

Molecular evidence that human immunodeficiency virus type 1 dissemination in a small Brazilian city was already taking place in the early 1990s

Walter A Eyer-Silva^{*/**/+}, Gonzalo Bello, Mariza G Morgado

Laboratório de Aids & Imunologia Molecular, Instituto Oswaldo Cruz-Fiocruz, Av. Brasil 4365, 21045-900 Rio de Janeiro, RJ, Brasil

^{*}Hospital Universitário Gaffrée e Guinle, Universidade Federal do Estado do Rio de Janeiro, Rio de Janeiro, RJ, Brasil

^{**}Programa Municipal de HIV/Aids de Miracema, Miracema, RJ, Brasil

We recently performed a molecular epidemiology survey of human immunodeficiency virus type 1 (HIV-1) infection in Miracema, a small city in Southeast Brazil, and found multiple monophyletic clusters, consistent with independent introductions and spread of different viral lineages in the city. Here we apply Bayesian coalescent-based methods to the two largest subtype B clusters and estimate that the most recent common ancestors that gave rise to these two transmission chains were in circulation around 1991-1992. The finding that HIV-1 spread in this Brazilian small city was already taking place at a time Aids was considered a problem restricted to large urban centers may have important public health implications.

Key words: Aids - epidemiology - human immunodeficiency virus type 1 - Brazil

The true magnitude of the Aids epidemic in small Brazilian cities and rural areas is unknown. Small Brazilian cities are being challenged to provide care to an increasing number of human immunodeficiency virus type 1 (HIV-1)-infected patients, yet the clinical, epidemiological, and virological features of Aids in inner Brazil are scarcely studied. The advent of investigations that focus on the characteristics of HIV-1 infection in areas other than the large metropolitan centers of the country is a relatively recent achievement (Dhalia et al. 2000, Szwarcwald et al. 2000, Brito et al. 2001, Eyer-Silva & Morgado 2005, 2006, Eyer-Silva et al. 2005, 2007).

We recently reported on the molecular epidemiology of HIV-1 infection in the small city of Miracema, northwestern the state of Rio de Janeiro, Southeast Brazil (Eyer-Silva & Morgado 2006). In a setting primarily characterized by heterosexual viral transmission (Eyer-Silva et al. 2005), we performed phylogenetic analyses of HIV-1 strains recovered from a total of 63 adult patients and found a pattern consistent with the existence of multiple viral introductions in the region. We also found 29 samples (representing 46% of the cohort) forming six monophyletic clusters. The two largest clusters that emerged from this study were cluster I, which harbored subtype B strains recovered from ten patients, and cluster II, which encompassed subtype B sequences from seven additional cases (Figure). Intra-cluster patients probably took part in the same chain of viral spread, suggesting the existence of sexual networks and a high incidence molecular profile. Each of these clusters prob-

ably arose as the result of independent introductions of different HIV-1 subtype B lineages in the city. Patients from cluster I and II were diagnosed with HIV-1 infection between the years of 2000-2003 and 1996-2003, respectively. The earliest recorded cases of Aids among residents of the city date back to 1989.

In the present study we apply Bayesian coalescent-based methods to estimate the substitution rate and the time in which the most recent common ancestors of these two largest clusters originated. These estimates may thus stand as an indication of the date in which the founder strains were introduced in the city and began to spread to give rise to the phylogenetic clusters currently observed.

Partial sequences of C2V3 hypervariable region of HIV-1 *envelope* gene were generated from samples collected between 2001 and 2004 using outer and nested primer sets (available on request). Only one sequence per patient was included in the study. Nucleotide sequences were aligned using CLUSTAL X program (Thompson et al. 1997) and later hand edited. All positions with alignment gaps were excluded from the analyses. An alignment of around 215-bp spanning the V3 region of the *envelope* gene was generated for cluster I (nucleotides 7071 to 7289 relative to HXB2 clone) and cluster II (nucleotides 7027 to 7254 relative to HXB2 clone). Alignments are available from the authors upon request. The data sets had no evidence of substitution saturation, hypermutation or intersubtype recombination. Full details of the phylogenetic analyses are described elsewhere (Eyer-Silva & Morgado 2006). The phylogenetic tree is shown in the Figure.

Estimates of the evolutionary rate (μ , nucleotide substitutions per site per year, *subs/site/year*) and the time of the most recent common ancestor (T_{mrca} , years) for each cluster were obtained by using Bayesian Markov Chain Monte Carlo (MCMC) inference as implemented in BEAST v1.3 software (Drummond et al. 2002,

⁺Corresponding author: walter-eyer@unirio.br

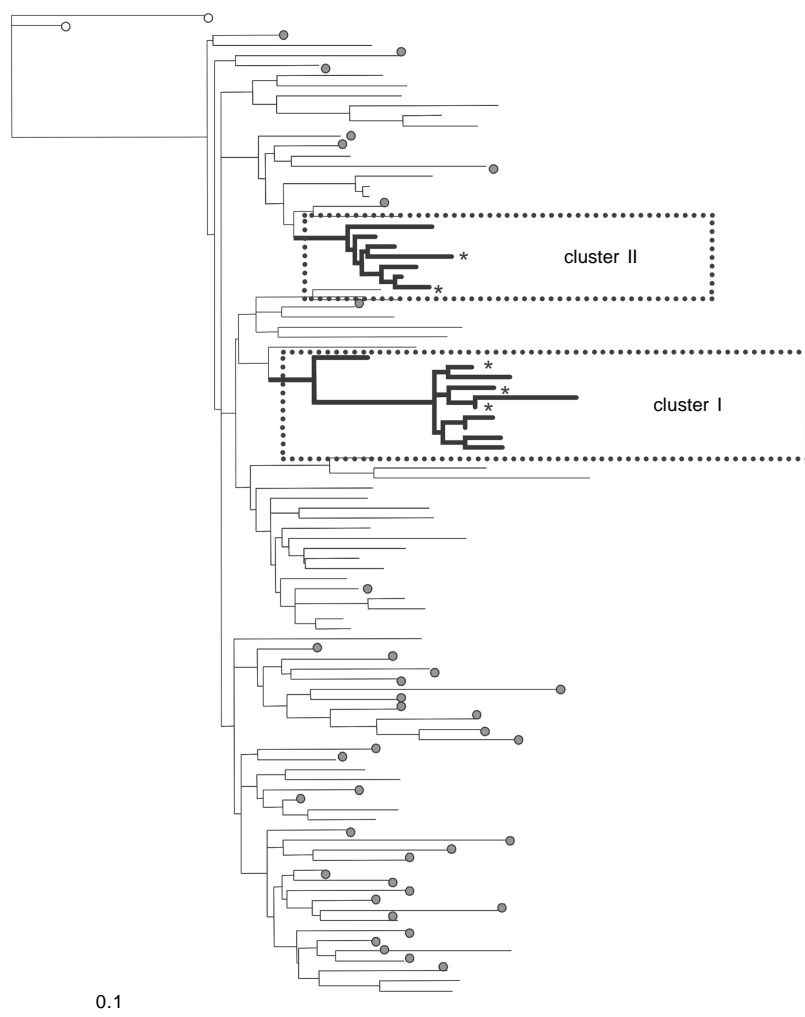
Received 16 January 2007

Accepted 28 May 2007

Drummond & Rambaut 2003). The number and the time span covered (i.e., four years) by clusters I and II *envelope* sequences were not sufficient to reliably estimate μ or T_{mrca} using sample dates. Thus, we used as an empirical prior distribution the posterior distribution of μ (3.8×10^{-3} to 6.2×10^{-3} *subs/site/year*) previously estimated by Bayesian MCMC inference from an independent data set of 145 subtype B C2V3 *envelope* sequences sampled between 1989 and 2004 in Brazil (Bello et al. 2007). Analyses were performed using the Hasegawa, Kishino, and Yano (HKY) model of nucleotide substitution (Hasegawa et al. 1985) with gamma-distributed rates across sites and a proportion of sites assumed to be invariable. Three separate MCMC chains were run for 5×10^6 genera-

tions for each data set, with a burn-in of 5×10^5 . Effective Sampling Sizes (ESS) for the estimates were above 100.

Estimates of substitution rate and T_{mrca} were obtained separately for each cluster data set. The mean estimated HIV-1 evolutionary rate for both clusters was 5.2×10^{-3} *subs/site/year*, similar to that previously estimated for subtype B in Brazil (Bello et al. 2007). The similar mean values and the completely overlap of highest posterior density (HPD) intervals for each rate estimate (Table), suggest no major differences in evolution rate between the two clusters. However, 95% HPD intervals of evolution rate estimates for both clusters (4.1×10^{-3} - 6.2×10^{-3} *subs/site/year*) almost coincide with empirical prior interval (3.8×10^{-3} - 6.2×10^{-3} *subs/site/year*), indicat-



Maximum likelihood optimality criterion phylogenetic reconstruction of 58 subtype B *envelope* hypervariable C2V3 sequences generated from samples collected between 2001 and 2004 from the Miracema study. Clusters I and II are indicated. The tree was constructed using two subtype C sequences as outgroups (white circles) and applying the general time-reversible model of nucleotide substitution with gamma-distributed rates across sites and a fraction of sites assumed to be invariable. Brazilian and foreign subtype B reference strains are indicated as gray circles. The robustness of the tree was evaluated by bootstrap analysis with 1000 rounds of replication. Cluster I and II sequences generated from patients on an Aids-defining clinical stage are indicated with an asterisk. Full details of this molecular epidemiology survey are given in Eyer-Silva and Morgado (2006).

TABLE
 Bayesian estimates of the mean evolutionary rate (μ , substitutions site⁻¹ year⁻¹) and median time of the most recent common ancestor (T_{mrca} , year) for clusters I and II

	Cluster I	Cluster II
μ (95% HPD)	5.2×10^{-3} (4.1×10^{-3} - 6.2×10^{-3})	5.2×10^{-3} (4.1×10^{-3} - 6.2×10^{-3})
T_{mrca} (95% HPD)	1991 (1986 - 1995)	1992 (1988 - 1996)

ing that not much information was added by the data. With these rates of nucleotide substitutions, the mean estimated T_{mrca} for the different clusters was situated around the years 1991 and 1992 for clusters I and II, respectively. For both clusters, 95% HPD intervals around envelope T_{mrca} estimates displayed great overlap (Table), supporting similar emergence dates of these clusters.

In summary, we estimate that the founder strains that triggered the transmission chains resulting in the recently observed clusters I and II in Miracema date back to the early 1990s. Given that all intra-cluster cases of these monophyletic clusters are (or have been) residents of the city, and that some intra-cluster cases had advanced HIV-1 disease in the early 2000s (Figure), it seems plausible that most recent common ancestors of clusters I and II were already circulating in the region around the years of 1991 and 1992. These results suggest that HIV-1 dissemination in this small Brazilian city was probably taking place by the time Aids was considered a problem restricted to large urban centers of the country. It is conceivable that similar chains of viral spread have taken place and may currently be going on in many other small Brazilian cities. Although these results should certainly be interpreted taking into account the limitations of small sample size and narrow time span covered by the sequences, they highlight the urgent need to recognize Aids as an important public health problem in small Brazilian cities and to design strategies to halt epidemic spread towards inner Brazil.

REFERENCES

- Bello G, Eyer-Silva WA, Couto-Fernandez JC, Guimaraes ML, Chequer-Fernandez SL, Teixeira SL, Morgado MG 2007. Demographic history of HIV-1 subtypes B and F in Brazil. *Infect Genet Evol* 7: 263-270.
- Brito AM, Castilho EA, Szwarwald CL 2001. AIDS and HIV infection in Brazil: a multifaceted epidemic. *Rev Soc Bras Med Trop* 34: 207-217.
- Dhalia C, Barreira D, Castilho EA 2000. A AIDS no Brasil: situação atual e tendências. *Boletim epidemiológico Aids, XIII, no. 01*.
- Drummond AJ, Nicholls GK, Rodrigo AG, Solomon W 2002. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics* 161: 1307-1320.
- Drummond AJ, Rambaut A 2003. Beast version 1.0, Available from http://beast.bio.ed.ac.uk/Main_Page/.
- Eyer-Silva WA, Morgado MG 2005. A genotyping study of human immunodeficiency virus type-1 drug resistance in a small Brazilian municipality. *Mem Inst Oswaldo Cruz* 100: 869-873.
- Eyer-Silva WA, Morgado MG 2006. Molecular epidemiology of HIV-1 infection in a small Brazilian county: usefulness of envelope and polymerase sequences to epidemiologic studies. *J Acquir Immune Defic Syndr* 41: 664-670.
- Eyer-Silva WA, Basilio-de-Oliveira CA, Morgado MG 2005. HIV infection and AIDS in a small municipality in Southeast Brazil. *Rev Saúde Pública* 39: 950-955.
- Eyer-Silva WA, Couto-Fernandez JC, Morgado MG 2007. Molecular epidemiology of HIV type 1 in Inner Rio de Janeiro State, Brazil. *AIDS Res Hum Retroviruses* 32: 303-308.
- Hasegawa M, Kishino H, Yano T 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol* 22: 160-174.
- Szwarwald CL, Bastos FI, Esteves MA, de Andrade CL 2000. The spread of the AIDS epidemic in Brazil from 1987 to 1996: a spatial analysis. *Cad Saúde Pública* 16: 7-19.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25: 4876-4882.

