

# Identifying HIV-1 Transmission Clusters and Presence of Drug Resistance within Washington, DC

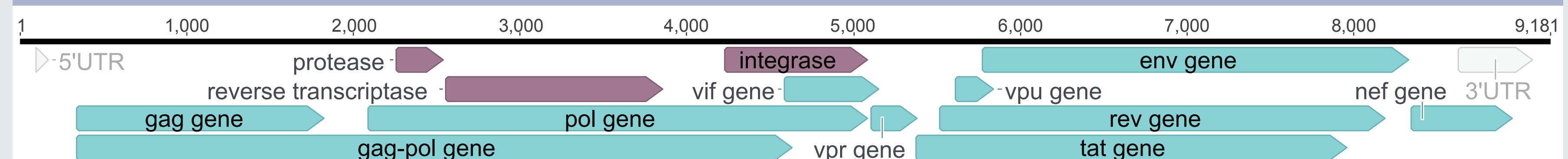
Margaret Steiner<sup>1</sup>, Keylie M Gibson<sup>1</sup>, Seble Kassaye<sup>2</sup>, Frank Maldarelli<sup>3</sup>, Marcos Pérez-Losada<sup>1</sup>, and Keith A Crandall<sup>1</sup>

<sup>1</sup>Computational Biology Institute, The Milken Institute of Public Health, The George Washington University, Washington, DC, USA 20052; <sup>2</sup>Department of Medicine, Georgetown University, Washington, DC, US 20057; <sup>3</sup>HIV DRP Host-Virus Interaction Branch, National Institute of Health, Bethesda, MD, USA 20892

## Background

- DC has the highest HIV prevalence in the US.
- To target HIV intervention, we propose to identify transmission clusters via HIV sequencing.
- By mapping epidemiological variables (sex, race, risk factor, geography) on estimated transmission clusters, we propose to better target intervention strategies.
- By identifying drug resistance variants, HIV infection can be better treated.

## The HIV Genome



## Objective

- To use viral sequence to identify drug resistant variants and transmission clusters in addition to epidemiological associates to better treat and prevent the spread of HIV in Washington DC.

## Phylogenetic and Transmission Cluster Analyses

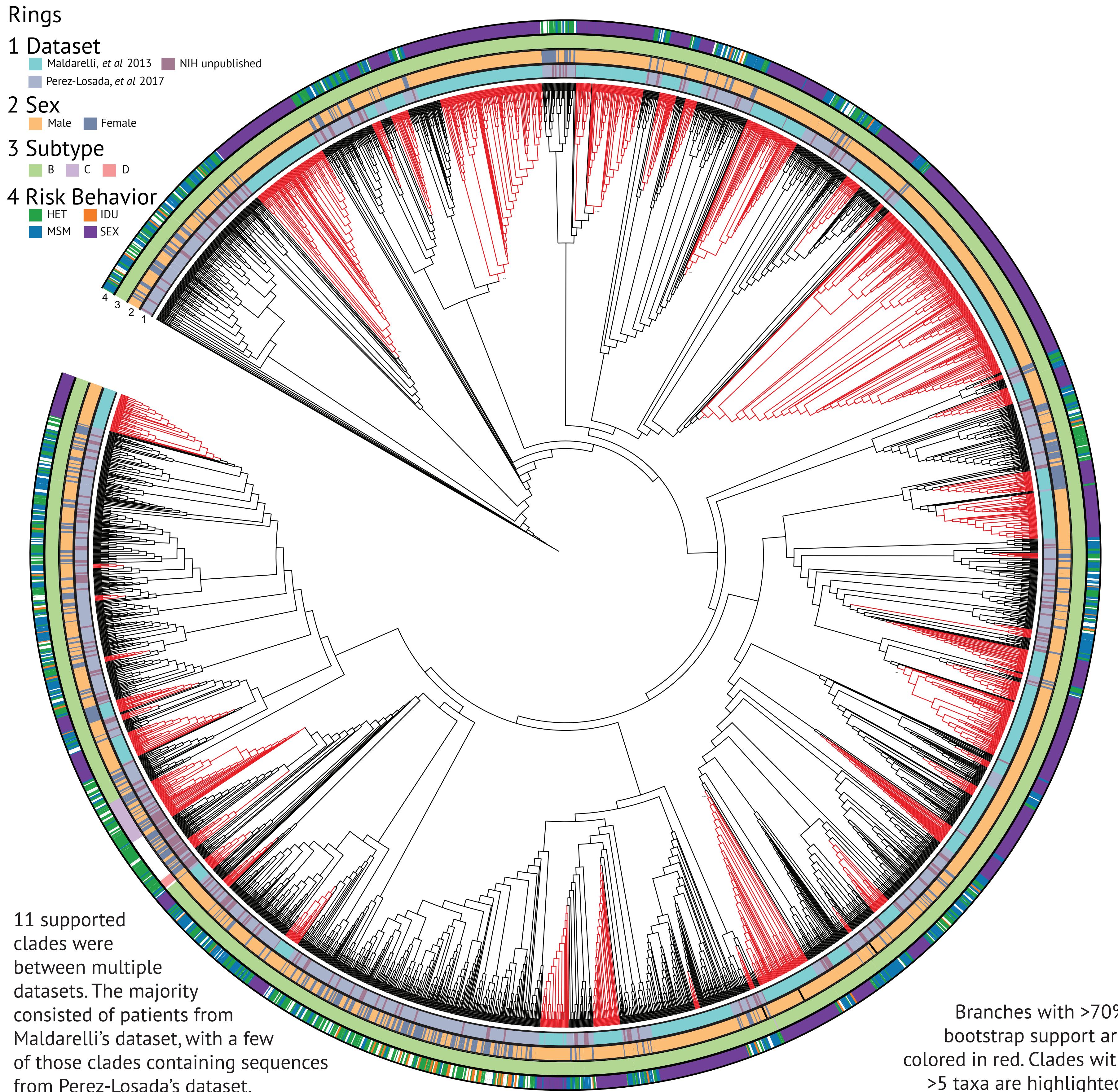
### Rings

1 Dataset  
Maldarelli, et al 2013 NIH unpublished  
Perez-Losada, et al 2017

2 Sex  
Male Female

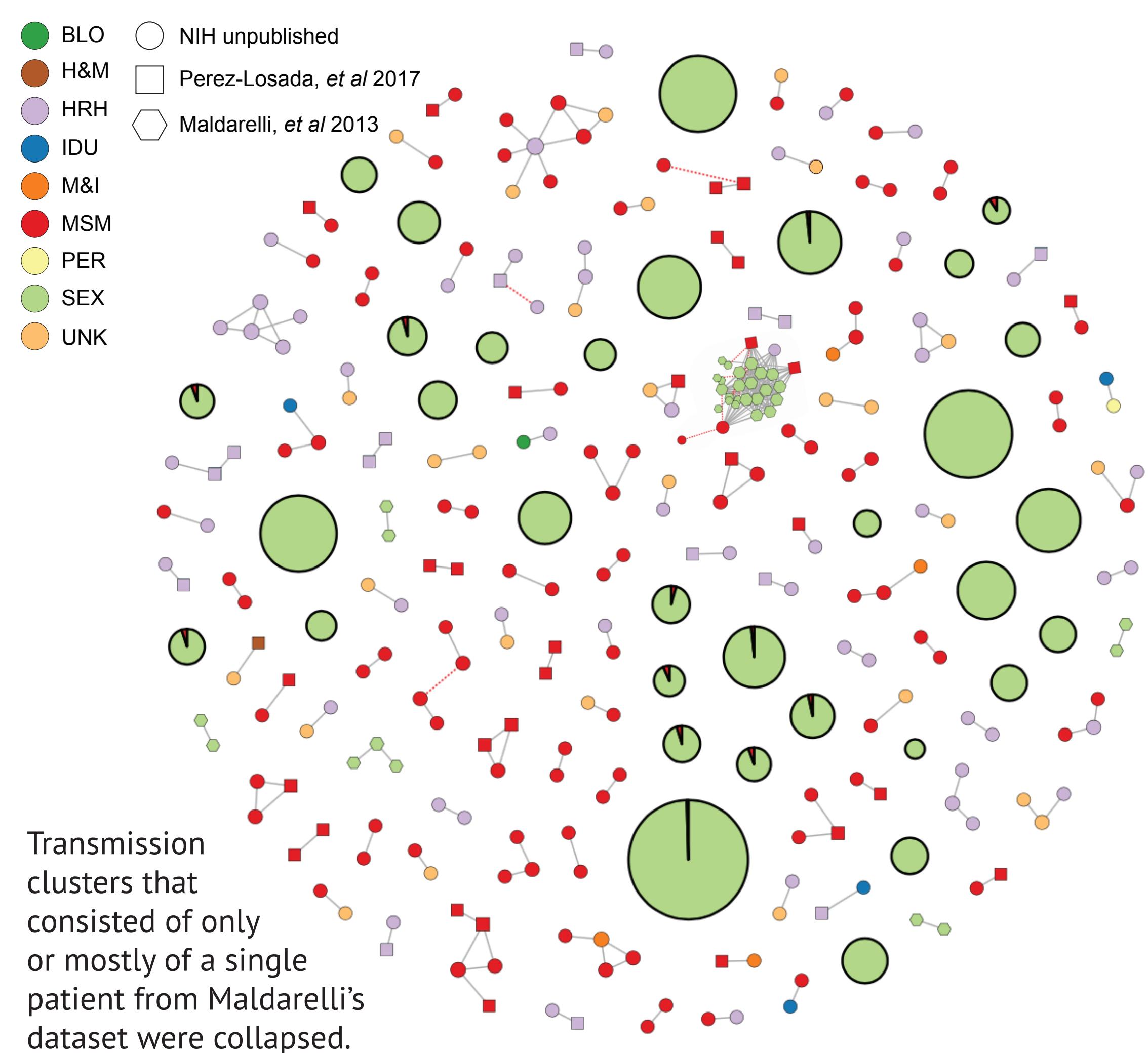
3 Subtype  
B C D

4 Risk Behavior  
HET IDU  
MSM SEX



## Methods

- Collected ~3,500 HIV sequences from three separate datasets.
- Focused specifically on the pol gene.
- Phenotypic associations: sex, risk factor, viral load, subtype, and anti-retroviral regimen type
- Aligned all sequences using MAFFT (Multiple Alignment using Fast Fourier Transform).
- Constructed phylogenetic trees using both maximum likelihood (RAxML) and Bayesian inference (MrBayes).
- Transmission clusters were estimated with HIV-TRACE.
- Identified drug resistant mutations (DRMs) using the Stanford University HIV Drug Resistance Database.



## Conclusions and Future Work

- The identification of new transmission clusters shows that combining similar datasets provides a more robust representation of the HIV epidemic in the DC area.
- The majority (62.5%) of clusters found between different datasets were between men who have sex with men (MSM).
- DRMs caused changes in 24 amino acids in Protease and 61 amino acids in Reverse Transcriptase across all datasets. A total of \_\_\_\_\_ amino acids were found to be under positive selection.
- The results of this study will be combined with surveillance study results to aid future HIV prevention efforts in Washington, DC.



Computational  
Biology Institute  
THE GEORGE WASHINGTON UNIVERSITY

Milken Institute School  
of Public Health  
THE GEORGE WASHINGTON UNIVERSITY