HPTN 061 Research Proposal

1. Rationale:

HPTN 061 enrolled 1,553 Black MSM in six U.S. cities. Key study findings included a high prevalence of prior and incident incarceration among the study participants, and the association of prior incarceration with substance use, HIV risk behaviors, and poverty (1, 2).


Main questions:

Do the HIV sequences of men who share experiences (e.g., involvement with criminal justice systems) or behavioral characteristics (e.g., IV drug use) cluster in specific phylogenetic analyses?

Are HIV sequences from men who share these experiences or behavioral characteristics close in genetic distance (without regard to phylogeny)?

Importance and filling gaps in knowledge:

Genetic clustering implies the presence of circulating viral variants among subnetworks of transmission networks—although precise transmission chains are not identifiable from such data. Knowledge of the presence of transmission subnetworks within larger networks can potentially aid in targeting of prevention resources.

Much work has been done on genetic clustering of HIV infections, but to our knowledge, no one has looked at the role of criminal justice involvement, or other behavioral risk factors.

2. Objective:

Evaluation of the degree of HIV genetic clustering associated with subject involvement in the criminal justice system or shared behavioral risks.

3. Hypothesis:

Men in HPTN 061 who have current or prior involvement in the criminal justice system or who engage in similar behaviors such as injection drug use will tend to have HIV sequences that cluster genetically as observed through phylogenetic analyses (which are based on reconstructing lineages) and through analyses of genetic distance (which are based only on pairwise distance of sequences from pairs of subjects).
4. Product:

Report and presentation of project for Harvard School of Public Health Biostat doctoral student, Octavious Talbot. Focus is on identification of a thesis topic for development of new methods in this research area, but there is a possibility of publication or other presentation of substantive findings, should any arise from this research.

5. Lead:

Octavious Talbot, PhD candidate. Others include Susan Eshleman, Iris Chen, Ken Mayer, Victor DeGruttola. Other members of the HPTN Laboratory Center (LC) and HPTN 061 study team will be included as authors if the work results in publication or presentation at a meeting.

6. Perspectives of community:

If approved to move forward, the authors will solicit co-authors from among the HPTN 061 study team and/or HPTN Black Caucus who can ensure that the perspectives of the Black MSM community are considered in the analysis, interpretation and presentation of this research.”

7. Approach

This work will be performed using HIV pol sequences from men in HPTN 061. This includes 169 sequences from men who were HIV infected at study enrollment and 23 sequences from men who seroconverted during the study. These sequences were generated at the HPTN LC using the ViroSeq HIV-1 Genotyping System. Additional sequencing (of other HIV regions or samples collected at end of study) may be performed at the HPTN LC if HPTN LC CORE funding is available (no additional funding is requested for laboratory work for this ancillary study).

To assess the effect of involvement in criminal justice (CJ) systems or other behavioral risk factors on viral genetic linkage, we will begin by using APE (a package in R) to obtain pairwise genetic distances between all pairs of sequences. An estimate of the overall proportion of clustering for an individual (e.g., the proportion of sequences that link to at least one other sequence), will be calculated using the pairwise distance matrix of translated amino acid residues from phylogenetic analysis. This distance quantifies the degree of similarity between any two sequences. Several methods exist for calculating the distance matrix, including Jukes-Cantor, Dayhoff, and JTT, among others. These methods assume different relationships among transition probabilities and different equilibrium frequencies. We will investigate the effect of choice of distance measure on the extent of clustering.

To investigate whether subjects with a history of involvement in CJ systems (who who have other common other behavioral risk factors) are more likely to be clustered with each other, we will take two different approaches. The first defines a genetic cluster on the basis of a threshold, below which sequences are considered to be clustered. To be specific, let $D_{ij}$ indicates the pairwise distance between individuals $i$ and $j$, where $N$ is the total number of individuals in the dataset. If there is at least one $j$ for which $D_{ij}$ is less than the threshold, then individual $i$ is considered to be in a cluster. We can also investigate whether that individual $i$ is in a cluster that includes someone who has CJ history. We
can definite a test statistic as the proportion of people within any group who cluster with at least one
other person in that or a different group. Then we can do a permutation test to assess whether group
members are more often clustered with each other than would arise by chance.