Sequence Note

Rapid Spread and Genetic Diversification of HIV Type 1 Subtype C in a Rural Area of Southern Mozambique

MARIA LAHUERTA,[1,*] ESTER APARICIO,[2,*] AZUCENA BARDAJI,[1,3] SANDRA MARCO,[2]
JAHIT SACARLAL,[3,4] INACIO MANDOMANDO,[3] PEDRO ALONSO,[1,3] MIGUEL ANGEL MARTINEZ,[2]
CLARA MENENDEZ,[1,3] and DENISE NANICHE[1]

ABSTRACT

In this study, we analyzed the human immunodeficiency type 1 (HIV-1) viruses circulating between 1999 and 2004 in antiretroviral-naive women from a rural area of southern Mozambique. Nucleotide sequencing of the HIV-1 long terminal repeat (LTR) U3, envelope (*env*) C2V3C3, and protease (*pr*) genomic regions was performed from women sera samples collected in 1999 and 2004. Phylogenetic analysis revealed that all amplified sequences belonged to subtype C. Although *env* sequences were predominantly CCR5-tropic (R5), CXCR4-tropic (X4) variants were also identified (13%). Both 1999 and 2004 sequences were widely dispersed across multiple clusters and were related to different reference sequences from neighboring countries. Sequences from 2004 showed significantly more nucleotide genetic diversity than sequences from 1999. Importantly, genetic diversification was also observed at the *pr* and *env* amino acid level, suggesting that positive selection forces were implicated in the viral diversification. These results indicate the rapid spread and diversification of subtype C virus in Mozambique where HIV-1 prevalence in the Manhiça antenatal clinic reached 23% in 2004.

INTRODUCTION

SUB-SAHARAN AFRICA is the region most severely affected by HIV-1, with 24.7 million (60%) of the 39.5 million people living with HIV/AIDS.[1] This region is characterized by the circulation of several genotypes and circulating recombinant forms (CRFs).[2] However, subtype C viruses are rapidly spreading in southern Africa and becoming the predominant variant of the epidemic, accounting for more than 56% of all global infections.[3,4] There is evidence that suggests that different subtypes may have distinct biological features that may have an impact on transmissibility and on disease progression.[5,6] The high genetic variability of HIV-1 and its potential for recombination are clearly important to escape from the immune system and also play a major role in the development of resistance to highly active antiretroviral therapy (HAART).[7,8]

There has been tremendous development of new antiretroviral drugs for the treatment of HIV-1 subtype B viruses, but information on their efficacy in nonsubtype B viruses is scarce. Recently, the genetic characterization of the protease and reverse transcriptase genes from non-B subtypes has revealed that minor mutations are often present in natural variants.[9–11] However, it is unknown whether the presence of these polymorphisms facilitates the generation of resistance during treatment with protease inhibitors (PI).[12]

Although Mozambique is ranked as having the tenth highest HIV-1 prevalence in the world, relatively little information on the molecular epidemiology of HIV-1 in Mozambique is avail-

[1]Barcelona Center for International Health Research (CRESIB), Hospital Clinic, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Universtitat de Barcelona, Barcelona, Spain.
[2]Fundació irsiCaixa, Hospital Universitari Germans Trias i Pujol, Universitat Autonoma de Barcelona, 08916 Badalona, Spain.
[3]The Manhiça Health Research Center (CISM), Manhiça, Mozambique.
[4]Faculdade de Medicina da Universidade Eduardo Mondlane, Maputo, Mozambique.
*Both authors contributed equally to this study.

able. Mozambique suffered a long civil war until 1992, and since then, the spread of HIV-1 has been increasing rapidly, with a shift in overall reported prevalence from 8.2% in 1998 to 16.2% in 2004.[13] The number of people living with HIV-1 is estimated to be approximately 1.8 (1.4–2.2) million, with women representing up to 58% of those infected.[1] In Mozambique, it is estimated that only 3.4% of pregnant women who need antiretroviral therapy (ART) for prevention of mother-to-child transmission actually receive it. Furthermore, of those HIV-1-infected adults in Mozambique who require HAART, only approximately 9% receive treatment.[1]

To date, there are only two published studies assessing both HIV-1 subtypes and genetic diversity in Mozambique. One study evaluated drug resistance in drug-naive HIV-1 patients in two hospitals near Maputo.[14] The second study performed genetic characterization of HIV-1 circulating in Beira, a central region of the country.[15] Both studies determined that subtype C was the dominant HIV-1 subtype. However, in Mozambique longitudinal studies analyzing HIV-1 genetic diversification and evolution are lacking. Large-scale implementation of HAART has begun in many southern and eastern African countries, including Mozambique. Therefore, studies including phylogenetic analysis, drug resistance polymorphisms, genetic diversity, and molecular evolution of HIV-1 may improve the treatment success in this area.

To gain a better understanding of genetic diversity, molecular evolution and epidemiological patterns of HIV-1 spread, and drug resistance polymorphisms, we studied regions in the HIV U3 long terminal repeat (LTR), the envelope *(env)* C2V3C3 and protease *(pr)* gene. This study focused on HIV-1 sequences obtained in 1999 and in 2004, from women living in a semi-rural area of southern Mozambique.

MATERIALS AND METHODS

Study area and population

The study was undertaken at the Centro de Investigação em Saúde da Manhiça (CISM) in Manhiça District, southern Mozambique. Blood samples were obtained from HIV-positive women at the time of recruitment from two different studies performed in 1999 and 2004.[16,17] After obtaining written informed consent, blood samples were collected in EDTA anticoagulant tubes and plasma was stored at −80°C until analyzed. HIV-1 disease status was not available for these women. Since HAART was not available in Manhiça during these studies, all women were considered to be drug naive at recruitment. The 1999 study included 300 women, of whom 30 were HIV$^+$, aged 14–61 years and attending antenatal and family planning clinics. The 2004 study included 1030 pregnant women attending the antenatal clinic of whom 207 were HIV$^+$. Both studies were approved by the Ethics Review Board of Mozambique and Hospital Clinic of Barcelona.

PCR amplification and sequencing

Extraction of viral RNA from the 1999 plasma samples was performed using the Qiamp viral RNA mini kit (Qiagen) and viral load was determined using the Cobas Amplicor HIV-1 Monitor Test v1.5 (Roche Diagnostics). For the 2004 samples, viral load was determined using the Amplicor HIV-1 Monitor Test v1.5 (Roche Diagnostics) and due to the small volume of samples, RNA extracted during this process was reverse transcribed to cDNA with Superscript II and random hexamer primers (Invitrogen). The RNA template and random primers (300 ng) were heated to 65°C for 10 min, chilled on ice, and reverse transcribed in a 20 μl reaction volume containing 1× reaction buffer, 109 mM dithiothreitol, 0.5 mM each deoxynucleoside triphosphate, and 200 U of Superscript reverse transcriptase (Invitrogen) at 37°C for 1 h, followed by 15 min at 95°C. The cDNA was shipped to Fundació Irsi-Caixa (Barcelona, Spain) where amplification and sequencing were performed.

Amplification of the U3 LTR, *env* C2V3C3 region, and *pr* was performed for the 1999 and 2004 samples. The LTR and *env* C2V3C3 regions were amplified following the reaction conditions and primers described in Ibanez *et al.* and Tapia *et al.*, respectively.[18,19] LTR primers used for first-round amplification were NI25 (HXB2 positions 57–77) and NI23 (HXB2 positions 389–408) and for second-round amplification were NI33 (HXB2 positions 350–372) and NI35 (HXB2 positions 81–100). The outer pair of primers for *env* amplification were ENV1 (HXB2 positions 6858–6878) and ENV4 (HXB2 positions 7520–7539) and the inner primers were ENV2 (HXB2 positions 6885–6904) and ENV3 (HXB2 positions 7365–7385). The *pr* region was amplified from the cDNA with 5PROT1 (5′-ATTTTTTA GGGAARATYTGGCCT-3′; positions 2084–2106) and 3PROT1 (5′-YTGGAGTR TTRTATGGATTTT-CAGG-3′; positions 2703–2727) as outer primers and 5PROT2 (5′-ATTTTTTAGGGAARATYTGGCCT-3′; positions 2118–2140) and 3PROT2 (5′-CTTTTATTTTYTCTTCTGTYAA-TGG-3′; positions 2622–2646) as inner primers. The following amplification conditions were used for the first-round and nested polymerase chain reaction (PCR): 2 min at 95°C, followed by 35 cycles of 30 s at 95°C, 30 s at 55°C, 40 s at 72°C, and a final extension step at 72°C for 7 min.

Both strands of the PCR fragments were purified with Exo-SAP-IT (Applied Biosystems) and sequenced directly using internal (nested) PCR primers and the ABI Prism Dye Terminator Cycle Sequencing Reaction kit (Perkin Elmer). The products of the reactions were then analyzed on an ABIprism 3100 sequencer (Perkin Elmer). Sequence editing was performed using the program SEQUENCER, version 4.6 (GeneCodes).

Subtyping and phylogenetic analysis

The new sequences were manually edited with BioEdit version 7 and compared to subtype reference strains downloaded from the Los Alamos subtype database (http://hiv-web.lanl.gov/content/hiv-db/SUBTYPE_REF/align.html). To examine intrasubtype variations, a subset of subtype C from neighboring countries was downloaded from the Los Alamos BLAST search database (http://hiv-web.lanl.gov). Multiple alignments were performed with Clustal W 1.7.[20] Sequence and phylogenetic analyses were performed as previously described.[21,22] Briefly, phylogenetic trees were created with PAUP* version 4.0b 10 with likelihood settings by selecting the best-fitting evolutionary models with the Akaike identification system, implemented in Modeltest 3.6.[23,24] The TVM+I+G model was used

for LTR, the GTR+I+G model for the protease region, and the K81uf+I+G for the envelope region. The robustness of the trees was evaluated by bootstrap analysis with 1000 rounds of replication. Trees were viewed with Treeview.[25] The sequence ID used to name the study sequences is the ID number used in the CISM to track samples, while the sequence ID for the reference strains has been simplified with the initials of the country.

Mean genetic distances were measured with the corresponding model for each region implemented in PAUP* version 4.0b.[23] SNAP (http://hiv-web.lanl.gov) was used to calculate the synonymous-to-nonsynonymous substitution ratio.[26,27] The V3 loop amino acid sequences were introduced into the WebPSSM program (http://ubik.microbiol. washington.edu/ computing/pssm) to determine the coreceptor usage and the net charge of the V3 loop.[28–30] The V3 loop of the virion's envelope glycoprotein gp120 has been identified as a major determinant for viral tropism and coreceptor usage, depending on its net charge and the polarity of the amino acids located at positions 11 and 25.[31] A lower V3 net charge facilitates a tighter interaction of gp120 with CCR5 while a higher V3 net charge facilitates the interaction with CXCR4.[32]

Drug resistance genotyping

The Stanford University HIV Drug Resistance database (aIg Program) (http://hivdb.stanford.edu) was used to determine resistance-associated mutations and polymorphisms present in the protease (1999 and 2004 sequences). Known mutations and/or polymorphisms were assessed in the protease sequences from 1999 ($n = 20$) and from 2004 ($n = 31$).

Statistical analyses

The nonparametric Mann–Whitney test was used to test for differences between intrasubtype genetic distances, synonymous and nonsynonymous mutations, and distance to ancestor between 1999 and 2004 sequences. A p value lower than 0.05 was considered statistically significant. Analyses were performed using the Graph Pad software package, version 4 (GraphPad Software, Inc.).

RESULTS

Characteristics of the study population

Manhiça district, a semirural area of southern Mozambique, has an estimated population of 36,000 inhabitants under demographic surveillance prior to 2004. The prevalence of HIV-1 infection in pregnant women attending the antenatal clinic was 12% in 1999 and 23% in 2004.[16,17] Both in 1999 and 2004, the two populations were women, with a median age of 33 (IQR 25–48) and 23 (IQR 20–30) years, respectively. All women from the 2004 study and 60% of women from the 1999 study were pregnant at the time of sample collection. The two groups had similar viral load levels, with median values of 12,200 copies/ml (IQR 3650–36,100) for the 1999 study population and 18,620 (IQR 10,388–49,091) copies/ml for the 2004 study population ($p = 0.11$).

Subtyping and phylogenetic analysis

To subtype and determine the phylogenetic relationships between sequences, the LTR, *env,* and *pr* regions were amplified from available sera samples collected in 1999 and 2004. Of the 30 HIV$^+$ blood samples available from 1999, amplification was successfully performed for LTR ($n = 21$), *env* ($n = 18$), and *pr* ($n = 20$). Of the 207 HIV$^+$ women in the 2004 study, 51 were randomly chosen for this study. Due to small quantities of cDNA, out of the 51 samples available from 2004, amplification was performed for LTR ($n = 37$), *env* ($n = 32$), and *pr* ($n = 31$).

Phylogenetic analysis showed that all sequences from LTR, *env,* and *pr* clustered into the subtype C clade (data not shown). To investigate within-subtype C clustering, phylogenetic trees were constructed with published LTR, envelope, and protease C sequences from central Mozambique and several sequences from nine different African neighboring countries (South Africa, Zimbabwe, Tanzania, Malawi, Ethiopia, Kenya, Botswana, Somalia, and Senegal) from the sequences available at the Los Alamos HIV sequence database (http://www.hiv.lanl. gov/components/hiv-db/combined_ search_s_tree/search.html).

Both 1999 and 2004 envelope sequences from Manhiça were widely dispersed across multiple clusters (Fig. 1A). Similar tree topologies were obtained for LTR and protease (Fig. 1B and C), although the envelope tree tended to have higher bootstrap values reflecting a higher robustness of the clusters.

Different phylogenetic associations were observed with other subtype C sequences from neighboring countries. However, distance matrix-based phylogenetic analysis did not reveal potential clusters and couples of genetically related sequences (bootstrap values higher than 70%) between our sequences and any of the reference sequences from neighboring countries. Distinct sublineages consisting only of closely related Mozambican sequences from 1999 and 2004 with bootstrap values higher than 70% were identified in LTR, envelope, and protease, suggesting an epidemiological link between these samples (marked with brackets in Fig 1).

Genetic divergence and selection pressure

To assess the intrasubtype genetic diversity, the genetic distances between each set of sequences were calculated. Significantly higher median genetic distances were observed for sequences from 2004 as compared to sequences from 1999 in all regions assessed: LTR, *env*, and *pr* (Table 1). Larger median genetic distances reflecting increased genetic diversity were most pronounced in the envelope region.

Sequences were compared to assess the mutational pattern of protease and envelope in the absence of drug therapy (Table 1). Both genes were found to be under purifying (negative) selection pressure with a ratio of *ds/dn* greater than 1. The average *ds/dn* for both 1999 and 2004 sequences was 1.979 and 2.047 for the envelope region and 6.123 and 7.196 for the protease region, respectively. Protease showed very low rates of nonsynonymous mutations reflecting different pressure constraints between *pr* and *env* genes.

Coreceptor usage determined by V3 loop sequences

The average net charge in this group of sequences was +4.3. The characteristic GPGQ crown found in most non-B subtypes
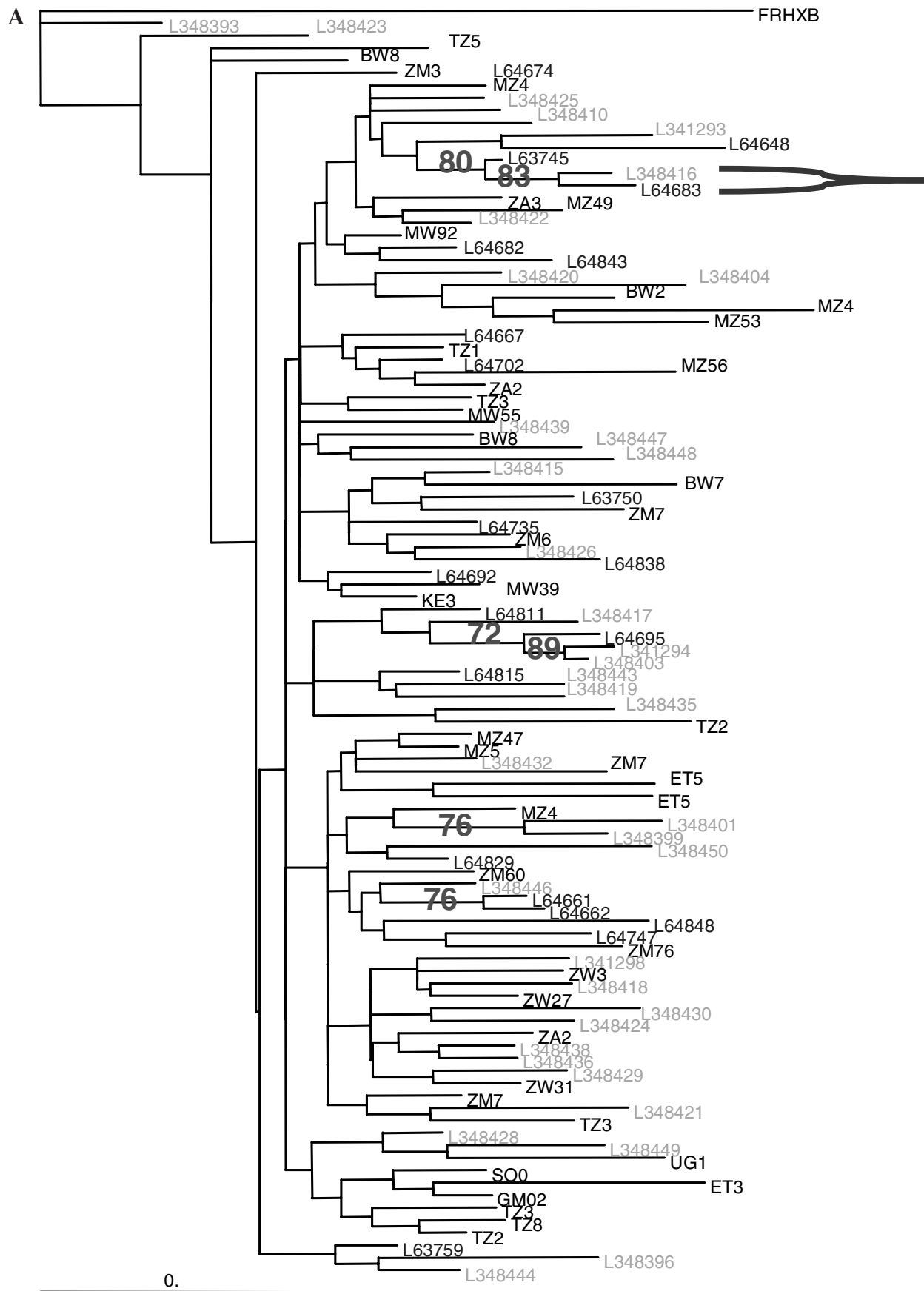
A

FRHXB
L348393   L348423
TZ5
BW8
ZM3   L64674
MZ4
L348425
L348410   L341293   L64648
80   L63745
83   L348416
L64683
ZA3   MZ49
L348422
MW92
L64682
L64843
L348420   L348404
BW2
MZ4
MZ53
L64667
TZ1
L64702   MZ56
ZA2
TZ3
MW55
L348439   L348447
BW8   L348448
L348415
BW7
L63750
ZM7
L64735
ZM6
L348426
L64838
L64692
MW39
KE3   L64811   L348417
72   L64695
89   L341294
L348403
L64815   L348443
L348419
L348435
TZ2
MZ47
MZ5
L348432   ZM7
ET5
ET5
76   MZ4   L348401
L348399
L348450
L64829
ZM60
L348446
76   L64661
L64662   L64848
L64747
ZM76
L341298
ZW3
L348418
ZW27
L348430
ZA2   L348424
L348438
L348436   L348429
ZW31
ZM7
TZ3   L348421
L348428   L348449
UG1
SO0
ET3
GM02
TZ3
TZ8
TZ2
L63759   L348396
L348444

0.

FIG. 1.   (A) Envelope sequences from Manhiça (both 1999 and 2004) were widely dispersed across multiple clusters. Similar tree topologies were obtained for LTR (B) and protease (C).
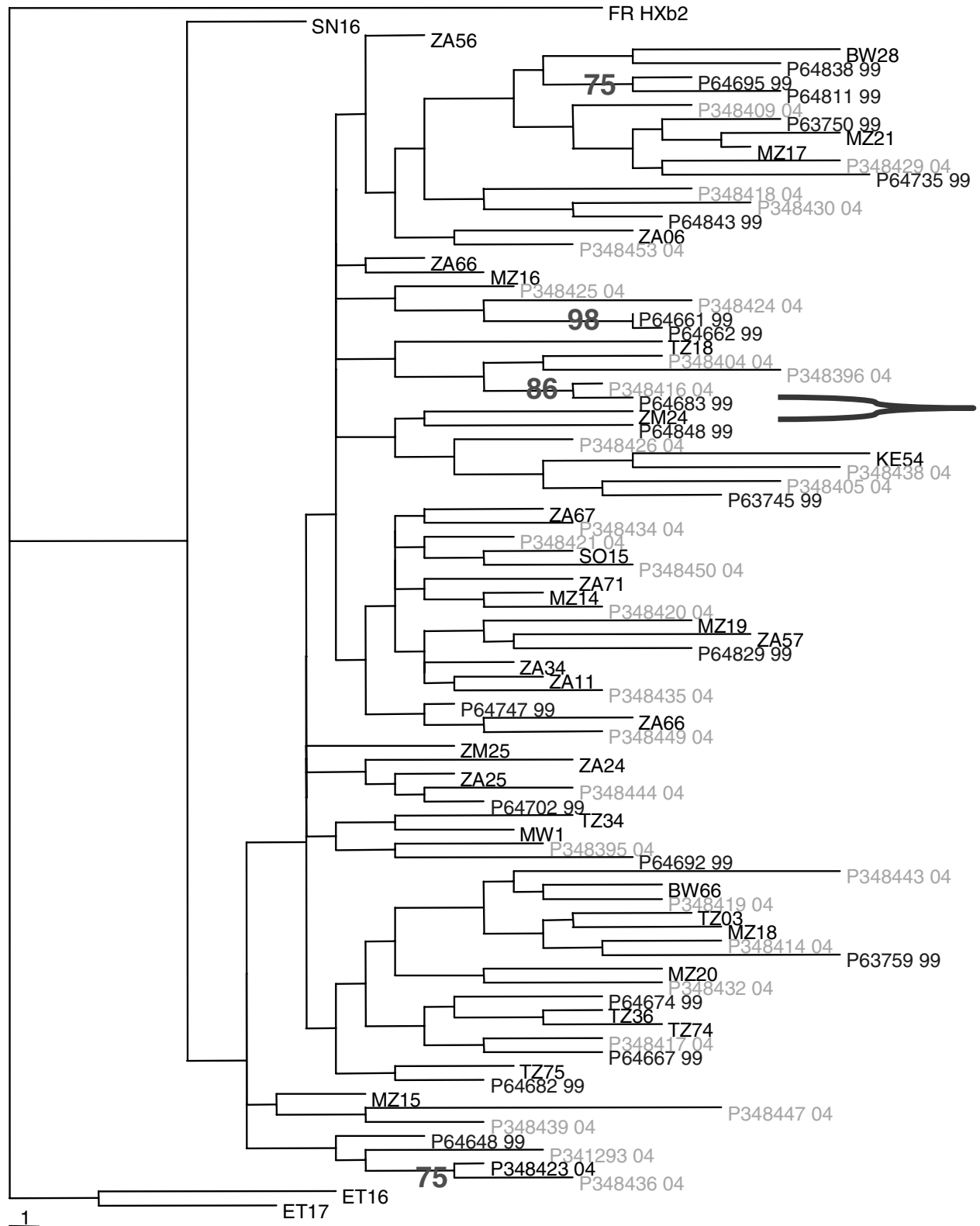
FIG. 1. (*Continued*).

C

FIG. 1. (*Continued*).

TABLE 1. GENETIC DISTANCES, RATIO OF NONSYNONYMOUS TO SYNONYMOUS AMINO ACID SUBSTITUTIONS

| | 1999 sequences[a] | | 2004 sequences[b] | | |
	Median	SD	Median	SD	p-value
Genetic distances					
LTR	0.1345	0.035	0.157	0.036	0.0085[c]
Protease	0.0664	0.0189	0.071	0.0179	0.0085[c]
Envelope	0.1805	0.046	0.2	0.037	0.0001[c]
Envelope					
Synonymous substitutions	0.2204	0.0583	0.2431	0.07264	0.0022[c]
Nonsynonymous substitutions	0.115	0.02651	0.1236	0.02295	0.0001[c]
Ratio *ds/dn*	1.979	0.589	2.047	0.7758	0.6855
Protease					
Synonymous substitutions	0.1551	0.0431	0.1703	0.04397	0.0001[c]
Nonsynonymous substitutions	0.02826	0.01171	0.0297	0.0116	0.2202
Ratio *ds/dn*	6.123	7.521	7.196	5.265	0.3129

[a]1999 samples: LTR (*n* = 21), *env* (*n* = 18), and *pr* (*n* = 20).
[b]2004 samples: LTR (*n* = 37), *env* (*n* = 32), and *pr* (*n* = 31).
[c]*p*-values <0.05 were considered statistically significant.

was conserved in 98% (49/50) of the samples. According to the sequence analysis with WebPSSM, 88% (44/50) of the viruses potentially used CCR5, with six sequences reflecting potential R5X4 or X4 receptor usage (three from 1999 and three from 2004). Although these results indicated a predominance of R5 viruses, they also showed that potential X4 receptor usage is not prohibited in subtype C envelope regions.

Mutations associated with drug resistance to protease inhibitors

No primary resistance mutations to protease inhibitors were detected in any of the 51 samples. However, a high prevalence of minor mutations was observed at the following positions: K20R (21.6%), M36I (74.5%), L63P (27.5%), and I93L (94.1%). Other polymorphisms present at lower frequency are reported in Table 2.

DISCUSSION

Despite the high prevalence of HIV-1 in Mozambique, so far only two studies have examined the genetic diversity of Mozambican HIV-1 viruses.[14,15] Thus, the aim of this study was to characterize the HIV-1 epidemic at a molecular level in women from a rural area of southern Mozambique and to assess viral genetic diversity in 1999 and in 2004.

Subtype analysis showed that all the LTR, envelope, and protease regions studied were subtype C, in accordance with other studies that found all HIV-1 isolates in Maputo, Mozambique to be subtype C and all but one to be subtype C in Beira.[14,15] Overall, these results indicate that the HIV-1 epidemic in Mozambique is quite homogeneous with regard to subtype since it is composed almost exclusively of clade C viruses. This is similar to the situation found in neighboring countries such as South Africa, but different from countries such as Tanzania that also have A and D viruses circulating.[19,33] Phylogenetic analyses revealed that Mozambican sequences did not form a coun-

try-specific cluster. Instead, sequences were randomly spread throughout the trees without showing different homology from other subtype C sequences from neighboring countries. The wide genetic variation of these sequences suggests multiple introductions of subtype C viruses in Mozambique and thus confirms the existence of multiple circulating clade C sublineages in southern Africa as described.[34] Studies of population migrations suggest that HIV-1 introduction into Mozambique may have been associated with the frequent temporary labor migrations of adult males from southern Mozambique to South Africa.[35]

Nucleotide substitutions in genes coding for proteins can be either synonymous (no change of amino acid) or nonsynonymous (change of amino acid). Investigating the number of syn-

TABLE 2. FREQUENCY OF KNOWN MUTATIONS/POLYMORPHISMS DETECTED IN HIV PROTEASE FROM ARV DRUG-NAIVE WOMEN

Protease		
Mutation	Number of strains	Frequency of mutation (%) (N = 51)
L10V	1	2
L10I	1	2
K20R	11	21.6
M36I	38	74.5
M36V	1	2
M36L	2	3.9
L63P	14	27.5
L63A	2	3.9
L63S	1	2
L63V	1	2
A71T	2	3.9
V77I	8	15.7
V77L	1	2
I93L	48	94.1

onymous and nonsynonymous substitutions may therefore provide information on the degree of selection operating on the virus. Both in *env* and *pr*, the number of synonymous substitutions increased between 1999 and 2004. Nonsynonymous substitutions also increased in *env*, indicating that higher variability in this region of *env* may be caused by the evolutionary force of genetic drift in addition to external selective pressure. An interesting finding of the present study is the higher synonymous and nonsynonymous diversity found within the 2004 sequences as compared to the 1999 sequences. Moreover, median genetic distances were also higher for the 2004 sequences. Although we cannot exclude that differences observed are due to small sample size, it can be hypothesized that the higher heterogeneity of the 2004 viruses may be due to both (1) the introduction of new viruses in Mozambique between 1999 and 2004 as well as (2) the diversification of the viruses introduced prior to 1999. Whether this diversification affects the transmissibility or virulence of subtype C viruses remains to be elucidated. It has been suggested that the genetic bottlenecks created by the transmission of few clones of HIV-1 and pressure from the immune system could reduce the overall fitness of the virus. Furthermore, with the dramatic global spread of subtype C, the question has been raised as to whether this dominance is related to an attenuation of virulence, but this has yet to be investigated.[36]

Primary resistance mutations to protease inhibitors were not found in either the 1999 or 2004 sequences. Nevertheless, a high prevalence of secondary resistance mutations was observed in our sequences, predominantly the M36I and I93L mutations. Although the prevalence of polymorphisms is higher than that described for subtype B viruses, these results are consistent with the high prevalence described in two studies conducted in Mozambique and with other studies performed on non-B clades.[10,14,15,37,38] The implications of a high prevalence of secondary resistance mutations in subtype C are not known, although the mutation at M36 in the B clade has been associated with greater risk of virological failure and with the appearance of the major resistance mutation L90M.[12,39]

The prediction of CCR5 or CXCR4 receptor usage based on the net charge of the V3 loop and the presence of the GPGQ crown indicated that most viruses in this study potentially used CCR5, with 13% potentially displaying X4 or mixed R5X4 usage. Little is known in Africa regarding CCR5 and CXCR4 usage as a reflection of disease progression. It is generally accepted that viruses present in an individual early in HIV-1 infection are predominantly R5, but the timing of the R5-X4 switch for subtype C viruses is unclear.[40–43] Nevertheless, as described by Coetzer *et al.*, our results show that CXCR4 usage is not a rare event in subtype C viruses as previously suggested.[44–46] Unfortunately, CD4[+] T cell counts were not available for the samples analyzed and consequently we were unable to assess correlations between CXCR4 usage and disease stage.

In conclusion, this study describes the genetic diversity, molecular evolution, and epidemiological patterns of HIV-1 spread in women from a rural area of southern Mozambique. As implementation of HAART has begun in this area, the description of drug resistance polymorphisms in antiretroviral drug-naive women will be relevant to future evaluations of antiretroviral therapy programs.

SEQUENCE DATA

The GenBank accession numbers of the sequences reported in this paper are EF407648–EF407665 and EF407667–EF407809. The GenBank accession numbers of the reference sequences from neighboring countries used in the phylogenetic trees are AF023427, AF023431, AF127568, AF023430, AF096778, AF096780, AF096782, U63536, U63539, AY452654, AF196734, AF196733, AF196732, AY452658, AY162225, AF290029, AY162224, AF239628, AF239638, AF239633, AF239654, AF254734, AF254686, AF254723, AF254711, AB192571, AF102218, AF102202, AF102207, AB191674, AB191676, AB191660, AB191679, U08455, U51292, AM076841, AM076843, AM076847, and AM076846.

ACKNOWLEDGMENTS

REFERENCES

1. UNAIDS/WHO: AIDS Epidemic Update: December 2006. UNAIDS/WHO, Geneva, 2006.
2. Leitner T, Korber B, Daniels M, Calef C, *et al.*: HIV-1 subtype and circulating recombinant form (CRF) reference sequences, 2005. Sequence Compendium. http://www.hiv.lanl.gov/content/hiv-db/REVIEWS/reviews.html. Los Alamos National Laboratories, Los Alamos, NM, 2005.
3. Julg B and Goebel FD: HIV genetic diversity: Any implications for drug resistance? Infection 2005;33:299–301.
4. Butler IF, Pandrea I, Marx PA, and Apetrei C: HIV genetic diversity: Biological and public health consequences. Curr HIV Res 2007;5:23–45.
5. Thomson MM and Najera R: Molecular epidemiology of HIV-1 variants in the global AIDS pandemic: An update. AIDS Rev 2005;7:210–224.
6. Geretti AM: HIV-1 subtypes: Epidemiology and significance for HIV management. Curr Opin Infect Dis 2006;19:1–7.
7. Geretti AM: Clinical implications of HIV drug resistance to nucleoside and nucleotide reverse transcriptase inhibitors. AIDS Rev 2006;8:210–220.
8. Kantor R and Katzenstein D: Polymorphism in HIV-1 non-subtype B protease and reverse transcriptase and its potential impact on drug susceptibility and drug resistance evolution. AIDS Rev 2003;5:25–35.
9. Montes B, *et al.*: Comparison of drug resistance mutations and their interpretation in patients infected with non-B HIV-1 variants and matched patients infected with HIV-1 subtype B. J Acquir Immune Defic Syndr 2004;35:329–336.

10. Vergne L, *et al.*: Genetic diversity of protease and reverse transcriptase sequences in non-subtype-B human immunodeficiency virus type 1 strains: Evidence of many minor drug resistance mutations in treatment-naive patients. J Clin Microbiol 2000;38:3919–3925.

11. Vergne L, *et al.*: Natural polymorphism in protease and reverse transcriptase genes and in vitro antiretroviral drug susceptibilities of non-B HIV-1 strains from treatment-naive patients. J Clin Virol 2006;36:43–49.

12. Perno CF, *et al.*: Minor mutations in HIV protease at baseline and appearance of primary mutation 90M in patients for whom their first protease-inhibitor antiretroviral regimens failed. J Infect Dis 2004;189:1983–1987.

13. UNAIDS/WHO. Uniting the world against AIDS. http://www.unaids.org/en/Regions_Countries/Countries/mozambique.asp. UN Technical Working Group, Mozambique, 2006.

14. Bellocchi MC, *et al.*: Subtype analysis and mutations to antiviral drugs in HIV-1-infected patients from Mozambique before initiation of antiretroviral therapy: Results from the DREAM programme. J Med Virol 2005;76:452–458.

15. Parreira R, *et al.*: Genetic characterization of human immunodeficiency virus type 1 from Beira, Mozambique. Microbes Infect 2006;8:2442–2451.

16. Menendez C, *et al.*: Prevalence and risk factors of sexually transmitted infections and cervical neoplasia in women from a rural area of Southern Mozambique (submitted for publication).

17. Menendez C, *et al.*: Combination of interventions for malaria prevention in pregnancy. P105 One (in press).

18. Ibanez A, Clotet B, and Martinez MA: Absence of genetic diversity reduction in the HIV-1 integrated proviral LTR sequence population during successful combination therapy. Virology 2001;282:1–5.

19. Tapia N, *et al.*: Influence of human immunodeficiency virus type 1 subtype on mother-to-child transmission. J Gen Virol 2003;84:607–613.

20. Thompson JD, Higgins DG, and Gibson TJ: CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–4680.

21. Fernandez G, *et al.*: Purifying selection of CCR5-tropic HIV-1 variants in AIDS subjects that have developed syncytium-inducing CXCR4-tropic viruses. J Gen Virol 2006;87:1285–1294.

22. Fernandez G, *et al.*: Fitness landscape of HIV-1 protease quasispecies. J Virol 2007;81:2485–2496.

23. Swofford DL: *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Version 4. Sinauer Associates, Sunderland, MA, 2003.

24. Akaike H: A new look at statistical model identification. IEEE Trans Automatic Control 1997;19:716–723.

25. Page RDM: TREEVIEW: An application to display phylogenetic trees on personal computers. Comput Appl Biosci 1996;12:357–358.

26. Aylor DL, Price EW, and Carbone I: SNAP: Combine and map modules for multilocus population genetic analysis. Bioinformatics 2006;22:1399–1401.

27. Price EW and Carbone I: SNAP: Workbench management tool for evolutionary population genetic analysis. Bioinformatics 2005;21:402–404.

28. Brumme ZL, *et al.*: Clinical and immunological impact of HIV envelope V3 sequence variation after starting initial triple antiretroviral therapy. AIDS 2004;18:F1–F9.

29. Jensen MA, Coetzer M, van't Wout AB, Morris L, and Mullins JI: A reliable phenotype predictor for human immunodeficiency virus type 1 subtype C based on envelope V3 sequences. J Virol 2006;80:4698–4704.

30. Jensen MA, *et al.*: Improved coreceptor usage prediction and genotypic monitoring of R5-to-X4 transition by motif analysis of HIV-1 env V3 loop sequences. J Virol 2003;77:13376–13388.

31. Xiao L, *et al.*: CCR5 coreceptor usage of non-syncytium-inducing primary HIV-1 is independent of phylogenetically distinct global HIV-1 isolates: Delineation of consensus motif in the V3 domain that predicts CCR-5 usage. Virology 1998;240:83–92.

32. Pollakis G, *et al.*: N-linked glycosylation of the HIV type-1 gp120 envelope glycoprotein as a major determinant of CCR5 and CXCR4 coreceptor utilization. J Biol Chem 2001;276:13433–13441.

33. Arroyo MA, *et al.*: HIV-1 diversity and prevalence differ between urban and rural areas in the Mbeya region of Tanzania. AIDS 2005;19:1517–1524.

34. Gordon M, *et al.*: Molecular characteristics of human immunodeficiency virus type 1 subtype C viruses from KwaZulu-Natal, South Africa: Implications for vaccine and antiretroviral control strategies. J Virol 2003;77:2587–2599.

35. Nhacolo AQ, *et al.*: Levels and trends of demographic indices in southern rural Mozambique: Evidence from demographic surveillance in Manhica district. BMC Public Health 2006;6:291.

36. Arien KK, Vanham G, and Arts EJ: Is HIV-1 evolving to a less virulent form in humans? Nat Rev Microbiol 2007;5:141–151.

37. Marechal V, *et al.*: Increasing HIV type 1 polymorphic diversity but no resistance to antiretroviral drugs in untreated patients from Central African Republic: A 2005 study. AIDS Res Hum Retroviruses 2006;22:1036–1044.

38. Perno CF, *et al.*: Low prevalence of primary mutations associated with drug resistance in antiviral-naive patients at therapy initiation. AIDS 2002;16:619–624.

39. Perno CF, *et al.*: Secondary mutations in the protease region of human immunodeficiency virus and virologic failure in drug-naive patients treated with protease inhibitor-based therapy. J Infect Dis 2001;184:983–991.

40. Connor RI, Sheridan KE, Ceradini D, Choe S, and Landau NR: Change in coreceptor use correlates with disease progression in HIV-1-infected individuals. J Exp Med 1997;185:621–628.

41. Pastore C, Ramos A, and Mosier DE: Intrinsic obstacles to human immunodeficiency virus type 1 coreceptor switching. J Virol 2004;78:7565–7574.

42. Johnston ER, *et al.*: High frequency of syncytium-inducing and CXCR4-tropic viruses among human immunodeficiency virus type 1 subtype C-infected patients receiving antiretroviral treatment. J Virol 2003;77:7682–7688.

43. Ping LH, *et al.*: Characterization of V3 sequence heterogeneity in subtype C human immunodeficiency virus type 1 isolates from Malawi: Underrepresentation of X4 variants. J Virol 1999;73:6271–6281.

44. Coetzer M, *et al.*: Longitudinal analysis of HIV type 1 subtype C envelope sequences from South Africa. AIDS Res Hum Retroviruses 2007;23:316–321.

45. Morris L, Cilliers T, Bredell H, Phoswa M, and Martin DJ: CCR5 is the major coreceptor used by HIV-1 subtype C isolates from patients with active tuberculosis. AIDS Res Hum Retroviruses 2001;17:697–701.

46. Peeters M, *et al.*: Evidence for differences in MT2 cell tropism according to genetic subtypes of HIV-1: Syncytium-inducing variants seem rare among subtype C HIV-1 viruses. J Acquir Immune Defic Syndr 1999;20:115–121.

Address reprint requests to:
Denise Naniche
Barcelona Center for International Health (CRESIB)
IDIBAPS/Hospital Clinic
Villarroel, 170 Barcelona, E-08036 Spain

E-mail: dsuzanne@clinic.ub.es