PRIORITIZING HIV-1 TRANSMISSION CLUSTERS FOR INTERVENTION: A PHYLOGENETIC APPROACH

R04

Rachel L. Miller[1,2], Angela McLaughlin[1,2], Richard H. Liang[1], Delphine Duhirwe[1], Anh Q. Le[1], Chanson J. Brumme[1,3], Julio S.G. Montaner[1,3], Jeffrey B. Joy[1,2,3]

[1]BC Centre for Excellence in HIV/AIDS, Vancouver, Canada, [2]Bioinformatics Program, University of British Columbia , Vancouver, Canada, [3]Department of Medicine, University of British Columbia, Vancouver, Canada

BACKGROUND

Public health officials faced with many transmission clusters require rapid, scalable, and unbiased ways to prioritize distribution of limited resources. Current non-phylogenetic prioritization methods can be subjective, inconsistent and reliant on historical data, but these shortcomings could be eliminated by instead considering phylogenetically-derived lineage-level diversification rates.

We evaluate the hypothesis that transmission cluster prioritization based on phylogenetically-derived lineage-level diversification rates will be at least equally as effective as commonly used non-phylogenetic prioritization measures in stratifying clusters by transmission activity, without need for historical data or subjective interpretation.

METHODS

- 9,822 HIV *pol* sequences, the earliest collected per individual during routine drug resistance genotyping between 1996-2019, were used to infer approximate maximum-likelihood phylogenetic trees rooted by root-to-tip regression.
- Transmission clusters were inferred using a patristic distance threshold of 0.02 substitutions/site, such that lineage-level diversification rates calculated for all tips could be summarized by cluster.
- The same methods were applied to simulated epidemic sequence data generated in the FAVITES[1] framework via 7 parameter sets, 1 designed to mimic the last 10 years of the BC epidemic and 6 plausible variations.

$$DR_i = \left(\sum_{j=1}^{N_i} \frac{l_j}{2^{j-1}}\right)^{-1}$$

N_i = number of branches between tip i and the root
l_j = length of branch j Jetz et al. 2012[2]

Lineage-level Diversification Rate (DR)

- Tracks the historical branching rate associated with a tip, with emphasis on more recent events, thus approximating transmission rate.
- Tips sharing very recent common ancestry with many other lineages will have higher diversification rates, providing evidence for rapid transmission.

REFERENCES

1. Moshiri N, Ragonnet-Cronin M, Wertheim JO, et al.; FAVITES: simultaneous simulation of transmission networks, phylogenetic trees and sequences. *Bioinformatics* 2019;**35**(11):1852-1861. doi: 10.1093/bioinformatics/bty921.
2. Jetz W, Thomas GH, Joy JB, et al.; The global diversity of birds in space and time. *Nature* 2012;**491**(7424):444-8. doi: 10.1038/nature11631.

Lineage-level diversification rate-based measures outperform non-phylogenetic measures in prioritizing transmission clusters with the greatest growth potential.

RESULTS

Empirical data showed that diversification rate-based measures displayed relatively similar ability to delineate prioritized versus less urgent clusters relative to non-phylogenetic measures overall. Simulated data revealed a stronger relationship between the majority of diversification rate-based measures and future cluster growth than for non-phylogenetic measures.

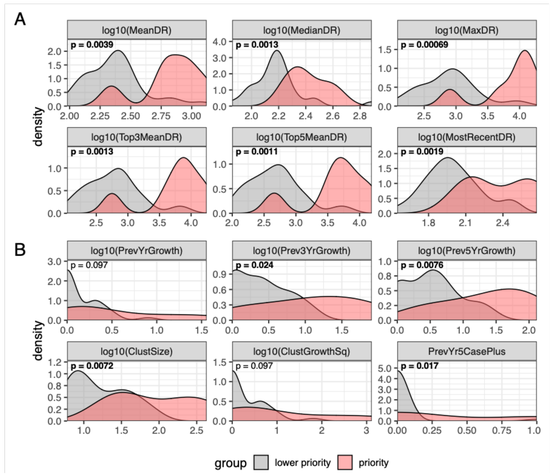


Figure 1. Comparison of prioritization measures between priority and non-priority groups, using empirical 2018 data. Density plots showing the difference in A) phylogenetic diversification rate-based measures or B) non-phylogenetic measures between clusters defined as "priority" by the current public health protocol and the remainder of the clusters to next be addressed, marked here for the purpose of comparison as "not priority". Only clusters that had newly diagnosed cases added in the previous year are shown. In 2018, 6/6 DR measures were significantly different between priority and lower priority groups, as were 4/6 non-phylogenetic measures.

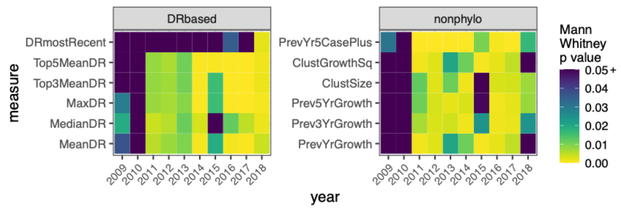


Figure 2. Heatmap of significance (Mann-Whitney U) of separation between priority and lower priority groups, as produced by each measure. Diversification rate-based measures often show greater consistency capturing statistically significant differences between top priority and lower priority clusters across the ten year period investigated relative to non-phylogenetic measures.

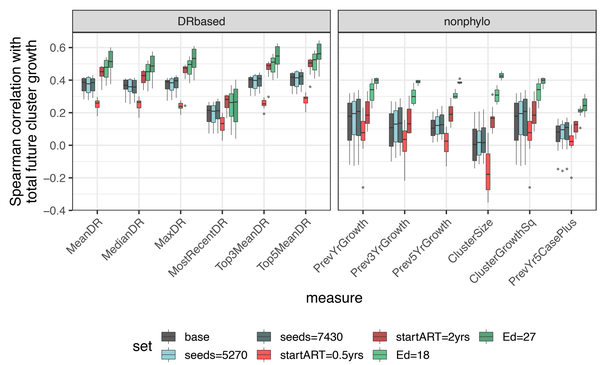


Figure 3. Comparison of correlation between prioritization measures and total future growth, using simulated data. Each box represents the Spearman correlation between a given prioritization measure and parameter set combination and total future cluster growth. In general, diversification rate-based measures showed stronger correlations with total future cluster growth than non-phylogenetic measures. Diversification rate measure correlations were also relatively robust to downsampling or showed some weakening, but non-phylogenetic correlations often weakened, sometimes to the point of reversal (not shown).

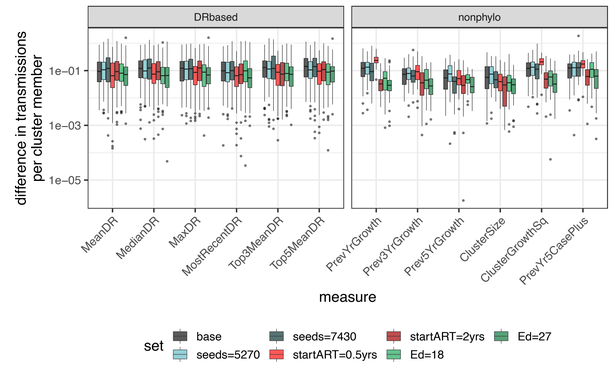


Figure 4. Difference in future transmissions across all prioritization measure and parameter set combinations. Each box represents the median and IQR of the difference in mean direct transmissions per cluster member between the top prioritized clusters, up to inclusion of 100 individuals (or the size of the top cluster, if its size exceeds 100), according to a given prioritization measure and a random sample of lower-ranking clusters containing the same number of individuals, across all clusters in the simulation time remaining. The difference in mean direct transmissions was, in general, slightly higher for diversification rate-based measures than for non-phylogenetic measures.

CONCLUSIONS

Diversification rate-based measures not only frequently outperform non-phylogenetic measures in their ability to identify groups with the highest potential for future growth, particularly long-term, but are also less likely to be perturbed by missing data and remain free of need for historical data and subjective interpretation. In contrast, although non-phylogenetic measures can sometimes be equally effective, they have the potential to become misleading in cases of lower sampling proportion. In combination with phylogenetic clustering, phylogenetically-derived lineage-level diversification rates can provide a simple, widely applicable and robust solution to focus prioritization of transmission clusters contributing the most to an ongoing epidemic.

ACKNOWLEDGEMENTS

We are grateful to the clinical laboratory, bioinformatics pipeline, and Drug Treatment Program teams at the BC Centre for Excellence in HIV/AIDS. We thank Providence Health Care, St. Paul's Hospital, Natural Sciences and Engineering Research Council, Canadian Institutes of Health Research, Public Health Agency of Canada, Genome Canada, and Genome BC for funding and support.

Please direct any questions to **rmiller@bccfe.ca.**

BRITISH COLUMBIA CENTRE for EXCELLENCE in HIV/AIDS

Public Health Agency of Canada | Agence de la santé publique du Canada

UBC | Providence Health Care | NSERC CRSNG | CIHR IRSC | GenomeCanada | Genome BritishColumbia