HIV Transmission Networks Among the USA, Mexico and Central America

Santiago Avila-Ríos[1*], Joel O. Wertheim[2*], Ann Dennis[8*], Gustavo Reyes-Terán[1], Carlos Mejía-Villatoro[3], Elsa Palou[4], Guillermo Porras-Cortés[5], Juan M Pascale[6], Marvin Manzanero[7], Sanjay R. Mehta[9*]

1 Center for Research in Infectious Diseases, National Institute of Respiratory Diseases, Mexico City, Mexico; 2 Department of Medicine, University of California San Diego, San Diego, CA, USA; 3 Infectious Diseases Clinic, Roosevelt Hospital, Guatemala City, Guatemala; 4 Hospital Escuela Universitario, Tegucigalpa, Honduras; 5 Gorgas Memorial Institute for Health Studies, Panama City, Panama; 6 Vivian Pellas Metropolitan Hospital, Managua, Nicaragua; 7 Ministry of Health, Belmopan, Belize; 8 University of North Carolina, Chapel Hill, NC, USA; 9 Division of Infectious Diseases, University of California San Diego; San Diego, CA, USA.

*santiago.avila@cieni.org.mx; jwertheim@ucsd.edu; adennis@med.unc.edu; srmehta@ucsd.edu

242

Background

- The role of migration in shaping genetic diversity of viral populations has been well established.
- We show results of a large collaborative effort to assess epidemiological networks and dispersion routes between the HIV epidemics in the United States, Mexico and Central America, including Guatemala, Belize, Honduras, El Salvador, Nicaragua and Panama.

Methods

- We analyzed 12,595 HIV-1 pol sequences collected from Mexican and Central American national HIV sequence databases and local databases located at the University of North Carolina and the University of California, San Diego.
- Sequences were aligned to the reference sequence HXB2. We then estimated Tamura-Nei (TN) 93 pairwise genetic distances among all sequences.
- A putative transmission linkage was inferred when two sequences were found to be ≤1.5% divergent calculated using the TN93 distance.
- We then constructed transmission clusters by linking sequences through shared potential transmission partners.
- Clustering information and demographic information were plotted geographically using ArcGIS (ESRI, Redlands, CA).
- Maximum clade credibility phylogeny of selected clusters was inferred using the BEAST software package (V1.8.1).
- Differences between groups were analyzed using the Pearson's X^2 test for categorical variables, and the Kruskal-Wallis Test for continuous variables. We used logistic regression models to evaluate characteristics associated with cluster membership.

Results

- In total, 2,866 (23%) sequences had an inferred potential transmission partner in the network, comprising 1,031 distinct transmission clusters.
- Clusters were predominately observed within a single country (986 clusters, 95.6%).
- The majority of clusters were located within highly represented countries (USA 39.2%, Mexico 33.8%).
- Most clusters were pairs (71.2%) while only four clusters included more than 20 members.
- Only 45 (4.5%) clusters included sequences isolated from at least two individuals in different countries.
- Of sequences in clusters, 878 (31%) were linked to a sequence sampled from another city/state, with highest rates of these "trans-regional" links in Guatemala (47%) and Mexico (43%), compared to the US (17%) [P<0.001]
- Fourteen clusters included linkages between sequences isolated in San Diego and in Mexico (mainly Tijuana).
- Factors associated with cluster membership, both in bivariate and multivariate models, included residence in Mexico or Guatemala, MSM risk factor, younger age, higher CD4 T cell counts, higher HIV VL and more recent sampling date.
- Using a multivariate model, factors associated with membership in a transnational cluster included residence in Mexico, younger age, more recent sampling year and higher HIV VL.
- Similar findings were seen for trans-regional clusters.
- Risk factors for HIV acquisition were mostly heterosexual risks in Guatemala, Honduras, Nicaragua and Panama.
- MSM transmission predominated in most Mexican states (mainly Veracruz, Tabasco, Campeche, Yucatán, Michoacán, Mexico City and Puebla) and at USA sites.

Table 1. Description of Study Population

Variable	N (%)
Country	
USA	6,563 (52)
Mexico	3,292 (26)
Central America	
Guatemala	1,177 (9)
Honduras	700 (6)
Panama	363 (3)
El Salvador	163 (1)
Belize	97 (<1)
Sex	
Male	9,497 (77)
Female	2,763 (23)
Risk Behavior	
MSM	4,488 (36)
Heterosexual	4,130 (33)
IDU	627 (5)
MSM/IDU	537 (4)
Other	204 (2)
Unknown/Missing	2,609 (21)
Age, median years (IQR)	37 (29-46)
CD4 Lymphocyte count, median (IQR)	267 (97-466)
Log RNA Viral load, median (IQR)	4.7 (3.9-5.3)
Sample Year*, median (IQR)	2009 (2007-2013)

*Missing for 1024

Table 2. Number of clusters by country

Country	Frequency	Percent
USA	404	39.2
Mexico	348	33.8
Guatemala	134	13.0
Mixed*	45	4.4
Honduras	36	3.5
Nicaragua	23	2.2
Panama	19	1.8
El Salvador	16	1.6
Belize	6	0.6
Total	1031	100

*More than one country represented

Table 3. Cluster size distribution

Cluster size	Frequency	Percent
2	734	71.2
3	137	14.3
4	66	6.4
5	28	2.7
6	16	1.6
7	9	0.9
8	5	0.5
9	9	0.9
10	4	0.4
11	4	0.4
12	2	0.2
13	2	0.2
16	1	0.1
24	1	0.1
25	1	0.1
29	1	0.1
45	1	0.1
Total	1031	100

Table 4. Bivariable associations with clustering

Variable	Not in cluster	In any cluster		In a transnational cluster	
		n (%)	OR (95% CI)	n (%)	OR (95% CI)
Region of sampling [n (%)]					
Guatemala	755 (64)	422 (46)	**2.6 (2.2-3.1)***	12 (1)	0.8 (0.4-1.6)
Mexico	2,310 (70)	982 (30)	**2.0 (1.7-2.3)***	110 (3)	**2.7 (1.6-4.3)***
USA	5,377 (82)	1,186 (18)	1.0 (0.9-1.2)	40 (<1)	0.5 (0.3-0.8)
Other	1,287 (82)	276 (17)	Ref	20 (1)	Ref
Sex and risk factor [n (%)]					
Male – MSM	3,698 (75)	1,231 (25)	**1.3 (1.2-1.5)***	51 (1)	0.7 (0.4-1.0)
Male – Other	2,198 (80)	541 (20)	1.0 (0.9-1.1)	39 (1.)	0.9 (0.6-1.4)
Female	2,209 (80)	554 (20)	Ref	43 (2)	Ref
Age [years; median (IQR)]	38 (30-47)	31 (25-40)	**0.6 (0.59-0.65)***	31 (24-40)	0.7 (0.6-0.8)*
CD4 T cell count [cells/µL; median (IQR)]	241 (84-453)	323 (145-500)	**1.08 (1.06-1.1)***	297 (128-465)	1.0 (0.9-1.1)
RNA Viral Load [log copies/mL; median (IQR)]	4.6 (3.8-5.2)	4.8 (4.2-5.4)	**1.3 (1.2-1.3)***	5.0 (4.2-5.5)	**1.3 (1.1-1.5)***
Sample Year [median (IQR)]	2009 (2006-2012)	2010 (2008-2013)	**1.1 (1.05-1.07)***	2012 (2010-2013)	**1.2 (1.1-1.2)***

*P<0.001 in bivariable model; *Per 10 years increase; *Per 100 cells/mL increase; **BOLD** indicates variables significant in multivariable model.
In a multivariable model outcome of membership in any clusters, all the variables significant in the bivariable model remain significant. Factors associated with cluster membership included residence in Mexico or Guatemala, MSM risk factor, younger age, high CD4 count, higher HIV RNA, and more recent sampling date. In the multivariable model outcome of membership in a transnational cluster (compared to individuals not in a transnational cluster), sampling in Mexico (OR 2.95), more recent sampling year, and higher HIV RNA remain significant.

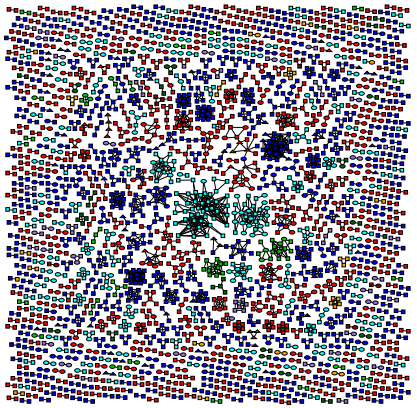
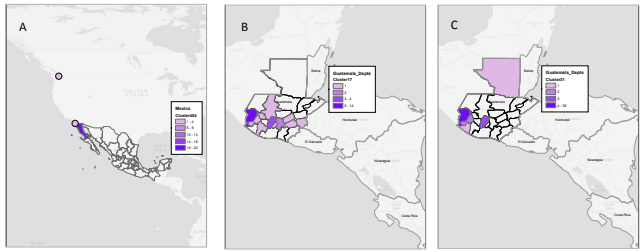


Figure 1. Cluster analysis for 12,595 *pol* sequences from the USA, Mexico and Central America. Tamura-Nei (TN) 93 pairwise genetic distances were calculated between all sequences and linkages were inferred between those ≤1.5% apart. Multiple linkages were resolved into clusters. 1,031 clusters were identified. Nodes are colored by country of origin. Squares=males; ovals=females; triangles=unreported gender.



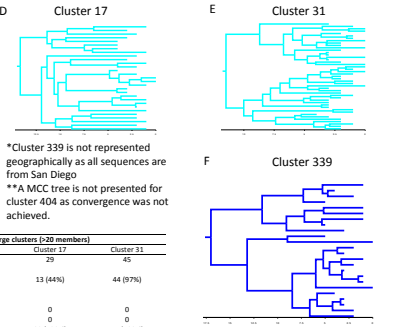


Figure 3. Geographic distribution and MCC trees of the largest inferred transmission clusters. A. Cluster 404 (n=25); B. Cluster 17 (n=29); C. Cluster 31 (n=45). Cluster 404 represents a USA-Mexico transnational cluster and Clusters 17 and 31 represent Guatemalan national clusters. Below maps are MCC trees of clusters with diversification timeline. D. Cluster 17 (n=29); E. Cluster 31 (n=45); F. Cluster 339 (n=24), USA is shown separately. Timeline in years.

*Cluster 339 is not represented geographically as all sequences are from San Diego
**A MCC tree is not presented for cluster 404 as convergence was not achieved.

Table 5. Description of Large Clusters

Variable	Characteristics of large clusters (>20 members)			
	Cluster 339	Cluster 404	Cluster 17	Cluster 31
Number of members	24	25	29	45
Male, n (%)	16 (67%)	11 (44%)	13 (44%)	44 (97%)
Country, n (%)				
USA	24 (100%)	2 (8%)*	0	0
Mexico	0	23 (92%)**	0	0
Guatemala	0	0	29 (100%)	45 (100%)

* 1 Latina sampled in San Diego, and 1 Latino male sampled at UW.
**22 sampled in Baja California (mainly Tijuana), 1 in Distrito Federal.

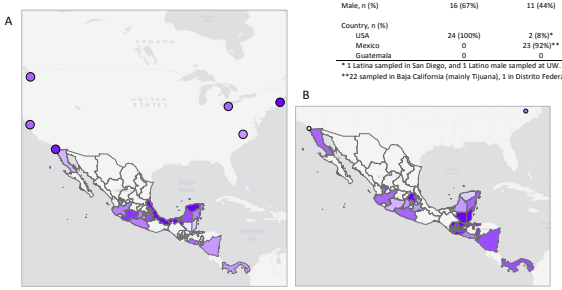


Figure 2. HIV risk factor maps for A) MSM and B) Heterosexual transmission. HIV risk factor was normalized to sampling. Colors represent 10% intervals in proportionate representation from 10% to 100% where the lightest shade of purple represents 0-10% of sampled individuals having the risk factor, and the darkest shade of purple represents 90-100% of sampled individuals having the risk factor. Grey areas represent unsampled districts/states or countries.

Conclusions

- Our work underlines the importance of political and cultural barriers in HIV transmission among the USA, Mexico and Central America, with most transmission events occurring within countries.
- Nevertheless, the presence of international connections was observed at rates (4%) similar to other published work. These viral migrations may be important in influencing the spread and diversity of HIV in the region.

Funding

This work was supported by grants from the Mexican Government (Comisión de Equidad y Género de las Legislaturas LX-LXI y Comisión de Igualdad de Genero de la Legislatura LXII de la H. Cámara de Diputados de la República Mexicana), CONACYT (SALUD-2013-01-202475), and the US National Institutes of Health AI093163 (SRM), AI110181 (JOW), AI112432 (AMD).