zorro3 mRNA during the initial step of reverse transcription (48). The remaining clone (#3) contained 113 base pairs of sequence upstream of the 5' poly(A) tract, preceded by an untemplated G, suggesting that this clone was derived from an mRNA that initiated transcription upstream of the 5' poly(A) tract.

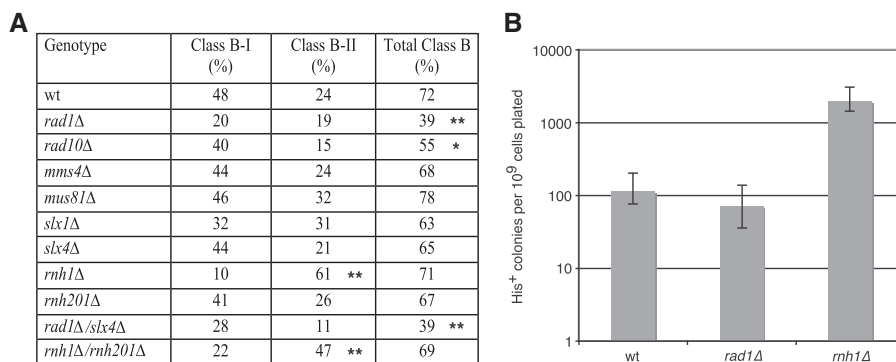
The junctions of class B-II products derived from Zorro3 and Zorro3pA<sup>-</sup> were essentially the same, with



**Figure 5.** Pre-existing homology is not required for class B product formation. (A) Structure of Zorro3 donor and view of Zorro3 regions where most class B junctions occur. Class B-I circles typically form from a joining between the 3' poly(A) region and the 5' poly(A) region, whereas class B-II circles typically form from a joining between the 3' poly(A) region and the interORF poly(A) region. Zorro3pA<sup>-</sup> sequences are identical to Zorro3 except the stretch of 20 (A)s in the 5' poly(A) region is deleted. (B–E) Sequences of cloned class B product junctions derived from wild-type Zorro3 or Zorro3pA<sup>-</sup> donors. Untemplated nucleotides are indicated in green. Yeast genome sequences are indicated in brown.

variations in poly(A) length being the only difference between clones (Figure 5C and E). One clone (Zorro3pA<sup>-</sup> #5) was missing two Gs at the 3' end, but otherwise identical to the rest. The similarities between Zorro3 and Zorro3pA<sup>-</sup> class B-II clones were expected, as no modifications were made to the interORF poly(A) tract in Zorro3pA<sup>-</sup>.

The aforementioned data suggests that, in most cases, stretches of homology [poly(A) tracts] between the 5' and 3' ends of the element serve as the preferred connecting junction for class B circles. As other notable LINE elements (e.g. human and mouse L1) do not contain pre-existing homology between the 5' and 3' ends, we analysed Zorro3pA<sup>-</sup> class B-I products to ask whether



**Figure 6.** Mutations affecting class B frequency and ratio. (A) Zorro3 was induced for retrotransposition in the indicated mutants. In all, 94 independently derived clones were picked for each mutant and assayed by cPCR for the presence of Class B products. **\*\*** $P < 0.001$ . **\*** $P < 0.05$ . (B) Loss of *RNH1* leads to an increase in total Zorro3 retrotransposition. A quantitative retrotransposition assay was performed on the indicated strains. Frequency is presented as the average of five independent cultures. Black bars indicate the range of frequencies obtained.

the pre-existing 5' and 3' homology in Zorro3 was an absolute requirement for class B circle formation. When the 5' poly(A) tract was deleted, class B-I products still formed, but with greater heterogeneity and non-templated nucleotides at the junction (Figure 5D). Non-templated nucleotides are similarly seen at mammalian L1 5' junctions (36,37). It has been proposed that these non-templated nucleotides are added onto the 3' end of the minus strand by reverse transcriptase, facilitating microhomology-mediated template switching (49). Two of the truncation points (clones #2, 4) are upstream of the beginning of the element [where the 5' poly(A) tract would normally be]. This upstream sequence may represent transcription initiating upstream of the Zorro3 start. Two of the truncation points (clones #1, 5) are downstream in the element. The final clone 5' truncates downstream in Zorro3 sequence, but template switches to 35S rRNA sequence before template switching to Zorro3 3' sequence. In a prior analysis of class A clones, we saw similar cases of template switching to endogenous cellular sequences (40). There is no appreciable pre-existing homology at the junction in any of the Zorro3pA<sup>-</sup> class B-I clones. In 3/5 cases, there were untemplated bases at the junction point before the template switch. This demonstrates that pre-existing element junction homology is not a requirement for the formation of class B products.

#### Mutations affecting class B frequency and ratio

Zorro3 retrotransposition events (which, as shown earlier, are predominantly class B products) are dependent on conserved residues in ORF1, ORF2 endonuclease and ORF2 reverse transcriptase (40). Thus, it is likely that class B products are formed via a variation of TPRT. The endonuclease requirement, in particular, suggests that for class B circles, the reaction requires nicked chromosomal DNA as the initiating primer for reverse transcription. This would mean that even retrotransposition intermediates destined to become circles would be initially attached to a chromosome.

Based on this assumption, we explored how a non-LTR retrotransposition intermediate could be excised from a

chromosome. TPRT predicts the formation of 3' flaps on the top and bottom strands of a retrotransposition intermediate (see Figure 1A, bottom for an example of the bottom strand flap). Cleavage of both of these flaps by a cellular endonuclease(s) could lead to excision of the L1 intermediate. Previous work has shown that the XPF/ERCC1, a 3' flap endonuclease involved in nucleotide excision repair and homologous recombination, can restrict human L1 retrotransposition (50). This suggests a potential interaction between L1 intermediates and 3' flap endonucleases.

To test whether the loss of 3' flap endonucleases in yeast affects the formation of class B products, we knocked-out various genes involved in recognition and cleavage of branched nucleic acid structures. Rad1/Rad10 is the yeast ortholog of the mammalian XPF/ERCC1 heterodimer mentioned earlier. Mus81/Mms4 is an endonuclease involved in Holiday junction resolution with a preference for 3' flap substrates (51,52). Slx4 is thought to recognize DNA structures and serves as a scaffold to recruit endonucleases, such as Rad1/Rad10, Mus81/Mms4 and Slx1 (a 5' flap endonuclease) (53). In each of these mutant strains, we induced Zorro3 retrotransposition and assayed the percentage and ratio of class B events (Figure 6A). The most striking effects on formation of class B products were found in *rad1Δ* and *rad10Δ* strains, which showed a significant decrease in total percentage of class B products ( $P = 9 \times 10^{-6}$  and  $P = 0.022$ , respectively; Fisher's exact test). Although Slx4 can form a complex with Rad1/Rad10, deletion of *SLX4* did not lead to dramatic changes in class B frequency. In addition, a *rad1Δ/slx4Δ* double mutant showed a similar phenotype as the *rad1Δ* single mutant. Together, these data suggest that Rad1/Rad10 can act independently of Slx4. Although the percentage of class B product formation decreased in *rad1Δ* mutants, overall levels of retrotransposition remained largely unchanged (Figure 6B). Our data suggest that Rad1/Rad10 may play a role flap in cleavage during class B circle release from the chromosome.

Although LINE reverse transcription is usually initiated by endonuclease mediated nicks, in some cases, DNA



breaks formed by other means can serve as a primer for reverse transcription (54,55). We sought to genetically induce DNA breaks by deleting RNaseH. *rnh1Δ* strains are expected to lead to a relative persistence of RNA/DNA hybrids. Multiple studies suggest that persistent RNA/DNA hybrids lead to DNA breaks (56–59). We predicted that an environment with preexisting DNA breaks would provide more potential chromosomal primers for LINE minus and/or plus strand synthesis. In the case of Zorro3, which has a low rate of retrotransposition ( $\sim 1 \times 10^{-6}$  under the conditions of our assay), this may lead to an increase in retrotransposition frequency. Indeed, we found that in an *rnh1Δ* strain, overall Zorro3 retrotransposition frequency increases by greater than 10-fold (Figure 6B). Importantly, class B products still comprised  $\sim 70\%$  of these *rnh1Δ* retrotransposition events, indicating that class A and class B products are equally increased by introducing more potential chromosomal primers. This further strengthens our hypothesis that class B products are initiated at the chromosome.

At some point during TPRT, cleavage of the top strand must occur. The newly synthesized minus strand can then jump templates from retrotransposon RNA to the newly cleaved top strand. We hypothesized that in an *rnh1Δ* strain, the abundant availability of chromosomal breaks to serve as a primer (especially if a break is made at the target site) could allow premature minus strand template jumping and plus strand synthesis. Earlier minus strand template jumping [e.g. from the interORF poly(A) tract instead of the 5' UTR poly(A) tract] would be expected to give rise to a greater proportion of class B-II products. Consistent with this hypothesis, in *rnh1Δ* strains, we observe a dramatic shift in the ratio of class B-I to class B-II products. Wild-type class B products are 33% class B-II, whereas *rnh1Δ* class B products are 85% class B-II ( $P = 3 \times 10^{-10}$ ; Fisher's exact test). Again, these results are consistent with minus and plus strand synthesis being initiated from the chromosome.

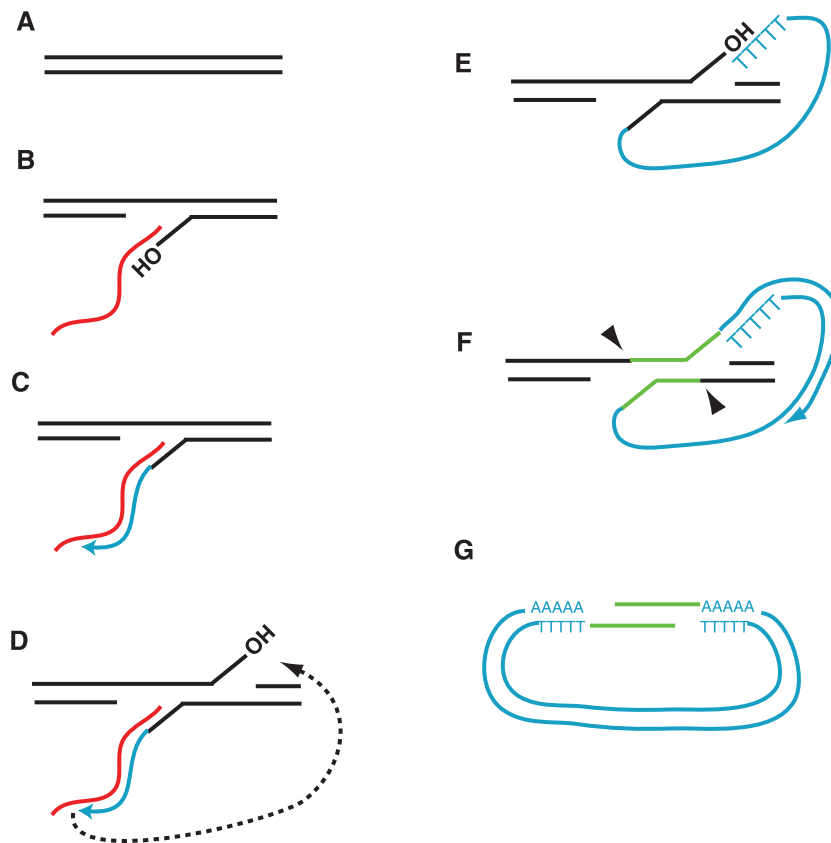
## DISCUSSION

We have shown that in addition to the expected integrated retrotransposition events, a LINE element can also generate episomal retrotransposition events. Previously, circular products derived from retroviruses/LTR-retrotransposons have been observed (60–62). However, retroviruses/LTR-retrotransposons replicate in a fundamentally different manner when compared with non-LTR retrotransposons. Retroviruses/LTR-retrotransposons complete linear cDNA synthesis in the cytoplasm. The formation of retroviral/LTR-retrotransposon circles is postulated to occur by various mechanisms, including homologous recombination between the 5' and 3' LTRs, non-homologous end joining of the linear cDNA ends, errors in reverse transcription and auto-integration (63–69). As these mechanisms involve retroviral intermediates, they seem unlikely to explain the formation of LINE circles. A less-specific class of circular DNA, small dispersed circular DNA

(spcDNA), has been found to be present in essentially all mammalian cells (70,71). spcDNA consists of heterogeneous circles of sizes from hundreds to thousands of base pairs, and spcDNA is increased in cells with genome instability (70,71). spcDNA appears to be derived from chromosomal DNA and contains unique sequence, non-coding sequence and repetitive elements (including Alus and L1s). As short interspersed nuclear elements (SINEs) and LINEs are among the most abundant sequences in the genome, it is not surprising that circularized chromosomal fragments would contain these elements owing to their sheer percentage of genome mass. There is a variety of proposed mechanisms for the formation of spcDNA (72), the most likely being 'pop-out' of existing genomic DNA. Thus, the work described here is the first example of LINE circle formation via *bona fide* retrotransposition.

How are class B products formed? As mentioned earlier, class B products are normally endo-dependent. This suggests that chromosomal nicks are used as primers for reverse transcription, and thus class B products initiate attached to a chromosome. The requirement of ORF1, ORF2 endo and ORF2 reverse transcriptase strongly suggests that class B products are formed by TPRT, with a variation in retrotransposition intermediate resolution. The massive increase in both class A and class B retrotransposition in *rnh1Δ* strains further supports a TPRT-based model for class B formation, as *rnh1Δ* strains are expected to contain excess DNA breaks, a substrate for TPRT (54–59). Based on our data, we present a speculative model for circle formation in Figure 7. The steps in Figure 7A–E are identical to a standard TPRT reaction. We speculate that a structure-specific endonuclease can cleave the top and bottom strands in the vicinity of the 3' flaps. As deletion of the Rad1/Rad10 3' flap endonuclease significantly reduces class B product formation, Rad1/Rad10 may play a role in such a cleavage. However, Rad1/Rad10 cannot be solely responsible for this step, as loss of Rad1/Rad10 does not lead to the complete elimination of class B products. It is possible that other cellular nucleases play a redundant role, or even that the Zorro3 endonuclease may perform the recleavages. Further work is needed to distinguish between these possibilities. Once cleaved from the chromosome, the ends of the predicted linear excision product may contain genomic DNA corresponding to the genomic target site (highlighted in green in Figure 7F and G). These ends could anneal and be repaired by DNA repair pathways to form a circle. As Zorro3 targets poly(A) tracts, the highlighted green sequence in Figure 7F and G is likely to be poly(A), giving the appearance of a pure poly(A) junction. If this model were extrapolated to other LINE elements, the highlighted green segment of DNA would correspond to part of the target site duplication, and the size would be dependent on the distance between the top and bottom strand recleavages.

An obvious alternative to the model in Figure 7 is the 'pop-out' of stably integrated Zorro3 retrotransposition products owing to homologous recombination between the flanking poly(A) tracts. Several lines of evidence argue against this possibility. First, class B products



**Figure 7.** A speculative model for the formation of class B products. (A) Genomic target site. (B) Bottom chromosome strand nick and LINE mRNA annealing. (C) Minus strand synthesis. (D–E) Top strand chromosomal nick and template jump to top strand. (F) Recleavage of top and bottom strands (at arrowheads) to release the retrotransposition intermediate. The timing and placement of recleavages do not necessarily need to occur as indicated in the figure. Chromosomal target site DNA between the sites of recleavage have been highlighted in green for emphasis. As Zorro3 integrates into poly(A) tracts, the DNA highlighted in green would likely be poly(A) sequence. For other LINE elements, the DNA in green would likely correspond to part of the target site duplication. (G) Resolution of the excised circle by annealing of complementary strands. Red lines indicate RNA. Blue lines indicate newly synthesized retrotransposon DNA.

were initially observed in *rad52Δ* strains, which are deficient in homologous recombination (Figures 1–5). Second, the portion of retrotransposition products that are class B is similar between a *rad52Δ* strain (Figure 4D) and *Rad52<sup>+</sup>* strain (Figure 6A, wt). Third, we do not see clones that contain mixtures of class A/class B events or class B-I/class B-II events. We also have never witnessed the conversion of class A to class B products, or conversion of class B-I to the smaller class B-II products. This indicates that these retrotransposition products are independently generated and not derived from each other.

We do not currently know whether class B products can be formed by other L1 clade members. The majority of investigation on the L1 clade has focused on human and mouse L1, and thus far circular retrotransposition products have not been observed for these elements. However, the current assay for mammalian retrotransposition is not designed to detect circular retrotransposition products. Generally, unless a plasmid is specifically engineered to be maintained as an episome, clones derived from mammalian plasmid transfection and selection are the result of plasmid integration into

the chromosome. Thus, if class B products are formed during mammalian retrotransposition assays, they are likely lost or integrated into the chromosome under the current assay schemes. We detected class B products in yeast owing to the fortuitous ability for Zorro3 class B products to be maintained in yeast under selection. This suggests that some subset of Zorro3 sequence can act as an autonomous replicating sequence in yeast. The ability for foreign sequence to replicate in yeast is not unprecedented (73), and we view the Zorro3 autonomous replicating sequence activity as serendipity. Even if other L1 clade members form class B products in their native hosts, we do not necessarily expect these products to contain an origin of replication.

Although we would expect any particular non-LTR retrotransposition-derived circle to exist only transiently, class B products may not necessarily be ‘dead ends’, as some LINES (notably human and mouse L1) encode a promoter on their mRNA. Thus, if a mammalian full-length class B product is formed in a cell, this circle could in principle be a source for future retrotransposition events. As the circles would be newly synthesized DNA divorced from the chromosomal context, they may escape

silencing and could be even more retrotranspositionally active (on a per copy basis) than chromosomal copies. This would be an example of how a transposon, in this case, a LINE element, could circumvent apparent 'death by excision', to potentially jump again. Further studies are needed to assess whether class B products can be formed in other organisms, especially in evolutionarily relevant cellular niches (e.g. the germ line of mammals).

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–2 and Supplementary Figures 1–2.

## ACKNOWLEDGEMENTS

We thank members of the Han lab for critical reading of the manuscript, the Koshland lab for discussion and data sharing, Allison Pinder for sequencing, and Super Billy for miscellaneous support. The authors declare no financial conflicts of interest.

## FUNDING

National Institutes of Health [R01GM090192]; the Carnegie Institution for Science. Funding for open access charge: Carnegie Institution for Science.

*Conflict of interest statement.* None declared.

## REFERENCES

- Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860–921.
- Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P. *et al.* (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature*, **420**, 520–562.
- Chimpanzee Sequencing and Analysis Consortium. (2005) Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature*, **437**, 69–87.
- Gibbs, R.A., Weinstock, G.M., Metzker, M.L., Muzny, D.M., Sodergren, E.J., Scherer, S., Scott, G., Steffen, D., Worley, K.C., Burch, P.E. *et al.* (2004) Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature*, **428**, 493–521.
- Warren, W.C., Hillier, L.W., Marshall Graves, J.A., Birney, E., Ponting, C.P., Grutzner, F., Belov, K., Miller, W., Clarke, L., Chinwalla, A.T. *et al.* (2005) Genome analysis of the platypus reveals unique signatures of evolution. *Nature*, **438**, 175–183.
- Gentles, A.J., Wakefield, M.J., Kohany, O., Gu, W., Batzer, M.A., Pollock, D.D. and Jurka, J. (2007) Evolutionary dynamics of transposable elements in the short-tailed opossum *Monodelphis domestica*. *Genome Res.*, **17**, 992–1004.
- Pontius, J.U., Mullikin, J.C., Smith, D.R., Lindblad-Toh, K., Gnerre, S., Clamp, M., Chang, J., Stephens, R., Neelam, B., Volfvsky, N. *et al.* (2007) Initial sequence and comparative analysis of the cat genome. *Genome Res.*, **17**, 1675–1689.
- Lindblad-Toh, K., Wade, C.M., Mikkelsen, T.S., Karlsson, E.K., Jaffe, D.B., Kamal, M., Clamp, M., Chang, J.L., Kulbokas, E.J. 3rd, Zody, M.C. *et al.* (2005) Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature*, **438**, 803–819.
- Dewannieux, M., Esnault, C. and Heidmann, T. (2003) LINE-mediated retrotransposition of marked Alu sequences. *Nat. Genet.*, **35**, 41–48.
- Wallace, N.A., Belancio, V.P. and Deininger, P.L. (2008) L1 mobile element expression causes multiple types of toxicity. *Gene*, **419**, 75–81.
- Gasior, S.L., Wakeman, T.P., Xu, B. and Deininger, P.L. (2006) The human LINE-1 retrotransposon creates DNA double-strand breaks. *J. Mol. Biol.*, **357**, 1383–1393.
- Kazazian, H.H. Jr., Wong, C., Youssoufian, H., Scott, A.F., Phillips, D.G. and Antonarakis, S.E. (1988) Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. *Nature*, **332**, 164–166.
- Babushok, D.V. and Kazazian, H.H. Jr. (2007) Progress in understanding the biology of the human mutagen LINE-1. *Hum. Mutat.*, **28**, 527–539.
- Korbel, J.O., Urban, A.E., Affourtit, J.P., Godwin, B., Grubert, F., Simons, J.F., Kim, P.M., Palejev, D., Carriero, N.J., Du, L. *et al.* (2007) Paired-end mapping reveals extensive structural variation in the human genome. *Science*, **318**, 420–426.
- Kidd, J.M., Cooper, G.M., Donahue, W.F., Hayden, H.S., Sampas, N., Graves, T., Hansen, N., Teague, B., Alkan, C., Antonacci, F. *et al.* (2008) Mapping and sequencing of structural variation from eight human genomes. *Nature*, **453**, 56–64.
- Ewing, A.D. and Kazazian, H.H. Jr. (2010) High-throughput sequencing reveals extensive variation in human-specific L1 content in individual human genomes. *Genome Res.*, **20**, 1262–1270.
- Beck, C.R., Collier, P., Macfarlane, C., Malig, M., Kidd, J.M., Eichler, E.E., Badge, R.M. and Moran, J.V. (2010) LINE-1 retrotransposition activity in human genomes. *Cell*, **141**, 1159–1170.
- Huang, C.R., Schneider, A.M., Lu, Y., Niranjana, T., Shen, P., Robinson, M.A., Steranka, J.P., Valle, D., Civin, C.I., Wang, T. *et al.* (2010) Mobile interspersed repeats are major structural variants in the human genome. *Cell*, **141**, 1171–1182.
- Iskow, R.C., McCabe, M.T., Mills, R.E., Torene, S., Pittard, W.S., Neuwald, A.F., Van Meir, E.G., Vertino, P.M. and Devine, S.E. (2010) Natural mutagenesis of human genomes by endogenous retrotransposons. *Cell*, **141**, 1253–1261.
- Akagi, K., Li, J., Stephens, R.M., Volfvsky, N. and Symer, D.E. (2008) Extensive variation between inbred mouse strains due to endogenous L1 retrotransposition. *Genome Res.*, **18**, 869–880.
- Hohjoh, H. and Singer, M.F. (1996) Cytoplasmic ribonucleoprotein complexes containing human LINE-1 protein and RNA. *Embo. J.*, **15**, 630–639.
- Hohjoh, H. and Singer, M.F. (1997) Sequence-specific single-strand RNA binding protein encoded by the human LINE-1 retrotransposon. *Embo. J.*, **16**, 6034–6043.
- Kolosha, V.O. and Martin, S.L. (1997) In vitro properties of the first ORF protein from mouse LINE-1 support its role in ribonucleoprotein particle formation during retrotransposition. *Proc. Natl. Acad. Sci. USA*, **94**, 10155–10160.
- Kolosha, V.O. and Martin, S.L. (2003) High-affinity, non-sequence-specific rna binding by the open reading frame 1 (ORF1) protein from long interspersed nuclear element 1 (LINE-1). *J. Biol. Chem.*, **278**, 8112–8117.
- Martin, S.L. and Bushman, F.D. (2001) Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon. *Mol. Cell Biol.*, **21**, 467–475.
- Feng, Q., Moran, J.V., Kazazian, H.H. Jr. and Boeke, J.D. (1996) Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell*, **87**, 905–916.
- Hattori, M., Kuhara, S., Takenaka, O. and Sakaki, Y. (1986) L1 family of repetitive DNA sequences in primates may be derived from a sequence encoding a reverse transcriptase-related protein. *Nature*, **321**, 625–628.
- Fanning, T. and Singer, M. (1987) The LINE-1 DNA sequences in four mammalian orders predict proteins that conserve homologies to retrovirus proteins. *Nucleic Acids Res.*, **15**, 2251–2260.
- Mathias, S.L., Scott, A.F., Kazazian, H.H. Jr., Boeke, J.D. and Gabriel, A. (1991) Reverse transcriptase encoded by a human transposable element. *Science*, **254**, 1808–1810.

30. Martin,S.L., Li,J., Epperson,L.E. and Lieberman,B. (1998) Functional reverse transcriptases encoded by A-type mouse LINE-1: defining the minimal domain by deletion analysis. *Gene*, **215**, 69–75.
31. Kulpa,D.A. and Moran,J.V. (2005) Ribonucleoprotein particle formation is necessary but not sufficient for LINE-1 retrotransposition. *Hum. Mol. Genet.*, **14**, 3237–3248.
32. Martin,S.L., Cruceanu,M., Branciforte,D., Wai-Lun Li,P., Kwok,S.C., Hodges,R.S. and Williams,M.C. (2005) LINE-1 retrotransposition requires the nucleic acid chaperone activity of the ORF1 protein. *J. Mol. Biol.*, **348**, 549–561.
33. Moran,J.V., Holmes,S.E., Naas,T.P., DeBerardinis,R.J., Boeke,J.D. and Kazazian,H.H. Jr. (1996) High frequency retrotransposition in cultured mammalian cells. *Cell*, **87**, 917–927.
34. Luan,D.D., Korman,M.H., Jakubczak,J.L. and Eickbush,T.H. (1993) Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell*, **72**, 595–605.
35. Cost,G.J., Feng,Q., Jacquier,A. and Boeke,J.D. (2002) Human L1 element target-primed reverse transcription in vitro. *Embo. J.*, **21**, 5899–5910.
36. Gilbert,N., Lutz-Prigge,S. and Moran,J.V. (2002) Genomic deletions created upon LINE-1 retrotransposition. *Cell*, **110**, 315–325.
37. Symer,D.E., Connelly,C., Szak,S.T., Caputo,E.M., Cost,G.J., Parmigiani,G. and Boeke,J.D. (2002) Human L1 retrotransposition is associated with genetic instability in vivo. *Cell*, **110**, 327–338.
38. Gilbert,N., Lutz,S., Morrish,T.A. and Moran,J.V. (2005) Multiple fates of L1 retrotransposition intermediates in cultured human cells. *Mol. Cell Biol.*, **25**, 7780–7795.
39. Han,J.S. and Boeke,J.D. (2004) A highly active synthetic mammalian retrotransposon. *Nature*, **429**, 314–318.
40. Dong,C., Poulter,R.T. and Han,J.S. (2009) LINE-Like Retrotransposition in *Saccharomyces cerevisiae*. *Genetics*, **181**, 301–311.
41. Goodwin,T.J., Ormandy,J.E. and Poulter,R.T. (2001) L1-like non-LTR retrotransposons in the yeast *Candida albicans*. *Curr. Genet.*, **39**, 83–91.
42. Goodwin,T.J., Busby,J.N. and Poulter,R.T. (2007) A yeast model for target-primed (non-LTR) retrotransposition. *BMC Genomics*, **8**, 263.
43. Khazina,E. and Weichenrieder,O. (2009) Non-LTR retrotransposons encode noncanonical RRM domains in their first open reading frame. *Proc. Natl. Acad. Sci. USA*, **106**, 731–736.
44. Baudin,A., Ozier-Kalogeropoulos,O., Denouel,A., Lacroute,F. and Cullin,C. (1993) A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. *Nucleic Acids Res.*, **21**, 3329–3330.
45. Deutschbauer,A.M. and Davis,R.W. (2005) Quantitative trait loci mapped to single-nucleotide resolution in yeast. *Nat. Genet.*, **37**, 1333–1340.
46. Amberg,D.C., Burke,D.J. and Strathern,J.N. (2005) *Methods in Yeast Genetics*. CSHL Press, Cold Spring Harbor, NY.
47. Curcio,M.J. and Garfinkel,D.J. (1991) Single-step selection for Ty1 element retrotransposition. *Proc. Natl. Acad. Sci. USA*, **88**, 936–940.
48. Kulpa,D.A. and Moran,J.V. (2006) Cis-preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. *Nat. Struct. Mol. Biol.*, **13**, 655–660.
49. Bilbillo,A. and Eickbush,T.H. (2004) End-to-end template jumping by the reverse transcriptase encoded by the R2 retrotransposon. *J. Biol. Chem.*, **279**, 14945–14953.
50. Gasior,S.L., Roy-Engel,A.M. and Deininger,P.L. (2008) ERCC1/XPF limits L1 retrotransposition. *DNA Repair (Amst)*, **7**, 983–989.
51. Ciccia,A., Constantinou,A. and West,S.C. (2003) Identification and characterization of the human mus81-emel endonuclease. *J. Biol. Chem.*, **278**, 25172–25178.
52. Constantinou,A., Chen,X.B., McGowan,C.H. and West,S.C. (2002) Holliday junction resolution in human cells: two junction endonucleases with distinct substrate specificities. *EMBO J.*, **21**, 5577–5585.
53. Fekairi,S., Scaglione,S., Chahwan,C., Taylor,E.R., Tissier,A., Coulon,S., Dong,M.Q., Ruse,C., Yates,J.R., Russell,P. 3rd, et al. (2009) Human SLX4 is a Holliday junction resolvase subunit that binds multiple DNA repair/recombination endonucleases. *Cell*, **138**, 78–89.
54. Morrish,T.A., Gilbert,N., Myers,J.S., Vincent,B.J., Stamato,T.D., Taccioli,G.E., Batzer,M.A. and Moran,J.V. (2002) DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. *Nat. Genet.*, **31**, 159–165.
55. Morrish,T.A., Garcia-Perez,J.L., Stamato,T.D., Taccioli,G.E., Sekiguchi,J. and Moran,J.V. (2007) Endonuclease-independent LINE-1 retrotransposition at mammalian telomeres. *Nature*, **446**, 208–212.
56. Huertas,P. and Aguilera,A. (2003) Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol. Cell*, **12**, 711–721.
57. Li,X. and Manley,J.L. (2006) Cotranscriptional processes and their influence on genome stability. *Genes Dev.*, **20**, 1838–1847.
58. Li,X. and Manley,J.L. (2005) Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell*, **122**, 365–378.
59. Wahba,L., Amon,J.D., Koshland,D. and Vuica-Ross,M. (2011) RNase H and multiple RNA biogenesis factors cooperate to prevent RNA:DNA hybrids from generating genome instability. *Mol. Cell*, **44**, 978–988.
60. Stanfield,S. and Helinski,D.R. (1976) Small circular DNA in *Drosophila melanogaster*. *Cell*, **9**, 333–345.
61. Flavell,A.J. and Ish-Horowicz,D. (1981) Extrachromosomal circular copies of the eukaryotic transposable element copia in cultured *Drosophila* cells. *Nature*, **292**, 591–595.
62. Whitcomb,J.M. and Hughes,S.H. (1992) Retroviral reverse transcription and integration: progress and problems. *Annu. Rev. Cell Biol.*, **8**, 275–306.
63. Li,L., Olvera,J.M., Yoder,K.E., Mitchell,R.S., Butler,S.L., Lieber,M., Martin,S.L. and Bushman,F.D. (2001) Role of the non-homologous DNA end joining pathway in the early steps of retroviral infection. *EMBO J.*, **20**, 3272–3281.
64. Kilzer,J.M., Stracker,T., Beitzel,B., Meek,K., Weitzman,M. and Bushman,F.D. (2003) Roles of host cell factors in circularization of retroviral dna. *Virology*, **314**, 460–467.
65. Jeanson,L., Subra,F., Vaganay,S., Hervy,M., Marangoni,E., Bourhis,J. and Mouscadet,J.F. (2002) Effect of Ku80 depletion on the preintegrative steps of HIV-1 replication in human cells. *Virology*, **300**, 100–108.
66. Varmus,H. and Brown,P. (1989) Retroviruses. In: Berg,D.E. and Howe,M.M. (eds), *Mobile DNA*. American Society for Microbiology, Washington, DC, pp. 53–108.
67. Miller,M.D., Wang,B. and Bushman,F.D. (1995) Human immunodeficiency virus type 1 preintegration complexes containing discontinuous plus strands are competent to integrate in vitro. *J. Virol.*, **69**, 3938–3944.
68. Lee,Y.M. and Coffin,J.M. (1990) Efficient autointegration of avian retrovirus DNA in vitro. *J. Virol.*, **64**, 5958–5965.
69. Farnet,C.M. and Haseltine,W.A. (1991) Circularization of human immunodeficiency virus type 1 DNA in vitro. *J. Virol.*, **65**, 6942–6952.
70. Schmidt,H., Taubert,H., Lange,H., Kriese,K., Schmitt,W.D., Hoffmann,S., Bartel,F. and Hauptmann,S. (2009) Small polydispersed circular DNA contains strains of mobile genetic elements and occurs more frequently in permanent cell lines of malignant tumors than in normal lymphocytes. *Oncol. Rep.*, **22**, 393–400.
71. Cohen,S., Regev,A. and Lavi,S. (1997) Small polydispersed circular DNA (spcDNA) in human cells: association with genomic instability. *Oncogene*, **14**, 977–985.
72. Gaubatz,J.W. (1990) Extrachromosomal circular DNAs and genomic sequence plasticity in eukaryotic cells. *Mutat. Res.*, **237**, 271–292.
73. Stinchcomb,D.T., Thomas,M., Kelly,J., Selker,E. and Davis,R.W. (1980) Eukaryotic DNA segments capable of autonomous replication in yeast. *Proc. Natl. Acad. Sci. USA*, **77**, 4559–4563.