Poster# 246
Session# P-B3
22nd Conference on Retroviruses and Opportunistic Infections
February 23-26, 2015
Seattle, Washington

HIV Transmission Network Structure Reveals Characteristics of Bridging Individuals

Sanjay R. Mehta[1,2], Joel O. Wertheim[1], Konrad Scheffler[1], Susan Little[1], Richard Garfein[1], Sergei Kosakovsky Pond[1], and Davey M. Smith[1,2]

[1]University of California San Diego, La Jolla, California, USA, and [2]Veterans Administration San Diego Healthcare System, San Diego, California, USA

Sanjay Mehta, M.D.
Department of Medicine
University of California, San Diego
9500 Gilman Drive
San Diego, CA 92093-8208, USA
E-mail: srmehta@ucsd.edu
Phone: (619) 543-3150
Fax: (619) 543-5094

Background

- HIV transmission networks can be inferred by linking nodes (individuals) that have highly similar genetic sequences via an edge (putative transmission link).
- The characteristics of members and the structure of their networks can be used to better understand HIV local transmission dynamics.
- Some individuals (nodes) who are located centrally within networks can act as 'bridges' to link previously disconnected parts of the network.
- These 'bridging' individuals can have a disproportionate impact on the spread of an epidemic.

Objective

To characterize individuals who serve as 'bridges' within the San Diego HIV transmission network.

Methods

- We inferred a putative molecular transmission network from a annotated collection of 1024 partial *pol* sequences, representing 713 San Diego Primary Infection Resource Consortium (SDPIRC) participants sampled between 1996 and 2013, and 311 sequences from chronically infected individuals residing locally.
- Two individuals (nodes) were linked if their sequences were <1.5% genetically distant (TN93 evolutionary model). Multiple connections were resolved into transmission clusters.
- Three measures of centrality were measured within each cluster for each node (**Figure 1**): a) Network degree, b) Mean path length, and c) Betweenness centrality. We then analyzed relationships between sociodemographic factors and network properties.
- These measures were used to identify "bridging" individuals, who by virtue of a central position in a cluster, linked together two sub-clusters **(Figure 2)**.
- We then created a "uniqueness" score **(Table 1)** from 5 demographic attributes (age, race, ethnicity, HIV risk factor, and location of residence). This score was used to identify the most 'unique' person(s) in each cluster. We then analyzed this "uniqueness score" in relation to measures of centrality.

Figure 1. Definitions of Centrality Measures

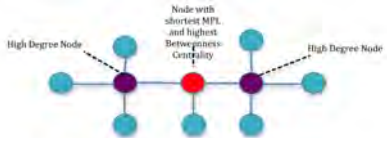


Figure 1: Example Network (cluster). Purple nodes have the highest Network Degree (4 edges). The red node has both the shortest Mean Path Length (MPL) to all other nodes in the cluster (i.e. fewest edges to each other node). The red node also has the highest Betweeness Centrality as 69% of paths from one node to another have to pass through the red node. Blue nodes are not central in this cluster.

Figure 2. Definition of Bridging



Figure 2: Bridging. Example A shows two unlinked transmission clusters. In example B, the addition of the red node has bridged (linked) these two separate clusters into one. Note that the red node has the shortest MPL and highest Betweenness centrality of all nodes in cluster B.

Table 1. Components of Uniqueness Score

Attribute	Uniqueness score of 1 if:
Race, Ethnicity, HIV Risk Factor	Individual in a cluster is different from ≥2/3 of cluster members.
Age	Individual's age is more than one standard deviation from mean age of the overall cluster.
Zip Code of Residence	Individual's zip code is entirely outside the area defined as one standard deviation away from the midpoint of the cluster (measured using zip code centroids) (see Figure 3).

Table 1: These five socio-demographic characteristics were used to determine the most unique individual within each transmission sub-network. Sex was not considered, since all clustering subjects were male.(Centroid=geographic midpoint)

Figure 3. Defining a Unique Zip Code

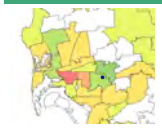


Figure 3: The blue star represents the geographic midpoint of a cluster. The circle is the area in which two-thirds of cluster members reside (one standard deviation). A zip code of residence is only considered unique when it lies entirely outside this circle (eg. red zip code, but not orange zip code).

Table 2. Uniqueness Score is Associated with Centrality

Centrality Measures	Individuals with Highest Uniqueness Score	Other Cluster Members	P value (Chi-Sq/ Mann-Whitney Test)
Shortest Mean Path Length	12/27 (44.4%)	9/66 (13.6%)	0.001
Highest Degree	13/27 (48.1%)	40/66 (60.6%)	0.271
>10% Betweenness Centrality	11/27 (40.7%)	7/66 (10.6%)	0.001

Table 2: Individuals identified as most central by shortest Mean Path Length, and those having Betweenness Centrality >10% were significantly more likely to have the highest uniqueness score within that cluster. In other words the most unique individuals in a cluster were more likely to be the most central in their cluster.

Results

Figure 4. The inferred San Diego HIV Transmission Network

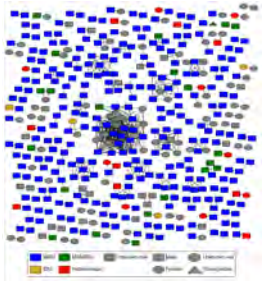


Figure 4. Map of inferred HIV-1 transmission clusters within the sampled population coded by risk factor (color), and sex (shape).

Younger Individuals are More Likely to Cluster

- 431 (42.1%) of all sequences clustered, including 356 (50%) of sequences from the SDPIRC
- Clustering individuals did not differ by race, ethnicity, HIV risk factor or geography of residence from non-clustering individuals, but were younger (32.5 +/- 8.9 vs. 35 +/-10.1 years, p=0.01, t-test).

No Associations between Socio-demographic Factors and Network Centrality

- 203 members of the transmission clusters with >2 nodes were analyzed.
- Network centrality calculations of all nodes identified 118 individuals as 'central' by mean path length, 18 individuals with a Betweenness Centrality of >0.10%, and 83 individuals as high degree nodes (>3 connections) in their respective clusters.
- No associations were found between any of the three measures of centrality and race, ethnicity, sex, HIV risk factor, injection drug use, methamphetamine use, or sexually transmitted infections.

Conclusions

- In a cohort made up predominantly of individuals with recent HIV infection, ~50% had virus genetically related to a virus from another individual in the network **(Figure 4)**
- Socio-demographic factors (age, HIV risk factors, race, ethnicity, and sexually transmitted infections) were not associated with centrality in the transmission network as measured by degree distribution, mean path length, and Betweenness Centrality.
- Individuals who were most 'unique' within their cluster, and therefore involved in disassortative transmission links, were more likely to be most central by mean path length and betweenness measurements **(Table 2).** In other words, disassortative mixing from individuals who are 'unique' in a cluster was most likely the reason for the bridging of sub-networks.

Acknowledgments
This work was supported by grants from the National Institutes of Health: DA034978, DA031074, AI69432, AI96113, AI93163, AI43638, AI90970, AI96189, AI67039, AI106039, AI100665, AI80353, MH97520, MH62512, MH083552, AI36214, TW008908, AI69432, AI47745, AI74621, GM093939, UCSD CFAR (AI306214), Department of Veterans Affairs, James B. Pendleton Charitable Trust, and the California HIV-1 Research Program (CHRP) RN07-SD-702.