

RESEARCH

Open Access



The origin and dissemination of HIV-1 subtype B on Hainan Island, South China, 2007–2024

Chunyun Lu^{1†}, Kaokao Zhu^{2†}, Nengde Zhang^{1†}, Mu Li³, Shanmei Zhong³, Cai Qin³, Tong Luo³, Rongjing Zhang³, Jing Zhang¹, Yi Ning¹, Hao Liang⁴, Li Ye^{4*}, Bingyu Liang^{3*} and Dee Yu^{1*}

Abstract

Background HIV-1 subtype B, introduced to mainland China at an early stage, rapidly spread among commercial plasma donors and heterosexuals (HETs) during the 1990s. It was first identified in Hainan, which has a population of 10 million, in 2007. Nevertheless, there is a paucity of data about the molecular-epidemiological characteristics, spatial origins, and transmission patterns of this strain on the island of China.

Methods *Pol* sequences for HIV-1 subtype B from Hainan Island (2007–2024) were obtained from our research and global databases. An analysis of phylogenetic data, molecular clock models, and a Bayesian coalescent-based approach was performed to investigate how HIV-1 subtype B was introduced and disseminated on this island.

Results A total of 76 sequences were obtained from Hainan Island, of which 52.63% were collected from men who have sex with men (MSM). Phylogenetic analysis demonstrated that HIV-1 subtype B circulating on the island was closely related to HIV-1 sequence lineages in Guangdong (22.84%, 127/556), Hubei (12.59%, 70/556), and Beijing (14.03%, 78/556), indicating multiple introductions from various cities in China. Bayesian time-scaled phylogenetic reconstructions revealed that 27 Hainan sequences were together in clusters [posterior probability (*PP*) > 0.90], and 18 Hainan sequences were intermixed in larger clades containing sequences from other provinces (*PP* > 0.90). Bayes factors (BF) demonstrated that subtype B on the Island originated from Guangdong Province (BF > 100) in 1999.9 (95% HPD:1996.7, 2002.7) and Hubei Province (BF > 100) in 2007.9 (95% HPD:2004.4, 2011.2). Subsequently, HIV-1 subtype B was introduced to Beijing from the island (BF > 100). The analysis of viral migration patterns among risk subgroups revealed that subtype B was introduced to the island by MSMs (BF > 100). Within the island, this strain was from MSM to HETs (BF > 100).

Conclusion Using Hainan Island as a case study, our findings shed light on the intricate evolution and transmission dynamics of HIV-1 subtype B on the island, revealing crucial insights for HIV prevention efforts in the future.

[†]Chunyun Lu, Kaokao Zhu and Nengde Zhang contributed equally to this work.

*Correspondence:

Li Ye

yeli@gxmu.edu.cn

Bingyu Liang

liangbingyu@gxmu.edu.cn

Dee Yu

yudee@muh.n.edu.cn

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Keywords HIV-1, Subtype B, Origin, Phylogenetic analysis, Hainan Island

Introduction

The phylogenetic diversity of HIV-1, encompassing groups M, N, O, and P, underscores the complexity and interspecies transmission dynamics of certain HIV-1 strains from SIV to humans [1–3]. The genetic complexity of the HIV-1 pandemic, driven by group M, is highlighted by the presence of numerous subtypes and the dynamic nature of circulating recombinant forms (CRFs) and unique recombinant forms (URFs) [4, 5]. The recombination event that led to CRFs is primarily responsible for the Southeast Asian outbreak [6, 7]. According to phylogenetic analyses, HIV-1 subtype B originated in Kinshasa, Africa, and likely spread to the Caribbean, specifically Haiti, during 1966. Subsequent localized dissemination facilitated the virus's arrival in the United States, primarily via homosexual and bisexual contact, approximately in 1969 [8, 9].

Meanwhile, subtype B strains were identified as constituting a significant proportion of HIV infections in Southeast Asia, particularly in Indonesia, Malaysia, Myanmar, and East India. This prevalence was plausibly attributable to the widespread incidence of injecting drug use observed in these regions [10–12]. The B' variant, also referred to as B' Thailand or Thai-B, was initially identified in 1988 among people who inject drugs (PWIDs) in the Golden Triangle, a region that encompasses the borders of Laos, Thailand, Myanmar, and China [8].

In China, the distribution of HIV-1 subtypes is characterized by notable diversity and complexity. The most recent nationwide molecular epidemiological survey identified eighteen HIV-1 genetic variants in China. Among these, CRF01_AE, CRF07_BC, CRF08_BC, and subtype B collectively accounted for 92.8% of HIV-1 infections [13]. Hainan, China's southernmost gateway province, exhibits unique epidemiological features of HIV due to its ethnic diversity and the influx of seasonal migrants. In 2007, only five subtypes, with subtype B ranking second, were detected in Hainan Province. Blood transmission and IDU were identified as the primary routes, with no evidence of homosexual transmission observed at that time [14]. However, a total of 4209 individuals had contracted HIV in Hainan Province by the end of 2021. And sexual contact (82.5%) was the primary risk factor of HIV-1, with 42.9% of newly diagnosed cases attributed to homosexual contact and 39.5% to heterosexual contact. To 2020, a total of 14 subtypes have been identified, including 12 CRFs [15].

Although subtype B strains not being the predominant strain in Hainan Province, they still play a significant role for two reasons: firstly, their presence since their discovery in 2007 partially reflects the evolution of the

HIV epidemic; secondly, during our continuous surveillance, it was discovered that the cases of subtype B have been increasing annually, which is in contrast to that in the cities of mainland [15]. Nevertheless, there is no clear understanding of the genetic evolution, the geographic origin, and the transmission dynamics of HIV-1 subtype B in Hainan.

The densely populated Hainan Island is home to a diverse population of both domestic and foreign residents. Conducting research on the transmission dynamics of HIV-1 subtype B on Hainan Island can offer valuable observations into the evolving patterns of HIV-1 subtype B transmission in China and potentially even Southeast Asia. This research has the potential to contribute significantly to our understanding of the spread and impact of HIV-1 subtype B in this region. In our study, we conducted molecular epidemiological, phylogenetic, and migration analyses utilizing polymerase (*pol*) sequences from HIV-1 infected subjects to deduce the geographical and temporal evolutionary trajectory of subtype B in Hainan.

Methods

Study design and participants

A cross-sectional study aimed at monitoring the molecular epidemiology and drug resistance of HIV was conducted at the Fifth People's Hospital of Hainan Province, the largest clinical treatment center for HIV/AIDS in the region, from 2014 to 2024. The eligibility criteria for inclusion in this study were as follows: (1) individuals aged over 18 years; (2) individuals with a confirmed diagnosis of HIV-1 infection, verified through both enzyme-linked immunosorbent assay (ELISA) and Western blot testing; (3) the participants voluntarily participated in this study and provided informed consent. At the same time, we accessed the system of Hainan Skin and Venereal Disease Prevention Center to inquire about the information we needed to collect, including age, ethnicity, education, time and location of diagnosis, and route of infection. The information data was matched with the blood sample information according to the patient coding. Finally, the study recruited a total of 1742 HIV-positive individuals.

Specimen preparation and laboratory testing

In this study, we obtained 10 mL of peripheral blood from each participant, which was collected into EDTA tubes and subsequently centrifuged at 3000 rpm for 15 min to isolate the plasma. The plasma samples were then stored at -80°C. Viral RNA extraction and HIV-1 *pol* gene amplification were performed at the Guangxi

Key Laboratory of AIDS Prevention and Treatment. Subsequently, the 1300 base-pair sequence of the HIV-1 *pol* gene, which includes the regions encoding the HIV protease and reverse transcriptase (RT, amino acids 1–335), was sequenced by Sangon Biotech. The detailed methodologies have been previously described in our earlier research [16].

HIV subtyping analysis

In this study, the *pol* gene sequences were assembled utilizing Sequencher v5.4.6 (<http://www.genecodes.com>) and aligned with an HXB2 sequence using the MAFFT algorithm, implemented in the HIV sequence database (<http://www.hiv.lanl.gov/content/s/VIRALIGN/viralgn.html>). The *pol* segments located between positions 2253–3821 of HXB2 were extracted using BioEdit v7.2.5 (<http://thalljscience.github.io/>). Subsequently, the quality control of the sequences was conducted utilizing the online tool available on the Los Alamos National Laboratory (LANL) platform, then the sequences with mixed bases $\geq 5\%$ or length < 1000 bp were removed.

The primary determination of HIV-1 subtypes was conducted utilizing two online tools: COMET (<https://comet.lnh.lu/>) [17] and REGA v3.46 (<https://www.genome.detective.com/app/typingtool/hiv>). This initial classification was subsequently validated through the construction of a maximum likelihood (ML) phylogenetic tree, incorporating reference sequences (subtypes A–K and CRFs) obtained from LANL database (<http://www.hiv.lanl.gov/>) (Figure S1, Table S1). The ML tree was assembled utilizing the GTR + F + R10 model, within the framework of IQ tree version 1.6.12. The optimal model was determined in accordance with the Akaike Information Criterion (AIC) [18]. Clusters with a bootstrap value exceeding 0.9 with corresponding reference sequences were classified as the same subtype. Lastly, the ML tree was visualized using Figtree version 1.4.4 [19].

Sequence datasets

For phylogenetic analysis, all obtainable HIV-1 subtype B sequence data were retrieved from LANL database. One sequence was downloaded for each patient. For all 76 Hainan sequences, 9 were retrieved from web databases and 67 from local databases. With 76 Hainan sequences as index, Blast was performed at a ratio of 1:50 to retrieve the sequences most similar from the HIV sequence database for further analysis. Quality control protocols were instituted to assess the purity of subtypes and to detect any potential sequence gaps. Redundant sequences were eliminated, culminating in a refined dataset (Dataset 1) comprising 556 sequences, of which 76 originated from Hainan (Table S2). The sequences in Dataset 1 adhere to the criteria delineated in our prior research [20].

Dataset-1 was employed to deduce a preliminary maximum likelihood (ML) phylogeny using IQ-TREE, with the aim of examining the evolutionary connections between Hainan-derived sequences and a range of sequences gathered worldwide. As the Hainan strains did not group into a single monophyletic clade (Fig. 1), this indicates that there were multiple transmission events from external provinces occurring at various time points. Subsequently, Database 2 was then constructed for further analysis by adjusting the ratio of local BLAST to 1:30 while removing sequences from the same region but with < 10 sequences. The regression analysis was performed on the ML tree using TempEst 1.5.3 (<http://tree.bio.ed.ac.uk/software/tempest/>). The outliers with poor temporal signals were subsequently removed from the final dataset. A total of 328 sequences, with 75 of them originating from Hainan, were compiled in Database 2 to conduct an in-depth analysis of the phylodynamic patterns of HIV-1 subtype B using the Bayesian coalescent approach available in BEAST version 1.10.5 [21].

Bayesian phylogeny and molecular clock calibration

The Bayesian phylogeny for Dataset-2 was estimated using the Bayesian Markov chain Monte Carlo (MCMC) method implemented in BEAST version 1.10.5. The analysis employed the GTR + invariant + Gamma model, which was determined to be the best fit according to the hierarchical likelihood ratio test conducted by jModelTest 2 [22]. Bayes factors (BF) were utilized to compare two demographic priors and a constant-size model under a relaxed molecular clock. The MCMC recombination process was executed over 10 billion generations, with sampling occurring every 10 million generations. To ensure sufficient recombination, the effective sample size (ESS) was calculated using Tracer version 1.5 [23]. An Effective Sample Size (ESS) exceeding 200 for the estimated parameters was considered acceptable following a 10% burn-in period.

The maximum clade credibility (MCC) tree was obtained utilizing Tree Annotator version 1.8.3 within the BEAST software package. This process incorporated a 10% burn-in and utilized median node heights. The posterior probability (PP) as clade support was calculated [24]. The 95% highest posterior density (95% HPD) interval was used to assess the uncertainty in the estimations. The MCC tree was subsequently refined and visualized using FigTree version 1.4.4. Additionally, the Bayesian Stochastic Search Variable Selection (BSSVS) method and the discrete Bayesian phylogeographic approach were employed for further analysis. The sequences were assigned a geographic location based on the sampling city, and viral migration routes were analyzed and summarized using Spread3 version 0.9.7.1 [25]. The BSSVS method was utilized to determine the Bayes Factor

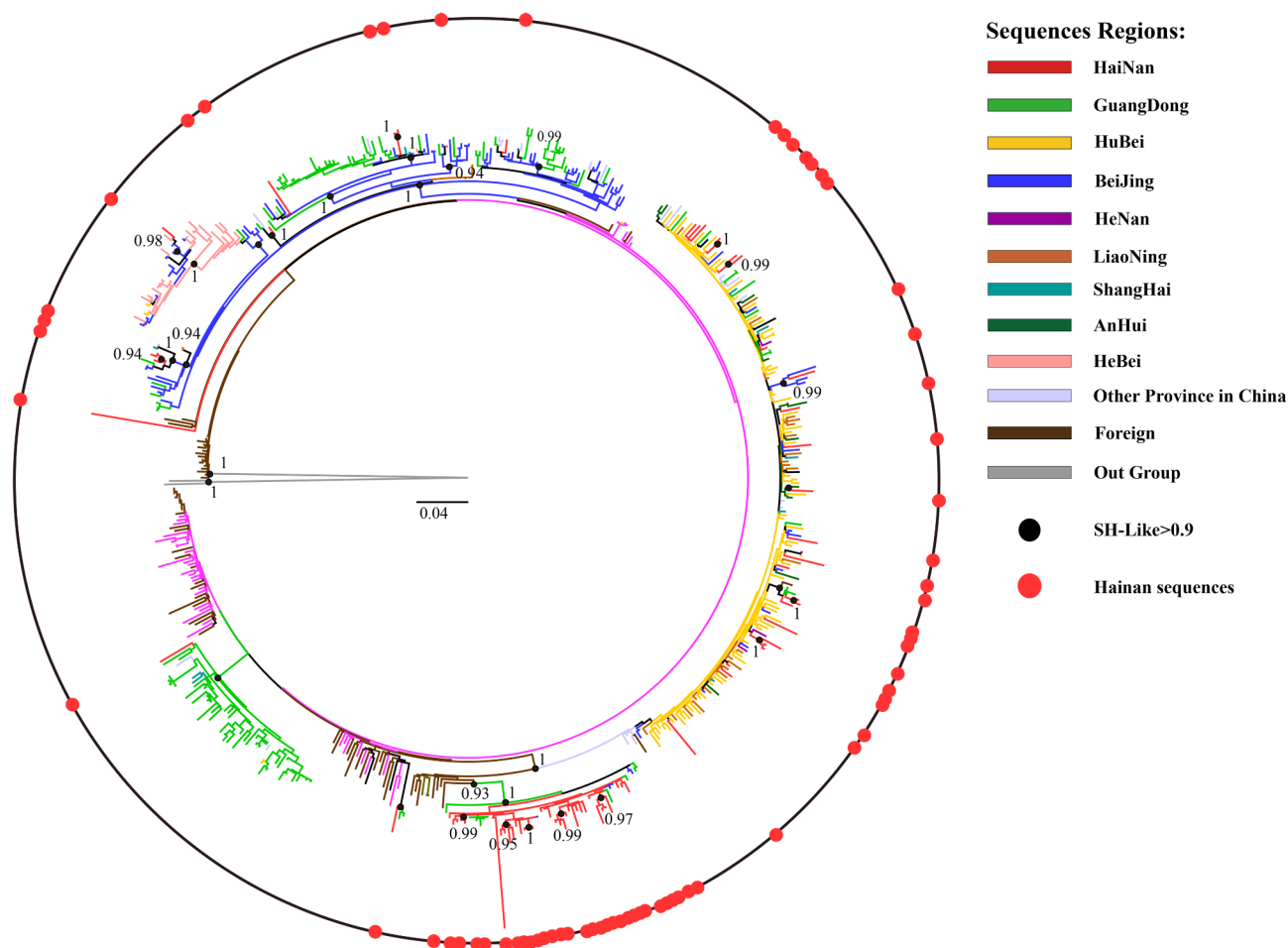


Fig. 1 Global maximum likelihood phylogenetic tree of HIV-1 subtypes B on Hainan Island. The tree reconstruction was inferred using Dataset-1 containing 76 Hainan sequences and 556 selected HIV-1 subtype B sequences from the global search. The red dots on the circle indicate the position of the Hainan sequences on the phylogenetic tree. Branch support values ≥ 90 are indicated by black dots. Tree branches are color-coded according to the cities of isolation.

(BF), providing an accurate depiction of the viral spread dynamics. To elaborate, a BF value below 3 suggests insufficient evidence, a range of 3–6 implies modest evidence, 6–10 represents substantial evidence, and a BF above 10 denotes exceedingly strong evidence [26].

Results

The demographic characteristics of participants

In this study, a total of 76 individuals infected with HIV-1 subtype B between 2007 and 2024 on Hainan Island were analyzed. Of the 76 participants, the mean age was 35.13 years (range: 17–75 years), with males accounting for 90.79%. The majority (52.63%) of infections were attributed to MSM, followed by heterosexual contacts (28.95%). Approximately two-thirds (63.16%) of patients were residents of Haikou city, whereas 31.58% were from Sanya city, and only a small proportion (5.28%) came from other cities in Hainan province. The majority (86.85%) of diagnoses occurred after 2012. Additionally, 25.00% of

the respondents had completed high school or technical secondary school education, and 30.26% had obtained a college degree or higher. (Table 1).

Multiple introductions of HIV-1 subtype B on Hainan Island

To estimate the potential origin of HIV-1 subtype B on Hainan island, an ML tree was constructed using Dataset 1 to identify the sequences from global reference strains that exhibit close genetic similarity to the Hainan strains (Fig. 1). Sequences from Hainan were dispersed across the ML tree, eliminating the hypothesis of a single founder event followed by regional dissemination, but supporting multiple introductions from different countries. The findings revealed that Hainan sequences have close genetic relationships with sequences from various regions in China, including Guangdong sequences (22.84%, 127/556), Beijing (14.03%, 78/556), Hubei (12.59%, 70/556), and other provinces sequences in database 1 (55.58%, 309/556), suggesting multiple

Table 1 Respondents' sociodemographic characteristics

Characteristics	The number of patients (n)	Percentage (%)
Total	76	100
Sex		
Male	69	90.79
Female	7	9.21
Diagnosis age (years)		
< 20	4	5.26
20–29	10	13.16
30–39	27	35.53
40–49	10	13.16
> 49	15	19.74
NA	14	18.42
Mean age(years)	35.13	
Likely location of infection		
Haikou	48	63.16
Sanya	24	31.58
Wanning	1	1.32
Chenmai	1	1.32
Tunchang	1	1.32
Wuzhishan	1	1.32
Risk factor		
HET	22	28.95
MSM	40	52.63
MTC	1	1.32
IDU	1	1.32
NA	12	15.79
Diagnosis year		
2007–2010	13	17.11
2011–2014	10	13.16
2015–2018	14	18.42
2019–2021	18	23.68
2021–2024	21	27.63
Education		
Illiteracy	1	1.32
Primary school	2	2.63
Junior high school	14	18.42
High school or technical secondary school	19	25
University or above	23	30.26

Abbreviation: HET, heterosexuality; MSM, men who have sex with men; MTC, mother to child; NA, not Record

introductions and subsequent transmission to a limited number of individuals. The ML tree demonstrated the presence of 13 clusters containing Hainan sequences with strong statistical support (bootstrap value > 90%). More than half of these clusters consisted of mixtures of Hainan and non-local sequences.

The time-scale of HIV-1 subtype B introduction on Hainan Island

Bayesian time-scaled phylogenies were inferred from Dataset 2, and the MCC tree was constructed from the posterior distribution (Fig. 2). As expected, the sequences from Hainan were scattered across the phylogenetic tree, suggesting multiple incursions from various geographical areas within China. The MCC tree traced its roots back to the United States, with the tMRCA dating back to 1973.7 (95% HPD: 1968.96, 1979.70). Moreover, the MCC tree revealed a robust connection among 44 Hainan sequences ($PP > 0.9$). Among these, 26 sequences exclusively showed linkage to other Hainan sequences, while the remaining 18 sequences exhibited connections with external sequences. These interconnections formed a network of 17 clusters, each containing Hainan sequences. Notably, ten clusters solely comprised local Hainan sequences.

There were seven clusters comprising sequences from Hainan and other provinces. Specifically, two clusters originated from Beijing and spread to Hainan, three clusters moved from Guangdong to Hainan, one cluster came from Hubei to Hainan, and one cluster traveled from Hainan to Beijing. The earliest transmission event of tMRCA from Beijing to Hainan is 2011.97 (95% HPD: 2008.44, 2014.27), it was transmitted to the HET population in Hainan by an unknown transmission route in Beijing. While the earliest tMRCA transmission from Hainan to Beijing is 2013.83 (95% HPD: 2012.37, 2015.30), the transmission was from MSM in Hainan to MSM in Beijing. The tMRCA of transmission from Hubei to Hainan is 2007.87 (95% HPD: 2004.37, 2011.17), the transmission was from the MSM population in Hubei to the MSM population in Hainan. The earliest introduction to Hainan was traced back to Guangdong tMRCA in 1999.84 (95% HPD: 1996.71, 2002.66), the transmission route was unknown in Guangdong province to HETs in Hainan province.

Overall, from Bayesian analysis, it was determined that multiple introductions of HIV-1 subtype B into Hainan originated in Guangdong, Hubei (BF > 100), Beijing (BF = 20), and Liaoning provinces (BF = 47). Most migrant events originated from Hubei ($n = 15.28$), followed by Guangdong ($n = 10.98$). The Bayesian geographic analysis also supported the importation of Hainan subtype B strains into Beijing (BF > 100, $n = 2.72$) (Fig. 3A). The transmission route within Hainan Island extended from Haikou to multiple cities including Wanning (BF = 7.12), Chengmai (BF = 7.04), Tunchang (BF = 6.10), and Wenchang (BF = 22.73), Wuzhishan (BF = 21.53), as well as Sanya (BF > 100). Additionally, the transmission also occurred from Sanya to Haikou (BF > 100) (Fig. 3B). The MCC tree on Hainan Island originated in 1981.81 (95% HPD: 1962.99, 2006.44). The MCC tree was

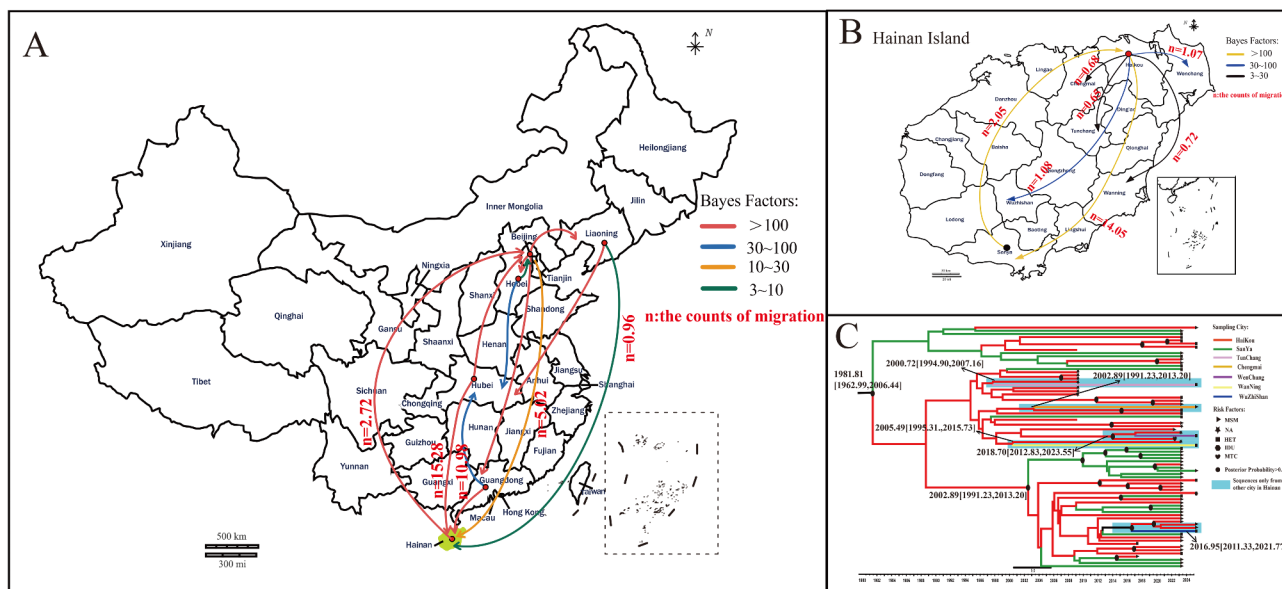


Fig. 2 Bayesian molecular clock analysis of HIV-1 subtype B on Hainan Island. Bayesian maximum clade credibility (MCC) tree and the time of the most recent common ancestor (tMRCA) estimation. The MCC tree was reconstructed using Dataset-2 consisting of 328 sequences, 75 of which were from Hainan, and 253 were selected from further analysis by adjusting the ratio of local BLAST to 1:30. Hainan and non-local mixed clusters were marked with a purple background, and Hainan local clusters were marked with a yellow background. Black dots indicate the posterior probability of the node greater than 0.9. The colors of the branches represent the sampled cities. Scale years are reported at the bottom of the figure.

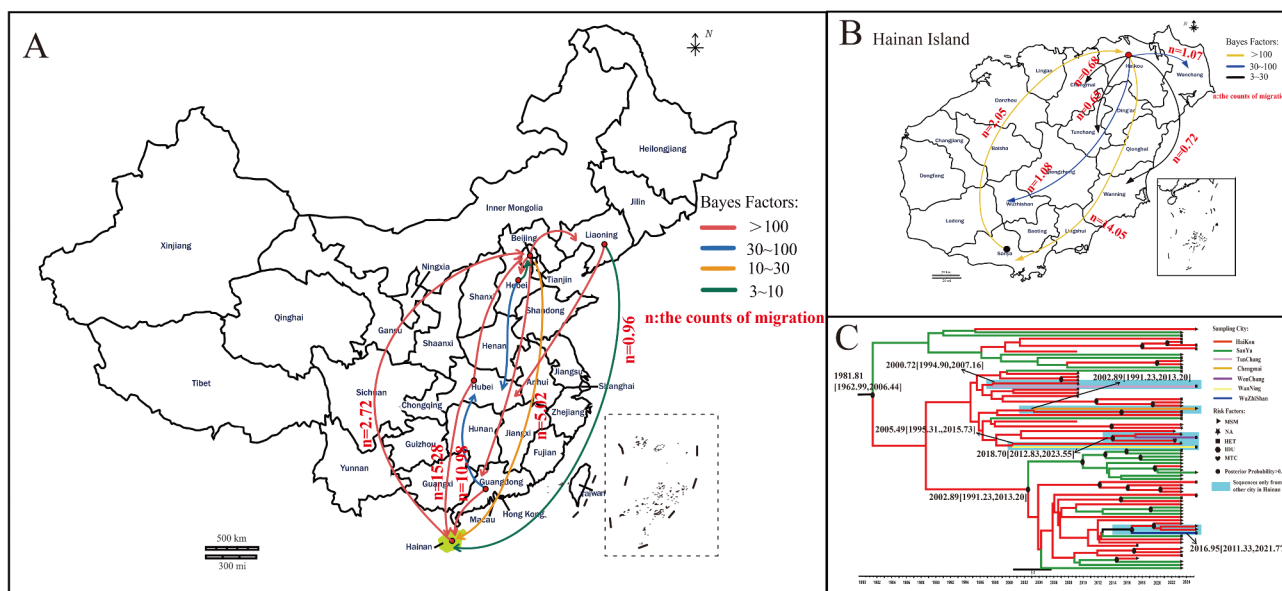


Fig. 3 (A):The introduction and dissemination of HIV-1 subtype B in China.The arrow direction represents the propagation direction. Line colors represent the range of Bayesian factors (BF): red represents BF >100, blue represents from 30 to 100, yellow represents from 10 to 30, and green represents 3-10.n, the counts of migrations.(B):The introduction and dissemination of HIV-1 subtype B within Hainan Island. The arrow direction represents the propagation direction. Line colors represent the range of Bayesian factors (BF): yellow represents BF >100, blue represents from 30 to 100, and black represents 3-10.n, the counts of migrations. (C):Bayesian molecular clock analysis of HIV-1 subtype B within Hainan Island.Bayesian maximum clade credibility (MCC) tree and the time of the most recent common ancestor (tMRCa) estimation.The MCC tree was reconstructed using 76 hainan local sequences.Black dots indicate the posterior probability of the node greater than 0.9. The colors of the branches represent the sampled cities. Scale years are reported at the bottom of the figure.

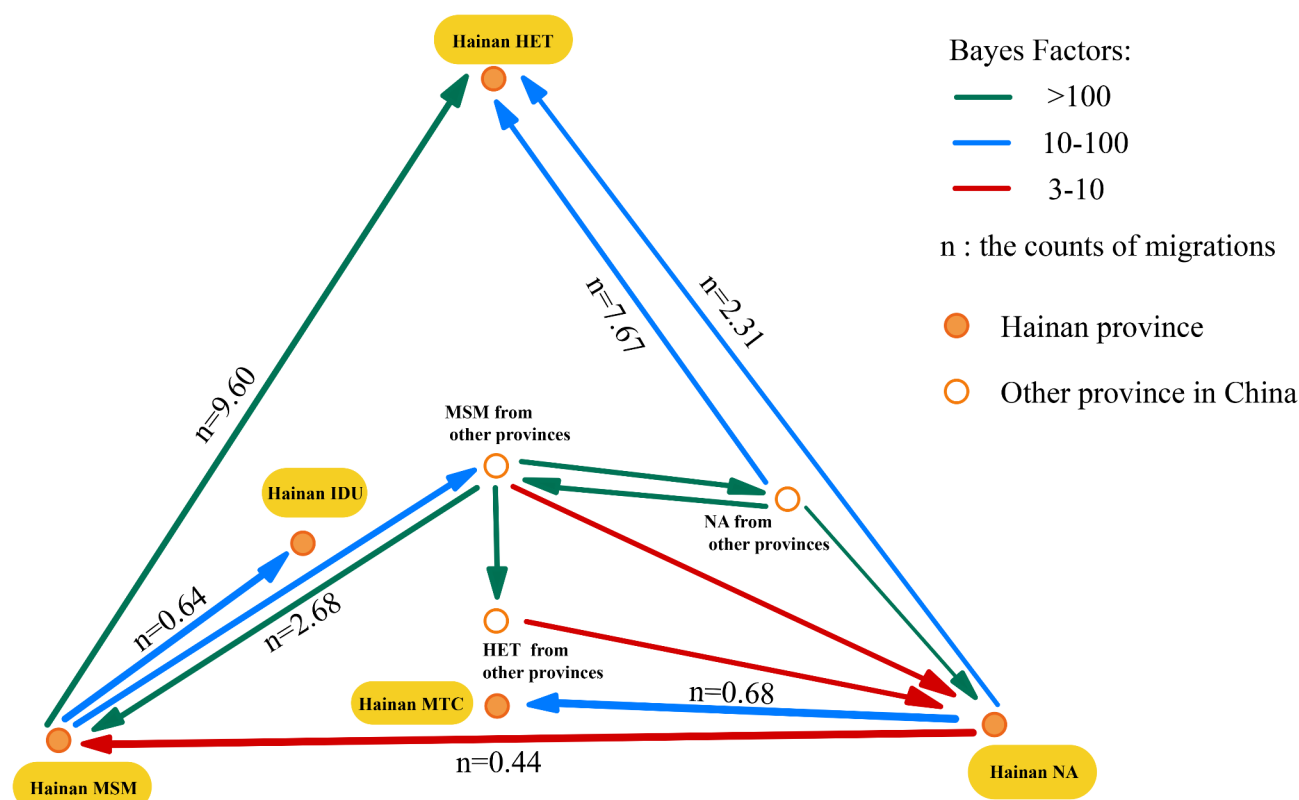


Fig. 4 Dynamic evolution of HIV-1 subtype B transmission mode on Hainan Island. The arrow direction represents the propagation direction. Line colors represent the range of Bayesian factors (BF): green represents $BF > 100$, blue represents from 10 to 100, and red represents from 3 to 10. n , the counts of migrations.

transmitted from an infected person with an unknown transmission route to a heterosexual infected person in Tunchang in Haikou city in 2000.72. The transmission time from Haikou to Chengmai was 2002.89, from heterosexual infected persons in Haikou to MSM population in Chengmai, and from Haikou to Wenchang was 2018.70, from infected persons with unknown transmission route in Haikou to heterosexual infected persons in Wenchang, and from Haikou to Wanning was 2005.49. The transmission route was from an infected person in Haikou to a heterosexual infected person in Wanning, and the transmission time was from Haikou to Wuzhishan on 2016.95, and the transmission was from MSM in Haikou to a heterosexual infected person in Wuzhishan (Fig. 3C).

The transmission pattern dynamic of risk factors for HIV-1 subtype B on Hainan Island

Bayesian phylogenetic analysis unveiled the complexity and inter-population transmission of sexually transmitted infections on Hainan Island. The HIV infections among MSM in Hainan originated from MSM from other provinces ($BF > 100$, $n = 2.68$) and individuals with unknown transmission routes within the province ($BF = 3.64$, $n = 0.44$). Moreover, the HIV strains among

heterosexual individuals in Hainan were introduced by three distinct groups: MSM within the province ($BF > 100$, $n = 9.60$), unknown transmission routes within Hainan province ($BF = 23.78$, $n = 2.309$), and individuals with unknown transmission routes outside the province ($BF = 60.57$, $n = 7.67$). In Hainan, cases of mother-to-child transmission may be due to unknown transmission routes within Hainan province ($BF = 32.68$, $n = 0.68$). Hainan IDU originated from MSM within the province ($BF = 17.01$, $n = 0.64$). (Fig. 4).

Discussion

Phylogenies reconstructed using Bayesian time-scales revealed the transmission pattern of subtype B in Hainan province was complex, involving multiple imported cases from various regions rather than a single source. The initial imported cases originated from Guangdong Province around 2000, followed by repeated introductions from Guangdong and Hubei to Hainan Island, and subsequently from Hainan Island to Beijing. The transmission routes exhibited diversity, initially including blood transfusion and heterosexual transmission from other provinces to Hainan island, which later shifted towards homosexual transmission.

The Chinese government provides free antiretroviral therapy and implements measures, including prevention of mother-to-child transmission and needle exchange programs, to effectively control the spread of HIV [27–29]. Although subtype B is not the predominant strain, its transmission pattern has shifted from blood and heterosexual contact to homosexual contact within the MSM population, characterized by multiple sexual partners and high mobility. Therefore, continuous HIV molecular surveillance is essential for HIV prevention and control.

The Bayesian geographic evolutionary analysis revealed the multiple sources of HIV transmission on Hainan Island, indicating frequent population movements across various regions over a prolonged duration. The HIV-1 subtype B strains were introduced to Hainan Province from the provinces of Hubei and Guangdong. Conversely, Beijing was introduced from Hainan province. In 1985, subtype B was introduced into Yunnan through intravenous drug users in Southeast Asia. Consequently, it spread from Yunnan to central China (including Henan and Hubei) during the 1990s via former plasma donors [30]. The rapid dissemination of subtype B within central China resulted in its expansion towards Beijing, Shandong, Gansu, and 21 other neighboring provinces through both blood and sexual transmissions [31]. As a result of stringent measures implemented against illicit commercial blood sampling activities along with intensified efforts aimed at combating drug-related practices nationwide, sexual transmission has gradually emerged as the predominant mode for HIV-1 B subtype propagation [30, 31]. This study, for the first time, identified the connection between subtype B in Hainan and Hubei, with Hubei serving as a transportation hub that facilitates HIV transmission to other locations. Additionally, Guangdong's developed economy attracts additional migrant workers and its geographical proximity to Hainan creates favorable conditions for the transmission of HIV subtype B. Previous studies have shown that subtype B accounted for 5% of treatment failure in HIV-1 infected patients in Guangdong Province [32], while its prevalence was 5.13% among newly diagnosed HIV/AIDS cases in Shenzhen from 2011 to 2018 due to elevated population mobility [33]. The combination of high prevalence and geographical proximity facilitated the spread of subtype B strain from Guangdong to Hainan, subsequently disseminating to Beijing, underscoring its widespread distribution across multiple provinces and emphasizing the need for vigilant surveillance. The prolonged and frequent transmission of the virus in these two provinces was identified, which is closely associated with the convenient transportation between Hainan and Guangdong, including the railway system and maritime routes. Additionally, the close cultural and folk exchange activities serve as driving factors that may facilitate the mutual transmission

of cases between the two provinces. The connection and communication between the two regions will likely accelerate the further spread of HIV-1 infection. These findings underscore the necessity to establish molecular surveillance systems and share HIV-1 infection information in Hainan and Guangdong provinces.

Compared with other cities in the Chinese mainland, the import of subtype B in Hainan mainly depends on the MSM and HET populations. In contrast to the traditional blood transmission mode in cities of the Chinese mainland, subtype B in Hainan is imported through sexual transmission. This might be associated with the natural transportation barrier of Hainan Island, indicating the lag of virus import on the island. Simultaneously, it reflects that compared with traditional blood transmission, sexual transmission has a stronger transmission capacity. This presents distinct challenges for the prevention and management of subtype B HIV in Hainan, as compared to the control of blood-borne transmission. Additionally, this study reveals that men who have sex with men (MSM) demonstrate higher levels of mobility and transmission potential relative to other transmission groups [34]. MSM exhibited greater mobility and transmissibility compared to other transmission groups [34], as also evidenced by this study. Hainan MSM not only facilitates inter-provincial dissemination of HIV-1 subtype B within their community but also propagates it similarly to other high-risk groups such as heterosexual sexual networks. This finding suggested the necessity of revising the prevention and control strategy for subtype B, highlighting the significant role of MSM in HIV transmission in addition to conventional methods aimed at preventing bloodborne transmission. Moreover, the high rate of population mobility and clustering of ethnic minority groups in Hainan Province poses a potential risk of dual infection for the heterosexual population, both locally and beyond. Additionally, patriarchal beliefs may lead MSM to engage in covert marriages [35, 36], and research also revealed the presence of minority female sex workers in this area, with a low level of awareness regarding AIDS knowledge [37]. Consequently, it is imperative to implement health education initiatives and bolster the self-protection skills of heterosexual communities in Hainan Province.

In summary, this study has revealed the diversity of transmission sources of HIV-1 subtype B on Hainan Island. The virus has been imported from multiple provinces, mainly through sexual behavior from Guangdong and Hubei. Additionally, it reflects a shift in the transmission pattern of HIV-1 subtype B on Hainan Island, from heterosexual contact and blood transmission to homosexual contact. This study provided insights into the current transmission situation of HIV-1 subtype B on Hainan Island and offers an important reference for

the formulation of AIDS surveillance and prevention strategies.

The limitations of this study primarily revolve around the initially limited sample size and the absence of full-length sequences. Specifically, only *pol* region sequences were available, and the database-provided sequences exhibited incomplete coverage. This may have introduced potential biases in our results, which should be addressed in future research. However, it is important to emphasize that these limitations do not diminish the significance of this study, as it represents the first attempt to investigate the origin of subtype B on Hainan Island using a phylogenetic approach.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12879-025-10703-1>.

Supplementary Material 1

Acknowledgements

Not applicable.

Author contributions

DY, BL, NY and KZ designed the study; LC and ZN participated in the process of sequence editing and phylogenetic analyses; DY and LC collected samples and the demographic data; DY, JZ, LC, and ZN participated in the writing process. All authors read and approved the final manuscript.

Funding

This study was supported by Hainan Provincial Natural Science Foundation of China (824RC519); a postgraduate innovation program of Hainan Medical University (HYYS2022B13).

Data availability

The newly collected sequences of this study have been uploaded to GenBank. Accession Number(s): PV021501–PV021513, PV021542–PV021564. The remaining Hainan subtype B subtype sequences can be downloaded from the HIV database (<https://www.hiv.lanl.gov/components/sequence/HIV/search/search.html/>), detailed accession numbering can be obtained from Table S2.

Declarations

Ethics approval and consent to participate

This study received approval from the Ethical Committee of Hainan Medical University, Hainan, China (HYLL-2022-422). The methods employed adhered to pertinent guidelines and regulations. All participants were fully informed about the study's content and objectives and provided their informed consent prior to participation. This study adhered to the Declaration of Helsinki to this effect.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹School of Public Health, Heinz Mehlhorn Academician Workstation, Key Laboratory of Tropical Translational Medicine of Ministry of Education, Hainan Medical University, Haikou, Hainan 571199, China

²The Fifth People's Hospital of Hainan Province, Affiliated Dermatology Hospital of Hainan Medical University, Haikou, Hainan 570102, China

³Guangxi Key Laboratory of AIDS Prevention and Treatment & Guangxi Colleges and Universities Key Laboratory of Prevention and Control of Highly Prevalent Diseases, School of Public Health, Guangxi Medical University, Nanning, Guangxi 530021, China

⁴Collaborative Innovation Centre of Regenerative Medicine and Medical BioResource Development and Application Co-constructed by the Province and Ministry, Life Sciences Institute, Guangxi Medical University, Nanning, Guangxi 530021, China

Received: 23 August 2024 / Accepted: 21 February 2025

Published online: 04 March 2025

References

1. Van Heuverswyn F, Li Y, Bailes E, Neel C, Lafay B, Keele BF, et al. Genetic diversity and phylogeographic clustering of SIVcpzPtt in wild chimpanzees in Cameroon. *Virology*. 2007;368:155–71.
2. Sharp PM, Hahn BH. Origins of HIV and the AIDS pandemic. *Cold Spring Harb Perspect Med*. 2011;1:a006841.
3. Hemelaar J. The origin and diversity of the HIV-1 pandemic. *Trends Mol Med*. 2012;18:182–92.
4. Chan PA, Hogan JW, Huang A, DeLong A, Salemi M, Mayer KH, et al. Phylogenetic investigation of a statewide HIV-1 epidemic reveals ongoing and active transmission networks among men who have sex with men. *J Acquir Immune Defic Syndr*. 2015;70:428–35.
5. Hemelaar J, Gouws E, Ghys PD, Osmanov S. WHO-UNAIDS network for HIV isolation and characterisation. Global trends in molecular epidemiology of HIV-1 during 2000–2007. *AIDS*. 2011;25:679–89.
6. Bokazhanova A, Rutherford GW. The epidemiology of HIV and AIDS in the world. *Coll Antropol*. 2006;30(Suppl 2):3–10.
7. Rife B, Salemi M. On the early dynamics and spread of HIV-1. *Trends Microbiol*. 2015;23:3–4.
8. Junqueira DM, Almeida SE, De M. HIV-1 subtype B: traces of a pandemic. *Virology*. 2016;495:173–84.
9. Magiorkinis G, Angelis K, Mamais I, Katzourakis A, Hatzakis A, Albert J, et al. The global spread of HIV-1 subtype B epidemic. *Infect Genet Evol*. 2016;46:169–79.
10. Dokubo EK, Kim AA, Le L-V, Nadol PJ, Prybylski D, Wolfe MI. HIV incidence in Asia: a review of available data and assessment of the epidemic. *AIDS Rev*. 2013;15:67–76.
11. Lau KA, Wang B, Saksena NK. Emerging trends of HIV epidemiology in Asia. *AIDS Rev*. 2007;9:218–29.
12. Sugimoto SP, Techasrivichien T, Musumari PM, El-saadi C, Lukhele BW, Ono-Kihara M, et al. Changing patterns of HIV epidemic in 30 years in East Asia. *Curr HIV/AIDS Rep*. 2014;11:134–45.
13. He X, Xing H, Ruan Y, Hong K, Cheng C, Hu Y, et al. A comprehensive mapping of HIV-1 genotypes in various risk groups and regions across China based on a nationwide molecular epidemiologic survey. *PLoS ONE*. 2012;7:e47289.
14. Deng W, Fu P, Bao L, Vidal N, He Q, Qin C, et al. Molecular epidemiological tracing of HIV-1 outbreaks in Hainan Island of Southern China. *AIDS*. 2009;23:977–85.
15. Yu DE, Xu YJ, Li M, Yang Y, Liang HY, Zhong SM, et al. HIV-1 subtype diversity and factors affecting drug resistance among patients with virologic failure in antiretroviral therapy in Hainan Province, China, 2014–2020. *Biomed Environ Sci*. 2023;36:800–13.
16. Chen R, Liang B, Wen B, Huang G, Ning C, Lao C, et al. No difference in prevalence of transmitted drug resistance between injection drug users and Non-Injection drug users: A Cross-Sectional study among antiretroviral Treatment-Naïve HIV patients. *Intervirology*. 2018;61:281–91.
17. Struck D, Lawyer G, Ternes A-M, Schmit J-C, Bercoff DP. COMET: adaptive context-based modeling for ultrafast HIV-1 subtype identification. *Nucleic Acids Res*. 2014;42:e144.
18. Trifinopoulos J, Nguyen L-T, von Haeseler A, Minh BQ. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res*. 2016;44:W232–235.
19. Rambaut A. (2009). FigTree. Tree Figure Drawing tool. <http://tree.bio.ed.ac.uk/software/figtree/>. Accessed 12 May 2024.
20. Yu D, Zhu K, Li M, et al. The origin, dissemination, and molecular networks of HIV-1 CRF65_cpx strain in Hainan Island, China. *BMC Infect Dis*. 2024;24(1):269.

21. Drummond AJ, Rambaut A. BEAST: bayesian evolutionary analysis by sampling trees. *BMC Evol Biol*. 2007;7:214.
22. Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods*. 2012;9:772.
23. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. Posterior summarization in bayesian phylogenetics using tracer 1.7. *Syst Biol*. 2018;67:901–4.
24. Baele G, Lemey P, Vansteelandt S. Make the most of your samples: Bayes factor estimators for high-dimensional models of sequence evolution. *BMC Bioinformatics*. 2013;14:85.
25. Bielejec F, Baele G, Vrancken B, Suchard MA, Rambaut A, Lemey P. Spred3: interactive visualization of spatiotemporal history and trait evolutionary processes. *Mol Biol Evol*. 2016;33:2167–9.
26. Kass RE, Raftery AE. (1995). Bayes Factors. *Journal of the American Statistical Association*, 90(430), 773–795. https://sites.stat.washington.edu/courses/stat527/s14/readings/KassRaftery_jasa1995.pdf. Accessed 12 May 2024.
27. Cao W, Hsieh E, Li T. Optimizing treatment for adults with HIV/AIDS in China: successes over two decades and remaining challenges. *Curr HIV/AIDS Rep*. 2020;17:26–34.
28. Wang L, Guo W, Li D, Ding Z, McGoogan JM, Wang N, et al. HIV epidemic among drug users in China: 1995–2011. *Addiction*. 2015;110(Suppl 1 0 1):20–8.
29. Zhao D-C, Wen Y, Ma Y, Zhao Y, Zhang Y, Wu Y-S, et al. Expansion of China's free antiretroviral treatment program. *Chin Med J (Engl)*. 2012;125:3514–21.
30. Ye J, Lu H, Su X, Xin R, Bai L, Xu K, et al. Phylogenetic and Temporal dynamics of human immunodeficiency virus type 1B in China: four types of B strains circulate in China. *AIDS Res Hum Retroviruses*. 2014;30:920–6.
31. Li Z, He X, Wang Z, Xing H, Li F, Yang Y, et al. Tracing the origin and history of HIV-1 subtype B' epidemic by near full-length genome analyses. *AIDS*. 2012;26:877–84.
32. Lan Y, Deng X, Li L, Cai W, Li J, Cai X, et al. HIV-1 drug resistance and genetic transmission networks among MSM failing antiretroviral therapy in South China 2014–2019. *Infect Drug Resist*. 2021;14:2977–89.
33. Zhang D, Zheng C, Li H, Li H, Liu Y, Wang X, et al. Molecular surveillance of HIV-1 newly diagnosed infections in Shenzhen, China from 2011 to 2018. *J Infect*. 2021;83:76–83.
34. Dong M-J, Peng B, Liu Z-F, Ye Q-N, Liu H, Lu X-L, et al. The prevalence of HIV among MSM in China: a large-scale systematic analysis. *BMC Infect Dis*. 2019;19:1000.
35. Chow EPF, Wilson DP, Zhang L. What is the potential for bisexual men in China to act as a Bridge of HIV transmission to the female population? Behavioural evidence from a systematic review and meta-analysis. *BMC Infect Dis*. 2011;11:242.
36. Li X, Zhang B, Wang J, Li Y, Li X, Yu P, et al. Sexual health status of women who have regular sexual relations with men who have sex with men in Mainland China. *BMC Public Health*. 2017;17:168.
37. Weeks MR, Abbott M, Liao S, Yu W, He B, Zhou Y, et al. Opportunities for woman-initiated HIV prevention methods among female sex workers in Southern China. *J Sex Res*. 2007;44:190–201.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.