Published in final edited form as:

Infect Genet Evol. 2011 July ; 11(5): 975–979. doi:10.1016/j.meegid.2011.03.009.

Using Phylogeography to Characterize the Origins of the HIV-1 Subtype F Epidemic in Romania

Sanjay R. Mehta, MD^{a,*}, Joel O. Wertheim, PhD^{b,*}, Wayne Delport, PhD^a, Luminita Ene^c, Gratiela Tardei^c, Dan Duiculescu^c, Sergei L. Kosakovsky Pond, PhD^a, and Davey M. Smith, MD, MAS^{a,c}

Sanjay R. Mehta: srmehta@ucsd.edu; Joel O. Wertheim: jwertheim@ucsd.edu; Wayne Delport: wdelport@ucsd.edu; Sergei L. Kosakovsky Pond: spond@ucsd.edu; Davey M. Smith: d13smith@ucsd.edu

^a Department of Medicine, University of California, San Diego, CA 92103

^b Department of Pathology, University of California, San Diego, CA 92103

^c Victor Babes Hospital for Infectious Diseases, Bucharest Romania

^d Veterans Affairs San Diego Healthcare System, San Diego, CA 92161

Abstract

Background—During the late 1980s and early 1990s, an estimated 10,000 Romanian children were infected with HIV-1 subtype F nosocomially through contaminated needles and blood transfusions. However, the geographic source and origins of this epidemic remain unclear.

Methods—Here we used phylogenetic inference and “relaxed” molecular clock dating analysis to further characterize the Romanian HIV-1 subtype F epidemic.

Results—These analyses revealed a major lineage of Romanian HIV sequences consisting nearly entirely of virus sampled from adolescents and children and a distinct cluster that included a much higher ratio of adult sequences. Divergence time estimates inferred the time of most recent common ancestor of subtype F1 sequences to be 1973 (1966–1980) and for all Angolan sequences to 1975 (1968–1980). The most common ancestor of the Romanian sequences was dated to 1978 (1972–1983) with pediatric and adolescent sequences interspersed throughout the lineage. The phylogenetic structure of the entire subtype F epidemic suggests that multiple introductions of subtype F into Romania occurred either from the Angolan epidemic or from more distant ancestors. Since the historical records note that the Romanian pediatric epidemic did not begin until the late 1980s, the inferred time of most recent common ancestor of the Romanian lineage of 1978 suggests that there were multiple introductions of subtype F occurred into the pediatric population from HIV already circulating in Romania.

Conclusions—Analysis of the subtype F HIV-1 epidemic in an historical context allows for a deeper appreciation of how the HIV pandemic has been influenced by socio-political events.

© 2011 Elsevier B.V. All rights reserved.

Correspondence: Sanjay R. Mehta MD, University of California, San Diego, 9500 Gilman Drive 0711, La Jolla, CA 92093-0711 USA, Tel: +1 858 822 4092, Alt tel: +1 858 822 0333, Fax: +1 858 822 5362, srmehta@ucsd.edu.

*These authors contributed equally to this work

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

Phylogeography; Romania; Subtype F; Socio-political; HIV

Introduction

The poor fidelity of HIV-1's reverse transcriptase leads to significant diversity within individuals and in populations over time¹. Although this diversity has impeded the progress of a vaccine against HIV, it can be used to reconstruct the evolutionary history of extant viruses. Reconstruction of the molecular evolutionary history and the spatial diffusion of a virus, or phylogeography, can provide clues to the dynamics of viral outbreaks². These insights could direct intervention and prevention strategies to reduce transmission of ongoing outbreaks, or prevent future ones. Phylogeographic analysis of the pandemic lineage of HIV-1 group M suggests that the origin of this epidemic lies in central West Africa³. Group M is classified into nine subtypes (A–D, F–H, J, and K) and a variety of circulating and unique recombinant forms which often are concentrated in specific geographic areas⁴. Of the nine subtypes, subtype F causes less than 1% of all HIV-1 infections globally⁵ and makes up less than 1% of sampled HIV infections in Africa⁴. However, more than 99% of sampled HIV infections in Romania are of descendants of a major lineage within subtype F, known as F1⁴.

The source of the subtype F epidemic in Romania has been debated^{6–9}, although there is some understanding as to how it spread¹⁰. The first reported case of HIV in Romania was published in 1985, and by December of 1988 only 22 additional cases had been reported¹¹. By late 1989, it was apparent that there was a burgeoning pediatric outbreak of HIV taking place¹¹, however, the extent of the outbreak remained unknown. After the overthrow of the Ceausescu regime in 1989, Romanian physicians and researchers were able to perform epidemiologic studies^{9–12} which revealed that one of the world's largest nosocomial disasters was taking place¹¹. During the late 1980s and early 1990s, an estimated 10,000 Romanian children were infected^{8–10} with predominantly subtype F HIV-1^{6, 13}, presumably through contaminated needles and blood products at orphanages and pediatric hospitals.

Early molecular epidemiology studies identified a shared ancestry between subtype F sequences in Romania and Brazil¹³, but as more sequences became available, subsequent phylogenetic analyses have suggested that the Romanian epidemic stemmed from Angola⁷. It is unclear whether the Romanian pediatric epidemic arose as the result of expansion from a single founder event (suggested by the relatively low diversity within the epidemic¹⁴), or as a result of multiple introductions of subtype F from the Romanian HIV infected population into the blood supply and injectable medicines used to treat children⁸. Here we use phylogeographic methods to evaluate a large dataset of subtype F *pol* sequences to reconstruct the global evolutionary history of subtype F and to determine how it became the predominant subtype of HIV-1 present in Romania.

Methods

Maximum likelihood phylogenetic inference

We downloaded all available HIV-1 Group M subtype F1 and F2 *pol* sequences with length greater than 900 nucleotides sampled between 1987 and 2007 from the Los Alamos National Laboratory (LANL) HIV sequence database (www.hiv.lanl.gov)⁴ on 4/1/2010. Recombinant sequences were identified using the Subtype Classification Using Evolutionary Algorithms (SCUEAL) application¹⁵ and removed. To this collection of 531 subtype F sequences (see

Supplementary Table 1 for details), we included additional HIV-1 Group M subtype sequences (one sequence per subtype per sampling year), which resulted in an alignment of 606 HIV sequences; recombinant subtypes (including subtype G) were not included¹⁶. Subject antiretroviral treatment status was unknown for the majority of sequences. Alignment was trivial and performed manually (available from the authors upon request). The maximum likelihood phylogeny was inferred using PHYML 3.0 using a sub-tree pruning and re-grafting algorithm under a generalized time-reversible (GTR) + Γ_4 substitution model^{17, 18}. Node support was assessed using an approximate likelihood-ratio test¹⁹.

Dating analysis

To estimate the time of most recent common ancestor (tMRCA) for groups within subtype F, we employed a Bayesian Markov chain Monte Carlo (BMCMC) framework implemented in BEAST v1.5.3²⁰. To create a dataset suitable for BMCMC analysis, which unlike PHYML cannot easily accommodate hundreds of sequences, we pruned our dataset. First, non-subtype F sequences were excluded. Second, heavily sampled Romanian and Brazilian lineages were sub-sampled to include a maximum of five sequences per sampling country per year. Sequences associated with the pediatric outbreak in Romania were kept in the alignment. This procedure produced an alignment containing 156 sequences sampled between 1987 and 2007.

Four independent replicate BMCMC analyses of 50 million generations (5 million burn-in) were performed under a GTR + Γ_4 substitution model. Tracer v1.5 was used to check for mixing and convergence with an estimated sample size > 200 for all parameters. Three different coalescent demographic scenarios were explored: constant population size, exponential population growth, and a Bayesian skyline plot, which places the least prior constraint on the demography. All analyses were performed using an uncorrelated lognormal relaxed molecular clock. The posterior trees were summarized using the maximum clade credibility tree. We identified tMRCAs on highly supported nodes using mean and 95% highest posterior density (HPD) values.

In addition, to better appreciate the evolutionary dynamics of the Romanian epidemic, BMCMC analysis was also performed on subtype F sequences from Romania. To obtain a computationally tractable dataset, a maximum of 10 sequences per sampling year were included in the analysis. Two independent replicate BMCMC analyses of 25 million generations (2.5 million burn-in) were performed under an HKY + Γ_4 substitution model using a Bayesian skyline plot. An uncorrelated lognormal relaxed molecular clock was implemented.

Results

Phylogenetic reconstruction of subtype F1 pol sequence relationships

The maximum likelihood tree (Figure 1) and the Bayesian maximum clade credibility tree (Supplemental Figure 1) had similar topologies. Within the Romanian subtype F1 lineage, adolescent and pediatric sequences were found interspersed with adult sequences in the other Romanian clusters. The nesting of an Angolan lineage within the entire Romanian lineage (Figure 1 – shown in red) is in contrast to previous work reporting reciprocal monophyly between subtype F1 sequences from Angola and Romania⁷. These Angolan sequences, which are nested in the Romanian lineage make up the main Angolan lineage. Other sampled Angolan sequences were found deep within the F1 lineage suggesting that either Angola was a repeated source of the Romanian F1 epidemic, or that multiple introductions of F1 back into Angola occurred early in the HIV-1 pandemic. Of note, a

Cuban F1 sequence was also closely related to the Angolan sequences (Figure 1 – shown in orange).

The Brazilian subtype F1 sequences clustered separately from the Romania-Angola group. Interspersed within this Brazilian cluster were sequences from nine other nations, including Italy (7), Switzerland (2) and one each from France, Luxembourg, Belgium, Portugal, Japan, Argentina and Finland (Figure 1 – shown in green). Our phylogenetic analysis included subtype F2 sequences, predominantly sampled from Cameroon, and confirmed that the F2 lineage is reciprocally monophyletic to the F1 lineage.

Time to most recent common ancestor dating

Reconstruction of the sub-sampled subtype F phylogeny using a BMCMC approach (Supplemental Figure 1) allowed us to infer the tMRCA for many nodes in the subtype F tree. As part of this analysis, we explored the fit of three population size models: exponential growth, constant population size, and a Bayesian skyline plot. The Bayesian skyline plot showed a dramatic increase in population size of the subtype F HIV-1 epidemic in the late 1970s and 1980s followed by a plateau, which indicated that neither a constant population size nor an exponential population size model was appropriate (Supplemental Figure 2). The tMRCAs inferred using the skyline model were generally in accordance with historical documentation^{3, 7, 11, 21–23} and reported here (Table 1). These dates were not markedly different from those obtained using constant population size and exponential population size models (Supplementary Table 2).

The tMRCA inferred for all of the subtype F sequences was dated to 1967 (95% HPD: 1956–1976), whereas the MRCA for subtype F1 dated to 1973 (95% HPD: 1966–1980). The tMRCA for the Angolan subtype F1 epidemic was dated to 1975 (95% HPD: 1968–1980). The entire Romanian lineage had a tMRCA at 1978 (95% HPD: 1972–1983). According to historical accounts, the Romanian pediatric epidemic took place in the late 1980s and early 1990s¹¹. A Bayesian skyline encompassing only subtype F Romanian sequences is consistent with this historical observation, indicating an exponential expansion of the epidemic between 1985 and 1990 (Figure 2). A sub-cluster of the Romanian F1 lineage that was found to contain sequences that were nearly entirely of pediatric origin had a tMRCA of 1984 (95% HPD: 1980–1988), more consistent with the dates of the nosocomial epidemic; however, these sequences were sampled from only two locations in Romania. When the diversity of all the pediatric and adolescent nosocomial isolates was examined, it encompassed all of the Romanian subtype F diversity, suggesting that the nosocomial outbreak was seeded by several circulating strains and not a single source sometime in the late 1980s.

We also inferred the tMRCA of these same lineages using the sub-sampled subtype F sequences and additional HIV-1 reference sequences (i.e., 226 sequences). BMCMC analysis using this dataset produced a tMRCA of HIV-1 group M that was in accordance with previous estimates³. However, the tMRCA estimates for divergence events within subtype F were substantially older than those inferred from the sub-sampled subtype F-only analysis (Supplementary Table 3) and not in agreement with historical records. For example, the tMRCA of the Romanian lineages was estimated to be 1955 (95% HPD: 1942–1966). This discrepancy can likely be accounted for by an increased substitution rate in subtype F: 3.15×10^{-3} substitutions/site/year (95% HPD: $2.45 \times 10^{-3} - 3.88 \times 10^{-3}$) for subtype F compared to 1.76×10^{-3} substitutions/site/year (95% HPD: $1.31 \times 10^{-3} - 2.22 \times 10^{-3}$) for HIV-1 group M. Alternatively, this observed discrepancy in substitution rates might be due to an imbalance in the number of representative sequences from each subtype and differences in the demography of the sampled viruses in each subtype. Although we present

these data to be complete, we have little confidence in the tMRCA's produced by the analysis on the full HIV-1 dataset.

Discussion

Africa

Our phylogenetic analyses demonstrate that the migration of HIV-1 subtype F around the world may be more complicated than previously thought. The bulk of sampled Angolan F1 sequences forms a major lineage that clusters together, while other sampled sequences lie deep within the F1 lineage. Taken together, our dating analyses of these sequences suggest that the tMRCA of subtype F in Angola is 1975 (95% HPD: 1968–1980). A recent study by Guimaraes et al. suggested that sampled subtype F *env* sequences from Romania were reciprocally monophyletic with sequences from Angola⁷, but our analysis of *pol* revealed an Angolan sub-cluster segregating within the Romanian sequences. The finding of a monophyletic subset of Angolan sequences with a tMRCA of 1984 (95% HPD: 1980–1988) within the Romanian sequences could suggest that an Angolan sub-epidemic stemmed from the Romanian epidemic. A more historically consistent interpretation, however, would be that several introductions of HIV-1 subtype F occurred from Angola to Romania, and that the Angolan epidemic has not been sampled as densely as the Romanian epidemic. We also found that a Cuban isolate clustered within Angolan sequences, which suggests that the Angolan HIV-1 epidemic is the source of both the Romanian and at least part of the Cuban HIV-1 subtype F epidemics. Interestingly, this is consistent with the strong military ties between all three of these nations between 1975 and 2002^{24,25}. We speculate that the Angolan civil war played a major role in the dissemination of subtype F into Cuba and Romania.

North and South America

Recent studies have estimated the date of introduction of subtype F1 into South America to be in the late 1970s²⁶ or early 1980s²² and likely originating as a single introduction in southeast Brazil^{21, 22}. Close socio-political ties between Brazil and Angola during the late 1970s to early 1990s fueled speculation that the F1 epidemic in Brazil stemmed from transmissions in relation to Angola⁴. In fact, from the mid 1970s to the early 1990s, an estimated 15,000 Angolans fleeing the civil war settled in Brazil²⁷. However, recent phylogenetic analyses from this study and others⁷ do not explicitly support this scenario. A closer look at our phylogenetic tree reveals sequences from several western European nations (as well as Argentina and Japan) nested within this Brazilian cluster that extend to the base of the lineage. These interspersed sequences, of which Italian sequences make up the majority, suggest that subtype F virus most likely migrated frequently between these nations.

Our analysis of North American isolates suggests that at least part of the Cuban epidemic was also derived from the Angolan epidemic. Specifically, of the two available Cuban subtype F *pol* sequences, one falls within an Angolan lineage, and the other appears to be situated more basally in the F1 clade. In the historical context, there was increasing Cuban involvement with Angola beginning in the 1960s as part of Fidel Castro's attempts to bring Marxism to Africa²⁴. Additionally, we found a Mexican subtype F1 isolate that was also closely related to Angolan sequences.

Europe

Although HIV-1 subtype B is the predominant circulating virus in Europe⁴, subtype F1 sequences have been sampled in several European nations (including Russia, Bulgaria, and Belgium). Despite these examples, Romania has the highest prevalence of subtype F in

Europe. Efforts to pinpoint the source of the Romanian subtype F1 epidemic have been contradictory, with early work suggesting a relationship to the Brazilian epidemic¹³ and more recent work suggesting a single transmission event between the Romanian and Angolan epidemics⁷. Our analysis suggests a more complex scenario in which there were several transmission events of subtype F1 between Angola and Romania. The Romanian F1 sequences form a paraphyletic group, containing a lineage of predominantly Angolan sequences (Figure 1). Socio-political ties between Angola and Romania increased in the 1970s during the Angolan civil war^{23–25}, which corresponds with the inferred tMRCA for all for all Romanian subtype F1 sequences in 1978 (95% HPD: 1972–1983). Although it is likely that Angola was the source of the original Romanian epidemic, the direction of subsequent transmissions between these nations remains obscure.

Although widespread sampling did not occur in Romania until 1990, only a limited number of cases of HIV were detected in the preceding years, suggesting that an explosion in the subtype F1 epidemic occurred in the late 1980s and early 1990s, in agreement with our molecular dating and demographic analyses. During this period, many Romanian children were being cared for in institutions, because governmental policy had led to a dramatic increase in children whose families were unable to care for them^{9, 12, 23}. The 10 year delay between the inferred tMRCA of the entire Romanian subtype F epidemic and the observed Romanian nosocomial pediatric HIV epidemic suggests that circulating subtype F virus was responsible for multiple introductions of HIV into the blood supply or the injectables provided to the institutionalized children, the presumed cause of the Romanian nosocomial adolescent and pediatric HIV-1 subtype F1 epidemic^{8, 9, 12, 23}. Although we identified a sub-cluster of Romanian sequences that was made up almost entirely of pediatric sequences with a tMRCA dated to 1984 (95% HPD: 1980–1988), other pediatric and adolescent sequences were interspersed throughout the rest of the Romanian lineage providing further evidence for the multiple introduction theory. These results are consistent with the findings of Op de Coul et al.⁸ who, using a much smaller dataset, also found that the Romanian HIV subtype F1 pediatric epidemic was not the result of a single point source infection, but rather multiple transmissions.

Although the Los Alamos database provides the most extensive dataset of subtype F *pol* sequences available, there is inherent sampling bias in using this dataset. Sequences within this dataset are mostly from a handful of studies each of which used different sampling techniques, and occurred at different times. Thus, our dataset might not adequately represent the temporal, geographic, and demographic characteristics of the subtype F HIV-1 epidemic.

Summary

The unique worldwide distribution of HIV-1 subtype F has allowed us to use phylogeography to trace its spread from its central west African origins and correlate this spread with historical events. Although the histories of the more rampant HIV-1 subtype B and subtype C epidemics are more complex, socio-political relationships between individuals and nations likely guided those epidemics as well. Analysis of the subtype F HIV-1 epidemic in the context of history provides an improved understanding of how HIV sub-epidemics can be influenced by socio-political events.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We appreciate the assistance of Cris Achim MD, PhD on this manuscript.

This work was supported by grants from the National Institutes of Health and the National Science Foundation: (SRM) 1R01AI087164-01, P30 AI036214; (SLP) Joint DMS/NIGMS Mathematical Biology Initiative through Grant NSF-0714991, the National Institutes of Health (AI47745), by the National Institutes of Health through grants AI077304, AI36214, AI047745 and AI074621;(JOW) NIH-AI07384; (DMS) AI69432, AI043638, MH62512, MH083552, AI077304, AI36214, AI047745, AI74621, AI080353, and the James B. Pendleton Charitable Trust.

References

1. Holmes EC. Error thresholds and the constraints to RNA virus evolution. *Trends Microbiol.* Dec; 2003 11(12):543–546. [PubMed: 14659685]
2. Lemey P, Rambaut A, Drummond AJ, Suchard MA. Bayesian phylogeography finds its roots. *PLoS Comput Biol.* Sep.2009 5(9):e1000520. [PubMed: 19779555]
3. Worobey M, Gemmel M, Teuwen DE, et al. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature.* Oct 2; 2008 455(7213):661–664. [PubMed: 18833279]
4. HIV Sequence Database. Los Alamos National Laboratory; <http://www.hiv.lanl.gov>
5. Hemelaar J, Gouws E, Ghys PD, Osmanov S. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS.* Oct 24; 2006 20(16):W13–23. [PubMed: 17053344]
6. Apetrei C, Necula A, Holm-Hansen C, et al. HIV-1 diversity in Romania. *AIDS.* Jun 18; 1998 12(9):1079–1085. [PubMed: 9662205]
7. Guimaraes ML, Vicente AC, Otsuki K, et al. Close phylogenetic relationship between Angolan and Romanian HIV-1 subtype F1 isolates. *Retrovirology.* 2009; 6:39. [PubMed: 19386115]
8. Op De Coul E, van den Burg R, Asjo B, et al. Genetic evidence of multiple transmissions of HIV type 1 subtype F within Romania from adult blood donors to children. *AIDS Res Hum Retroviruses.* Mar 1; 2000 16(4):327–336. [PubMed: 10716370]
9. Patrascu IV, Dumitrescu O. The epidemic of human immunodeficiency virus infection in Romanian children. *AIDS Res Hum Retroviruses.* Jan; 1993 9(1):99–104. [PubMed: 8427718]
10. Hersh BS, Popovici F, Jezek Z, et al. Risk factors for HIV infection among abandoned Romanian children. *AIDS.* Dec; 1993 7(12):1617–1624. [PubMed: 8286071]
11. Hersh BS, Popovici F, Apetrei RC, et al. Acquired immunodeficiency syndrome in Romania. *Lancet.* Sep 14; 1991 338(8768):645–649. [PubMed: 1679471]
12. Patrascu IV, Constantinescu SN, Dublanchet A. HIV-1 infection in Romanian children. *Lancet.* Mar 17.1990 335(8690):672. [PubMed: 1969055]
13. Bandea CI, Ramos A, Pieniazek D, et al. Epidemiologic and evolutionary relationships between Romanian and Brazilian HIV-subtype F strains. *Emerg Infect Dis.* Jul-Sep; 1995 1(3):91–93. [PubMed: 8903171]
14. Dumitrescu O, Kalish ML, Kliks SC, Bandea CI, Levy JA. Characterization of human immunodeficiency virus type 1 isolates from children in Romania: identification of a new envelope subtype. *J Infect Dis.* Feb; 1994 169(2):281–288. [PubMed: 8106761]
15. Kosakovsky Pond SL, Posada D, Stawiski E, et al. An evolutionary model-based algorithm for accurate phylogenetic breakpoint mapping and subtype prediction in HIV-1. *PLoS Comput Biol.* Nov.2009 5(11):e1000581. [PubMed: 19956739]
16. Abecasis AB, Lemey P, Vidal N, et al. Recombination confounds the early evolutionary history of human immunodeficiency virus type 1: subtype G is a circulating recombinant form. *J Virol.* Aug; 2007 81(16):8543–8551. [PubMed: 17553886]
17. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol.* Oct; 2003 52(5):696–704. [PubMed: 14530136]
18. Hordijk W, Gascuel O. Improving the efficiency of SPR moves in phylogenetic tree search methods based on maximum likelihood. *Bioinformatics.* Dec 15; 2005 21(24):4338–4347. [PubMed: 16234323]
19. Anisimova M, Gascuel O. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Syst Biol.* Aug; 2006 55(4):539–552. [PubMed: 16785212]
20. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol.* 2007; 7:214. [PubMed: 17996036]

21. Bello G, Eyer-Silva WA, Couto-Fernandez JC, et al. Demographic history of HIV-1 subtypes B and F in Brazil. *Infect Genet Evol.* Mar; 2007 7(2):263–270. [PubMed: 17150415]
22. Bello G, Guimaraes ML, Morgado MG. Evolutionary history of HIV-1 subtype B and F infections in Brazil. *AIDS.* Mar 21; 2006 20(5):763–768. [PubMed: 16514307]
23. Rothman, SM.; Rothman, DJ. *The New York Review of Books.* November 8. 1990 How AIDS Came to Romania.
24. Spikes, D. *Angola and the Politics of Intervention.* Jefferson: McFarland & Company; 1993.
25. Saferworld. Saferworld. Saferworld.org.uk. London: Saferworld; 2002. Arms Production, exports and decision -making in Central and Eastern Europe.
26. Aulicino PC, Bello G, Rocco C, et al. 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.* Oct; 2007 23(10):1176–1182. [PubMed: 17961101]
27. Baeninger. *Africanos no Brasil: dubiedade e estereótipos.* Brasil: Migracoes internacionais e identidade; 2000. <http://www.comciencia.br/reportagens/migracoes/migr11.htm>

Figure 1.

Maximum likelihood phylogeny depicting the relationships between subtype F sequences. Large clades are collapsed into triangles for ease of viewing. Romanian sequences are shown in blue, Angolan in red, Cuban in orange, Brazilian in green, and F2 in purple. Sequences from other countries are shown in black. Other HIV-1 group M subtypes are not depicted. Nodes with support greater than 0.70 are identified with an asterisk.

Figure 2. Bayesian skyline plot for Romanian subtype F sequences sampled between 1993 and 2007. Median effective population size is shown in black, and 95% HPD intervals are shown in gray.

Table 1

Time of most recent common ancestor (tMRCA) estimates for subtype F lineages.

Taxon	Mean tMRCA	(95% HPD)
Subtype F	1967	(1956–1976)
F1	1973	(1966–1980)
F2	1979	(1973–1985)
Angolan subtype F	1975	(1968–1980)
Romanian subtype F	1978	(1972–1983)
Brazilian subtype F	1980	(1975–1985)