
Evolution and Diversity of HIV in Southeastern Michigan

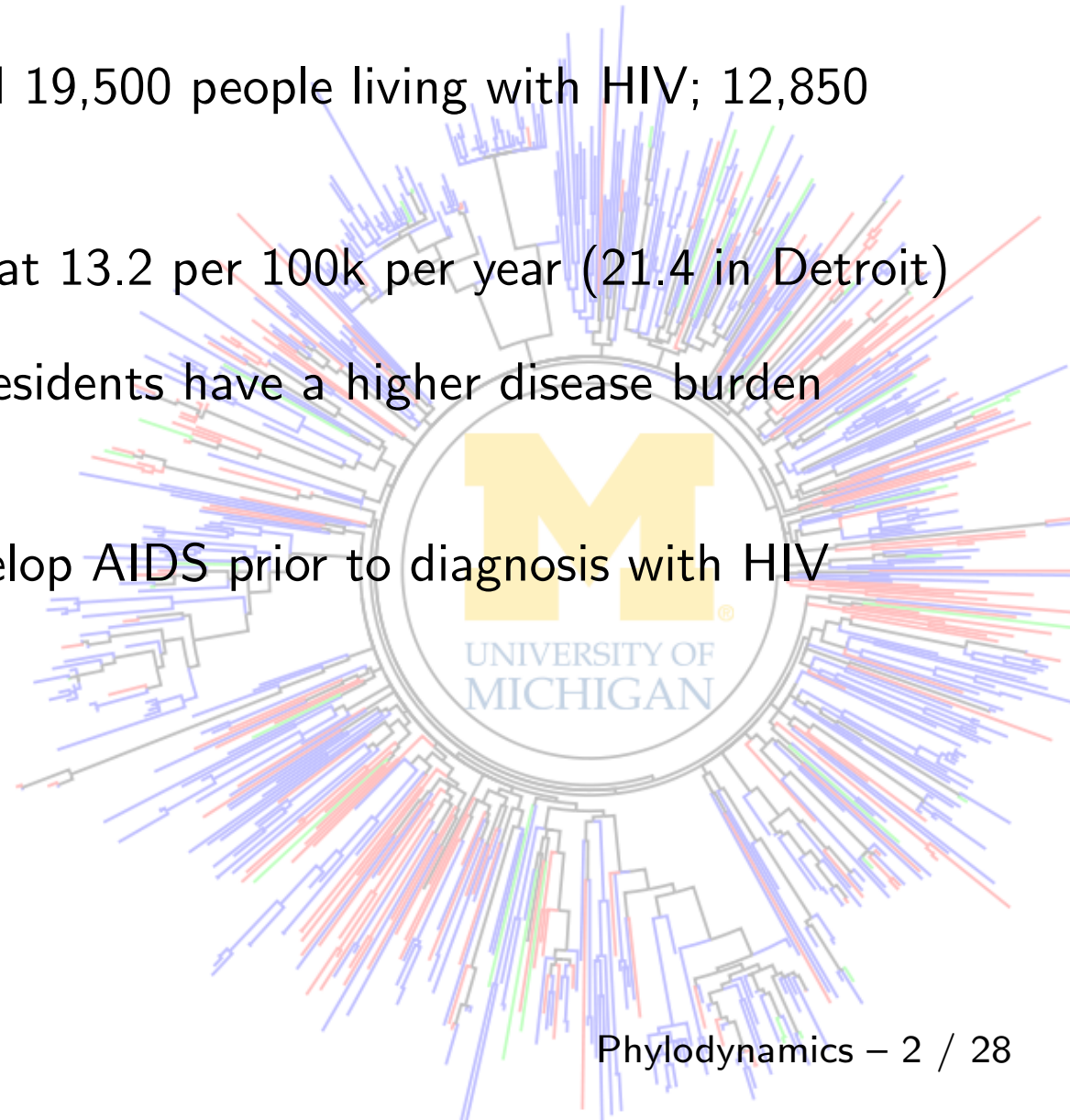
Erik M Volz

Department of Epidemiology, University of Michigan– Ann Arbor

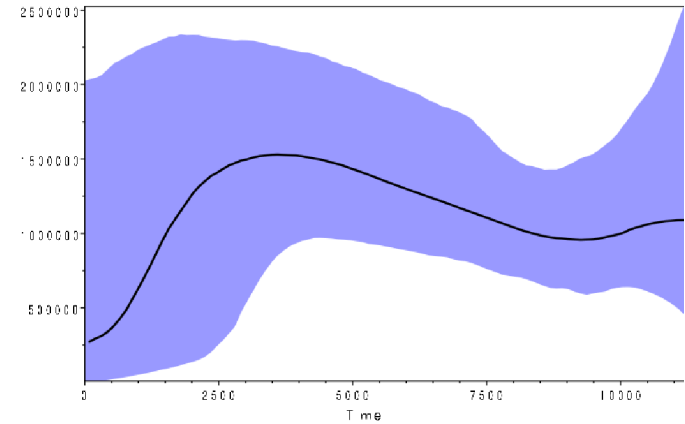
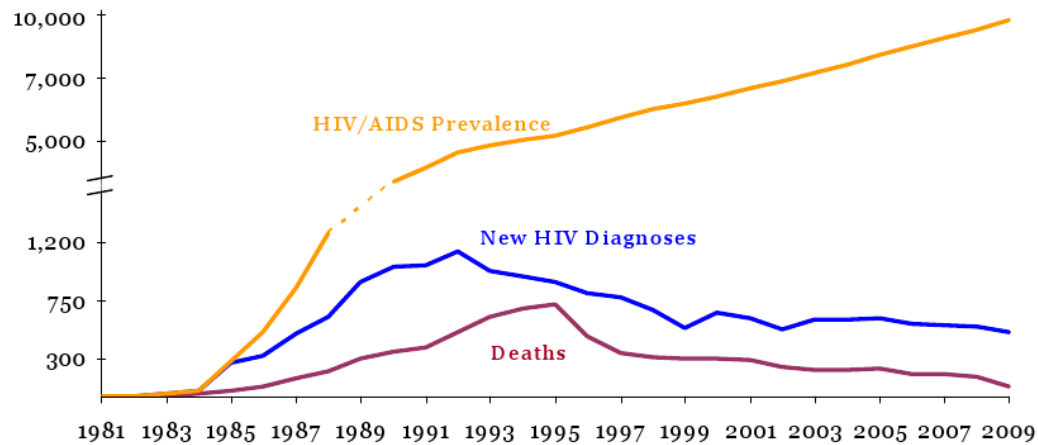
April 20, 2011

The HIV epidemic in Southeastern Michigan

- Statewide there are an estimated 19,500 people living with HIV; 12,850 infections in the Detroit MSA
- Rates of infection remain stable at 13.2 per 100k per year (21.4 in Detroit)
- African Americans and Detroit residents have a higher disease burden (11.1×)
- 20 % of infected individuals develop AIDS prior to diagnosis with HIV



Estimated prevalence in SE Michigan



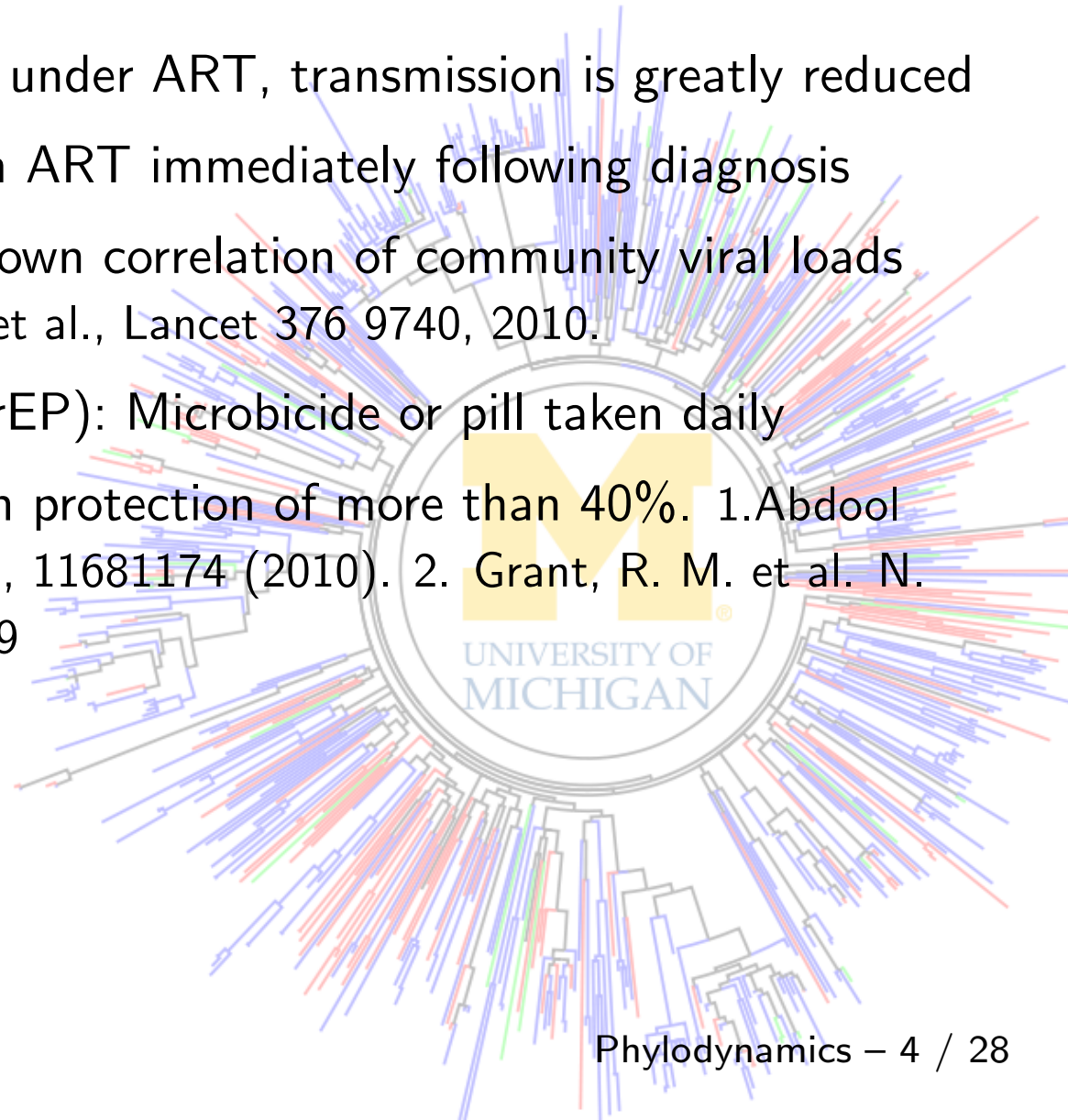
Source: Michigan Dept of Comm Health, Downloaded Jan 29

http://www.michigan.gov/documents/mdch/2010_Detroit_EMA-ALL_bm_335497_7.pdf

- HAART introduced in 1995
- It is difficult to estimate incidence in the early epidemic

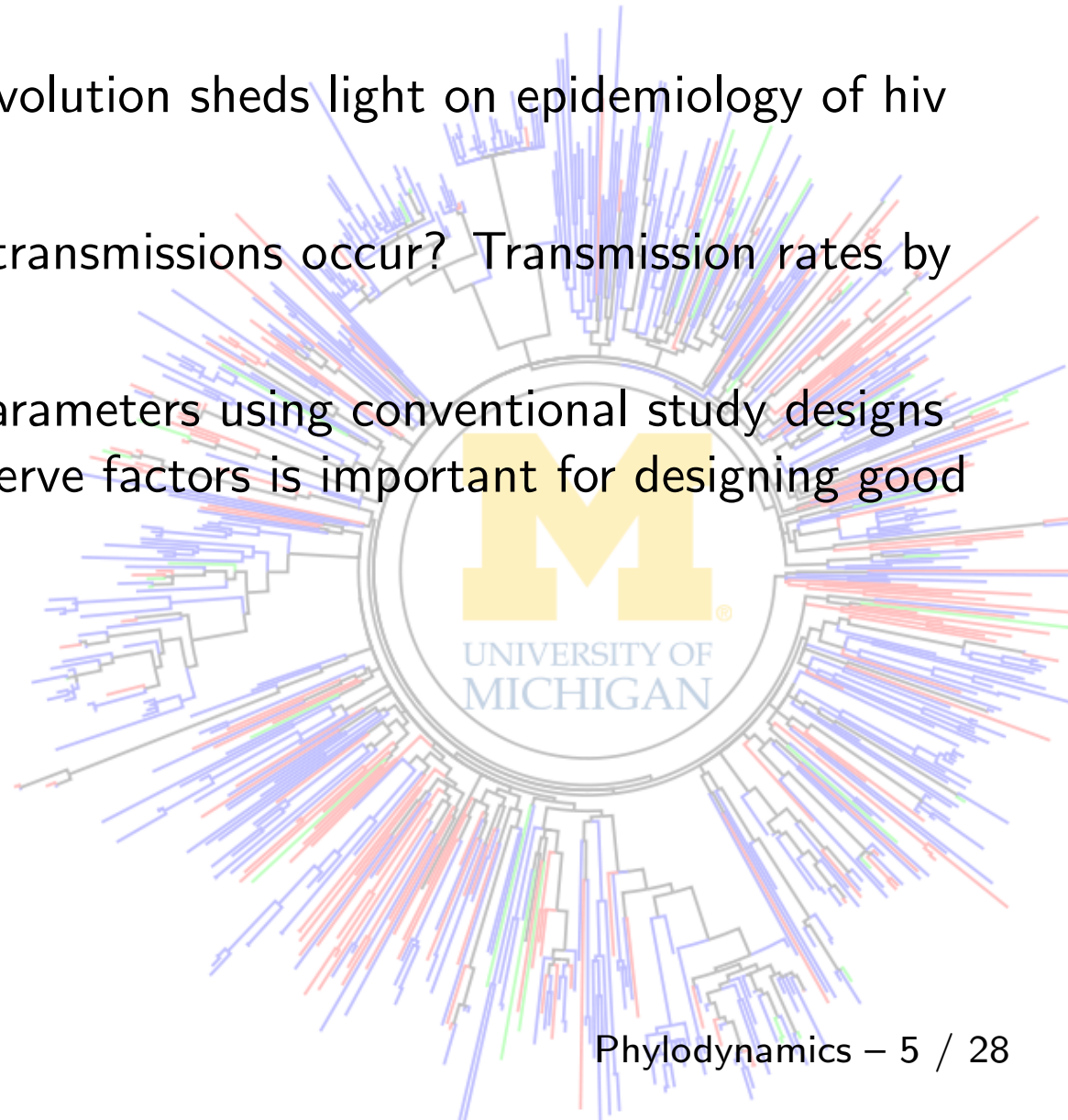
PHIs based on HAART

- Two public health interventions have been proposed based on the use of ARVs
- Viral loads become undetectable under ART, transmission is greatly reduced
 - Test and Treat (T&T): Begin ART immediately following diagnosis
 - ▷ Ecological studies have shown correlation of community viral loads and incidence 1.Montaner et al., Lancet 376 9740, 2010.
 - Pre-Exposure Prophylaxis (PrEP): Microbicide or pill taken daily
 - ▷ Recent studies have shown protection of more than 40%. 1.Abdool Karim, Q. et al. Science 329, 11681174 (2010). 2. Grant, R. M. et al. N. Engl. J. Med. 363, 25872599



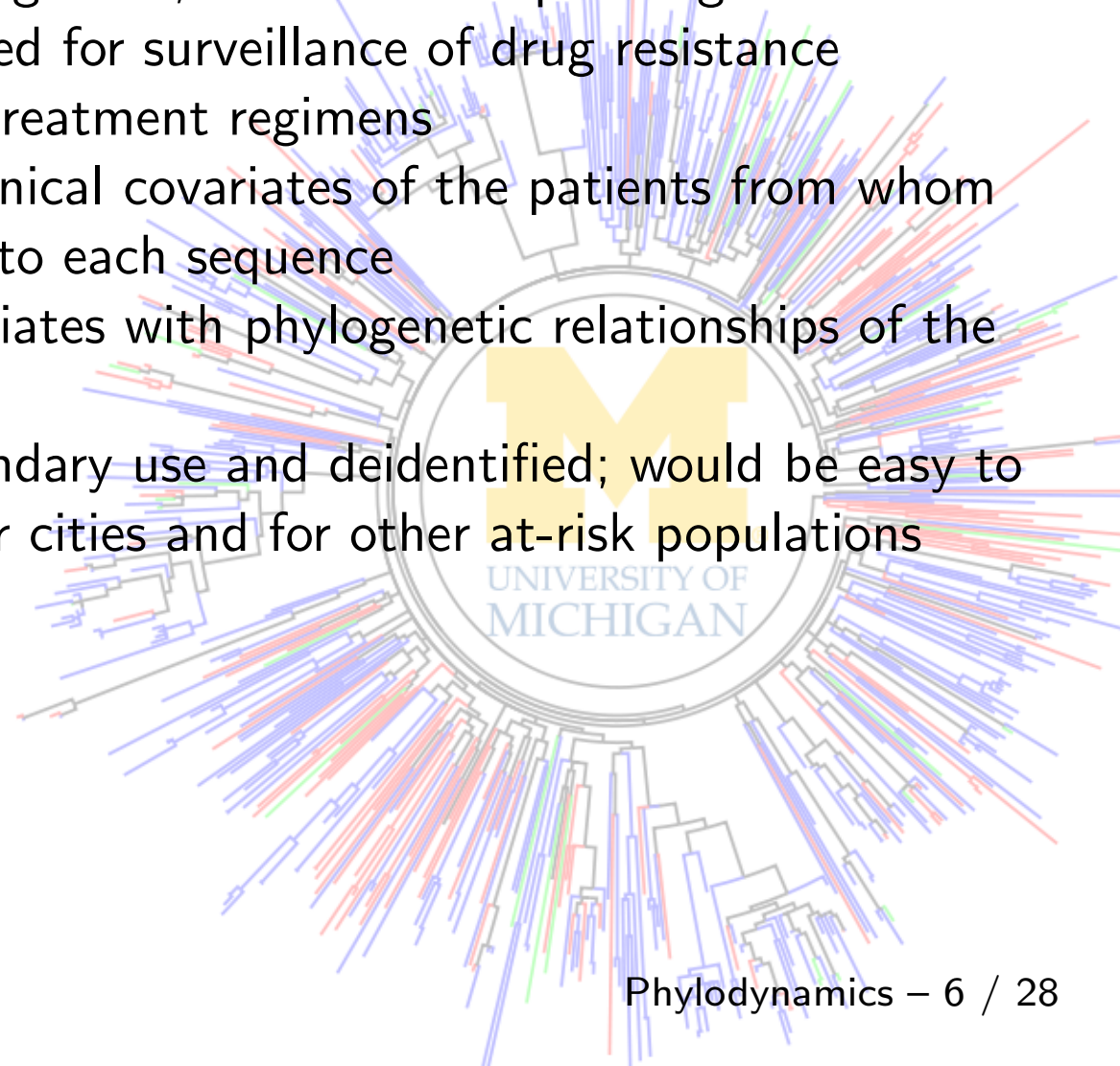
Our study...

- Motivation: Viral diversity and evolution sheds light on epidemiology of HIV that is hard to observe directly
 - Who infects who? When do transmissions occur? Transmission rates by stage of infection?
- It is difficult to estimate these parameters using conventional study designs
- Understanding these hard-to-observe factors is important for designing good PHIs



Our study...

- 2004-2010 Michigan Department of Community Health has collected HIV genetic sequences from newly diagnosed; the rate of sequencing has increased year over year. Collected for surveillance of drug resistance mutations and optimizing ARV treatment regimens
- Behavioral, demographic, and clinical covariates of the patients from whom the virus was isolated are linked to each sequence
- Correlation of patient-level covariates with phylogenetic relationships of the virus
- All data used in this study are 2ndary use and deidentified; would be easy to reproduce these analyses in other cities and for other at-risk populations



Methodology: Initial focus on MSM in Detroit MSA

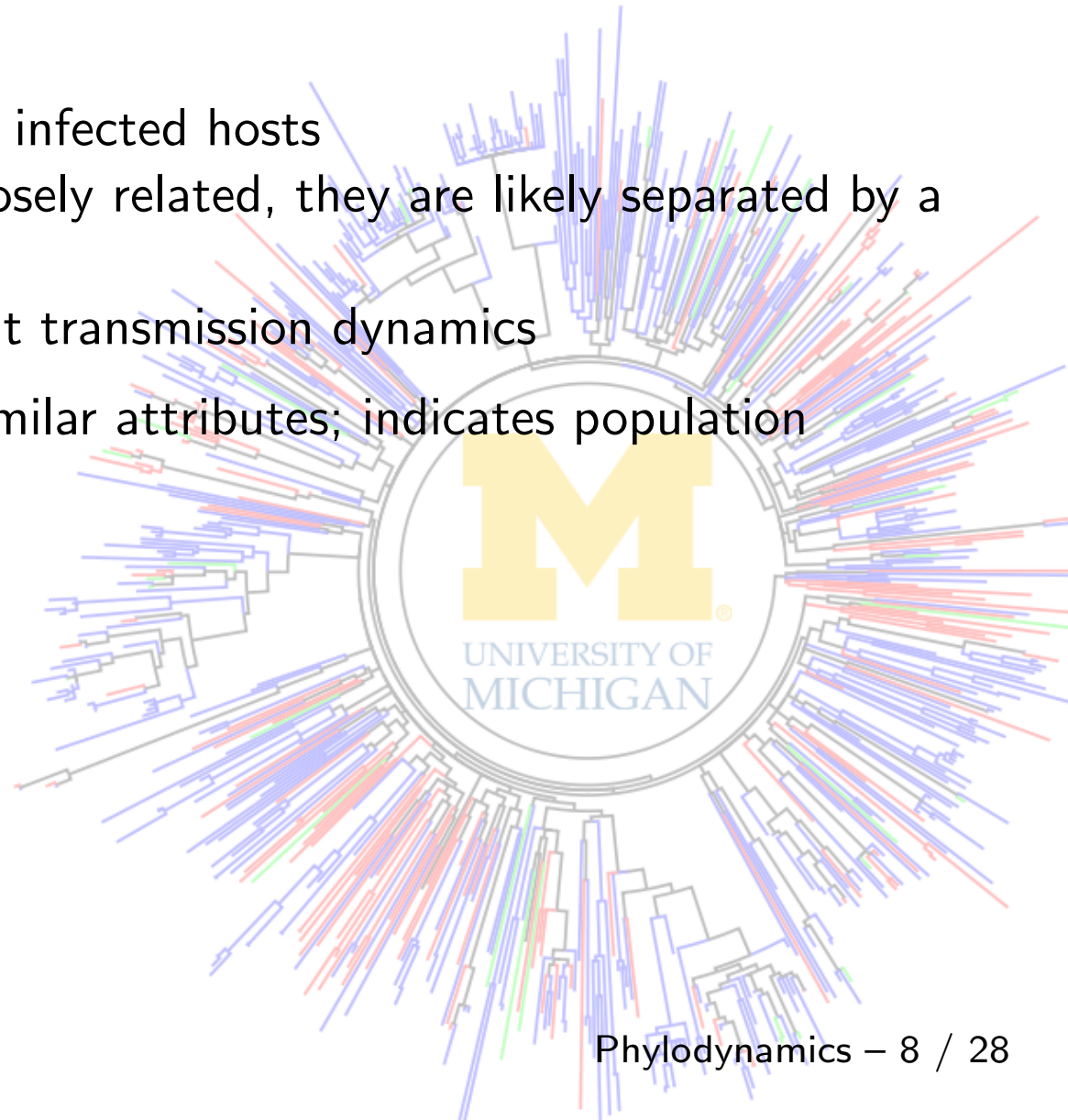
- Approx. half of new infections occur in those with confirmed MSM risk factor
- Approx. two thirds of new infections occur in Detroit and surrounding counties

Total: 3525 → MSM: 1881 → Detroit MSA: 1153
→ Tested for recent/acute infection: 711 → HIV-1 Subtype-B, high-quality: 612



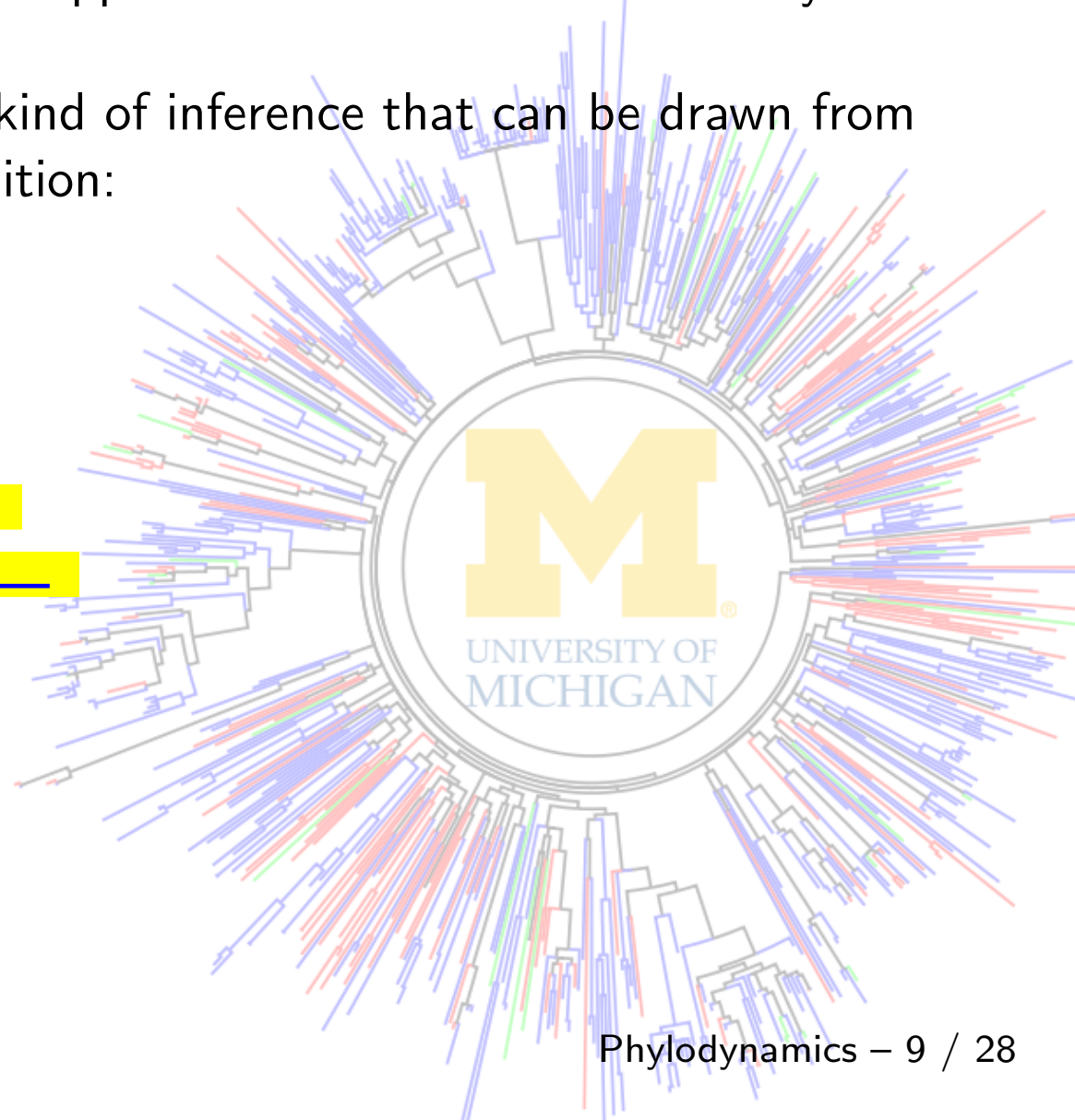
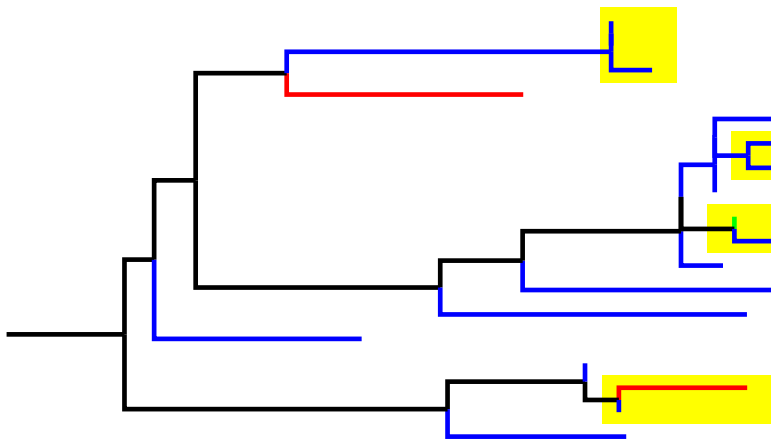
Phylogenetic clusters (Who infects who?)

- Sets of closely related virus from infected hosts
- If virus from two hosts is very closely related, they are likely separated by a short chain of transmissions
- Clusters can be informative about transmission dynamics
 - Clustered hosts often have similar attributes; indicates population structure

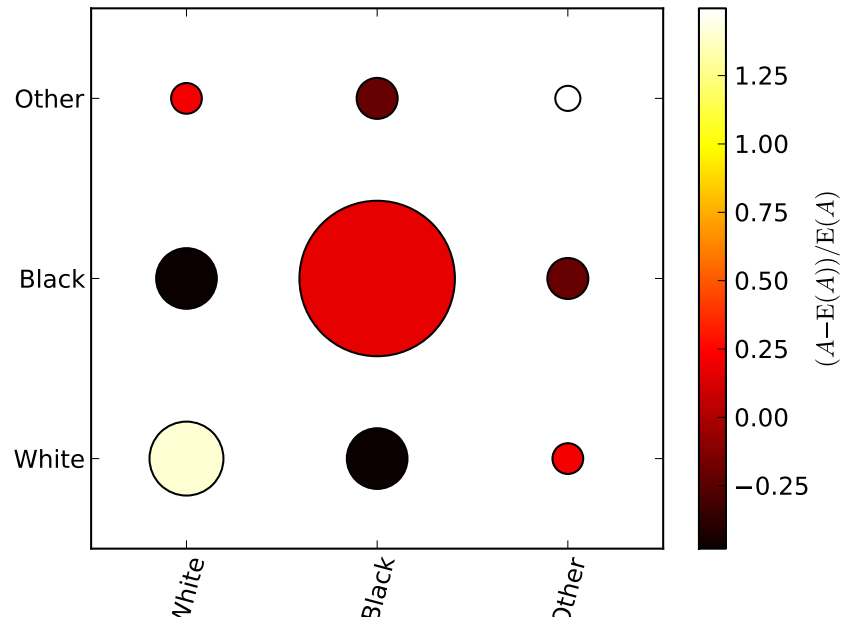


A simple definition of a cluster

- Our goal is to characterize clustering patterns as a function of demographic & behavioral variables; statistical support for a cluster is of secondary interest
- Straightforward to illustrate the kind of inference that can be drawn from clusters using a very simple definition:
 - Clades of size 2 (“cherries”)

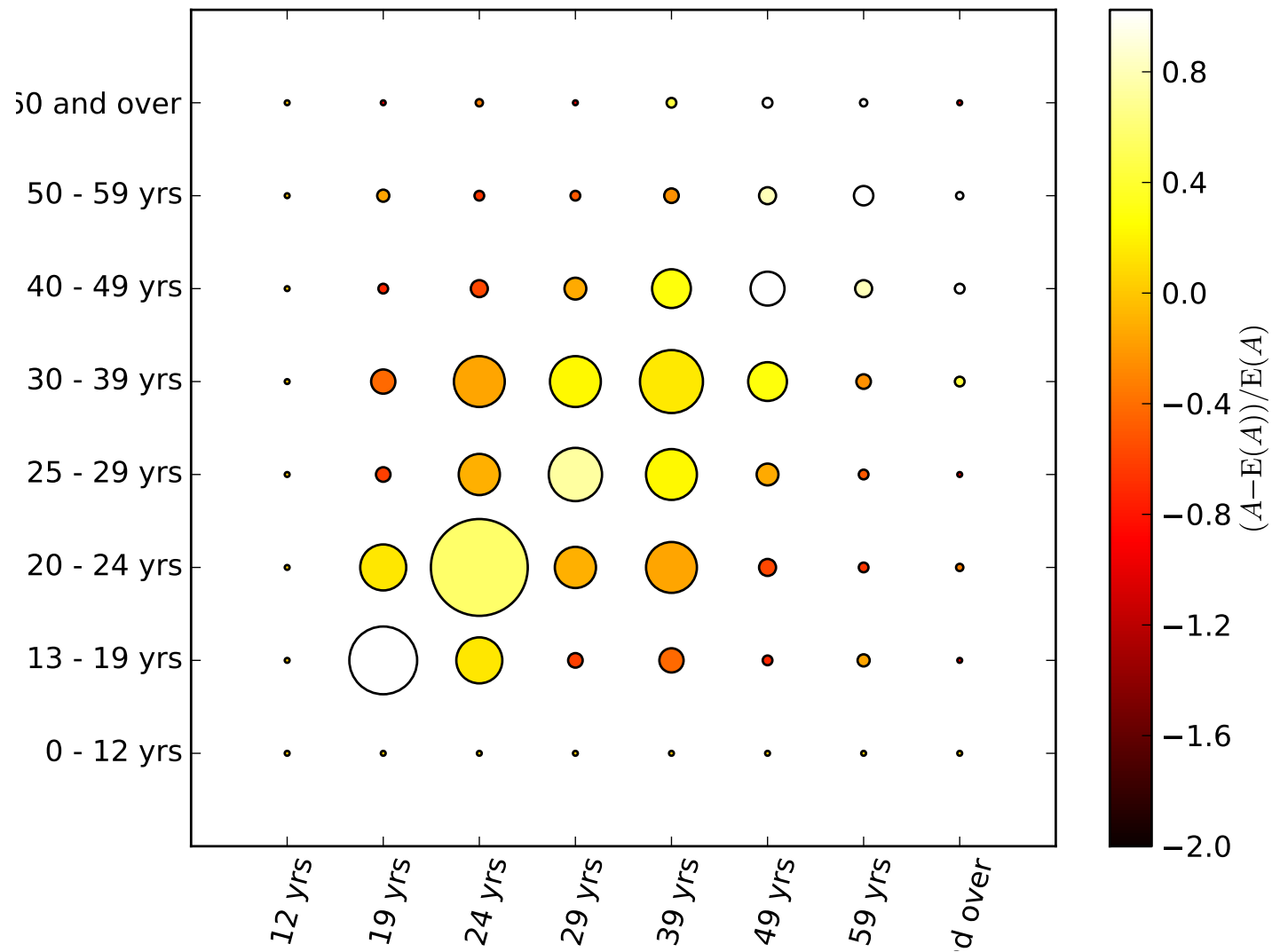


Clustering and race matrix

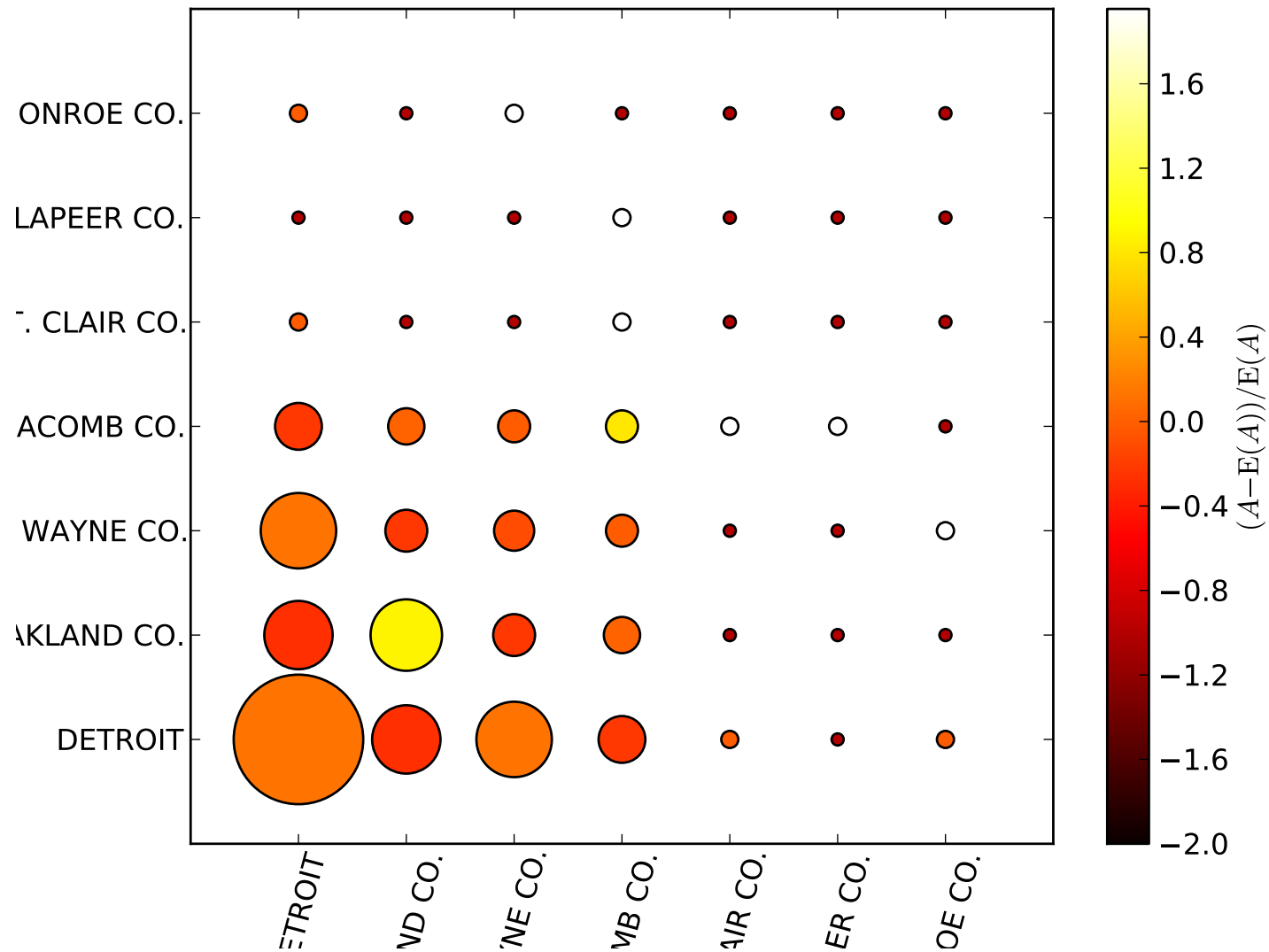


- Size indicates strength of phylogenetic linkage
- Color indicates deviation from null
- Sexual assortativity by race compounds differential disease burden (Morris et al., *AJPH*, 69, 2009)

Clustering and age

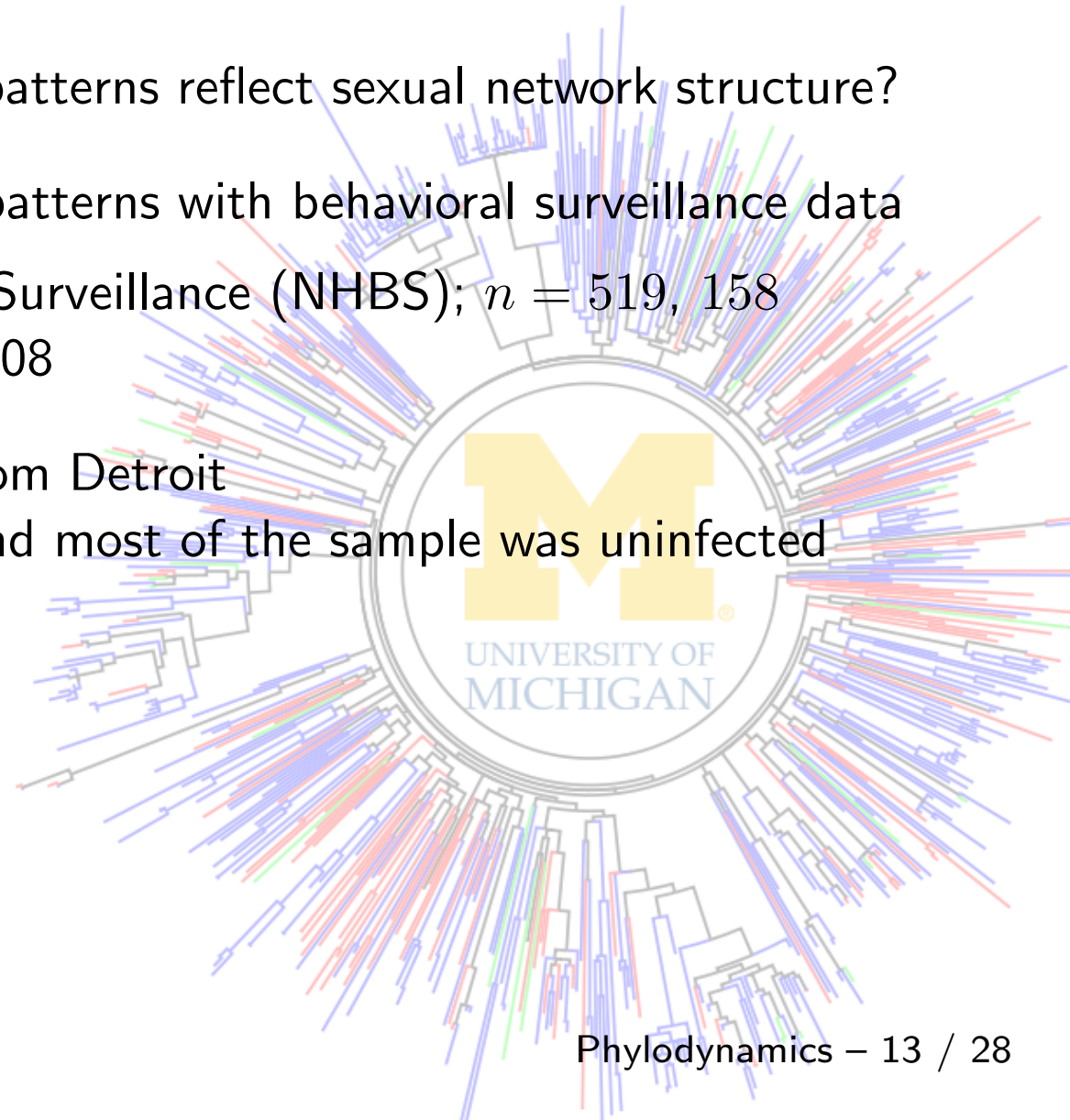


Clustering and location

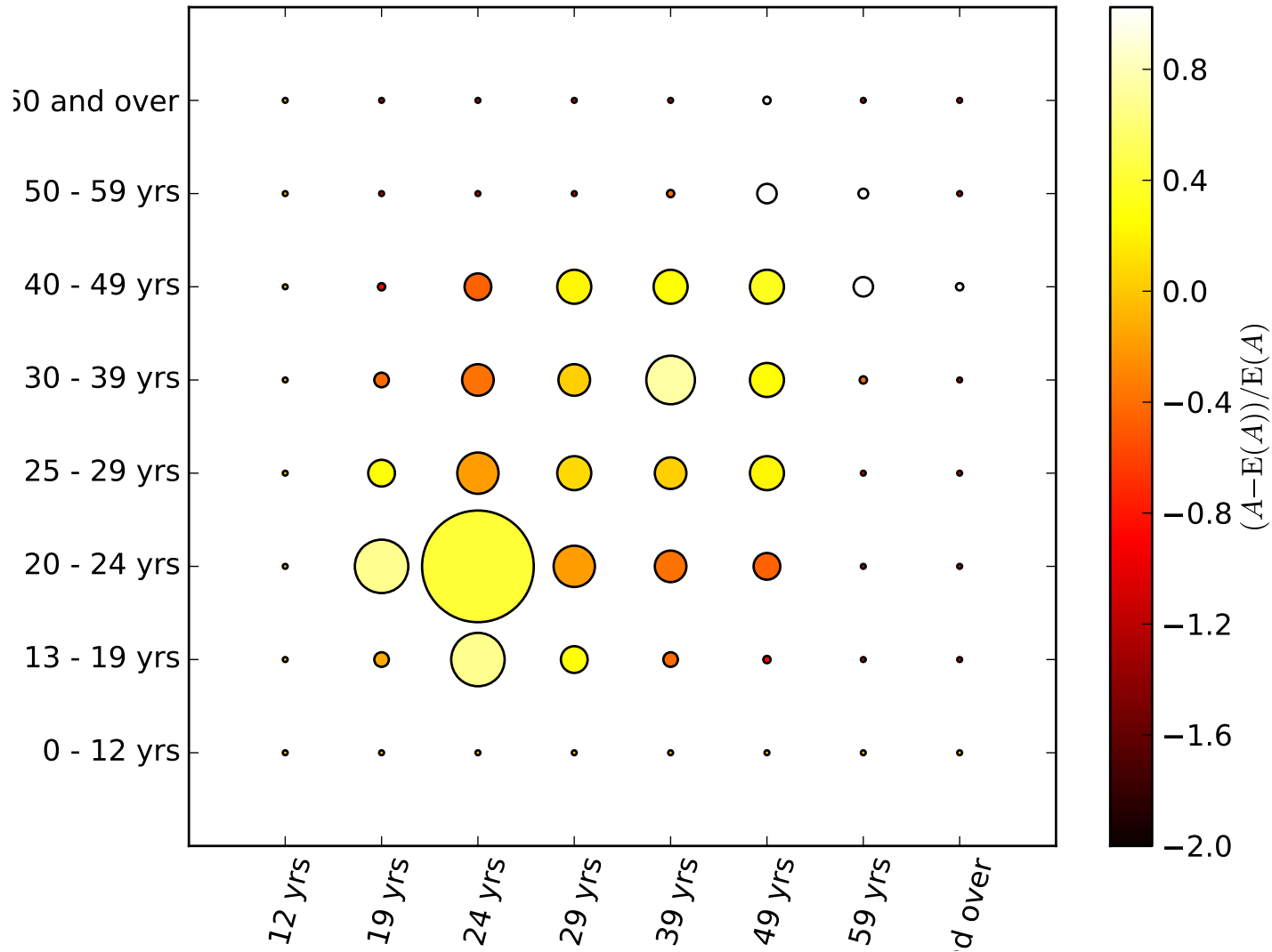


Clustering and age: behavioral surveillance

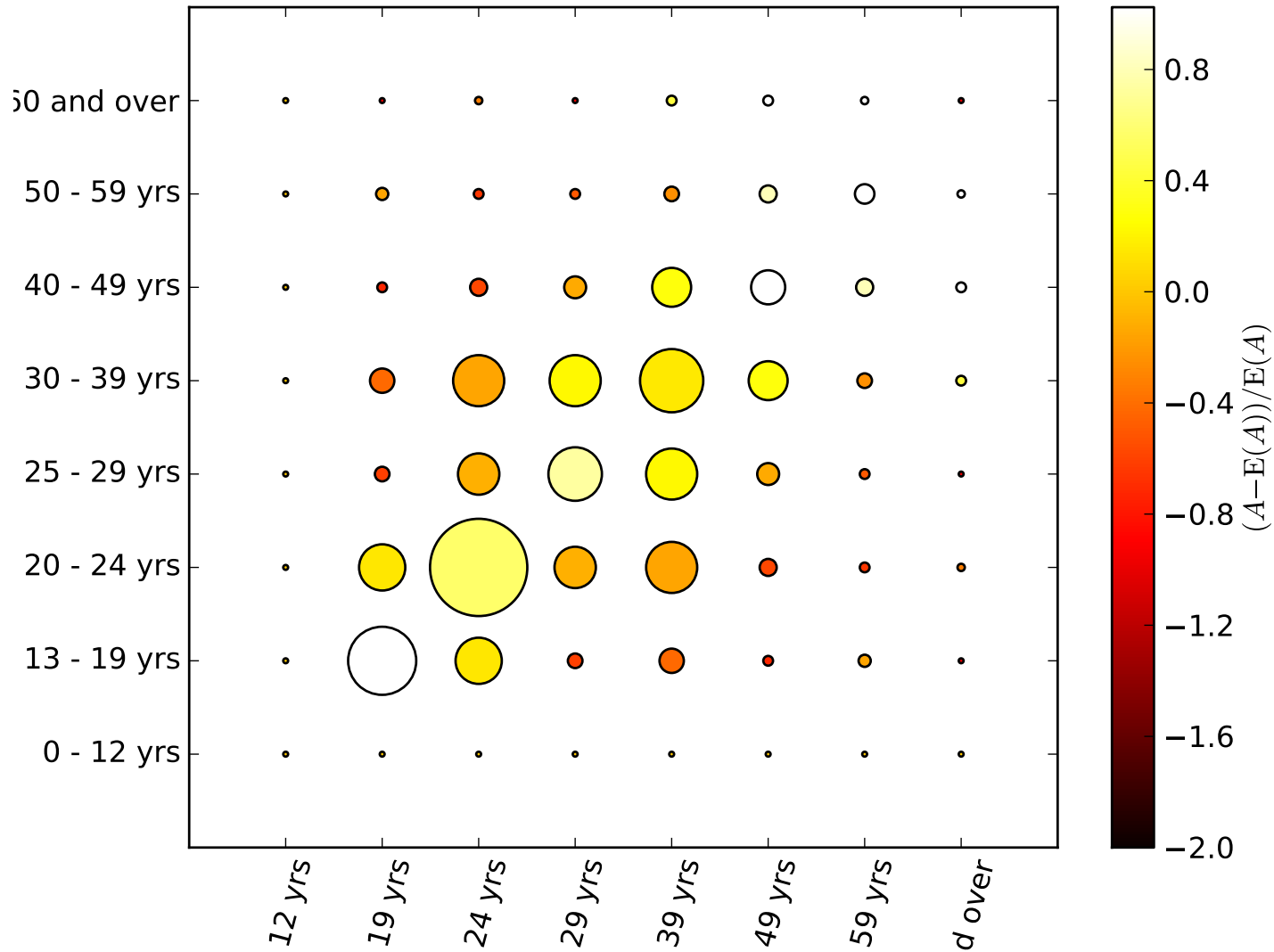
- To what extent does clustering patterns reflect sexual network structure?
- We can compare the clustering patterns with behavioral surveillance data
- National Health and Behavioral Surveillance (NHBS); $n = 519, 158$ partnerships reported, MSM , 2008
 - Sample was only collected from Detroit
 - Different age composition, and most of the sample was uninfected



Clustering and age: behavioral surveillance



Clustering and age: behavioral surveillance



Comparison to survey data

Newman's *coefficient of assortativity*:

$$r = \frac{\sum_i e_{ii} - \sum_i a_i b_i}{1 - \sum_i a_i b_i}$$

$$r_{race} = 38.2\%$$

$$r_{county} = 6.9\%$$

$$r_{age} = 16.1\%$$

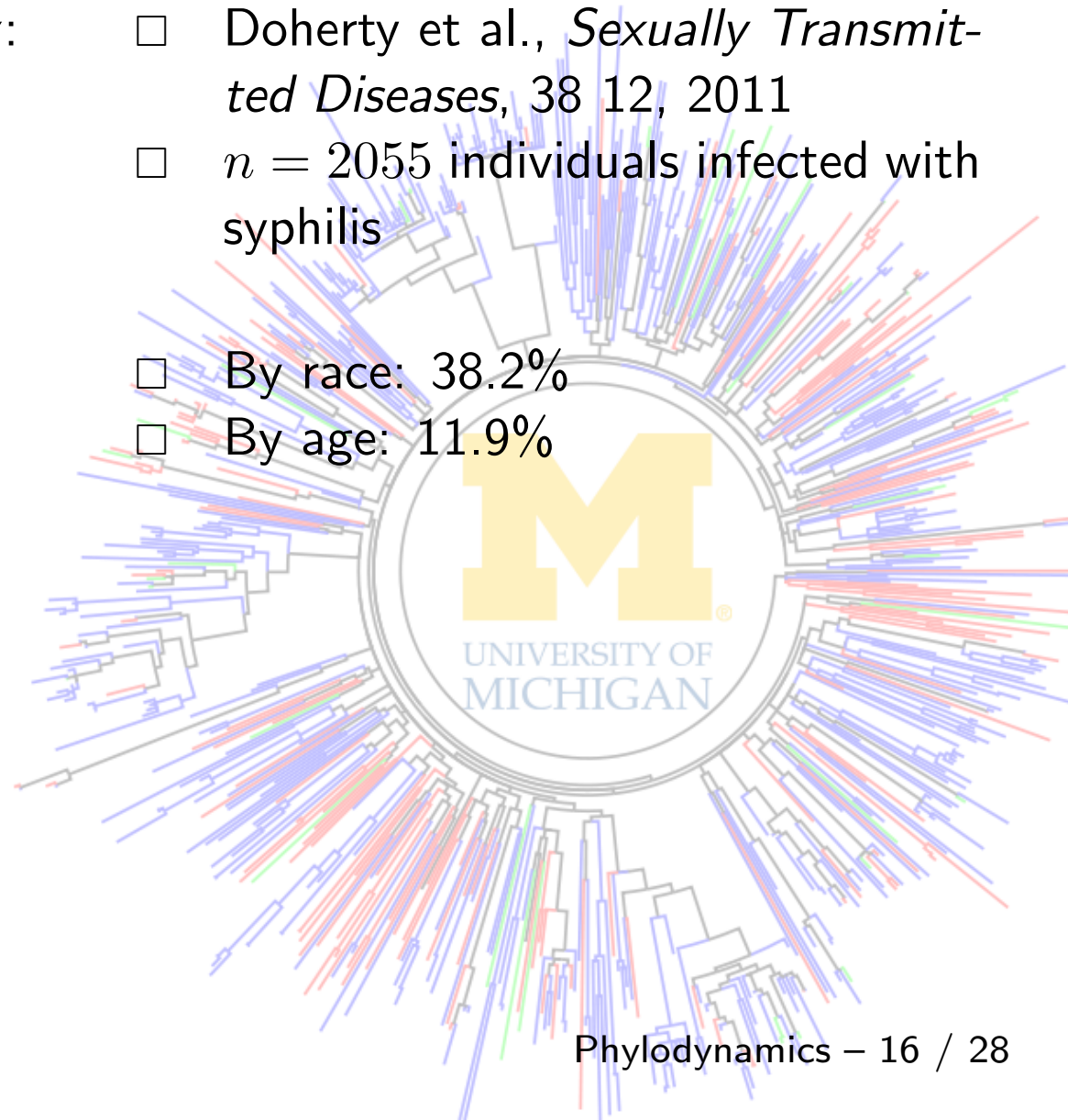
$$\text{NHBS } r_{age} = 6.5\%$$

□ Doherty et al., *Sexually Transmitted Diseases*, 38 12, 2011

□ $n = 2055$ individuals infected with syphilis

□ By race: 38.2%

□ By age: 11.9%



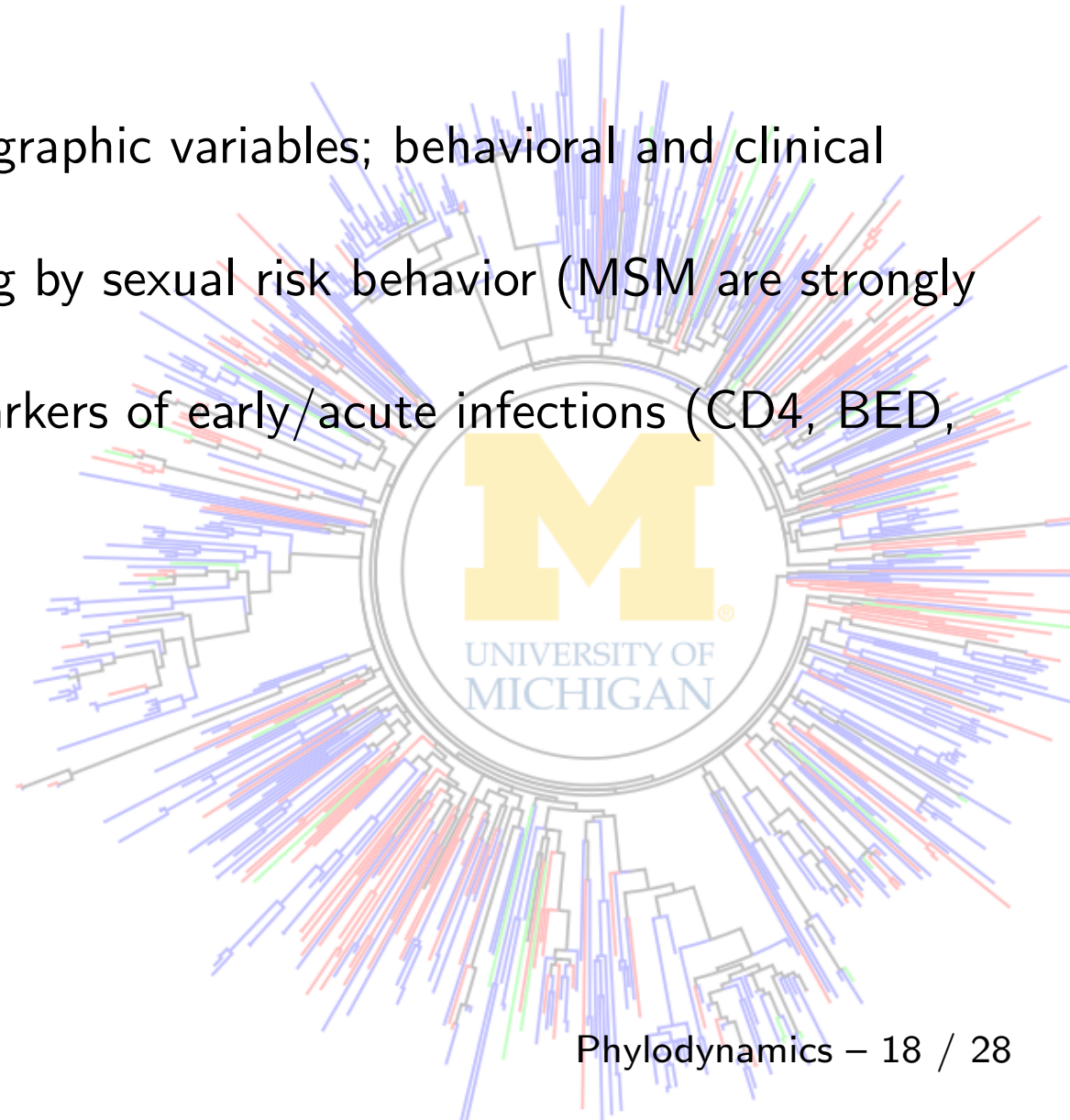
Comparison to survey data

| <i>Study</i> | <i>Assortativity*</i> | | |
|--------------|-----------------------|------------------|-------------------|
| | Degree | Age | Ethnicity |
| Antiviral | -- | 0.08 | 0.48 |
| Baltimore | 0.08 | 0.11 | 0.34 |
| Bushwick | 0.08 | 0.11 | 0.62 |
| Chlamydia | 0.13 | 0.22 | 0.22 |
| Flagstaff | -0.03 | 0.18 | 0.35 |
| GC1981 | 0.03 | 0.14 | 0.38 |
| HIV | 0.11 | 0.14 | 0.27 |
| Houston | 0.15 | 0.10 | 0.73 |
| Manitoba | 0.18 | 0.20 | -- |
| PPNG | 0.16 | 0.12 | 0.05 |
| Project90 | 0.03 | 0.10 | 0.18 |
| Rockdale | -0.11 | -0.08 | -0.43 |
| Syph318 | 0.03 | -0.08 | -- |
| Urban | 0.05 | 0.07 | 0.02 |
| Urban2 | -0.08 | 0.07 | 0.10 |
| <i>mean</i> | 0.06 | 0.10 (0.0986) | 0.25 (0.25432) |

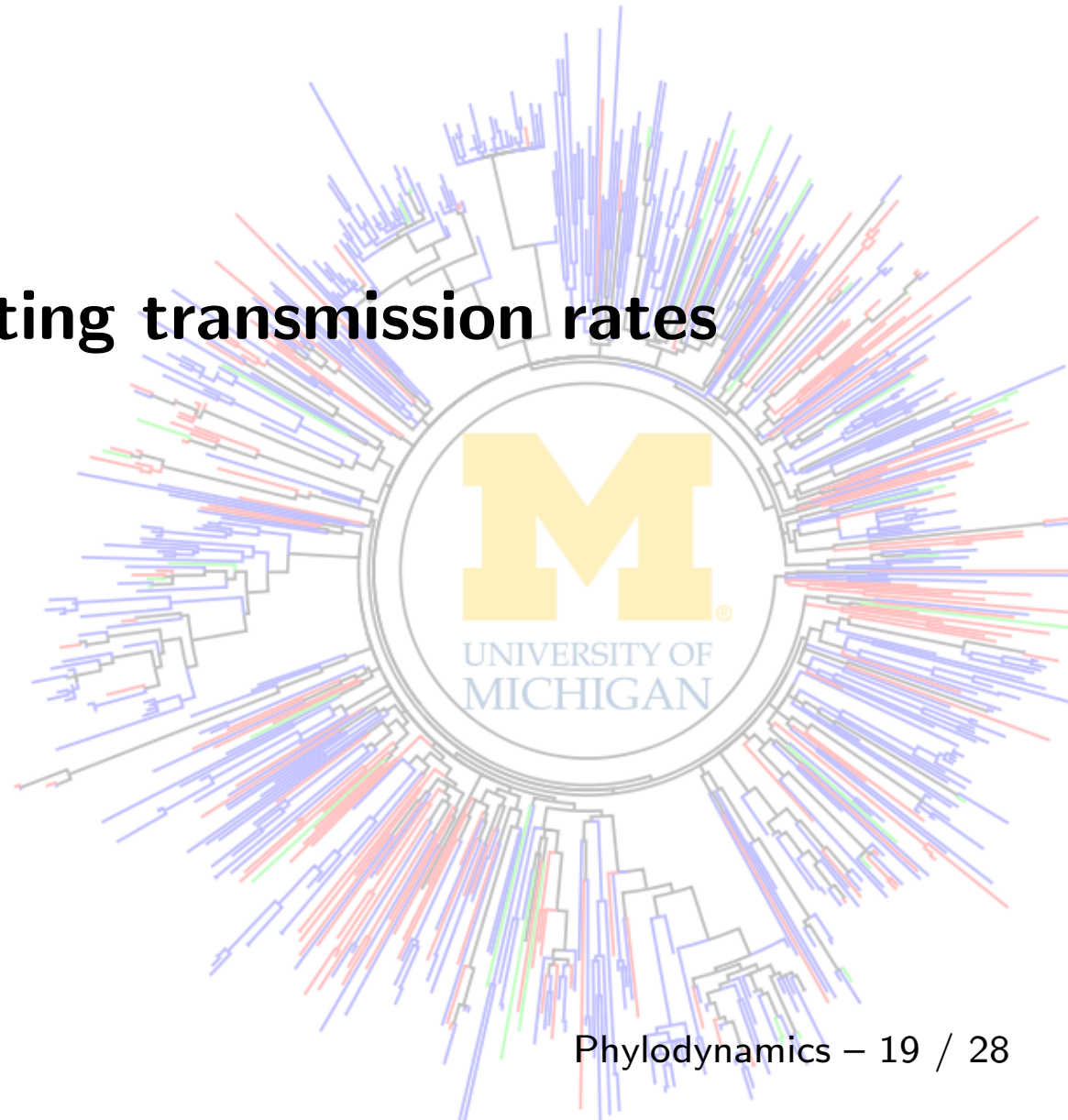
- Muth et al. have conducted a meta-analysis of many social network surveys
- Assortativity is almost always greater for race than age
- We found relative high assortativity by both race and age

Clustering: future work

- We have mostly examined demographic variables; behavioral and clinical variables are also important!
- We also observe strong clustering by sexual risk behavior (MSM are strongly clustered)
- And we observe clustering by markers of early/acute infections (CD4, BED, viral diversity within host)



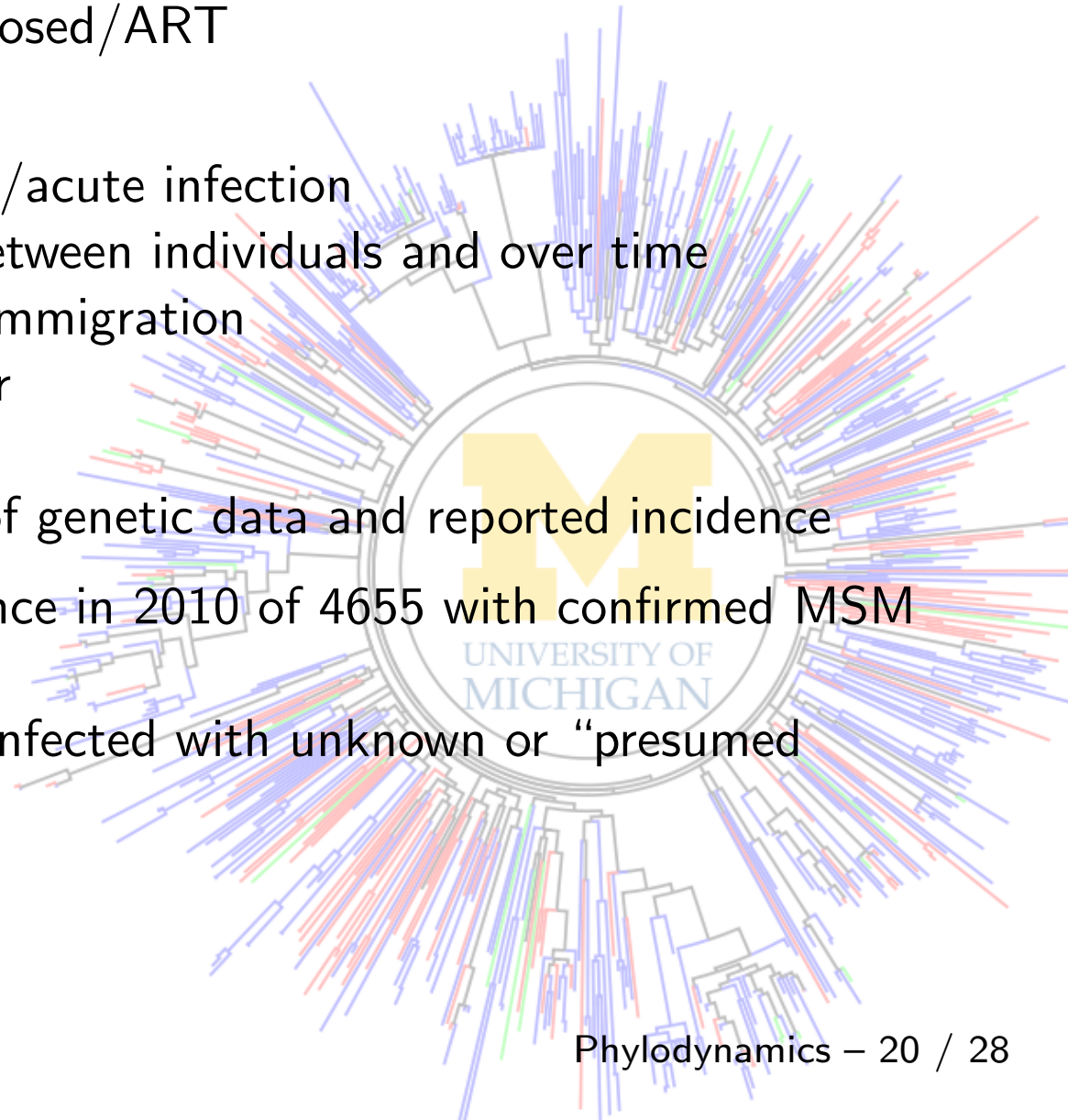
Estimating transmission rates



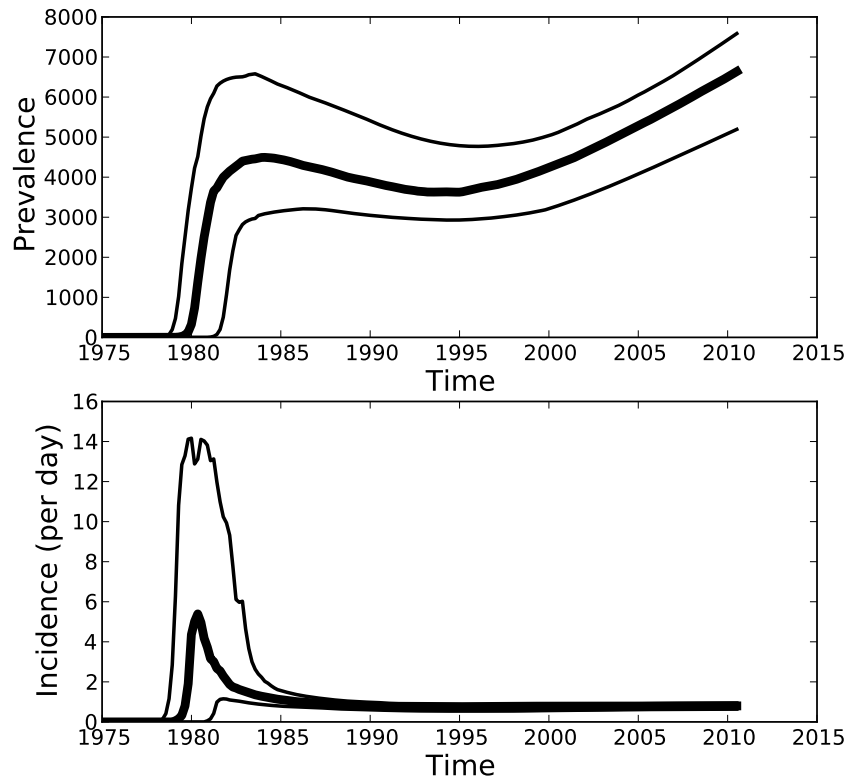
Building a model for HIV in SE Michigan

$S \rightarrow$ early/acute infection \rightarrow chronic infection \rightarrow AIDS/death
chronic infection \rightarrow diagnosed/ART

- High transmissibility during early/acute infection
- Heterogeneity of risk behavior between individuals and over time
- Population growth and decline; immigration
- Variable duration of risk behavior
- Introduction HAART in mid-90s
- Estimation used a combination of genetic data and reported incidence
 - MDCH data indicate prevalence in 2010 of 4655 with confirmed MSM risk factor
 - An additional 2030 men are infected with unknown or “presumed heterosexual” risk factors



Estimating transmission rates:results



- We estimate prevalence and incidence over the entire course of the epidemic
- Estimates based only on genetic data and MDCH estimates for 2010
- What this model gets right: Rapid rise during the 80s, plateau during the 90s, and gradual rise with HAART
- What it gets wrong: peak incidence is 81-82; mostly likely occurs 83-84; probably over estimates prevalence \sim 1985

Parameter estimates

| Variable | Median | 5pc | 95pc |
|------------------------------------|---------|---------|---------|
| β_1^{-1} (days) | 32.05 | 30.25 | 36.50 |
| β_2^{-1} (days) | 1373.61 | 1000.42 | 1896.85 |
| Fraction early/acute transmissions | 0.56 | 0.47 | 0.65 |
| Start of epidemic (years) | 31.70 | 29.70 | 32.43 |
| α | 7.36 | 4.39 | 11.77 |
| R_0 | 3.31 | 2.74 | 3.99 |

Acute-stage transmission rate is $\approx 43\times$ chronic-stage transmission

Using genetic data increases precision

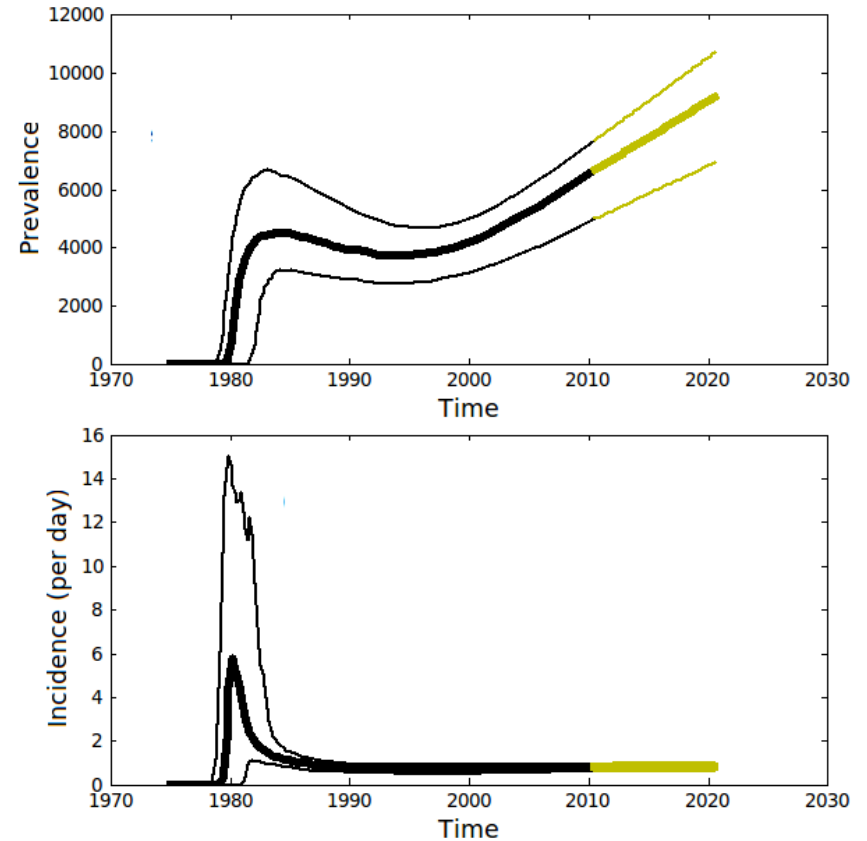
- Inference was based on both genetic data and independent estimate of prevalence in 2010
- What is the relative contribution of the genetic data and feature matching to the precision of parameter estimates?

| Variable | Median (FM) | Median (FM+G) | Design Effect |
|-----------------------|-------------|---------------|---------------|
| β_1^{-1} (days) | 38.04 | 32.05 | 0.47 |
| β_2^{-1} (days) | 1726.64 | 1373.61 | 0.75 |
| α | 8.24 | 7.36 | 0.92 |
| R_0 | 2.76 | 3.31 | 0.83 |

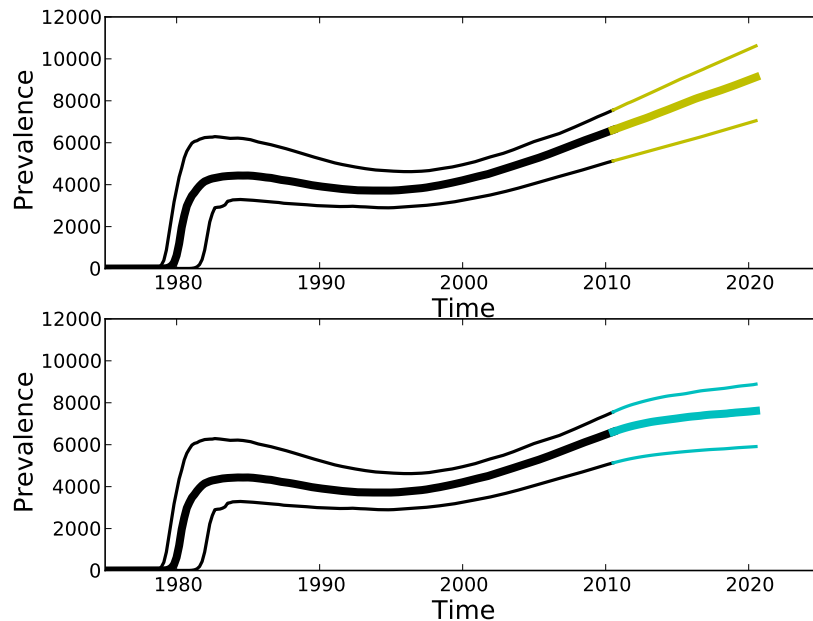
- Increase in precision is greatest for transmission rates

Predicted prevalence and incidence

- Fitting nonlinear dynamical models makes it easy to extrapolate prevalence over next decade
- ≈ 2800 new infections in MSM over the next 10 years; incidence stable

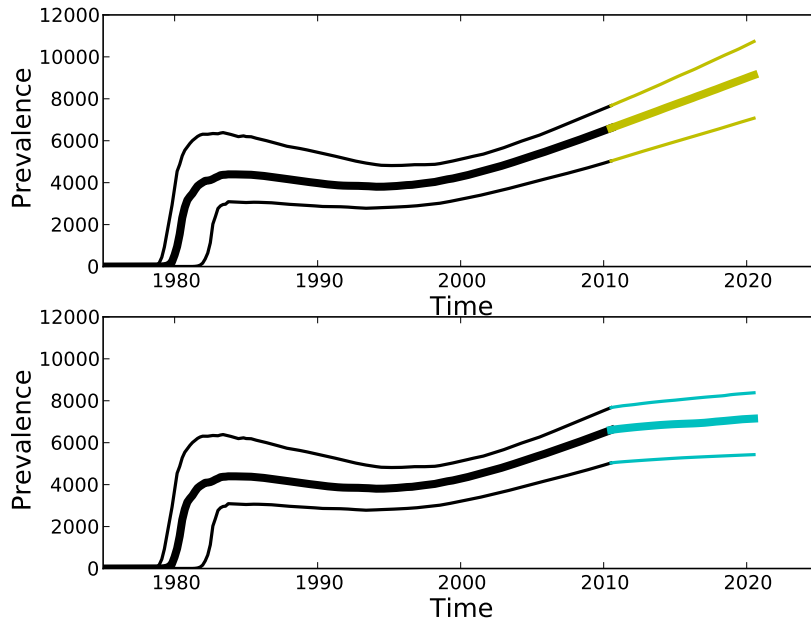


Predicted prevalence and T&T



- Fitting nonlinear dynamical models also makes it easy to explore potential interventions
- Test and Treat, optimistic scenario: 80% diagnosed within 2 years, immediately placed on ART, zero transmission while on ART
- ≈ 1500 infections in MSM over next 10 years
- 1300 infections averted

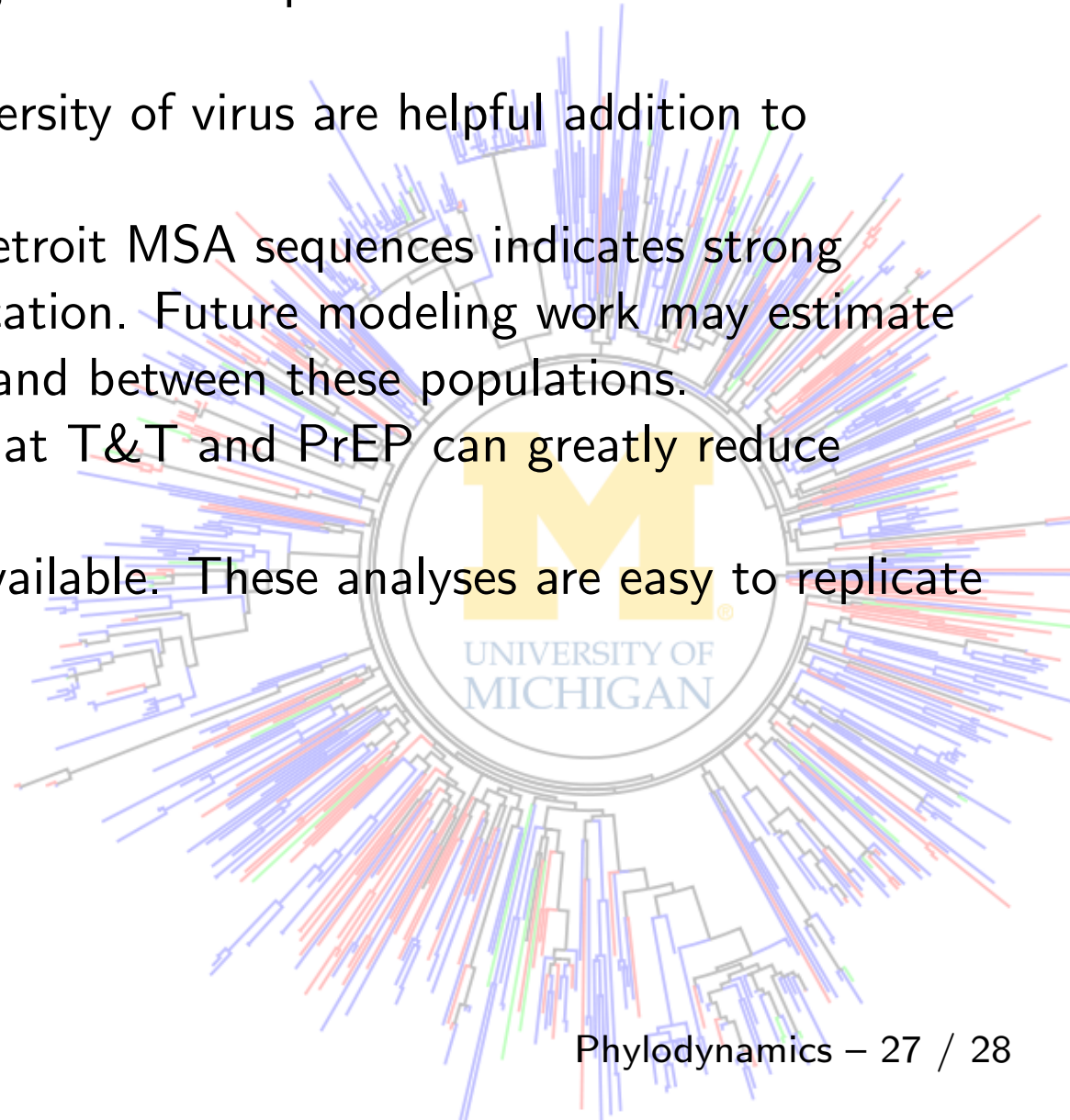
Predicted prevalence and PrEP



- PrEP, optimistic scenario: transmission rates reduced by 40%
- Only around 300 new infections over 10 years, 2500 averted
- Still not enough to get prevalence to fall

Conclusions

- The HIV epidemic remains a major threat to public health in the Detroit MSA
- Measurements of the genetic diversity of virus are helpful addition to epidemiological surveillance
- Phylogenetic clustering of the Detroit MSA sequences indicates strong assortativity by race, age and location. Future modeling work may estimate the rates of transmission within and between these populations.
- Initial modeling work suggests that T&T and PrEP can greatly reduce incidence
- HIV sequences are now widely available. These analyses are easy to replicate for other at-risk populations.



Thanks

- James S. Koopman, University of Michigan
- Ethan Romero Sevenson, University of Michigan
- Eve Mokotoff, MDCH
- Mary-Grace Brandt, MDCH
- Emily Higgins, MDCH
- Rieza Soelaeman, MDCH
- Simon Frost, University of Cambridge
- Andrew Leigh Brown, University of Edinburgh
- Melissa J Ward, University of Edinburgh

