

Extremely Rapid Spread of Human Immunodeficiency Virus Type 1 BF Recombinants in Argentina[∇]

Paula C. Aulicino,^{1†} Edward C. Holmes,^{2,3†} Carlos Rocco,¹ Andrea Mangano,¹ and Luisa Sen^{1*}

Laboratorio de Biología Celular y Retrovirus-CONICET, Hospital de Pediatría J. P. Garrahan, Buenos Aires, Argentina¹; Center for Infectious Disease Dynamics, Department of Biology, the Pennsylvania State University, University Park, Pennsylvania 16802²; and Fogarty International Center, National Institutes of Health, Bethesda, Maryland 20892³

Received 4 July 2006/Accepted 10 October 2006

The epidemic of human immunodeficiency virus type 1 (HIV-1) in Argentina is distinctive in that many infections are caused by subtype BF recombinant viruses. To determine their demographic history, we estimated the evolutionary rate, mode of population growth, and age of genetic diversity among 40 BF *vpu* sequences. This revealed one of the highest substitution rates reported for HIV-1, at 10.793×10^{-3} substitutions per site per year, and a very rapid rate of population growth, with an initial mean epidemic doubling time of 3.72 months. This rapid population growth is compatible with an elevated fitness for subtype BF compared to that for “pure” B and F viruses.

The number of people living with human immunodeficiency virus (HIV) in Latin America is estimated to have reached 1.8 million (15). After Brazil, Argentina has the highest number of HIV-1 infections in the region, with the first AIDS case diagnosed in 1982 (5). Similar to what was found in North America and Europe, the AIDS infections reported in Argentina until 1985 were in young homosexual or bisexual men, with the first case of pediatric AIDS documented in 1986 (10). Almost simultaneously, adult female HIV-1 infection, mainly through heterosexual contact and/or intravenous drug use, was reported. The number of HIV-1-infected children then increased rapidly, reaching a peak in incidence during 1991 to 1996. Although the administration of zidovudine has reduced the rate of mother-to-child transmission (MTCT), MTCT still accounts for 7.4% of all HIV-1 infections in Argentina, a much higher rate than those observed in developed countries.

Initially, the genetic composition of HIV-1 in Argentina was thought to resemble that of HIV-1 in North America, with a predominance of subtype B strains (7). However, full-length genome sequencing led to the identification of a variety of subtype BF recombinants (2, 14), with the circulating recombinant form CRF12_{BF} representing nearly 25% of those described. Currently, approximately 80% of HIV-1 isolates from Argentina are BF recombinants, with almost all the remaining 20% constituting subtype B viruses and scattered cases of subtypes A, C, BC, and F (1, 3).

To infer the demographic history of the BF-associated HIV-1 epidemic in Argentina, we applied Bayesian coalescent methods to a data set of 40 sequences of nearly 500 bp with a conserved BF mosaic pattern, comprising the *vpu* gene and the 5' end of *env*. The sequences were obtained after proviral DNA amplification (6) from peripheral blood mononuclear

cells from 79 HIV-1-infected children born between 1986 and 2003 and followed up at the Garrahan Pediatric Hospital, Buenos Aires, Argentina. Samples were from the 1997-to-2004 period, with informed consent obtained in all cases. Sequences were aligned using Clustal X (13). The HIV-1 subtype was investigated by neighbor-joining phylogenetic analysis incorporated in the MEGA3 package (9), and recombination breakpoints were identified by boot-scanning analysis using Simplot (version 2.5) with a 200-bp sliding window and a step size of 50 bp. Of the total 79 *vpu* sequences obtained, 59 were BF recombinants. Among them, 35 shared a single B-to-F recombination breakpoint homologous to CRF12_{BF}, located at position 6230 of HXB-2 (GenBank accession number K03455). Five additional HIV-1 *vpu* sequences from Argentina with the same characteristics as CRF12_{BF} were collected from the Los Alamos HIV Sequence Database (<http://hiv-web.lanl.gov>), resulting in a final data set of 40 BF HIV-1 *vpu* sequences. The monophyletic clustering of this data set in phylogenetic trees confirmed their common ancestry, while separate trees constructed on each side of the breakpoint confirmed their mixed B and F parentage (these data are available from the authors upon request). In each case, the exact day of sampling was recorded, ranging from 11 May 1997 to 22 April 2004.

Rate of nucleotide substitution, mode and rate of population growth, and age of most recent common ancestor were estimated using a Bayesian MCMC (Markov Chain Monte Carlo) method implemented in the BEAST program (<http://evolve.zoo.ox.ac.uk/beast/>) (4). Four models of demographic history were compared—constant population size, exponential growth, logistic growth, and expansion growth—assuming either a constant (strict) or a variable (relaxed) molecular clock. In each case, we employed the general time-reversible model of nucleotide substitution incorporating invariant sites and a gamma distribution for among-site rate variation (with four rate categories). Akaike's information criterion was used to determine which model best fit the data, with uncertainty in parameter estimates reflected in the 95% highest probability density (HPD) values. All runs were checked for convergence by using the TRACER program (<http://evolve.zoo.ox.ac.uk/software>

* Corresponding author. Mailing address: Laboratorio de Biología Celular y Retrovirus, Hospital de Pediatría Garrahan, Combate de los Pozos 1881, (1245) Capital Federal, Argentina. Phone: 54 11 4308 1998. Fax: 54 11 4308 5325. E-mail: lsen@garrahan.gov.ar.

† Both authors contributed equally to this work.

∇ Published ahead of print on 18 October 2006.

TABLE 1. Bayesian estimates of population dynamics and evolutionary parameters of HIV-1 subtype BF in Argentina

Parameter	Estimate
Sample size.....	40
Sample date range.....	11 May 1997–22 April 2004
Best-fit demographic model.....	Logistic growth (relaxed molecular clock)
MCMC chain length	70,000
Mean substitution rate ^a	10.793×10^{-3}
95% HPD substitution rate.....	3.318×10^{-3} – 16.093×10^{-3}
Mean age (yr)	11.350
95% HPD age.....	7.255–21.981
Mean population growth rate ^b	2.237
95% HPD population growth rate.....	0.207–4.559
Mean epidemic doubling time (mo).....	3.720
95% HPD epidemic doubling time.....	1.824–40.188

^a Number of substitutions per site per year.

^b Number of new infections per individual per year.

.html?id=tracer). The epidemic doubling time (λ) in years was calculated using the relation $\lambda = \ln(2)/r$, where r is the population growth rate estimated in BEAST.

The best-fit model of demographic history for subtype BF in Argentina was that of logistic population growth under a relaxed molecular clock (Table 1 and data not shown). In this model, an initially rapid phase of exponential growth is followed by a reduction in growth rate toward the present, as reflected in the branching pattern of the maximum a posteriori tree of these data (Fig. 1). The mean rate of population growth during the initial growth phase was 2.237 new infections per individual per year, which equates to a mean epidemic doubling time of 3.72 months (but with a large associated error). Although logistic population growth has been observed for subtype B in industrialized nations (11), this rate is the highest

reported for any HIV-1 subtype by using coalescence methods. More specifically, previous coalescent analyses of *env*, *gag* and reverse transcriptase genes revealed growth rates of 0.386 to 0.609 year⁻¹ for the parenterally transmitted subtype B population of Western Europe and North America during the 1980s (16), roughly one-fifth of the rate observed here. However, despite this initially high rate of exponential population growth, BF recombinants seem to have reached the population carrying capacity in relation to MTCT, as opposed to subtypes B and C in adults from Brazil, which may still be spreading at exponential rates (12).

A demographic pattern of explosive population growth followed by a “steady state” in the number of infections was also reflected in epidemiological data for MTCT in Argentina. Data collected since the start of the HIV-1 epidemic by the National Ministry of Health (10) show that the incidence of HIV infection in children due to vertical transmission has the same dynamics as the growth of the “CRF12-like” viral population, with an initial exponential growth followed by a plateau in 1991 (Fig. 2). In the last 5 years, however, a decline in the rate of MTCT is apparent, most likely related to public health policies offering HIV-1 testing to all pregnant women. Overall, these data suggest that the BF recombinants have played an important part in shaping the demographic history of HIV-1 in the pediatric population.

A second notable feature of the spread of the HIV-1 BF recombinants in Argentina is its high rate of evolutionary change; the rate for our “CRF-like” *vpv* BF recombinant data set was 10.793×10^{-3} nucleotide substitutions per site per year. This is among the highest evolutionary rates estimated for HIV-1 to date (17), although the lower HPD value encompasses the values obtained in other studies (8). Whether these differences in substitution rate are due to differences in the

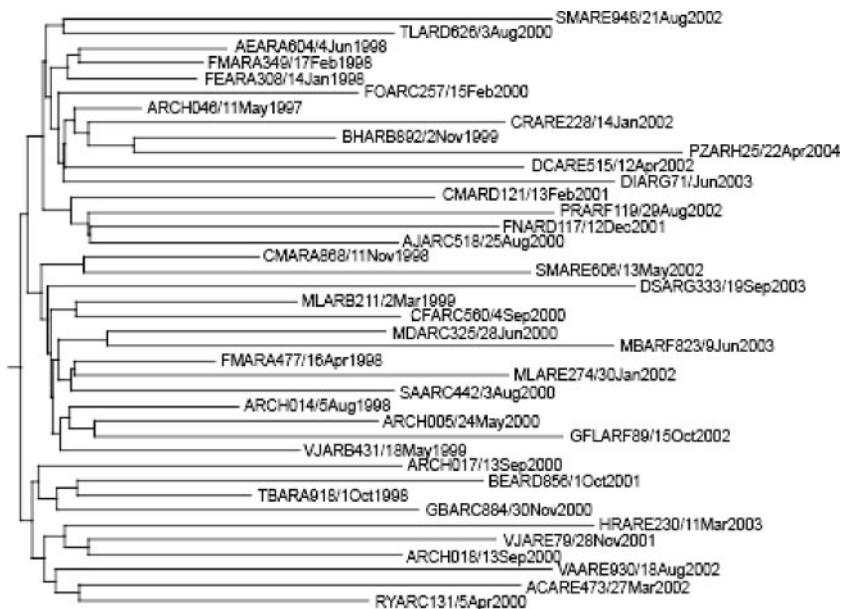


FIG. 1. Maximum a posteriori phylogenetic tree of 40 CRF12_BF sequences from Argentina comprising the *vpv* gene and the 5' end of the *env* gene. The precise day of sampling is given in the sequence identifier of each viral isolate. For isolate DIARG71/Jun2003, the exact day of sampling was unavailable so the mid-month date of 15 June was chosen to minimize error. In all cases, tip times correspond to the dates of virus sampling.

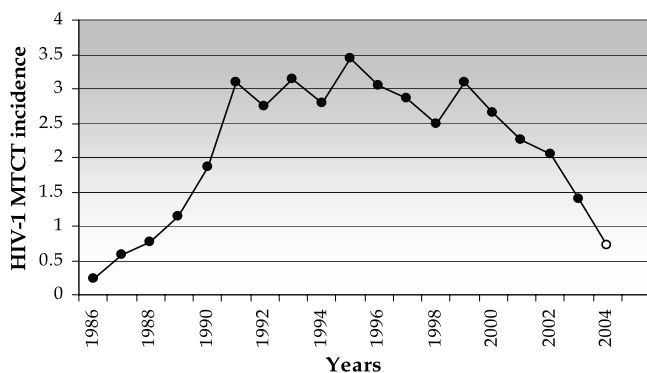


FIG. 2. Incidence of MTCT in Argentina from 1986 to 2004. Incidence is expressed as the number of new HIV-1-infected newborns/10,000 children born alive. The source was the National AIDS Bulletin of Argentina (2005) (10). The data for 2004 (open circle) are preliminary.

gene regions studied, the incorporation of a relaxed molecular clock, or the use of daily, rather than yearly, sampling is unclear. However, although the inferred substitution rate is high, the estimated age of the most recent common ancestor of the viruses under this substitution rate—11.350 years according to the most recent sample, from 1992 (95% HPD = 1981 to 1996)—is compatible with the “known” history of the “CRF12-like” BF recombinants in Argentina, in which pediatric AIDS and BF recombinant strains in children were first reported in 1986 (6).

To our knowledge, this is the first time that the population dynamics of a recombinant HIV-1 population have been analyzed. That the initial spread of the BF variant, predominantly by mother-to-child transmission, was so rapid indicates that this recombinant strain was also one of high fitness, such that it was transmitted at a high frequency when it entered the susceptible Argentine population. These rapid population dynamics are especially notable in that the BF recombinant is now sampled more frequently than the subtype B viruses in Argentina, and only one pure HIV-1 subtype F infection has been documented in this country (1). As BF recombinants are present at appreciable frequencies in most Latin American countries, determining the factors that have facilitated their rapid spread is of utmost importance.

Nucleotide sequence accession numbers. The data set of 40 BF recombinant *vpu* sequences used here comprises 35 new sequences (GenBank accession numbers AY284979.1 to AY284988.1, AY284992.1, and DQ767694 to DQ767717) and five collected from the Los Alamos sequence database (accession numbers AF454481, AF454482, AF454485, AF454487, and AY037266).

This work was supported in part by Agencia Nacional de Promoción Científica y Tecnológica award no. 990506326 and CONICET PIP no. 6057/05.

We gratefully thank Carmen Gálvez, Rodrigo Cánepa, and Andrés Peralta for technical assistance.

REFERENCES

- Aulicino, P. C., J. Kopka, A. M. Mangano, C. Rocco, M. Iacono, R. Bologna, and L. Sen. 2005. Novel HIV-1 subtypes A, B/C and F circulate in Argentina. *AIDS Res. Hum. Retrovir.* **21**:158–164.
- Carr, J. K., M. Avila, M. Gomez Carrillo, H. Salomon, J. Hierholzer, V. Watanaveradej, M. A. Pando, M. Negrete, K. L. Russell, J. Sanchez, D. L. Birs, R. Andrade, J. Vinales, and F. E. McCutchan. 2001. Diverse B/F recombinants have spread widely since the introduction of HIV-1 into South America. *AIDS* **15**:F41–F47.
- Carrion, G., L. Eyzaguirre, S. R. Montano, V. Laguna-Torres, M. Serra, and N. Aguayo. 2004. Documentation of subtype C HIV type 1 strains in Argentina, Paraguay, and Uruguay. *AIDS Res. Hum. Retrovir.* **20**:1022–1025.
- Drummond, A. J., S. Y. Ho, M. J. Phillips, and A. Rambaut. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4**:e88.
- Estevez, M. E., S. Bruno, L. Sen, C. Scaglione, R. A. Diez, and A. M. Musso. 1983. Síndrome de inmunodeficiencia adquirida (AIDS) con sarcoma de Kaposi en homosexuales en la Argentina. *Medicina (Buenos Aires)* **43**:477.
- Gomez Carrillo, M., M. Avila, J. Hierholzer, M. Pando, P. L. Martinez, F. E. McCutchan, and J. K. Carr. 2002. Mother-to-child HIV type 1 transmission in Argentina: BF recombinants have predominated in infected children since the mid-1980s. *AIDS Res. Hum. Retrovir.* **18**:477–483.
- Gomez Carrillo, M., C. Piccardo, and O. Libonatti. 1992. Molecular analysis of the principal neutralization epitope (V3 loop) of human immunodeficiency virus type 1 in Argentina. *Rev. Argent. Microbiol.* **24**:91–101.
- Hué, S., D. Pillay, J. Clewley, and O. Pybus. 2005. Genetic analysis reveals the complex structure of HIV-1 transmission within defined risk groups. *Proc. Natl. Acad. Sci. USA* **102**:4425–4429.
- Kumar, S., K. Tamura, and M. Nei. 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief. Bioinform.* **5**:150–163.
- Ministerio de Salud y Ambiente de la Nación. September 2005. Boletín sobre el HIV/SIDA en la Argentina. Year XI, number 24. Programa Nacional de Lucha contra los Retrovirus del Humano, SIDA y ETS. <http://www.msal.gov.ar/htm/site/sida/site/datos-epid-boletines.asp>.
- Pybus, O. G., A. Rambaut, and P. H. Harvey. 2000. An integrated framework for the inference of viral population history from reconstructed genealogies. *Genetics* **155**:1429–1437.
- Salemi, M., T. de Oliveira, M. A. Soares, O. G. Pybus, A. T. Dumans, A. M. Vandamme, A. Tanuri, S. Cassol, and W. M. Fitch. 2005. Different epidemic potentials of the HIV-1 B and C subtypes. *J. Mol. Evol.* **60**:598–605.
- Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The CLUSTAL_X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**:4876–4882.
- Thomson, M. M., E. Delgado, I. Herrero, M. L. Villahermosa, E. Vazquez-de Parga, M. T. Cuevas, R. Carmona, L. Medrano, L. Perez-Alvarez, L. Cuevas, and R. Najera. 2002. Diversity of mosaic structures and common ancestry of human immunodeficiency virus type 1 B/F recombinant viruses from Argentina revealed by analysis of near full-length genome sequences. *J. Gen. Virol.* **83**:107–119.
- UNAIDS. December 2005. AIDS epidemic update. <http://www.unaids.org>.
- Walker, P. R., O. G. Pybus, A. Rambaut, and E. C. Holmes. 2005. Comparative population dynamics of HIV-1 subtypes B and C: subtype-specific differences in patterns of epidemic growth. *Infect. Genet. Evol.* **5**:199–208.
- Wolfs, T. F. W., J.-J. de Jong, H. Van Den Berg, J. M. G. H. Tijnagel, W. J. A. Krone, and J. Goudsmit. 1990. Evolution of sequences encoding the principal neutralisation epitope of human immunodeficiency virus 1 is host dependent, rapid, and continuous. *Proc. Natl. Acad. Sci. USA* **87**:9938–9942.