

# Genetic and Biologic Characterization of HIV Type 1 Subtype C Isolates from South Brazil

JOANA P. MONTEIRO,<sup>1</sup> GERALDO A. FERRARO,<sup>1,2</sup> TÚLIO OLIVEIRA,<sup>3</sup> LUCIANO Z. GOLDANI,<sup>4</sup>  
SIMONE KASHIMA,<sup>5</sup> LUIZ C.J. ALCANTARA,<sup>1,2</sup> MARIZA G. MORGADO,<sup>6</sup>  
DUMITH CHEQUER BOU-HABIB,<sup>7</sup> and BERNARDO GALVÃO-CASTRO<sup>1,2</sup>

## ABSTRACT

The molecular and biological properties of HIV-1 subtype C strains from South Brazil were investigated. We sequenced *gag* and *env* fragments of viruses from 22 HIV-1-infected individuals from Porto Alegre City, which has the highest frequency of subtype C in the country. The sequences were then compared with other subtype B, C, and F strains isolated in Brazil and other countries using phylogenetic methods. Amino acid signatures were identified and correlated with phenotypic characteristics. We identified six strains with subtype C (27.3%), eight subtype B (36.4%), one subtype F (4.5%), six C/B recombinants (27.3%), and one B/F recombinant (4.5%). The Brazilian subtype C sequences formed a unique phylogenetic group and presented 6 and 18 specific amino acid signatures in *gag* and *env*, respectively. Three distinct patterns of C/B recombinants presented characteristic Brazilian amino acid substitutions. Subtype C viruses were predominantly R5 and non-syncytium-inducing, while C/B recombinants were R5/X4 and syncytium-inducing viruses. These findings suggest that subtype C viruses circulating in Brazil are the result of a unique introduction into the country. Recombination events between subtypes B and C have been occurring frequently for more than 10 years in South Brazil. Biological characterization confirms the hypothesis that subtype C is distinct from the others in the evolution of coreceptor utilization.

## INTRODUCTION

THE RELATIONSHIP BETWEEN THE GENETIC VARIABILITY, biological properties, and pathogenicity of HIV-1 is not well established, in part because the majority of research efforts have focused on subtype B, which is the most prevalent subtype in the western world. However, subtype C is currently the most prevalent genetic form of HIV-1, accounting for more than 56% of all infections worldwide.<sup>1</sup> The rapid spread of subtype C has been related to host, viral, and socioeconomic factors. There is growing evidence that molecular and biological characteristics distinguish subtype C viruses from other subtypes, and that such differences may lead to increased fitness, possibly affecting transmission and pathogenesis.<sup>2–6</sup>

Although the incidence of HIV/AIDS in Brazil began to decline in 1996, infection rates have increased in recent years (<http://www.aids.gov.br>). The HIV/AIDS epidemic in Brazil is characterized by the predominance of subtype B; however, the increased prevalence of non-B subtypes<sup>7,8</sup> has contributed to the emergence of variant viruses that may eventually become resistant and/or undetectable with current diagnostic tests. Therefore, HIV-1 molecular surveillance programs are of paramount importance to systematically monitor the distribution of HIV-1 and document the introduction of new variants.

The subtype C epidemic has been increasing in Brazil since its introduction around 1990,<sup>9</sup> particularly in the South Region, where prevalence varies between 25% and 45%.<sup>10,11</sup> However, the few studies that have characterized subtype C strains cir-

<sup>1</sup>Laboratório Avançado de Saúde Pública (LASP), Centro de Pesquisa Gonçalo Moniz (CPqGM), Fundação Oswaldo Cruz (FIOCRUZ), Salvador, Bahia, Brasil.

<sup>2</sup>Escola Bahiana de Medicina e Saúde Pública (EBMSP), Fundação Bahiana para o Desenvolvimento das Ciências (FBDC), Salvador, Bahia, Brasil.

<sup>3</sup>Department of Zoology, University of Oxford, Oxford, UK.

<sup>4</sup>Unidade de Medicina Interna do Hospital de Clínicas de Porto Alegre/UFRS, Porto Alegre, RS.

<sup>5</sup>Hemocentro de Ribeirão Preto/USP, Ribeirão Preto, São Paulo, Brazil.

<sup>6</sup>Laboratório de AIDS e Imunologia Molecular and <sup>7</sup>Laboratório de Imunologia Clínica, Departamento de Imunologia, Instituto Oswaldo Cruz/FIOCRUZ, Rio de Janeiro, RJ, Brazil.

culating in Brazil have focused on the *pol* region, so the biological properties of Brazilian subtype C are largely unknown.<sup>12,13</sup> In this work, we assessed HIV-1 subtypes in Porto Alegre, a city in the South Region of Brazil with at least three circulating HIV-1 subtypes and the highest rate of subtype C infection in the country.<sup>10</sup> This study represents the first investigation taking into account the association of molecular and biological features of a panel of subtype C viruses in Brazil and contributes to the understanding of both its evolutionary history in South Brazil and the HIV-1 epidemic in this region.

## MATERIALS AND METHODS

### *Patient population*

In 2003, blood samples were obtained from 22 HIV-1-infected patients followed up at the Hospital de Clínicas in Porto Alegre (Rio Grande do Sul, Brazil). Subjects gave informed consent to participate in the study. Peripheral blood mononuclear cells (PBMCs) were isolated from blood samples by Ficoll-Hypaque density gradient centrifugation (Sigma Chemical Co., St Louis, MO) for viral isolation and DNA extraction.

### *Virus isolation and biological characterization*

HIV-1 isolation and expansion were performed according to WHO/UNAIDS guidelines.<sup>14</sup> To investigate each isolate's ability to replicate in monocyte-derived macrophages, these cells were infected with HIV-1-positive cell-free supernatants, as previously described.<sup>15</sup> The preferential coreceptor usage of the primary viruses was assessed by infecting U87 cells transfected with CD4 and chemokine receptors.<sup>15</sup> The isolate's ability to induce syncytium formation was evaluated by infecting MT-2, Sup-T1, and PM-1 CD4<sup>+</sup> T cell lines with these viruses, as described previously.<sup>15</sup> The CCR5-dependent isolate Ba-L and the CXCR4-using virus III-B were used as controls in all assays.

### *DNA isolation, polymerase chain reaction (PCR), and sequencing*

DNA was extracted using the Qiagen extraction kit (QIAGEN, Valencia, CA). A nested PCR assay was conducted for individual amplification of the *gag* (~450 bp) and *env* (~550 base pairs) regions using a Perkin Elmer 9600 Thermal Cycle. The *gag* amplification used H1G777/H1P202 and H1Gag1584/G17 as outer and inner primer sets, while the *env* amplification used ED5/ED12 and ED31/ED33, respectively, as outer and inner primers. PCR conditions were as described elsewhere.<sup>16,17</sup> PCR products were purified in Qiagen columns (QIAGEN, Valencia, CA) and sequenced in an ABI 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA) using a Big Dye Terminator kit (Applied Biosystems, Foster City, CA) and the same PCR inner primers. Sequences were assembled using SeqMan software (DNASStar, Madison, WI) and were reported to the GenBank database under accession numbers DQ358756 to DQ358798 and DQ394366.

### *Phylogenetic analysis*

For subtype determination, sequences were aligned with an HIV-1 subtype reference set from the Los Alamos database

(<http://hiv-web.lanl.gov>) in CLUSTAL X<sup>18</sup> and manually edited in GENEDOC.<sup>19</sup> Phylogenetic trees were constructed using the neighbor-joining method with the TVM + I + G (*env*) and GTR + I + G (*gag*) models of substitution implemented in PAUP\*4.0b2a.<sup>20</sup> The reliability of each cluster was determined using 1000 bootstrap replicates. Trees were drawn with the TreeView program<sup>21</sup> and the subtyping results were confirmed using the REGA HIV-1 subtyping tool.<sup>22</sup>

To examine intrasubtype C relationships, a subset of *gag* and *env* subtype C sequences (including our new Brazilian sequences, previously characterized Brazilian sequences, and multiple subtype C sequences from other countries) was submitted to phylogenetic analysis as described above, using the GTR + I + G model of substitution. Mean genetic distances within and between groups were measured with the Kimura two-parameter model implemented in MEGA 2.0,<sup>23</sup> using 500 bootstrap replicates to calculate the standard deviations. To identify amino acid patterns, nucleotide sequences were translated in GENEDOC, and amino acid sequences were compared to the world consensus subtype C using the BIOEDIT program.<sup>24</sup> Sequences were screened for the presence of biologically important sites with the Prosite database.<sup>25</sup>

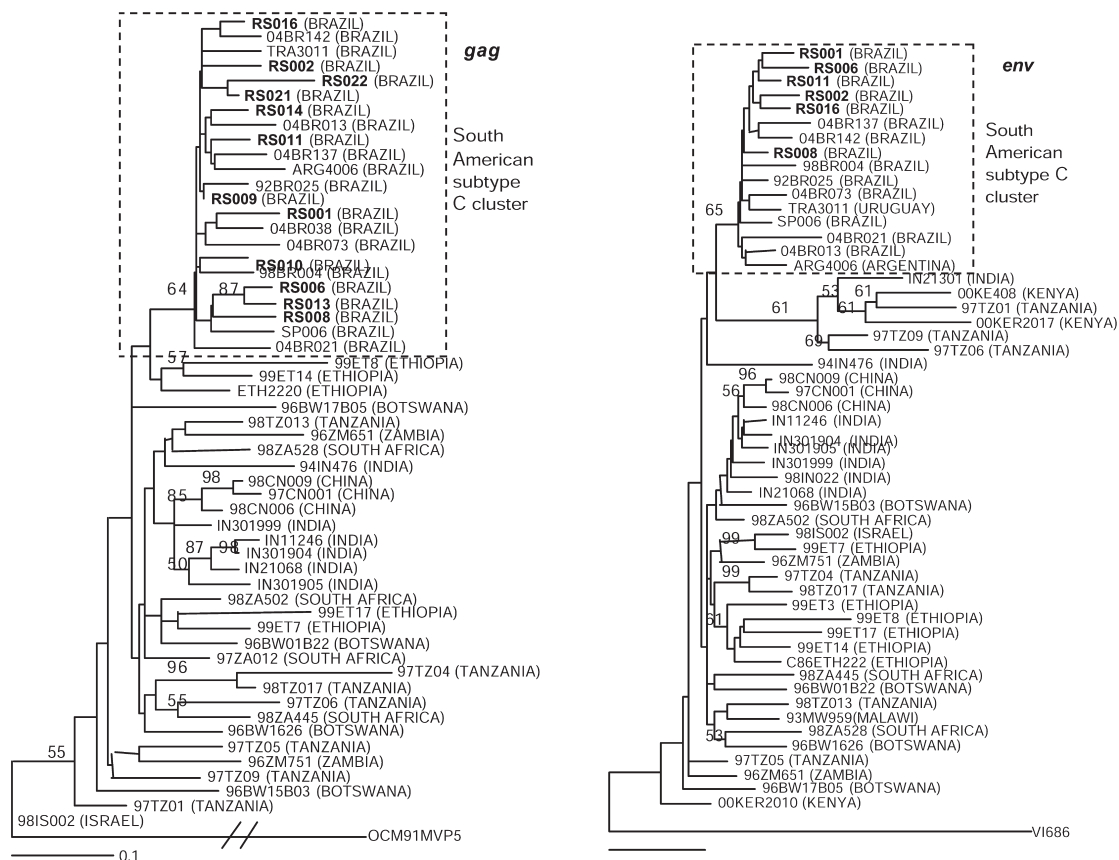
Samples with discordant subtype classifications had their *gag* and *env* sequences attached (1000 bp) and analyzed using the bootscanning method implemented in SIMPLOT software.<sup>26</sup> Based on these 1000 base pair sequences, a phylogenetic tree was constructed using the methodology described above. To further investigate the relationship between mosaic viruses circulating in Brazil, each recombinant 1000-bp fragment was also submitted to bootscanning analysis, using reference strains from subtypes A–K and one of the other Brazilian recombinants.

## RESULTS

Among the 22 HIV-1-infected individuals, 9 (41%) were women (male-to-female ratio of 1.44), 68% reported heterosexual behavior, and 41% were intravenous drug users (IDU). Mean age was 36 years, and mean income was \$293/month. Most of the patients (72.7%) had opportunistic disease and/or CD4 counts typical of AIDS status (Centers for Disease Control and Prevention classification—<http://www.cdc.gov/hiv/pubs/guidelines.htm>). Except for individuals RS021 and RS022, who were twin brothers and IDU, the patients were not epidemiologically related.

Phylogenetic analysis of *env* and *gag* regions showed that 15 specimens (68%) had concordant subtypes in both genomic regions: 8 subtype B (36.4%), 6 subtype C (27.3%), and 1 subtype F (4.5%). Discordant subtypes were observed for 7 (32%) samples, suggesting the possibility of recombinant viruses. Through recombination analysis using bootscanning, we classified 6 (27.3%) of these viruses as subtype C in *gag* and subtype B in *env* and one (4.5%) as subtype B in *gag* and subtype F in *env* (data not shown).

To examine the intrasubtype relationships between subtype C strains, trees were constructed using published C sequences from different countries (Fig. 1). All new and previously described Brazilian subtype C (C<sub>BR</sub>) sequences, including strains from Porto Alegre isolated between 1992 and 2003, as well as one sequence from Argentina and one sequence from Uruguay,



**FIG. 1.** *gag* and *env* phylogenetic relationships among subtype C HIV-1 viruses from different countries. Group O sequences were used as the outgroup. All new and previously described Brazilian subtype C sequences formed a monophyletic group shown within the box. Trees were constructed using the neighbor-joining method and the bootstrap values for 1000 replicates are indicated. Branches highly supported by the ML method are indicated (\*\**p* < 0.001). HIV-1 sequences generated in the present study are indicated in bold and those presenting discordant subtyping are underlined.

grouped as a single monophyletic group in both *gag* and *env* neighbor-joining and maximum-likelihood analyses. Contrasting, sequences from all other countries were widely dispersed across multiple clusters. Concerning the levels of viral diversity within each geographic group (Table 1), C<sub>BR</sub> sequences exhibited lower levels of divergence than those from other countries. Brazilian subtype B presented greater intrasubtype

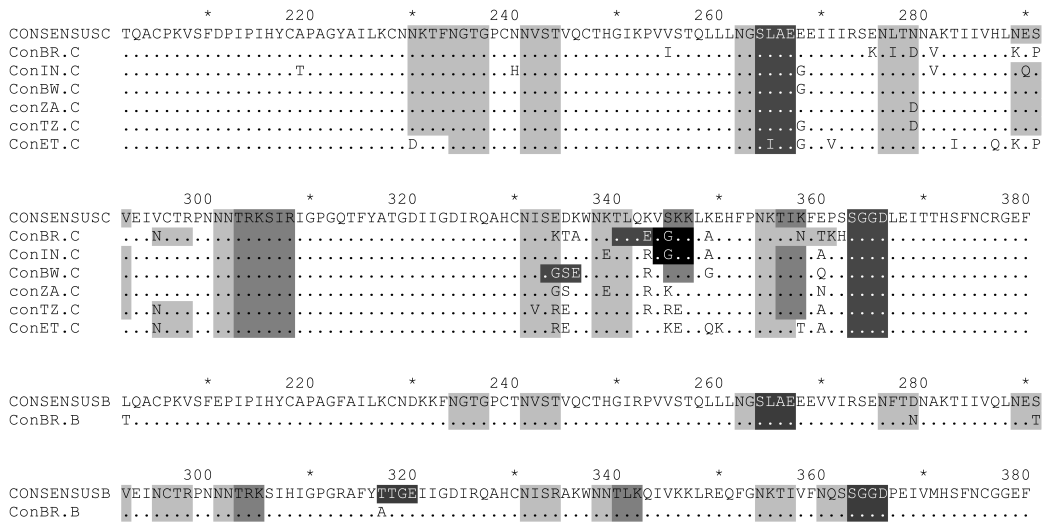
diversity than Brazilian subtype C in both *gag* and *env* regions. No significant differences in the mean distances between groups were observed.

We next investigated whether C<sub>BR</sub> sequences had characteristic amino acid substitutions (Fig. 2). In the 180 amino acid fragment of envelope (positions 202–382 in HXB2), considering only the substitutions with frequency equal to or greater

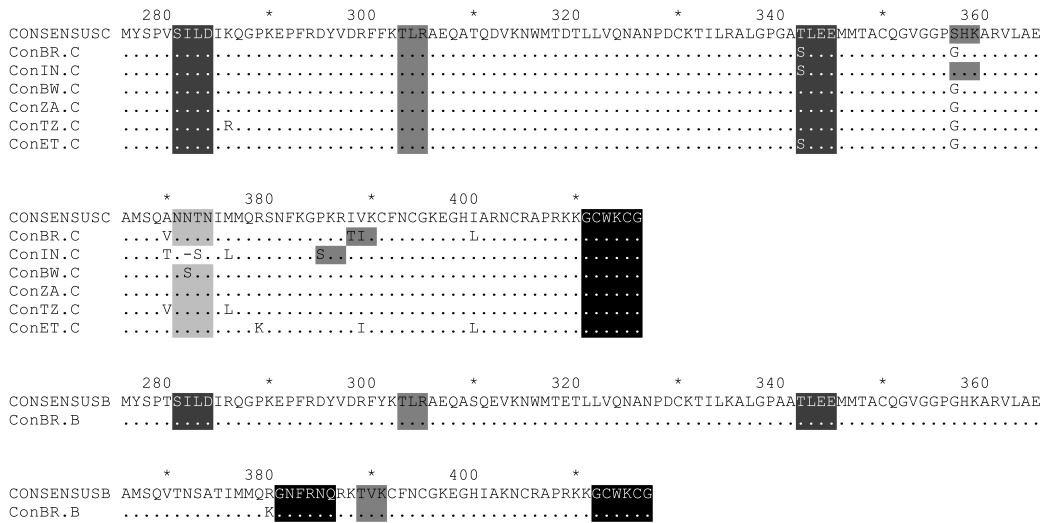
TABLE 1. DNA DISTANCES OF DIFFERENT GEOGRAPHIC GROUPS OF SUBTYPE C SEQUENCES AND BRAZILIAN SUBTYPE B SEQUENCES

| Country and ethnic group | Subtype | Mean distance % (no. of viruses) [standard deviation] |                 |                |
|--------------------------|---------|---|-----------------|----------------|
|                          |         | <i>env</i> (C2–C3)                                    | <i>env</i> V3   | <i>gag</i>     |
| Brazil                   | C       | 7.6 (14) [0.7]  | 4.4 (14) [1.2]  | 3.7 (20) [0.4] |
| India                    | C       | 8.6 (12) [0.7]  | 6.4 (12) [1.3]  | 5.0 (14) [0.6] |
| Botswana                 | C       | 12.4 (13) [0.9]                                       | 10.5 (13) [1.7] | 6.5 (15) [0.6] |
| South Africa             | C       | 12.1 (15) [0.8]                                       | 12.2 (15) [2.0] | 5.7 (14) [0.6] |
| Tanzania                 | C       | 15.8 (7) [1.3]  | 12.8 (7) [2.7]  | 7.0 (7) [0.8]  |
| Ethiopia                 | C       | 12.0 (8) [1.1]  | 13.1 (8) [2.4]  | 6.0 (7) [0.7]  |
| Brazil                   | B       | 14.4 (14) [1.1]                                       | 15.3 (14) [2.3] | 8.0 (10) [0.8] |

(a) *env*



(b) *gag*



(c) V3 loop

|             | 11           | 18                      | 25     | 29        |                |
|-------------|--------------|-------------------------|--------|-----------|----------------|
| CONSENSUS C | CTRPNNNTRKSI | RIGPGQTFYATGDIIGDIRQAHC |        |           |                |
| CONSENSUS B |              | H...RA..T..E.....       |        |           |                |
| RS001 (C)   |              |                         | E..... |           | +3 R5/X4 SI MT |
| RS002 (C)   |              |                         | R..... | K.....    | +5 R5 NSI MT   |
| RS006 (C)   |              |                         | E..... | E.....Y.. | +1 R5 NSI MT   |
| RS008 (C)   |              |                         |        |           | +3 R5 NSI MT   |
| RS016 (C)   |              |                         |        | E.....    | +2 R5 NSI MT   |
| RS021 (C/B) |              | P..F.RVL                |        |           | +4 R5/X4 SI MT |
| RS022 (C/B) |              | P..F.RVL                |        | K.....    | +6 R5/X4 SI MT |

**FIG. 2.** Multiple alignment of consensus subtype C amino acid sequences for different geographic groups of *gag* (a) and *env* (b) compared with the world consensus C sequence, and signature patterns for each of these groups. Brazilian subtype B sequences were also aligned against the world consensus B. Functional motifs are shaded differently: N-linked glycosylation (light gray); protein kinase C phosphorylation (medium gray); casein kinase II phosphorylation (dark gray); amidation site (black). (c) Summary of mutations in the V3 region of the seven HIV-1 isolates studied. Amino acid positions above the alignment are numbered according to their positions in the V3 loop. Net charges and the biological properties of each isolate are indicated at the right of the sequences. R5, tropic for CCR5 coreceptor; R5X4, tropic for both CCR5 and CXCR4 coreceptors; NSI, non-syncytium-inducing when exposed to MT-2 and Sup-T1 cell lines; SI, syncytium-inducing when exposed to MT-2 and Sup-T1 cell lines; MT, able to replicate in macrophages.

than 50%, C<sub>BR</sub> sequences ( $n = 12$ ) differed at 18 positions from the world consensus C (consC). Of these, 11 substitutions were exclusive of C<sub>BR</sub> sequences: V255I (50%), E275K (92%), L277I (58%), E334K (58%), D335T (58%), K336A (83%), K343E (75%), K358N (50%), E360T (75%), P361K (58%), and S362H (92%). In *gag*, C<sub>BR</sub> sequences ( $n = 18$ ) showed 6 changes in 141 amino acids (positions 276–417 in HXB2) when compared to the world consensus, one of which (I388T, 83%) represented exclusive substitution of Brazilian sequences.

To evaluate whether the C<sub>BR</sub> amino acid substitutions were restricted to Brazilian subtype C sequences, we looked for these signatures in our set of subtype B (B<sub>BR</sub>) sequences. In *env* ( $n = 13$ ), only the V295N (92%) substitution was present among B<sub>BR</sub> sequences, while in *gag* ( $n = 10$ ), three signatures were also present in subtype B: S357G (80%), A370V (70%), and I388T (50%). When we compared B<sub>BR</sub> with the world consensus subtype B to investigate a possible signature pattern, four significant substitutions were found in *env*: L202T (85%), D279N (61%), S291T (54%), and T317A (77%), and one specific signature (R380K, 50%) was observed in *gag*.

The significance of C<sub>BR</sub> amino acid substitutions was evaluated through their impact on functional motifs. In the consensus C, 9 potential N-linked glycosylation sites were found in the *env* fragment, 8 of which were conserved among C<sub>BR</sub> sequences. Substitutions N289K and S291P were associated with elimination of the N-linked glycosylation site at positions 289–293; however, the V295N amino acid change led to a new adjacent N-linked site (positions 295–298). Another alternative N-linked site (positions 358–361) appeared as a consequence of substitutions K358N, E360T, and P361K. C<sub>BR</sub> substitutions S345G and K358N resulted in the loss of two potential protein kinase C phosphorylation (PKC) sites (positions 345–347 and 356–358, respectively), while K343E and S345G resulted in (respectively) a new casein kinase II phosphorylation motif (positions 340–343) and a new amidation site (positions 344–347). In *gag*, two C<sub>BR</sub> substitutions were associated with changes in functional motifs; S357G resulted in a loss of a PKC site (positions 357–359) and I388T led to the emergence of a new PKC site (positions 388–390).

Since we observed that 27.3% of our samples were potential C/B recombinants, we further characterized the phylogenetic relationships of these viruses. A phylogenetic tree was constructed based on the attachment of *gag* and *env* (1000 bp) sequences (Fig. 3a). All six recombinants clustered as a monophyletic group (bootstrap = 88) that fell between subtype C and subtype B clusters and included a C/B recombinant from Argentina (ARE195FL). Bootscanning analysis (Fig. 3b) showed that all recombinant C sequences (*gag*) were more similar to those of other recombinants than to reference strain sequences, which did not always occur with the recombinant subtype B region (*env*). However, the RS014, RS021, and RS022 isolates were more similar to each other in both regions than to other reference strains. The same was observed for isolates RS009 and RS010. The B region of isolate RS013 was always more similar to a subtype B reference strain than to other Brazilian recombinants. All recombinants presented the six *gag* amino acid substitutions observed in C<sub>BR</sub> sequences, except RS013, which did not have the T342S substitution. Recombinant sequences were found to contain B<sub>BR</sub> *env* substitutions L202T (83%), D279N (67%), and T317A (67%), and the V295N (83%)

substitution characteristic of B<sub>BR</sub> and C<sub>BR</sub> sequences. In addition, RS021 and RS022 were probably epidemiologically linked viruses. Their *gag* sequences presented 13 nucleotide mutations that resulted in five amino acid substitutions (3.1% of nucleotide pairwise distance), while their *env* fragments were characterized by 29 mutations that resulted in 16 amino acid substitutions (5.6% of mean distance).

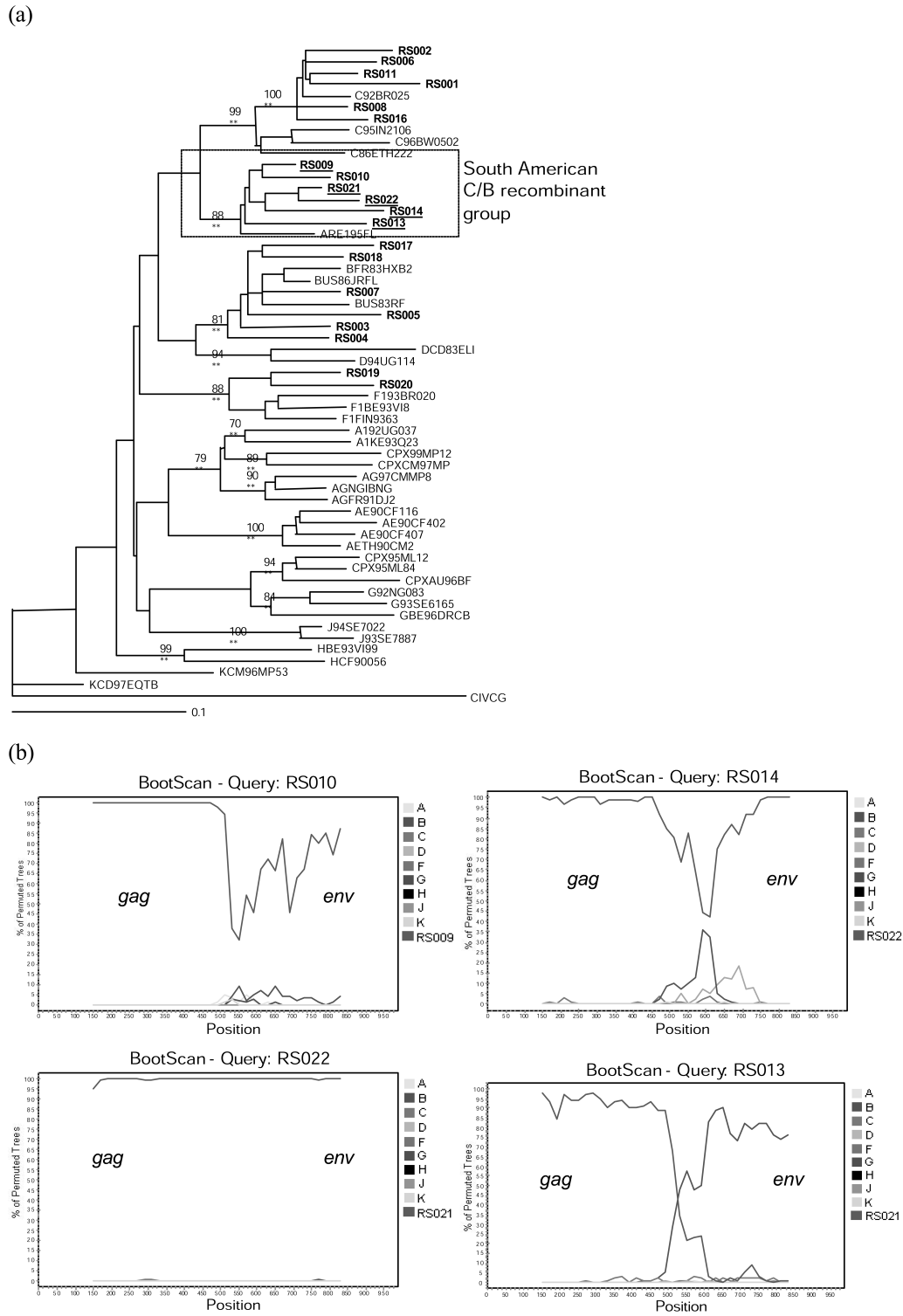
We isolated and investigated the biological properties of seven viruses from our cohort. All isolates were macrophage tropic in culture. Four of the five subtype C strains infected only U87 CCR5<sup>+</sup> cells (R5 viruses), while the fifth (RS001) was also able to use the CXCR4 coreceptor, albeit with lower cytopathogenicity relative to the use of CCR5. Both recombinants were able to use both CCR5 and CXCR4 coreceptors (R5/X4 viruses). The four subtype C R5 isolates did not form syncytium when exposed to MT-2 and Sup-T1 cell lines (NSI viruses), while the three dual tropic viruses induced syncytium in these cells (SI viruses). As expected, all isolates were able to form syncytia in CCR5<sup>+</sup> CXCR4<sup>+</sup> PM-1 cells.

Since coreceptor and T cell line tropisms are associated with basic amino acids in the V3 loop in subtype B and CRF01\_AE viruses, we characterized the V3 region of our isolates (Fig. 2c). None of the subtype C isolates had positively charged residues at positions 11, 25, and 29. However, isolate RS002 presented positive residues at positions 18 and 32, resulting in an overall charge increase of +5. A neutral serine residue was always present at position 11, while a negatively charged amino acid (either D or E) was present at position 25. The characteristic subtype C GPGQ motif at the tip of the loop was conserved in all but one subtype C isolate (RS002). In contrast, the C/B recombinants had a basic residue at position 18 (arginine, characteristic of B viruses) and RS022 also had another residue at position 25, increasing the overall charge to +6, which is characteristic of SI viruses. The potential N-linked glycosylation site NXT at positions 6–8, highly conserved in most HIV-1 isolates, was present in all of our isolates, and in all other Brazilian sequences except RS014. Overall, C<sub>BR</sub> V3 sequences were three times more conserved than B<sub>BR</sub> sequences (Table 1).

## DISCUSSION

Despite the small number of individuals enrolled in this present study, they presented features similar to the current tendencies of the Brazilian HIV/AIDS epidemic: (1) patients were primarily female (here, the sex ratio was 1.44, whereas a ratio of 4.0 was registered in 1993<sup>27</sup>), (2) there was a marked frequency in reported heterosexual behavior (68.2%), and (3) patients were of low socioeconomic status (on average).

The prevalence of subtype C was 27.3% for both *gag* and *env* genes and 55% when including recombinants involving subtype C. The high proportion (27.3%) of recombinant subtypes B and C viruses in our cohort is similar to that found by a study based on *pol* region analysis (25%).<sup>10</sup> The circulation of BC recombinant strains in Brazil was reported previously; however, the majority of these reports were based on *pol* analysis.<sup>10,11,28</sup> Other studies have found different patterns of BC recombination between different fragments of *gag* and *env*.<sup>29,30</sup> The C/B strains identified here presented a genetic structure similar to



**FIG. 3.** *gag* and *env* C/B recombinant viruses among Brazilian HIV-1 samples. Analyses were performed based on the attachment of *gag* and *env* sequences. (a) NJ tree showing the phylogenetic relationships of the Brazilian C/B recombinant viruses relative to other subtypes in the group M data set. The recombinant cluster is shown in the gray box and the bootstraps for the cluster is indicated. Branches highly supported by the ML method are indicated (\*\* $p < 0.001$ ). (b) Bootscan analysis showing the relationship between the Brazilian C/B recombinant sequences. In this analysis isolates (RS010 and RS009, RS022 and RS021, and RS022 and RS014) are more similar to each other than to reference sequences, while isolate RS013 is more similar to a subtype B reference strain in its *env* region than to other Brazilian recombinants.

an Argentine recombinant.<sup>31</sup> The monophyletic group and the observation of Brazilian signature amino acids in their *gag* and *env* sequences suggest local recombination events. Comparing the six C/B recombinants, we observed at least three distinct patterns of recombination, suggesting that recombination events are very frequent. Furthermore, we could speculate that recombination events between these two subtypes have been occurring for more than 10 years in Brazil, since patients RS021 and RS022 were seropositive in 1998, and the first BC recombinant in Brazil was detected by 1992–1993.<sup>29</sup>

The monophyletic group formed by C<sub>BR</sub> sequences in both *gag* and *env* confirms that this subtype was recently introduced in South Brazil from one or more similar founder strains, as previously suggested by *pol* evaluation.<sup>28</sup> This hypothesis is also supported by the fact that mean distances of C<sub>BR</sub> sequences are lower than those of subtype C viruses from Africa and of Brazilian subtype B viruses, which represent longer duration AIDS epidemics caused by multiple virus introductions.<sup>25,32,33</sup> C<sub>BR</sub> viruses had average distances more similar to those of Indian C viruses, which, as suggested, represent a recent and unique introduction of this subtype.<sup>33</sup> In addition, one virus from Argentina and another from Uruguay clustered inside the Brazilian group. Considering that subtype C was first detected in Argentina by 1999 and in Uruguay by 2001,<sup>34</sup> subtype C probably spread from Brazil.

The close relationship among C<sub>BR</sub> viruses is also supported by a pattern of signature amino acids in both genomic regions. These amino acid sites may have evolutionary, structural, and phenotypic significance. For instance, out of 18 sites differentially conserved in C<sub>BR</sub> sequences, 5 were proposed to be positively selected (positions 281, 291, 335, 336, and 343 of gp120).<sup>35</sup> We could also associate some C<sub>BR</sub>-specific signatures with lost or emergent N-linked glycosylation sites in the envelope. Carbohydrates in the gp120 limit the humoral immune response and help shield the virus from immune recognition.<sup>36,37</sup> Moreover, removal of N-linked glycans enhances CTL responses to HIV-1 *env*.<sup>38</sup> We observed, for instance, that Brazilian C sequences lacked the N-linked site at positions 289–293; however, a new site was found closely adjacent. A similar phenomenon was observed in the V1 region by Reitter *et al.*,<sup>36</sup> indicating a potential strategy for maintenance of these N-linked sites within gp120.

Although the five subtype C isolates were obtained from AIDS patients, these viruses preferentially used CCR5, similar to previous reports demonstrating that this subtype is distinct in the evolution of coreceptor utilization.<sup>39</sup> In association, V3 sequences of subtype C viruses were highly conserved and had lower overall positive charge than subtype B sequences, which is consistent with the NSI phenotype of subtype C viruses, as reported previously.<sup>40</sup> Indeed, none of C<sub>BR</sub> specific amino acid substitutions occurred in this genomic region. However, Brazilian exclusive substitution V295N created a new N-linked glycosylation site at position –1 of the V3 loop that is characteristically absent in most C viruses.<sup>25</sup> The absence of this site was associated with the presence of a serine phosphorylation site at position 11, which could contribute to an efficient use of CCR5 in subtype C viruses.<sup>25</sup> However, C<sub>BR</sub> viruses, which are effective in utilizing CCR5, have both N-linked and phosphorylation sites. Thus, further analysis may reveal the true importance of these sites for viral entry. In addition, the dual tropic

subtype C isolate (RS001) did not present basic amino acids in V3 sequences, and its net charge (+3) was similar to that of NSI viruses. This suggests that other regions of *env* may contribute to X4 tropism, as previously suggested.<sup>41,42</sup>

Among our samples, recombination events always occurred with subtype C in *gag* and subtype B in *env*. Since these patients presented advanced stages of disease, and considering that the R5-to-X4 switch occurs in at least 50% of subtype B-infected individuals,<sup>43</sup> the selection of these variants may be related to biological features of the subtype B envelope. Indeed, the two C/B recombinants were dual tropic and syncytium-inducing viruses, features more frequently observed among subtype B strains.

Our results point to a high genetic diversity of HIV-1 in South Brazil. One contributing factor is the elevated rate of recombinant events over a long period of time. However, since no circulating recombinant forms (CRF) containing Brazilian subtype C viruses have been described, studies involving the complete genome of Brazilian subtype C strains are needed. We also can conclude that subtype C isolates from South of Brazil are phenotypically similar to those in other regions.

## SEQUENCE DATA

The sequences were reported to the GenBank database under accession numbers DQ358756 to DQ358798 and DQ394366.

## ACKNOWLEDGMENTS

The authors are grateful to Hemocenter of Bahia and to Hemotherapy Service of the Hospital Clementino Fraga Filho (Federal University of Rio de Janeiro, RJ) for providing buffy coats from normal donors and to the Bioinformatic Laboratory of the LASP/CPqGM/FIOCRUZ. We thank Mrs. Elisabeth Deliege Vasconcelos and Mr. Craig Milroy for editing and revising this manuscript. This work was supported by grants from the Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB), the Brazilian Ministry of Health STD/AIDS Program, CNPq, and FIOCRUZ.

## REFERENCES

1. Esparza J and Bhamarapravati N: Accelerating the development and future availability of HIV-1 vaccines: Why, when, where, and how? *Lancet* 2000;355:2061–2066.
2. Salminen MO, Johansson B, Sonnerborg A, Ayejunie S, Gotte D, Leinikki P, *et al.*: Full-length sequence of an ethiopian human immunodeficiency virus type 1 (HIV-1) isolate of genetic subtype C. *AIDS Res Hum Retrovirus* 1996;12:1329–1339.
3. Carr JK, Salminen MO, Koch C, Gotte D, Artenstein AW, Hegerich PA, *et al.*: Full-length sequence and mosaic structure of a human immunodeficiency virus type 1 isolate from Thailand. *J Virol* 1996;70:5935–5943.
4. Ndung'u T, Renjifo B, and Essex M: Construction and analysis of an infectious human immunodeficiency virus type 1 subtype C molecular clone. *J Virol* 2001;75:4964–4972.
5. Montano MA, Novitsky VA, Blackard JT, Cho NL, Katzenstein DA, and Essex M: Divergent transcriptional regulation among ex-

- panding human immunodeficiency virus type 1 subtypes. *J Virol* 1997;71:8657–8665.
6. Blackard JT, Renjifo BR, Mwakagile D, Montano MA, Fawzi WW, and Essex M: Transmission of human immunodeficiency type 1 viruses with intersubtype recombinant long terminal repeat sequences. *Virology* 1999;254:220–225.
  7. Brazilian Network for HIV Isolation and Characterization: HIV-1 diversity in Brazil: Genetic, biologic, and immunologic characterization of HIV-1 strains in three potential HIV vaccine evaluation sites. *J Acquir Immune Defic Syndr* 2000;23:184–193.
  8. Morgado MG, Guimarães ML, and Galvão-Castro B: HIV-1 polymorphism: A challenge for vaccine development. *Mem Inst Oswaldo Cruz* 2002;97:143–150.
  9. Salemi M, de Oliveira T, Soares MA, Pybus O, Dumans AT, Vandamme AM, *et al.*: Different epidemic potentials of the HIV-1B and C subtypes. *J Mol Evol* 2005;60:598–605.
  10. Brindeiro RM, Diaz RS, Sabino EC, Morgado MG, Pires IL, Brígido L, *et al.*: Brazilian Network for Drug Resistance Surveillance (HIV-BResNET): A survey of chronically infected individuals. *AIDS* 2003;17:1063–1069.
  11. Rodrigues R, Scherer LC, Oliveira CM, Franco HM, Sperhacker RD, Paula Ferreira JL, *et al.*: Low prevalence of primary antiretroviral resistance mutations and predominance of HIV-1 clade C at polymerase gene in newly diagnosed individuals from south Brazil. *Virus Res* 2005;116:201–207.
  12. WHO Network for HIV Isolation and Characterization: HIV type 1 variation in World Health Organization-sponsored vaccine evaluation sites: Genetic screening, sequence analysis, and preliminary biological characterization of selected viral strains. *AIDS Res Hum Retroviruses* 1994;10:1327–1343.
  13. De Wolf F, Hogervorst E, Goudsmit J, Fenyo EM, Rubsamen-Waigmann H, Holmes H, *et al.*: Syncytium-inducing and non-syncytium-inducing capacity of human immunodeficiency virus type 1 subtypes other than B: Phenotypic and genotypic characteristics. WHO Network for HIV Isolation and Characterization. *AIDS Res Hum Retroviruses* 1994;10:1387–1400.
  14. Rubsamen-Waigmann H, von Briesen H, Holmes H, Bjorndal A, Korber B, Esser R, *et al.*: Standard conditions of virus isolation reveal biological variability of HIV type 1 in different regions of the world. *AIDS Res Hum Retroviruses* 1994;11:1401–1408.
  15. Ferraro GA, Mello MA, Suttmoller F, Van Weyenbergh J, Shindo N, Galvao-Castro B, *et al.*: Biological characterization and chemokine receptor usage of HIV type 1 isolates prevalent in Brazil. *AIDS Res Hum Retroviruses* 2001;17:1241–1247.
  16. Delwart EL, Shpaer EG, Louwagie J, McCutchan FE, Grez M, Rubsamen-Waigmann H, *et al.*: Genetic relationships determined by a DNA heteroduplex mobility assay: Analysis of HIV-1 env genes. *Science* 1993;262:1257–1261.
  17. Heyndriks L, Janssens W, Zekeng L, Musinda R, Anagonou S, Van Der Awerda G, *et al.*: Simplified strategy for detection of recombinant HIV-1 group M isolates by gag/env heteroduplex mobility assay. *J Virol* 2000;74:363–370.
  18. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins D: The CLUSTAL\_X Windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 1997;25:4876–4882.
  19. Nicholas KB, Nicholas HBJ, and Deerfield DW in GeneDoc: Analysis and visualization of genetic variation (EMBNEW NEWS 30), 1997.
  20. Swofford D: PAUP 4.0: Phylogenetic analysis using parsimony (and other methods), 4.0b2a. Sinauer Associates, Inc., Sunderland, MA, 1999.
  21. Page RD: TreeView: An application to display phylogenetic trees on personal computers. *Comput Appl Biosci* 1996;12:357–358.
  22. de Oliveira T, Deforche K, Cassol S, Salminen M, Paraskevis D, Seebregts C, *et al.*: An automated genotyping system for analysis of HIV-1 and other microbial sequences. *Bioinformatics* 2005;21:3797–3800.
  23. Kumar S, Tamura K, Jakobsen IB, and Nei M: MEGA2: Molecular evolutionary genetic analysis software. *Bioinformatics* 2001;17:1244–1245.
  24. Hall TA: BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *NAS* 1999;41:95–98.
  25. Gordon G, De Oliveira T, Bishop K, Coovadia HM, Madurai L, Engelbrecht S, *et al.*: Molecular characteristics of human immunodeficiency virus type 1 subtype C viruses from KwaZulu-Natal, South Africa: Implications for vaccine and antiretroviral control strategies. *J Virol* 2003;77:2587–2599.
  26. Salminen MO, Carr JK, Burke DS, and McCutchan FE: Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. *AIDS Res Hum Retroviruses* 1995;11:1423–1425.
  27. MINISTÉRIO DA SAÚDE: AIDS no Brasil: um esforço conjunto governo—sociedade/Coordenação Nacional de DST e AIDS, 1998.
  28. Soares MA, De Oliveira T, Brindeiro RM, Diaz RS, Sabino EC, Brígido L, *et al.*: A specific subtype C of human immunodeficiency virus type 1 circulates in Brazil. *AIDS* 2003;17:11–21.
  29. Cornelissen M, Kampinga G, Zorgdrager F, Goudsmit J, and UNAIDS Network for HIV Isolation and Characterization: Human immunodeficiency virus type 1 subtypes defined by env show high frequency of recombinant gag genes. *J Virol* 1996;70:8209–8212.
  30. Guimarães ML, Moreira AS, Loureiro R, Galvão-Castro B, The Brazilian Network for HIV Isolation and Characterization, and Morgado MG: High frequency of recombinant genomes in HIV type 1 samples from Brazilian southeastern and southern regions. *AIDS Res Hum Retroviruses* 2002;18:1261–1269.
  31. Aulicino PC, Kopka J, Rocco C, Mangano A, and Sen L: Sequence analysis of a South American HIV type 1 BC recombinant. *AIDS Res Hum Retroviruses* 2005;21:894–896.
  32. Brindeiro R, Vanderborght B, Caride E, Correa L, Oravec RM, Berro O, *et al.*: Sequence diversity of the reverse transcriptase of human immunodeficiency virus type 1 from Brazilian untreated individuals. *Antimicrob Agents Chemother* 1999;43:1674–1680.
  33. Shankarappa R, Chatterjee R, Learn GH, Neogi D, Ding M, Roy P, *et al.*: 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* 2001;75:10479–10487.
  34. Carrion G, Eyzaguirre L, Montano SM, Laguna-Torres V, Serra M, Aguayo N, *et al.*: Documentation of subtype C HIV type 1 strains in Argentina, Paraguay, and Uruguay. *AIDS Res Hum Retroviruses* 2004;20:1022–1025.
  35. Yamaguchi-Kabata Y and Gojobori T: Reevaluation of amino acid variability of the human immunodeficiency virus type 1 gp 120 envelope glycoprotein and prediction of new discontinuous epitopes. *J Virol* 2000;74:4335–4350.
  36. Reitter JN, Means RE, and Desrosiers RC: A role for carbohydrates in immune evasion in AIDS. *Nat Med* 1998;4:679–684.
  37. Chackerian B, Rudensey LM, and Overbaugh J: Specific N-linked and O-linked glycosylation modifications in the envelope V1 domain of simian immunodeficiency virus that evolve in the host alter recognition by neutralizing antibodies. *J Virol* 1997;71:7719–7727.
  38. Doe B, Steimer KS, and Walker CM: Induction of HIV-1 envelope (gp120)-specific cytotoxic T lymphocyte responses in mice by recombinant CHO cell-derived gp120 is enhanced by enzymatic removal of N-linked glycans. *Eur J Immunol* 1994;24:2369–2376.
  39. Cecilia D, Kulkarni SS, Tripathy SP, Gangakhedkar RR, Paranjape RS, and Gadkari DA: Absence of coreceptor switch with disease progression in human immunodeficiency virus infections in India. *Virology* 2000;271:253–258.



40. Korber BT, MacInnes K, Smith RF, and Myers G: Mutational trends in V3 loop protein sequences observed in different genetic lineages of human immunodeficiency virus type 1. *J Virol* 1994;68:6730–6744.
41. Carrillo A and Ratner L: Human immunodeficiency virus type 1 tropism for T-lymphoid cell lines: Role of the V3 loop and C4 envelope determinants. *J Virol* 1996;70:1301–1309.
42. Hu Q, Barry AP, Wang Z, Connolly SM, Peiper SC, and Greenberg ML: Evolution of the human immunodeficiency virus type 1 envelope during infection reveals molecular corollaries of specificity for coreceptor utilization and AIDS progression. *J Virol* 2000;74:11858–11872.
43. Connor RI, Sheridan KE, Ceradini D, Choe S, and Landau NR: Ccr5 in coreceptor use correlates with disease progression in HIV-1-infected individuals. *J Exp Med* 1997;185:621–628.

Address reprint requests to:

*Bernardo Galvão-Castro*

*LASP/CPqGM/FIOCRUZ*

*Rua Waldemar Falcão 121*

*Candeal, Salvador, Bahia, Brazil 40296-710*

*E-mail: bgalvao@cpqgm.fiocruz.br*