Molecular characterization of a complex, recombinant human immunodeficiency virus type 1 (HIV-1) isolate (A/G/J/K/?): evidence to support the existence of a novel HIV-1 subtype

D. Paraskevis,1 E. Magiorkinis,1 G. Magiorkinis,1 C. Anastassopoulou,1 M. Lazanas,2 G. Chrysos,3 A. M. Vandamme4 and A. Hatzakis1

1National Retrovirus Reference Center, Department of Hygiene and Epidemiology, Athens University Medical School, 75 Mikras Asias Street, GR-115 27 Athens (Goudi), Greece
2Red Cross Hospital, Athens, Greece
3Tzanio General Hospital, Piraeus, Greece
4Rega Institute for Medical Research and University Hospitals, Katholieke Universiteit Leuven, Belgium

Recombination is one of several factors that contribute to the great genetic diversity of human immunodeficiency virus type 1 (HIV-1). In the current study, analysis of the full-length genome of a novel complex mosaic HIV-1 isolate (99GR303) from a Greek sailor who was possibly infected in Sierra Leone, Africa is presented. The 99GR303 isolate was found to comprise genomic regions belonging to subtypes A, G, J and K as well as of regions of a subtype that remains unclassified. For a partial region of env as well as vpr, no apparent similarity to the known HIV-1 subtypes or to any of the circulating recombinant forms was found. In fact, in the partial env gene, including the C2-V3 region, the 99GR303 isolate formed a new clade, suggesting the existence of an additional HIV-1 subtype. Thus, novel recombinants embody partial genomic regions which may have originated either from subtypes that existed in the past and became extinct or from contemporary subtypes that are extremely rare.

A hallmark of human immunodeficiency virus (HIV) is its extensive genetic diversity. HIV type 1 (HIV-1) genomes have thus been phylogenetically divided into three main groups, M, O and N (Kuiken et al., 1999). The most widespread viruses of group M have been further divided into nine distinct subtypes, A–D, F–H, J and K, and into 11 circulating recombinant forms (CRFs) (Kuiken et al., 1999). Since the first description of mosaic HIV-1 viruses in 1994 (Robertson et al., 1995a, b), a significant proportion of circulating HIV-1 strains have been found to comprise intersubtype recombinants (Kuiken et al., 1999; Quíñones-Mateu & Arts, 1999). Since then, several recombinant viruses with either a simple chimeric pattern or a more complex mosaic pattern have been described (Kuiken et al., 1999; Quíñones-Mateu et al., 1999). CRF04 cpx, for instance, which represents a recombinant form of HIV-1 found in Cyprus and Greece, consists of at least four distinct subtypes as well as of unclassified regions (Kostrikis et al., 1995; Gao et al., 1998; Nasioulas et al., 1999; Paraskevis et al., 2001). Several other mosaic HIV-1 sequences have also been described to consist of partial genomic regions that do not fall within any of the HIV-1 subtypes characterized previously (Kuiken et al., 1999; Quíñones-Mateu et al., 1999). These observations suggest that mosaic genomes may be looked upon as the only contemporary windows through which one can catch a glimpse of additional HIV-1 subtypes that existed in the past and were driven to extinction or rarity.

In the current study, we present the analysis of the complete full-length HIV-1 sequence (99GR303) isolated from a Greek sailor who was probably infected in Africa, a location to which he had travelled several times; it is thought that the HIV-1 infection was acquired in Sierra Leone. For this patient, HIV-1 seropositivity was first documented in 1996; plasma samples have been obtained on a regular basis since then for follow-up. For the purposes of the current study, RNA obtained from cryopreserved plasma samples from 1999 was utilized.

RNA was extracted from plasma using the Total RNA Isolation kit (Ambion). Catalysed by reverse transcriptase (RT), cDNA was synthesized with random hexamers and oligo(dT) primers from the GeneAmp RNA PCR kit (Roche), according to manufacturer’s recommendations.

All genomic regions were amplified by nested PCR using HIV-1-specific primers, as described previously (Paraskevis et al., 2000). More specifically, the complete genome of the 99GR303 isolate was obtained by nested PCR amplification of
overlapping DNA fragments of approximately 500 bp in length. All second-round PCR products were cloned into the pT-Adv vector (Clontech) and, for each PCR fragment, an individual clone was sequenced directly on both complementary strands using a VGI automated DNA sequencer (Visible Genetics).

Phylogenetic analysis of two fragments of the partial RT region of pol, which was initially examined for monitoring genotypic resistance to anti-retroviral drugs, revealed that the phylogenetic position of the 99GR303 isolate differed in these fragments, suggesting that it may contain an intersubtype recombinant (data not shown). To further investigate the recombination pattern of this isolate, the full-length genome was sequenced and analysed in detail.

To examine for any potential relationships of the 99GR303 isolate with any of the HIV-1 subtypes A–D, F–H, J and K characterized previously, we performed a bootscanning plot using the SimPlot software (Fig. 1) (Ray, 1998). More specifically, trees were constructed for a window of 400 bp moving in steps of 50 bp. Bootstrap values obtained supported the clustering of the 99GR303 isolate with the different subtypes and were plotted across the alignment (Gao et al., 1996). The bootscanning plot suggested that the 99GR303 isolate comprised genomic regions belonging to subtypes A, G, K and J, as well as of regions that did not show any close relationship with any of the HIV-1 subtypes characterized previously (Fig. 1). Subtype classification for each region was confirmed by phylogenetic analysis using the neighbour-joining method with Kimura's two-parameter correction (Kimura, 1980) and programs of the PHYLIP package (Felsenstein, 1993). The reliability of the phylogenetic trees produced was estimated by bootstrapping for 100 replicates. Representative trees of the analysis are shown in Fig. 2. For the region spanning nucleotides 2500–2950, the clustering of the 99GR303 isolate with subtype K was not significantly supported (Fig. 1). Phylogenetic analysis including an additional sequence of subtype K (accession no. AJ249235) showed that the 99GR303 isolate clusters significantly with subtype K for the above region (Fig. 2). More specifically, the p17/p24 gag region, the 5’ RT gene, a middle region of pol and most of env classified as subtype A. The 3’ gag/pol region and the 3’ portion of pol clustered with subtype G. The 3’ pol/vif region, the first exons of tat and rev, as well as vpu/env and a middle portion of pol were classified as subtypes J and K, respectively. Nevertheless, the remaining genomic regions of vpr as well as the C2-V3 region of env could not be classified with any of the HIV-1 subtypes characterized previously and, thus, remain unclassified. Since all known HIV-1 subtypes comprise distinct clades in the C2-V3 region, the existence of a new branch represented by the 99GR303 isolate suggests the existence of an additional HIV-1 subtype (Fig. 2).

Based on these analyses, the recombination pattern of the full-length genomic sequence of the 99GR303 isolate is displayed in Fig. 3. The sequence of the 99GR303 isolate presents an unusually complex mosaic pattern, consisting of subtypes A, G, J and K as well as of unclassified regions. For the C2-V3 region of env, results suggest that it probably originated from an additional HIV-1 subtype.

To examine further the C2-V3 region, all recombinant sequences for which nucleotide sequence data were available,
A complex A/G/J/K/? recombinant HIV-1 strain

Fig. 2. Representative phylogenetic trees in different genomic regions. Positions in the alignments are shown above each tree. Bootstrap values are indicated at the nodes of each branch. Only values greater than 75 are shown.

at least for the complete env, were re-analysed. Our analysis revealed that none of the above sequences, which comprise all known recombinant env sequences to date, include any recombination breakpoints within the C2-V3 region. Similarly, for none of the ten CRFs characterized previously was a recombination breakpoint observed within this region. To examine further the degree of similarity of the C2-V3 region of the 99GR303 isolate to other HIV-1 sequences, a BLAST search against all sequences was performed using the default settings. Interestingly, the 99GR303 isolate best matched the partially characterized HIM389775 (accession no. AJ389775) sequence isolated in Nigeria (Peeters et al., 2000). Phylogenetic analysis revealed that the 99GR303 and HIM389775 isolates cluster together in the C2-V3 region, suggesting that additional HIV-1 isolates contain partial genomic regions of this novel sequence (Fig. 3). Both sequences have been isolated from Western Africa, but due to the fact that the HIM389775 isolate has been characterized only in a partial region of env, it is not possible to conclude whether these sequences display a similar recombination pattern.

Analysis of the full-length HIV-1 sequence isolated from a Greek subject who was probably infected in Africa revealed a complex mosaic genomic organization comprising subtypes A,
D. Paraskevis and others

Fig. 3. (a) Mosaic structure of the 99GR303 isolate genomic sequence. Partial regions classified as subtypes A, G, J, K and unclassified are shown. Open boxes indicate positions with uncertain classification. (b) Phylogenetic analysis of the 99GR303 isolate in the C2-V3 region of env, including the HIM389775 isolate as well as sub-subtype A₂ sequences.

G, J, K and regions of indeterminate classification. Interestingly, the 99GR303 isolate formed an additional cluster in the C2-V3 region of env, where all HIV-1 subtypes characterized previously constitute distinct clades, suggesting that this isolate represents a novel HIV-1 subtype.

The C2-V3 region has been sequenced extensively (Kuiken et al., 1999) and constitutes one of the most highly recommended regions of sequence for use in subtyping HIV-1. The fact that several CRFs, such as CRF01 AE and CRF04 cpx, constitute distinct clades in this region renders it appropriate for the documentation of additional subtypes. Interestingly, analysis of HIV-1 mosaic sequences available to date revealed that there were no recombination breakpoints in this specific genomic region. The absence of recombination breakpoints
may be explained by the fact that this region is the most variable region of the HIV-1 genome (Kuiken et al., 1999) and, thus, according to the proposed models for recombination, it is not an ideal region for RT jumps (Hu & Temin, 1990).

The re-analysis of all previously partially characterized HIV-1 recombinants in env revealed that none of them was closely related to the 99GR303 isolate in the C2-V3 region. Nevertheless, the 99GR303 isolate clustered together with the HIM389775 isolate in the above region. Interestingly, HIM389775 was isolated from Nigeria, suggesting that an unrecognized HIV-1 subtype found in the partial genomic regions of the 99GR303 and HIM389775 isolates possibly existed in Western Africa or spread from other African regions.

The genome of the 99GR303 isolate is a complex recombinant comprising at least five distinct subtypes. Subtypes J and K are extremely rare in Western Africa (Heyndrickx et al., 1999). An unrecognized HIV-1 subtype found in the partial genomic regions of the 99GR303 and HIM389775 isolates possibly existed in Western Africa or spread from other African regions. The genome of the 99GR303 isolate is a complex recombinant comprising at least five distinct subtypes. Subtypes J and K are extremely rare in Western Africa (Heyndrickx et al., 1999; Ishikawa et al., 1996; Takehisa et al., 1997; Ellenberger et al., 1999; McCutchan et al., 1999), suggesting that putative recombination events involving these subtypes could have occurred in other geographical regions. The identification of additional mosaic HIV-1 viruses consisting of partial regions that show similarity to this novel sequence could provide further information on how these isolates originated.

Identification of this partial genomic region in a mosaic sequence, which probably originated from an additional and previously unrecognized HIV-1 subtype, suggests that complex recombinants embody partial genomic regions originating from HIV-1 clades that either existed in the past and became extinct or are extremely rare. Not surprisingly, unclassified regions have been found more frequently in complex recombinants isolated mainly from Central Africa where the greatest genetic heterogeneity of the virus has been documented (Louwagie et al., 1995; Takehisa et al., 1998; Candotti et al., 1999; Kuiken et al., 1999; Mboudjeka et al., 1999; Mokili et al., 1999; Müller-Trutwin et al., 2001; Kalish et al., 2001). This observation may be explained by several reasons. Firstly, Central Africa is the geographical region where HIV-1 originated and evolved and, consequently, several other subtypes have existed in the past and probably became extinct or are extremely rare (Korber et al., 2000; Salemi et al., 2001). Secondly, the greatest genetic variability of the virus has been documented in this region and, therefore, several recombination events among different subtypes in this region could have given rise to complex recombinants (Paraskevis & Hatzakis, 1999; Kalish et al., 2001).

The extinction of additional subtypes which may have existed in the past need not necessarily be explained by altered biological properties of the viruses, but may also have resulted from epidemiological circumstances. Further studies are required to clarify the issue of whether the various HIV-1 subtypes or CRFs differ significantly in their biological properties, such as infectivity, disease progression or response to therapy.

This work was supported by the Hellenic Scientific Society for the Study of AIDS and Sexually Transmitted Diseases. D.P. and C.A. were supported by the Hellenic Center for Infectious Diseases Control (HCIDC) of the Ministry of Health and Welfare.

References


and Human Retrovirology

AIDS Reviews

Central African Republic.

virus type 1 isolates of African origin.

diversity of the envelope glycoprotein from human immunodeficiency


Molecular characterization of a recombinant HIV

Paraskevis, D., Magiorkinis, M., Vandamme, A. M., Kostrakis, L. G. &


1 isolates from Cyprus and Greece, initially designated as ‘subtype I’,

reveals a unique complex A/G/H/K/? mosaic pattern. Journal of General

Virology 82, 575–580.

Peeters, M., Esu-Williams, E., Vergne, L., Montavon, C., Mulanga-

Kabeya, C., Harry, T., Ibirone, A., Lesage, D., Patel, D. & Delaporte,

E. (2000). Predominance of subtype A and G HIV type 1 in Nigeria,

with geographical differences in their distribution. AIDS Research and

Human Retroviruses 16, 315–325.

Quiñones-Mateu, M. L. & Arts, E. J. (1999). Recombination in HIV-1:

update and implications. AIDS Reviews 1, 89–100.


Hopkins University School of Medicine, Baltimore, MD, USA. http://

www.med.jhu.edu/deptmed/sray/download/.


Robertson, D. L., Sharp, P. M., McCutchan, F. E. & Hahn, B. H.


Salemi, M., Strimmer, K., Hall, W. W., Duffy, M., Delaporte, E., Mboup,


ancestor of SIVcpz and HIV-1 group M and the origin of HIV-1 subtypes

using a new method to uncover clock-like molecular evolution. FASEB

Journal 15, 276–278.

Takehisa, J., Osei-Kwasi, M., Ayisi, N. K., Hishida, O., Miura, T.,

Igarashi, T., Brandful, J., Ampofo, W., Netty, V. B., Mensah, M.,


of human immunodeficiency virus 1 in Ghana. Acta Viroligica 41,

51–54.

Takehisa, J., Zekeng, L., Ido, E., Mbondjeka, I., Moriyama, H., Miura,


Various types of HIV mixed infections in Cameroon. Virology 243,

1–10.

Received 11 April 2001; Accepted 25 June 2001