Predicting HIV cluster growth using phylodynamic reconstruction in Los Angeles County

Manon Ragonnet-Cronin[1], Yunyin W. Hu[2], and Joel O. Wertheim[1]

background

- 60,000 people in Los Angeles County (LAC) are living with HIV-1
- HIV-1 pol is sequenced in newly diagnosed patients and sequences are reported to Departments of Public Health
- In LAC *pol* sequences have been reported for new HIV diagnoses and for treatment failures since 2005 (n=22,398)
- HIV-1 clusters of genetically similar viruses are attractive targets for public health interventions
- The assumption of targeting clusters is that past growth predicts future growth.

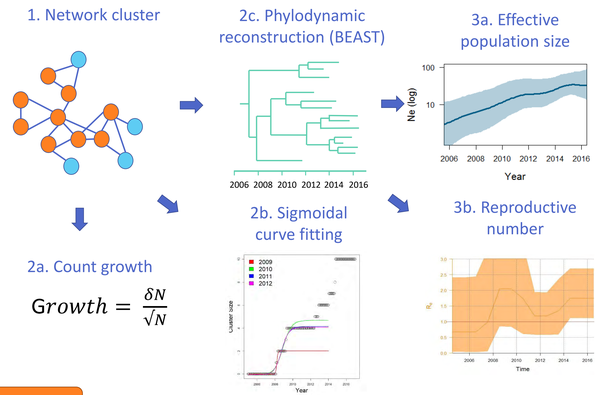
objective

- We aimed to
- → Implement different algorithms for measuring cluster growth
- → Test whether different growth measures are better at predicting future growth.

methods

1. Measure past growth

- We reconstructed the genetic network for LAC using HIV-TRACE [1], linking together genetic sequences with a pairwise genetic distance ≤0.015 substitutions/site
- For each cluster, we measured growth by:
  - (2a) Counting the number of diagnoses in the most recent year [2]
  - (2b) Fitting a sigmoidal curve to growth [3]
  - Reconstructing (3a) effective population size and (3b) reproductive number through time using BEAST [4].



2. Test which method best predicts future growth

- For each year and each method, we calculated growth for each cluster in the network
- We then identified the highest growth clusters based on each algorithm (~200 nodes in total)
- We constructed the network the following year and counted the number of nodes added to each cluster, to compare prediction methods
- We compared growth-based interventions to prioritizing clusters at random.

conclusions

- Past cluster growth (in the previous year) is predictive of future cluster growth (the following year)
- Cluster growth measures can be used to select clusters for network-based interventions
- Growth measures based on phylodynamic reconstruction (change in effective population size and reproductive number) did not perform better than count growth, but were much more cumbersome to compute
- In addition, phylodynamic reconstructions converged poorly for clusters <8 sequences, which make up 89% of clusters in LAC in 2016
- When we limited our analysis to large clusters, no method performed better than prioritizing large clusters at random
  - However large clusters did not grow more than smaller clusters.

results

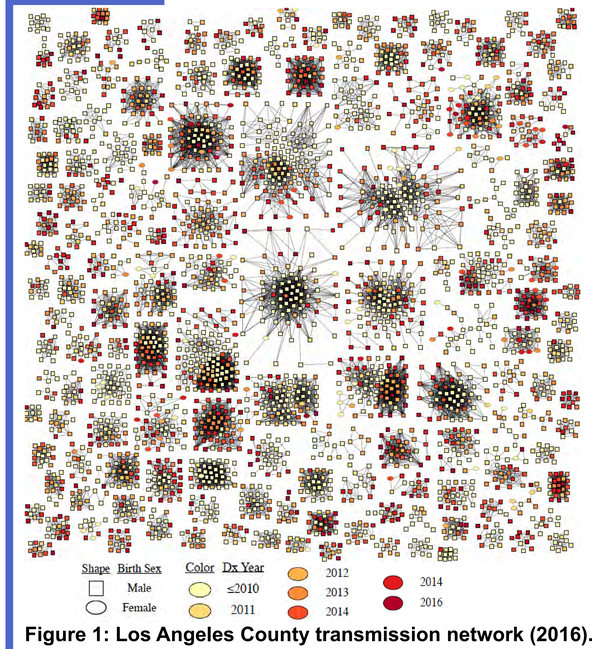


Figure 1: Los Angeles County transmission network (2016). Only clusters ≥ 10 nodes are shown.

1. The proportion of sequences clustering increased over time

- By 2016, 8,133/22,398 (36.3%) of individuals clustered (Figure 1, Figure 2).

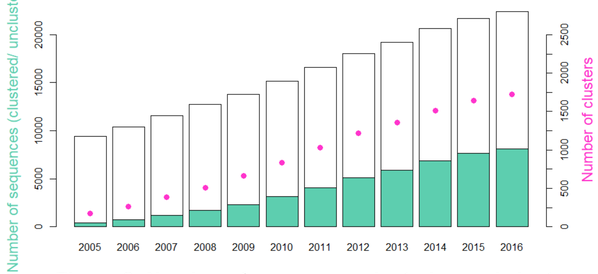


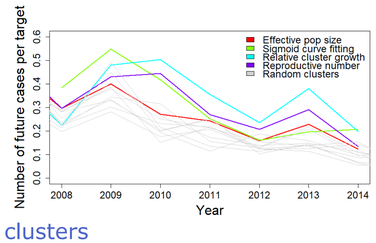
Figure 2: Number of sequences clustering and clusters over time in the Los Angeles County dataset

2. Past growth predicts future growth

- Cases targeted based on growth measures were associated with 0.30 new diagnoses on average, compared to 0.15 for random clusters (Figure 3)
- No single algorithm performed better than the others
- Phylodynamic reconstruction imputes missing cases but takes much longer to run, and estimates are difficult to obtain for clusters <8 sequences.

All algorithms prioritized clusters that were more likely to grow than randomly selected clusters.

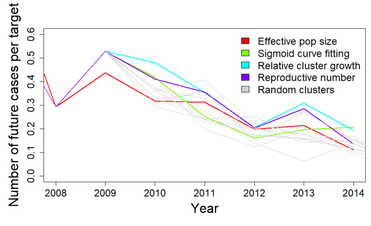
Figure 3: Average number of future cases for each case prioritized (compared to prioritizing clusters at random).



3. Prioritizing large clusters

When we limited our analysis to clusters ≥10 people (which run best in BEAST), no method worked better than selecting large clusters at random.

Figure 4: Average number of future cases for each case prioritized (clusters ≥ 10)



- Large clusters did not grow more than smaller clusters, so clusters should be targeted based on growth rather than size.

references

1. Kosakovsky Pond et al. HIV-TRACE (Transmission Cluster Engine): a tool for large scale molecular epidemiology of HIV-1 and other rapidly evolving pathogens, MBE 2018
2. Wertheim et al. Growth of HIV-1 Molecular Transmission Clusters in New York City (in preparation)
3. Poon et al. The Impact of Clinical, Demographic and Risk Factors on Rates of HIV Transmission: A Population-based Phylogenetic Analysis in British Columbia, Canada, JID 2015
4. Drummond et al. Bayesian phylogenetics with BEAUti and the BEAST 1.7, MBE 2012

acknowledgements

1. Department of Medicine, University of California, San Diego
2. Division of HIV and STD Programs, Los Angeles County Department of Public Health

CFAR P30 AI036214
NIH-NIAID K01AI110181
CHRP ID15-SD-052