Retroviral Recombination and Reverse Transcription

WEI-SHAU HU AND HOWARD M. TEMIN

Recombination occurs at a high rate in retroviral replication, and its observation requires a virion containing two different RNA molecules (heterodimeric particles). Analysis of retroviral recombinants formed after a single round of replication revealed that (i) the nonselected markers changed more frequently than expected from the rate of recombination of selected markers; (ii) the transfer of the initially synthesized minus strand strong stop DNA was either intramolecular or intermolecular; (iii) the transfer of the first synthesized plus strand strong stop DNA was always intramolecular; and (iv) there was a strong correlation between the type of transfer of the minus strand strong stop DNA and the number of template switches observed. These data suggest that retroviral recombination is ordered and occurs during the synthesis of both minus and plus strand DNA.

HOMOLOGOUS RECOMBINATION MAKES IT POSSIBLE TO (i) increase the variation in a population by combining different variants, and (ii) repair damaged genes (*1*). Different organisms have different, intricate mechanisms for homologous recombination (*2*). We have studied homologous recombination in retroviruses–a family of RNA viruses that replicate through a DNA intermediate (*3*). Upon infection of host cells, the retroviral RNA is reverse-transcribed into double-stranded DNA, which then integrates into the host cell chromosome where it is designated a provirus. The host RNA polymerase II transcribes the provirus to produce viral RNA, which is then packaged into retroviral particles (virions) (Fig. 1).

Two retroviral RNA molecules are packaged in one virion. One consequence of this packaging of two RNA molecules in one retroviral particle is a high rate of homologous recombination (*4, 5*).

Recombination between two genetically marked retroviruses is not observed after coinfection with these two viruses, but only after infection with viruses produced from cells coinfected with these viruses. Besides the two parental viruses, the coinfected cells also produce a new population of virions that contain two different genomic RNA's (heterodimeric virion particles) (*5, 6*). Thus, recombination occurs between the two co-packaged RNA's or the DNA products reverse-transcribed from these RNA's.

Proviral DNA is terminally redundant. The repeats on the ends are named the long terminal repeats (LTR's). The LTR's contains numerous regulatory elements, including promoter sequences, and they comprise U3, R, and U5 (unique 3′, repeat, unique 5′) sequences (Fig. 2A). After transcription, however, viral genomic RNA's do not contain the promoter sequences at their 5′ ends. The process of reverse transcription regenerates the promoter at the 5′ end of the provirus by combining the U3 region from the 3′ end of the RNA and the U5 region from the 5′ end of the RNA with the R region. A strand transfer mechanism is used by retroviruses to achieve this goal (*7*) (Fig. 2B).

Mechanisms of retroviral recombination. Two models have been proposed for the mechanism of retroviral recombination: forced copy-choice (*8*) and strand displacement-assimilation (*9*).

The forced copy-choice model (Fig. 3A) proposes that the genomic RNA in the virion is damaged. When reverse transcriptase encounters a break in the viral RNA, it switches to the other copy of genomic RNA to salvage the encoded genetic information. Thus, the forced copy-choice model predicts that recombination occurs during the synthesis of minus strand DNA. However, breaks in the RNA's may not be necessary to promote this switching of templates (*10*). Thus, this model can be generalized to include all recombination that occurs during the synthesis of minus strand DNA.

The strand displacement-assimilation model (Fig. 3B) proposes that both copies of viral RNA are intact and that two minus strand DNA's are made by one virion. Since plus strand DNA synthesis is initially discontinuous (*11*), a fragment of product DNA might be displaced by the continuous DNA synthesis initiated from the transferred plus strand of the strong stop DNA (the first plus strand DNA synthesized, which acts after transfer as a primer for the synthesis of the rest of the plus strand DNA). This displaced DNA fragment could then hybridize to the minus strand DNA synthesized with the other molecule of viral RNA as template. If the assimilated fragment remains as part of the plus strand DNA, a recombinant emerges after mismatch repair. Thus, the strand displacement-assimilation model predicts that recombination occurs

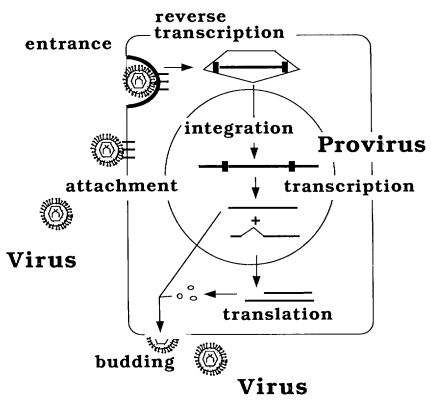

Fig. 1. Retrovirus life cycle. Retroviruses are RNA viruses that replicate through integrated DNA intermediates. A viral particle (virion) attaches to the cell surface and enters the cell. The viral RNA genome is reverse transcribed into DNA; this viral DNA then integrates into the host chromosome, forming a provirus. Host RNA polymerase II transcribes the provirus to produce viral RNA, which is either translated to produce viral proteins or packaged as the viral genome. Virus is released from the cell by budding.

The authors are at the McArdle Laboratory for Cancer Research, 1400 University Ave., University of Wisconsin–Madison, Madison, WI 53706.

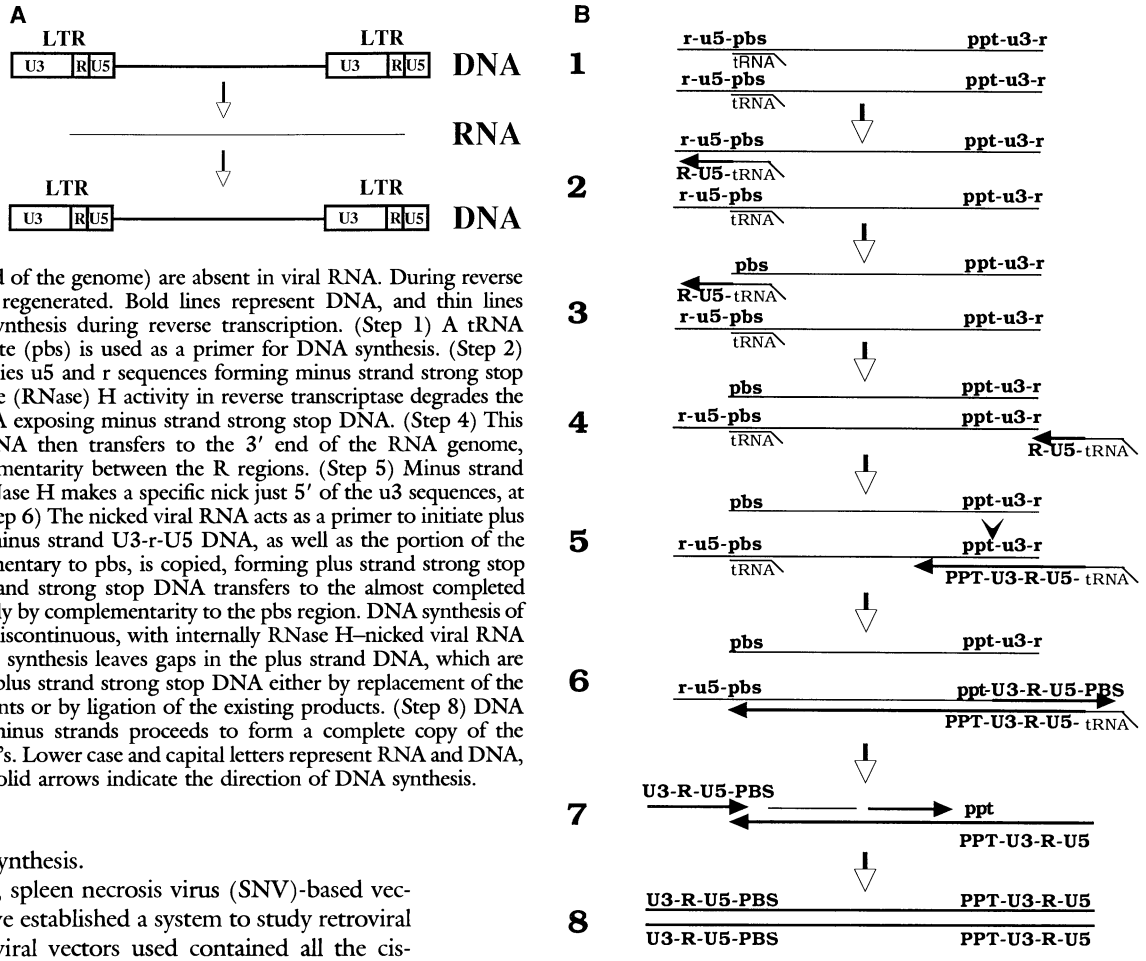

Fig. 2. Reverse transcription of retroviral genomes. (**A**) Structure of the terminal sequences of proviral DNA and viral RNA. Long terminal repeats (LTR's) are located at the two ends of the proviral DNA. After transcription by host RNA polymerase II, parts of the terminal sequences (including the promoter at the 5′ end of the genome) are absent in viral RNA. During reverse transcription, the LTR's are regenerated. Bold lines represent DNA, and thin lines represent RNA. (**B**) DNA synthesis during reverse transcription. (Step 1) A tRNA annealed to primer binding site (pbs) is used as a primer for DNA synthesis. (Step 2) Reverse transcription first copies u5 and r sequences forming minus strand strong stop DNA. (Step 3) A ribonuclease (RNase) H activity in reverse transcriptase degrades the u5 and r of the template RNA exposing minus strand strong stop DNA. (Step 4) This minus strand strong stop DNA then transfers to the 3′ end of the RNA genome, presumably using the complementarity between the R regions. (Step 5) Minus strand DNA synthesis continues. RNase H makes a specific nick just 5′ of the u3 sequences, at the polypurine tract (ppt). (Step 6) The nicked viral RNA acts as a primer to initiate plus strand DNA synthesis. The minus strand U3-r-U5 DNA, as well as the portion of the primer tRNA that is complementary to pbs, is copied, forming plus strand strong stop DNA. (Step 7) This plus strand strong stop DNA transfers to the almost completed minus strand DNA, presumably by complementarity to the pbs region. DNA synthesis of the plus strand is sometimes discontinuous, with internally RNase H–nicked viral RNA as primers; this discontinuous synthesis leaves gaps in the plus strand DNA, which are filled after the transfer of the plus strand strong stop DNA either by replacement of the previously synthesized fragments or by ligation of the existing products. (Step 8) DNA synthesis of both plus and minus strands proceeds to form a complete copy of the retroviral RNA with two LTR's. Lower case and capital letters represent RNA and DNA, respectively; bold lines with solid arrows indicate the direction of DNA synthesis.

during plus strand DNA synthesis.

Using C-type retrovirus, spleen necrosis virus (SNV)-based vectors and helper cell lines, we established a system to study retroviral recombination. The retroviral vectors used contained all the cis-acting elements required for retroviral replication, but they lacked the genes that encode viral proteins. Helper cell lines express all viral proteins essential for retroviral replication. Thus, retroviral vectors can form progeny virus in helper cell lines. The two vectors we used expressed the neomycin resistance gene (*neo*) (*12*) in direct transcripts from the LTR and expressed the hygromycin phosphotransferase B gene (*hygro*) (*13*) from spliced mRNA's (*14, 15*). In each of the vectors, either the *neo* or the *hygro* gene was disabled by a frameshift mutation such that a different selectable gene remained functional in the two vectors (*5*). Helper cells were infected with these two vectors, and doubly resistant cell clones containing both vectors were isolated. Viruses were harvested from these helper cell clones and were used to infect fresh D17 cells, a dog osteosarcoma cell line that is permissive for SNV infection (*16*). D17 cells do not express viral proteins. Thus, retroviral vectors integrate and form proviruses in D17 cells, but they cannot produce new viruses to infect other cells. From the proviruses in the helper cells to the proviruses in the target cells, the retrovirus vectors undergo one round of replication (*17*). Recombinant proviruses were not observed in the helper cell clones, but they were observed in the D17 target cells; thus, recombination occurred during this one round of retroviral replication (*18*).

Target D17 cells containing one of the parental viruses were only resistant to a single selective drug, whereas D17 cells containing recombinant viruses that have both genes functional were resistant to selection with both drugs (double selection). By subjecting target cells to different selections, we distinguished between the cells containing parental proviruses (singly resistant titer) and the cells containing recombinant proviruses (doubly resistant titer). The ratio of doubly resistant titer to singly resistant titer gave the recombination frequency. With such a scheme and vectors containing two mutations 1 kb apart, we observed recombination at a frequency of 2 percent [equivalent to a rate of 40 percent for a wild-type retrovirus, or about one in two to three wild-type viruses undergoes recombination (*19*)].

Frequency of change of nonselectable markers. WH13 and WH204 are two SNV-based splicing vectors that contain both *neo* and *hygro* genes (Fig. 4A). Eight differences in restriction enzyme sites distinguish these two otherwise identical vectors. Two of these eight sites are located in the coding regions of the drug resistance genes and inactivate one of the two genes in each vector. Thus, WH13 contains a mutant *neo* gene and a functional *hygro* gene, whereas WH204 contains a functional *neo* gene and a mutant *hygro* gene. These two vectors were introduced by infection into DSDh helper cells (*5*). The DSDh cells are derived from D17 cells and contain two constructs that express the essential trans-acting viral proteins for retroviral replication, one construct expressing *gag-pol* gene products and the other expressing the *env* gene product. Both constructs are driven by promoters not homologous to the vectors to eliminate the possibility of generating replication-competent viruses through homologous recombination. Viruses harvested from these cells were used to infect fresh target D17 cells. After one round of retroviral replication, as described above, doubly resistant D17 cell clones were isolated, and DNA from these cell clones was analyzed by DNA hybridization (*20*).

Each recombinant provirus was analyzed for one or the other of the paired marker restriction enzyme sites by detection of one fragment and the absence of the other (Fig. 4B). In no case were both of the markers present in the same clone.

Because double selection was used to distinguish cells containing recombinant proviruses from cells containing nonrecombinant pro-

viruses, the two Nco I markers located in the coding regions of the two drug resistance genes were selected. The other markers did not affect the expression of the *neo* or *hygro* genes and were therefore unselected markers.

A template switch is defined by the change of the growing point of DNA synthesis from one parent to the other co-packaged parent. Because only eight sets of markers were used, the number of template switches was determined by the maps of these markers. Of the 22 recombinant proviruses analyzed, 10 appeared to have one template switch and 12 appeared to have more than one template switch (Fig. 4C) (*21*).

Using results from a homologous set of vectors that differed only in two markers 1 kb apart, we calculated that one of seven vector viruses experienced at least one recombination event (*5*). If recombination events were independent, only one of seven recombinants would be expected to contain two recombination events. However, we found that 12 of the 22 recombinants had more than one template switch. Thus, either recombination events are not independent or a single recombination event often leads to more than one template switch.

The two models for retroviral recombination predict recombinants with different patterns of template switches. In copy choice recombination, each time the growing point of minus strand DNA changes template, the resulting recombinant reflects one template switch. However, in the strand displacement-assimilation model, one displacement event during plus strand DNA synthesis results in a recombinant that appears to have switched template twice (Fig. 3, A and B). Thus, either the strand switches during synthesis of the minus strand DNA were associated or some recombinants were generated by strand displacement-assimilation recombination.

If all recombinants resulted from plus strand recombination, then all recombinants would appear to have more than one template switch. Since 10 of 22 recombinant proviruses apparently contained only one template switch, plus strand recombination did not seem to be responsible for the generation of all the recombinants.

Number of template switches and transfer of strong stop DNA. Two sets of restriction enzyme site markers in the parental vectors were located in the LTR's. One set of markers was located in the U3 sequences (Sac I for WH13 and Not I for WH204), and another set of markers was located in the U5 sequences (Bam HI for WH13 and Cla I for WH204). By examining the distribution of these markers, we determined whether the transfer of the minus strand strong stop DNA was intermolecular or intramolecular (Fig. 5A). For example, the minus-strand DNA synthesis initiated from

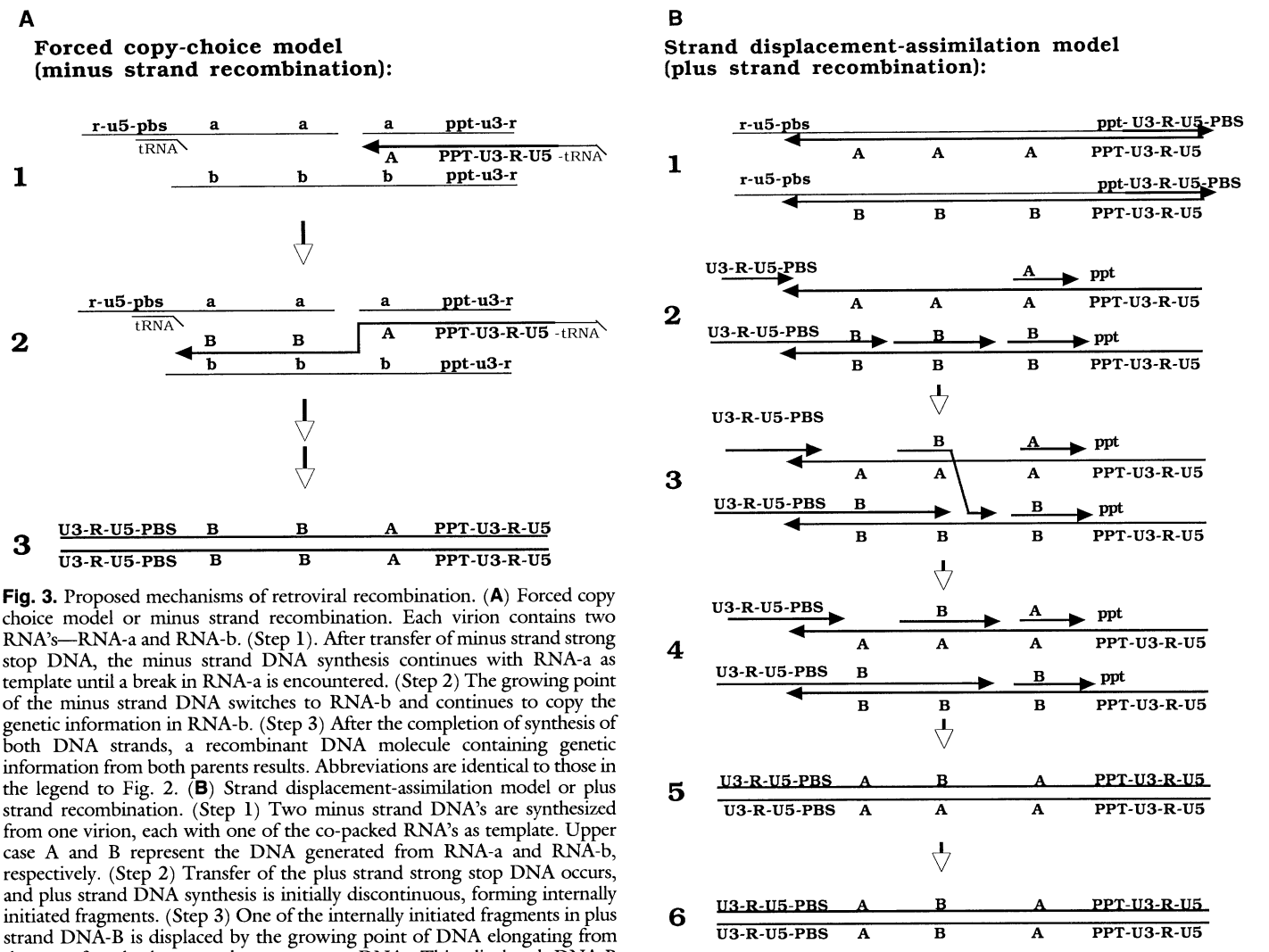

Fig. 3. Proposed mechanisms of retroviral recombination. (**A**) Forced copy choice model or minus strand recombination. Each virion contains two RNA's—RNA-a and RNA-b. (Step 1). After transfer of minus strand strong stop DNA, the minus strand DNA synthesis continues with RNA-a as template until a break in RNA-a is encountered. (Step 2) The growing point of the minus strand DNA switches to RNA-b and continues to copy the genetic information in RNA-b. (Step 3) After the completion of synthesis of both DNA strands, a recombinant DNA molecule containing genetic information from both parents results. Abbreviations are identical to those in the legend to Fig. 2. (**B**) Strand displacement-assimilation model or plus strand recombination. (Step 1) Two minus strand DNA's are synthesized from one virion, each with one of the co-packed RNA's as template. Upper case A and B represent the DNA generated from RNA-a and RNA-b, respectively. (Step 2) Transfer of the plus strand strong stop DNA occurs, and plus strand DNA synthesis is initially discontinuous, forming internally initiated fragments. (Step 3) One of the internally initiated fragments in plus strand DNA-B is displaced by the growing point of DNA elongating from the transferred plus strand strong stop DNA. This displaced DNA-B fragment begins to hybridize to the minus strand DNA-A. (Step 4) The displaced DNA-B fragment is assimilated into the DNA-A structure. (Step 5) A complete viral DNA molecule is formed. A region of this DNA molecule contains genetic information from the two parents, the minus strand from parent a and the plus strand from parent b. (Step 6) Mismatch repair, probably before cell replication, corrects the different sequences in this region. When the genetic information from parent b remains, a recombinant is formed.

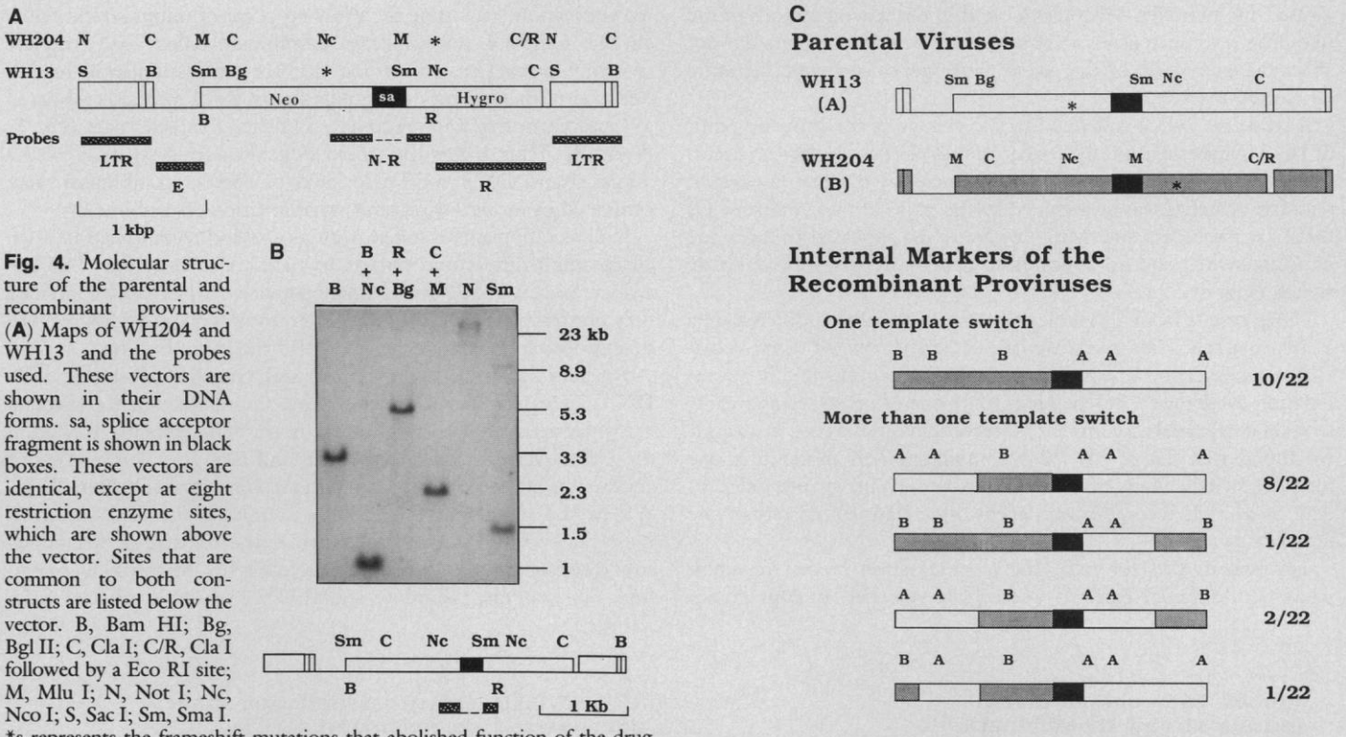

Fig. 4. Molecular structure of the parental and recombinant proviruses. (**A**) Maps of WH204 and WH13 and the probes used. These vectors are shown in their DNA forms. sa, splice acceptor fragment is shown in black boxes. These vectors are identical, except at eight restriction enzyme sites, which are shown above the vector. Sites that are common to both constructs are listed below the vector. B, Bam HI; Bg, Bgl II; C, Cla I; C/R, Cla I followed by a Eco RI site; M, Mlu I; N, Not I; Nc, Nco I; S, Sac I; Sm, Sma I. *s represents the frameshift mutations that abolished function of the drug resistance genes. The markers are small insertions of 4 to 8 bp. All probes used to map the recombinants are shown as dotted boxes. (**B**) Mapping the recombinant provirus in D17 cell clone 15D1. The + symbol indicates digestion with both indicated enzymes. N-R probe (dotted boxes) was used. A preliminary map generated from this Southern blot is shown below the autoradiograph. Lane B, a 3.3-kb band indicating the presence of two B sites shown in the map. B + Nc, a 1-kb band indicating the presence of both Nc sites, one from each parental vector. R + Bg, a 5.3-kb band indicating the absence of the Bg site from WH13. C + M, a 2.2-kb band indicating the presence of the C site in the *neo* fragment from WH204; N, a band >23 kb indicating the absence of at least one and possibly both of the N markers from WH204; Sm, a 1.5-kb band, indicating the presence of both Sm sites from WH13, and a 8.9 kb band with a weaker intensity, indicating a DNA fragment from the Sm site 5′ to the *hygro* gene to the flanking host sequences. (**C**) Internal markers of the recombinant proviruses. Open boxes, WH13 and WH13-derived sequences; shadowed boxes, WH204 and WH204-derived sequences; A and B, markers derived from WH13 and WH204, respectively. Recombinants with similar maps of internal markers are shown together. The number following each recombinant pattern represents the frequency of that pattern among the recombinant proviruses.

WH204 contained the Cla I marker in the U5 region. This minus strand strong stop DNA could either be transferred to the same RNA molecule(intramolecular transfer) or to the other co-packaged RNA molecule (intermolecular transfer). When the virion contained two different viral RNA's, the product of the two types of transfer were different (*22*). In such cases, intramolecular transfer resulted in an LTR identical to the parent WH204 (Not I in U3 and Cla I in U5), and intermolecular transfer resulted in an LTR different from either of the two parents (Sac I in U3 and Cla I in U5).

We observed three types of LTR's among the 22 recombinants analyzed; two of them were identical to the parents, and one was different from the two parents (Fig. 5B), an indication that the minus strand strong stop DNA transfer was either intermolecular or intramolecular. This result is in contrast to that previously described (*22, 23*). The fourth possible type of LTR (Not I in U3 from WH204 and Bam HI in U5 from WH13) was not observed (*24, 25*).

Suppose two minus strand DNA's were copied from the two RNA molecules in one virus, then the plus strand strong stop DNA's could either be transferrred to the same DNA molecule (intramolecular) or to the other DNA molecule (intermolecular). If the transfer is intramolecular, then the two LTR's in the resulting provirus are identical. If the transfer is intermolecular, then the two LTR's in the resulting provirus are different. Among all 22 recombinant proviruses, the two U3 regions within each provirus had the same markers. This result indicates that the transfer of the plus strand strong stop DNA was always intramolecular; a result in agreement with an earlier one (*22*).

One provirus with different markers in the two U5 regions was also found (Figs. 5B and 6). However, because recombinants were generated from heterodimeric virions and the two U3 markers of this provirus were both from WH13 RNA, the transfer of the plus strand strong stop DNA was most likely intramolecular as for all the other proviruses. Several mechanisms that generated this provirus are considered possible (*26*).

Correlation of the number of template switches with the type of transfer of the minus strand strong stop DNA. Examination of the restriction enzyme maps of the entire recombinant proviruses showed that among 22 recombinants, ten patterns of distribution of the markers were observed (Fig. 6). Two types of transfer of the minus strand strong stop DNA (minus strand transfer) were observed: intermolecular and intramolecular. Most (7 of 10) of the recombinants that had only one template switch between the LTR's also contained LTR's that resulted from intermolecular minus strand transfer. In contrast, most (10 of 11) of the recombinants that had more than one template switch between the LTR's also contained LTR's resulting from intramolecular minus strand transfer. Fisher's exact test (*27*) indicated that the probability of this distribution being random is less than 0.01. Thus, the type of minus strand DNA transfer correlated with the number of template switches.

This correlation supports the hypothesis that retroviral recombination occurs during both minus strand and plus strand DNA synthesis and shows that each mechanism produces a group of phenotypically different recombinants. Intermolecular transfer of

the minus strand strong stop DNA generally leads to recombination during the synthesis of the minus strand DNA, whereas intramolecular transfer of the minus strand strong stop DNA is more likely to lead to recombination during the synthesis of the plus strand DNA.

Recombination and evolutionary strategy. To elucidate the mechanisms of retroviral recombination, we used two almost homologous vectors that differed at eight markers, and we studied recombinants after one round of retroviral replication. Analysis of the recombinants showed that the unselected markers changed more frequently than expected from the rate of recombination of two selected markers determined with two similar vectors. More than half of the recombinants contained more than one template switch; this frequency is higher than the expected one in seven recombinants containing two recombination events.

Analysis of the distribution of markers in the U3 and U5 regions of the LTR's in the recombinant proviruses indicated that the transfer of minus strand strong stop DNA was both intermolecular and intramolecular, whereas the transfer of plus strand strong stop DNA was exclusively intramolecular.

Furthermore, most of the recombinants in which a single switch between the templates occurred had intermolecular transfer of the minus strand strong stop DNA. Conversely, most recombinants in which more than one switch between the templates occurred had intramolecular transfer of the minus strand strong stop DNA. This correlation suggests that two mechanisms are involved in retroviral recombination. We propose that retroviral recombination takes place during the synthesis of both the minus strand and the plus strand DNA. When the viral RNA genomes are damaged, intermolecular transfer of minus strand DNA occurs; recombination during the synthesis of minus strand DNA generates DNA molecules from virions containing damaged genomes (*8*). When the viral RNA genomes are not damaged, intramolecular transfer and plus strand recombination occurs. The advantage of having recombination during the synthesis of the plus strand DNA may be efficiency. Each of the plus strand recombination events results in a recombinant that contains two template switches; thus, the effect of one recombination event is maximized. Alternatively, plus strand recombination may be a consequence of the enzymology or conformation of the reverse transcription complex.

Recombinants with more than one template switch were observed frequently. This result indicates that the genetic information in two copackaged retroviral RNA's can be shuffled rapidly, and recombinants with mosaic patterns can form frequently in one round of replication. Thus, the diversity of a retroviral population increases quickly, especially with a hypermutated virus (*25*) that has multiple mutations in one viral genome. In addition, the correlation of type of strong stop DNA transfer with the number of template switches

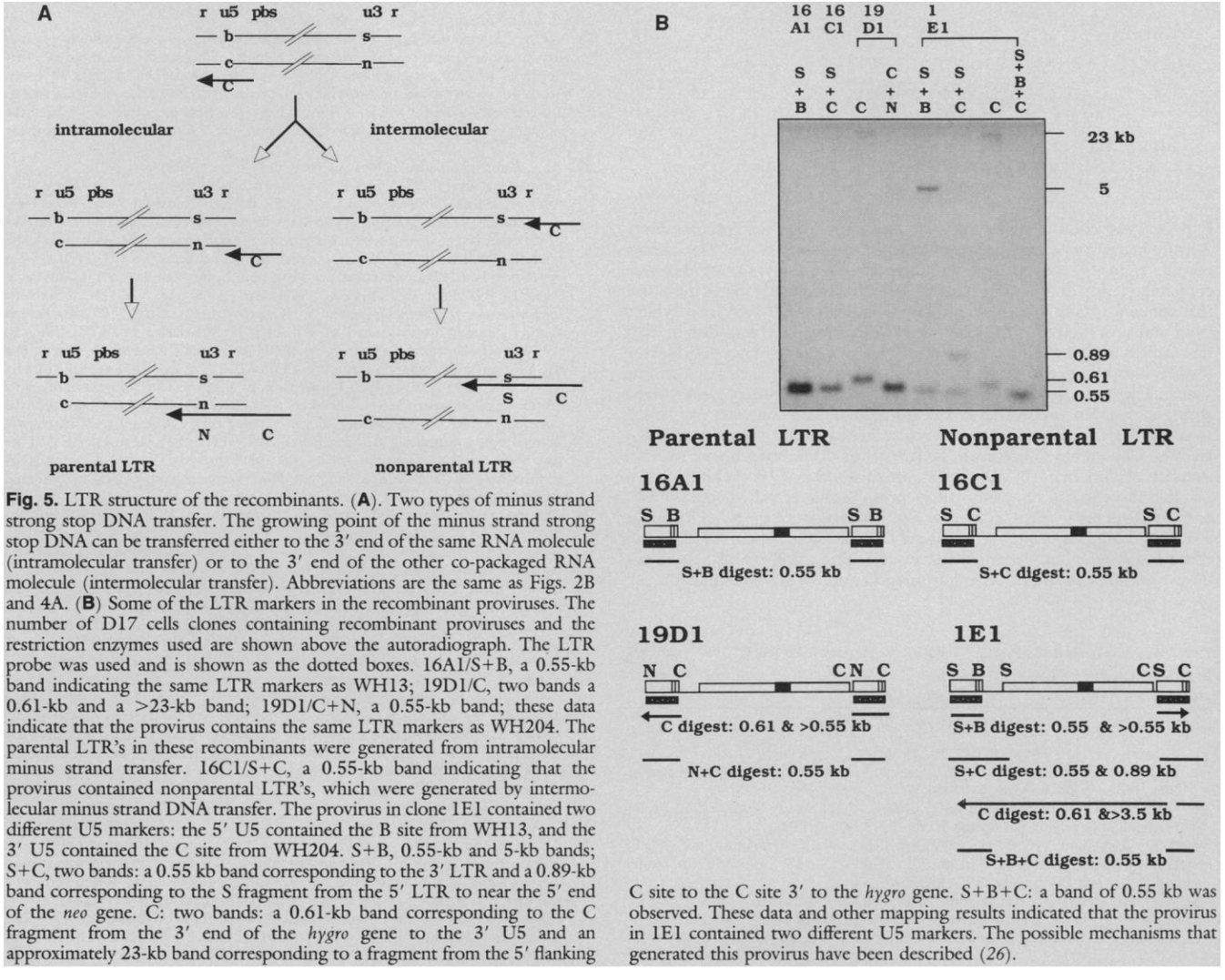

Fig. 5. LTR structure of the recombinants. (**A**). Two types of minus strand strong stop DNA transfer. The growing point of the minus strand strong stop DNA can be transferred either to the 3′ end of the same RNA molecule (intramolecular transfer) or to the 3′ end of the other co-packaged RNA molecule (intermolecular transfer). Abbreviations are the same as Figs. 2B and 4A. (**B**) Some of the LTR markers in the recombinant proviruses. The number of D17 cells clones containing recombinant proviruses and the restriction enzymes used are shown above the autoradiograph. The LTR probe was used and is shown as the dotted boxes. 16A1/S+B, a 0.55-kb band indicating the same LTR markers as WH13; 19D1/C, two bands a 0.61-kb and a >23-kb band; 19D1/C+N, a 0.55-kb band; these data indicate that the provirus contains the same LTR markers as WH204. The parental LTR's in these recombinants were generated from intramolecular minus strand transfer. 16C1/S+C, a 0.55-kb band indicating that the provirus contained nonparental LTR's, which were generated by intermolecular minus strand DNA transfer. The provirus in clone 1E1 contained two different U5 markers: the 5′ U5 contained the B site from WH13, and the 3′ U5 contained the C site from WH204. S+B, 0.55-kb and 5-kb bands; S+C, two bands: a 0.55 kb band corresponding to the 3′ LTR and a 0.89-kb band corresponding to the S fragment from the 5′ LTR to near the 5′ end of the *neo* gene. C: two bands: a 0.61-kb band corresponding to the C fragment from the 3′ end of the *hygro* gene to the 3′ U5 and an approximately 23-kb band corresponding to a fragment from the 5′ flanking C site to the C site 3′ to the *hygro* gene. S+B+C: a band of 0.55 kb was observed. These data and other mapping results indicated that the provirus in 1E1 contained two different U5 markers. The possible mechanisms that generated this provirus have been described (*26*).

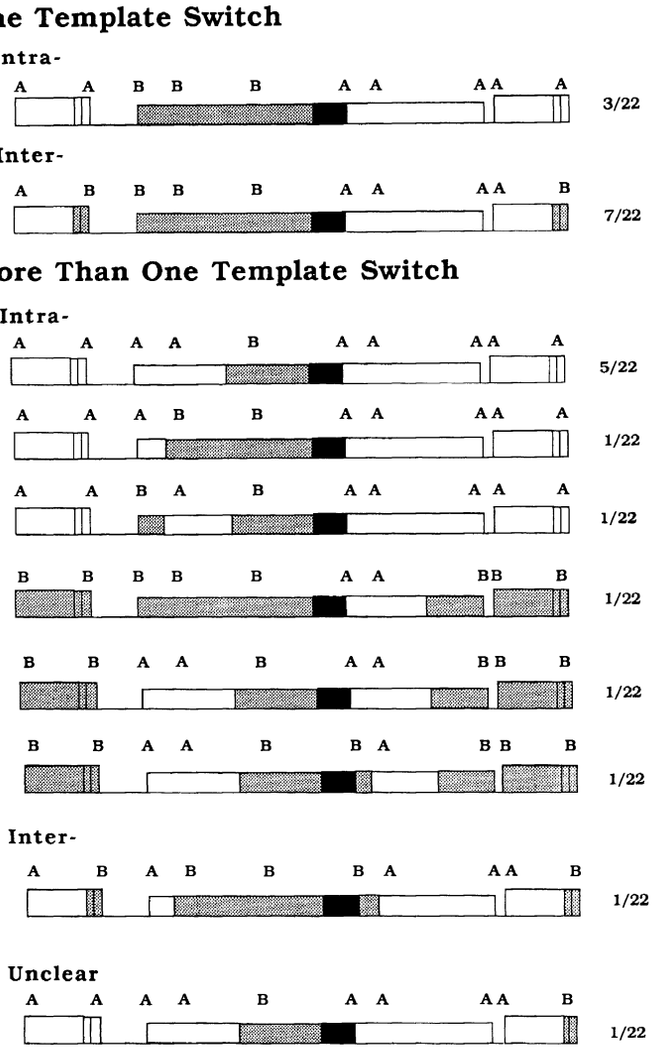

Fig. 6. Restriction enzyme maps of the entire recombinant proviruses. Recombinants containing identical markers are shown together. The number on the right of each pattern indicates the frequency of recombinants containing that pattern. Some patterns appeared frequently. Since the clones that contained these proviruses resulted from infection by viruses harvested from different helper cell clones and from different petri dishes, they represented independent recombination events. Recombinant proviruses are divided into two groups, one template switch and more than one template switch. These two groups were further divided by the types of minus strand strong stop DNA transfer used. Intra, intramolecular minus strand strong stop DNA transfer; inter, intermolecular minus strand strong stop DNA transfer. The number of template switches was determined by the minimal number of switches required for a particular pattern. For example, it is possible that the one template switch recombinants actually contained two template switches by having the second switches between the 5′ LTR and the 5′ end of the *neo* gene. However, it is unlikely that all of the ten recombinants contained a second unselected switch event within a 350-bp distance. Thus, these recombinants are presented as one template switch recombinants.

indicates that the retrovirus virion is organized to maximize viability and variation (*28*).

REFERENCES AND NOTES

1. F. Stahl, *Sci. Am.* **256**, 90 (1987).
2. R. Kucherlapati and G. R. Smith, *Genetic Recombination* (American Society of Microbiology, Washington, DC, 1988).
3. R. Weiss, N. Teich, H. Varmus, J. Coffin, *RNA Tumor Viruses* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, ed. 2, 1982).
4. P. K. Vogt, *Virology* **46**, 947 (1971); S. Kawai and H. Hanafusa, *ibid.* **49**, 37 (1972); J. A. Wyke and J. A. Beamand, *J. Gen. Virol.* **43**, 349 (1979); M. Linial and S. Brown, *J. Virol.* **31**, 257 (1979).
5. W.-S. Hu and H. M. Temin, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 1556 (1990).
6. R. A. Weiss, W. S. Mason, P. K. Vogt, *Virology* **52**, 535 (1973).
7. E. Gilboa, S. W. Mitra, S. Goff, D. Baltimore, *Cell* **18**, 93 (1979); H. M. Temin, *ibid.* **27**, 1 (1981).
8. P. K. Vogt, in *Possible Episomes in Eukaryotes*, L. Sylvestri, Ed. (North-Holland, Amsterdam, 1973), pp. 35–41; J. M. Coffin, *J. Gen. Virol.* **42**, 1 (1979).
9. E. Hunter, *Current Topics Microbiol. Immunol.* **79**, 295 (1978); A. M. Skalka, L. Boone, R. Junghans, D. Luk, *J. Cell. Biochem.* **19**, 293 (1982).
10. H. Xu and J. D. Boeke, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 8553 (1987).
11. H. E. Varmus *et al.*, *J. Mol. Biol.* **120**, 55 (1978); T. W. Hsu and J. Taylor, *J. Virol.* **44**, 47 (1982).
12. R. A. Jorgenson *et al.*, *Mol. Gen. Genet.* **177**, 65 (1979).
13. L. Gritz and J. Davis, *Gene* **25**, 179 (1983).
14. J. P. Dougherty and H. M. Temin, *Mol. Cell. Biol.* **6**, 4387 (1986).
15. When JD216NeoHy viruses were used to infect D17 cells, similar numbers of G418, hygromycin, and G418 plus hygromycin–resistant colonies resulted (*5* and *14*); G418 is an analog of neomycin.
16. J. L. Riggs, R. M. McAllister, E. H. Lennette, *J. Gen. Virol.* **25**, 21 (1974).
17. Precautions were taken to ensure that only one round of retroviral replication occurred. All helper cells were propagated in the presence of antibody to SNV. Helper cell clones containing both parental viruses were also subjected to DNA hybridization. No evidence of reinfection was observed.
18. Homologous recombination could occur between proviruses integrated into the host chromosomes like cellular genes. With a set of modified retroviral vectors having mutations similar to the vectors that we used, the frequency of such events was measured as 10^{-5} (S. Yang, W.-S. Hu, H. M. Temin, unpublished data). This frequency is three orders of magnitude lower than that of the homologous recombination events described in this article.
19. The frequency of recombinants was calculated by comparing the number of doubly resistant clones to the smaller of the two populations of singly resistant clones. This frequency was 2 percent. If compared to the number of both parents, the frequency was about 1 percent. The percentage of recombination events that can be observed is 25 percent. Half of the virions are homodimers, containing the same RNA's; thus recombination events cannot be observed. The other half of the population consists of heterodimers. However, 50 percent of the recombination events in heterodimers resulted in recombinants containing two mutant genes and thus cannot be scored. The recombination rate for two markers that are 1 kb apart is 4 percent (1 percent × 4). For a virus with a genome of 10 kb, the rate of recombination is 40 percent (4 percent × 10). For details of calculations, see discussion of (*5*).
20. E. M. Southern, *J. Mol. Biol.* **98**, 503 (1975).
21. It can be argued that those recombinants that appeared to have one template switch may actually have three, five, or seven switches, whereas the recombinants that appeared to have two template switches may have four, six, or eight switches. However, since the rate of recombination predicts that one in seven vector virions has one recombination event, it is more likely for these recombinants to have the number of switches they appear to have, that is, one or two switches in the template used.
22. A. T. Panganiban and D. Fiore, *Science* **241**, 1064 (1988).
23. There were several differences in the systems used in the papers. (i) Our vectors were larger than those used in (*22*). (ii) The relative expression of the two vectors was different; in (*22*), the titer of one of the vectors was 1000 times higher than that of the other vector; in our system the titers of the two vectors were similar. (iii) The methods for introducing the vectors into the cells differed; in (*22*) both vectors were introduced by transfection. In our system both vectors were introduced by infection. (iv) Helper cell line C3A2 was used in (*22*) and we used DSDh. Both cell lines were derived from D17 cells. (v) Location of one of the LTR markers was slightly different in these two experiments. LTR markers in pJD214Hy and pWH13 were the same; the markers in pAPHneo and pWH204 were different. For the U5 markers, pAPHneo contained a Hind III linker insertion in an Rsa I site, whereas WH204 contained a 4-bp insertion in a Bam HI site; these two markers are 20 to 30 bp apart. The U3 markers of these two vectors were located in the same Sac I site, but the nature of the sequences was different. pAPHneo was derived from a mutant that lacked the Sac I site, while pWH204 was derived by the insertion of a Not I linker into the blunted Sac I site. (vi) The quality of the RNA in the virions might have been different. However, both reports indicated ordered reverse transcription.
24. The expected LTR from intermolecular transfer from WH13 to WH204 was not observed. SNV vectors with these LTR's are functional and viable (*25*). We propose later that intermolecular transfer of minus strand strong stop DNA was correlated with forced copy-choice recombination. Thus, proviruses containing these types of LTR's probably underwent minus strand recombination. In such a provirus, the growing point of the minus strand DNA synthesis would have to change template twice between the two LTR's: once between the U3 sequences and the selected marker in the *hygro* gene and again between the two selected markers to form a doubly wild-type virus. With two required recombination events, this recombinant is expected to be observed in less than 1 of 31 recombinants (the recombination rate between two markers that are 1 kb apart is 4 percent; the distance between the U3 markers and the selected marker in the *hygro* gene is 0.8 kb; thus, the rate of recombination between these two markers is 3.2 percent; since all the recombinants observed in this system have the recombination required to obtain both selected markers, the probability of such a recombinant having another recombination event between the U3 and *hygro* markers would be 3.2 percent or 1 in 31 recombinants).
25. V. K. Pathak and H. M. Temin, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6019 (1990).
26. Although the provirus in 1E1 contained two different U5 markers, it also contained two identical U3 markers indicating that the plus strand strong stop DNA transfer was intramolecular. Thus, this provirus could have been generated through three routes: (i) deletion, (ii) mismatch repair of the 5′ U5 marker, or (iii) mismatch repair of the 3′ U5 marker. The Cla I site in the U5 sequences of WH204 was generated by inserting 4 bp at the Bam HI site. A 4-bp deletion could revert this Cla I site into a Bam HI site and result in different markers in the two U5

sequences. However, since the reversion rate in retroviral replication of such a 4-bp insertion is lower than 10^{-7} per insertion (*5*), this explanation is unlikely. Different markers in the two U5 regions may also result from mismatch repair of either the 5′ U5 or the 3′ U5 region. If the 5′ U5 region experienced the repair and resulted in a different marker than the 3′ U5 region, then the minus strand transfer had to be intermolecular (from WH204 RNA to WH13 RNA) to form the 3′ LTR. Suppose the minus strand strong stop DNA synthesis of WH13 RNA never initiated. As a result u5 and pbs were intact. Then it was possible for the minus strand DNA synthesis to continue and copy the U5 marker of WH13 (Bam HI). After completion of DNA synthesis and repair, the Bam HI marker remained in the 5′ U5 and resulted in different markers in the two U5 regions. It is also possible that the 3′ U5 region was repaired to a different marker. Then the minus strand strong stop transfer had to be intramolecular (from WH13 to WH13) to result in the 5′ WH13-like LTR. Suppose a strand displacement event occurred at the end of the 3′ LTR and formed a mismatch at the 3′ U5: the minus strand DNA contained a Bam HI site and the plus strand contained a Cla I site from the displaced fragment. After repair, the Cla I marker remained and resulted in the two different U5 markers. It is most likely that the different markers in the two U5 sequences were generated from mismatch repair of one of the LTR's. However, it is not clear which pathway generated this recombinant provirus.

27. Fisher's exact test was calculated with 21 recombinants. The recombinant with two different markers in the U5 was not included because it is not clear which type of minus strand transfer was used. Calculation including this sample was also performed, and the result indicated the distribution still is not random.
28. H. M. Temin, *Trends Genet.*, in press.
29. We thank Rebecca Wisniewski, Bonnie Fritz, and Kevin Krebsbach for technical assistance; N. Drinkwater, W. Sugden, and M. Susman for critical reading of the manuscript; M. Hannink, K. Iwasaki, V. Pathak, and G. Pulsinelli for discussion and helpful comments. Supported by Public Health Service grants CA-22443 and CA-07175 from the National Cancer Institute; a postdoctoral fellowship from the American Cancer Society (W.-S.H.); and an American Cancer Society Research Professorship (H.M.T.)

13 June 1990; accepted 19 September 1990

Germ Line p53 Mutations in a Familial Syndrome of Breast Cancer, Sarcomas, and Other Neoplasms

DAVID MALKIN, FREDERICK P. LI, LOUISE C. STRONG, JOSEPH F. FRAUMENI, JR., CAMILLE E. NELSON, DAVID H. KIM, JAYNE KASSEL, MAGDALENA A. GRYKA, FARIDEH Z. BISCHOFF, MICHAEL A. TAINSKY, STEPHEN H. FRIEND*

Familial cancer syndromes have helped to define the role of tumor suppressor genes in the development of cancer. The dominantly inherited Li-Fraumeni syndrome (LFS) is of particular interest because of the diversity of childhood and adult tumors that occur in affected individuals. The rarity and high mortality of LFS precluded formal linkage analysis. The alternative approach was to select the most plausible candidate gene. The tumor suppressor gene, p53, was studied because of previous indications that this gene is inactivated in the sporadic (nonfamilial) forms of most cancers that are associated with LFS. Germ line p53 mutations have been detected in all five LFS families analyzed. These mutations do not produce amounts of mutant p53 protein expected to exert a trans-dominant loss of function effect on wild-type p53 protein. The frequency of germ line p53 mutations can now be examined in additional families with LFS, and in other cancer patients and families with clinical features that might be attributed to the mutation.

In 1969, LI AND FRAUMENI REVIEWED MEDICAL RECORDS and death certificates of 648 childhood rhabdomyosarcoma patients and identified four families in which siblings or cousins had a childhood sarcoma (*1*). These four families also had striking histories of breast cancer and other neoplasms, suggesting a new familial cancer syndrome of diverse tumors (Li-Fraumeni syndrome; LFS). Recently completed prospective studies have confirmed the high risk in family members of the tumor types that comprise LFS (*2*). Since the original description of the syndrome, systematic studies and anecdotal reports have confirmed its existence in various geographic and ethnic groups (*3*). The spectrum of cancers in the syndrome (Table 1) has been determined to include breast carcinomas, soft tissue sarcomas, brain tumors, osteosarcoma, leukemia, and adrenocortical carcinoma. Possible component tumors of LFS are melanoma, gonadal germ cell tumors, and carcinomas of the lung, pancreas, and prostate (*4, 5*). These diverse tumor types in family members characteristically develop at unusually early ages, and multiple primary tumors are frequent.

To test the hypothesis that the Li-Fraumeni syndrome has a genetic etiology, Williams and Strong (*5*) applied segregation analysis and demonstrated that the observed cancer distribution in families best fit a rare autosomal dominant gene model. This model also predicted that the probability, for the families at risk, of developing any invasive cancer (excluding carcinomas of the skin)

D. Malkin, C. E. Nelson, D. H. Kim, J. Kassel, M. A. Gryka, and S. H. Friend are in the Division of Molecular Genetics, Massachusetts General Hospital Cancer Center, MGH East, Charlestown MA 02129. F. P. Li and C. E. Nelson are in the Clinical Epidemiology Branch, Division of Cancer Etiology, NCI, NIH, Bethesda MD, 20892 and in the Division of Biostatistics and Epidemiology, Dana-Farber Cancer Institute, Boston, MA, 02115. L. C. Strong is in the Division of Pediatrics, University of Texas M. D. Anderson Cancer Center, Houston, TX 77030. J. F. Fraumeni, Jr., is in the Epidemiology and Biostatistics Program, Division of Cancer Etiology, NCI, NIH, Bethesda MD, 20892. F. Z. Bischoff and M. A. Tainsky are in the Department of Tumor Biology, University of Texas M. D. Anderson Cancer Center, Houston TX, 77030. S. H. Friend is also in the Division of Hematology-Oncology, The Children's Hospital, Dana-Farber Cancer Institute, Department of Pediatrics, Harvard Medical School, Boston MA, 02114.

*To whom correspondence should be addressed.