Identification of CRF66_BF, a new HIV-1 circulating recombinant form of South American origin.

Joan Bacqué[1], Elena Delgado[1], Sonia Benito[1], María Moreno-Lorenzo[1], Vanessa Montero[1], Horacio Gil[1], Mónica Sánchez[1], María Carmen Nieto-Toboso[2], Josefa Muñoz[2], Miren Z. Zubero-Sulibarria[2], Estíbaliz Ugalde[2], Elena García-Bodas[1], Javier E. Cañada[1], Jorge del Romero[3], Carmen Rodríguez[3], Iciar Rodríguez-Avial[4], Luis Elorduy-Otazua[5], José J. Portu[6], Juan García-Costa[7], Antonio Ocampo[8], Jorge J. Cabrera[8], Michael M. Thomson[1*].

[1]HIV Biology and Variability Unit, Centro Nacional de Microbiología, Instituto de Salud Carlos III, Majadahonda, Madrid, Spain.
[2]Hospital Universitario de Basurto, Bilbao, Spain.
[3]Centro Sanitario Sandoval, Madrid, Spain.
[4]Hospital Clínico San Carlos, Madrid, Spain.
[5]Hospital Universitario de Cruces, Barakaldo, Bizkaia, Spain.
[6]Hospital Universitario de Araba, Vitoria, Spain.
[7]Complejo Hospitalario Universitario de Ourense, Ourense, Spain.
[8]Complejo Hospitalario Universitario de Vigo, Vigo, Pontevedra, Spain.

***Corresponding author:**
Michael M. Thomson
mthomson@isciii.es

Abstract

Circulating recombinant forms (CRFs) are important components of the HIV-1 pandemic. Among 108 reported in the literature, 17 are BF1 intersubtype recombinant, most of which are of South American origin. Among these, all 5 identified in the Southern Cone and neighboring countries, except Brazil, derive from a common recombinant ancestor related to CRF12_BF, which circulates widely in Argentina, as deduced from coincident breakpoints and clustering in phylogenetic trees. In a HIV-1 molecular epidemiological study in Spain, we identified a phylogenetic cluster of 20 samples from 3 separate regions which were of F1 subsubtype, related to the Brazilian strain, in protease-reverse transcriptase (Pr-RT) and of subtype B in integrase. Remarkably, 14 individuals from this cluster (designated BF9) were Paraguayans and only 4 were native Spaniards. HIV-1 transmission was predominantly heterosexual, except for a subcluster of 6 individuals, 5 of which were men who have sex with men. Ten additional database sequences, from Argentina (n=4), Spain (n=3), Paraguay (n=1), Brazil (n=1), and Italy (n=1), branched within the BF9 cluster. To determine whether it represents a new CRF, near full-length genome (NFLG) sequences were obtained for 6 viruses from 3 Spanish regions. Bootscan analyses showed a coincident BF1 recombinant structure, with 5 breakpoints, located in $p17^{gag}$, integrase, gp120, gp41-*rev* overlap, and *nef*, which was identical to that of two BF1 recombinant viruses from Paraguay previously sequenced in NFLGs. Interestingly, none of the breakpoints coincided with those of CRF12_BF. In a maximum likelihood phylogenetic tree, all 8 NFLG sequences grouped in a strongly supported clade segregating from previously identified CRFs and from the CRF12_BF 'family' clade. These results allow us to identify a new HIV-1 CRF, designated CRF66_BF. Through a Bayesian coalescent analysis, the most recent common ancestor of CRF66_BF was estimated around 1984 in South America, either in Paraguay or Argentina. Among Pr-RT sequences obtained by us from HIV-1-infected Paraguayans living in Spain, 14 (20.9%) of 67 were of CRF66_BF, suggesting that CRF66_BF may be one of the major HIV-1 genetic forms circulating in Paraguay. CRF66_BF is the first reported non-Brazilian South American HIV-1 CRF_BF unrelated to CRF12_BF.

1. Introduction

One of the most distinctive features of HIV-1 evolution is its high recombinogenic potential, possibly the greatest among human pathogens, which is reflected in the high frequency of unique recombinant forms (URFs), each generated in a dually- or multiply-infected individual, found wherever different genetic forms circulate in the same population (Nájera et al., 2002). Some of the HIV-1 recombinant forms have spread beyond a group of epidemiologically-linked individuals, in which case they are designated circulating recombinant forms (CRFs) (Robertson et al., 2000). Currently, 108 CRFs have been reported in the literature and their number is increasing incessantly, due to both the generation of new CRFs and the identification of old previously undocumented CRFs. The proportion of CRFs in the HIV-1 pandemic has increased over time, representing around 17% infections in 2010-2015 (Hemelaar et al., 2020). Among CRFs, the most numerous are those derived from subtype B and subusbtype F1, 17 of which have been identified, most of them originated in South America, derived from the F1 variant circulating in Brazil (Louwagie et al., 1994). The first CRF_BF identified in South America was CRF12_BF, which circulates widely in Argentina and Uruguay, where URFs related to CRF12_BF are frequently found (Thomson et al., 2000; Thomson et al., 2002; Carr et al., 2001). Subsequently, 4 CRF_BFs related to CRF12_BF, as evidenced by shared breakpoints and phylogenetic clustering, were identified in the Southern Cone of South America or neighboring countries, CRF17_BF (Aulicino et al., 2012), CRF38_BF (Ruchansky et al., 2009), CRF44_BF (Delgado et al., 2010), and CRF89_BF (Delgado et al., 2021), the last three having clear country associations, with Uruguay, Chile, and Bolivia and Peru, respectively. Due to their common ancestry, these 5 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, all CRF_BFs identified in Brazil are unrelated to CRF12_BF (De Sá Filho et al., 2006; Guimaraes et al., 2008; Sanabani et al., 2010; Pessoa et al., 2014; Reis et al., 2017; Reis et al., 2019). Here we report the identification of a new CRF_BF, found mainly in Paraguayan immigrants in Spain and also identified in Paraguay and Argentina. Interestingly, unlike all South

American CRF_BFs identified to date outside of Brazil, it has no relationship with CRF12_BF.

2. Materials and methods

2.1. Samples

Plasma samples from HIV-1-infected individuals were collected in 14 Spanish regions 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. The study 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.

2.2. RNA extraction, RT-PCR amplification, and sequencing

An ~1.4 kb *pol* fragment in protease-reverse transcriptase (Pr-RT) was amplified from plasma RNA 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. Some samples were also subject to amplification and sequencing of integrase. Near full-length genome (NFLG) sequences were obtained for selected samples by RT-PCR/nested PCR amplification from plasma RNA in four 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 MK298150, OK011530-OK011552.

2.3. 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 assessment of node support with

Shimodaira-Hasegawa (SH)-like local support values (Guindon et al., 2010). Subsequent maximum likelihood (ML) trees with sequences of interest were constructed with W-IQ-Tree (Trifinopoulos et al., 2016), using the best-fit substitution model selected by ModelFinder program (Kalyaanamoorthy et al., 2017), with assessment of node support with the ultrafast bootstrap approximation approach (Hoang et al., 2018). Trees were visualized with MEGA v7.0 (Kumar et al., 2016).

Mosaic structures were analyzed by bootscanning (Salminen et al., 1995) with SimPlot v1.3.5 (Lole et al., 1999). In these analyses, trees were constructed using the neighbor-joining method with the Kimura 2-parameter model and a window width of 400 nucleotides. Recombinant segments identified with SimPlot were further phylogenetically analyzed via ML with W-IQ-Tree. Intersubtype breakpoint locations were also determined with jpHMM (Schultz et al., 2009).

2.4. Temporal and geographical estimations

The time and the location of the most recent common ancestor (MRCA) of the identified CRF was 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). Prior to the BEAST analysis, the existence of temporal signal in the dataset was assessed with TempEst v1.5.3 (Rambaut et al., 2016), which determines the correlation of genetic divergence among sequences (measured as root-to-tip distance) with time. The BEAST analysis was performed using the SRD06 codon-based evolutionary model (with two codon position partitions: 1st+2nd and 3rd) (Shapiro et al., 2006). We also specified an uncorrelated lognormal relaxed clock and a Bayesian SkyGrid coalescent tree prior (Gill et al., 2013). The MCMCs were run for 50 million generations. We performed runs in duplicate, combining the posterior tree files with LogCombiner v1.10.4. Mixing and convergence were checked 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 clade credibility (MCC) tree with TreeAnnotator v1.10.4, after removal a 10% burn-in. MCC trees were visualized with FigTree v1.4.2 (Rambaut,

http://tree.bio.ed.ac.uk/software/figtree/). Parameter uncertainty was summarized in 95% highest posterior density (HPD) intervals.

3. Results

3.1. Identification of a BF recombinant cluster and epidemiological associations

In a molecular epidemiology study on HIV-1 in Spain we identified a cluster of 20 viruses of F1 subsubtype in Pr-RT, that in integrase, sequenced in 4 samples, were of subtype B, which was designated BF9. Inclusion in the phylogenetic analyses of Pr-RT sequences of all viruses in the Los Alamos HIV Sequence Database (Los Alamos National Laboratory, 2021) classified as being of F1 subsubtype or BF1 recombinant identified 10 additional viruses belonging to BF9, from Argentina (n=4), Spain (n=3), Paraguay (n=1), Brazil (n=1), and Italy (n=1) (Fig. 1). Pr-RT sequences of the BF9 cluster were most closely related to F1 viruses of the Brazilian variant. Epidemiological data of the 20 samples of the BF9 cluster from Spain processed by us are shown in Table 1. Remarkably, 14 individuals were from Paraguay and all 3 remaining database sequences from samples collected in Spain were from Latin Americans, one each from Paraguay, Argentina, and an unspecified country. Transmission was predominantly heterosexual, but 7 were men who have sex with men (MSM), the sequences of 5 of whom branched in a subcluster (Fig. 1).

3.2. Analyses of NFLG sequences and identification of a new CRF

To determine whether viruses from the BF9 cluster represent a new CRF, we obtained NFLG sequences from 6 samples from 3 Spanish regions and analyzed their mosaic structures by bootscanning. Two additional NFLG sequences of BF recombinant viruses from databases were also analyzed by bootscanning, both from Paraguay: 02PY_PSP0094, that branched in the BF9 cluster in Pr-RT, and 02PY_PSP0093, that showed high similarity to NFLGs of the BF9 cluster in BLAST searches of the Los Alamos database. All 8 analyzed genomes showed coincident mosaic structures, with 5 breakpoints, located in $p17^{gag}$, integrase, gp120, gp41-*rev* overlap, and *nef* (Fig. 2). Breakpoints were more precisely located using the midpoint

of B-F1 transitions, according to the positions where 75% consensuses of subtype B and the F1 Brazilian strain genomes differ, in HXB2 positions 950, 4327, 6486, 8498, and 9161. Breakpoint locations were also determined with jpHMM (Supplementary Table), which also found 5 breakpoints for each virus in intervals overlapping those of the other analyzed viruses and the 75% consensus B-F1 transition intervals in all cases except the breakpoint interval in *nef* of MD497796, that did not overlap the consensus B-F1 transition interval, and that in $p17^{gag}$ of PV106451, that was not detected by jpHMM. ML phylogenetic trees constructed with each interbreakpoint fragment confirmed the subtype assignation determined with bootscanning (Fig. 3).

In an ML tree constructed with the 7 NFLG genomes of the BF9 cluster and 02PY_PSP0093, all 8 genomes grouped in a strongly supported clade segregating away from all other CRF_BFs and of the clade formed by the 5 CRF_BFs of the CRF12_BF family (Fig. 4). It should be pointed out that 02PY_PSP0093 did not branch in the BF9 cluster in the tree of Pr-RT, which suggests that the Pr-RT segment of this virus could derive from secondary recombination with an F1 strain different from the parental F1 strain of all other BF9 viruses.

These results, therefore, allow to define a new CRF, which was designated CRF66_BF, whose mosaic structure is shown in Fig. 5.

3.3. Prevalence of CRF66_BF among HIV-1-infected Paraguayans residing in Spain

Among 67 HIV-1-infected Paraguayans residing in Spain studied by us, CRF66_BF was the most common non-subtype B genetic form, representing 20.9% (14 of 67) of total infections, 48.3% (14 of 29) of non-subtype B infections, and 60.1% (14 of 23) of F1/BF1 infections.

3.4. Temporal and geographical estimations of CRF66_BF origin

To estimate the time and place of origin of CRF66_BF, Pr-RT sequences where analyzed with a Bayesian coalescent method with BEAST 1.10.4. Prior to this analysis, TempEst analysis determined that there was an adequate temporal signal in the dataset ($r^2 = 0.5871$). In the BEAST analysis, for the sequences corresponding

to South American individuals residing in Spain, the assigned location trait was their country of origin, rather than their place of residence. This was done because most individuals harboring CRF66_BF identified in Spain were of South American origin (mostly from Paraguay) and because we found no definitive evidence of the local circulation of CRF66_BF in Spain, as reflected in clusters mainly comprising Spanish individuals. Therefore, we assumed that South Americans harboring CRF66_BF viruses had acquired HIV-1 in their countries of origin. In this analysis, the time of the MRCA of CRF66_BF was estimated around 1984 (95% HPD, 1971-1996), and its most probable location was in Paraguay (PP=0.77), with Argentina in second place (PP=0.20) (Fig. 6). Considering the possibility that local subclusters each found in one city could represent local transmissions, we performed a second analysis in which we assigned the country of location of the most recent diagnoses of such clusters to Spain, irrespective of the countries of origin of the individuals. In this analysis, Paraguay was also estimated as the most probable location of the MRCA of CRF66_BF, although with a lower support (PP=0.55), and the support for Argentina increased to a PP=0.42 (Supplementary Figure 1).

4. Discussion

The results of this study allow to define a new HIV-1 CRF, designated CRF66_BF, which is the 18th reported CRF derived from subtypes B and F. Samples harboring CRF66_BF were collected in 5 countries, in South America (Argentina, Paraguay, and Brazil) and Western Europe (Spain and Italy), with a majority collected in Spain. However, of samples collected in Spain, a great majority were from Paraguayan individuals. Bayesian coalescent analyses (performed with the assumption that South American individuals living in Spain harboring CRF66_BF viruses had acquired them in their countries of origin), pointed to a most probable origin of CRF66_BF in Paraguay (PP=0.77), with Argentina being the second most probable location (PP=0.20). When the analysis was performed assigning the most recently diagnosed samples of clusters found in a single Spanish city to Spain as the location trait, irrespective of the country of origin of the individual, the PPs for a MRCA in

Paraguay or Argentina were not very different (0.55 vs. 0.42, respectively). Therefore, the results point to a South American origin of CRF66_BF, either in Paraguay or Argentina, without a definitive support for either country. However, given the great predominance of Paraguayans among CRF66_BF-infected individuals living in Spain, we cannot rule out that the same could happen in Argentina, where Paraguayans represent the largest immigrant national group (Instituto Nacional de Estadísticas y Censos, República Argentina, 2021). If this was the case, and information on country of origin of the sampled individuals living in Argentina was included in the analyses, it is possible that the support for a root of the CRF66_BF tree in Paraguay would increase.

Among HIV-1-infected Paraguayans residing in Spain studied by us, there was relatively high prevalence (21%) of CRF66_BF infections, which suggests that CRF66_BF could be one of the major HIV-1 genetic forms circulating in Paraguay. A better knowledge of the current prevalence of CRF66_BF in Paraguay would require sequencing a representative sample of recent HIV-1 diagnoses in the country. However, HIV-1 sequences from only 24 patients sampled in Paraguay are available at the Los Alamos HIV Sequence database, and the most recent molecular epidemiological study published to date involves the analysis of sequences from 55 samples collected 18 to 19 years ago (Aguayo et al., 2008), which are not available in public databases.

Notably, CRF66_BF, unlike all other non-Brazilian CRF_BFs identified to date in South America (CRF12_BF, CRF17_BF, CRF38_BF, CRF44_BF, and CRF89_BF, all circulating in the Southern Cone or neighboring countries), is unrelated to CRF12_BF, as deduced from the lack of breakpoint coincidence and of phylogenetic clustering with CRF12_BF. This implies that CRF66_BF originated independently from viruses of the CRF12_BF family, with a presumable ancestry in Brazil, where B and F1 viruses are circulating (Louwagie et al., 1994).

CRF66_BF is the 5th CRF of South American ancestry originally identified in Western Europe [after CRF42_BF (Struck et al., 2015), CRF47_BF (Fernández-García et al., 2010), CRF60_BC (Simonetti et al., 2014), and CRF89_BF (Delgado

et al., 2021)], which, together with the reported propagation of HIV-1 variants of South America origin among the European population (Delgado et al., 2015; de Oliveira et al., 2010; Collaço Verás et al., 2012; Thomson et al., 2012; Vinken et al., 2019; Lai et al., 2014; Fabeni et al., 2015; Fabeni et al., 2020; Carvalho et al., 2015), points to an increasing relationship between the HIV-1 epidemics in both continents. This reflects migratory fluxes, most notably in Spain, where around 2.5 million South Americans live, representing nearly 40% of the migrant population (Instituto Nacional de Estadística, 2021a), and immigration from South America has increased greatly in recent years (Instituto Nacional de Estadística, 2021b) (Supplementary Figure 2). Considering the large and increasing South American immigrant population in Europe and the scarcity of HIV-1 sequences from most South American countries available in public databases, studies on HIV-1 genetic diversity and molecular epidemiology among South American immigrants living in Europe could provide novel insights into the HIV-1 epidemics in their countries of origin, as well as on the diffusion of South American HIV-1 variants in Europe.

It is interesting to note that although transmission of CRF66_BF is predominantly via heterosexual contact, most individuals in a cluster are MSM (Fig. 1), which suggests diffusion from a heterosexual to a MSM network. HIV-1 propagation between heterosexual and MSM networks has also been reported for CRF89_BF (Delgado et al., 2021) and for a large CRF02_AG cluster in Spain (Delgado et al., 2019), although in the latter case the direction of propagation was from MSM to heterosexuals.

One of the essential tasks of Biology is naming and classifying organisms. In this work, we have accomplished this task by identifying a new HIV-1 circulating recombinant form, derived from subtypes B and F1, named CRF66_BF. CRF66_BF most likely originated in South America, either in Paraguay or Argentina, and, unlike all non-Brazilian South American CRFs identified to date, is unrelated to CRF12_BF. The identification and genetic characterization of HIV-1 variants is the first and necessary step for molecular epidemiological studies examining their geographic dissemination, growth dynamics, and epidemiological associations, as well as for analyzing their biological properties, such as pathogenic and transmissibility

296 potentials, response to antiretroviral therapies, and susceptibility to immune
297 responses inducible by vaccines.

298

299 **Acknowledgments**

300 We thank José Antonio Taboada, from Consellería de Sanidade, Xunta de Galicia,
301 and Daniel Zulaika, from Osakidetza-Servicio Vasco de Salud, for their support of
302 this study, and the personnel at the Genomic Unit, Instituto de Salud Carlos III, for
303 technical assistance in sequencing.

304

305 **Author contributions**

306 MT and ED conceived the study and supervised the experimental work. JB, MT, and
307 ED processed sequences and performed phylogenetic analyses. MT performed
308 phylodynamic analyses. HG performed data curation. JB, SB, MM-L, VM, MS, EG-
309 B, and JEC performed experimental work. MCN-T, JM, MZZ-S, EU, JdR, CR, IR-A,
310 LE-O, JJP, JG-C, AO, and JJC obtained samples and epidemiological data from
311 patients. MT, ED, and HG wrote the manuscript. All authors read and approved the
312 manuscript.

313

314 **Funding**

315 This work was funded through Acción Estratégica en Salud Intramural (AESI),
316 Instituto de Salud Carlos III, projects PI16CIII/00033 and PI19CIII/00042; Red de
317 Investigación en SIDA (RIS), Instituto de Salud Carlos III, Subdirección General de
318 Evaluación y Fondo Europeo de Desarrollo Regional (FEDER), Plan Nacional I+D+I,
319 project RD16ISCIII/0002/0004; and scientific agreements with Consellería de
320 Sanidade, Government of Galicia (MVI 1004/16) and Osakidetza-Servicio Vasco de
321 Salud, Government of Basque Country (MVI 1001/16).

322

323

Conflict of interest statement

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.

REFERENCES

Aguayo, N., Laguna-Torres, V. A., Villafane, M., Barboza, A., Sosa, L., Chauca, G., et al. (2008). Epidemiological and molecular characteristics of HIV-1 infection among female commercial sex workers, men who have sex with men and people living with AIDS in Paraguay. Rev. Soc. Bras. Med. Trop. 41, 225-231. doi: 10.1590/s0037-86822008000300001.

Aulicino, P. C., Gómez-Carrillo, M., Bello, G., Rocco, C., Mangano, A., Carr, J., et al. (2012). Characterization of full-length HIV-1 CRF17_BF genomes and comparison to the prototype CRF12_BF strains. Infect. Genet. Evol. 12, 443-447. doi:10.1016/j.meegid.2012.01.003.

Cañada, J. E., Delgado, E., Gil, H., Sánchez, M., Benito, S., García-Bodas, E., et al (2021). Identification of a new HIV-1 BC intersubtype circulating recombinant form (CRF108_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. (2001). Diverse BF recombinants have spread widely since the introduction of HIV-1 into South America. AIDS 15, F41-F47. doi:10.1097/00002030-200110190-00002.

Carvalho, A., Costa, P., Triunfante, V., Branca, F., Rodrigues, F., Santos, C. L., et al. (2015). Analysis of a local HIV-1 epidemic in Portugal highlights established transmission of non-B and non-G subtypes. J. Clin. Microbiol. 53, 1506-1514. doi: 10.1128/JCM.03611-14.

Collaço Verás, N. M., Gray, R. R., de Macedo Brígido, L. F., Rodrigues, R., and Salemi, M. (2012). High-resolution phylogenetics and phylogeography of human immunodeficiency virus type 1 subtype C epidemic in South America. J. Gen. Virol. 92, 1698-1709. doi: 10.1099/vir.0.028951-0.

de Oliveira, T., Pillay, D., and Gifford, R. J. (2010). The HIV-1 subtype C epidemic in South America is linked to the United Kingdom. PLoS One 5, e9311. doi: 10.1371/journal.pone.0009311.

De Sá Filho, D. J., Sucupira, M. C., Caseiro, M. M., Sabino, E. C., Diaz, R. S., and Janini, L. (2006). M. Identification of two HIV type 1 circulating recombinant forms in Brazil. AIDS Res. Hum. Retroviruses 22, 1-13. doi: 10.1089/aid.2006.22.1.

Delgado, E., Benito, S., Montero, V., Cuevas, M. T., Fernández-García, A., Sánchez-Martínez, M., et al. (2019). Diverse large HIV-1 non-subtype B clusters are spreading among men who have sex with men in Spain. Front. Microbiol. 10, 655. doi: https://doi.org/10.3389/fmicb.2019.00655.

Delgado, E., Cuevas, M. T., Domínguez, F., Vega, Y., Cabello, M., Fernández-García, A., et al. (2015). Phylogeny and phylogeography of a recent HIV-1 subtype F outbreak among men who have sex with men in Spain deriving from a cluster with a wide geographic circulation in Western Europe. PLoS One 10, e0143325. doi: 10.1371/journal.pone.0143325.

Delgado, E., Fernández-García, A., Pérez-Losada, M., Moreno-Lorenzo, M., Fernández-Miranda, I., Benito, S., et al. (2021). Identification of CRF89_BF, a new member of an HIV-1 circulating BF intersubtype recombinant form family widely spread in South America. Sci. Rep. 11, 11442. doi: 10.1038/s41598-021-90023-x.

Delgado, E., Ríos, 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 in Galicia, Spain, which exhibits a pseudotype-like virion structure. J. Acquir. Immune Defic. Syndr. 29, 536-543. doi:10.1097/00126334-200204150-00016.

Fabeni, L., Alteri, C., Orchi, N., Gori, C., Bertoli, A., Forbici, F., et al. (2015). Recent transmission clustering of HIV-1 C and CRF17_BF strains characterized by NNRTI-related mutations among newly diagnosed men in Central Italy. PLoS One 10, e0135325. doi: 10.1371/journal.pone.0135325.

Fabeni, L., Santoro, M. M., Lorenzini, P., Rusconi, S., Gianotti, N., Costantini, A., et al. (2020). Evaluation of HIV transmission clusters among natives and foreigners living in Italy. Viruses 12, 791. doi: 10.3390/v12080791.

Fernández-García, A., Pérez-Alvarez, L., Cuevas, M. T., Delgado, E., Muñoz-Nieto, M., Cilla, G., et al. (2010). Identification of a new HIV type 1 circulating BF intersubtype recombinant form (CRF47_BF) in Spain. AIDS Res. Hum. Retroviruses 26, 827-832. doi: 10.1089/aid.2009.0311.

Gill, M., Lemey, P., Faria, N., Rambaut, A., Shapiro, B., and Suchard M. (2013). Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. Mol. Biol. Evol. 30, 713-724. doi: 10.1093/molbev/mss265.

Guimarães, M. L., Eyer-Silva, W. A., Couto-Fernandez, J. C., and Morgado, M. G. (2008). Identification of two new CRF_BF in Rio de Janeiro State, Brazil. AIDS 22, 433-435. doi: 10.1097/QAD.0b013e3282f47ad0.

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307-321. doi:10.1093/sysbio/syq010.

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/S1473-3099(18)30647-9.

Hoang, D. T., 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.

Instituto Nacional de Estadística (2021a). Principales series de población desde 1998. https://www.ine.es/jaxi/Datos.htm?path=/t20/e245/p08/l0/,andfile=01006.px#!tabs-tabla. Retrieved 17 August 2021.

Instituto Nacional de Estadística (2021b). Migraciones exteriores. https://www.ine.es/jaxiT3/Tabla.htm?t=24295 Retrieved 17 August 2021.

Instituto Nacional de Estadísticas y Censos, República Argentina (2021). Censo 2010. https://www.indec.gob.ar/indec/web/Nivel4-Tema-2-41-135. Retrieved 17 August 2021.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589. doi: https://doi.org/10.1038/nmeth.

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772-780. doi: 10.1093/molbev/mst010.

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.

Lai, A., Bozzi, G., Franzetti, M., Binda, F., Simonetti, F. R., Micheli, V., et al. (2014). Phylogenetic analysis provides evidence of interactions between Italian heterosexual and South American homosexual males as the main source of national HIV-1 subtype C epidemics. J. Med. Virol. 86, 729-736. doi: 10.1002/jmv.23891.

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. https://www.hiv.lanl.gov/content/sequence/HIV/mainpage.html. Retrieved 17 August 2021.

Louwagie, J., Delwart, E. L., Mullins, J. I., McCutchan, F. E., Eddy, G., and Burke, D. S. (1994). Genetic analysis of HIV-1 isolates from Brazil reveals presence of two distinct genetic subtypes. AIDS Res. Hum. Retroviruses 10:561-5677. doi: 10.1089/aid.1994.10.561.

441 Nájera, R., Delgado, E., Pérez-Álvarez, L., and Thomson, M. M. (2002). Genetic
442 recombination and its role in the development of the HIV-1 pandemic. AIDS 16, S3-
443 S16. doi: 10.1097/00002030-200216004-00002.

444 Pessôa, R., Watanabe, J. T., Calabria, P., Felix, A. C., Loureiro, P., Sabino, E. C.,
445 et al. (2014). Deep sequencing of HIV-1 near full-length proviral genomes identifies
446 high rates of BF1 recombinants including two novel circulating recombinant forms
447 (CRF) 70_BF1 and a disseminating 71_BF1 among blood donors in Pernambuco,
448 Brazil. PLoS One 9, e112674. doi: 10.1371/journal.pone.0112674.

449 Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2--approximately
450 maximum-likelihood trees for large alignments. PLoS One 5, e9490. doi:
451 10.1371/journal.pone.0009490.

452 Rambaut, A., Lam, T. T., Max, C. L., and Pybus, O. G. (2016). Exploring the temporal
453 structure of heterochronous sequences using TempEst (formerly Path-O-Gen). Virus
454 Evol. 2, vew007. doi: 10.1093/ve/vew007.

455 Reis, M. N. G., Bello, G., Guimarães, M. L., and Stefani, M. M. A. (2017).
456 Characterization of HIV-1 CRF90_BF1 and putative novel CRFs_BF1 in Central
457 West, North and Northeast Brazilian regions. PLoS One 12:e0178578. doi:
458 10.1371/journal.pone.0178578.

459 Reis, M. N. G., Guimarães, M. L., Bello, G., and Stefani, M. M. A. (2019).
460 Identification of new HIV-1 Circulating Recombinant Forms CRF81_cpx and
461 CRF99_BF1 in Central Western Brazil and of Unique BF1 Recombinant Forms.
462 Front Microbiol 10, 97. doi: 10.3389/fmicb.2019.00097.

463 Robertson, D. L., Anderson, J. P., Bradac, J. A., Carr, J. K., Foley, B., Funkhouser,
464 R. K., et al. (2000). HIV-1 nomenclature proposal. Science 288, 55-56. doi:
465 10.1126/science.288.5463.55d.

466 Ruchansky, D., Casado, C., Russi, J. C., Arbiza, J. R., and López-Galíndez, C.
467 (2009). Identification of a new HIV type 1 circulating recombinant form (CRF38_BF1)
468 in Uruguay. AIDS Res. Hum. Retroviruses 25, 351-356. doi: 10.1089/aid.2008.0248.

469 Salminen, M. O., Carr, J. K., Burke, D. S., and McCutchan, F. E. (1995). Identification
470 of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. AIDS
471 Res. Hum. Retroviruses 11, 1423-1425. doi: 10.1089/aid.1995.11.1423.

472 Sanabani, S. S., Pastena, E. R., Neto, W. K., Martinez, V. P., and Sabino, E. C.
473 (2010). Characterization and frequency of a newly identified HIV-1 BF1 intersubtype
474 circulating recombinant form in São Paulo, Brazil. Virol. J. 7, 74. doi:10.1186/1743-
475 422X-7-74.

476 Schultz, A. K., Zhang, M., Bulla, I., Leitner, T., Korber, B., Morgenstern, B., et al.
477 (2009). jpHMM: improving the reliability of recombination prediction in HIV-1. Nucleic
478 Acids Res. 37, W647-W651. doi: 10.1093/nar/gkp371.

479 Shapiro, B., Rambaut, A., and Drummond, A. J. (2006). Choosing appropriate
480 substitution models for the phylogenetic analysis of protein-coding sequences. Mol.
481 Biol. Evol. 23, 7-9. doi: 10.1093/molbev/msj021.

482 Sierra, M., Thomson, M. M., Ríos, M., Casado, G., Ojea de Castro, R., Delgado E,
483 et al. (2005). The analysis of near full-length genome sequences of human
484 immunodeficiency virus type 1 BF intersubtype recombinant viruses from Chile,
485 Venezuela and Spain reveals their relationship to diverse lineages of recombinant
486 viruses related to CRF12_BF. Infect. Genet. Evol. 5, 209-217. doi:
487 10.1016/j.meegid.2004.07.010.

488 Simonetti, F.R., Lai, A., Monno, L., Binda, F., Brindicci, G., Punzi, G., et al. (2014).
489 Identification of a new HIV-1 BC circulating recombinant form (CRF60_BC) in Italian
490 young men having sex with men. Infect. Genet. Evol. 23, 176-181. doi:
491 10.1016/j.meegid.2014.02.007.

492 Struck, D., Roman, F., De Landtsheer, S., Servais, J. Y., Lambert, C., Masquelier,
493 C., et al. (2015). Near full-length characterization and population dynamics of the
494 human immunodeficiency virus type 1 circulating recombinant form 42 (CRF42_BF)
495 in Luxembourg. AIDS Res. Hum. Retroviruses 31, 554-558. doi:
496 10.1089/AID.2014.0364.

497 Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J., and Rambaut,
498 A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST
499 1.10. Virus Evol. 4, vey016. doi: 10.1093/ve/vey016.

500 Thomson M. M., Fernández-García, A., Delgado, E., Vega, Y., Díez-Fuertes, F.,
501 Sánchez-Martínez, M., et al. (2012). Rapid expansion of a HIV-1 subtype F cluster
502 of recent origin among men who have sex with men in Galicia, Spain. J. Acquir.
503 Immune Defic. Syndr. 59, e49-e51. doi: 10.1097/QAI.0b013e3182400fc4.

504 Thomson, M. M., and Nájera, R. (2005). Molecular epidemiology of HIV-1 variants
505 in the global AIDS pandemic: an update. AIDS Rev. 7, 210-224.

506 Thomson, M. M., Herrero, I., Villahermosa, M. L., Vázquez de Parga, E., Cuevas, M.
507 T., Carmona, R., et al. (2002). Diversity of mosaic structures and common ancestry
508 of human immunodeficiency virus type 1 BF intersubtype recombinant viruses from
509 Argentina revealed by analysis of near full-length genome sequences. J. Gen. Virol.
510 83, 107-119. doi:10.1099/0022-1317-83-1-107.

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

515 Trifinopoulos, J., Nguyen, L. T., von Haeseler, A., and Minh, B. Q. (2016). W-IQ-
516 TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids
517 Res. 44(W1):W232-W235. doi: 10.1093/nar/gkw256.

518 Vinken, L., Fransen, K., Cuypers, L., Alexiev, I., Balotta, C., Debaisieux, L., et al.
519 (2019). Earlier initiation of antiretroviral treatment coincides with an initial control of
520 the HIV-1 sub-subtype F1 outbreak among men-having-sex-with-men in Flanders,
521 Belgium. Front. Microbiol. 10, 613. doi: 10.3389/fmicb.2019.00613.

522 Zhang, M., Foley, B., Schultz, A. K., Macke, J. P., Bulla, .I, Stanke, M., et al. (2010).
523 The role of recombination in the emergence of a complex and dynamic HIV
524 epidemic. Retrovirology 7, 25. doi: 10.1186/1742-4690-7-25.

Figure legends

Figure 1. Maximum likelihood tree of Pr-RT sequences of BF9 cluster. Names of sequences obtained by us, all collected in Spain, are in bold type. In database sequences, the country of sample collection is indicated before the virus name with the 2-letter ISO country code. After the names of viruses of the BF9 cluster, the 2-letter ISO code of country of origin of the patient and/or the transmission route, when known, are shown in parentheses. F1 sequences from different countries are included in the analysis, together with two F2 sequences used as outgroups. Only ultrafast bootstrap values ≥80% are shown. PY: Paraguay; AR: Argentina; ES: Spain; BR: Brazil; IT: Italy; GQ: Equatorial Guinea; MSM: man who has sex with men; HT: heterosexual; SX-M: male with unspecified sexual acquisition of HIV-1.

Figure 2. Bootscan analyses of 6 NFLG sequences of viruses of the BF9 cluster obtained by us and of two BF1 database NFLG sequences from Paraguay, 02PY_PSP0093 and 02PY_PSP0094. The horizontal axis represents the position in the HXB2 genome of the midpoint of a 400 nt window moving in 20 nt increments and the vertical axis represents bootstrap values supporting clustering with subtype reference sequences.

Figure 3. Phylogenetic trees of interbreakpoint genome segments of the BF recombinant viruses analyzed by bootscanning. HXB2 positions delimiting the analyzed segments are indicated on top of the trees. Sequence names of BF viruses are in bold type. Names of subtype references are preceded by the corresponding subtype name. Only ultrafast bootstrap values ≥80% are shown.

Figure 4. Maximum likelihood tree of NFLG sequences of viruses of the BF9 cluster and PY02_PSP0094. References of all published CRF_BFs and of HIV-1 subtypes are also included in the analysis. The tree is rooted with SIVcpz virus MB66. Names of sequences obtained by us are in bold type. In reference sequences, the subtype or CRF is indicated before the virus name. Only ultrafast bootstrap values ≥90% are shown.

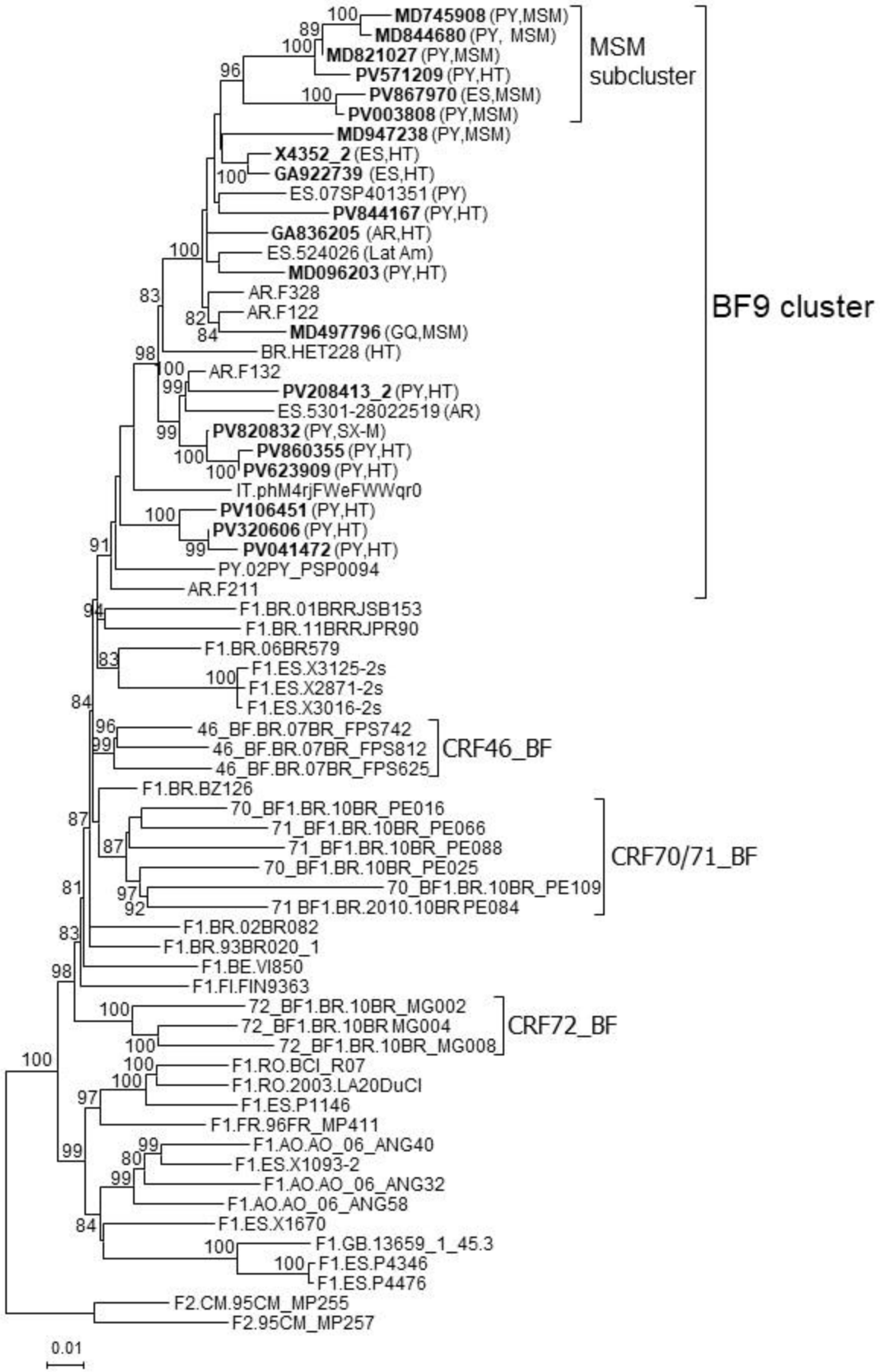
Figure 5. Mosaic structure of CRF66_BF. Breakpoint positions are numbered as in the HXB2 genome. The drawing was made using the Recombinant HIV-1 Drawing Tool https://www.hiv.lanl.gov/content/sequence/DRAW_CRF/recom_mapper.html

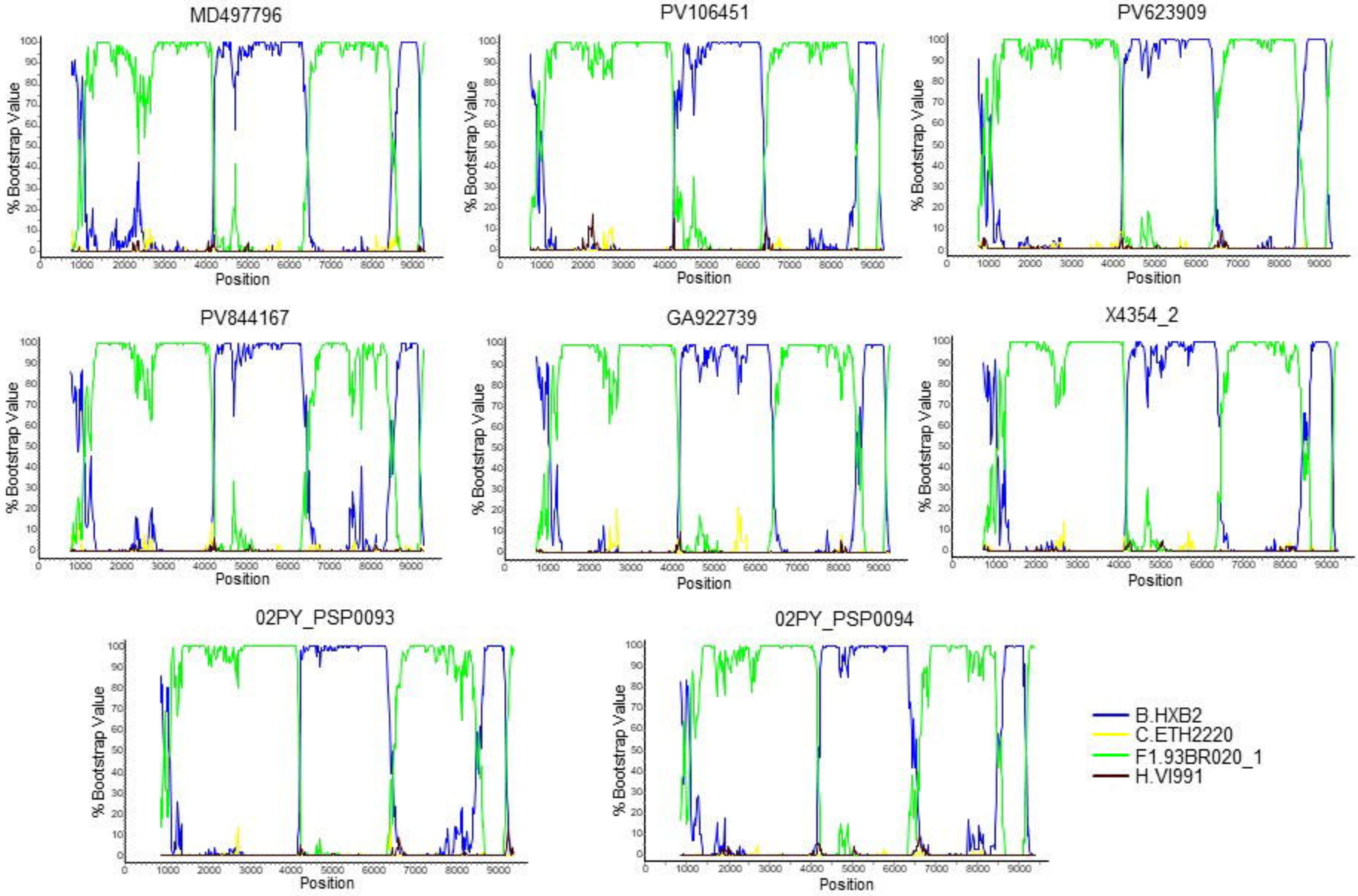
Figure 6. Maximum clade credibility tree of CRF66_BF Pr-RT sequences. Branch colors indicate, for terminal branches, country of sample collection or, for South American individuals residing in Spain, of origin of the individual, which was used as location trait (see Methods), and for internal branches, the most probable location country of the subtending node, according to the legend on the upper left. For database sequence 524026, from a sample collected in Spain, location was assigned to Paraguay as the most probably country of origin, although the only available information in the GenBank entry is that the individual was from Latin America, because 15 (88.2%) of 17 Latin Americans with CRF66_BF sampled in Spain were from Paraguay. Nodes supported by PP≥0.95 and PP 0.9-0.949 are indicated with filled and unfilled circles, respectively. The two most probable countries at the root of the tree are indicated, together with the PPs supporting each location and the time of the MRCA (mean value, with 95% HPD interval in brackets).

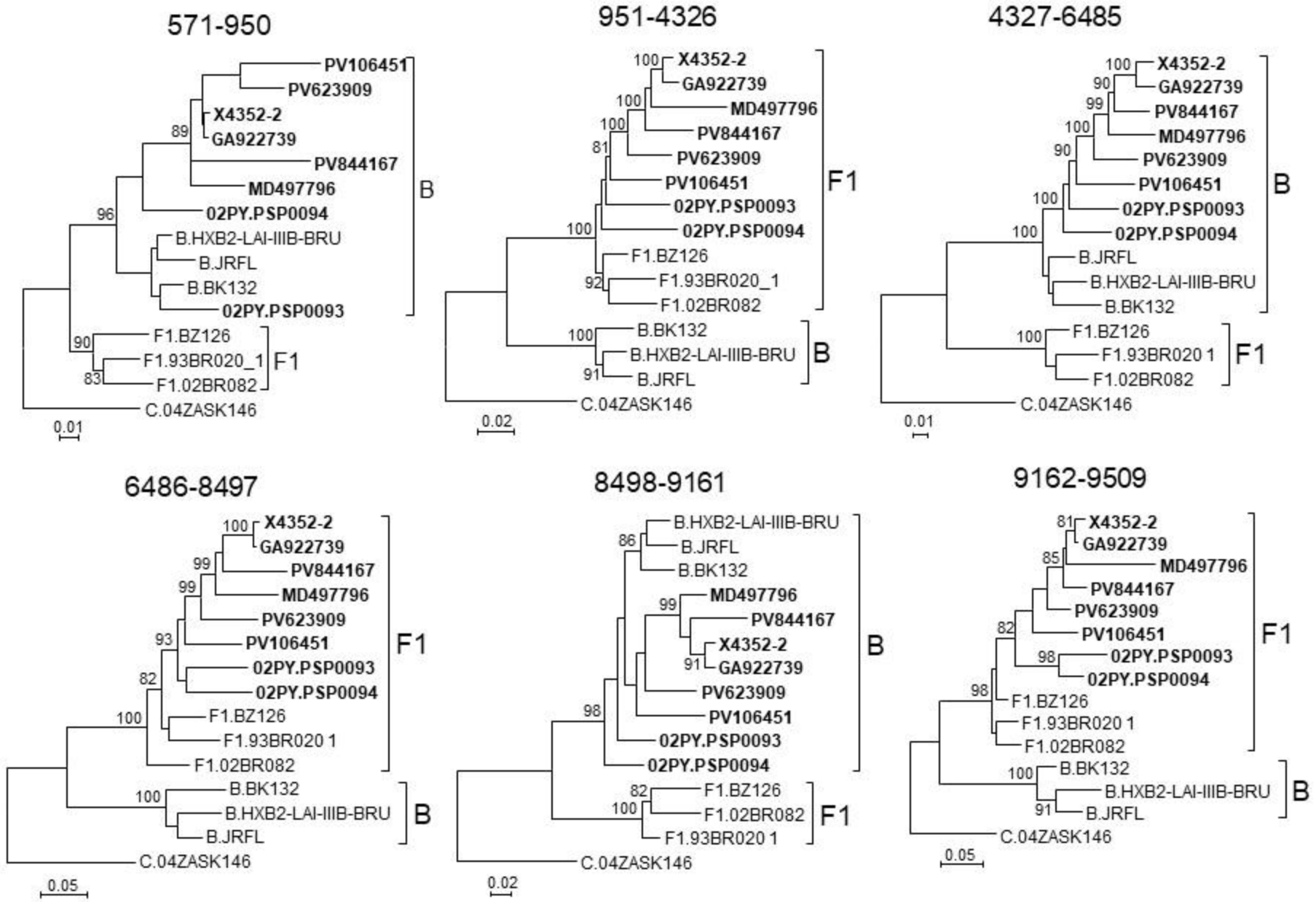
Table 1. Epidemiological data of patients and GenBank accessions of sequences.

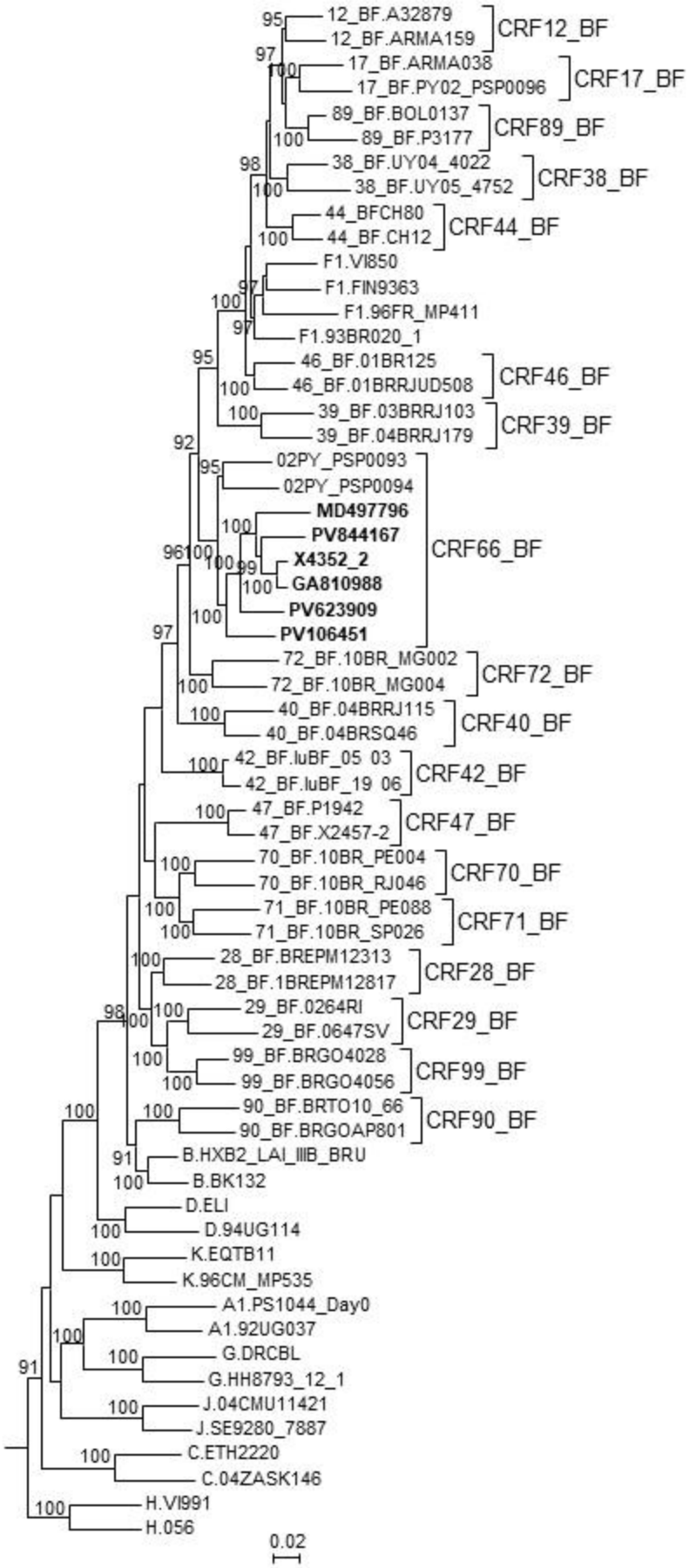
SAMPLE ID	City of sample collection	Region of sample collection	Year of sample collection	Year of HIV diagnosis	Gender	Transmission route*	Country of origin	GenBank accessions
X4352_2	Vigo	Galicia	2017	2017	M	HT	Spain	MK298150 (NFLG)
GA836205	Vigo	Galicia	2020	2020	M	HT	Argentina	OK011531 (Pr-RT) OK011530 (integrase)
GA922739	Ourense	Galicia	2018	2017	M	HT	Spain	OK011532 (NFLG)
MD096203	Madrid	Madrid	2017	2011	F	HT	Paraguay	OK011534 (Pr-RT) OK011533 (integrase)
MD497796	Madrid	Madrid	2017	2017	M	MSM	Equatorial Guinea	OK011534 (NFLG)
MD745908	Madrid	Madrid	2019	2019	M	MSM	Spain	OK011536 (Pr-RT)
MD821027	Madrid	Madrid	2018	2018	M	MSM	Paraguay	OK011537 (Pr-RT)
MD844680	Madrid	Madrid	2020	2020	M	MSM	Paraguay	OK011538 (Pr-RT)
MD947238	Madrid	Madrid	2018	2016	M	MSM	Paraguay	OK011539 (Pr-RT)
PV003808	Bilbao	Basque Country	2020	2020	M	MSM	Paraguay	OK011541 (Pr-RT) OK011540 (integrase)
PV041472	Bilbao	Basque Country	2014	2014	F	HT	Paraguay	OK011542 (Pr-RT)
PV106451	Bilbao	Basque Country	2010	2010	F	HT	Paraguay	OK011543 (NFLG)
PV208413_2	Bilbao	Basque Country	2009	2009	M	HT	Paraguay	OK011544 (Pr-RT)
PV320606	Bilbao	Basque Country	2014	2014	M	HT	Paraguay	OK011545 (Pr-RT)
PV571209	Bilbao	Basque Country	2013	2013	M	HT	Paraguay	OK011546 (Pr-RT)
PV623909	Bilbao	Basque Country	2011	2011	F	HT	Paraguay	OK011547 (NFLG)
PV820832	Bilbao	Basque Country	2008	2008	M	Sexual	Paraguay	OK011548 (Pr-RT)
PV844167	Vitoria	Basque Country	2016	2016	M	HT	Paraguay	OK011549 (NFLG)
PV860355	Bilbao	Basque Country	2011	2011	M	HT	Paraguay	OK011550 (Pr-RT)
PV867970	Bilbao	Basque Country	2020	2020	M	MSM	Spain	OK011552 (Pr-RT) OK011551 (integrase)

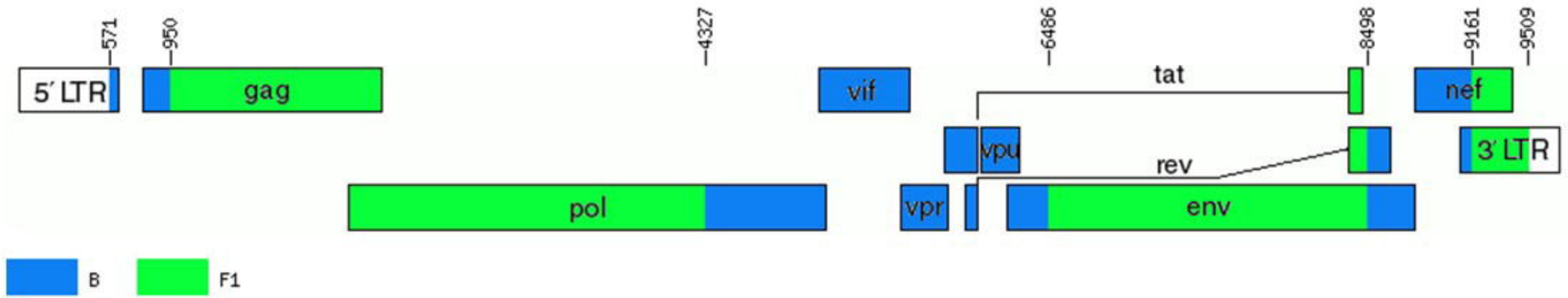
*HT: heterosexual; MSM: man who has sex with men.

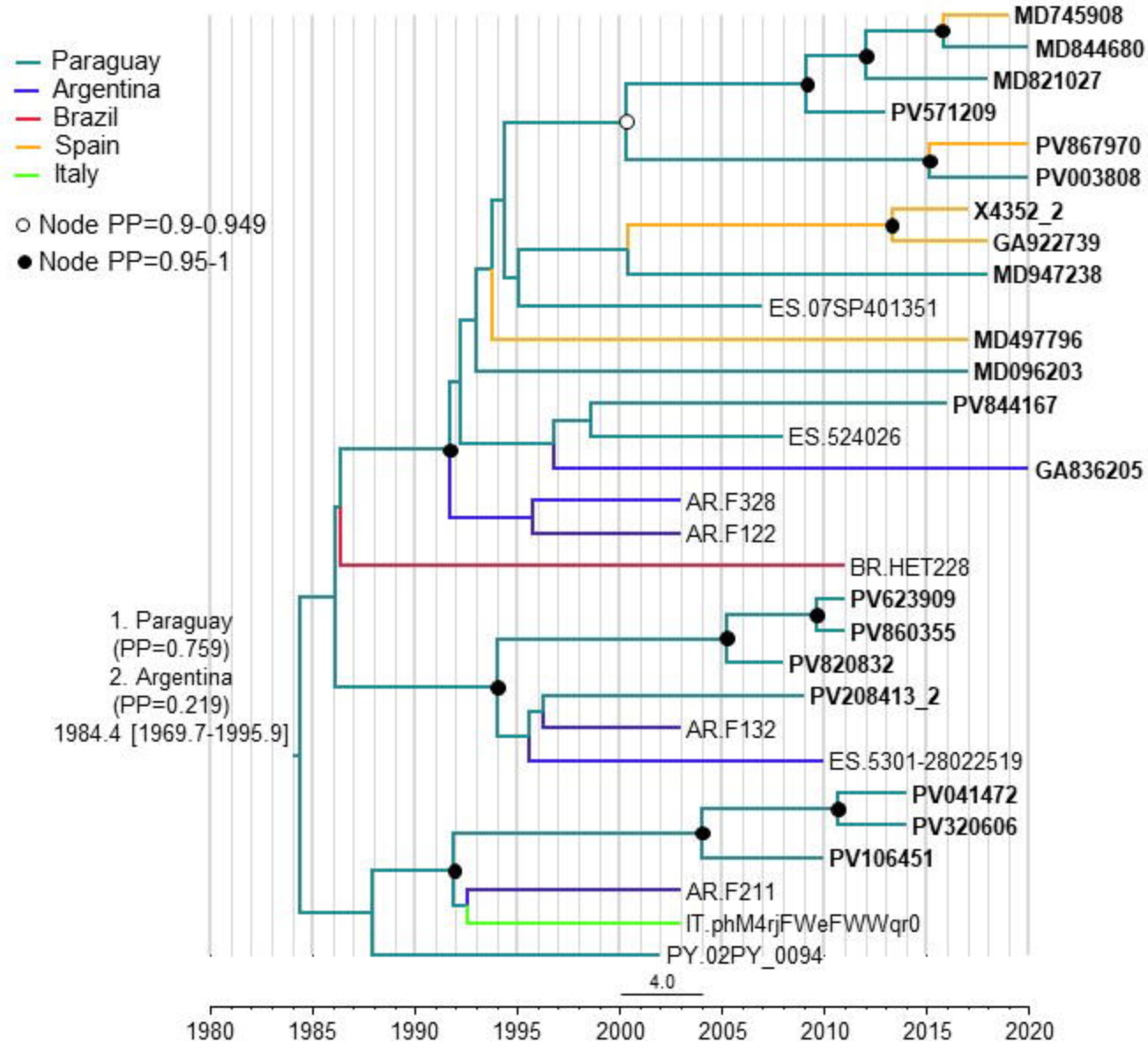

MD745908 (PY,MSM)
MD844680 (PY, MSM)
MD821027 (PY,MSM)
PV571209 (PY,HT)
PV867970 (ES,MSM)
PV003808 (PY,MSM)
MSM subcluster
MD947238 (PY,MSM)
X4352_2 (ES,HT)
GA922739 (ES,HT)
ES.07SP401351 (PY)
PV844167 (PY,HT)
GA836205 (AR,HT)
ES.524026 (Lat Am)
MD096203 (PY,HT)
AR.F328
AR.F122
MD497796 (GQ,MSM)
BR.HET228 (HT)
AR.F132
PV208413_2 (PY,HT)
ES.5301-28022519 (AR)
PV820832 (PY,SX-M)
PV860355 (PY,HT)
PV623909 (PY,HT)
IT.phM4rjFWeFWWqr0
PV106451 (PY,HT)
PV320606 (PY,HT)
PV041472 (PY,HT)
PY.02PY_PSP0094
AR.F211
BF9 cluster
F1.BR.01BRRJSB153
F1.BR.11BRRJPR90
F1.BR.06BR579
F1.ES.X3125-2s
F1.ES.X2871-2s
F1.ES.X3016-2s
46_BF.BR.07BR_FPS742
46_BF.BR.07BR_FPS812
46_BF.BR.07BR_FPS625
CRF46_BF
F1.BR.BZ126
70_BF1.BR.10BR_PE016
71_BF1.BR.10BR_PE066
71_BF1.BR.10BR_PE088
70_BF1.BR.10BR_PE025
70_BF1.BR.10BR_PE109
71 BF1.BR.2010.10BR PE084
CRF70/71_BF
F1.BR.02BR082
F1.BR.93BR020_1
F1.BE.VI850
F1.FI.FIN9363
72_BF1.BR.10BR_MG002
72_BF1.BR.10BR MG004
72_BF1.BR.10BR_MG008
CRF72_BF
F1.RO.BCI_R07
F1.RO.2003.LA20DuCI
F1.ES.P1146
F1.FR.96FR_MP411
F1.AO.AO_06_ANG40
F1.ES.X1093-2
F1.AO.AO_06_ANG32
F1.AO.AO_06_ANG58
F1.ES.X1670
F1.GB.13659_1_45.3
F1.ES.P4346
F1.ES.P4476
F2.CM.95CM_MP255
F2.95CM_MP257
0.01


MD497796
PV106451
PV623909
PV844167
GA922739
X4354_2
02PY_PSP0093
02PY_PSP0094
% Bootstrap Value
Position
B.HXB2
C.ETH2220
F1.93BR020_1
H.VI991


571-950
PV106451
PV623909
X4352-2
GA922739
PV844167
MD497796
02PY.PSP0094
B.HXB2-LAI-IIIB-BRU
B.JRFL
B.BK132
02PY.PSP0093
F1.BZ126
F1.93BR020_1
F1.02BR082
C.04ZASK146
B
F1
0.01
951-4326
X4352-2
GA922739
MD497796
PV844167
PV623909
PV106451
02PY.PSP0093
02PY.PSP0094
F1.BZ126
F1.93BR020_1
F1.02BR082
B.BK132
B.HXB2-LAI-IIIB-BRU
B.JRFL
C.04ZASK146
F1
B
0.02
4327-6485
X4352-2
GA922739
PV844167
MD497796
PV623909
PV106451
02PY.PSP0093
02PY.PSP0094
B.JRFL
B.HXB2-LAI-IIIB-BRU
B.BK132
F1.BZ126
F1.93BR020 1
F1.02BR082
C.04ZASK146
B
F1
0.01
6486-8497
X4352-2
GA922739
PV844167
MD497796
PV623909
PV106451
02PY.PSP0093
02PY.PSP0094
F1.BZ126
F1.93BR020 1
F1.02BR082
B.BK132
B.HXB2-LAI-IIIB-BRU
B.JRFL
C.04ZASK146
F1
B
0.05
8498-9161
B.HXB2-LAI-IIIB-BRU
B.JRFL
B.BK132
MD497796
PV844167
X4352-2
GA922739
PV623909
PV106451
02PY.PSP0093
02PY.PSP0094
F1.BZ126
F1.02BR082
F1.93BR020 1
C.04ZASK146
B
F1
0.02
9162-9509
X4352-2
GA922739
MD497796
PV844167
PV623909
PV106451
02PY.PSP0093
02PY.PSP0094
F1.BZ126
F1.93BR020 1
F1.02BR082
B.BK132
B.HXB2-LAI-IIIB-BRU
B.JRFL
C.04ZASK146
F1
B
0.05


12_BF.A32879
12_BF.ARMA159
CRF12_BF
17_BF.ARMA038
17_BF.PY02_PSP0096
CRF17_BF
89_BF.BOL0137
89_BF.P3177
CRF89_BF
38_BF.UY04_4022
38_BF.UY05_4752
CRF38_BF
44_BFCH80
44_BF.CH12
CRF44_BF
F1.VI850
F1.FIN9363
F1.96FR_MP411
F1.93BR020_1
46_BF.01BR125
46_BF.01BRRJUD508
CRF46_BF
39_BF.03BRRJ103
39_BF.04BRRJ179
CRF39_BF
02PY_PSP0093
02PY_PSP0094
MD497796
PV844167
X4352_2
GA810988
PV623909
PV106451
CRF66_BF
72_BF.10BR_MG002
72_BF.10BR_MG004
CRF72_BF
40_BF.04BRRJ115
40_BF.04BRSQ46
CRF40_BF
42_BF.luBF_05 03
42_BF.luBF_19 06
CRF42_BF
47_BF.P1942
47_BF.X2457-2
CRF47_BF
70_BF.10BR_PE004
70_BF.10BR_RJ046
CRF70_BF
71_BF.10BR_PE088
71_BF.10BR_SP026
CRF71_BF
28_BF.BREPM12313
28_BF.1BREPM12817
CRF28_BF
29_BF.0264RI
29_BF.0647SV
CRF29_BF
99_BF.BRGO4028
99_BF.BRGO4056
CRF99_BF
90_BF.BRTO10_66
90_BF.BRGOAP801
CRF90_BF
B.HXB2_LAI_IIIB_BRU
B.BK132
D.ELI
D.94UG114
K.EQTB11
K.96CM_MP535
A1.PS1044_Day0
A1.92UG037
G.DRCBL
G.HH8793_12_1
J.04CMU11421
J.SE9280_7887
C.ETH2220
C.04ZASK146
H.VI991
H.056
0.02


571
950
4327
6486
8498
9161
9509
5′ LTR
gag
vif
tat
nef
vpu
rev
3′ LTR
pol
vpr
env
B
F1


Paraguay
Argentina
Brazil
Spain
Italy
Node PP=0.9-0.949
Node PP=0.95-1
MD745908
MD844680
MD821027
PV571209
PV867970
PV003808
X4352_2
GA922739
MD947238
ES.07SP401351
MD497796
MD096203
PV844167
ES.524026
GA836205
AR.F328
AR.F122
BR.HET228
PV623909
PV860355
PV820832
PV208413_2
AR.F132
ES.5301-28022519
PV041472
PV320606
PV106451
AR.F211
IT.phM4rjFWeFWWqr0
PY.02PY_0094
1. Paraguay
(PP=0.759)
2. Argentina
(PP=0.219)
1984.4 [1969.7-1995.9]
4.0
1980
1985
1990
1995
2000
2005
2010
2015
2020