Virology Reports 5 (2015) 63–74

Contents lists available at ScienceDirect

Virology Reports

journal homepage: www.elsevier.com/locate/virep

The origin of HTLV-1 in southern Bahia by phylogenetic, mtDNA and β-globin analysis

Milena Magalhães Aleluia [a], Marco Antônio Gomes Mello [b], Luiz Carlos Junior Alcântara [b], Filipe Ferreira Almeida Rego [b], Lucas Pereira de Souza Santos [a], Bernardo Galvão-Castro [b,c], Marilda de Souza Gonçalves [b], Túlio de Oliveira [d], Lauro Juliano Marin [a], Sandra Mara Bispo Sousa [e], Sandra Rocha Gadelha [a,*]

[a] Universidade Estadual de Santa Cruz, Ilhéus, Bahia, Brazil
[b] Centro de Pesquisa Gonçalo Moniz/Fundação Oswaldo Cruz, Salvador, Bahia, Brazil
[c] Escola Baiana de Medicina e Saúde Pública/Fundação Bahiana para Desenvolvimento das Ciências, Salvador, Bahia, Brazil
[d] Trust-Africa Centre for Health and Population Studies and Southern African Treatment Research Network, South Africa
[e] Universidade Estadual do Sudoeste da Bahia, Vitória da Conquista, Bahia, Brazil

ARTICLE INFO

Article history:
Received 28 November 2014
Received in revised form 17 April 2015
Accepted 5 May 2015
Available online 12 May 2015

Keywords:
HTLV-1a-A
HTLV-1 in Brazil
Ancestry
Phylogeny
Viral spread

ABSTRACT

Different hypotheses have been elaborated to explain how the HTLV spread throughout the world. It has been proposed that the virus was introduced in Bahia, during the slave-trade period from the 16th century to 19th century. However, there is no information about the HTLV evolutionary history in southern Bahia. The phylogeny is fundamental in order to clarify its introduction and dispersion. The DNA of 29 samples was extracted, followed by nested-PCR assay for the LTR and DNA sequencing. These sequences were analyzed by phylogenic methods. The mtDNA ancestry markers and β^A-globin haplotypes were analyzed by PCR/RFLP. In relation to HTLV subtyping, all samples were classified as cosmopolitan subtype and transcontinental subgroup. Results suggest an ancient post-Columbian introduction of HTLV-1a-A associated with the slave trade between the XVI and late XIX centuries in southern Bahia. As regards the ethnicity of HTLV-infected women, the haplotype characterization of β-globin gene and the mtDNA ethnicity of HTLV-infected women, we have detected a major

* Corresponding author at: Universidade Estadual de Santa Cruz, Departamento de Ciências Biológicas, Rodovia Ilhéus-Itabuna, km 16, Salobrinho, Ilhéus, BA Cep 45.662-900, Brazil.
E-mail address: srgmello@uesc.br (S.R. Gadelha)

http://dx.doi.org/10.1016/j.virep.2015.05.002

African contribution, with a predominance of Benin and Bantu types. HTLV-1 infection is spread in Bahia and the point of origin was possibly Salvador.

1. Introduction

The Human T-cell lymphotropic virus type 1 (HTLV-1) infects approximately 5 to 10 million people in the world (Gessain and Cassar, 2012). The reported endemic areas for this virus are sub-Saharan Africa, Central and South America, Caribbean, Japan, Melanesia and the Middle East (Mueller, 1991 and review by; Proietti et al., 2005; Gonçalves et al., 2010). The phylogenetic analysis of LTR region classifies the HTLV-1 into seven subtypes: a, or cosmopolitan (Miura et al., 1994); b, or Central Africa (Vandamme et al., 1994); c, or Melanesia (Gessain et al., 1991); d, or Cameroon (Chen et al., 1995); e, or Democratic Republic of Congo (Salemi et al., 1998); f, of a Gabon individual (Salemi et al., 1998); and g, identified in Cameroon (Wolfe et al., 2005). The Cosmopolitan subtype, which is the most prevalent in Latin America, is divided into 5 subgroups: A, or transcontinental, B, or Japanese, C, or East Africa, D, or North Africa (Miura et al., 1994) and E, or Peru (Van Dooren et al., 1998).

The virus can be transmitted from mother-to-child, sexually and by parenteral exposure (Gessain and Cassar, 2012; Proietti et al., 2005). Some data in the literature reported low prevalence of HTLV-1 in pregnant women in Brazil (Ydy et al., 2009; Figueiró-Filho et al., 2007; Sequeira et al., 2012; Souza et al., 2012). However, in the southern Bahia, a prevalence of 1.05% has been detected among pregnant women (Mello et al., 2014). In fact, other studies conducted in Bahia state have found similar HTLV prevalences (around to 1.0%) among pregnant women (Santos et al., 1995; Bittencourt et al., 2001; Magalhães et al., 2008). Besides, analyses from the general population from Salvador-Bahia, have demonstrated that this city has the highest HTLV prevalence in the country (1.7%) (Dourado et al., 2003). Information based in the LTR phylogenetic analysis and historical data served as a basis to propose that the HTLV-1 was introduced in Bahia during the slave trade period, i.e., from the 16th to the 19th century (Alcantara et al., 2003; Rego et al., 2008). In fact, Bahia is the state with the greatest African influence in the country, as confirmed by the high number of African descendants (Krieger et al., 1965; Gattás et al., 2004).

The Brazilian population stems from the fusion of genetic diversity from three ethnic groups: Europeans, Amerindians and Africans, besides the influence of other groups (Gattás et al., 2004; Callegari-Jacques and Salzano, 1999). The native Amerindians, roughly 896.9 thousand, lived in Brazil up to 1500, when the European colonization started (they were mostly male and Portuguese). In addition to Portuguese, millions of immigrants such as Italian, Spaniards, Germans, Syrians, Lebanese, Japanese and others arrived in Brazil since 1820 (Instituto Brasileiro de Geografia e Estatística — IBGE Censo, 1938). The Africans came mainly from the 16th to the 19th century. The slave trade brought 4 million Africans to Brazil, especially from Sub-Saharan Africa. This region covers territories even from Senegal to Nigeria, which nowadays belongs to Angola, Congo and Mozambique (Curtin, 1969). Indeed, the Brazilian population is characterized as an admixed population.

Studies aimed at evaluating genetic ancestry in different HTLV-infected populations have been conducted in order to understand the prevalence of this virus and its history. The skin color, phenotypically evaluated, has a thin correlation with the ancestry, whereas the determination of ethnic contribution has been done by analysis of molecular markers (for example: mtDNA and β-globin haplotypes). Haplotypes of the 5′ region of the β-globin gene cluster have been used as an important tool to trace the origin, evolution and migration of humanity (Wainscoat et al., 1986). Besides that, the β-globin gene grouping reveals haplotypes associated with the presence of hemoglobin S in different ethnic and geographical origins: the type Benin (BEN), originated in west-central Africa; Bantu type (CAR), in south-central and eastern Africa; type (SEN) Senegal, in west Atlantic Africa; type (SAUDI), in Saudi Arabia, in the Indian subcontinent and east of the Arabian Peninsula; and the Cameroon type (CAM), along the west coast of Africa (Nagel, 1984; Hattori et al., 1986). Salvador has a high rate of ethnic admixture with strong African contribution, reflecting into an unusual haplotype distribution of the βS-globin gene when compared with those described in other Brazilian states, predominantly of heterozygous CAR/BEN (Azevedo et al., 1981; Gonçalves et al., 2003).

Due to the lack of information on the evolutionary history of HTLV in southern Bahia and the high prevalence of this virus in our state, information about its phylogeny is important so as to clarify its introduction and dispersion. The haplotype characterization of the β-globin gene and mtDNA complements the study as it generates solid and intrinsic data useful in establishing the geographic origin of the virus and the ethnicity of HTLV-infected women.

2. Materials and methods

2.1. Ethical statement

The Research Ethics Committee of the *Universidade Estadual de Santa Cruz* (UESC) has approved this study (Protocol 196/08). Express written consent was informed and obtained from all study participants.

2.2. Study area

To investigate the molecular characteristics of HTLV in southern Bahia, samples of 29 HTLV-1 seropositive women were analyzed. These samples are only part of a larger one which has been used to analyze HTLV-seroprevalence, risk factors associated with infection and maternal–fetal transmission in 2.759 pregnant women from São José Hospital/Santa Helena Maternidade, in Ilhéus, and Santa Casa de Misericórdia Manoel Novaes Hospital, in Itabuna (Mello et al., 2014). These are the two most important hospitals situated in the two biggest cities of southern Bahia. Besides that, these health facilities are the only ones which provide public health care, performing from 70% to 90% of child-births on SUS-provided hospitals. In addition, the analysis of mtDNA, required 29 HTLV-negative samples randomly selected using the SPSS program version 20.0.0. An error rate of 5% and a confidence level of 95% have been considered.

2.3. Interview questionnaire

Before blood collection, all of the women have answered a standardized questionnaire with socio-demographic and behavioral questions. It has used the information of self-reported skin color (black, brown, white, yellow and Indigenous), according to Instituto Brasileiro de Geografia e Estatística (IBGE) criteria.

2.4. Collection of samples

All samples were previously submitted to ELISA (ORTHO HTLV-1/HTLV-2) and the HTLV infection was confirmed by Western blot (HTLV WB 2.4; Genelabs Diagnostics, Singapore). After that, the DNA was extracted by a QIAgen Kit (QIAamp DNA Blood Kit) and a nested-PCR was conducted for the Long Terminal Repeat (LTR) region on HTLV-1. Two HTLV-1 overlapping fragments were amplified: a LTR-gag 473 bp and a LTR-tax 479 bp, as previously described by Alcantara et al. (2003). The PCR product was purified using the Qiagen PCR Purification Kit. DNA sequencing was performed using the Taq FS Dye terminator cycle sequencing kit (Applied Biosystems, Foster City, CA) on an automated 3100 genetic analyzer (Applied Biosystems Inc.) applying the identical inner primers of nested-PCR.

2.5. Phylogenetic analysis

The new HTLV-1 LTR sequences were submitted to the Laboratório Avançado de Saúde Pública (LASP) HTLV-1 Automated Subtyping Tool (Alcantara et al., 2011). These sequences and the reference sequences selected in the database GenBank/EMBL were aligned using the ClustalX software (Thompson et al., 1997). Alignment was manually edited using the programs GeneDoc and Se-Al (Nicholas et al., 1997). The first phylogenetic analyses were generated using the neighbor-joining (NJ) and maximum-likelihood (ML) methods implemented in the PAUP software version 4.0b10.19 (Swofford et al., 1996). The evolutionary model of Tamura–Nei with gamma distribution was selected as the best adaptive model for the data (parameter alpha 0.814563), using the software Modeltest 3.7 (Posada and Crandall, 1998). NJ trees were constructed with an optimized nucleotide substitution rate matrix and with parameter gamma distribution, employing empirical frequencies. The reliability of the NJ trees was evaluated by 1000 bootstrap pseudo-replicates.

The bootstrap values equal or above 60% were considered significant because the low genetic diversity on HTLV-1 LTR sequences. ML tree construction consisted in a heuristic search using the NJ tree, including its optimized parameters. The likelihood ratio test (LRT) was used to calculate the statistical support. Trees were visualized with the TreeView software, version 1.6.6 (Kumar et al., 1994).

Brazilian HTLV-1 LTR sequences available were downloaded from the GenBank database, including the recently sequenced strains from Ilhéus and Itabuna, for the Bayesian analysis. The dataset was aligned using the ClustalX software and manually edited using the Se-Al program. We have performed the Bayesian tree in duplicate and using the MrBayes program to verify the posterior probability (PP) statistical parameter. The time of the most recent common ancestor (TMRCA) from the sequences with the sampling year provided from GenBank and the new LTR sequences was estimated for the two main Latin American clusters (Latin American cluster A [LA_A] and Latin American cluster B [LA_B]) using the Beast package (Drummond and Rambaut, 2007). For the Beast analysis, we have assumed the previously described evolutionary rate (2×10^{-5} substitutions/site/year for one fixed mutation) (Alcantara et al., 2006). The analysis in 4 independent MCMC runs was carried out so as to enhance the result reliability, using the strict molecular clock with the constant growth tree priors, the effective sample size (ESS) was considered if above 200.

2.6. mtDNA marker analysis

The mtDNA ancestry markers were analyzed by PCR/RFLP. The amplification primers were used according to the conditions described by (Torroni et al., 1992; Torroni et al., 1993; Torroni et al., 1996) and digestions were performed with appropriate restriction endonucleases according to haplogroup: L3a/+2349 *MboI*; L3b/−8616 *MboI*; L3c/+10084 *TaqI*; L3d/−10394 *DdeI*; A/+663 *HaeII*; B/9-bp Deletion; C/−13259 *HincII*; and D/−5176 *AluI*. The haplogroups used are typical of sub-Saharan Africa: L3a, L3b, L3C and L3d, and original Amerindian haplogroups: A, B, C and D. The single-nucleotide polymorphism profile was determined using the previously established criteria (Chen et al., 1995; Torroni et al., 1996).

2.7. Analysis of nuclear DNA polymorphisms (AIMs — ancestry informative markers)

The AIMs presented a high differentiation of allele frequency (δ) between parental populations, therefore, they can be used to characterize the genetic composition of admixed populations (Parra et al., 2003). The autosomal markers represented by AIMs, LPL, APO (with values of $\delta > 40\%$ between African and European or Amerindian) and PV92 (with values of $\delta > 40\%$ between Amerindian and African or European) from autosomal DNA were analyzed. The amplification primers were used according to the conditions described by (Parra et al., 1998; Shriver et al., 2003). The LPL polymorphism was detected by RFLP and the APO and PV92 by Alu insertions discrimination.

2.8. β-Globin haplotype analysis

β^A-Globin haplotypes were amplified as previously described, generating fragments from the A-globin gene cluster (5′ε/3′β, 5′γ^A/3′β, 5′ψβ/3′β, ψβ, 3′ψβ). These fragments were purified by the Promega Wizard PCR prep system (Madison, WI), and a DNA fragment was digested with an appropriated restriction endonuclease (*XmnI*, *HindIII*, *HincII*) used for each site (Sutton et al., 1989). The fragments were analyzed by 2% agarose gel electrophoresis with SYBR green or gel red staining under ultraviolet light.

3. Results

Of the 29 HTLV-1 seropositive women, it was only possible to obtain samples from 21. All of them were successfully amplified and sequenced. The LASP HTLV-1 Automated Subtyping Tool has classified all the sequences as subtype “a” subgroup “A” with a bootstrap support of 100%.

3.1. Phylogenetic analysis

For phylogenetic analysis, the trees were reconstructed on the basis of two datasets. The first dataset contains the 713 base pair fragments of the LTR region (Fig. 1). The second dataset contains all the sequences

from the previous dataset and those recently identified in Mozambique (Fig. 2) with a fragment of 514 base pairs (Vicente et al., 2011). The first analysis was used to confirm the statistical significance for group nodes, using the neighbor-joining (NJ) and maximum-likelihood (ML) methods.

The entire paraphyletic cluster observed in the tree (Fig. 1) was also present in the analysis. It was observed that one sequence from southern Bahia (IL1787) had grouped with a sequence from Salvador (FNN159). Furthermore, another sequence from our study (IT196) had grouped with a sequence from São Francisco River Village (VSF287), and has been analyzed by (Rego et al., 2008) (Fig. 1). Besides that, it was observed that other sequences from southern Bahia (IL 945, IL245, IL174) had also grouped with other sequences from Salvador (FNN 158) (Fig. 1). We can therefore infer that HTLV-1 is spread over several regions of Bahia and, possibly, the infection came from the city of Salvador, since the harbor of All Saints' Bay had been one of the most important harbors since the beginning to the end of slave trading. All these analyses have high statistical support from bootstrap values (above 60% in 1000 replicates) and ML values, expressed as $^{*}p < 0.001$ (Fig. 1). In this same tree, in the Latin America group B, it was possible to observe a sequence from South Africa (HTLV 24) to group together with the Brazilian and the Latin American sequences (Fig. 1). In effect, the literature has demonstrated that Brazilian sequences are grouped with South African sequences (Van Dooren et al., 1998; Mota et al., 2007). Furthermore, data from our analyses corroborate the hypothesis that the Bantu population has migrated from Central Africa to South West Africa, as there have been sequences from Cameroon (2472LE) and Chile (CH26) with a common ancestor, denoting a high statistical support (Fig. 1).

In the second tree (Fig. 2), we can verify that some Mozambican sequences were grouped with Brazilian strains, but without a bootstrap support. The isolated IT190 was seen to have formed a paraphyletic group with a sequence of Mozambique (GU194508) in the Latin American cluster B with statistical significance for ML analysis ($p < 0.001$), but not statistically significant by bootstrap analysis. Besides the relationship among Brazilian and South African sequences, it was first presented the grouping among Bahia, while Mozambican sequences were recently identified.

The TMRCA analysis of the two main Latin American groups shows the ancient post Columbian introduction of the HTLV-1a-A in Latin America, being LA_A strains introduced in Brazil between 140 and 392 years ago with an ESS of 17676 and the LA_B strains introduced between 95 and 315 years ago with an ESS of 15430. The analyzed data provided high PP support (> 0.9) to the two subject clusters.

3.2. mtDNA marker analysis

The two groups, HTLV positive and negative controls (NC), have self-reported the color brown as the most frequent. These data were complemented with the mtDNA and β-globin analysis. mtDNA analysis has demonstrated that the most frequent haplogroup was the African (L3A — 27.5%, L3B — 6.8%, L3C — 37.9%, L3D — 10.3%), followed by Amerindian contribution (10.2% — haplogroup A — 3.4%, B — 3.4% and C — 3.4%) (Table 2). The analyzed African haplogroup characterizes Sub-Saharan African populations. Three of the HTLV positive samples were classified in the Amerindian haplogroup. All samples were HTLV-1. It was detected no HTLV-2. This fact reveals that the two types of HTLV 1 and 2 may have had different origins in our country.

3.3. Analysis of nuclear DNA polymorphisms (AIMs — ancestry informative markers)

As an Amerindian contribution of 10.2% ($n = 3$) was detected, an additional analysis was performed using autosomal ancestry informative markers (LPL, APO — African, PV92 — Amerindian). In this analysis, 33.3% were related to the marker LPL and 66.7% to the PV92, showing that two samples have Amerindian mitochondrial lineage in the nuclear and mitochondrial DNA.

3.4. β-Globin haplotype analysis

The analyses of β-globin were done in HTLV-1 positive samples of the 32 β-globin chromosomes studied: 14 (43.7%) BEN, 13 (40.6%) CAR, 3 (9.4%) SEN and 1 (3,1%) CAM. Only one sample was found to be Atypical (Atp). The genotype showed: BEN/BEN (31.25%), BEN/CAR (25.00%), CAR/CAR (12.50%), CAR/CAM (18.75%), CAR/SEN (6.25%), and CAR/Atypical (6.25%). Haplotype identities are indicated by combinations of plus and minus signs representing, respectively, the presence and absence of restriction sites (Table 1).

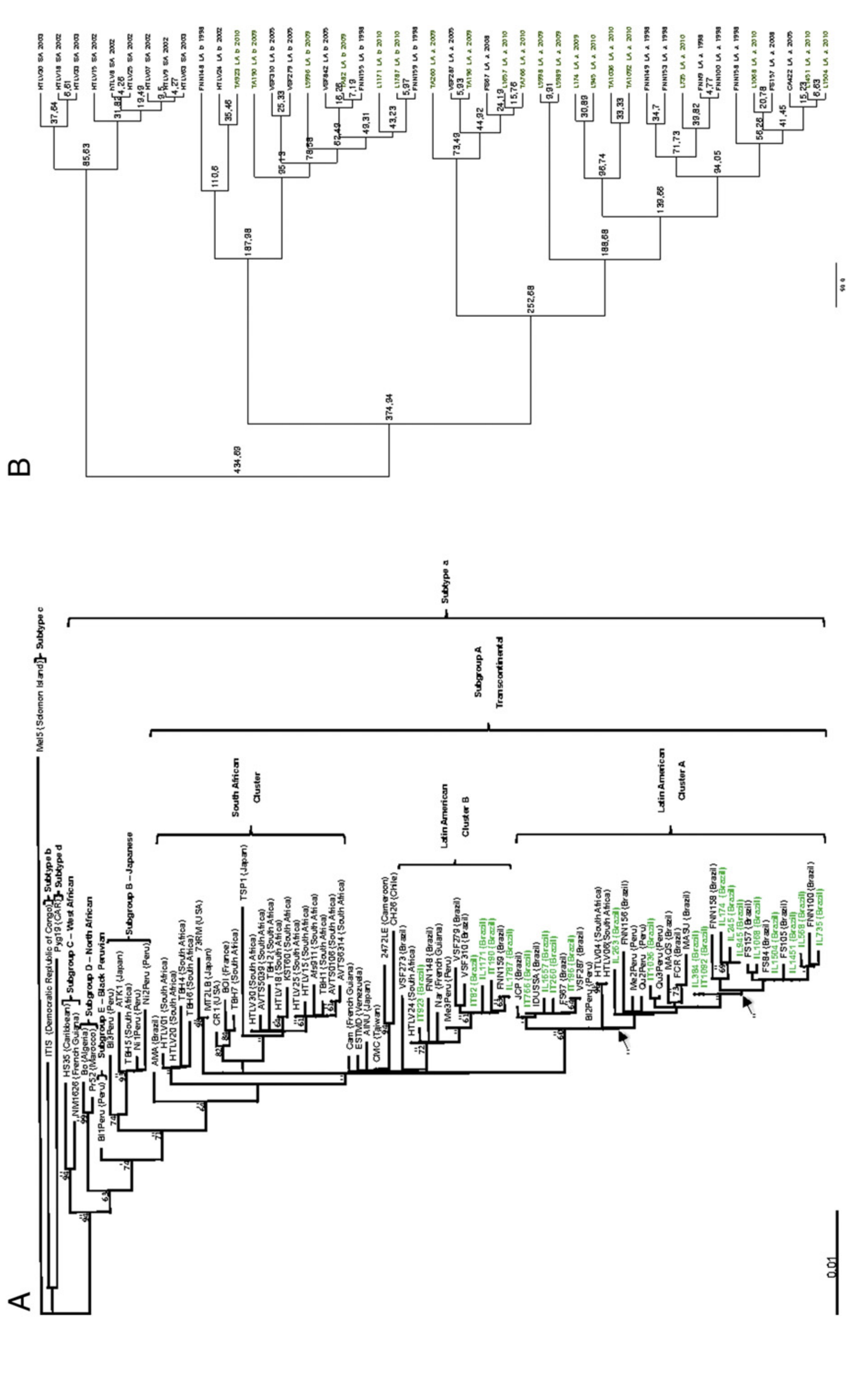
A
B
Subtype a
Subtype b
Subtype c
Subtype d
Subgroup A
Transcontinental
South African
Cluster
Latin American
Cluster A
Latin American
Cluster B
Subgroup B – Japanese
Subgroup C – West African
Subgroup D – North African
Subgroup E – Black Peruvian
0.01

Fig. 2. Rooted neighbor-joining tree of 21 HTLV-1 strains based on 514-bp fragment of the LTR region. NJ tree with 21 new sequences of HTLV and 96 references based on 514 bp fragment of LTR region of HTLV-1. The geographic origin is in parentheses. The new LTR sequences included in this analysis are in bold and green. The bootstrap values (above 60% in 1000 pseudo-replicates) represent the significance level for each branch corresponding to the number of times that the group has occurred in 1000 repetitions, providing statistical support to built tree. The isolated Mel5 was used as an outgroup to root the tree. The LRT method was considered highly statistically significant if $p < 0.001$ (**) and statistically significant if $p < 0.005$ (*).

4. Discussion

This study is part of a larger study that has analyzed the HTLV seroprevalence, risk factors related to the HTLV infection and maternal–fetal transmission in 2759 pregnant women from the São José Hospital/Santa Helena Maternity, in Ilhéus, and Santa Casa de Misericórdia Manoel Novaes Hospital, in Itabuna. The detected prevalence was 1.05% (Mello et al., 2014). Some data in the literature report lower prevalence in pregnant women in Brazil (Ydy et al., 2009; Figueiró-Filho et al., 2007; Sequeira et al., 2012; Souza et al., 2012).

Fig. 1. A — Rooted neighbor-joining tree of 21 HTLV-1 strains based on 713-bp fragment of the LTR region. B — Bayesian tree of HTLV-1 strains with collection date available on GenBank.A — Rooted neighbor-joining tree of 21 HTLV-1 strains based on 713-bp fragment of the LTR region. The geographic origin is in parentheses. The new LTR sequences included in this analysis are in bold and green. The bootstrap values (above 60% in 1000 pseudo-replicates) represent the significance level for each branch corresponding to the number of times that the group has occurred in 1000 repetitions, providing statistical support to built tree. The isolated Mel5 was used as an outgroup to root the tree. The LRT method was considered highly statistically significant if $p < 0.001$ (**) and statistically significant if $p < 0.005$ (*). B — Bayesian tree of HTLV-1 strains with collection date available on GenBank. The number in the nodes indicates the median TMRCA for the clade. The new LTR sequences are in green.

Table 1
Frequency of the haplotypes associated with the β-globin gene cluster in the patients studied.

Gene β-globin	5′ε/3′β	5′γ^A/3′β	5′ψβ/3′β	ψβ	3′ψβ	n (%)
Haplotypes	*Xmn I*	*Hind III*	*Hind III*	*Hinc II*	*Hinc II*	
CAR/Atp	−/−	+/+	−/−	−/−	+/−	6.25
BEN/BEN	−/−	−/−	−/−	−/−	+/+	31.25
BEN/CAR	−/−	+/−	−/−	−/−	+/−	25.00
CAR/CAR	−/−	+/+	−/−	−/−	−/−	12.50
CAR/CAM	−/−	+/+	+/−	−/−	+/−	18.75
CAR/SEN	+/−	+/+	−/−	+/−	+/−	6.25
Total						100

BEN (Benin), CAR (Central African Republic), CAM (Cameroon), SEN (Senegal) and Atp (Atypical).

However, prevalence rates of 0.9 and 1.0% were observed in two other cities from Bahia: Salvador and Cruz das Almas (Recôncavo), respectively (Santos et al., 1995; Bittencourt et al., 2001; Magalhães et al., 2008). These data suggest that the prevalence of HTLV-1 is an important public health problem in Bahia.

As regards the virus subtyping using the LASP HTLV-1 Automated Subtyping Tool (Version 1.0), only HTLV-1aA (cosmopolitan subtype, transcontinental subgroup) has been found. This subtype is the most prevalent in America. In addition, the phylogenetic analysis demonstrated that clusters grouped in Latin America (cosmopolitan subtype and transcontinental subgroup) had a common ancestor with sequences from South Africa (Van Dooren et al., 1998; Magalhães et al., 2008; Rego et al., 2008; Mota et al., 2007). The results were confirmed by phylogenetic analysis. Furthermore, all isolates of this study belonged to the cosmopolitan subtype (63% ML bootstrap $p < 0.001$) and the transcontinental subgroup (74% ML and bootstrap $p < 0.001$) (Fig. 1). Due to the grouping of sequences in the two Latin America clusters, it is suggested that there have been multiple HTLV post-Columbian introductions. This post-Columbian introduction is indicated by the presence of sequences in both African and Brazilian clusters. These sequences can be grouped together directly in the topology tree or through a common ancestor (Van Dooren et al., 1998; Mota et al., 2007) (Fig. 1).

Following phylogenetic analysis, it could be observed, even in the first tree (Fig. 1), that southern Bahia sequences grouped with sequences from Salvador city, as well as those with a sequence belonging to a small village located in the countryside of Bahia state (Rego et al., 2008). Furthermore, it was found that a sequence from Feira de Santana city (100 km away from Salvador) has also been grouped with sequences from southern Bahia and Salvador (Fig. 1). These analyses have strong statistical support from ML and bootstrap values (Fig. 1). These results allow for inferring that HTLV-1 infection is spread over several regions of Bahia and this infection has possibly originated from the slave trading in Salvador (Alcantara et al., 2006).

Concerning the introduction of HTLV in the Americas and Brazil, some studies reveal that the virus has been through human migrations, both during migrations of ancestral populations in the pre-Columbian era – from 15,000 to 35,000 years, through the Bering Strait (Greenberg et al., 1986; Bonatto and Salzano,

Table 2
Frequency of the haplotypes associated with the mtDNA in the patients studied.

Haplotypes	Haplotypes of mtDNA								Freq.
	A	B*	C	D	L3A	L3B	L3C	L3D	
	Amerindian				African				
1	−	−	−	−	+	−	−	−	10
2	−	−	−	−	−	−	+	−	11
3	−	−	−	−	−	+	−	−	3
4	−	−	−	−	−	−	−	+	2
5	+	−	−	−	−	−	−	−	1
6	−	+	−	−	−	−	−	−	1
7	−	−	+	−	−	−	−	−	1
N° of the patients									29
N° of the haplotypes									7

* Absent in haplogroup L3B.

1997) – as during the slave trade and/or Asian population migration in the post-Colombian era (Miura et al., 1994; Van Dooren et al., 1998; Yamashita et al., 1999). In fact, a population study analyzing the β-globin gene in samples of Salvador has served as a basis for the hypothesis that the virus was introduced from Africa to Salvador during the slave-trading period in the eighteenth century. Likewise, recent data suggest that the African Bantu ethnic groups were brought to Bahia in this period (Alcantara et al., 2006). In addition, there are reports that a migration of the Bantu population occurred in Central Africa to South West Africa about 3000 years ago (Curtin, 1969). Once many slaves from this area came to Salvador, it is possible that Bantu Africans had been to Brazil in that period. Therefore, it has been suggested that the HTLV-1 arrived in the city of Salvador in the post-Columbian time (Verger, 1976). Although the majority of the Africans brought to Bahia during the slave trade came from West Africa (Benin, Nigeria, and northern Angola), there is evidence that slaves were also brought from other different regions of Africa. Therefore, infected South Africans could have introduced the HTLV-1 Cosmopolitan subtype of the Transcontinental subgroup into Salvador and possibly into other areas in the Latin American continent. These data suggest the hypothesis that the Bantu population migrated from Central Africa to South West Africa.

As was previously demonstrated (Rego et al., 2008), the South African sequence was grouped with the Brazilian sequence (Fig. 2). Hence, we can observe the grouping with sequences from Mozambique. As one can see, the HTLV-positive sequence forms a paraphyletic group with a sequence from Mozambique in the Latin American cluster B (Fig. 2), thus demonstrating that, aside from the grouping of Brazilian and South African sequences, the grouping with sequences from Mozambique was also observed. It was explained with evidences that Africans were likewise brought from the southern regions of the continent (currently: southern Angola, South Africa and Mozambique) (Verger, 1976; Rego et al., 2008). It is also known that the departures of slave ships from ports were not necessarily related to the ethnic origins of Africans. Indeed, during the colonization of South Africa by the British in the seventeenth and eighteenth centuries, many Africans migrated to neighboring regions currently known as Angola, Madagascar and Mozambique, from where they were captured and transported in the slave trade to Salvador city. Bayesian analysis confirms the previous analysis, thus suggesting an ancient post-Columbian introduction of HTLV-1a-A related to the slave trade between the XVI and late XIX centuries.

In respect to the ethnic characterization of the analyzed population, mtDNA markers that characterize African populations and whose geographical distribution is in sub-Saharan Africa have been used (Soares et al., 2011). The used L3 haplogroup was implicated in Bantu expansion (Salas et al., 2002). The HTLV-1 Transcontinental sub-group strains from South Africa and Mozambique suggest a common origin of HTLV-1 in both countries. This can be explained by their common people and migration patterns, as well as their intense commerce (Melo et al., 2000). In the period from the fourteenth to the nineteenth century, approximately 4 million Africans from this region arrived in Brazil, mostly of Bantu and Sudanese origins (Gonçalves et al., 2003; Alcantara et al., 2006). The latter covers areas ranging from Senegal to Nigeria; regions that nowadays belong to Angola, Congo and Mozambique (Curtin, 1969).

In ancestry-related studies using mtDNA markers in samples, urban northeastern Brazilian, Amerindian (22%) and African (44%) contributions have been verified (Alves-Silva et al., 2000). As for the ethnic contribution in HTLV positive samples (Brucato et al., 2010), there has been a great African contribution (99.3%) in patients from the French Guiana region. We have detected a major African contribution similar to that from Salvador (Table 2) (Alcantara et al., 2003). Yet, we have only found African contribution and a small Amerindian contribution. In Brazil, a significant Amerindian contribution (10.2%) was only reported in HTLV-1 patients from the Amazon region, where the Amerindian contribution is much higher (Table 2) (Vallinoto et al., 2002). These data suggest that the HTLV may have had different origins.

Three samples with Amerindian contribution have been detected. In these cases, additional analyses were carried out using autosomal ancestry informative markers (AIMs) (LPL and APO — African and PV92 — Amerindian). The Amerindian contribution was confirmed, but there was no European contribution, corroborating previous studies which have not detected European contribution in the analysis of samples from urban and rural areas of Bahia (Mello et al., 2014).

Besides the African contribution detected by mtDNA haplotype analysis, the presence of β-globin haplotypes (nuclear DNA) demonstrating: 43.7% Benin type (BEN), 40.6% Bantu type (CAR), 9.4% Senegal type (SEN) and 3.1% Cameroon type (CAM) was verified. The BEN type is closely related to west-central Africa in the same way CAR is associated with south-central and eastern Africa; SEN is linked to West Africa, and CAMER is connected to the west coast of Africa (Table 1). These results complement data found in mtDNA

haplotypes. Benin type is related to the west-central African region and is of Bantu and Sudanese origins. In fact, Benin and its bordering regions, such as Congo, Angola, Senegal and Nigeria were part of the slave trade (Curtin, 1969; Gonçalves et al., 2003; Alcantara et al., 2006).

Regarding the history of Ilhéus, it is known that the Captaincy of "São Jorge dos Ilhéus" had a slavery period when Dom Pedro II, in 07/26/1534, donated a vast expanse of land, beginning the colonization of the region (Vinhães, 2001). After being a city, the region received people from Sergipe state and also slaves from different African regions (Carneiro, 1991). In 1924, cocoa farmers began the construction of the port, allowing the export of cocoa and the cultural and foreign exchange (Vinhães, 2001). Itabuna city, next to Ilhéus, began to be populated by cowboys from Sergipe state whose destination was Vitória da Conquista city (Andrade and Rocha, 2005). Sergipean immigrants initiated holdings on Cachoeira river banks. Jesuits helped to catechize the native indians *Pataxó*, *Guerren* and *Camacã*. In 1730 and 1790, pioneers braved the wilderness and the natives (Andrade and Rocha, 2005). Ilhéus and Itabuna used to have Amerindians from different tribes as slaves, reflecting the possible Amerindian influence detected in this study.

In summary, all virus analyzed in this study were classified as cosmopolitan subtype and transcontinental subgroup (HTLV-1aA). In addition, the phylogeny showed multiple introductions of the virus in Brazil while the TMRCA analysis confirms the previous analyses, suggesting an ancient post-Columbian introduction of HTLV-1a-A related to the slave trade between the XVI and late XIX centuries. Additionally, the haplotype characterization of the β-globin gene and mtDNA has complemented the study, generating solid data about the ethnicity of HTLV-infected women. These characterizations allowed for identifying the African ethnicity with a predominance of Benin and Bantu types. There are widespread expectations that more sequences will be identified and thus enable the construction of further phylogenetic analyses, including the characterization of the ethnic population so as to elucidate the virus' arrival hypotheses already drawn in Brazil.

HTLV-1 infection is spread in Bahia and the point of origin was possibly Salvador. The phylogenetic analysis suggests an ancient post-Columbian introduction of HTLV-1a-A in southern Bahia related to the slave trade. The major African contribution in the population, similar to that seen in Salvador, corroborates and strengthens these hypotheses. Besides, the grouping of Brazilian sequence with sequences of Mozambique suggests a possible route of the slave trade, coming from Mozambique.

Sequence data

The GenBank accession numbers of the new HTLV-1 fragments included in phylogenetic study were as follows: KF202307; KF202308; KF202309; KF202310; KF202311; KF202312; KF202313; KF202314; KF202315; KF202316; KF202317; KF202318; KF202319; KF202320; KF202321; KF202322; KF202323; KF202324; KF202325; KF202326; and KF202327.

Acknowledgments

Research reported in this publication was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB) and Universidade Estadual de Santa Cruz (UESC). These sequence analyses and bioinformatics were performed in the LASP/CPqGM/FIOCRUZ. The authors are grateful to Noilson Lázaro de Souza Gonçalves and Camylla Vilas Boas Figueiredo for their technical assistance in the HTLV serology and the β-globin haplotype analyses, respectively.

References

Alcantara, J.R.L.C., Van Dooren, S., Gonçalves, M.S., et al., 2003. Globin haplotypes of human T-cell lymphotropic virus type I-infected individuals in Salvador, Bahia, Brazil, suggest a post-Columbian African origin of this virus. J. Acquir. Immune Defic. Syndr. 33, 536–542.

Alcantara, L.C., Oliveira, T., Gordon, M., et al., 2006. Tracing the origin of Brazilian HTLV-1 as determined by analysis of host and viral genes. AIDS 20, 780–782.

Alcantara, L.C.J., Dooren, S.V., Vandamme, A.M., Galvao-Castro, B., Oliveira, T., 2011. LASP HTLV-1 Automated Subtyping Tool (Version 1.0). http://www.bahiana.edu.br/bioinfo/virus-genotype/html/subtypinghtlv.html (Accessed 08 November 2011).

Alves-Silva, J., Santos Da, S.M., Guimaraes, P.E., et al., 2000. The ancestry of Brazilian mtDNA lineages. Am. J. Hum. Genet. 67, 444–461.

Andrade, P.M., Rocha, B.L., 2005. De Tabocas a Itabuna. Um Estudo Histórico –Geográfico/Concepção e Organização. Editus, Itabuna.

Azevedo, E.S., Silva, K.M. Da, Silva, M.C., et al., 1981. Genetic and anthropological studies in the island of Itaparica, Bahia, Brazil. Hum. Hered. 31, 353–357.

Bittencourt, A.L., Dourado, I., Filho, P.B., 2001. Human T-cell lymphotropic virus type 1 infection among pregnant women in northeastern Brazil. J. Acquir. Immune Defic. Syndr. 26, 490–494.
Bonatto, S.L., Salzano, F.M., 1997. A single and early migration for the peopling of the Americas supported by mitochondrial DNA sequence data. Proc. Natl. Acad. Sci. 94, 1866–1871.
Brucato, N., Cassar, O., Tonasso, L., et al., 2010. The imprint of the Slave Trade in an African American population: mitochondrial DNA, Y chromosome and HTLV-1 analysis in the Noir Marron of French Guiana. BMC Evol. Biol. 10, 314.
Callegari-Jacques, S.M., Salzano, F.M., 1999. Brazilian Indian/non-Indian interactions and their effects. Cienc. Cult. 51, 166–174.
Carneiro, E., 1991. Religiões negras e negros bantos. Civilização brasileira 3rd ed. (Ilhéus).
Chen, J., Zekeng, L., Yamashita, M., 1995. HTLV isolated from a Pygmy in Cameroon is related but distinct from the known Central African type. AIDS Res. Hum. Retrovir. 11, 1529–1531.
Curtin, P.D., 1969. The Atlantic slave trade. A Census. University of Wisconsin Press, Madison.
Dourado, I., Alcântara, L.C., Barreto, M.L., Teixeira, M.G., Galvao-Castro, B., 2003. HTLV-1 in the general population of Salvador, Brazil: a city with African ethnic and sociodemographic characteristics. J. Acquir. Immune Defic. Syndr. 34, 527–531.
Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.
Figueiró-Filho, E.A., Senefonte, F.R., de Lopes, A.H., Morais, O.O. Souza, Júnior, V.G., Maia, T.L., et al., 2007. Frequency of HIV-1, rubella, syphilis, toxoplasmosis, cytomegalovirus, simple herpes virus, hepatitis B, hepatitis C, Chagas disease and HTLV I/II infection in pregnant women of State of Mato Grosso do Sul. Rev. Soc. Bras. Med. Trop. 40, 181–187.
Gattás, G.J.F., Gomes, L., Kohler, P., 2004. Ethnicity and glutathione S-transferase (GSTM1/GSTT1) polymorphisms in a Brazilian population. Braz. J. Med. Biol. Res. 37, 451–458.
Gessain, A., Cassar, O., 2012. Epidemiological aspects and world distribution of HTLV-1 infection. Front. Microbiol. 3, 388.
Gessain, A., Yanagihara, R., Francini, G., 1991. Highly divergent molecular variants of human T-lymphotropic virus type from isolated populations in Papua New Guinea and the Solomon Islands. Proc. Natl. Acad. Sci. U. S. A. 88, 7694–7698.
Gonçalves, M.S., Bomfim, G.C., Maciel, E., et al., 2003. β^S-Haplotypes in sickle cell anemia patients from Salvador-Bahia-Brazil. Braz. J. Med. Biol. Res. 36, 1283–1288.
Gonçalves, D.U., Proietti, F.A., Ribas, R.J.G., et al., 2010. Epidemiology, treatment, and prevention of human T-cell leukemia virus type 1-associated diseases. Clin. Microbiol. 23, 577–589.
Greenberg, J., Turner, C., Zegura, S., 1986. The settlement of the Americas: a comparison of the linguistic, dental and genetical evidence. Curr. Anthropol. 27, 477–497.
Hattori, Y., Kutlar, F., Kutlar, A., et al., 1986. Haplotypes of beta S chromosomes among patients with sickle cell anemia from Georgia. Hemoglobin 10, 623–642.
Instituto Brasileiro de Geografia e Estatística — IBGE Censo, 1938. www.censo2010.ibge.gov.br (Accessed 05 December 2012).
Krieger, H., Morton, N.E., Mi, M.P., Azevedo, E., Freire-Maia, A., Yasuda, N., 1965. Racial admixture in Northeastern Brazil. Ann. Hum. Genet. 19, 113–125.
Kumar, S., Tamura, K., Nei, M., 1994. MEGA: Molecular Evolutionary Genetics Analysis software for microcomputers. Comput. Appl. Biosci. 10, 189–191.
Magalhães, T., Mota-Miranda, A.C., Alcantara, L.C., 2008. Phylogenetic and molecular analysis of HTLV-1 isolates from a medium sized town in northern of Brazil: tracing a common origin of the virus from the most endemic city in the country. J. Med. Virol. 80, 2040–2045.
Mello, M.A., Conceição, F.A., Sousa, S.M., Alcântara, L.C., Marin, L.J., Raiol, M.R.S., Boa-Sorte, N., Santos, L.P., Almeida, M., Galvão, T.C., Bastos, R.G., Lázaro, N., Galvão-Castro, B., Gadelha, S.R., 2014. HTLV-1 in pregnant women from the Southern Bahia, Brazil: a neglected condition despite the high prevalence. Virol. J. 11, 28.
Melo, J., Beby-Defaux, A., Faria, C., et al., 2000. HIV and HTLV prevalences among women seen for sexually transmitted diseases or pregnancy follow-up in Maputo, Mozambique. J. Acquir. Immune Defic. Syndr. 23, 203–204.
Miura, T., Fukunaga, T., Igarashi, T., 1994. Phylogenetic subtypes of human T-lymphotropic virus type I and their relations to the anthropological background. Proc. Natl. Acad. Sci. 91, 1124–1127.
Mota, A.C., Van Dooren, S., Fernandes, F.M., et al., 2007. The close relationship between South Africa and Latin American HTLV strains corroborated in a molecular epidemiological study of the HTLV-I isolates from a blood donor cohort. AIDS Res. Hum. Retrovir. 23, 503–507.
Mueller, N., 1991. The epidemiology of HTLV-I infection. Cancer Causes Control 2, 37–52.
Nagel, R.L., 1984. The origin of the hemoglobin S gene: clinical, genetic and anthropological consequences. Einstein Q. J. Biol. Med. 2, 53–62.
Nicholas, K.B., Nicholas, H.B.J., Deerfield, D.W., 1997. GeneDoc: analysis and visualization of genetic variation. Embnew News 30.
Parra, E.J., Marcini, A., Akey, J., Martinson, J., Batzer, M.A., Cooper, R., Forreester, D.B., Deka, R., Ferrell, R.E., Shriver, M.D., 1998. Estimating African American admixture proportions by use of population-specific alleles. Am. J. Hum. Genet. 63, 1839–1851.
Parra, F.C.Amado, Lambertucci, R.C., Rocha, J.R., Antunes, J., Pena, C.M., 2003. Color and genomic ancestry in Brazilians. Proc Natl Acad Sci USA 100, 177–182.
Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.
Proietti, F.A., Carneiro-Proietti, A.B.F., Catalan-Soares, B.C., Murphy, E.L., 2005. Global epidemiology of HTLV-1 infection and associated diseases. Oncogene 24, 6058–6068.
Rego, F.F., Alcantara, L.C., Moura Neto, J.P., et al., 2008. HTLV type 1 molecular study in Brazilian villages with African characteristics giving support to the post-Columbian introduction hypothesis. AIDS Res. Hum. Retrovir. 24, 673–677.
Salas, A., Richards, M., De La Fe, T., et al., 2002. The making of the African mtDNA landscape. Am. J. Hum. Genet. 71, 1082–1111.
Salemi, M., Van Dooren, S., Audenaert, E., 1998. Two new human T-lymphotropic virus type I phylogenetic subtypes in seroindeterminates, a Mbuti Pygmy and a Gabonese, have closest relatives among African STLV-I strains. Virology 246, 277–287.
Santos, J.I., Lopes, M.A.A., Deliége-Vasconcelos, E., et al., 1995. Seroprevalence of HIV, HTLV-I/II and other perinatally-transmitted pathogens in Salvador, Bahia. Rev. Inst. Med. Trop. Sao Paulo 37, 343–348.
Sequeira, C.G., Tamegão-Lopes, B.P., Santos, E.J.M., Ventura, A.M.R., Moraes-Pinto, M.I., Succi, R.C.M., 2012. Descriptive study of HTLV infection in a population of pregnant women from the state of Pará, Northern Brazil. Rev. Soc. Bras. Med. Trop. 45, 453–456.
Shriver, M.D., Parra, E.J., Dios, S., Bonilla, C., Norton, H., Jovel, C., Pfaff, C., Jones, C., Massac, A., Cameron, N., 2003. Skin pigmentation, biogeographical ancestry and admixture mapping. Hum. Genet. 112, 387–399.
Soares, P., Alshamali, F., Pereira, J.B., et al., 2011. The expansion of mtDNA haplogroup L3 within and out of Africa. Mol. Biol. Evol. 29, 915–927.

Souza, V.G., Martins, M.L., Carneiro-Proietti, A.B., Januário, J.N., Ladeira, R.V., Silva, C.M., et al., 2012. High prevalence of HTLV-1 and 2 viruses in pregnant women in São Luis, state of Maranhão, Brazil. Rev. Soc. Bras. Med. Trop. 45, 159–162.
Sutton, M., Bouhassira, E.E., Nagel, R.L., 1989. Polymerase chain reaction amplification applied to the determination of β-like globin gene cluster haplotypes. Am. J. Hematol. 32, 66–69.
Swofford, D.L., Olsen, G.J., Waddell, P.J., Hillis, D.M., 1996. Phylogenetic inference. Mol. Syst. Biol. 2, 407–414.
Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.
Torroni, A., Schurr, T.G., Yang, C.C., et al., 1992. Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations. Genetics 130, 153–162.
Torroni, A., Schurr, T.G., Cabell, M.F., et al., 1993. Asian affinities and continental radiation of the four founding Native American mtDNAs. Am. J. Hum. Genet. 53, 563–590.
Torroni, A., Huoponen, K., Francalacci, P., et al., 1996. Classification of European mtDNAs from an analysis of three European populations. Genetics 144, 1835–1850.
Vallinoto, A.C., Ishak, M.O., Azevedo, V.N., et al., 2002. Molecular epidemiology of human T-lymphotropic virus type II infection in Amerindian and urban populations of the Amazon region of Brazil. Hum. Biol. 74, 633–644.
Van Dooren, S., Gotuzzo, E., Salemi, M., et al., 1998. Evidence for a post-Columbian introduction of human T-cell lymphotropic virus in Latin America. J. Gen. Virol. 79, 2695–2708.
Vandamme, A.M., Liu, H.F., Goubau, P., Desmyter, J., 1994. Primate T-lymphotropic virus type I LTR sequence variation and its phylogenetic analysis: compatibility with an African origin of PTLV-I. Virology 202, 212–223.
Verger, P., 1976. Trade relations between the Bight of Benin and Bahia, 17th to 19th century. In: Crawford, Evelyn (Ed.), Trans. Ibadan. Ibadan University Press, pp. 24–26.
Vicente, A.C.P., Gudo, S.E., Iniguez, A.M., et al., 2011. Genetic characterization of human T-cell lymphotropic virus type 1 in Mozambique: transcontinental lineages drive the HTLV-1 endemic. PLoS Negl. Trop. Dis. 5, 1038.
Vinháes, J.C., 2001. São Jorge dos Ilhéus. Da capitania ao fim do século XX. Editus, Ilhéus.
Wainscoat, J.S.S., Hill, A.V., Boyce, A.L., et al., 1986. Evolutionary relationships of human populations from an analysis of nuclear DNA polymorphisms. Nature 319, 491–493.
Wolfe, N.D., Heneine, W., Carr, K.J., et al., 2005. Emergence of unique primate T-lymphotropic viruses among central African bushmeat hunters. Proc. Natl. Acad. Sci. U. S. A. 102, 7994–7999.
Yamashita, M., Veronesi, R., Menna-Barreto, M., 1999. Molecular epidemiology of human T-cell leukemia virus type I (HTLV-1) Brazil: the predominant HTLV-Is in South America differ from HTLV-Is of Japan and Africa, as well as those of Japanese immigrants and their relatives in Brazil. Virology 261, 59–69.
Ydy, R.R., Ferreira, D., Souto, F.J., Fontes, C.J., 2009. Prevalence of human T-cell lymphotropic virus (HTLV-1/2) infection among puerperae in Cuiabá, Mato Grosso, 2006. Rev. Soc. Bras. Med. Trop. 42, 28–32.