

# Coalescent inference of HIV transmission history

Raymond Heil

T-6: Theoretical Biology and Biophysics

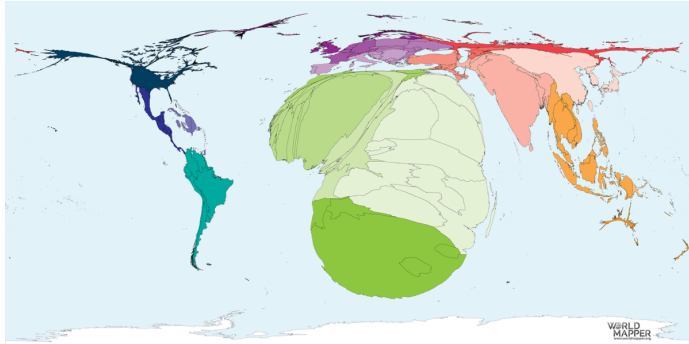
Mentors

Emma Goldberg, Thomas Leitner, Ethan Romero-Severson

20 July 2022

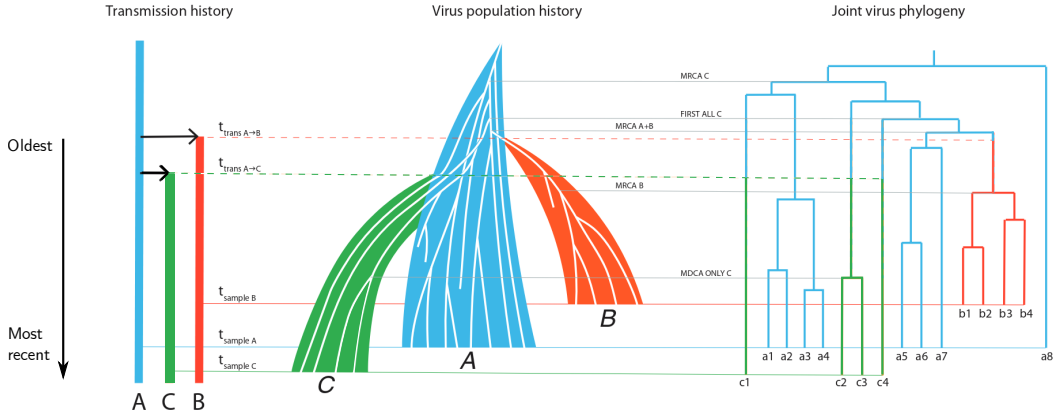


## Why this project?

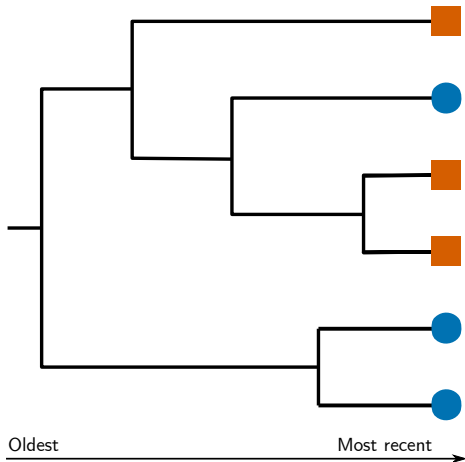


- \* HIV is globally widespread
- \* Despite effective therapy, it is still spreading
- \* How can we understand the way it spreads?

## Setting up the problem

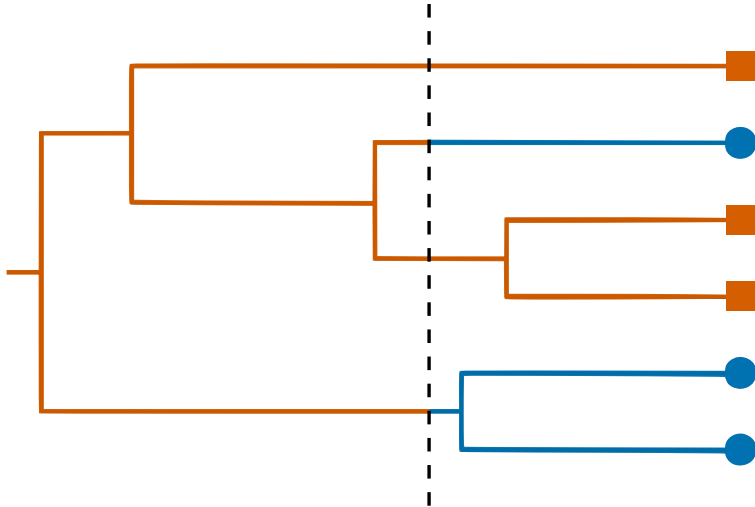


## Inferring information from a tree

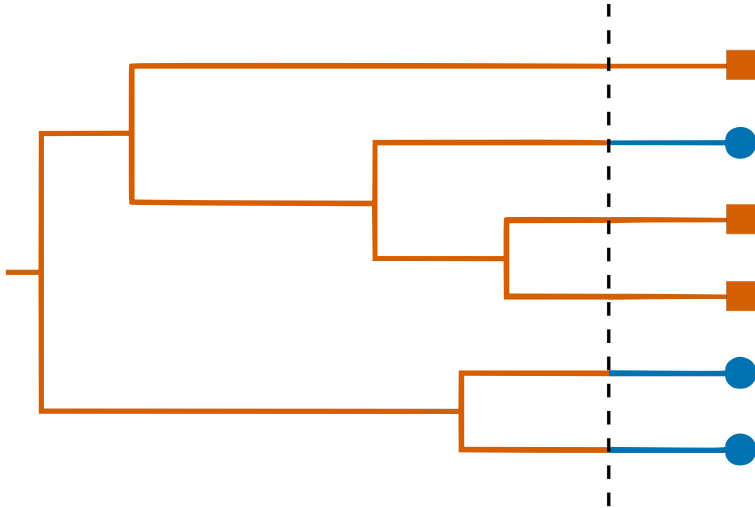


- \* Tree tips represent **individual viral sequences**
- \* Three samples from each individual
- \* If one infected the other, when?

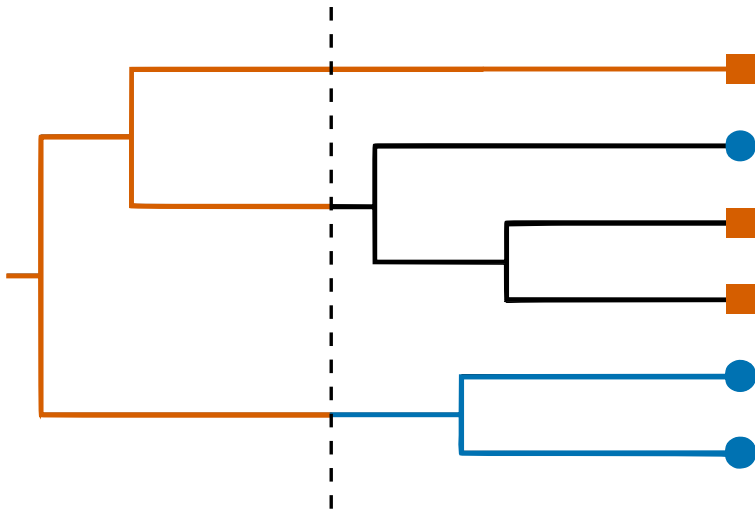
## Inferring information from a tree



## Inferring information from a tree



## Inferring information from a tree



# Coalescent theory

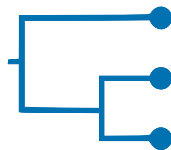
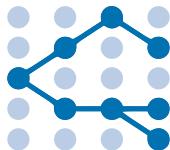
Node times as a function of population size



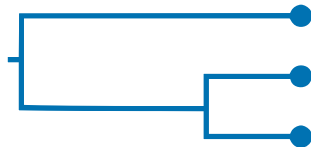
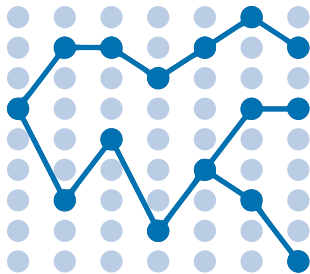
## Relationship between population and samples

Large  $N$  causes node times to be further apart, stretching the tree

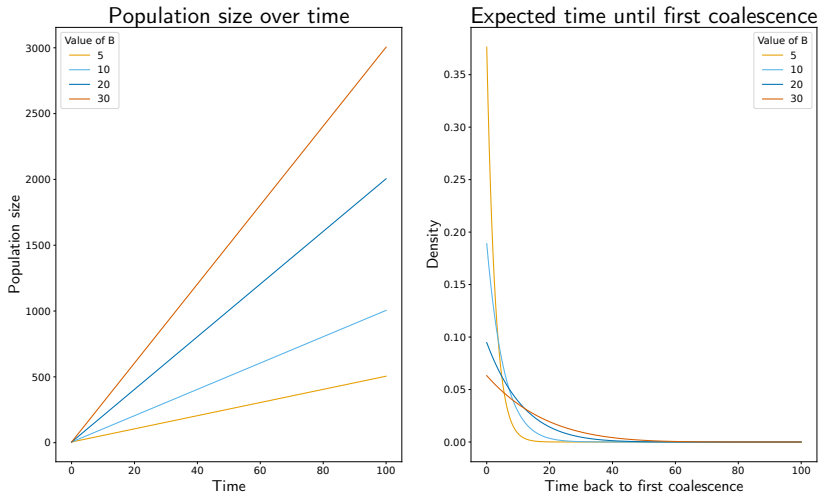
$N = 5$



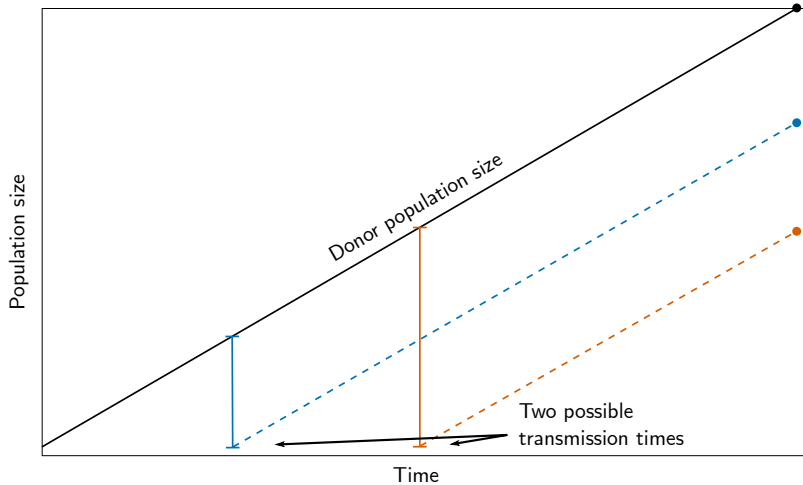
$N = 9$



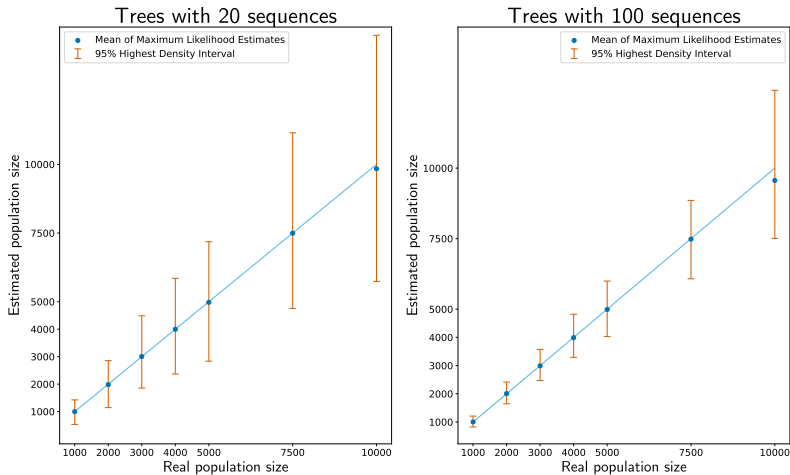
## Effect of different linear populations



## Predicting transmission time on a changing population

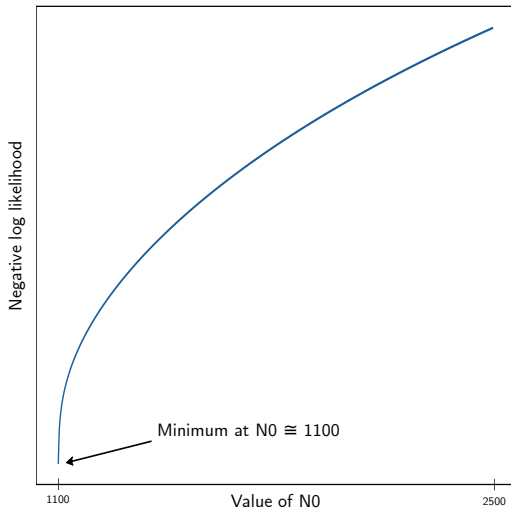


## Results: Single Host with Constant Population

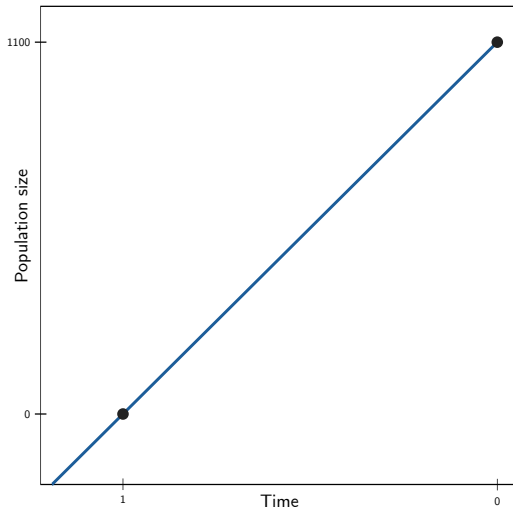


## Results: Single Host with Linear Population

Likelihood surface of  $N_0$  with  $b$  fixed at 1100



Population size of a tree with  $b$  fixed at 1100

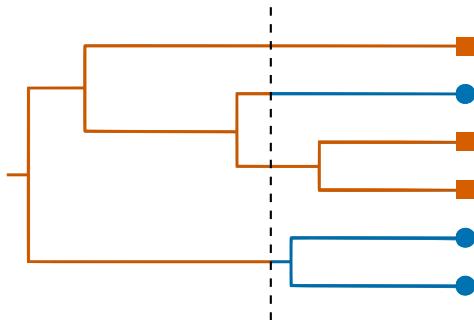


## Next steps: Two Hosts with Linear Population

**Immediately:** Solve numerical issues with multi-parameter optimization

**Immediately:** Extend my current work to trees with multiple hosts

- \* Split tree by host
- \* Isolate hosts until a transmission occurs



**Overall:** Find the most likely time of transmission for phylogenetic trees under a range of conditions

Thank you!

