A Phylogenetic Approach to Prioritizing HIV Transmission Clusters

Jeffrey Joy[1,2,3], Angela McLaughlin[1,2,3], Richard H. Liang[1], Jinny Choi[1], Chanson J. Brumme[1], Julio S.G. Montaner[1,4]

[1]British Columbia Centre for Excellence in HIV/AIDS, St. Paul's Hospital, [2]Bioinformatics, University of British Columbia (UBC), [3]School of Population and Public Health, UBC, [4]Department of Medicine, UBC

Background

- Identifying areas that are at a high risk for ongoing HIV transmission is critical for prioritization of limited public health resources to support people living with HIV and prevent new cases.
- Despite advancements in testing and treating, foci of high transmission remain even in well managed epidemics in developed countries.
- Detection and tracking of phylogenetic transmission clusters is now becoming common practice in many jurisdictions.
- When new cases are observed to join a cluster, these groups are candidates for targeted public health interventions such as enhanced testing, offering of PrEP to uninfected individuals within a network, and delivery of enhanced partner care.
- Public health agencies may need to prioritize cluster intervention when faced with many actively growing clusters and limited resources.
- Since transmission of HIV to a new host is equivalent to the formation of a new lineage, diversification rates inferred from viral phylogenetic trees can serve as estimates for transmission rates.
- **Lineage-level diversification rates (Figure 1) are an alternative phylogenetic measure which can be rapidly computed and combined with traditional phylogenetic clustering to highlight clusters which are undergoing the most rapid transmission.**

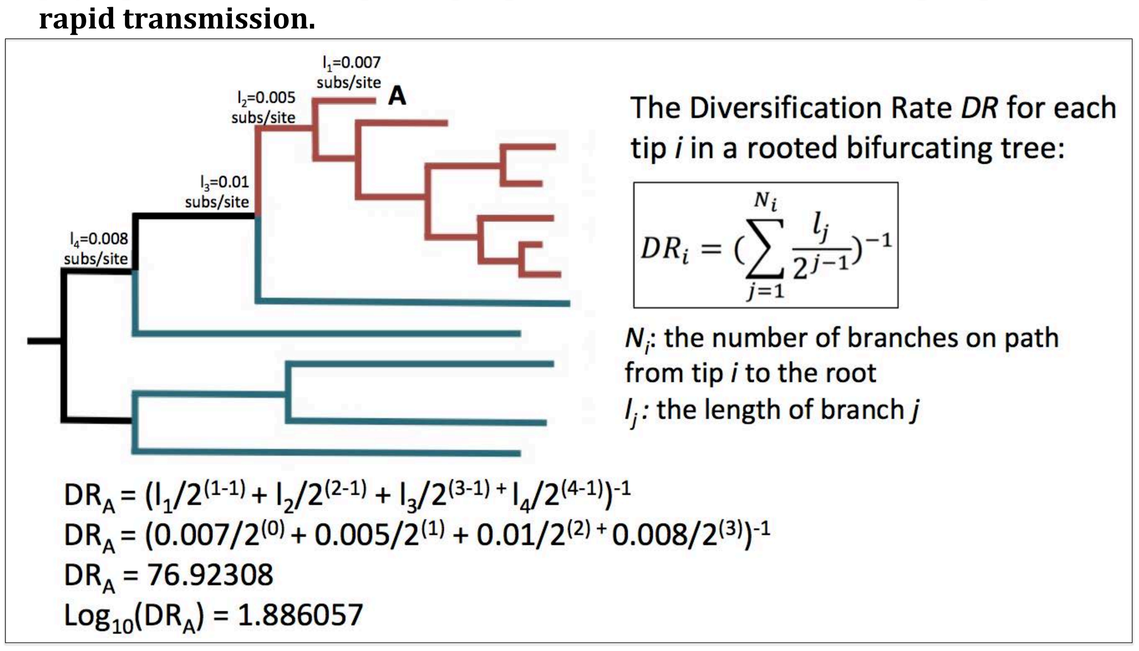

Figure 1. The lineage-level diversification rate for each tip on a viral phylogenetic tree estimates the between-host transmission rate based on length of its branches[1].

Methods

- A total of 36,271 HIV-1 resistance genotype tests (HIV *protease* and partial *reverse transcriptase* genes, partial *pol*) – were run for 9,630 participants enrolled in the British Columbia (BC), Canada (Figure 2) Drug Treatment Program (DTP) at the BC-CfE between May 1996 and March 2018.
- All sequences were aligned to HXB2 reference genome using MAFFT.
- Insertions and deletions relative to HXB2, as well as amino acids corresponding to known drug resistance mutation sites were removed prior to tree inference.
- A set of shuffled bootstrap alignments were generated to infer 100 approximate maximum likelihood phylogenetic trees in FastTree2.1.
- Trees were pruned to include each patient's oldest sample and then rooted using root-to-tip regression in the R package ape (Figure 4).
- Transmission clusters were inferred using a patristic (tip-to-tip) distance threshold of 0.02 substitutions/site (95th percentile of within-host patristic distances) and had to contain a minimum of five individuals (Figure 3)[2].
- For each tip on each bootstrap tree, the viral lineage-level diversification rate was calculated and the mean diversification rate across 100 bootstrap trees for each tip was calculated.
- We then aggregated mean, median, and maximum diversification rates for each cluster to identify which clusters displayed the highest diversification (transmission) rates and compared it with public health data (Figure 5).

Figure 2. The study area is based in British Columbia (BC), Canada.

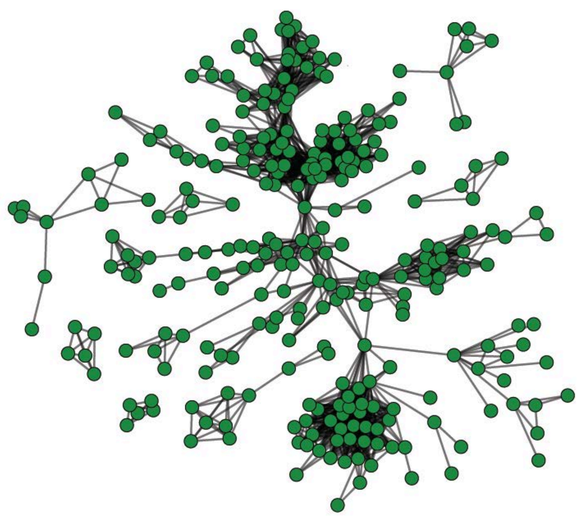

Figure 3. Exemplary clusters in the BC epidemic.

Results

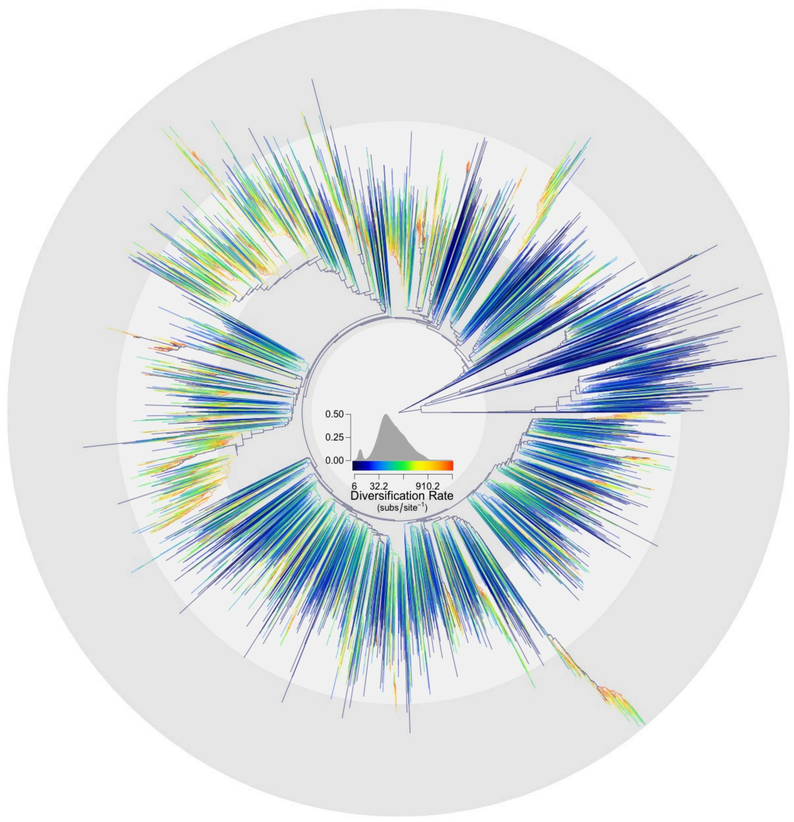

Figure 4. Representative approximate maximum likelihood bootstrap phylogenetic tree for 2018 coloured by lineage-level diversification rate. Cooler (blue) colours represent slower diversification rates while warmer (redder) colours represent rapid diversification.

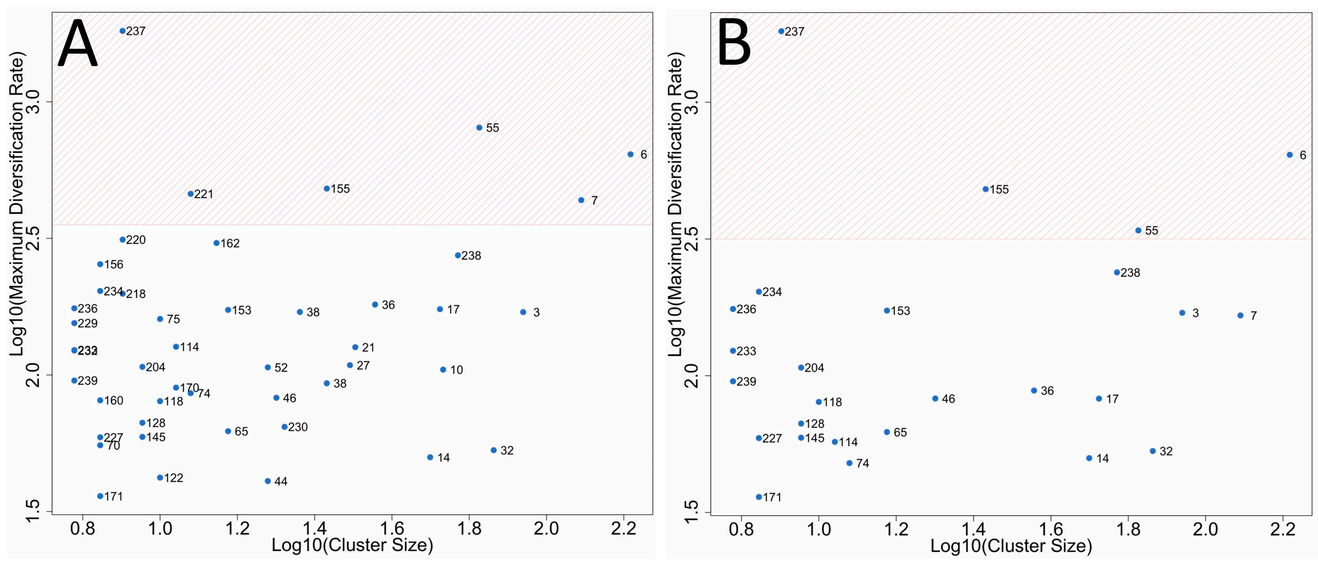

Figure 5. British Columbia phylogenetic transmission clusters for **A.** the preceding 1 year and **B.** the preceding 6 months. Clusters are plotted by log_{10}(cluster size) (x-axis) and log_{10}(cluster maximum diversification rate) (y-axis). Numbers refer to cluster designation. The hatched area highlights clusters that were also clusters of current urgent public health concern during the time period.

Conclusions

- **The maximum diversification (transmission) rate for a cluster was found to be the best predictor of clusters independently identified as being of public health concern compared with mean or median diversification rate.**
- The combination of phylogenetic clustering and lineage level diversification rates, which are both feasible to compute in a short time scale even with large datasets, may allow public health agencies to increase the specificity with which they provide interventions to communities and groups who most urgently need them and may also be predictive of clusters undergoing further rapid growth.
- Phylogenetic data can complement traditional epidemiological data by providing insight into temporally-informed between-host evolution.
- Such data and analyses can not account for the ever present possibility of unsampled infections and can not be utilized to determine transmission directionality.

Acknowledgments and Financial Disclosure

Thank you to the BC Centre for Excellence in HIV/AIDS at St. Paul's Hospital for coordinating the Drug Treatment Program and to the participants in the program for participating in HIV research in British Columbia. Jeffrey Joy is funded by a Genome Canada and Genome BC Bioinformatics and Computational Biology grant - 287PHY and the Public Health Agency of Canada. Angela McLaughlin has received funding through the BC Centre for Excellence in HIV/AIDS, Canadian Institute of Health Research, and the British Columbia Health Officers Council.

References

1. Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K. & Mooers, A. O. The global diversity of birds in space and time. Nature 491, 444–448 (2012).
2. Poon, A. F. Y. et al. The impact of clinical, demographic and risk factors on rates of HIV transmission: a population-based phylogenetic analysis in British Columbia, Canada. J. Infect. Dis. 211, 926–935 (2015).

Providence HEALTH CARE How you want to be treated.
BRITISH COLUMBIA CENTRE *for* EXCELLENCE *in* HIV/AIDS
UBC
Public Health Agency of Canada
Agence de la santé publique du Canada
GenomeCanada
Genome BritishColumbia Leading > Investing > Connecting
CIHR IRSC Canadian Institutes of Health Research / Instituts de recherche en santé du Canada