

RESEARCH ARTICLE

# Short-Term Dynamic and Local Epidemiological Trends in the South American HIV-1B Epidemic

Dennis Maletich Junqueira<sup>1,2,3\*</sup>, Rubia Marília de Medeiros<sup>2,3</sup>, Tiago Gräf<sup>3,4</sup>, Sabrina Esteves de Matos Almeida<sup>2,3,5</sup>

**1** Centro Universitário Ritter dos Reis—Uniritter, Departamento de Ciências da Saúde, Porto Alegre, RS, Brazil, **2** Programa de Pós-Graduação em Genética e Biologia Molecular, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil, **3** Centro de Desenvolvimento Científico e Tecnológico (CDCT), Fundação Estadual de Produção e Pesquisa em Saúde (FEPPS), Porto Alegre, RS, Brazil, **4** Programa de Pós-graduação em Biotecnologia e Biociências, Universidade Federal de Santa Catarina, Florianópolis, SC, Brazil, **5** Instituto de Ciências da Saúde, Universidade FEEVALE, Novo Hamburgo, RS, Brazil

\* [dennismaletich@hotmail.com](mailto:dennismaletich@hotmail.com); [dennis\\_junqueira@uniritter.edu.br](mailto:dennis_junqueira@uniritter.edu.br)



**OPEN ACCESS**

**Citation:** Junqueira DM, de Medeiros RM, Gräf T, Almeida SEdM (2016) Short-Term Dynamic and Local Epidemiological Trends in the South American HIV-1B Epidemic. PLoS ONE 11(6): e0156712. doi:10.1371/journal.pone.0156712

**Editor:** Kok Keng Tee, University of Malaya, MALAYSIA

**Received:** February 1, 2016

**Accepted:** May 18, 2016

**Published:** June 3, 2016

**Copyright:** © 2016 Junqueira et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** DMJ, TG, and RMM were funded by a fellowship from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

The human displacement and sexual behavior are the main factors driving the HIV-1 pandemic to the current profile. The intrinsic structure of the HIV transmission among different individuals has valuable importance for the understanding of the epidemic and for the public health response. The aim of this study was to characterize the HIV-1 subtype B (HIV-1B) epidemic in South America through the identification of transmission links and infer trends about geographical patterns and median time of transmission between individuals. Sequences of the protease and reverse transcriptase coding regions from 4,810 individuals were selected from GenBank. Maximum likelihood phylogenies were inferred and submitted to ClusterPicker to identify transmission links. Bayesian analyses were applied only for clusters including  $\geq 5$  dated samples in order to estimate the median maximum inter-transmission interval. This study analyzed sequences sampled from 12 South American countries, from individuals of different exposure categories, under different antiretroviral profiles, and from a wide period of time (1989–2013). Continentally, Brazil, Argentina and Venezuela were revealed important sites for the spread of HIV-1B among countries inside South America. Of note, from all the clusters identified about 70% of the HIV-1B infections are primarily occurring among individuals living in the same geographic region. In addition, these transmissions seem to occur early after the infection of an individual, taking in average 2.39 years (95% CI 1.48–3.30) to succeed. Homosexual/Bisexual individuals transmit the virus as quickly as almost half time of that estimated for the general population sampled here. Public health services can be broadly benefitted from this kind of information whether to focus on specific programs of response to the epidemic whether as guiding of prevention campaigns to specific risk groups.

## Introduction

Today, nearly one million individuals are infected by HIV-1 in South America and approximately 35,000 individuals die per year victims of the symptoms of AIDS [1]. A quick answer to the high mortality rate, morbidity and transmission related to HIV in certain specific regions of the South America may require a broad and efficient collection of data on the virus and the host [2]. In this context, phylogenetic analyses have been used to investigate the origin, spread and transmission of HIV between individuals and can be a powerful method in the understanding of the social, demographic and geographical issues in the epidemic [3–11].

Transmission clusters are identified here as phylogenetic inferences of the transmission links between different HIV-infected individuals. Through the genetic and evolutionary relationship these epidemiological chains allow the inference of social contacts between individuals and provide a way to reconstruct the main transmission links involved in the spread of the virus in a certain place [6,12,13]. These findings improve the correlation of local HIV epidemic with transmission pathway, drug resistance, risk behavior and cluster size and may influence the direction focus of public campaigns to specific populations aiming to reduce the rate of transmissions and consequently delaying the increase in the number of new infection cases [14]. While several countries in Europe, Asia, Africa, Australia and North America have already impactful information about the transmission between individuals, South American countries are clearly deficient for this kind of information [6,15–25]. Addressing the specific issues within local epidemics is crucial to a greatly improvement in the HIV-1 epidemic [2].

Several studies have exploited the HIV-1 high rates of mutation and its intrinsic rapid evolution to understand the spread of the virus in a certain population of a specific place [4,26–28]. Here we used phylogenetic analysis to identify and characterize the transmission links and characterize among HIV-1 South American sequences available in Genbank. As South America has an epidemic primarily based on HIV-1 subtype B (HIV-1B), all analysis were performed using *pol* sequences of this subtype.

## Methods

### Alignments

All available South American HIV-1 sequences from *pol* gene including the protease and partial segment of the reverse transcriptase (nucleotides 2253–3252 relative to strain HXB2) were selected from the Los Alamos HIV Sequence Database (<http://www.hiv.lanl.gov/>). Additionally, PubMed (<http://www.ncbi.nlm.nih.gov/pubmed>) was used to consult all studies of HIV-1 conducted in South American countries through the use of the search key “HIV-1 AND country” (where “country” was substituted for the name of each South American country) to select sequences deposited in GenBank (<http://www.ncbi.nlm.nih.gov/nucleotide/>) but not included in the Los Alamos Database. When available, geographical, demographic and clinical information were also retrieved from Los Alamos or directly from the study describing the sequence.

All sequences were submitted to REGA Subtyping Tool v2.0 [29], and RIP 3.0 [30] to confirm the subtype. Sequences presenting discordances in the subtype assignment were submitted to bootscanning analysis using Simplot 3.5.1 [31]. Final alignment were generated using MUSCLE [32] and it excluded sequences presenting premature stop codons, replicates from the same patient, and intersubtype recombinants. Four sequences of subtype D were used as outgroup.

In order to compare the influence of convergent evolution of resistance mutations on the transmission cluster detection, two versions of the final alignment were built. A set including

complete sequences (complete set, 1000bp), and a codon-stripped dataset (codon-stripped set) in which the main sites associated with major antiretroviral drug resistance for protease and reverse transcriptase were excluded, resulting in a 901bp alignment [33]. All alignments are available from the authors upon request.

## Identification of Transmission Clusters

Maximum Likelihood (ML) phylogenies were inferred in RAxML [34] on the CIPRES Science Gateway [35] incorporating the best-fitted nucleotide substitution model (GTR+I+ $\Gamma$ ) as determined in MEGA6 [36]. The approximate likelihood-ratio test (aLRT) based on Shimodaira-Hasegawa-like procedure were used to assess confidence in topology [37].

Transmission links were identified by using Cluster Picker [38] with a SH-aLRT support threshold of  $\geq 90$ . Different maximum pairwise genetic distances (1.0%, 1.5%, 2.0%, 2.5%, 3.0%, 3.5%, 4.0%, 4.5%, 5.0%, 5.5%, 6.0%, 6.5%, 7.0%, and 7.5%) within the clusters were evaluated and transmission cases previously known (09 transmission pairs and 01 transmission triads) were used as controls for the genetic distance evaluation [39–41]. Resultant transmission clusters were classified in local (clusters involving sequences sampled in the same state for Brazilian sequences or in the same country for non-Brazilian sequences), interstate (clusters involving sequences sampled in different states within Brazil), or international types (clusters involving sequences sampled in different countries) according to the sampling region of the sequences involved in the cluster.

## Time-scaled phylogenies

In order to estimate the median time of HIV-1 transmission between individuals, dated phylogenies were reconstructed using a Bayesian MCMC method implemented in BEAST v1.8 [6,22,42]. Only clusters including  $\geq 5$  dated samples were selected for further analysis.

The Bayesian analyses assumed an uncorrelated lognormal relaxed molecular clock under the GTR+I+ $\Gamma$  nucleotide substitution model and five different demographic models were tested: constant population size, expansion, exponential growth, logistic growth, and the non-parametric Bayesian skyline plot. The best fitted model were evaluated through marginal likelihood estimation. Previous estimates of the evolutionary rate for subtype B were used as normal mean prior [43]. The MCMC chain were run for  $5.0 \times 10^8$  chain steps and the convergence was evaluated in TRACER v1.5 excluding an initial 10% for burn-in [44]. Maximum clade credibility trees (MCC) were summarized using TreeAnnotator v1.8.0 in BEAST package after 50% of the burn-in was discarded and the resulting tree was visualized with FigTree v.1.4.2.

## Resistance Mutation Analysis

Complete sequences alignment were submitted to Calibrated Population Resistance tool (CPR) to detect the presence of surveillance Drug Resistance Mutations (DRM) [45]. Clusters that included sequences presenting the same antiretroviral drug resistance mutations were used to compose a new dataset in order to better understand the transmission of drug resistance among different individuals within a transmission group. The new dataset were submitted to maximum likelihood phylogeny inference in PhyML incorporating the best-fitted nucleotide substitution model (GTR+I+ $\Gamma$ ). The resulting maximum likelihood tree were used to reconstruct the ancestral nucleotide sequence of each cluster using the FastML web server software [46].

## Statistical Analysis

Statistical comparisons between datasets (complete *versus* codon-stripped dataset, including geographic distribution, resistance mutation analysis and time between infections) were made using Pearson's  $\chi^2$ -test and Fisher's exact test when appropriate. Statistical analyses were performed using WinPepi v.11.22 and the significance level was set at  $P < 0.05$ .

## Results

### Dataset

Among 7,600 *pol* sequences available in public databases, 4,810 HIV-1 subtype B isolates from South American countries were selected to set up a dataset for the identification of transmission links (S1 Table). Twelve of the 13 South American countries were represented in the 26 different regions included in this study. Due to the large area occupied by Brazil in South America, samples isolated in this country were identified by specific state of sampling (Fig 1).

Brazil was by far the most represented country in our dataset, exhibiting 2,826 sequences (58.7%) from 14 geographical states (Fig 1). Due to the unavailability of sampling region information, 780 sequences from Brazil could not have an attribution for the sampling state. Argentina was the second best represented country in our dataset with 1,490 sequences. Bolivia, Ecuador, Guyana, Paraguay, Suriname, and Uruguay were misrepresented regions with less than 10 sequences per country.

### Alignments

Traditionally, transmission clusters studies have used datasets including codon-stripped sequences from which codons associated with major resistance in protease and reverse transcriptase are removed [6,16,18,20,38]. As one of the goals of this study was to evaluate the influence and transmission of drug resistance mutations in transmission clusters, we constructed two different datasets (complete and codon-stripped datasets) to test whether the analysis of the complete set could distort the results. Despite the differences in the composition of sequences inside the transmission clusters, all results obtained in this study are relatively similar for both datasets, including geographic distribution, resistance mutation and time between infections (S1 Fig, S2 and S4 Tables).

### Within-cluster maximum genetic distance thresholds

We first evaluated the effect of different maximum genetic distances on cluster identification among HIV-1 sequences from South America. The branch support threshold was fixed at 90 (SH-aLRT) with a within-cluster maximum genetic distance varying between 1.5% and 7.5% (using intervals of 0.5%). The number of clusters detected increased with the genetic distance threshold, reaching a maximum at 6.5% (Fig 2 and S1 Fig). At genetic distances of 7.0% and 7.5%, the number of detected clusters was smaller than for 6.5%. Beyond the within-cluster maximum genetic distance of 4.5%, the proportion of new transmission clusters gradually decreases.

In the other way, the proportion of sequences included in the clusters enlarged with the increase in the genetic distance threshold (Fig 2 and S1 Fig). At the genetic distance of 7.5%, more than 80% of the sequences analyzed were included in transmission clusters. These results indicate that as the maximum genetic distance cut-off is relaxed, more sequences are being added to the clusters.

In order to understand the HIV-1 epidemic patterns, transmission links were identified using Cluster Picker in a within-cluster maximum genetic distance of 4.5%. This cut-off effectively detected all transmission links included as controls in this study and it was previously used in other studies [6,22,38].

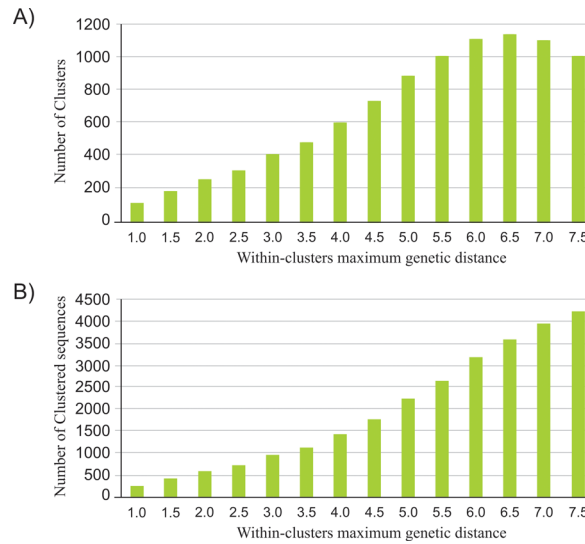
### HIV-1 Subtype B transmission links

A total of 729 transmission clusters were identified in a SH-aLRT support threshold of  $\geq 90$  and a maximum genetic distance of 4.5% (Table 1 and S2 Table). Around 71% of the transmissions links identified here are clustering sequences from the same geographical region (Table 1 and S2 Table). As expected, transmissions links including individuals from different states (within Brazil) or countries were detected to a lesser extent (8,6% and 5,5% respectively).



**Fig 1. Geographic distribution and proportion of HIV-1 subtype B pol sequences from South American countries.** Map shows locations of the HIV-1 subtype B sequences included in the dataset. The proportion (green bars) and the total number of sequences analyzed from each location is indicated. A compilation of all sequences from the Brazilian set and its respective state of sampling is indicated at the table included in the figure. Black dots indicate the cities sampled in this study. Sequences from Venezuela were sampled at Caracas (n = 213), from Colombia were sampled at Medellín (n = 32) and Bogotá (n = 7), the unique sequence from Bolivia was sampled at La Paz, sequences from Argentina were sampled at Mendoza (n = 2) and Buenos Aires (n = 1238), and sequences from Uruguay were sampled at Montevideo. For the rest of the countries the sequences had no identification of sampling region. Gray-shaded areas indicate regions not included in this study.

doi:10.1371/journal.pone.0156712.g001



**Fig 2. Number of transmission clusters and clustered sequences among 4,810 HIV-1 Subtype B codon-stripped *pol* sequences from South America.** (A) Number of transmission clusters identified using Cluster Picker with a SH-aLRT support threshold of  $\geq 90$  and under different within-maximum genetic distances. (B) Number of clustered sequences under different within-cluster genetic distances.

doi:10.1371/journal.pone.0156712.g002

Brazil, Argentina and Venezuela seemed to play an important role in the HIV-1 subtype B epidemic inside the continent, presenting the higher number of intercountry connections

**Table 1. Geographical type of HIV-1 Subtype B transmissions among clusters identified within South America for the codon-stripped dataset (901bp).**

Clustered Individuals	Geographical Type of Transmission	Number of Clusters Identified	%
2	Local Transmission*	392	70,6
	Interstate Transmission (Brazil)	44	7,93
	International Transmission	31	5,59
	Unidentified	88	15,9
3	Local Transmission*	81	68,6
	Interstate Transmission (Brazil)	13	11
	International Transmission	6	5,08
	Unidentified	18	15,3
4	Local Transmission*	28	73,7
	Interstate Transmission (Brazil)	4	10,5
	International Transmission	2	5,26
	Unidentified	4	10,5
5	Local Transmission*	9	81,8
	Interstate Transmission (Brazil)	1	9,09
	International Transmission	1	9,09
6	Local Transmission*	1	100
7	Local Transmission*	1	50
	Interstate Transmission (Brazil)	1	50
$\geq 8$	Local Transmission*	4	100
<b>Total</b>		<b>729</b>	<b>-</b>

\* Transmission Clusters involving sequences sampled in the same state for Brazilian sequences or in the same country for non-Brazilian sequences

doi:10.1371/journal.pone.0156712.t001

(S3 Table). In Brazil, the states of São Paulo, Goiás, Rio de Janeiro, Mato Grosso, and Paraná posed as potential sites for the international link within the HIV-1B epidemic in South America as seen by the number of links with other sequences (S3 Table).

Four transmission clusters linked more than seven individuals (Table 1). It is important to highlight that two of these larger clusters were associated exclusively to MSM transmission, and one included both MSM and heterosexual individuals. We verified that approximately 41% of all clusters including more than 4 individuals were made up entirely of MSM individuals.

### Dating Transmission events

As each tip of the tree represents a single patient, the internal nodes represent the most recent common ancestors of these infections and include at least one transmission event. Therefore, the internode intervals are used to estimate the maximum inter-transmission times between patients scaled by calendar years. It is important to highlight that these time intervals are maximum estimates and the inclusion of more individuals in the cluster would lead to shorter average transmission intervals.

We determined the average period in which the clustered transmissions occurred for all groups including  $\geq 5$  individuals (82 sequences, 14 clusters) under a BSP demographic model (Table 2 and S4 Table). Analysis of the overall distribution of the internode intervals in South America revealed a median of 1.98 years (95% confidence interval [CI] 1.04–2.92). The great majority of the transmission clusters analyzed with the Bayesian approach were from Brazil allowing the determination of a country-specific average transmission interval of 2.65 years (95% CI 1.3–4.0).

Clusters primarily composed of MSM individuals were separately analyzed in order to calculate the average time of transmission between MSM individuals in South America. For MSM clusters, we found a median time of transmission between individuals of 0.8 years (7 clusters, 45 sequences, 95% CI 0–2.16).

**Table 2. Average time of HIV-1 subtype B transmission among South American individuals for the codon-stripped dataset (901bp).**

Cluster	Taxa	Geographical Type	Exposure Category	Median Internode Intervals (years)
1	4	Local	-	6.410
2	7	Interstate	HET/MSM	4.468
3	5	Local	MSM	0.411
4	5	Local	MSM	0.302
5	4	Local	HET	2.247
6	6	Local	MSM	0.394
7	5	Local	MSM	1.623
8	5	International	-	1.117
9	5	Local	-	1.540
10	5	Local	MSM	0.757
11	13	Local	MSM	1.596
12	6	Local	-	3.361
13	5	Local	MSM	0.517
14	7	Local	HET/MSM	2.970
				<b>Median 1.98 (95% CI 1.04–2.92)</b>

Abbreviations: HET: Heterosexual individual, MSM: men who have sex with men individual

doi:10.1371/journal.pone.0156712.t002

### Resistance Mutation Analysis

The overall DRM prevalence of the samples included in clusters was 44.1% (720/1633; 95% CI 41.7–46.5; [Table 3](#)). The prevalence of sequences presenting only drug resistance mutations against protease inhibitors (PI) was 1.53% (25/1633; 95% CI 0.94–2.13), against nucleoside RT inhibitors (NRTI) was 6.43% (105/1633; 95% CI 5.24–7.62), and against non-NRTI (NNRTI) was 3.74% (61/1633; 95% CI 2.82–4.66). In addition, 170 sequences (10.4%; 95% CI 8.93–11.89) presented mutations that conferred resistance to the three classes of antiretroviral drugs, 171 sequences (10.5%; 95% CI 8.99–11.96) presented resistance mutations against PI and NRTI, 10 sequences (0.6%; 95% CI 0.23–0.99) presented PI and NNRTI-resistance mutations, and 178 sequences (10.9%; 95% CI 9.39–12.41) were identified as carrying resistance mutations against NRTI and NNRTI ([Table 3](#)). Within protease, L90M (17.5%) was the most frequent mutation, followed by M46I (13%), I54V (9.1%), V82A (9%), and I84V (9%) ([S5 Table](#)). The revertants at RT position 184 (M184V) were the most prevalent (18%), followed by M41L (15%), T215Y (13%), L210W (9.7%), and D67N (9.6%) ([S6](#) and [S7 Tables](#)). We found 400 clusters presenting samples with DRM.

We also analyzed the prevalence of DRM in the set of sequences not included in any cluster in order to understand the influence of the clustering in the transmission of drug resistance mutations. From 3,177 sequences not included in transmission clusters, 2,287 samples (72%; 95% CI 70.42–73.55) were identified as carrying drug resistance mutations ([Table 3](#)). Due to the high prevalence of resistance mutation in this set, we analyzed all sequences by searching for the patient antiretroviral therapy status directly from the study describing each sequence. The great majority of the sequences not included in any cluster (1,537 samples, 48.4%) could not have its antiretroviral status attributed ([S8 Table](#)). The remaining 1,640 sequences were categorized in three groups: (i) sequences sampled from patients failing antiretroviral therapy (1271 samples, 40%), (ii) sequences from treatment-naive patients (236 samples, 7.4%), and (iii) sequences from antiretroviral treated individuals (133 samples, 4.2%; [S8 Table](#)). Statistical comparison revealed that clustered sequences included more samples from treatment-naive patients and fewer samples from patients failing antiretroviral therapy ( $P < 0.001$ ) than no clustered sequences. Analysis of the most frequent DRM in protease and reverse transcriptase genes revealed a similar pattern between the sequences not included in any cluster and those included in clusters ([S5–S7 Tables](#)).

**Table 3. Drug resistance mutations identified among 4,810 sequences clustered or not clustered in transmission clusters within South America.**

Drug Resistance Mutation (DRM)	Number of Sequences presenting DRM (n = 4810)	Clustered Sequences (n = 1633)		Not Clustered Sequences (n = 3177)	
		N	%	N	%
PI	64	25	1.53	39	1.23
NRTI	346	105	6.43	241	7.59
NNRTI	180	61	3.74	119	3.75
PI + NRTI	751	171	10.5	580	18.3
PI + NNRTI	31	10	0.61	21	0.66
NRTI + NNRTI	788	178	10.9	610	19.2
PI + NRTI + NNRTI	847	170	10.4	677	21.3
<b>Total</b>	<b>3007</b>	<b>720</b>	<b>44.1</b>	<b>2287</b>	<b>72</b>

Abbreviations: NRTI: nucleoside reverse transcriptase inhibitor, NNRTI: non- nucleoside reverse transcriptase inhibitor, PI: protease inhibitor

doi:10.1371/journal.pone.0156712.t003



In addition, we examined the ancestral states at all sites associated with drug resistance for 28 clusters including more than three individuals (129 total individuals) (S9 Table). In all cases, the reconstructed ancestral sequence of the transmission cluster harbored the same drug resistance mutation as the sequences within the cluster. However, the selection of the clusters to evaluate the ancestral amino acids may create a bias in this analysis since we only chose clusters that included more than three sequences presenting the same resistance mutation.

## Discussion

The heterogeneity of the dataset used in this study brings a new challenge to the phylodynamic field. Due to the inexistence of a consensus methodology to identify transmission clusters, here we tested the effect of different maximum genetic distances on cluster identification among HIV-1B samples from South America. We observed that from the within-cluster maximum genetic distance of 4.5% the proportion of new transmission clusters identified by Cluster Picker gradually decreases. This result was expected since the inclusion of more sequences in each cluster and/or the combination of smaller ones will allow the detection of increased size clusters. Based on this result we judged the within-cluster maximum genetic distance of 4.5% to be appropriate in detecting transmission clusters [47].

At a genetic distance threshold of 4.5%, more than 720 individuals had a link to at least one other subject (Table 1). It is important to note that the vast majority of these transmission links (71%) included individuals living in the same geographic region. These results may suggest that HIV-1B infections primarily occur among individuals living in the same geographic region. This estimative represents the minimum proportion since several sequences in our dataset presented missing data for geographic origin. If extrapolated, this result may indicate that in average 71% of the epidemic from a particular location in South America can be explained only by local epidemiological trends. We also observed a tendency relating geographical distance to epidemiological connection most likely linking individuals from closest states or countries. These estimates might be a reflection of the poor infrastructure integration relative to accessibility to and among urban centers in the South America. The isolation of countries and cities in the region therefore might be viewed as potential factors intensifying the local trends in the HIV-1B epidemic. Mir et al. (2015) in a study including 6789 HIV-1B sequences from seven different Latin American countries found a much greater geographic compartmentalization of the HIV-1B epidemic in Latin American and Caribbean regions than in Europe [25]. The study supports that the epidemic is mainly driven by transmissions within countries and among neighboring countries than transmissions among countries without a common border. These results most likely reflect an intraregional difference in human mobility and spatial accessibility in Latin American countries.

In relation to the Brazilian interstate links, the geographic distribution revealed that Goiás, São Paulo, Rio de Janeiro, Paraná, and Mato Grosso are important sites of transmission within this country (Fig 1 and S4 Table). Despite the increased number of samples from São Paulo, Rio de Janeiro, and Goiás in our dataset, these results support such conclusions because we chose to identify transmission clusters based in two assumptions: (i) the support threshold of >90 and (ii) the within-clade maximum genetic distance of 4.5%. These assumptions do not prevent the increase in probability of link one sequence of a densely sampled state to another of different state in the same clade; however, it avoids the identification of such clade as a transmission cluster. In addition, demographic data from the Official Brazilian Demographic Data Center—IBGE (<http://www.ibge.gov.br/home/>) show that these states receive the highest number of internal migrants in the country corroborating the idea that migratory relationship between different states creates potential networks to the diffusion of HIV in the human

population (IBGE, 2010). Moreover, São Paulo and Rio de Janeiro are important sites for an international connection of the HIV-1B epidemic within Brazil (Fig 1). Demographic data can support our results since São Paulo and Rio de Janeiro are among the states in Brazil to receive the highest number of immigrants from other countries (<http://www.ibge.gov.br/home/>).

Our results revealed that Brazil, Argentina and Venezuela are important sites for the spread of HIV-1B within South America. Although most South American countries are misrepresented regions in our dataset, creating a potential sampling bias to conclude about their relative importance in the continental spread of HIV-1B, our findings are in agreement with migratory data revealing that Brazil, Argentina and Venezuela are primary points for regional migration (<http://www.iom.int/world-migration>). Supposedly, the spread of HIV-1 in South America may be associated to the migratory pathways established by the human population in search of economic or social change. These results highlight the importance of association between demographic data and epidemiological information to construct a full scenario of the current epidemic in a certain region [48].

We also analyzed the dynamic of dissemination within clusters to infer the time between transmissions from an individual to another. Bayesian analysis revealed an average time period of HIV-1B transmissions within South America of 1.98 years (Table 2). Our results propose a short-term dynamic of the epidemic among HIV infected individuals in which after approximately 2 years being infected patients will transmit the virus to other individual. A previous study analyzing heterosexual transmission in a group of 11,071 individuals from the United Kingdom epidemic of HIV-1 non-B subtypes found similar results [22]. The agreement with our findings may indicate a similar pattern in the transmission of HIV-1 irrespective of the population analyzed. This information is of utmost importance to understand the dynamics of infection and suggest that the early stage of infection potentially influence the onward spread of HIV [49,50].

The identification of clusters including only MSM individuals in our results allowed the estimation of an average time of transmission in this group of 0.8 years. We observed that MSM individuals transmit the virus to another subject in almost half time of that for the general population sampled here. The short time interval between transmissions can be explained by the stage of infection since within the first year after transmission, HIV infection in MSM individuals are eight times as infectious as during chronic phase [49]. Additionally, we encountered a larger number of transmission clusters (41%) including more than four individuals comprising only MSM individuals. It suggests a difference in the dynamics of the epidemics in different risk groups most likely reflecting social and sexual behavior of the human population [22]. These results should be taken as a warning data to the public health services since the transmission of the virus within the group is faster than for others groups and that in most countries the population of MSM individuals is still the most affected by HIV infection [51,52]. Prevention measures and information campaigns could absorb this information to guide the reallocation of resources and manage new investments to focus in populations in which the epidemic is concentrated.

An elevated proportion (62%) of sequences presented drug resistance mutation (DRM) in our dataset (Table 3 and S8 Table). The vast majority of sequences assigned for a therapy status were sampled from patients failing antiretroviral therapy (37%) which might explain the increased prevalence of drug resistance mutations in our sample [53–55]. When stratifying the data, we found lower prevalence of DRM in the set of sequences included in transmission clusters (Table 3). This result is consistent with the decreased replication fitness of viruses harboring DRM [11,56,57]. In addition, statistical comparisons based on Pearson's  $\chi^2$ -test revealed an increased number of sequences from treatment-naive patients included in the linked transmissions suggesting that naive individuals might be transmitting the virus in higher rates than

treated or failing antiretroviral individuals. The reasons for this pattern may be related to the lack of knowledge about HIV status and the high viral load characteristic of these untreated patients.

The ancestral reconstructions within transmission clusters revealed a tendency to the circulation of transmitted drug resistance mutations in South America since we found all inferred ancestral sequences harboring the most prevalent mutation within each cluster (S7 Table). In addition, among the resistance mutations found in this study, M184V (found in 38% of the sequences) and M41L (found in 28.5% of the sequences) were the most frequent mutations found in our samples (S2 and S4 Tables). Both mutations reflect resistance to NNRTI drugs and are described as the most common in patients failing antiretroviral therapy [58].

## Conclusion

Our results show that the HIV-1B epidemic is largely influenced by regional trends and suggest a higher probability of HIV transmission between individuals from the same geographic origin. This compartmentalized analysis of the epidemic (by states or countries) and the consequent influence of local, interstate or international transmissions had never been defined for the HIV-1 epidemic in South America. In addition, we propose a short-term dynamics within clusters in which the mean time of transmissions is two years. Information regarding the regional dynamic of HIV infections allied to statistical methods may contribute to the understanding on dissemination and expansion of new variants in certain places, including the entry of new subtypes and the circulation of drug resistant strains.

## Supporting Information

**S1 Fig. Number of transmission clusters and clustered sequences for the complete and codon-stripped datasets among 4,810 HIV-1 Subtype B *pol* sequences from South America.** (A) Absolute number of transmission clusters identified with a SH-aLRT support threshold of  $\geq 90$ . (B) Absolute number of clustered sequences under different within-cluster genetic distances.

(TIF)

**S1 Table. GenBank accession number and country of origin of 4,810 HIV-1 Subtype B sequences isolated in South America.**

(DOCX)

**S2 Table. Geographical type of HIV-1 Subtype B transmissions among clusters identified within South America for the complete dataset (1000bp).**

(DOCX)

**S3 Table. Number of links identified for the states within Brazil involved in interstate transmissions and for the countries involved in international transmissions at South America (including results from transmission pairs and transmission clusters).**

(DOCX)

**S4 Table. Average time of HIV-1 subtype B transmission among South American individuals for the complete dataset (1000bp).**

(DOCX)

**S5 Table. Amino acid substitutions in HIV-1 Protease gene related to drug resistance to protease inhibitors (PI) identified among 4,810 sequences clustered or not clustered in transmission clusters within South America.**

(DOCX)

**S6 Table. Amino acid substitutions in HIV-1 Reverse transcriptase gene related to drug resistance to nucleoside reverse transcriptase inhibitors (NRTI) identified among 4,810 sequences clustered or not clustered in transmission clusters within South America.**

(DOCX)

**S7 Table. Amino acid substitutions in HIV-1 reverse transcriptase gene related to drug resistance to non-nucleoside reverse transcriptase inhibitors (NNRTI) identified among 4,810 sequences clustered or not clustered in transmission clusters within South America.**

(DOCX)

**S8 Table. Antiretroviral therapy (ART) status of the patients and clustering behavior of the respective sequences.**

(DOCX)

**S9 Table. Ancestral reconstruction analysis of the sequences included in clusters and presenting resistance drug mutations.**

(DOCX)

## Author Contributions

Conceived and designed the experiments: DMJ SEMA. Performed the experiments: DMJ RMM TG. Analyzed the data: DMJ RMM TG SEMA. Contributed reagents/materials/analysis tools: DMJ RMM TG SEMA. Wrote the paper: DMJ RMM TG SEMA.

## References

1. UNAIDS. Report on the global AIDS epidemic [Internet]. 2013.
2. UNAIDS. Local Epidemics Issues Brief [Internet]. 2014.
3. Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ, et al. The early spread and epidemic ignition of HIV-1 in human populations. *Science* (80-). 2014; 346: 56–61. doi: [10.1126/science.1256739](https://doi.org/10.1126/science.1256739) PMID: [25278604](https://pubmed.ncbi.nlm.nih.gov/25278604/)
4. Junqueira DM, de Medeiros RM, Matte MCC, Araújo LAL, Chies JAB, Ashton-Prolla P, et al. Reviewing the History of HIV-1: Spread of Subtype B in the Americas. Martin DP, editor. *PLoS One*. 2011; 6: e27489. doi: [10.1371/journal.pone.0027489](https://doi.org/10.1371/journal.pone.0027489) PMID: [22132104](https://pubmed.ncbi.nlm.nih.gov/22132104/)
5. Bello G, Guimarães ML, Morgado MG, Guimara ML. Evolutionary history of HIV-1 subtype B and F infections in Brazil. *AIDS*. 2006; 20: 763–8. PMID: [16514307](https://pubmed.ncbi.nlm.nih.gov/16514307/)
6. Lewis F, Hughes GJ, Rambaut A, Pozniak A, Leigh Brown AJ, Brown AJL. Episodic sexual transmission of HIV revealed by molecular phylodynamics. *PLoS Med*. 2008; 5: e50. doi: [10.1371/journal.pmed.0050050](https://doi.org/10.1371/journal.pmed.0050050) PMID: [18351795](https://pubmed.ncbi.nlm.nih.gov/18351795/)
7. Almeida SE, de Medeiros RM, Junqueira DM, Gräf T, Passaes CPB, Bello G, et al. Temporal dynamics of HIV-1 circulating subtypes in distinct exposure categories in southern Brazil. *Virology*. 2012; 9: 306. doi: [10.1186/1743-422X-9-306](https://doi.org/10.1186/1743-422X-9-306) PMID: [23234345](https://pubmed.ncbi.nlm.nih.gov/23234345/)
8. Rothenberg RB, Potterat JJ, Woodhouse DE, Muth SQ, Darrow WW, Klovdahl AS. Social network dynamics and HIV transmission. *AIDS*. 1998; 12: 1529–36. PMID: [9727575](https://pubmed.ncbi.nlm.nih.gov/9727575/)
9. Frost SDW, Pillay D. Understanding Drivers of Phylogenetic Clustering in Molecular Epidemiological Studies of HIV. *J Infect Dis*. 2015; 211: 856–8. doi: [10.1093/infdis/jiu563](https://doi.org/10.1093/infdis/jiu563) PMID: [25312038](https://pubmed.ncbi.nlm.nih.gov/25312038/)
10. Poon AFY, Joy JB, Woods CK, Shurgold S, Colley G, Brumme CJ, et al. The impact of clinical, demographic and risk factors on rates of HIV transmission: a population-based phylogenetic analysis in British Columbia, Canada. *J Infect Dis*. 2015; 211: 926–35. doi: [10.1093/infdis/jiu560](https://doi.org/10.1093/infdis/jiu560) PMID: [25312037](https://pubmed.ncbi.nlm.nih.gov/25312037/)
11. Pennings PS, Holmes SP, Shafer RW. HIV-1 transmission networks in a small world. *J Infect Dis*. 2014; 209: 180–2. doi: [10.1093/infdis/jit525](https://doi.org/10.1093/infdis/jit525) PMID: [24151310](https://pubmed.ncbi.nlm.nih.gov/24151310/)
12. Hué S, Pillay D, Clewley JP, Pybus OG. Genetic analysis reveals the complex structure of HIV-1 transmission within defined risk groups. *Proc Natl Acad Sci U S A*. 2005; 102: 4425–9. PMID: [15767575](https://pubmed.ncbi.nlm.nih.gov/15767575/)
13. Zehender G, Ebranati E, Lai A, Santoro MM, Alteri C, Giuliani M, et al. Population dynamics of HIV-1 subtype B in a cohort of men-having-sex-with-men in Rome, Italy. *J Acquir Immune Defic Syndr*. 2010; 55: 156–60. doi: [10.1097/QAI.0b013e3181eb3002](https://doi.org/10.1097/QAI.0b013e3181eb3002) PMID: [20703157](https://pubmed.ncbi.nlm.nih.gov/20703157/)

14. Brenner BG, Wainberg M. Future of phylogeny in HIV prevention. *J Acquir Immune Defic Syndr*. 2013; 63 Suppl 2: S248–54. doi: [10.1097/QAI.0b013e3182986f96](https://doi.org/10.1097/QAI.0b013e3182986f96) PMID: [23764643](https://pubmed.ncbi.nlm.nih.gov/23764643/)
15. Ragonnet-cronin M, Ofner-agostini M, Merks H, Pilon R, Rekart M, Archibald CP, et al. Longitudinal Phylogenetic Surveillance Identifies Distinct Patterns of Cluster Dynamics. *J Acquir Immune Defic Syndr*. 2010; 55: 102–108. doi: [10.1097/QAI.0b013e3181e8c7b0](https://doi.org/10.1097/QAI.0b013e3181e8c7b0) PMID: [20622676](https://pubmed.ncbi.nlm.nih.gov/20622676/)
16. Bezemer D, Faria NR, Hassan A, Hamers RL, Mutua G, Anzala O, et al. HIV Type 1 transmission networks among men having sex with men and heterosexuals in Kenya. *AIDS Res Hum Retroviruses*. 2014; 30: 118–26. doi: [10.1089/AID.2013.0171](https://doi.org/10.1089/AID.2013.0171) PMID: [23947948](https://pubmed.ncbi.nlm.nih.gov/23947948/)
17. Yerly S, Junier T, Gayet-Ageron A, Amari EB El, von Wyl V, Günthard HF, et al. The impact of transmission clusters on primary drug resistance in newly diagnosed HIV-1 infection. *AIDS*. 2009; 23: 1415–23. doi: [10.1097/QAD.0b013e32832d40ad](https://doi.org/10.1097/QAD.0b013e32832d40ad) PMID: [19487906](https://pubmed.ncbi.nlm.nih.gov/19487906/)
18. Kaye M, Chibo D, Birch C. Phylogenetic investigation of transmission pathways of drug-resistant HIV-1 utilizing pol sequences derived from resistance genotyping. *J Acquir Immune Defic Syndr*. 2008; 49: 9–16. doi: [10.1097/QAI.0b013e318180c8af](https://doi.org/10.1097/QAI.0b013e318180c8af) PMID: [18667928](https://pubmed.ncbi.nlm.nih.gov/18667928/)
19. Brenner BG, Roger M, Routy J, Moisi D, Ntemgwa M, Matte C, et al. High rates of forward transmission events after acute/early HIV-1 infection. *J Infect Dis*. 2007; 195: 951–9. PMID: [17330784](https://pubmed.ncbi.nlm.nih.gov/17330784/)
20. Chalmet K, Staelens D, Blot S, Dinakis S, Pelgrom J, Plum J, et al. Epidemiological study of phylogenetic transmission clusters in a local HIV-1 epidemic reveals distinct differences between subtype B and non-B infections. *BMC Infect Dis*. 2010; 10: 262. doi: [10.1186/1471-2334-10-262](https://doi.org/10.1186/1471-2334-10-262) PMID: [20822507](https://pubmed.ncbi.nlm.nih.gov/20822507/)
21. Brenner BG, Roger M, Moisi DD, Oliveira M, Hardy I, Turgel R, et al. Transmission networks of drug resistance acquired in primary/early stage HIV infection. *AIDS*. 2008; 22: 2509–15. doi: [10.1097/QAD.0b013e3283121c90](https://doi.org/10.1097/QAD.0b013e3283121c90) PMID: [19005274](https://pubmed.ncbi.nlm.nih.gov/19005274/)
22. Hughes GJ, Fearnhill E, Dunn D, Lycett SJ, Rambaut A, Leigh Brown AJ. Molecular phylodynamics of the heterosexual HIV epidemic in the United Kingdom. *PLoS Pathog*. 2009; 5: e1000590. doi: [10.1371/journal.ppat.1000590](https://doi.org/10.1371/journal.ppat.1000590) PMID: [19779560](https://pubmed.ncbi.nlm.nih.gov/19779560/)
23. Audelin AM, Cowan S, Obel N, Nielsen C, Jørgensen LB, Gerstoft J. Phylogenetics of the Danish HIV epidemic: the role of very late presenters in sustaining the epidemic. *J Acquir Immune Defic Syndr*. 2013; 62: 102–8. doi: [10.1097/QAI.0b013e318276becc](https://doi.org/10.1097/QAI.0b013e318276becc) PMID: [23075917](https://pubmed.ncbi.nlm.nih.gov/23075917/)
24. Mehta SR, Delport W, Brouwer KC, Espitia S, Patterson T, Pond SK, et al. The relatedness of HIV epidemics in the United States-Mexico border region. *AIDS Res Hum Retroviruses*. 2010; 26: 1273–7. doi: [10.1089/aid.2010.0021](https://doi.org/10.1089/aid.2010.0021) PMID: [20977301](https://pubmed.ncbi.nlm.nih.gov/20977301/)
25. Mir D, Cabello M, Romero H, Bello G. Phylodynamics of major HIV-1 subtype B pandemic clades circulating in Latin America. *AIDS*. 2015; 29: 1863–1869. doi: [10.1097/QAD.0000000000000770](https://doi.org/10.1097/QAD.0000000000000770) PMID: [26372392](https://pubmed.ncbi.nlm.nih.gov/26372392/)
26. Lemey P, Rambaut A, Bedford T, Faria N, Bielejec F, Baele G, et al. Unifying viral genetics and human transportation data to predict the global transmission dynamics of human influenza H3N2. *PLoS Pathog*. 2014; 10: e1003932. doi: [10.1371/journal.ppat.1003932](https://doi.org/10.1371/journal.ppat.1003932) PMID: [24586153](https://pubmed.ncbi.nlm.nih.gov/24586153/)
27. Paraskevis D, Pybus O, Magiorkinis G, Hatzakis A, Wensing AMJ, Vijver DA, et al. Tracing the HIV-1 subtype B mobility in Europe: a phylogeographic approach. *Retrovirology*. 2009; 6: 49. doi: [10.1186/1742-4690-6-49](https://doi.org/10.1186/1742-4690-6-49) PMID: [19457244](https://pubmed.ncbi.nlm.nih.gov/19457244/)
28. Gilbert MTP, Rambaut A, Wlasiuk G, Spira TJ, Pitchenik AE, Worobey M. The emergence of HIV/AIDS in the Americas and beyond. *Proc Natl Acad Sci U S A*. 2007; 104: 18566–70. PMID: [17978186](https://pubmed.ncbi.nlm.nih.gov/17978186/)
29. Alcantara LCJ, Cassol S, Libin P, Deforche K, Pybus OG, Van Ranst M, et al. A standardized framework for accurate, high-throughput genotyping of recombinant and non-recombinant viral sequences. *Nucleic Acids Res*. 2009; 37: 1–9.
30. Kosakovsky Pond SL, Posada D, Stawiski E, Chappey C, Poon AFY, Hughes G, et al. An evolutionary model-based algorithm for accurate phylogenetic breakpoint mapping and subtype prediction in HIV-1. *PLoS Comput Biol*. 2009; 5: e1000581. doi: [10.1371/journal.pcbi.1000581](https://doi.org/10.1371/journal.pcbi.1000581) PMID: [19956739](https://pubmed.ncbi.nlm.nih.gov/19956739/)
31. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, et al. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol*. 1999; 73: 152–60. PMID: [9847317](https://pubmed.ncbi.nlm.nih.gov/9847317/)
32. McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, et al. Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res*. 2013; 41: W597–600. doi: [10.1093/nar/gkt376](https://doi.org/10.1093/nar/gkt376) PMID: [23671338](https://pubmed.ncbi.nlm.nih.gov/23671338/)
33. Johnson V, Calvez V, Günthard HF, Paredes R, Pillay D, Shafer RW, et al. Update of the drug resistance mutations in HIV-1: March 2013. *Top Antivir Med*. 2013; 21: 6–14. PMID: [23596273](https://pubmed.ncbi.nlm.nih.gov/23596273/)
34. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014; 30: 1312–3. doi: [10.1093/bioinformatics/btu033](https://doi.org/10.1093/bioinformatics/btu033) PMID: [24451623](https://pubmed.ncbi.nlm.nih.gov/24451623/)

35. Miller MA, Pfeiffer W, Schwartz T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Proceedings of the Gateway Computing Environments Workshop (GCE). New Orleans; 2010. pp. 1–8.
36. Tamura K, Stecher G, Peterson D, Filipksi A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol.* 2013; 30: 2725–9. doi: [10.1093/molbev/mst197](https://doi.org/10.1093/molbev/mst197) PMID: [24132122](https://pubmed.ncbi.nlm.nih.gov/24132122/)
37. Anisimova M, Gascuel O. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Syst Biol.* 2006; 55: 539–52. PMID: [16785212](https://pubmed.ncbi.nlm.nih.gov/16785212/)
38. Ragonnet-Cronin M, Hodcroft E, Hué S, Fearnhill E, Delpuch V, Brown AJL, et al. Automated analysis of phylogenetic clusters. *BMC Bioinformatics.* 2013; 14: 317. doi: [10.1186/1471-2105-14-317](https://doi.org/10.1186/1471-2105-14-317) PMID: [24191891](https://pubmed.ncbi.nlm.nih.gov/24191891/)
39. Eyer-Silva WA, Morgado MG. Autochthonous horizontal transmission of a CRF02\_AG strain revealed by a human immunodeficiency virus type 1 diversity survey in a small city in inner state of Rio de Janeiro, Southeast Brazil. *Mem Inst Oswaldo Cruz.* 2007; 102: 809–815. PMID: [17992366](https://pubmed.ncbi.nlm.nih.gov/17992366/)
40. Eyer-Silva WA, Couto-Fernandez JC, Morgado MG. Molecular epidemiology of HIV type 1 in inner Rio De Janeiro State, Brazil. *AIDS Res Hum Retroviruses.* 2007; 23: 303–8. PMID: [17331037](https://pubmed.ncbi.nlm.nih.gov/17331037/)
41. Eyer-Silva WA, Morgado MG. Molecular epidemiology of HIV-1 infection in a small Brazilian county: usefulness of envelope and polymerase sequences to epidemiologic studies. *J Acquir Immune Defic Syndr.* 2006; 41: 664–70. PMID: [16652042](https://pubmed.ncbi.nlm.nih.gov/16652042/)
42. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol.* 2012; 29: 1969–73. doi: [10.1093/molbev/mss075](https://doi.org/10.1093/molbev/mss075) PMID: [22367748](https://pubmed.ncbi.nlm.nih.gov/22367748/)
43. Bello G, Eyer-silva W, Couto-Fernandez JC, Guimarães ML, Chequer-Fernandez SL, Teixeira SLM, et al. Demographic history of HIV-1 subtypes B and F in Brazil. *Infect Genet Evol.* 2007; 7: 263–70. PMID: [17150415](https://pubmed.ncbi.nlm.nih.gov/17150415/)
44. Rambaut A, Drummond A. Tracer v1.6 [Internet]. 2007.
45. Gifford RJ, Liu TF, Rhee S-Y, Kiuchi M, Hue S, Pillay D, et al. The calibrated population resistance tool: standardized genotypic estimation of transmitted HIV-1 drug resistance. *Bioinformatics.* 2009; 25: 1197–8. doi: [10.1093/bioinformatics/btp134](https://doi.org/10.1093/bioinformatics/btp134) PMID: [19304876](https://pubmed.ncbi.nlm.nih.gov/19304876/)
46. Ashkenazy H, Penn O, Doron-Faigenboim A, Cohen O, Cannarozzi G, Zomer O, et al. FastML: a web server for probabilistic reconstruction of ancestral sequences. *Nucleic Acids Res.* 2012; 40: W580–4. doi: [10.1093/nar/gks498](https://doi.org/10.1093/nar/gks498) PMID: [22661579](https://pubmed.ncbi.nlm.nih.gov/22661579/)
47. Leigh Brown AJ, Lycett SJ, Weinert L, Hughes GJ, Fearnhill E, Dunn DT. Transmission network parameters estimated from HIV sequences for a nationwide epidemic. *J Infect Dis.* 2011; 204: 1463–9. doi: [10.1093/infdis/jir550](https://doi.org/10.1093/infdis/jir550) PMID: [21921202](https://pubmed.ncbi.nlm.nih.gov/21921202/)
48. Beyrer C, Baral SD, Weir BW, Curran JW, Chaisson RE, Sullivan PS. A call to action for concentrated HIV epidemics. *Curr Opin HIV AIDS.* 2014; 9: 95–100. doi: [10.1097/COH.000000000000043](https://doi.org/10.1097/COH.000000000000043) PMID: [24499807](https://pubmed.ncbi.nlm.nih.gov/24499807/)
49. Volz EM, Ionides E, Romero-Severson EO, Brandt M-G, Mokotoff E, Koopman JS. HIV-1 Transmission during Early Infection in Men Who Have Sex with Men: A Phylodynamic Analysis. Hallett TB, editor. *PLoS Med.* 2013; 10: e1001568. doi: [10.1371/journal.pmed.1001568](https://doi.org/10.1371/journal.pmed.1001568) PMID: [24339751](https://pubmed.ncbi.nlm.nih.gov/24339751/)
50. Hallett TB. Early HIV infection in the United States: a virus's eye view. *PLoS Med.* 2013; 10: e1001569. doi: [10.1371/journal.pmed.1001569](https://doi.org/10.1371/journal.pmed.1001569) PMID: [24339752](https://pubmed.ncbi.nlm.nih.gov/24339752/)
51. Beyrer C, Sullivan P, Sanchez J, Baral SD, Collins C, Wirtz AL, et al. The increase in global HIV epidemics in MSM. *AIDS.* 2013; 27: 2665–78. doi: [10.1097/01.aids.0000432449.30239.fe](https://doi.org/10.1097/01.aids.0000432449.30239.fe) PMID: [23842129](https://pubmed.ncbi.nlm.nih.gov/23842129/)
52. Jia Y, Aliyu MH, Jennifer Huang Z. Dynamics of the HIV epidemic in MSM. *Biomed Res Int.* 2014; 2014: 497543. doi: [10.1155/2014/497543](https://doi.org/10.1155/2014/497543) PMID: [25093171](https://pubmed.ncbi.nlm.nih.gov/25093171/)
53. Sinha S, Shekhar RC, Ahmad H, Kumar N, Samantaray JC, Sreenivas V, et al. Prevalence of HIV drug resistance mutation in the northern Indian population after failure of the first line antiretroviral therapy. *Curr HIV Res.* 2012; 10: 532–8. PMID: [22716105](https://pubmed.ncbi.nlm.nih.gov/22716105/)
54. Varella RB, Ferreira SB, Castro MB de, Tavares MD, Zalis MG. Prevalence of resistance-associated mutations in Human Immunodeficiency Virus type 1-positive individuals failing HAART in Rio de Janeiro, Brazil. *Brazilian J Infect Dis.* 2008; 12: 380–384.
55. WHO. The HIV drug resistance report. 2012.
56. Wertheim JO, Leigh Brown AJ, Hepler NL, Mehta SR, Richman DD, Smith DM, et al. The global transmission network of HIV-1. *J Infect Dis.* 2014; 209: 304–13. doi: [10.1093/infdis/jit524](https://doi.org/10.1093/infdis/jit524) PMID: [24151309](https://pubmed.ncbi.nlm.nih.gov/24151309/)
57. Biesinger T, Kimata JT. HIV-1 Transmission, Replication Fitness and Disease Progression. *Virology (Auckl).* 2008; 2008: 49–63.

58. van de Vijver DAMC, Wensing AMJ, Åsjö B, Bruckova M, Jorgensen LB, Camacho R, et al. HIV-1 drug-resistance patterns among patients on failing treatment in a large number of European countries. *Acta dermatovenerologica Alpina, Pannonica, Adriat.* 2010; 19: 3–9.