View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Harvard University - DASH

DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY
Office for Scholarly Communication

Genomic Sites of Human Immunodeficiency Virus Type 2 (HIV-2) Integration: Similarities to HIV-1 In Vitro and Possible Differences In Vivo

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation	MacNeil, A., J.-L. Sankale, S. T. Meloni, A. D. Sarr, S. Mboup, and P. Kanki. 2006. "Genomic Sites of Human Immunodeficiency Virus Type 2 (HIV-2) Integration: Similarities to HIV-1 In Vitro and Possible Differences In Vivo." Journal of Virology 80 (15) (July 12): 7316–7321. doi:10.1128/jvi.00604-06.
Published Version	doi:10.1128/JVI.00604-06
Citable link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:26879442
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

JOURNAL OF VIROLOGY, Aug. 2006, p. 7316–7321 Vol. 80, No. 15
0022-538X/06/$08.00+0 doi:10.1128/JVI.00604-06
Copyright © 2006, American Society for Microbiology. All Rights Reserved.

Genomic Sites of Human Immunodeficiency Virus Type 2 (HIV-2) Integration: Similarities to HIV-1 In Vitro and Possible Differences In Vivo

Adam MacNeil,[1] Jean-Louis Sankalé,[1] Seema Thakore Meloni,[1] Abdoulaye Dieng Sarr,[1] Souleymane Mboup,[2] and Phyllis Kanki[1]*

Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, Massachusetts,[1] and Laboratoire de Bacteriologie et Virologie, Universite Cheikh Anta Diop, Dakar, Senegal[2]

Received 24 March 2006/Accepted 19 May 2006

Retroviruses have distinct preferences in integration site selection in the host cell genome during in vitro infection, with human immunodeficiency virus type 1 (HIV-1) integration strongly favoring transcriptional units. Additionally, studies with HIV-1 have shown that the genomic site of proviral integration may impact viral replication, with integration in heterochromatin associated with a block in viral transcription. HIV-2 is less pathogenic than HIV-1 and is believed to have a lower replication rate in vivo. Although differences in integration site selection between HIV-2 and HIV-1 could potentially explain the attenuated pathogenicity of HIV-2, no studies have characterized integration site selection by HIV-2. In this study, we mapped 202 HIV-2 integration sites during in vitro infection of peripheral blood mononuclear cells with a primary HIV-2 isolate. In addition, we assayed for in vivo proviral integration within heterochromatin in 21 HIV-1-infected subjects and 23 HIV-2-infected subjects, using an alphoid repeat PCR assay. During in vitro infection, HIV-2 displayed integration site preferences similar to those previously reported for HIV-1. Notably, 82% of HIV-2 integrations mapped to Refseq genes, and integration strongly favored regions of the genome with high gene density and high GC content. Though rare, the proportion of HIV-2 subjects with evidence of proviral integration within heterochromatin in vivo was higher than that of HIV-1-infected subjects. It is therefore possible that integration site selection may play a role in the differences in HIV-1 and HIV-2 in vivo pathogenesis.

Integration of proviral DNA into the host cell genome is catalyzed by the viral integrase protein and is necessary for replication of retroviruses (14). Although integration site selection does not occur in a sequence-specific manner, recent studies have indicated that retrovirus integration favors palindromic sequences and that distinct genomic sequences are favored in a retrovirus species-specific manner (15, 42). At the genomic level, it is now clear that retroviruses have discrete preferences in integration site selection. For instance, studies have shown that during in vitro infection, human immunodeficiency virus type 1 (HIV-1) has a preference for proviral integration into coding regions of the genome and in actively transcribed host genes (27, 34). Simian immunodeficiency virus (SIV) has been shown to have a similar preference for integration site selection, both in cell lines and in hematopoietic stem cells of rhesus monkeys (10, 13). In contrast, murine leukemia virus (MLV) integration has a preference for integration near the start of transcriptional units and in close proximity to CpG islands (27, 41). Finally, integration site selection for avian sarcoma-leukosis virus (ASLV) has been shown to be the closest to random, with only a minor preference for integration in transcriptional units during infection in both human cell lines (27, 28) and chicken embryo fibroblasts (4).

* Corresponding author. Mailing address: Department of Immunology and Infectious Diseases, Harvard School of Public Health, 651 Huntington Ave., Boston, MA 02115. Phone: (617) 432-1267. Fax: (617) 432-3575. E-mail: pkanki@hsph.harvard.edu.

Human immunodeficiency virus type 2 (HIV-2) is a lentivirus closely related to SIV and more distantly related to HIV-1. Like HIV-1, HIV-2 causes AIDS in humans. However, disease progression occurs much more slowly in HIV-2 infection, and spread of the virus is primarily limited to West Africa (19, 25, 26, 40). Previous studies have shown that proviral loads are similar between people infected with HIV-1 and those with HIV-2 (5, 29, 30). In contrast, plasma viral loads are significantly higher in people infected with HIV-1 (1, 2, 30, 31, 36). Based on these findings, it has been suggested that viral replication rates are lower in people infected with HIV-2 than in people infected with HIV-1 (30).

In people infected with HIV-1, a subset of infected cells have integrated proviral genomes that fail to undergo replication (23). Studies in vitro have shown that a small proportion of cells infected by HIV-1 develop a similar latent state, with the absence of transcription driven from the viral promoter in nonstimulated cells (16, 17). Importantly, within these cells, viral integration occurs commonly within or close to alphoid repeats in heterochromatin. In contrast, integration in heterochromatin is disfavored in cells that had active transcription from the viral promoter, suggesting that the genomic site of proviral integration may affect viral replication (16). Similarly, a recent integration mapping study found that the genomic site of HIV-1 integration has a major effect on viral transcription, and low viral transcription is associated with integration in gene deserts, centromeric heterochromatin, and very highly expressed genes (24).

These studies underscore the effect that integration site selection has on retroviral replication. No studies to date have

examined integration site selection by HIV-2. In order to characterize integration site selection by HIV-2, we infected primary peripheral blood mononuclear cells (PBMCs) with a primary HIV-2 isolate and mapped 202 integration sites within the human genome. To examine if integration site selection may account for differences between HIV-1 and HIV-2 in in vivo viral replication, we assayed for the presence of proviral integration within heterochromatin in PBMCs of people infected with HIV-1 and people infected with HIV-2.

MATERIALS AND METHODS

Sample acquisition. PBMC samples were obtained from a cohort of registered female sex workers in Dakar, Senegal, that have been followed since 1985. Informed consent was obtained for all study participants. Epidemiologic and clinical aspects of this cohort have been described previously (18). $CD4^+$ T-cell counts were determined, and serum samples were diagnosed for HIV-1- and HIV-2-specific antibodies as described previously (18). DNA was extracted from PBMC samples (Blood and Cell Culture DNA Midi kit; QIAGEN, Valencia, CA) and resuspended in 125 μl H_20. DNA concentrations were determined by optical density readings at 260 nm.

In vitro viral infection. The HIV-2 viral strain used in this study (p1629) was previously isolated from a subject from our female sex worker cohort (18). For infection, PBMCs from a normal U.S. blood donor were separated and stimulated for 72 h with 5 μg/ml phytohemagglutinin in complete interleukin 2 (IL-2) medium (RPMI medium 1640 supplemented with 20% [vol/vol] fetal bovine serum, 1% antibiotics, and 50 units/ml IL-2), as previously described (21). Phytohemagglutinin was removed, and cells were infected with p1629 at a high multiplicity of infection. Twenty-four hours postinfection, cells were washed to remove residual supernatant virus and resuspended in fresh complete IL-2 medium.

Sequencing of HIV-2 integration sites. A library of DNA fragments containing the terminal 5′ end of integrated HIV-2 long terminal repeat (LTR) and upstream human genomic DNA corresponding to the integration site junction was created by linker-mediated nested PCR, using the Universal Genome Walker kit (BD Biosciences Clontech, Palo Alto, CA), according to the manufacturer's recommendations, with modifications described below. Two micrograms of DNA extracted 7 days postinfection was digested with DraI in a 20-μl reaction at 37°C for 12 h and inactivated by incubation at 65°C for 20 min. Digested DNA was ligated to an adaptor provided by the manufacturer in a 40-μl reaction at 16°C for 12 h and inactivated by incubation at 70°C for 5 min. Sequences specific to the viral integration site junction were amplified in a nested PCR using the adaptor-specific primers AP1 and AP2 provided by the manufacturer and HIV-2-specific primers link1 (TCCTGCCGCCCTTACTGCCTTCACTCA) and link2 (GGTA CCTTACTCCTGGCCCATGAGTATA) for the first- and second-round PCRs, respectively, using commercially available PCR reagents (BD Advantage 2 PCR kit; BD Biosciences Clontech, Palo Alto, CA). The first-round PCR was performed at a volume of 100 μl, and 2 μl of the first round product was used in the second-round PCR in a volume of 50 μl. Both PCRs were performed under the following reaction conditions: 7 cycles of 94°C for 25 s and 72°C for 3 min; 32 cycles of 94°C for 25 s and 67°C for 3 min; 7-min extension at 67°C. Amplified products were cloned into an expression vector (pCR2.1; Invitrogen, Carlsbad, CA), and transformed into competent cells (TOP10; Invitrogen, Carlsbad, CA). Individual colonies were selected; cloned plasmids were purified (SNAP MiniPrep kit; Invitrogen, Carlsbad, CA) and sequenced using M13 primers (Invitrogen, Carlsbad, CA).

Mapping integration sites. The BLAT program (University of California, Santa Cruz, Human Genome Project working draft May 2004 freeze; http://www.genome.ucsc.edu/) (20) was used to map integration sites. Sequences that met the following criteria were considered authentic integration sites: (i) contained the terminal 5′ end of the HIV-2 LTR; (ii) had matching genomic DNA within five bp of the end of the viral LTR; (iii) had at least 95% homology to human genomic sequence across the entire sequenced region; (iv) matched a single human genetic locus with at least 95% homology across the entire sequenced region. A total of 202 sequences met the above criteria and were used in this study.

Development of real-time PCR standards. A 292-bp fragment containing a portion of the HIV-1 LTR and *gag* gene was amplified from DNA of a patient infected with HIV-1 subtype A in a heminested PCR using the first-round primers AM1LTRsd (CCTCAATAAAGCTTGCCTTGAG) and G75 (CTTCT ATTACTTTTACCCATGC) and second-round primers AM1LTRsd and AM1C1r (CTTAATACTGACGCTCTCGCACCC). A 307-bp fragment of the HIV-2 LTR was amplified from a previously cloned HIV-2 DNA sample in a single-round PCR using the primers AM2LTRsd (TCTGTATAAATGTACCC GCTT) and AM2C1r (AAGGGTCCTAACAGACCAGGGTCT). The HIV-1 and HIV-2 amplicons were cloned into an expression vector (pCR2.1; Invitrogen, Carlsbad, CA) and transformed into competent cells (TOP10; Invitrogen, Carlsbad, CA). Cloned plasmids were purified (SNAP MiniPrep kit; Invitrogen, Carlsbad, CA), and concentrations of purified plasmids were determined by optical density readings at 260 nm.

Quantification of proviral load. The Applied Biosystems Taqman model 7000 thermocycler and a standardized master mix containing buffer, nucleotides, and enzyme (Taqman Universal PCR Master Mix; Applied Biosystems, Foster City, CA) were used for all real-time PCRs. For quantification of HIV-1 samples, a 158-bp fragment containing a portion of the HIV-1 LTR and *gag* gene was amplified using primers AM1C2f (ATCTCTAGCAGTGGCGCCCGA) and AM1C2r (CCTTCTAGCCTCCGCTAGTCA) and detected using the probe INT-1 (ACGCAGGACTCGGCTTGCTG). For quantification of HIV-2 samples, a 155-bp fragment of the HIV-2 LTR was amplified using primers AM2C2f (GCAGGTAGAGCCTGGGTGTTC) and AM2C2r (CAGGCGGCGACTAG GAGAGAT) and detected using the probe INT-2 (AGACGGCTCCACGCTT GCTT). Both probes contained a 5′ 6-carboxyfluorescein fluorescent reporter and a 3′ MGB nonfluorescent quencher. Twenty picomoles of each primer and 10 picomoles of probe were used for the respective HIV-1 and HIV-2 real-time PCRs, and all reactions were performed in a final volume of 50 μl. PCR conditions for both HIV-1 and HIV-2 were as follows: initial incubation at 50°C for 2 min; denaturation at 94°C for 5 min; and 45 PCR cycles of 95°C for 15 s and 60°C for 1 min. New HIV-1 and HIV-2 standards were made for each real-time PCR run from stock aliquots at a concentration of 10^5 copies/μl and were diluted by 10-fold serial dilutions in HIV-negative human DNA at a concentration of 10 ng/μl. Ten microliters of standard was used for each real-time PCR run, and standards were used in triplicate. For PBMC samples, 1 μg of extracted DNA was used to quantify the proviral load, and reactions were performed in duplicate. Because of differences in the lengths of the HIV-1 and HIV-2 LTR regions, the real-time PCR assay for HIV-1 used a 5′ primer specific to the viral LTR and a 3′ primer specific to the viral *gag* gene, which amplified a single fragment within the viral genome, whereas both primers for the HIV-2 real-time PCR assay were specific to the viral LTR. Since proviral DNA contains copies of the LTR on both ends of the viral genome, the actual HIV-2 proviral load was determined by dividing by 2 the proviral load quantified using real-time PCR.

Detection of proviral integration in heterochromatin in vivo by alphoid repeat PCR. To assay for proviral integration within heterochromatin, we modified an assay described previously (16) that uses PCR to selectively amplify proviral sequences within or near alphoid repeats, using a virus-specific primer and primers designed to alphoid repeat consensus sequence. Alphoid repeats are highly enriched within heterochromatin (33). For the PCR, we used the alphoid repeat primers α1 (AGACAGAAGCATTCTSAGAA) and α4 (AAAGAGTGT TTCMAANCTGCTCW) (16), which amplify in opposite directions within alphoid repeats, to maximize the likelihood of detecting proviral DNA, and the viral primers AM1C1r (CTTAATACTGACGCTCTCGCACCC) and AM2C1r (AAGGGTCCTAACAGACCAGGGTCT) for HIV-1-infected and HIV-2-infected subjects, respectively. PCR amplification was performed in a 50-μl, 40-cycle PCR using 1 μg of DNA under reaction conditions described previously (16). The amount of proviral sequence amplified by alphoid repeat-specific PCR was quantified by real-time PCR using 5 μl of amplified product in duplicate, as described above. Samples with proviral sequence amplified at least 10 times relative to the amount of nonamplified proviral DNA carryover from the alphoid repeat-specific PCR were defined as being positive for the presence of proviral integration within heterochromatin.

Nucleotide sequence accession numbers. Nucleotide sequences of integration site junctions were deposited in GenBank (accession numbers DQ632388 to DQ632563).

RESULTS

Genomic features of HIV-2 integration site selection. HIV-1 integration strongly favors transcriptional units (27, 34). We assessed the tendency of HIV-2 to integrate within transcriptional units by measuring the proportion of HIV-2 integrations occurring in Refseq genes. Refseq genes represent a nonredundant collection of transcripts that are collected on the basis of known mRNA transcripts and manually curated to ensure

TABLE 1. Characteristics of HIV-2 integration sites and comparison with the human genome[a]

Parameter	Value for HIV-2 integrations	Value for human genome	P value[f]
% Located:			
Within transcriptional unit[b]	82.2	35.9	<0.0001
Within ±1 kb of a CpG island	2.0	2.1[d]	0.9055
Within ±5 kb of a transcriptional start site[b]	16.3	5.9[e]	<0.0001
GC content (%) in flanking 1-Mb window	44.6	40.9	<0.0001
Gene density (%) in flanking 1-Mb window[c]	51.2	35.9	<0.0001

[a] Two hundred two HIV-2 integration sites were analyzed. Human genome data are from the Human Genome Project working draft May 2004 freeze (http://www.genome.ucsc.edu/).
[b] Transcriptional units are Refseq genes.
[c] Gene density is defined as the percentage of overall bases falling within Refseq genes.
[d] Data from simulation of random integrations by Wu et al. (41).
[e] Expected value from simulation of random integrations by Wu et al. (41), extrapolated to correct for increased numbers of Refseq transcripts.
[f] P value for χ^2 test for difference in proportions between HIV-2 integrations and the human genome.

TABLE 2. Percentages of HIV-2 integrations within repetitive elements and percentages observed in the human genome[a]

Repetitive element	% of HIV-2 integrations	% of human genome	P value[b]
SINE	25.3	13.7	<0.0001
Alu element	22.3	10.8	<0.0001
LINE	10.9	21.1	0.0004
LTR	3.0	8.7	0.0038
DNA	5.0	3.0	0.1065

[a] Two hundred two HIV-2 integrations were analyzed. Percentages of repetitive elements in the human genome are taken from the Human Genome Project working draft May 2004 freeze (http://www.genome.ucsc.edu/).
[b] P value for χ^2 test for difference in proportions between HIV-2 integrations and the human genome.

accuracy (32). At the time of our analysis, 24,722 Refseq genes had been identified within the human genome, accounting for 35.9% of the human genome. Of 202 HIV-2 integration sites mapped, 166 landed with Refseq genes (82.2%) (Table 1), a proportion much higher than expected by chance ($P < 0.0001$). Interestingly, we noted that among integration sites within Refseq genes, a high proportion of HIV-2 genomes integrated in the reverse direction relative to the direction of transcription (60.1%; $P = 0.0122$). This finding contrasts with findings for HIV-1, which has been reported to have no bias in the direction of integration (34).

CpG islands are regions of the genome enriched in the rare CpG dinucleotide and are commonly located near transcriptional start sites. MLV has been shown to have a strong tendency towards integration near CpG islands and near sites of transcriptional initiation (27, 41). Of the HIV-2 integration sites mapped, 2.0% were within 1 kb of a CpG island. As a comparison, previous simulations have estimated 2.1% of random integrations to fall within ±1 kb of a CpG island in the human genome (41), which is not statistically different from the proportion observed for HIV-2 ($P = 0.9055$). In order to examine HIV-2 integration near transcriptional start sites, the proportion of integrations expected to randomly land within ±5 kb of a transcriptional start site based on a simulation published by Wu et al. (41) were extrapolated to correct for increased numbers of Refseq transcripts at the time of this analysis, as described previously (28). The proportion of HIV-2 integrations located within ±5 kb of transcriptional start sites was significantly higher than expected for random integration (16.3%; $P < 0.0001$); however, this tendency was not as high as previously reported for MLV (27, 41). Because HIV-2 did not have a bias in integration near CpG islands, the high proportion of HIV-2 integrations near transcription start sites is likely due to the strong preference of HIV-2 for integration within transcriptional units and not due to a direct tendency toward integration near transcriptional start sites.

We examined the GC content in the flanking 1-Mb window (±500 kb from the integration site) surrounding each integration site. The mean GC content in this window was 44.6%, a value significantly higher than expected by chance ($P < 0.0001$). Similarly, the average gene density (the percentage of bases that are in Refseq genes) in the 1-Mb window surrounding each integration site was 51.2%, which was higher ($P < 0.0001$) than the overall gene density of the human genome.

Repetitive elements account for close to half of the sequence in the human genome (22), with the majority falling into four classes of transposable elements: short interspersed elements (SINEs), long interspersed elements (LINEs), LTR elements, and DNA elements (3, 37). Compared to the proportion of the genome represented by SINEs, HIV-2 had a significantly increased probability for integration within SINEs (25.5%; $P < 0.0001$) (Table 2), as well as integration within Alu elements, the most common type of SINE (22.3%; $P < 0.0001$). Conversely, integration by HIV-2 was disfavored within LINEs (10.9%; $P = 0.0004$) as well as within LTR elements (3.0%; $P = 0.0038$). The proportion of HIV-2 integrations within DNA elements did not differ statistically from the proportion of the genome (5.0%; $P = 0.1065$). Other repetitive elements represent a minor proportion of overall genomic sequence, and we did not observe a significant number of HIV-2 integrations within any other repetitive elements. Of note, alphoid repeats are highly enriched in heterochromatin (35). Of 202 HIV-2 integration sites mapped during in vitro infection, no integrations occurred within alphoid repeats.

Integration of proviral DNA within heterochromatin in vivo. Integration of HIV-1 within heterochromatin can result in a block in viral transcription (16, 24); however, integration within heterochromatin is a rare event during in vitro HIV-1 infection (7, 34). In this study, similar results were observed during in vitro HIV-2 infection. In order to assay for evidence of in vivo viral integration within heterochromatin, we modified a previously described alphoid repeat PCR assay (16) that amplifies proviral DNA through the use of a virus-specific primer and primers specific to alphoid repeats. To avoid study bias due to unequal amplification between HIV-1 and HIV-2 samples, we used virus-specific reverse primers for HIV-1 and HIV-2 that were located equal distances from the 5′ end of the respective viral genomes and were previously observed to effectively amplify integrated proviral DNA (P. Kanki, unpublished observations). We assayed for the presence of proviral integration in heterochromatin in PBMC DNA samples from

TABLE 3. Proportions of HIV-1-infected and HIV-2-infected subjects with evidence of proviral integration within heterochromatin, based on alphoid repeat PCR

Virus	No. of alphoid repeat PCR-positive subjects/no. of subjects tested (P value)[a]
HIV-1	0/21 (0.0497)
HIV-2	5/23

[a] Two-sided P value for Fisher's exact test for a difference in the proportions of HIV-1- and HIV-2-infected subjects that showed evidence of proviral integration in heterochromatin.

21 HIV-1-infected and 23 HIV-2-infected subjects. All subjects were antiretroviral therapy naive and had detectable proviral loads. There was no difference in mean CD4$^+$ T-cell counts or mean proviral loads between subjects infected with HIV-1 and those with HIV-2. Proviral integration in heterochromatin was detected in 0 of 21 HIV-1-infected subjects and 5 of 23 HIV-2-infected subjects (Table 3). Although integration within heterochromatin was not detected for the majority of HIV-2-infected subjects, the proportion testing positive was statistically higher than that of HIV-1-infected subjects (Fisher's exact test; two-sided P value = 0.0497).

DISCUSSION

HIV-2 has been shown to be less pathogenic and is believed to have a lower in vivo rate of replication than HIV-1. Studies with HIV-1 have shown that integration site selection has an effect on viral transcription (16, 24) and possibly viral replication. To examine whether HIV-2 has a tendency to integrate in regions of the genome that may affect its replicative capacity relative to that of HIV-1, we mapped HIV-2 integration sites during in vitro infection of primary PBMCs. Previously it has been shown that cell type-specific differences in transcription have a modest but significant effect on HIV-1 integration targeting (27). Additionally, differences in integration site selection by HIV-1 have been reported between growth-arrested and dividing cells (9). In order to minimize bias in the interpretation of our results, we used a primary HIV-2 isolate and used primary PBMCs as the target cell type for this study.

We used a linker-mediated nested PCR approach to map HIV-2 integration sites. This technical approach has been used previously and has been demonstrated to effectively map retroviral integration sites without significant bias (34, 41). In our study, we used a first-round reverse viral primer that was specific to the HIV-2 *gag* gene, immediately downstream of the HIV-2 5′ LTR, and a second-round reverse viral primer located near the 5′ end of the viral 5′ LTR. This effectively prevented amplification of internal HIV-2 genomic fragments primed from the 3′ viral LTR. The recognition sequence for the enzyme we selected, DraI, occurs frequently in the human genome but is not found in the LTR of circulating strains of HIV-2. Of concern was the possibility that the recognition sequence for this enzyme (TTT/AAA) would favor AT-rich regions of the genome. However, the average GC content in the flanking 1 Mb surrounding the HIV-2 integration sites was significantly higher than the average GC content of the human genome, demonstrating that our approach did not significantly bias the results towards AT-rich regions of the genome.

HIV-2 strongly favored integration within transcriptional units. Similar integration targeting preferences have been reported for HIV-1 (27, 34) as well as for SIV (10, 13). For instance, in a recent report, Crise et al. mapped 148 SIV integration sites in CEMx174 cells and reanalyzed data from 334 HIV-1 integrations in SupT1 cells. The frequency of integrations landing within Refseq genes was 74% for SIV and 72% for HIV-1 (10). Although these proportions are slightly less than we observed for HIV-2 (82%), differences are likely because the above data were analyzed based on the July 2003 freeze of the human genome, and the data set of Refseq genes would have been slightly smaller than the number currently available for our analysis.

Within transcriptional units, HIV-2 had a preference to integrate in the reverse direction, a finding that differed from that for HIV-1 (34). When located within transcriptional units, human endogenous retrovirus (HERV) sequences occur more commonly in the reverse direction of the transcript, although this is believed to be due to selection against transcriptional termination caused by the HERV polyadenylation sequence (3, 37). However, the possibility remains that the direction of proviral integration may have an effect on viral transcription for HERVs, and similarly, proviral integration in the reverse direction of the cellular transcript could possibly increase the likelihood of HIV-2 establishing latent infection.

HIV-2 is now the third lentivirus (along with HIV-1 and SIV) reported to have a strong preference towards integration within transcriptional units. Only a slight preference towards transcriptional units has been reported for MLV and ASLV. In contrast to MLV and ASLV, HIV-1, HIV-2, and SIV all infect primates as their natural hosts; it is possible that differences in integration targeting preference by MLV and ASLV are due to the fact that these viruses have been studied in human cells. However, Barr et al. recently examined integration targeting by HIV-1 and ASLV in chicken embryo fibroblasts and found integration preferences similar to those found previously in human cells (4). It is also possible that lentiviruses have an integration targeting mechanism that differs from that of other retroviruses. For instance, the p75 isoform of the transcriptional coactivator lens epithelium-derived growth factor (LEDGF/p75) is believed to promote tethering of HIV-1 integrase to DNA (38). In addition to HIV-1, LEDGF/p75 has been shown to interact with the integrase of HIV-2 and that of the lentivirus feline immunodeficiency virus but does not interact with the integrases of human T-cell lymphotrophic virus type 2, Moloney murine leukemia virus, and Rous sarcoma virus (6). Additionally, HIV-1 integration within transcriptional units occurs less frequently in cells that are depleted of LEDGF/p75 than control cells (8).

HIV-2 integration was strongly favored in SINEs, and similar preferences have been reported for HIV-1 integration (34). In contrast, integration within LINEs and LTR elements was disfavored by HIV-2. Areas of high gene expression in the human genome are associated with high gene density, high GC content, high SINE density, and low LINE density (39). These characteristics are all consistent with preferences in integration site targeting observed for HIV-2. This preference could possibly be explained by an interaction between the HIV-2 preintegration complex and the transcription factors that are commonly found at these domains. Indeed, HIV-1 maintains a

preference of integrating within actively transcribed genes (27, 34). It is also possible that these regions are favored because they are relatively accessible, making it easier for an interaction to occur between cellular DNA and the preintegration complex.

Since only a small proportion of cells carry integrated proviral DNA in vivo, studies examining retroviral integration sites in vivo are extremely tedious and have been limited. In the most extensive study, Han et al. mapped 74 HIV-1 integration sites in resting $CD4^+$ T cells from 16 patients on highly active antiretroviral therapy and found that the majority of integrations were located within transcriptional units (11). Because integration within heterochromatin results in low viral transcription (16, 24), differences between HIV-1 and HIV-2 in the proportion of proviral integrations within heterochromatin in vivo could account for the difference in replication and pathogenesis observed between these viruses. As an alternative to mapping individual integration sites in vivo, we used alphoid repeat PCR to broadly assay for evidence of proviral integration in heterochromatin in subjects infected with HIV-1 and subjects infected with HIV-2.

We acknowledge that alphoid repeat PCR had distinct limitations. Importantly, amplification was contingent on proviral integration in relatively close proximity to alphoid repeats. The absence of amplification does not definitively represent the absence of integration within heterochromatin. Additionally, although this approach allowed us to assay for a rare event more efficiently than can be detected by integration mapping, the overall proviral DNA copy number in vivo is relatively small (median = 87 copies/μg DNA for samples used in this study), limiting our overall ability to effectively quantify the number of integration events in heterochromatin in vivo. It should be noted that these limitations exist for both HIV-1 and HIV-2 samples, yet evidence of proviral integration within heterochromatin existed for a statistically higher proportion of HIV-2-infected subjects than for HIV-1-infected subjects. Although this observation is consistent with the notion that HIV-2 replicates to a lesser degree than HIV-1 in vivo, this difference was modest and cannot fully account for the substantially higher viral loads observed in HIV-1 infection.

In this study we found in vitro integration site selection for HIV-2 to be similar to what has previously been reported for HIV-1, yet we found a possible difference in the integration site profile in vivo. Two possible explanations may account for this discordance. First, it is possible that in vitro infection studies do not properly model in vivo retroviral integration site selection and there are cellular or physiologic factors involved in proviral integration in vivo that are absent during in vitro viral infection. Second, it is possible that HIV-1 and HIV-2 do not have inherent differences in integration site selection per se but that differences in the integration profile arise due to selective pressures in vivo and are not captured during acute in vitro infection. For instance, the LTRs of HIV-1 and HIV-2 contain different regulatory elements and respond differently to cellular stimuli (12). If integration in heterochromatin were to result in repression of transcription more effectively for HIV-2 than for HIV-1, then the result would likely be a strong selection against integration within heterochromatin in vivo for HIV-1 due to cytopathic effects of viral replication and cell-mediated immune responses against infected cells.

To our knowledge, this is the first study to examine integration site selection by HIV-2. We examined integration site selection during in vitro infection using a primary viral isolate to infect primary PBMCs and observed HIV-2 to have integration tendencies similar to those previously reported for HIV-1. We observed evidence of integration in heterochromatin in vivo more commonly with HIV-2-infected subjects than with HIV-1-infected subjects. However, evidence of integration in heterochromatin occurred in only a small proportion of HIV-2 samples. These findings suggest that the impact of integration site selection may play a small or negligible role in the differences in pathogenesis between HIV-1 and HIV-2. Further studies are needed to fully characterize the viral determinants involved in the pathogenicity of HIV-1 and HIV-2.

ACKNOWLEDGMENTS

We thank Shaun Rodriguez for helpful suggestions and Beth Chaplin and Christopher Mullins for technical assistance.

This work was supported by National Institute of Allergy and Infectious Disease Grant AI46274-04 from the National Institutes of Health.

REFERENCES

1. **Alabi, A. S., S. Jaffar, K. Ariyoshi, T. Blanchard, M. Schim van der Loeff, A. A. Awasana, T. Corrah, S. Sabally, R. Sarge-Njie, F. Cham-Jallow, A. Jaye, N. Berry, and H. Whittle.** 2003. Plasma viral load, CD4 cell percentage, HLA and survival of HIV-1, HIV-2, and dually infected Gambian patients. AIDS **17:**1513–1520.
2. **Andersson, S., H. Norrgren, Z. da Silva, A. Biague, S. Bamba, S. Kwok, C. Christopherson, G. Biberfeld, and J. Albert.** 2000. Plasma viral load in HIV-1 and HIV-2 singly and dually infected individuals in Guinea-Bissau, West Africa: significantly lower plasma virus set point in HIV-2 infection than in HIV-1 infection. Arch. Intern. Med. **160:**3286–3293.
3. **Bannert, N., and R. Kurth.** 2004. Retroelements and the human genome: new perspectives on an old relation. Proc. Natl. Acad. Sci. USA **101**(Suppl. 2)**:**14572–14579.
4. **Barr, S. D., J. Leipzig, P. Shinn, J. R. Ecker, and F. D. Bushman.** 2005. Integration targeting by avian sarcoma-leukosis virus and human immunodeficiency virus in the chicken genome. J. Virol. **79:**12035–12044.
5. **Berry, N., K. Ariyoshi, O. Jobe, P. T. Ngum, T. Corrah, A. Wilkins, H. Whittle, and R. Tedder.** 1994. HIV type 2 proviral load measured by quantitative polymerase chain reaction correlates with CD4+ lymphopenia in HIV type 2-infected individuals. AIDS Res. Hum Retrovir. **10:**1031–1037.
6. **Busschots, K., J. Vercammen, S. Emiliani, R. Benarous, Y. Engelborghs, F. Christ, and Z. Debyser.** 2005. The interaction of LEDGF/p75 with integrase is lentivirus-specific and promotes DNA binding. J. Biol. Chem. **280:**17841–17847.
7. **Carteau, S., C. Hoffmann, and F. Bushman.** 1998. Chromosome structure and human immunodeficiency virus type 1 cDNA integration: centromeric alphoid repeats are a disfavored target. J. Virol. **72:**4005–4014.
8. **Ciuffi, A., M. Llano, E. Poeschla, C. Hoffmann, J. Leipzig, P. Shinn, J. R. Ecker, and F. Bushman.** 2005. A role for LEDGF/p75 in targeting HIV DNA integration. Nat. Med. **11:**1287–1289.
9. **Ciuffi, A., R. S. Mitchell, C. Hoffmann, J. Leipzig, P. Shinn, J. R. Ecker, and F. D. Bushman.** 2006. Integration site selection by HIV-based vectors in dividing and growth-arrested IMR-90 lung fibroblasts. Mol. Ther. **13:**366–373.
10. **Crise, B., Y. Li, C. Yuan, D. R. Morcock, D. Whitby, D. J. Munroe, L. O. Arthur, and X. Wu.** 2005. Simian immunodeficiency virus integration preference is similar to that of human immunodeficiency virus type 1. J. Virol. **79:**12199–12204.
11. **Han, Y., K. Lassen, D. Monie, A. R. Sedaghat, S. Shimoji, X. Liu, T. C. Pierson, J. B. Margolick, R. F. Siliciano, and J. D. Siliciano.** 2004. Resting CD4+ T cells from human immunodeficiency virus type 1 (HIV-1)-infected individuals carry integrated HIV-1 genomes within actively transcribed host genes. J. Virol. **78:**6122–6133.
12. **Hannibal, M. C., D. M. Markovitz, N. Clark, and G. J. Nabel.** 1993. Differential activation of human immunodeficiency virus type 1 and 2 transcription by specific T-cell activation signals. J. Virol. **67:**5035–5040.
13. **Hematti, P., B. K. Hong, C. Ferguson, R. Adler, H. Hanawa, S. Sellers, I. E. Holt, C. E. Eckfeldt, Y. Sharma, M. Schmidt, C. von Kalle, D. A. Persons, E. M. Billings, C. M. Verfaillie, A. W. Nienhuis, T. G. Wolfsberg, C. E. Dunbar, and B. Calmels.** 2004. Distinct genomic integration of MLV and SIV vectors in primate hematopoietic stem and progenitor cells. PLoS Biol. **2:**e423.

14. **Hindmarsh, P., and J. Leis.** 1999. Retroviral DNA integration. Microbiol. Mol. Biol. Rev. **63:**836–843.
15. **Holman, A. G., and J. M. Coffin.** 2005. Symmetrical base preferences surrounding HIV-1, avian sarcoma/leukosis virus, and murine leukemia virus integration sites. Proc. Natl. Acad. Sci. USA **102:**6103–6107.
16. **Jordan, A., D. Bisgrove, and E. Verdin.** 2003. HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. EMBO J. **22:**1868–1877.
17. **Jordan, A., P. Defechereux, and E. Verdin.** 2001. The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. EMBO J. **20:**1726–1738.
18. **Kanki, P., S. M'Boup, R. Marlink, K. Travers, C. C. Hsieh, A. Gueye, C. Boye, J. L. Sankale, C. Donnelly, W. Leisenring, et al.** 1992. Prevalence and risk determinants of human immunodeficiency virus type 2 (HIV-2) and human immunodeficiency virus type 1 (HIV-1) in west African female prostitutes. Am. J. Epidemiol. **136:**895–907.
19. **Kanki, P. J.** 1999. Human immunodeficiency virus type 2 (HIV-2). AIDS Rev. **1:**101–108.
20. **Kent, W. J., C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle, A. M. Zahler, and D. Haussler.** 2002. The human genome browser at UCSC. Genome Res. **12:**996–1006.
21. **Kokkotou, E. G., J. L. Sankale, I. Mani, A. Gueye-Ndiaye, D. Schwartz, M. E. Essex, S. Mboup, and P. J. Kanki.** 2000. In vitro correlates of HIV-2-mediated HIV-1 protection. Proc. Natl. Acad. Sci. USA **97:**6797–6802.
22. **Lander, E. S., L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J. P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J. C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R. H. Waterston, R. K. Wilson, L. W. Hillier, J. D. McPherson, M. A. Marra, E. R. Mardis, L. A. Fulton, A. T. Chinwalla, K. H. Pepin, W. R. Gish, S. L. Chissoe, M. C. Wendl, K. D. Delehaunty, T. L. Miner, A. Delehaunty, J. B. Kramer, L. L. Cook, R. S. Fulton, D. L. Johnson, P. J. Minx, S. W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J. F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, et al.** 2001. Initial sequencing and analysis of the human genome. Nature **409:** 860–921.
23. **Lassen, K., Y. Han, Y. Zhou, J. Siliciano, and R. F. Siliciano.** 2004. The multifactorial nature of HIV-1 latency. Trends Mol. Med. **10:**525–531.
24. **Lewinski, M. K., D. Bisgrove, P. Shinn, H. Chen, C. Hoffmann, S. Hannenhalli, E. Verdin, C. C. Berry, J. R. Ecker, and F. D. Bushman.** 2005. Genome-wide analysis of chromosomal features repressing human immunodeficiency virus transcription. J. Virol. **79:**6610–6619.
25. **Marlink, R., P. Kanki, I. Thior, K. Travers, G. Eisen, T. Siby, I. Traore, C. C. Hsieh, M. C. Dia, E. H. Gueye, et al.** 1994. Reduced rate of disease development after HIV-2 infection as compared to HIV-1. Science **265:**1587–1590.
26. **Marlink, R. G., D. Ricard, S. M'Boup, P. J. Kanki, J. L. Romet-Lemonne, I. N'Doye, K. Diop, M. A. Simpson, F. Greco, M. J. Chou, et al.** 1988. Clinical, hematologic, and immunologic cross-sectional evaluation of individuals exposed to human immunodeficiency virus type-2 (HIV-2). AIDS Res. Hum Retrovir. **4:**137–148.
27. **Mitchell, R. S., B. F. Beitzel, A. R. Schroder, P. Shinn, H. Chen, C. C. Berry, J. R. Ecker, and F. D. Bushman.** 2004. Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. PLoS Biol. **2:**E234.
28. **Narezkina, A., K. D. Taganov, S. Litwin, R. Stoyanova, J. Hayashi, C. Seeger, A. M. Skalka, and R. A. Katz.** 2004. Genome-wide analyses of avian sarcoma virus integration sites. J. Virol. **78:**11656–11663.
29. **Norrgren, H., S. Marquina, T. Leitner, P. Aaby, M. Melbye, A. G. Poulsen, O. Larsen, F. Dias, D. Escanilla, S. Andersson, J. Albert, and A. Naucler.** 1997. HIV-2 genetic variation and DNA load in asymptomatic carriers and AIDS cases in Guinea-Bissau. J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. **16:**31–38.
30. **Popper, S. J., A. D. Sarr, A. Gueye-Ndiaye, S. Mboup, M. E. Essex, and P. J. Kanki.** 2000. Low plasma human immunodeficiency virus type 2 viral load is independent of proviral load: low virus production in vivo. J. Virol. **74:**1554–1557.
31. **Popper, S. J., A. D. Sarr, K. U. Travers, A. Gueye-Ndiaye, S. Mboup, M. E. Essex, and P. J. Kanki.** 1999. Lower human immunodeficiency virus (HIV) type 2 viral load reflects the difference in pathogenicity of HIV-1 and HIV-2. J. Infect. Dis. **180:**1116–1121.
32. **Pruitt, K. D., T. Tatusova, and D. R. Maglott.** 2005. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res. **33:**D501–D504.
33. **Rudd, M. K., and H. F. Willard.** 2004. Analysis of the centromeric regions of the human genome assembly. Trends Genet. **20:**529–533.
34. **Schroder, A. R., P. Shinn, H. Chen, C. Berry, J. R. Ecker, and F. Bushman.** 2002. HIV-1 integration in the human genome favors active genes and local hotspots. Cell **110:**521–529.
35. **She, X., J. E. Horvath, Z. Jiang, G. Liu, T. S. Furey, L. Christ, R. Clark, T. Graves, C. L. Gulden, C. Alkan, J. A. Bailey, C. Sahinalp, M. Rocchi, D. Haussler, R. K. Wilson, W. Miller, S. Schwartz, and E. E. Eichler.** 2004. The structure and evolution of centromeric transition regions within the human genome. Nature **430:**857–864.
36. **Simon, F., S. Matheron, C. Tamalet, I. Loussert-Ajaka, S. Bartczak, J. M. Pepin, C. Dhiver, E. Gamba, C. Elbim, J. A. Gastaut, et al.** 1993. Cellular and plasma viral load in patients infected with HIV-2. AIDS **7:**1411–1417.
37. **Smit, A. F.** 1999. Interspersed repeats and other mementos of transposable elements in mammalian genomes. Curr. Opin. Genet. Dev. **9:**657–663.
38. **Van Maele, B., K. Busschots, L. Vandekerckhove, F. Christ, and Z. Debyser.** 2006. Cellular co-factors of HIV-1 integration. Trends Biochem. Sci. **31:**98–105.
39. **Versteeg, R., B. D. van Schaik, M. F. van Batenburg, M. Roos, R. Monajemi, H. Caron, H. J. Bussemaker, and A. H. van Kampen.** 2003. The human transcriptome map reveals extremes in gene density, intron length, GC content, and repeat pattern for domains of highly and weakly expressed genes. Genome Res. **13:**1998–2004.
40. **Whittle, H., J. Morris, J. Todd, T. Corrah, S. Sabally, J. Bangali, P. T. Ngom, M. Rolfe, and A. Wilkins.** 1994. HIV-2-infected patients survive longer than HIV-1-infected patients. AIDS **8:**1617–1620.
41. **Wu, X., Y. Li, B. Crise, and S. M. Burgess.** 2003. Transcription start regions in the human genome are favored targets for MLV integration. Science **300:**1749–1751.
42. **Wu, X., Y. Li, B. Crise, S. M. Burgess, and D. J. Munroe.** 2005. Weak palindromic consensus sequences are a common feature found at the integration target sites of many retroviruses. J. Virol. **79:**5211–5214.