

Autochthonous horizontal transmission of a CRF02_AG strain revealed by a human immunodeficiency virus type 1 diversity survey in a small city in inner state of Rio de Janeiro, Southeast Brazil

Walter A Eyer-Silva/*/**/+, 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-1/Aids de Saquarema, Saquarema, RJ, Brasil

As part of an ongoing study on the features of AIDS spread towards small cities and rural areas, we present a molecular survey of human immunodeficiency virus type 1 (HIV-1) polymerase sequences recovered between 2004 and 2006 from 71 patients receiving care in the city of Saquarema, inner state of Rio de Janeiro. Phylogenetic reconstructions found the two prevalent lineages in the state (subtypes B [59 strains, 83.1%], F1 [6 strains; 8.4%], and BF1 recombinants [four strains; 5.6%]), as well as two (2.8%) CRF02_AG strains, which seems to be an emerging lineage in the capital. These CRF02_AG sequences were recovered from a married heterosexual couple who never traveled abroad, thus providing the first molecular evidence of autochthonous horizontal transmission of this lineage of major global importance. Also, three phylogenetic clusters of strains recovered from a total of 18.3% of the cohort were uncovered. Their close genetic relatedness suggests they were recovered from patients who probably took part in the same chain of viral spread. In conjunction with our previous surveys from inner Rio de Janeiro, these results suggest that although small cities harbor unique molecular features of HIV-1 infection, they also clearly reflect and may rapidly absorb the diversity recorded in large urban centers.

Key words: Brazil - CRF02_AG - human immunodeficiency virus type 1 - molecular epidemiology - transmission clusters

The extremely high genetic diversity of human immunodeficiency virus type 1 (HIV-1) has been successfully used to study the regional and global distribution of different viral groups and subtypes (Hu et al. 1996) and makes it possible to trace patterns of viral spread between populations, groups and people (Kuiken et al. 2000). Phylogenetic analysis allows classification of HIV-1 strains into three groups: M, N and O. The M (main) group, the widest spread, has been further subdivided into at least nine subtypes (A-D, F-H, J, and K) and 32 circulating recombinant forms (CRF), whereas a high diversity of unique intersubtype recombinant forms has also been demonstrated (Leitner et al. 2005). In Brazil, molecular analyses of HIV-1 nucleotide sequences have shown an epidemic driven mainly by three group M subtypes: B, C, and subs subtype F1 (heretofore designated subtype F1), as well as a myriad of unique BF1 intersubtype recombinants (Morgado et al. 1998, Tanuri et al. 1999, Vicente et al. 2000, Brindeiro et al. 2003, Soares et al. 2003). Isolated cases of other subtypes, such as subtype D (Morgado et al. 1998, Couto-Fernandez et al. 2006), subtype A (Caride et al. 2000), CRF02_AG (Pires et al. 2004, Couto-Fernandez et al. 2005), CRF28_BF and CRF29_BF (De Sa Filho et al. 2006) have also been reported.

Phylogenetic analyses of HIV-1 nucleotide sequences are being increasingly used to better understand epidemiological patterns of viral spread. Such studies have proven to be invaluable to clarify the routes of viral transmission in community settings (Holmes et al. 1995, Brown et al. 1997), to establish (Ou et al. 1992, Blanchard et al. 1998) or rule out (Holmes et al. 1993, Jaffe et al. 1994) iatrogenic viral transmission, to study the pathways of viral spread among clusters of patients (Leitner et al. 1996, Robbins et al. 2002), to trace the geographic origin of recently introduced subtypes (Gao et al. 1996, Kato et al. 1999, Yu et al. 2001, Couto-Fernandez et al. 2006), to infer the demographic history and date of origin of epidemic spread (Robbins et al. 2003, Bello et al. 2006), to investigate unusual forms of viral transmission (Goujon et al. 2000, Andreo et al. 2004), and has also found forensic applications (Albert et al. 1994, Banaschak et al. 2000, Metzker et al. 2002, Lemey et al. 2005). Phylogenetic analyses of HIV-1 sequences have also been used to provide evidence of a high incidence molecular profile, as evidenced by multiple transmission clusters, low level of sequence diversity and signature amino acid substitutions (Liitsola et al. 1998, Oelrichs et al. 2000, Shankarappa et al. 2001, Nguyen et al. 2002), and of an old and mature epidemic pattern, given the high sequence diversity and the absence of transmission clusters (Vidal et al. 2000, Trask et al. 2002).

It is known that the AIDS epidemic in Brazil is spreading from the large urban centers towards small cities and the innermost parts of the country (Szwarcwald et al. 2000). Therefore, in an effort to better understand HIV-1 diversity, the molecular epidemiological patterns, and

+Corresponding author: walter-eyer@unirio.br

Received 27 February 2007

Accepted 25 September 2007

to assess the existence of micro-epidemic lineages associated with HIV-1 spread towards inner Brazil, our group is carrying out cross-sectional surveys in selected, small inner cities of the state of Rio de Janeiro. We have recently reported on the molecular epidemiology of HIV-1 infection in two small, neighboring cities in the north-western region of the state: Miracema and Santo Antonio de Pádua (Eyer-Silva & Morgado 2006, Eyer-Silva et al. 2007a, b). We found both a polyphyletic pattern suggestive of multiple viral introductions in the region, and that more than 40% of the recovered sequences fell within clusters of genetically related strains, suggesting the existence of sexual networks and a high incidence molecular profile. Additionally, phylogenetic clusters from one city were unrelated to those found in the other, suggesting that each of these two neighboring cities harbors its own set of micro-epidemic lineages (Eyer-Silva et al. 2007a). We now expand our molecular characterization of HIV-1 strains to another city in inner state of Rio de Janeiro: Saquarema, in the Sun Coat Region. The data presented herein shall broaden our understanding of the genetic diversity and the molecular epidemiological profile associated with HIV-1 spread in inner state of Rio de Janeiro.

PATIENTS, MATERIALS AND METHODS

Patients and setting - Saquarema is situated 116 km north from the city of Rio de Janeiro and counted, as of July 2005, 61,591 inhabitants distributed in a geographic area of 355 km² (<http://www.ibge.gov.br/cidadesat/>). Bathed by the Atlantic, the city enjoys a nationwide reputation for its natural beauty and attracts thousands of domestic and foreign visitors and vacationers every year. A municipal HIV/AIDS program was established in 2000 and patients are offered regular medical appointments. The present study protocol was reviewed and approved by the Ethics Review Board at Instituto de Pesquisa Clínica Evandro Chagas, Fundação Oswaldo Cruz, Rio de Janeiro.

Nucleotide sequencing and phylogenetic reconstructions - After signed informed consent, blood samples were collected, DNA was extracted from peripheral blood mononuclear cells, and the sequence of a single fragment (around 1000 nucleotides) of HIV-1 polymerase gene encompassing the whole protease (PR) and reverse transcriptase (RT) amino acids 1-230 was generated using a previously described in-house protocol (Eyer-Silva & Morgado 2005). In spite of being a relatively conserved region of the HIV-1 genome, the polymerase gene has been shown to contain enough phylogenetic signal for subtyping analysis (Cornelissen et al. 1997) and to allow studies of epidemiologic relationships between populations, groups and people (Hue et al. 2004, Eyer-Silva & Morgado 2006). Sequences were deposited on the GenBank database under accession numbers DQ899721 through DQ899721 and EF042627 through EF042690.

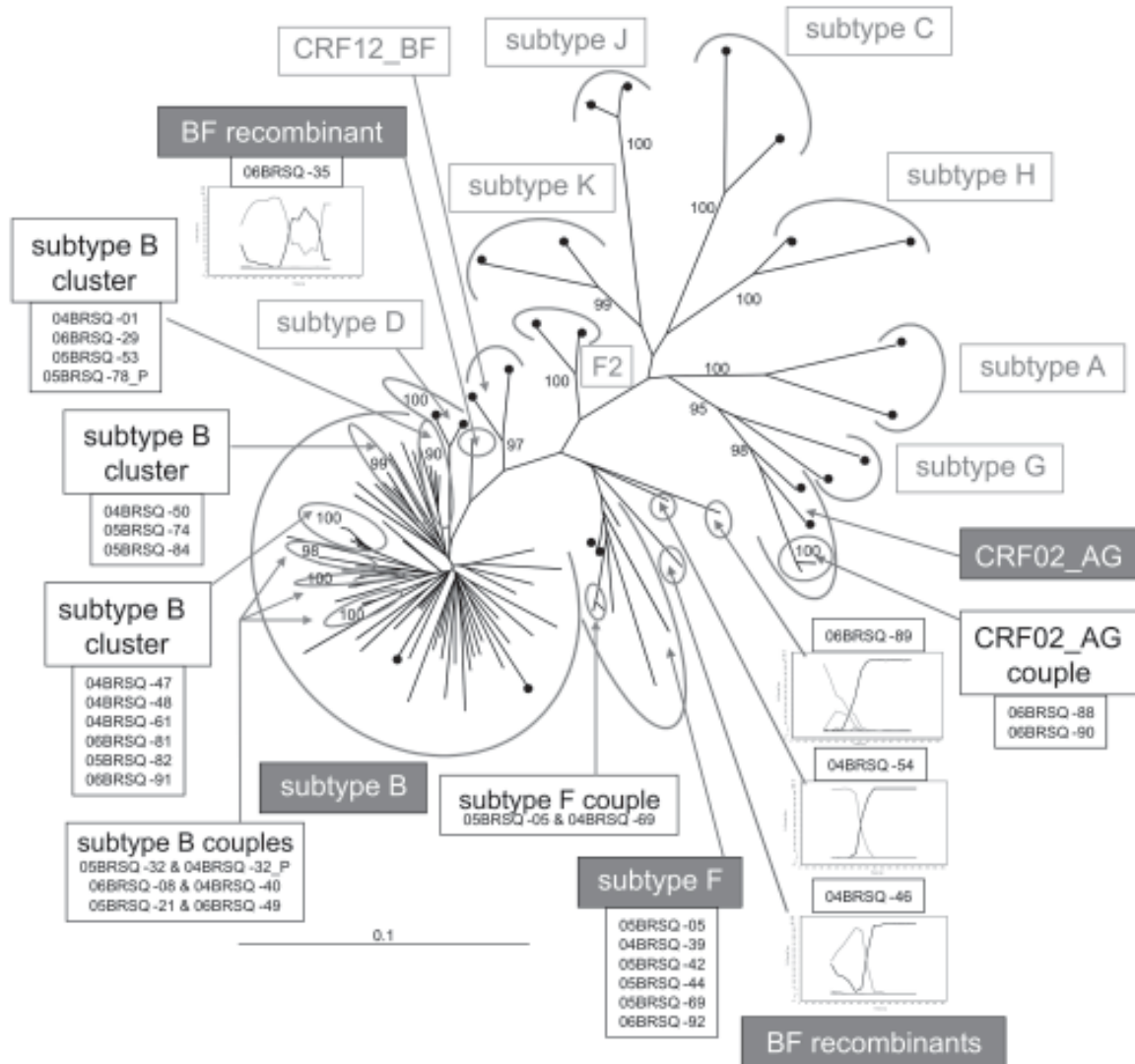
Sequences were aligned against a set of reference strains from HIV-1 group M subtypes and trimmed to equivalent lengths by using CLUSTAL X (Thompson et al. 1997). Minor adjustments were performed using

BioEdit (Hall 1999). An alignment of 901 nucleotides, corresponding to positions 2313 through 3213 relative to HXB2 genome, was obtained. Phylogenetic inferences were performed by applying the maximum likelihood (ML) optimality criterion (Felsenstein 1973) as implemented in PAUP* version 4.0b10 (Swofford 1999). The best fitting nucleotide substitution model was selected by using the Akaike criterion as implemented in MODELTEST version 3.06 (Posada & Crandall 1998). We found the GTR model with gamma-distributed rates across sites and a fraction of sites assumed to be invariable to be the best fitting. The parameters of the chosen model are indicated in the Figure. ML trees were run with PAUP* using heuristic searches based on a subtree pruning and regrafting (SPR) branch-swapping algorithm (Felsenstein 2004) and 1000 rounds of bootstrap replication (Felsenstein 1985). The robustness of the tree was evaluated by bootstrap analysis with 1,000 rounds of replication and the phylogenetic tree was schematically represented with the TreeView program (Page 1996). The bootscanning method was used to detect and study recombination, as implemented in the SIMPLOT software, version 3.5 (Salminen et al. 1995). We found no evidence of hypermutation in the dataset when sequence CD83ELI (GenBank accession no. K02454) was used as the reference strain in the HYPERMUT program (Rose & Korber 2000). The presence of nucleotide substitution saturation at different codon positions was analyzed by plotting the transitions and transversions versus the F84 model of substitution by using the DAMBE software (Xia & Xie 2001). No nucleotide substitution saturation could be observed on inspection of the plots. This implies that the dataset was phylogenetically informative.

Statistical analyses - Patients whose samples yielded sequences that fell within a phylogenetic cluster were compared with those whose samples yielded non-clustered sequences. Depending on the variables, the Wilcoxon rank sum test, and the Fisher's exact test were used for the univariable analysis. Multivariate logistic regression was used to identify variables independently associated with being a patient whose sample yielded clustered sequences (response variable). Age, gender, place of birth and residence, marital status, exposure category, AIDS-defining CDC stage, having a known direct epidemiologic relationship within the cohort, having always lived in Saquarema, previous use of snorted cocaine, and level of education were studied as explanatory variables. Statistical analyses were performed using the software R (Ihaka & Gentleman 1996).

RESULTS

Assignment to subtypes - Between October 2001 and May 2006 a total of 93 HIV-infected adult patients received care at a local ambulatory facility. A detailed clinical and epidemiological characterization of this cohort will be presented elsewhere. Between the years of 2004 and 2006 blood samples were collected from 73 enrolled patients. A total of 71 strains were recovered from 68 adult and three pediatric patients (including one mother and daughter couple). We failed to obtain se-



Maximum likelihood optimality criterion phylogenetic study of 71 polymerase sequences recovered from patients followed in the Municipal HIV/AIDS Program of Saquarema, inner state of Rio de Janeiro, between 2004 and 2006. Alignments were prepared with reference human immunodeficiency virus type 1 (HIV-1) group M subtypes gathered from the Los Alamos HIV database (<http://hiv-web.lanl.gov>). Reference strains are represented as black dots. Phylogenetically related clusters and couples of sequences supported by a bootstrap value greater than 80 are circled and the corresponding strains are listed. The first two digits of the sequence name indicate the year of sample collection, whereas the suffix _P indicates a pediatric patient. Trees were constructed by using the GTR model with Γ -distributed rates across sites and a fraction of sites assumed to be invariable. The robustness of the trees was evaluated by bootstrap analysis with 1,000 rounds of replication. The parameters of the model were as follows: equilibrium nucleotide frequencies, $f_A=0.4125$, $f_C=0.1310$, $f_G=0.2117$, $f_T=0.2448$; R matrix values, $R_{A\leftrightarrow C}=2.8404$, $R_{A\leftrightarrow G}=9.2035$, $R_{A\leftrightarrow T}=0.7296$, $R_{C\leftrightarrow G}=1.7363$, $R_{C\leftrightarrow T}=15.0106$, $R_{G\leftrightarrow T}=1$. The proportion of invariable sites was 46.16% and the shape parameter of the gamma distribution used for the reconstruction of the ML tree was 1.04. Bootscanning plots depicting the relationship of BF1 intersubtype recombinant strains to representatives of group M subtypes (F1 in black, B in light gray) are also shown. The analysis was performed on a sliding window of 400 nucleotides of the query sequences moving along an alignment of reference sequences of group M subtypes by increments of 40 nucleotides.

quences from two samples. The Figure presents the ML phylogenetic reconstruction. A total of 59 sequences (83.1%) were assigned group M subtype B. The remaining sequences were assigned group M subtypes F1 (6 strains; 8.4%), CRF02_AG (2 strains; 2.8%), and unique BF1 recombinants (4 strains; 5.6%). Three unique BF1 recombinants (04BRSQ-46, 04BRSQ-54, 06BRSQ-89) harbored a subtype B fragment in the PR region and a subtype F1 fragment in the RT gene. The remaining BF1

mosaic strain (06BRSQ-35) displayed a predominant subtype B pattern with an F1 fragment in the RT region (Figure). Phylogenetic analyses of these BF1 recombinants found that they are not related to previously described CRF_BF (data not shown).

Identification of phylogenetically related clusters and couples - ML phylogenetic reconstruction yielded three subtype B clusters encompassing a total of 13 strains (Figure), corresponding to 18.3% of all gener-

ated sequences. The largest cluster harbored six sequences and was supported by a bootstrap value of 100. Two other subtype B clusters harbored four and three strains and were supported by bootstrap figures of 90 and 99, respectively. The former included a pediatric patient (05BRSQ-78_P) whose mother had died of complications of AIDS before the study period. Additionally, five phylogenetically-related couples of sequences were also seen: three couples were assigned subtype B, whereas the other two were assigned subtype F and CRF02_AG. These CRF02_AG strains (06BRSQ-88 and 06BRSQ-90) branched with a bootstrap value of 100 and were recovered from a married heterosexual couple who reported never have traveled abroad. Patient 06BRSQ-88, a 28-year-old male, attributed the acquisition of HIV-1 infection to unprotected sexual contacts in the city of Rio de Janeiro. His 29-year-old wife (patient 06BRSQ-90) reported having always lived in Saquarema.

Known epidemiologic relationships in the cohort - All sequences recovered from patients whose epidemiologic relationships were known a priori were found to be phylogenetically associated within one of the observed clusters or couples of related strains. These were the five phylogenetically-related couples (patients SQ-32/SQ-32_P, the mother and child couple, SQ-88/SQ90, a married couple infected with CRF02_AG strains, as well as couples SQ-01/SQ-53 and SQ-81/SQ-82 and the triad SQ-47/SQ-48/SQ-61), all of whom found to be infected with clustered sequences. The other patients whose sequences were found to be phylogenetically associated had no known direct epidemiologic relationship.

Demographic, social, clinical and behavioral associations within phylogenetic clusters - Samples from 12 (17.6%) out of 68 adult patients yielded clustered sequences. On univariable analysis, these patients were found to be younger, more likely to have a known epidemiologic link within the cohort, to always have lived in Saquarema and to reside within the geographic limits of Saquarema as of sampling. Multivariable logistic regression analysis identified having a known direct epidemiologic relationship (odds ratio [OR] 9.44, 95% confidence interval [CI] 7.83-11.05, $P = 0.006$) and having always lived in Saquarema (OR 9.11, 95% CI 7.29-10.93, $p = 0.017$) as independent predictors of belonging to a cluster.

DISCUSSION

The present molecular survey of HIV-1 polymerase sequences from the city of Saquarema uncovered the occurrence of three HIV-1 subtypes (B, F1, and CRF02_AG) and four BF1 mosaic forms. It is important to highlight that these subtype assignments have been made based on the study of a fragment of the polymerase gene and can not be extrapolated to other viral genomic regions. The finding of such a genetic diversity in a small city in inner state of Rio de Janeiro further highlights the complexity of the molecular epidemiology of HIV-1 in Brazil. Additionally, we present the first molecular evidence of horizontal transmission of subtype CRF02_AG in Brazil. This is a rare HIV-1 variant in the country and only two cases have been published in peer-

reviewed journals (both from the metropolitan area of Rio de Janeiro) (Pires et al. 2004, Couto-Fernandez et al. 2005). However, the recent identification of novel cases of HIV-1 infection due to CRF02_AG at our laboratory (unpublished observations) suggests that this might prove to be an emerging lineage in the city of Rio de Janeiro. The present molecular evidence of autochthonous heterosexual transmission of CRF02_AG strains, an HIV-1 variant of major global importance, emphasize the importance of continued monitoring of viral diversity in Brazil.

Phylogenetic analyses found evidence of three clusters of subtype B strains recovered from a total of 12 adult patients and one pediatric case whose mother had died of complications of AIDS before the study period. The largest cluster harbored six strains, whereas two other clusters contained four and three sequences (Figure). These intracluster strains shared close genetic relatedness and were phylogenetically associated with high bootstrap values, suggesting that they were recovered from patients who probably took part in the same chain of viral spread. Similar phylogenetic clusters have also been identified in our previous molecular epidemiology surveys of HIV-1 strains from two small cities in northwestern state of Rio de Janeiro. In the present study, however, we found a lower prevalence of clustered sequences. Whereas in the previously reported analyses more than 40% of the strains were found to belong to a phylogenetic cluster (Eyer-Silva & Morgado 2006, Eyer-Silva et al. 2007a), in the present Saquarema study 13 (18.3%) of the strains fell within a cluster. The lower prevalence of phylogenetically related cases probably is due to differences in the characteristics of the cohorts. When compared to the two northwestern cities, patients from Saquarema were significantly more likely to have been born elsewhere and to have initiated HIV follow up in another city (data not shown). For instance, 60.2% of patients from Saquarema were born elsewhere, whereas the corresponding figures for Santo Antonio de Pádua and Miracema were 20.6% and 11.1%, respectively ($p < 0.01$). In agreement with our previous studies (Eyer-Silva & Morgado 2006), patients whose sequences fell within a cluster were significantly more likely to have a known epidemiologic relationship within the cohort and to always have lived in the studied city.

In Brazil relatively little attention is being paid to the study of the features of the spread of the AIDS epidemic towards small cities and rural areas. This molecular analysis of HIV-1 sequences from the small city of Saquarema, in conjunction with our previous surveys from the inner cities of Miracema and Santo Antonio de Pádua, demonstrate the cocirculation in inner state of Rio de Janeiro of the two prevalent subtypes in Southeast Brazil (B, F1, and BF1 recombinants), as well as two minor variants previously recorded in the capital: CRF02_AG, presented in this report, and the recently studied subtype D (Morgado et al. 1998, Couto-Fernandez et al. 2006, Eyer-Silva et al. 2007a). Moreover, we present the first molecular evidence of autochthonous horizontal transmission of CRF02_AG strains in Brazil. These results suggest that although small

cities of inner state of Rio de Janeiro harbor unique molecular features of HIV-1 infection, they also clearly reflect and may rapidly absorb the diversity recorded in large urban centers. The finding of such a complex pattern of viral subtypes and recombinant forms in small cities of Rio de Janeiro also underscore the need for a continued study of the diversity and routes of HIV-1 spread in inner Brazil. These surveys will lead to a better understanding of the dynamics of the Brazilian HIV-1 epidemic, will be invaluable to monitor the prevalence and emergence of different viral lineages and may eventually be important to geographically related vaccine development.

ACKNOWLEDGEMENTS

To the patients followed at the Municipal HIV/ AIDS Program of Saquarema, for agreeing to participate in this study. To the many dedicated health care professionals in Saquarema, without whom this study would not be possible.

REFERENCES

- Albert J, Wahlberg J, Leitner T, Escanilla D, Uhlen M 1994. Analysis of a rape case by direct sequencing of the human immunodeficiency virus type 1 pol and gag genes. *J Virol* 68: 5918-5924.
- Andreo SM, Barra LA, Costa LJ, Sucupira MC, Souza IE, Diaz RS 2004. HIV type 1 transmission by human bite. *AIDS Res Hum Retroviruses* 20: 349-350.
- Banaschak S, Werwein M, Brinkmann B, Hauber I 2000. Human immunodeficiency virus type 1 infection after sexual abuse: value of nucleic acid sequence analysis in identifying the offender. *Clin Infect Dis* 31: 1098-1100.
- Bello G, Guimaraes ML, Morgado MG 2006. Evolutionary history of HIV-1 subtype B and F infections in Brazil. *AIDS* 20: 763-768.
- Blanchard A, Ferris S, Chamaret S, Guetard D, Montagnier L 1998. Molecular evidence for nosocomial transmission of human immunodeficiency virus from a surgeon to one of his patients. *J Virol* 72: 4537-4540.
- Brindeiro RM, Diaz RS, Sabino EC, Morgado MG, Pires IL, Brigido L, Dantas MC, Barreira D, Teixeira PR, Tanuri A 2003. Brazilian Network for HIV Drug Resistance Surveillance (HIV-BResNet): a survey of chronically infected individuals. *AIDS* 17: 1063-1069.
- Brown AJ, Lobidel D, Wade CM, Rebus S, Phillips AN, Brettell RP, France AJ, Leen CS, McMenamin J, McMillan A, Maw RD, Mulcahy F, Robertson JR, Sankar KN, Scott G, Wyld R, Peutherer JF 1997. The molecular epidemiology of human immunodeficiency virus type 1 in six cities in Britain and Ireland. *Virology* 235: 166-177.
- Caride E, Brindeiro RM, Hertogs K, Larder B, Dehertogh P, Machado E, de Sa CA, Eyer-Silva WA, Sion FS, Passioni LF, Menezes JA, Calazans AR, Tanuri A 2000. Drug-resistant reverse transcriptase genotyping and phenotyping of B and non-B subtypes (F and A) of human immunodeficiency virus type 1 found in Brazilian patients failing HAART. *Virology* 275: 107-115.
- Cornelissen M, van den Burg R, Zorgdrager F, Lukashov V, Goudsmit J 1997. pol gene diversity of five human immunodeficiency virus type 1 subtypes: evidence for naturally occurring mutations that contribute to drug resistance, limited recombination patterns, and common ancestry for subtypes B and D. *J Virol* 71: 6348-6358.
- Couto-Fernandez JC, Eyer-Silva WA, Guimaraes ML, Chequer-Fernandez SL, Grinsztejn B, Delaporte E, Peeters M, Morgado MG 2006. Phylogenetic analysis of Brazilian HIV type 1 subtype D strains: tracing the origin of this subtype in Brazil. *AIDS Res Hum Retroviruses* 22: 207-211.
- Couto-Fernandez JC, Silva-de-Jesus C, Veloso VG, Rachid M, Gracie RS, Chequer-Fernandez SL, Oliveira SM, Arakaki-Sanchez D, Chequer PJ, Morgado MG 2005. Human immunodeficiency virus type 1 (HIV-1) genotyping in Rio de Janeiro, Brazil: assessing subtype and drug-resistance associated mutations in HIV-1 infected individuals failing highly active antiretroviral therapy. *Mem Inst Oswaldo Cruz* 100: 73-78.
- De Sa Filho DJ, Sucupira MC, Caseiro MM, Sabino EC, Diaz RS, Janini LM 2006. Identification of two HIV type 1 circulating recombinant forms in Brazil. *AIDS Res Hum Retroviruses* 22: 1-13.
- Eyer-Silva WA, Bello G, Morgado MG 2007b. Molecular evidence that human immunodeficiency virus type 1 dissemination in a small Brazilian city was already taking place in the early 1990s. *Mem Inst Oswaldo Cruz* 102: 647-649.
- Eyer-Silva WA, Couto-Fernandez JC, Morgado MG 2007a. Molecular epidemiology of HIV type 1 in Inner Rio de Janeiro State, Brazil. *AIDS Res Hum Retroviruses* 23: 303-308.
- 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.
- Felsenstein J 1973. Maximum-likelihood estimation of evolutionary trees from continuous characters. *Am J Hum Genet* 25: 471-492.
- Felsenstein J 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783-791.
- Felsenstein J 2004. Finding the best tree by heuristic search. *Infering phylogenies*. Sinauer associates, Inc., Sunderland, MA.
- Gao F, Robertson DL, Morrison SG, Hui H, Craig S, Decker J, Fultz PN, Girard M, Shaw GM, Hahn BH, Sharp PM 1996. The heterosexual human immunodeficiency virus type 1 epidemic in Thailand is caused by an intersubtype (A/E) recombinant of African origin. *J Virol* 70: 7013-7029.
- Goujon CP, Schneider VM, Grofti J, Montigny J, Jeantils V, Astagneau P, Rozenbaum W, Lot F, Frocrain-Herchkovitch C, Delphin N, Le Gal F, Nicolas JC, Milinkovitch MC, Deny P 2000. Phylogenetic analyses indicate an atypical nurse-to-patient transmission of human immunodeficiency virus type 1. *J Virol* 74: 2525-2532.
- Hall TA 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41: 95-98.
- Holmes EC, Zhang LQ, Robertson P, Cleland A, Harvey E, Simmonds P, Leigh Brown AJ 1995. The molecular epidemiology of human immunodeficiency virus type 1 in Edinburgh. *J Infect Dis* 171: 45-53.
- Holmes EC, Zhang LQ, Simmonds P, Rogers AS, Brown AJ 1993. Molecular investigation of human immunodeficiency virus (HIV) infection in a patient of an HIV-infected surgeon. *J Infect Dis* 167: 1411-1414.

- Hu DJ, Dondero TJ, Rayfield MA, George JR, Schochetman G, Jaffe HW, Luo CC, Kalish ML, Weniger BG, Pau CP, Schable CA, Curran JW 1996. The emerging genetic diversity of HIV. The importance of global surveillance for diagnostics, research, and prevention. *JAMA* 275: 210-216.
- Hue S, Clewley JP, Cane PA, Pillay D 2004. HIV-1 pol gene variation is sufficient for reconstruction of transmissions in the era of antiretroviral therapy. *AIDS* 18: 719-728.
- Ihaka R, Gentleman R 1996. R: A Language for Data Analysis and Graphics. *J Comp Graph Stat* 5: 299-314.
- Jaffe HW, McCurdy JM, Kalish ML, Liberti T, Metellus G, Bowman BH, Richards SB, Neasman AR, Witte JJ 1994. Lack of HIV transmission in the practice of a dentist with AIDS. *Ann Intern Med* 121: 855-859.
- Kato K, Shiino T, Kusagawa S, Sato H, Nohtomi K, Shibamura K, Nguyen TH, Pham KC, Truong XL, Mai HA, Hoang TL, Bunyaraksyotin G, Fukushima Y, Honda M, Wasi C, Yamazaki S, Nagai Y, Takebe Y 1999. Genetic similarity of HIV type 1 subtype E in a recent outbreak among injecting drug users in northern Vietnam to strains in Guangxi Province of southern China. *AIDS Res Hum Retroviruses* 15: 1157-1168.
- Kuiken C, Thakallapalli R, Esklid A, de Ronde A 2000. Genetic analysis reveals epidemiologic patterns in the spread of human immunodeficiency virus. *Am J Epidemiol* 152: 814-822.
- Leitner T, Escanilla D, Franzen C, Uhlen M, Albert J 1996. Accurate reconstruction of a known HIV-1 transmission history by phylogenetic tree analysis. *Proc Natl Acad Sci USA* 93: 10864-10869.
- Leitner T, Foley B, Hahn BH, Marx P, McCutchan F, Mellors JW, Wolinsky S, Korber BT (eds) 2005. *HIV Sequence Compendium 2005*. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM, 698 pp.
- Lemey P, Van Dooren S, Van Laethem K, Schrooten Y, Derdelinckx I, Goubau P, Brun-Vezinet F, Vaira D, Vandamme AM 2005. Molecular testing of multiple HIV-1 transmissions in a criminal case. *AIDS* 19: 1649-1658.
- Liitsola K, Tashkinova I, Laukkanen T, Korovina G, Smolskaja T, Momot O, Mashkilleysen N, Chaplinskias S, Brummer-Korvenkontio H, Vanhatalo J, Leinikki P, Salminen MO 1998. HIV-1 genetic subtype A/B recombinant strain causing an explosive epidemic in injecting drug users in Kaliningrad. *AIDS* 12: 1907-1919.
- Metzker ML, Mindell DP, Liu XM, Ptak RG, Gibbs RA, Hillis DM 2002. Molecular evidence of HIV-1 transmission in a criminal case. *Proc Natl Acad Sci USA* 99: 14292-14297.
- Morgado MG, Guimaraes ML, Gripp CB, Costa CI, Neves I, Jr, Veloso VG, Linhares-Carvalho MI, Castello-Branco LR, Bastos FI, Kuiken C, Castilho EA, 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.
- Nguyen L, Hu DJ, Choopanya K, Vanichseni S, Kitayaporn D, van Griensven F, Mock PA, Kittikraisak W, Young NL, Mastro TD, Subbarao S 2002. Genetic analysis of incident HIV-1 strains among injection drug users in Bangkok: evidence for multiple transmission clusters during a period of high incidence. *J Acquir Immune Defic Syndr* 30: 248-256.
- Oelrichs RB, Shrestha IL, Anderson DA, Deacon NJ 2000. The explosive human immunodeficiency virus type 1 epidemic among injecting drug users of Kathmandu, Nepal, is caused by a subtype C virus of restricted genetic diversity. *J Virol* 74: 1149-1157.
- Ou CY, Ciesielski CA, Myers G, Bandea CI, Luo CC, Korber BT, Mullins JI, Schochetman G, Berkelman RL, Economou AN, Witte JJ, Furman LJ, Satten GA, Macinnes KA, Curran JW, Jaffe HW, Laboratory Investigation Group, Epidemiologic Investigation Group 1992. Molecular epidemiology of HIV transmission in a dental practice. *Science* 256: 1165-1171.
- Page RD 1996. TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* 12: 357-358.
- Pires IL, Soares MA, Speranza FA, Ishii SK, Vieira MC, Gouvea MI, Guimaraes MA, de Oliveira FE, Magnanini MM, Brindeiro RM, Tanuri A 2004. Prevalence of human immunodeficiency virus drug resistance mutations and subtypes in drug-naive, infected individuals in the army health service of Rio de Janeiro, Brazil. *J Clin Microbiol* 42: 426-430.
- Posada D, Crandall KA 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14: 817-818.
- Robbins KE, Lemey P, Pybus OG, Jaffe HW, Youngpairaj AS, Brown TM, Salemi M, Vandamme AM, Kalish ML 2003. U.S. Human immunodeficiency virus type 1 epidemic: date of origin, population history, and characterization of early strains. *J Virol* 77: 6359-6366.
- Robbins KE, Weidle PJ, Brown TM, Saekhou AM, Coles B, Holmberg SD, Folks TM, Kalish ML 2002. Molecular analysis in support of an investigation of a cluster of HIV-1-infected women. *AIDS Res Hum Retroviruses* 18: 1157-1161.
- Rose PP, Korber BT 2000. Detecting hypermutations in viral sequences with an emphasis on G → A hypermutation. *Bioinformatics* 16: 400-401.
- Salminen MO, Carr JK, Burke DS, McCutchan FE 1995. Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. *AIDS Res Hum Retroviruses* 11: 1423-1425.
- Shankarappa R, Chatterjee R, Learn GH, Neogi D, Ding M, Roy P, Ghosh A, Kingsley L, Harrison L, Mullins JI, Gupta P 2001. Human immunodeficiency virus type 1 env sequences from Calcutta in eastern India: identification of features that distinguish subtype C sequences in India from other subtype C sequences. *J Virol* 75: 10479-10487.
- Soares MA, De Oliveira T, Brindeiro RM, Diaz RS, Sabino EC, Brígido L, Pires IL, Morgado MG, Dantas MC, Barreira D, Teixeira PR, Cassol S, Tanuri A 2003. A specific subtype C of human immunodeficiency virus type 1 circulates in Brazil. *AIDS* 17: 11-21.
- Swofford DL 1999. PAUP* - Phylogenetic analysis using parsimony (*and other methods). Version 4.0. Sinauer Associates, Inc, Sunderland, MA.
- Szwarcwald 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 Saude Pública* 16: 7-19.
- Tanuri A, Swanson P, Devare S, Berro OJ, Savedra A, Costa LJ, Telles JG, Brindeiro R, Schable C, Pieniazek D, Rayfield M 1999. HIV-1 subtypes among blood donors from Rio de Janeiro, Brazil. *J Acquir Immune Defic Syndr Hum Retrovirol* 20: 60-66.
- 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.
- Trask SA, Derdeyn CA, Fideli U, Chen Y, Meleth S, Kasolo F, Musonda R, Hunter E, Gao F, Allen S, Hahn BH 2002. Molecular epidemiology of human immunodeficiency virus type 1 transmission in a heterosexual cohort of discordant couples in Zambia. *J Virol* 76: 397-405.
- Vicente AC, Otsuki K, Silva NB, Castilho MC, Barros FS, Pieniazek D, Hu D, Rayfield MA, 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, 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.
- Xia X, Xie Z 2001. DAMBE: software package for data analysis in molecular biology and evolution. *J Hered* 92: 371-373.
- Yu XF, Liu W, Chen J, Kong W, Liu B, Yang J, Liang F, McCutchan F, Piyasirisilp S, Lai S 2001. Rapid dissemination of a novel B/C recombinant HIV-1 among injection drug users in southern China. *AIDS* 15: 523-525.