In-Depth, Longitudinal Analysis of Viral Quasispecies from an Individual Triply Infected with Late-Stage Human Immunodeficiency Virus Type 1, Using a Multiple PCR Primer Approach

M. Gerhardt,^{1*} D. Mloka,^{1,2} S. Tovanabutra,³ E. Sanders-Buell,³ O. Hoffmann,^{1,4} L. Maboko,⁵
D. Mmbando,⁶ D. L. Birx,³ F. E. McCutchan,³ and M. Hoelscher¹

Department of Infectious Diseases and Tropical Medicine, University of Munich, Munich, Germany¹; Muhimbili University College of Health Sciences, Dar-es-Salaam, Tanzania²; U.S. Military HIV Research Program, Rockville, Maryland³; Centre for Population Studies, London School of Hygiene and Tropical Medicine, London, United Kingdom⁴; Mbeya Medical Research Programme, Mbeya, Tanzania⁵; and Mbeya Regional Medical Office, Ministry of Health, Mbeya, Tanzania⁶

Received 11 November 2004/Accepted 15 February 2005

Coinfections with more than one human immunodeficiency virus type 1 (HIV-1) subtype appear to be the source of new recombinant strains and may be commonplace in high-risk cohorts exposed to multiple subtypes. Many potential dual infections have been identified during the HIV Superinfection Study in Mbeya, Tanzania, where 600 female bar workers who are highly exposed to subtypes A, C, and D have been evaluated every 3 months for over 3 years by use of the MHAacd HIV-1 genotyping assay. Here we describe an in-depth, longitudinal analysis of the viral quasispecies in a woman who was triply infected with HIV-1 and who developed AIDS and passed away 15 months after enrollment. The MHA results obtained at 0, 3, 6, 9, and 12 months revealed dual-probe reactivities and shifts in subtype over time, indicating a potential dual infection and prompting further investigation. The multiple infection was confirmed by PCR amplification of three genome regions by a multiple primer approach, followed by molecular cloning and sequencing. A highly complex viral quasispecies was found, including several recombinant forms, with *vpu/gp120* being the most diverse region. A significant fluctuation in molecular forms over time was observed, showing that the serial sample format is highly desirable, if not essential, for the identification of multiple infections. In a separate experiment, we confirmed that the detection of coinfections is more efficient with the use of multiple amplification primers to overcome the primer bias that results from the enormous diversity in the HIV-1 genome.

The high level of diversity of the human immunodeficiency virus type 1 (HIV-1) genome is caused by the ability of the virus to mutate, recombine, and replicate with a short generation time. The high mutation rate of 0.3 nucleotide changes per genome per replication cycle (16) is the result of inaccurate viral DNA synthesis by the viral enzyme reverse transcriptase, which is error-prone partly due to its lack of 3'→5' exonuclease proofreading activity (4, 6, 14). The enzyme also switches between the two RNA templates that are packaged in each virion approximately three times per genome per replication cycle (1, 16). When a cell is simultaneously infected with two or more HIV-1 strains, recombination during replication leads to a novel viral genome which contains genetic information from both proviruses. Some mosaic HIV-1 genomes circulate broadly in the human population and are now designated "circulating recombinant forms" (CRFs), 16 of which have been described to date (2, 28, 34). The number of CRFs will continue to increase as HIV-1-infected populations are evaluated completely at the molecular level.

HIV-1 exists *in vivo* as a quasispecies, a population of highly related but genetically distinct viruses (8, 33). This rapid vari-

ability provides the virus with the capability to adapt its genome to escape selective pressures from the immune system and antiretroviral therapy. Several reports have shown correlations between HIV-1 evolution in an infected individual and the progression to AIDS (7, 32), the development of drug resistance (5, 17, 24), and vertical transmission (23).

Adequate sampling of the viral quasispecies is challenging, yet necessary, for the detection of multiple, molecularly distinct infections. The full recovery and characterization of strains from a multiply infected individual over the course of HIV-1 infection require new approaches and sensitive methods.

HIV-1 subtypes A, C, D, and CRF10_CD and many unique recombinant forms are cocirculating in Tanzania, a setting that facilitates the evaluation of multiple infections with different subtypes and the recombination between them (12, 15, 18, 19, 26). In this report, we describe an in-depth, longitudinal analysis of the viral quasispecies in a woman with late-stage HIV-1 infection who was 1 of 600 female bar workers constituting a high-risk cohort in a longitudinal HIV Superinfection Study (HISIS) in Mbeya Region, Tanzania, that is described elsewhere (13, 27). The study participants were evaluated every 3 months for up to 4 years. HIV-1-positive samples were screened by a multiregion hybridization assay (MHAacd) (11). The objective of this work was to characterize the viral quasispecies in a triply infected HIV-1-positive individual as accu-

* Corresponding author. Mailing address: Department of Infectious Diseases and Tropical Medicine, Ludwig-Maximilians-University, Leopoldstr. 5, 80802 Munich, Germany. Phone: 49 89 2180-3925. Fax: 49 89 33 60 38. E-mail: martina.gerhardt@lrz.uni-muenchen.de.

FIG. 1. Locations of second-round PCR amplicons in the HIV-1 genome. Region 1 refers to *gag/pol*, region 2 refers to *vpu/gp120*, and region 3 refers to *gp41/nef*. Each fragment was generated by a multiple-primer PCR approach and was between 1.1 and 1.4 kb long. Numbers below the fragments indicate HXB2 locations.

rately as possible by using a multiple primer approach for nested PCR amplification of three 1.4-kb parts of the genome. Two primer sets used in four pair-wise combinations for each PCR were used to minimize bias in the recovery of strains. This approach was employed to give us a more complete picture of the viral population within the patient, and based on our results, this approach does increase the likelihood of detecting dual infections.

MATERIALS AND METHODS

Study population. Six hundred women from a high-risk population in Mbeya, Tanzania, were enrolled between September and November 2000 for a prospective HIV Superinfection Study (HISIS) within the Mbeya Medical Research Programme. The women have been working in bars along the Trans-African highway and are highly exposed to subtypes A, C, and D, the three predominant HIV-1 subtypes circulating in East Africa. After giving informed consent, each participant provided blood samples at enrollment and every 3 months thereafter for a period of up to 4 years. In 2001, the HISIS project provided health care to all participants, which included the treatment of all acute infectious diseases, screening for and treatment of sexually transmitted diseases, and since 2003, cotrimoxazole prophylaxis for women with CD4 counts below 200 (27). Since September 2004, all women with low CD4 counts have been referred for antiretroviral treatment within the Southern Highland Care and Treatment Programme.

Sample processing and MHAacd. HIV-1 antibody testing and the isolation of peripheral blood mononuclear cells (PBMCs) were done in Mbeya, Tanzania. The HIV-1 status of each patient was determined by using two diagnostic HIV enzyme-linked immunoassay tests (Enzygnost Anti HIV1/2 Plus [Dade Behring, Liederbach, Germany] and Determine HIV 1/2 [Abbott, Wiesbaden, Germany]). Discordant results were resolved by Western blotting (Genelabs Diagnostics, Geneva, Switzerland). PBMCs were isolated by Histopaque-1077-1 (Sigma Diagnostics, Taufkirchen, Germany) density gradient centrifugation and shipped in liquid nitrogen to the Tropical Institute in Munich, Germany, where DNA extraction, MHAacd, PCR amplification, and cloning were performed. For HIV-positive samples, DNAs were typically extracted from 200,000 PBMCs (High Pure viral nucleic acid kit; Roche Diagnostics, Mannheim, Germany), and screening for the HIV-1 subtype was done by a multiregion hybridization assay (MHAacd) (11) on an ABI 7700 sequence detection system (Applied Biosystems, Darmstadt, Germany). Plasma viral loads were determined with RNAs extracted from plasma by use of an Amplicor HIV-1 MONITOR test, version 1.5 (Roche Molecular Systems, Branchburg, N.J.).

In-depth analysis of participant 507. One participant who appeared to be multiply infected, as assessed by MHAacd, was studied at 0, 3, 6, 9, and 12 months.

To prove that the serial samples came from the same individual, we used an AmpFESTR Profiler kit (Applied Biosystems, Foster City, Calif.), based on a method using length variations of short tandem repeat loci in the human genome for unique human identification. With this PCR amplification kit, the repeat regions of nine short tandem repeat loci and a segment of the X-Y homologous gene amelogenin were coamplified, with one primer of each locus-specific primer pair being labeled with a fluorescent dye (35). Electrophoresis of the PCR products to separate the alleles according to size and fluorescence detection were done on an ABI PRISM 3100 genetic analyzer (Applied Biosystems, Foster City, Calif.). The multiple infection was confirmed by sequencing multiple clones at each of three genome regions for each time point.

PCR amplification. Using nested PCR, we generated amplicons of 1.1 to 1.4 kb from regions 1 (*gag/pol*), 2 (*vpu/gp120*), and 3 (*gp41/nef*). Region 1 stretches

from the p17 protein gene in *gag* to the beginning of the *pol* gene, region 2 starts in the middle of *vpu* and covers all of the *gp120* gene, and genome region 3 contains the *gp41* gene and half of the *nef* coding region. To increase the sensitivity for recovering all viral quasiespecies, we used a multiple primer approach. Two outer primer pairs and two inner primer pairs were chosen and used in four different combinations in four separate PCRs. The PCR mixtures contained a 200 μ M concentration of each deoxynucleoside triphosphate, 1.5 mM $MgCl_2$, 15 mM Tris-HCl, pH 8.0, 50 mM KCl, 0.2 μ M (each) primers, and 3.75 U of AmpliTaq Gold (Applied Biosystems, Darmstadt, Germany). Five microliters (for genome regions 1 and 2) or 10 μ l (for genome region 3) of extracted DNA was added to the first-round PCRs, while 1 μ l (for regions 1 and 3) or 3 μ l (for region 2) of the first-round products was enough for the second round of amplification. The thermocycler protocol was as follows: 95°C for 10 min; 35 cycles of 95°C for 10 s, annealing for 30 s, and 72°C for 2 min; and one cycle of 72°C for 10 min. Figure 1 shows the locations of the second-round amplicons, and the primers and annealing temperatures are described in Table 1. Five microliters of each second-round PCR product was subjected to agarose gel electrophoresis, and the products were visualized by ethidium bromide staining.

Cloning and sequencing. The four PCR products for each region were pooled, and 4 μ l of the mix was cloned into the pCR4-TOPO vector according to the instructions of a TOPO TA cloning kit for sequencing (Invitrogen, Karlsruhe, Germany). At least 24 clones were picked and grown overnight in 3.3 ml of Luria-Bertani broth (Invitrogen, Karlsruhe, Germany) plus kanamycin at 37°C. After purification of the plasmid DNA (QIAprep spin miniprep kit; QIAGEN, Hilden, Germany), 3 μ l was used for enzymatic digestion with 8 U of EcoRI (New England Biolabs, Frankfurt am Main, Germany) and 10 $\times$ NEB (5 mM NaCl, 10 mM Tris-HCl, 1 mM $MgCl_2$, 0.0025% Triton X-100, pH 7.5). Incubation was performed at 37°C for 2 h, and agarose gel electrophoresis was performed with 10 μ l of the digestion products. Clones with inserts of the expected size were shipped to the Henry M. Jackson Foundation, Rockville, Md., where sequencing and phylogenetic analysis took place. Between 17 and 25 positive clones for each region and time point were sequenced by using fluorescent dye terminators (Applied Biosystems, Foster City, Calif.), and electrophoresis and data collection were done on an ABI Prism 3100 genetic analyzer (Applied Biosystems, Foster City, Calif.).

Phylogenetic analysis. The sequences were assembled with Sequencher, aligned with reference sequences of all HIV-1 subtypes, and phylogenetically analyzed with the MEGA, version 2.1, program (20) and the SEQBOOT, DNAPARS, DNADIST, NEIGHBOR, and CONSENSE modules of the PHYLIP package incorporated in the GDE interface (30). Distance analysis was performed with the Kimura two-parameter model and a transition/transversion ratio of 2.0. Bootscanning (25, 29) and informative site analysis software, which tabulates subtype associations in 10-bp increments, were used to identify recombinant breakpoints, which were confirmed by visual inspections of the sequence alignment and/or by subregion trees.

Evaluation of multiple-primer PCR approach. The last available sample from the patient (at 12 months) and the *vpu/gp120* genome region were chosen for a separate experiment to determine whether the multiple-primer PCR approach was more sensitive than a single nested PCR for the recovery of the viral quasiespecies. Extracted DNAs were separately amplified with the four different primer combinations by nested PCR as described for the first set of experiments, but the four second-round amplicons were screened by MHA and cloned individually. At least 20 clones from each amplification were sequenced and phylogenetically analyzed by the same procedures as those used previously.

GenBank accession numbers. All sequences described in this paper are available under GenBank accession no. AY753734 to AY753739 and AY753746 to AY753837 (*gag/pol*), AY775581 to AY775676 (*gp41/nef*), and AY821308 to AY821493 (*vpu/gp120*).

TABLE 1. PCR primer sequences with HXB2 locations and annealing temperatures for regions 1, 2, and 3

Target or primer set	Primer name (reference)	Sequence (5'-3')	HXB2 numbering	<i>T</i> _{anneal} (°C)
<i>gag/pol</i> (region 1)				
outer primer pair 1	F2NST (25)	→ GCGGAGGCTAGAAGGAGAGAGATGG	769–793	58
	SP3AS	← CCTCCAATTCCTCCCTATCATTTTTGG	2382–2407	
outer primer pair 2	MSF12B (25)	→ AAATCTCTAGCAGTGGCGCCCGAACAG	623–649	58
	BJPOL3 (31)	← GTTGACAGGTGTAGGTCCTAC	2481–2501	
Inner primer pair 1	MHGAG1 (11)	→ AGTATGGGCAAGCAGGGA	891–908	55
	MHPOL4 (11)	← TGCCAAAGAGGTGATTTGAGG	2253–2272	
inner primer pair 2	DD	→ GTATGGGCAAGCAGGGAGCTAGAA	892–915	55
	MHPOL2 (11)	← ACTGTATCATCTGCTCCTGT	2328–2347	
<i>vpu/gp120</i> (region 2)				
outer primer pair 1	MHVPU1 (11)	→ CCTATGGCAGGAAGAAGCGG	5967–5986	52
	GP120 3' (22)	← AGTGCTTCCTGCTGCTCC	7794–7811	
outer primer pair 2	ED3	→ TTAGGCATCTCCTATGGCAGGAAGAAGCGG	5957–5986	60
	AA1570 (3)	← GCCCATAGTGCTTCTGCTGCTCC	7794–7817	
inner primer pair 1	MHVPU3 (11)	→ GCAGAAGAYAGTGGCAATGAG	6209–6229	54
	EDS8	← CACTTCTCCAATTGTCCTCA	7648–7668	
inner primer pair 2	GP120 5' (22)	→ AGAGCAGAAGACAGTGGCAATGA	6206–6228	54
	AA1400 (3)	← TTTCTCTCTCCAGGTCTGAAG	7625–7645	
<i>gp41/nef</i> (region 3)				
outer primer pair 1	AES9	→ GAGGGACAATTGGAGAAGTG	7649–7668	60
	JL90	← GTGTGTGGTAGACCCACAGATCAAGGAT	9121–9148	
outer primer pair 2	AES9	→ GAGGGACAATTGGAGAAGTG	7649–7668	60
	JL89 (21)	← TCCAGTCCCCCTTTTCTTTTAAAAA	9064–9089	
inner primer pair 1	JL110.ta	→ GGGCAAAGAGAAGAGTGGTG	7723–7742	60
	JL90	← GTGTGTGGTAGACCCACAGATCAAGGAT	9121–9148	
inner primer pair 2	ZKF	→ GTGGGAATAGGAGCTGTGTTCTTGGG	7761–7787	60
	JL88.ta	← GTCATTGGTCTTAAAGGTACCTG	9013–9035	

RESULTS

Within the HIV Superinfection Study (HISIS) in Mbeya, Tanzania, 407 of 600 (67.8%) female bar workers tested HIV-1 positive at baseline and were screened for their HIV-1 subtype every 3 months by use of a multiregion hybridization assay (MHAacd). This genotyping assay is based on subtype-specific fluorescent probes binding to HIV DNA in five genome regions (11). Among the many potential dual infections that have been identified during this study, one participant with late-stage HIV-1 infection was selected for analysis. Serological assays at enrollment indicated that the woman had a cured syphilis infection and was, as were 86% of all bar workers, infected with herpes simplex virus 2. She was healthy at enrollment, reported short episodes of fever and diarrhea at the 6- and 9-month follow-up visits, and presented with a very reduced general condition and AIDS-defining symptoms such as weight loss, chronic fever, and cough (with a duration of longer than 1 month) at the 12-month follow-up. At all time points, her viral load was above 500,000 copies/ml; her CD4 count only became available later in the study. The woman missed the subsequent follow-up at month 15. A verbal autopsy conducted among relatives revealed an AIDS-related opportunistic infection as the cause of death.

The initial MHAacd screening of the five available time points showed a dual infection with subtypes A and C. It was interesting that the recognition of the different subtypes was inconsistent in the longitudinal follow-up (data not shown).

For confirmation of the dual infection, three fragments from each time point, each of which was approximately 1.4 kb, were amplified, cloned, and subsequently sequenced. The three

fragments, for the *gag/pol* (1), *vpu/gp120* (2), and *gp41/nef* (3) regions, covered all genome regions probed by the MHA. In total, 298 sequences were derived, with between 17 and 25 sequences from each time point and genome region. A phylogenetic analysis identified a highly complex viral quasispecies in all three regions, including several different recombinant forms. The details of the phylogenetic analysis are described below.

Region 1 (*gag/pol*). Ninety-eight sequences of the *gag/pol* genome region were drawn in equal numbers from the 0-, 3-, 6-, 9-, and 12-month samples. As shown in Fig. 2, three molecular forms were identified, namely, a pure subtype A form (form I) and two AC recombinant forms (forms II and III). Supported by a significant bootstrap value, the A portions of the three forms were similar, indicating that the A strain was one of the parental strains of the two AC recombinants. Both recombinants had similar structures but differed in the position of the first breakpoint. A subregion tree (bootstrap 100) (Fig. 3) indicated that the C portions originated from the same parental C strain, although it was not detected as a pure subtype C strain among the analyzed sequences. Further confirmation of these findings was done by a distance analysis of each subregion. The mean distances between the A portions of each form and between the C portions of forms II and III did not differ significantly from the heterogeneity within the sequence clusters of each form and were all below 3.0%. Form II was the predominant strain, representing 60% of all sequences, followed by form I, the pure A strain, at 34%. Form III was observed more rarely than the other forms (6% of all sequences), and only at the 6- and 9-month follow-ups. No ob-

FIG. 2. (A) Neighbor-joining (NJ) phylogenetic tree of HIV-1 region 1 (*gag/pol*) nucleotide sequences taken 0 (circles), 3 (diamonds), 6 (squares), 9 (triangles), and 12 (pentagons) months after enrollment. Numbers at branch nodes refer to bootstrap values; only values of >70% are shown. (B) Structures of the three molecular forms confirmed by bootscanning and subregion trees as described in the legend to Fig. 3. (C) Histogram representing the proportion of each molecular form in all sequences of each time point.

vious trend in the proportions of the three forms over time could be observed, although their proportions did fluctuate between time points (Fig. 2C).

Region 2 (*vpu/gp120*). Figure 4 shows the viral diversity in *vpu/gp120*. Among the 104 sequences, at least 12 different molecular forms were identified, making *vpu/gp120* the most

diverse region studied for this individual. The patient was apparently triply infected, with one strain being subtype C (form I) and with two divergent subtype A strains, A₁ (form II) and A₂ (form III). The mean distance between the two A clusters was 9.2% higher than the heterogeneity within the sequence of each A variant (about 2%), confirming the existence of two

FIG. 3. (A) Bootscans of the three molecular forms in region 1 (*gag/pol*), with the y axis referring to the bootstrap values and with recombinant breakpoints being marked on the x axis. (B) NJ phylogenetic subtrees of the genome regions between the breakpoints with a subset of analyzed sequences. Forms I (circles), II (triangles), and III (squares) cluster together differently according to their subtypes in each subregion. Letters next to the sequence clusters refer to HIV-1 subtypes.

FIG. 4. (A) NJ phylogenetic tree of HIV-1 region 2 (*vpu/gp120*) nucleotide sequences taken 0 (circles), 3 (diamonds), 6 (squares), 9 (triangles), and 12 (pentagons) months after enrollment. Numbers at branch nodes refer to bootstrap values; only values of >70% are shown. (B) Structures of the molecular forms confirmed by bootscanning and subregion trees (see Fig. 5). (C) Histogram representing the proportion of each molecular form in all sequences of each time point and its fluctuation with time.

distinct A strains. However, intrasubtype recombination events following infection probably led to less divergence between the two A forms compared to the distance between A reference strains and the two A clusters (about 15%). The three pure subtypes were found to be the dominating strains at all time points, with A₂ representing 43%, C representing 28%, and A₁ representing 20% of all sequences. Furthermore, four inter-subtype recombinants (forms IV to VII) and five different A₁A₂ recombinants (combined within form VIII) were identified. Only one sequence could be identified for each of these recombinants. The construction of subregion trees (Fig. 5) with high bootstrap values and visual inspection confirmed that all recombinants emerged through a single crossover event between two of the three parental strains (forms I to III). Fluctuations in the frequencies of the molecular forms were high. Forms I to III were present at visit 0, 1, and 3, but at visit 2 only form III (subtype A₂) persisted and represented almost all sequences. At visit 4, only forms I and III could be detected. The different A₁C and A₂C recombinant forms (IV to VII) emerged and disappeared at various visits, suggesting that none had a distinct selective advantage. The A₁A₂ recombinant strains could be detected at all visits except for the last one.

Region 3 (gp41/nef). The analysis of the fragment spanning gp41 and the beginning of *nef* included 96 sequences. Figure 6 illustrates the structures and temporal fluctuations of the four molecular forms detected for this genome region. The most abundant was form I, a subtype A strain, but three different AC recombinant forms (forms II, III, and IV) were also found. The A and C portions of all of these molecular forms were from the same A and C parental strains, as confirmed by using subregion trees (Fig. 7) and distance analysis. The sequence heterogeneity within each form, as well as the distance between the A portions of the viral variants and between the C portions of forms II, III, and IV, was below 3.0%. In contrast to the other genome regions studied, form I, which was the A subtype, dominated at all time points and represented 84% of all sequences. Form II could represent the parental form for forms III and IV, which would require one and two crossovers, respectively. These were the most rare forms in the quasispecies. The hatched fragment in form IV could not be classified, probably because it was quite short.

Evaluation of multiple PCR primer approach. A separate experiment was set up to explore whether the multiple primer approach had actually augmented the recovery of the diverse viral quasispecies from this multiply infected individual. After amplification of the *vpu*/gp120 fragment from the sample taken at 12 months with the four different primer combinations (a, b, c, and d), the PCR products were cloned separately to identify the strains that were amplified by each of the four reactions. At least 20 clones from each reaction were sequenced. As shown in Fig. 8, the analysis of 103 sequences revealed that the sample contained at least three different molecular forms. Subtype A contributed 81% and subtype C contributed 18% of all analyzed sequences. An AC recombinant of the parental strains was found only once.

The separate analysis of each of the four primer combinations revealed a significant amplification bias. With setups b and d, subtype C could not be detected at all, and with setups a and c and the original experiment, in which all PCR products

were pooled before cloning, subtype C was recovered and represented at least 20% of the analyzed sequences.

Since the number of sequences that can be analyzed by cloning and sequencing is limited, the more sensitive MHAacd assay, which can detect unequal mixtures of 1 in 3,000 (11), was used to prove the presence or absence of the two molecular forms in the four setups. Similar to the sequencing results, the MHA confirmed the presence of strains A and C in setups a and c and the absence of strain C in setups b and d (data not shown).

DISCUSSION

An evaluation of blood samples collected every 3 months during an HISIS participant's last year of life revealed multiple concurrent infections with at least three different HIV-1 strains. MHAacd indicated an AC dual infection. Cloning and sequencing confirmed this finding and, furthermore, distinguished two distinct A strains in addition to the subtype C strain and several recombinant forms. However, the three parental strains could only be recovered by using genome region 2. In this and other multiple infection cases from HISIS studied to date, it has often been observed that the complexity of molecular forms varies by the genome region examined (data not shown), and the fact that more viral diversity was recovered for *vpu*/gp120 than for the other genome regions studied is not unusual: it may be that the recombination and selection of strains "purify" some genome regions more than others in the setting of dual infection. In any case, a highly complex viral quasispecies did emerge during this late-stage infection, including inter- and intrasubtype recombination, with *vpu*/gp120 being the most diverse genomic region. The tendency towards more homogeneity among strains in the late stage of AIDS, as reported earlier (7), could not be observed in this case.

Whether the patient was simultaneously or sequentially infected with the three HIV-1 strains cannot be determined. However, the woman was highly exposed to multiple partners for several years, which makes a superinfection a distinct possibility. Another case of superinfection in this cohort, in a participant who seroconverted during the study, has already been confirmed (McCutchan et al., submitted for publication).

The high viral load of >500,000 copies/ml in all follow-ups was most likely the result of the late-stage HIV-1 infection but may also have been caused by the concurrent infection with multiple HIV-1 strains, as shown earlier (10).

This report confirms that multiple HIV-1 infections can occur and shows that several HIV-1 strains can persist over time, without any single strain gaining predominance. This phenomenon may be more likely when immune responses are waning and there is minimal immune pressure. However, we do not claim to have determined the exact proportions of the circulating forms, and the existence of undetected viral strains is even within the bounds of possibility. The exact amplification of all viral quasispecies can be hampered by several factors, as follows. (i) Every PCR utilizes a limited number of proviral copies. In our case, 27 to 66 proviral DNA copies extracted from 400,000 to 1,000,000 PBMCs were used for the primary PCR (9). (ii) Primer bias exists, which we have tried to minimize but have definitely not eliminated. (iii) Finally, the num-

FIG. 5. Phylogenetic subtrees for recombinant fragments IV, V, VI, and VII (see Fig. 4) of the molecular forms found in region 2 (*vpu/gp120*). Only one or two sequences of each form, the outgroup, and A, C, and D reference sequences are included in the trees. Numbers at branch nodes refer to bootstrap values. The HXB2 location is given for each fragment.

FIG. 6. (A) NJ phylogenetic tree of HIV-1 gp41/nef nucleotide sequences taken 0 (circles), 3 (diamonds), 6 (squares), 9 (triangles), and 12 (pentagons) months after enrollment. Numbers at branch nodes refer to bootstrap values; only values of >70% are shown. Sequence “h” was hypermutated. (B) Structures of the four molecular forms confirmed by bootscanning and subregion trees (see Fig. 7). Numbers below the breakpoints and the edges of fragments indicate HXB2 locations. (C) Histogram representing the proportion of each molecular form in all sequences of each time point and its fluctuation with time.

FIG. 7. Phylogenetic subtrees for the recombinant fragments (see Fig. 6) of the molecular forms in region 3 (gp41/nef). Only one or two sequences of each form, the outgroup, and A, C, and D reference sequences are included in the trees. Numbers at branch nodes refer to bootstrap values. The HXB2 location is given for each fragment.

FIG. 8. (A) NJ phylogenetic tree of HIV-1 *vpu/gp120* nucleotide sequences obtained from visit 4 only. A (diamonds), B (squares), C (triangles), and D (pentagons) refer to the sequences obtained from the corresponding PCR products generated by the four different primer combinations. For the fifth group, all PCR products were pooled (circles) before cloning and sequencing. (B) Structures of the three molecular forms and primer combinations that recovered them.

ber of clones analyzed can have an effect on the amplification of strains.

The choice of primer sequences can bias the recovery of strains and may preclude the detection of multiple infections altogether. To our knowledge, this is the first systematic evaluation of single versus multiple primer approaches for the recovery of diverse quasiespecies. We have demonstrated that two of the primer combinations used for the *vpu/gp120* genome region only amplified the subtype A strain, whereas the other two were able to detect a dual infection by recovering subtypes A and C. Due to the rapid HIV-1 evolution in individuals and the community enhancing viral diversity, mismatches between sample and primer sequences may occur frequently, and thus the multiple primer approach is an efficient method for increasing the likelihood of detecting multiple infections compared to the conventional, single nested PCR.

To ensure a more accurate reproduction of the circulating quasiespecies, even more than two primer combinations per round could be employed in future PCR approaches.

The serial sampling format may also provide better detection of coinfections than a cross-sectional sampling frame, taking into account the high fluctuation and recombination of the molecular forms over time that we described in this case.

In addition, our data show once more that HIV-1 recombination is a common phenomenon in multiply infected individuals. Within individuals, recombination can lead to radically different genomic combinations and will have a much more dramatic impact on viral evolution than do nucleotide substitutions.

We believe that the characterization of the viral quasiespecies emerging within an HIV-1-infected individual during the course of infection and their correlation with host immune responses will be informative in the quest for an HIV-1 vaccine that protects against multiple HIV-1 subtypes.

ACKNOWLEDGMENTS

This work was supported by the European Commission (DG XII, INCO-DC) and by a cooperative agreement between the Henry M. Jackson Foundation for the Advancement of Military Medicine and the U.S. Department of Defense.

The views and opinions expressed herein do not necessarily reflect those of the U.S. Army or the Department of Defense.

We thank the excellent staff at the Mbeya Medical Research Programme who conducted the HISIS study, especially Vera Kleinfeldt, Frowin Nichombe, Weston Assisiya, and Clemence Konkamkula, and all participants in the study.

REFERENCES

- An, W., and A. Telesnitsky. 2002. HIV-1 genetic recombination: experimental approaches and observations. *AIDS Rev.* 4:195–212.
- Anonymous. Los Alamos Databases. [Online.] <http://www.hiv.lanl.gov/>.
- Artenstein, A. W., T. C. VanCott, J. R. Mascola, J. K. Carr, P. A. Hegerich, J. Gaywee, E. Sanders-Buell, M. L. Robb, D. E. Dayhoff, S. Thitvichianlert, et al. 1995. Dual infection with human immunodeficiency virus type 1 of distinct envelope subtypes in humans. *J. Infect. Dis.* 171:805–810.
- Bebenek, K., J. Abbotts, J. D. Roberts, S. H. Wilson, and T. A. Kunkel. 1989. Specificity and mechanism of error-prone replication by human immunodeficiency virus-1 reverse transcriptase. *J. Biol. Chem.* 264:16948–16956.
- Birk, M., S. Aleman, U. Visco-Comandini, and A. Sonnerborg. 2000. Proviral HIV-1 dynamics and evolution in patients receiving efficient long-term antiretroviral combination therapy. *HIV Med.* 1:205–211.
- Boyer, J. C., K. Bebenek, and T. A. Kunkel. 1992. Unequal human immunodeficiency virus type 1 reverse transcriptase error rates with RNA and DNA templates. *Proc. Natl. Acad. Sci. USA* 89:6919–6923.
- Delwart, E. L., H. Pan, H. W. Sheppard, D. Wolpert, A. U. Neumann, B. Korber, and J. I. Mullins. 1997. Slower evolution of human immunodeficiency virus type 1 quasiespecies during progression to AIDS. *J. Virol.* 71:7498–7508.
- Eigen, M. 1993. Viral quasiespecies. *Sci. Am.* 269:42–49.
- Gibellini, D., F. Vitone, P. Schiavone, C. Ponti, M. La Placa, and M. C. Re. 2004. Quantitative detection of human immunodeficiency virus type 1 (HIV-1) proviral DNA in peripheral blood mononuclear cells by SYBR green real-time PCR technique. *J. Clin. Virol.* 29:282–289.
- Grobler, J., C. M. Gray, C. Rademeyer, C. Seioighe, C. Ramjee, S. A. Karim, L. Morris, and C. Williamson. 2004. Incidence of HIV-1 dual infection and its association with increased viral load set point in a cohort of HIV-1 subtype C-infected female sex workers. *J. Infect. Dis.* 190:1355–1359.
- Hoelscher, M., W. E. Dowling, E. Sanders-Buell, J. K. Carr, M. E. Harris, A. Thomschke, M. L. Robb, D. L. Bix, and F. E. McCutchan. 2002. Detection of HIV-1 subtypes, recombinants, and dual infections in East Africa by a multi-region hybridization assay. *AIDS* 16:2055–2064.
- Hoelscher, M., B. Kim, L. Maboko, F. Mhalu, F. von Sonnenburg, D. L. Bix, and F. E. McCutchan. 2001. High proportion of unrelated HIV-1 intersubtype recombinants in the Mbeya region of southwest Tanzania. *AIDS* 15:1461–1470.
- Hoffmann, O., B. Zaba, B. Wolff, E. Sanga, L. Maboko, D. Mmbando, F. von Sonnenburg, and M. Hoelscher. 2004. Methodological lessons from a cohort study of high risk women in Tanzania. *Sex. Transm. Infect.* 80(Suppl. 2):69–73.
- Hu, W. S., and H. M. Temin. 1990. Retroviral recombination and reverse transcription. *Science* 250:1227–1233.
- Janssens, W., A. Buve, and J. N. Kengasong. 1997. The puzzle of HIV-1 subtypes in Africa. *AIDS* 11:705–712.
- 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.
- Kijak, G. H., V. Simon, P. Balfe, J. Vanderhoeven, S. E. Pampuro, C. Zala, C. Ochoa, P. Cahn, M. Markowitz, and H. Salomon. 2002. Origin of human immunodeficiency virus type 1 quasiespecies emerging after antiretroviral treatment interruption in patients with therapeutic failure. *J. Virol.* 76:7000–7009.
- Kiwelu, I. E., B. Renjifo, B. Chaplin, N. Sam, W. M. Nkya, J. Shao, S. Kapiga, and M. Essex. 2003. HIV type 1 subtypes among bar and hotel workers in Moshi, Tanzania. *AIDS Res. Hum. Retrovir.* 19:57–64.
- Koulinska, I. N., T. Ndung'u, D. Mwakagile, G. Msamanga, C. Kagoma, W. Fawzi, M. Essex, and B. Renjifo. 2001. A new human immunodeficiency virus type 1 circulating recombinant form from Tanzania. *AIDS Res. Hum. Retrovir.* 17:423–431.
- Kumar, S., K. Tamura, I. B. Jakobsen, and M. Nei. 2001. MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* 17:1244–1245.
- Louwagie, J., W. Janssens, J. Mascola, L. Heyndrickx, P. Hegerich, G. van der Groen, F. E. McCutchan, and D. S. Burke. 1995. Genetic diversity of the envelope glycoprotein from human immunodeficiency virus type 1 isolates of African origin. *J. Virol.* 69:263–271.
- Michael, N. L., G. Chang, P. K. Ehrenberg, M. T. Vahey, and R. R. Redfield. 1993. HIV-1 proviral genotypes from the peripheral blood mononuclear cells of an infected patient are differentially represented in expressed sequences. *J. Acquir. Immune Defic. Syndr.* 6:1073–1085.
- Nowak, P., A. C. Karlsson, L. Naver, A. B. Bohlin, A. Piasek, and A. Sonnerborg. 2002. The selection and evolution of viral quasiespecies in HIV-1 infected children. *HIV Med.* 3:1–11.
- Paolucci, S., F. Baldanti, G. Campanini, M. Zavattoni, E. Cattaneo, L. Dossena, and G. Gerna. 2001. Analysis of HIV drug-resistant quasiespecies in plasma, peripheral blood mononuclear cells and viral isolates from treatment-naïve and HAART patients. *J. Med. Virol.* 65:207–217.
- Piyasirisilp, S., F. E. McCutchan, J. K. Carr, E. Sanders-Buell, W. Liu, J. Chen, R. Wagner, H. Wolf, Y. Shao, S. Lai, C. Beyrer, and X. F. Yu. 2000. A recent outbreak of human immunodeficiency virus type 1 infection in southern China was initiated by two highly homogeneous, geographically separated strains, circulating recombinant form AE and a novel BC recombinant. *J. Virol.* 74:11286–11295.
- 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.
- Riedner, G., M. Rusizoka, O. Hoffmann, F. Nichombe, E. Lyamuya, D. Mmbando, L. Maboko, P. Hay, J. Todd, R. Hayes, M. Hoelscher, and H. Grosskurth. 2003. Baseline survey of sexually transmitted infections in a cohort of female bar workers in Mbeya Region, Tanzania. *Sex. Transm. Infect.* 79:382–387.
- Robertson, D. L., P. M. Sharp, F. E. McCutchan, and B. H. Hahn. 1995. Recombination in HIV-1. *Nature* 374:124–126.
- Salminen, M. O., J. K. Carr, D. S. Burke, and F. E. McCutchan. 1995. Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. *AIDS Res. Hum. Retrovir.* 11:1423–1425.
- Smith, S. W., R. Overbeek, C. R. Woese, W. Gilbert, and P. M. Gillevet. 1994.

- The genetic data environment: an expandable GUI for multiple sequence analysis. *Comput. Appl. Biosci.* **10**:671–675.
31. **Vesanen, M., M. Salminen, M. Wessman, H. Lankinen, P. Sistonen, and A. Vaheri.** 1994. Morphological differentiation of human SH-SY5Y neuroblastoma cells inhibits human immunodeficiency virus type 1 infection. *J. Gen. Virol.* **75**:201–206.
32. **Visco-Comandini, U., S. Aleman, Z. Yun, and A. Sonnerborg.** 2001. Human immunodeficiency virus type 1 variability and long-term non-progression. *J. Biol. Regul. Homeost. Agents* **15**:299–303.
33. **Wain-Hobson, S.** 1992. Human immunodeficiency virus type 1 quasispecies in vivo and ex vivo. *Curr. Top. Microbiol. Immunol.* **176**:181–193.
34. **Zhuang, J., A. E. Jetzt, G. Sun, H. Yu, G. Klarmann, Y. Ron, B. D. Preston, and J. P. Dougherty.** 2002. Human immunodeficiency virus type 1 recombination: rate, fidelity, and putative hot spots. *J. Virol.* **76**:11273–11282.
35. **Ziegler, J. S., Y. Su, K. P. Corcoran, L. Nie, P. E. Mayrand, L. B. Hoff, L. J. McBride, M. N. Kronick, and S. R. Diehl.** 1992. Application of automated DNA sizing technology for genotyping microsatellite loci. *Genomics* **14**:1026–1031.