

Re-mapping the molecular features of the human immunodeficiency virus type 1 and human T-cell lymphotropic virus type 1 Brazilian sequences using a bioinformatics unit established in Salvador, Bahia, Brazil, to give support to the viral epidemiology studies

Artur Trancoso Lopo de Queiroz^{/*}, Aline Cristina Andrade Mota-Miranda^{/*},
Tulio de Oliveira^{**}, Domingos Ramon Moreau, Caroline de Carvalho Urpia,
Chandra Mara Carvalho, Bernardo Galvão-Castro^{/*}, Luiz Carlos Junior Alcantara^{*/+}

Laboratório Avançado de Saúde Pública, Centro de Pesquisa Gonçalo Moniz-Fiocruz, Rua Valdemar Falcão 121, 40295-001 Salvador, BA, Brasil ^{*}Escola Bahiana de Medicina e Saúde Pública, Fundação para o Desenvolvimento das Ciências, Salvador, BA, Brasil ^{**}Zoology Department, Oxford University, Oxford, United Kingdom

The analysis of genetic data for human immunodeficiency virus type 1 (HIV-1) and human T-cell lymphotropic virus type 1 (HTLV-1) is essential to improve treatment and public health strategies as well as to select strains for vaccine programs. However, the analysis of large quantities of genetic data requires collaborative efforts in bioinformatics, computer biology, molecular biology, evolution, and medical science. The objective of this study was to review and improve the molecular epidemiology of HIV-1 and HTLV-1 viruses isolated in Brazil using bioinformatic tools available in the Laboratório Avançado de Saúde Pública (Lasp) bioinformatics unit. The analysis of HIV-1 isolates confirmed a heterogeneous distribution of the viral genotypes circulating in the country. The Brazilian HIV-1 epidemic is characterized by the presence of multiple subtypes (B, F1, C) and B/F1 recombinant virus while, on the other hand, most of the HTLV-1 sequences were classified as Transcontinental subgroup of the Cosmopolitan subtype. Despite the high variation among HIV-1 subtypes, protein glycosylation and phosphorylation domains were conserved in the pol, gag, and env genes of the Brazilian HIV-1 strains suggesting constraints in the HIV-1 evolution process. As expected, the functional protein sites were highly conservative in the HTLV-1 env gene sequences. Furthermore, the presence of these functional sites in HIV-1 and HTLV-1 strains could help in the development of vaccines that pre-empt the viral escape process.

Key words: human immunodeficiency virus type 1 - human T-cell lymphotropic virus type 1 - bioinformatics - molecular epidemiology

Recent advances have led to an unprecedented increase in sequence data for human immunodeficiency virus type 1 (HIV-1) and human T-cell lymphotropic virus type 1 (HTLV-1). For example, the number of HIV-1 sequences in GenBank had increased from approximately 42,000 in September 2000 (Gaschen et al. 2001) to nearly 115,000 sequences on September 2004. These sequences are also a valuable source of data for genetic analyses, such as the HIV-1 drug resistance study in which mutations into protease and reverse transcriptase genes can be used to better manage antiretroviral treatment (Shafer et al. 2000). Vaccine initiatives use sequences from the immunodominant regions of the viruses to design artificial peptides that stimulate the immune system and control the viral replication (Addo et al. 2001). Other research projects are evaluating whether the HTLV-1 envelope protein immunodominant region

could elicit neutralizing antibody responses against HTLV-1. Such information could be used in the development of vaccines and to improve diagnostic methods (Sundaram et al. 2004).

The post-translational modifications in the virus proteins are essential for HIV-1 and HTLV-1 fitness, assembly, and immune escape. The most frequent modifications are N-glycosylation, N-myristylation, and phosphorylation by protein kinases (Adachi et al. 1992, Reitter et al. 1998, Ono et al. 2000, Bouamir et al. 2003, Grasmann et al. 2005)

The HIV-1 epidemic in Brazil was initially dominated by HIV-1 subtype B (Morgado et al. 1994) and the virus spread to all the states in the country through different transmission routes. HIV-1 subtype F1 was first identified in Salvador and recombinants of subtypes B and F1 were identified in Rio de Janeiro (Sabino et al. 1994). A more recent heterosexual epidemic, characterized by the presence of subtype C viruses, was identified in South Brazil (Guimarães et al. 2002, Soares et al. 2003). Other reports have also demonstrated that distinct HIV-1 subtypes (A, B, C, D, F1) and recombinant forms (B/C, B/F1) are actively participating in the Brazilian Aids epidemic (Morgado et al. 1998, Caride et al. 2001, Soares et al. 2005, Couto-Fernandez et al. 2005, Barreto et al. 2006, De Sa Filho et al. 2006).

Financial support: Fapesp, PN-DST/AIDS
+Corresponding author: lalcan@cpqgm.fiocruz.br
ATLQ and ACAM-M contribute equally to this paper.
Received 13 July 2006
Accepted 14 February 2007

The investigation of Brazilian HTLV-1 isolates from different regions and ethnic groups showed that the epidemic is more homogeneous than that of HIV-1, and is characterized by the presence of HTLV-1 Cosmopolitan subtype, Transcontinental subgroup, throughout Brazil (Liu et al. 1996, Yamashita et al. 1999, Alcantara et al. 2003, Zehender et al. 2004, Vallinoto et al. 2004, Laurentino et al. 2005, Alcântara et al. 2006).

In short, the molecular epidemiology of HIV-1 and HTLV-1 has been evolving rapidly. The dynamic changes in these epidemics need to be monitored, since there are important implications for treatment response and for the selection of vaccine candidates.

In this report, we analyze the Brazilian HIV-1 and HTLV-1 sequences, previously deposited in the GenBank, to review the molecular epidemiology of both viruses so as to better characterize the functional domains of HIV-1 and HTLV-1 proteins using the tools available at the bioinformatics unit of Lasp/CPqGM/Fiocruz, in Salvador, BA, Brazil.

MATERIALS AND METHODS

Brazilian HIV-1 and HTLV-1 sequences genotyping

To review the molecular epidemiology of HIV-1 in Brazil, we organized three main datasets from the *pol*, *gag*, and *env* genes imported from the GenBank database. We randomly selected 44 Brazilian HIV-1 *pol* sequences (1035 bp) from GenBank, and performed phylogenetic analyses. The subtyping for all Brazilian HIV-1 *gag* (n = 223; 400 pb) and randomly *env* (126; 400) sequences were found in previously published papers. In addition, we confirmed the subtyping by performing the phylogenetic analyses using the Clustal X software to align (Jeanmougin et al. 1998), and the BioEdit and GeneDoc programs to edit the alignment (Nicholas et al. 1997, Hall 1999). The nucleotide substitution model was chosen using Modeltest version 3.06, and PAUP* version 4.0b10 (Swofford 1998) was used to obtain the neighbor joining (NJ) and maximum likelihood (ML) trees. However, to review the molecular epidemiology of HTLV-1 in Brazil, we also organized a dataset with all (n = 243; 700) LTR sequences (total and partial) from Brazilian isolates published in the Genbank database before

October 2006. The selected evolution model for the *pol* dataset was TVM+I+G with the following parameters: $f_A = 0.3676$, $f_C = 0.1751$, $f_G = 0.1967$, $f_T = 0.2328$; *R* matrix values, $R_{A \rightarrow C} = 2.4067$, $R_{A \rightarrow G} = 9.9966$, $R_{A \rightarrow T} = 0.8926$, $R_{C \rightarrow G} = 1.2052$, $R_{C \rightarrow T} = 9.9966$, $R_{G \rightarrow T} = 1.0000$, proportion of invariable sites = 0.33 and heterogeneous variable site distribution (gamma) with alpha shape = 0.74. The phylogenetic analysis of all these sequences was performed according to the methods used for HIV-1 subtyping to confirm their genotyping.

HIV-1 and HTLV-1 protein domain analysis - The p17 (n = 131) and p24 (n = 92) aminoacid sequences from the *gag* gene, and gp120 (n = 85) and gp41 (n = 102) from the *env* gene, were analyzed for potential protein sites using the GeneDoc software with the Prosite tool. The same analysis was applied to the gp46 and gp21 protein domains of 15 Brazilian HTLV-1 *env* gene sequences from GenBank (Schulz et al. 1991, Alcantara et al. 2006).

Diversity of HIV-1 and HTLV-1 sequences - Mean inter-patient genetic distances of the HIV-1 *pol*, *gag*, and *env* genes, and HTLV-1 LTR region were measured using the Kimura 2- α -parameter model with the distance matrix implemented in the MEGA 3.0 package.

RESULTS

HIV-1 and HTLV-1 genotyping - The phylogenetic analysis of representative Brazilian HIV-1 *pol* sequences identified three main subtypes: B, C, and F1 (data not shown), confirming the subtyping previously published: B (n = 20), F (n = 17), and C (n = 7) subtypes. The phylogenetic tree also showed that Brazilian sequences were related to sequences from Europe, the United States, Asia, and Africa. The Brazilian subtype B cluster was previously characterized by the presence of multiple lineages that clustered with isolates from the US and Europe, and as reviewed, the Brazilian subtype F1 isolates clustered with isolates from Europe and Argentina (Morgado et al. 2002). The Brazilian subtype C lineage was monophyletic, with all Brazilian sequences (n = 7) clustering tightly together. The phylogenetic results were supported by genetic distance analysis (Table I).

TABLE I
DNA distances between human immunodeficiency virus type 1 subtypes B, F, and C from Brazil

	<i>pol</i> (1035 bp)			<i>gag</i> p17 (250 bp)			<i>gag</i> p24 (170 bp)			<i>env</i> gp41 (315 bp)			<i>env</i> gp120 (134 bp)		
	Mean % distance			Mean % distance			Mean % distance			Mean % distance			Mean % distance		
	N	Nucl. (%)	A (%)	N	Nucl. (%)	A (%)	N	Nucl. (%)	A (%)	N	Nucl. (%)	A (%)	N	Nucl. (%)	A (%)
Subtype B	20	5.6	5.9	100	8.6	13.6	58	4.7	4.5	77	10.9	15.9	58	11.0	20.2
Subtype F1	16	4.2	5.2	8	10.1	19.4	22	4.6	3.6	14	9.4	15.6	10	5.9	11.9
Subtype C	7	3.8	4.7	14	6.2	13.4	6	4.3	3.3	9	6.8	9.8	11	8.4	12.7
Inter-subtype	BxF	BxC	CxF	BxF	BxC	CxF	BxF	BxC	CxF	BxF	BxC	CxF	BxF	BxC	CxF
nucleotide	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)
genetic distance	9.8	8.8	11.9	15.9	17.2	21.9	7.7	9.9	11.9	16.3	16.5	16.2	38.2	38.5	15

Nucl.: nucleotide; A: amino-acid; n: number of strains.

The Brazilian subtype B showed the highest level of diversity in the country in the *pol* gene (5.6%) followed by Brazilian subtype F1 (4.2%), and subtype C (3.8%). The inter-subtype genetic distance between Brazilian HIV-1 subtypes varied from 8.8 to 11.9% at nucleotide level within the *pol* gene, from 15 to 21.9% in *gag* and around 16% in *env* (gp41). The available Brazilian HIV-1 sequences were from the states of São Paulo, Rio de Janeiro, Mato Grosso do Sul, Minas Gerais, Amazonas, and Bahia. The prevalences of the B, F1, and C subtypes were 83, 11, and 4% respectively. A and D subtypes and B/F1 recombinants were the least common in this phylogenetic analysis, 1% (A and D) and 0.98% (B/F1).

As for HTLV-1 molecular epidemiology, we chose to analyze the LTR region because it is impossible to identify subgroups using other regions, despite the low variability of the HTLV-1 genome. All sequences were confirmed to be of the Cosmopolitan subtype, and only eight (3.3%) were classified as Japanese subgroup (B), while the other 235 sequences (96.7%) were of the Transcontinental subgroup (A). The Transcontinental subgroup isolates were identified in the North, Northeast, Southeast, and South region of the country while the Japanese subgroup has not yet been identified in the South. Five of the Japanese subgroup isolates were from the Southeast region, two from the North, and one from the Northeast region. The HTLV-1 genome is characterized by low variability confirmed by the diversity rates within the Transcontinental and Japanese subgroups which were 3 and 0.5% respectively. Moreover, the inter-subgroup genetic distance was 2.5%.

HIV-1 and HTLV-1 protein domain analysis - The protein domain site analysis of the *pol* gene (protease and reverse transcriptase proteins), *gag* gene (p17 and p24 proteins), and envelope gene (gp41 and gp120 proteins) are summarized in Table II. The following sites were mapped: protein kinase C phosphorylation (PKC), N-glycosylation, N-myristylation, and casein kinase 2 (CK2) phosphorylation sites. The PKC and CK2 phosphorylation sites were found in all proteins of all subtypes, with the exception of the envelope gp41 which was found to lack CK2 sites. N-myristylation sites were also found in all proteins except in the envelope gp120 proteins of subtype B and C. N-glycosylation sites were only found in three (gp41, gp120, p17) of the six proteins, but remained conserved in all subtypes of gp41 and p17. The gp120 protein showed the highest degree of N-glycosylation site variation, with six in subtype B, five in subtype F1, and four in subtype C (Table II).

To investigate the molecular characterization of HTLV-1 in Brazil, we also analyzed the potential sites in the *env* gene. We identified N-glycosylation sites located at 222-225, 244-247, and 272-275aa in 100, 86.6, and 100% of the sequences that belong to gp46 respectively. We found one N-myristylation site located at 327-338aa in 100% of the sequences within gp21. Two sites of phosphorylation by protein kinases were identified: one residue of CK2 phosphorylation, at 194-197aa in 80% in gp46 and two residues of PKC phosphorylation at 310-312 and 342-344aa in 100% of the sequences that belong to gp21.

DISCUSSION

The genetic analysis of HIV-1 and HTLV-1 showed distinct patterns. The Brazilian HIV-1 genetic analyses involved 44 representative *pol* sequences and 342 sequences from the HIV-1 *gag* and *env* genes. The results showed great heterogeneity, with multiple subtypes (A, B, C, D, F1) and recombinant viruses from subtypes B and F1 (B/F1) as previously described. The *pol* gene phylogenetic tree confirmed the relationship between Brazilian subtypes B, C, and F1 and representative isolates from other countries (data not shown). The Brazilian subtype B isolates were introduced early in the HIV/AIDS epidemic (Morgado et al. 1994) and it is characterized by the presence of multiple lineages and high genetic variation (5.6%) within the *pol* gene. Brazilian subtype F1 forms a cluster with sequences from Europe and Argentina, and probably also appeared early on the epidemic, as demonstrated in studies of HIV evolutionary history (Bello et al. 2006). Prior introductions of both subtype B and F1 are also supported by the detection of B/F1 recombinants in Brazil. The subtype C epidemic has appeared more recently (Soares et al. 2003). The Brazilian subtype C epidemic is characterized by the presence of a monophyletic cluster and low levels of diversity in the *pol*, *gag*, and *env* genes (Table I). As expected, the *gag* and *env* proteins showed a higher level of intra-subtype (4.7 to 11%) and inter-subtype diversity (7.7 to 22%). Despite the fact that diversity between subtypes was almost 20% at the amino acid level in the *gag* and *env* genes (Louwagie et al. 1993), the majority of the protein domain sites remained conserved among the different HIV-1 subtypes found in Brazil. In contrast, the vast majority (96.7%) of the HTLV-1 isolates were classified as subgroup A, a finding which agrees with previous reports (Alcantara et al. 2003), while the remaining strains (3.3%) belonged to subgroup B previously characterized in Japan. Considering the intense Japanese immigration at the beginning of the 1900s, it is not surprising to find this subtype in Brazil. Three of the Japanese subgroup sequences were from Japanese immigrants living in the state of São Paulo and the only Japanese subgroup sequence from the Northeast region was isolated from a pregnant woman whose husband had lived for a few years in São Paulo. It has been suggested that the genomic variability of this virus is related to ethnic background as well as the geographical distribution of carriers of the virus. Therefore, we could suggest that because of the very low variation of the Japanese subgroup sequences, they were probably introduced into the country recently. In contrast, the Transcontinental subgroup has a higher variation and a wide distribution across Brazilian regions. Certain evolutionary studies suggest that the first introduction of the virus in the continent occurred with the migration of the Mongol population through Bhering Straits around 40,000 to 10,000 years ago. Other studies suggest that the virus was brought along with the post-Columbian African slave trade around 400 years ago, or more recently, due to migrations of HTLV-1 infected Japanese immigrants to Latin America as previously reviewed (Laurentino et al. 2005).

The functional regions responsible for phosphorylation, myristylation, and glycosylation sites mapped in the HIV-1 proteins in this study were highly conserved. These sites are potentially important for the functioning of these proteins and their exclusion could potentially produce fewer functional virus particles. The myristylation sites are associated with viral assembly, and their loss can block particle assembly and viral replication, and particularly mutations on the N-terminal of structural precursor polyprotein Pr55^{gag} (Bouamir et al. 2003). In our study we did not analyze the N-terminal region, but, in our results, we show three conservative myristylation sites in different regions of the the *gag* proteins, and functional studies would be important to describe the role of this assign in these regions. The glycosylation sites are associated with virus protection against the immune system (Reitter et al. 1998) and they were frequently found in our results suggesting that the virus could have maintained these sites to escape from the immune system. An ideal vaccine should contain epitopes to create immune pressure on the functional regions of the viruses that cannot escape or produce fewer functional virus escape variants (De Oliveira et al. 2004). In addition, as recently suggested (Borgetz et al. 2000, Morgado et al. 2002), it is very important to continue mapping the genetic information of the HIV-1

strains circulating in Brazil, because any eventual vaccine will need to be related to the epidemic at a national level. The presence of N-glycosylation and protein kinase phosphorylation sites in the HTLV-1 *env* gene sequences could indicate a possible escape mechanism that would prevent the recognition of antibodies in the infected cells, and avoid syncytium formation as well as the presentation of the virus proteins to the major histocompatibility complex (MHC) and consequently activate the immune system via CD4⁺ and CD8⁺ cells.

A more detailed analysis of the protein domains of HIV-1 and HTLV-1, as well as genetic analyses of the complete Brazilian HIV-1 sequence dataset (3156 sequences), is underway.

In conclusion, to facilitate the effective use of all available HIV-1 and HTLV-I genetic data and to train local expertise, we have developed a specialized viral bioinformatics unit at the Advanced Public Health Laboratory of the Gonçalo Moniz Research Center-Oswaldo Cruz Foundation. The work of this unit will enable the analysis of all Brazilian viral sequences deposited in public database through the constructing of new datasets and make use of all molecular biology programs available. Furthermore, this unit could provide important information to be used in HIV-1 evolutionary studies, vaccine and antiretroviral treatment design, as well as fa-

TABLE II
Frequency of Brazilian human immunodeficiency virus type 1 protein domains from subtypes B, F1, and C

	CK2 site			Phosphorylation			N-myristoylation site			N-glycosylation site			PKC site		
	n	Position	%	n	Position	%	n	Position	%	n	Position	%	n	Position	%
<i>pol</i> protease															
Subtype B	2	26-29 31-24	100 100	3	27-32 48-53 92-97	100 95 100	NF			1	12-15	85			
Subtype F1	2	26-29 31-24	100 100	3	27-32 48-53 92-97	100 100 80	NF			1	12-15	85			
Subtype C	2	26-29 31-24	100 100	3	27-32 48-53 92-97	100 100 100	NF			1	12-15	94			
<i>pol</i> RT															
Subtype B	5	3-6 39-42 107-110 182-185 213-126	100 80 100 100 65	1	194-199	75	NF			1	68-70	95			
Subtype F1	5	3-6 39-42 107-110 182-185 213-126	95 60 95 100 60	1	194-199	88	NF			1	68-70	100			
Subtype C	4	3-6 107-110 182-185 213-126	100 85 85 100	1	194-199	85	NF			1	68-70	88			
<i>gag</i> p17															
Subtype B	1	70-73	100	1	49-54	97	1	109-110	90	1	111-113	83			
Subtype F1	1	70-73	100	1	49-54	50	1	109-110	87.5	1	111-113	25			
Subtype C	NF			1	49-54	93	1	109-110	78.6	1	111-113	7			

	CK2 site			Phosphorylation			N-myristoylation site			N-glycosylation site			PKC site		
	n	Position	%	n	Position	%	n	Position	%	n	Position	%	n	Position	%
<i>gag p24</i>															
Subtype B	4	48-51	95	2	60-65	100	NF				1	16-18	24		
		72-75	93		101-111	100									
		110-113	66												
		149-152	73												
Subtype F1	4	48-51	95	2	60-65	100	NF				1	16-18	23		
		72-75	95		101-111	100									
		110-113	63												
		149-152	77												
Subtype C	5	48-51	100	2	60-65	100	NF				2	16-18	85		
		72-75	85		101-111	100						171-173	85		
		110-113	100												
		149-152	100												
		210-213	100												
<i>env gp41</i>															
Subtype B	NF			1	1095-1100	100	1	1112-1115	100	1	1101-1102	100			
Subtype F1	NF			1	1095-1100	100	1	1112-1115	100	1	1101-1102	100			
Subtype C	NF			1	1095-1100	100	1	1112-1115	100	1	1101-1102	100			
<i>env gp120</i>															
Subtype B	1	682-685	81	NF				6	680-681	81	1	721-723	96		
									694-697	91					
									707-710	70					
									713-716	86					
									718-719	89					
									749-752	94					
Subtype F1	2	682-685	80	1	752-760	70	5	694-697	100	1	721-723	100			
		735-738	80					707-710	100						
								713-716	80						
								718-719	100						
								749-752	100						
Subtype C	1	757-760	63	NF			4	680-681	63	2	721-723	72			
								694-697	91		724-726	100			
								718-719	100						
								749-752	100						

NF: not found; n: number of sites; %: frequency of the corresponding site.

cilitate the transfer of bioinformatics and genetic analysis expertise to Brazilian researchers. Our efforts coincide with the purpose of DBCollHIV database (Araújo et al. 2006) which is to support Brazilian research groups for data storage and sequence analysis.

ACKNOWLEDGMENTS

To Dr Anne-Mieke Vandamme who inspired us to initiate this project during the first Brazilian workshop on viral evolution and molecular epidemiology.

REFERENCES

Adachi Y, Copelandj TD, Takahashi C, Nosakal T, Ahmedj A, Oroszlanj S, Hatanaka M 1992. Phosphorylation of the ex protein of human T-cell leukemia virus type I. *J Biol Chem* 267: 21977-21981.

Addo MM, Altfeld M, Rosenberg ES, Eldridge RL, Philips MN, Habeeb K, Khatri A, Brander C, Robbins GK, Mazzara GP, Goulder PJ, Walker BD 2001. HIV Controller Study Collaboration. The HIV-1 regulatory proteins Tat and Rev are frequently targeted by cytotoxic T lymphocytes derived from HIV-1-infected individuals. *Proc Natl Acad Sci USA* 98: 1783-1796.

Alcantara LCJ, De Oliveira T, Gordon M, Pybus O, Mascarenhas RE, Seixas MO, Gonçalves M, Hlela C, Cassol S, Galvão-Castro B 2006. Tracing the origin of Brazilian HTLV-1 as determined by analysis of host and viral genes. *AIDS* 20: 780-782.

Alcantara LCJ, Van Dooren S, Gonçalves MS, Kashima S, Costa MCR, Santos FLN, Bittencourt AL, Dourado I, Andrade Filho A, Covas DT, Vandamme A-M, Galvão-Castro B 2003. Globin haplotypes of human T-cell lymphotropic virus type I (HTLV-I) infected individuals in Salvador, Bahia, Brazil suggest a post-Columbian African origin of this virus. *J Acquir Immune Defic Syndr* 33: 536-542.

Araújo LV, Soares MA, Oliveira SM, Cequer P, Tanuri A, Sabino EC, Ferreira JE 2006. DBCollHIV: a database system for collaborative HIV analysis in Brazil. *Gen Mol Res* 5: 203-215.

Barreto CC, Nishyia A, Araujo LV, Ferreira JE, Busch MP, Sabino EC 2006. Trends in antiretroviral drug resistance and clade distributions among HIV-1 infected blood donors in São Paulo, Brazil. *J Acquir Immune Defic Syndr* 41 : 338-341.

Bello G, Guimarães ML, Morgado MG 2006. Evolutionary history of HIV-1 subtype B and F infections in Brazil. *AIDS* 20: 763-768.

- Bongertz V, Bou-Habib DC, Brigido LF, Caseiro M, Chequer PJ, Couto-Fernandez JC, Ferreira PC, Galvão-Castro B, Greco D, Guimarães ML, Linhares de Carvalho MI, Morgado MG, Oliveira CA, Osmanov S, Ramos CA, Rossini M, Tanuri A, Ueda M 2000. HIV-1 diversity in Brazil: genetic, biologic and immunologic characterization of HIV-1 strains in three potential HIV vaccine evaluation sites. Brazilian Network for HIV isolation and characterization. *J Acquir Immune Defic Syndr* 23: 184-193.
- Bouamir F, Scarlata S, Carter C 2003. Role of myristylation in HIV-1 gag assembly. *Biochemistry* 42: 6408-6417.
- Caride E, Brindeiro R, Hertogs K, Larder B, Dehertogh P, Machado E, Sá CA, Eyer-Silva WA, Sion FS, Passioni LF, Menezes JA, Calazans A, Tanuri A 2001. Drug-resistant reverse transcriptase genotyping and phenotyping B and non-B subtypes (F and A) of human immunodeficiency virus type 1 found in Brazilian patients failing HAART. *Virology* 275: 107-115.
- Couto-Fernandez JC, Silva-de-Jesus C, Veloso VG, Rachid M, Gracie RS, Chequer-Fernandez SL, Oliveira SM, Arakaki-Sanchez D, Chequer PJ, Morgado MG 2005. Human immunodeficiency virus type 1 (HIV-1) genotyping in Rio de Janeiro, Brazil: assessing subtype and drug-resistance associated mutations in HIV-1 infected individuals failing highly active antiretroviral therapy. *Mem Inst Oswaldo Cruz* 100: 73-78.
- De Oliveira T, Salemi M, Gordon M, Vandamme AM, Van Rensburg EJ, Engelbrecht S, Cassol S 2004. Codon-substitutions models as indicators of positive selection and amino acid diversification: an alternative approach to HIV-1 vaccine development? *Genetics* 167: 1047-1058.
- De Sá Filho DJ, Sucupira MC, Caseiro MM, Sabino EC, Diaz RS, Janini LM 2006. Identification of two HIV type 1 circulating recombinant forms in Brazil. *AIDS Res Hum Retroviruses* 22: 1-13.
- Gaschen B, Kuiken C, Korber B, Foley B 2001. Retrieval and on-the-fly alignment of sequence fragments from the HIV database. *Bioinformatics* 17: 415-418.
- Grassmann R, Aboud M, Jeang KT 2005. Molecular mechanisms of cellular transformation by HTLV-1 tax. *Oncogene* 24: 5976-5985.
- Guimarães ML, dos Santos Moreira A, Loureiro R, Galvão-Castro B, Morgado MG 2002. Brazilian Network for HIV isolation and characterization high frequency of recombinant genomes in HIV type 1 samples from Brazilian southeastern and southern regions. *AIDS Res Hum Retroviruses* 18: 1261-1269.
- Hall TA 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* 41: 95-98.
- Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Gibson TJ 1998. Multiple sequence alignment with Clustal X. *Trends Biochem Sci* 23: 403-405.
- Laurentino RV, Lopes IL, Azevedo VN, Machado LFA, Moreira MRC, Lobato L, Ishak MG, Ishak R, Vallinoto AC 2005. Molecular characterization of human T-cell lymphotropic virus (HTLV) coinfecting HIV-1 infected patients in the Amazon region of Brazil. *Mem Inst Oswaldo Cruz* 100: 371-376.
- Liu HF, Goubau P, Van Brussel M, Van Laethem K, Chen YC, Desmyter J, Vandamme AM 1996. The three human T-lymphotropic virus type I subtypes arose from three geographically distinct simian reservoirs. *J Gen Virol* 77: 359-368.
- Louwagie J, Mccutchan FE, Peeters M, Brennan TP, Sanders-Buell E, Eddy GA, Van Der Groen G, Franssen K, Gershky-Damet GM, Deleys R 1993. Phylogenetic analysis of gag genes from 70 international HIV-1 isolates provides evidence for multiple genotypes. *AIDS* 7: 769-780.
- Morgado MG, Guimarães ML, Galvão-Castro B 2002. HIV-1 polymorphism: a challenge for vaccine development – A Review. *Mem Inst Oswaldo Cruz* 97: 143-150.
- Morgado MG, Guimarães ML, Gripp CBG, Neves Jr. I, Costa CI, Santos VGV, Linhares-de-Carvalho MI, Galvão-Castro B, Bongertz V, The Hospital Evandro Chagas AIDS Clinical Research Group 1998. High prevalence of HIV-1 subtype B and identification of an HIV-1 subtype D infection in the City of Rio de Janeiro, Brazil. *J AIDS Hum Retroviruses* 18: 488-494.
- Morgado MG, Sabino EC, Shpaer EG, Bongertz V, Brigido L, Guimarães MD, Castilho EA, Galvão-Castro B, Mullins JI, Hendry RM. 1994. V3 region polymorphisms in HIV-1 from Brazil: prevalence of subtype B strains divergent from North American/European prototype and detection of subtype F. *AIDS Res Hum Retroviruses* 10: 569-576.
- Nicholas KB, Nicholas HBJ, Deerfield DW 1997. GeneDoc: analysis and visualization of genetic variation. *EMB news* 14: 30.
- Ono A, Orenstein JM, Freed EO 2000. Role of the Gag matrix domain in targeting human immunodeficiency virus type 1 assembly. *J Virol* 74: 2855-2866.
- Reitter JN, Robert EM, Desrosiers RC 1998. A role for carbohydrates in immune evasion in AIDS. *Nat Med* 4: 679-684.
- Sabino EC, Shpaer EG, Morgado MG, Korber BT, Diaz RS, Bongertz V, Cavalcante S, Galvão-Castro B, Mullins JI, Mayer A 1994. Identification of human immunodeficiency virus type 1 envelope genes recombinant between subtypes B and F in two epidemiologically linked individuals from Brazil. *J Virol* 68: 6340-6346.
- Schulz TF, Calabro ML, Hoad JG, Carrington CV, Matutes E, Catovsky D, Weiss RA 1991. HTLV-1 envelope sequences from Brazil, the Caribbean, and Romania: clustering of sequences according to geographic origin and variability in an antibody epitope. *Virology* 184: 483-491.
- Shafer RW, Kantor R, Gonzales MJ 2000. The genetic basis of HIV-1 resistance to reverse transcriptase and protease inhibitors. *AIDS Reviews* 2: 211-228.
- Soares MA, De Oliveira T, Brindeiro RM, Diaz RS, Sabino EC, Brigido L, Pires IL, Morgado MG, Dantas MC, Barreira D, Teixeira PR, Cassol S, Tanuri A, The Brazilian Network for Drug Resistance Surveillance 2003. A specific subtype C of human immunodeficiency virus type 1 circulates in Brazil. *AIDS* 17: 11-21.
- Soares EA, Martinez AM, Souza TM, Santos AF, Da Hora V, Silveira J, Bastos FI, Tanuri A, Soares MA 2005. HIV-1 subtype C dissemination in southern Brazil. *AIDS* 19 (Suppl. 4): S81-86.
- Sundaram R, Lynch MP, Rawale Sun Y, Kazanji M, Kaumaya PTP 2004. De novo design of peptide immunogens that mimic the coiled coil region of human T-cell leukemia virus type-1 glycoprotein 21 transmembrane subunit for induction of na-

- tive protein reactive neutralizing antibodies. *J Biol Chem* 279: 24141-24151.
- Swofford DL 1998. PAUP*: phylogenetic analysis using parsimony (*and other methods). Version 4.0b5. Sinauer Associates, Sunderland, Mass.
- Vallinoto AC, Muto NA, Pontes GS, Machado LFA, Azevedo VN, Santos SB, Santos AR, Ishak MG, Ishak R 2004. Serological and molecular evidence of HTLV-I infection among Japanese immigrants living in the Amazon Region of Brazil. *J Infect Dis* 57: 156-159.
- Yamashita M, Veronesi R, Menna Barreto M, Harrington Jr WJ, Sampaio C, Brites C, Badaró R, Andrade-Filho AS, Onkura S, Igarashi T, Takehisa J, Miura T, Chamone D, Bianchini O, Jardim C, Sonoda S, Hayami M 1999. Molecular epidemiology of human T-cell leukemia virus type I (HTLV-I) in Brasil: the predominant HTLV-Is in South America differ from HTLV-I of Japan and Africa, as well as those of Japanese immigrants and their relatives in Brazil. *Virology* 261: 59-69.
- Zehender G, Colasante C, De Maddalena C, Bernini F, Savasi V, Persico T, Merli S, Ridolfo A, Santambrogio S, Moroni M, Galli M 2004. High prevalence of human T-lymphotropic virus type 1 (HTLV-1) in immigrant male-to-female transsexual sex workers with HIV-1 infection. *J Med Virol* 74: 207-215.