

Effects of Varying Sequence Similarity on the Frequency of Repeat Deletion during Reverse Transcription of a Human Immunodeficiency Virus Type 1 Vector

Wenfeng An and Alice Telesnitsky*

Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, Michigan 48109-0620

Received 19 December 2001/Accepted 29 April 2002

Genetic recombination contributes to human immunodeficiency virus type 1 (HIV-1) diversity, with homologous recombination being more frequent than nonhomologous recombination. In this study, HIV-1-based vectors were used to assay the effects of various extents of sequence divergence on the frequency of the recombination-related property of repeat deletion. Sequence variation, similar in degree to that which differentiates natural HIV-1 isolates, was introduced by synonymous substitutions into a gene segment. Repeated copies of this segment were then introduced into assay vectors. With the use of a phenotypic screen, the deletion frequency of identical repeats was compared to the frequencies of repeats that differed in sequence by various extents. During HIV-1 reverse transcription, the deletion frequency observed with repeats that differed by 5% was 65% of that observed with identical repeats. The deletion frequency decreased to 26% for repeats that differed by 9%, and when repeats differed by 18%, the deletion frequency was about 5% of the identical repeat value. Deletion frequencies fell to less than 0.3% of identical repeat values when genetic distances of 27% or more were examined. These data argue that genetic variation is not as inhibitory to HIV-1 repeat deletion as it is to the corresponding cellular process and suggest that, for sequences that differ by about 25% or more, HIV-1 recombination directed by sequence homology may be no more frequent than that which is homology independent.

Human immunodeficiency virus type 1 (HIV-1) is characterized by its extraordinary genetic diversity. Factors which contribute to this diversity include the low fidelity of HIV-1 DNA synthesis (approximately 3×10^{-5} mutations per base per replication cycle [19]), a high recombination frequency between copackaged RNAs (two to three recombination events per genome per replication cycle [12, 35, 42]), a high level of virus production (10^8 to 10^{10} virions per day [8, 29, 39]), a large number of viral replication cycles (140 to 300 cycles per year [4, 29]), and a large pool of infected individuals. Due to the high rate of HIV-1 reverse transcription errors, the virus exists within its host as a pool of closely related genetic variants, known collectively as a quasispecies. Within infected individuals, both genetic divergence (differences from the founder strain) and diversity (breadth of genetic variation within the viral population) increases by $\sim 1\%$ per year during the early phase of viral infection, with increases in divergence typically continuing for the duration of the infection (33). Thus, differences between two HIV-1 DNAs isolated from a singly infected individual can be used to estimate the number of viral generations that separate the two isolates. In contrast to point mutations, which accumulate gradually, recombination shuffles the genetic information between two genomes, providing the potential for evolutionary leaps with far more dramatic genetic consequences than the incremental accumulation of individual mutations due to base misinsertion errors (18).

Recombination for retroviruses such as HIV-1 results from template switching between two RNAs during reverse transcription. Frequencies of recombination can be determined experimentally if virions that contain two different vector RNAs engineered to contain selectable markers are produced (41). To generate heterozygous virions, two genetically marked viral RNAs are coexpressed in virus-producing cells, and the resulting virions are used to infect susceptible cells. Cells with integrated reverse transcription products for which the heterozygous virions serve as templates are then tested for expression of each vector's marker either separately or in combination with the second marker. Frequencies of homologous (between identical donor and acceptor sequences) and nonhomologous recombination (between unrelated sequences) have been compared in this way for gamma retroviruses, and it was determined that homologous recombination is at least 10^3 times more frequent than nonhomologous recombination (45).

Among global isolates, the sequence divergence of HIV-1 viruses is great. Based on phylogenetic analysis, three distinct groups of HIV-1 (termed groups M, N, and O) have been identified, with group M further divided into at least nine subtypes (A to D, F to H, J, and K) (32). Among isolates within a given subtype, extents of sequence variation are relatively modest (0 to 15% in the *env* nucleotide sequence). Different subtypes differ from each other by roughly 15 to 30% in the *env* nucleotide sequence (31). Between two given groups, significant sequence divergence exists: for example, the homology of group M and O viruses in various genome regions ranges from 55 to 76% and is about 65% on average (36).

Despite this variation, the frequent observation of natural recombinants suggests that recombination is a relatively com-

* Corresponding author. Mailing address: Department of Microbiology and Immunology, University of Michigan Medical School, 1150 W. Medical Ctr. Dr., Rm. 5641, Ann Arbor, MI 48109-0620. Phone: (734) 936-6466. Fax: (734) 764-3562. E-mail: ateles@umich.edu.

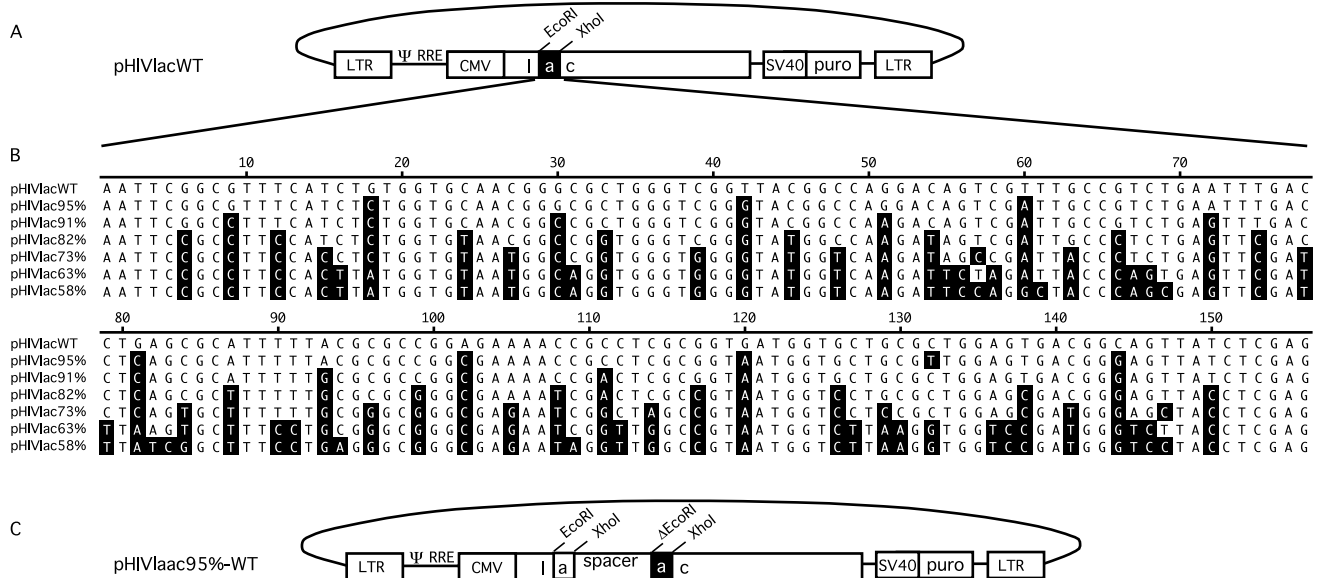


FIG. 1. HIV-1-transducing vectors. (A) Structure of pHIVlacWT. LTR, long terminal repeat; Ψ , packaging signal; RRE, Rev response element; CMV, cytomegalovirus immediate-early promoter; SV40, simian virus 40 promoter; puro, puromycin resistance gene. (B) Nucleotide sequence alignment of 156-nt *EcoRI-XhoI lacZ* fragments from pHIVlacWT, pHIVlac95%, pHIVlac91%, pHIVlac82%, pHIVlac73%, pHIVlac63%, and pHIVlac58%. Nucleotides of the variants that differed from wild type sequence are highlighted (white type on black background). (C) Structure of pHIVlac95%-WT, a representative repeat vector. Its 5' repeat was derived from the 156-nt *lacZ* fragment of pHIVlac95%, and its 3' repeat was derived from pHIVlacWT. Abbreviations are as defined for panel A.

mon occurrence among different strains of HIV-1 (27, 31). The majority of HIV-1 recombinants characterized thus far contain sequences from members of different group M subtypes. Some of these are classified as circulating recombinant forms and have been found to contribute to the global HIV-1 pandemic (27, 32). Intrasubtype recombinants have been identified within individual patients and may be more common than has been recognized due to the lower sequence divergence among parental strains. Recombination can apparently occur even among distantly related strains, because a limited number of intergroup M/O recombinants have been reported (28, 36). However, whether the generation of viable in-frame M/O recombinants is a chance event, detectable only after further virus replication, and how the frequency of recombination is affected by sequence divergence such as that among natural HIV-1 isolates have not been determined.

In this work, we addressed the effects of systematic variation in the extents of sequence homology on recombination rates, using HIV-1-based assay vectors designed to measure the recombination-related property of repeated sequence deletion. Repeat deletion has frequently been used to measure retroviral recombination properties (6, 11, 14, 26, 30). We have previously studied HIV-1 repeat deletion with vectors engineered to contain coding region internal duplications within *lacZ*, so that a functional *lacZ* gene would result only if faithful deletion of the repeated sequence occurred during reverse transcription. We showed that 117-, 284-, and 971-base identical direct repeats were deleted in 6.8, 19.9, and 87%, respectively, of all single-cycle reverse transcription products, which was consistent with previous gamma retrovirus work showing that repeat deletion is roughly proportional to repeat length (1, 6, 14, 30). In the present work, we designed a series of vectors in which

the two repeats ranged from 100% identical to only 58% homologous, which provided sequence variation similar to that of intrasubtype, intersubtype, and intergroup HIV-1 strains. We determined that the frequency of template switching decreased dramatically as sequence divergence between repeated sequences increased.

Establishing assay vectors that contain sequence degeneracy within *lacZ*. We generated an HIV-1-based transducing vector (pHIVlacWT) that contained expression cassettes for both the puromycin resistance selectable marker gene and for the β -galactosidase (*lacZ*) gene (Fig. 1A). pHIVlacWT was a derivative of pHIVlac (1), with silent mutations that generated *EcoRI* and *XhoI* restriction sites flanking a 156-bp region of *lacZ*. Derivatives of pHIVlacWT in which the 156-bp sequence was 95, 91, 82, 73, 63, or 58% homologous to the wild-type sequence were then constructed. For each derivative, a pair of lengthy sense and antisense oligonucleotides with complementary 3'-terminal ends were annealed, and standard procedures for *Escherichia coli* DNA polymerase I Klenow fragment were used to mutually prime and fill out the oligonucleotides. *EcoRI* and *XhoI* restriction sites engineered into the 5' ends of the oligonucleotides were used to generate replacement fragments, forming pHIVlac95%, pHIVlac91%, pHIVlac82%, pHIVlac73%, pHIVlac63%, and pHIVlac58% (Fig. 1B). Sequence modifications introduced into this region followed three criteria, as follows: (i) the amino acid coding was not altered; (ii) the mutations were distributed evenly over the region; and (iii) mutations in the variants were introduced based on the codon frequency of the parental *lacZ* gene (to minimize any effect of codon usage on translational efficiency) wherever possible. The integrity of the *lacZ* variants was confirmed by transfecting individual vectors into ET cells—the ET

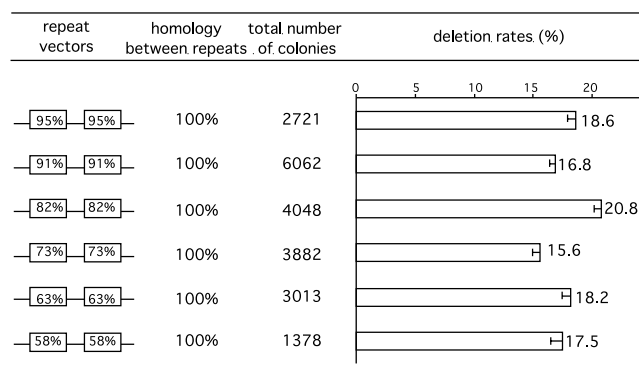


FIG. 2. Repeat deletion rates for vectors with identical *lacZ* variant repeats. The repeat vector structures are illustrated with boxed percentages (e.g., the first repeat vector structure listed represents pHIVlac95%-95%). The homology between repeats was 100% for each vector in this series. The repeat deletion rates were calculated as described in the text (means [bars] minus standard errors of the means [error bars] are shown).

cell line is a 293T derivative that constitutively expresses ectopic envelope (30)—and examining *lacZ* expression with 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-Gal; Sigma) staining as previously described (30) (results not shown).

Repeat vectors were then constructed by joining a plasmid half that contained the 156-bp portion of *lacZ* and upstream sequences from one pHIVlac derivative with a plasmid half that included the 156-bp segment and downstream sequences from a second vector, using the *Xho*I and *Eco*RI sites that flanked the varied portion of *lacZ* and the plasmid backbone site, *Bcl*I. These two larger fragments were ligated with a 1.5-kb *Xho*I-*Mfe*I linker fragment that contained murine leukemia virus coding sequence from pNCA (5). In this way, repeat vectors with various combinations of 5' and 3' repeats were constructed. For example, pHIVlac95%-WT contained a 156-bp segment that differed by 5% from that of the wild type, followed by the linker segment and then a wild type 156-bp segment (Fig. 1C). It has been shown for murine leukemia virus that deletion rates increase when repeated sequences are separated by a spacer sequence, presumably because the distance between repeats increases the chance of template switching (6).

Assaying HIV-1 template switching by deletion between identical repeats engineered into *lacZ*. It has been suggested that reverse transcriptase template switches do not occur at uniform frequencies and are highly context dependent (2, 40). For example, sequences that potentially form secondary structures have been shown to be hot spots for both mutation and recombination (12, 13, 25). Thus, to address potential sequence-specific effects of the introduced variation on template switching rates, the initial deletion vectors tested contained each of the 156-nucleotide (nt) region sequence variants in tandem with an identical repeat of itself (Fig. 2). Because the repeats in each of these vectors were identical to one another, differences among vectors' deletion rates would reveal whether or not particular 156-nt variants displayed significant differences in intrinsic template switching properties.

HIV-1 vector virions were generated by transiently cotrans-

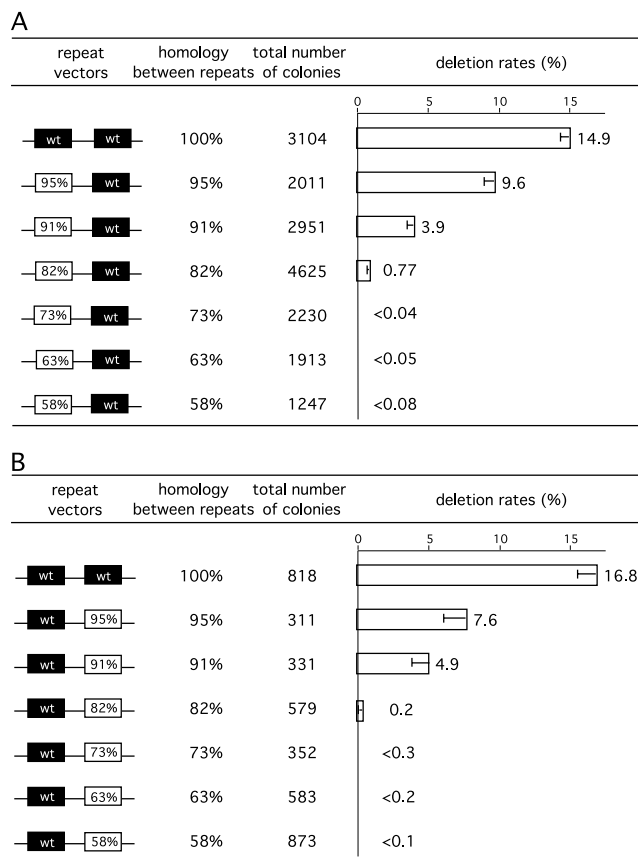


FIG. 3. Effects of various extents of sequence identity on repeat deletion frequencies. (A) Repeat deletion rates for vectors with wild-type *lacZ* sequence in the 3' repeat. The repeat vector structures are illustrated with boxed percentages (the 5' and 3' repeats are as indicated in the left and right boxes, respectively). For example, the second repeat vector structure listed represents pHIVlac95%-WT, where the homology between repeats was 95%. Deletion rates were calculated as described in the text (means [bars] minus standard errors of the means [error bars] are shown). (B) Repeat deletion rates for vectors with wild-type *lacZ* sequence in the 5' repeat. The methods for illustrating repeat vector structures and calculation of deletion rates are as described for panel A.

fecting ET cells with the HIV-1 helper plasmid, pCMV Δ R8.2 (23), and each of the identical repeat vectors (Fig. 2). The resulting virions were harvested, and mATRC1/293 cells (17) were infected, selected in puromycin, and assayed for β -galactosidase by blue-white screening as previously reported (30). For each vector, three independent infections were performed in triplicate to assess deletion rates during single replication cycles. The numbers of blue and white colonies were summed for each infection experiment, and deletion rates were calculated as ratios of blue colonies to the total number of puromycin-resistant colonies. To correct for the *lacZ* inactivation rate (5.0% for pHIVlacWT [data not shown]), the deletion rates calculated as described above were divided by 0.95 (i.e., $1 - 0.05$). The corrected deletion rates were then averaged from three independent infections for each vector. For the 156-nt wild-type tandem repeat (wt-wt), the measured deletion rate was 15.6% (average calculated from data shown in Fig. 3A and B). Deletion rates for vectors containing identical repeats

of 95, 91, 82, 73, 63, and 58% homology to the wild-type sequence were all similar in value and were observed to be 18.6, 16.8, 20.8, 15.6, 18.2, and 17.5%, respectively (Fig. 2). Note that pairwise χ^2 tests suggest that deletion rates for certain pairs of vectors, such as the 82%-82% and 73%-73% vectors, may differ significantly ($P < 0.05$). However, most pairwise comparisons did not result in statistically different values; thus, observed deletion rate variation among identical repeat vectors should not affect interpretation of the trends reported here.

Deletion rates between nonidentical repeats. Vectors with two different repeats were constructed to test the effects of sequence divergence on repeat deletion (Fig. 3). It is believed that repeat deletion takes place mostly during retroviral DNA minus strand synthesis when the growing point of the newly synthesized minus strand DNA dissociates from the 3'-repeat template RNA (template switch donor) and reassociates with the homologous 5' repeat (template switch acceptor) (44). Thus, deletion rates were first determined for vector viruses containing the wild-type *lacZ* sequence as their 3' repeat, because identical sequences would serve as the template switch donor for each vector under these conditions, provided that template switching occurred during minus strand recombination. As shown in Fig. 3A, deletion rates were 9.6, 3.9, 0.77, <0.04, <0.05, and <0.08% for repeats of 95, 91, 82, 73, 63, and 58% homologies, respectively.

An additional set of vectors was constructed which differed from the first in that the orientation of the repeats was inverted. Thus, the new vectors contained different variants as template switching donors and identical wild-type sequences as template switching acceptors (Fig. 3B). When assayed, deletion rates were determined to be 7.6, 4.9, 0.2, <0.3, <0.2, and <0.1% for repeats of 95, 91, 82, 73, 63, and 58% homologies, respectively (Fig. 3B). Although the total number of colonies examined for each of the vectors in this second set was lower than that of the previous set, deletion rate trends associated with sequence modulation were consistent with those observed for the set of vectors analyzed in Fig. 3A. Statistical analysis revealed no significant difference between the deletion rates between the two sets of vectors (χ^2 test or Fisher exact test, $P > 0.05$).

The work in this study tested effects of altering the genetic distance between repeated sequences on HIV-1 template switching. HIV-1 deletion frequencies for repeats of 95% homology were about 65% as high as values for identical repeats. The frequency decreased to 26% of the identical repeat rate for repeats with 91% homology and was about 5% for repeats with 82% homology. Failure to detect any recombinants among thousands of products that resulted from vectors with repeats that differed by 27% (73% homology) suggests that this level of divergence decreased recombination frequency more than 300-fold. Considering that nonhomologous recombination reportedly occurs roughly 1,000-fold less frequently than homologous retroviral recombination (45), this level of recombination suppression appears similar in magnitude to that predicted for two unrelated sequences. The effects of sequence context on repeat deletion rates were tested by placing each variant repeat in tandem with itself. Whether or not there were hotspots for template switching in our 156-nt sequences was not addressed directly. However, the level of similarity in de-

letion rates observed for the different identical repeat vectors suggests that the magnitude of any possible hotspots was too low to alter the above conclusion.

These results suggest that retroviral recombination is more permissive of genetic variation than is bacterial or mammalian DNA recombination. It has previously been established that homologous recombination is far more frequent during retroviral replication than during cellular replication (10, 15): our new findings suggest that greater levels of sequence divergence are tolerated during retroviral recombination as well. Studies on cellular homology requirements during recombination usually take one of two approaches: either to determine the rate of recombination as a function of the length of shared uninterrupted homology or to test the effects of base pair mismatch on the rate of recombination (see reference 37 and citations therein). The system used here, whereby mismatch was introduced into the 156-base *lacZ* gene fragment, resulting in variants with differing degrees of sequence homology, could be considered as following the second approach. In bacteria, a 16% mismatch between sequences reduces phage-plasmid recombination rates by a factor of 100 (34), and in mammalian cells, 19% mismatch reduced intrachromosomal recombination over 1,000-fold (38). In comparison, our results show 18% mismatch reduced HIV-1 repeat deletion by roughly 20-fold (comparing the observed deletion rate of 0.77% reported here for 82% homologous repeats with the 14.9% deletion rate determined for identical repeats). We realize that this comparison has its limitation because the corresponding rate of cellular recombination was not measured for the 156-bp sequence used in this study. Nevertheless, the findings presented here suggest that HIV-1 recombination may be less sensitive to sequence variation than is the cellular replication machinery. Recent research has demonstrated that DNA mismatch repair systems play critical roles in inhibiting recombination between nonidentical sequences in both bacteria and eukaryotes (7). Mutational inactivation of components of the mismatch repair system can significantly reduce the block to homeologous recombination between similar but nonidentical sequences (22). Whether or not such factors as the host cells' mismatch repair machinery affected our read-out of apparent HIV-1 deletion frequencies was not addressed here.

Here, repeat deletion was used to measure the effect of sequence divergence on HIV-1 template switching, whereas genetic recombination as commonly defined occurs between sequences on different copackaged RNAs. Repeat deletion appears to occur primarily through intramolecular template switching, at least for gamma retroviruses (9, 43). In contrast, recombination between different strains of HIV-1 is an intermolecular event, and our assays did not address parameters that differ between inter- and intramolecular template switching. These parameters include the physical interaction which has been implicated in promoting recombination at the region flanking the dimerization initiation sequence (3, 16, 20, 21). Indirect evidence from melting profiles of nicked genomic RNAs suggests that extensive RNA-RNA interactions exist at other regions of the RNA genome as well (24), and these may also contribute to intermolecular recombination.

Another way our approach differed from intermolecular recombination is that tandem repeats provide a total of four possible templates for the assayed region (two on each copack-

aged RNA) instead of the two present for intermolecular recombination. Thus, in our system, competitive effects of the homologous repeat on the copackaged RNA may have affected deletion rates. However, since such competition exists for both nonidentical and identical repeat vectors, the observed decreases in deletion clearly reflect effects of decreasing homology on template switching, though they may misrepresent the magnitude of such effects during intermolecular switching.

Not surprisingly, our results suggest that regions with high sequence similarity are more likely to serve as HIV-1 recombination substrates than sequences that differ more significantly. However, as evidenced by the natural occurrence of M/O recombinants, whose parent strains differ by an average of roughly 35%, recombination between highly divergent sequences can take place, even though scattered 27% sequence differences reduced recombination frequencies greater than 300-fold in our assays. Selective pressures lead to differences in extents of sequence variation among isolates in different portions of the HIV-1 genome, with some regions more conserved than others, and it is possible that local regions of similarity serve disproportionately as recombination crossover sites. However, due to the very high replication rate of HIV-1, even recombination products that arise at frequencies no higher than those between unrelated sequences likely contribute to viral populations if they confer a selective advantage (4).

This work was supported by National Institutes of Health grant GM64479.

REFERENCES

- An, W., and A. Telesnitsky. 2001. Frequency of direct repeat deletion in a human immunodeficiency virus type 1 vector during reverse transcription in human cells. *Virology* **286**:475–482.
- Anderson, J. A., E. H. Bowman, and W.-S. Hu. 1998. Retroviral recombination rates do not increase linearly with marker distance and are limited by the size of the recombining subpopulation. *J. Virol.* **72**:1195–1202.
- Balakrishnan, M., P. J. Fay, and R. A. Bambara. 2001. The kissing hairpin sequence promotes recombination within the HIV-1 5' leader region. *J. Biol. Chem.* **276**:36482–36492.
- Coffin, J. M. 1995. HIV population dynamics in vivo: implications for genetic variation, pathogenesis and therapy. *Science* **267**:483–489.
- Colicelli, J., and S. P. Goff. 1988. Sequence and spacing requirements of a retrovirus integration site. *J. Mol. Biol.* **199**:47–59.
- Delviks, K. A., and V. K. Pathak. 1999. Effect of distance between homologous sequences and 3' homology on the frequency of retroviral reverse transcriptase template switching. *J. Virol.* **73**:7923–7932.
- Harfe, B. D., and S. Jinks-Robertson. 2000. DNA mismatch repair and genetic instability. *Annu. Rev. Genet.* **34**:359–399.
- Ho, D., A. Neumann, A. Perelson, W. Chen, J. Leonard, and M. Markowitz. 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* **373**:123–126.
- Hu, W.-S., E. H. Bowman, K. A. Delviks, and V. K. Pathak. 1997. Homologous recombination occurs in a distinct retroviral subpopulation and exhibits high negative interference. *J. Virol.* **71**:6028–6036.
- Hu, W.-S., V. Pathak, and H. M. Temin. 1993. Role of reverse transcriptase in retroviral recombination, p. 251–274. *In* A. M. Skalka and S. P. Goff (ed.), *Reverse transcriptase*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Hwang, C. K., E. S. Svarovskaia, and V. K. Pathak. 2001. Dynamic copy choice: steady state between murine leukemia virus polymerase and polymerase-dependent RNase H activity determines frequency of in vivo template switching. *Proc. Natl. Acad. Sci. USA* **98**:12209–12214.
- Jetzt, A. E., H. Yu, G. J. Klarmann, Y. Ron, B. D. Preston, and J. P. Dougherty. 2000. High rate of recombination throughout the human immunodeficiency virus type 1 genome. *J. Virol.* **74**:1234–1240.
- Jones, J. S., R. W. Allan, and H. M. Temin. 1994. One retroviral RNA is sufficient for synthesis of viral DNA. *J. Virol.* **68**:207–216.
- Julias, J. G., D. Hash, and V. K. Pathak. 1995. E- vectors: development of novel self-inactivating and self-activating retroviral vectors for safer gene therapy. *J. Virol.* **69**:6839–6846.
- Li, T., and J. Zhang. 2000. Determination of the frequency of retroviral recombination between two identical sequences within a provirus. *J. Virol.* **74**:7646–7650.
- Lund, A. H., J. G. Mikkelsen, J. Schmidt, M. Duch, and F. S. Pedersen. 1999. The kissing-loop motif is a preferred site of 5' leader recombination during replication of SL3–3 murine leukemia viruses in mice. *J. Virol.* **73**:9614–9618.
- Malhotra, S., A. G. Scott, T. Zavorotinskaya, and L. Albritton. 1996. Analysis of the murine ecotropic leukemia virus receptor reveals a common biochemical determinant on diverse cell surface receptors that is essential to retrovirus entry. *J. Virol.* **70**:321–326.
- Malim, M. H., and M. Emerman. 2001. HIV-1 sequence variation: drift, shift, and attenuation. *Cell* **104**:469–472.
- Mansky, L. M., and H. M. Temin. 1995. Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *J. Virol.* **69**:5087–5094.
- Mikkelsen, J. G., A. H. Lund, M. Duch, and F. S. Pedersen. 2000. Mutations of the kissing-loop dimerization sequence influence the site specificity of murine leukemia virus recombination in vivo. *J. Virol.* **74**:600–610.
- Mikkelsen, J. G., A. H. Lund, M. Duch, and F. S. Pedersen. 1998. Recombination in the 5' leader of murine leukemia virus is accurate and influenced by sequence identity with a strong bias toward the kissing-loop dimerization region. *J. Virol.* **72**:6967–6978.
- Myung, K., A. Datta, C. Chen, and R. D. Kolodner. 2001. SGS1, the *Saccharomyces cerevisiae* homologue of BLM and WRN, suppresses genome instability and homologous recombination. *Nat. Genet.* **27**:113–116.
- Naldini, L., U. Blomer, P. Gallay, D. Ory, R. Mulligan, F. H. Gage, I. M. Verma, and D. Trono. 1996. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* **272**:263–267.
- Ortiz-Conde, B. A., and S. H. Hughes. 1999. Studies of the genomic RNA of leukemia viruses: implications for RNA dimerization. *J. Virol.* **73**:7165–7174.
- Pathak, V. K., and H. M. Temin. 1992. 5-azacytidine and RNA secondary structure increase the retrovirus mutation rate. *J. Virol.* **66**:3093–3100.
- 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: substitutions, frameshifts, and hypermutations. *Proc. Natl. Acad. Sci. USA* **87**:6019–6023.
- Peeters, M. 2000. Recombinant HIV sequences: their role in the global epidemic, p. 139–154. *In* C. L. Kuiken, B. Foley, B. Hahn, B. Korber, F. McCutchan, P. A. Marx, J. W. Mellors, J. I. Mullins, J. Sodroski, and S. Wolinsky (ed.), *Human retroviruses and AIDS 2000*. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, N.Mex.
- Peeters, M., F. Liegeois, N. Torimiro, A. Bourgeois, E. Mpoudi, L. Vergne, E. Saman, E. Delaporte, and S. Saragosti. 1999. Characterization of a highly replicative intergroup M/O human immunodeficiency virus type 1 recombinant isolated from a Cameroonian patient. *J. Virol.* **73**:7368–7375.
- Perelson, A. S., A. U. Neumann, M. Markowitz, J. M. Leonard, and D. D. Ho. 1996. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science* **271**:1582–1586.
- Pfeiffer, J. K., R. Topping, N.-H. Shin, and A. Telesnitsky. 1999. Altering the intracellular environment increases the frequency of tandem repeat deletion during Moloney murine leukemia virus reverse transcription. *J. Virol.* **73**:8441–8447.
- Quinones-Mateu, M. E., and E. J. Arts. 1999. Recombination in HIV-1: update and implications. *AIDS Rev.* **1**:89–100.
- Robertson, D. L., J. P. Anderson, J. A. Bradac, J. K. Carr, B. Foley, R. K. Funkhouser, F. Gao, B. H. Hahn, M. L. Kalish, C. Kuiken, G. H. Learn, T. Leitner, F. McCutchan, S. Osmanov, M. Peeters, D. Pieniazek, M. Salminen, P. M. Sharp, S. Wolinsky, and B. Korber. 1999. HIV-1 nomenclature proposal, p. 492–505. *In* C. L. Kuiken, B. Foley, B. Hahn, B. Korber, F. McCutchan, P. A. Marx, J. W. Mellors, J. I. Mullins, J. Sodroski, and S. Wolinsky (ed.), *Human retroviruses and AIDS: a compilation and analysis of nucleic acid and amino acid sequences*. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, N.M.
- Shankarappa, R., J. B. Margolick, S. J. Gange, A. G. Rodrigo, D. Upchurch, H. Farzadegan, P. Gupta, C. R. Rinaldo, G. H. Learn, X. He, X. L. Huang, and J. I. Mullins. 1999. Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. *J. Virol.* **73**:10489–10502.
- Shen, P., and H. V. Huang. 1986. Homologous recombination in *Escherichia coli*: dependence on substrate length and homology. *Genetics* **112**:441–457.
- St. Louis, D. C., D. Gotte, Sanders-Buell, D. W. Ritchey, M. O. Salminen, J. K. Carr, and F. E. McCutchan. 1998. Infectious molecular clones with the nonhomologous dimer initiation sequences found in different subtypes of human immunodeficiency virus type 1 can recombine and initiate a spreading infection in vitro. *J. Virol.* **72**:3991–3998.
- Takehisa, J., L. Zekeng, E. Ido, Y. Yamaguchi-Kabata, I. Mboudjeka, Y. Harada, T. Miura, L. Kaptue, and M. Hayami. 1999. Human immunodeficiency virus type 1 intergroup (M/O) recombination in Cameroon. *J. Virol.* **73**:6810–6820.

37. **Waldman, A. S., and R. M. Liskay.** 1988. Dependence of intrachromosomal recombination in mammalian cells on uninterrupted homology. *Mol. Cell. Biol.* **8**:5350–5357.
38. **Waldman, A. S., and R. M. Liskay.** 1987. Differential effects of base-pair mismatch on intrachromosomal versus extrachromosomal recombination in mouse cells. *Proc. Natl. Acad. Sci. USA* **84**:5340–5344.
39. **Wei, X., S. K. Gosh, M. E. Taylor, V. A. Johnson, E. A. Emini, P. Deutsch, J. D. Lifson, S. Bonhoeffer, M. A. Nowak, B. H. Hahn, M. S. Saag, and G. M. Shaw.** 1995. Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* **373**:117–122.
40. **Wooley, D. P., L. A. Bircher, and R. A. Smith.** 1998. Retroviral recombination is nonrandom and sequence dependent. *Virology* **243**:229–234.
41. **Yu, H., A. E. Jetzt, and J. P. Dougherty.** 1997. Use of single-cycle analysis to study rates and mechanisms of retroviral mutation. *Methods* **12**:325–336.
42. **Yu, H., A. E. Jetzt, Y. Ron, B. D. Preston, and J. P. Dougherty.** 1998. The nature of human immunodeficiency virus type 1 strand transfers. *J. Biol. Chem.* **273**:28384–28391.
43. **Zhang, J., and Y. Ma.** 2001. Evidence for retroviral intramolecular recombinations. *J. Virol* **75**:6348–6358.
44. **Zhang, J., L. Y. Tang, T. Li, Y. Ma, and C. M. Sapp.** 2000. Most retroviral recombinations occur during minus-strand DNA synthesis. *J. Virol.* **74**:2313–2322.
45. **Zhang, J., and H. M. Temin.** 1993. Rate and mechanism of nonhomologous recombination during a single cycle of retroviral replication. *Science* **259**:234–238.