Genetic Recombination of Human Immunodeficiency Virus Type 1 in One Round of Viral Replication: Effects of Genetic Distance, Target Cells, Accessory Genes, and Lack of High Negative Interference in Crossover Events†

Terence D. Rhodes,[1,2] Olga Nikolaitchik,[1] Jianbo Chen,[1] Douglas Powell,[3] and Wei-Shau Hu[1]*

HIV Drug Resistance Program[1] and Data Management Services, Inc.,[3] National Cancer Institute at Frederick, Frederick, Maryland, and Department of Microbiology, Immunology, and Cell Biology, West Virginia University, Morgantown, West Virginia[2]

Recombination is a major mechanism that generates variation in populations of human immunodeficiency virus type 1 (HIV-1). Mutations that confer replication advantages, such as drug resistance, often cluster within regions of the HIV-1 genome. To explore how efficiently HIV-1 can assort markers separated by short distances, we developed a flow cytometry-based system to study recombination. Two HIV-1-based vectors were generated, one encoding the mouse heat-stable antigen gene and green fluorescent protein gene (*GFP*), and the other encoding the mouse *Thy-1* gene and *GFP*. We generated derivatives of both vectors that contained nonfunctional *GFP* inactivated by different mutations. Recombination in the region between the two inactivating mutations during reverse transcription could yield a functional *GFP*. With this system, we determined that the recombination rates of markers separated by 588, 300, 288, and 103 bp in one round of viral replication are 56, 38, 31, and 12%, respectively, of the theoretical maximum measurable recombination rate. Statistical analyses revealed that at these intervals, recombination rates and marker distances have a near-linear relationship that is part of an overall quadratic fit. Additionally, we examined the segregation of three markers within 600 bp and concluded that HIV-1 crossover events do not exhibit high negative interference. We also examined the effects of target cells and viral accessory proteins on recombination rate. Similar recombination rates were observed when human primary CD4+ T cells and a human T-cell line were used as target cells. We also found equivalent recombination rates in the presence and absence of accessory genes *vif*, *vpr*, *vpu*, and *nef*. These results illustrate the power of recombination in generating viral population variation and predict the rapid assortment of mutations in the HIV-1 genome in infected individuals.

Genetic recombination plays an important role in the evolution of human immunodeficiency virus type 1 (HIV-1) (35). Recombination shuffles viral genomes and redistributes the mutations generated from reverse transcription, leading to increased variation within the infected host and, ultimately, the viral populations distributed throughout the world (35). Of the more than 70,000 sequences that are catalogued in the HIV sequence database at the Los Alamos National Laboratory (http://www.hiv.lanl.gov), approximately 8% (>5,500) are classified as recombinants. Furthermore, recombinant strains of HIV-1 have been observed worldwide (13, 16, 26, 28, 38, 46). The inherent ability of HIV-1 to recombine poses a constant problem for effective anti-HIV-1 treatment because multidrug-resistant variants can be generated by recombining the genome of singly or weakly resistant viruses. The increased variation caused by recombination also hinders the development of effective vaccines; the induced host immune response has to combat not only the variants generated by the rapidly changing virus genome but also all the different subtypes of HIV-1.

Therefore, rapid recombination of the HIV-1 genome creates a vast advantage for the evolution of the virus and an enormous difficulty for the host.

The ability of HIV-1 to carry out frequent recombination is the result of a unique feature of the retrovirus family: retroviruses package two copies of viral RNA into each virion (15, 23). During reverse transcription, reverse transcriptase can use portions of the genomes from each RNA as templates to generate a recombinant viral DNA (7, 19). Although recombination can occur in all virions, a genetically different progeny can only be generated from virions with two different RNAs (heterozygous virions), not from virions with two identical RNAs (homozygous virions) (18). Heterozygous virions are only generated from cells infected with more than one retrovirus (double infection) (18). We have previously demonstrated that double infection occurs frequently in HIV-1 infection of cultured T cells and primary T cells (8), providing the basis for the generation of heterozygous virions that allows the observed frequent recombination.

Frequent HIV-1 recombination has been evident from studies with different approaches. Studies of individuals infected with more than one genetically distinct HIV-1 revealed that these patients often also harbor hybrid viruses in their viral populations (13, 16, 26, 28, 38, 46). Intersubtype recombinants that cause epidemics in certain geographical areas have multiple break-off points along the genome (35). Additionally,

some circulating forms of recombinants have mosaic genomes that were derived from five or more genetically distinct HIV-1 variants, emphasizing the common occurrence of double infection and frequent recombination (43).

In addition to the viral populations in infected individuals, recombination events have also been observed in experimental systems. It has been demonstrated that recombination can occur with purified reverse transcriptase and nucleic acid templates (5, 10–12, 30); these studies have also yielded interesting data on different aspects of recombination, such as the effects of the viral nucleocapsid protein (9, 30, 33, 36, 37) and the templates (5, 11, 12, 33, 36, 37). With cell culture-based systems, it has been demonstrated that HIV-1 recombination is frequent and occurs throughout the viral genome with potential localized hot spots (6, 20, 22, 24, 25, 47). Comparisons between HIV-1 and the simple retrovirus murine leukemia virus revealed that HIV-1 recombines more frequently than murine leukemia virus (31, 34).

In a previous study, we measured HIV-1 recombination rates for three genetic distances (1.9, 1.3, and 1.0 kb) (34). We observed that HIV-1 recombines at exceedingly high rates, because sequences separated by 1.3 kb segregated as unlinked alleles in a single round of viral replication. These results also led to questions that could not be easily addressed with the same system, such as the relationship between recombination rate and marker distances shorter than 1.0 kb. Many of the mutations that cause drug resistance or immune evasion are clustered within 1.0 kb of the viral genome (21, 41). Therefore, it is important to measure how frequently these mutations can be assorted. Furthermore, because drug selection was used in the previous system to measure recombination, we could not easily adapt this strategy to measure recombination in cells that grow in suspension, such as T cells, which are natural target cells for HIV-1.

In this report, we describe the development of a new experimental system that can measure recombination rates when markers are separated by short distances (0.6 kb or less). Furthermore, since this system is based on the detection of green fluorescent protein (GFP) and cell surface proteins that can be labeled with fluorescent antibodies, we can measure the recombination rates in cultured T cells and primary cells. With this new system, we have measured the recombination rates at four genetic distances. In addition, we measured the frequency of double recombination to test whether HIV-1 recombination exhibits interference.

MATERIALS AND METHODS

Nomenclature and plasmid construction. Plasmids were constructed with standard molecular cloning techniques (39). The names of all plasmids used in this study begin with p, but the names of the viruses derived from these plasmids do not. Plasmid pHIV-HSA-IRES-GFP (a kind gift from Derya Unutmaz, Vanderbilt University) is a pNL4-3-based construct that encodes *gag*, *pol*, *tat*, and *rev* but contains inactivating mutations in *vif*, *vpu*, *vpr*, and *env*. Additionally, this plasmid has an insertion in the *nef* reading frame that contains a mouse heat-stable antigen gene (*HSA*) followed by an internal ribosome entry site (IRES) from encephalomyocarditis virus and the green fluorescent protein gene (*GFP*).

Plasmid pON-fHIG is identical to pHIV-HSA-IRES-GFP except that an NcoI site in the plasmid backbone was eliminated by partial NcoI digestion and fill-in reaction with the Klenow fragment of *Escherichia coli* DNA polymerase I. Plasmid pON-fTIG was identical to pON-fHIG except that *HSA* was replaced with the mouse CD90.2 gene, also known as *Thy-1*. *Thy-1* was PCR amplified from pSRalphaLthy (32) (a kind gift from Irvin S. Y. Chen, University of California at

Los Angeles) with primers containing SacII and XhoI sites. The PCR product was digested with SacII plus XhoI, and the resulting DNA fragment was inserted into SacII- plus XhoI-digested pHDV-eGFP (42) to generate pHIV-Thy1. *Thy-1* was amplified by PCR from pHIV-Thy1; the PCR products were digested with NotI plus EcoRI and inserted into NotI- plus EcoRI-digested pON-fHIG to generate pfTIG-SL-delta. The sequence between *env* and the 5′ end of *Thy-1* in pfTIG-SL-delta was replaced with the sequence from pHIV-Thy1 by substituting the BamHI-BstEII DNA fragment to generate pON-fTIG.

HIV-1 vectors with mutated *GFP* were constructed by generating PCR products of mutated *GFP* and then subcloning the PCR DNA fragment into pON-fHIG or pON-fTIG. Mutated *GFP* was generated by single-round or overlapping PCR with pON-fHIG as the template. The PCR product containing the mutated *GFP* was digested with NcoI plus XhoI and cloned into NcoI- plus XhoI-digested pON-fHIG or pON-fTIG. The names of these plasmids contain a letter indicating the encoded functional marker (H for HSA and T for Thy-1) and a number indicating the position of the mutation in *GFP* from the translational start codon. Plasmids pON-H0, pON-H5, and pON-H6 contained mutations 15, 500, and 603 bp, respectively, downstream from the translation start codon, whereas pON-T3 and pON-T6 contained mutations 303 and 603 bp, respectively, downstream from the GFP start codon.

Helper plasmid pCMV-dGag was derived from pCMVΔ8.2 (29) in two steps. First, a 4-bp frameshift mutation was introduced into the SpeI site in *gag* to generate pCMV-Spe*. A second mutation in *gag*, an in-frame stop codon, was introduced 37 bp downstream of the *gag* AUG by site-directed mutagenesis to generate pCMV-dGag. The expression and/or function of the accessory proteins were confirmed by Western analyses and/or functional assays.

All constructs were characterized by restriction digestion, and the PCR-amplified regions were verified by DNA sequencing to avoid inadvertent mutations. The phenotypes of the constructs were characterized by flow cytometry analyses of cells transfected with these plasmids and cells infected with viruses derived from these plasmids.

Cells, transfections, and infections. The modified human embryonic kidney cell line 293T (14) was maintained in Dulbecco's modified Eagle's medium. The human T-cell line Hut/CCR5, derived from Hut78 cells to express chemokine receptor CCR5 (44), was maintained in RPMI medium. The media for both cell lines were supplemented with 10% fetal calf serum, penicillin (50 U/ml), and streptomycin (50 U/ml). Puromycin (1 μg/ml) and G418 (500 μg/ml) were also added to the medium for Hut/CCR5 cells. All cultured cells were maintained in humidified 37°C incubators with 5% CO_2.

Primary blood lymphocytes were isolated from healthy donors through Histopaque (Sigma) gradients and activated by phytohemagglutinin (2 μg/ml) for 3 days. Activated cells were maintained in RPMI medium supplemented with 10% fetal calf serum and 200 U of recombinant interleukin-2 per ml for 3 to 4 days. $CD4^+$ T cells were isolated with the Dynabeads CD4 positive isolation kit, which typically generated >99% pure $CD4^+$ T cells, as determined by flow cytometry analyses.

DNA transfections were performed by the calcium phosphate method (39) with an MBS mammalian transfection kit (Stratagene). Cells were plated at a density of 4×10^6 per 100-mm-diameter dish and transfected 18 h later. Viral supernatants were harvested 24 h later, clarified through a 0.45-μm-pore-size filter to remove cellular debris, and used immediately or stored at −80°C prior to infection.

For infection of 293T cells, cells were plated in a 100-mm-diameter dish at a density of 10^6 cells and infected 18 h later. Serial dilutions were generated from each viral stock and used for infection in the presence of Polybrene at a final concentration of 50 μg/ml. Viruses were removed 1 h later, and fresh medium was added to the cells; 48 h postinfection, the cells were processed, and flow cytometry analyses were performed. For infection of Hut/CCR5 cells, 2.5×10^5 or 1×10^6 cells were plated in a 24- or 6-well plate, respectively; infection was performed without Polybrene, and infected cells were analyzed 72 h postinfection.

Antibody staining and flow cytometry. Cells were stained with phycoerythrin-conjugated anti-HSA antibody (Becton Dickinson Biosciences) and allophycocyanin-conjugated anti-Thy-1 antibody (eBioscience). The concentrations of antibodies used to stain 293T cells were 0.8 μg/ml (phycoerythrin-labeled HSA) and 3.6 μg/ml (allophycocyanin-labeled Thy1.2), whereas the concentrations of antibodies used to stain Hut/CCR5 cells were 0.48 μg/ml (phycoerythrin-labeled HSA) and 3.6 μg/ml (allophycocyanin-labeled Thy-1). Cells that were used only for flow cytometry analyses were fixed with paraformaldehyde (1%, final concentration). Flow cytometry analyses were performed on a FACSCalibur (BD Biosciences), whereas cell sorting was performed on a FACSVantage SE system with the FACSDiVa digital option (BD Biosciences). Data obtained from flow cytometry analyses were analyzed with FlowJo software (Tree Star).
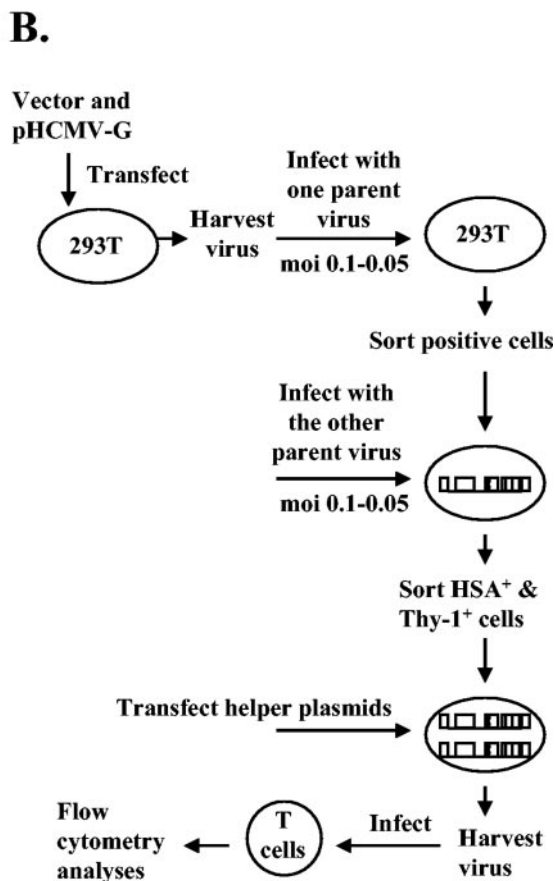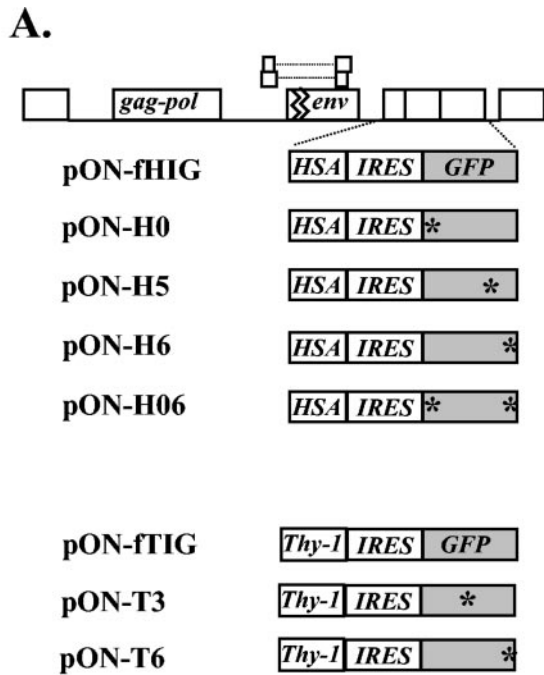
A.

B.

FIG. 1. Viral vectors and protocol used to measure HIV-1 recombination rates. (A) General structures of the vectors. All listed vectors have similar structures but differ in the encoded marker genes. Asterisk, inactivating mutation in *GFP*. (B) Protocol used to measure the recombination rates of HIV-1.

TABLE 1. Sequence comparison between wild-type and mutant *GFP*

Muta-tion	DNA sequence in *GFP*[a]	
	Wild type	Mutant
H0	ATGGTGAGCAAGGGCGAG	ATGGT**AGTT**AA**CT**G**A**GAG
T3	CATCTTCTTC–AAGGACGACG	CATCTTCTTC**G**AAGGACGACG
H5	TGAACTTCAA–GATCCGCCAC	TGAACTTCAA**G**GATCCGCCAC
H6/T6	ACAACCACTAC–CTGAGCACC	ACAACCACTA**GT**CTGAGCACC

[a] Substituted and inserted nucleotides are in bold.

Calculation of the recombination rate. Multiplicity of infection (MOI) was calculated as follows. Y represents the number of total live cells analyzed during flow cytometry analyses; Z_i and Z_g represent the number of infected cells and the number of GFP$^+$ cells, respectively. Infection MOI was calculated as log $(1 - Z_i/Y)/\log[(Y - 1)/Y]/Y$, whereas GFP MOI was calculated as log $(1 - Z_g/Y)/\log[(Y - 1)/Y]/Y$. The percentage of theoretical maximum measurable recombination rate for measuring markers 588, 300, 288, and 103 bp apart was calculated as $[(GFP^+ \ MOI/infection \ MOI)/12.5\%] \times 100\%$.

RESULTS

Experimental system used to examine HIV-1 recombination. We developed a flow cytometry-based system to measure the recombination rates at marker distances shorter than 1.0 kb in a single round of HIV-1 replication. This system uses three markers: HSA, Thy-1, and GFP. HSA and Thy-1 are cell surface proteins that can be detected with specific fluorochrome-conjugated antibodies and flow cytometry, whereas GFP is a fluorescent protein. We constructed two pNL4-3-based HIV-1 vectors, pON-fHIG and pON-fTIG (Fig. 1A). These vectors contained all the *cis*-acting elements necessary for virus replication and encoded *gag*, *pol*, *tat*, and *rev*. In addition, each vector contained two marker genes in the *nef* reading frame: pON-fHIG had *HSA* and *GFP*, whereas pON-fTIG had *Thy-1* and *GFP*. *HSA* and *Thy-1* were expressed by spliced mRNAs, whereas the translation of *GFP* was facilitated by an IRES from encephalomyocarditis virus in both vectors.

Three vectors were derived from pON-fHIG by introducing inactivating mutations in *GFP*. Plasmid pON-H0 contained 7-nucleotide substitutions between nucleotides 6 and 15 that introduced a stop codon in each reading frame, with the translation start site as nucleotide 1 (Table 1). Plasmids pON-H5 and pON-H6 contained a +1 frameshift between nucleotides 500 and 501 and between nucleotide 603 and 604 of *GFP*, respectively (Table 1). Similarly, two vectors were derived from pON-fTIG: pON-T3 and pON-T6 contained a +1 frameshift between nucleotides 303 and 304 and between nucleotides 603 and 604 of *GFP* (Table 1). None of the five plasmids with a mutation in the *GFP* gene could express functional GFP; however, upon recombination, a functional *GFP* gene could be generated and its gene products could be scored by flow cytometry. By combining different pairs of vectors, we could measure the recombination rates for different marker distances.

Experimental procedure used to generate producer cell lines and measure recombination rates. To measure the recombination rates in a single round of HIV-1 replication, we elected to produce the virus for the recombination assay from established cell lines containing HIV-1 vector proviruses instead of transiently cotransfecting the two HIV-1 vectors along with helper plasmids into cells. Although more laborious, our pro-

cedure eliminated possible DNA recombination between the two HIV-1 vectors during transfection and avoided overexpression of transfected plasmids, which would not reflect the true expression of the integrated proviruses. Furthermore, our procedure allowed us to measure the recombination events in one complete round of virus replication—from a provirus in the producer cells to a provirus in the target cells.

The experiments were performed in the following manner (Fig. 1B). First, 293T cells were transfected with an HIV-1 vector and pHCMV-G (45), a plasmid that expresses vesicular stomatitis virus G protein that can pseudotype HIV-1. Viral supernatants were harvested, clarified through a filter, serially diluted, and used to infect fresh 293T cells. A portion of the infected cells was stained with antibodies and analyzed by flow cytometry. Although this procedure could easily infect more than 80% of the cells, we selected cell populations that were infected with an MOI of 0.05 to 0.1 for further experiments. This approach allowed us to avoid cell populations in which a large proportion of the cells contained more than one virus; selecting such populations would complicate later analyses (see Discussion). The infected cells were enriched by cell sorting (Fig. 1B) and infected by a second virus at an MOI of 0.05 to 0.1; the doubly infected cells were then sorted until more than 95% of the cells expressed both HSA and Thy-1 to ensure the expression of both parental viruses. During the generation of the cell lines, cells were sorted no more than a total of four times. To avoid biases, each cell line consisted of a large number (20,000 to 60,000) of independently infected cells.

To measure the recombination rate, cell lines expressing both *HSA*- and *Thy-1*-encoding vectors were transfected with helper plasmids, and viruses harvested from these cells were used to infected the human T-cell line Hut/CCR5; the infected cells were then analyzed by flow cytometry to determine the numbers of cells expressing HSA, Thy-1, and GFP. Two helper plasmids were used in these experiments: pIIINL(AD8)env, which expressed HIV-1 envelope from the AD8 strain (8), and pCMV-dGag, which expressed HIV-1 accessory genes *vif*, *vpr*, *vpu*, *tat*, *rev*, and *nef*. Because 293T cells did not express CD4, virus generated from these cells could not reinfect the producer cells; additionally, once transferred into Hut/CCR5 cells, the HIV-1 vectors did not express Env and thus could not reinfect target cells. Therefore, these experiments measured recombination events that occurred between the provirus in the producer cells to the provirus in the target cells—one complete cycle of HIV-1 replication.

Experimental controls for the flow cytometry analyses. Before this new system could be used for recombination studies, we performed various control experiments to establish the background of the experimental system and our ability to detect the marker genes. Flow cytometry analyses were performed on mock-infected 293T cells and the producer cells. Uninfected 293T cells stained with anti-HSA and anti-Thy-1 antibodies had negligible numbers of cells that were positive for any of the three markers (Fig. 2A and B). An example of a producer cell line doubly infected with ON-H0 and ON-T6 is shown in Fig. 2C and D; although more than 95% of the cells were positive for both HSA and Thy-1 expression, none of the cells were positive for GFP expression (0 in 11,698 live events, or less than 0.008%).

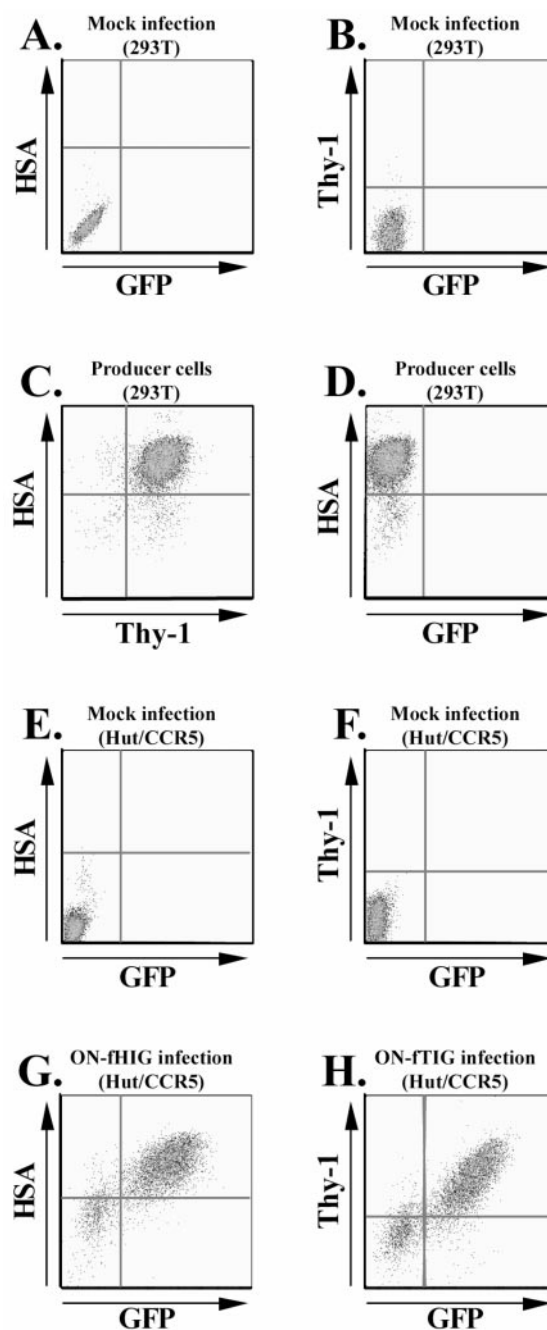To determine the experimental background, we mock-in-

FIG. 2. Representative flow cytometry analyses of mock-infected cells, producer cells, and cells infected with control plasmids. (A and B) Analyses of mock-infected 293T cells stained with anti-HSA and anti-Thy-1 antibodies. (C and D) Analyses of a producer cell line used for virus production. This cell line was sequentially infected with ON-H0 and ON-T3 at low MOIs; HSA[+] and Thy-1[+] cells were enriched by sorting, stained with antibodies, and analyzed. (E and F) Analyses of uninfected Hut/CCR5 target cells stained with anti-HSA and anti-Thy-1 antibodies. (G) Analysis of Hut/CCR5 target cells infected with ON-fHIG virus and stained with anti-HSA antibody. (H) Analysis of Hut/CCR5 cells infected with ON-fTIG virus and stained with anti-Thy-1 antibody. In all panels, the *x* and *y* axes denote the expression of a particular marker as indicated.

TABLE 2. Recombination between two markers separated by 588 bp

Accessory gene products present and cell line	Total no. of live events	No. of infected cells	No. of GFP⁺ cells	Infection MOI	GFP MOI	GFP MOI/ infection MOI	% of TMMRR[a]
With all accessory gene products							
Cell line 1	302,115	132,787	11,866	0.58	0.040	0.069	55.4
Cell line 2	311,294	108,238	7,896	0.43	0.026	0.060	48.1
Cell line 3	134,998	53,493	5,366	0.50	0.041	0.080	64.3
Mean ± SD							55.9 ± 8.1
Without Vif, Vpr, Vpu, and Nef							
Cell line 1	244,153	92,266	7,543	0.47	0.031	0.066	52.9
Cell line 2	323,309	102,981	7,682	0.38	0.024	0.063	50.2
Cell line 3	153,351	44,530	3,800	0.34	0.025	0.073	58.5
Mean ± SD							53.9 ± 4.3

[a] TMMRR, theoretical maximum measurable recombination rate.

fected Hut/CCR5 cells, stained these cells with anti-HSA and anti-Thy-1 antibodies, and analyzed them by flow cytometry. An example is shown in Fig. 2E and 2F; very few cells were positive for HSA, Thy-1, or GFP expression. Because recombination events were scored by the reconstitution of a functional *GFP* gene, we further defined the background of GFP detection. In multiple independent experiments, we scored a total of more than 2.3 million mock-infected cells and observed 18 GFP⁺ cells, indicating a background of ≈0.0008%. We generated viruses derived from either pON-fHIG or pON-fTIG, with transient transfection of 293T cells along with helper plasmids, and infected Hut/CCR5 cells. As shown in the representative analyses (Fig. 2G for ON-fHIG virus infection and Fig. 2H for ON-fTIG virus infection), most of the cells were either double negative or double positive, indicating equivalent detection of the two marker genes in both HIV-1 vectors.

It was also important that the *GFP* mutants used in these experiments were indeed negative for GFP expression; we performed multiple experiments to confirm that all four mutations in *GFP* inactivated the gene product (Fig. 2D and data not shown).

Detection of recombination events for markers separated by 588 bp. Three independent cell lines doubly infected with the ON-H0 and ON-T6 viruses were generated (shown as cell lines 1, 2, and 3 in Table 2); the *GFP* mutations in ON-H0 and ON-T6 were 15 and 603 bp downstream of the AUG codon, respectively, creating a genetic distance of 588 bp. Helper plasmids were transfected into producer cell lines, viruses were harvested and used to infect Hut/CCR5 cells, and the infected cells were analyzed by flow cytometry. Representative mock-infected Hut/CCR5 cells are shown in Fig. 3A and B, and infected Hut/CCR5 cells are shown in Fig. 3C and D. As shown in Fig. 3C and D, cell populations that were negative, single positive, or double positive for the markers could be easily detected and scored. The data generated from these analyses are summarized in Table 2. For example, viruses produced from cell line 1 were used to infect Hut/CCR5 cells; of the 302,115 live events scored, 132,787 cells expressed at least one marker (infected cells), and 11,866 cells expressed GFP. To more accurately calculate the recombination events, we converted the cell numbers into MOIs (see Materials and Methods for formula for conversion); the MOIs for infection and GFP⁺ virus were 0.58 and 0.04, respectively, indicating that 6.9% of the infection events generated GFP⁺ phenotypes.

Calculation of theoretical maximum GFP⁺ phenotype. In this experimental system, we could measure the virus titers of the two parental phenotypes and the GFP⁺ recombinant phenotype. We made the following calculation to obtain the theoretical maximum proportion of cells that could have the GFP⁺ phenotype. Assuming that RNA expression of the two parental vectors was equal in the producer cell and that RNA packaging was random, theoretically 50% of the virions produced would contain a copy of RNA from each parent (heterozygotes), 25% would contain two copies of RNA from one parent, and 25% would contain RNAs from the other parent (both homozygotes). The GFP⁺ phenotype could only be generated from reverse transcription of the heterozygous virions. With two mutations in *GFP*, four *GFP* genotypes could be generated during recombination (Fig. 4); of these, only the
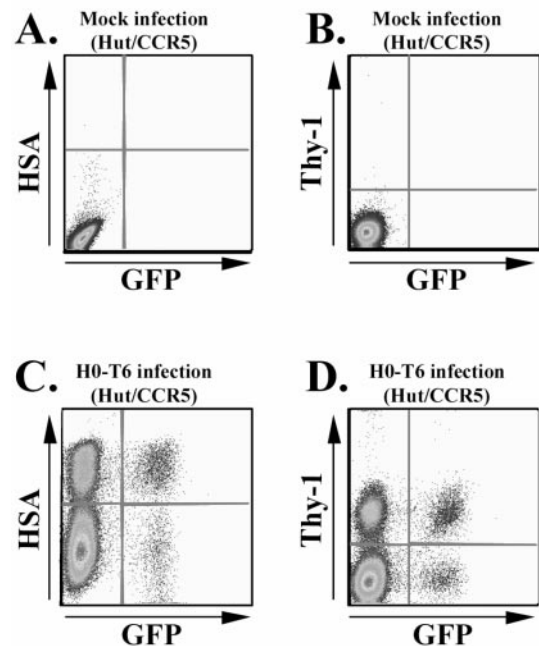

FIG. 3. Representative flow cytometry analyses of mock-infected and infected target cells. (A and B) Analyses of mock-infected Hut/CCR5 target cells stained with anti-HSA and anti-Thy-1 antibodies. (C and D) Analyses of Hut/CCR5 target cells infected with virus harvested from a producer cell line harboring both ON-H0 and ON-T6.
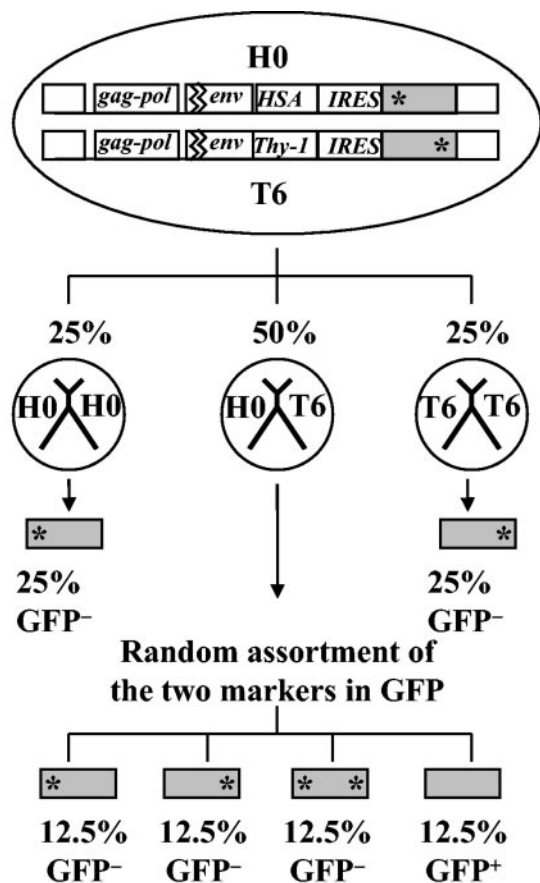
FIG. 4. Theoretical distribution of GFP genotypes and phenotypes in the progeny generated from doubly infected cells after one round of viral replication. Vectors ON-H0 and ON-T6 are used as examples (shown as H0 and T6, respectively). Several assumptions were made in the calculated theoretical frequency. First, the 25%:50%:25% distribution of the virion content was based on the assumptions that H0 and T6 were expressed at similar levels in the producer cells and virion RNAs were packaged randomly. Second, the 12.5% distribution of each genotype was based on the assumption that the H0 and T6 markers segregated as unlinked markers. GFP⁻, cells that did not express functional GFP; GFP⁺, cells expressing functional GFP.

GFP without any mutation could express functional proteins. At the maximum recombination rate, the two markers in *GFP* assorted randomly; of all the virions generated from the producer cells, only 12.5% of the progeny were expected to re-

constitute a functional *GFP*. Therefore, at most, 12.5% of the infection events should yield the GFP⁺ phenotype. When we observed that 6.9% of the infection events had the GFP⁺ phenotype, recombination between these two markers in *GFP* separated by 588 bp occurred at 55.4% of the theoretical maximum measurable rate (6.9%/12.5% × 100%) (Table 2).

Recombination rates for markers separated by 300, 288, and 103 bp. We generated cell lines containing both ON-H0 and ON-T3 proviruses and cell lines containing both ON-T3 and ON-H6 proviruses. The distance between the H0 and T3 mutations in *GFP* was 288 bp, whereas the distance between the T3 and H6 mutations was 300 bp. Despite the similar genetic distance between these two sets of mutations, they contained completely different nucleotide sequences: H0 to T3 comprised the sequences from the 5′ portion of *GFP*, whereas T3 to H6 encompassed the 3′ portion of *GFP* (Fig. 1). We measured the frequencies at which GFP⁺ phenotypes were generated among infection events in cell lines containing either set of mutations; these results are summarized in Tables 3 and 4. The recombination rate and standard deviation (SD) between H0 and T3 (288 bp) was 30.7% ± 4.3% of the theoretical maximum measurable rate; the recombination rate between T3 and H6 (300 bp) was 38.2% ± 5.6%. There were overlaps (SD) among the three sets of measurements of the two distances, indicating that the two measured sequences yielded similar recombination rates.

We also generated cell lines containing ON-H5 and ON-T6 proviruses and measured the recombination rate when the two markers were separated by 103 bp. The data generated from three independent experiments indicated that 1.4% of the infection events were GFP⁺, or 11.5% ± 2.4% (SD) of the theoretical maximum measurable recombination rate (Table 5).

Effect of accessory genes on the recombination rate. The recombination experiments described above were performed in the presence of the accessory gene products. It has been shown that some accessory genes can affect the HIV-1 mutation rate or the process of reverse transcription (1, 27, 40). To examine the effect of accessory genes on the recombination rate, we also performed parallel experiments without the expression of *vif*, *vpr*, *vpu*, and *nef* by omitting the helper construct pCMV-dGag. The results from these experiments are summarized in Tables 2, 3, 4, and 5; these results indicated that although the presence of the accessory gene products might have affected the infectivity of the virus as previously de-

TABLE 3. Recombination between two markers separated by 300 bp

Accessory gene products present and cell line	Total no. of live events	No. of infected cells	No. of GFP⁺ cells	Infection MOI	GFP MOI	GFP MOI/ infection MOI	% of TMMRR[a]
With all accessory gene products							
Cell line 1	169,742	68,118	4,751	0.51	0.028	0.055	44.3
Cell line 2	457,070	125,496	6,071	0.32	0.013	0.042	33.3
Cell line 3	377,710	142,939	8,228	0.48	0.022	0.046	37.1
Mean ± SD							38.2 ± 5.6
Without Vif, Vpr, Vpu, and Nef							
Cell line 1	177,088	43,332	2,513	0.28	0.014	0.051	40.7
Cell line 2	543,741	133,692	6,768	0.28	0.013	0.044	35.5
Cell line 3	710,696	157,916	7,463	0.25	0.011	0.042	33.6
Mean ± SD							36.6 ± 3.7

[a] TMMRR, theoretical maximum measurable recombination rate.

TABLE 4. Recombination between two markers separated by 288 bp

Accessory gene products present and cell line	Total no. of live events	No. of infected cells	No. of GFP⁺ cells	Infection MOI	GFP MOI	GFP MOI/ infection MOI	% of TMMRR[a]
With all accessory gene products							
Cell line 1	174,404	72,866	4,140	0.54	0.024	0.044	35.5
Cell line 2	540,508	195,323	8,225	0.45	0.015	0.034	27.4
Cell line 3	271,641	127,694	6,218	0.64	0.023	0.036	29.2
Mean ± SD							30.7 ± 4.3
Without Vif, Vpr, Vpu, and Nef							
Cell line 1	176,526	41,014	1,777	0.26	0.010	0.038	30.6
Cell line 2	469,566	144,843	6,371	0.37	0.014	0.037	29.6
Cell line 3	585,638	203,767	8,356	0.43	0.014	0.034	26.9
Mean ± SD							29.0 ± 1.9

[a] TMMRR, theoretical maximum measurable recombination rate.

scribed, these gene products did not affect the overall recombination rate significantly (two-way analysis of variance, $P = 0.435$; analysis of covariance, $P = 0.514$) (Fig. 5).

HIV-1 recombination in primary CD4⁺ activated T cells. We also measured the HIV-1 recombination rate in activated human primary CD4⁺ T cells. In order to make direct comparisons between the recombination rates in different target cells, we performed infection of the primary cells with the same virus stocks that were used to generate the data for Tables 2, 4, and 5. Only viruses propagated in the presence of all accessory genes were used in these experiments. Three days postinfection, these cells were processed and analyzed by flow cytometry. For each recombination rate, we performed experiments with cells derived from three different donors and viruses generated from three independent cell lines. The ratios of GFP MOI/infection MOI for markers separated by 588, 288, and 103 bp were 7.1% ± 1.1% (SD), 4.7% ± 0.5% (SD), and 2.0% ± 0.3% (SD), respectively. These data correspond to 56.8, 37.6, and 16% of the theoretical maximum measurable recombination rate, and these values were similar to those determined in Hut/CCR5 cells (Fig. 5).

Assortment of the HSA, Thy-1, and GFP markers in target cells. The two parental vectors used in this study had different marker genes upstream of *GFP*. Therefore, we could also analyze the distribution of the GFP⁺ cells in HSA⁺ and Thy-1⁺ cells to monitor additional recombination events. The distances from the 3′ end of *HSA*/*Thy-1* to the H0, T3, and H5 mutations in *GFP* were 0.6, 0.9, and 1.1 kb, respectively. In order for the GFP⁺ provirus in the H0-T6 and H0-T3 experiment to have *HSA*, reverse transcriptase would have to switch

templates within the 0.6 kb separating the *HSA* and the first 4 nucleotides of *GFP* on the *Thy-1*-containing RNA. Similarly, in order for the GFP⁺ provirus to have *Thy-1* or *HSA* in the T3-H6 or H5-T6 experiments, reverse transcriptase would have to switch templates within 0.9 or 1.1 kb, respectively. We also analyzed the distribution of GFP⁺ cells in HSA⁺ and Thy-1⁺ cells (Fig. 6).

In both the H0-T6 and H0-T3 experiments, the ratios of GFP⁺ in HSA⁺ cells were lower than those of GFP⁺ in Thy-1⁺ cells, whereas in T3-H6 and H5-T6 experiments, the ratios of GFP⁺ in HSA⁺ cells were similar to those of GFP⁺ in Thy-1⁺ cells. The marker distributions in these experiments were similar to the expected distribution based on the distance. For example, we expected that more GFP⁺ viruses would have Thy-1 markers in the H0-T6 experiment because the recombination rate was under the maximum rate when markers were separated by 0.6 kb. These data suggested that HIV-1 crossover events do not have high negative interference, in contrast to the previous proposal by others (47). Although able to provide useful information, the distribution of the markers could be complicated by the doubly infected cell population and the expression of the two parental viruses. We performed further experiments to address whether HIV-1 recombination has high negative interference.

Generating GFP⁺ phenotype by double-crossover events. High negative interference is defined by double-crossover events generated at a frequency far higher than predicted from the rates of the single crossover events. To directly test whether HIV-1 recombination exhibits high negative interference, we generated pON-H06 (Fig. 1A), which is identical to pON-H0

TABLE 5. Recombination between two markers separated by 103 bp

Accessory gene products present and cell line	Total no. of live events	No. of infected cells	No. of GFP⁺ cells	Infection MOI	GFP MOI	GFP MOI/ infection MOI	% of TMMRR[a]
With all accessory gene products							
Cell line 1	155,972	63,480	1,455	0.52	0.009	0.018	14.3
Cell line 2	1,067,637	405,475	6,241	0.48	0.006	0.012	9.8
Cell line 3	1,049,218	321,544	5,021	0.37	0.005	0.013	10.5
Mean ± SD							11.6 ± 2.4
Without Vif, Vpr, Vpu, and Nef							
Cell line 1	171,884	44,695	841	0.30	0.005	0.016	13.0
Cell line 2	1,094,731	387,098	5,960	0.44	0.005	0.013	10.0
Cell line 3	1,319,480	372,408	5,565	0.33	0.004	0.013	10.2
Mean ± SD							11.1 ± 1.7

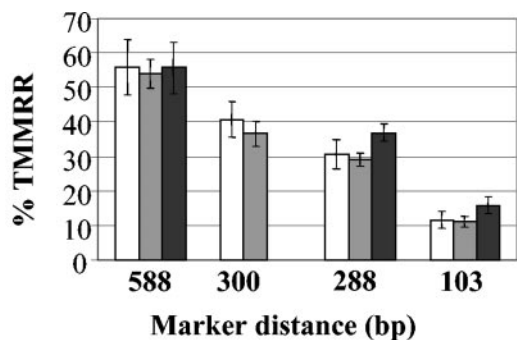[a] TMMRR, theoretical maximum measurable recombination rate.

FIG. 5. Effects of accessory proteins and target cells on HIV-1 recombination. The y axis represents the percentage of the theoretical maximum measurable recombination rate (TMMRR); the x axis represents distances between markers. White and gray bars represent the average recombination rates measured in the presence or absence, respectively, of accessory genes *vif*, *vpr*, *vpu*, and *nef* in the experimental system with Hut/CCR5 cells as target cells. Black bars represent the recombination rate in the presence of accessory genes with activated primary T cells as the target cells. All of the histograms show the average of three independent experiments; standard deviations are shown as error bars.

except that it also has the H6 mutation in *GFP*. With the protocol described for Fig. 1B, we generated two cell lines doubly infected with ON-H06 and ON-T3 and measured the rates at which GFP$^+$ viruses were generated after one round of HIV-1 replication in the presence or absence of the accessory genes. These results are summarized in Table 6. Based on our data from the H0-T3 and T3-H6 experiments, we could calculate the theoretical frequency at which GFP$^+$ viruses would be generated if the recombination events were independent.

In the H0-T3 and T3-H6 experiments, the ratios of the GFP$^+$ phenotype in the total infection events were 3.8 and 4.8%, respectively. Because heterozygous viruses were present in only half of the viruses and only half of the recombinant genotypes could be scored in this system, the recombination rate between H0-T3 and T3-H6 was 15.2% (3.8% × 2 × 2) and 19.2% (4.8% × 2 × 2), respectively. Therefore, the predicted double recombination rate was 2.9% (15.2% × 19.2%). Given the inability to score one of the double-crossover recombinant genotypes and
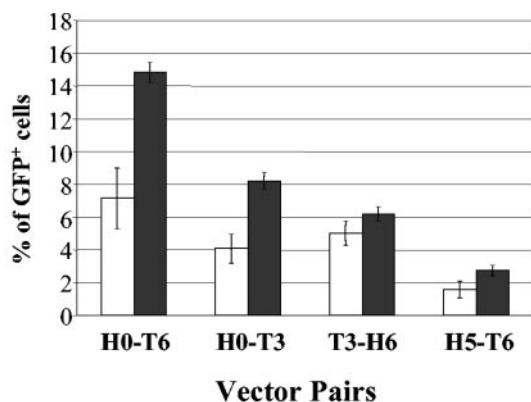
TABLE 6. Recombination between three markers with the ON-H06 and ON-T3 viruses

Accessory gene products present and cell line	Total no. of live events	Infected cells	GFP$^+$ cells	Infection MOI	GFP MOI	GFP MOI/ infection MOI
With all accessory gene products						
Cell line 1	919,357	223,561	2,379	0.28	0.003	0.0093
Cell line 2	1,034,680	343,027	3,868	0.40	0.004	0.0093
Without Vif, Vpr, Vpu, and Nef						
Cell line 1	930,100	262,201	2,946	0.33	0.003	0.0096
Cell line 2	813,916	326,331	3,955	0.51	0.005	0.0095

the presence of the homozygous viruses, the expected ratio of the GFP$^+$ phenotype in infected events was 0.73% (2.9%/4). The predicted numbers based on single recombination rates are very similar to our observed rate: the average ratio of GFP$^+$ in infected events was 0.93%, which is within the 90% confidence interval of the expected ratio. This is equivalent to not being able to reject the null hypothesis that recombination events are independent at the 0.10 level of significance. Therefore, our data do not support the hypothesis that HIV-1 recombination events exhibit high negative interference.

DISCUSSION

Development of a versatile system for HIV-1 recombination studies. In this report, we describe the development of a system to examine factors that influence the HIV-1 recombination rate. Detection of infected cells and recombinants is based on the expression of proteins that can be scored by flow cytometry. The advantages of this system are the ability to detect recombination events in T cells and primary cells and the ease and speed of collecting data from samples with large numbers of cells. Additionally, our system is versatile; we can directly measure recombination rates at various distances up to 600 bp using mutants with defined genotypes and phenotypes to generate recombinants. Furthermore, we collected viruses used in recombination studies from infected, sorted producer cells, thereby eliminating background problems such as the detection of DNA recombination events of the transfected parental vectors and allowing us to measure events that occur at lower frequencies. Recently, another flow cytometry-based assay for studying HIV-1 recombination was described, which uses derivatives of GFP that fluoresce at different wavelengths as markers and reconstitution of GFP* as a means to measure recombination (25). Although based on similar principles, the system is limited to mutations that conferred alteration of the fluorescence wavelength and could be complicated by the presence of more than one mutation that could affect the protein excitation wavelength and the multiple recombinant genotypes that could have varied phenotypes.

Calculation of HIV-1 recombination rate. Previously, we measured the recombination rates of spleen necrosis virus, murine leukemia virus, and HIV-1 (2, 18, 34). In these studies, we used two parental vectors, each encoding a functional and a nonfunctional drug resistance gene; recombination events were scored by the generation of recombinants with two functional drug resistance genes. In these previous experiments, three drug selection regimens—two single and one double

FIG. 6. Distribution of the GFP$^+$ phenotype in HSA$^+$ and Thy-1$^+$ cells. The y axis denotes the percentage of GFP$^+$ cells; the x axis denotes the vector pairs used. White bars represent GFP$^+$ cells within the HSA$^+$ populations; black bars represent GFP$^+$ cells within the Thy-1$^+$ populations.

—were used to measure the titers of the two parents and the recombinant, respectively. The recombination rate was calculated by dividing the double drug resistance titer with the lower of the two parental titers; the resulting number was then doubled to account for the recombinant with two nonfunctional genes, which could not be detected.

In our current system, flow cytometry analyses can simultaneously measure the expression of all three markers. Therefore, in addition to measuring the cell population infected by viruses with the parental phenotypes and by viruses with the recombinant phenotype (GFP$^+$), we can directly determine the total infected cell population. Based on Poisson distribution, the number of infected cells is not always in linear proportion to the number of infection events, especially at higher MOIs. Additionally, the number of GFP$^+$ cells is always lower than the total number of infected cells. Therefore, to estimate the recombination rate more accurately, we converted the numbers of infected cells to MOIs prior to calculation. This concern is less of an issue for previous studies because the numbers of drug-resistant colonies are always counted on cells infected with virus at low MOIs.

To better appreciate that the true frequency with which markers separated by a given distance segregate in one round of viral replication, we compared the rate at which GFP$^+$ was generated during infection to the theoretical maximum measurable recombination rate. We calculated the theoretical maximum measurable recombination rate in this experimental design (Fig. 4) based on the assumption that the two parents had the same level of RNA expression. In all our experiments, the expression of HSA and Thy-1 indicated that the titers of the two parental viruses were similar, generally within twofold of each other. We elected to generate virus from pools of producer cells to better mimic the events in infected patients. Because all the producer cells were pools of infected cells, it was possible that the two parents were expressed at very different levels in each cell yet generated similar titers when viruses were harvested from a large pool of cells. To reduce this bias, we infected the cells with low MOIs (0.1 to 0.05) during the generation of our cell lines to decrease the probability of the presence of multiple proviruses from the same parent in one cell. Therefore, the estimated value of 12.5% GFP$^+$ cells (Fig. 4) is likely to reflect the maximum that can be observed in our experimental condition. On the other hand, if we convert the twofold difference in the parental titer in some experiments to differences in RNA expression in the producer cells, the theoretical maximum measurable recombination rate would be 11.1%, similar to the 12.5% that we estimated in Fig. 4.

Near-linear relationship between recombination rate and marker distance at the range of 0.1 to 0.6 kb that is part of an overall quadratic fit. The recombination rates measured in this study are shown in Fig. 7A. We performed statistical analyses on the recombination rates measured at the aforementioned four distances. We found that distance is a significant factor that affects recombination rates (two-way analysis of variance, $P < 0.001$; analysis of covariance, $P < 0.001$). Further analyses indicated that at this range, the relationship between recombination rate and marker distance is nearly linear and fits a regression model involving both a linear component (first power of distance) and a quadratic component (second power of distance), with both components being significant ($P <$
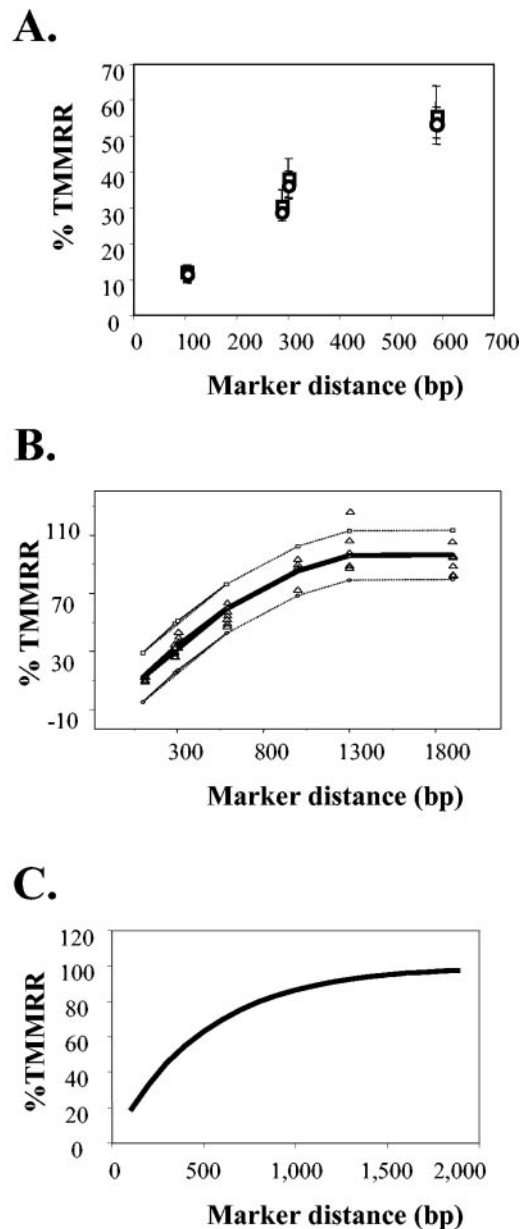
FIG. 7. Relationship between HIV-1 recombination rate and marker distance. (A) Near-linear relationship between recombination rate and marker distance in the 0.1 to 0.6 kb range. (B) The relationship between HIV-1 recombination rate and marker distances has a quadratic fit. All data points are shown as triangles; the quadratic fit is shown as a black line, and the 95% confidence limits are shown as dotted lines. (C) Simulation of observed recombination rate. This simulation is based on the assumptions that the frequency of crossover event is proportional to the marker distance, crossover occurs randomly throughout the genome, and crossovers are independent events. The x axis represent marker distance; the y axis represent the percentage of the theoretical maximum measurable recombination rate (TMMRR).

0.001 for the linear and $P = 0.015$ for the quadratic component) and an R^2 value of 0.985.

To gain insight into the overall relationship between recombination rate and marker distance, we also converted our previous measurements of recombination rates at 1.0, 1.3, and 1.9 kb into theoretical maximum measurable recombination rates.

We then added these rates to the rates determined in this report and generated a model to fit these data. We note that this approach has a caveat in that data from two independent studies are combined, thereby creating a potential weakness. The resulting model illustrates that the relationship between recombination rate and marker distance has a quadratic fit, with the linear and quadratic components being the first and second powers, respectively (Fig. 7B). This model has an outstanding fit to the data, with an R^2 value of 0.984; of the 41 data points from the experiments, only one point was outside the 95% confidence level of the model.

We also generated a simple simulation to predict the relationship between marker distance and recombination rate. This simulation has the following assumptions: the frequency of crossover between two markers is directly proportional to the distance, and the crossover events are independent, occur randomly throughout the distance, and have a Poisson distribution. Furthermore, we took into the account that between the two markers, only odd numbers of crossover events will generate a recombinant genotype, whereas even numbers of crossovers will not. The stimulation is shown in Fig. 7C and resembles the graph shown in Fig. 7B. Therefore, our data are consistent with the hypothesis that the frequency of crossover events is proportional to the distance between the markers and independent of one another.

Lack of high negative interference in HIV-1 crossover events. Previously, we observed high negative interference in spleen necrosis virus (4, 17). Based on the recombination rate and mapping of recombinant viruses, double or multiple crossovers in the viral genome occurred far more frequently than expected: we observed viruses containing two, three, four, and five crossovers at a 1.4-, 10-, 65-, and 50-fold higher frequency than expected, respectively (4). Our studies also revealed that murine leukemia virus recombination had high negative interference (3). In sharp contrast, the data presented in this report demonstrated that the directly measured HIV-1 double-crossover frequency was within the 90% confidence interval of the expected frequency calculated from the single-crossover events. This result is distinctly different from the previously observed large increase in double or multiple crossovers in simple retroviruses (3, 4). Therefore, we conclude that HIV-1 crossover events lack high negative interference. Although our data clearly indicate that HIV-1 does not exhibit the same level of high negative interference as murine leukemia virus and spleen necrosis virus, the number of experiments performed was limited; therefore, at this time we cannot completely rule out the possibility that a very low level of negative interference exists in HIV-1 crossover events.

It is important to point out that although HIV-1 crossover events lack high negative interference, HIV-1 recombination can exhibit high negative interference in certain given circumstances. In our experimental system, we have made every effort to eliminate potential biases, such as using vectors derived from the same HIV-1 strain and adapting protocols to reduce uneven distribution of the parental proviruses in virus-producing cells. However, high negative interference could occur if an experimental system has biological biases, such as the perturbation of heterozygote formation.

If heterozygote formation is reduced, the overall recombination rate will be decreased due to the lower number of observable recombinants. Because the expected frequency of double-crossover events is the product of the two single-crossover frequencies, the same bias will be applied twice, generating an expected double-crossover frequency that is lower than the observed frequency. For example, if heterozygote formation was reduced by threefold, then each of the single recombination rates would be reduced by threefold, which would yield a ninefold reduction in the expected double-crossover frequency. However, the observed double-crossover frequency would only decrease by threefold (reflecting the threefold reduction in heterozygotes). The discrepancy between the observed and the expected double-crossover frequency generates the high negative interference in recombination. Therefore, it is possible for HIV-1 to exhibit high negative interference even though HIV-1 crossover events do not.

A report by Dougherty and colleagues suggests that HIV-1 recombination has high negative interference (47). In their system, two HIV-1 vectors were derived from closely related but different strains of HIV-1, and the two parental vectors had a two-log difference in viral titers. It is possible that in the previous study, the viral genomes that did not have crossovers came from homozygous viruses. It is also possible that other biological differences in the systems created circumstances that allowed high negative interference to be observed.

The detailed mechanisms that caused the observed high negative interference in spleen necrosis virus and murine leukemia virus are not known. Similar to our HIV-1 studies, we used highly homologous vectors and producer cells that express the two parental viruses at similar levels. Thus, biological differences are most likely the causes of the observed high negative interference in murine leukemia virus and spleen necrosis virus but not in HIV-1. Although HIV-1 recombination rates can approach the theoretical maximum rates when markers are separated by 1.3 kb or more (34), the murine leukemia virus recombination rates reach a plateau far below the maximum rate even when the distances between markers are increased to 7.1 kb (2). We have previously observed in spleen necrosis virus that intramolecular template-switching events occurred far more frequently than intermolecular template-switching events. We proposed two hypotheses to explain this preference: the existence of viral populations that allow recombination (intermolecular template switching), and nonrandom RNA packaging that causes the reduction of heterozygote formation (17). Others have also hypothesized that nonrandom RNA packaging may be responsible for reduced recombination rate in murine leukemia virus (31). From the ability of HIV-1 to achieve the theoretical maximum recombination rate, we concluded that RNA copackaging between highly homologous vectors is random and heterozygotes are formed according Hardy-Weinberg distribution (34). Additional experiments are needed to determine the nature of murine leukemia virus heterozygote formation.

Implications of the high recombination frequency. One of the major hindrances in the treatment of HIV-1 infection is viral variation. The high rate of HIV-1 recombination contributes to viral variation, which hinders effective treatments. Through recombination, multidrug-resistant variants can be generated quickly in a host; for example, sequences separated by 100 bp can be assorted at a frequency of 11% per replication cycle, and double-crossover events occur frequently without

interference. With the high rates of HIV-1 replication in infected individuals, recombination can assort the genome in a short period of time. The ability of HIV-1 to evolve presents a daunting challenge for HIV-1 treatment and vaccine development. However, if we can further define and understand the pathways that HIV-1 uses to generate variation, we can better estimate the evolutionary potential and limitations of HIV-1. This understanding can help us design better therapeutic strategies for the treatment of HIV-1 infection. To continue building the scaffold needed to understand the evolution potential of HIV-1, we are currently performing further studies to elucidate the mechanisms and factors important for recombination.

ACKNOWLEDGMENTS

REFERENCES

1. **Aiken, C., and D. Trono.** 1995. Nef stimulates human immunodeficiency virus type 1 proviral DNA synthesis. J. Virol. **69:**5048–5056.
2. **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.
3. **Anderson, J. A., V. K. Pathak, and W. S. Hu.** 2000. Effect of the murine leukemia virus extended packaging signal on the rates and locations of retroviral recombination. J. Virol. **74:**6953–6963.
4. **Anderson, J. A., R. J. Teufel 2nd, 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.
5. **Balakrishnan, M., B. P. Roques, P. J. Fay, and R. A. Bambara.** 2003. Template dimerization promotes an acceptor invasion-induced transfer mechanism during human immunodeficiency virus type 1 minus-strand synthesis. J. Virol. **77:**4710–4721.
6. **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.
7. **Coffin, J. M.** 1979. Structure, replication, and recombination of retrovirus genomes: some unifying hypotheses. J. Gen. Virol. **42:**1–26.
8. **Dang, Q., J. Chen, D. Unutmaz, J. M. Coffin, V. K. Pathak, D. Powell, V. N. KewalRamani, F. Maldarelli, and W. S. Hu.** 2004. Nonrandom HIV-1 infection and double infection via direct and cell-mediated pathways. Proc. Natl. Acad. Sci. USA **101:**632–637.
9. **Derebail, S. S., M. J. Heath, and J. J. DeStefano.** 2003. Evidence for the differential effects of nucleocapsid protein on strand transfer in various regions of the HIV genome. J. Biol. Chem. **278:**15702–15712.
10. **DeStefano, J. J.** 1994. Kinetic analysis of the catalysis of strand transfer from internal regions of heteropolymeric RNA templates by human immunodeficiency virus reverse transcriptase. J. Mol. Biol. **243:**558–567.
11. **DeStefano, J. J., L. M. Mallaber, L. Rodriguez-Rodriguez, P. J. Fay, and R. A. Bambara.** 1992. Requirements for strand transfer between internal regions of heteropolymer templates by human immunodeficiency virus reverse transcriptase. J. Virol. **66:**6370–6378.
12. **DeStefano, J. J., B. Roberts, and D. Shriner.** 1997. The mechanism of retroviral recombination: the role of sequences proximal to the point of strand transfer. Arch. Virol. **142:**1797–1812.
13. **Diaz, R. S., E. C. Sabino, A. Mayer, J. W. Mosley, and M. P. Busch.** 1995. Dual human immunodeficiency virus type 1 infection and recombination in a dually exposed transfusion recipient. Transfusion Safety Study Group. J. Virol. **69:**3273–3281.
14. **DuBridge, R. B., P. Tang, H. C. Hsia, P. M. Leong, J. H. Miller, and M. P. Calos.** 1987. Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. Mol. Cell. Biol. **7:**379–387.
15. **Duesberg, P. H.** 1968. Physical properties of Rous Sarcoma Virus RNA. Proc. Natl. Acad. Sci. USA **60:**1511–1518.
16. **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.
17. **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.
18. **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.
19. **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.
20. **Iglesias-Sanchez, M. J., and C. Lopez-Galindez.** 2002. Analysis, quantification, and evolutionary consequences of HIV-1 in vitro recombination. Virology **304:**392–402.
21. **Iversen, A. K., R. W. Shafer, K. Wehrly, M. A. Winters, J. I. Mullins, B. Chesebro, and T. C. Merigan.** 1996. Multidrug-resistant human immunodeficiency virus type 1 strains resulting from combination antiretroviral therapy. J. Virol. **70:**1086–1090.
22. **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.
23. **Kung, H. J., J. M. Bailey, N. Davidson, M. O. Nicolson, and R. M. McAllister.** 1975. Structure, subunit composition, and molecular weight of RD-114 RNA. J. Virol. **16:**397–411.
24. **Kuwata, T., Y. Miyazaki, T. Igarashi, J. Takehisa, and M. Hayami.** 1997. The rapid spread of recombinants during a natural in vitro infection with two human immunodeficiency virus type 1 strains. J. Virol. **71:**7088–7091.
25. **Levy, D. N., G. M. Aldrovandi, O. Kutsch, and G. M. Shaw.** 2004. Dynamics of HIV-1 recombination in its natural target cells. Proc. Natl. Acad. Sci. USA **101:**4204–4209.
26. **Liu, S. L., J. E. Mittler, D. C. Nickle, T. M. Mulvania, D. Shriner, A. G. Rodrigo, B. Kosloff, X. He, L. Corey, and J. I. Mullins.** 2002. Selection for human immunodeficiency virus type 1 recombinants in a patient with rapid progression to AIDS. J. Virol. **76:**10674–10684.
27. **Mansky, L. M.** 1996. The mutation rate of human immunodeficiency virus type 1 is influenced by the vpr gene. Virology **222:**391–400.
28. **McCutchan, F. E., P. A. Hegerich, T. P. Brennan, P. Phanuphak, P. Singharaj, A. Jugsudee, P. W. Berman, A. M. Gray, A. K. Fowler, and D. S. Burke.** 1992. Genetic variants of HIV-1 in Thailand. AIDS Res. Hum. Retroviruses **8:**1887–1895.
29. **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.
30. **Negroni, M., and H. Buc.** 1999. Recombination during reverse transcription: an evaluation of the role of the nucleocapsid protein. J. Mol. Biol. **286:**15–31.
31. **Onafuwa, A., W. An, N. D. Robson, and A. Telesnitsky.** 2003. Human immunodeficiency virus type 1 genetic recombination is more frequent than that of Moloney murine leukemia virus despite similar template switching rates. J. Virol. **77:**4577–4587.
32. **Planelles, V., A. Haislip, E. S. Withers-Ward, S. A. Stewart, Y. Xie, N. P. Shah, and I. S. Chen.** 1995. A new reporter system for detection of retroviral infection. Gene Ther. **2:**369–376.
33. **Raja, A., and J. J. DeStefano.** 2003. Interaction of HIV reverse transcriptase with structures mimicking recombination intermediates. J. Biol. Chem. **278:**10102–10111.
34. **Rhodes, T., H. Wargo, and W. S. Hu.** 2003. High rates of human immunodeficiency virus type 1 recombination: near-random segregation of markers one kilobase apart in one round of viral replication. J. Virol. **77:**11193–11200.
35. **Robertson, D. L., P. M. Sharp, F. E. McCutchan, and B. H. Hahn.** 1995. Recombination in HIV-1. Nature **374:**124–126.
36. **Roda, R. H., M. Balakrishnan, M. N. Hanson, B. M. Wohrl, S. F. Le Grice, B. P. Roques, R. J. Gorelick, and R. A. Bambara.** 2003. Role of the reverse transcriptase, nucleocapsid protein, and template structure in the two-step transfer mechanism in retroviral recombination. J. Biol. Chem. **278:**31536–31546.
37. **Roda, R. H., M. Balakrishnan, J. K. Kim, B. P. Roques, P. J. Fay, and R. A. Bambara.** 2002. Strand transfer occurs in retroviruses by a pause-initiated two-step mechanism. J. Biol. Chem. **277:**46900–46911.
38. **Saksena, N. K., B. Wang, Y. C. Ge, S. H. Xiang, D. E. Dwyer, and A. L. Cunningham.** 1997. Coinfection and genetic recombination between HIV-1 strains: possible biological implications in Australia and South East Asia. Ann. Acad. Med. Singapore **26:**121–127.
39. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
40. **Schwartz, O., V. Marechal, O. Danos, and J. M. Heard.** 1995. Human immunodeficiency virus type 1 Nef increases the efficiency of reverse transcription in the infected cell. J. Virol. **69:**4053–4059.
41. **Shirasaka, T., M. F. Kavlick, T. Ueno, W. Y. Gao, E. Kojima, M. L. Alcaide, S. Chokekijchai, B. M. Roy, E. Arnold, R. Yarchoan, et al.** 1995. Emergence of human immunodeficiency virus type 1 variants with resistance to multiple

dideoxynucleosides in patients receiving therapy with dideoxynucleosides. Proc. Natl. Acad. Sci. USA **92:**2398–2402.

42. **Unutmaz, D., V. N. KewalRamani, S. Marmon, and D. R. Littman.** 1999. Cytokine signals are sufficient for HIV-1 infection of resting human T lymphocytes. J. Exp. Med. **189:**1735–1746.

43. **Vidal, N., C. Mulanga-Kabeya, N. Nzilambi, E. Delaporte, and M. Peeters.** 2000. Identification of a complex env subtype E HIV type 1 virus from the Democratic Republic of Congo, recombinant with A, G, H, J, K, and unknown subtypes. AIDS Res. Hum. Retroviruses **16:**2059–2064.

44. **Wu, L., T. D. Martin, R. Vazeux, D. Unutmaz, and V. N. KewalRamani.** 2002. Functional evaluation of DC-SIGN monoclonal antibodies reveals DC-SIGN interactions with ICAM-3 do not promote human immunodeficiency

45. **Yee, J. K., A. Miyanohara, P. LaPorte, K. Bouic, J. C. Burns, and T. Friedmann.** 1994. A general method for the generation of high-titer, pantropic retroviral vectors: highly efficient infection of primary hepatocytes. Proc. Natl. Acad. Sci. USA **91:**9564–9568.

46. **Zhu, T., N. Wang, A. Carr, S. Wolinsky, and D. D. Ho.** 1995. Evidence for coinfection by multiple strains of human immunodeficiency virus type 1 subtype B in an acute seroconverter. J. Virol. **69:**1324–1327.

47. **Zhuang, J., A. E. Jetzt, G. Sun, H. Yu, G. Klarmann, Y. Ron, B. D. Preston, and J. P. Dougherty.** 2002. Human immunodeficiency virus type 1 recombination: rate, fidelity, and putative hot spots. J. Virol. **76:**11273–11282.

virus type 1 transmission. J. Virol. **76:**5905–5914.