Dissemination Dynamics of HIV-1 Subtype B Pandemic and Non-pandemic Lineages Circulating in Amazonas, Brazil

Ighor Arantes[1], Tiago Gräf[2], Paula Andrade[1], Yury Oliveira Chaves[3], Monick Lindenmeyer Guimarães[1] and Gonzalo Bello[1]*

[1] Laboratório de AIDS e Imunologia Molecular, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz (FIOCRUZ), Rio de Janeiro, Brazil, [2] Instituto Gonçalo Moniz, Fundação Oswaldo Cruz (FIOCRUZ), Salvador, Brazil, [3] Laboratório de Diagnóstico e Controle de Doenças Infecciosas na Amazônia, Instituto Leônidas e Maria Deane, Fundação Oswaldo Cruz (FIOCRUZ), Manaus, Brazil

The HIV-1 epidemic in the Amazonas state, as in most of Brazil, is dominated by subtype B. The state, nonetheless, is singular for its significant co-circulation of the variants B_{CAR}, which can mostly be found in the Caribbean region, and B_{PAN}, a clade that emerged in the United States and aggregates almost the totality of subtype B infections world-wide. The Amazonian HIV-1 epidemic provides a unique scenario to compare the epidemic potential of B_{PAN} and B_{CAR} clades spreading in the same population. To reconstruct the spatiotemporal dynamic and demographic history of both subtype B lineages circulating in Amazonas, we analyzed 1,272 HIV-1 *pol* sequences sampled in that state between 2009 and 2018. Our phylogeographic analyses revealed that while most B_{CAR} infections resulted from a single successful founder event that took place in the Amazonas state around the late 1970s, most B_{PAN} infections resulted from the expansion of multiple clusters seeded in the state since the late 1980s. Our data support the existence of at least four large clusters of the pandemic form in Amazonas, two of them nested in Brazil's largest known subtype B cluster (B_{BR-I}), and two others resulting from new introductions detected here. The reconstruction of the demographic history of the most prevalent B_{PAN} ($n = 4$) and B_{CAR} ($n = 1$) clades identified in Amazonas revealed that all clades displayed a continuous expansion [effective reproductive number (R_e) > 1] until most recent times. During the period of co-circulation from the late 1990s onward, the R_e of Amazonian B_{PAN} and B_{CAR} clusters behaved quite alike, fluctuating between 2.0 and 3.0. These findings support that the B_{CAR} and B_{PAN} variants circulating in the Brazilian state of Amazonas displayed different evolutionary histories, but similar epidemic trajectories and transmissibility over the last two decades, which is consistent with the notion that both subtype B variants display comparable epidemic potential. Our findings also revealed that despite significant advances in the treatment of HIV infections in the Amazonas state, B_{CAR} and B_{PAN} variants continue to expand and show no signs of the epidemic stabilization observed in other parts of the country.

Keywords: HIV-1, subtype B, Brazil, Amazonas, phylodynamics

INTRODUCTION

The HIV-1 subtype B pandemic started when the ancestral virus arrived and first established itself in the Caribbean region during the mid-1960s (Gilbert et al., 2007). The subsequent subtype B spread generated a set of local clades, designated as B_{CAR}, that remained mostly confined to the Caribbean region (Gilbert et al., 2007; Cabello et al., 2014). One of those viruses, however, migrated from the Caribbean to the United States around the late 1960s and established a pandemic clade, called B_{PAN}, that was then disseminated worldwide (Worobey et al., 2016). The remarkable dissemination of the B_{PAN} clade was probably shaped by ecological, rather than virological factors (Gilbert et al., 2007; Arantes et al., 2020); but there are no studies comparing the epidemic potential of B_{CAR} and B_{PAN} clades co-circulating outside the Caribbean region.

The HIV-1 subtype B epidemic in Brazil is mainly driven by the B_{PAN} clade, with a few notable exceptions, like the Northern state of Amazonas, which is characterized by the co-circulation of the B_{PAN} (75%) and B_{CAR} (25%) variants at a high prevalence (Divino et al., 2016; Crispim et al., 2019; Gräf et al., 2021). Among regional HIV-1 epidemics in Brazil, the one in the state of Amazonas stands out as the second largest AIDS detection rate (34.8 cases per 100,000 inhabitants) in the country, well above the mean national rate (17.8 cases per 100,000 inhabitants) (Ministério da Saúde, 2020). Furthermore, despite important advances in HIV diagnosis and treatment, the Amazonas HIV/AIDS epidemic is not stabilized and has displayed a rising AIDS incidence over the last 10 years (Ministério da Saúde, 2020).

The Amazonian HIV-1 epidemic thus provides a great opportunity to compare the epidemic dynamics of the B_{PAN} and B_{CAR} clades spreading in the same population outside the Caribbean region. Previous studies identified four major B_{PAN} ($B_{PAN-BR-I}$ to $B_{PAN-BR-IV}$) and four major B_{CAR} ($B_{CAR-BR-I}$ to $B_{CAR-BR-IV}$) clusters circulating in Brazil (Mir et al., 2015; Divino et al., 2016). The lineage $B_{CAR-BR-I}$ aggregates the majority (51%) of non-pandemic subtype B sequences from Brazil (Divino et al., 2016) and its most recent common ancestor (MRCA) could be traced back to Amazonas in the late 1970s, from a viral migration probably originated in the French Guiana (Divino et al., 2016; Arantes et al., 2019). The cluster composition and the evolutionary history of the B_{PAN} clade in the Amazonas state are currently unknown.

This work aims to characterize the spatiotemporal dynamics of the major B_{PAN} clusters circulating in the state of Amazonas and to compare the evolutionary and demographic history of major pandemic and non-pandemic subtype B lineages that drive the expanding HIV-1 subtype B epidemic in this Northern Brazilian state.

METHODOLOGY

HIV-1 Subtype B *pol* Sequence Dataset From Amazonas State

In this study, we used a total of 1,272 HIV-1 subtype B *pol* sequences (nucleotides 2,253–3,275 of reference strain HXB2) from Amazonas state sampled between 2009 and 2018 that were either available at the Los Alamos HIV Database[1] by March 2021 or were recently published (Chaves et al., 2021; Gräf et al., 2021) and made available in GenBank[2]. Only one sequence per subject was selected. Sequences were aligned using the ClustalW program (Larkin et al., 2007) and all sites associated with major antiretroviral drug resistance in protease and reverse transcriptase were excluded.

Phylogenetic Classification of Amazonian HIV-1 Subtype B Sequences

Amazonian HIV-1 subtype B sequences were initially classified as B_{CAR} or B_{PAN} and next within major B_{CAR} or B_{PAN} Brazilian clusters by using an evolutionary placement algorithm (EPA) available in RAxML v.8.0.0 (Stamatakis, 2014) for the rapid assignment of query sequences to edges of a reference phylogenetic tree under a maximum-likelihood (ML) model. This analysis allowed us to classify sequences within 10 different clades: $B_{CAR-BR-I}$ to $B_{CAR-BR-IV}$, $B_{PAN-BR-I}$ to $B_{PAN-BR-IV}$, other B_{CAR} lineages, and other B_{PAN} lineages.

Selection of Brazilian and Global Reference HIV-1 Subtype B *Pol* Datasets

After the initial cluster assignment, the HIV-1 subtype B *pol* Amazonian sequences were aligned with different sub-sets of non-Amazonian subtype B *pol* reference sequences (covering nucleotides 2,253–3,260 relative to HXB2 genome). Amazonian subtype B sequences classified within major B_{CAR} ($B_{CAR-BR-I}$ to $B_{CAR-BR-IV}$) or B_{PAN} ($B_{PAN-BR-I}$ to $B_{PAN-BR-IV}$) Brazilian clusters were aligned with Brazilian sequences from other states that were previously classified within those major lineages (Mir et al., 2015; Divino et al., 2016). Amazonian subtype B sequences classified as "others B_{CAR} lineages" were aligned with sequences representative of the B_{CAR} diversity in the Caribbean region ($n = 228$) that were also described previously (Cabello et al., 2014; Mendoza et al., 2014). Finally, Amazonian subtype B sequences classified as "others B_{PAN} lineages" were aligned with: (i) one subset of closely related B_{PAN} sequences from Brazil ($n = 687$) selected from a large dataset of Brazilian sequences ($n = 88,441$) described previously (Gräf et al., 2021) and (ii) one subset of closely related B_{PAN} sequences from different countries ($n = 1,700$) selected from a large dataset of worldwide sampled sequences ($n = 71,160$) recovered from Los Alamos HIV Database. We used the basic local alignment search tool (BLAST)[3] to select the 10 subtype B reference sequences (Brazilian and worldwide) with the highest similarity score (>95%) to each Amazonian subtype B sequence.

Detection of Major Amazonian HIV-1 Subtype B Clades

Amazonian and non-Amazonian subtype B sequences were subject to ML phylogeographic analyses to identify the B_{CAR} and B_{PAN} sub-clusters that probably originated in Amazonas.

---

[1]http://www.hiv.lanl.gov

[2]https://www.ncbi.nlm.nih.gov/genbank/

[3]www.ncbi.nlm.nih.gov/BLAST

ML phylogenetic trees were inferred with the PhyML program (Guindon et al., 2010) using an online web server (Guindon et al., 2005) under the GTR + I + $\Gamma 4$ nucleotide substitution model, as selected by the jModelTest program (Posada, 2008), and the SPR branch-swapping algorithm of heuristic tree search. The reliability of the obtained tree topology was estimated with the approximate likelihood-ratio test (aLRT) (Anisimova and Gascuel, 2006) based on the Shimodaira-Hasegawa-like procedure. Trees were rooted using subtype D sequences and visualized using the FigTree v1.4.0 software (Rambaut, 2009). The ML trees were employed for the ancestral character state reconstruction (ACR) of epidemic locations with PastML (Ishikawa et al., 2019), using the maximum likelihood Joint (Pupko et al., 2000) and marginal posterior probabilities approximation (MPPA) methods with an F81-like model. Beyond Amazonas, remaining Brazilian states were aggregated in five discrete locations according to their geographic regions and sequences from other countries were aggregated in a single "non-Brazilian" location. Amazonian B_{CAR} and B_{PAN} clades were defined as those monophyletic clusters with high support (aLRT $\geq$ 0.85) that were mostly composed of sequences from Amazonas (>85%) and displayed Amazonas as the most probable ($P \geq 0.85$) state location of its MRCA. Amazonian clades were further subdivided according to size into large ($n > 30$), medium ($n = 10$–30), and small ($n = 2$–9) clades.

Phylodynamic Analysis

The study of epidemiological and evolutionary parameters of major B_{CAR} and $B_{PANDEMIC}$ clusters from Amazonas was done by Bayesian inference using coalescent and birth-death tree priors implemented, respectively, in BEAST v1.10 (Drummond et al., 2002; Suchard et al., 2018) with BEAGLE (Suchard and Rambaut, 2009) to improve run-time, and in BEAST 2.6 (Bouckaert et al., 2019) software packages. Dated phylogenies were inferred with the flexible Bayesian skyline coalescent model (Drummond, 2005). Changes across time in their effective sample size (N_e) were estimated using the coalescent Bayesian Skygrid (BSKG) model (Gill et al., 2012), and in their effective reproductive number (R_e), using the Birth-death Skyline (BDSKY) model (Stadler et al., 2012). For BDSKY, the sampling rate (δ) was set to zero for the period before the oldest sample and estimated afterward. The R_e was modeled in a piecewise manner in equidistant intervals from the most recent sample up to the root of the tree with a lognormal prior (mean = 0; standard deviation = 1), and the becoming non-infectious rate with a lognormal prior (mean = 0.25; standard deviation = 0.5). All Bayesian MCMC analyses were performed using the GTR + I + $\Gamma 4$ nucleotide substitution model, and a relaxed uncorrelated lognormal molecular clock model (Drummond et al., 2006) with a uniform prior distribution on the substitution rate that encompasses mean values previously estimated for the subtype B pol gene (2.0–3.0×10^{-3} subst./site/year) (Hue et al., 2005; Zehender et al., 2010; Chen et al., 2011; Mendoza et al., 2014; Bello et al., 2018). MCMC chains were run for 50–100×10^6 generations and convergence and uncertainty of parameter estimates were assessed by calculating the effective sample size (ESS) and 95% highest probability density (HPD)

values, respectively, after excluding the initial 10% of each run with Tracer v1.7.1 (Rambaut et al., 2018). The convergence of parameters was considered when ESS $\geq$ 200.

Statistical Analysis

Demographic information of age group and gender of individuals with samples included in the present study were compared using Pearson's chi-squared test as implemented in R version 3.6.3 (R Core Team, 2018), with 10,000 replicates. The false discovery rate (FDR) method was used to correct for multiple hypothesis testing and to reduce false positives. Statistical significance was defined as p-values <0.05.
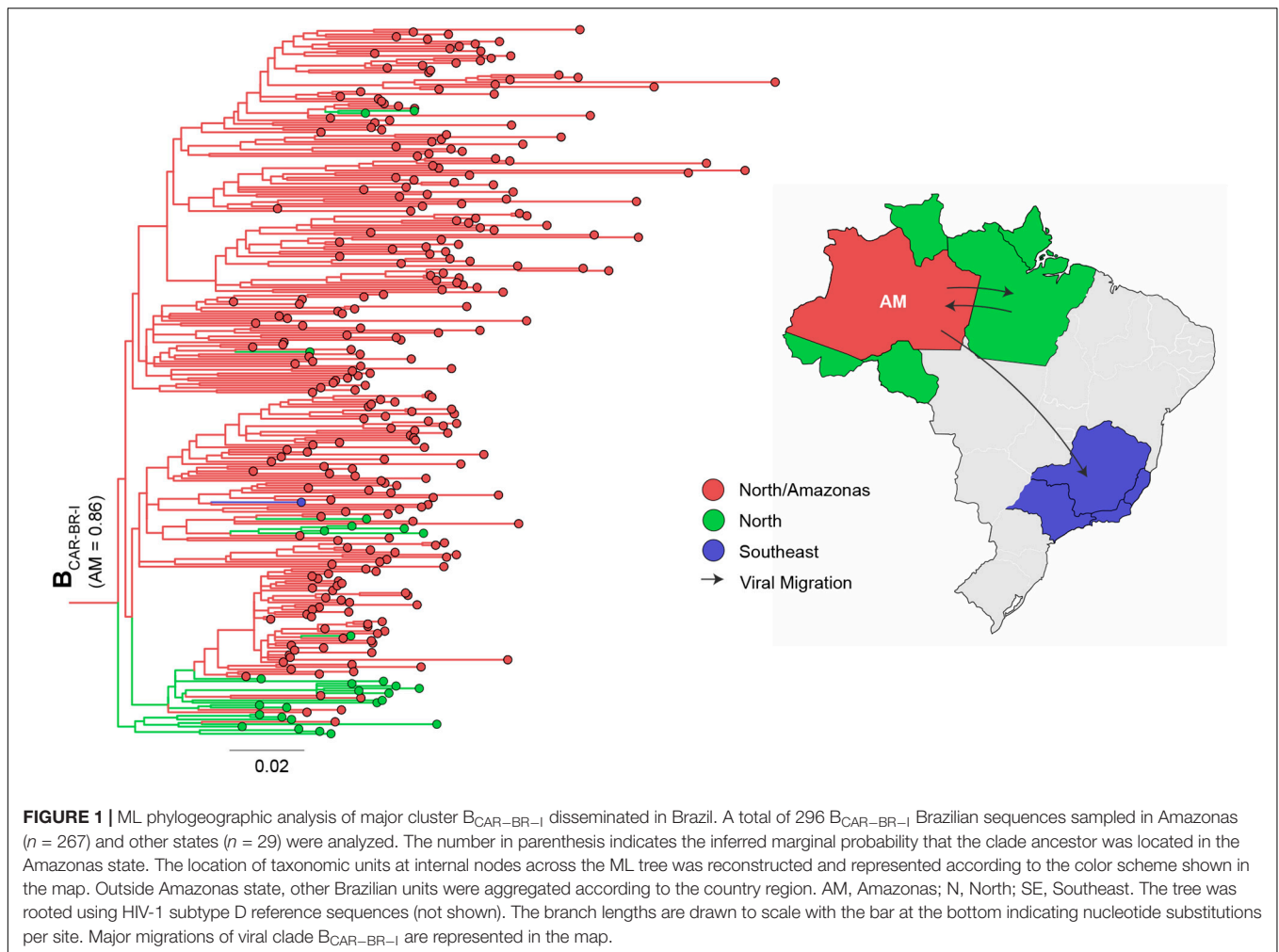
RESULTS

The 1,272 HIV-1 subtype B pol sequences from Amazonas were assigned to either B_{CAR} (23%) or B_{PAN} (77%) lineages (**Table 1**). The sub-lineage assignment reveals that most B_{CAR} Amazonian sequences belong to the major Brazilian clade $B_{CAR-BR-I}$ (89%), while the remaining sequences were classified within clades $B_{CAR-BR-II}$ (1%), $B_{CAR-BR-III}$ (1%), or branched outside known Brazilian clades (8%) (**Table 1**). Although a high proportion of B_{PAN} Amazonian sequences also branched within major Brazilian B_{PAN} clades (45%), particularly the $B_{PAN-BR-I}$ (37%), most pandemic sequences (55%) branched outside known countrywide Brazilian clades (**Table 1**). To identify the major subtype clades circulating in Amazonas, we next conducted independent ML phylogeographic analyses for: (i) sequences that branched within major Brazilian B_{CAR} or B_{PAN} clades, by combining Brazilian sequences sampled in Amazonas and other states, and (ii) sequences that branched outside major Brazilian clades, by combining Amazonian sequences with either B_{CAR} sequences of Caribbean origin or B_{PAN} sequences sampled in

TABLE 1 | Lineage classification of HIV-1 subtype B pol sequences from Amazonas state.

Lineage	Sub-lineage	N (%)	Sampling range
B_{PAN}	$B_{PAN-BR-I}$	360 (37%)	2009–2018
	$B_{PAN-BR-II}$	38 (4%)	
	$B_{PAN-BR-III}$	23 (2%)	
	$B_{PAN-BR-IV}$	23 (2%)	
	Others	530 (55%)	
	Total	**974 (100%)**	
B_{CAR}	$B_{CAR-BR-I}$	267 (90%)	2009–2018
	$B_{CAR-BR-II}$	2 (1%)	2016–2017
	$B_{CAR-BR-IV}$	4 (1%)	2015–2017
	Others	25 (8%)	2009–2018
	Total	**298 (100%)**	

The table details the distribution of 1,272 HIV-1 subtype B pol sequences (nucleotides 2,253–3,275 of reference strain HXB2) from Brazil's Amazonas state across the known clusterization profile of the pandemic (B_{PAN}, n = 974, 77%) and non-pandemic (B_{CAR}, n = 298, 23%) forms. For each cluster, its absolute and relative frequency of samples from Amazonas state, as well as their distribution in time, are indicated. The class "Other" aggregates sequences not clustered among established major Brazilian clusters.

FIGURE 1 | ML phylogeographic analysis of major cluster $B_{CAR-BR-I}$ disseminated in Brazil. A total of 296 $B_{CAR-BR-I}$ Brazilian sequences sampled in Amazonas ($n = 267$) and other states ($n = 29$) were analyzed. The number in parenthesis indicates the inferred marginal probability that the clade ancestor was located in the Amazonas state. The location of taxonomic units at internal nodes across the ML tree was reconstructed and represented according to the color scheme shown in the map. Outside Amazonas state, other Brazilian units were aggregated according to the country region. AM, Amazonas; N, North; SE, Southeast. The tree was rooted using HIV-1 subtype D reference sequences (not shown). The branch lengths are drawn to scale with the bar at the bottom indicating nucleotide substitutions per site. Major migrations of viral clade $B_{CAR-BR-I}$ are represented in the map.

Brazil and worldwide. The ML phylogeographic analyses confirm multiple introductions of B_{CAR} ($n = 18$) and B_{PAN} ($n = 291$) variants into the Amazonas state (**Supplementary Table 1**). The major B_{CAR} founder event resulted in the $B_{CAR-BR-I}$ clade (**Figure 1**) while the remaining B_{CAR} sequences were distributed among a few local clusters of small size (2–9 sequences, 7%) or singletons (4%) (**Supplementary Figures 1A–C**). A few B_{PAN} introductions ($n = 4$) originated highly supported (aLRT > 0.85) Amazonian B_{PAN} clades of large size that were mostly composed by sequences from Amazonas (>85%) and most probably arose in Amazonas ($P > 0.90$). Two major clusters, $B_{PAN-BR-I-AM-I}$ ($n = 39$) and $B_{PAN-BR-I-AM-II}$ ($n = 35$), were nested within the large Brazilian clade $B_{PAN-BR-I}$ (**Figure 2**). The other two clusters, $B_{PAN-AM-I}$ ($n = 86$) and $B_{PAN-AM-II}$ ($n = 60$), branched outside the major B_{PAN} Brazilian clades (**Figure 3**). Because the phylogenetic placement of some Amazonian basal sequences in clusters $B_{PAN-AM-I}$ and $B_{PAN-AM-II}$ changed according to the reference (Brazilian or worldwide) dataset, we define the final size of those clusters by the monophyletic groups supported by both analyses. The four major Amazonian B_{PAN} clades, together, comprise 23% of all B_{PAN} Amazonian sequences analyzed and the remaining sequences branched within Amazonian clades of

medium (24%) or small (35%) size or appeared as singletons that branched with non-Amazonian sequences (18%) (**Figures 2**, **3** and **Supplementary Figures 1D–F**).

To study in more detail the evolutionary and demographic history of lineages B_{CAR} and B_{PAN} spreading in Amazonas, we selected the five major clades that display both local epidemic importance – as, combined, they comprise 38% of HIV-1 subtype B infections in the state – and adequate sample sizes to give reliable demographic estimates. Time-scaled trees were reconstructed using a Bayesian coalescent model with an informative clock rate prior due to the weak temporal structure of Amazonian subtype B *pol* datasets (**Supplementary Figure 2**). Posterior estimates, that were, as expected, largely influenced by the selected clock rate prior, traced the median T_{MRCA} of major Amazonian clades to the late 1970s for $B_{CAR-BR-I}$, the late 1980s for $B_{PAN-AM-I}$, the mid-1990s for $B_{PAN-AM-II}$, and the late 1990s for $B_{PAN-BR-I-AM-I}$ and $B_{PAN-BR-I-AM-II}$ (**Table 2**). These findings support that the major B_{CAR} clade was successfully spreading in Amazonas for about 10 years before the emergence of the major B_{PAN} clades, which may explain the singular high prevalence of non-pandemic subtype B variants in Amazonas with respect to most other Brazilian states. The

FIGURE 2 | ML phylogeographic analysis of major $B_{PAN-BR-I}$ cluster disseminated in Brazil. A total of 844 $B_{PAN-BR-I}$ Brazilian sequences sampled in Amazonas (n = 360) and other states from the Northern and Southeastern regions (n = 484) were analyzed together. We selected sequences from Northern and Southeastern states because were the regions more strongly connected with the Amazonas state. Major Amazonian clades $B_{PAN-BR-I-AM-I}$ and $B_{PAN-BR-I-AM-II}$ are indicated by colored shaded boxes along with cluster aLRT support (number at basal branch) and the inferred marginal probability that cluster ancestor was located in Amazonas state (in parentheses). The location of taxonomic units at internal nodes across the ML tree was reconstructed and represented according to the color scheme shown in the map. Sequences were grouped in three discrete locations: Amazonas state (AM), Northern region (N), and Southeastern region (SE). The tree was rooted using HIV-1 subtype D reference sequences (not shown). The branch lengths are drawn to scale with the bar at the bottom indicating nucleotide substitutions per site. Major migrations of viral clades $B_{PAN-BR-I-AM-I}$ and $B_{PAN-BR-I-AM-II}$ are represented in the map.

BSKG model supports that the N_e of lineages $B_{CAR-BR-I}$ and $B_{PAN-BR-I-AM-I}$ steadily increased until recent years, while the N_e of lineages $B_{PAN-AM-I}$, $B_{PAN-AM-II}$, and $B_{PAN-BR-I-AM-II}$ increased until the late-2000s, but then stabilized in more recent years (**Figures 4A–E**). The temporal trajectories of the R_e estimated using the Bayesian BDSKY model, however, support that all major Amazonian HIV-1 subtype B clades continuously expanded (median $R_e > 1$) over all the studied period, with some temporal fluctuations in the rate of expansion (**Figures 4A–E**). The clade $B_{CAR-BR-I}$ reached the highest median R_e (2.5–2.6) between the late 1970s and the early 1990s, while the B_{PAN} clades reached the highest median R_e (2.9–3.4) between the mid-1990s and the mid-2000s (**Table 2**). Despite those differences in the early phase of spread, all major Amazonian subtype B clades

converge to the roughly similar median growth rate (R_e = 1.6–2.3) at the most recent time period analyzed (2010–2018), with no evidence of recent epidemic stabilization ($R_e > 1$) (**Figure 4F**).

DISCUSSION

Previous studies demonstrate that the expanding HIV-1 subtype B epidemic in the Northern Brazilian state of Amazonas was driven by both pandemic (B_{PAN}) and non-pandemic (B_{CAR}) viral variants (Cabello et al., 2015; Divino et al., 2016; Arantes et al., 2019; Crispim et al., 2019; Chaves et al., 2021; Gräf et al., 2021), thus creating a great opportunity to compare the epidemic dynamics of both subtype B forms spreading in
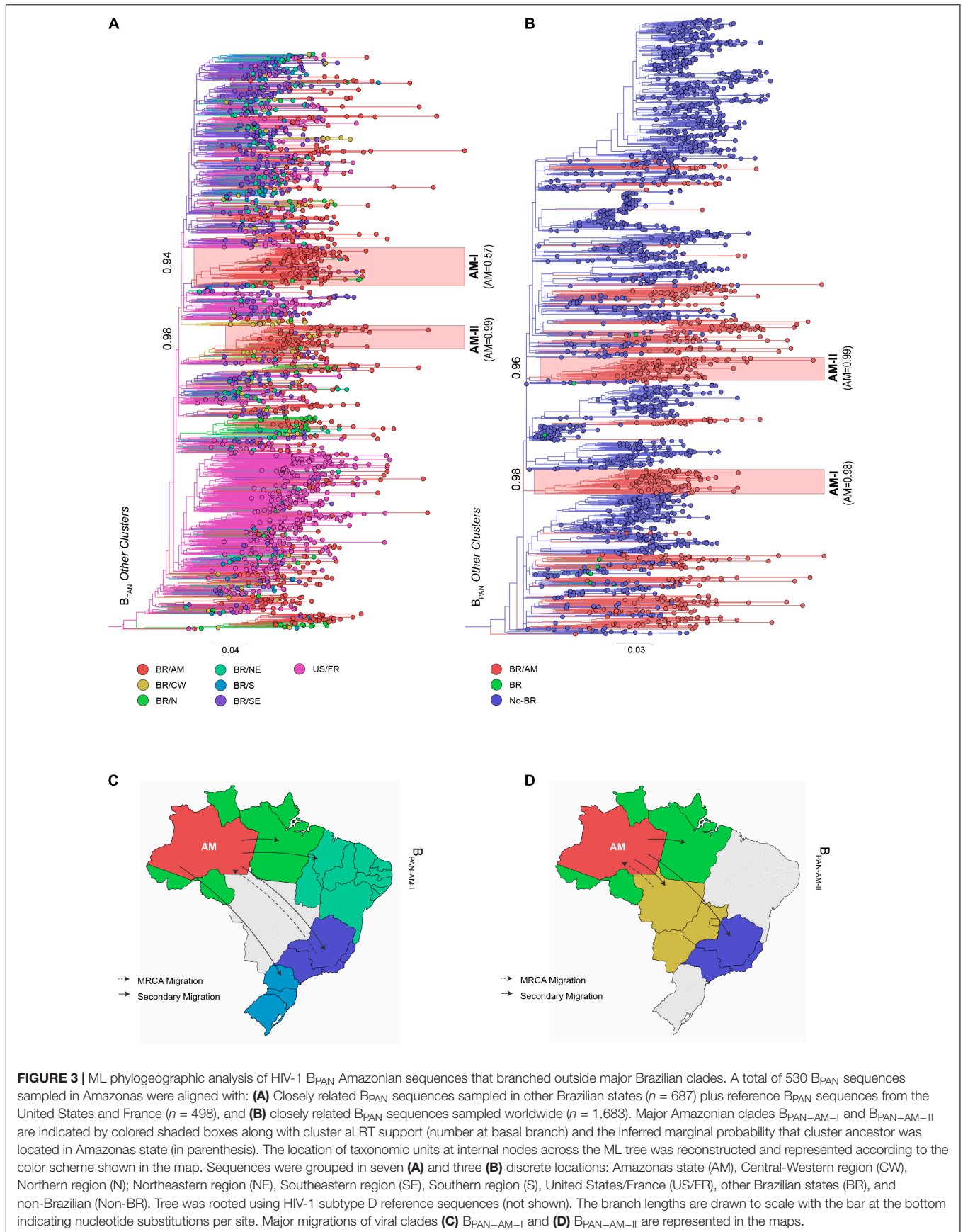
FIGURE 3 | ML phylogeographic analysis of HIV-1 B_{PAN} Amazonian sequences that branched outside major Brazilian clades. A total of 530 B_{PAN} sequences sampled in Amazonas were aligned with: **(A)** Closely related B_{PAN} sequences sampled in other Brazilian states (n = 687) plus reference B_{PAN} sequences from the United States and France (n = 498), and **(B)** closely related B_{PAN} sequences sampled worldwide (n = 1,683). Major Amazonian clades $B_{PAN-AM-I}$ and $B_{PAN-AM-II}$ are indicated by colored shaded boxes along with cluster aLRT support (number at basal branch) and the inferred marginal probability that cluster ancestor was located in Amazonas state (in parenthesis). The location of taxonomic units at internal nodes across the ML tree was reconstructed and represented according to the color scheme shown in the map. Sequences were grouped in seven **(A)** and three **(B)** discrete locations: Amazonas state (AM), Central-Western region (CW), Northern region (N); Northeastern region (NE), Southeastern region (SE), Southern region (S), United States/France (US/FR), other Brazilian states (BR), and non-Brazilian (Non-BR). Tree was rooted using HIV-1 subtype D reference sequences (not shown). The branch lengths are drawn to scale with the bar at the bottom indicating nucleotide substitutions per site. Major migrations of viral clades **(C)** $B_{PAN-AM-I}$ and **(D)** $B_{PAN-AM-II}$ are represented in the maps.

TABLE 2 | Bayesian estimates of evolutionary and demographic parameters of major HIV-1 subtype B clades originated in the Amazonas state.

Clade	N	T_{MRCA}[a] (95% HPD)	R_e[b] (95% HPD)	
$B_{CAR-BR-I}$	267	1978 (1970–1987)	1978–1985	2.6 (0.2–4.6)
			1986–1993	2.5 (1.5–4.2)
			1994–2001	1.9 (1.2–3.0)
			2002–2009	1.9 (1.2–3.0)
			2010–2018	2.3 (1.5–3.5)
$B_{PAN-AM-I}$	86	1988 (1981–1994)	1988–1997	2.0 (0.5–4.1)
			1998–2007	2.9 (1.6–4.8)
			2008–2018	1.6 (1.1–2.6)
$B_{PAN-AM-II}$	60	1995 (1990–2000)	1995–2006	2.9 (1.6–5.4)
			2007–2018	2.2 (1.4–3.6)
$B_{PAN-BR-I-AM-I}$	39	1998 (1992–2002)	1998–2007	3.1 (1.6–5.6)
			2008–2018	2.1 (1.1–3.7)
$B_{PAN-BR-I-AM-II}$	35	1998 (1992–2003)	1998–2007	3.4 (1.7–5.9)
			2008–2018	1.9 (1.1–3.3)

The table details the evolutionary and epidemiological parameters of the major HIV-1 subtype B clusters of pandemic (B_{PAN}) and non-pandemic (B_{CAR}) forms circulating in Brazil's Amazonas state.
[a]*Median value and 95% HPD interval of the time to the most recent common ancestor (T_{MRCA}).*
[b]*Median value and 95% HPD interval of the effective reproductive number (R_e) inferred in a bird-death statistical model.*

the same population. This study revealed that the B_{PAN} and B_{CAR} epidemics in Amazonas have been shaped by different evolutionary histories, but displayed very similar transmissibility and expansion dynamics at most recent times.

Our analysis confirmed that variants B_{CAR} and B_{PAN} were introduced multiple times in the Amazonas state, although the estimated number of B_{PAN} introductions was 16 times higher than that of B_{CAR}. The founder event that originated the clade $B_{CAR-BR-I}$ occurred in the late 1970s (Divino et al., 2016; Arantes et al., 2019) and gave rise to 89% of B_{CAR} and 19% of total subtype B infections in Amazonas. In sharp contrast, most B_{PAN} sequences from Amazonas branched into multiple state-specific clusters of medium/small size (57%) or appeared as unclustered infections (19%). The four largest B_{PAN} Amazonian clades identified probably arose between the late 1980s and late 1990s and, together, comprise 23% of the B_{PAN} and 17% of all subtype B infections in the state. These findings support that the early introduction (late 1970s) of the $B_{CAR-BR-I}$ ancestor in Amazonas from neighboring Caribbean countries probably drove its successful establishment and wide dissemination in the state. Although the B_{PAN} strains arrived in Amazonas later, they reached a high prevalence because they were introduced at much higher numbers and spread through more transmission networks than B_{CAR} strains.

The AIDS detection rate increased ~10% in the Amazonas state between 2009 and 2019 (Ministério da Saúde, 2020). This finding is consistent with our BDSKY analyses that support a continuous expansion ($R_e > 1$) of major B_{CAR} and B_{PAN} Amazonian clades over all the studied period. The BSKG model indicates a recent stabilization of some Amazonian B_{PAN} clades since the late 2000s, and a previous study conducted by our

group also indicated a recent epidemic stabilization of the clade $B_{CAR-BR-I}$ since the late 2000s (Arantes et al., 2019). Although the median estimated R_e of the B_{CAR} and B_{PAN} Amazonian clades was somewhat lower between 2010 and 2018 (1.6–2.3) than during the previous decades (2.5–3.4), we found no solid evidence of epidemic stabilization or reduction in the BDSKY analyses. A previous study pointed out that the BSKG model requires strongly informative data to prevent erroneous estimates of N_e stabilization (Volz and Didelot, 2018). Thus, we hypothesize that the much larger number of recent (2009–2018) $B_{CAR-BR-I}$ sequences used in the present study ($n = 267$) compared to the previous one ($n = 45$) allowed us to obtain a more accurate demographic reconstruction of the epidemic pattern in the last two decades.

It is interesting to note that the B_{CAR} (2.5–2.6) and B_{PAN} (2.9–3.4) Amazonian clades reached similar highest median R_e values. Furthermore, the highest median R_e estimated here using a birth-death approach was comparable to the previous ones estimated for the $B_{CAR-BR-I}$ clade (3.8) using a coalescent-based approach (Arantes et al., 2019), but lower than those estimated for major B_{PAN} Brazilian lineages spreading in the Southeastern region (5.0–7.9) (Mir et al., 2015). These findings support that differences in the spreading dynamics of subtype B lineages may reflect discrepancies in the connectivity of underlying transmission networks across different Brazilian states/regions, rather than intrinsic differences in viral transmissibility. A preliminary analysis of the available demographic data (sex and age) of HIV-infected subjects from Amazonas revealed no significant differences between major B_{CAR} and B_{PAN} clades (**Supplementary Table 2**), supporting that both viral lineages are possibly spreading through networks with similar epidemiological properties. This observation is also consistent with a previous study that revealed comparable epidemic growth rates of B_{CAR} and B_{PAN} lineages circulating in the French Guiana (Bello et al., 2018).

Our study has some limitations. First, inferences about potential sources, total number of viral introductions, and local clade size in Amazonas were limited by both the incomplete sampling of local population and the limited number of non-Amazonian reference sequences included in each ML phylogenetic analysis as revealed by the variable phylogenetic and phylogeographic placement of some Amazonian sequences that branched basal to each local clade. The bulk of B_{CAR} and B_{PAN} sequences that compose each major Amazonian clade, however, remained constant across analyses, and our major phylogeographic conclusions were robust to sampling bias. Second, time-scale reconstructions were largely influenced by the selected clock rate prior due to the weak temporal structure of Amazonian HIV-1 subtype B *pol* datasets. Despite this, the T_{MRCA} here obtained were fully consistent with the overall time-scale of dissemination of the HIV-1 B_{CAR} and B_{PAN} lineages in the Americas and Brazil described in previous studies (Gilbert et al., 2007; Cabello et al., 2014; Mir et al., 2015; Worobey et al., 2016; Arantes et al., 2019; Bello et al., 2019). Finally, the lack of epidemiological data regarding the mode of transmission of individuals analyzed reduced the power of our
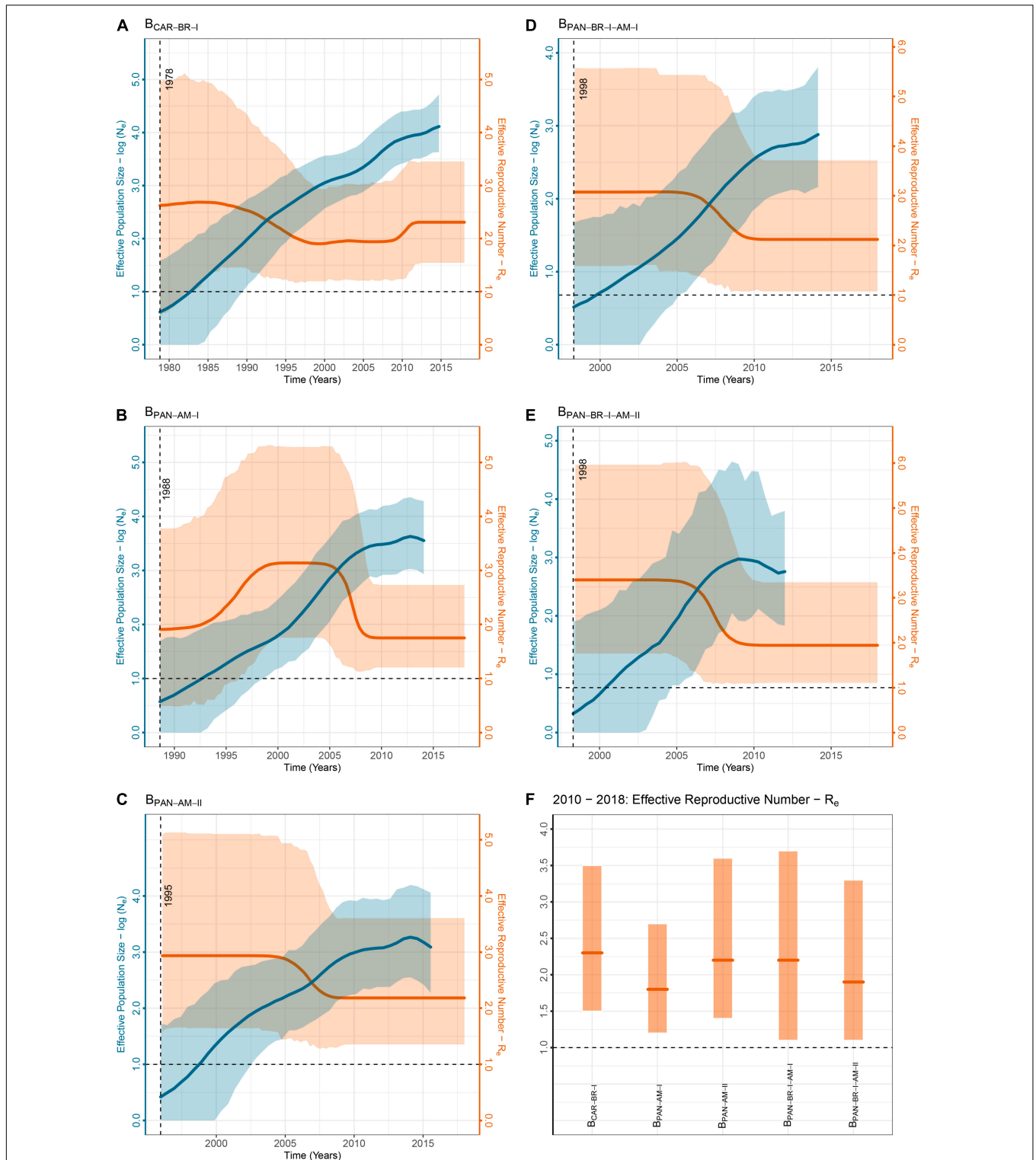
FIGURE 4 | Demographic history of subtype B cluster in Brazil's Amazonas state. Each plot **(A–E)** details the demographic history of one subtype B large cluster ($n \geq 30$) in Amazonas from the pandemic (B_{PAN}) and non-pandemic (B_{CAR}) forms. The graphs exhibit their effective number of HIV-1 infections under the Bayesian Skygrid (BSKG) model in blue (N_e, y-left-axis), and their effective reproductive number under the Birth-death Skyline (BDSKY) model in orange (N_e, y-right-axis). For both parameters are indicated their median (solid lines) and 95% HPD intervals (pale areas) estimates. A dashed vertical line indicates the TMRCA of the clades, accompanied by its median value. The last graph **(F)** compares the R_e obtained for the five clusters in the last period of analysis (2010–2018). For each cluster, its median R_e (solid line) and 95% HPD interval (pale area) inferred values are represented.

study to confirm any association between the inferred rate of viral spread and the ecological characteristics of local transmission networks in Amazonas.

In summary, this study highlights that the HIV-1 epidemic in the Amazonas state mostly results from the local expansion of one B_{CAR} strain ($B_{CAR-BR-I}$) introduced around the late 1970s and of multiple B_{PAN} viral strains introduced since the late 1980s. Albeit the earlier introduction of the $B_{CAR-BR-I}$ clade granted a much prolonged period of local spread than that of the B_{PAN} strains, this was compensated by a much higher number of independent introductions and the concurrent establishment of multiple B_{PAN} local transmission networks. Despite significant differences in the pattern of early establishment, major B_{CAR} and B_{PAN} clades circulating in Amazonas have been spreading at a quite similar rate over the last two decades, arguing against the hypothesis of significant differences in their intrinsic transmissibility. Our analyses also demonstrate that major Amazonian B_{CAR} and B_{PAN} clades continued to spread and showed no clear signs of recent epidemic stabilization, supporting the relevance of designing more effective strategies to prevent HIV transmission in the region.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: GenBank accession numbers: KEXV01000001 to KEXV01046877; HQ127524 to HQ127607; KU762066 to KU762066; MH673055 to MH673280; and MW545333 to MW545424; Los Alamos HIV Sequence Database (http://www.hiv.lanl.gov).

ETHICS STATEMENT

Ethical review and approval were not required for the study on human participants in accordance with the Local Legislation and Institutional Requirements. Written informed consent for participation was not required for this study in accordance with the National Legislation and the Institutional Requirements.

AUTHOR CONTRIBUTIONS

GB conceived and designed the study and supervised the experiments. IA conducted the experiments and analyzed the data. YO, TG, and MG provided HIV-1 sequence data and intellectual input. IA and GB wrote the first draft of the manuscript. All authors assisted with the writing and approved the final manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.835443/full#supplementary-material

REFERENCES

Anisimova, M., and Gascuel, O. (2006). Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Syst. Biol.* 55, 539–552. doi: 10.1080/10635150600755453

Arantes, I., Esashika Crispim, M. A., Nogueira da Guarda Reis, M., Martins Araújo Stefani, M., and Bello, G. (2019). Reconstructing the dissemination dynamics of the major HIV-1 subtype B non-pandemic lineage circulating in brazil. *Viruses* 11:909. doi: 10.3390/v11100909

Arantes, I., Ribeiro-Alves, M., de Azevedo, S. S. D., Delatorre, E., and Bello, G. (2020). Few amino acid signatures distinguish HIV-1 subtype B pandemic and non-pandemic strains. *PLoS One* 15:e0238995. doi: 10.1371/journal.pone.0238995

Bello, G., Arantes, I., Lacoste, V., Ouka, M., Boncy, J., Césaire, R., et al. (2019). Phylogeographic Analyses Reveal the Early Expansion and Frequent Bidirectional Cross-Border Transmissions of Non-pandemic HIV-1 Subtype B Strains in Hispaniola. *Front. Microbiol.* 10:1340. doi: 10.3389/fmicb.2019.01340

Bello, G., Nacher, M., Divino, F., Darcissac, E., Mir, D., and Lacoste, V. (2018). The HIV-1 subtype B epidemic in french guiana and suriname is driven by ongoing transmissions of pandemic and non-pandemic lineages. *Front. Microbiol.* 9:1738. doi: 10.3389/fmicb.2018.01738

Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., et al. (2019). BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* 15:e1006650. doi: 10.1371/journal.pcbi.1006650

Cabello, M., Junqueira, D. M., and Bello, G. (2015). Dissemination of nonpandemic Caribbean HIV-1 subtype B clades in Latin America. *AIDS* 29, 483–492. doi: 10.1097/qad.0000000000000552

Cabello, M., Mendoza, Y., and Bello, G. (2014). Spatiotemporal dynamics of dissemination of non-pandemic HIV-1 subtype B clades in the caribbean region. *PLoS One* 9:e106045. doi: 10.1371/journal.pone.0106045

Chaves, Y. O., Pereira, F. R., de Souza Pinheiro, R., Batista, D. R. L., da Silva Balieiro, A. A., de Lacerda, M. V. G., et al. (2021). High detection rate of HIV drug resistance mutations among patients who fail combined antiretroviral therapy in manaus, brazil. *BioMed. Res. Int.* 2021:5567332. doi: 10.1155/2021/5567332

Chen, J. H. K., Wong, K. H., Chan, K. C. W., To, S. W. C., Chen, Z., and Yam, W. C. (2011). Phylodynamics of HIV-1 subtype B among the men-having-sex-with-men (MSM) population in Hong Kong. *PLoS One* 6:e25286. doi: 10.1371/journal.pone.0025286

Crispim, M. A. E., Reis, M. N. D. G., Abrahim, C., Kiesslich, D., Fraiji, N., Bello, G., et al. (2019). Homogenous HIV-1 subtype B from the Brazilian Amazon with

infrequent diverse BF1 recombinants, subtypes F1 and C among blood donors. *PLoS One* 14:e0221151. doi: 10.1371/journal.pone.0221151

Divino, F., de Lima Guerra Corado, A., Gomes Naveca, F., Stefani, M. M. A., and Bello, G. (2016). High prevalence and onward transmission of non-pandemic HIV-1 subtype B clades in northern and northeastern brazilian regions. *PLoS One* 11:e0162112. doi: 10.1371/journal.pone.0162112

Drummond, A. J. (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* 22, 1185–1192. doi: 10.1093/molbev/msi103

Drummond, A. J., Ho, S. Y. W., Phillips, M. J., and Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4:e88. doi: 10.1371/journal.pbio.0040088

Drummond, A. J., Nicholls, G. K., Rodrigo, A. G., and Solomon, W. (2002). Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics* 161, 1307–1320. doi: 10.1093/genetics/161.3.1307

Gilbert, M. T. P., Rambaut, A., Wlasiuk, G., Spira, T. J., Pitchenik, A. E., and Worobey, M. (2007). The emergence of HIV/AIDS in the Americas and beyond. *Proc. Natl. Acad. Sci.* 104, 18566–18570. doi: 10.1073/pnas.0705329104

Gill, M. S., Lemey, P., Faria, N. R., Rambaut, A., Shapiro, B., and Suchard, M. A. (2012). Improving bayesian population dynamics inference: a coalescent-based model for multiple loci. *Mol. Biol. Evol.* 30, 713–724. doi: 10.1093/molbev/mss265

Gräf, T., Bello, G., Andrade, P., Arantes, I., Pereira, J. M., da Silva, A. B. P., et al. (2021). HIV-1 molecular diversity in Brazil unveiled by 10 years of sampling by the national genotyping network. *Sci. Rep.* 11:15842. doi: 10.1038/s41598-021-94542-5

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of phyml 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010

Guindon, S., Lethiec, F., Duroux, P., and Gascuel, O. (2005). PHYML Online–a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Res.* 33, W557–W559. doi: 10.1093/nar/gki352

Hue, S., Pillay, D., Clewley, J. P., and Pybus, O. G. (2005). Genetic analysis reveals the complex structure of HIV-1 transmission within defined risk groups. *Proc. Natl. Acad. Sci.* 102, 4425–4429. doi: 10.1073/pnas.0407534102

Ishikawa, S. A., Zhukova, A., Iwasaki, W., and Gascuel, O. (2019). A Fast Likelihood Method to Reconstruct and Visualize Ancestral Scenarios. *Mol. Biol. Evol.* 36, 2069–2085. doi: 10.1093/molbev/msz131

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948. doi: 10.1093/bioinformatics/btm404

Mendoza, Y., Martínez, A. A., Castillo Mewa, J., González, C., García-Morales, C., Avila-Ríos, S., et al. (2014). Human immunodeficiency virus type 1 (HIV-1) subtype B epidemic in Panama is mainly driven by dissemination of country-specific clades. *PLoS One* 9:e95360. doi: 10.1371/journal.pone.0095360

Ministério da Saúde (2020). *Boletim Epidemiológico HIV/AIDS 2020*. Available online at: http://www.aids.gov.br/system/tdf/pub/2016/67456/boletim_hiv_aids_2020_com_marcas_2.pdf?file=1&type=node&id=67456&force=1 (accessed December 14, 2021).

Mir, D., Cabello, M., Romero, H., and Bello, G. (2015). Phylodynamics of major HIV-1 subtype B pandemic clades circulating in Latin America. *AIDS* 29, 1863–1869. doi: 10.1097/qad.0000000000000770

Posada, D. (2008). jModelTest: phylogenetic model averaging. *Mol. Biol. Evol.* 25, 1253–1256. doi: 10.1093/molbev/msn083

Pupko, T., Pe, I., Shamir, R., and Graur, D. (2000). A fast algorithm for joint reconstruction of ancestral amino acid sequences. *Mol. Biol. Evol.* 17, 890–896. doi: 10.1093/oxfordjournals.molbev.a026369

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: https://www.R-project.org

Rambaut, A. (2009). *FigTree v1.4: Tree Figure Drawing Tool*. Available online at: http://tree.bio.ed.ac.uk/software/figtree/ (accessed on 25 Nov, 2018)

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., and Suchard, M. A. (2018). Posterior summarization in bayesian phylogenetics using tracer 1.7. *Syst. Biol.* 67, 901–904. doi: 10.1093/sysbio/syy032

Stadler, T., Kuhnert, D., Bonhoeffer, S., and Drummond, A. J. (2012). Birth-death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV). *Proc. Natl. Acad. Sci.* 110, 228–233. doi: 10.1073/pnas.1207965110

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033

Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J., and Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Vir. Evol.* 4:vey016. doi: 10.1093/ve/vey016

Suchard, M. A., and Rambaut, A. (2009). Many-core algorithms for statistical phylogenetics. *Bioinformatics* 25, 1370–1376. doi: 10.1093/bioinformatics/btp244

Volz, E. M., and Didelot, X. (2018). Modeling the growth and decline of pathogen effective population size provides insight into epidemic dynamics and drivers of antimicrobial resistance. *Syst. Biol.* 67, 719–728. doi: 10.1093/sysbio/syy007

Worobey, M., Watts, T. D., McKay, R. A., Suchard, M. A., Granade, T., Teuwen, D. E., et al. (2016). 1970s and 'Patient 0' HIV-1 genomes illuminate early HIV/AIDS history in North America. *Nature* 539, 98–101. doi: 10.1038/nature19827

Zehender, G., Ebranati, E., Lai, A., Santoro, M. M., Alteri, C., Giuliani, M., et al. (2010). Population dynamics of HIV-1 subtype B in a cohort of men-having-sex-with-men in Rome, Italy. *JAIDS* 55, 156–160. doi: 10.1097/qai.0b013e3181eb3002