In Vitro Intersubtype Recombinants of Human Immunodeficiency Virus Type 1: Comparison to Recent and Circulating In Vivo Recombinant Forms

Miguel E. Quiñones-Mateu,^{1†} Yong Gao,¹ Sarah C. Ball,¹ Andre J. Marozsan,²
Awet Abraha,¹ and Eric J. Arts^{1,2*}

Division of Infectious Diseases, Department of Medicine,¹ and Department of Pharmacology,² Case Western Reserve University, Cleveland, Ohio 44106

Received 7 March 2002/Accepted 10 June 2002

The increased prevalence of human immunodeficiency virus type 1 (HIV-1) intersubtype recombinants (ISRs) is shaping HIV-1 evolution throughout the world and will have an impact on both therapeutic and vaccine strategies. This study was designed to generate and compare in vitro ISRs to those isolated from HIV-infected individuals throughout the world. Human peripheral blood mononuclear cells were dually infected with seven pairs of HIV-1 isolates from different subtypes (i.e., A to F). Recombinant crossover sites were mapped to specific regions in the envelope (*env*) gene by using a cloning-hybridization technique and subtype-specific probes. In vitro intersubtype recombination was at least twofold more frequent in the V1-to-V3 region than in any other *env* fragment, i.e., C1 to V1, V3 to V5, or V5 to gp41. Sequence and recombination site analyses suggested the C2 *env* domain as a “hot region” for recombination and selection of replication-competent ISRs during the 15-day incubation. In addition to these regional preferences for *env* recombination, homopolymeric nucleotide tracts, i.e., sequences known to pause reverse transcriptase and promote template switching, were found in most in vitro crossover sites. ISRs, originating from recent dual infections and limited transmission events, partly retained this in vitro regional or sequence preference for recombination sites. However, a shift to crossover sites flanking the gp120-coding sequence was evident in the stable circulating recombinant forms of HIV-1. Based on these findings, HIV-1 recombinants generated from these dual infections may be used as a model for in vivo intersubtype recombination and for the design of various diagnostic assays and vaccine constructs.

High genetic variability is inherent to all RNA viruses but has been best characterized with human immunodeficiency virus (HIV) (25, 59). The extensive heterogeneity observed in the worldwide HIV-AIDS epidemic originates from the rapid viral turnover (10^{10} viral particles/day) in an HIV-infected individual (27, 70) and the high rate of incorrect nucleotide introductions during HIV reverse transcription (10^{-4} per nucleotide) (48). In addition to this rapid accumulation of minor genotypic changes, different HIV type 1 (HIV-1) strains can also recombine to generate larger genetic alterations (30, 66; S. Wain-Hobson, 6th Annu. Disc. Meet. HIV Dynamics Evol., 1999). Recombination between two genetically distinct isolates of the same retrovirus species was first described in the 1970s (35, 42, 69). Prior to a recombination event, heterodiploid virus must be produced from cells coinfecting with two different viruses (Fig. 1A). Strand displacement and template switching during minus- or plus-strand DNA synthesis in the heterodiploid virus results in recombination (10, 33). Although retroviral recombination was initially an in vitro observation, it is now apparent that recombination between HIV-1 quasiespecies results in large genetic shifts within the inpatient population

and promotes the rapid selection of variants resistant to HIV-specific drug and immune pressure (17, 24, 26, 28, 50; S. Butto, C. Argentini, A. M. Mazzella, M. P. Iannotti, P. Leone, P. Leone, A. Nicolosi, and G. Rezza, Letter, AIDS 11:694-696). In addition, different recombinant forms of HIV-1 have recently appeared in the HIV-1 epidemic without direct links to a co-HIV-infected individual (47).

Recombination can be distinguished from an accumulation of minor genotypic changes. In general, a similarity of two contiguous sequence to different HIV-1 subtypes or groups separated by a distinct breakpoint is characteristic of intersubtype recombinant viruses (ISRs) (56, 57). However, recombinants in the HIV-1 population could be identified only after an extensive accumulation of HIV-1 sequence data necessary for the phylogenetic classification of different HIV types (1 and 2), groups (M, N, and O), and subtypes (40). HIV-1 group M strains, which are responsible for the worldwide epidemic and over 90% of current and new infections, can now be subdivided into 10 different subtypes or clades (A to J) (39) sharing 70 to 80% nucleotide sequence identity in the envelope (*env*) gene (40). Increases in travel and migration have resulted in the cocirculation of multiple HIV-1 subtypes in several regions throughout the world. For example, clades A, C, D, and, to a lesser extent, F and G have been identified in HIV-infected Ugandans (7, 12, 40). Full genome sequencing has also identified 14 circulating recombinant forms (CRFs) of HIV-1 with defined breakpoints and subtype-specific regions in the genome (reviewed in references 40, 47, and 50).

* Corresponding author. Mailing address: Division of Infectious Diseases, BRB 1029, Case Western Reserve University, 10900 Euclid Ave., Cleveland, OH 44106. Phone: (216) 368-8904. Fax: (216) 368-2034. E-mail: eja3@po.cwru.edu.

† Present address: Department of Virology, Lerner Research Institute, Cleveland Clinic Foundation, Cleveland, OH 44195.

FIG. 1. Schematic representation of in vitro recombination experiments and methods for selection and detection of HIV-1 recombinants. Dual infections of PBMC with the B-HXB2 and E-CMU06 HIV-1 isolates were used to illustrate the experiments performed in this study. (A) Pairs of HIV-1 isolates were added at three different MOI ratios (0.01:0.1, 0.1:0.1, and 0.1:0.01) as described in Materials and Methods. After the first round of replication, coinfecting cells can produce both parental and heterodiploid viruses. Infection of new cells with heterodiploid virions can lead to intersubtype recombination. (B) Scheme for PCR amplification of intersubtype HIV-1 *gag* and *env* fragments. Universal primers were used to PCR amplify the *gag* (GS1 and GA4 primers) and *env* (envB and E15 primers) DNAs from the various dual and monoinfections. Nested PCR amplification with subtype-specific primer pairs internal to the external pair was then used to select recombinants in the *gag* (MA-p6-coding region) and *env* (gp120-gp41- or V1-V5-coding region) genes. (C) Products amplified by the different combinations of subtype-specific primer pairs (e.g., *b-envL* and *b-env41*, *b-envL* and *e-env41*, *e-envL* and *b-env41*, and *e-envL* and *e-env41*) were then identified on 1% agarose gels. These products were then denatured, annealed to a radiolabeled subtype E *env* probe (HIV-1 E-TH22), and separated on an 8% nondenaturing polyacrylamide gel. (D) An autoradiograph of the HTA. Recombined B-E *env* heteroduplexes migrate to a position between the heteroduplexes of each parental *env* DNA annealed to the probe.

In contrast to the stable CRFs, ISRs, as defined on this paper, appear to be less stable in the population, have ill-defined crossover sites, and have been generated from more recent dual infections than CRFs (13, 14, 32, 46, 47, 52, 59). All HIV-1 recombinants are generated, replicate, and spread after initial co- or superinfection of single target cells in a human host. The incidence of new HIV infections, the prevalence of cocirculating clades, and limited prevention measures in several geographic regions predict that coexposures must occur at a much higher frequency than actual coinfections. Several mechanisms may limit superinfection, including a broad HIV-specific CD8⁺ cytotoxic-T-cell response in the mucosal layers of an HIV-infected individual (3, 34). To date, ISRs have been identified in nearly every region of the world where two or more subtypes cocirculate, and they may account for over 10%

of new HIV-1 infections (39, 47). ISRs are undoubtedly contributing to HIV-1 evolution and may ultimately result in a complete dissolution of defined HIV-1 subtypes (50). In addition, recombination between subtypes could represent major antigenic shifts in the HIV-1 population and hamper the effectiveness of antiretroviral therapy and vaccine strategies (50). We have developed a method to generate ISRs in vitro by first coinfecting peripheral blood mononuclear cells (PBMC) with two primary HIV-1 isolates of different subtypes. Replication-competent HIV-1 isolates containing a recombinant *env* gene were then PCR amplified with subtype-specific primers. Recombination, or crossover sites in the *env* gene, was mapped by a cloning-hybridization technique using subtype-specific probes. Finally, we compared the exact crossover sequences found in ISRs in vitro to those found in the HIV-infected

population. In general, ISRs selected in coinfecting PBMC resembled those generated *in vivo* but were significantly different from the stable CRFs.

MATERIALS AND METHODS

Cell cultures. PBMC from HIV-seronegative blood donors were obtained by Ficoll-Hypaque density gradient centrifugation of heparin-treated venous blood and cultured as described previously (51). Prior to HIV-1 infection, cells were stimulated with 2 μ g of phytohemagglutinin (PHA) (Gibco BRL) per ml for 3 to 4 days and maintained in RPMI 1640–2 mM L-glutamine medium (Cellgro) supplemented with 10% fetal bovine serum (Cellgro), 10 mM HEPES buffer (Cellgro), 1 ng of interleukin-2 (IL-2) (Gibco BRL) per ml, 100 U of penicillin (Cellgro) per ml, and 100 μ g of streptomycin (Cellgro) per ml.

Viruses. Six syncytium-inducing (SI) HIV-1 isolates (laboratory-adapted strain B-HXB2 and the five primary isolates A-92UG029, D-92UG021, D-93UG067, E-CMU06, and F-93BR020) and three non-syncytium-inducing (NSI) HIV-1 strains (laboratory-adapted strain B-BaL and the two primary isolates A-92RW009 and C-92BR025) were obtained from the AIDS Research and Reference Reagent Program. For most of the strains listed above, the letter before the dash indicates the subtype of the viral envelope and precedes the year of isolation, country of origin, and strain number; e.g., A-92RW009 is a clade A HIV-1 strain isolated in Rwanda in 1992. All viral stocks were propagated and expanded in PHA-stimulated, IL-2 treated PBMC as described previously (51). The 50% tissue culture infective dose was determined for each isolate by using the method of Reed and Muench (53), and titers were expressed as infectious units per milliliter.

HIV-1 dual-infection assay. Different pairs of HIV-1 isolates were used to simultaneously infect PBMC as described previously (51) (Fig. 1A). We performed three separate dual infections of PHA- and IL-2 treated PBMC with two HIV-1 isolates at different multiplicities of infection (MOIs) (i.e., 0.01 and 0.1, 0.1 and 0.1, and 0.1 and 0.01). One million PBMC were incubated with these virus mixtures for 2 h at 37°C with 5% CO₂, washed, and then resuspended in complete medium (10⁶/ml). Supernatants and two aliquots of cells were harvested at day 15, resuspended in dimethyl sulfoxide-fetal bovine serum, and then stored at –80°C for subsequent analysis.

PCR strategy to select HIV-1 ISRs. For all dual-infection experiments, proviral DNA was extracted from lysed PBMC by using the QIAamp DNA Blood Kit (Qiagen). Segments of the *gag* and *env* genes of the HIV-1 genome were PCR amplified by using the universal primers GS1 (TAAACATATAGTATGGGC AAGC) and GA4 (TTGCCAAAGAGTGACCTGAGGGAA) for a 1.3-kbp fragment in the *gag* gene and envB (22) and E15 (61) for a 2.17-kbp fragment encoding gp160 of *env*. Subtype-specific primers, internal to the previous *gag* and *env* products, were then used to PCR amplify subtype B and E recombinants in the following viral fragments: the *b-gag*S1 (CTGGGACAGCTACAACCATCC CTT) or *e-gag*S1 (CACACTTGTGGAAATGGGTGACTT) primer was paired with *b-gag*A1 (ACTTCGGACTCATTGTTCATTT) or *e-gag*A1 (AGGTCTG TTAGTATTATTAAG) to amplify the MA-p6-coding region of *gag* (1.2 kbp), *b-env*L (nucleotides nt 6592 to 6612) and *e-env*41 (nt 7606 to 7628) were used for the *env* gp120- and partial gp41-coding regions (1.8 kbp), and *b-env*V1 (nt 6259 to 6284) and *e-env*V5 (nt 8109 to 8141) were used for the V1-to-V5 *env* region (1.4 kbp) (Fig. 1B). Subtype-specific primers such as *b-gag*S1 are designated by a letter for the specific subtype followed by the target gene and orientation (sense or antisense to the plus-strand RNA). Table 1 provides a list of all oligonucleotide primers used to specifically amplify the *env* gene of an HIV-1 subtype or isolate. Although some primers are isolate specific, these primers are still referred to as subtype specific to avoid confusion. In addition, *e-gag*S1 and *e-gag*A1 (or *e* in *gag*) primers refer to subtype A-specific *gag* primers annealing to the ancestral subtype A *gag* gene found in CRF01_AE HIV-1 isolates. For the same reason we refer to CRF01_AE as subtype E. A compilation of the *gag* and *env* primer pairs used in subtype-specific amplifications is shown in Fig. 1B. Both external and nested PCRs were carried out in a 100- μ l reaction mixture with defined cycling condition (51). PCR-amplified products were separated on agarose gels and then purified by using the QIAquick PCR purification kit (Qiagen). Control PCR amplifications were performed with subtype-specific DNA templates to rule out the possibility of *Taq*-generated recombinants (6). Briefly, proviral DNAs from the B-HXB2 and E-CMU06 HIV-1 isolates were mixed and used directly as templates for PCR amplification with universal (external) and clade B and E subtype-specific (nested) primers (E. J. Arts, H. Baird, G. P. Caleodis, and M. E. Quinones Mateu, submitted for publication). These PCR products were then analyzed as described above (Fig. 1C).

HTA for detection of ISRs. The ISR PCR products (V1 to V5 of the HIV-1 *env* gene) were analyzed by using heteroduplex tracking analysis (HTA) (16, 51). Probes for HTA were amplified from plasmids containing *env* DNAs of six HIV-1 strains (A-pRW20, B-pSF162, C-pMA959, D-pUG46, E-pTH22, and F-pBZ162) (51). For this amplification, the universal ED5 (61) primer was radiolabeled using T4 polynucleotide kinase and 2 μ Ci of [γ -³²P]ATP and paired with ED12 (61). Radiolabeled PCR-amplified probes were separated on 1% agarose gels and then purified with a QIAquick gel extraction kit (Qiagen). Ten microliters of unlabeled PCR-amplified intersubtype DNA and 0.1 pmol of the radiolabeled *env* probe were then denatured and annealed in a buffer containing 100 mM NaCl, 10 mM Tris-HCl (pH 7.8), and 2 mM EDTA. Heteroduplexes were resolved on a 5% nondenaturing polyacrylamide gel and analyzed as described previously (51).

Cloning and probe hybridization for mapping HIV-1 intersubtype recombination breakpoints. Crossovers in the *env* gene were mapped by using a cloning and probe hybridization technique. Briefly, *env* ISRs (V1 to V5 and gp120-gp41 fragments) were PCR amplified with subtype-specific primers (see above) and then cloned into pCRII-TOPO vector (Invitrogen). Two or three *env* products from separate PCR amplifications were mixed to avoid potential resampling artifacts during clone selection. Approximately 100 individual bacteria colonies containing an *env* recombinant plasmid were transferred to nylon membranes and lysed with 10% sodium dodecyl sulfate. Bacterial DNA covalently linked to the membrane was then denatured and hybridized to γ -³²P-labeled clade-specific *env* oligonucleotides as described previously (51). Thirty oligonucleotides, between 20 to 29 nt in length and annealing to specific HIV-1 subtypes or isolates (A, B, C, D, E, and F) in V1 (nt position 6653), V3 (nt 7139), and V5 (nt 7606) of the *env* gene, were used (Table 1). The sequences and approximate genomic position of these primers are shown in Table 1 and Fig. 2, respectively. Filters were autoradiographed and analyzed to determine the subtype identity and to estimate the crossover location in each HIV-1 *env* recombinant clone.

Nucleotide sequencing, phylogenetic analysis, and recombination analysis. In vitro HIV-1 intragene recombinants (*gag* and *env*) were sequenced by using an ABI Prism BigDye terminator cycle sequencing ready reaction kit (Perkin-Elmer). The primers used in the sequencing reactions have been previously described (51). Nucleotide sequence alignments were performed using the CLUSTAL X version 1.63b program (67). Sites of intersubtype recombination were verified by using the Recombinant Identification Program (RIP) (63) as described previously (<http://hiv-web.lanl.gov/>). Sixty-five ISRs and eighty CRFs were identified in the Los Alamos HIV-1 sequence database and compared to in vitro ISRs generated in this study. Recombination sites in the *env* gene of each ISR and CRF were confirmed by using RIP and manual analysis of sequence alignments. Each crossover site was then defined by a 50-nt sequence overlapping the putative intersubtype recombination site. More than one isolate of CRFs 1 to 7, 10 to 12, and 14 were employed only if crossover sites were unique and not overlapping. A list of isolates and accession numbers from GenBank is available upon request.

RESULTS

Identifying HIV-1 *env* ISRs in dual infections. Non-CRF HIV-1 recombinants have been previously identified in HIV-infected individuals residing in geographical regions where two or more subtypes cocirculate (references 47 and 50 and references therein). However, *in vivo* intersubtype recombination is a rare event that follows (i) nearly simultaneous coinfection with two isolates of different subtypes or (ii) superinfection of an HIV-positive individual with a heterologous HIV-1 subtype (11). This study was designed to detect and characterize intersubtype recombination in the HIV-1 *env* and *gag* genes and to compare *in vitro* *env* recombinants to the CRFs identified in HIV-infected individuals.

We have previously described a method for infecting PBMCs with two or more primary HIV-1 isolates of different subtypes (51). Using HTA, we were able to measure dual virus production and to derive relative fitness values of each isolate, but we were unable to detect HIV-1 recombination in *env* (51). In a parallel study, we compared the actual frequency of recombination with the predicted frequency of HIV-1 recombi-

TABLE 1. Nomenclature and sequences of subtype- and isolate-specific oligonucleotide *env* primers

Oligonucleotide	Subtype specificity	Isolate specificity	<i>env</i> region	Direction	Sequence (5'→3')
<i>d1-envV1</i>	D	92UG021	V1	Sense (+)	CTGAAGGAAGATCCATGGGC
<i>d2-envV3</i>			V3	Sense	GTTAGCAAGGGTAAGGGCAAGG
<i>d3-envV5</i>			V5	Antisense (-)	CACAAAATTCAATGACACAGAATTC
<i>f-envV1</i>	F	93BR020	V1	Sense	GCACTGAATGGAAGAATGCC
<i>f-envV3</i>			V3	Sense	GGTAAGCTACAGGACACCTATAGGGG
<i>f-envV5</i>			V5	Antisense	AGTACGATAGGGGCAAATGG
<i>b-envL</i>	B	HXB2	C1	Sense	ATGGGGGTACCTGTGTGGAAGGAAGCAACC
<i>b-envV1</i>			V1	Sense	TAGCGGGAGAATGATAATGG
<i>b-envV3</i>			V3	Sense	AAGAATCCGTATCCAGAGAG
<i>b-envV5</i>			V5	Antisense	AAGATCTCGGACTCATTGTTGC
<i>b-env41</i>			5' gp41	Antisense	TGGCCTGTACCGTCAGCGTCATTGAGGCT
<i>e-envL</i>	E	CMU06	C1	Sense	AAGCATATGAGACAGAAGTG
<i>e-envV1</i>			V1	Sense	CCAAGGGTAGAGCACAAACC
<i>e-envV3</i>			V3	Sense	GTAAGAACAAGGATGACTATGGGACC
<i>e-envV5</i>			V5	Antisense	GTCTCGTTAGTATTATTAAGACC
<i>e-env41</i>			5' gp41	Antisense	CCTGTACCGTCAGCGTTATT
<i>a-envL</i>	A	92UG029	C1	Sense	GATGGGGTACTATGATCTTGG
<i>a-envV1</i>			V1	Sense	GGGTAGAGACAACATAACGGG
<i>a-envV3</i>			V3	Sense	TTTCAGTTCCATTTGTATTGTTACCAACC
<i>a-envV5</i>			V5	Antisense	TAACAACAATATAAGCAGAAGAGG
<i>a-env41</i>			5' gp41	Antisense	TGGCCTGTACCGTCAGCGTTACTGATGTC
<i>a-env41-2</i>			Mid-gp41	Antisense	TCGAGACCCCTTGGGTTTGGGGTAT
<i>d-envL</i>	D	93UG067	C1	Sense	CTCCTTGGGAATGTTGATGATG
<i>d-envV1</i>			V1	Sense	TGAAAGTGTCACTAACTCTCC
<i>d-envV3</i>			V3	Sense	CTCATTCTGACTAGTATTATTTGTGC
<i>d-envV5</i>			V5	Antisense	CTACAGGAAAGTAAGAATAAAACACC
<i>d-env41</i>			5' gp41	Antisense	GTCTCGTTAGTATTATTAAGACC
<i>d-env41-2</i>			Mid-gp41	Antisense	TCGGGTCCCCTCGGGGCTGGGAGGA
<i>a3-envV1</i>	A	92RW009	V1	Sense	TAGAGTGTAACAACATCACC
<i>a3-envV3</i>			V3	Sense	AGTTGTAGTATTATTATTTCC
<i>a3-envV5</i>			V5	Antisense	GTCACACTTTGAGAACATTAC
<i>c-envV1</i>	C	92BR025	V1	Sense	TAACAGAACTGATAACATGGGG
<i>c-envV3</i>			V3	Sense	TGTGTCATGCATACCTGTACC
<i>c-envV5</i>			V5	Antisense	GCAGAGCACTTCCTTAATAAAG
<i>b2-envV1</i>	B	BaL	V1	Sense	TGCACTGATTGAGGAATGC
<i>b2-envV3</i>			V3	Sense	CCTCGGTCTTGTGTCTCCTCAGG
<i>b2-envV5</i>			V5	Antisense	AGAACAATTGGGAATAAAC

nants based on single-cycle infection studies (Arts et al., submitted). The actual frequency of recombination after four rounds of replication was approximately 3 to 5%/kbp, or at least 5- to 10-fold less than predicted. This level of recombination is below the limit of HTA detection. Coinfected cells are capable of producing heterodiploid virus containing an RNA genome from each HIV-1 isolate (Fig. 1A). Actual recombination occurs during de novo infection with this heterodiploid HIV-1 and is mediated by reverse transcriptase (RT) jumping from one plus-strand RNA (or minus-strand DNA) template to the other during DNA synthesis (11).

For this study, we have performed seven dual infections with two laboratory strains (B-HXB2 and B-BaL) and seven primary HIV-1 isolates of different subtypes (A, B, C, D, E, and F). MOIs of 0.1 and 0.01 of each of two isolates were added to primary blood mononuclear cells treated with PHA and IL-2. Dual infection was monitored by HTA after 6 and 15 days

postinfection. Detailed analyses focused on recombinations between the subtype B laboratory isolate, B-HXB2, and the primary subtype E isolate, E-CMU06, as well as between the primary Ugandan isolates, A-92UG029 and D-93UG067. All four of these HIV-1 isolates had an SI phenotype and utilized the CXCR4 coreceptor for entry. However, we also performed dual infections and analyzed recombinants involving three NSI, CCR5-tropic HIV-1 isolates (B-BaL, A-92RW009, and C-92BR025). This study was designed over 4 years ago, when emerging data showed that subtypes E and B cocirculate in Thailand whereas subtypes A and D predominate in Uganda and much of sub-Saharan Africa (40). Interestingly, subtype A-D intergenic (between HIV-1 coding regions) and intragenic (within a coding region) recombinants have been identified throughout sub-Saharan Africa (4, 13, 14, 46, 54, 57), but subtype B-E recombinants have only recently been identified in Thailand (68). It is important to note that all subtype E

FIG. 2. Mapping of crossover sites to different regions in the *env* gene. Recombined and parental *env* fragments were cloned into the pCRII-TOPO vector (Invitrogen) and transformed into *Escherichia coli*. Ampicillin-resistant, white colonies were transferred to nitrocellulose, lysed, and probed with subtype-specific oligonucleotides annealing to the V1, V3, V5, or hypervariable region in gp41. (A) Summary of the crossover mapping analyses in the *env* genes of multiple B-HXB2/E-CMU06 and A-92UG029/D-93UG067 recombined *env* clones. Crossovers were mapped to envelope regions, depicted as boxes of gradient shading. Frequencies of crossover sites in ISRs and CRFs were also plotted as crossover frequency per nucleotide (i.e., number of crossovers/length of the region defined by probes; see Fig. 3). A sixth region of recombination found at the end of the gp41-coding sequence (gp41-end) was not defined by probe hybridization analyses in Fig. 3. Twenty-two of 60 crossovers that occurred in gp41-end were excluded to generated the ISR (C1-gp41) crossover frequency. (B) Recombinants from seven dual infections (A-C, A-B, A-E, D-E, D-E, B-E, and A-D) were used to determine the frequency of crossovers in the V1-V3 region compared to the V3-V5 region of *env*. Only the V1-V5 *env* fragments were PCR amplified and cloned for this hybridization analyses with probes specific to the V3 region. The number of *env* recombinant clones analyzed for each dual infection is shown in Table 2. Twenty-one of 60 in vivo ISR crossovers were found in the V1-V5 region and are plotted.

strains (currently CRF01_AE) are derived from an old recombination event between an extinct “subtype E” *env* and a early ancestor of subtype A *gag* and *pol* genes (8, 21).

PCR amplification of intersubtype *env* and *gag* recombinants. As mentioned above, we were unable to detect recombinants in the 480-bp *env* fragment used in an HTA to measure dual virus production. Thus, a PCR method was devised to detect and amplify possible *env* and *gag* recombinants in these seven HIV-1 dual infections. We designed a set of *gag*- and *env*-derived oligonucleotides that shared perfect identity with the consensus sequence of only one subtype or isolate (A through F) (Table 1). The primers *e-gag*S1 and *e-gag*A1 anneal to the ancestral subtype A *gag* gene found in subtype E HIV-1 isolates. Amplification of recombinant *gag* and *env* genes with

subtype-specific primers was preceded by an external PCR amplification using conserved *env* or *gag* primer pairs. Figure 1C clearly indicates that the subtype B-specific primer pair (*b-envS1* and *b-envA1*) or subtype E specific primer pair (*e-envS1* and *e-envA1*) could amplify *env* product in mono- or dual infections containing the B-HXB2 (lanes I, II, III, and IV) or E-CMU06 (lanes II, III, IV, and V) isolate, respectively. A primer pair containing a subtype B- and E-specific oligonucleotide (*b-envS1/e-envA1* or *e-envS1/b-envA1*) could amplify *env* gene products only from B-HXB2 plus E-CMU06 dual infections (lanes II, III, and IV). However, nonspecific PCR products were obtained by using heterogeneous primer pairs (lanes I and V). The inability of these primer pairs (*b-e* or *e-b*) to amplify specific *env* (or *gag*) products in the mono-infections provides strong evidence that mainly B-E or E-B *env* recombinants were amplified in the dual infections. Products of the incorrect size were amplified due to nonspecific annealing of recombination-specific primers to cellular DNA extracted from mono-infections (Fig. 1). We were unable to amplify B-E or E-B recombinants in mock infections. In this experiment, equivalent amounts of plasmids containing entire *env* (or *gag*) gene of B-HXB2 or E-CMU06 were mixed and then added to the external PCRs. Subsequent nested amplifications with the subtype-specific primer pairs failed to amplify recombinant *env* (or *gag*) DNA from the mock coinfections (Fig. 1C, lane C). Thus, template switching and recombination mediated by *Taq* polymerase is likely orders of magnitude less than that occurring during reverse transcription in vivo (Arts et al., submitted).

Identifying HIV-1 ISRs by HTA. The detection of *env* or *gag* DNA following PCR amplification with heterologous pairs of subtype-specific primers (i.e., *b-e* or *e-b*) is still not proof that intersubtype recombination arose in dual HIV-1 infections. B-E and E-B *env* products (Fig. 1C) were employed in HTA to confirm the presence of recombinants. An *env* recombinant was easily identified when the heteroduplex containing the B-E (or E-B) DNA migrated between the parental heteroduplexes, i.e., B-HXB2 (Fig. 1D, lane I, top of gel) or E-CMU06 heteroduplexes (Fig. 1D, lane V, bottom of gel), both derived from mono-infections. Following PCR amplification from each dual infection with subtype-specific primers (data not shown), HTA was used as initial screen for intersubtype *env* recombinants. As described below, probe hybridization assays and a PCR-sequencing technique were then used to map the breakpoint and crossover sites in these intersubtype *env* recombinant clones. It is important to note that *env* recombinants were PCR amplified from both dual infections involving two NSI/R5 or two SI/X4 isolates. However, no HIV-1 recombinants were generated in dual infections involving an SI/X4 isolate and an NSI/R5 HIV-1 isolate. All of the SI/X4 isolates had outcompeted the NSI/R5 HIV-1 isolates, reducing the probability of subsequent recombination (data not shown). Head-on competitions in PBMC suggest that SI/X4 isolates are generally less fit than NSI/R5 HIV-1 isolates (51). In addition, the frequency of coinfection with an NSI/R5 isolate and an SI/X4 isolate would be significantly less than coinfection with two isolates of the same tropism and phenotype, since few cell types in a PBMC population express both CCR5 and CXCR4. Finally, a reduced MOI (0.01 versus 0.1) also decreased the ratio of

HIV-1 recombinants to parental HIV-1 isolates produced from the dual infection (Arts et al., submitted).

Mapping intersubtype recombination to specific regions in *env*. *env* ISRs amplified with subtype-specific primers (see above) were cloned into the pCRII-TOPO vector, resulting in a disruption of the β -galactosidase gene (e.g., pCR-B-E *env*). Ampicillin-resistant, white bacterial colonies carrying the pCR-recombinant *env* vector were streaked onto fresh plates, transferred to nitrocellulose, lysed, and probed with a series of subtype-specific, radiolabeled oligonucleotides. Using these probes annealing to the hypervariable regions in *env* (V1, V3, V5, and gp41 nt 8109), we were able to map recombination or crossover sites to a 330-bp sequence in the gp120 C1-V1 region, a 510-bp sequence in the V1-V2-C2-V3 (V1-V3) region, a 470-bp sequence in the V3-C3-V4-C4-V5 (V3-V5) region, a 170-bp sequence in the gp120 V5-gp120/gp41 interface (V5-gp41), and/or a 280-nt region in gp41. Figure 2A summarizes the probe hybridization and *env* recombination analysis using 70 B-E, 62 E-B, 40 A-D, and 22 D-A colonies. This probe hybridization technique identified regions of recombination for subsequent sequencing analysis and avoided over 2,000 sequencing reactions. Approximately 40% of the crossovers in the *env* gene, derived from the B-HXB2 plus E-CMU06 or the A-92UG029 plus D-92UG067 dual infections, occurred in a region containing the end of V1 as well as the V2 and C2 domains. In contrast, the V3-V5 region supported the fewest recombination events (<13%). Differences in the lengths of these regions (delineated by the subtype-specific probes) may influence the frequency of recombination (Fig. 3 and Table 2). For example the V1-V3 fragment is 510 bp, whereas the V3-V5 fragment is only 470 bp. However, increased recombination ($>1.1 \times 10^{-3}/\text{nt}$) was still observed in the V1-V3 region with correction for length (Fig. 3 and Table 2). Approximately <5% of the clones contained two or three crossover sites. However, it is important to note that we were unable to identify those recombinants with two crossover events in a single *env* region defined by these probes.

Increased frequency of crossovers in the V1-V3 region compared to the V3-V5 region may be due in part to (i) differences in sequence identity between the pair of isolates, (ii) the sequence homology between the regions, and/or (iii) the ability of the two regions to accommodate significant genetic changes and still remain replication competent. We have initiated studies to examine the replication efficiency or fitness of these intersubtype *env* recombinants by subcloning them into a neutral HIV-1 genome. To test the first hypothesis, we have selected and analyzed intersubtype *env* recombinants in the V1-V3 and V3-V5 regions from five dual infections involving nine HIV-1 isolates of six different subtypes (A, B, C, D, E, and F) (Fig. 2B). Preference for intersubtype recombination in the V1-V3 (70% of the recombinants) over the V3-V5 *env* region was even greater with these five pairs of primary HIV-1 isolates than with B-HXB2 plus E-CMU06 or A-92UG029 plus D-92UG067 dual infections (Fig. 2). With correction for fragment lengths, the frequency of crossovers in the V1-V3 region was $1.35 (\pm 0.20) \times 10^{-3}/\text{nt}$, or twofold greater than that observed in the V3-V5 region ($P < 0.001$) (Fig. 3 and Table 2). Finally, we have compared the *env* sequence heterogeneity between each pair of HIV-1 isolates used in these dual infections. The V3-V5 region was slightly more heterogeneous than

TABLE 2. Crossover frequency in HIV-1 *env* recombinants

Dual infection ^a	<i>env</i> recombinants ^b	No. of clones ^c	Recombination frequency (crossovers/nucleotide) (10 ³)/genetic distance (substitutions/nucleotide) ^d					
			C1-V1	V1-V3	V3-V5	V5-gp41	gp41-gp41	<i>env</i> end
A-92RW009 C-92BR025	A-C + C-A	54	ND ^e	1.22/0.19	0.65/0.26	ND	ND	ND
A-92RW009 B-BAL	A-B + B-A	47	ND	1.27/0.19	0.38/0.23	ND	ND	ND
A-92UG029 E-CMU06	A-E + E-A	88	ND	1.60/0.18	0.45/0.25	ND	ND	ND
D-92UG021 E-CMU06	D-E + E-D	64	ND	1.45/0.19	0.57/0.24	ND	ND	ND
D-92UG021 F-93BR020	D-F + F-D	39	ND	1.63/0.20	0.80/0.25	ND	ND	ND
B-HXB2 E-CMU06	B-E + E-B	132	0.29/0.15	1.13/0.21	0.91/0.22	1.03/0.09	ND	ND
A-92UG029 D-92UG067	A-D + D-A	62	0.25/0.23	1.15/0.20	0.90/0.27	0.62/0.16	0.57/0.11	ND
ISR	A-C, A-D, A-G, B-D, B-G, C-A, C-D, D-A, D-B, E-A, E-H, F-B, G-H, H-A, H-E, H-G	28 (60 crossovers)	0.37/0.15	0.49/0.18	0.27/0.23	0.13/0.10	0.17/0.14	0.59/0.18
CRF 01-14	A-C, A-G, A-H, B-A, B-C, B- F, B-G, C-D, C-F, D-C, D-F, E-G, F-A, F-B, F-D, G-A, G-H, G-J, H-A, H-C, J-G	21 (45 crossovers)	0.62/0.15	0.18/0.19	0.20/0.24	0.06/0.11	0.47/0.13	0.56/0.19

^a Two HIV-1 isolates were added at different MOI ratios (1:10, 1:1, and 10:1) as previously described (51).

^b *env* PCR recombinant fragments were selected by using subtype-specific primers as described in Materials and Methods (Fig. 1B). In each case, both types of recombinants (e.g., A-C and C-A) were selected and analyzed.

^c Total number of *env* recombinant clones analyzed for each dual infection. In contrast, 28 ISRs with a total of 60 crossover sites in *env* and 21 CRFs with a total of 45 crossover sites were obtained from the Los Alamos HIV-1 Sequence Database and used to estimate the frequency of in vivo recombinants in specific *env* regions (40).

^d Recombination frequency corresponds to the total number of recombinant breakpoints or crossover sites per genomic region (crossovers) divided by the total number of recombined clones analyzed and by the size of the fragment. The genetic distances between each HIV-1 pair and within the specific *env* region were compared by using the Kimura two-parameter model as previously described (49). HIV-1 *env* genomic regions: C1-V1, 330 bp; V1-V3, 510 bp; V3-V5, 470 bp; V5-gp41, 170 bp; gp41-gp41, 280 bp; and *env* end, 640 bp.

^e ND, not determined.

the V1-V3 region (Table 2). However, this difference was quite modest (1.2-fold; $P < 0.001$), but it may still contribute to the enhanced recombination in the V1-V3 region (2-fold; $P < 0.001$).

Comparison of in vitro and in vivo intersubtype recombination in the *env* gene. The results described above suggest that the V1-V3 region of *env* could be a “hot region,” or selected region for intersubtype recombination in replication-competent viruses in vitro. To examine whether preferential *env* recombination occurs in vivo, we have collected nearly all of the full-length intersubtype *env* recombinants (19) and CRFs (38) available in the Los Alamos HIV sequence database (40). Intersubtype HIV-1 *env* recombinants and CRFs containing crossovers in *env* were compared to the ISRs generated in vitro. Although HIV-1 recombination is not limited to the *env* gene, a significant proportion of the intersubtype HIV-1 recombinants (34%) have been identified in the *env* gene (40, 50). This likely reflects a greater number of HIV-1 *env* sequences submitted to GenBank rather than increased recombination frequency in *env*. The

RIP (63) was used to confirm the exact site of intersubtype *env* recombination in these primary HIV-1 isolates (data not shown). Most of the ISRs are not or have yet to be classified as stable CRFs and are likely the result of a recent dual infection. Very few of the ISRs isolated from HIV-infected individuals share identical crossover sequences. However, there was a higher frequency of multiple crossover sites in ISR isolates than in those generated in vivo. We did exclude ISRs with identical crossover sites from these analyses. Likewise, analysis of *env* crossover sites in CRFs was limited to representatives of the 15 classified types. Most isolates of a specific CRF contained the same crossover site in the *env* gene. However, we did include six isolates of CRF 12 and three isolates of CRF 11 due to different crossover sites (defined by a 50-nt segment) in *env* regions.

Crossover sites in the CRF and ISR sequences were then mapped to specific regions in HIV-1 *env* defined by our subtype-specific probes and for comparisons to the intersubtype recombination in vitro (Fig. 3 and Table 2). Surprisingly, the likelihood that an in vivo ISR crossover site mapped to a

FIG. 3. Frequency of recombination breakpoints in the *env* gene. Crossover sites were mapped in Fig. 2, and the frequency of recombination within a specific region was calculated in Table 2. (A) The frequencies of region-specific recombination with the *env* ISRs selected in vitro (B-HXB2/E-CMU06 and A-92UG029/D-93UG067) were compared to those of 60 ISR and 45 CRF crossovers reported in the HIV database (40). As described in Fig. 2, 42% of 28 ISR isolates and 75% of 21 CRF isolates contained more than one crossover site in *env*. (B) The crossover frequencies (10^{-3} per base pair) in the V1-V3 and V3-V5 gp120-coding regions of seven recombined pairs were also compared to those of the in vivo ISRs and CRFs. $P < 0.001$, Pearson product moment correlation.

specific region in *env* nearly matched that observed in vitro. When corrected for length (crossovers per 10^3 nt), the highest frequency of in vivo ISR crossovers mapped to the V1-V3 region of *env* (Fig. 3B; Table 2). A more defined analysis was obtained following the removal of *x* crossovers that mapped to *env* regions outside of those defined by the hybridization probes. Preference for recombination in the V1-V3 region was even more striking in the in vivo ISRs than in those generated in vitro. As described below, most of the in vivo and in vitro recombination within this V1-V3 region could be further mapped to the conserved C2 domain. A further exclusion of in vivo ISR crossover sites outside the V1-V5 region confirmed a significant increase in V1-V3 crossovers over V3-V5 crossovers, similar to that observed with the in vitro ISRs. By contrast, very few crossover sites from CRFs mapped to the gp120-

coding region of *env*. In fact, there appears to be an exclusion or selection against C2 crossover sites in CRFs. Crossover sites are common in the C1-V1 *env* domains of CRFs 01, 02, 04, 07, 10, 11, 13, and 14 (i.e., near the end of the Vpu open reading frame) as well as in gp41 intracellular domains of CRF 01, 02, 03, 12, and 14. Unfortunately, limited sequence diversity in the gp120 leader-C1 domain and in the transmembrane-intracellular gp41 domains hampered the design of subtype- or isolate-specific primers to detect in vitro recombination in these regions. However, the lack of C2 crossover sites in CRFs is quite apparent and is in contrast to what is observed in ISRs isolated in vivo.

Comparison of the nucleotide sequences at the sites of intersubtype recombination in vitro and in vivo. For comparison of the nucleotide sequences at the sites of intersubtype recom-

A.

Clone	Position	Region	Crossover and adjacent sequences (5' - 3')
B/E-I	1757	p24	<u>AAAAAATGGATGACAGAAACCTTGTGCA</u> <u>TCCAAAATGCGAA</u> TCCAGACTGTAAGTCCATTTTAAAGCATTAA
B/E-II	1757	p24	<u>AAAAAATGGATGACAGAAACCTTGTGCA</u> <u>TCCAAAATGCGAA</u> TCCAGACTGTAAGTCCATTTTAAAGCATTAA
B/E-III	1381	p24	TGCTAAACATAGTGGGGGACATCAGGCAGCA <u>ATGCAAATGTTAAAGA</u> GACCATCAATGAGGAAGCTGCAG
B/E-IV	1702	p24	<u>AAAAGACTATGTAGATAGTTCTATAAACTCTC</u> <u>AGAGC</u> CGAGCAAGCTTCACAGGAGGTAAAAAATGGATGA

B.

Clone	Position	Region	Crossover and adjacent sequence (5' - 3')
A/D-1	6858	C2	<u>AAAGGTATCCTTTGAA</u> <u>CCAATCCCATACATTATGTGCCCCAGCTGG</u> <u>ATTTCGATTCTAAAATGTAAC</u>
A/D-2	7096	C3	<u>AAAACCATATAGTACAACCTTACCAAGC</u> <u>CTGTAACAATTAATTG</u> <u>CACAAGGCCCTACAAGAAACAAGAATAA</u>
A/D-3	6810	V2	<u>AAATAATGGTAGTCTGTATAGATTAATAAAG</u> <u>TGTAATAC</u> <u>CACAGCCATTACACAGGCGTTC</u> <u>CAAAGGTAACCT</u>
A/D-4	7039	C2	<u>AGTCTAGCAGAAAAAGAGGTAAGAA</u> <u>TAGATCTGAAAAT</u> <u>CTCACAAATAATGCTAAAATCATAATAGTACAGC</u>
A/D-5	7360	C3	<u>AAATTACAATGCATCATTTTAATTGTA</u> <u>AAGGGGAATTTTCTACT</u> <u>GCAATACATCAGGACTGTTTAAACAACA</u>
A/D-6	7693	C5	<u>AAGTGAATTATATAAATATAAGTAGTAA</u> <u>AAATTGAACCA</u> <u>CTAGGAATAGCACCACAGGCAAGAGAAGA</u>
A/D-7	7744	C5-gp41	<u>AGAAGATGGTGG</u> <u>AGAGAGAAAAAGAGCAGTGGGAATAGGAGC</u> <u>TTTGTTCCTTGGGTTCTTGGGAGC</u>
B/E-1	6936	C2	<u>TAATGATAAGAATTTCAATGGGGCAGGG</u> <u>CCATGTA</u> <u>CAAATGTCAGCACAGTACAATGTACATGGGATTAGG</u>
B/E-2	6930	C2	<u>TTTTAAAGTGTAATGATAAGAATTTCAATGGG</u> <u>ACAGG</u> <u>ACCATGTACAAATGTCAGCACAGTACAATGAACACAT</u>
B/E-3	6906	C2	<u>TTATTGTGCTCCAGCTGGTTATGCGATTTTAAAG</u> <u>TGTAATGATAA</u> <u>AAAGTTCAATGGGACAGGTCCATGTAATAA</u>
B/E-4	6936	C2	<u>AATGATAAGAATTTCAATGGGACAGGG</u> <u>CCATGTA</u> <u>AAAAATGTCAGC</u> <u>ACAGTACAGTGACACATGGGATTAAAGC</u>
B/E-5	7794	gp41	<u>TTTGTTCCTTGGGTTCTTG</u> <u>GGAGCAGCAGGAAGCACTATGGGCGC</u> <u>GGCGTCAATAACGCTGACGGAAGG</u>
B/E-6	7619	V5	<u>TATTAACAAGAGATGGTGTCTTAATAATACTAA</u> <u>CGAGA</u> <u>TCTTCAGACCTGGAGGAGGATATGAGGGACAAT</u>
B/E-7	7652	C5	<u>ATATGAG</u> <u>GGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAC</u> <u>AAATTGAACCACTAGGAATAG</u>

FIG. 4. Mapping of intersubtype recombination sites to specific regions in *gag* or *env*. Four *gag* (A) and 14 *env* (B) B-E or A-D recombinant clones were sequenced to identify crossover sites by using RIP. The *env* sequences from 28 in vivo *env* ISRs (C) and 21 CRFs (D) were also analyzed by RIP. Although 105 crossover sites were identified (Table 2), only 11 ISR and 12 CRF crossover sequences are shown in panels C and D, respectively. The sequences shown in all panels are not aligned and are found in different regions of *env* or *gag*. Clone refers to the intersubtype *gag* or *env* clone selected in the dual infection, whereas isolate refers to the ISR found in the HIV-1 sequence database (<http://hiv-web.lanl.gov>). The HXB2 sequence was used to reference the nucleotide position of each crossover site. The crossover sites have also been localized to specific regions in the *env* and *gag* genes. Homopolymeric tracts are underlined near or within the putative crossover sequence (in italics). The thickness of the line represents the strength of the putative RT pause site, i.e., A_5 , U/T_5 , or $G_{4/5} > C_{4/5}$, G_3 , A_4 , or U/T_4 (37).

bination in vitro and in vivo, we selected and sequenced seven A-D and B-E HIV-1 *env* recombinant clones. The *env* recombinant clones described in Fig. 4B are based on the relative number of recombinants that mapped to the various *env* regions, i.e., V1-V3 (7 of the 14 recombinant *env* clones), V3-V5 (2 of 14), and V5-gp41 (5 of 14). Recombination sites in *gag* were identified only by the RIP program following DNA sequencing and were not mapped by the probe hybridization technique. The consensus sequences of various subtypes and

the RIP program were used to define the site of intersubtype recombination in the patient HIV-1 isolates (data not shown). Due to sequence heterogeneity between isolates within the same subtype, the putative crossover sequence is considerably longer and less defined in vivo than in vitro. As indicated in Fig. 2 and 3 and Table 2, the V1-V3 region was a preferred, or hot, region for intersubtype HIV-1 recombination and viral selection in both in vitro and in vivo ISRs but not CRFs. Preferential recombination in the V1-V3 *env* region was mostly

Isolate	Position	Region	Crossover sequence (5' - 3')
A/D-92UG035	6469	C1	GCCTGTGTACCAACAGACCCTAACCCACAAGAAATAATATTAGAAATGTCACAGAAAAC
C/D-91BU0908	6551	C1	ATATGGTGGATCAGATGCATGAGGATATAATCAGTTTATGGGATCAAGGCCTAAAACCAT
A/C-97TZ06	6666	V1	TCCTACTACCAATAACTGTAACTTACAAAAAGAAATGACAAATTGCTCGTCCAACA
A/C/D-95SE8603	6763	V2	CAAAGAGTATATTCACTTTTTTATAAACTTGATGTAGTACAAATTAACAGCAATCAAAAT
A/C-95IN21301	6767	V2	AGGTATATTCACTTTTTTATAGGCTTGATGTAGTACCAATTGAGGAAGGTCAAGGTAATA
A/C-SE9488	6926	C2	CGATTCTAAAGTGTAAGGATGCAAAAGTTCAATGGAATAGGGCCATGCAATAATGTCAGCA
A/C/D-95SE8603	7061	C2	AGATAATAGTTAGATCTGAAAAATCTCACAATAATGCTAAAAATCATAATAATACAGCTTA
A/C/G-93VI1035	7166	V3	CAAGAAAAAGTGATCGTATAGGACCAGGACAAACATTCTTCAACAGGAGACATAATAG
A/D-UG.C6080	7676	C5	ATATGAGAGACAAATTGGAGAAGTGAATTATACAAATATAAAGTAGTAAAAATTGAACCAC
A/C-97TZ01	8268	gp41	TTGGTTTGACATATCAAAATGGCTGTGGTATATAAAAAATTTTATAATGATAGTAGGAGG
A/C-92RW009	8294	gp41	ATATAAAAAATTTTATAATGATAGTAGGAGGCTTAATAGGTTTAAGAAATAATTTTTGCTG

Isolate	CRF	Position	Region	Crossover sequence (5' - 3')
C/D-96TZBF061	10	6414	C1	CAGATGCTAAAGCCTACAAAGCAGAGGCACATAATATCTGGGCTACACATGCCTGCGTA
cpx-94CY032	4	6428	C1	TATGAGAAAGAAGTACATAATATCTGGGCTACACATGCCTGTGTACCCACAGACCCCAA
B/C-97CN001	7	6700	V2	CACATGCCTGTGTATCCCGCAGACCCCAACCCACAAGAAATGGTTTTGGAAAAATGTAACA
cpx-CM.CA1	11	6878	C2	AGGTAACTTTGAGCCAATTTCCATACATTATTGTGCTCCAGCTGGTTATGCGATTCTA
cpx-97PVMY	4	6895	C2	AGAAGGTATATTCACTTTTTTATAGGCTTGATGTAGTACCAATTGAGGAAGGTCAAGGTT
cpx-97PVMY	4	7460	V4	AAACAAATAGTACAAACACGACAGATGAAAAATTTATCAGGCTTCCATGCAGAAATAAAA
cpx-NG3670b	11	7587	C4	AACATGTAACTCAAACATTACAGGGCTACTATTAAACAAGAGATGGTGGGAAAAATAACA
cpx-97PVMY	11	7825	gp41	TTGGGAGCAGCAGGAAGCACTATGGGCGCAGCGTCAATAACGCTGACGGTACAGGCCAG
cpx-GR17	11	7899	gp41	AAGAGATAATAGTTAGATCTGAAAAATCTCACAATAATGCTAAAAATCATAATAATACAG
B/G-ES.X397	14	8249	gp41	ATAAATGGGCAAGTTTGTGGAATTGGTTTAGCATAACAAAATGGCTGTGGTATATAAAA
D/F-VI1310	5	8344	gp41	AGAAATAGTTTTTACTGTGCTTTCCATAGTGAATAGAGTTAGGAAGGGATACTCACCTCT
B/F-99ARMA159	12	8501	gp41	TAGTGAACGGATTATTGACACTTATCTGGGACGACCTGAGGAGCCTGTGCTTTTCAGC

FIG. 4—Continued.

attributable to crossovers within the C2 domain (six of seven V1-V3 recombinants) (Fig. 4). Crossover sites were scattered throughout C2 domain, and there was no evidence of a specific hot sequence for recombination. However, several homopolymeric tracts were present in the 50-nt sequence encompassing both the in vitro and in vivo recombination sites (Fig. 4). For example, five consecutive adenosine (A) or thymidine (T) residues appear fewer than four times in this region of *env* (2,000

bp) from these four HIV-1 isolates (A-92UG029, B-HXB2, D-93UG067, and E-CMU06), and yet 6 of the 14 recombination sites were mapped to 50-nt sequences containing a tract of five A or T residues. This represents a 10- to 20-fold increase in recombination at a 50-bp sequence containing an A₅ or T₅ tract. In addition, there was also enrichment of poly(G₄) and poly(C₄) tracts at the site of recombination. HIV-1 RT has been shown to pause or dissociate at homopolymeric se-

TABLE 3. Frequency of homopolymeric tracts in consensus *env* gene sequences or in the crossover sequences of in vitro ISRs, in vivo ISRs, and CRFs

Homopolymeric tract	Consensus <i>env</i> sequence ^a		In vitro ISR ^b		In vivo ISR ^b		CRF ^b	
	Frequency/50-nt sequence	No. of 50-nt sequences	Frequency/50-nt sequence	No. of crossovers	Frequency/50-nt sequence	No. of crossovers	Frequency/50-nt sequence	No. of crossovers
GGG	0.40	204	0.83	18	0.22	60	0.34	45
CCCC/TTTT/AAAA	0.39		1.0		0.45		0.38	
≥CCCC	0		0.06		0		0	
≥GGGG	0.20		0.17		0.07		0.09	
≥TTTTT/AAAAA	0.22		0.50		0.36		0.18	

^a Homopolymeric sequences were identified in the consensus sequence of the HIV-1 *env* genes of subtypes A, B, C, and D. This analysis represents the frequency of homopolymeric tracts in 206 50-nt segments (51 *env* segments per subtype).

^b The frequency of homopolymeric tracts was analyzed in the crossover sites of 18 ISRs generated in vitro, 28 ISRs with 60 crossover sites, and 21 CRFs with 45 crossover sites. A 50-nt sequence surrounding the exact crossover sequence was selected for these analyses.

quences in the template (e.g., tracts of three or more Gs, four or more Cs, or of five or more As and Ts) (2, 37). RT pausing or dissociation has also been shown to enhance the efficiency of template switching and subsequent recombination (20).

We also examined the frequency of homopolymeric tracts in 60 ISR and 45 CRF crossover sites defined by a 50-nt sequence in *env* (Table 3). As described above, all CRF crossover sites were distinct in the 11 classified forms used in this study (see Materials and Methods). For comparison, we calculated the frequency of homopolymeric tracts (per 50 nt) in the consensus *env* gene sequences of subtypes A, B, C, and D (204 50-nt segments). With exception of the poly(G_{>4}) tract, there appears to be at least a two- to threefold enrichment of in vitro ISR crossovers adjacent to all homopolymeric tracts shown to promote RT pausing and template switching. There is no enrichment of homopolymeric tracts in the CRF crossover sites. However, in vivo ISR crossovers appear again to be an intermediate of a recent recombination event (e.g., in vitro ISRs) and selection of the stable CRF crossover sites. This observation is most apparent when comparing the frequencies of the strong RT pause-dissociation site at the crossover sites (≥A₅ or T₅). In the HIV-1 *env* genes of subtypes A, B, C, and D, five or more A or T residues appear in approximately 22% of every 50-nt *env* sequence, whereas 50 and 36% of all in vitro and in vivo ISR crossover sites (respectively) contain an A_{>5} or T_{>5} tract (Table 3). The frequency of these A and T tracts in CRF crossovers is approximately 18%. It is important to note that not all in vitro and in vivo ISRs contain homopolymeric tracts (Fig. 4C and D). In these instances it is possible that other factors such as RNA secondary structure may induce template switching.

DISCUSSION

The implications of retroviral recombination have come to light only with the recent advances in HIV molecular epidemiology (47, 50, 56, 57). Recombination is now considered a hallmark feature of retroviruses and has been the subject of hundreds of studies. Partial and full genome sequencing of HIV-1 isolates from around the world clearly indicates that at least 20% are chimeras of different HIV-1 subtypes or clades (13, 14, 18, 40, 43, 46, 47, 50, 52, 57, 60, 62). Recombination between HIV-1 subtypes or groups can result in major antigenic shifts and hamper both vaccine and drug development

(50). At another level, rapid inpatient evolution may be driven by recombination between quasispecies and promote changes in tissue-cell tropism, immune avoidance, and multi-drug resistance (17, 24, 26, 28, 36, 45, 50; Butto et al., letter). However, the production and selection of an ISR are difficult to examine in vivo or in an animal model.

We have adopted a primary tissue culture system to establish dual infections in PBMC with primary HIV-1 isolates and to screen for intersubtype HIV-1 recombination (51). In a previous study, we employed an HTA to measure dual virus production and to determine the relative fitnesses of primary HIV-1 isolates in these competitions (51). Subsequent analyses revealed that the frequency of intersubtype recombination was approximately 3 to 5%/1 kbp, or below the limits of a direct HTA screen (Arts et al., submitted). This study focuses on mapping, analyzing, and then comparing the crossover sites found in intersubtype HIV-1 *env* (or *gag*) recombinants generated in vitro to those previously identified in vivo. Intersubtype HIV-1 recombinants were selectively amplified with subtype-specific primers and then identified by HTA. These recombinant *env* products were then cloned and analyzed by probe hybridization assays and DNA sequencing. A rough mapping of recombination sites in *env* revealed similar in vitro and in vivo ISR preferences for crossovers in the V2-C3 region. As described below, preferential sites of recombination are difficult to identify in single-cycle recombination studies employing defective retroviral vectors (1, 23, 30, 71). We suspect that increased recombination in the C2 region is likely due to (i) higher conservation of this sequence compared to most of the other conserved *env* domains (i.e., C1, C3, C4, and C5) and (ii) selection of ISRs with proper *env* function following recombination in this C2 region versus other *env* regions.

Our studies on ex vivo HIV-1 fitness and recombination were planned soon after the initial identification of intersubtype recombinants (56, 57). In 1997, HIV-1 subtypes were defined by thousands of HIV-1 sequences from around the world, but few isolates were classified as ISRs (40). Recombination is now evident between nearly every HIV-1 subtype and group (47, 50). Many of these ISRs emerged shortly after the divergence of group M subtypes and now exist as CRFs of HIV-1 (8, 29, 31, 47, 55, 57, 64, 73; C. Montavon, V. Nicole, L. Vergne, K. Toure, E. Esu-Williams M. Souleymane, N. Nzilambi, M. Eitel, P. Delphine, L. Florian, M. Claire, S. Saragosti, E. Delaporte, and M. Peeters, 6th Annu. Int. Disc. Meet. HIV

Dynamics Evol., 1999). In addition, increased travel, trade, migration, tourism, and wars have resulted in cocirculation of multiple subtypes outside Africa. For example, the Brazilian population now harbors stable and cocirculating subtype B, C, and F epidemics (5). The early predominance of subtypes B and E in Southeast Asia or subtype A in southern Africa has been overwhelmed by the rapid emergence of subtype C infections (58). This study has developed a model of intersubtype recombination in vitro to compare with in vivo HIV-1 recombination and to identify sites for stable recombination and subsequent ISR survival. We utilized a total of nine HIV-1 isolates of different subtypes and performed seven dual infections to screen for intersubtype recombination. Remarkably, the patterns of ISR recombination between different pairs of HIV-1 isolates were quite similar regardless of the *env* subtype or sequence diversity between the pairs. The large evolutionary jumps associated with these intersubtype recombination events will undoubtedly affect the HIV-1 epidemic. Selection pressures on ISRs in the human population are likely complex and involve efficiency of transmission, variations in host pressure (immune response and genetic factors), and replication efficiency. Thus, CRF recombinants will emerge from this heterogeneous pool of ISRs. CRF selection is likely based on fitness in the human population rather than the frequency of ISRs with favored crossover sites.

Retroviral recombination is preceded by the production of heterodiploid virus from a coinfecting cell. The frequency of recombination is dependent on multiple factors that affect virus replication (Arts et al., submitted). However, the actual recombination event is controlled solely by the ability of RT to jump between plus-strand RNA or minus-strand DNA templates (23, 30). Aside from the first- and second-strand switch that occurs during every round of retroviral reverse transcription, there is little evidence of selective recombination at other regions in the retroviral genome. Recent studies suggest that the dimer linkage and/or dimer initiation sequence is responsible for genomic RNA dimerization as well as for promoting a strand transfer event during reverse transcription (15, 44). Our study clearly indicates that the C2 domain of *env* is a hot region for both in vitro and in vivo intersubtype recombination and further selection of replication-competent viruses. However, specific crossover sites were not found throughout the C2 region. When the *env* genes of different HIV-1 subtypes are compared, the variability in C2 is slightly less than that in other conserved domains (with the exception of the short C5 domain). On the other hand, a twofold increase in C2 versus C4 recombination may not be attributable to a 2 to 7% difference in nucleotide sequence identity. Previous studies employing in vitro assays or defective retroviruses in single-cycle infections clearly showed that increased sequence homology between the two genomic templates resulted in a higher frequency of recombination (41, 72). However, we have actually observed higher frequencies of recombination in the variable *env* gene than in the more conserved *gag* gene (Arts et al., submitted). Evidence of a preferred recombination site within *env* (e.g., C2) and the latter result are contrary to those previously published (41, 72). This discrepancy is likely due to variations in the assays rather than differences in the mechanism of recombination. Unlike in earlier studies employing defective retroviral particles and packaging cell lines (1, 23, 30, 71), recom-

bination in our assay was not limited to a single replication cycle. Replication of both parental and recombined HIV-1 in these dual infections would lead to a rapid selection of replication-competent viruses (Arts et al., submitted).

Most of the HIV-1 ISRs generated in vitro (or within a coinfecting individual) are likely defective or dead and would not survive a 5- to 15-day competition with the parental isolates. Surprisingly, the majority of HIV-1 *env* recombinants that survive this selection and competition contain a crossover in the C2 domain and not in the hypervariable regions. In contrast to the hot "region" of intersubtype *env* recombination and viral selection, there was no defined hot "spot" for the actual crossovers or breakpoints. This may be due in part to the extreme sequence variability among the different pairs of HIV-1 isolates used in the various dual infections. In addition, any sequence specificity for strand displacements and crossover events may be difficult to identify after multiple rounds of infection. It is important to stress that this selection of replication-competent ISRs is likely unrelated to the mechanisms involved in strand displacement and template switching (i.e., recombination) (20). Therefore, it was somewhat surprising that a majority of in vitro and in vivo ISRs contained runs of homopolymeric sequences within or adjacent to crossover sites. A 50-nt scan across the entire *env* gene suggests an enhancement of recombination near poly(A₅) or poly(U/T₅) tracts. As described above, very few ISRs evolved into CRFs. Thus, the nucleotide composition of crossover sites in CRFs resembles that observed throughout the *env* gene; i.e., there is no enrichment of homopolymeric tracts.

Although this preference for homopolymeric tracts may be maintained in replication-competent ISRs, the homopolymeric regions are never located at a defined site within or adjacent to the crossover sequence. Several studies have now shown that HIV-1 RT pauses or even dissociates at homopolymeric sequences in the template (2, 37). The strength of the pause site is dependent on the length and base comprising these tracts; i.e., a G₄ tract induces more RT pausing than a C₄, A₄, or U/T₄ tract (37). A homopolymeric sequence of five or more A or U/T residues induces minor groove compressions, a bend in the nucleic acid duplex (38), and a termination of DNA synthesis from an RNA or DNA template by HIV-1 RT. In fact, plus-strand DNA synthesis from the central polypurine tract primer is terminated at the center of the HIV-1 genome by a phased poly(A)-poly(T) tract known as the central termination sequence (9, 65). The relationship between pausing or termination during DNA synthesis and template switching or recombination has been the focus of several studies and is the basis of three models for retroviral recombination (10, 11, 20, 33). Dissociation of the RT-primer-template complex at homopolymeric tracts or regions of RNA secondary structure increases the possibility that RT and the primer can reassociate with the other template in the diploid virus. Even though there appear to be stable G-C-rich duplexes or stem-loops at some crossover sites, RNA secondary structures are difficult to predict and are not static during reverse transcription.

In a future study, single-cycle, in vitro reverse transcription-recombination assays will be used to examine the possible sequence specificity for crossover sites [e.g., poly(A) or poly(U/T) tracts] and in the absence of selection for replication-competent ISRs (e.g., in the C2 region). Although RT

pausing at homopolymeric sequences or regions of RNA secondary structure increases the frequency of template switching, the displaced strand and RT may not always anneal to the same sequence on the alternative template. In addition, sequence variability between the two templates in the heterodiploid retrovirus may also result in heterogeneous breakpoints, i.e., originating from the same pause site. Most of these crossover events will result in defective or dead HIV-1 particles, but a few particles will be capable of replicating and competing with the parental HIV-1 isolates. This may explain the variable location or homopolymeric stretches within a crossover sequence. Certain domains in *env* (e.g., C2) may accommodate these major genetic shifts better than other regions and lead to production of functional glycoproteins. Thus, the possible interplay between (i) preferential crossovers at pause or termination sites during HIV-1 DNA synthesis and (ii) subsequent selection of replication-competent viruses appears to influence the generation of ISRs both in vitro and in the HIV-1 epidemic.

ACKNOWLEDGMENTS

E.J.A. was supported by research grants from the National Institute of Child Health and Development, NIH (NO1-HD-0-3310-502-02), and from the National Institute of Allergy and Infectious Diseases, NIH (AI49170). M.E.Q.-M. was supported by a Pulmonary Pathogen Defense Mechanisms training grant from the National Heart, Lung, and Blood Institute, NIH (HL07889). Support was also provided by the cores of the NIH Center for AIDS Research (AI36219) at Case Western Reserve University.

REFERENCES

- 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.
- Arts, E. J., J. P. Marois, Z. Gu, S. F. Le Grice, and M. A. Wainberg. 1996. Effects of 3'-deoxynucleoside 5'-triphosphate concentrations on chain termination by nucleoside analogs during human immunodeficiency virus type 1 reverse transcription of minus-strand strong-stop DNA. *J. Virol.* **70**:712–720.
- Barker, E., K. N. Bossart, C. P. Locher, B. K. Patterson, and J. A. Levy. 1996. CD8+ cells from asymptomatic human immunodeficiency virus-infected individuals suppress superinfection of their peripheral blood mononuclear cells. *J. Gen. Virol.* **77**:2953–2962.
- 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.
- Bongertz, V., D. C. Bou-Habib, L. F. Brigido, M. Caseiro, P. J. Chequer, J. C. Couto-Fernandez, P. C. Ferreira, B. Galvao-Castro, D. Greco, M. L. Guimaraes, M. I. Linhares de Carvalho, M. G. Morgado, C. A. Oliveira, S. Osmanov, C. A. Ramos, M. Rossini, E. Sabino, A. Tanuri, M. Ueda, et al. 2000. HIV-1 diversity in Brazil: genetic, biologic, and immunologic characterization of HIV-1 strains in three potential HIV vaccine evaluation sites. *J. Acquir. Immune Defic. Syndr.* **23**:184–193.
- Bradley, R. D., and D. M. Hillis. 1997. Recombinant DNA sequences generated by PCR amplification. *Mol. Biol. Evol.* **14**:592–593.
- Bruce, C., C. Clegg, A. Featherstone, J. Smith, B. Biryahawaho, R. Downing, and J. Oram. 1994. Presence of multiple genetic subtypes of human immunodeficiency virus type 1 proviruses in Uganda. *AIDS Res. Hum. Retroviruses* **10**:1543–1550.
- Carr, J. K., M. O. Salminen, C. Koch, D. Gotte, A. W. Arntstein, P. A. Hegerich, D. St. Louis, D. S. Burke, and F. E. McCutchan. 1996. Full-length sequence and mosaic structure of a human immunodeficiency virus type 1 isolate from Thailand. *J. Virol.* **70**:5935–5943.
- Charneau, P., M. Alizon, and F. Clavel. 1992. A second origin of DNA plus-strand synthesis is required for optimal human immunodeficiency virus replication. *J. Virol.* **66**:2814–2820.
- Coffin, J. M. 1979. Structure, replication, and recombination of retrovirus genomes: some unifying hypotheses. *J. Gen. Virol.* **42**:1–26.
- Coffin, J. M. 1990. Genetic variation in retroviruses, p. 11–33. *In* E. Kurstak, R. G. Marusyk, R. A. Murphy, and M. H. V. Van Regenmortel (ed.), *Applied virology research*. Plenum Publishing, New York, N.Y.
- Collins, K. R., H. Mayanja-Kizza, B. A. Sullivan, M. E. Quinones-Mateu, Z. Toossi, and E. J. Arts. 2000. Greater diversity of HIV-1 quasiespecies in HIV-infected individuals with active tuberculosis. *J. Acquir. Immune Defic. Syndr.* **24**:408–417.
- Cornelissen, M., G. Kampinga, F. Zorgdrager, J. Goudsmit, and the UNAIDS Network for HIV Isolation and Characterization. 1996. Human immunodeficiency virus type 1 subtypes defined by *env* show high frequency or recombinant *gag* genes. *J. Virol.* **70**:8209–8212.
- Cornelissen, M., R. van den Burg, F. Zorgdrager, V. Lukashov, and J. Goudsmit. 1997. *pol* gene diversity of five human immunodeficiency virus type 1 subtypes: evidence for naturally occurring mutations that contribute to drug resistance, limited recombination patterns, and common ancestry for subtypes B and D. *J. Virol.* **71**:6348–6358.
- Das, A. T., B. Klaver, and B. Berkhout. 1998. The 5' and 3' TAR elements of human immunodeficiency virus exert effects at several points in the virus life cycle. *J. Virol.* **72**:9217–9223.
- Delwart, E. L., E. G. Shpaer, J. Louwagie, F. E. McCutchan, M. Grez, H. Rübsamen-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.
- Diaz, R. S., E. C. Sabino, A. Mayer, J. W. Mosley, M. P. Busch, and the Transfusion Safety Study Group. 1995. Dual human immunodeficiency virus type 1 infection and recombination in a dually exposed transfusion recipient. *J. Virol.* **69**:3273–3281.
- Douglas, N. W., A. I. Knight, A. Hayhurst, W. Y. Barrett, M. J. Kevany, and R. S. Daniels. 1996. An efficient method for the rescue and analysis of functional HIV-1 *env* genes: evidence for recombination in the vicinity of the *tat/rev* splice site. *AIDS* **10**:39–46.
- Emini, E. A., and S. D. Putney. 1992. Human immunodeficiency virus. *Bio/Technology* **20**:309–326.
- Figlerowicz, M., and A. Bibillo. 2000. RNA motifs mediating in vivo site-specific nonhomologous recombination in (+) RNA virus enforce in vitro nonhomologous crossovers with HIV-1 reverse transcriptase. *RNA* **6**:339–351.
- 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.
- Gao, F., L. Yue, S. Craig, C. L. Thornton, D. L. Robertson, F. E. McCutchan, J. A. Bradac, P. M. Sharp, B. H. Hahn, and W. H. O. Network for HIV Isolation and Characterization. 1994. Genetic variation of HIV type 1 in four World Health Organization-sponsored vaccine evaluation sites: generation of functional envelope (glycoprotein 160) clones representative of sequence subtypes A, B, C, and E. *AIDS Res. Hum. Retrovir.* **10**:1359–1368.
- Goodrich, D. W., and P. H. Duesberg. 1990. Retroviral recombination during reverse transcription. *Proc. Natl. Acad. Sci. USA* **87**:2052–2056.
- Groenink, M., A. C. Andeweg, R. A. M. Fouchier, S. Broersen, R. C. M. van der Jagt, H. Schuitemaker, R. E. Y. de Goede, M. L. Bosh, H. G. Huisman, and M. Tersmette. 1992. Phenotype-associated *env* gene variation among eight related human immunodeficiency virus type 1 clones: evidence for in vivo recombination and determinants of cytotropism outside the V3 domain. *J. Virol.* **66**:6175–6180.
- Hahn, B. H., M. A. Gonda, G. M. Shaw, M. Popovic, J. A. Hoxie, R. C. Gallo, and F. Wong-Staal. 1985. Genomic diversity of the acquired immune deficiency syndrome virus HTLV-III: different viruses exhibit greatest divergence in their envelope genes. *Proc. Natl. Acad. Sci. USA* **82**:4813–4817.
- Halapi, E., T. Leitner, M. Jansson, G. Scarlatti, P. Orlandi, A. Plebani, L. Romiti, J. Albert, H. Wigzell, and P. Rossi. 1997. Correlation between HIV sequence evolution, specific immune response and clinical outcome in vertically infected infants. *AIDS* **11**:1709–1717.
- Ho, D. D., A. U. Neumann, A. S. Perelson, W. Chen, J. M. Leonard, and M. Markowitz. 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* **373**:123–126.
- Howell, R. M., J. E. Fitzgibbon, M. Noe, Z. Ren, D. J. Gocke, T. A. Schwartz, and D. T. Dubin. 1991. *In vivo* sequence variation of the human immunodeficiency virus type 1 *env* gene: evidence for recombination among variants found in a single individual. *AIDS Res. Hum. Retrovir.* **7**:869–876.
- Hoxie, J. A., J. D. Alpers, J. L. Rackowski, K. Huebner, B. S. Haggarty, A. J. Cedarbaum, and J. C. Reed. 1986. Alterations in T4 (CD4) protein and mRNA synthesis in cells infected with HIV. *Science* **234**:1123–1127.
- 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.
- Janini, L. M., A. Tanuri, M. Schechter, J. M. Peralta, A. C. Vicente, T. N. Dela, N. J. Pieniazek, C. C. Luo, A. Ramos, V. Soriano, G. Schochetman, M. A. Rayfield, and D. Pieniazek. 1998. Horizontal and vertical transmission of human immunodeficiency virus type 1 dual infections caused by viruses of subtypes B and C. *J. Infect. Dis.* **177**:227–231.
- 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.

33. Junghans, R. P., L. R. Boone, and A. M. Skalka. 1982. Retroviral DNA H structures: displacement-assimilation model of recombination. *Cell* **30**:53–62.
34. Kannagi, M., T. Masuda, T. Hattori, T. Kanoh, K. Nasu, N. Yamamoto, and S. Harada. 1990. Interference with human immunodeficiency virus (HIV) replication by CD8⁺ T cells in peripheral blood leukocytes of asymptomatic HIV carriers in vitro. *J. Virol.* **64**:3399–3406.
35. Kawai, S., and H. Hanafusa. 1972. Genetic recombination with avian tumor virus. *Virology* **49**:37–44.
36. Kellam, P., and B. A. Larder. 1995. Retroviral recombination can lead to linkage of reverse transcriptase mutations that confer increased zidovudine resistance. *J. Virol.* **69**:669–674.
37. Klarmann, G. J., C. A. Schaubert, and B. D. Preston. 1993. Template-directed pausing of DNA synthesis by HIV-1 reverse transcriptase during polymerization of HIV-1 sequences in vitro. *J. Biol. Chem.* **268**:9793–9802.
38. Koo, H. S., J. Drak, J. A. Rice, and D. M. Crothers. 1990. Determination of the extent of DNA bending by an adenine-thymine tract. *Biochemistry* **29**:4227–4234.
39. Korber, B., B. H. Hahn, B. Foley, J. Mellors, T. Leitner, G. Myers, F. E. McCutchan, and C. Kuiken. 1997. Human retroviruses and AIDS 1997. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, N.Mex.
40. Kuiken, C. L., B. Foley, B. Hahn, B. Korber, F. McCutchan, P. Marx, J. Mellors, J. I. Mullins, J. Sodroski, and S. Wolinsky. 1999. Human retroviruses and AIDS 1999: a compilation and analysis of nucleic acid and amino acid sequences. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, N.Mex.
41. 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.
42. Linial, M., and S. Brown. 1979. High-frequency recombination within the *gag* gene of Rous sarcoma virus. *J. Virol.* **31**:257–260.
43. Lole, K. S., R. C. Bollinger, R. S. Paranjape, D. Gadkari, S. S. Kulkarni, N. G. Novak, R. Ingersoll, H. W. Sheppard, and S. C. Ray. 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J. Virol.* **73**:152–160.
44. Mikkelsen, J. G., A. H. Lund, M. Duch, and F. S. Pedersen. 2000. Mutations of the kissing-loop dimerization sequence influence the site specificity of murine leukemia virus recombination in vivo. *J. Virol.* **74**:600–610.
45. Moutouh, L., J. Corbeil, and D. D. Richman. 1996. Recombination leads to the rapid emergence of HIV-1 dually resistant mutants under selective drug pressure. *Proc. Natl. Acad. Sci. USA* **93**:6106–6111.
46. Neilson, J. R., G. C. John, J. K. Carr, P. Lewis, J. K. Kreiss, S. Jackson, R. W. Nduati, D. Mbori-Ngacha, D. D. Panteloeff, S. Bodrug, C. Giachetti, M. A. Bott, B. A. Richardson, J. Bwayo, J. Ndinya-Achola, and J. Overbaugh. 1999. Subtypes of human immunodeficiency virus type 1 and disease stage among women in Nairobi, Kenya. *J. Virol.* **73**:4393–4403.
47. Peeters, M. 2000. Recombinant HIV sequences: their role in the global epidemic, p. 1–39–I–54. In C. Kuiken, B. Foley, B. Hahn, B. Korber, F. McCutchan, P. Marx, J. Mellors, J. I. Mullins, J. Sodroski, and S. Wolinsky (ed.), *Human retroviruses and AIDS 2000*. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, N.Mex.
48. Preston, B. D., B. J. Poiesz, and L. A. Loeb. 1988. Fidelity of HIV-1 reverse transcriptase. *Science* **242**:1168–1171.
49. Quinones-Mateu, M. E., J. L. Albright, A. Mas, V. Soriano, and E. J. Arts. 1998. Analysis of *pol* gene heterogeneity, viral quasispecies, and drug resistance in individuals infected with group O strains of human immunodeficiency virus type 1. *J. Virol.* **72**:9002–9015.
50. Quinones-Mateu, M. E., and E. J. Arts. 1999. Recombination in HIV-1: update and implications. *AIDS Rev.* **1**:89–100.
51. Quinones-Mateu, M. E., S. C. Ball, A. J. Marozsan, V. S. Torre, J. L. Albright, G. Vanham, G. G. van der, R. L. Colebunders, and E. J. Arts. 2000. A dual infection-competition assay shows a correlation between ex vivo human immunodeficiency virus type 1 fitness and disease progression. *J. Virol.* **74**:9222–9233.
52. Ramos, A., A. Tanuri, M. Schechter, M. A. Rayfield, D. J. Hu, M. C. Cabral, C. I. Bandea, J. Baggs, and D. Pieniazek. 1999. Dual and recombinant infections: an integral part of the HIV-1 epidemic in Brazil. *Emerg. Infect. Dis.* **5**:65–74.
53. Reed, L. J., and H. Muench. 1938. A simple method of estimating fifty percent endpoints. *Am. J. Hyg.* **27**:493–497.
54. 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. Retrovir.* **14**:635–638.
55. Robertson, D. L., J. P. Anderson, J. A. Bradac, J. K. Carr, B. Foley, R. K. Funkhouser, F. Gao, B. H. Hahn, M. L. Kalish, C. Kuiken, G. H. Learn, T. Leitner, F. McCutchan, S. Osmanov, M. Peeters, D. Pieniazek, M. Salminen, P. M. Sharp, S. Wolinsky, and B. Korber. 2000. HIV-1 nomenclature proposal. *Science* **288**:55–56.
56. Robertson, D. L., B. H. Hahn, and P. M. Sharp. 1995. Recombination in AIDS viruses. *J. Mol. Evol.* **40**:249–259.
57. Robertson, D. L., P. M. Sharp, F. E. McCutchan, and B. H. Hahn. 1995. Recombination in HIV-1. *Nature* **374**:124–126.
58. Rodenburg, C. M., Y. Li, S. A. Trask, Y. Chen, J. Decker, D. L. Robertson, M. L. Kalish, G. M. Shaw, S. Allen, B. H. Hahn, and F. Gao. 2001. Near full-length clones and reference sequences for subtype C isolates of HIV type 1 from three different continents. *AIDS Res. Hum. Retrovir.* **17**:161–168.
59. Saag, M. S., B. H. Hahn, J. Gibbons, Y. Li, E. S. Parks, W. P. Parks, and G. M. Shaw. 1988. Extensive variation of human immunodeficiency virus type-1 *in vivo*. *Nature* **334**:440–444.
60. 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.
61. Sanders-Buell, E., M. Salminen, and F. McCutchan. 1995. Sequencing primers for HIV-1, p. III-15–III-21. In B. Korber, B. D. Walker, J. P. Moore, G. Myers, C. Brander, R. A. Koup, and B. Haynes (ed.), *Human retroviruses and AIDS*. Theoretical Biology and Biophysics, Los Alamos, N.Mex.
62. Sherefa, K., B. Johansson, M. Salminen, and A. Sonnerborg. 1998. Full-length sequence of human immunodeficiency virus type 1 subtype A, recombined with subtype C in the env V3 domain. *AIDS Res. Hum. Retrovir.* **14**:289–292.
63. Siepel, A. C., A. L. Halpern, C. Macken, and B. T. M. Korber. 1995. A computer program designed to screen rapidly for HIV type 1 intersubtype recombinant sequences. *AIDS Res. Hum. Retrovir.* **11**:1413–1416.
64. Soriano, V., M. Gutierrez, G. Garcia-Lerma, O. Aguilera, A. Mas, R. Bravo, M. L. Pérez-Labad, M. Baquero, and J. González-Lahoz. 1996. First case of HIV-1 group O infection in Spain. *Vox Sang* **71**:66.
65. Stetor, S. R., J. W. Rausch, M. J. Guo, J. P. Burnham, L. R. Boone, M. J. Waring, and S. F. Le Grice. 1999. Characterization of (+) strand initiation and termination sequences located at the center of the equine infectious anemia virus genome. *Biochemistry* **38**:3656–3667.
66. 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.
67. Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**:4673–4680.
68. Tovanabutra, S., V. Polonis, M. De Souza, R. Trichavaroj, P. Chanbancherd, B. Kim, E. Sanders-Buell, S. Nitayaphan, A. Brown, M. R. Robb, D. L. Birx, F. E. McCutchan, and J. K. Carr. 2001. First CRF01_AE/B recombinant of HIV-1 is found in Thailand. *AIDS* **15**:1063–1065.
69. Vogt, P. K. 1971. Genetically stable reassortment of markers during mixed infection with avian tumor viruses. *Virology* **46**:947–952.
70. Wei, X., S. K. Ghosh, M. E. Taylor, V. A. Johnson, E. A. Emini, P. Deutsch, J. D. Lifson, S. Bonhoeffer, M. A. Nowak, B. H. Hahn, M. S. Saag, and G. M. Shaw. 1995. Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* **373**:117–122.
71. Yin, P. D., V. K. Pathak, A. E. Rowan, R. J. Teufel, and W. S. Hu. 1997. Utilization of nonhomologous minus-strand DNA transfer to generate recombinant retroviruses. *J. Virol.* **71**:2487–2494.
72. Zhang, J., and H. M. Temin. 1993. Rate and mechanism of nonhomologous recombination during a single cycle of retroviral replication. *Science* **259**:234–238.
73. 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.