

High Rate of Recombination throughout the Human Immunodeficiency Virus Type 1 Genome

AMANDA E. JETZT,^{1,2} HONG YU,^{1†} GEORGE J. KLARMANN,³ YACOV RON,¹
BRADLEY D. PRESTON,³ AND JOSEPH P. DOUGHERTY^{1*}

Department of Molecular Genetics and Microbiology, Robert Wood Johnson Medical School, Piscataway, New Jersey 08854¹; Graduate Program in Microbiology and Molecular Genetics, Rutgers University, New Brunswick, New Jersey 08903²; and Departments of Biochemistry and Radiation Oncology, Eccles Institute of Human Genetics, University of Utah, Salt Lake City, Utah 84112³

Received 9 August 1999/Accepted 27 October 1999

The diploid nature of human immunodeficiency virus type 1 (HIV-1) indicates that recombination serves a central function in virus replication and evolution. Previously, while examining the nature of obligatory primer strand transfers during reverse transcription, a high rate of recombination was observed at the ends of the viral genome within the viral long terminal repeats, prompting the following question: does recombination occur at a high rate throughout the genome? To address this question, two vectors based upon different strains of HIV-1 were utilized. The vectors were comprised predominantly of autologous HIV-1 sequence and were approximately the same size as the parental genome. The proviral progeny of heterodimeric virions were analyzed after a single cycle of replication, and the sequence heterogeneity between the two strains allowed direct examination of recombination crossovers. The results obtained indicate that HIV-1 undergoes approximately two to three recombination events per genome per replication cycle. These results imply that both HIV-1 RNAs are typically utilized during reverse transcription and that recombination is an important aspect of HIV-1 replication.

Retroviruses, including human immunodeficiency virus type 1 (HIV-1), are diploid, containing two genomic RNA molecules per virion. This viral RNA serves as the template for proviral DNA synthesis by the virus-encoded enzyme reverse transcriptase (RT). HIV-1 is characterized by its rapid genetic evolution. Genetic diversity in HIV-1 is well evidenced from the large number of different HIV-1 strains isolated around the world, which have been divided into three groups. The major group has been further divided into 10 nucleotide sequence-defined subtypes (24, 37, 39). Due to the rapid genetic changes, HIV-1 *in vivo* is defined as a quasispecies, that is, a population of highly related yet genetically distinct viruses within the same individual (11, 46). Sequential analysis of the quasispecies from infected patients shows substantial variation of genetic information over the course of infection (4, 8, 28, 45). This rapid genetic variation provides HIV-1 with maximum adaptation efficiency and poses serious challenges for chemotherapy and vaccine development for HIV-1 infection.

One mechanism that contributes to HIV-1 genetic variation is its high mutation rate. The RT of HIV-1 is error prone, in part due to its lack of proofreading activity. Mutation rate studies indicate that the HIV-1 genome averages 0.3 nucleotide change per cycle of virus replication (25; P. K. O'Neil, G. Sun, B. D. Preston, and J. P. Dougherty, unpublished data).

Another means to generate genetic changes is through recombination. Since two RNA molecules are packaged in each virion, RT may switch from one template to another during reverse transcription. If two RNAs with sequence differences are copackaged in one virion, a mosaic HIV-1 genome con-

taining genetic information from both RNAs could be generated, yielding novel viral genomes. Recombination can also salvage genome damage or detrimental mutations in one RNA molecule by adopting genetic information from the intact template (5, 38). Thus, recombination may not only introduce genetic diversity, it may also serve as a repair mechanism for the HIV-1 genome.

Accumulating evidence has confirmed the existence of recombinant HIV-1 in nature (33). Phylogenetic analyses have detected recombinant HIV-1 by exploiting the substantial nucleotide sequence differences between different subtypes (3, 12, 13, 27, 32, 36). Results indicate that approximately 10% of sequenced HIV-1 strains are mosaic genomes with genetic material from different subtypes (34, 35). In addition, recombination between two highly divergent groups of HIV-1 has recently been identified (40). Although the efficiency and the circumstances fostering the generation of recombination are largely unknown, the phylogenetic studies clearly indicate that recombination events contribute significantly to HIV-1 diversity. Further information regarding the rate, spectrum, and mechanism of HIV-1 recombination is needed for a thorough understanding of the HIV-1 recombination process.

A system employing two HIV-1-derived vectors confined to a single cycle of replication was used to examine the rate and mechanism of HIV-1 recombination throughout the entire genome using the heteroduplex tracking assay (HTA). The vectors were based on different strains of HIV-1 and were used to generate heterozygous virions. In order to closely reflect natural virus replication, these vectors were similar to the wild type in both size and sequence content. Sequence differences between the two strains of HIV-1 were exploited to monitor recombination events. The results from this study indicate that, on average, HIV-1 recombines approximately two to three times in every cycle of replication. This overall rate of recombination is similar to the rate previously observed during minus-strand synthesis (49), indicating that recombination occurs mainly during minus-strand DNA synthesis. It also suggests

* Corresponding author. Mailing address: UMDNJ, Robert Wood Johnson Medical School, Department of Molecular Genetics and Microbiology, 675 Hoes Ln., Piscataway, NJ 08854. Phone: (732) 235-4588. Fax: (732) 235-5223. E-mail: doughejp@umdnj.edu.

† Present address: Section of Immunobiology, Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, CT 06510.

that both virion RNAs are utilized during replication. Furthermore, a higher than average number of recombination events was observed in two regions of the genome, suggesting that sequence context can play a role in HIV-1 recombination.

MATERIALS AND METHODS

Plasmid construction. HIV-gpt_{HXB2}, pSG3.1, and pNL4-3 were obtained from the AIDS Research and Reference Reagent Program (1, 14, 29). The construction of HIV-puro_{BCSG3}, previously referred to as SG3puro_{BCSG3}, has been described elsewhere (49). To construct HIV-puro_{NL4-3}, a 1.2-kbp deletion was made in the envelope gene of pNL4-3 from the *Nde*I site at nucleotide 6400 to the *Bgl*II site at nucleotide 7611. The deletion was replaced with a 1.3-kbp *Eco*RI-to-*Nco*I fragment from pAJpac, which contains the spleen necrosis virus (SNV) U3 promoter and puromycin resistance gene coding sequence (map available upon request).

Cell culture. 293 cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum. HeLaT4 cells were grown in DMEM supplemented with 10% fetal bovine serum and 0.1 mg of hygromycin per ml. #69TRevEnv cells were grown in DMEM supplemented with 10% fetal bovine serum, 0.2 mg of G418 per ml, 0.1 mg of hygromycin per ml, and 2 µg of tetracycline per ml (50).

To create producer cell clones, HIV-gpt_{HXB2} and HIV-puro_{BCSG3} were separately cotransfected with an amphotropic murine leukemia virus *env* expression vector, pEnvAm (26), into 293 cells. Pseudotyped HIV-gpt_{HXB2} vector virus produced from 293 cells was then used to inoculate #69TRevEnv cells; this was followed by selection, isolation, and expansion. These cell clones were subsequently inoculated with pseudotyped HIV-puro_{BCSG3} vector virus produced from 293 cells; this was followed by selection, isolation, and expansion.

Transfection and infection. Transfections were performed by using a modified calcium phosphate precipitation method (15). To perform infections, 2×10^5 target cells were treated with 2 µg of Polybrene per ml for 30 min. Treated cells were subsequently inoculated with virus in 0.3 ml of medium and incubated for 2 h. At 24 h after infection, the cells were selected for resistance to either GPT (xanthine at 250 µg/ml, hypoxanthine at 15 µg/ml, mycophenolic acid at 7 µg/ml) or puromycin (1 µg/ml).

PCR and sequencing analysis of proviral DNA. Proviral DNA was amplified from genomic DNA of infected cell clones by PCR using *Taq* DNA polymerase (Promega, Madison, Wis.). The sequences of the primer pairs used to amplify different segments of proviral DNA generated from HIV-gpt_{HXB2}/HIV-puro_{BCSG3} producer cell clones and their nucleotide positions according to the HXB2 provirus sequence number (in parentheses) are as follows: segment 1, 5'-AGTCAGTGTGGAAAATCTCT-3' (611 to 632) and 5'-GCTGAAAGCCTTCTCTTACTA-3' (1261 to 1283); segment 2, 5'-TAGTAGAAGAGAAGGCTTCA-3' (1261 to 1283) and 5'-GTCCTCTTCCACATTTCC-3' (2028 to 2049); segment 3, 5'-TGGAAATGTGGAAAGGAAGAC-3' (2028 to 2049) and 5'-TCCCTGGTGTCTCATTGTTA-3' (2801 to 2820); segment 4, 5'-CAAGACTTCTGGGAAGTT-3' (2801 to 2820) and 5'-GCACTGCTCTGTGTTAATGT-3' (3646 to 3665); segment 5, 5'-GCAAGGCCAATGGACATATCAA-3' (3547 to 3568) and 5'-GTTGGCATATTCCTGGACTAC-3' (4394 to 4415); segment 6, 5'-GATGGAATAGATAAGGCCCAAGA-3' (4271 to 4290) and 5'-CATCTGTCTACTTGGCACACA-3' (5047 to 5066); segment 7, 5'-GGAAAGACCAGCAAAGCT-3' (4932 to 4940) and 5'-CTGCTATGTCGACACC CAAT-3' (5778 to 5797); segment 8, 5'-ATTGGGTGTCGACATAGCAG-3' (5778 to 5797) and 5'-GATGCACAAAATAGAGTGGTGGT-3' (6368 to 6390); segment 9, 5'-GAGGAGGAGATATGAGGGACAA-3' (7635 to 7656) and 5'-CTGTCTGTCTCTCTCTC CA-3' (8434 to 8454); segment 10, 5'-GTGAATAGAGTTAGGACGGAT-3' (8336 to 8357) and 5'-AGTAGCCTTGTGTGGTAG-3' (9137 to 9156). The above primers were also used to amplify proviral DNA generated from HIV-gpt_{HXB2}/HIV-puro_{NL4-3} producer cell clones, with the following exceptions: segment 3, 3' end, 5'-CCTAATTGAACTTCCCAGAA-3' (2807 to 2826); segment 4, 5'-TTCTGGGAAGTTCAATTAGG-3' (2807 to 2826) and 5'-ACTGCCTCTGTTAATTGTTT-3' (3644 to 3663); segment 5, 3' end, 5'-CTACAGTCTACTTGTCCATGCA-3' (4377 to 4398); segment 9, 5' end, 5'-GATATGAGGGACAATTGGAGAA-3' (7643 to 7664). To sequence the PCR products, the amplified DNA was first separated from its primers through a 1% agarose gel. The DNA was then purified from the gel by using a QIAquick gel extraction kit (QIAGEN Inc., Chatsworth, Calif.). Purified DNA was cloned by using the pGEM-T Easy Vector System (Promega). The cloned DNA was subjected to automated sequencing using T7 or SP6 primers or the primers used for the proviral DNA amplification.

HTA. The HTA is described elsewhere (51). Briefly, a ³²P-labeled single-stranded probe was prepared from HIV-puro_{BCSG3} plasmid DNA by using an asymmetrical PCR utilizing the primer pairs mentioned above. The ratio of primer concentrations used to amplify the single-stranded probe was 1:100 (0.2/20 mM), and the reaction was carried out in a mixture of 0.1 mM each deoxynucleoside triphosphate, 2 mM MgCl₂, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 0.02% Nonidet P-40, 2 U of *Taq* polymerase (Promega), and 10 µCi of [³²P]dCTP. A 2-µl aliquot of ³²P-labeled probe was mixed with 8 to 10 µl of symmetric PCR products amplified with the corresponding primers from genomic DNA of target cell clones and 1.5 µl of 10× annealing buffer (1 M NaCl,

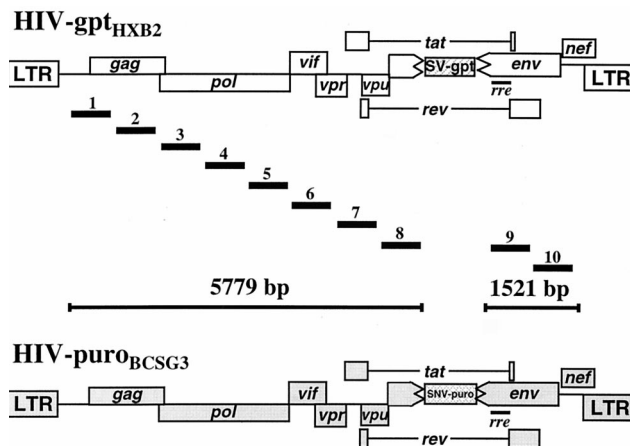


FIG. 1. HIV-1 vectors and segments of provirus DNA amplified to examine recombination rate. HIV-gpt_{HXB2} is based upon HIV-1 strain HXB2 (29). HIV-puro_{BCSG3} is based upon HIV-1 strain BCSG3 (14, 49). SV-gpt represents the SV40 early promoter driving expression of the *gpt* gene. SNV-puro represents the SNV U3 promoter controlling *puro* gene expression. Segments 1 to 10 are different segments of amplified provirus DNA. The coordinates of the amplified segments are as follows: 1, 611 to 1283; 2, 1261 to 2049; 3, 2028 to 2820; 4, 2801 to 3665; 5, 3547 to 4415; 6, 4271 to 5066; 7, 4932 to 5797; 8, 5778 to 6390; 9, 7635 to 8454; 10, 8336 to 9156 (sequence coordinates according to the HXB2 provirus sequence).

100 mM Tris-HCl [pH 7.8], 20 mM EDTA). In a PCR machine, the DNA mixtures were denatured at 98°C for 2 min, annealed by decreasing the temperature 1°C/10 s, incubated at 22°C for 10 min, and stored on ice. Homoduplex and heteroduplex DNAs were differentiated by electrophoresis in a nondenaturing 6% polyacrylamide gel, dried onto Whatman 3MM filter paper, and exposed to X-ray film.

RESULTS

Generation of heterozygous virions. Recombination events may be observed from progeny of heterozygous virions after a single cycle of virus replication. Two different HIV-1 vectors were used to generate heterozygous vector virions. HIV-gpt_{HXB2} and HIV-puro_{BCSG3} are *env*-defective HIV-1 vectors (Fig. 1) and are based on the HXB2 and BCSG3 strains of HIV-1, respectively (29, 49). The sequence difference between the HXB2 and BCSG3 strains of HIV-1 is approximately 5%.

The protocol employed to measure HIV-1 recombination is outlined in Fig. 2. The approach is predicated upon confining replication to a single cycle because the provirus in the target cell provides a "fossil record" allowing one to deduce conclusions about dynamic events that occurred during reverse transcription. Eight independent producer cell clones containing single copies of HIV-gpt_{HXB2} and HIV-puro_{BCSG3} vector proviruses were established as previously described (49). Upon removal of tetracycline, producer cell clones were induced to express the HIV-1 Env protein and infectious vector virus was produced. Supernatant containing vector virus was used to inoculate CD4-positive HeLaT4 cells. Since the producer cells do not express CD4, vector virus cannot spread among them. Since the CD4-positive HeLaT4 cells do not express the HIV-1 Env protein, the vector virus cannot further propagate after infection. Therefore, from the producer cell to the target cell, the vector virus is restricted to a single cycle of replication.

Only progeny derived from heterozygous virions are informative for the study of recombination, since crossovers between homozygous genomes cannot be scored. The puromycin-resistant titers were typically 100-fold lower than the GPT-resistant titers, so that the proviruses in puromycin-resistant

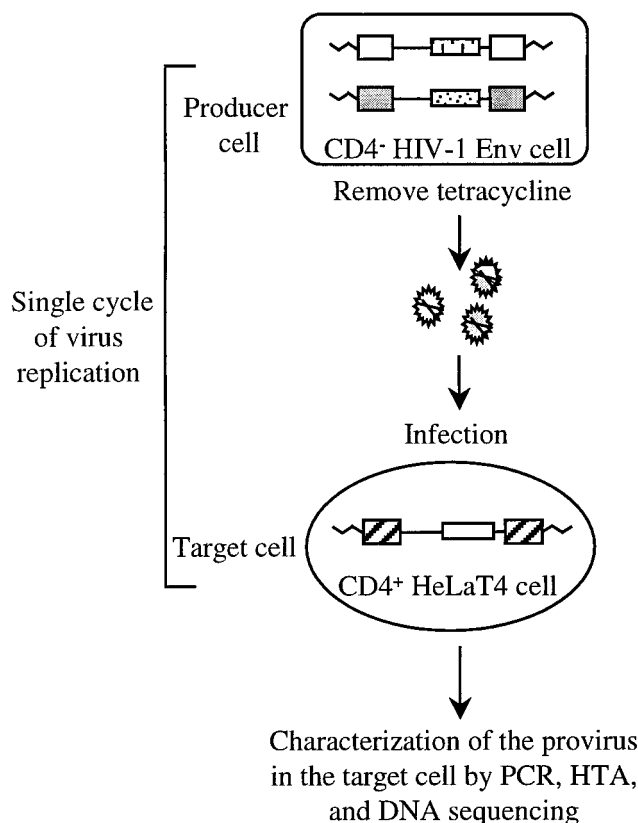


FIG. 2. Protocol for the study of HIV-1 recombination. Individual cell clones containing a single copy of both HIV-gpt_{HXB2} and HIV-puro_{BCSG3} were established (49). Upon induction, vector virus was harvested and used to inoculate CD4-positive HeLaT4 target cells; this was followed by selection and isolation of target cell clones. Going from a provirus in a producer cell to a provirus in a target cell comprised a single cycle of virus replication (see text). Provirus in each target cell clone were further analyzed by PCR, the HTA, and DNA sequencing in order to identify recombinants and crossover points.

target cells are likely to be progeny of heterozygous virions, in that the difference in titer is likely to reflect a difference in the amount of RNA transcribed and packaged. This prediction was subsequently borne out, since all of the proviral clones analyzed were shown to be recombinants (see below). From 8 independent producer cell clones, 86 puromycin-resistant target cell clones were established and analyzed for recombination. These 86 clones were sensitive to GPT selection, thus ruling out the possibility of double infection.

Detection of recombinants by HTA. HTA was used to screen the progeny proviruses for sequence changes caused by recombination (51). Based on the mobility differences between homoduplex and heteroduplex DNAs upon electrophoresis, HTA can differentiate sequence heterogeneity. In HTA, a ³²P-labeled single-stranded probe is amplified first from wild-type plasmid DNA. This probe is then annealed to both wild-type and mutant DNAs. Upon annealing, a homoduplex forms between the probe and the wild-type DNA whereas sequence differences between the probe and the mutant DNA cause heteroduplex formation. These DNA complexes, when subjected to electrophoresis on a nondenaturing polyacrylamide gel, exhibit different mobilities. Wild-type and mutant DNAs can therefore be differentiated. The HTA is capable of detecting sequence divergence when the degree of mismatch exceeds approximately 1% (9, 51). Since the sequence heterogeneity

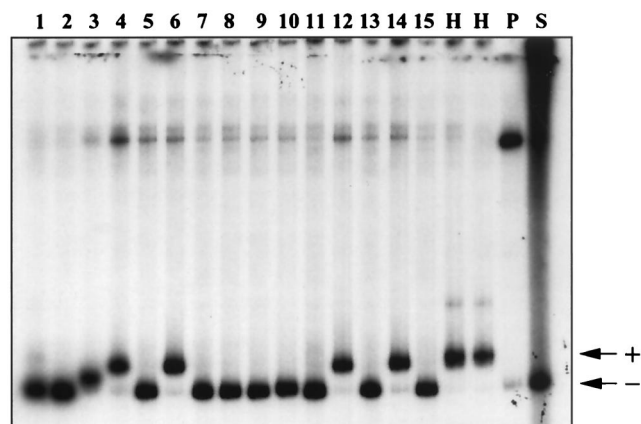


FIG. 3. Representative example of an HTA. A radioactively labeled single-stranded probe was prepared from HIV-puro_{BCSG3} plasmid DNA by asymmetric PCR using the segment 9 primer pair (lane P). Segment 9 is 819 bp long. The sequence difference between the two strains within this segment is 4.4%. The probe was annealed with PCR products amplified from HIV-puro_{BCSG3} plasmid DNA (lane S, homoduplex control), HIV-gpt_{HXB2} plasmid DNA (lanes H, heteroduplex controls), and genomic DNAs from different target cell clones (lanes 1 to 15). The plus sign indicates a band shifted due to heteroduplex formation. The minus sign indicates a nonshifted homoduplex band. Lanes 3, 7, 9, 16, and 18 are shifted compared to the band for HIV-puro_{BCSG3} DNA and were therefore scored as recombinants.

between the HXB2 and BCSG3 strains of HIV-1 is approximately 5% throughout the viral genome, the HTA can be utilized to detect crossovers between HIV-gpt_{HXB2} and HIV-puro_{BCSG3} that introduce >1% mismatch.

To carry out HTA screening, primers common to both HIV-gpt_{HXB2} and HIV-puro_{BCSG3} were designed to amplify segments of provirus DNA. Figure 1 shows the distribution of primers over the provirus genomes. The size of each amplified DNA fragment ranged from 612 to 865 bp, and the sequence differences between HIV-gpt_{HXB2} and HIV-puro_{BCSG3} in each DNA segment ranged from 3.2 to 6.2%. Figure 3 shows a typical HTA gel. A ³²P-labeled single-stranded DNA probe corresponding to HIV-puro_{BCSG3} nucleotides 7168 to 7987 (according to the BCSG3 RNA sequence) (14) was prepared by asymmetric PCR (Fig. 3, lane P). The same primer pair was also used in a standard PCR to amplify provirus DNA from different progeny cell clones and from both HIV-gpt_{HXB2} and HIV-puro_{BCSG3} plasmids. After the probe had been annealed to the amplified DNA, homoduplex and heteroduplex DNAs were distinguished through gel electrophoresis. The probe formed a homoduplex when annealed to HIV-puro_{BCSG3} plasmid DNA (Fig. 3, lane S), whereas it formed a heteroduplex when annealed to HIV-gpt_{HXB2} plasmid DNA (Fig. 3, lane H). DNAs from progeny cell clones 1, 2, 5, 7 to 11, 13, and 15 formed homoduplexes when annealed to the probe, indicating no sequence change or less than 1% sequence mismatch. On the other hand, DNAs from clones 3, 4, 6, 12, and 14 formed heteroduplexes when annealed to the probe, indicating more than 1% mismatch and divergence from the original HIV-puro_{BCSG3} sequence. By the same approach, sequence changes were identified in the other segments of proviruses of the 86 progeny cell clones.

To confirm that the heteroduplex shifts observed in the HTA analysis were produced by recombination and not mutation, several different fragments from several different clones were subjected to sequencing analysis. A total of 47 DNA fragments were sequenced. Fifteen of the 21 DNA fragments that had caused heteroduplex formation in HTA displayed an

TABLE 1. Rates of HIV-1 recombination

Producer cell clone	No. of progeny cell clones analyzed	No. of crossovers	Avg no. of crossovers
1	26	68	2.6
2	18	60	3.3
3	13	46	3.5
4	4	17	4.3
5	6	17	2.8
6	8	23	2.9
7	3	4	1.3
8	8	25	3.1
Total	86	260	3.0

HIV-gpt_{HXB2} sequence pattern, while the remaining 6 DNA fragments had part of the sequence from HIV-puro_{BCSG3} and part from HIV-gpt_{HXB2} (data not shown), allowing the crossover regions to be identified in them. Furthermore, 26 of the DNA fragments that had formed homoduplexes in the HTA showed an HIV-puro_{BCSG3} pattern (data not shown). Although it could not be ruled out that some gel shifts were caused by insertions, deletions, or mutations, sequencing analysis clearly indicated that the majority of DNA fragments that caused a gel shift had part or all of their sequences changed from HIV-puro_{BCSG3} to HIV-gpt_{HXB2} as a result of recombination.

Recombination occurs at a high rate during virus replication. To assess recombination over the entire provirus genome, sequence information obtained by HTA for the 10 provirus segments (Fig. 1) and by restriction enzyme analysis for the 5' and 3' long terminal repeats (LTRs) (49) were combined. The results for 86 provirus clones are summarized in Table 1. The average number of recombination events that occurred was three crossovers per genome per replication cycle. Although the majority of progeny proviruses experienced two to three recombination events, the number of crossovers per progeny provirus ranged from one to seven (Table 2).

Sequence context may play a role in HIV-1 recombination. To evaluate the effect of sequence context upon the generation of HIV-1 recombination, the frequency of recombination events that occurred in different DNA segments was compared and plotted in Fig. 4. Recombinants can be identified when two consecutive DNA segments show different sequence origins, as ascertained by HTA. The recombination event might have occurred in the DNA segment that caused heteroduplex formation in HTA and caused more than a 1% sequence difference from the original HIV-puro_{BCSG3} sequence. Alternatively, the crossover point might have occurred in the adjacent DNA segment that displayed homoduplex formation in the HTA and introduced less than a 1% sequence difference from the original HIV-puro_{BCSG3} sequence.

Although recombinations were observed across the entire HIV-1 provirus genome, a high crossover rate was observed in two regions; one region is between segment 8 and the SNV-puro expression cassette, and the other is in the HIV-1 *pol* region. Eighty of 86 clones recombined between segment 8 and the exogenous SNV-puro sequences. DNA sequencing of HIV-

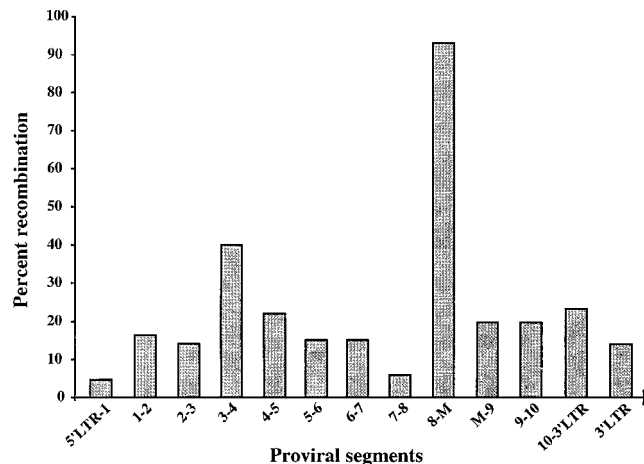


FIG. 4. Comparison of recombination events across the entire HIV-1 genome. The values are percentages of progeny proviruses in which a crossover occurred between the two specified segments. M corresponds to the puromycin resistance gene.

puro_{BCSG3} revealed a 111-bp region between the SNV U3 promoter and the puromycin resistance marker gene (*puro*) that is homologous to the 3' end of the simian virus 40 (SV40) early promoter in HIV-gpt_{HXB2}. This region includes the AT-rich region that contains the TATA box-like element. Sequencing analysis of six recombinant clones revealed that all had crossed over in that region and used the SV40 early promoter from the SV-gpt expression cassette to drive *puro* expression. Thus, there was a high rate of recombination between sequences of nonretroviral origin in this region. The other region of the genome that seemed to be a hot spot for crossovers was at the 5' end of the *pol* region. Thirty-four of 86 clones recombined in this region. Twenty-one of these clones were sequenced such that the crossover regions could be visualized. The results showed that these 21 clones had not crossed over at exactly the same spot but that the crossover regions spanned both segments 3 and 4.

To determine whether the observed hot spot in nonretroviral sequences significantly affects the overall recombination rate, the sequence containing the hot spot was deleted and the analysis was repeated. For this purpose, HIV-puro_{NL4-3} was constructed. It is also *env* defective with the deleted sequences replaced with an SNV-puro expression cassette which does not contain the nonretroviral crossover hot spot. Moreover, HIV-puro_{NL4-3} was derived from HIV-1 strain NL4-3, allowing measurement of recombination rates with a different viral strain. HIV-puro_{NL4-3} was used in conjunction with HIV-gpt_{HXB2}. The sequence difference between the HXB2 and NL4-3 strains of HIV-1 is approximately 3%, providing sufficient heterogeneity for HTA analysis.

Following the protocol described earlier and previously outlined (49), producer cell clones containing a single copy of both HIV-gpt_{HXB2} and HIV-puro_{NL4-3} were isolated. Since the puromycin-resistant titers were again significantly lower than the GPT-resistant titers after one cycle of replication, puromycin-resistant target cell clones were isolated and analyzed by HTA (data not shown). The 1,473-bp viral sequence immediately upstream of the marker cassette for each of 12 progeny proviral clones was analyzed. The results indicated that five crossovers had occurred, yielding a recombination rate of 2.8×10^{-4} /bp per replication cycle. This rate corresponds to approximately 2.8 crossovers per genome per replication cycle.

TABLE 2. Recombination events across the HIV-1 genome

No. of crossovers/clone	1	2	3	4	5	6	7
No. of progeny clones corresponding to no. of crossovers	12	22	26	11	10	4	1

DISCUSSION

In this report, we describe a system to examine the rate of HIV-1 recombination across the entire HIV-1 genome after a single cycle of viral replication. The vectors utilized were similar in size and sequence content to the authentic HIV-1 genome. The recombination rate obtained was approximately two to three crossovers per genome per replication cycle. Crossovers were identified throughout the viral genome.

A single-cycle approach has been used previously to study retroviral recombination in the avian oncoretrovirus SNV. The recombination rate obtained for HIV-1, a primate lentivirus, is approximately 5- to 15-fold higher than that obtained for SNV (16, 17, 21). Although experimental differences between these studies may be a factor and accessory proteins may influence recombination, we believe that the likeliest reason for the discrepancy is inherent differences between the HIV-1 and SNV RTs.

It is also noteworthy that a positive correlation was previously noted between intermolecular minus-strand primer transfers and recombination for SNV (21). However, this does not seem to be the case for HIV-1. The nature of HIV-1 minus-strand primer transfers, that is, whether they are inter- and/or intramolecular, for the same 86 proviral clones was analyzed in a previous study (49). Of the 260 recombination events, 143 (55%) occurred after intramolecular minus-strand transfer while 117 (45%) occurred after intermolecular minus-strand transfer. Thus, there does not appear to be a strong correlation between the nature of the minus-strand primer transfer and the rate of recombination. We believe that this again reflects an intrinsic difference between the HIV-1 and SNV RTs.

Two regions exhibited higher-than-average recombination rates. A particularly strong hot spot was the nonretroviral SV40 sequence that was part of the expression cassettes. This hot spot contains an 89-base palindrome that can theoretically form a stable hairpin structure, as predicted by the Genetics Computer Group program FoldRNA (10, 52). This might induce RT to pause, promoting recombination at the palindrome. Hairpin structures have been shown to promote strand transfer in a cell-free system (23). The 5' end of the *pol* region also seemed to be a relatively strong hot spot for recombination. Nearly 40% of the proviral clones had crossed over in this region, which spans approximately 1 kbp of the viral sequence. Further study is needed to ascertain whether RNA secondary structure affected recombination in this region.

The results show that the presence of a nonviral hot spot did not affect the overall recombination rate in other parts of the genome. The calculated recombination rate, including the nonviral hot spot, is 3×10^{-4} /bp per replication cycle, whereas omission of the hot spot crossover events from the calculation produces a rate of 2.4×10^{-4} bp per cycle. Deletion of this hot spot sequence from the vector significantly reduced the recombination rate in that region, suggesting that sequence context and RNA secondary structure can affect recombination, which is not surprising in light of previous reports suggesting that they can affect the frequency of mutation by RT (21, 30, 31). However, although the frequency of recombination decreased in this segment, recombination events did occur at a rate of 2.8 crossovers per genome per cycle. Thus, even in the absence of this hot spot, the recombination rate remained extremely high. In addition, the six clones that had not recombined at the hot spot retained the SNV U3 promoter and experienced crossover events at a rate of 2.3 per genome per replication cycle, providing further evidence that recombination at the nonviral hot spot is not a prerequisite for a high rate of recombination.

Two models of retroviral recombination have been proposed (41). The strand displacement assimilation model proposes that DNA fragments are displaced during plus-strand DNA synthesis from one template and subsequently assimilated by the plus-strand DNA synthesized from the other template, causing recombination during plus-strand synthesis (20, 22). The copy choice model proposes that recombination occurs during minus-strand DNA synthesis. The original "forced copy choice" model hypothesizes that RT switches templates when it encounters RNA breaks, promoting recombination during minus-strand DNA synthesis (6, 44). The copy choice model has recently been broadened to include recombination that occurs during minus-strand synthesis but without the requirement of breaks in viral RNA. This model suggests that the low processivity of RT causes the enzyme to dissociate from the template, allowing the short DNA-RNA hybrid to be disrupted so that the growing DNA strand is displaced to the other RNA template (7, 48).

The results presented here and previously provide evidence that HIV-1 recombination occurs mainly during minus-strand DNA synthesis, supporting a copy choice model as the predominant mechanism of recombination. We heretofore reported that the rate of recombination in U3 of the viral LTR was 3×10^{-4} /bp per replication cycle, which when extrapolated to the entire genome, indicates a rate of approximately three crossovers per genome per replication cycle. Because the rate was obtained for the viral LTRs, it was possible to determine whether the crossovers occurred during minus-strand or plus-strand synthesis. The results indicated that recombination occurred primarily during minus-strand DNA synthesis (6:1 minus-strand-to-plus-strand ratio) (49). The fact that the rate obtained during minus-strand synthesis in U3 is similar to that obtained for the entire genome provides support for the theory that the majority of crossovers throughout the genome occur during minus-strand DNA synthesis. Previous studies using SNV-based vectors have also indicated that retroviral recombination occurs during minus-strand DNA synthesis (2, 18). Moreover, given that the rate of recombination during minus-strand synthesis is high at least early during reverse transcription, the frequency of recombination events should be further skewed toward minus-strand DNA synthesis, particularly late during reverse transcription. If a crossover occurs during minus-strand synthesis, then RNase H activity degrades the second RNA template, preventing further synthesis of a second minus-strand DNA, obviating further recombination during plus-strand synthesis because of the lack of a second template. Furthermore, given the high rate of recombination, the forced copy choice mechanism would require that the majority of virion RNAs have multiple breaks. This is possible, but given the low degree of processivity exhibited by HIV-1 RT (19), it seems more plausible that crossovers can occur without RT encountering strand breaks. Taken together, these results suggest that HIV-1 recombination occurs primarily during minus-strand synthesis via a simple copy choice mechanism.

The high rate of recombination, taken together with the previous finding that 50% of all minus-strand primer transfers are intermolecular (43, 47, 49), implies that both HIV-1 RNAs are typically utilized during reverse transcription. Nevertheless, this does not exclude the possibility that a single virion RNA can act as the sole template for reverse transcription, as was previously reported for SNV (21). The high capacity of HIV-1 RT to generate recombinant progeny proviruses seems to support the contention that the reason retroviruses are diploid is to provide a recombination partner (42), implying that there are strong selective pressures promoting recombination. This indicates that recombination is an integral aspect

of HIV-1 replication and is likely to play an important role in the generation of viral diversity.

ACKNOWLEDGMENTS

A.E.J. and H.Y. contributed equally to this study.

This work was supported by National Institutes of Health grants CA50777 and AI34834. A.E.J. is supported by National Institutes of Health Training Grant 5T32 AI07403.

We thank Malvika Kaul, Annmarie Pacchia, and Amariliz Rivera for constructive comments on the manuscript.

REFERENCES

- Adachi, A., H. E. Gendelman, S. Koenig, T. Folks, R. Willey, A. Rabson, and M. A. Martin. 1986. Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. *J. Virol.* **59**:284–291.
- Anderson, J. A., R. J. Teufel II, P. D. Yin, and W.-S. Hu. 1998. Correlated template-switching events during minus-strand DNA synthesis: a mechanism for high negative interference during retroviral recombination. *J. Virol.* **72**:1186–1194.
- Blackard, J. T., B. R. Renjifo, D. Mwakagile, M. A. Montano, W. W. Fawzi, and M. Essex. 1999. Transmission of human immunodeficiency type 1 viruses with intersubtype recombinant long terminal repeat sequences. *Virology* **254**:220–225.
- Cichutek, K., H. Merget, S. Norley, R. Linde, W. Kreuz, M. Gabr, and R. Kurth. 1992. Development of a quasispecies of human immunodeficiency virus type 1 in vivo. *Proc. Natl. Acad. Sci. USA* **89**:7365–7369.
- Clavel, F., M. D. Hoggan, R. L. Willey, K. Strebel, M. A. Martin, and R. Repaske. 1989. Genetic recombination of human immunodeficiency virus. *J. Virol.* **63**:1455–1459.
- Coffin, J. M. 1979. Structure, replication, and recombination of retrovirus genomes: some unifying hypotheses. *J. Gen. Virol.* **42**:1–26.
- Coffin, J. M. 1996. *Retroviridae: the viruses and their replication*, p. 1767–1847. In B. N. Fields, D. M. Knipe, and P. M. Howley (ed.), *Virology*. Lippincott-Raven Publishers, Philadelphia, Pa.
- Delwart, E. L., H. Pan, H. W. Sheppard, D. Wolpert, A. U. Neumann, B. Korber, and J. I. Mullins. 1997. Slower evolution of human immunodeficiency virus type 1 quasispecies during progression to AIDS. *J. Virol.* **71**:7498–7508.
- Delwart, E. L., E. G. Shpaer, J. Louwagie, F. E. McCutchan, M. Grez, H. Rubsamen-Waigmann, and J. I. Mullins. 1993. Genetic relationships determined by a DNA heteroduplex mobility assay: analysis of HIV-1 env genes. *Science* **262**:1257–1261.
- Devereux, J., P. Haeblerli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**:387–395.
- Eigen, M. 1993. Viral quasispecies. *Sci. Am.* **269**:42–49.
- Gao, F., D. L. Robertson, C. D. Carruthers, Y. Li, E. Bailes, L. G. Kostrikis, M. O. Salminen, F. Bibollet-Ruche, M. Peeters, D. D. Ho, G. M. Shaw, P. M. Sharp, and B. H. Hahn. 1998. An isolate of human immunodeficiency virus type 1 originally classified as subtype I represents a complex mosaic comprising three different group M subtypes (A, G, and I). *J. Virol.* **72**:10234–10241.
- Gao, F., D. L. Robertson, S. G. Morrison, H. Hui, S. Craig, J. Decker, P. N. Fultz, M. Girard, G. M. Shaw, B. H. Hahn, and P. M. Sharp. 1996. The heterosexual human immunodeficiency virus type 1 epidemic in Thailand is caused by an intersubtype (A/E) recombinant of African origin. *J. Virol.* **70**:7013–7029.
- Ghosh, S. K., P. N. Fultz, E. Keddie, M. S. Saag, P. M. Sharp, B. H. Hahn, and G. M. Shaw. 1993. A molecular clone of HIV-1 tropic and cytopathic for human and chimpanzee lymphocytes. *Virology* **194**:858–864.
- Gorman, C. 1985. High efficiency gene transfer into mammalian cells, p. 143–190. In D. M. Glover (ed.), *DNA cloning*. IRL Press, Oxford, England.
- Hu, W. S., and H. M. Temin. 1990. Retroviral recombination and reverse transcription. *Science* **250**:1227–1233.
- Hu, W. S., and H. M. Temin. 1990. Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination. *Proc. Natl. Acad. Sci. USA* **87**:1556–1560.
- Hu, W. S., and H. M. Temin. 1992. Effect of gamma radiation on retroviral recombination. *J. Virol.* **66**:4457–4463.
- Huber, H. E., J. M. McCoy, J. S. Seehra, and C. C. Richardson. 1989. Human immunodeficiency virus 1 reverse transcriptase. Template binding, processivity, strand displacement synthesis, and template switching. *J. Biol. Chem.* **264**:4669–4678.
- Hunter, E. 1978. The mechanism for genetic recombination in the avian retroviruses. *Curr. Top. Microbiol. Immunol.* **79**:295–309.
- 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.
- Junghans, R. P., L. R. Boone, and A. M. Skalka. 1982. Retroviral DNA H structures: displacement-assimilation model of recombination. *Cell* **30**:53–62.
- Kim, J. K., C. Palaniappan, W. Wu, P. J. Fay, and R. A. Bambara. 1997. Evidence for a unique mechanism of strand transfer from the transactivation response region of HIV-1. *J. Biol. Chem.* **272**:16769–16777.
- Leitner, T., B. Korber, D. Robertson, F. Gao, and B. Hahn. 1997. Updated proposal of reference sequences of HIV-1 genetic subtypes, p. III-19–III-24. In B. Korber, B. Hahn, B. Foley, J. W. Mellors, L. Leitner, G. Myers, F. McCutchan, and C. Kuiken (ed.), *Human retroviruses and AIDS: a compilation and analysis of nucleic acid and amino acid sequences*. Los Alamos National Laboratory, Los Alamos, N.Mex.
- 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.
- Markowitz, D., S. Goff, and A. Bank. 1988. Construction and use of a safe and efficient amphotropic packaging cell line. *Virology* **167**:400–406.
- McCutchan, F. E., J. K. Carr, M. Bajani, E. Sanders-Buell, T. O. Harry, T. C. Stoekli, K. E. Robbins, W. Gashau, A. Nasidi, W. Janssens, and M. L. Kalish. 1999. Subtype G and multiple forms of A/G intersubtype recombinant human immunodeficiency virus type 1 in Nigeria. *Virology* **254**:226–234.
- McDonald, R. A., D. L. Mayers, R. C. Chung, K. F. Wagner, S. Ratto-Kim, D. L. Bix, and N. L. Michael. 1997. Evolution of human immunodeficiency virus type 1 env sequence variation in patients with diverse rates of disease progression and T-cell function. *J. Virol.* **71**:1871–1879.
- Page, K. A., N. R. Landau, and D. R. Littman. 1990. Construction and use of a human immunodeficiency virus vector for analysis of virus infectivity. *J. Virol.* **64**:5270–5276.
- Parthasarathi, S., A. Varela-Echavarría, Y. Ron, B. D. Preston, and J. P. Dougherty. 1995. Genetic rearrangements occurring during a single cycle of murine leukemia virus vector replication: characterization and implications. *J. Virol.* **69**:7991–8000.
- Pathak, V. K., and H. M. Temin. 1992. 5-Azacytidine and RNA secondary structure increase the retrovirus mutation rate. *J. Virol.* **66**:3093–3100.
- Renjifo, B., B. Chaplin, D. Mwakagile, P. Shah, F. Vannberg, G. Msamanga, D. Hunter, W. Fawzi, and M. Essex. 1998. Epidemic expansion of HIV type 1 subtype C and recombinant genotypes in Tanzania. *AIDS Res. Hum. Retroviruses* **14**:635–638.
- Robertson, D. L., F. Gao, B. H. Hahn, and P. M. Sharp. 1997. Intersubtype recombinant HIV-1 sequences, p. III-25–III-30. In B. Korber, B. Hahn, B. Foley, J. W. Mellors, L. Leitner, G. Myers, F. McCutchan, and C. Kuiken (ed.), *Human retroviruses and AIDS: a compilation and analysis of nucleic acid and amino acid sequences*. Los Alamos National Laboratory, Los Alamos, N.Mex.
- Robertson, D. L., B. H. Hahn, and P. M. Sharp. 1995. Recombination in AIDS viruses. *J. Mol. Evol.* **40**:249–259.
- Robertson, D. L., P. M. Sharp, F. E. McCutchan, and B. H. Hahn. 1995. Recombination in HIV-1. *Nature* **374**:124–126.
- Salminen, M. O., J. K. Carr, D. L. Robertson, P. Hegerich, D. Gotte, C. Koch, E. Sanders-Buell, F. Gao, P. M. Sharp, B. H. Hahn, D. S. Burke, and F. E. McCutchan. 1997. Evolution and probable transmission of intersubtype recombinant human immunodeficiency virus type 1 in a Zambian couple. *J. Virol.* **71**:2647–2655.
- Simon, F., P. Maulele, P. Roques, I. Loussert-Ajaka, M. C. Muller-Trutwin, S. Saragosti, M. C. Georges-Courbot, F. Barre-Sinoussi, and F. Brun-Vezinet. 1998. Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. *Nat. Med.* **4**:1032–1037.
- Stahl, F. W. 1987. Genetic recombination. *Sci. Am.* **256**:90–101.
- Subbarao, S., and G. Schochetman. 1996. Genetic variability of HIV-1. *AIDS* **10**(Suppl. A):S13–S23.
- 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.
- Telesnitsky, A., and S. P. Goff. 1997. Reverse transcriptase and the generation of retroviral diversity, p. 121–160. In J. M. Coffin, S. H. Hughes, and H. E. Varmus (ed.), *Retroviruses*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Temin, H. M. 1993. Retrovirus variation and reverse transcription: abnormal strand transfers result in retrovirus genetic variation. *Proc. Natl. Acad. Sci. USA* **90**:6900–6903.
- van Wamel, J. L., and B. Berkhout. 1998. The first strand transfer during HIV-1 reverse transcription can occur either intramolecularly or intermolecularly. *Virology* **244**:245–251.
- Vogt, P. K. 1973. The genome of avian RNA tumor viruses: a discussion of four models, p. 35–41. In L. Sylvestri (ed.), *Possible episomes in eukaryotes*. North-Holland Publishing Co., Amsterdam, The Netherlands.
- Wahlberg, J., J. Albert, J. Lundberg, S. Cox, B. Wahren, and M. Uhlen. 1992. Dynamic changes in HIV-1 quasispecies from azidothymidine (AZT)-treated patients. *FASEB J.* **6**:2843–2847.
- Wain-Hobson, S. 1992. Human immunodeficiency virus type 1 quasispecies in vivo and ex vivo. *Curr. Top. Microbiol. Immunol.* **176**:181–193.
- Wilhelm, M., M. Boutabout, T. Heyman, and F. X. Wilhelm. 1999. Reverse transcription of the yeast Ty1 retrotransposon: the mode of first strand transfer is either intermolecular or intramolecular. *J. Mol. Biol.* **288**:505–510.
- Xu, H., and J. D. Boeke. 1987. High-frequency deletion between homologous

- sequences during retrotransposition of Ty elements in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **84**:8553–8557.
49. **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.
 50. **Yu, H., A. B. Rabson, M. Kaul, Y. Ron, and J. P. Dougherty.** 1996. Inducible human immunodeficiency virus type 1 packaging cell lines. *J. Virol.* **70**:4530–4537.
 51. **Zhu, T., N. Wang, A. Carr, D. S. Nam, R. Moor-Jankowski, D. A. Cooper, and D. D. Ho.** 1996. Genetic characterization of human immunodeficiency virus type 1 in blood and genital secretions: evidence for viral compartmentalization and selection during sexual transmission. *J. Virol.* **70**:3098–3107.
 52. **Zuker, M., and P. Stiegler.** 1981. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Res.* **9**:133–148.