FIOCRUZ
Fundação Oswaldo Cruz

Estimating the origin of the HIV-1 CRF02_AG lineages circulating in Brazil

Edson Delatorre[1], Carlos A. Velasco-de-Castro[2], José H. Pilotto[1,3], Gonzalo Bello[1], Mariza G. Morgado[1]

1. Laboratório de AIDS e Imunologia Molecular, Instituto Oswaldo Cruz, FIOCRUZ, Rio de Janeiro, Brazil. 2. Laboratório de Virologia, Departamento de Patologia Clínica, Instituto Fernandes Figueira – FIOCRUZ, Rio de Janeiro, Brazil. 3. Hospital Geral de Nova Iguaçu – Ministério da Saúde, Rio de Janeiro Brazil

edsonod@ioc.fiocruz.br

IOC
Instituto Oswaldo Cruz

LabAids
Laboratório de Aids e Imunologia Molecular

Introduction

- The HIV-1 CRF02_AG is one of the most prevalent recombinant forms in the world and responsible for at least 8% of the HIV-1 infections worldwide. This recombinant is distributed mainly in West Africa and, to a lesser extent, in the Middle East and North Africa. Due the migrations from these endemic regions, CRF02_AG has been reported recently in countries where this recombinant is not native, including Brazil.

Material and methods

- HIV-1 CRF02_AG *pol* sequences isolated in Brazil (*n* = 20, 18 from Rio de Janeiro and two from São Paulo) were compared with a large number (*n* = 1,505) of CRF02_AG *pol* sequences of African origin. Most of the Brazilian sequences and all those from Africa were obtained from Los Alamos HIV Sequence DB (www.hiv.lanl.gov).
- The Maximum Likelihood (ML) phylogenetic tree was reconstructed with PhyML program using an online web server .
- The evolutionary rate, the age of the most recent common ancestor (T_{MRCA}, years) and the most probably origin of the Brazilian CRF_AG sequences were jointly estimated using the Bayesian Markov Chain Monte Carlo (MCMC) approach implemented in the BEAST v1.8.

Results

- The ML analysis showed that the 20 CRF02_AG Brazilian sequences were distributed in five independent lineages. The lineages BR-I and BR-II were composed by 11 and 6 sequences respectively, all from Rio de Janeiro. The remaining three isolates formed single sequence lineages: BR-III, BR-IV and BR-V (Figure 1).
- The lineages BR-I and BR-II have spread outside the metropolitan region of Rio de Janeiro (Figure 2).
- The Bayesian phylogeographic analysis of the Brazilian sequences and their most closely related African sequences (*n* = 212) placed the origin of all Brazilian lineages in West Africa. The lineages BR-I and BR-II probably originated in Ghana [posterior state probability (PSP) = 1 and PSP = 0.94, respectively], the BR-III lineage probably originated in Ghana (PSP = 0.67) or Benin (PSP = 0.27), BR-IV in Senegal (PSP = 0.62) or Guinea-Bissau (PSP = 0.23) and BR-V lineage in Nigeria (PSP = 1) (Figure 3).
- The date of origin of lineages BR-I and BR-II were estimated at 1985 (95% highest posterior density: 1979 – 1992).

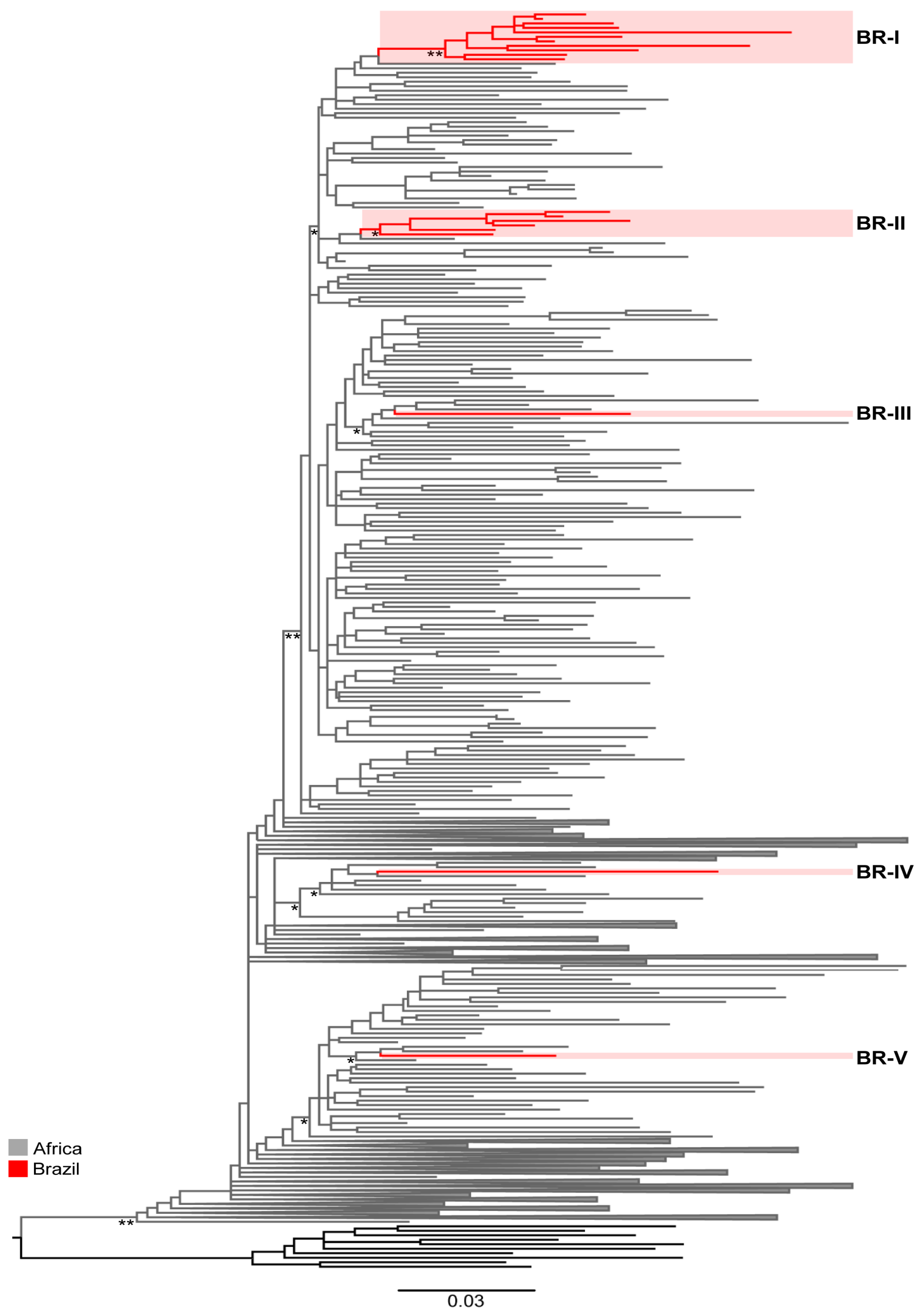

Figure 1. ML tree of HIV-1 CRF02_AG pol (~1,000 pb) sequences from Brazil (n = 20) and African countries (n = 1,505). The tree was rooted using HIV-1 subtype G sequences (black branches). The branches colors represent the geographic origin of the sequences. The red boxes highlight the position of the Brazilian CRF02_AG lineages (BR-I to BR-V). For visual clarity, only the clades most closely related to the Brazilian lineages and used in the subsequent analysis are showed. Asterisks point to key nodes with aLTR support values between 0.80 and 0.9 (*) and >0.9 (**). Horizontal branch lengths are drawn to scale with the bar at the bottom indicating nucleotide substitutions per site.

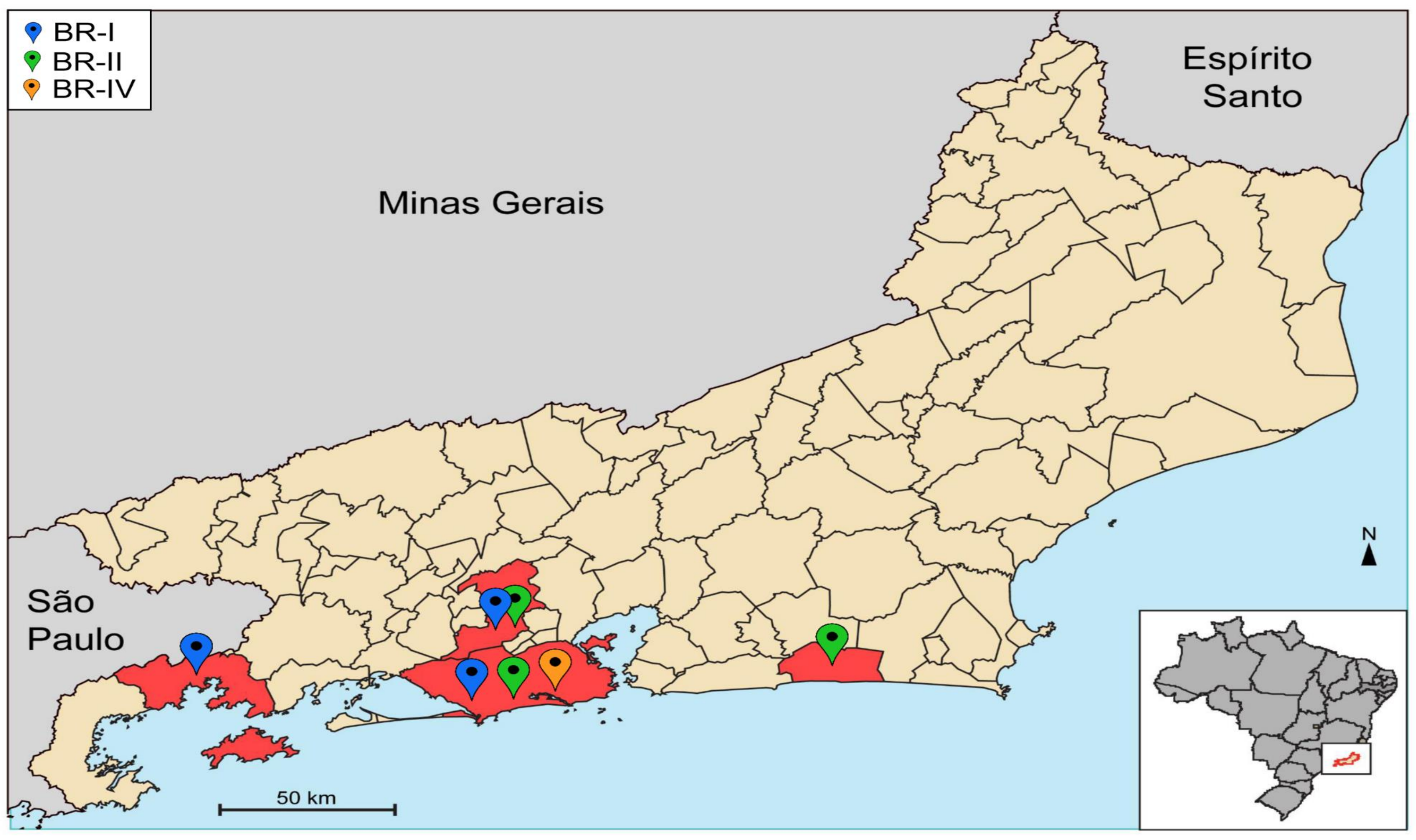

Figure 2. Political map of Rio de Janeiro state, showing the localization of the counties with HIV-1 CRF02_AG cases. Each affected county was colored red and each colored pin represents one of the HIV-1 CRF02_AG Brazilian lineage present.

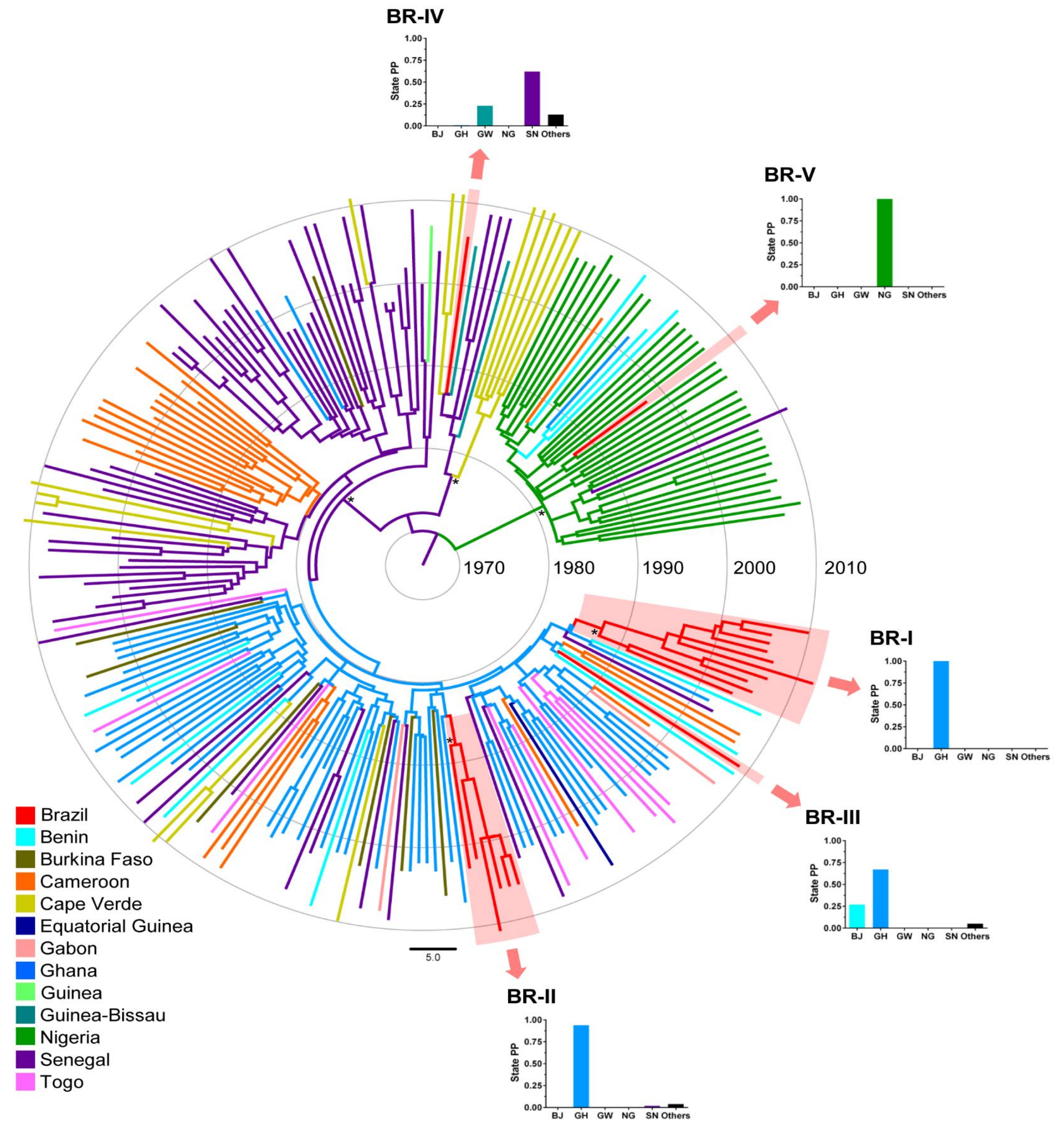

Figure 3. Time-scaled Bayesian MCC tree of the HIV-1 CRF02_AG Brazilian BR-I to BR-V lineages and the most closely related African sequences (*n* = 212). Branches are colored according to the most probable location state of their descendent nodes as indicated at the legend. The red boxes highlight the position of the Brazilian HIV-1 CRF02_AG lineages. Asterisks point to key nodes with high posterior probability support (*PP* > 0.9). The tree was automatically rooted under the assumption of a relaxed molecular clock and the branch lengths are drawn to scale with dates indicated in the concentric circles. Scale bar at the bottom indicates years. The *PSP* distributions at the first ancestral nodes of the Brazilian CRF02_AG lineages at the Bayesian MCC tree are indicated in the graphics. Countries represented are BJ (Benin), GH (Ghana), GW (Guinea-Bissau), NG (Nigeria), SN (Senegal) and other for countries with minor contributions.

Conclusions

- These results support the existence of at least five independent introductions of the CRF02_AG lineage from West Africa into Brazil and further indicate that at least two of these lineages have been disseminated in the Rio de Janeiro state for about 30 years.

- The influx of HIV-1 CRF02_AG strains of African origin into Brazil unveil the existence of unrecognized transmission networks linking this country to West Africa.

Financial support: FAPERJ Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro IOC Instituto Oswaldo Cruz