One Retroviral RNA Is Sufficient for Synthesis of Viral DNA

JEFFERY S. JONES,[†] ROBERT W. ALLAN, AND HOWARD M. TEMIN*

McArdle Laboratory for Cancer Research, University of Wisconsin, Madison, Wisconsin 53706

Received 25 August 1993/Accepted 15 October 1993

We used previously characterized spleen necrosis virus-based retroviral vectors and helper cells to study the strand transfers that occur during the reverse-transcription phase of a single cycle of retroviral replication. The conditions used selected only for formation of an active provirus rather than for expression of multiple drug resistance markers. In nonrecombinant proviruses the minus- and plus-strand DNA primer transfers were almost completely intramolecular. However, as previously reported, recombinant proviruses contained approximately equal proportions of inter- and intramolecular minus-strand DNA primer transfers. Thus, we conclude that in the absence of recombination, one molecule of retroviral RNA is sufficient for viral DNA synthesis. Large deletions and deletions with insertions were detected primarily at a limited number of positions which appear to be hot spots for such events, the primer binding site and regions containing multiple inverted repeats. At these hot spots, the rate of deletions and deletions with insertions visible with PCR was about 10% per genome per replication cycle. Other deletions and deletions with insertions (detectable with PCR) occurred at a rate of about 0.5%/kb per replication cycle. Crossovers occurred at a rate of about 6%/kb per replication cycle under single-selection conditions. This rate is comparable to the rate that we reported previously under double-selection conditions, indicating that retroviral homologous recombination is not highly error prone. The combined rates of deletions and deletions with insertions at hot spots (10% per genome per replication cycle) and other sites (0.5%/kb per replication cycle) and the rate of crossovers (6%/kb per replication cycle) indicate that on average, full-size (10-kb) type C retroviruses undergo an additional or aberrant strand transfer about once per cycle of infection.

Retroviruses are RNA viruses that replicate through a DNA intermediate. The process of copying the RNA form of the virus genome into the DNA form, called reverse transcription, is carried out by the virally encoded enzyme reverse transcriptase (RT) (3, 23, 30, 32). Many of the details of the mechanism of reverse transcription have not yet been elucidated; however, a widely accepted model of the process has been proposed (3, 6, 23, 27, 30, 32). As summarized in Fig. 1, this model requires that two strand transfers occur during the process of reverse transcription. Alterations such as base pair substitutions, frameshifts, deletions, deletions with insertions, nonhomologous recombination, and homologous recombination occur at high frequency during retroviral reverse transcription, resulting in high rates of retroviral variation (1, 32). The necessity for RT to be able to switch templates during reverse transcription may be the source of these alterations in the retrovirus genome (29).

A characteristic of retroviruses that makes them unique among viruses is that retrovirus particles contain two complete copies of the viral genetic material (3, 23, 30, 32). This genetic material is present as a dimer composed of two single strands of retroviral plus-strand RNA joined in a noncovalent linkage. Many other viruses recombine by accumulating within the host cell large numbers of viral genomes which act as substrates for recombination (21); however, the retroviral replication cycle necessitates copackaging for efficient recombination to occur (9, 25). In fact, the function of the dimer RNA may be to allow recombination to occur (16, 28). Homologous recombination is believed to be particularly important to retroviruses because it

also allows them to recover from damage to the viral RNA carried within the virus particle (2). The high rate of variation is also thought to be valuable to retroviruses because it allows retroviruses to alter or increase their host range, escape host defenses, and escape antiretroviral therapy. Characteristics that contribute to efficient retroviral recombination would be under strong positive selective pressure because recovery from damage and increased variation (by reassorting mutations) increase viral fitness. We are interested in the strand transfers required for reverse transcription (Fig. 1) as well as those that lead to recombination, deletions, or deletions with insertions.

Our laboratory has previously described retroviral vectors and helper cells based on spleen necrosis virus (SNV), an avian retrovirus similar to murine type C retroviruses (4, 5, 9–12). The helper cells provide retroviral proteins *in trans*. In our previous studies, we used two nearly identical retroviral vectors that contain all *cis*-acting sequences necessary for retroviral replication, the coding regions for two drug resistance markers, and a series of restriction site polymorphisms (Fig. 2). Because each vector contained only one functional resistance marker, the single-resistance (single^r) titer was compared to the double-resistance (double^r) titer, and the recombination rate for SNV was shown to be about 1 to 2%/kb per replication cycle (9–12). The recombinant proviruses contained in double^r cells were then characterized. Although our previous protocol allowed us to study relatively large numbers of recombinant proviruses, the majority of the genome was required to form a selectable provirus, and thus, most variants would have been lost. Furthermore, the recombinant proviruses in double^r cells are a special class of proviruses, which in those experiments account only for 1 to 2% of all proviruses, and thus, they may not be representative of the larger population of proviruses.

The experimental system used in this study was identical to that used previously (9–12) except that a single-selection protocol was used. As previously, we characterized the proviral DNA formed in a single cycle of retroviral replication. These

* Corresponding author. Mailing address: McArdle Laboratory, University of Wisconsin, 1400 University Ave., Madison, WI 53706. Phone: (608) 262-1209. Fax: (608) 262-2824. Electronic mail address: temin@oncology.wisc.edu.

[†] Present address: Department of Molecular Pathology, University of Texas M. D. Anderson Cancer Center, Houston, TX 77030.

FIG. 1. Reverse transcription. Thin lines and lowercase letters indicate regions of the retroviral RNA important for reverse transcription, and heavy lines and uppercase letters indicate DNA. r, the repeated region found at both ends of the retroviral RNA; ppt, polyuridine tract; u5 and u3, the unique regions found at the 5' and 3' ends of the RNA, respectively. (A) RT utilizes a cellularly derived tRNA annealed at the PBS as a primer to initiate minus-strand DNA synthesis. (B) Reverse transcription approaches the end of the template RNA. (RT also has RNase H activity, which specifically degrades RNA bound to DNA, thus degrading most or all of the r and u5 regions of the RNA template.) (C) The newly formed minus-strand DNA primer (formerly called the minus-strand strong-stop DNA) transfers to the 3' end of the retroviral RNA template by R region complementarity. Since retrovirus virions contain an RNA dimer, this jump can be either from the end of one RNA strand to the end of the other RNA strand, called an intermolecular minus-strand DNA primer transfer, or from the end of one RNA strand to the other end of the same RNA strand, called an intramolecular minus-strand DNA primer transfer. The primer transfer diagrammed is an intermolecular minus-strand DNA primer transfer, which is what usually occurs (see text). (D) Minus-strand DNA synthesis continues after the transfer, and RNase H makes a specific nick in the RNA (vertical arrow) in a region that contains a run of purine bases, the ppt, located 5' of the u3 region. (E) This nick acts as a primer to initiate plus-strand DNA synthesis. Plus-strand DNA synthesis continues through U3, R, and U5 and into the tRNA. (F) This plus-strand DNA, called the plus-strand DNA primer (formerly called the plus-strand strong-stop DNA), then transfers. As with the minus-strand DNA primer transfer, the plus-strand DNA primer transfer can be to the other end of the same minus-strand DNA (an intramolecular plus-strand DNA primer transfer) or, very rarely, to the minus-strand DNA of the other RNA template (an intermolecular plus-strand DNA primer transfer). (G) Following the plus-strand DNA primer transfer, DNA synthesis continues until both the minus- and plus-strand DNAs are completed.

new conditions allow (i) study of the strand transfers that occur during reverse transcription of the entire population of viral genomes, rather than just the 1 to 2% that yield selectable recombinants; (ii) study of nonselected regions of the genome, which are thus free to undergo variation; and (iii) detection of

FIG. 2. Important features of the retroviral vectors and locations of PCR primers. WH13, shown in white, expresses a functional hygromycin resistance gene (*hyg*), but the neomycin resistance gene (*neo*) is nonfunctional because of introduction of a frameshift mutation, indicated by the asterisk. WH204, shown in gray, expresses a functional *neo* gene, but *hyg* is nonfunctional because of introduction of a frameshift mutation, indicated by the asterisk. Approximate locations of restriction enzyme cleavage site polymorphisms, which are used to map recombinant proviruses, are indicated by vertical bars on WH204 (detailed maps are in references 10 and 11). Three sets of PCR primers were used to amplify the recombinant proviruses and are shown as arrows below the proviruses. White-tipped arrows and gray arrows at the 5' and 3' ends of the genome correspond to PCR primers that specifically recognize the different restriction sites in these locations in WH13 and WH204, respectively. Solid black arrows in the central regions of the genome represent PCR primers that recognize either vector.

recombination events that result in errors (such as deletions, deletions with insertions, and frameshift mutations) which would cause our previously employed double-selection protocol to underestimate the recombination rate.

MATERIALS AND METHODS

Nomenclature. Plasmids that contain proviral clones are identified by a "p" before the name of the construct. Viruses produced from these proviral clones do not have that letter. For example, pWH13 is the plasmid containing the proviral DNA clone that is used to produce virus WH13.

Retroviral vectors, cell culture, transfections, and infections. As previously described (10), proviral clones pWH13 and pWH204 were derived from a strain of reticuloendotheliosis virus, SNV, a murine leukemia virus-like avian retrovirus (15), and contain the *cis*-acting sequences necessary for retroviral replication. These vectors contain the bacterial neomycin resistance gene (*neo*) and the hygromycin B phosphotransferase gene (*hyg*) [4]; Fig. 2). As indicated by asterisks in Fig. 2, frameshift mutations have been introduced into *neo* of pWH13 and *hyg* of pWH204. Thus, pWH13 expresses only functional *hyg*, and pWH204 expresses only functional *neo*. Additional minor changes were introduced into pWH204 so that there are a total of 10 restriction enzyme cleavage site differences between retroviral vectors WH13 and WH204. These are indicated by vertical bars in Fig. 2.

The cell culture protocol, summarized in Fig. 3, has been previously described (9–12) except that single^r rather than double^r step 3 clones were characterized (see below). Briefly, dog osteosarcoma cell line D17 and helper cell derivatives of it were grown in Temin-modified Eagle's medium (Biologos, Naperville, Ill.) supplemented with 6% calf serum (Biologos) in humidified incubators at 6% CO₂ and at 37°C. In step 1, SNV-based retroviral helper cells, C3 (31), DAN (5), and DSDh (9) cells, were separately transfected with the proviral

FIG. 3. Protocol used to study retroviral replication and homologous recombination. In step 1, helper cells were separately transfected with the vector DNAs. In step 2, fresh helper cells were infected with vector virus produced by the step 1 cells. Step 2 cells that contained both WH13 and WH204 proviruses were isolated. In step 3, D17 target cells were infected with virus isolated from the step 2 cells. The infected D17 cells were selected for G418^r, Hyg^r, or both G418^r and Hyg^r.

clone pWH13 or pWH204 by the Polybrene-dimethyl sulfoxide method (13). In some cases plasmid pSV2α3.6, which confers ouabain resistance in mammalian cells (14), was cotransfected. One or two days posttransfection, 400 μg of G418 (a neomycin analog used in mammalian cell culture) per ml, 80 μg of hygromycin B per ml, or ouabain (10⁻⁷ M final concentration) was added to the culture media, as appropriate. At 1 to 2 weeks posttransfection, colonies were trypsinized, pooled, and replated in fresh medium lacking selective agents. Eighteen to 24 h later, virus pools were harvested and centrifuged at 3,000 × g for 10 min to remove cell debris. In step 2, the virus pools were used to infect fresh DSDh cells, 1 day postinfection selection was applied, and DSDh cell clones carrying WH13 and WH204 were isolated. These vector virus producer cell clones were maintained in the presence of chicken anti-SNV antiserum to prevent reinfection. In step 3, various dilutions of virus pools harvested from the producer clones were used to infect D17 target cells, which had been plated 1 day previously at 2 × 10⁵ cells per 60-mm-diameter dish, in the presence of 50 μg of Polybrene per ml. One day postinfection, cells were placed under selection for G418 resistance (G418^r), selection for hygromycin B resistance (Hyg^r), or selection for both G418 resistance and hygromycin B resistance (double^r). Several

weeks postinfection, colonies on the assay plates were counted to determine the titer for each of the selections. Large, well-isolated target colonies were picked from the selection plates, expanded, and further characterized.

Analysis of proviral DNA in target cells. As previously described (11, 12), the proviral DNA in the resistant D17 target cell clones was characterized by expanding the cells, lysing the cells (8), and subjecting the lysates to PCR analysis using PCR primers that have been described previously (11). The proviral DNA was amplified in three sections (as drawn in Fig. 2). Differential PCR was used in the amplification of both the 5' and 3' halves because the ends of the proviruses contained different restriction endonuclease recognition sites. The PCR products corresponding to the three sections of the provirus were mapped with restriction endonucleases.

Calculation of the composition of virion RNA. The single^r colonies could have arisen from a virion carrying two copies of one of the parental genomes (a homodimer) or one copy of each parental genome (a heterodimer). A homologous recombination event or an intermolecular strand transfer cannot be detected if the virion contains a homodimer, so we corrected the total number of proviruses assayed to remove from consideration those that were derived from virions containing homodimeric RNA. We calculated the composition of the genetic material in the virions using the Hardy-Weinberg equation (24): $G^2 + 2GH + H^2 = 1$, where G is the G418^r titer and H is the Hyg^r titer. Selection for G418^r or Hyg^r selects against the proviruses derived from virions containing the other homodimer. Because only one provirus is formed per virion (9, 16), half of the proviruses formed from heterodimers do not contain the selected drug resistance gene and thus are lost. For example, if the G418^r and Hyg^r titers were equal ($G = 0.5$ and $H = 0.5$), then the Hardy-Weinberg equation yields $0.25 + 0.5 + 0.25 = 1$. After selection with G418, the ratio of colonies arising from virions carrying homodimers to colonies arising from virions carrying heterodimers becomes 0.25 to 0.25. Thus, half of the colonies arose from homodimers, and half arose from heterodimers. The rates of recombination and of inter- or intramolecular strand transfers were determined by comparing the number of proviruses that we detected that had undergone these events to the number of proviruses that we calculated were derived from virions containing heterodimers.

RESULTS

The population of singly selected proviruses. In this study, vector virus pools from multiple DSDh producer cell clones containing one of each of the parental proviruses were used in several separate experiments to infect D17 target cells. One hundred twenty-seven G418^r and 56 Hyg^r single, well-isolated target cell colonies were cloned and expanded. Of the 183 clones characterized, 119 were calculated, as described in Materials and Methods, to have arisen from virions carrying heterodimers (virus titers and calculations available upon request). The proviral DNA was characterized by PCR amplification, restriction enzyme digestion, and agarose gel electrophoresis as previously described (11, 12). Maps representative of the proviruses detected in the G418^r and Hyg^r clones are shown in Fig. 4. The patterns of the restriction sites in the internal regions indicates that 13 of these proviruses were recombinant and 106 were not. This approach allowed us to determine in a singly selected population the pattern of the minus- and plus-strand DNA primer transfers and some genetic rearrangements, including recombination, large deletions, and large deletions with insertions. We could not detect base pair substitutions, frameshifts, and small deletions.

FIG. 4.

Minus- and plus-strand DNA primer transfers are primarily intramolecular. Analysis of the proviral DNA contained in the singly selected cell clones indicated that the minus-strand DNA primer transfers were almost always intramolecular in nonrecombinant proviruses (Fig. 4A and B). Of the 170 nonrecombinant proviruses characterized, the long terminal

repeat (LTR) markers of 169 were consistent with intramolecular minus-strand DNA primer transfer and 1 contained LTR markers consistent with intermolecular minus-strand DNA primer transfer (Fig. 4B, second from the top). However, as described above, some of these proviruses must have arisen from virions containing homodimers. We used the Hardy-

FIG. 4. Maps representative of the classes of proviruses for each of the selection conditions. (A) Hygromycin B single selection; (B) G418 single selection with no crossovers; (C) G418 single selection with crossovers. Parental vectors are shown at the top of each section. WH13 is shown in white, and WH204 is in gray. Maps illustrating the recombination patterns of representative proviruses are shown below the maps of the parental vectors. Colors in the maps correspond to the color of the parental vector that contributed the region. The numbers to the right indicate the fraction of proviruses in each class. Internal template switches are observed as color changes in the region between the LTRs, as is observed in the central regions of maps in panel C. LTRs that are all one color (U3 the same color as U5) result from intramolecular minus-strand DNA primer transfers, as in the top map in panel C. Intermolecular minus-strand DNA primer transfers are represented by LTRs that have two colors (U3 different from U5), as in the second map from the top in panel C. If plus-strand DNA primer transfers were intermolecular, then we would observe different 5' and 3' LTRs in a single provirus (for example, one gray and the other white); however, no proviruses with this type of LTR pattern were observed in these experiments. Thus, all of the plus-strand DNA primer transfers were intramolecular. Note that for these experiments, we generally used producer cell clones that had higher Hyg^r titers so that the G418^r colonies were more likely to arise from a heterodimer (see Materials and Methods). Thus, although we did examine proviruses in Hyg^r colonies, many of these were derived from virions containing Hyg homodimers, in which no crossovers could be detected.

Weinberg equation to remove from consideration the fraction of proviruses that arose from virions containing homodimers as described in Materials and Methods. Of the 106 nonrecombinant proviruses calculated to have arisen from virions containing heterodimers, 105 had undergone intramolecular minus-strand DNA primer transfers. The pattern of intramolecular minus-strand DNA primer transfers that we detected in these experiments is inconsistent with the data presented by Pangniban and Fiore (16). This difference will be further considered in the Discussion.

Consistent with our previous reports, the minus-strand DNA primer transfers in the recombinant proviruses were about equally divided between apparent intra- and intermolecular minus-strand DNA primer transfers (nine and four events, respectively; Fig. 4C [11, 12]). All of the plus-strand DNA primer transfers detected in these experiments were intramo-

lecular (Fig. 4), which is also consistent with earlier reports (10–12, 16, 25).

Large deletions and deletions with insertions are detected primarily at hot spots. In addition to studying the minus- and plus-strand DNA primer transfers by the single-selection protocol, we used this procedure to determine the rate of formation of large deletions or large deletions with insertions. Deletions and deletions with insertions were detected as DNA bands of unexpected sizes on agarose gels after PCR amplification of the proviral DNA. In contrast to other types of strand switching, such as minus-strand DNA primer transfer, we are able to detect deletions and deletions with insertions arising from virions carrying heterodimers or homodimers. Of the 183 proviruses characterized, 22 had detectable deletions or deletions with insertions (Fig. 4). (In addition, we were unable to PCR amplify the extreme 5' end of the 5' LTR of one provirus.

Since this is at most a 20-bp deletion and could also be a small substitution, we did not include it as a deletion or deletion with insertion for the analysis.) The lesions were localized by further PCR amplification and restriction enzyme mapping (data not shown). On the basis of the position and type of alteration, the mutant proviruses were divided into four groups (Fig. 4 and data not shown). The deletions and deletions with insertions clustered in several areas containing remnants of multiple linker cloning sites used in construction of the vectors, indicating that these regions of multiple inverted repeats may act as hot spots for deletions and deletions with insertions. Almost half of the lesions (9 of 22) were shown by DNA sequence analysis (Fig. 5A) to be approximately 300-bp deletions that removed the region from the primer binding site (PBS) to a region 5' of *neo* that contains the remnant of a polylinker used in construction of the vectors (4). In eight of these proviruses, the deletion junctions were at short regions of identity (Fig. 5), as is often observed in nonhomologous recombination (7, 26, 33, 34). We detected one additional deletion in this region. Similar to the other PBS deletions shown in Fig. 5A, the 3' endpoint of this deletion was also in the region of the polylinker remnant. However, the 5' endpoint of this deletion is 1 base 5' to the PBS and thus falls into the class defined by Pulsinelli and Temin (20) as pre-PBS (Fig. 5B). We isolated five other proviruses containing lesions in the 5' leader region (Fig. 4B), but rather than deletions they were insertions of approximately 50 to 100 bp, also near the polylinker remnant. Four other small insertions (Fig. 4B) were detected in the polylinker remnant that is between *neo* and *hyg*. Three proviruses that contain large insertions (200 to 600 bp) in the 3' LTR were isolated (Fig. 4A and B). Only one provirus contained both a crossover and a detectable deletion (Fig. 4C, bottom). We have not previously detected deletions or deletions with insertions in recombinant proviruses (9–12). However, this deletion removed the splice donor site so the provirus could not express *hyg* and thus would not have been detected in the double-selection protocol that was used previously. Thus, 19 of 22 deletions or deletions with insertions involved either the PBS, a polylinker remnant, or both and thus are at hot spots. Since these hot spots are specific sites within the genome, we calculated the rates of such events per genome rather than per kilobase. The rate of deletions or deletions with insertions that include the PBS is about 5% per genome (10 events per 183 proviruses characterized). The rate of deletions or deletions with insertions that include either of the polylinker remnants is also about 5% per genome (nine events per 183 proviruses characterized). Deletions and deletions with insertions not at obvious hot spots were calculated per kilobase since they could occur anywhere in the genome, and the rate for these events is about 0.5%/kb per replication cycle (three events per 183 proviruses characterized per 3.5-kb genome size).

Homologous recombination. Thirteen recombinant proviruses containing a total of 21 internal crossovers were detected in 119 (after correction for the number of heterodimers) proviruses characterized (Fig. 4C). Thus, the crossover rate was about 6%/kb per replication cycle (21 crossovers per 119 proviruses per 3-kb target), consistent with results of our previous studies (9–12). Comparable results were obtained with the vector combination WH13, which carries the encapsidation-dimerization signal at the normal 5' location, and a WH204 derivative (12) that carries the encapsidation-dimerization signal at an alternate 3' location after single selection (data not shown).

DISCUSSION

In this study, we used previously characterized SNV-based retroviral vectors and helper cells to study the strand transfers that occur during the reverse-transcription phase of a single cycle of retroviral replication. Minus-strand DNA primer transfers in the nonrecombinant proviruses detected by our single-selection protocol were primarily intramolecular. The pattern of intramolecular minus-strand DNA primer transfers that we detected in these experiments is inconsistent with the data presented by Panganiban and Fiore (16). They reported that minus-strand DNA primer transfers are 100% intermolecular. In preliminary experiments using a *neo*-containing vector (APHneo') similar to that used by Panganiban and Fiore (16), we were able to confirm their results. Briefly, we introduced APHneo' and JD214Hygro, the *hyg*-containing vector used by Panganiban and Fiore (16), into DSDh helper cells by transfection. The cells were plated separately under G418 selection and hygromycin B selection in the presence of anti-SNV antisera. Approximately 2 weeks following transfection more than 200 double^r colonies were trypsinized, pooled, and replated. Virus was harvested from the pool of transfected helper cells and was used to infect fresh D17 target cells which were subsequently selected for G418^r or Hyg^r. Single well-isolated target cell clones were isolated from each of the selections, and the proviruses were characterized. Most of the proviruses from the G418^r colonies that we characterized contained the U5 marker derived from JD214Hygro (APHneo' carries the *neo* marker), indicating that most of the minus-strand DNA primer transfers were intermolecular, which is consistent with the results reported by Panganiban and Fiore (16). The experimental design of the study by Panganiban and Fiore (16) intentionally skewed the G418^r and Hyg^r titers, resulting in an about 1,000-fold excess of the *hyg*-containing vector JD214Hygro. They did not characterize the proviruses in the Hyg^r colonies because most of those colonies would have been derived from virions containing homodimers and thus would have been uninformative, since the virus came from cells containing only a single vector. Because of differences in experimental design, the differences in G418^r and Hyg^r titers were much smaller in our study, so we also characterized the proviruses in Hyg^r colonies. The proviruses in the Hyg^r colonies (JD214Hygro carries the *hyg* marker) contained the wild-type U5 marker from JD214Hygro, indicating that intramolecular minus-strand DNA primer transfers occurred during formation of the proviruses in the Hyg^r cell clones. The intramolecular minus-strand DNA primer transfers detected in these experiments using JD214Hygro are consistent with the intramolecular minus-strand DNA primer transfers detected in nonrecombinant proviruses isolated under single-selection conditions with WH13 and WH204 as described in Results.

With the exception of the one provirus (of a total of 106 nonrecombinant proviruses calculated to have been derived from heterodimeric virions containing WH13 and WH204 genomic RNA) mentioned in Results, we detected intermolecular minus-strand DNA primer transfers in nonrecombinant proviruses only when we examined G418^r colonies arising from APHneo'. As described in the paragraph above, the other member of the pair of vectors used by Panganiban and Fiore (16) was JD214Hygro. The pattern of markers in proviruses in Hyg^r colonies that we detected using JD214Hygro was consistent with intramolecular minus-strand DNA primer transfers. JD214Hygro contains wild-type SNV LTRs.

APH neo' contains an 8-bp substitution in the inverted repeat adjacent to *attR* of U5 and consistently replicated at

FIG. 5. PBS and pre-PBS deletions. (A) Sequences of the region surrounding the deletion junctions of the nine proviruses containing PBS deletions; (B) sequence of the region surrounding the deletion junction of the provirus containing the pre-PBS deletion. The sequence of the PBS is shown at the top and underlined below. Boxes attached by lines indicate short direct repeats at the site of the deletion. Italicized bases (top sequence) are not in the original sequence of the provirus. Numbers above diamonds or dots indicate the position in the original provirus, and numbers following Δ indicate the size of the deletion. *, a pair of independently isolated proviruses that contain identical deletion junctions; **, a second pair of independently isolated proviruses that contain identical deletion junctions.

reduced titers compared with vectors containing wild-type LTRs. Ramsey and Panganiban (21a) have independently determined that APHneo' replicates at one-fifth the titer of the wild-type vector. Several explanations could account for

the decreased fitness and altered pattern of minus-strand DNA primer transfers that we detect with APHneo'. The 8-base substitution mutation present in APHneo' might decrease efficiency of formation of the minus-strand DNA primer or

decrease the efficiency of the minus-strand DNA primer transfer. JD214, when present in heterodimers with APHneo', could provide a wild-type minus-strand DNA primer *in trans*, thus bypassing either of the bottlenecks suggested. Further study is necessary to determine the cause or causes of the decreased fitness and altered pattern of minus-strand DNA primer transfers of APHneo' relative to those of vectors containing wild-type U5 regions. We conclude that the minus-strand DNA primer transfers are primarily intramolecular for wild-type SNV and are thus also likely to be intramolecular for other type C retroviruses.

The pattern of restriction sites in all of the proviruses detected in singly selected cell clones indicated that plus-strand DNA primer transfers were also primarily intramolecular (Fig. 4), consistent with previous reports (10–12, 16, 25). Thus, we conclude that one RNA template is sufficient for viral DNA synthesis in the absence of recombination and that the minus- and plus-strand DNA primer transfers are ordered rather than random events, because if the process were random, approximately equal proportions of inter- and intramolecular minus-strand DNA primer transfers would be expected.

An alternate explanation for our DNA primer transfer data is formally possible; however, as described below, we believe that this is unlikely. We have assumed here and in previous studies (9–12, 16, 25, 34) that encapsidation is random—that is, homodimers are not preferentially encapsidated. If RNA encapsidation is not random, then our assumptions about the composition of the dimer RNA in virions that led to formation of the single⁺ colonies are invalid (see Materials and Methods). We believe that this is an unlikely possibility because the retroviral vectors that we are using are nearly identical, with the RNA form differing only at eight restriction sites in a 3.5-kb genome. In a previous study (16), the results of which we independently confirmed in this study, the retroviral vectors were nonhomologous except for the LTR regions. Analysis of the proviruses arising from the nonhomologous pair is consistent with random encapsidation of RNAs that are almost completely different. Because these nonhomologous RNAs are encapsidated randomly, we believe that it is highly unlikely that the nearly identical genomes which we studied would be nonrandomly encapsidated.

In agreement with results of our previous studies, the minus-strand DNA primer transfers in the recombinant proviruses were about equally divided between apparent intra- and intermolecular minus-strand DNA primer transfers (Fig. 4C) (10–12). The consistent observation that recombinant proviruses contain approximately equal proportions of apparent intra- and intermolecular minus-strand DNA primer transfers suggests that this ordered process has been disrupted in some of those cases that result in recombination. Strand breaks in the template RNA have been identified as lesions that could be repaired by recombination (2), so it is possible that a strand break in one or both of the dimer RNAs disrupts a higher-order structure necessary for the ordered strand transfer. If reverse transcription were initiated on a template containing a strand break, then no true intramolecular minus-strand DNA primer transfer is possible. Thus, we describe the minus-strand DNA primer transfers in recombinant proviruses as apparently intra- or intermolecular. Alternatively or in addition, if a small fraction of the RT molecules, or reverse transcription complexes, is prone to intermolecular template switching, this could also result in formation of recombinant proviruses that had also undergone intermolecular minus-strand DNA primer transfers.

Deletions and deletions with insertions were detected primarily at a limited number of sites (PBS and polylinker

remnants), some of which had previously been identified as hot spots for such events. Of 22 deletions or deletions with insertions in this study, 9 had one end in the PBS (Fig. 5). Pulsinelli and Temin (20) isolated a series of proviruses that contain large deletions and deletions with insertions in the 5' half of the genome, and the PBS was an endpoint in many of them. One group of proviruses characterized by Pulsinelli and Temin (20), called PBS deletions, extended from the PBS to various positions in the center of the provirus. These lesions were slightly different from ours because they often contained insertions and usually lacked a short region of identity at the point of deletion. Another group of proviruses containing large deletions, which Pulsinelli and Temin called misalignment mutations (20), did not include the PBS but were at short regions of identity similar to ours. We also found one deletion that falls in the pre-PBS class (20). Nine proviruses contained deletions or deletions with insertions at two positions containing the remnants of multiple linker cloning sites that were used in construction of the retroviral vectors. These polylinker remnants contain multiple inverted repeats (palindromes). Pathak and Temin showed that multiple repeats of a single palindrome are a hot spot for small deletions with a deletion rate of 0.3% per genome per replication cycle (18). The hot spots that we detected and the hot spot constructed and described by Pathak and Temin (18) might function differently. Theirs was designed to form a large hairpin structure (it contains multiple identical palindromes), but ours contains multiple different palindromes. However, both sites do seem to be hot spots for deletions and deletions with insertions.

The total rate of deletions and deletions with insertions at hot spots was about 10% per genome per replication cycle. Deletions and deletions with insertions that were not at obvious hot spots occurred at a rate of about 0.5%/kb per replication cycle. Our determination of the rate of deletions and deletions with insertions may over- or underestimate the true rate. Factors that could contribute to our determination being an underestimate include the following: (i) a net increase or decrease in the size of the genome of greater than about 50 bp is required for deletions or deletions with insertions to be detected, and (ii) deletions or deletions with insertions that removed essential regions of the provirus could not be detected. The rate which we determined for deletions and deletions with insertions may be an overestimate of extra (internal) events because of the contribution of the hot spots for deletions and deletions with insertions discussed above. The rate which we determined for deletions and deletions with insertions is lower than the frequency of formation of defective proviruses by wild-type, replication-competent murine leukemia virus or SNV, which has been reported to be about 30 to 40% per genome. However, other lesions besides deletions and deletions with insertions contribute to those frequencies (15, 22). As described above, a rate of small deletions of 0.3% per genome per replication cycle has been determined at a position designed as a hot spot for deletion (18). Pathak and Temin also showed that small deletions and deletions with insertions occurred at a rate of 0.4%/kb per replication cycle in a region lacking both the PBS and significant secondary structure (17). Others have used selection to detect large deletions and deletions with insertions in the 5' half of a retroviral vector (the region that includes the PBS) and reported a rate of about 1% per genome per replication cycle (4, 20).

The differences between the spectrum of deletions and deletions with insertions detected here and those reported previously (18, 20) may reflect that (i) different vectors which would be expected to change which specific mutations are detected were used in the various studies, (ii) we detect

deletions or deletions with insertions by detecting differences in the size of PCR-amplified DNA, while the previous studies were specifically designed to genetically select (20) or screen (18) for such lesions, and (iii) the deletions and deletions with insertions that can be formed during retroviral DNA synthesis are probably a very broad continuum of which each study detects overlapping subsets.

Crossovers occurred at a rate of about 6%/kb per replication cycle. This rate is comparable to the rate which we reported previously under double-selection conditions (9–12), indicating that retroviral homologous recombination is not highly error prone. Recently Peliska and Benkovic used an *in vitro* assay to study template switching by the RT of human immunodeficiency virus type 1 (19). In that study, the vast majority of template switches resulted in insertion of an additional nucleotide at the site of the template switch. Thus, those results indicate that template switching may be highly error prone (19). However, if frameshifts were introduced at all sites of template switching in our experiments, then these mutations would be likely to disrupt one of the selectable markers required for detection of a functional provirus in the double-selection protocol. Alternatively, the single-selection protocol requires only one functional drug resistance marker, so even if frameshifts are introduced during template switching, a functional provirus could still be formed. Thus, a larger fraction of recombinant proviruses should be detected with the single-selection protocol than with the double-selection protocol. Recombinant proviruses were detected at comparable rates with either protocol, indicating that most retroviral homologous recombination is not highly error prone *in vivo*. Further studies will be necessary to resolve the question of whether some classes of retroviral homologous recombination are error prone. The results of Peliska and Benkovic indicate that homologous recombination at sites of damage such as strand breaks may be one such class (19).

The combined rates of large deletions and deletions with insertions at hot spots (10% per genome per replication cycle) and other sites (0.5%/kb per replication cycle) and the rate of crossovers (6%/kb per replication cycle) indicate that, on average, full-size (10-kb) type C retroviruses undergo an additional or aberrant strand transfer about once per cycle of infection.

ACKNOWLEDGMENTS

We thank Xiao-juan Bi, Ann Krueger, Jennifer Schoening, and Brad Seufzer for technical assistance. We thank Kathleen Boris-Lawrie, Louis M. Mansky, Antonito Panganiban, Gary Pulsinelli, Catherine Ramsey, Bill Sugden, and Shiaolan Yang for helpful discussions. We thank Ilse Riegel for comments on the manuscript.

This work was supported by Public Health Service grants CA22443 and CA07175. J.S.J. was supported by postdoctoral training grant CA09075 from the NCI. H.M.T. is an American Cancer Society research professor.

REFERENCES

1. **Bebenek, K., and T. A. Kunkel.** 1993. The fidelity of retroviral reverse transcriptases, p. 85–102. *In* A. M. Skalka and S. P. Goff (ed.), *Reverse transcriptase*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
2. **Coffin, J. M.** 1979. Structure, replication and recombination of retrovirus genomes: some unifying hypotheses. *J. Gen. Virol.* **42**:1–26.
3. **Coffin, J. M.** 1990. Retroviridae and their replication, p. 1437–1500. *In* B. N. Fields, D. M. Knipe, R. M. Chanock, M. S. Hirsch, J. L. Melnick, T. P. Monath, and B. Roizman (ed.), *Virology*, 2nd ed. Raven Press, New York.
4. **Dougherty, J. P., and H. M. Temin.** 1986. High mutation rate of a spleen necrosis virus-based retrovirus vector. *Mol. Cell. Biol.* **6**:4387–4395.
5. **Dougherty, J. P., R. Wisniewski, S. Yang, B. W. Rhode, and H. M. Temin.** 1989. New retrovirus helper cells with almost no nucleotide sequence homology to retrovirus vectors. *J. Virol.* **63**:3209–3212.
6. **Gilboa, E., S. W. Mitra, S. Goff, and D. Baltimore.** 1979. A detailed model of reverse transcription and tests of crucial aspects. *Cell* **18**:93–100.
7. **Hajjar, A. M., and M. Linial.** 1993. A model system for nonhomologous recombination between retroviral and cellular RNA. *J. Virol.* **67**:3845–3853.
8. **Higuchi, R.** 1989. Simple and rapid preparation of samples for PCR, p. 31–38. *In* H. A. Erlich (ed.), *PCR technology*. Stockton Press, New York.
9. **Hu, W.-S., and H. M. Temin.** 1990. Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination. *Proc. Natl. Acad. Sci. USA* **87**:1556–1560.
10. **Hu, W.-S., and H. M. Temin.** 1990. Retroviral recombination and reverse transcription. *Science* **250**:1227–1233.
11. **Hu, W.-S., and H. M. Temin.** 1992. Effect of gamma radiation on retroviral recombination. *J. Virol.* **66**:4457–4463.
12. **Jones, J. S., R. W. Allan, and H. M. Temin.** 1993. Alteration of location of dimer linkage sequence in retroviral RNA: little effect on replication or homologous recombination. *J. Virol.* **67**:3151–3158.
13. **Kawai, S., and M. Nishizawa.** 1984. New procedure for DNA transfection with polycation and dimethyl sulfoxide. *Mol. Cell. Biol.* **4**:1172–1174.
14. **Kent, R. B., J. R. Emanuel, Y. B. Neriah, R. Levenson, and D. E. Housman.** 1987. Ouabain resistance conferred by expression of the cDNA for a murine Na⁺.K⁺-ATPase α subunit. *Science* **237**:901–903.
15. **O'Rear, J. J., and H. M. Temin.** 1982. Spontaneous changes in nucleotide sequence in proviruses of spleen necrosis virus, an avian retrovirus. *Proc. Natl. Acad. Sci. USA* **79**:1230–1234.
16. **Panganiban, A. T., and D. Fiore.** 1988. Ordered interstrand and intrastrand DNA transfer during reverse transcription. *Science* **241**:1064–1069.
17. **Pathak, V. K., and H. M. Temin.** 1990. Broad spectrum of *in vivo* forward mutations, hypermutations, and mutational hotspots in a retroviral shuttle vector after a single replication cycle: deletions and deletions with insertions. *Proc. Natl. Acad. Sci. USA* **87**:6024–6028.
18. **Pathak, V. K., and H. M. Temin.** 1992. 5-Azacytidine and RNA secondary structure increase the retrovirus mutation rate. *J. Virol.* **66**:3093–3100.
19. **Peliska, J. A., and S. J. Benkovic.** 1992. Mechanism of DNA strand transfer reactions catalyzed by HIV-1 reverse transcriptase. *Science* **258**:1112–1118.
20. **Pulsinelli, G. A., and H. M. Temin.** 1991. Characterization of large deletions occurring during a single round of retrovirus vector replication: novel deletion mechanism involving errors in strand transfer. *J. Virol.* **65**:4786–4797.
21. **Ramig, R. F.** 1990. Principles of animal virus genetics, p. 95–122. *In* B. N. Fields, D. M. Knipe, R. M. Chanock, M. S. Hirsch, J. L. Melnick, T. P. Monath, and B. Roizman (ed.), *Virology*, 2nd ed. Raven Press, New York.
- 21a. **Ramsey, C., and A. Panganiban.** Personal communication.
22. **Shields, A., O. N. Witte, E. Rothenberg, and D. Baltimore.** 1978. High frequency of aberrant expression of Moloney murine leukemia virus in clonal infections. *Cell* **14**:601–609.
23. **Skalka, A. M., and S. P. Goff (ed.).** 1993. *Reverse transcriptase*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. **Strickberger, M. W.** 1976. *Genetics*. Macmillan Publishing Co. Inc., New York.
25. **Stuhlmann, H., and P. Berg.** 1992. Homologous recombination of copackaged retrovirus RNAs during reverse transcription. *J. Virol.* **66**:2378–2388.
26. **Swain, A., and J. M. Coffin.** 1992. Mechanism of transduction by retroviruses. *Science* **255**:841–845.
27. **Temin, H. M.** 1981. Structure, variation and synthesis of retrovirus long terminal repeat. *Cell* **27**:1–3.

28. **Temin, H. M.** 1991. Sex and recombination in retroviruses. *Trends Genet.* **7**:71–74.
29. **Temin, H. M.** 1993. Retrovirus variation and reverse transcription: abnormal strand transfers result in retrovirus genetic variation. *Proc. Natl. Acad. Sci. USA* **90**:6900–6903.
30. **Varmus, H., and P. Brown.** 1989. Retroviruses, p. 54–108. *In* D. E. Berg and M. M. Howe (ed.), *Mobile DNA*. American Society for Microbiology, Washington, D.C.
31. **Watanabe, S., and H. M. Temin.** 1983. Construction of a helper cell line for avian reticuloendotheliosis virus cloning vectors. *Mol. Cell. Biol.* **3**:2241–2249.
32. **Weiss, R., N. Teich, H. Varmus, and J. Coffin.** 1985. *RNA tumor viruses*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
33. **Zhang, J., and H. M. Temin.** 1993. 3' junctions of oncogene-virus sequences and the mechanisms for formation of highly oncogenic retroviruses. *J. Virol.* **67**:1747–1751.
34. **Zhang, J., and H. M. Temin.** 1993. Rate and mechanism of nonhomologous recombination during a single cycle of retroviral replication. *Science* **259**:234–238.