

Navigating the genetic variability of HIV with interdisciplinary tools.

Art FY Poon

BC Centre for Excellence in HIV/AIDS, British Columbia, Canada



British Columbia
Centre *for* Excellence
in HIV/AIDS

What an interdisciplinary background might look like:

3 years

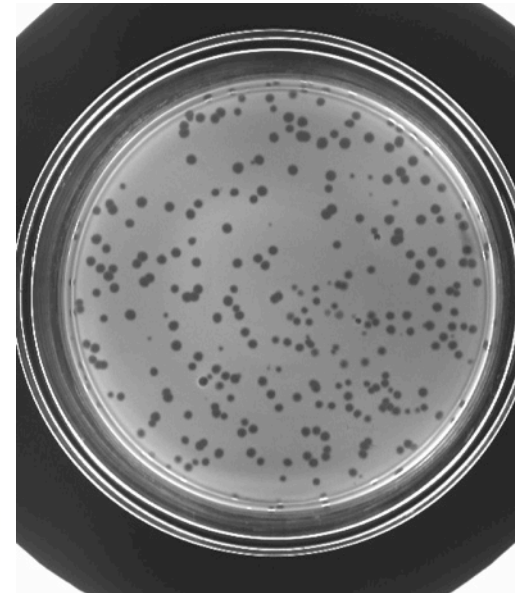


2 years

$$\lim_{\lambda \rightarrow 0} \frac{\lambda^n e^{-\lambda z'} (z')^{n-1}}{e^{-\lambda B} (\lambda B)^n \sum_{j=0}^{\infty} \frac{\Gamma(n)}{\Gamma(n+1+j)} (\lambda B)^j}$$

$$= \lim_{\lambda \rightarrow 0} \frac{e^{-\lambda z'} (z')^{n-1}}{e^{-\lambda B} (B)^n \left(\frac{1}{n} + \frac{\lambda B}{(n+1)n} + \dots \right)} = \frac{n(z')^{n-1}}{B^n}$$

5 years



5 years and counting..



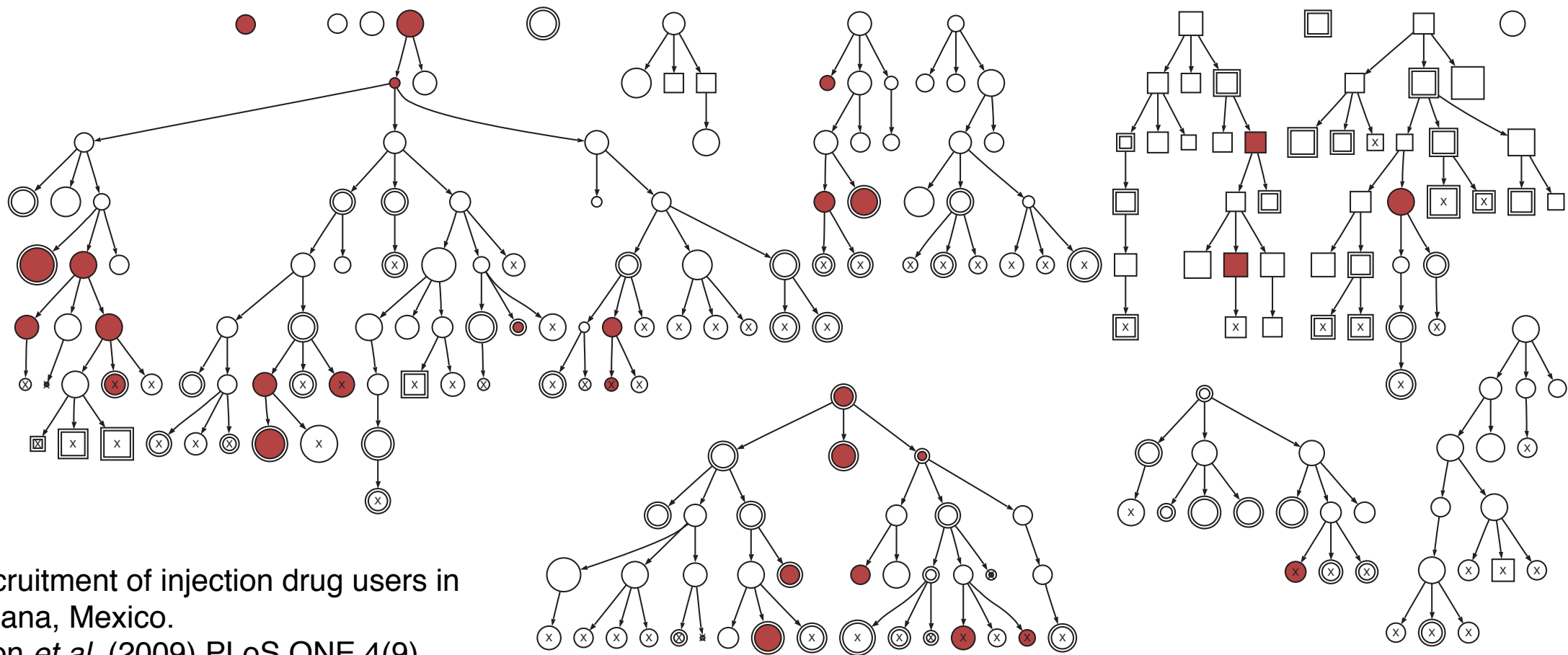
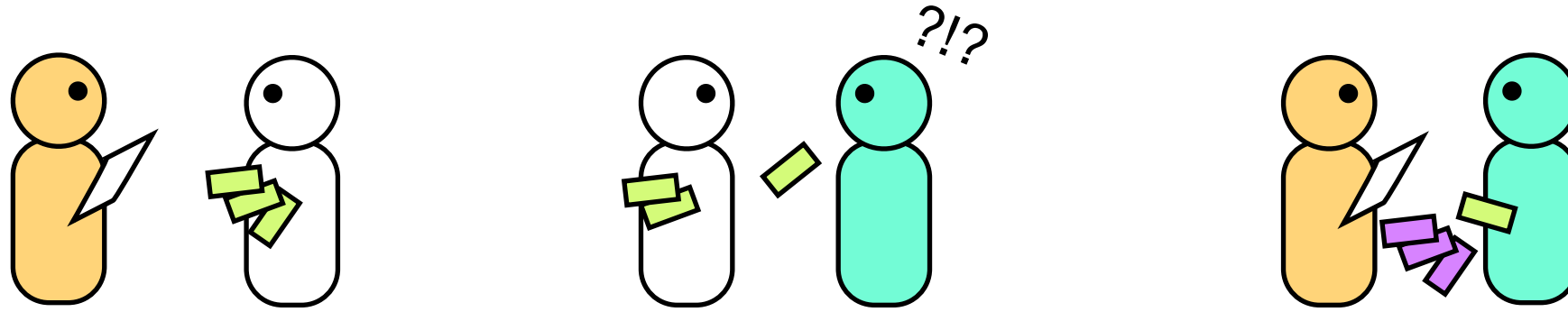
Quick outline

- 3 examples of using models from other disciplines to tackle problems in HIV.
 1. Using **linguistics** to analyze HIV study enrollment data.
 2. Identifying agents of selection using **artificial intelligence**.
 3. Inferring the date of HIV infection, and timing the emergence of adaptations, using **population genetics**.

Studying hidden populations.

- A **hidden population** is comprised of individuals sharing an attribute that is difficult to identify, carries a social stigma, or is illegal.
- *e.g.*, injection drug use.
- Such attributes can be associated with a **higher risk** of HIV infection.
- Can be sampled along **social networks** (peer referral).

Respondent-driven sampling (RDS).

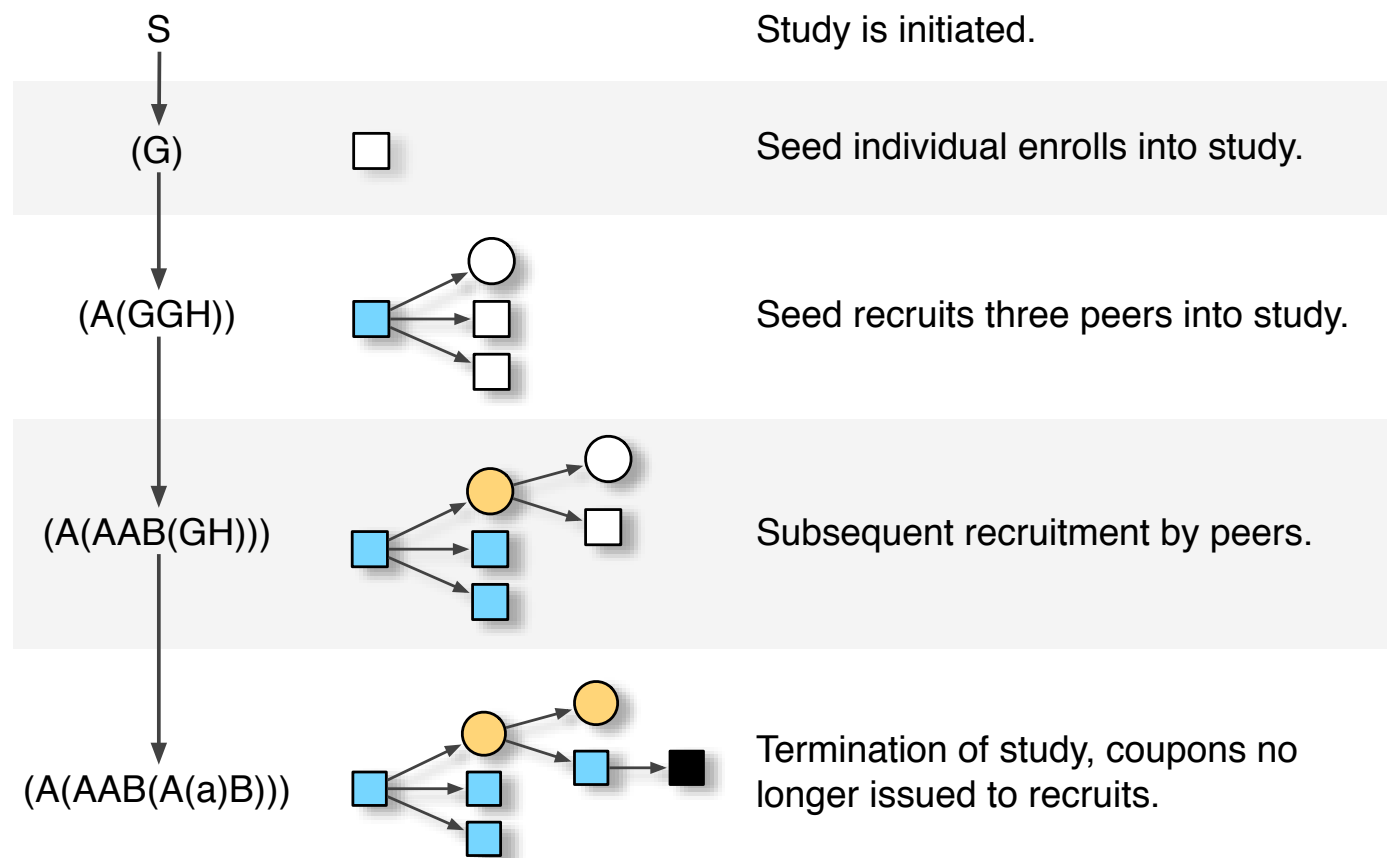
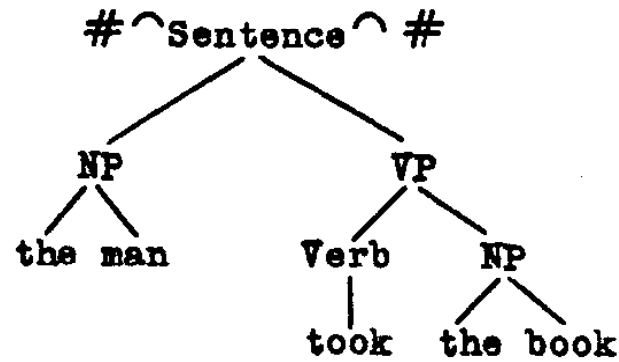


Recruitment of injection drug users in Tijuana, Mexico.
Poon *et al.* (2009) PLoS ONE 4(9).

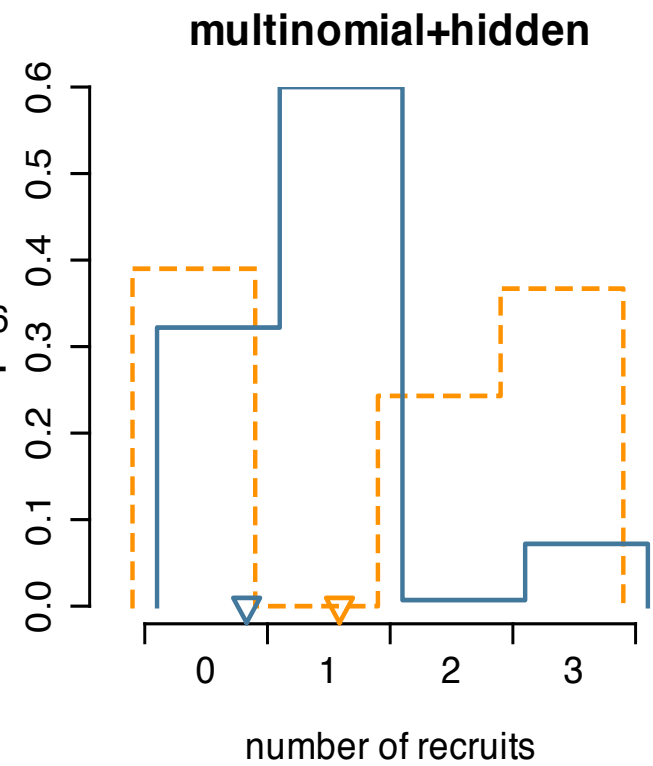
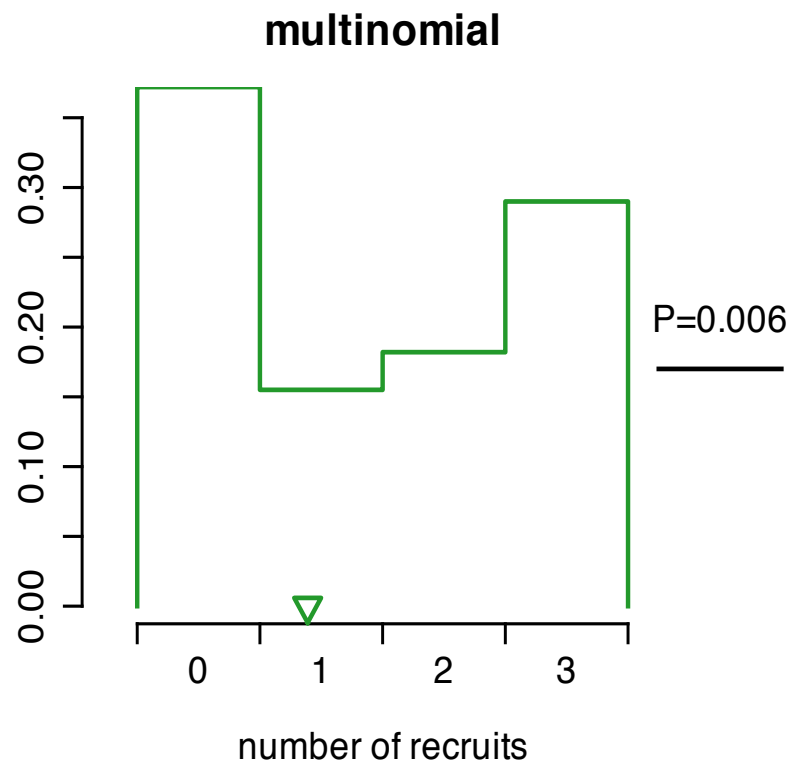
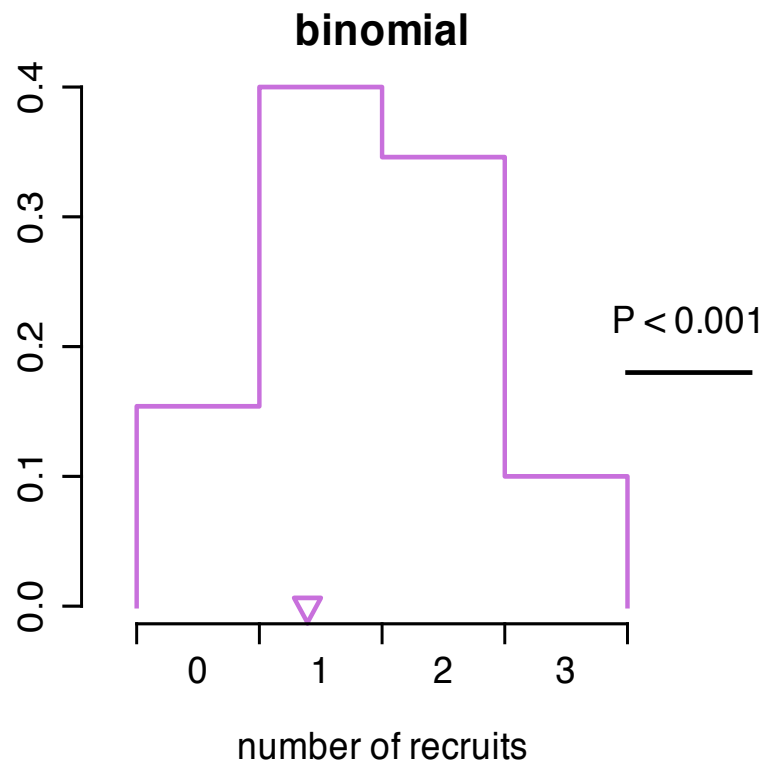
Interpreting RDS data.

- Study recruitment of injection drug user network is a “**branching process**” (tree-shaped).
- Is there variation in recruitment behaviour?
- Do recruits emulate the behaviour of their recruiters?
- A model from **linguistics** (Chomsky 1955) is useful for studying hidden variation in tree structures.

Using linguistics to analyze recruitment trees.



Recruits emulate the recruitment behaviour of their recruiters.



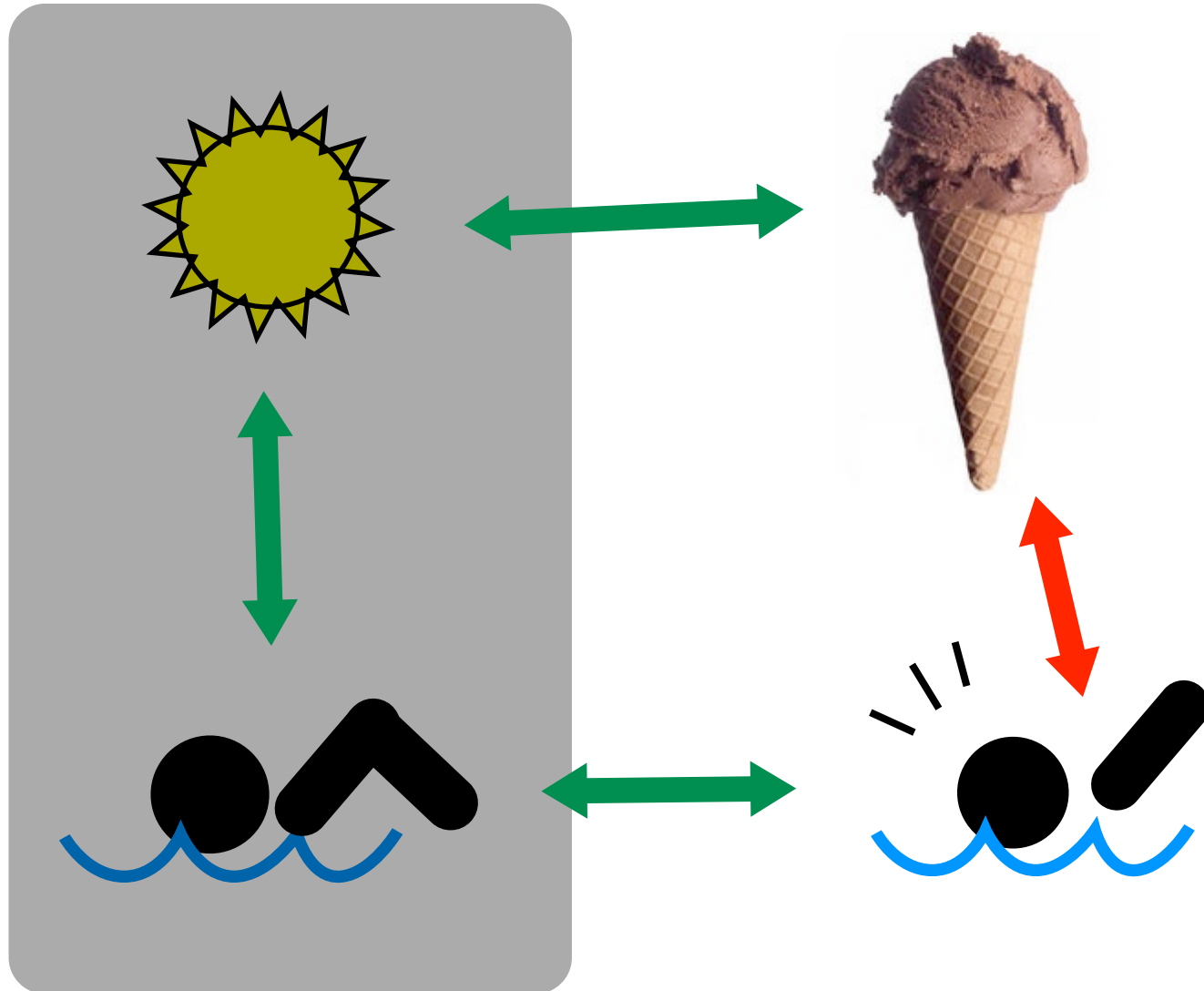
Agents of selection on HIV.

- Much of HIV evolution is driven by the **host-specific immune response**.
- *e.g.*, human leukocyte antigens (HLAs), antibodies
- encoded by highly variable regions of the human genome.
- Can offer significant protection from HIV infection (HLA B*5701 in elite controllers).

Identifying agents of selection.

- A statistical association between a polymorphism in the HIV genome and a factor in the environment can be due to:
 1. a “true” effect of the factor on HIV;
 2. confounded by a correlation between factors in the environment;
 3. confounded by an epistatic interaction between polymorphisms in HIV.
- **How can we tell these apart?**

Confounding - the “ice-cream” problem.



Consider a public swimming pool.

Ice-cream sales are correlated with drownings.

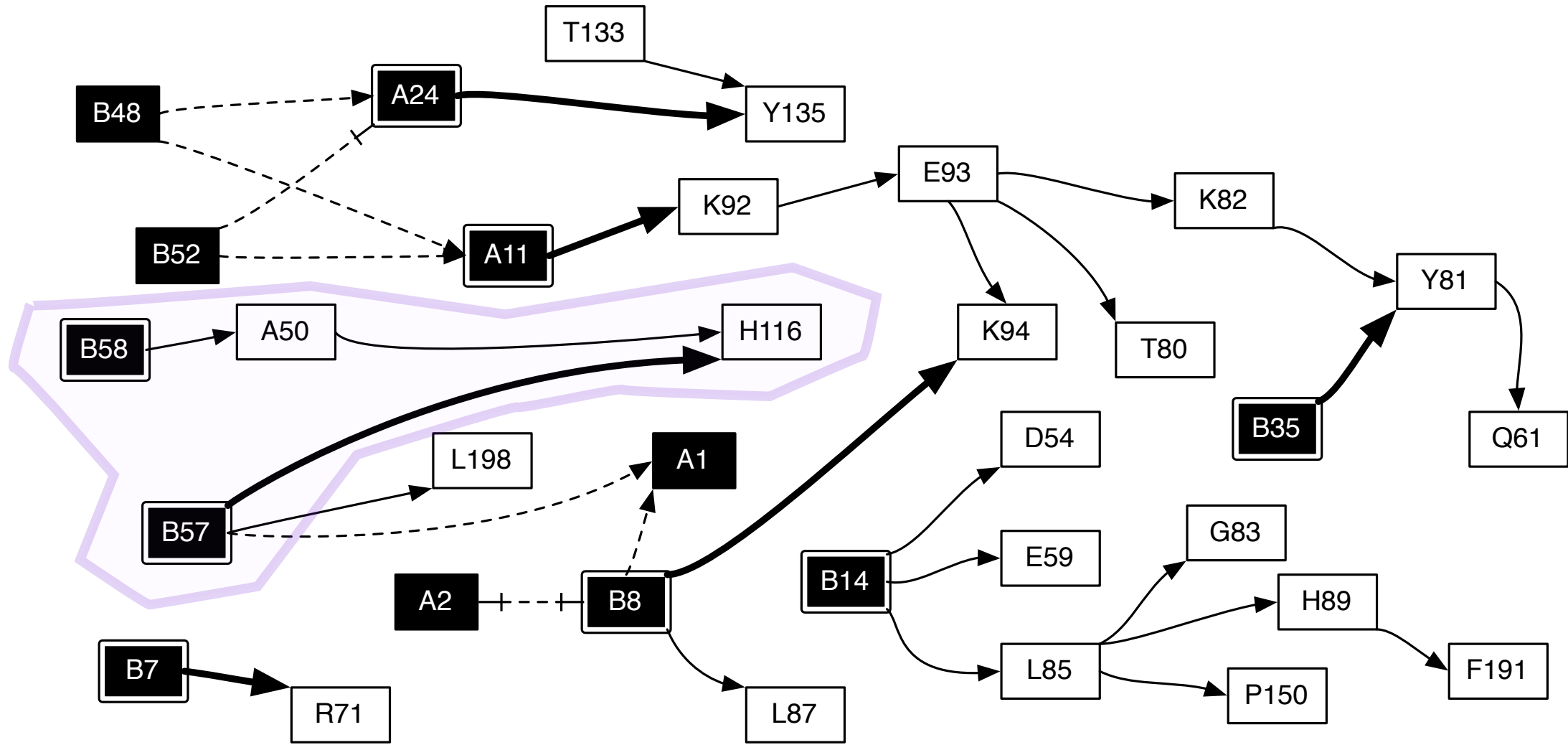
Bad A.I. (**red**) would conclude that eating ice-cream makes you drown.

Good A.I. (**green**) would recognize that this association is confounded by sunny days.

Bayesian networks.

- A class of models from artificial intelligence (machine learning).
- Variables are represented as “**nodes**” in a graph.
- Arrows (**edges**) are drawn to represent a conditional dependence of variable X on Y .
- Originates from Sewall Wright’s method of path coefficients.

Graph inferred from analysis of HIV-1 and HLA variation among $n=886$ patients in BC.



Studying within-host HIV evolution.

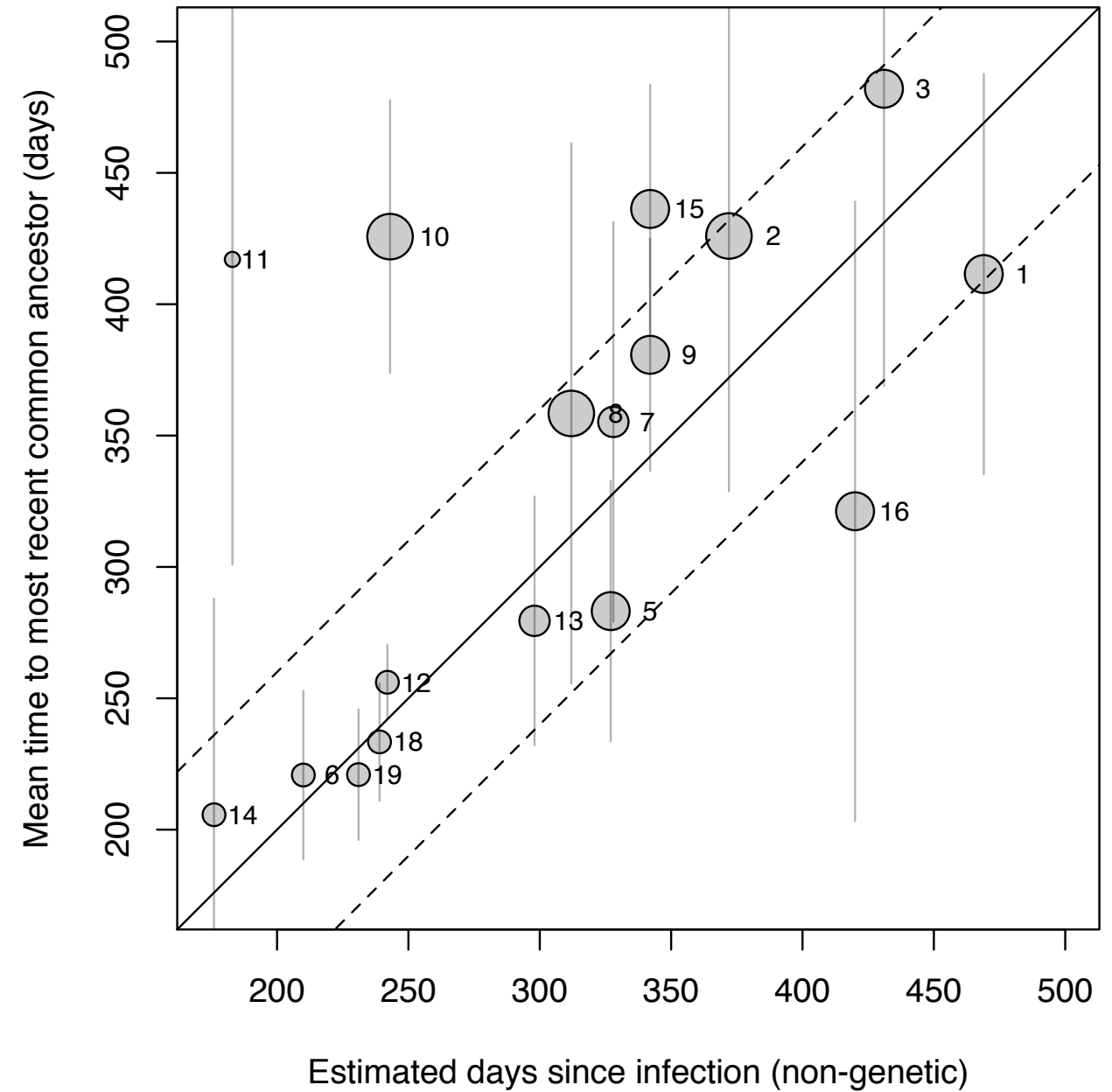
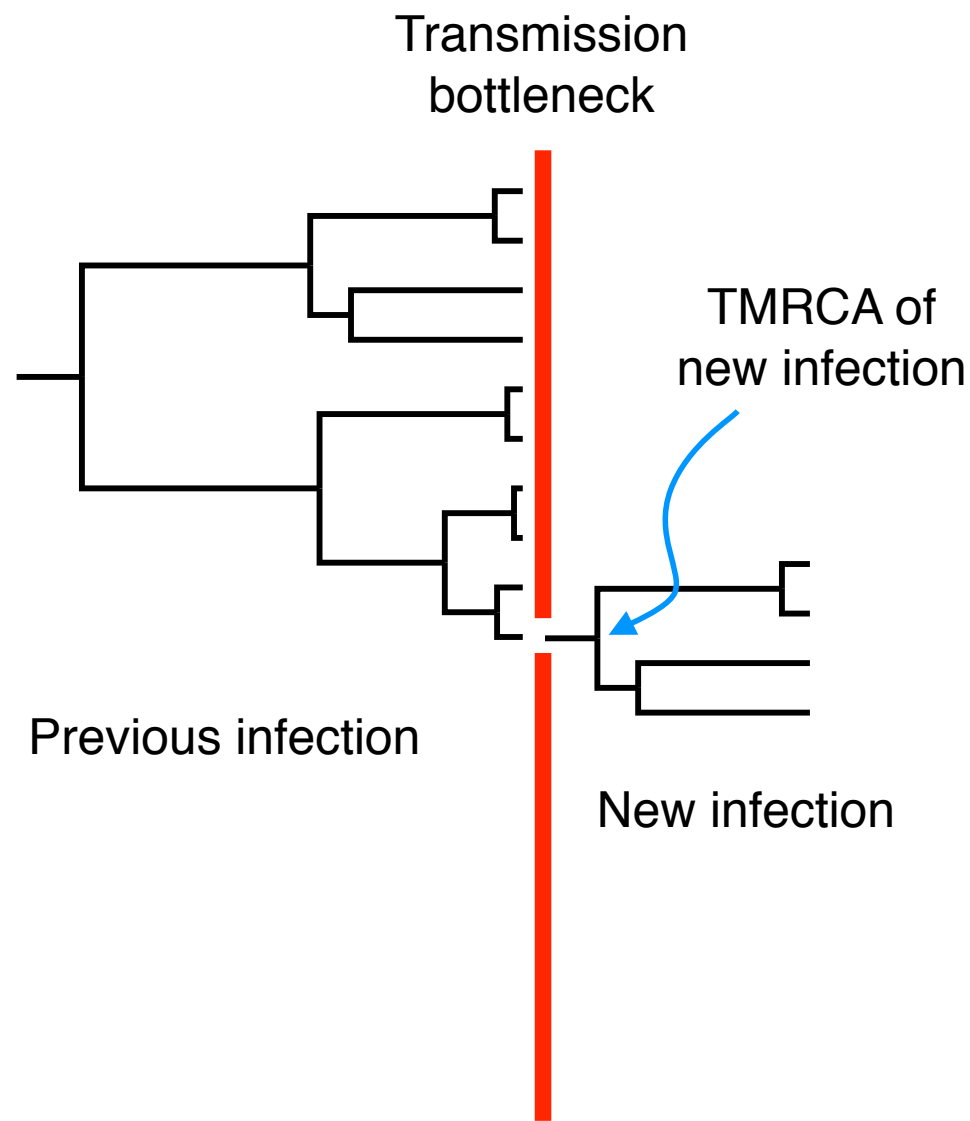
- **Next-generation sequencing** enables us to sequence 1,000's of HIV genomes from an infection in one run.
- Ideal raw material for analyzing how HIV evolution unfolds within a patient.
- No software - we have to make the tools ourselves.

The coalescent.

- A retrospective model from **population genetics**.
- Predicts the time since the **most recent common ancestor** of a random sample from a population.
- Combine with a model of molecular evolution (**molecular clock**) and serial samples.
- Can estimate the time of infection; when certain mutations emerged.

3. Timing HIV infections.

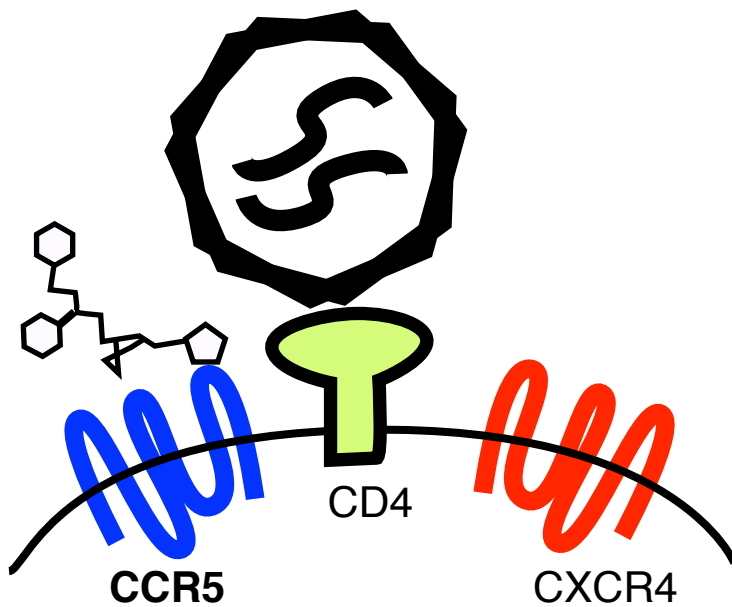
HIV transmission bottleneck causes time to MRCA to predict time of infection.



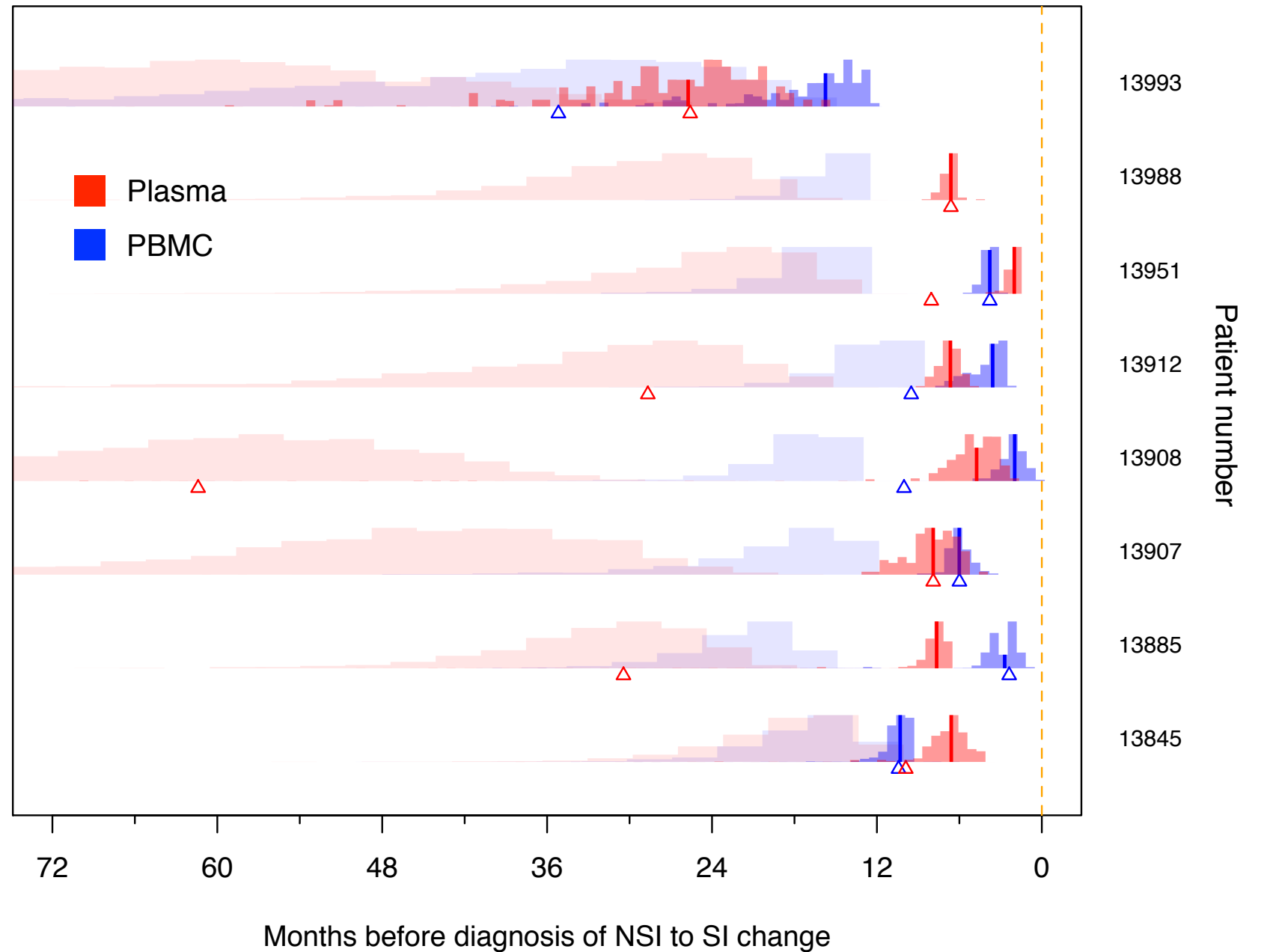
3. Timing HIV infections.

Reconstruct HIV evolution along coalescent tree to time emergence of adaptations.

HIV needs to bind one of two coreceptors to enter a host cell.



Some drugs block the CCR5 coreceptor, but HIV can switch to using CXCR4.



Acknowledgements



This work was supported by a **CIHR Fellowships Award** and by a **CIHR Operating Grant** (HIV/AIDS Research Initiative - Biomedical/Clinical Stream).



Julio Montaner
P. Richard Harrigan

BCCfE Molecular Lab
Luke Swenson
Rachel McGovern



Zabrina Brumme
Simon Fraser University



Sergei Pond
Joel Wertheim



Mark Wainberg
Lady Davis Institute, Jewish General Hospital, McGill University



Angélique B. van't Wout
Academisch Medisch Centrum, Universiteit van Amsterdam