

Phylogenetic clustering reveals some concordance with respondent-driven sampling recruitment chains among men who have sex with men

Jeffrey B. Joy^{1,2}, Ashleigh Rich^{1,3}, Nathan J. Lachowsky^{1,4}, Susan Shurgold¹, Art F.Y. Poon^{1,2}, P. Richard Harrigan^{1,2}, Julio S.G. Montaner^{1,2}, Eric A. Roth^{1,4}, David M. Moore^{1,2}, Robert S. Hogg^{1,3}

1. BC Centre for Excellence in HIV/AIDS, 2. University of British Columbia, 3. Simon Fraser University, 4. University of Victoria

Background

The patterns of transmission of sexually-transmitted infections (STIs), such as human immunodeficiency virus (HIV), could relate to the structure of the underlying sexual contact network, whose features are therefore of direct clinical and public health interest¹. HIV transmission is highly concentrated among certain groups such as men who have sex with men (MSM). Sexual contact networks among MSM are routinely traced using respondent driven sampling (RDS) and the results represented as a graph structure².

Recovering the relationships amongst HIV sequences using phylogenetic methods is useful for both inferring epidemic history and identifying communities or groups at high risk of HIV transmission. Phylogenetic analyses of HIV-1 epidemics among MSM have revealed large transmission clusters, likely resulting from within community transmission^{1,3}. Few studies, however, have evaluated the relationship between sequence based transmission clusters and RDS recruitment chains. Overlap between such social-sexual networks and HIV phylogenies could reveal important insights into HIV epidemic dynamics and help direct effective prevention efforts.

We compared RDS recruitment chains with HIV phylogenetic transmission clusters from MSM to understand better what factors contribute to ongoing HIV transmission amongst MSM in Vancouver, Canada.

Methods

We assembled the clinical, socio-demographic characteristics, RDS recruitment chain data and viral sequence data from HIV-positive MSM recruited into the Momentum Health Study^{2,4}. 719 participants were recruited into the Momentum Health Study as directly recruited "seeds" (n=119) or from other participants in the study (n=600)².

HIV pol sequences for all HIV positive participants recruited using RDS were aligned using MAFFT v7.154b⁴ and visually inspected using AliView v1.15⁵. Phylogenies were inferred using FastTree2⁶ and rooted using Root-To-Tip regression as implemented in TempEst⁸. Phylogenetic clusters of three or more participants were identified using a tip-to-tip (patristic) distance cutoff < 0.02 (Figures 1 and 2). Participant socio-demographic and clinical attributes were correlated with cluster membership using the R statistical framework.

Results

RDS recruitment chains from the Momentum Health Study are depicted in Figure 3. Phylogenetic tree of sequences from the 206 HIV-positive MSM participants is shown in Figure 4. Sequences from 32 (15%) grouped into 7 phylogenetic clusters (Figure 5). All phylogenetic clusters spanned more than 1 RDS chain (mean = 4, range = 2 - 8).

In 4 clusters, individuals (n = 2 per cluster) were also members of the same RDS recruitment chain. Nearly half of the phylogenetic clusters (n = 2/7) contained individuals who HIV seroconverted during the study period.

References

- [1] Villandre, L., et al. *PLoS One.*, DOI:10.1371/journal.pone.0148459, 2016.
- [2] Moore, D., et al. *Journal of Acquired Immune Deficiency Syndrome*, 72:87-95, 2016.
- [3] Poon, A.F.Y., et al. *Journal of Infectious Diseases*, 211:926-35, 2015.
- [4] Forrest, J.I., et al. *Culture, Health, & Sexuality*, 16:288-301, 2014.
- [5] Katoh, K., et al. *Methods in Molecular Biology*, 537:39-64, 2009.
- [6] A. Larsson. *Bioinformatics*, 30(22):3276-3278, 2014.
- [7] Price, M.N., et al. *PLoS One* 5.3:e9490, 2010.
- [8] Rambaut, A. et al. *Virus Evolution*, 2(1):1-7, 2016.

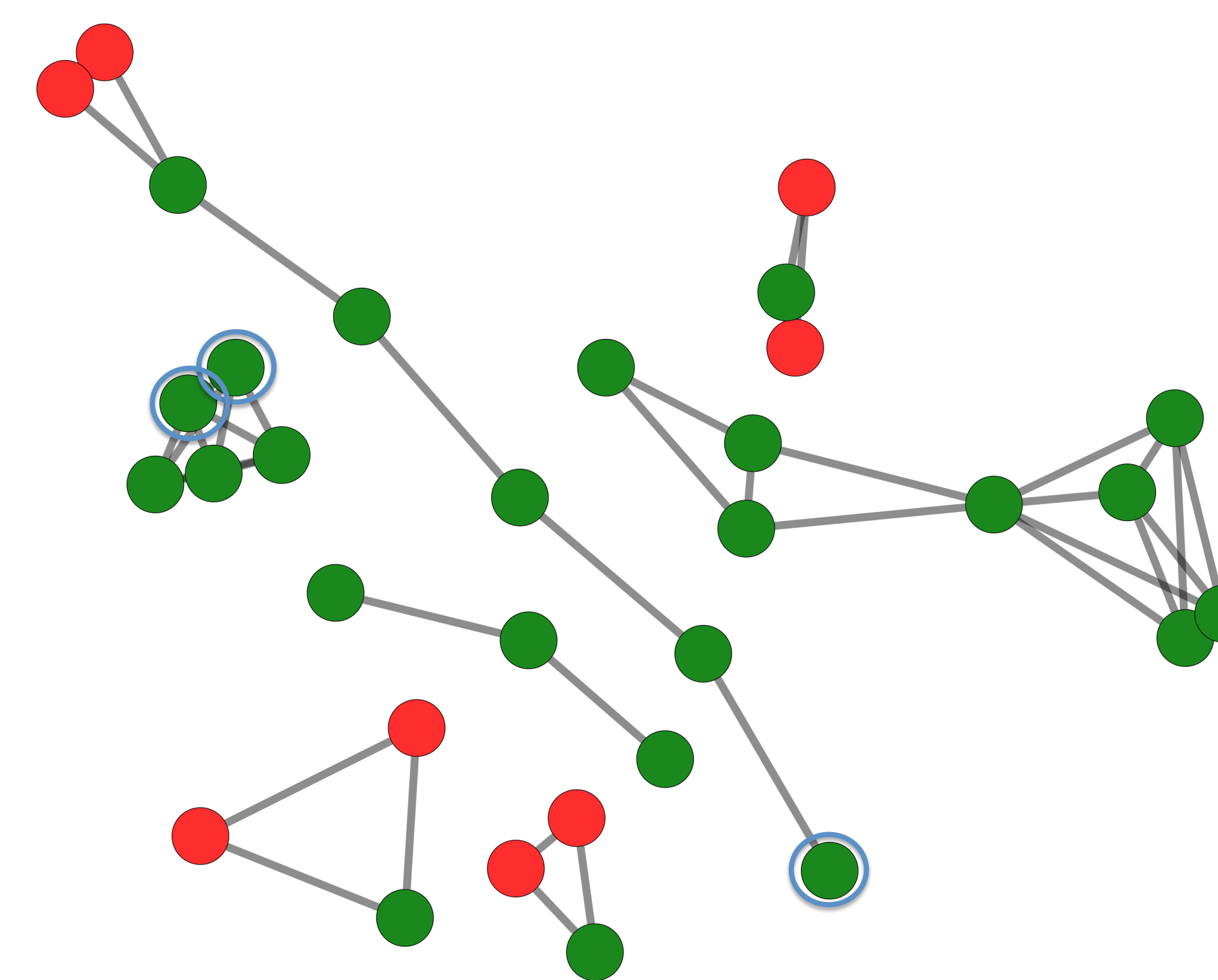


Figure 5 | Phylogenetic transmission clusters derived from 206 HIV-positive participants in the Momentum Health Study. Red indicates individuals that were also part of the same RDS recruitment chain. Circled individuals indicate individuals seroconverting for HIV-1 during the study period.

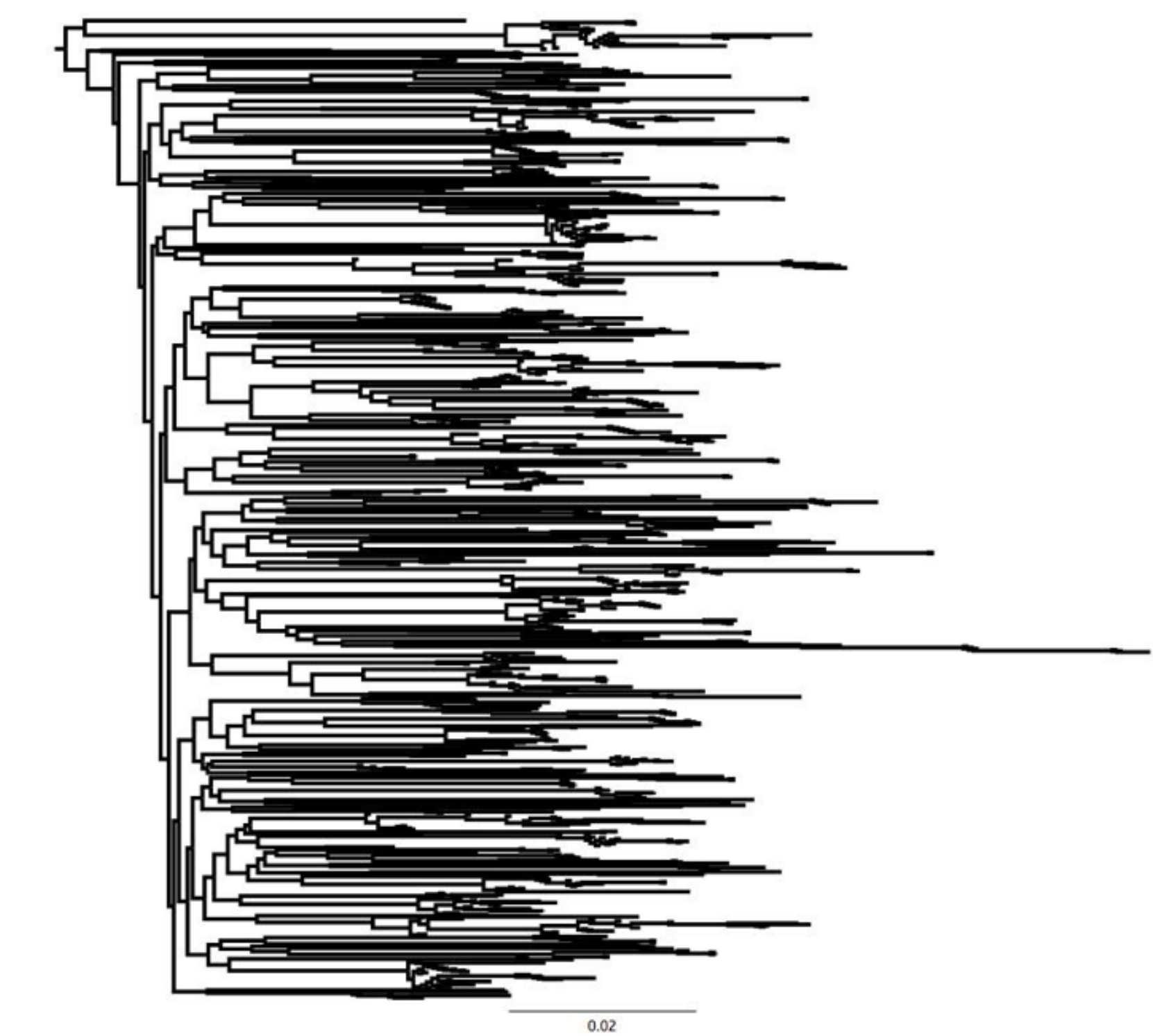


Figure 4 | Phylogenetic tree of sequences derived from HIV-positive participants in the Momentum Health Study. Tree was inferred based on pol sequences under approximate maximum likelihood with a GTR model as implemented in FastTree2.

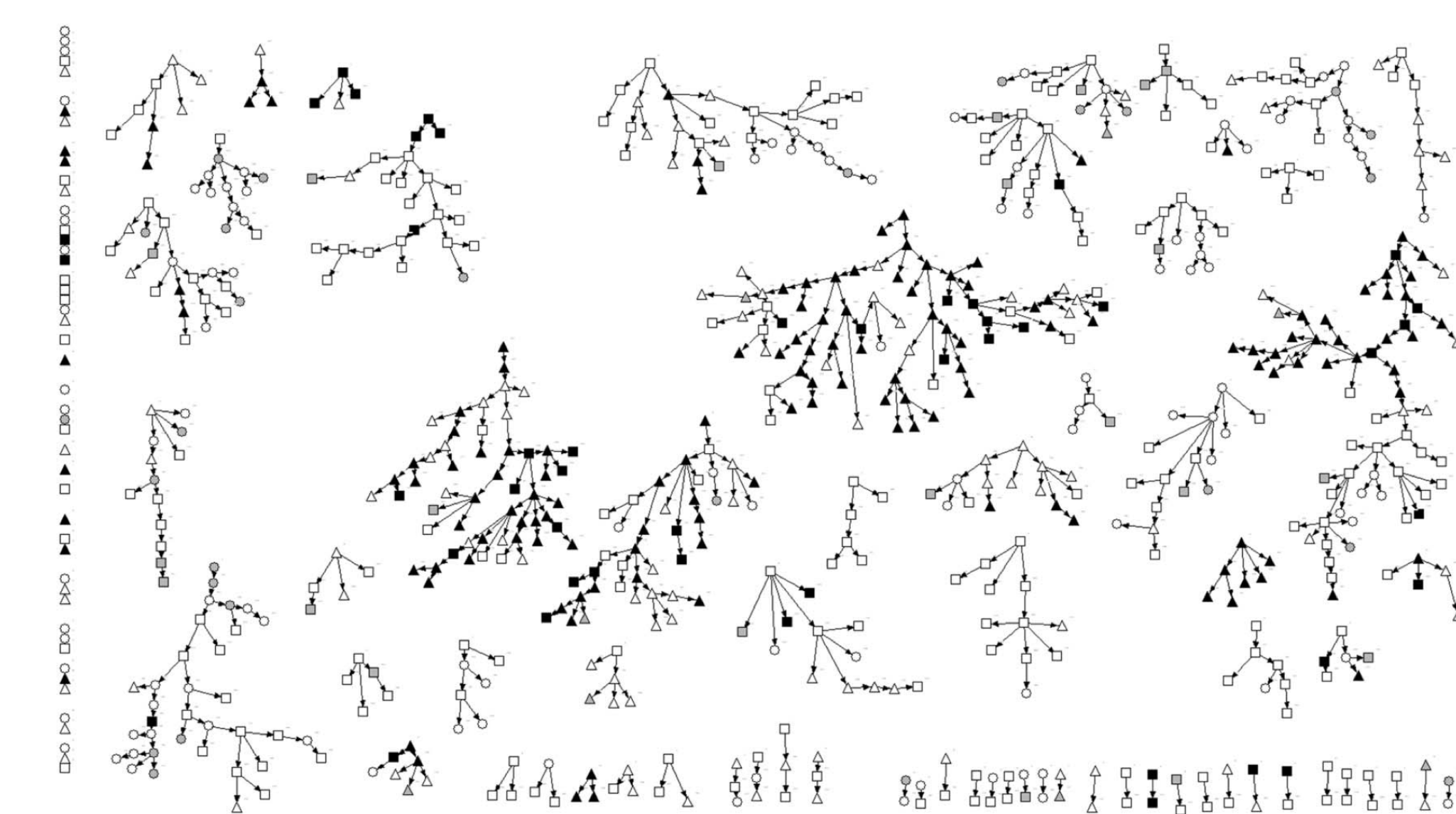


Figure 3 | Recruitment chains for the Momentum Health Study. White symbols = self reported HIV-negative, black = self-reported HIV-positive, grey = other; circle = 16-24 years of age; square = 25-39 years of age; and triangle = 40+ years of age. Seeds not giving rise to further recruitment are on the right of left (Moore et al. 2016).

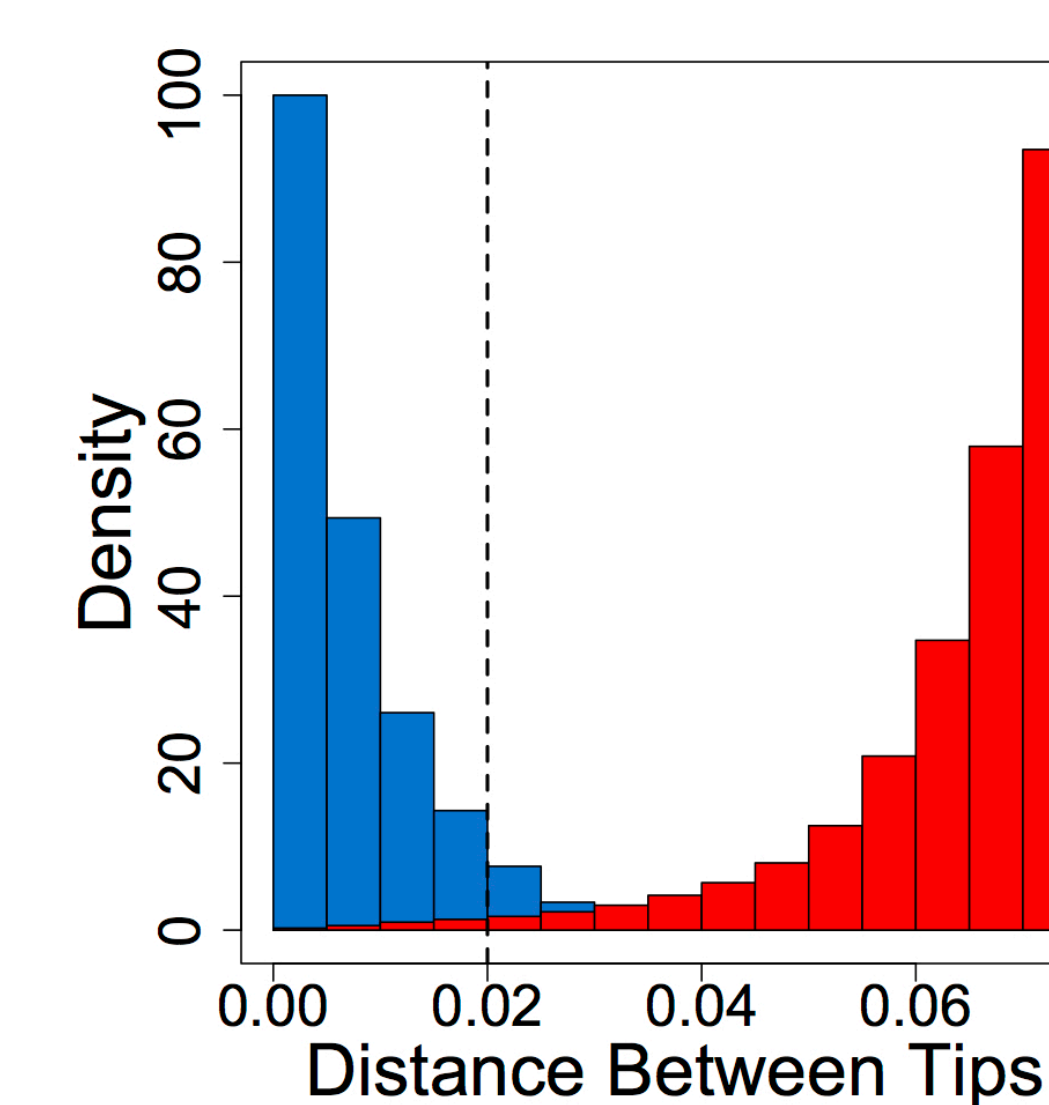


Figure 1 | Distribution of intra-patient patristic distances in blue and inter-patient patristic distances in red. The dashed line represents a quantitative break-point to determine inclusion in a cluster.

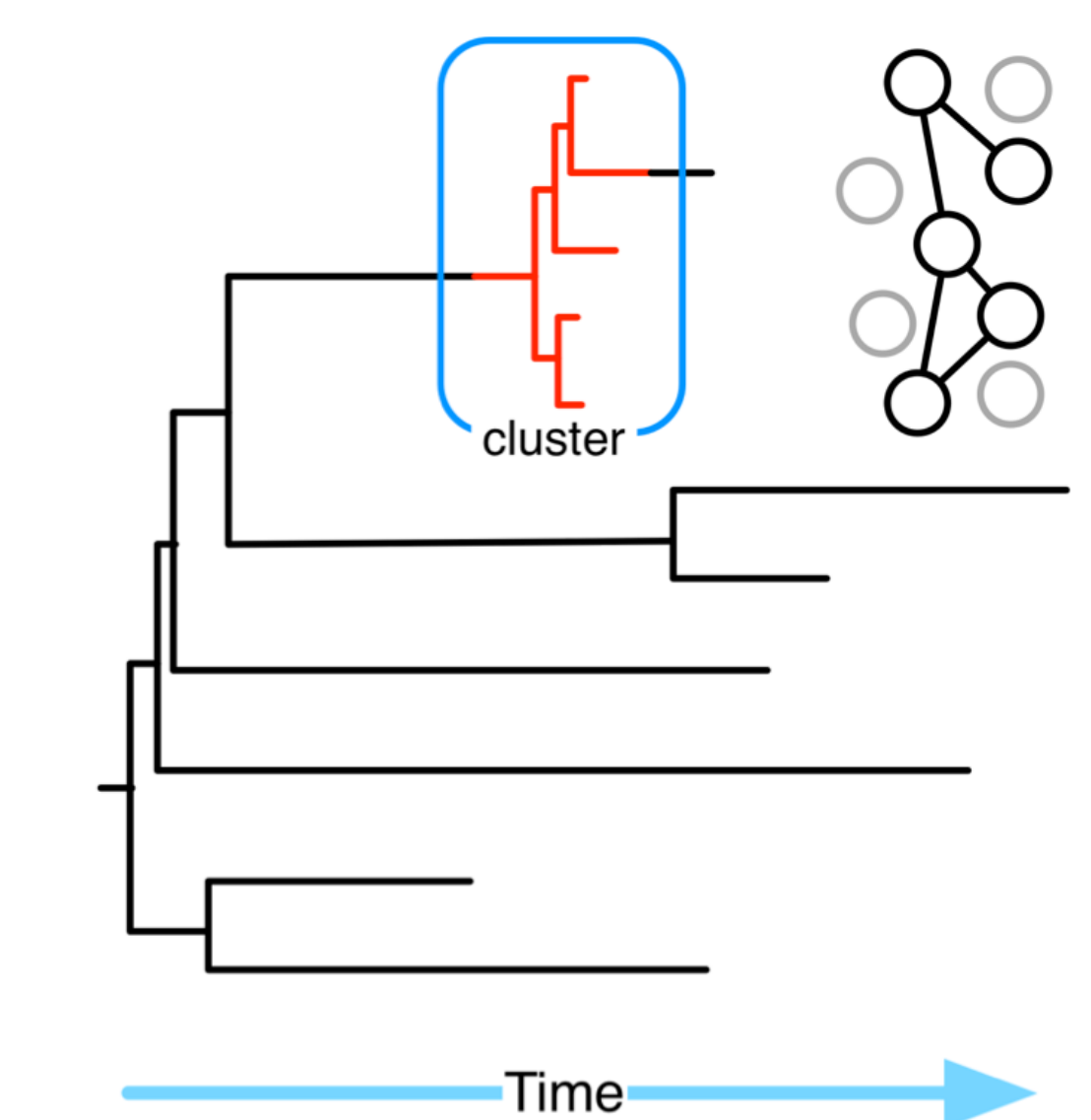


Figure 2 | Phylogenetic transmission cluster red sequences in the blue circle are related by a patristic distance cutoff of less than 0.02 and thus form a cluster.

Conclusion

We found relatively few HIV transmissions corresponding directly with the RDS recruitment chains. **However, we identified 3 seroconverters in phylogenetic clusters suggesting that RDS recruitment is successfully uncovering social-sexual networks with recent/ongoing HIV transmission. RDS sampling may be a useful tool when combined with phylogenetic monitoring for the targeted delivery of interventions to high risk populations.**