EBioMedicine 2 (2015) 1456–1463

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.ebiomedicine.com

Research Article

HIV Transmission Networks in the San Diego–Tijuana Border Region

Sanjay R. Mehta [a,b,*], Joel O. Wertheim [a], Kimberly C. Brouwer [c], Karla D. Wagner [d], Antoine Chaillon [a], Steffanie Strathdee [c], Thomas L. Patterson [e], Maria G. Rangel [f], Mlenka Vargas [a], Ben Murrell [a], Richard Garfein [c], Susan J. Little [a], Davey M. Smith [a,b]

[a] Division of Infectious Diseases, University of California San Diego, #8208 200 W. Arbor St. San Diego, CA 92103, United States
[b] Division of Infectious Diseases, San Diego Veterans Affairs Medical Center, 3350 La Jolla Village Drive, San Diego, CA 92161, United States
[c] Division of Global Public Health, University of California San Diego, #0507, La Jolla, CA 92093, United States
[d] School of Community Health Sciences, University of Nevada Reno, Lombardi Building, 203, MS 0274, Reno, NV 89557, United States
[e] Department of Psychiatry, University of California San Diego #0680, La Jolla, CA 92093, United States
[f] El Colegio de la Frontera Norte, San Antonio del Mar, Baja California, Mexico

ARTICLE INFO

Article history:
Received 24 June 2015
Accepted 16 July 2015
Available online 18 July 2015

Keywords:
HIV
Phylogeography
International border
Mexico
Transmission network

ABSTRACT

Background: HIV sequence data can be used to reconstruct local transmission networks. Along international borders, like the San Diego–Tijuana region, understanding the dynamics of HIV transmission across reported risks, racial/ethnic groups, and geography can help direct effective prevention efforts on both sides of the border.
Methods: We gathered sociodemographic, geographic, clinical, and viral sequence data from HIV infected individuals participating in ten studies in the San Diego–Tijuana border region. Phylogenetic and network analysis was performed to infer putative relationships between HIV sequences. Correlates of identified clusters were evaluated and spatiotemporal relationships were explored using Bayesian phylogeographic analysis.
Findings: After quality filtering, 843 HIV sequences with associated demographic data and 263 background sequences from the region were analyzed, and 138 clusters were inferred (2–23 individuals). Overall, the rate of clustering did not differ by ethnicity, residence, or sex, but bisexuals were less likely to cluster than heterosexuals or men who have sex with men ($p = 0.043$), and individuals identifying as white ($p \leq 0.01$) were more likely to cluster than other races. Clustering individuals were also 3.5 years younger than non-clustering individuals ($p < 0.001$). Although the sampled San Diego and Tijuana epidemics were phylogenetically compartmentalized, five clusters contained individuals residing on both sides of the border.
Interpretation: This study sampled ~7% of HIV infected individuals in the border region, and although the sampled networks on each side of the border were largely separate, there was evidence of persistent bidirectional cross-border transmissions that linked risk groups, thus highlighting the importance of the border region as a "melting pot" of risk groups.
Funding: NIH, VA, and Pendleton Foundation.

1. Introduction

Tijuana, Baja California is the fifth largest city in Mexico, and it shares an international border with San Diego, California, which is the eighth largest city in the United States (US). Tijuana's proximity to the US and the presence of large numbers of factories operating in the free trade zone between the US and Mexico have led to a large number of migrants to Tijuana from other parts of Mexico and from across Central and South America (Bronfman et al., 2002). In addition, California is home to the highest concentration of immigrants of Mexican-origin in the US, and many of these individuals commute regularly across the border to Mexico to visit family and access goods and services (Servin et al., 2012). Further, tourism draws large numbers of US residents south to Tijuana and the rest of Baja California (Latino, 2014). Hence, the land border crossing between San Diego and Tijuana is one of the busiest in the world, with over 50 million crossings each year (RITA, 2012). This extensive flow of people can lead to the rapid dissemination of infectious diseases, as demonstrated by the H1N1 influenza epidemic of 2009 (Dawood et al., 2009).

HIV is another important infectious disease with high potential for transmission along the San Diego–Tijuana border. The large migrant population and the border have fostered a thriving market for quasi-legal prostitution, which is permitted by Mexican authorities in a "tolerance" zone (aka *Zona Roja*), which is walking distance (<200 m) from the border with the US (Patterson et al., 2009). The low cost of transactional sex in this tolerance zone results in large

* Corresponding author at: 200 W. Arbor Drive #8208, San Diego, CA 92103-8208, United States.
E-mail address: srmehta@ucsd.edu (S.R. Mehta).

http://dx.doi.org/10.1016/j.ebiom.2015.07.024

numbers of US customers (Sirotin et al., 2010), and sex workers having US clients are offered more money for unprotected sex, have higher rates of other risky behaviors including drug use, and consequently have significantly higher rates of HIV infection than the general population (Patterson et al., 2009; Wagner et al., 2013; Syvertsen et al., 2014). Another important HIV risk group in the region is men who have sex with men (MSM). The prevalence of HIV is equally high among MSM in both San Diego and Tijuana, with estimated rates of 18% and 20%, respectively (MMWR, 2010; Pitpitan et al., 2015). Further, HIV risk factors often co-occur as ~47% of male injection drug users in Tijuana reported at least one lifetime male sexual partner (Deiss et al., 2008) and ~10% of San Diego MSM reported ever using injection drugs (Ghanem et al., 2011).

Along international borders, like the San Diego–Tijuana region, understanding the dynamics of HIV transmission across reported risks, racial/ethnic groups, and geography on each side of the border, and how the epidemics are linked, will be important in the implementation of effective prevention efforts (Mehta et al., 2010). Since HIV evolves rapidly, viral sequences isolated from infected individuals are typically unique, and transmission networks can be inferred by identifying sequences that are genetically closely related (Smith et al., 2009; Wertheim et al., 2014). In collaboration with multiple research groups in Tijuana and San Diego, we evaluated HIV sequence and socio-demographic data to characterize the HIV transmission network in the San Diego–Tijuana region.

2. Methods

2.1. Study Population and HIV Sequence Generation

The study population consisted of HIV-1 infected individuals drawn from ten studies located in the San Diego–Tijuana border region (n = 845) (Table 1 and Supplementary Tables 1 & 2). Subjects from the San Diego Primary Infection Resource Consortium (SD PIRC) were diagnosed and had sampling during acute or early infection, as previously described (Le et al., 2013). Subjects from all other studies were sampled at an unknown point during their infection. Partial HIV-1 *pol* sequences for all samples were obtained directly from the primary study or were obtained by sequencing of the HIV-1 (*prot/rt*) from RNA or DNA extracted from banked blood samples (Smith et al., 2009; Howlett et al., 2013). Only individuals who had a sequence available were included in the study. For individuals with sequences from multiple time points, only the sequence obtained from the earliest time point was included. When available, demographic, geographic and clinical data were collected from each study database of the participating program. This information was also used to screen genetically-linked individuals from different studies to ensure that the same person was not sampled twice.

2.2. Sequence Analysis and HIV Network Inference

All sequences that were not duplicated, contaminated or hypermutated by APOBEC were included in a single database, leaving 843 eligible sequences (Rose & Korber, 2000). An additional 263 sequences obtained from unique chronically infected individuals in San Diego but without associated socio-demographic data were included in the analysis to enhance detection of putative transmission clusters. All sequences were subtyped using SCUEAL (Kosakovsky Pond et al., 2009). Sequences were analyzed for genetic relatedness using a pairwise distance comparison using the Tamura-Nei 93 evolutionary model to correct for substitution biases and the unequal base composition found in HIV (Wertheim et al., 2014) (See Supplementary Material).

Putative transmission clusters were inferred by establishing linkage when two sequences had a Tamura-Nei 93 genetic distance (D) ≤1.5% separating them. This threshold is standard in the field, and was selected based on previous work from our group showing that within a mono-infected person, *pol* sequences typically do not diverge more than 1% from baseline over a decade (Hightower et al., 2013). Putative transmission clusters included all connected nodes such that all nodes (individuals) within a cluster would have a D of <1.5% from at least one other node in the cluster, but not necessarily within 1.5% of all nodes within the cluster.

Table 1
Descriptions of participating cohorts.

Cohort	Description	Dates	Subjects	Location
Detection of HIV-infected unaware Latinos in San Diego (San Diego Latino Study)	A pilot respondent-driven sampling (RDS) study to identify HIV-infected Latino population in San Diego border region.	3/2010–5/2010	11	San Diego South Bay
El Cuete*	An ongoing research program studying injection drug users in Tijuana, Mexico that used both RDS (2006–2009) and convenience sampling (2010–present) approaches.	2006 to present	29	Tijuana
Hombre Seguro*	An ongoing study of male clients of female sex workers working in the Tijuana, Mexico that used convenience sampling approaches.	9/2010–10/2012	12	Tijuana
Amigos*	Study of male clients of female sex workers in partnership with the *Hombre Seguro* study who were: ≥18 years old, reported heroin, methamphetamine, or cocaine use in the past four months, and having paid or traded something of value for sex with a female sex worker in Tijuana in the past four months.	6/2011–8/2012	3	Tijuana
Mujer Segura	A research program studying female sex workers in Tijuana, Mexico, some of who use injection drugs. The study used time location sampling and convenience sampling approaches.	2004–2009	25	Tijuana
Mujer Más Segura	A research program studying female sex workers in Tijuana, Mexico who use injection drugs. The study used targeted sampling approaches.	2008–2012	6	Tijuana
Proyecto Parejas	A research project studying female sex workers and their non-commercial partners in the Tijuana Mexico that used convenience sampling approaches.	2010–2013	9	Tijuana
San Diego Primary Infection Cohort	An ongoing study to recruit and characterize acute and early HIV infected individuals. This study has used venue based and RDS recruitment approaches.	6/1996–6/2013	693	San Diego
Study to Assess Hepatitis C Risk (STAHR)	An ongoing research program studying 18–40 year-old (STAHR) and 18+ year-old (STAHR-II) injection drug users in San Diego, CA recruited through RDS and convenience-based approaches.	3/2009–6/2010	17	San Diego
Study to Assess Hepatitis C Risk II (STAHR-II)	An ongoing research program studying 18–40 year-old (STAHR) and 18+ year-old (STAHR-II) injection drug users in San Diego, CA recruited through venue based, targeted and convenience sampling approaches.	6/2012–1/2014	40	San Diego

* El Cuete, Hombre Seguro, and Amigos studies all included at least one individual living in San Diego but sampled in Tijuana.

Table 2
Available socio-demographic data of subjects by participating study.

	Gender, % (n)			Race*, % (n)				Ethnicity, % (n)		Sexuality, % (n)			STI, % (n)		
	M	F	TG	Wh	AA	NAm	Other	Hisp	Non-Hisp	MSM	Bi	Het	GC	CT	Syp
Amigos	100 (3)	0 (3)	0 (3)	NA	NA	NA	NA	100 (3)	0 (3)	0 (3)	100 (3)	0 (3)	0 (3)	0 (3)	66.7 (3)
Detection and characterization of HIV in San Diego Latinos	54.6 (11)	36.4 (11)	9.1 (11)	NA	NA	NA	NA	100 (11)	0 (11)	18.2 (11)	9.1 (11)	72.7 (11)	NA	NA	NA
El Cuete	62.1 (29)	37.9 (29)	0 (29)	65.5 (29)	0 (29)	13.8 (29)	20.7 (29)	96.4 (28)	3.6 (28)	3.4 (29)	3.4 (29)	93.1 (29)	0 (1)	0 (1)	0 (1)
Hombre Seguro	100 (12)	0 (12)	0 (12)	NA	100 (1)	NA	NA	100 (12)	0 (12)	0 (12)	33.3 (12)	66.6 (12)	8.3 (12)	8.3 (12)	16.7 (12)
Mujer Segura and Mujer Más Segura	0 (32)	100 (32)	0 (32)	NA	NA	NA	NA	100 (32)	0 (32)	0 (32)	0 (32)	100 (32)	0 (6)	33.3 (6)	16.7 (6)
Proyecto Parejas	44.4 (9)	55.5 (9)	0 (9)	NA	NA	NA	NA	100 (9)	0 (9)	0 (9)	9.1 (9)	72.8 (9)	0 (7)	0 (7)	28.6 (7)
San Diego primary infection cohort	97.1 (653)	3.8 (653)	0.2 (653)	74.3 (650)	9.4 (650)	9.8 (650)	8.9 (650)	35.1 (502)	64.9 (502)	82.6 (642)	3.4 (642)	14.0 (642)	5.1 (430)	6.3 (431)	13.1 (389)
STAHR	94.1 (17)	5.9 (17)	0 (17)	NA	NA	NA	NA	29.4 (17)	70.6 (17)	64.7 (17)	29.4 (17)	5.9 (17)	44.4 (9)	0 (9)	0 (9)
STAHR-II	87.5 (40)	5.0 (40)	5.0 (40)	80.0 (35)	14.3 (35)	14.3 (35)	0	33.3 (39)	66.6 (39)	53.8 (39)	23.1 (39)	23.1 (39)	NA	NA	NA

Legend: Male (M), Female (F), Transgender (TG), White (Wh), African–American (AA), Native American (Nam), Hispanic (Hisp), Men who have sex with men (MSM), Bisexual (Bi), Heterosexual (Het), Gonorrhea (GC), Chlamydia (CT), Syphilis (Syp). n = total number in cohort for whom data were available.
* Subjects reporting more than one race between (White, African American, or Native American are counted more than once.

2.3. Correlates of Clustering

A standardized data extraction form was used to map variables from each participating study to a common set of variables used in this study. Extracted data was incorporated into a unified database of socio-demographic, clinical, geographic, and viral sequence data from all participating studies. The database was then hand-curated to ensure variables were translated and mapped appropriately and consistently. Given the variety of studies comprising the dataset and the differing goals associated with each of the participating studies, significant variation in the amount and type of data collected occurred between studies. Demographic information and clinical data collected included age, sex, self-reported race/ethnicity, and HIV risk factor(s) for all subjects. Additional data, including clinical information (e.g. CD4 T cell count, viral load, estimated date of infection [EDI], sexually transmitted infections [STI]), behavioral information (e.g. type and frequency of injection drug use, type and frequency of sexual activity), and geographic information varied in availability by study (Table 2 and Supplementary Tables 3 & 4). An individual was denoted to have heterosexual risk behavior, if he or she only reported sex only with the opposite gender. An individual was denoted to have bisexual risk behavior if he reported sex with both males and females; all bisexuals in this study were male.

To identify possible correlates of clustering, the socio-demographic characteristics of infected individuals whose viral sequences clustered were compared to non-clustering individuals. We also evaluated these factors in relation to cluster size, comparing large clusters (>3 members) to small clusters. Categorical variables were compared using the Chi-square or Fisher's exact test, and continuous variables were evaluated using *t*-tests.

2.4. Compartmentalization and Migration Analysis

To better understand the role of the San Diego–Tijuana border on HIV transmission networks, we evaluated the characteristics of the putative transmission clusters including individuals on both sides of the border. We also evaluated the relatedness of the epidemics in San Diego and Tijuana using estimates of population structure (*Fst* (Hudson et al., 1992)), nearest neighbor measurements (Hudson Snn Test (Hudson, 2000)), and cladistic measures of migration events (Slatkin–Maddison Test) (Slatkin & Maddison, 1989), followed by permutation testing in the eligible sequence dataset. We next applied phylogeographic analysis to evaluate the spatiotemporal dynamics of HIV migration between San Diego and Tijuana by analyzing sampled sequences in conjunction with background subtype B sequences obtained from the HIV Sequence Database at the Los Alamos National Laboratory (Kuiken et al., 2003). We employed a Bayesian discrete phylogeographic approach using Markov chain Monte Carlo (MCMC) sampling implemented in the BEAST v1.8.1 software package (Drummond & Rambaut, 2007). Details of the reconstruction are provided in the supplementary methods.

2.5. Drug Resistance Screening

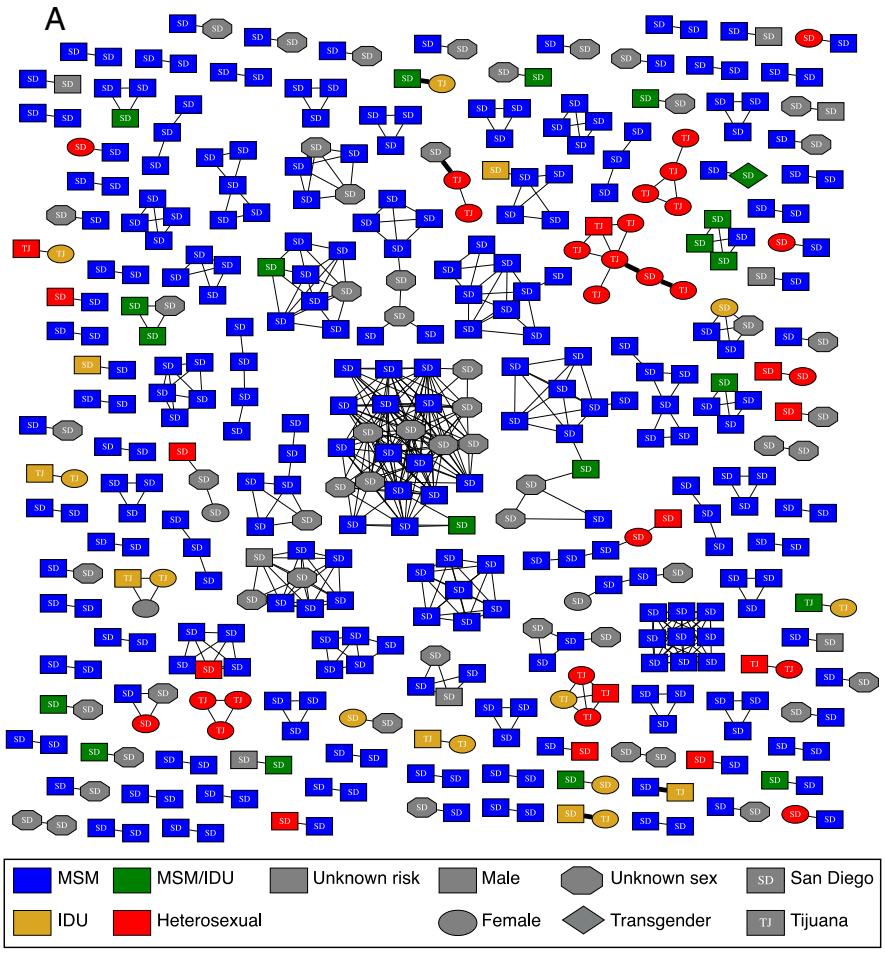
All sequences were screened for drug resistance mutations (DRM) utilizing the Calibrated Population Resistance tool (Gifford et al., 2009), and then evaluated in the context of the transmission cluster in which that individual was a member.

3. Results

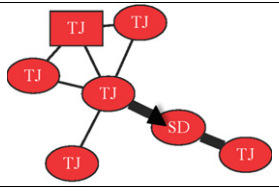
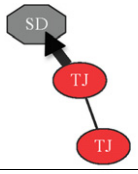
3.1. Study Population

A total of 843 eligible HIV-1 partial *pol* sequences were obtained from unique individuals participating in ten research studies from June 1996 through June 2013 (Table 1). Of the 843 sequences, 76 sequences were from individuals residing in Tijuana and 766 were

A



B

Cross-Border Clusters	Description	Estimated TMRCA for Cross Border Clusters*
	A predominantly female TJ heterosexual transmission cluster connected to a SD female	2000
	A cluster of two heterosexual females from TJ connected to a SD resident	2004
	A female IDU from TJ connected to a SD MSM/IDU	2006
	An SD MSM connected to a male IDU from TJ	2008
	A SD male IDU connected to a TJ female IDU	2010

***Timing of cross border clusters based on most recent common ancestor**

Table 3
Comparison of clustering and non-clustering individuals.

	Clustering	Non-clustering	p-Value
Country of Residence[*]			
Mexico	34 (44.7%)	42 (55.3%)	NS
US	348 (45.4%)	418 (54.6%)	
Race			
African American	19 (31.1%)	42 (68.9%)	NS
Native American	20 (33.9%)	39 (66.1%)	NS
Caucasian	248 (48.4%)	264 (51.6%)	0.03
Pacific Islander	9 (69.2%)	4 (30.8%)	0.09
Asian	8 (44.4%)	10 (55.6%)	NS
Mixed/other	9 (39.0%)	14 (61.0%)	NS
Ethnicity[*]			
Hispanic	113 (43.5%)	147 (56.5%)	NS
Non-Hispanic	158 (43.4%)	206 (56.6%)	
Sex[*]			
Male	327 (44.9%)	401 (55.1%)	NS
Female	38 (48.1%)	41 (51.9%)	
Sexual behavior			
Bisexual	14 (30.4%)	32 (69.6%)	0.04
MSM	287 (46.9%)	325 (53.1%)	NS
Heterosexual	58 (43.9%)	74 (56.1%)	NS
Injection drug use			
Yes	39 (33.3%)	78 (66.7%)	<0.01
No	306 (47.7%)	336 (52.3%)	
Research study			
SDPIC	328 (47.4%)	363 (52.6%)	0.01
San Diego South Bay	1 (9.0%)	10 (81.0%)	0.01
Tijuana based studies	39 (46.4%)	45 (53.6%)	NS
STAHR & STAHR-II	15 (26.3%)	42 (73.7.2%)	<0.01

[*] With a two-tailed alpha of 0.05, we had the power (0.8) to detect an absolute difference of 11.5% in the rate of clustering by ethnicity, 16.5% by residence, and 16% by sex.

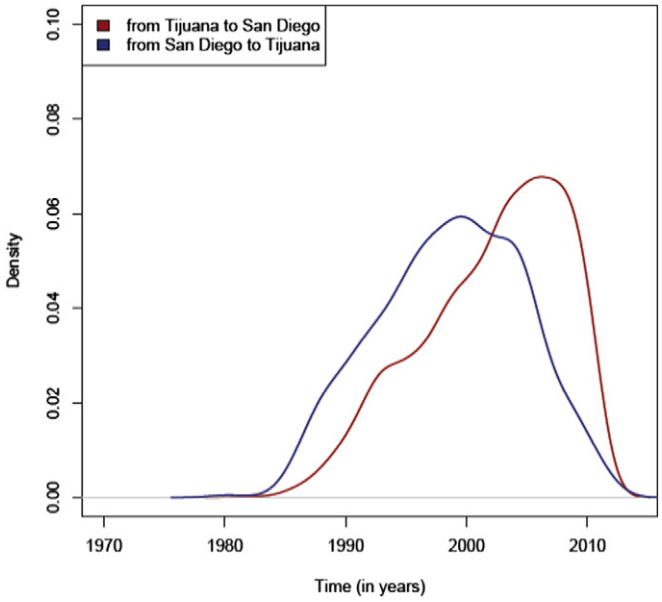


Fig. 2. Posterior probability density for the timing of HIV migration between San Diego and Tijuana. Using our available sequences, this plot extrapolates viral migration from the 1980s through 2014, and suggests that viral migration from San Diego to Tijuana was more likely during the 1990s, while viral migration from Tijuana to San Diego was more likely most recently, albeit there is considerable overlap in both directions.

from San Diego residents. Further, 702 sequences were from participants in the SD PIRC and most of these (64.7%) had acute or early HIV infection at the time of sampling (within six months of acquisition) (Morris et al., 2010). Socio-demographic data were available for 81.5% to 99.9% of the sample, depending upon the variable.

The study population was predominantly male (89.8%), and four individuals identified themselves as male-to-female transgender. Most participants reported their race as white (74.4%), and 41.6% reported their ethnicity as Hispanic. The mean age of participants was 35 years. More than three-quarters of participants were MSM. 15.4% of subjects were persons who inject drugs (PWID), and 61.2% of respondents who were PWID were also MSM. Overall, these demographics and risk behaviors were consistent with the HIV epidemic in San Diego County (County of San Diego HaHSA, 2012). All SD PIRC individuals were antiretroviral (ARV) naïve at the time of sampling, but the ARV status of other subjects was unconfirmed. Additional characteristics by study are listed in Table 2.

3.2. Network Cluster Inference and Correlates of Clustering

Putative transmission connections were inferred for 38.2% (423/1106) of the sequences sampled (Fig. 1). The likelihood that an individual clustered with at least one other individual did not differ by ethnicity, residence, or sex in our study cohort. Bisexuals were less likely to cluster than heterosexuals or MSM ($p = 0.043$), and individuals identifying as white ($p \leq 0.01$) were more likely to cluster than other races. Age was also significantly associated with clustering: the mean age of clustered individuals (32.8 years) was 3.5 years younger than non-clustered individuals (36.3 years) ($p < 0.001$) (Table 3). There was also a difference in clustering among the studies, likely related to differences in how each cohort was constituted, such as by HIV risk and sampling methods.

3.3. Cluster Analysis

Overall, 138 HIV putative transmission clusters were identified (ranging in size from 2 to 23 individuals), including 52 clusters that included three or more individuals (Fig. 1A). When we evaluated the socio-demographic characteristics of the 120 clusters that had those data available for more than one individual, we found five clusters composed predominantly of females, including two clusters in which only females were identified. Predominantly female clusters were mostly comprised of individuals living in Mexico. Concerning reported risk, nine clusters contained predominantly PWID, and these were significantly more likely to include bisexual males ($p = 0.0001$). Of the 120 clusters with sexual risk data on at least two individuals, 12 (10%) were predominantly comprised of heterosexuals, including six clusters in which the only reported risk factor was heterosexual risk. Interestingly, five heterosexual predominant clusters were comprised mainly of females, including one large cluster ($n = 7$) where 86% individuals were female. Of the remaining 108 clusters, 77 (71%) were predominately composed of MSM. The larger clusters did not differ

Fig. 1. (A) Inferred HIV Transmission clusters. HIV-1 transmission cluster diagrams illustrating the structure and demographics of the putative transmission clusters identified in the San Diego–Tijuana border region. Color indicates transmission risk factor, and shape denotes gender. All edges represent a genetic distance of ≤1.5% separating nodes. Edges connecting viruses from San Diego and Tijuana are shown in bold. (B) Estimated timing of cross border transmissions. A time-scaled phylogenetic tree was created using Bayesian inference implemented in the BEAST software package (Supplementary Fig. 2). This tree was used to infer the time of most recent common ancestor (TMRCA) of each inferred cross border transmission cluster. In this table, the network diagram of each of the clusters is illustrated in the column titled "Cross-Border Clusters". Within each of these clusters, putative cross border transmission events are highlighted with a bolded edge. Arrowheads represent directionality of these cross border transmission events where supported by the phylogenetic analysis. In the second column, the mean posterior probability of the TMRCA of the cluster as inferred from the time-scaled phylogenetic tree (Supplementary Fig. 2) is presented. TJ = Tijuana, SD = San Diego).

from the smaller clusters in terms of sex, sexual orientation risk, transactional sex, or country of residence, but trended towards having fewer individuals with syphilis ($p = 0.078$) and fewer PWID ($p = 0.047$) (Supplementary Table 5).

3.4. Compartmentalization and Viral Migration

Similar to our previous reports (Mehta et al., 2010), we found significant compartmentalization between the sampled HIV epidemics in San Diego and Tijuana (Fst $p < 0.001$). Nevertheless, of the 14 clusters that included individuals from Tijuana, five of these clusters included individuals who lived in San Diego. These international transmission clusters represent HIV transmission crossing the San Diego-Tijuana border. The four cross-border clusters with available demographic information had a significantly higher proportion of females ($p = 0.0001$), heterosexuals ($p = 0.0001$), and PWID ($p = 0.009$), as compared to clusters containing individuals only from either San Diego or Tijuana (Supplementary Table 6); differences were not seen with transactional sex or syphilis infection. Interestingly, in three of these four clusters, the members that were San Diego residents were enrolled in Tijuana.

Using phylogeographic methods, we then inferred the time of most recent common ancestor (TMRCA) for the cross border clusters. First, we created time-scale phylogenetic trees using sub-samples of our sequences using the BEAST software package (Drummond & Rambaut, 2007) (Supplementary Fig. 2). The mean estimated TMRCA of these clusters is reported in Fig. 1B, along with the structure of that particular cluster. Next, we used our time-scaled phylogenetic tree to reconstruct HIV migration patterns over time. Although limited by sampling, our results were consistent with historical observations that suggested earlier in the epidemic (1990s), most HIV infections in Tijuana were imported from San Diego (Valdespino-Gomez et al., 1995). Interestingly, our data suggest that there have been more viral migrations from Tijuana to San Diego more recently (2000s), but considerable overlap in the movement of these epidemics remains (Fig. 2).

3.5. Drug Resistance Analysis

Of all clustering individuals, 21.5% contained at least one member with a drug resistance mutation in their *pol* sequence, similar to the prevalence of drug resistance among the entire cohort (22.1%). All clustering remained even when codons associated with drug resistance mutations were stripped from the sequences. The presence of any drug resistance mutation was more common in members of San Diego only clusters ($p = 0.007$). A third of clusters included at least one individual with drug resistant virus. Twenty-five clusters (18.1%) had two or more individuals sharing one or more drug resistance mutations, accounting for 78.0% of the clustering individuals with resistance. The most common drug resistance mutation was K103N, which was represented in 11.5% of entire cohort and 10.9% of clustering individuals, and is similar to the prevalence of K103N seen in newly infected individuals in San Diego (Tilghman et al., 2014).

4. Discussion

Geography can either limit or facilitate the spread of infectious diseases, like HIV. The proximity and close cultural ties between populations in Mexico and the southwestern US, have linked these groups for many years. Thus, it is not surprising that in the early days of the HIV epidemic, nearly all identified HIV infections in Mexico were acquired in the US (Valdespino-Gomez et al., 1995). In this study, we explored the connectedness of the two cross-border epidemics over time and found, similar to our previous work (Mehta et al., 2010), which used a much smaller dataset and less detailed methods, that the two epidemics were mostly separate. In this new larger analysis we were able to better characterize viral migrations between San Diego and Tijuana, and we demonstrated that HIV may now also be flowing from Mexico to the US, likely because of mixing between high risk MSM (Ritieni et al., 2006) and PWID populations (Brouwer et al., 2009).

Given the estimated size of the HIV infected populations of San Diego (~12,000) (AIDSVu, 2014) and Tijuana (~1800–5500) (Brouwer et al., 2006), our bi-national dataset of 1106 unique viral sequences represented approximately 7% of infected individuals in the border region, and up to 12% of newly diagnosed infections in San Diego County (County of San Diego HaHSA, 2012). Although our analysis focused on the 843 sequence with associated demographic data collected from ten separate studies, each of which employed a different recruitment approach and was focused on a specific population (geographically and/or by risk factor), we still were able to obtain and include data across risk groups and geography. Although only 7.2% of our sequences were obtained from individuals residing in Tijuana, we identified 14 putative transmission clusters involving individuals from Tijuana and five clusters involving individuals reporting residence on both sides of the border. These cross-border clusters had higher proportions of females and heterosexuals than the rest of the transmission clusters, with >50% of members participating in transactional sex. This observation highlights the importance of the sex trade in HIV transmission across the region's border. Interestingly, two of the four cross border clusters included PWID from Tijuana and clients of female sex workers (FSW) from San Diego. Given that the focus of sampling in Tijuana was in high-risk FSW and PWID, this finding underscores the importance red light districts as a “melting pot” for risk groups, resulting in the bridging of different types of transmission networks.

Our phylogeographic analysis found that these viral migrations of HIV across the border stemmed from recently formed putative transmission clusters (TMRCAs between 2000 and 2010), suggesting ongoing cross-border transmission. Although the bidirectional migration we inferred is based on our focused sampling of high risk groups along the border in Tijuana, this observed viral flow also demonstrates the importance of transmission in this area to the predominantly MSM epidemic in central San Diego. The impact of policies and demographic shifts on this mixing needs to further elucidated, but it remains clear that the HIV epidemic on one side of the border remains important to at-risk communities on the other side.

As expected the majority of our putative transmission clusters were predominantly males who were MSM, since the majority of participants were from HIV research programs focused on the MSM population of San Diego. However, in transmission clusters comprised predominantly of heterosexuals, there was a predominance of females. This lack of males in putatively inferred transmission clusters has been demonstrated previously (Hue et al., 2014). Although injection drug use could lead to female–female transmission, unidentified infected male contacts of our sampled FSW are likely to be involved in these transmission clusters. Identification of these men will require different recruitment methods than those currently being used and underscores the importance of identifying and testing partners of high-risk individuals.

This cross-border molecular epidemiologic study has several limitations, the most important of which is sampling bias. Different methodologies were used to collect data across the different studies ranging from respondent driven sampling (Detection of HIV in Latinos, El Cuete, SD PIRC, and STAHR (Garfein et al., 2012)), passive enrollment (SD PIRC), venue based recruitment (SDPIC, Detection of HIV in Latinos, and STAHR-II (Robertson et al., 2014)), time location sampling (El Cuete and STAHR-II), partner identification (Amigos, SD PIRC), and convenience sampling (Hombre Seguro, Amigos, Proyecto Parejas, Mujer Segura and Mujer Más Segura, STAHR and STAHR-II). Variations in the depth of sampling across the region and collection of risk factor data also likely contributed to the study's overall sampling bias. A more comprehensive sample of the infected population in the border region would likely have resulted in more putatively identified transmission links, a higher rate of overall clustering, and a improved understanding of viral migration across the border. Despite this limitation, this study

was large and demonstrated linkage of transmission networks across different risk groups and geographically separate populations in the border region. Sampling across risk groups in the Tijuana border region also highlighted the role of such unique geographic areas in bridging transmissions across risk groups, and suggests the need for future research that is attentive to how social, sexual, and drug use networks interact with the built environment. Finally, this analysis highlights the potential usefulness of continuous molecular epidemiologic monitoring of HIV transmission networks, specifically by: 1) providing information about which sexual and PWID networks contribute disproportionately to new infections, and 2) identifying which important sociodemographic or risk groups might be missed in current identification strategies, resulting in improved targeting of prevention efforts.

Author Contributions

SRM: Conceived study, coordinated study, collected, interpreted, and analyzed data, and was the primary author on the manuscript.

JOW: Assisted with analysis and interpretation of data, and edited manuscript.

KCB: Provided data from subjects enrolled in Tijuana and edited manuscript.

KDW: Provided data from subjects enrolled in Tijuana and contributed to and edited manuscript.

AC: Assisted with analysis of data and generation of figures, and edited manuscript.

SS: Provided data from subjects enrolled in Tijuana and San Diego and contributed to and edited manuscript.

TLP: Provided data from subjects enrolled in Tijuana and San Diego and contributed to and edited manuscript.

MRG: Provided data from subjects enrolled in Tijuana and contributed to and edited manuscript.

MV: Performed DNA extraction and sequencing of HIV-1, and edited manuscript.

BM: Assisted with analysis of data and generation of figures, and edited manuscript.

RG: Provided data from subjects enrolled in San Diego and edited manuscript.

SJL: Provided data from subjects enrolled in San Diego and edited manuscript.

DMS: Conceived project, assisted with study coordination, data analysis and interpretation, and edited manuscript.

Declaration of Interests

SRM, JOW, KCB, KDW, AC, SAS, TLP, MGR, MV, BM, RG, SJL, declare no conflicts of interest. DMS has received grant support from ViiV and worked as a consultant for Hologic.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ebiom.2015.07.024.

Acknowledgments

This work was supported by grants from the National Institutes of Health: DA034978 (DMS), DA031074 (RG), AI93163 (SRM), AI43638 (SJL), AI106039 (SJL), AI100665 (DMS), AI110181 (JOW), AI74621 (SJL), DA031031 (KDW), DA019829 (SAS), DA027772 (SAS), DA023877 (SAS), DA029008 (TLP), MH065849 (TLP), and DA028692 (KCB), UCSD CFAR (AI036214); and the California HIV-1 Research Program (CHRP) RN07-SD-702, F13-SD-321 (AC). Additional support was provided by the Department of Veterans Affairs, James B. Pendleton Charitable Trust, and the Bettencourt-Schueller Foundation. Funders did not have any role in study design, data collection, data analysis, interpretation, or writing of the report.

References

AIDSVu, 2014. Persons Living with an HIV or AIDS Diagnosis, 2010. http://aidsvu.org (accessed 12/5/2014 2014).

Bronfman, M.N., Leyva, R., Negroni, M.J., Rueda, C.M., 2002. Mobile populations and HIV/AIDS in Central America and Mexico: research for action. AIDS 16 (Suppl. 3), S42–S49.

Brouwer, K.C., Strathdee, S.A., Magis-Rodriguez, C., et al., 2006. Estimated numbers of men and women infected with HIV/AIDS in Tijuana, Mexico. J. Urban Health 83 (2), 299–307.

Brouwer, K.C., Lozada, R., Cornelius, W.A., et al., 2009. Deportation along the U.S.–Mexico border: its relation to drug use patterns and accessing care. J. Immigr. Minor. Health 11 (1), 1–6.

County of San Diego HaHSA, 2012. HIV/AIDS Surveillance Program Epidemiology Report 2012. In: Health, P. (Ed.)County of San Diego, San Diego.

Dawood, F.S., Jain, S., Finelli, L., et al., 2009. Emergence of a novel swine-origin influenza A (H1N1) virus in humans. N. Engl. J. Med. 360 (25), 2605–2615.

Deiss, R.G., Brouwer, K.C., Loza, O., et al., 2008. High-risk sexual and drug using behaviors among male injection drug users who have sex with men in 2 Mexico–US border cities. Sex. Transm. Dis. 35 (3), 243–249.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.

Garfein, R.S., Rondinelli, A., Barnes, R.F., et al., 2012. HCV infection prevalence lower than expected among 18–40-year-old injection drug users in San Diego, CA. J. Urban Health 90 (3), 516–528.

Gifford, R.J., Liu, T.F., Rhee, S.Y., et al., 2009. The calibrated population resistance tool: standardized genotypic estimation of transmitted HIV-1 drug resistance. Bioinformatics 25 (9), 1197–1198.

Ghanem, A., Little, S.J., Drumright, L., Liu, L., Morris, S., Garfein, R.S., 2011. High-risk behaviors associated with injection drug use among recently HIV-infected men who have sex with men in San Diego, CA. AIDS Behav 15 (7), 1561–1569.

Hightower, G.K., May, S.J., Perez-Santiago, J., et al., 2013. HIV-1 clade B pol evolution following primary infection. PLoS One 8 (6), e68188.

Howlett, S.E., Castillo, H.S., Gioeni, L.J., Robertson, J.M., Donfack, J., 2013. Evaluation of DNA stable for DNA storage at ambient temperature. Forensic Sci. Int. Genet. 8 (1), 170–178.

Hue, S., Brown, A.E., Ragonnet-Cronin, M., et al., 2014. Phylogenetic analyses reveal HIV-1 infections between men misclassified as heterosexual transmissions. AIDS 28 (13), 1967–1975.

Hudson, R.R., Slatkin, M., Maddison, W.P., 1992. Estimation of levels of gene flow from DNA sequence data. Genetics 132 (2), 583–589.

Hudson, R.R., 2000. A new statistic for detecting genetic differentiation. Genetics 155 (4), 2011–2014.

Kosakovsky Pond, S.L., Posada, D., Stawiski, E., et al., 2009. An evolutionary model-based algorithm for accurate phylogenetic breakpoint mapping and subtype prediction in HIV-1. PLoS Comput. Biol. 5 (11), e1000581.

Kuiken, C., Korber, B., Shafer, R.W., 2003. HIV sequence databases. AIDS Rev. 5 (1), 52–61.

Latino, F.N., 2014. As Crime Drops, Americans Begin To Return To Mexico's Baja. Fox News, Fox News Latino, p. 1.

Le, T., Wright, E.J., Smith, D.M., et al., 2013. Enhanced CD4 + T-cell recovery with earlier HIV-1 antiretroviral therapy. N Engl J Med 368 (3), 218–230.

Mehta, S.R., Delport, W., Brouwer, K.C., et al., 2010. The relatedness of HIV epidemics in the United States–Mexico border region. AIDS Res. Hum. Retrovir. 26 (12), 1273–1277.

MMWR, 2010. Prevalence and awareness of hiv infection among men who have sex with men — 21 cities, United States, 2008. MMWR Morb Mortal Wkly Rep 59 (37), 1201–1207.

Morris, S.R., Little, S.J., Cunningham, T., Garfein, R.S., Richman, D.D., Smith, D.M., 2010. Evaluation of an HIV nucleic acid testing program with automated Internet and voicemail systems to deliver results. Ann Intern Med 152 (12), 778–785.

Patterson, T.L., Goldenberg, S., Gallardo, M., et al., 2009. Correlates of HIV, sexually transmitted infections, and associated high-risk behaviors among male clients of female sex workers in Tijuana, Mexico. AIDS 23 (13), 1765–1771.

Pitpitan, E.V., Goodman, D., Burgos, J.L., et al., 2015. Prevalence and correlates of HIV among men who have sex with men in Tijuana, Mexico. J. Int. AIDS Soc. 18 (1), 19304.

RITA, 2012. Research and Innovative Technology Administration: Bureau of Transportation Statistics. http://transborder.bts.gov/programs/international/transborder/TBDR_BC/TBDR_BCQ.html (accessed November 5, 2013 2013).

Ritieni, A., Facer, M., Ruiz, J., Meneses-Imay, M.C., Magis Rodriguez, C., Monitor, F., 2006. Prevalence of HIV infection and related risk behaviors among young Latino men who have sex with men: San Diego–Tijuana border region.

Robertson, A.M., Garfein, R.S., Wagner, K.D., et al., 2014. Evaluating the impact of Mexico's drug policy reforms on people who inject drugs in Tijuana, B.C., Mexico, and San Diego, CA, United States: a binational mixed methods research agenda. Harm Reduct. J. 11, 4.

Rose, P.P., Korber, B.T., 2000. Detecting hypermutations in viral sequences with an emphasis on G → A hypermutation. Bioinformatics 16 (4), 400–401.

Servin, A.E., Munoz, F.A., Strathdee, S.A., Kozo, J., Zuniga, M.L., 2012. Choosing sides: HIV health care practices among shared populations of HIV-positive Latinos living near the US-Mexico border. J Int Assoc Physicians AIDS Care (Chic) 11 (6), 348–355.

Sirotin, N., Strathdee, S.A., Lozada, R., et al., 2010. A comparison of registered and unregistered female sex workers in Tijuana, Mexico. Public Health Rep; 125 Suppl 4 101–109.

Slatkin, M., Maddison, W.P., 1989. A cladistic measure of gene flow inferred from the phylogenies of alleles. Genetics 123 (3), 603–613.
Smith, D.M., May, S.J., Tweeten, S., et al., 2009. A public health model for the molecular surveillance of HIV transmission in San Diego, California. AIDS 23 (2), 225–232.
Syvertsen, J.L., Robertson, A.M., Strathdee, S.A., Martinez, G., Rangel, M.G., Wagner, K.D., 2014. Rethinking risk: Gender and injection drug-related HIV risk among female sex workers and their non-commercial partners along the Mexico-U.S. border. Int J Drug Policy 25 (5), 836–844.
Tilghman, M.W., Perez-Santiago, J., Osorio, G., et al., 2014. Community HIV-1 drug resistance is associated with transmitted drug resistance. HIV Med 15 (6), 339–346.
Valdespino-Gomez, J.L., Garcia-Garcia Mde, L., del Rio-Zolezzi, A., Loo-Mendez, E., Magis-Rodriguez, C., Salcedo-Alvarez, R.A., 1995. The epidemiology of AIDS/HIV in Mexico: from 1983 to March 1995. Salud Publica Mex. 37 (6), 556–571.
Wagner, K.D., Pitpitan, E.V., Chavarin, C.V., Magis-Rodriguez, C., Patterson, T.L., 2013. Drug-using male clients of female sex workers who report being paid for sex: HIV/sexually transmitted infection, demographic, and drug use correlates. Sex Transm Dis 40 (8), 619–623.
Wertheim, J.O., Leigh Brown, A.J., Hepler, N.L., et al., 2014. The Global Transmission Network of HIV-1. J. Infect. Dis 209 (10), 1642–1652.