JOURNAL OF VIROLOGY, Oct. 1994, p. 6340–6346
0022-538X/94/$04.00+0

Vol. 68, No. 10

Identification of Human Immunodeficiency Virus Type 1 Envelope Genes Recombinant between Subtypes B and F in Two Epidemiologically Linked Individuals from Brazil

ESTER C. SABINO,[1,2] EUGENE G. SHPAER,[3]† MARIZA G. MORGADO,[1,4] BETTE T. M. KORBER,[5,6] RICARDO S. DIAZ,[1,7] VERA BONGERTZ,[4] SOLANGE CAVALCANTE,[8] BERNARDO GALVÃO-CASTRO,[9] JAMES I. MULLINS,[3] AND ALLEN MAYER[1]*

Irwin Memorial Blood Centers, San Francisco, California 94118[1]; Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, California 94305-5402[3]; Los Alamos National Laboratory, Los Alamos, New Mexico 87545[5]; Santa Fe Institute, Santa Fe, New Mexico 87501[6]; and Laboratorio de Retrovirus, Instituto Adolfo Lutz, São Paulo, São Paulo 01246,[2] Department of Immunology[4] and Hospital Evandro Chagas,[8] Instituto Oswaldo Cruz, Fiocruz, Rio de Janeiro 21045, Escola Paulista de Medicina, DIPA, São Paulo, São Paulo 04023,[7] and Lab Avancado de Saude Publica, Centro de Pesquisa Goncalo Moniz, Fiocruz, Bahia 45945,[9] Brazil

Received 9 May 1994/Accepted 19 July 1994

Sequence analysis of a human immunodeficiency virus type 1 *env* gene PCR amplified from a Brazilian woman's peripheral blood mononuclear cell DNA (sample RJIO1) showed that it was likely to have been derived from a double recombination event between human immunodeficiency virus type 1 subtypes B and F. The major portion of the gp120 coding sequence belonged to the B lineage, but a segment of the C2 to V3 region (approximately 135 nucleotides) clearly associated with sequences of the F lineage. The subtype F-like segment had 15 noncontiguous signature nucleotides in common with Brazilian subtype F sequences that were not found, or were rare, in subtype B sequences. In contrast, this same segment had only 3 signature nucleotides shared with subtype B sequences and not present in the Brazilian subtype F sequences. Phylogenetic analysis, amino acid signature pattern analysis, and the pattern of synonymous mutations all supported the hypothesis of a recombinational origin of the RJIO1 sequence. Related recombinant genes were also detected in peripheral blood mononuclear cell DNA obtained from the woman's recent sexual partner, indicating that the recombination event probably occurred at some previous time in the chain of virus transmission. Divergent viral sequences in the V3 region were found in the male sexual partner, while a relatively homogeneous viral population was detected in the woman, consistent with her recent infection.

Recombination of retroviral genomes is a well-documented phenomenon that can result in increased viral diversity (16). Recombination of human immunodeficiency virus type 1 (HIV-1) genomes occurring in vivo between different members of the viral population, or quasispecies, present within an infected individual has been inferred from analyses of nucleotide sequences (2, 11, 30). Recombination between different HIV-1 strains of the same subtype, and between different HIV-1 subtypes, could conceivably take place within an individual who had been exposed to and infected by HIV-1 more than once. We report here on envelope genes detected in two epidemiologically linked Brazilian individuals that appear to be recombinants between subtype B and subtype F HIV-1.

A prior analysis of 28 Brazilian HIV-1 strains had shown that, on the basis of the V3 loop and flanking sequences in *env*, most could be classified as subtype B but that one sample (RJIO3) contained a subtype F HIV-1 provirus (23). Proviral sequences within the peripheral blood mononuclear cells (PBMC) from another of these Brazilian samples, RJIO1, were found to have a V3 loop and flanking regions similar to subtype B sequences, except for a short region upstream of the V3 loop in the C2 region which bore a strong similarity to the corresponding region of the subtype F sequence. This suggested that the RJIO1 genome might be a recombinant between subtype B and subtype F and that what we were observing in the C2 region was the 3′ crossover point. We therefore embarked on more extensive analysis of the putative recombinant provirus (RJIO1) and proviruses from this individual's infected sexual partner, as well as Brazilian subtype F (RJIO3) and subtype B (SP3) proviruses, each of which is reported here.

MATERIALS AND METHODS

PBMC samples. Sample RJ548 was obtained from a seropositive man (CDC stage IV) in April 1992. At the same time, a sample (RJ549) was also obtained from this man's female sexual partner (CDC stage II/A). In April 1992, sexual contact between the two had been taking place for less than 6 months. Sample RJ549 was the initial seropositive sample from the female, in that its Western blot (immunoblot) was strongly positive only for core, negative for TM, and very weak for Gag and Env precursor proteins. Sample RJIO1 was obtained from the same woman in August 1992. Serum from this latter time

* Corresponding author. Mailing address: Hudson Valley Blood Services, Grasslands Rd., Valhalla, NY 10595. Phone: (914) 592-6000. Fax: (914) 592-1577.

† Present address: Perkin-Elmer Applied BioSystems, Foster City, CA 94404.

point showed strong reactivity for viral core and Gag and Env precursor proteins but weaker reactivity for viral SU and TM proteins by Western blot. Sample RJIO3, a representative of the Brazilian F subtype, was obtained from another woman (stage II) also infected by the heterosexual route. All three individuals were residents of Rio de Janeiro, Brazil. The SP3 sequence, a representative of the Brazilian B subtype, was obtained from an individual from São Paulo (23).

Sequencing of PCR products. DNA was extracted from Ficoll-Hypaque-purified PBMC by proteinase K digestion followed by phenol and chloroform extraction and ethanol extraction. DNA (50 to 100 ng) of samples RJIO1 (amplified in duplicate to yield clones 5 and 10), RJIO3, and SP3 was subjected to *env*-specific PCR using a nested protocol to amplify the V1 to V5 region of *env*, as described by Delwart et al. (4), using ED3 and ED14 primers in the first round (HXB2 positions 5956 to 5986 and 7936 to 7966, respectively). Second-round primers, ED5 (5′-CAUCAUCAUCAUAUGGGAU CAAAGCCUAAAGCCAUGUG; HXB2 positions 6562 to 6588) and ED12 (5′-CUACUACUACUAAGUGCUUCCUG CUGCUCCCAAGAACCCAAG; HXB2 positions 7792 to 7822), were modified as shown here to contain uracil instead of thymidine as well as additional nucleotides at the 5′ end for use with the CloneAmp system (GIBCO/BRL, Gaithersburg, Md.). PCR products of approximately 1,200 bp encompassing V1 through V5 of *env* were digested with uracil DNA glycosylase (to generate single-stranded ends and to digest primer dimers), annealed to the pAMP vector (GIBCO/BRL), and then directly transformed into DH5α bacteria (GIBCO/BRL). Plasmid DNA from one colony per transformation was prepared with Qiagen-tip 20 columns (Qiagen, Chatsworth, Calif.) and used as the template for dideoxy chain termination sequencing (Sequenase, version 2.0; U.S. Biochemicals, Cleveland, Ohio). For sample RJIO1, a second nested PCR round was also performed using primers SC27 (5′-GACAGTGGT CATGAGAGTGAGGGGGATCAGGAG; HXB2 positions 6214 to 6246) and ES4 (5′-TATGGGAATTGGCTCAAAGG; HXB2 positions 6849 to 6868) to obtain the gp120 sequence upstream of V1. This PCR product was ligated into plasmid pCRII (InVitrogen Corp., San Diego, Calif.) which was used to transform One Shot INV alpha F′ competent cells (InVitrogen). Plasmid DNA from a single colony was used for sequencing as described above.

The sequences of individual proviral genomes from samples RJ548 and RJ549 were obtained by end-point dilution cloning performed prior to PCR, as follows. PBMC DNAs (about 1 μg/ml) from samples RJ548 and RJ549 were serially diluted to the point that about one in five or fewer reactions yielded amplified product (29). First-round primers were 5′-ATAAGC TTCAATGTACACATGGAATT (HXB2 positions 6959 to 6976) and 5′-ATGAATTCATTACAGTAGAAAAATTCCC (HXB2 positions 7362 to 7381), and second-round primers spanning the V3 region of *env* (32) were 5′-CAUCAUCAUC AUGCAGUCUAGCAGAAGAAGA (HXB2 positions 7010 to 7029) and 5′-CUACUACUACUAUUCUGGGUCCCCU CCUGAGGA (HXB2 positions 7313 to 7333) modified as shown here for use with the GIBCO/BRL CloneAmp kit. Plasmid DNAs from single colonies were used as templates for sequencing as described above.

Phylogenetic analysis. DNA and protein alignments were generated and refined with the multiple aligned sequence editor MASE (7). Because of the highly skewed base composition of HIV and the asymmetrical substitutional frequencies of mutations from one base to another (for example, A-to-G substitutions were 10 times more frequent than C-to-G substitutions), we used weighted parsimony for generation of the phylogenetic trees presented in this paper (12, 15). This was particularly critical given the short region between the subtype F recombination points, as a small number of changes may have great impact on the structure of the tree. PAUP, version 3.1.1 (31), was used to generate these trees in conjunction with MacClade, version 3.03 (21). The frequencies of character state changes (f) were calculated for a phylogenetic tree based on the 18 taxa included in Fig. 2, over intact sequences encompassing three distinct regions (defined in Fig. 1) of the gp120 sequence, and produced from a preliminary parsimony analysis using PAUP and MacClade. The conversion $1/f$ was used to weight the possible nucleotide changes, with truncation to avoid violations of the triangle inequality (21). The resulting matrix was then included as a character type assumption for subsequent phylogenetic reconstructions using PAUP (31). The construction of new most parsimonious trees minimized the sum of the branch lengths calculated as the number of character changes multiplied by the weights of the respective character changes. The g_1 statistic for the phylogenetic analysis using parsimony was low for each of the three regions under consideration in this study (all three regions using the 18 taxa shown in Fig. 2 had g_1 statistic values less than -0.7, giving confidence limits greater than 99% that the data are nonrandom [14]). The g_1 statistic is a measure of the phylogenetic signal in the data (14), and a low negative value suggests that parsimony is an appropriate phylogenetic tool for application to this sequence set. Bootstrap proportions (8, 13) were calculated with weighted parsimony (31). All trees were based on alignments from which columns containing gaps inserted to maintain the alignment were deleted.

Phylogenetic trees were also constructed with PHYLIP, version 3.5 (9). Bootstrap analysis was done using programs SEQBOOT (to generate 100 reshuffled sequences), DNADIST (maximum-likelihood model), NEIGHBOR, and CONSENSE.

VESPA and synonymous site analyses. Viral amino acid signature pattern analysis (VESPA) was conducted as previously described (17). Synonymous substitution rates were calculated on the basis of the method of Nei and Gojobori (27). P_s is the number of observed synonymous substitutions divided by the number of possible synonymous substitutions.

Nucleotide sequence accession numbers. GenBank accession numbers for sequences presented here are U08955 to U08960, U08962 to U08975, and U10019 to U10026.

RESULTS

A 1,200-bp fragment spanning V1 through V5 of the HIV-1 *env* gene was amplified, cloned, and sequenced from PBMC DNA obtained from three Brazilian samples: (i) sample RJIO1, suspected to contain a provirus recombinant between subtype B and subtype F; (ii) sample RJIO3, containing a subtype F provirus (23); and (iii) sample SP3, containing a subtype B provirus. For RJIO1, the entire gp120 coding sequence was obtained. When these sequences were aligned with each other, with subtype B sequences from the database (25), and with additional Brazilian subtype B and F sequences (kindly provided by S. Osmanov and F. McCutchan), it became apparent that a stretch of approximately 135 bp in the C2 to V3 region of the RJIO1 sequence (from about positions 6958 to 7093 on the HXB2 map) corresponded closely to the subtype F sequences, while the surrounding regions corresponded to subtype B sequences. Figure 1, focusing on the C2 to V3 region, shows that 15 noncontiguous bases within the approximately 135-nucleotide F-like region are common between the subtype F and the RJIO1 sequences but are absent or rare in

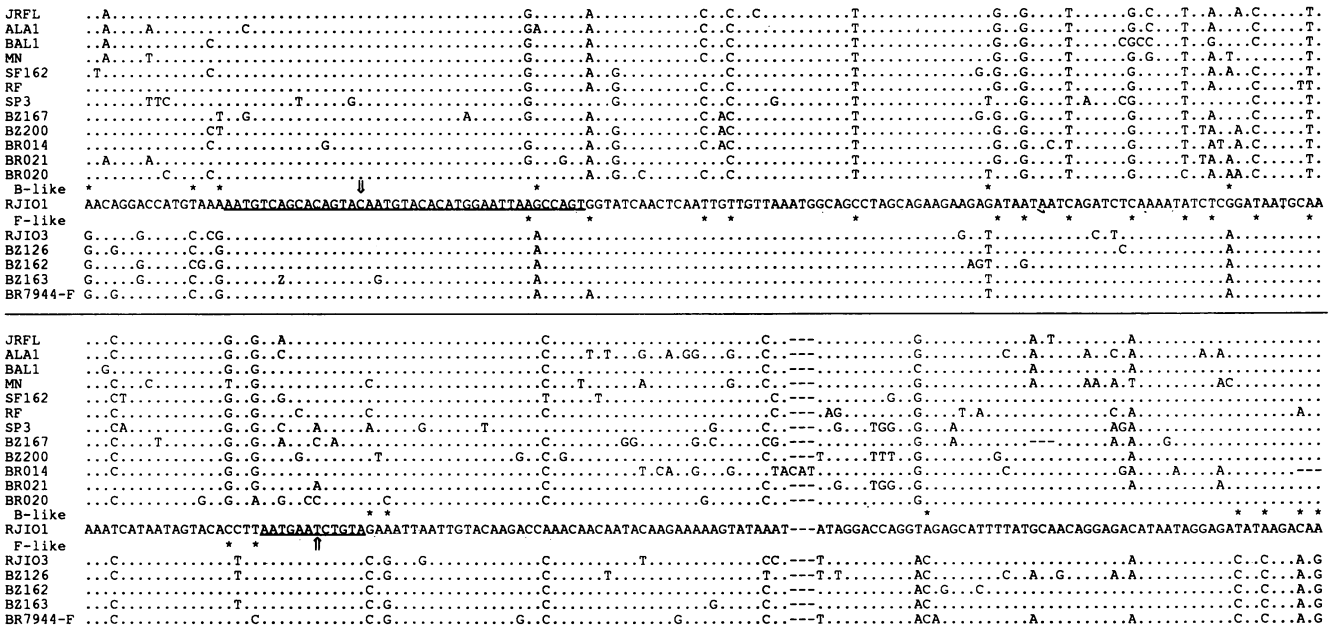

FIG. 1. Nucleotide alignment in the central, C2 to V3, region of HIV-1 *env* (HXB2 positions 6928 to 7207). The putative recombinant sequence RJIO1 is shown, with nucleotides from other sequences identical to RJIO1 represented by dots. Nucleotides in RJIO1 which are found in most subtype B or F sequences but not at all, or rarely, in the other subtypes are indicated by asterisks. Gaps introduced to maintain alignment are indicated by dashes. The underlined sequences indicate regions within which the recombination events likely occurred. The arrows within these regions (at positions 6958 and 7093 on the HXB2 map) separate the middle F-like segment of the RJIO1 sequence from the B-like upstream and downstream flanking segments and demarcate the boundaries used in the phylogenetic, VESPA, and synonymous site analyses shown in Fig. 2, 3, and 4. The major portions of the upstream and downstream flanking segments used in these other analyses are not shown.

subtype B *env* sequences from Brazil and elsewhere. In contrast, only 3 bases within this region are shared between the RJIO1 sequence and subtype B sequences and absent from the subtype F sequences.

The sequence of RJIO1 used in Fig. 1 was from a clone (clone 5) derived from a PCR that contained more than one target provirus, and since recombination has been demonstrated to occur during PCR and subsequent bacterial cloning when distinct target sequences are amplified together (3, 22), it was important to show that the observed recombinant genome was actually present in a provirus and did not arise artifactually during the PCR and/or the cloning process. With this in mind, RJIO1 PBMC DNA was subjected to an independent nested PCR. The V1 through V5 sequence of a clone (clone 10) obtained from this independent amplification experiment was identical to the first sequence, except for 7 scattered single-base differences that did not affect the subtype character of the sequence. This shows that the putative recombinant genome was indeed present in RJIO1 proviral DNA, and its detection in both sequences suggests that it was a prevalent sequence in the individual at the time of sampling.

Figure 2 shows three separate phylogenetic trees for three regions of the *env* gene: the subtype F-like central region in the putative recombinant (panel B) and its two flanking regions (panels A and C). The trees include the putative recombinant sequence, Brazilian subtype B and F sequences, and representatives of the subtype A, B, C, D, and E sequences. A simian immunodeficiency virus cpz strain was used to root the trees. Bootstrap proportions of greater than 50 of 100 replicates are shown.

The sequence RJIO1 clusters reliably (in 92 of 100 bootstrap repetitions) with subtype F genomes in the central region (Fig. 2B) and (in 81 and 90 of 100 bootstrap repetitions, respectively) with subtype B genomes in the upstream (Fig. 2A) and downstream (Fig. 2C) flanking regions. At the same time, the F branch in both panels A and C has a bootstrap value of 100. In the phylogenetic tree for the whole region (1,041 nucleotides), the RJIO1 provirus was closely associated (bootstrap value of 86) with an artificial recombinant sequence in which upstream and downstream flanking regions were from BZ167 (subtype B) and the central region was from BZ163 (subtype F) (data not shown). Although we show weighted parsimony analyses here, we have also performed bootstrap tests of neighbor-joining trees using a larger number of taxa, including 28 available B-type gp120 sequences (25) (data not shown). This method gave similarly strong support of an RJIO1 phylogenetic association with subtype B sequences in the upstream and downstream flanking regions (bootstrap values of 99 and 100, respectively) and with subtype F sequences in the middle segment (bootstrap value of 94). Phylogenetic analysis, therefore, strongly supports the hypothesis that the RJIO1 provirus was derived by recombination between viruses with subtype B and F *env* genes.

That the central region of RJIO1 derives from subtype F is supported by two additional very different approaches: by looking at amino acid signature patterns (Fig. 3) and by exclusively looking at synonymous (translationally silent) substitutions (Fig. 4). Both of these approaches were employed to examine the possibility that the observed F-like character and phylogenetic association of the region upstream of the V3 loop in RJIO1 resulted from parallel evolution or convergence over a relatively short domain of protein sequence due to structural or functional constraints and is not the result of a recombination event. As the putative recombination region is very short,

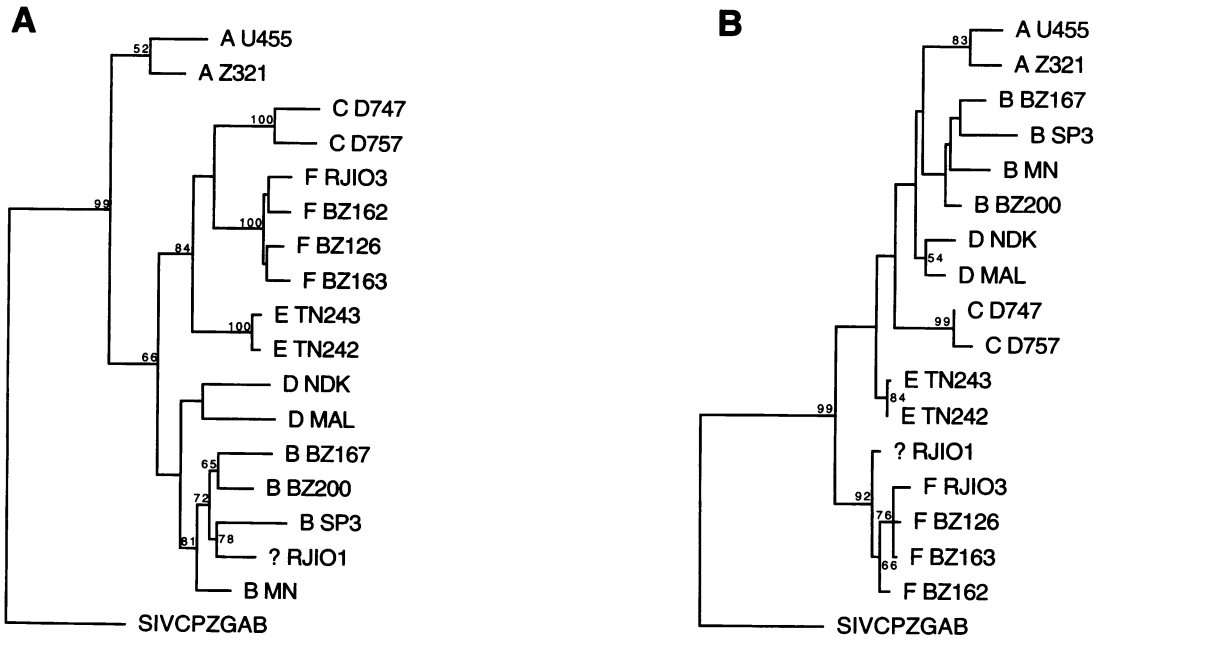

FIG. 2. Phylogenetic relationships of regions of the RJIO1 sequence and sequences representing the major HIV-1 clades defined for the envelope gene. Alignments of three gene segments, one corresponding to the middle F-like segment delineated in Fig. 1 and the other two representing the two flanking regions, were created and analyzed by weighted parsimony. Each tree shows branch lengths of the most parsimonious tree based on 10 randomized input orders of the taxa, using heuristic searches with the SPR swapping algorithm (30). The numbers shown at the nodes represent bootstrap proportions higher than 50 of 100 replicates. (A) Tree generated for the downstream flanking region, including 170 varied sites in 297 total; (B) tree generated for the middle segment, including 72 varied sites in 135 total; (C) tree generated for the downstream region, with 372 varied sites in 609 total. SIVCPZGAB, simian immunodeficiency virus strain used to root the trees.

homoplasy and phenotypic selection could conceivably be dominating the sequence conservation that is causing the phylogenetic association of the sequences. Both of the additional analyses, however, support the recombination hypothesis, as follows.

By analyzing amino acid signature patterns over the putative recombinant region upstream of the V3 loop (Fig. 3), we could quickly screen the available database of C2 to V3 region sequences to determine if any of the other non-F-subtype sequences had similar F-like amino acids in this proposed recombination region. A Brazilian subtype F amino acid signature pattern was defined on the basis of a comparison of four Brazilian subtype B sequences and five Brazilian subtype F sequences. (RJIO1 was excluded.) Signature amino acids were defined with the program VESPA (17) and were required to have 100% conservation among the subtype F sequences and not be the most common amino acid among subtype B sequences at a given position in the alignment. The program SPCOUNT (17) was then used to search the available database of V3 region amino acid sequences from 559 individuals (25). Of these sequences, 342 were associated with subtype B, and 10 closely related sequences from Romania belonged to subtype F (7). RJIO1 shared more of the subtype F amino acid signature sites than any of the 549 non-F-subtype sequences (Fig. 3). There were 6 subtype F signature amino acids in the region analyzed, and RJIO1 shared six of six amino acids with this signature pattern. The 10 Romanian subtype F sequences shared 5 of 6 amino acids. The next highest degree of similarity was found in a single subtype B sequence and two of the subtype A sequences, sharing 4 of 6 signature amino acids. Therefore, the amino acids that are particularly characteristic of the Brazilian subtype F are perfectly preserved in the region upstream of the V3 loop in RJIO1, and these amino acids are not preserved in any subtype B sequences. We therefore conclude that the signature does not represent a biologically favored F-like common form of the region that is preserved among a subset of subtype B sequences but rather that it is uniquely found in RJIO1 and subtype F sequences.

The number of synonymous substitutions in the three re-

A)

```
Brazilian F subtype signature:     ....Q.ISD.......H...S
RJ I01                             IIRSQNISDNAKIIIVHLNES
```

B)

Frequencies of matches	F subtype	B subtype
6/6	RJ I01	0
5/6	10 (Romania)	0
4/6	0	1
3/6	0	28
2/6	0	138
1/6	0	123
0/6	0	27

FIG. 3. Signature sequence analysis. The V3 region amino acid sequence compilation in the Human Retroviruses and AIDS database (24) derived from 559 individuals overlaps with the downstream portion of the putative F-like recombinant C2 to V3 region of the RJIO1 sequence. (A) The region of overlap, aligned with six Brazilian subtype F signature amino acids defined by contrasting subtype F and B sequences from Brazil; (B) numbers of subtype B and F sequences with a given frequency of shared amino acids with the signature. Not all sequences in the database extended far enough to cover the entire signature pattern. Only those that had sequence information at all six sites are included in this tally; none of the partial subtype B sequences had high levels of conservation of the subtype F signature amino acids.

gions (the central putative recombinant region and the two flanking regions) were then compared, contrasting B × F, B × B, F × F, F × RJIO1, and B × RJIO1 sequence comparisons (Fig. 4). Synonymous substitutions are generally considered to be under less stringent selective pressure than nonsynonymous substitutions (26), although they should not be considered entirely neutral (1, 5). Thus, in some circumstances they may be a more reliable indicator of phylogenetic relationships. It was particularly important to examine the rate of synonymous substitutions separately, because the recombinant portion of F-like genome found in RJIO1 was relatively short and convergence might be possible over such a limited region. The synonymous-substitution rate is a distance measure that is not subject to the influence of convergence at the protein level and thus helps differentiate between recombination and convergence or parallelism. RJIO1 showed synonymous-substitution frequencies typical for a subtype B sequence in the flanking regions and typical of a subtype F sequence in the central region in question, as expected.

Sample RJIO1 was obtained from a woman who entered and maintained a sexual relationship with an HIV-1-infected individual for approximately 10 months prior to this sampling. In order to determine whether recombination had occurred in the infected female or whether recombinant genomes had been transmitted to her at the time of infection, *env* V3 region DNA (corresponding to nucleotides 7029 to 7313 of the HXB2 genome), which includes the 3′ crossover point, was amplified and sequenced from PBMC DNA of samples RJ549 and RJ548. Sample RJ549 was obtained from the woman 4 months earlier than sample RJIO1, when her serological response was consistent with a recent seroconversion (see Materials and Methods). RJ548 was a sample from her male sexual partner obtained at the same time as RJ549. All of 17 sequences from the male and all of 7 sequences from the female were found to be recombinants in that they closely corresponded to the recombinant genome found in sample RJIO1 (Fig. 5). No fully B or F putative parental genomes were detected in this limited sampling.

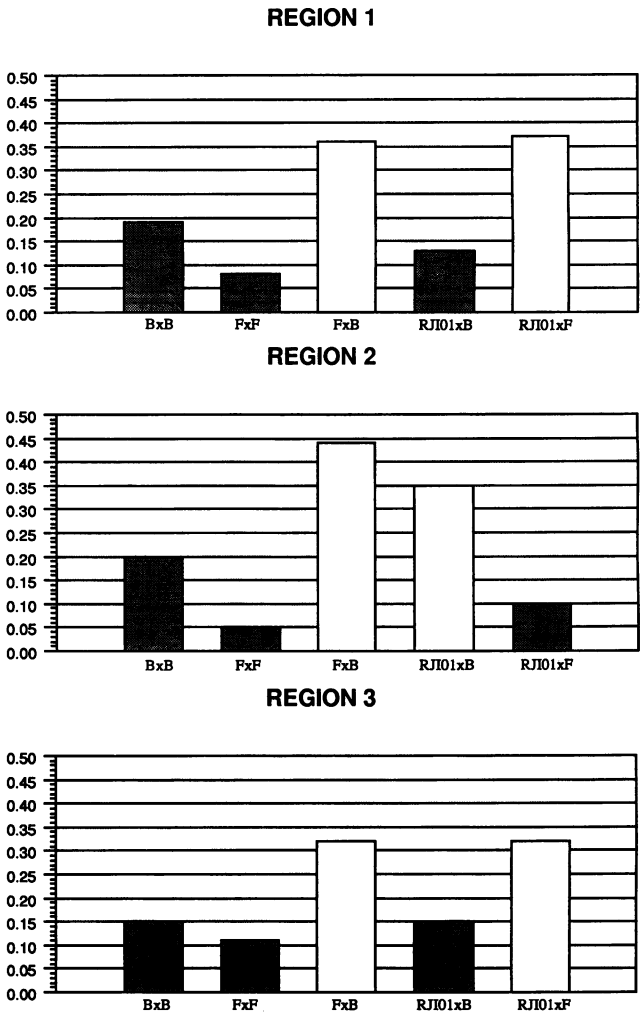

FIG. 4. Synonymous substitution calculations comparing RJIO1 and subtype B and F sequences. The sequences included here were the same subtype B and F sequences included in the phylogenetic analysis shown in Fig. 2. All pairwise sequence comparisons were made by calculating the observed number of synonymous substitutions divided by the possible number of synonymous substitutions (P_s) (26) over upstream flanking region 1, middle region 2, and downstream flanking region 3. The heights of the bars are the average P_s values for each set of comparisons, and the ranges for the bars from left to right are as follows: for region 1, 0.13 to 0.27, 0.03 to 0.11, 0.30 to 0.45, 0.09 to 0.18, and 0.33 to 0.41; for region 2, 0.06 to 0.32, 0.03 to 0.07, 0.37 to 0.51, 0.29 to 0.40, and 0.07 to 0.10; for region 3, 0.10 to 0.18, 0.07 to 0.14, 0.27 to 0.37, 0.10 to 0.17, and 0.30 to 0.33.

The sequence similarities between this male and female support the epidemiologic linkage of the pair. Furthermore, the relative homology of the sequences from the female in the initial RJ549 sample (average, 2.7%; range, 0 to 4.9%) compared to that from the male (average, 5.9%; range, 0.4 to 9.9%) (Fig. 5) supports the epidemiologic and serologic evidence that the male was the source of the female's infection, in that it is in agreement with other studies of linked samples close to the time of seroconversion (33, 34), when nearly homogeneous envelope viral sequences are present in the recipient soon after infection as compared to divergent sequences at later times of infection. Thus, our results indicate that

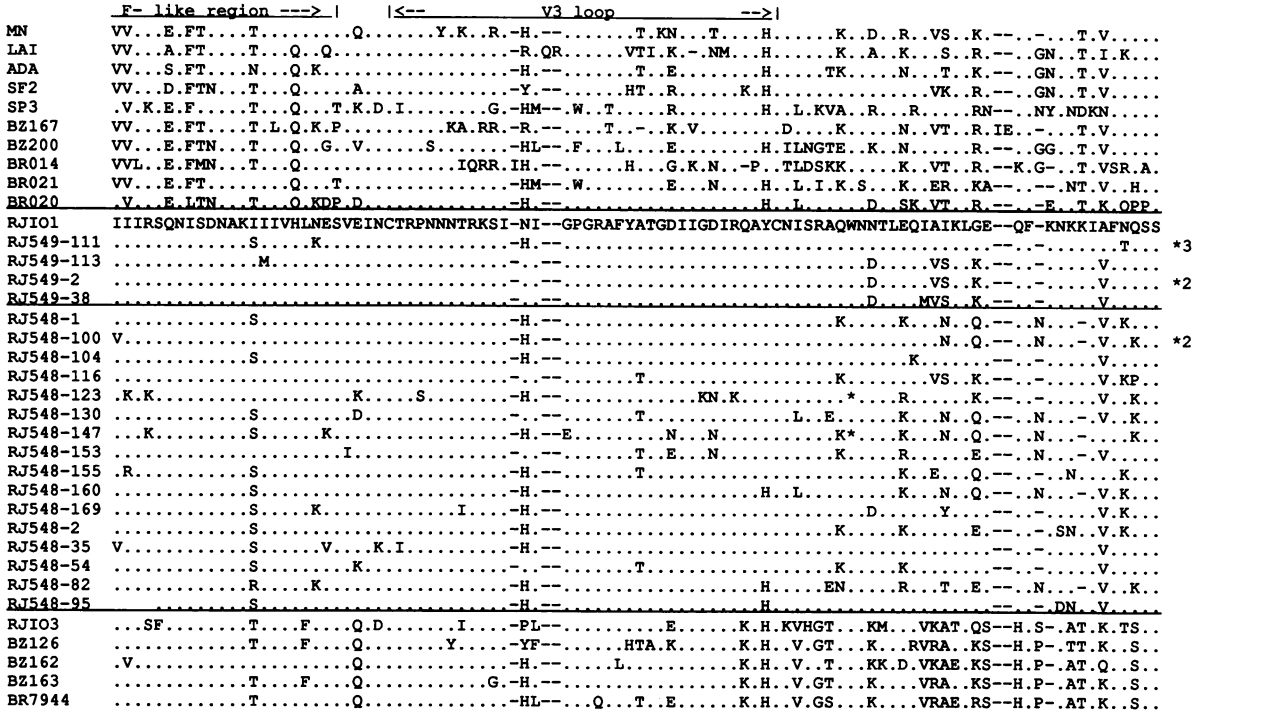

FIG. 5. Amino acid alignment for HIV-1 V3 region, as predicted by nucleotide sequence. Ten sequences representing *env* subtype B (including six from Brazil) are shown on the top; five Brazilian subtype F sequences are on the bottom. A set of 8 B-F recombinant sequences from the infected woman (samples RJIO1 and RJ549) and 17 from her male sexual partner (sample RJ548) are shown in the middle. Numbers to the right show the number of identical protein sequences. An asterisk in the sequence denotes a stop codon. Dots indicate identical amino acids; dashes indicate gaps introduced to maintain alignment.

the proposed recombination event took place at some point in the chain of infection prior to infection of the index female.

DISCUSSION

While only infrequently documented, recombination in vivo may be an important contributory factor to the rapid generation of sequence diversity seen in HIV-1 over the course of infection within an individual (2, 11) and also over the course of the epidemic (1). Recombination occurring between viruses from different subtypes could pose a further challenge to the already complex task of developing vaccines effective against the multiple subtypes of this rapidly evolving pathogen. In addition, recombination between viruses of distinct strains or lineages would suggest that serial HIV-1 infections in individuals can occur, which in itself would bode ill for vaccine efforts. We report here that recombination between subtypes is indeed occurring.

We have identified infection of two epidemiologically linked Brazilian individuals with virus derived from recombination between parental viruses harboring two different subtypes of the HIV-1 envelope gene (B and F). Recombinant proviruses were found in the index case, a female, as well as in the sexual partner who was the source of her infection. Extensive analysis of the recombinant genome in the index female showed that the major portion of its envelope gp120 belongs to the B lineage, but a portion of the gp120 central region is of subtype F origin. Recombinant genomes were the only viral sequences detected in both individuals, neither of the putative parental genomes being represented in this sampling from these two individuals. Presumably, dual infection with HIV-1 of both subtypes B and F, and subsequent recombination, took place in a person earlier in the chain of HIV-1 spread within the Brazilian population, or possibly at an earlier time in the male studied here.

These results provide evidence that (i) at least in rare cases individuals can be dually infected by genomes of HIV-1 subtypes B and F, (ii) individual cells within such a dually infected person can be superinfected by these two different HIV-1 subtypes and (iii) recombination can occur in vivo between HIV-1 genomes of these different subtypes.

There is at least one other case of a published HIV-1 sequence which seems to be recombinant between subtypes. The *gag* gene in the laboratory HIV_{MAL} strain belongs to subtype A, while the *env* gene belongs to subtype D (19). In addition, a genome recombinant between different HIV-2 subtypes has been detected in an individual in Africa (10). HIV-1 strains detected in Thailand, which are characterized as having the subtype E *env* gene but the subtype A *gag* gene, may also have originated from a recombination event or events between subtype A HIV-1 and a not-yet-detected HIV-1 strains having both a subtype E *env* gene and a putative subtype E *gag* gene (24, 25). An alternative explanation that does not involve recombination is that the Thai viruses evolved from subtype A virus very slowly in *gag* but diverged relatively rapidly in *env*, to the degree that the *env* gene is now classified as a different subtype.

Although these results provide evidence for dual infection by different HIV-1 subtypes, they do not address the question of whether dual infection with two HIV-1 strains of the same subtype can occur. It is still possible that humoral and/or cellular immunity interferes with dual infection of an individual by two different strains of the same HIV-1 subtype, or even

by different subtypes of HIV-1. Nevertheless, these findings do show that in areas of the world such as Brazil where two subtypes of HIV-1 are coprevalent (4, 20, 23, 28), dual infection and recombination between subtypes can and do occur.

ACKNOWLEDGMENTS

This work was supported in part by grant RO1-HL-48367 from the National Heart, Lung, and Blood Institute. Ricardo S. Diaz's stay in the United States was funded by a Fogarty Foundation International grant, through the School of Public Health, University of California, Berkeley.

We thank Francine E. McCutchan of the Henry M. Jackson Foundation, Rockville, Md., and Saladin Osmanov of the World Health Organization, Geneva, Switzerland, for supplying Brazilian sequences prior to publication.

REFERENCES

1. **Coffin, J. M.** 1992. Genetic diversity and evolution of retroviruses. Curr. Top. Microbiol. Immunol. **176:**143–164.
2. **Delassus, S., R. Cheynier, and S. Wain-Hobson.** 1991. Evolution of human immunodeficiency virus type 1 *nef* and long terminal repeat sequences over 4 years in vivo and in vitro. J. Virol. **65:**225–231.
3. **Delwart, E. L.; H. W. Sheppard, B. D. Walker, J. Goudsmit, and J. Mullins.** 1994. Human immunodeficiency virus type 1 evolution in vivo tracked by DNA heteroduplex mobility assays. J. Virol. **68:**●●●
4. **Delwart, E. L., E. G. Shpaer, F. E. McCutchan, J. Louwagie, M. Grez, H. Rübsamen-Waigmann, and J. I. Mullins.** 1993. Genetic relationships determined by a heteroduplex mobility assay: analysis of HIV *env* genes. Science **262:**1257–1261.
5. **Domingo, E., and J. J. Holland.** 1994. Mutation rates and rapid evolution of RNA viruses, p. 161–184. *In* S. Morse (ed.), The evolutionary biology of viruses. Raven Press, New York.
6. **Dumitrescu, O., M. Kalish, S. C. Kliks, C. I. Bandea, and J. A. Levy.** 1994. Characterization of HIV-1 strains isolated from children in Romania: identification of a new envelope subtype. J. Infect. Dis. **169:**281–288.
7. **Faulkner, D. V., and A. Jurka.** 1988. Multiple aligned sequence editor (MASE). Trends Biochem. Sci. **13:**321–322.
8. **Felsenstein, J.** 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39:**783–791.
9. **Felsenstein, J.** 1989. PHYLIP—phylogeny inference package. Cladistics **5:**164–166.
10. **Gao, F., L. Yue, D. L. Robertson, S. C. Hill, H. Hui, R. J. Biggar, A. E. Neequaye, T. M. Whelan, D. D. Ho, G. M. Shaw, P. M. Sharp, and B. H. Hahn.** Genetic diversity of human immunodeficiency virus type 2: evidence for distinct sequence subtypes with differences in virus biology. J. Virol., in press.
11. **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. Bosch, 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.
12. **Hillis, D. M., M. W. Allard, and M. M. Miyamoto.** 1993. Analysis of DNA sequences data: phylogenetic inference. Methods Enzymol. **224:**456–487.
13. **Hillis, D. M., and J. J. Bull.** 1993. An empirical test of bootstrapping as a method of assessing confidence in phylogenetic analysis. Syst. Biol. **42:**182–189.
14. **Hillis, D. M., and J. P. Huelsenbeck.** 1992. Signal, noise and reliability in molecular phylogenetic analysis. J. Hered. **83:**189–195.
15. **Hillis, D. M., J. P. Huelsenbeck, and C. W. Cunningham.** Application and accuracy of molecular phylogenies. Submitted for publication.
16. **Hu, W.-S., and H. M. Temin.** 1990. Retroviral recombination and reverse transcription. Science **250:**1227–1233.
17. **Korber, B., and G. Myers.** 1992. Signature pattern analysis: a method for assessing viral sequence relatedness. AIDS Res. Hum. Retroviruses **8:**1549–1560.
18. **Kusumi, K., B. Conway, S. Cunningham, A. Berson, C. Evans, A. K. N. Iversen, D. Colvin, M. V. Gallo, S. Coutre, E. G. Shpaer, D. V. Faulkner, A. DeRonde, S. Volkman, C. Williams, M. S. Hirsch, and J. I. Mullins.** 1992. Human immunodeficiency virus type 1 envelope gene structure and diversity in vivo and following cocultivation in vitro. J. Virol. **66:**875–885.
19. **Li, W.-H., M. Tanimura, and P. M. Sharp.** 1988. Rates and dates of divergence between AIDS virus nucleotide sequences. Mol. Biol. Evol. **5:**313–330.
20. **Louwagie, J., E. L. Delwart, J. I. Mullins, F. E. McCutchan, G. Eddy, and D. S. Burke.** 1994. Genetic analysis of HIV-1 isolates from Brazil reveals the presence of two distinct genotypes. AIDS Res. Hum. Retroviruses **10:**561–567.
21. **Maddison, W. P., and D. R. Maddison.** 1992. MacClade: analysis of phylogenetic and character evolution. Sinauer Associates, Inc., Sunderland, Mass.
22. **Meyerhans, A., J.-P. Vartanian, and S. Wain-Hobson.** 1990. DNA recombination during PCR. Nucleic Acids Res. **18:**1687–1691.
23. **Morgado, M. G., E. C. Sabino, E. G. Shpaer, V. Bongertz, L. Brigido, M. D. C. Guimaraes, E. A. Castilho, B. Galvao-Castro, J. I. Mullins, R. M. Hendry, and A. Mayer.** 1994. V3 region polymorphisms in HIV-1 from Brazil: prevalence of subtype B strains divergent from the North American/European prototype and detection of subtype F. AIDS Res. Hum. Retroviruses **10:**569–576.
24. **Murphy, E., B. Korber, M.-C. Georges-Courbot, B. You, A. Pinter, D. Cook, M.-P. Kieny, A. Georges, C. Mathiot, F. Barre-Sinoussi, and M. Girard.** 1993. Diversity of V3 region sequences of human immunodeficiency viruses type 1 from the Central African Republic. AIDS Res. Hum. Retroviruses **9:**997–1006.
25. **Myers, G., B. Korber, S. Wain-Hobson, R. F. Smith, and G. N. Pavlakis.** 1993. Human retroviruses and AIDS. Los Alamos National Laboratory, Los Alamos, N.Mex.
26. **Nei, M.** 1987. Molecular evolutionary genetics, p. 19–38. Columbia University Press, New York.
27. **Nei, M., and T. Gojorbori.** 1986. Simple method for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. **3:**418–426.
28. **Potts, K. E., M. L. Kalish, T. Lott, G. Orloff, C.-C. Luo, M.-A. Bernard, C. B. Alves, R. Badaro, J. Suleiman, O. Ferreira, G. Schochetman, W. D. Johnson, Jr., C.-Y. Ou, J. L. Ho, and the Brazilian Collaborative AIDS Research Group.** 1993. Genetic heterogeneity of the V3 region of the HIV-1 envelope glycoprotein in Brazil. AIDS **7:**1191–1197.
29. **Simmonds, P., P. Balfe, J. F. Peutherer, C. A. Ludlam, J. O. Bishop, and A. J. L. Brown.** 1990. Human immunodeficiency virus-infected individuals contain provirus in small numbers of peripheral mononuclear cells and at low copy numbers. J. Virol. **64:**864–872.
30. **Simmonds, P., L. Q. Zhang, F. McOmish, P. Balfe, C. A. Ludlam, and A. J. L. Brown.** 1991. Discontinuous sequence change of human immunodeficiency virus (HIV) type 1 *env* sequences in plasma viral and lymphocyte-associated proviral populations in vivo: implications for models of HIV pathogenesis. J. Virol. **65:**6266–6276.
31. **Swofford, D.** 1993. PAUP: phylogenetic analysis using parsimony, version 3.1. Center for Biodiversity, Illinois Natural History Survey, Champaign, Ill.
32. **Wolfs, T. F. W., G. Zwart, M. Bakker, and J. Goudsmit.** 1992. HIV-1 genomic RNA diversification following sexual and parenteral virus transmission. Virology **189:**103–110.
33. **Zhang, L. Q., P. MacKenzie, A. Cleland, E. C. Holmes, A. J. L. Brown, and P. Simmonds.** 1993. Selection for specific sequences in the external envelope protein of human immunodeficiency virus type 1 upon primary infection. J. Virol. **67:**3345–3356.
34. **Zhu, T., H. Mo, N. Wang, D. S. Nam, Y. Cao, R. A. Koup, and D. D. Ho.** 1993. Genotypic and phenotypic characterization of HIV-1 in patients with primary infection. Science **261:**1179–1181.