

Identification of a novel HIV-1 complex circulating recombinant form (CRF18_cpx) of Central African origin in Cuba

Michael M. Thomson^a, Gema Casado^a, David Posada^b,
María Sierra^a and Rafael Nájera^a

Background: Analysis of partial *pol* and *env* sequences have indicated a high diversity of HIV-1 genetic forms in Cuba, including two potential novel circulating recombinant forms (CRF): U^{pol}/H^{env} and D^{pol}/A^{env}.

Objectives: To determine whether U^{pol}/H^{env} recombinant viruses from Cuba, detected in 7% of samples, represent a novel HIV-1 CRF, and to identify non-Cuban viruses related to this recombinant form.

Methods: Near full-length genome amplification was carried out by nested polymerase chain reaction in four overlapping DNA segments of two epidemiologically unlinked viruses in uncultured peripheral blood mononuclear cells. The sequences were analysed phylogenetically. Recombinant structures and phylogenetic relationships were analysed by bootscanning and by maximum likelihood. Searches for related viruses in databases were initially based on sequence homology and sharing of signature nucleotides.

Results: Both Cuban viruses clustered uniformly in bootscans all along the genome with each other and with a virus from Cameroon, CM53379, indicating that all three represent the same recombinant form. Their genome comprised multiple segments clustering with subtypes A1, F, G, H and K, as well as segments failing to cluster with recognized subtypes. The newly defined CRF, designated CRF18_cpx, was phylogenetically related in partial segments to CRF13_cpx, CRF04_cpx and 36 additional viruses, most of them from Central Africa. One of the viruses from Cameroon, sequenced in the near full-length genome, was a CRF18_cpx/subtype G secondary recombinant.

Conclusions: A novel HIV-1 complex circulating recombinant form (CRF18_cpx) has been identified that is circulating in Cuba and Central Africa.

© 2005 Lippincott Williams & Wilkins

AIDS 2005, **19**:1155–1163

Keywords: HIV-1, genetic recombination, HIV sequence variability/subtypes, Cuba, Cameroon, Africa

Introduction

HIV-1 derives from at least three independent introductions into humans from chimpanzees in Central Africa, each originating one of the three phylogenetic groups of

HIV-1 [1]. Subsequently, HIV-1 strains belonging to the group responsible for the pandemic (main or M group) have diversified extensively into numerous variants through high rates of mutation [2], recombination [3–5] and viral turnover [6,7]. Some of the variants have propagated

From the ^aCentro Nacional de Microbiología, Instituto de Salud Carlos III, Majadahonda, Madrid, and the ^bDepartment of Biochemistry, Genetics and Immunology, University of Vigo, Vigo, Spain.

Correspondence to Dr. R. Nájera, Centro Nacional de Microbiología, Instituto de Salud Carlos III, Ctra. Majadahonda-Pozuelo, Km 2, 28220 Majadahonda, Madrid, Spain.

E-mail: rafael.najera@isciii.es

Received: 28 July 2004; revised: 21 December 2004; accepted: 3 March 2005.

epidemically and are known as subtypes and circulating recombinant forms (CRF) [8]. Currently, nine subtypes and 16 CRF are recognized [9]. In addition, unique recombinant forms (URF), generated in individuals infected with two or more HIV-1 clades (including subtypes, CRF or groups), are frequently found in areas where multiple variants are circulating [10]. In the Americas, subtype B is the predominant genetic form, although CRF12_BF and related URF are highly prevalent in Argentina and Uruguay [11–14], and subtypes F and C, as well as recombinants between these subtypes and subtype B, are relatively common in Brazil [15–18]. In the Caribbean area, subtype B is also predominant [19]; however, recently, we have identified high diversity of genetic forms in Cuba by analysis of partial *pol* and *env* sequences, including seven subtypes, two probably novel CRF and diverse URF [20]. One of the putative CRF, detected in 21 of 105 samples, grouped in *pol* with subtype D and in *env* with subtype A, and the other, detected in 7 of 105 samples, failed to group with known subtypes in *pol* and grouped with subtype H in *env*. Three URF derived from recombination between both mosaic forms, and six between either of them and subtypes B or G, were also identified. Several viruses from Central Africa grouped in *env* with the U^{pol}/H^{env} Cuban recombinants. One of them, CM53379, a complex mosaic virus from Cameroon [21], grouped with the Cuban recombinants both in *env* and *pol*, which suggests an African ancestry of the Cuban viruses.

The present study focused on the genetic analyses of near full-length genomes of U^{pol}/H^{env} recombinants. Previous analyses of partial segments already supported their common ancestry and their circulation in Cuba, based on their clustering with 100% bootstrap support and on secondary recombination, in at least six occasions, with other genetic forms circulating in this country [20]. Here we extend those findings by defining a novel HIV-1 CRF with a highly complex structure that circulates in Cuba and in Central Africa.

Materials and methods

Samples

Lysates from peripheral blood mononuclear cells of two Cuban HIV-1-infected individuals were used for polymerase chain reaction (PCR) amplification. CU14 was from a heterosexually infected woman from Havana City, diagnosed with HIV-1 infection in 1995. CU76 was from a homosexual man from the province of Villa Clara, who first tested HIV-1 seropositive in 1996. Samples were collected in 1999. Both subjects acquired HIV-1 in Cuba since they had not travelled outside the country prior to their HIV-1 diagnosis. The epidemiological questionnaire did not reveal a direct epidemiological link between the individuals.

Amplification and sequence analysis

Near full-length genome amplification was carried out by nested PCR in four overlapping segments and subsequent direct sequencing of the amplified products, as previously described [12,13]. Sequence electropherograms were corrected and assembled with Seqman (DNASTAR, Madison, Wisconsin, USA). Sequences were aligned with those of subtype reference viruses using Clustal X [22], and alignments were manually edited with Bioedit (Tom Hall, <http://www.mbio.ncsu.edu/BioEdit/bioedit.html>), considering the predicted amino acid sequences. Analyses of the phylogenetic relationships of the recombinants along the genome with each other and with African viruses, and of their mosaic structures, used bootscanning [23] with Simplot 3.2 beta [24]. In these analyses, the bootstrap values supporting the phylogenetic relationship (based on Kimura two-parameter distances and neighbour-joining trees) of the recombinants with each other or with subtype references within a window sliding along the alignment were plotted against the nucleotide position in the genome. Bootstrap values $\geq 70\%$ were considered definitive [25]. Phylogenetic relationships corresponding to the partitions resulting from the bootscan analysis were estimated independently under the maximum likelihood framework. The best-fit model of nucleotide substitution was selected using the Akaike Information Criterion in Modeltest v.3.6 [26]. Maximum likelihood trees were obtained under the best-fit model using the algorithm implemented in Phym1 v.2.4.1 [27], starting from a BIONJ tree [28]. One thousand bootstrap replicates were used to assess phylogenetic confidence [29]. In addition to bootstrap values, the recently defined parameter termed the branching index was used for subtype classification of the segments [30]. The branching index defines the point of divergence of the analysed sequence (sequence X) relative to that of the subtype references. If a is the distance between the origin of the subtype-defining branch and the node of sequence X, and b is the distance between this node and the node of the subtype references, the branching index is defined as the ratio $a/(a + b)$. Based on branching indices calculated for subtype references relative to other references of the same subtype, a cut-off value of ≥ 0.55 was proposed to classify a sequence as belonging to a given subtype [30].

Non-Cuban viruses related to the Cuban recombinants were searched in the Los Alamos HIV Sequence database [9]. Initial screening was carried out with similarity BLAST searches [31] of partial sequences and with searches for sequences containing signature nucleotides characteristic of the Cuban recombinants (nucleotides conserved in the Cuban recombinants but absent or highly unusual in reference viruses of HIV-1 subtypes or CRF) in alignments retrieved from the Los Alamos HIV Sequence Database. The phylogenetic relationships of the database sequences with the Cuban recombinants was analysed by bootscanning and by maximum likelihood as

described above. The newly derived sequences are deposited in Genbank under accession numbers AY586540 and AY586541.

Results

To determine whether the newly derived near full-length genome recombinant sequences from Cuba represent the same recombinant form, their phylogenetic relationship was examined by a bootscan analysis in which both viruses and reference viruses of all subtypes were included. This analysis demonstrated uniform phylogenetic clustering of the Cuban recombinants with each other all along the genome (not shown). Previously, we had reported that a complex recombinant virus from Cameroon, CM53379, sequenced in the near full-length

genome [22], clustered with the Cuban recombinants both in partial *pol* and *env* sequences [21]. In bootscan analysis, CM53379 clustered uniformly with CU14 and CU76 along the entire sequence (Fig. 1a). These results show that CU14, CU76 and CM53379 represent the same recombinant form.

The mosaic structure of this recombinant form was analysed by bootscanning with reference sequences of known subtypes. In this analysis, CU14, CU76 and CM53379, and the references of each subtype, were considered as groups, each group comprising viruses belonging to the same genetic form. The profile of the bootscan plot showed that the recombinants formed multiple segments that clustered with subtypes A, F, G, H and K, and other segments failing to cluster with recognized subtypes (Fig. 1b). The relationship of these segments with subtype references was confirmed in

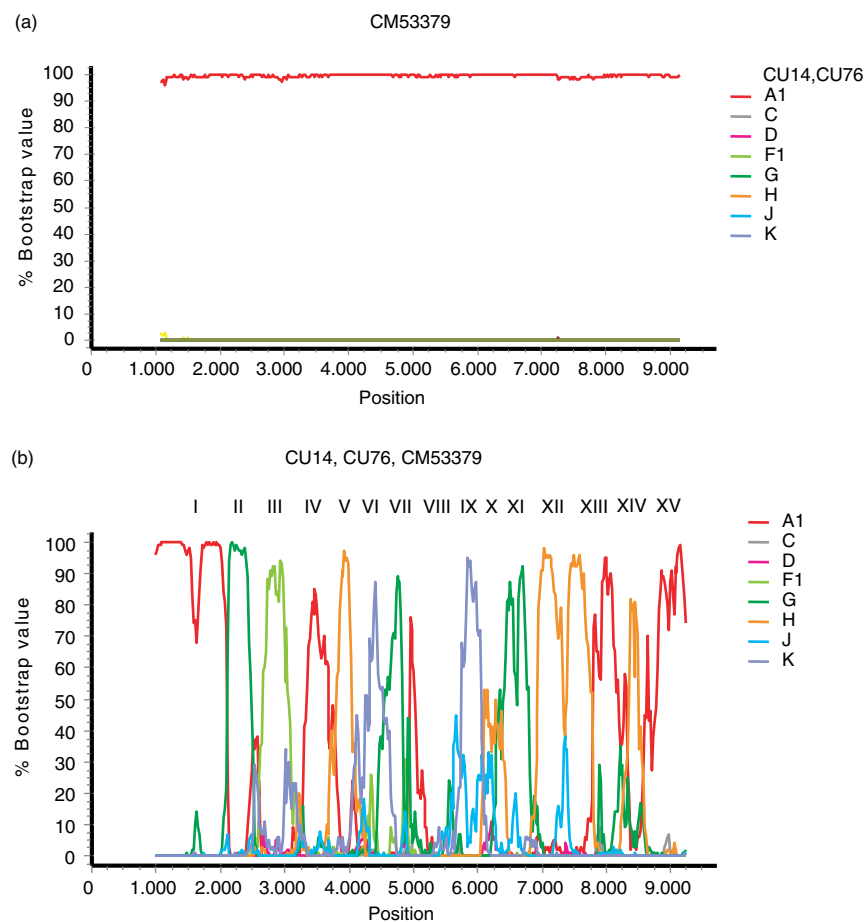


Fig. 1. Bootscan analyses of CU14, CU76 and CM53379. The horizontal axis represents the nucleotide position of the midpoint of a window sliding along the alignment. Nucleotide numeration in the horizontal axis coincides with that of the HXB2 isolate. The percentage bootstrap values were based on 500 replicates of the nodes joining the analysed recombinants (a) with each other or (b) with subtype references in each window. Bootscans were done with Simplot 3.2 beta. Trees were constructed with the neighbour-joining method, based on Kimura's two-parameter distances. Window width was 800 nucleotides (nt) in (a) and 400 nt in (b), moving in 20 nt increments. References used were those included in the subtype reference alignments at the Los Alamos HIV Sequence Database [9] and were considered as groups, each comprising viruses of the same subtype or subs subtype. CU14 and CU76 in (a) and both Cuban recombinants and CM53379 in (b) were analysed as groups.

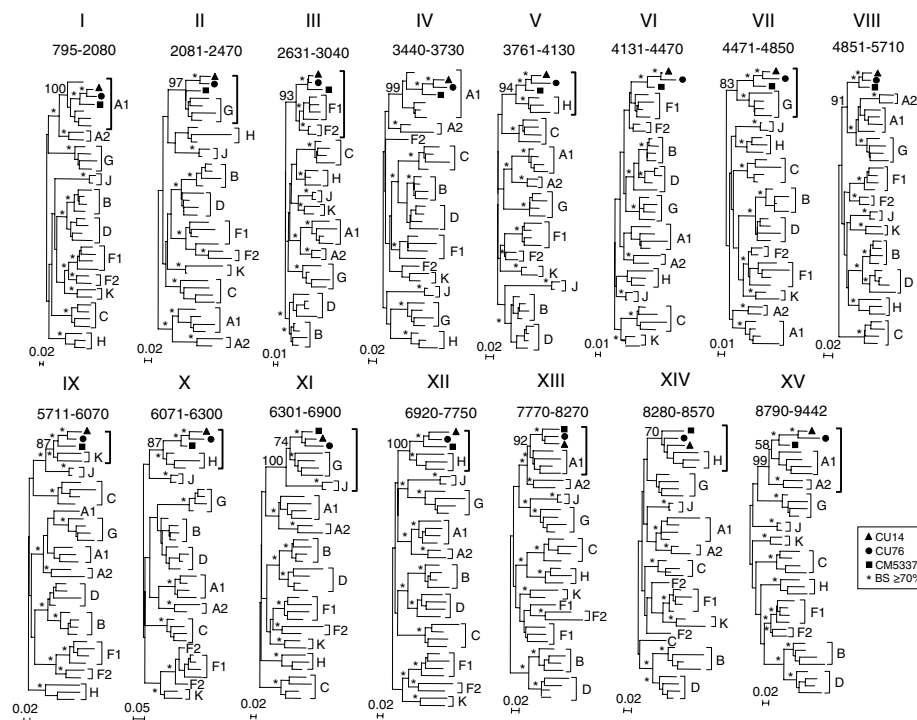


Fig. 2. Maximum likelihood trees for partial genome segments. Trees are numbered according to the partitions defined in the bootscan analysis (Fig. 1b). Nucleotide positions in the HXB2 genome delimiting each analysed segment is shown on top of each tree. Asterisks denote bootstrap values (BS) $\geq 70\%$ for the nodes to the right and below. Bootstrap values supporting the clustering of the analysed recombinants with the corresponding subtype or subsubtype are denoted with their percentage numerical values. Subtype references are those included in the reference alignment at the Los Alamos HIV Sequence Database [9]. Brackets denote subtype or subsubtype clusters. Thick brackets denote subtype or subsubtype clusters that include Cuban recombinants supported by $\geq 70\%$ bootstrap values.

maximum likelihood trees (Fig. 2), which supported the subtype assignments made by the bootscan analysis, except for a segment between positions 4131 and 4470 (partition VI), for which the subtype K affiliation suggested by bootscanning failed to be confirmed, and a segment between 6071 and 6300 (partition X), where the subtype H affiliation was supported with only 53% bootstrap value in bootscanning but with 87% in the maximum likelihood tree. Additional observations were derived from the maximum likelihood analyses: (a) those confirming the bootscan results (Fig. 1a) in that CU14, CU76 and CM53379 grouped in a monophyletic cluster in each of the partitions, except partitions II and XIV (in both partitions, however, several subtype and subsubtype references also failed to group in clusters well supported by bootstrap values); (b) the subtype A segments grouped with references of subsubtype A1, although for the segment between positions 8790 and 9442 in partition XV the bootstrap value supporting the relationship with subsubtype A1 references was only 58%; (c) the subtype F-related segment (partition VI) in *pol* branched outside both F1 and F2 radiations; (d) the segment between positions 4851 and 5710 in partition VIII, which could not be assigned with confidence to any subtype in the bootscan analysis, branched within a cluster comprising

subsubtypes A1 and A2 and subtype G, with a 91% bootstrap support; (e) the decrease in bootstrap values supporting branching with subtype H around the *env* V3 loop region (Fig. 1b) was much less pronounced (to 59%) in maximum likelihood trees estimated from segments of 400 nucleotides (nt) (in this region, the subtype H reference VI991 also failed to cluster consistently with the two other subtype H references in segments of 400 nt or less; therefore, it is possible that the decrease in bootstrap support for the clustering of the recombinants with subtype H in this highly variable *env* segment is a consequence of its derivation from a subtype H strain divergent from the references, rather than of recombination with another subtype). Only in two partitions, of subsubtype A1 in *gag* (partition I) and *pol* (partition IV), the cluster formed by CU14, CU76 and CM53379 branched interspersed within the reference viruses. In the remaining segments, the recombinants diverged basally to the subtype references (partitions II and XIV) or formed a sister clade with the corresponding subtype or subsubtype clusters (partitions V, VII, IX–XIII and XV). To define the subtype assignment for these segments further, the branching index was determined; this represents the point of divergence of the recombinants in the subtype-defining branch of the analysed segment relative to the subtype

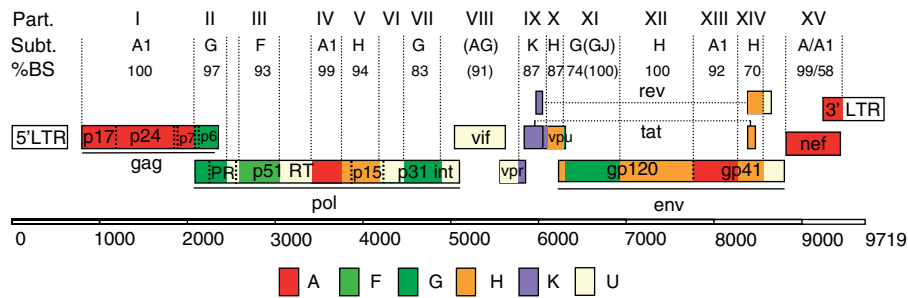


Fig. 3. Schematic depiction of the mosaic structure of CRF18_cpx, based on bootscan and maximum likelihood analyses (Figs 1b and 2). Numeration of segments analysed by maximum likelihood (Fig. 2) and bootstrap values (BS) supporting subtype or subs subtype classifications of each segment are shown at the top. Multisubtype clusters incorporating CRF18_cpx viruses, and BS supporting such clusters, are shown in parentheses.

references [30]. The branching index was > 0.55 in most segments, supporting the subtype affiliation indicated by bootstrap values, except in the subtype F-related segment (partition III; in which the recombinants branched

outside the F1 and F2 references and, therefore, the branching index could not be applied), in partition V (branching index, 0.52), and in partition VII (branching index, 0.53).

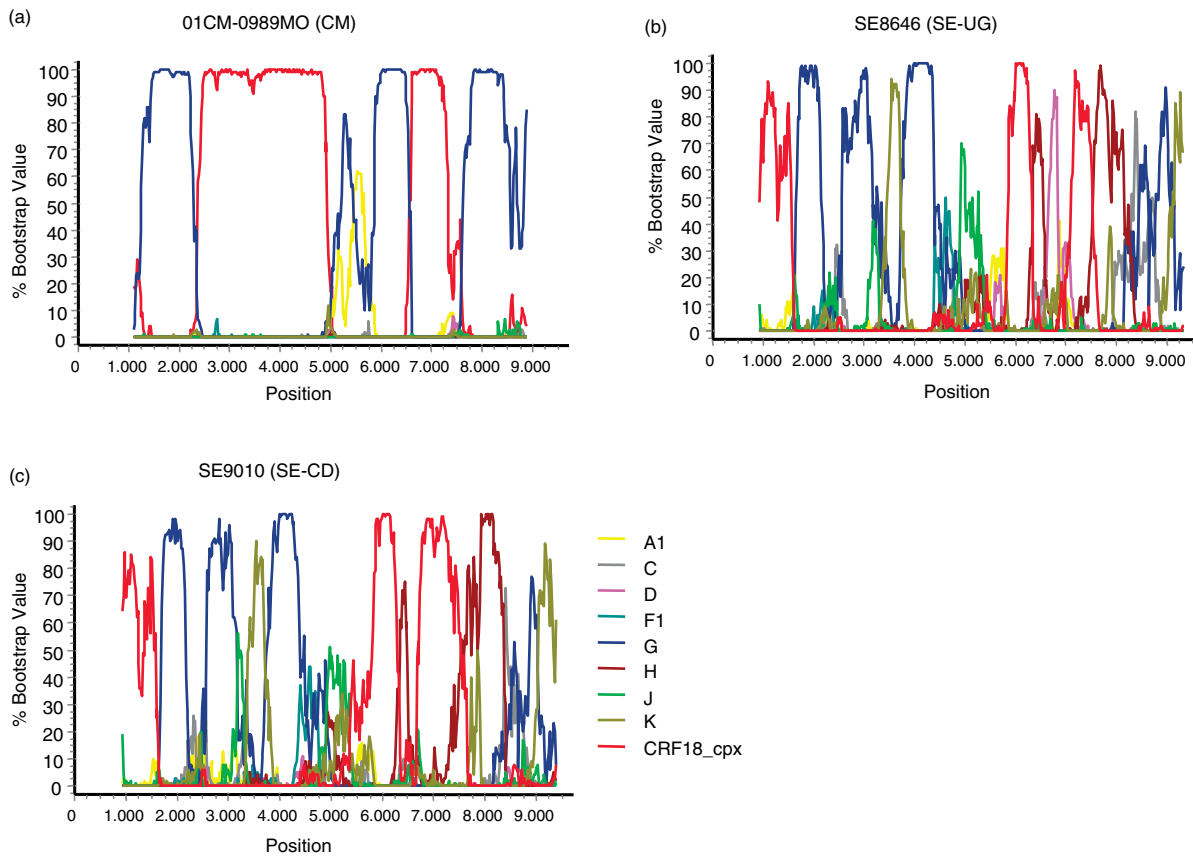


Fig. 4. Phylogenetic analyses showing the relationship of non-Cuban viruses to CRF18_cpx. (a–c) Bootscan analyses of near full-length genomes of three recombinant viruses. Bootscan settings are as in Fig. 1 using a window width of 500 nucleotides. (d–l) Maximum likelihood trees for partial genome segments. Only bootstrap values > 50%, are shown. Nucleotide positions in the HXB2 genome delimiting each analysed segment is shown on top of each tree. Brackets denote clusters of viruses related to CRF18_cpx in the corresponding segment. In (k), a subcluster of viruses from Cameroon more closely related to CM53379 is also bracketed. The two-letter code of countries of origin of the samples are shown in parentheses to the right of each isolate name. When country of sample collection and of probable HIV-1 acquisition differ, both are shown separated by a dash. AO, Angola; BE, Belgium; CD, Democratic Republic of Congo; CM, Cameroon; CY, Cyprus; FR, France; GA, Gabon; GR, Greece; LU, Luxembourg; SE, Sweden; SN, Senegal; UG, Uganda.

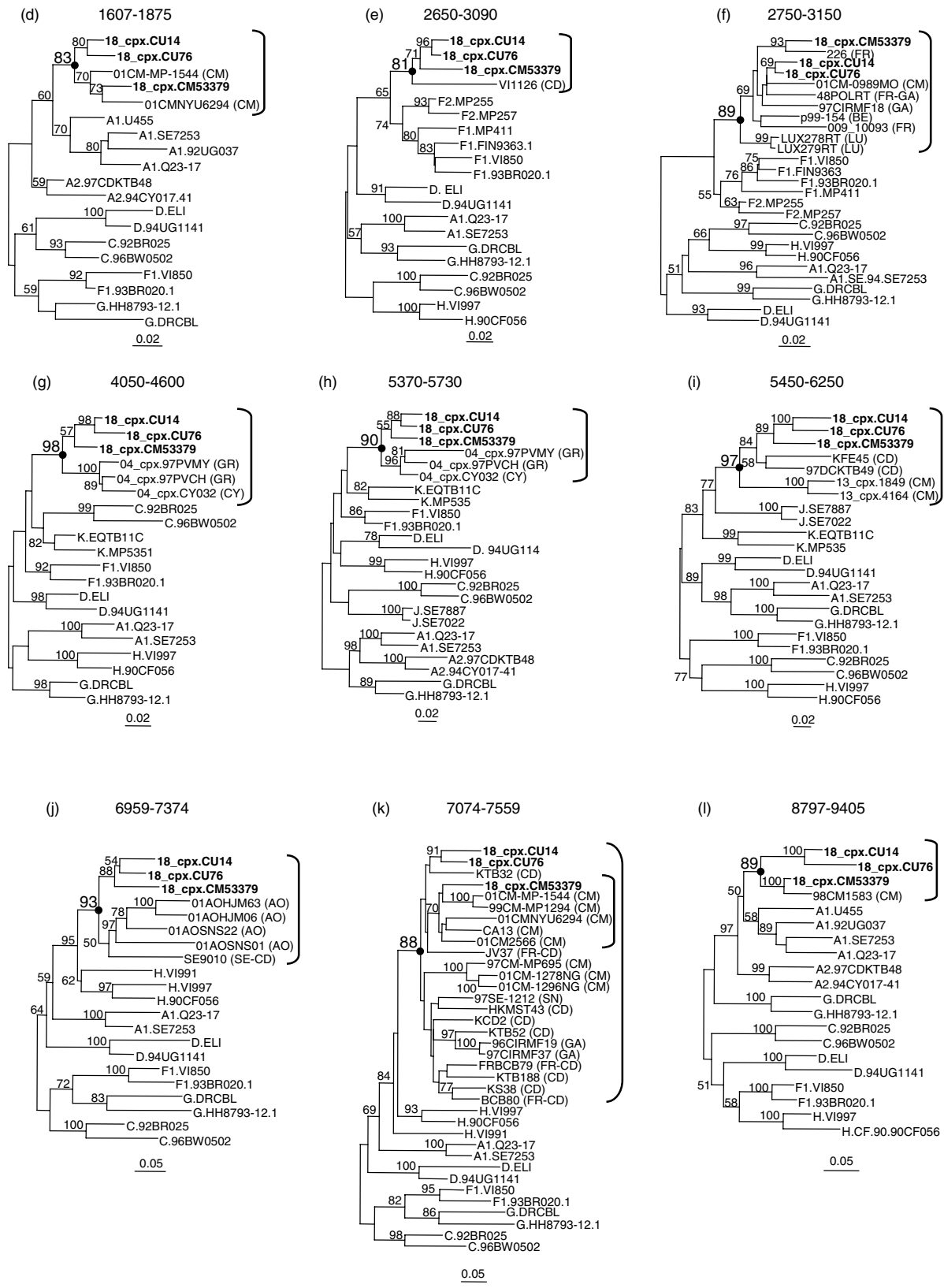


Fig. 4. (continued)

The results of the bootscan and maximum likelihood phylogenetic analyses allowed a new HIV-1 CRF to be defined [8]. According to the order of discovery and to its complex mosaic structure, it was designated CRF18_cpx. Its recombinant structure, together with bootstrap values supporting the subtype assignment for each partition, is schematically shown in Fig. 3.

The relationship of CRF18_cpx with other HIV-1 complex recombinant genomes (either CRF or URF) sequenced in the near full-length genome and deposited in Los Alamos HIV Sequence Database was analysed by bootscanning with Simplot and confirmed by maximum likelihood phylogenetic analysis. Segments of CRF13_cpx, CRF04_cpx and several other unique recombinants from Central Africa were phylogenetically related to CRF18_cpx (Fig. 4). One of these viruses, 01CM-0989MO, from Cameroon, is a CRF18_cpx/subtype G secondary recombinant in which most of *pol* and approximately 0.9 kb in *env* derive from CRF18_cpx (Fig. 4a).

Several searches for other sequences related to CRF18_cpx were completed. An initial screening was implemented by identifying 'signature' nucleotides that were conserved in the Cuban recombinants and CM53379 but were absent or highly unusual in HIV-1 subtypes and CRF references. The presence of those nucleotides was searched for in aligned sequences retrieved from the Los Alamos HIV Sequence Database. Similarities with sequences of different genes, as well as with protease-reverse transcriptase and *env* V3 segments, were also searched by using the HIV-BLAST program available at the Los Alamos HIV Sequence Database [9]. Sequences containing two or more consecutive CRF18_cpx signature nucleotides or having high BLAST similarity scores with CRF18_cpx viruses were subjected to bootscan analyses, and phylogenetic relationships were confirmed by maximum likelihood inference. In total, 40 viruses were found that clustered in partial segments with CRF18_cpx viruses with bootstrap values $\geq 70\%$ (Fig. 4); 30 of these were from Central Africa or of known Central African origin (13 from the Democratic Republic of Congo, nine from Cameroon, four from Gabon and four from Angola).

Discussion

The main objectives of this study were to determine whether HIV-1 Cuban recombinant viruses that grouped in clusters of unknown subtype in *pol* and of H subtype in *env* [20] represent a novel CRF, and to define their recombinant structures. A secondary objective was to identify non-Cuban viruses related to the Cuban recombinants. To accomplish these objectives, we sequenced

approximately 9 kb of each of two epidemiologically unlinked Cuban recombinants, CU14 and CU76. In bootscan analysis, both viruses clustered uniformly along the entire sequence with each other and with a complex recombinant from Cameroon, CM53379 [22] (Fig. 1a), confirming that all three viruses represent the same genetic form and supporting the Central African origin of the Cuban recombinants. Bootscan analysis and maximum likelihood phylogenetic trees indicated that CU14, CU76 and CM53379 are complex recombinants containing segments related to subtypes A (subsubtype A1), F, G, H and K, as well as segments of undefined subtype (Figs 1b and 2). The subtype classification of each segment was supported by bootstrap values $\geq 70\%$ joining the recombinants with alternative subtype references. It was also supported by branching indexes (see above) > 0.55 in most segments apart from three: the subtype F-related segment, for which a branching index could not be determined because it branched outside the F1 and F2 subclusters; and two short subtype G and H segments, for which branching indices were in the uncertainty zone (0.52 and 0.53, respectively). The latter two segments were classified as being of subtypes G and H, respectively, based on bootstrap supports of 94% and 83% and on the existence of other subtype G and H segments along the genome well supported both by bootstrap and branching index values. These results fulfil the requirements to define an HIV-1 CRF. According to the current nomenclature system, it has been designated CRF18_cpx [8].

The circulation of CRF18_cpx in Cuba is supported by epidemiological data. All 14 infections with viruses clustering at least in partial segments with CRF18_cpx [20] were probably acquired in Cuba, since the individuals harbouring them did not report travel outside of the country, and contact-tracing questionnaires failed to find a direct link between any of these individuals.

Searches for homologous sequences in databases yielded 40 viruses phylogenetically related in partial segments to CRF18_cpx, including CRF13_cpx and CRF04_cpx. Of these, 30 viruses were from samples collected in Central Africa or from individuals probably infected in this area (mostly the Democratic Republic of Congo and Cameroon, but also from Gabon and Angola). The remaining viruses were from samples collected in Senegal and five European countries. Except for the CRF04_cpx viruses, circulating among injecting drug users in Greece [32,33], epidemiological information is not available, and these viruses might also have been acquired in Central Africa. The bootscan analysis of one virus from Cameroon, 01CM-0989MO, sequenced in the near full-length genome, indicates that it is recombinant between CRF18_cpx and subtype G (Fig 4a).

Two viruses from Cameroon, 01CM-MP-1544 and 01CMNYU6294, had separate segments in *gag* and *env*

that clustered with CRF18_cpx with high bootstrap values, and interdigitated between CM53379 and the Cuban recombinants in the tree of the *gag* segments (Fig. 4d,k); therefore, these are strong candidates to be representatives of CRF18_cpx. Three other viruses that grouped with CM53379 within a subcluster in the *env* V3 region were from Cameroon, and another Cameroonian virus clustered with 100% bootstrap support with CM53379 in *nef*. These results suggest that CRF18_cpx is very probably circulating in Cameroon, a country with a high diversity of circulating HIV-1 genetic forms, including seven subtypes and four other CRF (01_AE, 02_AG, 11_cpx and 13_cpx) in addition to groups O and N [21,34–36].

In summary, we have identified CRF18_cpx as a novel HIV-1 circulating recombinant form: a complex mosaic virus with segments of A1, F, G, H and K, and undefined subtypes. CRF18_cpx is circulating in Cuba but originated in Central Africa, where we have detected its presence in Cameroon. We have also found multiple viruses phylogenetically related to CRF18_cpx, at least in part of their genomes, in the Democratic Republic of Congo, Gabon and Angola. Our previously published results indicate that, in addition to CRF18_cpx, at least three other genetic forms, subtypes B and G and another recombinant form (a novel CRF according to our preliminary data), are also circulating in Cuba [20]. In addition, we detected infections with five other subtypes and a recombinant virus probably acquired in Africa [20]. The high HIV-1 genetic diversity in Cuba is especially remarkable considering that the number of estimated HIV-1-infected persons in this country was only 3300 at the end of 2003 [37]. Its cause is presumably related to the large numbers of Cubans that have stayed in various sub-Saharan African countries, including soldiers involved in local armed conflicts in the 1980s and early 1990s and civilians participating in cooperation programmes [38,39].

The list of HIV-1 CRF has been incessantly growing in recent years and will certainly continue to increase, as a consequence of both the discovery of presently circulating, but still unrecognised, CRF and the generation of new CRF. These may emerge in areas where diverse HIV-1 genetic forms are cocirculating, where recombinant viruses are continually being generated in persons infected with two or more variants [10]. In addition to its value in tracking the origin and spread of HIV-1 variants, the identification of novel CRF may have practical implications in the clinical management of patients and in vaccine development, since differences between HIV-1 genetic forms have been reported with respect to the development of antiretroviral drug resistance mutations [40–43], performance in drug-resistance genotypic tests [44,45], accuracy of viral load determinations [46–48] and clade-specificity of immune responses [49,50].

Acknowledgements

We thank Ignacio Ruibal for providing clinical samples, and Alicia Ballester, Pablo Martínez and Aurora de Miguel, at the Genomic Unit of the Centro Nacional de Microbiología, Instituto de Salud Carlos III, for technical assistance.

Sponsorship: DP was funded by the Ramón y Cajal programme of the Spanish Government and by the Ministerio de Educación y Ciencia. This work was funded by grant 36205/01 from FIPSE, Spain.

Note: Michael M. Thomson and Gema Casado contributed equally to this paper.

References

- Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, Michael SF, *et al.* **Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*.** *Nature* 1999; **397**:436–441.
- Gao F, Chen Y, Levy DN, Conaway JA, Kepler TB, Hui H. **Unselected mutations in the human immunodeficiency virus type 1 genome are mostly nonsynonymous and often deleterious.** *J Virol* 2004; **78**:2426–2433.
- Rhodes T, Wargo H, Hu W-S. **High rates of human immunodeficiency virus type 1 recombination: near-random segregation of markers one kilobase apart in one round of viral replication.** *J Virol* 2003; **77**:11193–112000.
- Levy DN, Aldrovandi GM, Kutsch O, Shaw GM. **Dynamics of HIV-1 recombination in its natural target cells.** *Proc Natl Acad Sci USA* 2004; **101**:4204–4209.
- Shriner D, Rodrigo AG, Nickle DC, Mullins JI. **Pervasive genomic recombination of HIV-1 in vivo.** *Genetics* 2004; **167**:1573–1583.
- Wei X, Ghosh SK, Taylor ME, Johnson VA, Emini EA, Deutsch P, *et al.* **Viral dynamics in human-immunodeficiency-virus type-1 infection.** *Nature* 1995; **373**:117–122.
- Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. **HIV-1 dynamics in vivo: virion clearance rate, infected cell lifespan, and viral generation time.** *Science* 1996; **271**:1582–1586.
- Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, *et al.* **HIV-1 nomenclature proposal.** *Science* 2000; **288**:55–56.
- Theoretical Biology and Biophysics Group. *HIV Sequence Database*. Los Alamos: Los Alamos National Laboratory; <http://hiv-web.lanl.gov>.
- Nájera R, Delgado E, Pérez Álvarez L, Thomson MM. **Genetic recombination and its role in the development of the HIV-1 pandemic.** *AIDS* 2002; **16**:S3–S12.
- Thomson MM, Villahermosa ML, Vázquez de Parga E, Cuevas MT, Delgado E, Manjón N, *et al.* **Widespread circulation of a B/F intersubtype recombinant form among HIV-1-infected individuals in Buenos Aires, Argentina.** *AIDS* 2000; **14**:897–899.
- Thomson MM, Delgado E, Herrero I, Cuevas MT, Delgado E, Manjón N, *et al.* **Diversity of mosaic structures and common ancestry of human immunodeficiency virus type 1 BF intersubtype recombinant viruses from Argentina revealed by analysis of near full-length genome sequences.** *J Gen Virol* 2002; **83**:107–119.
- Sierra M, Thomson MM, Ríos M, Casado G, Ojea de Castro R, Delgado E, *et al.* **The analysis of near full-length genome sequences of human immunodeficiency virus type 1 BF intersubtype recombinant viruses from Chile, Venezuela and Spain reveals their relationship to diverse lineages of recombinant viruses related to CRF12_BF.** *Infect Genet Evol* 2005; in press.
- Carr JK, Ávila M, Gómez-Carrillo M, Salomon H, Hierholzer J, Watanaveeradej V, *et al.* **Diverse BF recombinants have spread widely since the introduction of HIV-1 into South America.** *AIDS* 2001; **15**:F41–F47.

15. Ramos A, Tanuri A, Schechter M, Rayfield MA, Hu DJ, Cabral MC, *et al.* **Dual and recombinant infections: an integral part of the HIV-1 epidemic in Brazil.** *Emerg Infect Dis* 1999; **5**:65–74.
16. Guimaraes ML, dos Santos Moreira A, Loureiro R, Galvao-Castro B, Morgado MG. **Brazilian Network for HIV Isolation and Characterization. High frequency of recombinant genomes in HIV type 1 samples from Brazilian southeastern and southern regions.** *AIDS Res Hum Retroviruses* 2002; **18**:1261–1269.
17. Soares MA, De Oliveira T, Brindeiro RM, Diaz RS, Sabino EC, Brigido L, *et al.* **A specific subtype C of human immunodeficiency virus type 1 circulates in Brazil.** *AIDS* 2003; **17**:11–21.
18. Thomson MM, Sierra M, Tanuri A, May S, Casado G, Manjón N, *et al.* **The analysis of near full-length genome sequences of HIV-1 BF intersubtype recombinant viruses from Brazil reveals their independent origins and their lack of relationship to CRF12_BF.** *AIDS Res Hum Retroviruses* 2004; **20**:1126–1133.
19. Vaughan HE, Cane P, Pillay D, Tedder RS. **Characterization of HIV type 1 clades in the Caribbean using pol gene sequences.** *AIDS Res Hum Retroviruses* 2003; **19**:929–932.
20. Cuevas MT, Ruibal I, Villahermosa ML, Díaz H, Delgado E, Delgado E, Vázquez de Parga E, *et al.* **High HIV-1 genetic diversity in Cuba.** *AIDS* 2002; **16**:1643–1653.
21. Carr JK, Torimiro JN, Wolfe ND, Eitel MN, Kim B, Sanders-Buell E, *et al.* **The AG recombinant IbNG and novel strains of group M HIV-1 are common in Cameroon.** *Virology* 2001; **286**:168–181.
22. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. **The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools.** *Nucl Acids Res* 1997; **25**:4876–4882.
23. Salminen MO, Carr JK, Burke DS, McCutchan FE. **Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning.** *AIDS Res Hum Retroviruses* 1995; **11**:1423–1425.
24. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, *et al.* **Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination.** *J Virol* 1999; **73**:152–160.
25. Hillis DM, Bull JJ. **An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis.** *Sys Biol* 1993; **42**:182–192.
26. Posada D, Crandall KA. **Modeltest: testing the model of DNA substitution.** *Bioinformatics* 1998; **14**:817–818.
27. Guindon S, Gascuel O. **A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood.** *Syst Biol* 2003; **52**:696–704.
28. Gascuel O. **BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data.** *Mol Biol Evol* 1997; **14**:685–695.
29. Felsenstein J. **Confidence limits on phylogenies: an approach using the bootstrap.** *Evolution* 1985; **39**:783–791.
30. Wilbe K, Salminen M, Laukkanen T, McCutchan F, Ray SC, Albert J, *et al.* **Characterization of novel recombinant HIV-1 genomes using the branching index.** *Virology* 2003; **316**:116–125.
31. Altschul SF, Gish W, Miller W, Myers E, Lipman DJ. **Basic local alignment search tool.** *J Mol Biol* 1990; **215**:403–410.
32. Gao F, Robertson DL, Carruthers CD, Li Y, Bailes E, Kostrikis LG, *et al.* **An isolate of human immunodeficiency virus type 1 originally classified as subtype I represents a complex mosaic comprising three different group M subtypes A, G, and I.** *J Virol* 1998; **72**:10234–10241.
33. Nasioulas G, Paraskevis D, Magiorkinis E, Theodoridou M, Hatzakis A. **Molecular analysis of the full-length genome of HIV type 1 subtype I: evidence of A/G/I recombination.** *AIDS Res Hum Retroviruses* 1999; **15**:745–758.
34. Wilbe K, Casper C, Albert J, Leitner T. **Identification of two CRF11-cpx genomes and two preliminary representatives of a new circulating recombinant form CRF13-cpx of HIV type 1 in Cameroon.** *AIDS Res Hum Retroviruses* 2002; **18**:849–856.
35. Zhong P, Burda S, Konings F, Urbanski M, Ma L, Zekeng L, *et al.* **Genetic and biological properties of HIV type 1 isolates prevalent in villagers of the Cameroon equatorial rain forests and grass fields: further evidence of broad HIV type 1 genetic diversity.** *AIDS Res Hum Retroviruses* 2003; **19**:1167–1178.
36. Vergne L, Bourgeois A, Mpoudi-Ngole E, Mougnotou R, Mbuagbaw J, Liegeois F, *et al.* **Biological and genetic characteristics of HIV infections in Cameroon reveals dual group M and O infections and a correlation between SI-inducing phenotype of the predominant CRF02_AG variant and disease stage.** *Virology* 2003; **310**:254–266.
37. UNAIDS. *Epidemiological Fact Sheet on HIV/AIDS And Sexually Transmitted Infections.* Geneva: UNAIDS; 2004 update; <http://www.unaids.org>.
38. Torres-Anjel MJ. **Macroepidemiology of the HIV-AIDS (HAIDS) pandemic. Insufficiently considered zoological and geopolitical aspects.** *Ann N Y Acad Sci* 1992; **653**:257–273.
39. Santana S, Faas L, Wald K. **Human immunodeficiency virus in Cuba: the public health response of a Third World country.** *Int J Health Serv* 1991; **21**:511–537.
40. Pérez-Álvarez L, Cuevas MT, Villahermosa ML, Pedreira JD, Manjón N, Herrero I, *et al.* **Prevalence of drug resistance mutations in B, non-B subtypes, and recombinant forms of human immunodeficiency virus type 1 in infected individuals in Spain (Galicia).** *J Hum Virol* 2001; **4**:35–38.
41. Ariyoshi K, Matsuda M, Miura H, Tateishi S, Yamada K, Sugijura W. **Patterns of point mutations associated with antiretroviral drug treatment failure in CRF01_AE subtype E infection differ from subtype B infection.** *J Acquir Immune Defic Syndr* 2003; **33**:336–342.
42. Grossman Z, Istomin V, Averbuch D, Lorber M, Risenberg K, Levi I, *et al.* **Genetic variation at NNRTI resistance-associated positions in patients infected with HIV-1 subtype C.** *AIDS* 2004; **18**:909–915.
43. Kantor R, Katzenstein D. **Drug resistance in non-subtype B HIV-1.** *J Clin Virol* 2004; **29**:152–159.
44. Beddows S, Galpin S, Kazmi SH, Ashraf A, Johargy A, Frater AJ, *et al.* **Performance of two commercially available sequence-based HIV-1 genotyping systems for the detection of drug resistance against HIV type 1 group M subtypes.** *J Med Virol* 2003; **70**:337–342.
45. Maes B, Schrooten Y, Snoeck J, Derdelinckx I, van Ranst M, Vandamme AM, *et al.* **Performance of ViroSeq HIV-1 Genotyping System in routine practice at a Belgian clinical laboratory.** *J Virol Meth* 2004; **119**:45–49.
46. Amendola A, Bordi L, Angeletti C, Visco-Comandini U, Abbate I, Cappiello G, *et al.* **Underevaluation of HIV-1 plasma viral load by a commercially available assay in a cluster of patients infected with HIV-1 A/G circulating recombinant form CRF02.** *J Acquir Immune Defic Syndr* 2002; **31**:488–494.
47. Antunes R, Figueiredo S, Bartolo I, Pinheiro M, Rosado L, Soares I, *et al.* **Evaluation of the clinical sensitivities of three viral load assays with plasma samples from a pediatric population predominantly infected with human immunodeficiency virus type 1 subtype G and BG recombinant forms.** *J Clin Microbiol* 2003; **41**:3361–3367.
48. Gottesman BS, Grosman Z, Lorber M, Levi I, Shitrit P, Mileguir F, *et al.* **Measurement of HIV RNA in patients infected by subtype C by assays optimized for subtype B results in an underestimation of the viral load.** *J Med Virol* 2004; **73**:167–171.
49. Mascola JR, Louwagie J, McCutchan FE, Fischer CL, Hegerich PA, Wagner KF, *et al.* **Two antigenically distinct subtypes of human immunodeficiency virus type 1: viral genotype predicts neutralization serotype.** *J Infect Dis* 1994; **16**:948–954.
50. Thomson MM, Pérez Álvarez L, Nájera R. **Molecular epidemiology of HIV-1 genetic forms and its significance for vaccine development and therapy.** *Lancet Infect Dis* 2002; **2**:461–471.