

Contents lists available at [SciVerse ScienceDirect](http://SciVerse.ScienceDirect.com)

# Infection, Genetics and Evolution

journal homepage: [www.elsevier.com/locate/meegid](http://www.elsevier.com/locate/meegid)

## Phylogenetics of HIV-1 subtype F1 in Angola, Brazil and Romania

Gonzalo Bello<sup>a,\*</sup>, Joana Morais Afonso<sup>a,b</sup>, Mariza G. Morgado<sup>a</sup><sup>a</sup>Laboratório de AIDS & Imunologia Molecular, Instituto Oswaldo Cruz – FIOCRUZ, Rio de Janeiro, Brazil<sup>b</sup>Faculdade de Medicina, Universidade Agostinho Neto, Luanda, Angola

### ARTICLE INFO

#### Article history:

Received 25 January 2012

Received in revised form 7 March 2012

Accepted 14 March 2012

Available online 29 March 2012

#### Keywords:

HIV-1

Subtype F1

Angola

Brazil

Romania

Demographic history

### ABSTRACT

The HIV-1 subtype F1 is exceptionally prevalent in Angola, Brazil and Romania. The epidemiological context in which the spread of HIV occurred was highly variable from one country to another, mainly due to the existence of a long-term civil war in Angola and the contamination of a large number of children in Romania. Here we apply phylogenetic and Bayesian coalescent-based methods to reconstruct the phylogenetic patterns of HIV-1 subtype F1 in such different epidemiological settings. The phylogenetic analyses of HIV-1 subtype F1 *pol* sequences sampled worldwide confirmed that most sequences from Angola, Brazil and Romania segregated in country-specific monophyletic groups, while most subtype F1 sequences from Romanian children branched as a monophyletic sub-cluster (Romania-CH) nested within sequences from adults. The inferred time of the most recent common ancestor of the different subtype F1 clades were as follow: Angola = 1983 (1978–1989), Brazil = 1977 (1972–1981), Romania adults = 1980 (1973–1987), and Romania-CH = 1985 (1978–1989). All subtype F1 clades showed a demographic history best explained by a model of logistic population growth. Although the expansion phase of subtype F1 epidemic in Angola (mid 1980s to early 2000s) overlaps with the civil war period (1975–2002), the mean estimated growth rate of the Angolan F1 clade ( $0.49 \text{ year}^{-1}$ ) was not exceptionally high, but quite similar to that estimated for the Brazilian ( $0.69 \text{ year}^{-1}$ ) and Romanian adult ( $0.36 \text{ year}^{-1}$ ) subtype F1 clades. The Romania-CH subtype F1 lineage, by contrast, displayed a short and explosive dissemination phase, with a median growth rate ( $2.47 \text{ year}^{-1}$ ) much higher than that estimated for adult populations. This result supports the idea that the AIDS epidemic that affected the Romanian children was mainly caused by the spread of the HIV through highly efficient parenteral transmission networks, unlike adult populations where HIV is predominantly transmitted through sexual route.

© 2012 Elsevier B.V. Open access under the [Elsevier OA license](http://www.elsevier.com/locate/elsevier).

### 1. Introduction

Phylogeographic analyses of the human immunodeficiency virus type 1 (HIV-1) group M, the pandemic branch of HIV, suggest that this clade originates in western-central Africa at around 1900–1930 and then spread to the rest of the world during the second half of the twentieth century (Korber et al., 2000; Worobey et al., 2008). The global dissemination of HIV-1 group M resulted from the random exportation of a few pandemic clades designated as subtypes (A–D, F–H, J and K) and inter-subtype recombinants. The most successful HIV-1 group M pandemic clades are subtype C (48% of all global infections), subtype A (12%), subtype B (11%), circulating recombinant form (CRF) 02\_AG (8%), CRF01\_AE (5%), subtype G (5%) and subtype D (2%); while subtypes F, H, J and K

together cause less than 1% of infections worldwide (Hemelaar et al., 2011).

Despite its low overall prevalence, subtype F1 is widely spread, being exceptionally prevalent in some specific countries of Central Africa, South America and Europe. Most of the HIV-1 subtype F1 infections described in Africa until date have been documented in Angola and the Democratic Republic of Congo (DRC). Molecular epidemiological data indicates a high prevalence of subtype F1 infections in Angola, ranging from 8% to 23%, being one of the most prevalent clades in the country (Abecasis et al., 2005; Bartolo et al., 2009; Castelbranco et al., 2010). By contrast, subtype F1 clade represent a small percentage (<5%) of the HIV-1 strains circulating in the DRC (Kalish et al., 2004; Mokili et al., 1999; Vidal et al., 2005, 2000; Yang et al., 2005). In South America, subtype F1 and BF1 recombinant variants reach a high prevalence (10–20%) in Brazil (Bongertz et al., 2000; Brindeiro et al., 2003). Other American countries from the Southern cone also displayed high prevalence of BF1 recombinants, but only sporadic cases of “pure” subtype F1 were described (Aulicino et al., 2007). In Europe, most subtype F1 infections are concentrated in Romania, where this subtype reaches a

\* Corresponding author. Address: Laboratório de AIDS & Imunologia Molecular, IOC/FIOCRUZ, Av. Brasil 4365, Leonidas Deane Building, Room #413, Rio de Janeiro 21040-900, Brazil. Tel.: +55 21 3865 8154; fax: +55 21 3865 8173.

E-mail address: [gbello@ioc.fiocruz.br](mailto:gbello@ioc.fiocruz.br) (G. Bello).

**Table 1**  
HIV-1 subtype F1 sequences.

Country	N	Sampling date
DRC + Belgium <sup>a</sup>	5	1987–2003
Angola	32	2001–2010
Brazil	49	1989–2009
Romania (AD) <sup>b</sup>	18	1997–2004
Romania (CH) <sup>b</sup>	18	1993–2004
Romania (CH-treated) <sup>b</sup>	333	2003–2007

<sup>a</sup> DRC, Democratic Republic of Congo.

<sup>b</sup> HIV-1 subtype F1 Romanian sequences derived from antiretroviral (ARV) therapy-naïve adults (AD) and children (CH), and from heavily treated adolescents infected during childhood (CH-treated).

prevalence >70% (Apetrei et al., 2003, 1998; Op De Coul et al., 2000; Paraschiv et al., 2007).

The epidemiological context in which subtype F1 dissemination occurred is highly variable among different countries. Heterosexual intercourse is the main epidemic's driving force in Angola, but the existence of a long-term civil war may have had profound effects on the HIV epidemic growth pattern. In Brazil, subtype F1 circulates among heterosexual, homosexual and intravenous drug users (IDU) (Guimaraes et al., 2001; Lacerda et al., 2007; Morgado et al., 1998; Raboni et al., 2010; Sabino et al., 1996; Teixeira et al., 2004; Vicente et al., 2000), indicating that both sexual and iatrogenic routes may have played an important role in viral dissemination. In Romania, subtype F1 seems to have been disseminated by the heterosexual route among adult population and by parenteral route in institutionalized children. The first pediatric case of HIV in Romania was reported in 1989 (Patrascu et al., 1990) and hundreds of cases were reported in the following years. Phylogenetic and epidemiologic evidences indicate that most Romanian children became horizontally infected by subtype F1 viruses from the adult population (Apetrei et al., 1997, 1998; Bandea et al., 1995; Op De Coul et al., 2000; Paraschiv et al., 2009), that probably entered in a health care environment and were subsequent disseminated through the use of contaminated needles and syringes and/or transfusion of unscreened blood or blood products (Hersh et al., 1991, 1993).

Such large differences in the epidemiological background may have a great impact in the patterns of epidemic growth of HIV-1 subtype F1 across diverse regions and populations. To test this hypothesis, we applied a Bayesian coalescent-based method to reconstruct the evolutionary and demographic history of the HIV-1 subtype F1 epidemic in Angola, Brazil, Romanian adults and Romanian children.

## 2. Material and methods

### 2.1. Sequence datasets

HIV-1 subtype F1 *pol* sequences from Angola, Brazil, Romania and the DRC, and some Belgian subtype F1 sequences probably linked to the DRC, that matched the selected genomic region (nucleotides 2550–3415 relative to HXB2) and from which the year of isolation was available were downloaded from the Los Alamos HIV Sequence Database (<http://www.hiv.lanl.gov>) by July 2011. The subtype assignment was confirmed using the REGA HIV subtyping tool v.2 (de Oliveira et al., 2005) and the Simplot software (Ray). All sequences with evidence of inter-subtype recombination, erroneous subtype F1 classification, stop codons and frame-shift mutations were removed. Sequences displaying major antiretroviral drug resistance mutations (DRM) were also initially excluded to avoid the effect of drug-induced convergent evolution on the phylogenetic analyses. Selected sequences were combined with 12

new subtype F1 sequences sampled in Angola between 2008 and 2010 (Afonso et al., unpublished results). This resulted in a final alignment of 122 subtype F1 sequences (Table 1). We also constructed a second alignment that further included 333 subtype F1 Romanian sequences with multiple DRM obtained from heavily treated adolescents infected during childhood (Paraschiv et al., 2009). Sequence alignments were created using the Clustal X program (Thompson et al., 1997) and are available from the authors upon request.

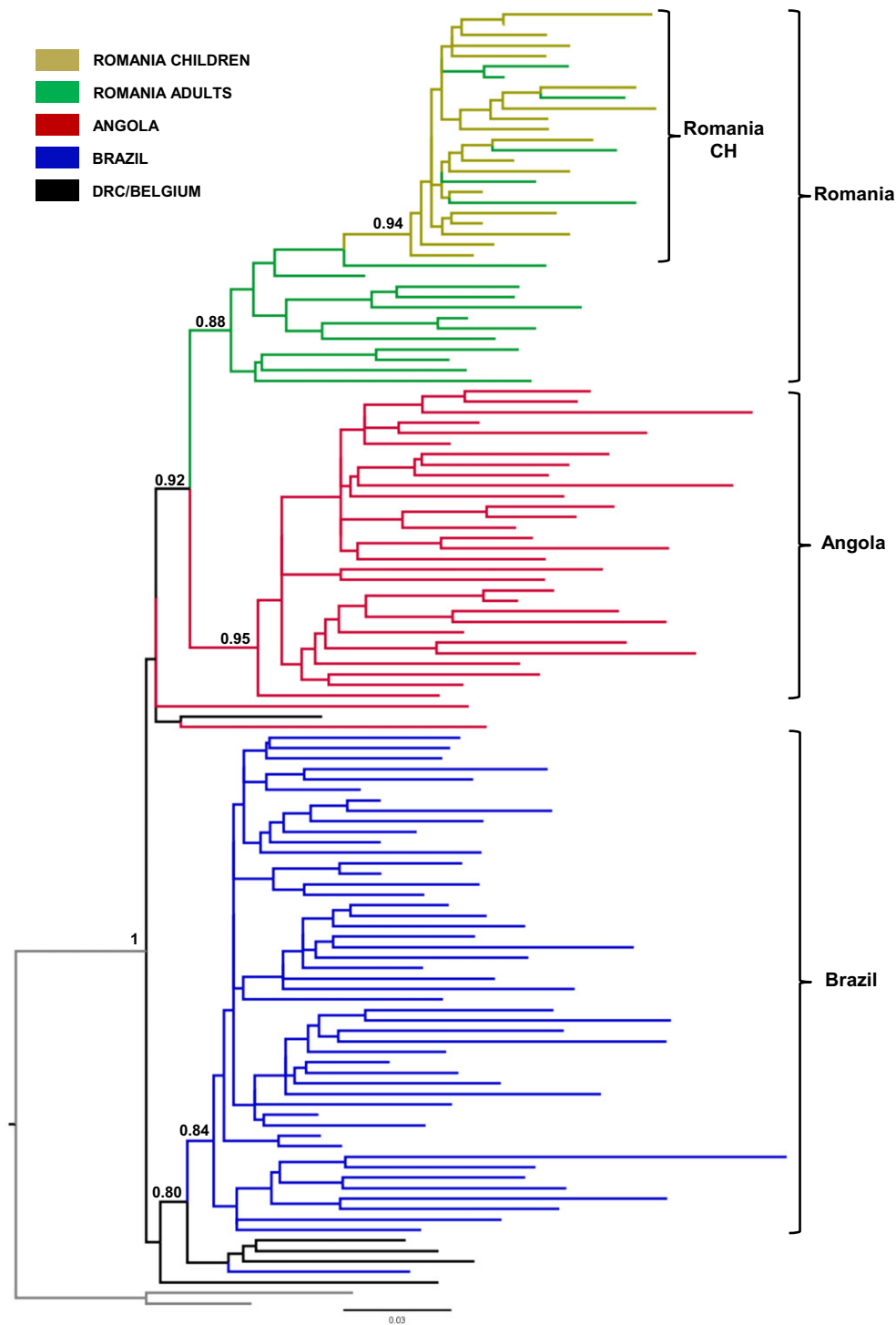
### 2.2. Phylogenetic analyses

Phylogenetic trees were inferred by the maximum likelihood (ML) and Bayesian methods under the GTR+I+ $\Gamma_4$  nucleotide substitution model, selected using the jModeltest program (Posada, 2008). ML trees were reconstructed with program PhyML (Guindon and Gascuel, 2003) using an online web server (Guindon et al., 2005). Heuristic tree search was performed using the SPR branch-swapping algorithm and the reliability of the obtained topology was estimated with the approximate likelihood-ratio test (*aLRT*) (Anisimova and Gascuel, 2006) based on the Shimodaira–Hasegawa-like procedure. Bayesian tree reconstructions were performed using MrBayes version 3.1.2 (Ronquist and Huelsenbeck, 2003). Two runs of four chains each were run for  $2 \times 10^7$  generations, with a burn-in of  $2 \times 10^6$  generations. Convergence of parameters was assessed by calculating the Effective Sample Size (ESS) using TRACER v1.4 (Rambaut and Drummond, 2007), after excluding an initial 10% for each run. All parameter estimates for each run showed ESS values >100. ML and Bayesian majority-rule consensus trees were visualized using the FigTree v1.3.1 program (Rambaut, 2009).

### 2.3. Estimation of the evolutionary and demographic history

The evolutionary rate ( $\mu$ , units are nucleotide substitutions per site per year, subst./site/year), the age of the most recent common ancestor ( $T_{\text{mrca}}$ , years), and the mode and rate ( $r$ , years<sup>-1</sup>) of population growth for the Angolan, Brazilian and Romanian subtype F1 epidemics were estimated using a Bayesian Markov Chain Monte Carlo (MCMC) approach as implemented in BEAST v1.6.2 (Drummond et al., 2002; Drummond and Rambaut, 2007). The temporal structure of subtype F1 *pol* datasets was not sufficient to reliably estimate the evolutionary rate under a chronological time-scale employing the dates of the sequences. Therefore, the interval of mean substitution rates at *pol* gene previously estimated for other HIV-1 group M subtypes ( $1.5 \times 10^{-3}$  to  $2.5 \times 10^{-3}$  subst./site/year) (Bello et al., 2008; Hue et al., 2005; Paraskevis et al., 2007; Passaes et al., 2009; Salemi et al., 2008) was incorporated as a prior probability distribution in our analyses. Analyses were performed using the GTR+I+ $\Gamma$  nucleotide substitution model assuming either a strict or a relaxed (uncorrelated Lognormal) molecular clock model (Drummond et al., 2006).

Estimations of evolutionary and demographic parameters involved two steps. First, the Bayesian skyline plot method (Drummond et al., 2005), was used to estimate  $\mu$ , the  $T_{\text{mrca}}$ , and the change in effective population size through time. Second, two different demographic models for each data set were compared: exponential and logistic growth; and estimates of the population growth rate were then obtained under the model that provided the best fit to the demographic signal in each data set. Model comparisons in a Bayesian framework were performed by calculating the Bayes Factor (BF) (Suchard et al., 2001) with TRACER v1.4. Two separate MCMC chains were run for  $1-5 \times 10^7$  generations for each data set, with a burn-in of  $1-5 \times 10^6$ . BEAST output was analyzed using TRACER v1.4, with uncertainty in parameter estimates reflected in the 95% Highest Probability Density (HPD) val-

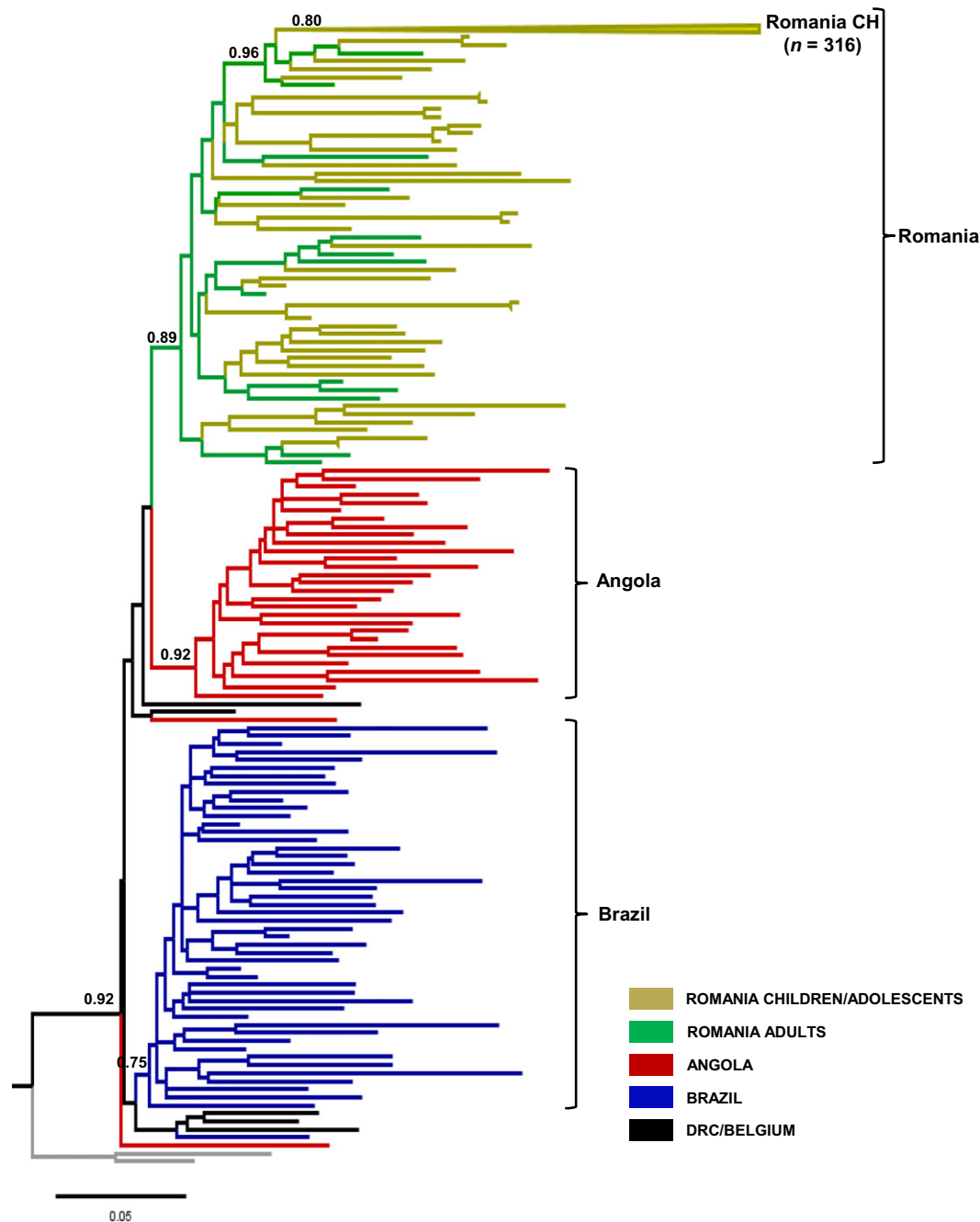


**Fig. 1.** Maximum likelihood tree of the *pol* gene of HIV-1 subtype F1 strains circulating in Angola, Brazil, Romania (adults and children), the DRC and Belgium (linked to the DRC). The aLRT support values are indicated only at key nodes. The color of each branch represents the country of origin of sequence corresponding to that branch, according to the legend in the figure. Brackets indicate the different monophyletic clusters identified. The tree was rooted using subtype F2 reference strains (green branches) as outgroups. Horizontal branch lengths are drawn to scale with the bar at the bottom indicating 0.03 nucleotide substitutions per site. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ues. Convergence of parameters was assessed through the ESS after excluding an initial 10% for each run. All parameter estimates for each run showed ESS values >200. A graphical representation of the effective number of infections through time was generated by using programs TRACER v1.4 and Prism 4 (GraphPad Software).

#### 2.4. GenBank accession numbers

The new Angolan HIV-1 subtype F1 *pol* sequences used in this study have been deposited in GenBank under Accession Nos. JN937026, JN937039, JN937044, JN937051, JN937064, JN937068,



**Fig. 2.** Maximum likelihood tree of the *pol* gene of HIV-1 subtype F1 strains circulating in Angola, Brazil, Romania (adults, adolescents and children), the DRC and Belgium (linked to the DRC). See legend of Fig. 1. The large Romania-CH clade made up almost entirely of sequences from adolescents and children was collapsed into a triangle for visual clarity and the number of sequences included in that clade is indicated. Horizontal branch lengths are drawn to scale with the bar at the bottom indicating 0.05 nucleotide substitutions per site.

JN937080, JN937089, JN937092, JN937111, JN937113 and JN937114.

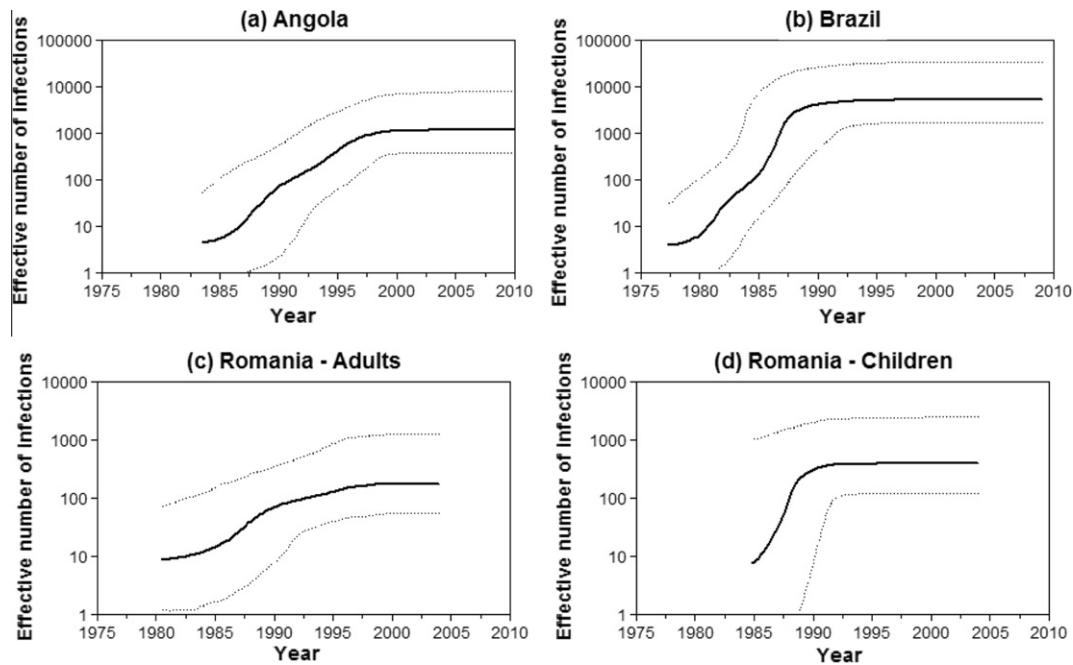
### 3. Results

#### 3.1. Phylogenetics analyses of HIV-1 subtype F1 sequences

The ML phylogenetic tree (Fig. 1) of 122 HIV-1 subtype F1 *pol* sequences sampled worldwide showed that isolates from the DRC and Belgium (probably linked to the DRC) occupy the most basal positions in subtype F1 phylogeny, while most subtype F1 sequences from Angola, Brazil and Romania segregated in

country-specific monophyletic groups ( $aLRT \geq 0.84$ ) nested among the basal strains from the DRC and Belgium. Two isolates sampled in Angola at 2001 (EU068351 and EU068372) and one isolate sampled in Brazil at 1994 (AY455781) were intermixed among strains from the DRC and Belgium, outside the major Angolan and Brazilian clades. This phylogenetic analysis also showed that Romanian and Angolan F1 clades formed a highly supported ( $aLRT = 0.92$ ) monophyletic lineage, suggesting that both epidemics arisen from closely related subtype F1 strains. The Bayesian tree displayed the same overall topology (Supplementary Fig. 1).

The observed tree topology is fully consistent with a previous study conducted by our group (Guimaraes et al., 2009), but is in contrast with a recent work that described the nesting of the major



**Fig. 3.** Bayesian skyline plots representing estimates of effective number of infections (y-axis;  $\log_{10}$  scale) through time (x-axis; calendar years) for HIV-1 subtypes F1 epidemics in Angola (a), Brazil (b), Romanian adults (c), and Romanian children (d). Median estimate of the effective number of infections (solid line) and 95% confidence limits of the estimate (dashed lines) are shown in each graphic.

Angolan lineage within the Romanian clade (Mehta et al., 2011). The study of Mehta et al. includes all Romanian subtype F1 *pol* sequences available at the Los Alamos HIV database, most of which (>90%) were retrieved from heavily treated adolescents (Paraschiv et al., 2009). Because all *pol* sequences from Romanian treated adolescents displayed multiple DRM, they were excluded from our previous analyses. In order to test whether the inclusion of such sequences may produce a different tree topology, the 122 subtype F1 *pol* sequences previously used were aligned with 333 subtype F1 *pol* sequences obtained from Romanian treated adolescents. The new ML phylogenetic tree (Fig. 2), however, confirmed the reciprocal monophyly between subtype F1 lineages from Angola and Romania.

A closer inspection of the Romanian subtype F1 clade revealed that adult sequences occupied the most basal positions within the lineage whereas sequences from drug-naïve children branched as a monophyletic sub-cluster (Romania-CH;  $aLRT = 0.94$ ) nested within adult sequences (Fig. 1). The Romania-CH clade also includes six sequences from adults (AF204051, HM191569, HM191570, HM191574, HM191575, HM191577); most of which were reported to be infected by the heterosexual route (Paraschiv et al., 2007). The analysis of subtype F1 sequences from heavily

treated adolescents infected during childhood showed that most (88%) of adolescent sequences branched within the Romania-CH clade together with pediatric sequences, while a minor proportion (12%) was interspersed with adult sequences at basal positions in the Romanian clade (Fig. 2). These results confirm the hypothesis that subtype F1 viruses circulating in Romanian children originated from the adult population and further suggest the existence of a major lineage circulating in the pediatric population.

### 3.2. Timing the emergence of country-specific subtype F1 clades

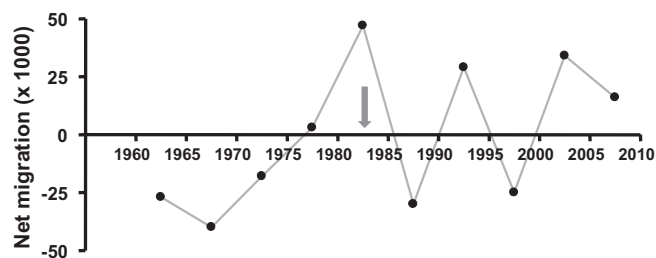
Bayesian MCMC analyses under a skyline tree prior were used to estimate the time-scale of the following country-specific subtype F1 clades: Angola ( $n = 30$ ), Brazil ( $n = 49$ ), Romanian adults (AD;  $n = 18$ ) and Romanian children (CH;  $n = 18$ ). For all subtype F1 data sets, the BF analysis favored the relaxed molecular clock over a strict molecular clock model ( $\ln BF > 10$ ), demonstrating a significant variation of substitution rate among branches (Supplementary Table 1). The median rate of evolution estimated under the relaxed molecular clock model was almost equal for all subtype F1 data sets ( $1.6\text{--}1.8 \times 10^{-3}$  subst./site/year) (Table 2). The 95% HPD intervals of such estimates (Table 2), however, almost coin-

**Table 2**

Time-scale and population dynamics estimates for Angolan, Brazilian and Romanian HIV-1 subtype F1 clades.

Data-set	Demographic model	$\mu$ (subst./site/year)	$T_{mrc}$	$r$ (year $^{-1}$ )	$\lambda$ (year)
Angola	Bayesian skyline	$1.7 \times 10^{-3}$ ( $1.5 \times 10^{-3}$ to $2.1 \times 10^{-3}$ )	1983 (1978–1989)	–	–
	Logistic growth	$1.7 \times 10^{-3}$ ( $1.5 \times 10^{-3}$ to $2.2 \times 10^{-3}$ )	1983 (1978–1989)	0.49 (0.30–0.77)	1.41 (0.90–2.31)
Brazil	Bayesian skyline	$1.6 \times 10^{-3}$ ( $1.5 \times 10^{-3}$ to $1.8 \times 10^{-3}$ )	1977 (1972–1981)	–	–
	Logistic growth	$1.6 \times 10^{-3}$ ( $1.5 \times 10^{-3}$ to $1.8 \times 10^{-3}$ )	1978 (1974–1982)	0.69 (0.43–1.02)	1.00 (0.68–1.61)
Romania (AD)	Bayesian skyline	$1.8 \times 10^{-3}$ ( $1.5 \times 10^{-3}$ to $2.4 \times 10^{-3}$ )	1980 (1973–1987)	–	–
	Logistic growth	$1.8 \times 10^{-3}$ ( $1.5 \times 10^{-3}$ to $2.4 \times 10^{-3}$ )	1981 (1974–1987)	0.36 (0.04–1.05)	1.92 (0.66–17.33)
Romania (CH)	Bayesian skyline	$1.7 \times 10^{-3}$ ( $1.5 \times 10^{-3}$ to $2.3 \times 10^{-3}$ )	1985 (1978–1989)	–	–
	Logistic growth	$1.7 \times 10^{-3}$ ( $1.5 \times 10^{-3}$ to $2.3 \times 10^{-3}$ )	1986 (1980–1989)	2.47 (0.57–7.16)	0.28 (0.10–1.22)

Median substitution rate ( $\mu$ ), median time of the most recent common ancestor ( $T_{mrc}$ ), median growth rate ( $r$ ) and median epidemic doubling time ( $\lambda$ ) estimated for the different HIV-1 subtype F1 clades. Ninety-five percent HPD intervals are shown in parentheses.



**Fig. 4.** Estimates of net number of migrants in Angola. The symbols represent the net number of migrants (number of immigrants minus the number of emigrants per thousands on the y-axis) for 5-year intervals (x-axis) from 1960 to 2010. Data were obtained from the United Nations World Population Prospects database (<http://esa.un.org/unpd/wpp/Excel-Data/migration.htm>). The arrow indicates the median age of the most recent common ancestor of Angolan subtype F1 clade.

cide with the informative prior interval used in our analyses ( $1.5\text{--}2.5 \times 10^{-3}$  subst./site/year), indicating that not much temporal information was added by the data. The median  $T_{\text{mrca}}$  for the different subtype F1 clades estimated under those substitution rates were as follows: Angola = 1983, Brazil = 1977, Romania-AD = 1980, and Romania-CH = 1985 (Table 2).

### 3.3. Demographic history of country-specific subtype F1 clades

The Bayesian skyline plot analyses suggest that all subtype F1 clades experienced an initial phase of fast exponential growth followed by a more recent decline in growth rate (Fig. 3). The precise time in which the growth rate starts to slow down, however, seems to vary among clades. While the growth rate of the Brazil and Romania-CH clades started to decrease around the early 1990s, subtype F1 continues to spread in Angola and Romanian adults until the early 2000s. To test the significance of such a recent decline in the epidemic growth rate, we compared the logistic and the exponential demographic models for each subtype F1 data set. According to the BF analysis, the model of logistic population growth was strongly supported over the exponential one for Angola, Brazil and Romania-CH datasets, and weakly supported for the Romania-AD dataset (Supplementary Table 1). A coalescent model of logistic growth was then used to estimate the initial growth rate of subtype F1 epidemics in the different populations, giving the following median values: Angola =  $0.49 \text{ year}^{-1}$ , Brazil =  $0.69 \text{ year}^{-1}$ , Romania-AD =  $0.36 \text{ year}^{-1}$  and Romania-CH =  $2.47 \text{ year}^{-1}$  (Table 2).

## 4. Discussion

Our phylogenetic analyses showed that most HIV-1 subtype F1 sequences from Angola, Brazil and Romania segregate in country-specific monophyletic groups nested within more basal subtype F1 sequences from the DRC and Belgium (probably linked to the DRC), indicating that the original diversification of the HIV-1 subtype F1 clade probably occurred within the DRC and subsequently spread to the other countries.

A recent study of Mehta et al. (2011) indicates that Angolan subtype F1 strains form a monophyletic cluster nested within the Romanian one, suggesting that an Angolan sub-epidemic stemmed from the Romanian epidemic. Alternatively, authors suggest that the Angolan epidemic has not been sampled as densely as the Romanian epidemic and that several introductions of HIV-1 subtype F1 occurred from Angola to Romania. This result contrast with a previous study conducted by our group that indicates that Angolan and Romanian subtype F1 isolates segregate in two reciprocal monophyletic clusters (Guimaraes et al., 2009). The study of Mehta et al. included a large number of Romanian subtype F1 sequences

retrieved from heavily treated adolescent containing multiple DRM, that were not considered in our previous work. In this study we have include those Romanian sequences from treated adolescent and we further expand the number of Angolan sequences analyzed. The new phylogenetic trees obtained, however, confirmed the reciprocal monophyly of Angolan and Romanian subtype F1 clades, supporting the hypothesis that both epidemics arose from single founder events involving closely related subtype F1 strains. Thus, factors other than the number of sequences and/or the presence of DRM should explain the differences between studies.

The  $T_{\text{mrca}}$  of Angolan, Brazilian and Romanian subtype F1 clades were estimated at 1983 (95% HPD: 1978–1989), 1977 (95% HPD: 1972–1981) and 1980 (95% HPD: 1973–1987), respectively; consistent with previous studies (Aulicino et al., 2007; Bello et al., 2007, 2006; Mehta et al., 2011). This indicates that the three subtype F1 sub-epidemics started to spread at around the same time, between the late 1970s and the early 1980s. This time-frame coincides with important socio-political changes that occurred in Angola after the beginning of the civil war in 1975. The Angolan civil war was not only associated to an important wave of emigration, but also received substantial support of several foreign powers including Cuba, countries from the Eastern bloc (including Romania), USA, South Africa and Zaire (current DRC). This situation gave an international dimension to the Angolan conflict that may have triggered the migration of subtype F1 viruses from Central Africa to Europe and South America as well as the dissemination of subtype F1 viruses within Angola. Of note, the estimated mean onset date of the Angolan subtype F1 clade coincides with a period of positive migration influx (1980–1985) that was preceded by a phase of negative migratory outflow (1960–1980), according to the estimates of the United Nations World Population Prospects for Angola (Fig. 4). This supports a scenario in which one Angolan individual that migrated to the DRC between 1960 and 1980, became infected with a subtype F1 strain and then returned to Angola in the early 1980s, initiating the local dissemination of the virus.

Our coalescent analyses suggest that subtype F1 infections in Angola and Brazil experienced an initial phase of fast exponential growth, followed by a deceleration of the rate of expansion in recent years. The expansion phase of the Angolan clade probably lasted from the mid 1980s to early 2000s, thus overlapping with the Angolan civil war period (1975–2002). While, the exponential growth of the Brazilian clade was probably from the late 1970s to early 1990s, in agreement with our previous estimations (Bello et al., 2007); covering a period during which no armed conflict occurred in the country. Despite such a difference in the socio-political context, the mean estimated growth rate of the Angolan subtype F1 epidemic ( $0.49 \text{ year}^{-1}$ ) was quite similar to that estimated for the Brazilian subtype F1 epidemic here ( $0.69 \text{ year}^{-1}$ ) and in a previous study ( $\sim 0.60 \text{ year}^{-1}$ ) (Bello et al., 2007). This result challenge the notion that HIV transmission is either hampered (Gisselquist, 2004; Strand et al., 2007) or accelerated (Hankins et al., 2002; Salama and Dondero, 2001) by the armed conflicts.

Romania was a special case in the global AIDS epidemic because it displayed of what may represent the world's largest population of HIV-infected children by parenteral transmission. Our phylogenetic analyses revealed that subtype F1 viruses from adults lie at the base of the Romanian clade together with some viruses from Romanian adolescents that were infected during childhood, supporting the hypothesis that the Romanian HIV pediatric epidemic was the result of multiple transmissions of subtype F1 from adult to children (Mehta et al., 2011; Op De Coul et al., 2000). One of those transmissions, however, was particularly successful in spreading among Romanian children because most subtype F1 viruses from adolescents and children included in this analysis branched in a monophyletic sub-cluster (Romania-CH) nested

within adult sequences. The Tmrca of the Romania-CH clade was estimated at 1985 (95% HPD: 1978–1989), five years after the estimated introduction of subtype F1 into the adult Romanian population. The existence of a Romanian subtype F1 sub-cluster made up almost entirely of pediatric sequences was recently described by two independent studies that also estimate the origin of this clade to around the middle 1980s (Mbisa et al., in press; Mehta et al., 2011).

The reconstruction of the demographic history of subtype F1 viruses showed that HIV epidemics in Romanian adults and children seem to have followed very different growth patterns. Subtype F1 epidemic in adults experienced a slow but steady increase until the early 2000s, after which the epidemic growth rate started to slow-down. The Romania-CH clade, by contrast, experienced a short period of extremely fast growth between 1985 and 1990, before stabilizing. A very similar demographic pattern was recently described for subtype F1 epidemic among Romanian children (Mbisa et al., in press). The reconstructed demographic patterns are fully consistent with the epidemiological data that shows that the number of pediatric AIDS cases in Romania reached a peak in 1990 and then registered a roughly continuous decline; while there was a slower but steady increase in number of adult AIDS cases up to the early 2000s (Ruta and Cerneanu, 2008). The sudden stabilization of subtype F1 expansion among Romanian children around 1990 further coincides with the first recognition of the AIDS pediatric epidemic in Romania and the subsequent implementation of prevention and control measures for avoid HIV transmission in health care settings (Danziger, 1996).

Coalescent-based analyses also indicate that the Romania-CH clade spread at a mean rate of  $2.47 \text{ year}^{-1}$ , which corresponds to an epidemic doubling time of only three months. Such a mean growth rate is much faster than that estimated for subtype F1 clades circulating in adult populations from Romania ( $0.36 \text{ year}^{-1}$ ), Angola ( $0.49 \text{ year}^{-1}$ ) and Brazil ( $0.69 \text{ year}^{-1}$ ), and is one of the highest ever estimated for HIV using coalescent methods. The extremely high rate of subtype F1 expansion estimated among children points to a very efficient route of viral spread, consistent with the presumed dissemination of the virus through the re-use of unsterilized needles and syringes and/or transfusion of unscreened blood or blood products (Hersh et al., 1991, 1993), in sharp contrast to adult populations where HIV is mainly transmitted through sexual contacts. Recovering growth rates estimates from Romanian adult and children with narrower confidence intervals will be crucial to confirm these results.

## 5. Conclusions

In summary, this study confirms that Angolan, Brazilian and Romanian HIV-1 subtype F1 epidemics resulted from single founder events that probably occurred between the late 1970s and the early 1980s. The DRC seems to be the epicenter of subtype F1 dissemination, but the precise pathways of migration from Central Africa to South America and Europe are not fully resolved. The pattern of growth of the HIV subtype F1 epidemic in adult populations of countries with (Angola) or without (Brazil and Romania) long-term armed conflicts was roughly similar. Subtype F1 epidemic in Romanian children, by contrast, displayed a more explosive and shorter period of growth than that observed in adult populations, consistent with the notion that HIV was primary spread in Romanian children through highly efficient parenteral transmission networks. A denser sampling of subtype F1 viruses from the DRC, Angola and Romania will be necessary to fully understand the world-wide migration routes of this subtype and to obtain more precise demographic estimates.

## Acknowledgments

This study was supported by CNPq and FAPERJ grants. J.M.A. was supported by the Organization for Women in Science for the Developing World (OWSDW) Program.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.meegid.2012.03.014>.

## References

- Abecasis, A., Paraskevis, D., Epalanga, M., Fonseca, M., Burity, F., Bartolomeu, J., Carvalho, A.P., Gomes, P., Vandamme, A.M., Camacho, R., 2005. HIV-1 genetic variants circulation in the North of Angola. *Infect. Genet. Evol.* 5, 231–237.
- Anisimova, M., Gascuel, O., 2006. Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Syst. Biol.* 55, 539–552.
- Apetrei, C., Descamps, D., Collin, G., Robertson, D.L., Pandrea, I., Groza, P., Prisecariu, L., Teodorescu, I., Luca, V., Brun-Vezinet, F., 2003. HIV type 1 diversity in northeastern Romania in 200–2001 based on phylogenetic analysis of pol sequences from patient failing antiretroviral therapy. *AIDS Res. Hum. Retroviruses* 19, 1155–1161.
- Apetrei, C., Lousert-Ajaka, I., Collin, G., Letourneur, F., Duca, M., Saragosti, S., Simon, F., Brun-Vezinet, F., 1997. HIV type 1 subtype F sequences in Romanian children and adults. *AIDS Res. Hum. Retroviruses* 13, 363–365.
- Apetrei, C., Necula, A., Holm-Hansen, C., Lousert-Ajaka, I., Pandrea, I., Cozmei, C., Streinu-Cercel, A., Pascu, F.R., Negut, E., Molnar, G., Duca, M., Pecet, M., Brun-Vezinet, F., Simon, F., 1998. HIV-1 diversity in Romania. *AIDS* 12, 1079–1085.
- Aulicino, P.C., Bello, G., Rocco, C., Romero, H., Mangano, A., Morgado, M.G., Sen, L., 2007. Description of the first full-length HIV Type 1 Subtype F1 strain in Argentina: implications for the origin and dispersion of this subtype in South America. *AIDS Res. Hum. Retroviruses* 23, 1176–1182.
- Banda, C.I., Ramos, A., Pieniazek, D., Pascu, R., Tanuri, A., Schochetman, G., Rayfield, M.A., 1995. Epidemiologic and evolutionary relationships between Romanian and Brazilian HIV-subtype F strains. *Emerg. Infect. Dis.* 1, 91–93.
- Bartolo, I., Rocha, C., Bartolomeu, J., Gama, A., Marcelino, R., Fonseca, M., Mendes, A., Epalanga, M., Silva, P.C., Taveira, N., 2009. Highly divergent subtypes and new recombinant forms prevail in the HIV/AIDS epidemic in Angola: new insights into the origins of the AIDS pandemic. *Infect. Genet. Evol.* 9, 672–682.
- Bello, G., Eyer-Silva, W.A., Couto-Fernandez, J.C., Guimaraes, M.L., Chequer-Fernandez, S.L., Teixeira, S.L., Morgado, M.G., 2007. Demographic history of HIV-1 subtypes B and F in Brazil. *Infect. Genet. Evol.* 7, 263–270.
- Bello, G., Guimaraes, M.L., Morgado, M.G., 2006. Evolutionary history of HIV-1 subtype B and F infections in Brazil. *AIDS* 20, 763–768.
- Bello, G., Passaes, C.P., Guimaraes, M.L., Lorete, R.S., Matos Almeida, S.E., Medeiros, R.M., Alencastro, P.R., Morgado, M.G., 2008. Origin and evolutionary history of HIV-1 subtype C in Brazil. *AIDS* 22, 1993–2000.
- Bongertz, V., Bou-Habib, D.C., Brigido, L.F., Caseiro, M., Chequer, P.J., Couto-Fernandez, J.C., Ferreira, P.C., Galvao-Castro, B., Greco, D., Guimaraes, M.L., Linhares de Carvalho, M.L., Morgado, M.G., Oliveira, C.A., Osmanov, S., Ramos, C.A., Rossini, M., Sabino, E., Tanuri, A., Ueda, M., 2000. HIV-1 diversity in Brazil: genetic, biologic, and immunologic characterization of HIV-1 strains in three potential HIV vaccine evaluation sites. Brazilian network for HIV isolation and characterization. *J. Acquir. Immune Defic. Syndr.* 23, 184–193.
- Brindeiro, R.M., Diaz, R.S., Sabino, E.C., Morgado, M.G., Pires, I.L., Brigido, L., Dantas, M.C., Barreira, D., Teixeira, P.R., Tanuri, A., 2003. Brazilian Network for HIV Drug Resistance Surveillance (HIV-BResNet): a survey of chronically infected individuals. *Aids* 17, 1063–1069.
- Castelbranco, E.P., da Silva Souza, E., Cavalcanti, A.M., Martins, A.N., de Alencar, L.C., Tanuri, A., 2010. Frequency of primary resistance to antiretroviral drugs and genetic variability of HIV-1 among infected pregnant women recently diagnosed in Luanda – Angola. *AIDS Res. Hum. Retroviruses* 26, 1313–1316.
- Danziger, R., 1996. An overview of HIV prevention in central and eastern Europe. *AIDS Care* 8, 701–707.
- de Oliveira, T., Deforche, K., Cassol, S., Salminen, M., Paraskevis, D., Seebregts, C., Snoeck, J., van Rensburg, E.J., Wensing, A.M., van de Vijver, D.A., Boucher, C.A., Camacho, R., Vandamme, A.M., 2005. An automated genotyping system for analysis of HIV-1 and other microbial sequences. *Bioinformatics* 21, 3797–3800.
- Drummond, A.J., Ho, S.Y., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4, e88.
- Drummond, A.J., Nicholls, G.K., Rodrigo, A.G., Solomon, W., 2002. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics* 161, 1307–1320.
- Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7, 214.
- Drummond, A.J., Rambaut, A., Shapiro, B., Pybus, O.G., 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* 22, 1185–1192.
- Gisselquist, D., 2004. Impact of long-term civil disorders and wars on the trajectory of HIV epidemics in sub-Saharan Africa. *SAHARA J* 1, 114–127.

- Guimaraes, M.L., Bastos, F.I., Telles, P.R., Galvao-Castro, B., Diaz, R.S., Bongertz, V., Morgado, M.G., 2001. Retrovirus infections in a sample of injecting drug users in Rio de Janeiro City, Brazil: prevalence of HIV-1 subtypes, and co-infection with HTLV-I/II. *J. Clin. Virol.* 21, 143–151.
- Guimaraes, M.L., Vicente, A.C., Otsuki, K., da Silva, R.F., Francisco, M., da Silva, F.G., Serrano, D., Morgado, M.G., Bello, G., 2009. Close phylogenetic relationship between Angolan and Romanian HIV-1 subtype F1 isolates. *Retrovirology* 6, 39.
- Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52, 696–704.
- Guindon, S., Lethiec, F., Duroux, P., Gascuel, O., 2005. PHYML Online – a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Res.* 33, W557–559.
- Hanks, C.A., Friedman, S.R., Zafar, T., Strathdee, S.A., 2002. Transmission and prevention of HIV and sexually transmitted infections in war settings: implications for current and future armed conflicts. *AIDS* 16, 2245–2252.
- Hemelaar, J., Gouws, E., Ghys, P.D., Osmanov, S., 2011. Global trends in molecular epidemiology of HIV-1 during 2000–2007. *AIDS* 25, 679–689.
- Hersh, B.S., Popovici, F., Apetrei, R.C., Zolotusca, L., Beldeanu, N., Calomfirescu, A., Jezek, Z., Oxtoby, M.J., Gromyko, A., Heymann, D.L., 1991. Acquired immunodeficiency syndrome in Romania. *Lancet* 338, 645–649.
- Hersh, B.S., Popovici, F., Jezek, Z., Satten, G.A., Apetrei, R.C., Beldeanu, N., George, J.R., Shapiro, C.N., Gayle, H.D., Heymann, D.L., 1993. Risk factors for HIV infection among abandoned Romanian children. *AIDS* 7, 1617–1624.
- Hue, S., Pillay, D., Clewley, J.P., Pybus, O.G., 2005. Genetic analysis reveals the complex structure of HIV-1 transmission within defined risk groups. *Proc. Natl. Acad. Sci. USA* 102, 4425–4429.
- Kalish, M.L., Robbins, K.E., Pieniazek, D., Schaefer, A., Nzilambi, N., Quinn, T.C., St Louis, M.E., Youngpairaj, A.S., Phillips, J., Jaffe, H.W., Folks, T.M., 2004. Recombinant viruses and early global HIV-1 epidemic. *Emerg. Infect. Dis.* 10, 1227–1234.
- Korber, B., Muldoon, M., Theiler, J., Gao, F., Gupta, R., Lapedes, A., Hahn, B.H., Wolinsky, S., Bhattacharya, T., 2000. Timing the ancestor of the HIV-1 pandemic strains. *Science* 288, 1789–1796.
- Lacerda, H.R., Barbosa de Medeiros, L., Salustiano Cavalcanti, A.M., de Alencar Ximenes, R.A., de Albuquerque, Militao, Mde, F., 2007. Comparison of the epidemiology, profile of mutations, and clinical response to antiretrovirals among subtypes B and F of the human immunodeficiency virus type 1. *Mem. Inst. Oswaldo Cruz* 102, 693–699.
- Mbisa, J.L., Hue, S., Buckton, A.J., Myers, R.E., Duiculescu, D., Ene, L., Oprea, C., Tardei, G., Rugina, S., Mardarescu, M., Floch, C., Notheis, G., Zoehrer, B., Cane, P.A., Pillay, D., in press. Phylogenetic and Phylogeographic Patterns of the HIV-1 Subtype F1 Parenteral Epidemic in Romania. *AIDS Res Hum Retroviruses*.
- Mehta, S.R., Wertheim, J.O., Delport, W., Ene, L., Tardei, G., Duiculescu, D., Pond, S.L., Smith, D.M., 2011. Using phylogeography to characterize the origins of the HIV-1 subtype F epidemic in Romania. *Infect. Genet. Evol.* 11, 975–979.
- Mokili, J.L., Wade, C.M., Burns, S.M., Cutting, W.A., Bopopi, J.M., Green, S.D., Peutherer, J.F., Simmonds, P., 1999. Genetic heterogeneity of HIV type 1 subtypes in Kimpese, rural Democratic Republic of Congo. *AIDS Res. Hum. Retroviruses* 15, 655–664.
- Morgado, M.G., Guimaraes, M.L., Gripp, C.B., Costa, C.I., Neves Jr., I., Veloso, V.G., Linhares-Carvalho, M.I., Castello-Branco, L.R., Bastos, F.I., Kuiken, C., Castilho, E.A., Galvao-Castro, B., Bongertz, V., 1998. Molecular epidemiology of HIV-1 in Brazil: high prevalence of HIV-1 subtype B and identification of an HIV-1 subtype D infection in the city of Rio de Janeiro, Brazil. *Evandro Chagas Hospital AIDS Clinical Research Group. J. Acquir. Immune Defic. Syndr. Hum. Retrovirol.* 18, 488–494.
- Op De Coul, E., van den Burg, R., Asjo, B., Goudsmit, J., Cupsa, A., Pascu, R., Usein, C., Cornelissen, M., 2000. Genetic evidence of multiple transmissions of HIV type 1 subtype F within Romania from adult blood donors to children. *AIDS Res. Hum. Retroviruses* 16, 327–336.
- Paraschiv, S., Otelea, D., Baicus, C., Tinischi, M., Costache, M., Neaga, E., 2009. Nucleoside reverse transcriptase inhibitor resistance mutations in subtype F1 strains isolated from heavily treated adolescents in Romania. *Int. J. Infect. Dis.* 13, 81–89.
- Paraschiv, S., Otelea, D., Dinu, M., Maxim, D., Tinischi, M., 2007. Polymorphisms and resistance mutations in the protease and reverse transcriptase genes of HIV-1 F subtype Romanian strains. *Int. J. Infect. Dis.* 11, 123–128.
- Paraskevici, D., Magiorkinis, E., Magiorkinis, G., Sympa, V., Paparizos, V., Lazanas, M., Gargalianos, P., Antoniadou, A., Panos, G., Chrysos, G., Sambatakou, H., Karafoulidou, A., Skoutelis, A., Kordossis, T., Koratzanis, G., Theodoridou, M., Daikos, G.L., Nikolopoulos, G., Pybus, O.G., Hatzakis, A., 2007. Increasing prevalence of HIV-1 subtype A in Greece: estimating epidemic history and origin. *J. Infect. Dis.* 196, 1167–1176.
- Passaes, C.P., Bello, G., Lorete, R.S., Matos Almeida, S.E., Junqueira, D.M., Veloso, V.G., Morgado, M.G., Guimaraes, M.L., 2009. Genetic characterization of HIV-1 BC recombinants and evolutionary history of the CRF31\_BC in Southern Brazil. *Infect. Genet. Evol.* 9, 474–482.
- Patrascu, I.V., Constantinescu, S.N., Dublanchet, A., 1990. HIV-1 infection in Romanian children. *Lancet* 335, 672.
- Posada, D., 2008. JModelTest: phylogenetic model averaging. *Mol. Biol. Evol.* 25, 1253–1256.
- Raboni, S.M., Almeida, S.M., Rotta, I., Ribeiro, C.E., Rosario, D., Vidal, L.R., Nogueira, M.B., Riedel, M., Winhescki Mda, G., Ferreira, K.A., Ellis, R., 2010. Molecular epidemiology of HIV-1 clades in Southern Brazil. *Mem. Inst. Oswaldo Cruz* 105, 1044–1049.
- Rambaut, A., 2009. FigTree v1.3.1: Tree Figure Drawing Tool. Available from: <<http://tree.bio.ed.ac.uk/software/figtree/>>.
- Rambaut, A., Drummond, A., 2007. Tracer v1.4. Available from: <<http://beast.bio.ed.ac.uk/Tracer>>.
- Ray, S., Simplot v2.5.0. Available from: <<http://www.med.jhu.edu/deptmed/sray/download/>>.
- Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574.
- Ruta, S., Cernescu, C., 2008. Influence of social changes on the evolution of HIV infection in Romania. *Int. J. Environ. Stud.* 65, 501–513.
- Sabino, E.C., Diaz, R.S., Brigido, L.F., Learn, G.H., Mullins, J.I., Reingold, A.L., Duarte, A.J., Mayer, A., Busch, M.P., 1996. Distribution of HIV-1 subtypes seen in an AIDS clinic in Sao Paulo City, Brazil. *AIDS* 10, 1579–1584.
- Salama, P., Dondero, T.J., 2001. HIV surveillance in complex emergencies. *AIDS* 15, S4–S12.
- Salemi, M., de Oliveira, T., Ciccozzi, M., Rezza, G., Goodenow, M.M., 2008. High-resolution molecular epidemiology and evolutionary history of HIV-1 subtypes in Albania. *PLoS One* 3, e1390.
- Strand, R.T., Fernandes Dias, L., Bergstrom, S., Andersson, S., 2007. Unexpected low prevalence of HIV among fertile women in Luanda, Angola. Does war prevent the spread of HIV? *Int. J. STD AIDS* 18, 467–471.
- Suchard, M.A., Weiss, R.E., Sinsheimer, J.S., 2001. Bayesian selection of continuous-time Markov chain evolutionary models. *Mol. Biol. Evol.* 18, 1001–1013.
- Teixeira, S.L., Bastos, F.I., Telles, P.R., Hacker, M.A., Brigido, L.F., de, F.O.C.A., Bongertz, V., Morgado, M.G., 2004. HIV-1 infection among injection and ex-injection drug users from Rio de Janeiro, Brazil: prevalence, estimated incidence and genetic diversity. *J. Clin. Virol.* 31, 221–226.
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25, 4876–4882.
- Vicente, A.C., Otsuki, K., Silva, N.B., Castilho, M.C., Barros, F.S., Pieniazek, D., Hu, D., Rayfield, M.A., Bretas, G., Tanuri, A., 2000. The HIV epidemic in the Amazon Basin is driven by prototypic and recombinant HIV-1 subtypes B and F. *J. Acquir. Immune Defic. Syndr.* 23, 327–331.
- Vidal, N., Mulanga, C., Bazepeo, S.E., Mwamba, J.K., Tshimpaka, J.W., Kashi, M., Mama, N., Laurent, C., Lepira, F., Delaporte, E., Peeters, M., 2005. Distribution of HIV-1 variants in the Democratic Republic of Congo suggests increase of subtype C in Kinshasa between 1997 and 2002. *J. Acquir. Immune Defic. Syndr.* 40, 456–462.
- Vidal, N., Peeters, M., Mulanga-Kabeya, C., Nzilambi, N., Robertson, D., Ilunga, W., Sema, H., Tshimanga, K., Bongo, B., Delaporte, E., 2000. Unprecedented degree of human immunodeficiency virus type 1 (HIV-1) group M genetic diversity in the Democratic Republic of Congo suggests that the HIV-1 pandemic originated in Central Africa. *J. Virol.* 74, 10498–10507.
- Worobey, M., Gemmel, M., Teuwen, D.E., Haselkorn, T., Kunstman, K., Bunce, M., Muyembe, J.J., Kabongo, J.M., Kalengayi, R.M., Van Marck, E., Gilbert, M.T., Wolinsky, S.M., 2008. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature* 455, 661–664.
- Yang, C., Li, M., Mokili, J.L., Winter, J., Lubaki, N.M., Mwandagaliwa, K.M., Kasali, M.J., Losoma, A.J., Quinn, T.C., Bollinger, R.C., Lal, R.B., 2005. Genetic diversification and recombination of HIV type 1 group M in Kinshasa, Democratic Republic of Congo. *AIDS Res. Hum. Retroviruses* 21, 661–666.