

Replication infidelity during a single cycle of Ty1 retrotransposition

ABRAM GABRIEL*[†], MICHÈLE WILLEMS*, EMILIE H. MULES*, AND JEF D. BOEKE[‡]

*Department of Molecular Biology and Biochemistry, Rutgers University, Piscataway, NJ 08855; and [‡]Department of Molecular Biology and Genetics, The Johns Hopkins School of Medicine, Baltimore, MD 21205

Communicated by Maxine F. Singer, Carnegie Institution of Washington, Washington, DC, January 11, 1996 (received for review November 3, 1995)

ABSTRACT Retroviruses undergo a high frequency of genetic alterations during the process of copying their RNA genomes. However, little is known about the replication fidelity of other elements that transpose via reverse transcription of an RNA intermediate. The complete sequence of 29 independently integrated copies of the yeast retrotransposon Ty1 (173,043 nt) was determined, and the mutation rate during a single cycle of replication was calculated. The observed base substitution rate of 2.5×10^{-5} bp per replication cycle suggests that this intracellular element can mutate as rapidly as retroviruses. The pattern and distribution of errors in the Ty1 genome is nonrandom and provides clues to potential *in vivo* molecular mechanisms of reverse transcriptase-mediated error generation, including heterogeneous RNase H cleavage of Ty1 RNA, addition of terminal nontemplated bases, and transient dislocation and realignment of primer-templates. Overall, analysis of errors generated during Ty1 replication underscores the utility of a genetically tractable model system for the study of reverse transcriptase fidelity.

Reverse transcription is a notoriously error-prone process. Mutations occurring during retroviral replication include simple base substitutions and frameshifts as well as complex deletions, deletions with insertions, and hypermutations (1–6). This low fidelity of replication is presumed to be the basis for the rapid genome evolution of retroviruses as well as for the ability of human immunodeficiency virus to evade its host immune system and develop drug resistance. Biochemical studies using purified retroviral reverse transcriptases (RTs) have elucidated multiple potential error-generating mechanisms, including extension past mismatched bases (7–11), poor discrimination of dNTPs (12–14), lack of associated 3' to 5' exonuclease activities (9, 15), slippage of primer-templates (16, 17), and addition of nontemplated bases at the ends of template strands (18, 19).

Retroviruses are the most highly visible members of a ubiquitous and varied confederation of genetic elements linked by their capacity to replicate by reverse transcription of an RNA intermediate (20, 21). While recent biochemical analyses have revealed novel priming mechanisms associated with several nonretroviral RTs (22–26), no study has addressed the rate or spectrum of errors occurring during a cycle of nonretroviral retrotransposition. Such studies are impeded by the polymorphic nature of multicopy elements in the genome (27, 28), their variable expression, and the difficulty of identifying events resulting from single rounds of replication. However, these problems could be obviated using a model system where transposition is limited to a single, marked element, where expression of the multiple endogenous copies is repressed, and where transposition events are trapped and prevented from secondary replication. The retrotransposon

Ty1, from the yeast *Saccharomyces cerevisiae*, provides just such a system.

Ty1 is one of the most intensively studied retrotransposons. Its basic long terminal repeat (LTR)-containing structure and mode of replication are strikingly similar to vertebrate retroviruses (29). Assay systems have been developed to overexpress marked plasmid-based copies of Ty1 under the inducible GAL1 promoter (i.e., pGAL-Ty1 constructs) and to detect their subsequent transposition (30–32). One such approach, the *his3Δ4* system, takes advantage of a promoterless allele of *HIS3* on a plasmid (see Fig. 1). In a yeast strain deleted for the *HIS3* gene and hence auxotrophic for histidine, a transposition event that inserts a copy of Ty1 close to the 5' end of *his3Δ4* on the plasmid results in histidine prototrophy, an easily selectable phenotype. Independent Ty1-containing, His⁺ recipient plasmids can be subsequently isolated. By performing this assay in a mutant strain of yeast that is unable to express endogenous Ty1 elements (*spt3*), all transpositions are derived from the marked donor element. We have made use of these genetic features of the Ty1 system to directly determine the rate and spectrum of errors generated during a single cycle of Ty1 replication. This work provides a basis for our long-term goal of understanding the genetic determinants of retrotransposition fidelity.

MATERIALS AND METHODS

Strains and Plasmids. Donor plasmid (pGTy1H3-*lacO*), recipient plasmid (pAB100), and host strain JB451 (*MATa spt3-101 lys2 his3Δ200, trp1-289, ura3-52*) have been described (30, 33). Twenty-nine independent recipient plasmids containing marked Ty1 elements were retransformed into *Escherichia coli* strain TG1, and large amounts of plasmids were prepared using Qiagen (Chatsworth, CA) columns for sequencing. The 29 plasmids have been designated AGE139–AGE161, AGE163, AGE164, and AGE174–AGE177.

Sequencing Strategy. Plasmid templates recovered from independent His⁺ yeast colonies were sequenced using 31 Ty1-specific oligonucleotide primers spaced at ≈ 200 -nt intervals. (Primer sequences and positions are available upon request from the authors.) Primers were 5' end-labeled with [γ -³²P]ATP and T4 polynucleotide kinase, and all sequencing reactions were done with *Taq* polymerase using cycle sequencing protocols recommended by BRL. Sequencing reactions were separated on 8 M urea/6% polyacrylamide gels using sharks tooth combs and 96 lanes per gel. For most gels, 20 lanes of adenine reactions, followed by 20 lanes each of cytosine, guanine and thymine reactions, were loaded so as to directly compare different independent transposition events. Additionally, for each gel, one template was loaded as ACGT to aid in reading the sequence. An average of 240 bases were clearly readable from each primer.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: LTR, long terminal repeat; pbs, primer-binding site; RT, reverse transcriptase.

[†]To whom reprint requests should be addressed.

In all cases where a mutation was identified, the original His⁺ yeast strain was regrown, DNA was reextracted, the target plasmid was recovered into *E. coli*, and the mutation site was resequenced to confirm that the observed mutation had not arisen subsequent to the original plasmid extraction from yeast.

Calculation of Mutation Rates. Rate determinations were based on the methods described by Drake (34, 35). The average mutation rate per base pair per Ty1 replication cycle (μ_b) equals the total number of base substitutions identified in the sample divided by the total number of base pairs replicated. This denominator was calculated as follows. The length of the marked Ty1 element is 5967 bases (5918-bp Ty1 plus 49-bp *lacO* insertion). Because RNA polymerase transcription begins inside the 5' LTR (consisting of regions designated U3, R, U5) at the U3-R border and ends at the R-U5 border of the 3' LTR, (see Fig. 4), 5679 bases are copied. RT regenerates duplicate LTRs during synthesis of both strands, equivalent to a total of 11,934 bases. Therefore, total bases copied is equal to 17,613 per transposition event multiplied by 29 transposition events, or 510,777. The mutation rate per Ty1 genome per retrotransposition cycle (μ_g) is equal to the base pair mutation rate multiplied by the length of the Ty1 genome. The mutation rate of $\approx 1.3 \times 10^{-6}$ per Ty1 element per cycle of cellular replication is calculated assuming an estimate of the yeast genomic mutation rate of 0.003 (34), multiplied by the length of a Ty1 element (5918 bp), and divided by the yeast-genome size (1.38×10^7 bp).

RESULTS AND DISCUSSION

Strategy. We previously recovered 29 independent Ty1 transposition events during a single cycle of replication using the strategy diagrammed in Fig. 1 (33). A yeast strain deleted for

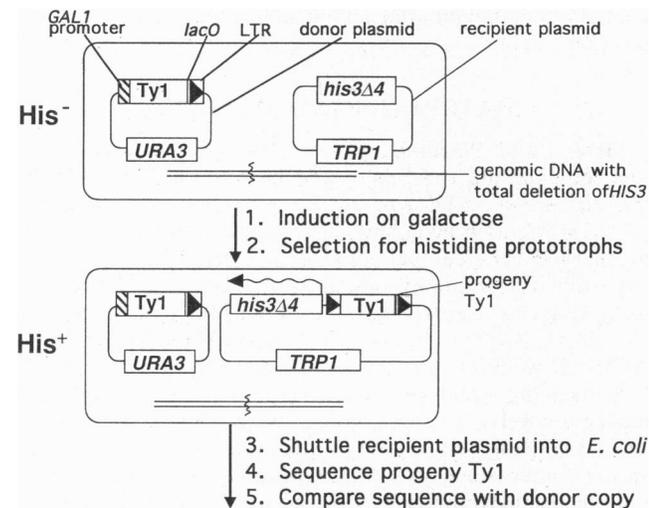


FIG. 1. Overall strategy for determining Ty1 mutation rate during a single cycle of replication. After induction of Ty1 transposition by growth of yeast on plates containing synthetic complete media (SC) lacking tryptophan and uracil, and supplemented with galactose, cells were replica-plated to SC-his/glucose, and His⁺ papillae were isolated. A fraction of the His⁺ papillae were due to transposition of Ty1 directly upstream of the *his3Δ4* allele, resulting in increased *his3Δ4* transcriptional activity (\sim). After segregating away the *URA3*-containing donor plasmid, total DNA from the His⁺ papillae was extracted and used to transform *lac⁺* *E. coli* strain HB101 to ampicillin resistance. Colonies which turned blue (due to repressor titration) on 5-bromo-4-chloro-3-indolyl β -D-galactosidase (i.e., X-gal) plates contained the *lacO* marker and were further analyzed by restriction mapping (detailed methods described in refs. 30 and 33). The Ty1 portion of each recipient plasmid was then sequenced in its entirety (see Fig. 2 for example) and analyzed for changes from the donor sequence.

the *HIS3* gene and hence auxotrophic for histidine and containing two plasmids (termed donor and target plasmids) was induced to undergo high levels of transposition of a *lacO*-marked Ty1 element driven by the *GAL1* promoter present on the pGTYH3-*lacO* donor plasmid. The target plasmid contains the promoterless, and consequently nonfunctional, *his3Δ4* allele. Insertion of Ty1 directly upstream of the plasmid-borne *his3Δ4* allele results in expression of this gene and is detected by selection for histidine prototrophy. Transposon-containing target plasmids were transformed into *E. coli*, and purified plasmids were subsequently used as sequencing templates.

These 29 independent transposition events derived from a *GAL1* promoter-driven, marked Ty1 construct (Fig. 1) in an *spt3* strain background and, therefore, represented ideal starting material for a study of events occurring during a single cycle of transposition. The *SPT3* gene encodes a factor presumed necessary for Ty1 transcription from its normal LTR promoter; in an *spt3* strain, endogenous Ty1 RNA levels are dramatically reduced (36, 37). Concomitantly, in *spt3* strains, spontaneous transposition of native Ty1s is decreased more than 100-fold (32, 33). Consequently, the vast majority of transposition events in this strain are *GAL*-Ty1-driven. In a previous study (33), much of the sequence variation that occurs during Ty1-*lacO* retrotransposition was shown to result from coexpression of chromosomal (LTR-driven) Ty1 elements, presumably from copackaging of marked and unmarked Ty1 transcripts, followed by RT template switching during replication. This source of heterogeneity is eliminated by introduction of an *spt3* mutation, which eliminates LTR-driven Ty1 transcription, but not *GAL*-Ty1 transcription (33). Further, in strain JB451, single rounds of transposition are assured because transposed copies of the marked Ty1 elements lack the *GAL1* promoter and are therefore prevented from further expression by the *spt3* mutation.

A further advantage to our assay system is that there is minimal selection bias against mutations generated during retrotransposition. While mutations might occur during RNA polymerase II-directed transcription of the *GAL*-Ty1 construct, only those resulting in deletions of the RNA packaging signal, the LTR sequences, or the primer-binding site (pbs) would interfere with transposition and be selected against in our screen. Assuming the presence of a faithful transcript, only sequence changes during reverse transcription that affect the ability of Ty1 to integrate into the plasmid or that affect our ability to detect a transposition event (e.g., changes in the Ty1 enhancer that no longer induce expression of *HIS3* or changes that inactivate the *lacO* marker) will potentially be selected against or go undetected. These sequences represent <5% of the Ty1 genome. Therefore, our derived rates may slightly underrepresent the actual mutation rate and should be considered a minimum estimate.

Determination of Mutation Rate and Spectrum. Using a collection of 31 oligonucleotide primers, we sequenced through the entire 5967 bases of 29 independent Ty1 transposition events, representing a sequence target of 173,043 bases. An example of our error-detection strategy is shown in Fig. 2, and a summary of the findings is shown in Fig. 3. We identified a total of 13 substitutions in 10 of the 29 independent transposition events by visually scanning the autoradiographs. No element had more than a single site of base substitution—i.e., in eight elements a single base was changed, whereas the other two involved changes of two and three consecutive bases. The substitutions resulted in seven missense mutations and three silent changes. Five of 13 substitutions were C \rightarrow A transversions, and transversions predominated over transitions by a ratio of 2.25:1. We observed no frameshifts, deletions, or insertions.

Drake (34, 35) has defined several parameters for mutation rates that facilitate comparison of data derived using a variety of estimation methods and from many different biological

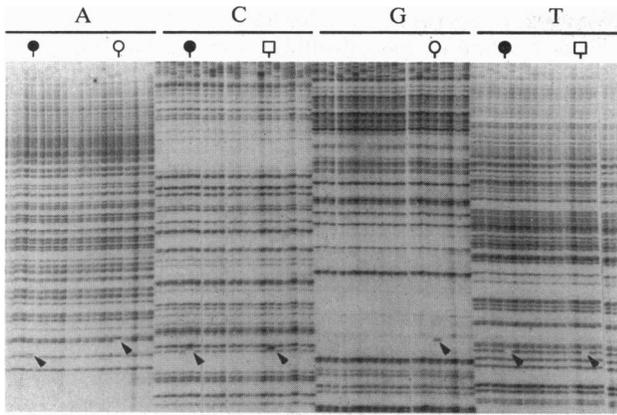


FIG. 2. Sequence scanning for mutations in transposed *Ty1* elements. Autoradiogram of 20 *Ty1* template sequences using a primer 120 bases downstream of the U5-pbs border. Base substitutions around the U5-pbs border are marked by arrowheads. Mutant templates correspond to those listed in Fig. 3: ●, AGE154; ○, AGE143; □, AGE144. Twenty adenine reactions, followed by cytosine, guanine, and thymine reactions were loaded so as to directly compare different independent transposition events and to identify compensatory base changes. Note the variable bands at the top of the gel that correspond to the heterogeneous plasmid insertion sites flanking the 5' end of *Ty1*.

systems. Using these definitions, we calculate the average mutation rate at 2.54×10^{-5} per bp per replication cycle and at 0.15 per *Ty1* genome per replication cycle. Four retroviruses have been analyzed for errors arising during single replication cycles (1–3, 5, 39–41). Based on Drake's compilation and analysis of these disparate results, the overall *Ty1* genomic mutation rate appears to agree well with those of the retroviruses (median of 0.2), despite the lack of *Ty1* errors other than base substitutions. This finding is not consistent with

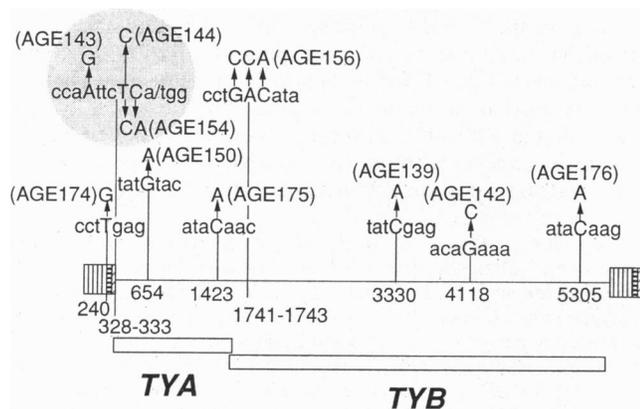


FIG. 3. Summary of observed substitution events. Scale drawing of *Ty1*, its coding domains, and its LTRs (vertical, horizontal, and diagonal stripes refer to U3, R, and U5, respectively). Sites of mutations are marked by vertical lines and precise nucleotide positions, based on the sequence coordinates of *Ty1*-H3 (38), are listed beneath each vertical line. Each inset contains the wild-type sequence with the altered base capitalized. An arrow points to the substituted base and lists the plasmid clone from which it was derived (e.g., AGE174). The shaded circle highlights the proposed mutational hot spot. The nucleotide positions of all other sequences in the 29 transposed copies were identical to the previously determined sequence of *Ty1*-H3 except that the 5' U3 sequence was identical to the 3' U3 sequence and a 49-base *lacO* marker was present in both donor and progeny *Ty1* elements. In two cases where mutations resulted in the gain or loss of a restriction site (AGE144 and AGE139) the preexisting presence or absence of the restriction site in genomic *Ty1* elements was checked by Southern blot analysis of the parental yeast strain. Neither mutation was present at levels within the detection limits of the system (≈ 1 copy in 20).

arguments that retroviruses have evolved high mutation rates specifically to evade their hosts, because the intracellular *Ty1* happily coexists with its host, yet generates a high level of variability relative to the host. An estimate for the rate at which a single *Ty1* element mutates per cellular replication cycle is on the order of 1.3×10^{-6} (see *Materials and Methods* for the calculation), which is 115,000 times lower than that obtained here during *Ty1* transposition. Rather, our findings suggest that both retroviruses and retrotransposons share a common strategy for rapid genome evolution.

The apparent lack of genetic alterations other than base substitutions may simply reflect the small number of independent events (i.e., 29 transpositions) analyzed. However, it may be a real phenomenon; interestingly, the frameshifts, deletions, and insertions associated with single cycles of retroviral replication have either been observed using artificial marker sequences such as *lacZ* α , *neo*, and *tk* or are present in noncoding retroviral sequences (1–3, 39, 41). Codon usage within coding regions may have subtly evolved to reduce the number of homopolymeric runs that might result in RT slippage or to minimize RNA secondary structure that could lead to RT pausing. Therefore, while the genomes of both *Ty1* and retroviruses may have coevolved with their RTs to minimize deleterious mutations, these types of errors would still be observable in artificial sequences. We are currently examining frameshift fidelity during retrotransposition using non-*Ty1*-related marker genes cloned into *Ty1* to analyze this phenomenon directly.

Insights from the Nature of the Base Substitution Hot Spot.

Substitutions are distributed nonrandomly along the *Ty1* sequence (Fig. 3). Four of the 10 mutational events are clustered in the 5' LTR; three of these occur within 7 bases of the 5' U5-pbs border, and the fourth mutation occurs at the 5' U3-R border (Fig. 3). This finding suggests that a mutational hot spot for *Ty1* replication exists in the 5' LTR and is supported by an examination of the U5-pbs border region for the 19 full-length *Ty1* and *Ty2* elements so far sequenced as part of the ongoing *S. cerevisiae* Systematic Chromosome Sequencing Project. The segment just upstream of the pbs is highly variable (only 6 of 23 nt are invariant) compared to the nearly homogeneous sequences present just upstream and downstream of this segment (20 and 22 of 23 nt invariant, respectively).

During which steps in the *Ty1* replication process are these errors made? While the detailed process of *Ty1* replication has not been fully dissected, available data indicate a general mechanism highly analogous to retroviral replication (42). In considering possible steps, it is noteworthy that all four of the substitutions in the 5' LTR are asymmetric—i.e., none are present in the 3' LTR. Further, all four of these events occur close to potential RT "pause sites," resulting from RT reaching a template end. It is most plausible that the mutations found in the 5' LTR are generated either during plus-strand strong-stop DNA synthesis (asterisk in Fig. 4E) or during completion synthesis of the 5' LTR minus-strand (asterisks in Fig. 4E and F). We can rule out the possibility that the errors are made by RT during the initial replication step of copying the U5 RNA template (minus-strand strong-stop DNA synthesis, Fig. 4C), since such errors would be copied symmetrically into both LTRs. It is also possible that the high frequency of errors is caused by RNA polymerase II while copying this region of the donor DNA (asterisk in Fig. 4B), but such a scenario would require that a specific, short region of sequence upstream of the pbs is particularly prone to a high level of unfaithful transcription.

Based on the specific clues provided by the distribution of *Ty1* errors, we believe that the mutational hot spot near the 5' U5-pbs border is most likely generated during the synthesis of the 5' LTR minus-strand (asterisks in Fig. 4E and F). During minus-strand strong-stop DNA synthesis, RNase H cleaves the RNA from the RNA-DNA duplex, leaving the U5-pbs border

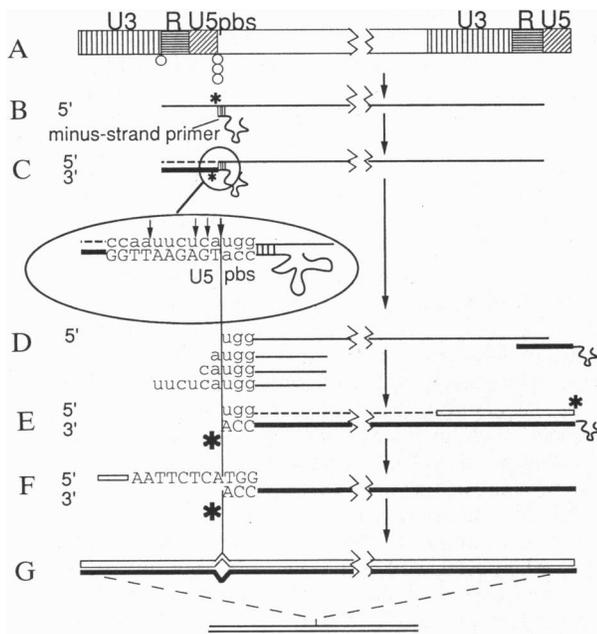


FIG. 4. The *Ty1* replication process and potential steps when base substitutions near the U5-pbs border could be generated. (A) Structure of the double-stranded *Ty1* DNA, emphasizing the symmetrical 5' and 3' LTR structure and the pbs next to the 5' U5 domain. Circles represent sites of mutations found in this study in the 5' LTR. (B) Structure of the *Ty1* full-length RNA (narrow solid line), bound to its $tRNA_{Met}$ minus-strand primer. Asterisk signifies that mutations may be generated during transcription of the full-length RNA (see text for details). (C) Formation of minus-strand strong-stop DNA (heavy solid line linked to $tRNA_{Met}$) together with RNase H cleavage of the RNA (dashed line). Asterisk signifies that mutations may be generated during copying of the RNA template (see text for details). (Inset) To explain the fact that substitutions occur at multiple sites upstream of the pbs, we propose that the initial cleavage site for *Ty1* RNase H is heterogeneous (arrows). (D) Intramolecular minus-strand transfer. Multiple potential RNA 5' ends are shown, resulting from heterogeneous RNase H cleavage. (E) Minus-strand cDNA synthesis with concomitant RNase H cleavage (dashed line) and plus-strand strong-stop DNA synthesis (heavy unfilled line). Smaller asterisk signifies that mutations may be generated during plus-strand strong-stop DNA synthesis; larger asterisk signifies that mutations may be generated as minus-strand DNA reaches the RNA template end (see text for details). (F) Intramolecular plus-strand transfer is followed by completion synthesis of the 5' LTR minus strand. Asterisk signifies that mutations may be generated during reinitiation of synthesis after strand transfer from the RNA template to the plus-strand cDNA (see text for details). (G) Completed double-stranded heteroduplex DNA integrates into host DNA where it undergoes repair or replication via host DNA polymerases. Size of asterisks corresponds to the likelihood of each step generating the observed mutations.

as the 5' end of the RNA genome (Fig. 4D). Before plus-strand strong-stop transfer, the border is the 3' terminus for minus-strand cDNA synthesis (Fig. 4E). After plus-strand transfer, the border is the point of reinitiation of minus-strand synthesis, now with a DNA template (Fig. 4F). The finding that the hot spot is a segment of several bases upstream of the U5-pbs border rather than a single base suggests that the positions of the termini are heterogeneous. This hypothesis could most simply be accounted for by imprecise cleavage of the 5' end of the template RNA by *Ty1* RNase H during minus-strand strong-stop synthesis [Fig. 4C (Inset) and D], a phenomenon that appears to occur naturally during *Ty1* replication (E.H.M. and A.G., unpublished data). An association between replication errors and RT pause sites (16, 43) or strand-transfer sites (18, 19, 44, 45) has been suggested for retroviruses based on *in vitro* studies and theoretical considerations but has not yet been detected *in vivo* (46).

What error-generating mechanisms could accommodate the terminal location of the substitutions at the hot spot? One possibility is that *Ty1* RT mediates terminal nontemplated base addition at this site during minus-strand synthesis. After plus-strand transfer, these nontemplated bases would form a terminally mismatched region that RT must extend past to complete the 5' LTR minus-strand. Biochemical data from retroviruses support the feasibility of this model; several retroviral RTs can extend a primer end one or more bases beyond a 5' template end *in vitro* (18, 19, 47), and the ability of RTs to extend past mismatched bases is a well-characterized *in vitro* phenomenon (7–11). Further, nontemplated base addition has been proposed to explain the occasional presence of extra bases observed at retroviral circle junctions (48–50).

An alternative explanation for two of the three mutations at the U5-pbs border (AGE154 and AGE144) is "dislocation mutagenesis" (51). This mechanism, which is usually associated with nucleotide repeats, involves the transient looping out (dislocation) of a base in one strand close to the primer end, followed by primer extension from an incorrect, neighboring, template base. After primer-template realignment, extension beyond the resulting mismatched primer end leads to base substitution in the growing strand (51–53).

Finally, the T \rightarrow G transversion in the region of the U3-R junction (Fig. 3, AGE174) could also be explained by either blunt-end addition of a nontemplated nucleotide or dislocation mutagenesis via copying of the preceding guanine residue. In this case the observed asymmetry could be explained if the initial minus-strand transfer was intermolecular, resulting in the 3' end of the minus-strand cDNA being at the U3-R border, before plus-strand transfer. In agreement with our model that substitutions are favored at termini is the fact that this same base at the U3-R junction is known to be highly polymorphic in natural *Ty1* elements (38).

Our study has revealed an *in vivo* hot spot for base substitutions, which could be explained by heterogeneous RNase H cleavage of the 5' RNA template, followed by either nontemplated base addition or terminal misincorporation via a dislocation mechanism. It will be important to determine whether such asymmetric hot spots are a unique consequence of *Ty1* RT-mediated replication in yeast cells or may be generalizable to additional reverse-transcribed genetic elements, including other retrotransposons, retroviruses, and hepadnaviruses.

We thank J. Drake, S. Goff, and V. Laueremann for helpful discussions; J. Dougherty and B. Preston for reviewing the manuscript; J. Leybenzon and R. Robichaud for technical assistance; and N. Yoshida and A. Kalogkhar for assistance with graphics. The work was supported in part by the Charles and Johanna Busch Endowment, the Lucille P. Markey Charitable Trust, and National Science Foundation Grant MCB9317281 (to A.G.) as well as by National Institutes of Health Grant RO1 GM36481 and an American Cancer Society Faculty Research Award (to J.D.B.). A.G. is a Lucille P. Markey Scholar.

1. Pathak, V. K. & Temin, H. M. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6019–6023.
2. Dougherty, J. P. & Temin, H. M. (1988) *J. Virol.* **62**, 2817–2822.
3. Pathak, V. K. & Temin, H. M. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6024–6028.
4. Vartanian, J. P., Meyerhans, A., Asjö, B. & Wain-Hobson, S. (1991) *J. Virol.* **65**, 1779–1788.
5. Leider, J. M., Palese, P. & Smith, F. I. (1988) *J. Virol.* **62**, 3084–3091.
6. Varela-Echavarria, A., Prorock, C. M., Ron, Y. & Dougherty, J. P. (1993) *J. Virol.* **67**, 6357–6364.
7. Perrino, F. W., Preston, B. D., Sandell, L. L. & Loeb, L. A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8343–8347.
8. Preston, B. D., Poiesz, B. J. & Loeb, L. A. (1988) *Science* **242**, 1168–1171.
9. Roberts, J. D., Bebenek, K. & Kunkel, T. A. (1988) *Science* **242**, 1171–1173.

10. Mendelman, L. V., Petruska, J. & Goodman, M. F. (1990) *J. Biol. Chem.* **265**, 2338–2346.
11. Yu, H. & Goodman, M. F. (1992) *J. Biol. Chem.* **267**, 10888–10896.
12. Mendelman, L. V., Boosalis, M. S., Petruska, J. & Goodman, M. F. (1989) *J. Biol. Chem.* **264**, 14415–14423.
13. Ricchetti, M. & Buc, H. (1990) *EMBO J.* **9**, 1583–1593.
14. Zinnen, S., Hsieh, J.-C. & Modrich, P. (1994) *J. Biol. Chem.* **269**, 24195–24202.
15. Battula, N. & Loeb, L. A. (1976) *J. Biol. Chem.* **251**, 982–986.
16. Bebenek, K., Abbotts, J., Roberts, J. D., Wilson, S. H. & Kunkel, T. A. (1989) *J. Biol. Chem.* **264**, 16948–16956.
17. Bebenek, K., Abbotts, J., Wilson, S. H. & Kunkel, T. A. (1993) *J. Biol. Chem.* **268**, 10324–10334.
18. Peliska, J. A. & Benkovic, S. J. (1992) *Science* **258**, 1112–1118.
19. Patel, P. H. & Preston, B. D. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 549–553.
20. Doolittle, R. F., Feng, D.-F., Johnson, M. S. & McClure, M. A. (1989) *Q. Rev. Biol.* **64**, 1–29.
21. Xiong, Y. & Eickbush, T. H. (1990) *EMBO J.* **9**, 3353–3362.
22. Luan, D. D., Korman, M. H., Jakubczak, J. L. & Eickbush, T. H. (1993) *Cell* **72**, 595–605.
23. Wang, G. H. & Seeger, C. (1992) *Cell* **71**, 663–670.
24. Tavis, J. E. & Ganem, D. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 4107–4111.
25. Levin, H. L. (1995) *Mol. Cell. Biol.* **15**, 3310–3317.
26. Zimmerly, S., Guo, H., Perlman, P. S. & Lambowitz, A. M. (1995) *Cell* **82**, 545–554.
27. Steinhauer, D. A. & Holland, J. J. (1987) *Annu. Rev. Microbiol.* **41**, 409–433.
28. Casacuberta, J. M., Vernhettes, S. & Grandbastien, M. (1995) *EMBO J.* **14**, 2670–2678.
29. Boeke, J. D. (1989) in *Mobile DNA*, eds Berg, D. E. & Howe, M. M. (Am. Soc. for Microbiol., Washington, DC), pp. 335–374.
30. Boeke, J. D., Garfinkel, D. J., Styles, C. A. & Fink, G. R. (1985) *Cell* **40**, 491–500.
31. Boeke, J. D., Xu, H. & Fink, G. R. (1988) *Science* **239**, 280–282.
32. Curcio, M. J. & Garfinkel, D. J. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 936–940.
33. Boeke, J. D., Styles, C. A. & Fink, G. R. (1986) *Mol. Cell. Biol.* **6**, 3575–3581.
34. Drake, J. W. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 7160–7164.
35. Drake, J. W. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 4171–4175.
36. Winston, F., Durbin, K. J. & Fink, G. R. (1984) *Cell* **39**, 675–682.
37. Winston, F. & Minehart, P. L. (1986) *Nucleic Acids Res.* **14**, 6885–6900.
38. Boeke, J. D., Eichinger, D., Castrillon, D. & Fink, G. R. (1988) *Mol. Cell. Biol.* **8**, 1432–1442.
39. Varela-Echavarría, A., Garvey, N., Preston, B. D. & Dougherty, J. P. (1992) *J. Biol. Chem.* **267**, 24681–24688.
40. Monk, R. J., Malik, F. G., Stokesberry, D. & Evans, L. H. (1992) *J. Virol.* **66**, 3683–3689.
41. Mansky, L. M. & Temin, H. M. (1994) *J. Virol.* **68**, 494–499.
42. Gabriel, A. & Boeke, J. D. (1993) in *Reverse Transcriptase*, eds Skalka, A. M. & Goff, S. P. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 275–328.
43. Klarmann, G. J., Schaubert, C. A. & Preston, B. D. (1993) *J. Biol. Chem.* **268**, 9793–9802.
44. Peliska, J. A. & Benkovic, S. J. (1994) *Biochemistry* **33**, 3890–3895.
45. Temin, H. M. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 6900–6903.
46. Zhang, J. & Temin, H. M. (1994) *J. Virol.* **68**, 2409–2414.
47. Clark, J. M. (1988) *Nucleic Acids Res.* **16**, 9677–9686.
48. Whitcomb, J. M., Kumar, R. & Hughes, S. H. (1990) *J. Virol.* **64**, 4903–4906.
49. Olsen, J. C. & Swanstrom, R. (1985) *J. Virol.* **56**, 779–789.
50. Smith, J. S., Kim, S. & Roth, M. J. (1990) *J. Virol.* **64**, 6286–6290.
51. Kunkel, T. A. & Alexander, P. S. (1986) *J. Biol. Chem.* **261**, 160–166.
52. Fresco, J. R. & Alberts, B. M. (1960) *Proc. Natl. Acad. Sci. USA* **46**, 311–321.
53. Fowler, R. G., Degnen, G. E. & Cox, E. C. (1974) *Mol. Gen. Genet.* **133**, 179–191.