Viruses Previously Identified in Brazil as Belonging to HIV-1 CRF72_BF1 Represent Two Closely Related Circulating Recombinant Forms, One of Which, Designated CRF122_BF1, Is Also Circulating in Spain

Javier E. Cañada-García1, Elena Delgado1, Horacio Gil1, Sonia Benito1, Mónica Sánchez1, Antonio Ocampo2, Jorge Julio Cabrera3,4, Celia Miralles5, Elena García-Bodas1, Ana Marino6, Patricia Ordóñez6, María José Gude7, Carmen Ezpeleta8 and Michael M. Thomson1*

1 HIV Biology and Variability Unit, Centro Nacional de Microbiología, Instituto de Salud Carlos III, Majadahonda, Spain, 2 Department of Internal Medicine, Complejo Hospitalario Universitario de Vigo, Vigo, Spain, 3 Department of Microbiology, Complejo Hospitalario Universitario de Vigo, Vigo, Spain, 4 Microbiology and Infectology Research Group, Galicia Sur Health Research Institute (IS Galicia Sur), SERGAS-UVigo, Vigo, Spain, 5 Infectious Diseases Unit, Complejo Hospitalario Universitario de Ferrol, Ferrol, Spain, 6 Department of Microbiology, Complejo Hospitalario Universitario de Ferrol, Ferrol, Spain, 7 Department of Microbiology, Hospital Universitario Lucus Augusti, Lugo, Spain, 8 Department of Clinical Microbiology, Complejo Hospitalario de Navarra, Pamplona, Spain

Circulating recombinant forms (CRFs) are important components of the HIV-1 pandemic. Those derived from recombination between subtype B and subsubtype F1, with 18 reported, most of them of South American origin, are among the most diverse. In this study, we identified a HIV-1 BF1 recombinant cluster that is expanding in Spain, transmitted mainly via heterosexual contact, which, analyzed in near full-length genomes in four viruses, exhibited a coincident BF1 mosaic structure, with 12 breakpoints, that fully coincided with that of two viruses (10BR_MG003 and 10BR_MG005) from Brazil, previously classified as CRF72_BF1. The three remaining Brazilian viruses (10BR_MG002, 10BR_MG004, and 10BR_MG008) previously identified as CRF72_BF1 exhibited mosaic structures highly similar, but not identical, to that of the Spanish viruses and to 10BR_MG003 and 10BR_MG005, with discrepant subtypes in two short genome segments, located in pol and gp120env. Based on these results, we propose that the five viruses from Brazil previously identified as CRF72_BF1 actually belong to two closely related CRFs, one comprising 10BR_MG002, 10BR_MG004, and 10BR_MG008, and the other, designated CRF122_BF1, comprising 10BR_MG003, 10BR_MG005, and the viruses of the identified Spanish cluster. Three other BF1 recombinant genomes, two from Brazil and one from Italy, previously identified as unique recombinant forms, were classified as CRF72_BF1. CRF122_BF1, but not CRF72_BF1, was associated with protease
INTRODUCTION

HIV-1 is characterized by high genetic diversity and rapid evolution, derived from elevated mutation and recombination rates. Through these mechanisms, the HIV-1 group M, the causative agent of the AIDS pandemic, has evolved into numerous circulating genetic forms, known as subtypes, of which 10 have been identified (A–D, F–H, J–L), subsubtypes (A1–A6, F1, and F2), and circulating recombinant forms (CRFs), 118 of which are currently listed in the Los Alamos HIV Sequence Database (Los Alamos National Laboratory, 2021). In addition, geographic variants and clusters, some representing substantial proportions of viruses in certain areas, have been identified through phylogenetic analyses within subtypes, subsubtypes, and CRFs (Thomson and Nájera, 2005; Delgado et al., 2015, 2019). Genetic characterization of HIV-1 variants is of public health relevance, as it allows tracking their geographic spread and estimating their population growth and the efficacy of preventive interventions (Paraskevis et al., 2016; Rífe et al., 2017; Vasilyeva et al., 2020). It has also biological and clinical relevance, as different biological properties have been associated with some HIV-1 variants (Kiwanuka et al., 2008; Pérez-Álvarez et al., 2014; Kouri et al., 2015; Venner et al., 2016; Cid-Silva et al., 2018; Song et al., 2019).

The number of CRFs is increasing incessantly, due to both the continuous generation of recombinant forms where diverse HIV-1 variants meet in the same population (Nájera et al., 2002), some of which become circulating through introduction into transmission networks, and the identification of old previously undocumented CRFs. The proportion of CRFs in the HIV-1 pandemic has increased over time, representing around 17% of infections in 2010–2015 (Hemelaar et al., 2020). Among CRFs, those derived from subtype B and subtype F1 are among the most numerous, 18 of which have been reported in the literature, most of them originated in South America. The most widely circulating CRF from South America is CRF12_BF, which circulates at high prevalences in Argentina and Uruguay, where unique recombinant forms (URFs) related to CRF12_BF are frequently found (Thomson et al., 2000, 2002; Carr et al., 2001). Four other CRF_BFs (numbers 17, 38, 44, and 89) related to CRF12_BF, as evidenced by shared breakpoints and phylogenetic clustering, were subsequently identified in different South American countries (Ruchansky et al., 2009; Delgado et al., 2010, 2021; Aulicino et al., 2012). Due to their common ancestry and similar structures, these five CRFs and related URFs have been proposed to constitute a “family” of recombinant viruses (Thomson and Nájera, 2005; Zhang et al., 2010; Delgado et al., 2021). By contrast, Brazilian CRF_BFs (De Sá Filho et al., 2006; Guimarães et al., 2008; Sanabani et al., 2010; Pessôa et al., 2014a,b, 2016; Reis et al., 2017, 2019) and CRF66_BF (the latter found mainly in Paraguay and Paraguayans living in Spain) (Bacqué et al., 2021) are unrelated to CRF12_BF. Similarly to the viruses of the CRF12_BF family, close relations have been reported between some Brazilian CRF_BFs: CRF28_BF and CRF29_BF (De Sá Filho et al., 2006) and CRF70_BF and CRF71_BF (Pessôa et al., 2014a).

In this study, we report the spread of a BF1 cluster in Spain whose viruses exhibit a mosaic structure identical to two Brazilian viruses previously identified as CRF72_BF1 (Pessôa et al., 2014b, 2016), which would represent a new CRF, with the three other viruses classified as CRF72_BF1 showing highly similar, but not identical, structures. We propose that viruses previously identified as CRF72_BF1 actually belong to two closely related CRFs that constitute a CRF family.

MATERIALS AND METHODS

Samples

Plasma and whole blood samples from HIV-1-infected individuals were collected in Spain for antiretroviral drug resistance tests and for a molecular epidemiological study. The study was approved by the Committee of Research Ethics of Instituto de Salud Carlos III, Majadahonda, Madrid, Spain. It did not require written informed consent by the study participants, as it used samples and data collected as part of routine clinical practice, and patients’ data were anonymized without retaining data allowing individual identification.

PCR Amplification and Sequencing

An ~1.4-kb pol fragment in protease–reverse transcriptase (Pr–RT) was amplified from plasma-extracted RNA or from whole blood-extracted DNA by (RT-)PCR followed by nested PCR, as described previously (Delgado et al., 2015), and sequenced with the Sanger method using a capillary automated sequencer. Near full-length genome (NFLG) sequences were obtained for selected samples by RT-PCR/nested PCR amplification from plasma RNA in five overlapping segments and sequenced by the Sanger method, as described (Delgado et al., 2002; Sierra et al., 2005; Cañada et al., 2021). Newly derived sequences are deposited in GenBank under accessions OL982311–OL982317 and OL982320–OL982323.

Phylogenetic Sequence Analyses

Sequences were aligned with MAFFT v7 (Katoh and Standley, 2013). Initial phylogenetic trees with all Pr–RT sequences obtained by us were constructed via approximate maximum likelihood with FastTree v2.1.10 (Price et al., 2010), using the general time-reversible evolutionary model with CAT approximation to account for among-site rate heterogeneity, with
The common ancestor (MRCA) of the newly defined CRF were
The time and the most probable location of the most recent
Clade Ancestors

Resistance Database (Rhee et al., 2003; Shafer, 2006).
HIVdb program of the Stanford University’s HIV Drug
Antiretroviral (ARV) drug resistance was analyzed with the
Determination
Antiretroviral Drug Resistance

org/phylo/treebase/phylows/study/TB2:S29595.
have been deposited in TreeBase, with accession URL http://purl.
genome segments.
node support of the clades of the ingroup, particularly in short
could be caused by an outgroup whose distance to the ingroup
prevent the long-branch attraction artifact (Bergsten, 2005) that
to the approach used in other studies (Travers et al., 2004;
use of a reconstructed ancestral sequence as outgroup is similar
and F subtypes (Zhu et al., 1998), obtained with IQ-Tree. The
outgroup, considering the phylogenetic relationship between B
and the Brazilian F1 strain), were further phylogenetically analyzed
via ML and Bayesian inference. These analyses were performed
with IQ-Tree; PhyML (Guindon et al., 2010), using the best-
fit evolutionary model selected by the SMS program (Lefort
et al., 2017) and node support assessment with the approximate
likelihood ratio test, SH–like procedure; and MrBayes v3.2
(Ronquist et al., 2012), using the GTR + G + I substitution
model, with two simultaneous independent runs and eight
chains 2–5 million generations long, ensuring that both runs
reached convergence, as determined by an average standard
deviation of split frequencies <0.01, discarding the first 50%
of the trees in the posterior distribution as burn-in. For these
analyses, we used a reconstructed BF1 ancestral sequence as
outgroup, considering the phylogenetic relationship between B
and F subtypes (Zhu et al., 1998), obtained with IQ-Tree. The
use of a reconstructed ancestral sequence as outgroup is similar
to the approach used in other studies (Travers et al., 2004;
Thomson and Fernández-García, 2011; Seager et al., 2014) to
prevent the long-branch attraction artifact (Bergsten, 2005) that
could be caused by an outgroup whose distance to the ingroup
is relatively long compared with the within-ingroup distances.
This artifact can result in collapse or a substantial decrease in
node support of the clades of the ingroup, particularly in short
genome segments.

Phylogenetic trees and alignments used for their construction
have been deposited in TreeBase, with accession URL http://purl.
org/phylo/treebase/phylows/study/TB2:S29595.

Antiretroviral Drug Resistance
Determination
Antiretroviral (ARV) drug resistance was analyzed with the
HIVdb program of the Stanford University’s HIV Drug
Resistance Database (Rhee et al., 2003; Shafer, 2006).

Temporal and Geographic Estimations of
Clade Ancestors
The time and the most probable location of the most recent
common ancestor (MRCA) of the newly defined CRF were
estimated using Pr–RT sequences with the Bayesian Markov
chain Monte Carlo (MCMC) coalescent method implemented in
BEAST v1.10.4 (Suchard et al., 2018). Before the BEAST analysis,
the existence of temporal signal in the dataset was assessed with
TempEst v1.5.3 (Rambaut et al., 2016). The BEAST analysis was
performed using the SRD06 codon-based evolutionary model
(where the third codon position is in a partition different from
the first and second positions) (Shapiro et al., 2006). We also
specified an uncorrelated lognormal relaxed clock and a Bayesian SkyGrid
coalicient tree prior (Gill et al., 2013). The MCMCs were run for
20 million generations. The runs were performed in duplicate,
and the posterior tree files were combined with LogCombiner
v1.10.4. Proper mixing of the chains was assessed with Tracer
v1.6, ensuring that effective sample size values of all parameters
were >200. The posterior distribution of trees was summarized
in a maximum cklade credibility (MCC) tree with TreeAnnotator
v1.10.4, after discarding 10% of the MCMC chain as burn-
in. MCC trees were visualized with FigTree v1.4.2 (Rambaut).1
Parameter uncertainty was summarized in 95% highest posterior
density (HPD) intervals.

RESULTS
Identification of a HIV-1 Cluster of F1
Subsubtype in Protease–Reverse
Transcriptase Propagating in Spain
In a molecular epidemiological study in Spain, based on Pr–
RT sequences, we detected frequent grouping in clusters, several
of which were of F1 subsubtype in Pr–RT (Thomson et al.,
2012; Delgado et al., 2015, 2019; Gil et al., 2022). One of
them, designated F1_2, which is the focus of the present study,
comprised 14 individuals, 13 of them from the region of Galicia,
northwest Spain (Table 1). Years of HIV-1 diagnoses were
between 2007 and 2019, and transmission was predominantly
heterosexual (n = 8), but there were three men who have sex
with men (two others had non-specified sexual transmission,
and no data on transmission route were available for another
individual) (Table 1). Most individuals were Spanish, but three
were Brazilian, one was Swiss, and one was Ukrainian. To
determine whether other sequences from databases belonged to
this cluster, we performed BLAST searches in the HIV Sequence
Database (Los Alamos National Laboratory, 2021), incorporating
the most similar sequences in the phylogenetic analyses. This
allowed identifying three additional sequences that belonged to
the F1_2 cluster, from Brazil, Portugal, and Germany (Figure 1).
All but two of the viruses collected in Spain and the virus from
Germany branched in a subcluster. Viruses from the F1_2 cluster
were most closely related to viruses of the Brazilian F1 strain and
to Brazilian CRF_BFs with Pr–RT derived from it.

Analyses of Near Full-Length Genome
Sequences
In order to determine whether the F1_2 cluster was of uniform
subtype or recombinant, we obtained NFLG sequences from

1http://tree.bio.ed.ac.uk/software/figtree/
three individuals from two cities through amplification from plasma RNA. A fourth NFLG sequence had been obtained previously from the virus culture supernatant (Sanchez et al., 2014). Preliminary analyses of the NFLG with Recombination Identification Program\(^2\) indicated that the genomes were BF1 recombinant. To determine whether they belonged to a known CRF, we constructed a phylogenetic tree in which genomes of all CRF_BFs were included. The tree showed that viruses of the F1_2 cluster grouped in a strongly supported clade with viruses classified as CRF72_BF1, with two of them, 10BR_MG003 and 10BR_MG005, being the most closely related to the viruses of the F1_2 cluster, and the other three, 10BR_MG002, 10BR_MG004, and 10BR_MG008, branching in a sister clade (Figure 2).

Bootscan analyses of NFLG sequences showed that the viruses of the F1_2 cluster were BF1 recombinant, exhibiting mosaic structures fully coincident with those of \(10\)BR_MG003 and 10BR_MG005, being the most closely related to the viruses of the F1_2 cluster, and the other three, 10BR_MG002, 10BR_MG004, and 10BR_MG008, branching in a sister clade (Figure 2).

Identification of CRF122_BF1

CRF122_BF1, which is closely related to, but different from, the three other viruses previously classified as CRF72_BF1, 10BR_MG002, 10BR_MG004, and 10BR_MG008, whose original CRF designation is maintained. The mosaic structures of both CRFs, as inferred from bootscan analyses, ML and Bayesian phylogenetic trees of partial sequences, and examination of intersubtype transitions of subtype-discriminating nucleotides, are shown in Figure 5.

Three additional BF1 recombinant NFLGs, originally identified as unique recombinant forms, two from Brazil, 99UFRJ-2 (Thomson et al., 2004) and BREPM1029 (Sa-Filho et al., 2007), and one from Italy, \(IT\)_BF PRIN_454 (Bruselles et al., 2009), in their published analyses, exhibited mosaic structures similar to CRF72_BF1 and CRF122_BF1. To determine whether they belonged to one of these CRFs, we constructed a phylogenetic tree with NFLG sequences including the three mentioned genomes, which showed that all of them grouped with CRF72_BF1 viruses (Supplementary Figure 2). Bootscan analyses showed mosaic structures of 99UFRJ-2 and \(IT\)_BF PRIN_454 coincident with that of CRF72_BF1; however, the bootscan plot of BREPM1029 failed to show clustering with the subtype B references in the protease–RT junction (HXB2 positions 2429–2618) (Supplementary Figure 3). Examination of subtype-discriminating nucleotides suggested that the 2429–2618 segment was of subtype B, as in CRF72_BF1, in 99UFRJ-2 and \(IT\)_BF PRIN_454, which was confirmed by phylogenetic analyses of this fragment (Supplementary Figure 4a).

However, in BREPM1029, the subtype B fragment in the protease–RT junction appeared to be slightly shorter, located between HXB2 positions 2479 and 2618, which was confirmed by phylogenetic trees (Supplementary Figure 4b). Phylogenetic analyses also showed that in all three genomes the 6432–6519 segment in gp120 was of subtype B, as in CRF72_BF1 and unlike CRF122_BF1 (Supplementary Figure 4c). These results allowed to confidently classify 99UFRJ-2 and \(IT\)_BF PRIN_454 as CRF72_BF1 viruses. As to BREPM1029, given its strong

---

\(^2\)https://www.hiv.lanl.gov/content/sequence/RIP/RIP.html

---

**TABLE 1 | Epidemiological data of patients and GenBank accessions of sequences.**

| Sample ID | City of sample collection | Region of sample collection | Country of origin | Year of HIV diagnosis | Year of sample collection | Gender* | Transmission route* | PR–RT GenBank accession | NFLG GenBank accession |
|-----------|---------------------------|----------------------------|-------------------|-----------------------|--------------------------|---------|-------------------|------------------------|-------------------------|
| X2592     | Vigo                      | Galicia                    | Spain             | 2008                  | 2008                     | F       | HT                | GU326146               | –                       |
| X2632     | Ferrol                    | Galicia                    | Spain             | 2009                  | 2009                     | M       | MSM               | GU326158               | KC113006                |
| X2657     | Vigo                      | Galicia                    | Spain             | 2009                  | 2009                     | F       | HT                | GU326163               | –                       |
| GA076319  | Vigo                      | Galicia                    | Ukraine           | 2010                  | 2010                     | F       | HT                | OL982311               | –                       |
| GA099170  | Vigo                      | Galicia                    | Switzerland       | 2011                  | 2011                     | F       | HT                | –                      | OL982312                |
| GA330265  | Vigo                      | Galicia                    | Brazil            | 2007                  | 2007                     | Trans   | Sexual            | OL982313               | –                       |
| GA486085  | Ferrol                    | Galicia                    | Spain             | 2018                  | 2018                     | F       | n.a.              | –                      | OL982314                |
| GA501952  | Vigo                      | Galicia                    | Spain             | 2014                  | 2014                     | F       | HT                | OL982315               | –                       |
| GA513250  | Vigo                      | Galicia                    | Spain             | 2012                  | 2012                     | M       | Sexual            | OL982316               | –                       |
| GA522821  | Lugo                      | Galicia                    | Brazil            | 2019                  | 2019                     | M       | MSM               | OL982317               | –                       |
| GA817166  | Vigo                      | Galicia                    | Spain             | 2008                  | 2008                     | F       | HT                | OL982320               | –                       |
| GA874035  | Vigo                      | Galicia                    | Spain             | 2012                  | 2012                     | M       | HT                | –                      | OL982321                |
| GA930064  | Vigo                      | Galicia                    | Spain             | 2012                  | 2012                     | F       | HT                | OL982322               | –                       |
| NAS84314  | Pamplona                  | Navarre                    | Brazil            | 2017                  | 2017                     | M       | MSM               | OL982323               | –                       |

* n.a., datum not available; Trans, transgender; HT, heterosexual; MSM, man who has sex with men; sexual, unspecified sexual transmission.

---
phylogenetic clustering with CRF72_BF1 references and its minimal difference in mosaic structure with CRF72_BF1, with a breakpoint displaced only around 50 nt relative to this CRF, it seems reasonable to also classify it as CRF72_BF1, although we cannot definitively discern whether its breakpoint displacement is due to a different recombination event or to mutations occurring near the CRF72_BF1 breakpoint.

Differences in Amino Acid Residues

We analyzed amino acid residues in viral proteins differing between CRF72_BF1 and CRF122_BF1 viruses and conserved within each CRF. We found 10 such amino acid residues (Table 2). One of them is in position 89 of protease, where CRF72_BF1 has leucine, which is the subtype B consensus, while CRF122_BF1 has methionine, which is the F1 subsubtype consensus. Protease L89M substitution has been reported to contribute, together with other protease mutations, to resistance to some protease inhibitor drugs (Calazans et al., 2005; Marcelin et al., 2008; Wensing et al., 2019).

Antiretroviral Drug Resistance Mutations

Primary ARV drug resistance mutations were detected in two CRF122_BF1 viruses, both from Brazil, one (10BR_MG003, collected in 2010) with L90M protease inhibitor (PI) resistance...
FIGURE 2 | Maximum likelihood tree of NFLG sequences of viruses of the F1_2 cluster. References of published CRF_BFs and of HIV-1 subtypes are also included in the analysis. Names of sequences obtained by us are in blue. In reference sequences, the subtype or CRF is indicated before the virus name. Only ultrafast bootstrap values ≥ 90% are shown. The scale indicates substitutions/site. *10BR_MG003 and 10BR_MG005 were originally identified as CRF72_BF1 (Pessôa et al., 2014b, 2016), but analyses described in this study have reclassified them as CRF122_BF1.

mutation and the other (BR05SP503, collected in 2005) with D30N PI resistance mutation and M41I, D67N, M184V, and T215Y mutations of resistance to nucleoside RT inhibitors.

Temporal and Geographic Estimation of CRF122_BF1 Origin

To estimate the time and place of origin of CRF122_BF1, PrñRT sequences were analyzed with the Bayesian coalescent method implemented in BEAST 1.10.4. Prior to this analysis, we performed TempEst analyses to determine whether there was an adequate temporal signal in the dataset. We found that the temporal signal, assessed by the correlation between root-to-tip distance and time, increased by masking the positions of codons with drug resistance mutations in any of the sequences ($r^2 = 0.5265$; Supplementary Figure 5). Therefore, the BEAST analysis was performed with a sequence alignment where these codons had been removed. In this analysis, the substitution rate was estimated at $1.829 \times 10^{-3}$ subs/site/year (95% HPD, $1.118 \times 10^{-3}$–$2.542 \times 10^{-3}$ subs/site/year) and the time of the MRCA of CRF122_BF1 was estimated around 1987 (95% HPD, 1976–1998), with its most probable location being Brazil (location PP = 0.89) (Figure 6). The introduction of CRF122_BF1 in Spain (according to the MRCA of the Spanish cluster) was estimated in the Galician city of Vigo (location PP = 0.992) around 2002 (95% HPD, 1998–2005).

DISCUSSION

The results presented here indicate that the five Brazilian viruses previously classified as CRF72_BF1 actually belong to two closely related CRFs, one of which is circulating in Spain. Consequently, the CRF comprising two Brazilian viruses previously classified as CRF72_BF1 and the four Spanish viruses with coincident mosaic structures is given a new designation, CRF122_BF1, while the
Identification of CRF122_BF1

FIGURE 3 | Bootscan analyses of NFLG sequences of viruses of the F1_2 cluster compared to those of viruses previously classified as CRF72_BF1. Bootscan plots of all four F1_2 viruses are shown, together with those of 10BR MG003 and 10BR MG002. The bootscan plot of 10BR MG005 is almost identical to that of 10BR MG003, and those of 10BR MG004 and 10BR MG008 are almost identical to that of 10BR MG002, and are shown in Supplementary Figure 1. The horizontal axis represents the position in the HXB2 genome of the midpoint of a 250-nt window moving in 20-nt increments, and the vertical axis represents bootstrap values supporting clustering with subtype reference sequences. The vertical dashed lines indicate BF1 breakpoints differing between the CRF72_BF1 viruses 10BR MG002, 10BR MG004, and 10BR MG008, on the one hand, and viruses of the F1_2 cluster and 10BR MG003 and 10BR MG005 (newly identified as CRF122_BF1), on the other. The bar on the top indicates the segments that were further analyzed with ML and Bayesian trees (Figure 4). Two genome segments that appear to group with different subtypes in 10BR MG002, 10BR MG004, and 10BR MG008 relative to 10BR MG003, 10BR MG005, and the F1_2 viruses are signaled with arrows above the bootscan plot of 10BR MG002 and were also analyzed via ML and Bayesian inference (Figure 4).

three other Brazilian viruses previously classified as CRF72_BF1 keep their original designation. Three other BF1 viruses analyzed in NFLGs originally classified as URFs, two from Brazil and one from Italy, were also classified on the basis of phylogenetic and bootscan analyses as CRF72_BF1. The close relationship between CRF122_BF1 and CRF72_BF1 is one more example of closely related CRFs, with precedents in South America. Other examples are CRF12_BF, CRF17_BF, and CRF89_BF (and more distantly related, CRF38_BF and CRF44_BF) (Delgado et al., 2021); CRF28 and CRF29_BF (De Sá Filho et al., 2006); and CRF70_BF and CRF71_BF (Pessôa et al., 2014a).

Failure to realize that the five viruses previously identified as CRF72_BF1 represent two different CRFs may derive from the short segments in which both CRFs differ in subtypes. These
FIGURE 4 | Phylogenetic trees of interbreakpoint genome segments of F1_2 viruses. Breakpoints were defined according to the bootscan analyses and to midpoints of transitions between subtype-discriminating nucleotides, here defined as those where the 75% consensus of subtype B and of the Brazilian variant of subtype F1 differ. HXB2 positions delimiting the analyzed segments and their numbers as indicated in Figure 3 are indicated on top of the trees. Sequence names of F1_2 viruses are in blue. Names of subtype reference sequences are preceded by the corresponding subtype name. Sequences of viruses previously classified as CRF72_BF1 were also included, with those reclassified in the present study as CRF122_BF1 (10BR_MG003 and 10BR_MG005) labeled with the new CRF designation. The trees are rooted with a reconstructed BF1 ancestor. Node supports for B and F1 clades are indicated, in this order, as ultrafast bootstrap value/aLRT SH–like support/posterior probability, which were obtained with IQ-Tree, PhyML, and MrBayes programs, respectively. For the other nodes, only ultrafast bootstrap values $\geq 80\%$ are indicated. Trees #14 and #15 (segments 2429–2618 and 6432–6519, respectively) correspond to the segments indicated with arrows in Figure 3, where F1_2 viruses and 10BR_MG003 and 10BR_MG005, on the one hand, and 10BR_MG002, 10BR_MG004, and 10BR_MG008, on the other, appeared to differ in subtype affiliations in the bootscan analyses. The scales indicate substitutions/site.

differences may be missed if bootscan analyses are performed using window widths much greater than the length of the recombinant fragment. We have also noticed that jPHMM (Schultz et al., 2009), used in a previous study to analyze CRF72_BF1 genomes (Pessôa et al., 2016), often fails to detect short recombinant segments (Delgado et al., 2021).

Given the close relationship and partial coincidence in mosaic structures of CRF72_BF1 and CRF122_BF1, we propose that they are members of a CRF family, similar to the CRF family of BF1 recombinant viruses from South America comprising CRFs numbers 12, 17, 38, 44, and 89 (Delgado et al., 2021). The grouping of some closely related HIV-1 recombinants derived from a common recombinant ancestor in families was proposed by Thomson and Nájera (2005) and Zhang et al. (2010). The proposed families of CRFs with close phylogenetic relations, shared parental strains, and partially coincident breakpoints are indicated in the phylogenetic tree shown in Supplementary Figure 6.

It is interesting to note that, in the Pr–RT tree, viruses from the Spanish CRF122_BF1 ("F1_2") cluster fail to group with the Brazilian CRF122_BF1 viruses. A similar phenomenon is observed with CRF66, CRF70, and CRF71_BF1 references, that fail to group in distinct clades with other references of the same CRF. This may be due to the relatively short length and high sequence conservation of this segment, together with the fact that 35 references from the Brazilian F1 strain or from CRF_BFs derived from it are included in the tree. This shows that exclusive phylogenetic analysis of Pr–RT may not be sufficient to phylogenetically classify an F1 sequence of the Brazilian strain as belonging or not to a given CRF_BF.

The estimated origin of CRF122_BF1 around 1987 is consistent with the estimated origin of the Brazilian F1 strain (around 1977) (Bello et al., 2007) and similar to those of other South American CRF_BFs (CRF12, CRF28/29, CRF38, CRF89, and CRF90) reported in the literature (Bello et al., 2010; Ristic et al., 2011; Reis et al., 2017; Delgado et al., 2021) but younger than some other estimates for CRF12_BF in the 1970s (Dilernia et al., 2011; Delgado et al., 2021) and older than the estimates for CRF99_BF, around 1993 (Reis et al., 2017).

The correct classification of HIV-1 genetic forms is important, since even relatively minor genetic differences in viral genomes may result in important biological differences. Examples in HIV-1 are frequent CXCR4 coreceptor usage in CRF14_BG, which is associated with only four amino acid residues in the Env V3 loop (Pérez-Álvarez et al., 2014), all or most of which are absent in viruses of the closely related CRF73_BG
FIGURE 5 | Mosaic structures of CRF72_BF1 and CRF122_BF1. Breakpoint positions are numbered as in the HXB2 genome. Vertical dashed lines indicate the BF1 breakpoint positions in the 5′ segment of gp120env differing between CRF122_BF1 and CRF72_BF1. The drawing was made using the Recombinant HIV-1 Drawing Tool (https://www.hiv.lanl.gov/content/sequence/DRAW_CRF/recom_mapper.html).

TABLE 2 | Differences in amino acid residues between CRF72_BF1 and CRF122_BF1.

| p17 gag | PR | RT | IN | Vif | Tat | Rev | Vpu | gp120 env |
|--------|-----|----|----|-----|-----|-----|-----|---------|
| 61     | 61  | 89 | 399| 84  | 151 | 68  | 22  | 80      |
| B      | L   | O  | L  | E   | I   | A   | S   | L       |
| CRF72_BF1 |     |   |    |     |     |     |     |         |
| 10BR_MG002 | I   | Q  | L  | E   | I   | A   | S   | L       |
| 10BR_MG004 | I   | Q  | L  | E   | L   | A   | S   | L       |
| 10BR_MG008 | I   | Q  | L  | D   | A   | S   | L   | D       |
| 99JFRJ-2  | I/M | Q  | L  | E   | L   | A   | P   | L       |
| IT_PRIN_454 | I   | Q  | L  | E   | L   | A   | S   | L       |
| 99JFRJ-2  | I/M | Q  | L  | E   | L   | A   | S   | L       |
| CRF122_BF1 |     |   |    |     |     |     |     |         |
| 10BR_MG003 | L   | N  | M  | D   | I   | T   | D   | I       |
| 10BR_MG005 | L   | N  | M  | D   | I   | T   | D   | I       |
| X2632_4  | L   | N  | M  | D   | L   | T   | D   | I       |
| GA099170 | L   | N  | M  | D   | I   | T   | D   | I       |
| GA486085 | L/I | N  | M  | D   | I/M | T   | D   | I       |
| GA874035 | L   | N  | M  | D   | I   | T   | D   | I       |

*Only amino acid residues conserved in at least five of six CRF72_BF1 viruses and in both Brazilian and at least three of four Spanish CRF122_BF1 viruses are included in the table.

(Fernández-García et al., 2016), which has a very similar mosaic structure, and differences in pathogenic potential or therapeutic response associated with clusters within HIV-1 CRF01_AE (Song et al., 2019) and F1 subtype (Cid-Silva et al., 2018). Here, we show that CRF122_BF1, but not CRF72_BF1, has the protease L89M substitution, that has been reported to contribute, together with other protease mutations, to resistance to tipranavir/ritonavir (Marcelin et al., 2008) and, within an F1 subtype background, to other protease inhibitor drugs (Calazans et al., 2005).

CRF122_BF1 represents one more example of a CRF of South American ancestry first identified in Western Europe. Others are CRF42_BF (Struck et al., 2015), CRF47_BF (Fernández-García et al., 2010), CRF60_BC (Simonetti et al., 2014), CRF66_BF (Bacqué et al., 2021), and CRF89_BF (Delgado et al., 2021). This may derive from the increasing migratory
flows from South America to Europe and from the relatively low number of HIV-1 sequences available in some South American countries (Bacqué et al., 2021). Therefore, HIV-1 molecular epidemiological studies in Europe may contribute to a better knowledge of the HIV-1 epidemics in South America.

In summary, we show that viruses of a BF1 recombinant cluster of Brazilian ancestry circulating in Spain exhibit a mosaic structure that is fully coincident with that of two Brazilian viruses previously classified as CRF72_BF1 and is highly similar, but not identical, to that of three other Brazilian viruses also classified as CRF72_BF1. Therefore, we propose a new CRF designation, CRF122_BF1, for the viruses of the Spanish cluster and the two Brazilian viruses with coincident structures, which together with CRF72_BF1 would constitute a CRF family. The accurate genetic characterization of HIV-1 variants is important to determine their associated biological features and to track their epidemic spread.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. Newly derived sequences are deposited in GenBank under accessions OL982311–OL982317 and OL982320–OL982323. Phylogenetic trees and alignments used for their construction have been deposited in TreeBase, with accession URL http://purl.org/phylo/treebase/phylows/study/TB2:S29595.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Committee of Research Ethics of Instituto de Salud Carlos III, Majadahonda, Madrid, Spain. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

MT and ED conceived the study and supervised the experimental work. JC-G, ED, and MT processed the sequences and performed phylogenetic and phylodynamic analyses. HG performed data curation. JC-G, SB, MS, and EG-B performed experimental work. AO, JC, CM, AM, PO, MG, and CE obtained the samples and epidemiological data from patients. MT, JC-G, ED, and HG wrote the manuscript. All authors read and approved the manuscript.

FUNDING

This work was funded through Acción Estratégica en Salud Intramural (AESI), Instituto de Salud Carlos III, projects “Estudios Sobre Vigilancia Epidemiológica Molecular del VIH-1 en España”, PI16CIII/00033, and “Epidemiología Molecular del VIH-1 en España y su Utilidad Para Investigaciones
ACKNOWLEDGMENTS

We would thank José Antonio Taboada from Consellería de Sanidade, Xunta de Galicia, for his support of this study, and the personnel at the Genomic Unit, Instituto de Salud Carlos III, for technical assistance in sequencing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.863084/full#supplementary-material

REFERENCES

Aulicino, P. C., Gómez-Carrillo, M., Bello, G., Rocco, C., Mangano, A., Carr, J., et al. (2012). Characterization of full-length HIV-1 CRF12_BF genomes and comparison to the prototype CRF12_BF strains. Infect. Genet. Evol. 12, 443–447. doi: 10.1016/j.meegid.2012.01.003

Bacqué, J., Delgado, E., Benito, S., Moreno-Lorenzo, M., Montero, V., Gil, H., et al. (2021). Identification of CRF66_BF, a new HIV-1 circulating recombinant form of South American origin. Front. Microbiol. 12:774386. doi: 10.3389/fmicb.2021.774386

Bello, G., Aulicino, P. C., Ruchansky, D., Guimaraes, M. L., López-Galán, C., Casado, C., et al. (2010). PhyloDynamics of HIV-1 circulating recombinant forms 12_BF and 38_BF in Argentina and Uruguay. Retrovirology 7:22. doi: 10.1186/1742-6900-7-22

Bello, G., Eyer-Silva, W. A., Couto-Fernandez, J. C., Guimaraes, M. L., Chequer-Fernandez, S. L., Teixeira, S. L., et al. (2007). Demographic history of HIV-1 subtypes B and F in Brazil. Infect. Genet. Evol. 7, 263–270. doi: 10.1016/j.meegid.2006.11.002

Bergsten, J. (2005). A review of long-branch attraction. Cladistics 21, 163–193. doi: 10.1111/j.1095-839X.2005.00593.x

Bruselles, A., Rozer, G., Bartolíni, B., Prosperi, M., Del Nonno, F., Narciso, P., et al. (2009). Use of massive parallel pyrosequencing for near full-length characterization of a unique HIV Type 1 recombinant associated with a fatal primary infection. AIDS Res. Hum. Retroviruses 25, 937–942. doi: 10.1089/aid.2009.0083

Calazans, A., Brindeiro, R., Brindeiro, P., Verli, H., Arruda, M. B., Gonzalez, L. M., et al. (2005). Low accumulation of L90M in protease from subtype F HIV-1 with resistance to protease inhibitors is caused by the L89M polymorphism. J. Infect. Dis. 191, 1961–1970. doi: 10.1086/430002

Castro-Relaño, E., Benito, S., García-Bodas, E., et al. (2021). Identification of a new HIV-1 BC intersubtype circulating recombinant form (CRF08_BC) in Spain. Viruses 13:93. doi: 10.3390/v13010093

Carr, J. K., Avila, M., Gomez-Carrillo, M., Salomon, H., Hierholzer, J., Watanaveeradej, V., et al. (2010). Diverse BF recombinants have spread widely since the introduction of HIV-1 into South America. AIDS 15, F41–F47. doi: 10.1097/00126334-200204150-00016

Delgado, E., Benito, S., Moreno-Lorenzo, M., Fernández-García, A., Sánchez, M., et al. (2019). Identification of two new HIV-1 circulating BF intersubtype recombinant forms widely spread in South America. Sci. Rep. 11:11442. doi: 10.1038/s41598-021-9003-x

Delgado, E., Rios, M., Fernández, J., Pérez-Alvarez, L., Nájera, R., and Thomson, M. M. (2010). Identification of a new HIV type 1 BF intersubtype circulating recombinant form (CRF44_BF) in Chile. AIDS Res. Hum. Retroviruses 26, 821–826. doi: 10.1089/aid.2010.0006

Delgado, E., Thomson, M. M., Villahermosa, M. L., Sierra, M., Ocampo, A., Miralles, C., et al. (2002). Identification of a newly characterized HIV-1 BG intersubtype circulating recombinant form (CRF73_BG), partially related to CRF14_BG, which is circulating in Portugal and Spain. PLoS One 7:e0148549. doi: 10.1371/journal.pone.0148549

Díaz, M., et al. (2018). Identification of a new HIV type 1 circulating BF intersubtype recombinant form (CRF47_BG) in Spain. AIDS Res. Hum. Retroviruses 24, 827–832. doi: 10.1089/aid.2018.0311

Gil, H., Delgado, E., Benito, S., Georgalis, L., Montero, V., Sánchez, M., et al. (2022). Transmission clusters, predominantly associated with men who have sex with men, play a main role in the propagation of HIV-1 in northern Spain (2013–2018). Front. Microbiol. 13:782609. doi: 10.3389/fmicb.2022.782609

Hemelaar, J., Elangovan, R., Yun, J., Dickson-Tetteh, L., Kirtley, S., Gouws-Williams, E., et al. (2020). Global and regional epidemiology of HIV-1 recombinants in 1990-2015: a systematic review and global survey. Lancet HIV 7, e772–e781. doi: 10.1016/S2352-3964(20)30201-X

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., Vinh, L. S. (2008). Phylogenetic analysis of HIV-1 multiple loci. Mol. Biol. Evol. 30, 713–724. doi: 10.1093/molbev/msx265

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermiin, L. S., Chernomor, O., von Haeseler, A., Minh, B. Q., and Vinh, L. S. (2018). UFBoot2: Improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522. doi: 10.1093/molbev/msx281

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermiin, L. S., Chernomor, O., von Haeseler, A., Minh, B. Q., and Vinh, L. S. (2018). UFBoot2: Improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522. doi: 10.1093/molbev/msx281

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermiin, L. S., Chernomor, O., von Haeseler, A., Minh, B. Q., and Vinh, L. S. (2018). UFBoot2: Improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522. doi: 10.1093/molbev/msx281
on disease progression in persons from Rakai, Uganda, with incident HIV-1 infection. *J. Infect. Dis.* 197, 707–713. doi: 10.1086/527416
Kouri, V., Khouri, R., Alemán, Y., Abrahantes, Y., Vercauteren, Y., Pineda-Peña, A. C., et al. (2015). CRF19_cpx is an evolutionary fit HIV-1 variant strongly associated with rapid progression to AIDS in Cuba. *ElhoMedicine* 2, 244–254. doi: 10.1016/j.elomed.2015.01.015
Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054
Lefort, V., Longueville, J. E., and Gascuel, O. (2017). SMS: smart model selection in PhyML. *Mol. Biol. Evol.* 34, 2422–2424. doi: 10.1093/molbev/msx149
Lole, K. S., Bollinger, R. C., Paranjape, R. S., Gadkari, D., Kulkarni, S. S., Novak, N. G., et al. (1999). Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J. Virol.* 73, 152–160. doi: 10.1128/JVI.73.1.152-160.1999
Los Alamos National Laboratory (2021). HIV Sequence Database. Available online at: https://www.hiv.lanl.gov/content/sequence/HIV/mainpage.html (accessed August 17, 2021).
Marcellin, A. G., Masquelier, B., Descamps, D., Irozpet, J., Charpentier, C., Allou, C., et al. (2008). Tipiranavir-ritonavir genotype score in protease inhibitor-experienced patients. *Antimicrob. Agents Chemother.* 52, 3237–3243. doi: 10.1128/AAC.00332–08
Nájera, R., Delgado, E., Pérez-Alvarez, L., and Thomson, M. M. (2002). Genetic recombination and its role in the development of the HIV-1 pandemic. *AIDS* 16, S3–S16. doi: 10.1097/00001883-200204000-00002
Paraskova, D., Nikolopoulos, K. G., Magiorkinis, G., Hodges-Mameletzis, I., and Hatzakis, A. (2016). The application of HIV molecular epidemiology to public health. *Infect. Genet. Evol.* 46, 159–168. doi: 10.1016/j.meegid.2016.06.021
Pérez-Alvarez, L., Delgado, E., Vega, Y., Montero, V., Cuevas, T., Fernández-Garcia, A., et al. (2014). Predominance of CXCR4 tropism in HIV-1 CRF14_BG strains from newly diagnosed infections. *J. Antimicrob. Chemother.* 31, 298–303. doi: 10.1093/nar/gkg100
Pessôa, R., Watanabe, J. T., Calabria, P., Felix, A. C., Loureiro, P., Sabino, E. C., et al. (2014a). Deep sequencing of HIV-1 near full-length proviral genomes identifies high rates of BF1 recombinants including two novel circulating recombinant forms (CRF) 70_BF1 and a disseminating 71_BF1 among blood donors in Pernambuco, Brazil. *PloS One* 9:e112674. doi: 10.1093/jac/dkt305.1371/journal.pone.0112674
Pessôa, R., Carneiro Proietti, A. B., Busch, M. P., and Sanabani, S. S. (2014b). Identification of a novel HIV-1 recombinating circulating form (CRF72_BF1) in deep sequencing data from blood donors in Southeastern Brazil. *Genome Announc.* 2:e00386–14. doi: 10.1128/genomeA.00386–14
Pessôa, R., Loureiro, P., Esther Lopes, M., Carneiro-Proietti, A. B., Sabino, E. C., Busch, M. P., et al. (2016). Ultra deep sequencing of HIV-1 near full-length and partial proviral genomes reveals high genetic diversity among Brazilian blood donors. *PloS One* 11:e0152499. doi: 10.1093/journal.pone.0152499
Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2—not approximately trees. *Virol. J.* 7:4. doi: 10.1186/1743-422X-7-74
Sanchez, A. M., DeMarco, C. T., Hora, B., Keinonen, S., Chen, Y., and Brinkley, C. (2014). Development of a contemporary globally diverse HIV viral panel by the EQAPOL program. *J. Immunol. Methods* 409, 117–130. doi: 10.1016/j.jim.2014.01.004
Schultz, A. K., Zhang, M., Bulla, I., Leitner, T., Korber, B., Morgenstern, B., et al. (2009). jPHMM: improving the reliability of recombination prediction in HIV-1. *Nucleic Acids Res.* 37, W647–W651. doi: 10.1093/nar/gkp371
Seager, I., Travers, S. A., Leeson, M. D., Crampin, A. C., French, N., Glynn, J. R., et al. (2014). Coreceptor usage, diversity, and divergence in drug-naïve and drug-exposed individuals from Malawi, infected with HIV-1 subtype C for more than 20 years. *AIDS Res. Hum. Retroviruses* 30, 975–983. doi: 10.1097/aid.2013.01.020
Shafer, R. W. (2006). Rationale and uses of a public HIV drug-resistance database. *J. Infect. Dis.* 194(Suppl. 1), S51–S58. doi: 10.1086/505356
Shapiro, B., Rambaut, A., and Drummond, A. J. (2006). Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Mol. Biol. Evol.* 23, 7–9. doi: 10.1093/molbev/msi021
Sierra, M., Thomson, M. M., Rios, M., Casado, G., Ojeda de Castro, R., Delgado, E., et al. (2005). The analysis of near full-length genome sequences of human immunodeficiency virus type 1 BF intersubtype recombinant viruses from Chile, Venenuela and Spain reveals their relationship to diverse lineages of recombinant viruses related to CRF12_BF. *Infect. Genet. Evol.* 5, 209–217. doi: 10.1016/j.meegid.2004.07.010
Simontelli, F. R., Lai, A., Monno, L., Binda, F., Brindicci, G., Punzi, G., et al. (2014). Identification of a new HIV-1 BC circulating recombinant form (CRF60_Bc) in Italian young men having sex with men. *Infect. Genet. Evol.* 23, 176–181. doi: 10.1016/j.meegid.2014.02.007
Song, H., Ou, W., Feng, Y., Zhang, J., Li, F., Hu, J., et al. (2019). Disparate impact on CD4 T cell count by two distinct HIV-1 phylogenetic clusters from the same clade. *Proc. Natl. Acad. Sci. U.S.A.* 116, 239–244. doi: 10.1073/pnas.1814714116
Suchard, M. A., Roman, F., De Landtsheer, S., Servais, J. Y., Lambert, C., Casas, M., et al. (2015). Near-full-length characterization and population dynamics of the human immunodeficiency virus type 1 circulating recombinant form 42 (CRF42_BF) in Luxembourg. *AIDS Res. Hum. Retroviruses* 31, 554–558. doi: 10.1097/AID.2014.0364
Thomson, M. M., and Fernández-Garcia, A. (2011). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Bioinformatics* 4:vey016. doi: 10.1093/bioinformatics/bey016
Thomson, M. M., and Fernández-Garcia, A. (2011). Phylogenetic structure in African HIV type 1 CRF populations and implications for the selective evolutionary forces driving HIV diversity. *Virology* 2, 30–38. doi: 10.1073/pnas.1103921107
Cañada-García et al. Identification of CRF122_BF1
Thomson, M. M., Fernández-García, A., Delgado, E., Vega, Y., Diez-Fuertes, F., Sánchez-Martínez, M., et al. (2012). Rapid expansion of a HIV-1 subtype F cluster of recent origin among men who have sex with men in Galicia, Spain. *J. Acquir. Immune Defic. Syndr.* 59, e49–e51. doi: 10.1097/QAI.0b013e31824006c4

Thomson, M. M., Herrero, I., Villahermosa, M. L., Vázquez de Parga, E., Cuevas, M. T., Carmona, R., et al. (2002). 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.* 83, 107–119. doi: 10.1099/0022-1317-83-1-107

Thomson, M. M., Sierra, M., Tanuri, A., May, S., Casado, G., Manjón, N., et al. (2004). Analysis of near full-length genome sequences of HIV type 1 BF intersubtype recombinant viruses from Brazil reveals their independent origins and their lack of relationship to CRF12_BF. *AIDS Res. Hum. Retroviruses* 20, 1126–1133. doi: 10.1089/aids.2004.20.1126

Thomson, M. M., Villahermosa, M. L., Vázquez de Parga, E., Cuevas, M. T., Delgado, E., Manjón, N., et al. (2000). Widespread circulation of a B/F intersubtype recombinant form among HIV-1-infected individuals in Buenos Aires, Argentina. *AIDS* 14, 897–899. doi: 10.1097/00002030-200005050-00020

Travers, S. A., Clewley, J. P., Glynn, J. R., Fine, P. E., Crampin, A. C., Sibande, F., et al. (2004). Timing and reconstruction of the most recent common ancestor of the subtype C clade of human immunodeficiency virus type 1. *J. Virol.* 78, 10501–10506. doi: 10.1128/JVI.78.19.10501-10506.2004

Vasylyeva, T. I., Zarebski, A., Smyrnov, P., Williams, L. D., Korobchuk, A., Liulchuk, M., et al. (2020). Phyldynamics helps to evaluate the impact of an HIV prevention intervention. *Viruses* 12:469. doi: 10.3390/v12040469

Venner, C. M., Nankya, I., Kyeyune, F., Demers, K., Kwok, C., Chen, P. L., et al. (2016). Infecting HIV-1 subtype predicts disease progression in women of Sub-Saharan Africa. *EBioMedicine* 13, 305–314. doi: 10.1016/j.ebiom.2016.10.014

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Cañada-García, Delgado, Gil, Benito, Sánchez, Ocampo, Cabrera, Miralles, García-Bodas, Martí, Ordóñez, Gade, Espeleta and Thomson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.