Ordered Interstrand and Intrastrand DNA Transfer During Reverse Transcription

ANTONITO T. PANGANIBAN AND DICCON FIORE

Retroviruses contain two copies of the plus stranded viral RNA genome. As a means of determining whether both of these RNA's are used in the reverse transcription reaction, cells were infected with heterozygous virus particles that varied in nucleotide sequence at two separate locations at the RNA termini. The DNA proviruses formed from a single cycle of reverse transcription were then examined. Of the 12 proviruses that were characterized, all exhibited long terminal repeats (LTR's) that would be expected to arise only if both RNA templates were used for the generation of minus strand DNA. In contrast, only a single minus strand DNA appeared to be used as template for the plus strand DNA in the generation of fully double-stranded viral DNA. These results indicate that the first strand transfer step in reverse transcription is an intermolecular event while that of the second transfer is intramolecular. Thus, retroviruses contain two functionally active RNA's, and both may be required for the generation of a single linear DNA molecule. Formation of heterozygotes during retrovirus infection would be expected to result in the efficient generation of LTR recombinants.

REVERSE TRANSCRIPTION, THE COPYING OF SINGLE-stranded retroviral RNA into double-stranded DNA, is catalyzed by the multifunctional, virally encoded enzyme, reverse transcriptase (*1, 2*). Proposed molecular models that describe this DNA synthesis reaction are characterized by several steps and uniformly invoke two distinct DNA strand transfer reactions (*3–5*). In the first transfer, a short nascent segment of minus strand DNA, initiated by a specific transfer RNA homologous to the primer binding site (pbs) (*6*), is translocated from a 5′ end of the plus strand viral RNA to the 3′ end of the same or a second RNA molecule. It is thought that this transfer is due, in part, to a direct repeat (r) present at both termini of the RNA. The second transfer step occurs later in the process of DNA synthesis and results in the transfer of a short plus strand DNA from the 5′ end of the minus strand DNA to the 3′ end of the same or another minus strand. In this case, plus strand initiation at the polypurine tract (ppt) (*7*) results in a short plus strand DNA that includes a primer binding site sequence complementary to the sequence present at the 3′ end of the minus strand DNA template. Since infectious particles each contain two copies of the RNA genome (*8*) exhibiting physical association ("dimer linkage") (*9, 10*), a quandary that arises as we consider these models is whether the strand transfer reactions are intra- or intermolecular events. A single RNA molecule contains all of the genetic information required for virus replication and thus may be sufficient for the generation of a complete viral DNA copy during reverse transcription by strand transfer from the 5′ to the 3′ end of the same template. Alternatively, one or both transfer steps might proceed from one template to a second template. Yet a third possibility is that both inter- and intrastrand transfer occur at equivalent frequency.

The double-stranded DNA that results from the reverse transcription process subsequently becomes inserted into the host cell genome to form the integrated "provirus" (*11–13*). The initial unintegrated linear DNA products of reverse transcription, as well as the provirus, both contain two long terminally repeated sequences (LTR's) derived from the termini of the virion RNA (*14, 15*). Each LTR can be subdivided into three regions on the basis of the origin of the sequence in the viral RNA. The R sequence is derived from the repeated sequence (r) at the termini of the viral RNA, U3 originates from a sequence (u3) at the 3′ end of the RNA adjacent to the 3′ r sequence, and U5 comes from a single copy RNA sequence (u5) juxtaposed to the 5′ r.

In order to examine the in vivo role of the two virion RNA molecules in the reverse transcription reaction, we generated heterozygous replication-defective virus stocks that contain two distinct RNA species differentially labeled with restriction enzyme sequence polymorphisms near the termini. These viruses were then used to

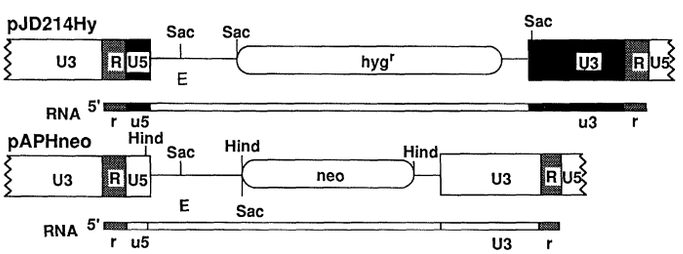

Fig. 1. Strategy for production of homozygous and heterozygous viral particles. The plasmids pJD214Hy (*17*) and pAPHneo each contain retroviral vector sequences that are transcriptionally expressed in transfected cells. Because transcription initiation and termination take place from the R sequence of the LTR, the u5 and u3 components in the viral RNA are both derived from the inside ends of the leftward and rightward LTR, respectively. Plasmid APHneo was constructed by inserting an 8-bp oligonucleotide containing a Hind III site into the U5 region; an LTR-containing plasmid (pSW210) (*20*) was partially digested with Rsa I, and Hind III "linkers" were inserted. This LTR was substituted for the upstream LTR of pJD215 (*17*) from an Eco RI site adjacent to the U3 region and a Kpn site in the E region to generate an intermediate plasmid pAPHneo′. The LTR of p60B-121, which lacks a U3 Sac I site (*20*), was then substituted for the downstream LTR of pAPHeno′ by joining, at one end, an Avr II site 38 bp upstream of the U3 terminus of p60B-121 with a similarly situated Cla I site in pAPHneo′. Ligation at the other end of the substituted region was achieved with an Ava I site at the U3-R junction.

The authors are at the McArdle Laboratory for Cancer Research, University of Wisconsin, 450 North Randall Avenue, Madison, WI 53706.

infect cells, and the resulting integrated proviruses were analyzed to determine the origin of the minus and plus strand DNA after a single round of DNA synthesis. The deduced origin of the DNA termini indicates that there is an operationally precise order to the strand transfer process. The first transfer is an intermolecular event involving both RNA molecules, and the second is an intramolecular process that requires only a single minus strand DNA template. We suggest that retroviruses contain two functionally active RNA's and that both are required for the generation of a single linear molecule.

Production of heterozygous virus particles. A derivative of the "helper" cell line (C3) (*16*) for an avian retrovirus, spleen necrosis virus, was used for the production of helper virus–free virus stocks harboring two different viral RNA species. An important feature of the helper line is that viral proteins are produced as a result of the presence of endogenously expressed viral genes. However, no replication-competent, wild-type virus is produced since those endogenous sequences lack a sequence (designated E) that must be present in cis for encapsidation of the RNA into virus particles. The introduction into these cells, and expression of, plasmid DNA encoding RNA with a functional E site results in the production of virus particles that arise from assembly of the functional viral proteins with the RNA derived from the introduced plasmid.

Plasmid JD214Hy encodes a retroviral vector that expresses a gene (*hyg*) conferring resistance to the compound hygromycin B and contains a downstream U3 region that has a Sac I restriction enzyme site (*17*). Since the RNA that results from transcription of pJD214Hy extends from the beginning of R in the upstream LTR and extends to the end of the R region in the downstream LTR, the Sac I sequence would be present in the RNA (Fig. 1). This plasmid was introduced into C3 cells by transfection (*18*), and cells with integrated pJD214 were selected by growing cells in the presence of hygromycin B at 400 μg/ml. A clonal population of cells (designated C3JD214Hy) was isolated and expanded for further use. This cell line produces approximately 8×10^5 helper-free virus particles per milliliter each containing JD214Hy RNA when quantified by infecting D17 cells and then determining the number of Hyg^r cell clones that arise from infection. The relatively high titer of this virus stock may be attributed to continual growth of the cell line in the presence of hygromycin resulting in the selection of only those cells that harbor and express JD214Hy. Moreover, these cells continually release virus particles carrying JD214Hy RNA which allows the reinfection and amplification of JD214Hy sequences in the cell population.

A second retroviral vector sequence (pAPHneo) was subsequently used to transfect 10^6 C3JD214Hy cells. This viral genome has an upstream U5 region containing a Hind III site, but lacks the Sac I site present in U3 of JD214Hy (Fig. 1). In addition, the vector contains the neo^r gene, which confers resistance to the drug G418 when expressed in eukaryotic cells. Eight days after the transfection, virus was harvested and used to infect D17 cells, G418-resistant cells were selected, and individual clonal cell populations were isolated.

After transfection of C3JD214Hy cells, the observed titer of APHneo-containing virus was only 3×10^2 particles per milliliter compared with a JD214Hy titer of 8×10^5. This difference in titer between neo^r- and hyg^r-containing viruses should be a direct consequence of differential RNA production. The consequence of such a difference in RNA production would be that packaging of one APHneo RNA into a virion should be accompanied by the packaging of a JD214Hy RNA. This would result in the generation of a pool of helper-free particles composed of abundant homozygous particles containing two copies of JD214Hy RNA and a lesser amount of heterozygous (APHneo-JD214Hy) particles. However, as the cells infected with this virus pool were selected for growth in the presence of G418, only those cells harboring a provirus derived from infection with a heterozygote would be expected to survive. Such clonal cell lines derived from infection with heterozygotes are designated with the letter B. A total of 12 lines were examined. In addition, homozygous APHneo-APHneo particles were generated by transfection of C3 cells. A control group of eight $G418^r$ cell lines that arose from infection with these particles was also examined. These are indicated with the letter A. Only a single infectious cycle is possible for these viruses since both the JD214Hy and APHneo genomes lack *gag*, *pol*, and *env* sequences.

To quantify directly the relative amounts of neo^r- and hyg^r-containing RNA in the heterozygous virion stock, we transfected C3JD214Hy cells with pAPHneo, purified the virus particles, isolated the RNA, and performed hybridization analysis with neo^r- and hyg^r-specific probes (*19*). Successive tenfold dilutions of the viral RNA preparation were bound to membrane filters and annealed with the neo^r DNA probe (specific activity = 3.6×10^7 cpm/μg) or the hyg^r gene probe (specific activity = 2.9×10^7 cpm/μg). As was expected, hyg^r-containing RNA was easily detected even at an RNA dilution of 10^{-3}. In contrast, there was no detectable neo^r-specific viral RNA, an indication that the amount of this RNA was much lower than that of the hyg^r gene RNA. This analysis confirmed that the amount of hyg^r RNA is several orders of magnitude more abundant than the neo^r RNA in the virus preparation in concordance with the results of the measured virus titer.

Since only a single round of infection and reverse transcription is possible for the heterozygous viral stock, instability of these restriction site polymorphisms would be highly improbable unless there is a profound effect on virus replication or gene expression. The particular mutations used in the analysis were chosen because they are located in segments of the LTR that are nonessential for viral DNA integration of transcription (*20*). However, some unforseen detrimental effect of these mutations might result in selection for the wild-type LTR and would complicate analysis of the G418-resistant colonies that arose through infection by heterozygous virus particles. To directly determine whether virus replication might be affected by the presence of the Hind III site in U5 or the absence of the Sac I site in U3, we examined replication competent retroviral variants that contain these same alterations.

Virus strain 60B is a wild-type stock that replicates to high titer [10^7 infectious units (IU) per milliliter]. Strains 60B-58 and 60B-121 are viral stocks that contain mutations identical to those of pAPHneo: 60B-58 contains the U5 Hind III site and 60B-121 lacks the U3 Sac I site. Plasmid DNA's bearing these viral derivatives were processed (*20*) and introduced into chicken embryo fibroblasts by transfection; virus was harvested from these cells 4 days later and used to infect fresh cells. The virus titer from these cells was determined 5 days later. In addition, we checked the stability of the restriction site markers through these many rounds of virus infection by isolating unintegrated viral DNA from cells infected in a manner identical to that used to determine the virus titer. These DNA's were then subjected to DNA hybridization analysis. All of the viral variants showed efficient replication and reached titers of 10^7. Moreover, in each case the restriction site polymorphism was maintained (*21*). Had the polymorphisms interfered with virus replication in some manner, a reduction in titer would have been expected. Similarly, if selection would have resulted in the loss or reestablishment of restriction sites to allow efficient replication, this would be reflected in a different restriction digestion pattern in the hybridization analysis. Thus, it appears that there is no effect of these mutations on viral replication and selection would not be expected to be a factor in analysis of the heterozygotes.

A different possible complication of the experiment could arise if there is a high rate of homologous recombination between the endogenous JD214Hy sequences present in C3JD214Hy cells and

the APHneo DNA that is introduced by transfection. However, such recombination only occurs at an efficiency of 0.1 to 1 percent (*22*). A further source of complexity could arise if more than one virus infected the cells that harbor a provirus that results from infection by heterozygous particles. As shown below, 6 of the 12 cell lines contain only a single provirus, whereas the other 6 contain two viral sequences, one of which appears to be derived from a heterozygous particle and a second that probably resulted from infection by a homozygous, JD214Hy-containing particle.

Strategy for determining the possible derivation of LTR's. Intra- or interstrand transfer can be detected by characterization of the proviral LTR's after infection by a heterozygous retrovirus. The LTR structure of *neo*^{r}-containing proviruses that arise from infection by an APHneo-JD214Hy heterozygote depends on whether the first and second DNA transfer steps of reverse transcription are inter- or intramolecular events (or both) (Fig. 2). (i) In the simplest case, where both transfer steps occur on the same template strand, both proviral LTR's contain U3 and U5 components identical to that of the parental APHneo RNA. Thus, the U5 region of each LTR would contain a Hind III site, but both U3 regions would be devoid of Sac I sites. (ii) If intrastrand transfer takes place in the first transfer reaction and the second transfer occurs between two different templates, then a provirus with two heterologous LTR's would be generated, the upstream LTR would be derived from JD214Hy RNA, and the downstream LTR would be derived from APHneo RNA. (iii) If the first transfer step is an intermolecular reaction and the second involves only a single template strand, then the U5 component of both LTR's would be expected to be derived from the JD214Hy RNA and contain no Hind III site, whereas the U3 component of both LTR's would originate from the U3 region of APHneo and lack Sac I sites. (iv) The final possibility is that both first and second transfers occur between the two template strands. The consequence of such double intermolecular reactions would be that U5 in the upstream LTR and U3 in the downstream LTR would be derived from APHneo and that U3 of the upstream LTR and U5 of the downstream LTR would arise from JD214Hy. This would yield an upstream LTR marked in U3 and U5 with a Sac I or Hind III, respectively, whereas the downstream LTR would not contain either restriction enzyme site.

In the two cases where the second transfer step would involve the movement of the short plus strand DNA segment to another minus strand DNA template, it would be necessary to infer that parallel minus strand DNA synthesis takes place for both RNA's of the virion. Alternatively, if the second translocation step is an intramolecular reaction, there would be no need to invoke the parallel minus strand DNA synthesis for both RNA's since a single minus strand DNA would be sufficient to generate a complete plus strand DNA copy.

Minus strand transfer is an intermolecular reaction and plus strand transfer is intramolecular. Examination of the provirus DNA in cells infected by heterozygous particles indicates that in minus strand DNA transfer both RNA templates are used. In contrast, plus strand DNA transfer requires a single minus strand DNA template. The twelve G418^r colonies that resulted from infection with heterozygous viruses containing APHneo and JD214Hy RNA were subjected to DNA hybridization analysis to determine whether the proviral LTR's contained Hind III and Sac I sites. In addition, the DNA from eight resistant colonies that came about through infection with homozygous particles containing APHneo RNA was analyzed in parallel. The presence or absence of Hind III and Sac I sites in one or both LTR's can be detected by digesting the proviral DNA with either of these enzymes and hybridizing with DNA probes specific to defined regions of the proviral DNA. If the upstream LTR of the *neo*^{r}-containing provirus

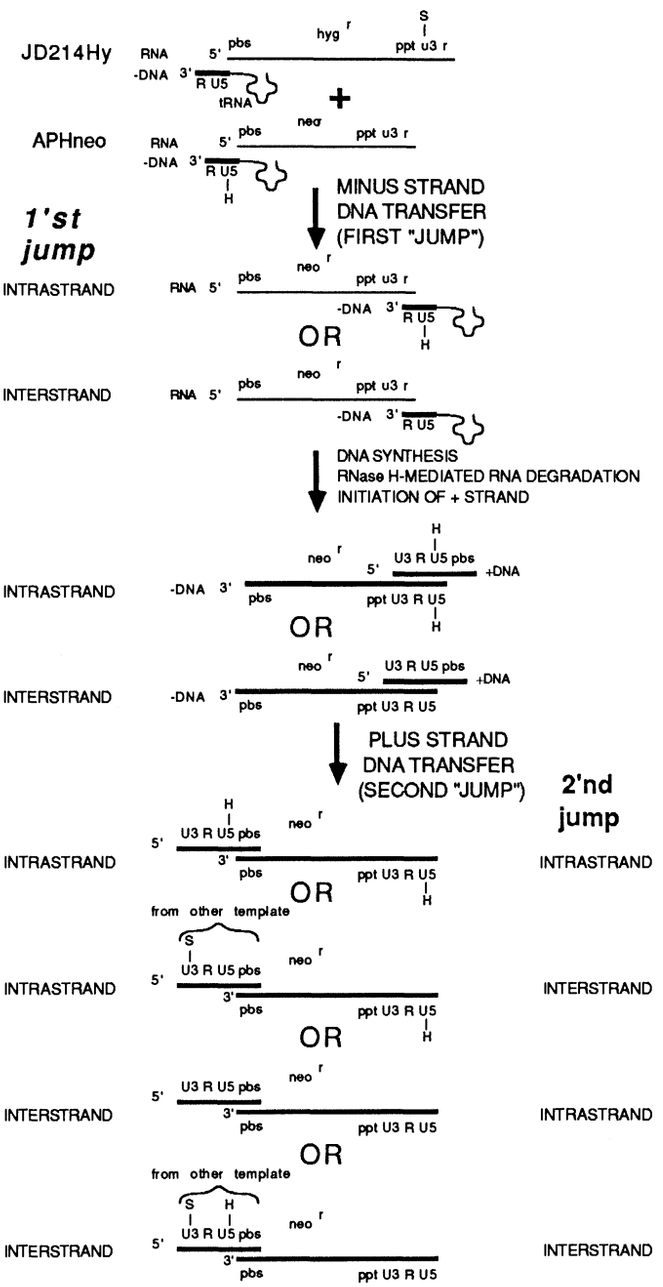

Fig. 2. Predicted derivation of the provirus U3 and U5 regions from heterozygous retroviruses if the two DNA transfer steps are intramolecular or intermolecular. At top are diagrammed JD214Hy and APHneo RNA's, upon completion of the short minus strand and before the first strand transfer reaction. The rest of the figure shows reverse transcription of only the *neo*^{r} template, but simultaneous, parallel reverse transcription of the *hyg*^{r} template may also occur. If transfer of the minus strand occurs on the same RNA template the transferred DNA contains a U5 Hind III site (H). Alternatively, if the minus strand DNA is transferred from the JD214Hy RNA, then U5 lacks a U5 Hind III site. The second transfer step then ensues. If this reaction is intramolecular, the transferred strand is exactly complementary to the downstream LTR sequence and that results in, after completion of DNA synthesis, two identical LTR's at both ends. If the first transfer is also intramolecular, then both LTR's will contain a U5 Hind III site and no U3 Sac I site (*5*), whereas neither site will be present if the first transfer is intermolecular. An intermolecular second transfer requires reverse transcription of both APHneo and JD214Hy templates. On the JD214Hy template (not diagrammed), U3 would have been derived from only that template and would contain a Sac I site. When this plus strand is transferred from this template to the *neo*^{r} template, upon completion of DNA synthesis the upstream LTR will contain a Sac I site.

contains the U5 region derived from APHneo RNA, then digestion with Hind III should result in the formation of a fragment 350 bp in length because of the presence of a Hind III site in U5 of the upstream LTR and a second Hind III site that flanks the *neo*^{r} gene (Fig. 3). Similarly, the presence of a Hind III site in the downstream LTR would yield a fragment of 520 bp resulting from the presence of yet another Hind III site located immediately downstream from the *neo*^{r} gene. The use of a DNA probe derived from the LTR and the adjacent E region should make it possible to detect both of these fragments and, in addition, to detect a third fragment comprised of the upstream LTR and adjacent cell DNA. If a Hind III site exists in only one of the LTR's, then only the 350- or the 520-bp fragment will be detected. If both LTR's are devoid of Hind III sites, then two fragments of variable size would be detected, depending on the location of Hind III sites in the adjacent cell DNA.

In the A series of clones that arise from infection by homozygous particles, each contains a U5 Hind III site as evidenced by the presence of both small Hind III fragments and a single clone-specific fragment (Fig. 3); the two small fragments were not resolved under the gel and electrophoresis conditions of Fig. 3A but were resolved and detected under those conditions described for Fig. 3B. In contrast, the proviruses that are derived from infection by heterozygous particles uniformly lack Hind III sites in either LTR. Neither the 350- or 520-bp fragments are present, but either two or three fragments of larger size are found (Fig. 3A). The presence of two such variable fragments is indicative of a single provirus lacking LTR Hind III sites, and this is the pattern found for clones B4, B5, B8, B9, B10, and B11. Each of the remaining six B clones (1, 2, 3, 6, 7, and 12) contain a third fragment of clone-specific size. This is consistent with the possibility that there are two proviruses in these clones, one containing the *neo*^{r} gene and lacking Hind III sites in the LTR (similar to those of the first group) and a second containing the *hyg*^{r} gene and also lacking Hind III sites.

To determine whether there might be a second provirus containing the *hyr*^{r} gene in some of the cell clones, we analyzed Hind III–digested DNA using a probe specific to the *hyg*^{r} gene (*23*). Since JD214Hy lacks internal Hind III sequences (Fig. 1) only a single Hind III fragment would be expected if a particular cell clone harbored a *hyg*^{r}-containing provirus. For each of the clones yielding three rather than two fragments in the previously described hybridization analysis, a single clone-specific Hind III fragment was detected with the *hyg*^{r} probe. Moreover, the size of each of the fragments detected by the *hyg*^{r} probe is identical to one of the three fragments detected by the probe used in Fig. 3, A and B (these fragments are indicated by the designation "h" in Fig. 3A). This further indicates that the other two fragments of clones B1, B2, B3, B6, B7, and B12 must arise from the *neo*^{r}-containing provirus (Fig. 3A). As is discussed above, the presence of two variable fragments is indicative of a single provirus lacking Hind III sites in the LTR's. Thus all clones of the B series infected with heterozygous particles

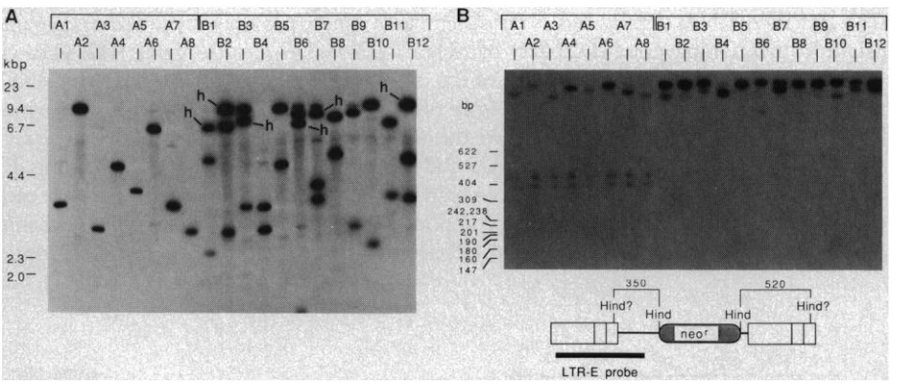

Fig. 3. Hybridization analysis of Hind III–digested DNA from *neo*^{r}-containing proviral DNA from cells infected with homozygous (**A**) and heterozygous (**B**) retroviral particles. (A) and (B) show the autoradiograms after hybridization with a ^{32}P-labeled DNA probe homologous to the LTR and E region. The DNA analyzed in (A) was subjected to electrophoresis on an 0.8 percent agarose gel and transferred to a nitrocellulose filter by blotting, whereas the DNA of (B) was processed on a 5 percent polyacrylamide gel and electrophoretically transferred to a gene screen filter. The LTR E-specific probe is also shown. This probe was derived from a 751-bp DNA fragment that resulted from digestion of cloned viral DNA containing a U3 Sac I cleavage site with Sac I. The internal 350- and 520-bp fragments derived from infection with homozygous particles expected to anneal with the probe are also shown. The question mark next to the Hind III sites of the LTR's indicates the Hind III sites analyzed; these sites are important in determining whether the first strand transfer is intra- or intermolecular (Fig. 2). In (A), those fragments that arise from the presence of a second provirus containing the *hyg*^{r} gene (*23*) are designated "h."

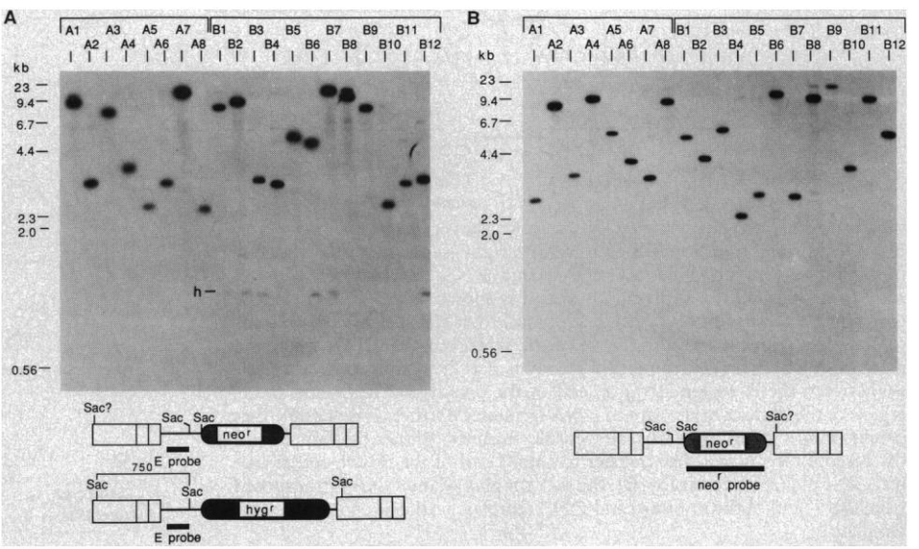

Fig. 4. Hybridization analysis of Sac I–digested *neo*^{r}-containing proviruses from cells infected with virus particles. (**A**) Digested DNA's were hybridized with a 185-bp fragment from the E region, which was isolated and labeled after digestion with Kpn I and Sac I. For those G418-resistant clones that arose from infection with a heterozygote and also contain a second provirus that probably arose from infection with a homozygote of JD214Hy, a 750-bp fragment would be expected. This fragment is also designated with an "h" in the autoradiogram. (**B**) DNA's have been detected by hybridization with a ^{32}P-labeled *neo*^{r}-specific probe. As in Fig. 3, the presence of a question mark denotes a Sac I site that is crucial for determining whether a strand transfer step in intra- or intermolecular. In this case, the presence or absence of a Sac I site is diagnostic for the second, rather than the first, transfer step.

contain only one *neo*^{r}-containing provirus and each of these proviruses lacks Hind III sites in either LTR. Thus, the U5 component of each *neo*^{r}-containing provirus is derived from JD214Hy, and the first transfer step (the translocation of minus strand DNA) occurs from one RNA template to the second.

The U3 Sac I site distribution for the *neo*^{r} proviruses that came about through infection with homo- or heterozygotes was determined by analyses similar to those for the Hind III sites (Fig. 4). For proviruses that arise through infection of homozygous APHneo clones (A series) no Sac I site would be expected in either the upstream (detected with an E-specific probe) or the downstream (detected with a *neo*^{r}-specific probe). Hybridization with both probes should reveal fragments of clone-specific variable size extending leftward from the Sac I site in E or rightward from the Sac I site just upstream of the *neo*^{r} gene. This was the observed result (Fig. 4). For the B series, the Sac I hybridization pattern was similar, an indication that there are no Sac I sites in either LTR. If a Sac I site had been present in the rightward LTR, a fragment 1.3 kb in length would have been seen in each clone. This leads to the conclusion that both U3 segments of the provirus are derived from APHneo rather than JD214Hy RNA and that the second strand transfer is an intramolecular reaction.

Each of the six G418^r cell clones that contain both a *hyg*^{r} and *neo*^{r} provirus, as determined by analysis of Hind III–cleaved DNA (B1, B2, B3, B6, B7, and B12), also yield a hybridizing 750-bp Sac I fragment with the E probe. This indicates that the upstream U3 region of this *hyg*^{r} provirus is derived from JD214Hy RNA rather than APHneo RNA. Again, this is consistent with a *hyg*^{r} provirus origin from a homozygous JD214Hy particle.

There are several important features of the proviruses containing the *neo*^{r} gene. Those DNA's that result from reverse transcription of homozygous APHneo viral RNA all have the conserved U5 and U3 marker that would be expected if both U3 and U5 are derived from APHneo. In contrast, the heterozygous particles all contain U5 components from JD214Hy and U3 components from APHneo RNA. These proviruses that arise from infection by heterozygotes are especially interesting since it appears that in each case the first transfer step occurred between the two RNA molecules carried in the heterozygote, whereas the second transfer step took place between the 5′ and 3′ ends of the same minus strand DNA template (Fig. 5). Thus, retroviruses have a "pseudodiploid" genome, that is, segments of a single DNA provirus are derived from both virion RNA's.

Characterization of RNA's from virion particles indicates that the RNA is in the form of a linked dimer (*9, 10*). Although the biological significance of this dimer structure remains to be assessed experimentally, this physical association between the two virion RNA's is likely to be important for minus strand transfer. Conversely, there is no reason to invoke mandatory physical association between the RNA's for transfer of the plus strand since this process appears to involve only a single template.

One provirus originates from one virion. For each virus particle infecting a single cell, the potential exists for the generation of two proviruses since there are two RNA molecules per virion. A model for such a scheme would require parallel first transfer steps and synthesis of a complete minus strand corresponding to both RNA's. All eight "A series" clones that resulted from infection by homozygous APHneo RNA-containing particles contained only a single provirus. Six of the cell clones resulting from infection with a heterozygous particle harbored only a single provirus. The remaining six clones had two proviruses, one containing a *neo*^{r} and one a *hyg*^{r} gene, but we believe that in each case the second provirus came about through infection by a second particle. Each of the *hyg*^{r} proviruses contains LTR's consistent with an origin from a homozygous JD214Hy virus, that is, the upstream U3 sequence contains the Sac I site, and both U5 regions lack the Hind III site. Therefore, if the *hyg*^{r} provirus arose from the same virus as the *neo*^{r} provirus, it would be necessary to postulate that the reverse transcription of the *hyg*^{r}-containing RNA proceeds through two intrastrand transfer steps and that each of the RNA's undergo reverse transcription by different mechanisms. If two such parallel pathways existed for reverse transcription, there should be an equal chance of a given

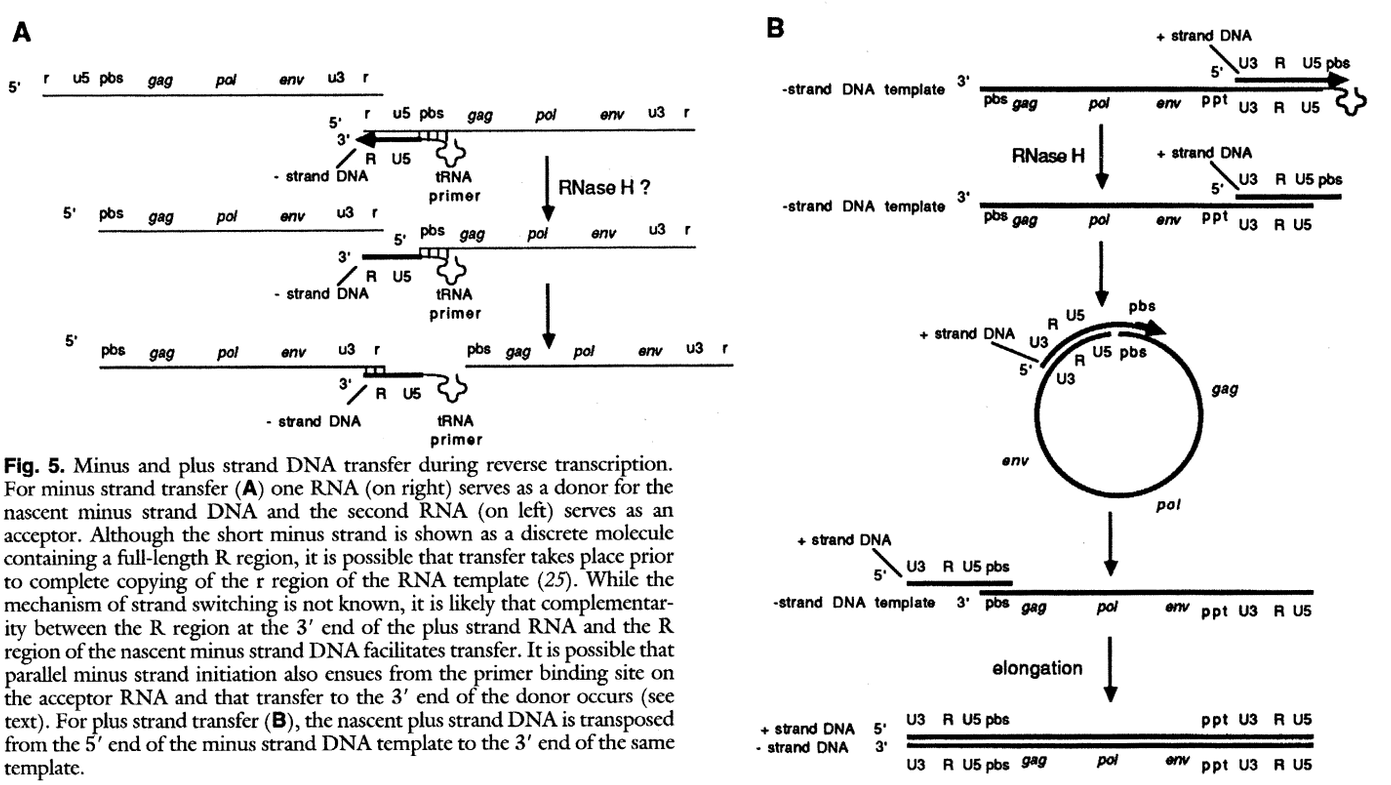

Fig. 5. Minus and plus strand DNA transfer during reverse transcription. For minus strand transfer (**A**) one RNA (on right) serves as a donor for the nascent minus strand DNA and the second RNA (on left) serves as an acceptor. Although the short minus strand is shown as a discrete molecule containing a full-length R region, it is possible that transfer takes place prior to complete copying of the r region of the RNA template (*25*). While the mechanism of strand switching is not known, it is likely that complementarity between the R region at the 3′ end of the plus strand RNA and the R region of the nascent minus strand DNA facilitates transfer. It is possible that parallel minus strand initiation also ensues from the primer binding site on the acceptor RNA and that transfer to the 3′ end of the donor occurs (see text). For plus strand transfer (**B**), the nascent plus strand DNA is transposed from the 5′ end of the minus strand DNA template to the 3′ end of the same template.

RNA proceeding along one pathway while the other proceeds along the second pathway. Since all 12 *neo*^{r} proviruses are apparently identical and appear to have arisen through successive inter- and intrastrand transfers and not through two intrastrand transfers, the possibility of alternative pathways of reverse transcription for each of the two RNA's of the virion is very small. On this basis, we conclude that a single virus gives rise to a single provirus and that the *hyg*^{r} proviruses arise from a second infection event. Since we examined a total of 26 proviruses, if two or more proviruses can arise from a single infection, the frequency of such an event must occur in less than 4 percent of infections.

The significance of this "one virus–one provirus" hypothesis for the overall retroviral replicative cycle remains ambiguous. From the deduced pattern of strand transfer, one of the RNA copies might only be required to supply a nascent minus DNA strand for the second RNA. Since the second transfer appears to be an intrastrand event, one RNA would then be dispensable for the remainder of the reverse transcription process. Alternatively, the efficiency of reverse transcription might be significantly higher than that of DNA integration. Two copies of unintegrated DNA may be generated from infection by a single particle, but if the integration process is not efficient then, at most, only one DNA copy might be ultimately inserted into the genome.

Two virion RNA's may be required for the generation of one DNA molecule. Analysis of the proviruses derived from single heterozygous particles indicates that the first transfer step is intermolecular. While this observation does not unequivocally prove that two RNA templates are absolutely required for this event under all circumstances, it does indicate that the usual mechanism of reverse transcription includes both RNA's of the virion. Electron microscopic examination of DNA's synthesized by reverse transcriptase in vitro revealed a prominent population of circularized molecules that may have been derived from intramolecular first strand transfer and subsequent DNA synthesis (*24*). This indicates that intrastrand transfer may take place in the in vitro reaction. However, from the results in vivo, we suggest that both RNA's are required for reverse transcription and, therefore, that both RNA's are required for function. For each linear double-stranded DNA resulting from reverse transcription, one RNA molecule serves as template for a short minus strand DNA providing an R and a U5 component, whereas the remaining complement of the viral genome is contributed by the second RNA molecule.

Although we used restriction site polymorphisms to follow the fate of only the u3 and u5 nucleic acid components, a prediction of the interstrand transfer is that the R segment of both LTR's would be at least partially derived from the donor RNA molecule that provides the U5 component. Studies on in vivo reverse transcription indicate that complete copying of the r region may not be required before template switching of the minus DNA strand (*25*). Thus, the relative contribution to the R region from the 5′ and 3′ ends of the RNA depends on the length of the r region copied prior to the first strand transfer. After transfer of the initiated minus strand to the 3′ end of the other RNA and subsequent DNA synthesis, the ribonuclease (RNase) H activity of reverse transcriptase would be expected to degrade the r region of the recipient RNA. The R region of the transferred DNA could then be copied. Since the second transfer step appears to be an intrastrand event, all or part of the R region as well as the entire U5 region will then have derived from the same donor RNA.

Interstrand transfer may result in increased retroviral variation. It is not yet clear whether two heterologous RNA's both containing the cis-acting encapsidation sequence are packaged into the same virus with the same efficiency as two identical RNA's. However, it is clear that heterozygotes do form. A consequence of intermolecular minus strand transfer is that the incorporation of two different RNA genomes with genotypic LTR differences would result in the efficient generation of an LTR variant in a single cycle of replication. Thus, in instances where more than one virus type infects a single cell, a high frequency of rearranged LTR derivatives may result from independent assortment during packaging of the two types of RNA's. Other circumstances that could result in heterozygote formation include the expression of endogenous proviruses during exogenous viral infection, the formation of a virus particle containing RNA from two distinct viral species, and the low-level incorporation of a cellular RNA along with a viral RNA. The packaging of these various heterologous RNA's has been demonstrated in several experiments (*26–28*).

Many of the cis-acting regulatory components of retroviruses reside in, and are immediately adjacent to, the LTR's. The mixing of LTR's during minus strand transfer could potentially result in the attainment of a different U5, R, and primer-binding sequence by a given retroviral genome. The U5 region contains sequences required for efficient integration of DNA (*20*). The R segment contains a sequence required for RNA termination and polyadenylation (*29*), and the primer binding site defines which specific cellular tRNA will act as primer for minus strand initiation. The general implication of interstrand transfer is that the process could, in some instances, result in the formation of new retroviral species that may have new biological consequences for the infected cell.

REFERENCES AND NOTES

1. H. M. Temin and S. Mizutani, *Nature* **226**, 1211 (1970).
2. D. Baltimore, *ibid.*, p. 1209.
3. E. Gilboa, S. W. Mitra, S. Goff, D. Baltimore, *Cell* **18**, 93 (1979).
4. H. M. Temin, *ibid.* **27**, 1 (1981).
5. H. E. Varmus, in *RNA Tumor Viruses*, R. Weiss *et al.*, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982), chap. 5, pp. 369–512.
6. J. M. Taylor, *Biochim. Biophys. Acta* **473**, 57 (1977).
7. S. W. Mitra *et al.*, *J. Biol. Chem.* **257**, 5983 (1982).
8. A. R. Bellamy, S. C. Gillies, J. D. Harvey, *J. Virol.* **14**, 1388 (1974).
9. H.-J. Kung *et al.*, *Cell* 7, 609 (1976).
10. W. Bender *et al.*, *J. Virol.* **25**, 888 (1978).
11. A. T. Khoury and H. J. Hanafusa, *Virology* **18**, 383 (1976).
12. H. E. Varmus, S. Heasley, J. Linn, K. Wheeler, *J. Virol.* **18**, 574 (1976).
13. N. Battula and H. M. Temin, *Cell* **13**, 387 (1978).
14. T. W. Hsu *et al.*, *J. Virol.* **28**, 810 (1978).
15. P. R. Shank *et al.*, *Cell* **15**, 1383 (1978).
16. S. Watanabe and H. M. Temin, *Mol. Cell. Biol.* **3**, 2241 (1983).
17. J. P. Dougherty and H. M. Temin, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 1197 (1987).
18. S. Kawai and H. Nishizawa, *Mol. Cell. Biol.* **4**, 1172 (1984).
19. RNA was prepared by phenol extraction of the pelleted virion preparation and ethanol precipitation of the nucleic acid. The RNA was subjected to dot blot hybridization analysis as described in the text [the undiluted RNA sample (10^0) contained about 7 μg of RNA] and the samples were analyzed by autoradiography. Autoradiography times ranged from 24 hours to 1 week. The 10^{-2} dilution of RNA contained RNA that annealed with the *hyg*^{r} probe visible after a 24-hour exposure time while the 10^{-3} dilution could be detected with the same probe after 72 hours. In contrast, no cross-hybridizing material was visible with the *neo*^{r} probe for the 10^0 dilution even after exposure for a week.
20. A. T. Panganiban and H. M. Temin, *Nature* **306**, 155 (1983).
21. DNA hybridization analysis showed that the restriction site polymorphisms of strains 60B-58 and 60B-121 were stable. Cells were infected as described, unintegrated viral DNA was prepared, and the DNA was digested with either Hind III or Sac I, and subjected to (Southern) hybridization. For 60B-58, all detectable viral DNA contained the Hind III site in both LTRs. Similarly, for 60B-121 all detectable viral DNA was devoid of Sac I sites in the LTRs. The relative amounts of unintegrated viral DNA were similar for 60B, 60B-58, and 60B-121.
22. P. K. Bandyopadhyay *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3476 (1984).
23. Cellular DNA was digested with Hind III and hybridized with a ^{32}P-labeled *hyg*^{r} gene probe.
24. R. P. Junghans, L. R. Boone, A. Skalka, *J. Virol.* **43**, 544 (1982).
25. L. I. Lobel and S. P. Goff, *ibid.* **53**, 447 (1985).
26. L. B. Crittenden, J. V. Motta, E. J. Smith, *Virology* 76, 90 (1977).
27. J. E. Embretson and H. M. Temin, *J. Virol.* **61**, 2675 (1987).
28. M. Linial, *Cell* **49**, 93 (1987).
29. J. R. Nevins, *Annu. Rev. Biochem.* **52**, 441 (1983).

30. We thank K. Talbot for technical assistance, J. Dougherty for providing pJD214Hy, members of our laboratory for helpful discussion, and Howard Temin and Rex Risser for evaluation of the manuscript. Supported by grants from the ACS and NIH, a Junior Faculty Research award (A.T.P.) from the ACS, and an award from the Shaw Scholars Fund.

8 December 1987; accepted 14 July 1988