DYNAMICS AND DISPERSAL OF LOCAL HIV EPIDEMICS WITHIN SAN DIEGO AND ACROSS THE SAN DIEGO-TIJUANA BORDER

Bram Vrancken[1,*], Sanjay R Mehta[2,*], Santiago Ávila-Ríos[3], Claudia García-Morales[3], Daniela Tapia-Trejo[3], Gustavo Reyes-Terán[4], Samuel Navarro-Álvarez[5], Susan J. Little[2], Martin Hoenigl[2], Heather A. Pines[2], Thomas Patterson[2], Steffanie A. Strathdee[2], Davey M. Smith[2], Simon Dellicour[1,6,†], Antoine Chaillon[2,†]

* contributed equally to this work

† contributed equally to this work

[1] Department of Microbiology, Immunology and Transplantation, Rega Institute, Laboratory for Computational and Evolutionary Virology, KU Leuven, Herestraat 49, 3000 Leuven, Belgium

[2] Division of Infectious Diseases and Global Public Health, University of California San Diego, CA.

[3] Centre for Research in Infectious Diseases, National Institute of Respiratory Diseases, Calzada de Tlalpan 4502, Colonia Sección XVI, CP 14080, Mexico City, Mexico

[4] Coordinating Commission of the Mexican National Institutes of Health, Periférico Sur 4809, Arenal Tepepan 14610, Mexico City, Mexico

[5] Hospital General de Tijuana, Av. Centenario 10851, Zona Río, CP 22320, Tijuana, BC, Mexico

[6] Spatial Epidemiology Lab (SpELL), Université Libre de Bruxelles, CP160/12 50, av. FD Roosevelt, 1050 Bruxelles, Belgium

Corresponding Author

Antoine Chaillon

University of California San Diego

Department of Medicine, 9500 Gilman Dr, MC 0697, La Jolla, CA 92093

achaillon@health.ucsd.edu

Summary: Discrete migration models using HIV sequences revealed that central San Diego was a major hub for HIV spread in the US-Mexico border region, with MSM playing an important role. Across the border, viral migration was more intense towards Tijuana.

ABSTRACT

Background. Evolutionary analyses of well-annotated HIV sequence data can provide insights into viral transmission patterns and associated factors. Here, we explored the transmission dynamics of the HIV-1 subtype B epidemic across the San Diego (US) – Tijuana (Mexico) border region to identify factors that could help guide public health policy.

Methods. HIV *pol* sequences were collected from people with HIV in San Diego County and from Tijuana between 1996-2018. A multistep phylogenetic approach was used to characterize the dynamics of spread. The contribution of geospatial factors and HIV risk group to the local dynamics were evaluated.

Results. Phylogeographic analyses of the 2,034 sequences revealed an important contribution of local transmission in sustaining the epidemic, as well as a complex viral migration network across the region. Geospatial viral dispersal between San Diego communities occurred predominantly among men-who-have-sex with-men with central San Diego being the main source (34.9%) and recipient (39.5%) of migration events. HIV migration was more frequent from San Diego county towards Tijuana than vice versa. Migrations were best explained by driving time between locations.

Conclusion. The US-Mexico border may not be a major barrier to the spread of HIV, which may stimulate coordinated transnational intervention approaches. Whereas a focus on central San Diego has the potential to avert most spread, the substantial viral migration independent of central San Diego shows that county-wide efforts will be more effective. Combined, this work shows that epidemiological information gleaned from pathogen genomes can uncover mechanisms that underlie sustained spread and, in turn, can be a building block of public health decision making.

Key words: HIV; Phylogeography; Bayesian Discrete Phylogeography; Generalized Linear Model

INTRODUCTION

San Diego County has the third highest number of HIV cases in California, with an estimated 13,200 persons living with HIV (PWH) [1, 2]. The state of Baja California in Mexico, which includes the city of Tijuana, has the eighth highest HIV prevalence in Mexico with 9300 PWH, nearly all living in and around Tijuana [3]. Together, these cities constitute the San Diego-Tijuana border region, one of the largest metropolitan areas in North America, with approximately 6 million residents. Further, the border between San Diego and Tijuana is the busiest land border-crossing area in the world with 43 million registered crossings in 2018 [4].

The HIV epidemic in the San Diego-Tijuana border region is complex. In contrast to other parts of Mexico and the United States that have epidemics in which the major risk group are men who have sex with men (MSM), the HIV epidemic along the border is more distributed among risk populations, including MSM, persons who inject drugs (PWID), persons who have transactional sex, and Mexican and Central American migrants [5-10]. Hence, the San Diego-Tijuana border region is at the crossroad of multiple risk groups, which leads to transmissions between risk groups [11], possibly enhancing the spread of HIV [12].

Genetic sequence data of pathogens are increasingly used to investigate the transmission dynamics of infectious diseases. This is possible because pathogens evolve as they spread, meaning that their genome contains a genetic imprint of past transmission events. By statistically analyzing this trail of mutations using phylogenetic models one can detect linkages among infections in time and space that may not be evident otherwise, and gain insights into the processes that govern the spatiotemporal spread among individuals. This holds the potential to support the planning, implementation and evaluation of public health practices and response. More specifically, understanding the transmission dynamics by identifying the geographic locations and risk groups associated with viral dispersal can help direct effective prevention and surveillance efforts [13-18]. Here, we use a discrete phylogeographic approach, leveraging the evolutionary signal present in HIV sequences, to explore the patterns and trends in the HIV epidemics of the San Diego-Tijuana border region.

METHODS

Compilation of data sets for phylogenetic inference

Our data set was compiled using all available HIV subtype B partial pol sequences, the predominant circulating subtype in the United States and Mexico [19, 20] and samples obtained in convenience based efforts from participants enrolled in (i) the San Diego Primary Infection Resource Consortium (SD PIRC) from 1996 to 2018 and (ii) in Tijuana as part of the Mexican HIV Drug Resistance Surveillance Network studies (76%) [10, 21], the Enlaces and El Cuete research studies (24%) [22] from 2008 to 2018. For all, sociodemographic information was collected by a counsellor or technician at the time of blood sample donation, and included date of collection, gender, age, HIV transmission risk (and communities for San Diego). See **Supplementary Figure S1** for a map of Tijuana and the communities where SD PIRC participants resided, and **Supplementary Material** for cohorts' description.

All studies were revised and approved by the UCSD Human Research Protections Program (San Diego) or the National Institute of Respiratory Diseases (INER) (Mexico City) Ethics Review Board.

Phylogenetic inference

Transmission networks that best approximate the epidemic dynamics were identified following Cuypers et al. [23]**.** For the identified networks, the spread across risk groups was reconstructed jointly with the geographical migration history in the BEAST 1.10.5 software package [24]. Transmission risks were defined as follows: heterosexual (HTS), men who have sex with men (MSM), heterosexual people who inject drugs (PWID-HTS), MSM who inject drugs (PWID-MSM), bisexual people, and other (**Supplementary Table S1**). The missing data for the risk group variable in the Tijuana cohort were accommodated for in the ancestral reconstructions as sampling uncertainty [25]. To accommodate for the different sampling periods of the cohorts, sensitivity analyses were performed in which all sequences from the San Diego PIRC cohort sampled prior to the inclusion of sequences from Tijuana were excluded (i.e. prior to 2008, n = 406). We refer to these analyses as 'time-filtered'. A GLM extension of the discrete trait model implemented in BEAST 1.10.5 [26] was used to investigate the potential contribution of location-associated variables to the dispersal rates among San Diego County communities and Tijuana.

Details of the methods are provided in **Supplementary Methods**.

RESULTS

Population characteristics

The San Diego - Tijuana data set included a total of 2,034 partial *pol* sequences and associated sociodemographic data collected from 1996-2018 in San Diego (n=806, including 49.7% sampled after 2008) and 2008-2018 in Tijuana (n=1,228). Enrolled PWH were predominantly male (83.5%, 1,445/1,730 PWH with available data), and reporting MSM risk (62%, 1,004/1,621 PWH with available data). Compared to PWH from Tijuana, PWH from San Diego were significantly more likely to be male (97% vs. 71.8%, $p<0.01$) and reporting MSM risk (including MSM also reporting injecting drug use 93% vs 31.2%, $p<0.01$, **Supplementary Table S1**). PWH from San Diego lived across 30 communities while participants from Mexico were assigned to Tijuana, as a single municipality (**Supplementary Table S2 and Supplementary Figure S1**).

Preliminary phylogenetic analysis and down-sampling

A set of 33,637 HIV-1 subtype B *pol* sequences from 60 countries across the world collected between 1992 and 2018 was combined with the San Diego - Tijuana data set. Using a branch support threshold of 0.9 based branch support based on a SH procedure, 104 highly supported clusters corresponding to independent introductions of HIV-1 B lineages into the San Diego/Tijuana area were identified (see **Supplementary Tables S2-S3**). Of these clades, 58 (55.8%) included only sequences from PWH living in Tijuana, 28 (26.9%) comprised only sequences from persons living in San Diego County and the remaining 18 clades (17.3%) included PWH living in both San Diego and Tijuana. Of these, clades that could not inform on the between-location movement (i.e. size < 3 or with sequences from a single location) were excluded from further analysis. This downsampling left 31 clades, of which 15 (48.4%) included sequences from both San Diego County and Tijuana. Using a SH support threshold ≥ 0.7, the final dataset consisted of 41 clades, 19 (46.3%) of which included sequences from both sides of the border (**Supplementary Figure S2**). Sequences within the SH ≥ 0.9 clades made up 75% of the sequences within clades identified the relaxed branch support threshold of 0.7.

Discrete phylogeographic inferences

Transmission dynamics across San Diego County and Tijuana

Phylogeographic analyses of the 2,034 sequences revealed a complex viral migration history across the region with support for links between San Diego communities and reciprocal migration between San Diego and Tijuana (**Table 1** and **Figures 1A-2A**). As the level of statistical support for a particular migration link does not inform on the relative importance of that pathway, the estimated number of expected migration events for the subset of well-supported movements was quantified (**Figure 1A-2A**). This analysis showed that, on average, 69.7% of migration events occurred within San Diego County while 30.3% of were cross-border events (**Figure 2A**). Central San Diego was a major hub of viral migration within San Diego County, as migration from central San Diego to other San Diego communities accounted for 34.9% of all migration events (BF≥3). Central San Diego was also the dominant destination of virus migration within San Diego County (39.5%), suggesting it was acting as the gravitational center of the San Diego epidemic. Moreover, our model also revealed that migration from San Diego County, mainly from the border communities of Chula Vista and San Ysidro, which are adjacent to and less than three miles from the international border, towards Tijuana was more frequent than migration from Tijuana to San Diego (mainly towards La Jolla and central San Diego). Using the more conservative SH clade support threshold of ≥ 0.7 yielded similar findings (**Supplementary Figure S3A and Supplementary Table S4**). To evaluate the potential impact of the discrepancies in sampling period for the San Diego PIRC cohort (starting in 1996) and Tijuana cohort (starting in 2008), we repeated the analyses after excluding sequences from the San Diego PIRC cohort collected prior to 2008 (n=406). The signal for more intense migration from San Diego towards Tijuana was robust to this time-filtering (**Supplementary Figure S4**).

Transmission dynamics between risk groups

As each sequence represents a unique PWH, the internal branches in the phylogeny can be assumed to encompasses at least one migration event [27]. Hence, the inferred risk at the start and end nodes of internal branches can be used to assess the spread within and between risk groups. Among the 31 clades included in the model, MSM accounted for 30.3% of individuals from Tijuana and 92% of PWH living in San Diego county. This was consistent with the local population characteristics (MSM representing 28.7% and 88.3% of PWH from Tijuana and San Diego respectively). Of the links between risk groups that were significant

(BF_{adj} ≥ 3, see Table 1 and **Supplementary Table S4**), those from HTS toward PWID-HTS and from PWID-HTS toward MSM were robust to the cluster identification threshold and the time-filtering. Combined, they represent on average between 52.0% and 60.8% of all transmissions between risk groups across these analyses (**Figures 2, S5, S6** and **Supplementary Table S5**). Transmission from bisexual people toward PWID-MSM was also consistently recovered, but this contributes only little to the mixing between risk groups (range: 1.4%-5.1%). MSM and people that reported 'other risk' are consistently identified as sources of spread, but to which risk group(s) varies with the cluster identification threshold and time-filtering. Their contribution to the overall spread also varies widely (the range for MSM and other risk is 11.9%-24.1% and 3.2%-20.9% respectively).

As the geographical and risk group migration processes were simultaneously inferred, the association between the patterns of spread among locations and risk groups can be probed. Considering only internal branches that accommodate a migration event for SH≥0.9 and BFs_{adj}≥3, we observed that MSM were the major source risk group (on average in 70% of the migration events between locations, **Figure 1B**). Viral migration from San Ysidro and Chula Vista toward Tijuana was among heterosexuals and MSM respectively. In contrast, migration events from Tijuana towards central San Diego were associated with transmission among MSM. These findings are similar for the model with SH threshold ≥0.7 (**Supplementary Figure S5 panel B).**

Estimating Correlates of Viral Migration

A phylogeography-based GLM analysis was used to investigate the association of potential explanatory variables to the dispersal of HIV across San Diego communities and Tijuana. Given the high degree of collinearity between the population size and number of HIV cases per administrative area, only population size was kept as a predictor in the final model. Here, shorter driving time (BF=39.4) and shared borders (BF=6.9) were both associated with the frequency of viral migration. These associations were robust to sampling imbalances (BF_{adj}=18.4 and 3.7 respectively), suggesting that movement of HIV between communities in the region is driven mainly by access and proximity (**Figure 3**). A preliminary analysis clearly indicated that no interaction is expected between both predictors.

DISCUSSION

This study focused on the San Diego-Tijuana border region, which encompasses the busiest land border crossing in the Western hemisphere. We found strong support for at least 104 independent transmission networks of HIV-1 subtype B within San Diego County and Tijuana. The majority (71/104, 68.3%) of these clades were comprised of sequences from a single community (central San Diego, n=13, Tijuana n=57) and did not contribute to cross-border viral migration. This suggests that HIV transmission in San Diego County and Tijuana is mainly sustained by local transmission.

Many HIV prevention programs in large cities are focused on the downtown areas as this is thought to most effectively contain viral spread [28], and this is also the first phase focus of the Ending the HIV Epidemic (EHE) program[29]. The high proportion of clades with sequences only from downtown San Diego indicate that this strategy makes sense. Furthermore, as transmissions towards other San Diego communities often originated from downtown San Diego (**Figure 2A**), this approach could also help reduce geographic spread at larger scales. On the flip side, the intensity of these intervention efforts is often significantly reduced in the surrounding suburban communities. San Diego County includes several larger population centers that each have varying numbers of prevalent and incident cases. Virus migration from these suburban communities towards downtown San Diego (**Figure 2A**) underscores the need for prevention programmes (e.g. EHE) to be vigilant for infections in all communities across a region, as infections in these peripheral communities can seed new chains of transmission in higher risk communities and populations [30].

We also uncovered bidirectional viral migration across the San Diego-Tijuana border, which is predominantly from San Diego toward Tijuana. This indicates that controlling infections in San Diego has the potential to also positively affect the HIV burden in Tijuana and, to a lesser extent, vice versa. Specifically, we found evidence for cross-border viral transmission linking the communities of San Ysidro and Chula Vista (San Diego County communities close to the US-Mexico border) and Tijuana. The San Ysidro port of entry is one of the busiest land border crossings in the world, with tens of thousands of daily commuters traveling from Tijuana to jobs in San Diego, and US residents working in maquiladoras, purchasing services, or seeking entertainment in Tijuana [31]. San Ysidro is tied closely to

Mexico, as 93% of San Ysidro residents were Hispanic and of those over 5 years old, 87% spoke Spanish [32]. Chula Vista, the second largest city in the San Diego County and also very close to the international border, was also identified as a source of HIV spread to Tijuana. These results show that in spite of anti-HIV programmes targeting the border/border towns and local cross-border collaborative initiatives[22, 33], the US-Mexico border does not act as a major barrier to the spread of HIV. This should be a stimulans for more comprehensive efforts, including more coordinated cross-border collaboration. The latter can take the form of a binational border registry to help providers on both sides of the border manage patients that often move between countries. Maintaining these individuals in care with viral suppression is likely to reduce transmission in the region[34]. Data sharing could also be aimed at improved epidemic monitoring, for example to allow for the more timely identification of growing transmission clusters [35, 36].

Viral migration between communities in the San Diego-Tijuana border region is negatively associated with driving time between communities. This indicates that progress in reducing infectivity of PWH in one area will be most felt in nearby locations. Further analyses evaluating how viral migration into and out of a community is impacted by interventions (e.g. screening for acute and early infection) delivered to those communities will be of great interest. With respect to this, promising developments are being made to accommodate such analyses in an online framework [37].

Understanding mixing patterns in transmission risk may also help understanding which groups are at a greater risk of disassortative transmission (i.e. transmission between different risk groups), potentially seeding new outbreaks [38, 39]. We found many links between risk groups (**Figures 2 and S5 and Supplementary Table S5**). This shows that the border region is a 'melting pot' where different types of transmission networks are bridged, which is in line with previous findings [40, 41]. Furthermore, the joint spatial and risk group ancestral reconstructions revealed that viral dispersal within San Diego county occurred almost exclusively among MSM and MSM that also report injecting drug use. Migration events across the San Diego-Tijuana border towards Tijuana were confined within MSM or heterosexual networks, and viral migration towards San Diego was between MSM (**panel B in Figures 1 and S5**). Combined, this indicates that the risk group intermixing occurs almost uniquely within communities, and suggests that non-MSM, non-HTS risk groups do not drive longer-distance spread in this border region.

Some aspects of the spread process in the border region could not be captured by our analyses. In particular, the absence of high-resolution sampling location information for the Tijuana municipality implies that the inter-neighborhood spread in Tijuana could not be investigated. Like the general SD PIRC cohort[42, 43], the geo-annotated subset of the SD PIRC cohort reflects the demographics of the HIV population in San Diego (**Supplementary Table S1**), with a large majority of MSM and a sampling by community that is proportional to the local HIV prevalence[44]. In contrast, determining how well the Tijuana cohort represents the general HIV population in Tijuana is more challenging (see also Supplementary) Nonetheless, the Tijuana cohort reflects the gender and risk group characteristics of the local HIV epidemic, including a larger proportion of women (**Supplementary Table S1)**. For the Tijuana general HIV population (and hence also our sample thereof), fear of discrimination and repression likely underlies reporting biases in disclosing risk behaviors; stigmatized risk groups (including MSM and PWID) are likely underreported and, consequently, our results likely underestimate their role in disassortative transmission. Bisexual people are grouped under 'other risk' in the San Diego cohort, and because of this their role too is likely underestimated.

CONCLUSION

In conclusion, our analysis of HIV migration across space and risk groups point to an uneven intensity of bi-directional viral migration across the US-Mexico border. We also identified Central San Diego as a central hub in regional geographic spread and corroborate the importance of MSM in the movement of HIV across the region. The approaches used to gain these insights have the potential to become standard instruments in the public health response toolbox.

NOTES

DISCLAIMER

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

FUNDING

This work was supported by grants from the NIH (San Diego Center for AIDS Research, CFAR) and the Mexican Government (Comisión de Equidad y Género de las Legislaturas LX-LXI y Comisión de Igualdad de Género de la Legislatura LXII de la Cámara de Diputados de la República Mexicana) to GRT, and Consejo Nacional de Ciencia y Tecnología (CONACyT SALUD-2017-01-289725, and CONACyT 303079) to SAR. BV is supported by a postdoctoral grant of the FWO (Fonds Wetenschappelijk Onderzoek – Vlaanderen, grant nr. 12U7118N). SJL acknowledges support from NIH Grants AI106039 (R24) and MH100974 (R01). SD was supported by the Fonds Wetenschappelijk Onderzoek (FWO, Belgium) and is currently funded by the Fonds National de la Recherche Scientifique (FNRS, Belgium). AC was supported by NIH Grant AI131971 (R21). AC was supported by NIH Grant AI131971 (R21). SRM is supported by NIH Grants AI106039 (R24) and R01 AI35992 (PI: Wertheim). HAP and TP are supported by NIH Grant R01DA037811. We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan V GPU used for this research.

Conflicts of Interest

SL reports non-financial support in the form of antiretroviral medication from Gilead Sciences during the conduct of the study. and personal fees and non-financial support from Gilead Sciences outside the submitted work. HP reports grants from the National Institute on Drug Abuse during the conduct of the study and grants from Gilead Sciences outside the submitted work. MH reports grants from Gilead and Pfizer outside the submitted work. All other authors have no potential conflicts.

REFERENCES

1. Health and Human Services Agency (HHSA) CoSD. HIV/AIDS Epidemiology Report 2015. San Diego, CA: County of San Diego, **2016**.
2. AIDSVu. AIDSVu: an interactive online mapping tool that visualizes the impact of the HIV epidemic on communities across the United States. Available at: https://aidsvu.org/. Accessed 03/2019.
3. Centro Nacional para la Prevención y el Control del VIH y el SIDA (CENSIDA). Vigilancia Epidemiológica de casos de VIH/SIDA en México, Registro Nacional de Casos de SIDA, Actualización al 11 de noviembre del 2019., **2019**.
4. United States Department of Transportation. Bureau of Transportaion Statistics (BTS). Available at: https://www.bts.gov/topics/national-transportation-statistics.
5. Goldenberg S, Silverman J, Engstrom D, Bojorquez-Chapela I, Strathdee S. "Right Here is the Gateway": Mobility, Sex Work Entry and HIV Risk Along the Mexico-U.S. Border. Int Migr **2014**; 52(4): 26-40.
6. Pitpitan EV, Rocha-Jimenez T, Salazar M, Chavarin C, Magis-Rodriguez C. A Mixed Methods Analysis of the Venue-Related Social and Structural Context of Drug Use During Sex Among Male Clients of Female Sex Workers in Tijuana, Mexico. AIDS Behav **2019**.
7. Zhang X, Martinez-Donate AP, Simon NE, et al. Risk behaviours for HIV infection among travelling Mexican migrants: The Mexico-US border as a contextual risk factor. Global public health **2017**; 12(1): 65-83.
8. Mehta SR, Wertheim JO, Brouwer KC, et al. HIV Transmission Networks in the San Diego-Tijuana Border Region. EBioMedicine **2015**; 2(10): 1456-63.
9. Marks C, Zuniga ML. CAM Practices and Treatment Adherence Among Key Subpopulations of HIV+ Latinos Receiving Care in the San Diego-Tijuana Border Region: A Latent Class Analysis. Frontiers in public health **2019**; 7: 179.
10. Garcia-Morales C, Tapia-Trejo D, Quiroz-Morales VS, et al. HIV pretreatment drug resistance trends in three geographic areas of Mexico. J Antimicrob Chemother **2017**; 72(11): 3149-58.
11. Le Vu S, Ratmann O, Delpech V, et al. HIV-1 Transmission Patterns in Men Who Have Sex with Men: Insights from Genetic Source Attribution Analysis. AIDS Res Hum Retroviruses **2019**; 35(9): 805-13.

12. Esbjörnsson J, Mild M, Audelin A, et al. HIV-1 transmission between MSM and heterosexuals, and increasing proportions of circulating recombinant forms in the Nordic Countries. Virus evolution **2016**; 2(1): vew010-vew.
13. Graf T, Vrancken B, Maletich Junqueira D, et al. Contribution of Epidemiological Predictors in Unraveling the Phylogeographic History of HIV-1 Subtype C in Brazil. J Virol **2015**; 89(24): 12341-8.
14. Perez AB, Vrancken B, Chueca N, et al. Increasing importance of European lineages in seeding the hepatitis C virus subtype 1a epidemic in Spain. Euro Surveill **2019**; 24(9).
15. Xia Q, Wertheim JO, Braunstein SL, Misra K, Udeagu CC, Torian LV. Use of molecular HIV surveillance data and predictive modeling to prioritize persons for transmission-reduction interventions. Aids **2019**.
16. Little SJ, Kosakovsky Pond SL, Anderson CM, et al. Using HIV networks to inform real time prevention interventions. PLoS ONE **2014**; 9(6).
17. Wertheim JO, Kosakovsky Pond SL, Little SJ, De Gruttola V. Using HIV transmission networks to investigate community effects in HIV prevention trials. PLoS ONE **2011**; 6(11): e27775.
18. Vrancken B, Cuypers L, Pérez AB, et al. Cross-country migration linked to people who inject drugs challenges the long-term impact of national HCV elimination programmes. J Hepatol **2019**; 71(6): 1270-2.
19. Hemelaar J, Gouws E, Ghys PD, Osmanov S, Isolation W-UNfH, Characterisation. Global trends in molecular epidemiology of HIV-1 during 2000-2007. AIDS (London, England) **2011**; 25(5): 679-89.
20. Avila-Ríos S, García-Morales C, Garrido-Rodríguez D, et al. National Prevalence and Trends of HIV Transmitted Drug Resistance in Mexico. PLoS ONE **2011**; 6(11): e27812.
21. Avila-Rios S, Garcia-Morales C, Valenzuela-Lara M, et al. HIV-1 drug resistance before initiation or re-initiation of first-line ART in eight regions of Mexico: a sub-nationally representative survey. J Antimicrob Chemother **2019**; 74(4): 1044-55.
22. Mehta SR, Chaillon A, Gaines TL, et al. Impact of Public Safety Policies on Human Immunodeficiency Virus Transmission Dynamics in Tijuana, Mexico. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America **2018**; 66(5): 758-64.
23. Cuypers L, Vrancken B, Fabeni L, et al. Implications of hepatitis C virus subtype 1a migration patterns for virus genetic sequencing policies in Italy. BMC Evol Biol **2017**; 17(1): 70.
24. Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol **2018**; 4(1): vey016.
25. Scotch M, Tahsin T, Weissenbacher D, et al. Incorporating sampling uncertainty in the geospatial assignment of taxa for virus phylogeography. Virus Evolution **2019**; 5(1).
26. Lemey P, Rambaut A, Bedford T, et al. Unifying viral genetics and human transportation data to predict the global transmission dynamics of human influenza H3N2. PLoS Pathog **2014**; 10(2): e1003932.
27. Romero-Severson E, Skar H, Bulla I, Albert J, Leitner T. Timing and order of transmission events is not directly reflected in a pathogen phylogeny. Molecular Biology and Evolution **2014**; 31(9): 2472-82.

28. Ávila-Ríos S, García-Morales C, Valenzuela-Lara M, et al. HIV-1 drug resistance before initiation or re-initiation of first-line ART in eight regions of Mexico: a sub-nationally representative survey. J Antimicrob Chemother **2019**; 74(4): 1044-55.
29. hiv.gov. What Is Ending the HIV Epidemic: A Plan for America? Available at: https://www.hiv.gov/federal-response/ending-the-hiv-epidemic/overview. Accessed 2020/09/09.
30. Mehta SR, Murrell B, Anderson CM, et al. Using HIV Sequence and Epidemiologic Data to Assess the Effect of Self-referral Testing for Acute HIV Infection on Incident Diagnoses in San Diego, California. Clin Infect Dis **2016**; 63(1): 101-7.
31. The General Services Administration G. San Ysidro Land Port of Entry. Available at: https://www.gsa.gov/about-us/regions/welcome-to-the-pacific-rim-region-9/land-ports-of-entry/san-ysidro-land-port-of-entry. Accessed August 2019.
32. U.S. Census Bureau. American FactFinder - Results from San Diego County - California. Available at: https://www.census.gov/quickfacts/table/PST045215/06073.
33. Strathdee SA, Magis-Rodriguez C, Mays VM, Jimenez R, Patterson TL. The emerging HIV epidemic on the Mexico-U.S. border: an international case study characterizing the role of epidemiology in surveillance and response. Annals of epidemiology **2012**; 22(6): 426-38.
34. Smith LR, Patterson TL, Magis-Rodriguez C, et al. Engagement in the HIV Care Continuum among Key Populations in Tijuana, Mexico. AIDS Behav **2016**; 20(5): 1017-25.
35. Poon AF, Gustafson R, Daly P, et al. Near real-time monitoring of HIV transmission hotspots from routine HIV genotyping: an implementation case study. The lancet HIV **2016**; 3(5): e231-8.
36. Ragonnet-Cronin M, Hodcroft EB, Wertheim JO. Understanding disclosed and cryptic HIV transmission risk via genetic analysis: what are we missing and when does it matter? Curr Opin HIV AIDS **2019**; 14(3): 205-12.
37. Gill MS, Lemey P, Suchard MA, Rambaut A, Baele G. Online Bayesian Phylodynamic Inference in BEAST with Application to Epidemic Reconstruction. Mol Biol Evol **2020**; 37(6): 1832-42.
38. Hoenigl M, Chaillon A, Kessler HH, et al. Characterization of HIV Transmission in South-East Austria. PLoS ONE **2016**; 11(3): e0151478.
39. Le Vu S, Ratmann O, Delpech V, et al. HIV-1 Transmission Patterns in Men Who Have Sex with Men: Insights from Genetic Source Attribution Analysis. AIDS Research and Human Retroviruses **2019**; 35(9): 805-13.
40. Mehta SR, Chaillon A, Gaines TL, et al. Impact of Public Safety Policies on Human Immunodeficiency Virus Transmission Dynamics in Tijuana, Mexico. Clin Infect Dis **2018**; 66(5): 758-64.
41. Mehta SR, Wertheim JO, Brouwer KC, et al. HIV Transmission Networks in the San Diego-Tijuana Border Region. EBioMedicine **2015**; 2(10): 1456-63.
42. Hoenigl M, Chaillon A, Morris SR, Little SJ. HIV Infection Rates and Risk Behavior among Young Men undergoing community-based Testing in San Diego. Sci Rep **2016**; 6: 25927.
43. County of San Diego HaHSAH. HIV/AIDS Epidemiology Report, **2015** 04/2016.

44. Chaillon A, Hoenigl M, Freitas L, et al. Optimizing Screening for HIV. Open forum infectious diseases **2020**; 7(2).
45. Lemey P, Rambaut A, Drummond AJ, Suchard MA. Bayesian phylogeography finds its roots. PLoS Computational Biology **2009**; 5(9).
46. Chaillon A, Gianella S, Dellicour S, et al. HIV persists throughout deep tissues with repopulation from multiple anatomical sources Journal of Clinical Investigation **2020**: In Press.
47. Kass RE, Raftery AE. Bayes Factors. Journal of the American Statistical Association **1995**; 90(430): 773-95.
48. INEGI (National Institute of Statistics and Geography). Mexico Population Census Available at: http://www3.inegi.org.mx/sistemas/scitel/Default?ev=5. Accessed 12/2018.
49. San Diego County Epidemiology Unit. HIV/AIDS Epidemiology Report Available at: http://www.sandiegocounty.gov/hhsa/programs/phs/hiv_aids_epidemiology_unit/reports_and_statistics.html. Accessed 08/25/2019.
50. Dellicour S, Vrancken B, Trovão NS, Fargette D, Lemey P. On the importance of negative controls in viral landscape phylogeography. Virus Evolution **2018**; 4(2).

TABLES AND FIGURE LEGENDS

TABLE 1. Overview of the well-supported migration links across locations and risk groups. The Bayes factors (BF) were obtained using a model averaging procedure (Bayesian stochastic search variable selection) [45]. We refer to Supplementary Methods for details on how the BF_{adj} - which is a more realistic measure of support compared to the 'default' BF [46] - is obtained. BF_{adj} support of 3 was considered as the lower bound for consideration[47]. Results based on clades identified with SH ≥ .7 are presented in Supplementary Table S4.

from	**to**	**BF**	**BF_{adj}**	**BF**	**BF_{adj}**
		complete San Diego cohort		time-filtered San Diego cohort	
location					
Carlsbad	San Diego (central)	64.04	27	110.74	20.04
Chula Vista	Tijuana	43.99	11.34	27.82	4.42
Encinitas	San Diego (central)	15.03	3.88		
San Ysidro	Tijuana	44.46	32.43	98.68	6.64
Tijuana	Vista	7.5	4.08	4.78	3.59
Jamul	Carlsbad	39.65	3.09	29.9	3.63
Tijuana	Carlsbad	4.29	3.11	5.47	3.89
Lamesa	Chula Vista	14.25	3.22	16.21	4.02
Tijuana	La Jolla	5.25	3.13		
San Diego (central)	Lakeside	6.39	3.08	4.45	3.50
San Diego (central)	Lamesa	26.41	6.81	35.71	5.83
San Diego (central)	National City	16.94	4.67		
Vista	Oceanside	21.21	3.99	19.66	3.41
Tijuana	San Diego (central)	12.5	10.76	18.57	5.24
Vista	San Diego (central)	26.41	4.97	34.54	4.47
Escondido	San Ysidro			5.12	3.50
National City	San Diego (central)			17.31	3.25
Tijuana	Bonita			4.57	3.91
Lakeside	Escondido			9.53	5.56

San Diego (central)	Escondido			5.98	4.64
Spring Valley	La Jolla			9.11	6.52
Lamesa	Lakeside			3.17	3.62

		risk group			
HTS	PWID-HTS	33.36	16.16	47.78	26.23
PWID-HTS	MSM	22.7	5.07	16.1	4.26
Other risk	HTS	13.19	5.7	40.04	5.15
Bisexual	PWID-MSM	8.68	3.18	7.67	3.06
MSM	Other risk	5.89	3.04		
MSM	Bisexual	3.78	4.06		
HTS	PWID-MSM			4.78	3.47
MSM	PWID-MSM			8.44	85.85
MSM	HTS			6.67	3.86
PWID-MSM	HTS			4.36	3.51
Other risk	PWID-MSM			3.86	4.28

FIGURE 1. Lineage dispersal events between locations (i.e. San Diego communities and the city of Tijuana) and between risk groups. A. The thickness of the arrows reflects the average number of inferred migration events between locations and the color of the arrows indicates the corresponding adjusted Bayes factor (BF_{adj}) support. **B.** For all migration events between locations with BF_{adj} ≥3, the thickness of the arrows reflects the average number of inferred migration events within or between risk groups and the color of the arrows indicates the group mixing patterns. Results were obtained from discrete models including clades with Shimodaira Hasegawa (SH) branch support ≥ 0.9. Tijuana is colored in darker grey. See also **Supplementary Figure S5** for results from discrete models including clades with SH branch support ≥ 0.7.

FIGURE 2. Relative contribution of the various migration links to the spread of HIV-1 subtype B in the San Diego-Tijuana area (A) and the relative contribution of risk groups to the spread of HIV-1 B in San Diego and the city of Tijuana (B). A: We present the results from the discrete phylogeographic analysis including clades with Shimodaira Hasegawa (SH) support ≥ 0.9. The Sankey plot represents the average proportion of migration events from each source location ('from') toward the recipient location ('to'). Left side of the plot shows the origin location and the right side of the plot shows the destination location. We here only report migration events associated with an adjusted Bayes factor (BF_{adj}) support ≥ 3. All corresponding BFs are presented in **Table 1**. **B:** Results based on the clade-identification using Shimodaira Hasegawa (SH) branch support ≥ 0.9 and accounting for migration links associated with an adjusted Bayes factor (BF_{adj}) ≥ 3. The Sankey plot represents the proportion of migration events from each source risk group ('from') toward the recipient risk group ('to'). Results from the discrete phylogeographic reconstruction based on clades with SH support ≥ 0.7 are presented in **Supplementary Figure S3**. All corresponding BFs are presented in **Table 1**. Colors were chosen to visually clearly distinguish the different types of migration events and have no specific meaning.

FIGURE 3. Predictors of migration rates between locations. The boxplots report the contribution of each predictor when included in the model. We also report BF support associated with each predictor considered in the GLM, and the BF_{adj} when ≥ 3.

Figure 1

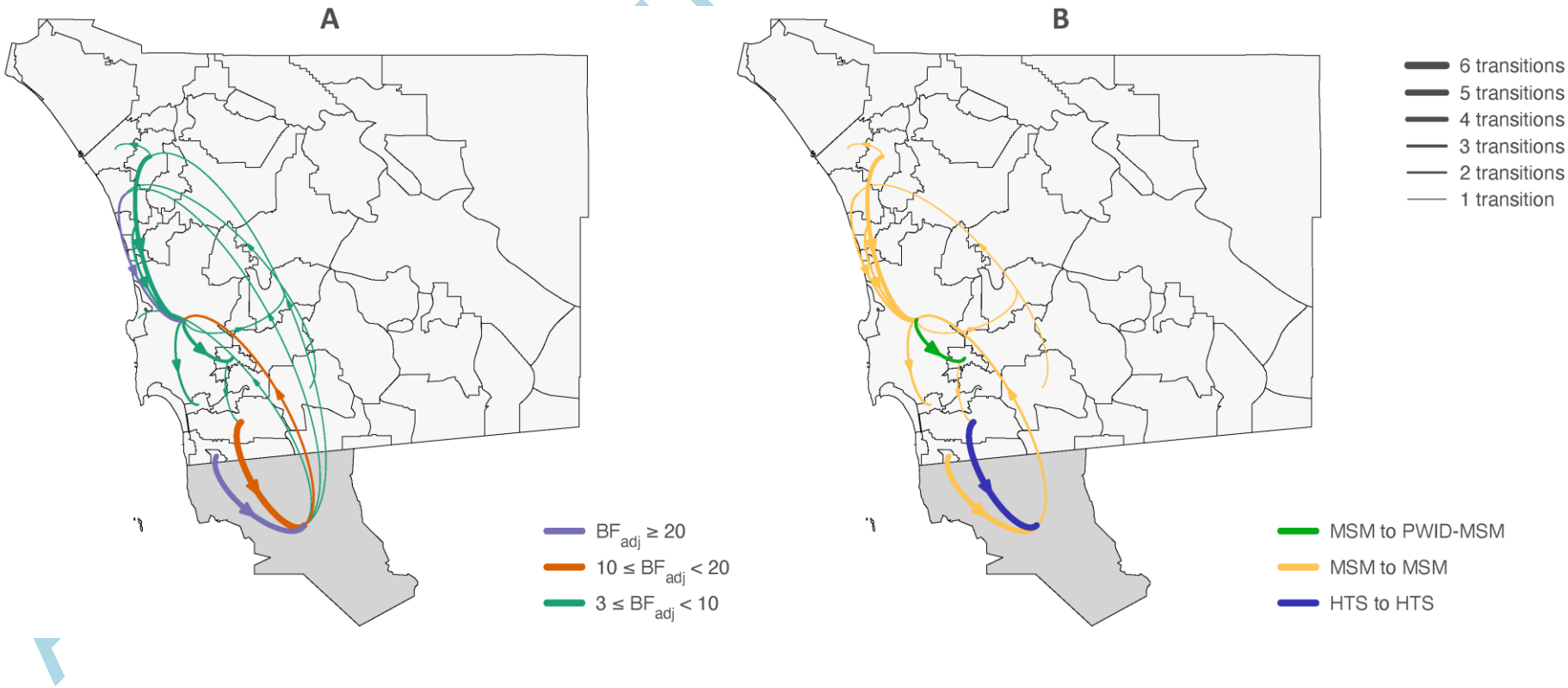

Figure 2

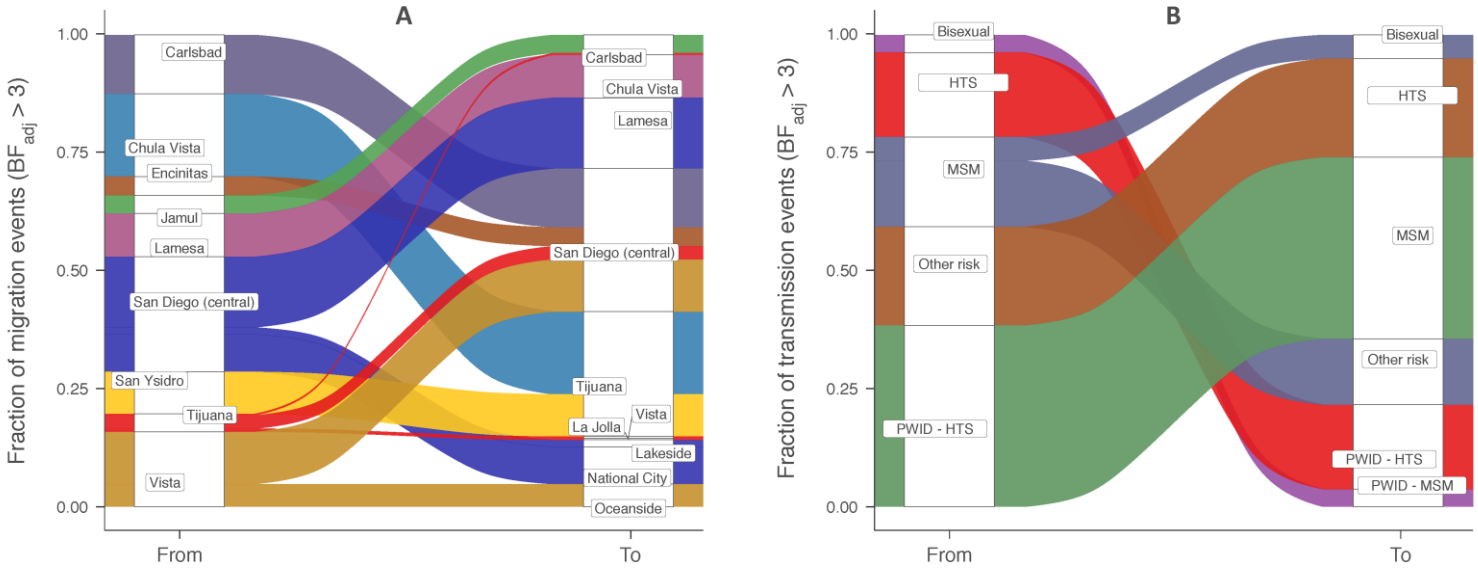

Figure 3