Intrapatient Variability of HIV Type 1 Group O ANT70 during a 10-Year Follow-up

WOUTER JANSSENS, JOHN NKENGASONG, LEO HEYNDRICKX, GERT VAN DER AUWERA, KATLEEN VEREECKEN, SANDRA COPPENS, BETTY WILLEMS, ELS BEIRNAERT, KATRIEN FRANSEN, MARTINE PEETERS, and GUIDO VAN DER GROEN

ABSTRACT

HIV-1 ANT70 is the first HIV-1 group O virus isolate obtained from a 25-year-old Cameroonian woman, who seroconverted in March 1987. This individual has remained asymptomatic and clinically healthy (clinical stage WHO I, CDC II) even though she did not receive any antiretroviral therapy for HIV-1 before 97 months postseroconversion. CD4+ T cell counts declined steadily to 200/μl at 70 months postseroconversion. The HIV-1 ANT70 nucleotide and amino acid sequence diversity of the V3C3-encoding env fragment within this individual was followed over a 10-year period. RT-PCR, cloning, sequencing, and genetic analyses were performed on eight plasma follow-up samples. Extensive increasing intra- and intersample variation was observed. This is the first long-term (>10 years) follow-up of the genetic variability of an HIV-1 group O-infected individual. As the course of the disease in the HIV-1 ANT70-infected woman was similar in many aspects to that of group M-infected individuals, it remains to be elucidated whether the changes observed in the V3 loop are critical for disease progression.

INTRODUCTION

PHYLOGENETICALLY, three different HIV-1 groups are distinguished: group M (major), group O (outlier), and group N. At the genetic level, the distance between HIV-1 group O and group M is about 50% in the Env protein. The human immunodeficiency virus type 1 (HIV-1) displays important genetic variability, which is driven by a high error rate of the reverse transcriptase, the presence of viral RNA as a dimer, the high turnover rate of HIV-1 in vivo, and selective immune responses. The study of the evolution of HIV-1 sequence variation during progression to AIDS may provide information on consequences of viral heterogeneity with respect to biologic implications and pathogenesis.

On the basis of the preliminary results of an 8-year virologic, immunologic, and clinical follow-up of the individual infected with the HIV-1 group O ANT70 virus and her partner, we previously reported a similar pathogenic course of infection due to group O as compared with group M viruses. The course of infection of the HIV-1 ANT70-infected individual was characterized by a healthy clinical status, a non-syncytium-inducing phenotype, and the presence of autologous neutralizing antibodies 70 and 94 months after the initial isolate. Here we report on the extension of this study to provide insight in the 10-year follow-up of the intrapatient env V3C3 sequence variability of HIV-1 ANT70.

MATERIALS AND METHODS

The HIV-1-infected individual

The Cameroonian woman infected with the prototype group O virus ANT70 seroconverted in March 1987. At diagnosis she was classified as a group A1 according to the Centers for Disease Control (CDC) staging system (Centers for Disease Control, 1992). The ANT70 virus was isolated from this individual and partially characterized; the full genome sequence...
of the ANT70 virus was described. Her spouse was infected with a related virus. The ANT70-infected Cameroonian woman was pregnant in 1988 and delivered a seronegative child. She started antiretroviral therapy in May 1995, when she was enrolled in a trial with combination therapy.

**Determination of CD4⁺ T cell count and viral loads**

HIV-1 ANT70 isolates A1 (the first sample from which the original ANT70 full genome sequence was documented), A2, A3, A4, and A5 (Table 1) were previously analyzed and described in terms of CD4⁺ T cell counts and viral loads. Additional isolates A2b, A6, and A7 were analyzed in this study. Viral loads were determined by a semiquantitative RNA polymerase chain reaction (PCR).

**Viral phenotypic analysis**

Virus was isolated from patient peripheral blood mononuclear cells (PBMCs) by cocultivation with phytohemagglutinin (PHA)-stimulated PBMCs (stimulated for 2 to 3 days) from seronegative blood donors. Cultures with virus supernatant positive for more than 3 weeks were considered positive. Cultures were scored as negative if they remained negative for HIV-1 antigen production after 28 days, using an in-house enzyme-linked immunosorbent assay (ELISA). A total of 10⁴ TCID₅₀ (50% tissue culture infective dose) of each of the viruses was used to infect 2 × 10⁶ MT-2 cells, as described previously.

The GHOST cell lines used were derived from HOS (human osteosarcoma) cells. HOS cells were transfected with a gene encoding human CD4, a gene encoding one of the chemokine receptors (CCR5, CXCR4) that function as HIV-1 coreceptors, and an indicator gene under the control of an HIV-2 promoter (inducible by Tat), which permits evaluation and quantification of HIV-1 infection by flow cytometry. Briefly, cells were seeded in 24-well plates at a concentration of 6 × 10⁴ cells/well. On the next day the cells were infected with undiluted virus stocks of ANT70 isolates, in combination with DEAE-dextran, to a final concentration of 8 µg/ml. After overnight incubation, the virus inoculum was removed and the cells were washed with phosphate-buffered saline (PBS). After 4 to 5 days, the cells were washed again with PBS, resuspended in 1 mM EDTA, and fixed in formaldehyde at a final concentration of 2%. The cells were then analyzed by fluorescence-activated cell sorting (FACS) analysis. The live cells were gated on the basis of forward and side scatter. The number of infected cells was determined by using the scattergram of fluorescence versus forward scatter after setting the gates with uninfected cells. More than 99 of 15,000 cells had to be present for a virus to be considered positive for infectivity in each of the GHOST cell lines tested.

**PCR, cloning, and sequencing of sequential V3 loop samples of ANT70**

The first sample, A1, was taken in March 1987. Eight plasma samples were taken at several time points over a period of nearly 10 years, as indicated in Table 1. RNA extraction and reverse transcriptase (RT)-PCR from these eight follow-up plasma samples were performed essentially as described previously. V3C3 (309 bp)-encoding fragments were amplified by hemi-nested PCR, using primers ENV5CAM and ENV3CAM in the first round, and AV3S (5’ ggcagctgAATCCTACCTAAA-CATGACC 3’; 5’ adaptor with a PvuII restriction site for cloning indicated in lower case) and ENV3CAM in the second-round amplification. For each sample, fragments obtained by one PCR were subsequently cloned and 3 to 12 clones were sequenced. Table 1 provides an overview of the samples taken at different time points. The name of the sample is given, together with the time of sampling, and the number of clones sequenced. In total, 71 sequences were determined. The nucleotide sequence data were deposited in the EMBL, GenBank, and DDBJ Nucleotide Sequence Databases under the following accession numbers: AF147884–AF147954.

**Analysis of nucleotide and amino acid V3C3 sequences**

The determined V3C3 sequences were included in an alignment containing the following previously documented group O sequences: MVP5180, VAU, and DUR; 193Ha, 267Ha, 341Ha, and 655Ha; and CA9 and VI686. Automatic alignment of the determined V3C3 sequences was performed with CLUSTAL W, and DCSE was used for manual editing. The DCSE (dedicated comparative sequence editor) software is a multiple alignment tool that can be used for manually editing amino acid and nucleotide alignments. The entire alignment contained 270 nucleotide positions. The phylogenetic tree was

---

**Table 1. Data from Each HIV-1 ANT70 Follow-up Sample**

<table>
<thead>
<tr>
<th>Virus isolate</th>
<th>Sampling time⁴</th>
<th>Number of clones analyzed</th>
<th>CD4⁺ T cell count (CD4⁺ cells/µl blood)</th>
<th>Viral load (RNA copies/ml plasma)³⁷</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>0</td>
<td>10</td>
<td>—</td>
<td>230</td>
</tr>
<tr>
<td>A2</td>
<td>12</td>
<td>10</td>
<td>—</td>
<td>230</td>
</tr>
<tr>
<td>A2b</td>
<td>27</td>
<td>7</td>
<td>427</td>
<td>2,300</td>
</tr>
<tr>
<td>A3</td>
<td>42</td>
<td>5</td>
<td>277</td>
<td>230</td>
</tr>
<tr>
<td>A4</td>
<td>70</td>
<td>3</td>
<td>186</td>
<td>230</td>
</tr>
<tr>
<td>A5</td>
<td>94</td>
<td>12</td>
<td>156</td>
<td>23,000</td>
</tr>
<tr>
<td>A6</td>
<td>110</td>
<td>12</td>
<td>248</td>
<td>2,300</td>
</tr>
<tr>
<td>A7</td>
<td>118</td>
<td>12</td>
<td>212</td>
<td>2,300</td>
</tr>
</tbody>
</table>

⁴Sampling time is given in months, relative to the first sample (A1).
³Viral loads were determined by a semiquantitative RNA PCR.
constructed with the software package TREECON,\textsuperscript{18} using the Kimura correction to convert dissimilarities into evolutionary distances, disregarding insertions and deletions. Trees were constructed by the neighbor-joining method, and 2000 bootstrap samples were analyzed. The software package MEGA\textsuperscript{19} was used for calculating the number of synonymous substitutions per potential synonymous site \((d_s)\) and the number of non-synonymous substitutions per potential non-synonymous site \((d_{ns})\), using the Jukes and Cantor correction. For the amino acid sequences the \(p\) distance was calculated, which is the proportion of different amino acids between two sequences. In all calculations, sites with alignment gaps were disregarded between each pair of sequences only, but were not deleted from the entire alignment.

**RESULTS**

**CD4\(^+\) T cell count and viral loads**

The changes in CD4\(^+\) cell counts and viral burden in sequential PBMC and plasma samples from the ANT70 individual (Table 1) indicate that this woman controlled her virus until 94 months (A5) after the first virus isolation. The CD4\(^+\) cell counts declined to 156 CD4\(^+\) cells/\(\mu l\) and the viral burden increased to 23,000 copies of RNA per milliliter of plasma at this time point.\textsuperscript{7} After antiviral therapy starting at month 97, CD4\(^+\) cell counts increased to 248 CD4\(^+\) cells/\(\mu l\), and the viral load decreased to 2300 copies of RNA per milliliter of plasma at month 110.

**Viral phenotypic analysis**

Culture of the ANT70 isolates was positive, except for A2b. None of the ANT70 isolates replicated in the MT2 cell line, indicating that viruses were of the slow/low or non-syncytium-inducing (NSI) biological phenotype. The coreceptor usage of the ANT70 isolates was examined on GHOST-3 cells, expressing CCR5 or CXCR4. Virus infection could be measured on GHOST-CCR5 cells for all ANT70 isolates. No virus infection was measured on the cells expressing the CXCR4 coreceptor.

**Relationship between ANT70 follow-up samples and other V3C3 sequences from group O**

Figure 1 shows a phylogenetic tree based on the \textit{env} V3C3 nucleotide sequences present in the alignment. Since it was previously shown on the basis of complete \textit{env} gene sequences that ANT70, the four Equatorial Guinea Ha isolates, and strains VI686 and CA9 belong to the same clade, which most probably represents a group O subtype,\textsuperscript{15} the tree was artificially rooted at this position. The ANT70 patient samples form a separate cluster, which is well supported by a bootstrap value of 97%. As can be seen in the tree, many intrapatient distances are larger than distances between sequences from different patients. The intrapatient nucleotide Kimura distances between the sequences of Fig. 1 range from 0 to 36%, while the interpatient distances range from 5 to 52%. At the amino acid level, the \(p\) distance ranges are 0 to 48% and 10 to 61%, respectively. If one assumes that isolates 655Ha, 341Ha, and 267Ha are epidemiologically linked, which may be supported by their close relationship, the minimum interpatient distances rise to 18% (nucleotides) and 27% (amino acids).

**Variability of the intrapatient ANT70 follow-up samples**

Figure 2a shows the nucleotide and amino acid distances observed in each sample (diversity). The intrasample distances decreased in the sample taken at 27 months, as compared with the samples taken at 0 and 12 months. The diversity again increased from the sample taken at 94 months, on to a level higher than seen in the first two samples.

In Fig. 2b, the average distances of each sample relative to sequences from the first sample is shown (divergence). The number of non-synonymous substitutions per potential non-synonymous site, and therefore, also the number of amino acid substitutions, increases with time. On the other hand, the number of synonymous substitutions per potential synonymous site shows a somewhat less consistent pattern.

**Patterns and phenotypic relevance of amino acid substitutions**

The predicted amino acid composition of the V3C3 region of the various isolates was compared with the reported ANT70 sequence.\textsuperscript{3} The average V3C3 amino acid distances for the different samples as compared with the A1 sequence are shown in Fig. 2b. The pattern of evolution of the V3C3 sequences is presented in Fig. 3.

During the 10-year follow-up various octamer sequences of the tip of the V3 loop were documented: RIGPMAWY, KIGPMAWY, KIGPMSWY, GIGPLAWY, GIGPMAWY, GIGPLSWY, RISSMGRY, and KIGPLSWY (Fig. 3).

A negatively charged amino acid at position 11, E (glutamic acid), is present in all clones of isolate A1 (Fig. 3). A positively charged amino acid at position 11, K (lysine), in isolates A2 (clones 1, 2, 6, 7, and 9) and A5 (clones 1, 3, 4, 6, 10, and 13); and K (lysine) at position 25 in clone 9 of A5, with a net V3 loop charge of 3 to 4 (A2) and 4 to 7 (A5) was observed. Amino acids at positions 25 (except for clone 9, isolate A5) and 37 had a neutral charge. Isolates A2b, A3, A4, A6, and A7 have a negatively charged amino acid, E (glutamic acid), at position 11, as was observed for A1; the amino acids at positions 25 and 37 remain neutral. For group O strains a high net positive charge of the V3 loop and a positively charged amino acid at position 11 or 25 are indicative of SI/MT2 tropic isolates, whereas a low net positive charge and a single positively charged amino acid at position 37 indicate NSI/MT2 tropic isolates.\textsuperscript{20} The results presented in this study confirmed these findings except for clone 9 of isolate A5, where the net positive charge of the V3 loop and the presence of a positive charge at position 25 were suggestive of an SI/MT2-tropic virus (Fig. 3).

**DISCUSSION**

In this study we provide for the first time insight concerning the increase in intrapatient V3C3 heterogeneity during the course of infection of an HIV-1 group O ANT70-infected individual. This individual has remained asymptomatic and clinically healthy (clinical stage WHO 1, CDC II) during the fol-
low-up period, which is consistent with the absence of a syncytium-inducing (SI) biological phenotype, and the presence of high-titered autologous neutralizing antibodies as well as heterologous neutralizing antibodies. The NSI biological phenotype of the ANT70 isolates as indicated by the absence of replication in MT2 cells correlated with virus infection in GHOST-3 cells expressing CCR5 but not CXCR4. In general the NSI status was also in agreement with a low net positive charge of the V3 loop.

The dynamics of the evolution of the viral population in this

**FIG. 1.** Phylogenetic tree based on the V3C3 region of group O env genes, constructed as described in Materials and Methods. Each sequence is named according to the sample taken (see Table 1), followed by the clone number. Bootstrap values are given in percentages at the internodes if they exceed the 70% level. The distance between two sequences is obtained by summing the lengths of the connecting horizontal branches, using the scale on top. The root of the tree is artificially placed at the branch leading to the ANT70 follow-up samples, the four Ha isolates, and isolates VI686 and CA9 (see text for details).
HIV-1 group O-infected individual are in agreement with what has been described for group M infections. A low V3C3 diversity was seen in the first postseroconversion sample. As compared with the initially reported ANT70 sequence, divergence increased in follow-up samples taken during disease progression. The intrasample amino acid diversity varied. A steep rise in intrasample variability was noted in sample 5, 94 months after the first sample was taken. Since the proviral copy num-

**FIG. 2.** Graphical representation of the average distances for each of the eight ANT70 follow-up samples as a function of time, i.e., months after the first sample (A1) was taken. Distances were calculated either by taking into account the entire V3C3 nucleotide alignment, the number of synonymous substitutions per potential synonymous site (ds), the number of nonsynonymous substitutions per potential nonsynonymous site (dn), or the derived amino acid sequences from the V3C3 region. All distances were calculated as described in Materials and Methods. (a) For each consecutive sample, the average distance between the sequences from that time point was calculated, representing the intrasample diversity. (b) The average distance between the sequences from each time point as compared with the initial A1 sequence (divergence) is given for each sample.
ber in the PCR was not quantified prior to cloning, the question remains whether the intrasample conservation of the sequences in samples A3 and A4 may be an artifact due to low copy number and the low number of clones analyzed rather than a reflection of conservation.

As indicated by the phylogenetic tree (Fig. 1), distances between the sequences of the AN70-infected individual are often higher than distances between group O samples from different patients. This illustrates that the intrapatient variation of the V3C3 region is rather high. However, despite this high vari-

FIG. 3. Amino acid alignments of the env V3C3 sequences derived from plasma during progressive infection of the HIV-1 AN70-infected individual. Amino acid identity with the HIV-1 AN70 sequence is represented by dashes; dots are introduced to align the sequences. Potential N-glycosylation sites are indicated by carets. Positions 11, 25, and 37 of the V3 loop are indicated by dollar symbols ($).
ability, the ANT70 sequences form a cluster that is well supported by bootstrap analysis (97%). Moreover, several bootstrap values within the ANT70 cluster are above the 70% level, which indicates that the V3C3 alignment contains sufficient information for investigating at least some relationships within this cluster. On the other hand, the relationships between the interpatient group O sequences are less clearly resolved, which is probably because the V3C3 region is too variable and too short (270 positions) for accurately assessing their phylogenetic position. The early branching of the sequences from sample A5 is somewhat troublesome, in contrast to the sequences from the other time points. The latter follow a pattern that is consistent with the time of sampling, branching off more or less consecutively. However, except for the clustering of samples, A4, A6, and A7, the relationships between the samples are not supported by bootstrap values, indicating that this information cannot be accurately deduced from the data set at hand. The early branching of sample A5 may be caused by the fact that long branches often have the tendency to be attracted to the tree base.

The high quasispecies diversity within the late-stage samples include different patterns of potential N-glycosylation sites as well as amino acid substitutions to proline residues, which may result in an altered presentation of epitopes and an altered immune response. A high quasispecies heterogeneity in two HIV-1 group O-infected individuals during an 18-month follow-up has been reported previously. When comparing inter- and intrahost variability among HIV-1 group O and group M the following observations can be made. The highest genetic distances between group O env sequences are comparable to the highest genetic distances in group M. However, study of the interpatient genetic characterization of the C2V3 region indicates that group O viruses are more diverse than group M viruses in this region, whereas the octameric tip of the V3 loop as well as the distribution of potential N-glycosylation sites are quite heterogeneous. Therefore, the observed intrapatient HIV-1 group O ANT70 heterogeneity pattern may reflect the previously documented HIV-1 group O interpatient diversity, and may explain the poorly supported phylogeny among group O C2V3 env sequences.

Estimation of synonymous and nonsynonymous substitution rates has provided an important tool with which to study the process of molecular sequence evolution. An excess of nonsynonymous substitutions relative to synonymous substitutions has been interpreted as an unambiguous indicator of positive natural selection at the molecular level in HIV-1 genes. In the course of HIV-1 group O ANT70 infection the number of synonymous and nonsynonymous substitutions at potential synonymous and nonsynonymous sites increased at similar rates. Similar observations have been described for group M infections, whereupon it is questioned to what extent a similar increase in synonymous and nonsynonymous substitutions may be considered in itself as evidence of selection for change in amino acid sequence. As different amino acid sites of the HIV-1 group M and O envelopes perform different functional and structural roles, they must be under different selection pressures. With respect to group M infections, Yamaguchi and Gojobori found in their study concentrated amino acid substitutions at five specific sites within the V3 region, which are known to be possibly responsible for production of antigenic variation and determination of the viral phenotypes. These observations indicate a strong possibility that positive selection is taking place by changing particular amino acid sites in the V3 region of HIV within a single host at one time or another after the infection.

The pattern seen in ANT70 is not necessarily representative of group O viruses. Many other factors such as viral load, CD4+ cell counts, and disease progression profiles may be determinant factors that govern evolutionary processes. The observed genetic diversity between HIV-1 group O and group M viruses also reflects antigenic diversity. However, some HIV-1 group M sera neutralized both group M and group O primary isolates, suggesting conserved neutralizing epitopes. Multivariate analysis of inter- and intraclade neutralization data revealed that the neutralization spectra of the HIV-1 group O isolates form a separate cluster. As the course of the disease in the HIV-1 ANT70-infected woman was similar in many ways to that of group M-infected individuals, the reason for the changes observed in the V3 loop is still unclear. Whether they are critical for disease progression remains to be elucidated. The pathogenic course of HIV-1 group O-infected individuals remains poorly documented, and only prospective studies will be able to provide information on the natural course of HIV-1 group O infections.

ACKNOWLEDGMENTS

This work was supported by Grants 3-3301-96 and G.0134.97 from the Fonds voor Wetenschappelijk Onderzoek (Brussels, Belgium); by EC Grant IC18-CT97-02476; and by a grant from the Human Science Foundation (Tokyo, Japan).

REFERENCES


Address reprint requests to:
Wouter Janssens
Department of Microbiology
Institute of Tropical Medicine
Nationalestraat 155
2000 Antwerp, Belgium