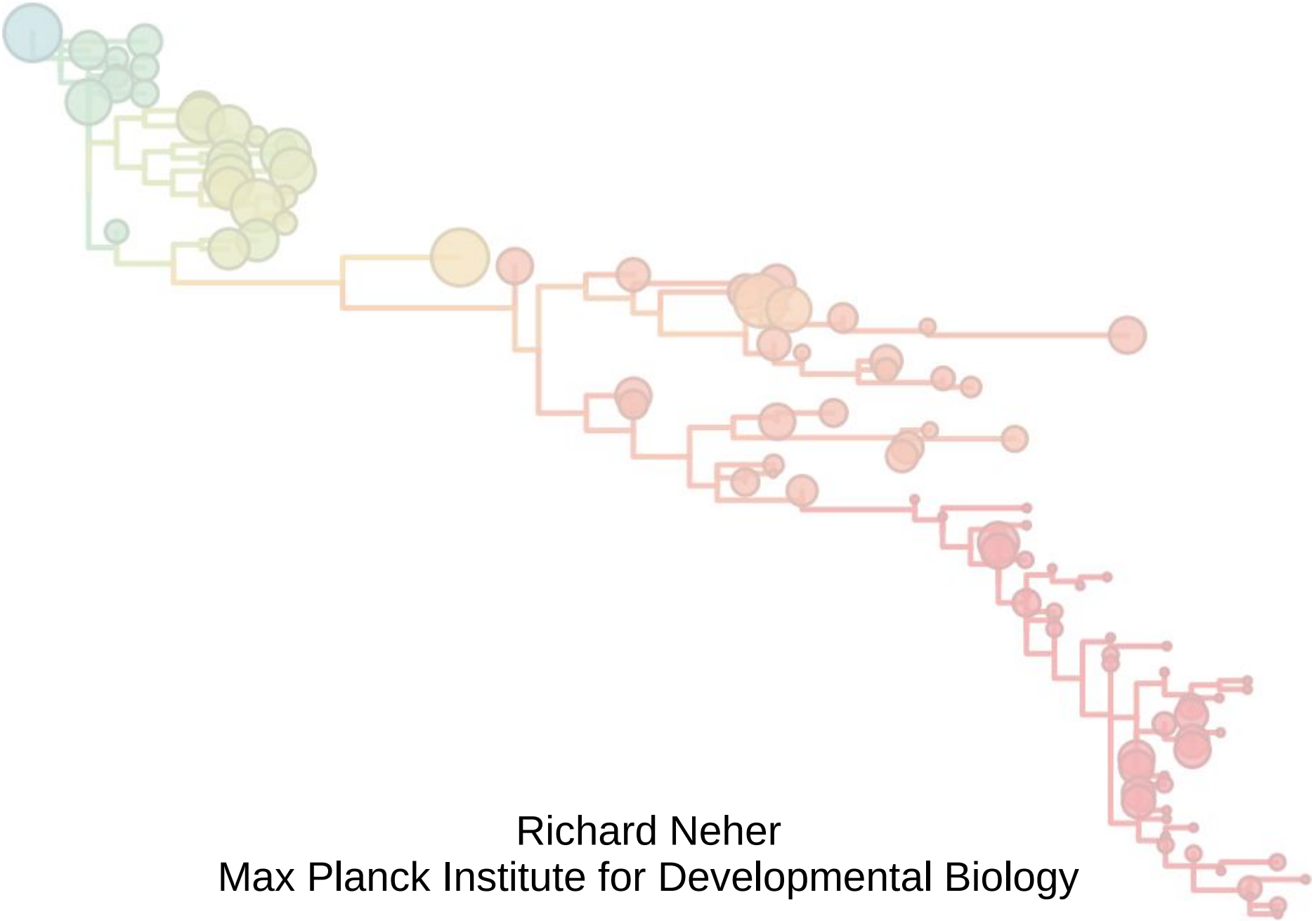
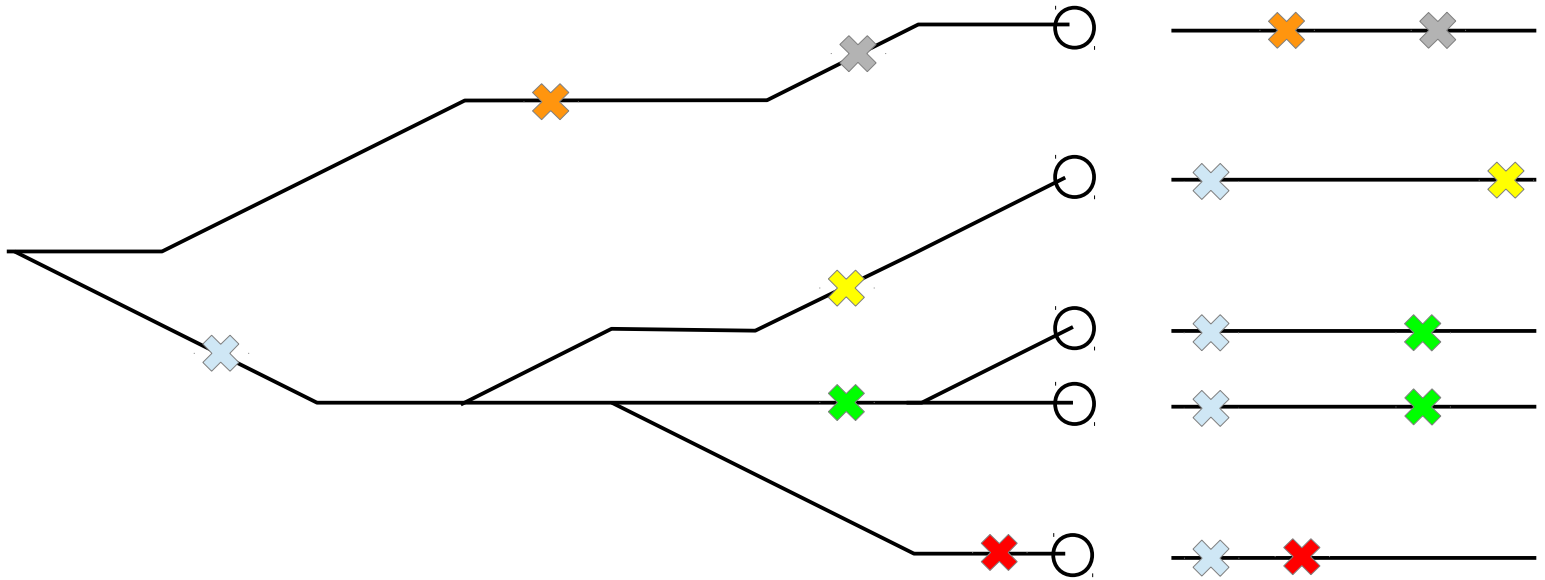
Quantifying and predicting the evolution of RNA viruses

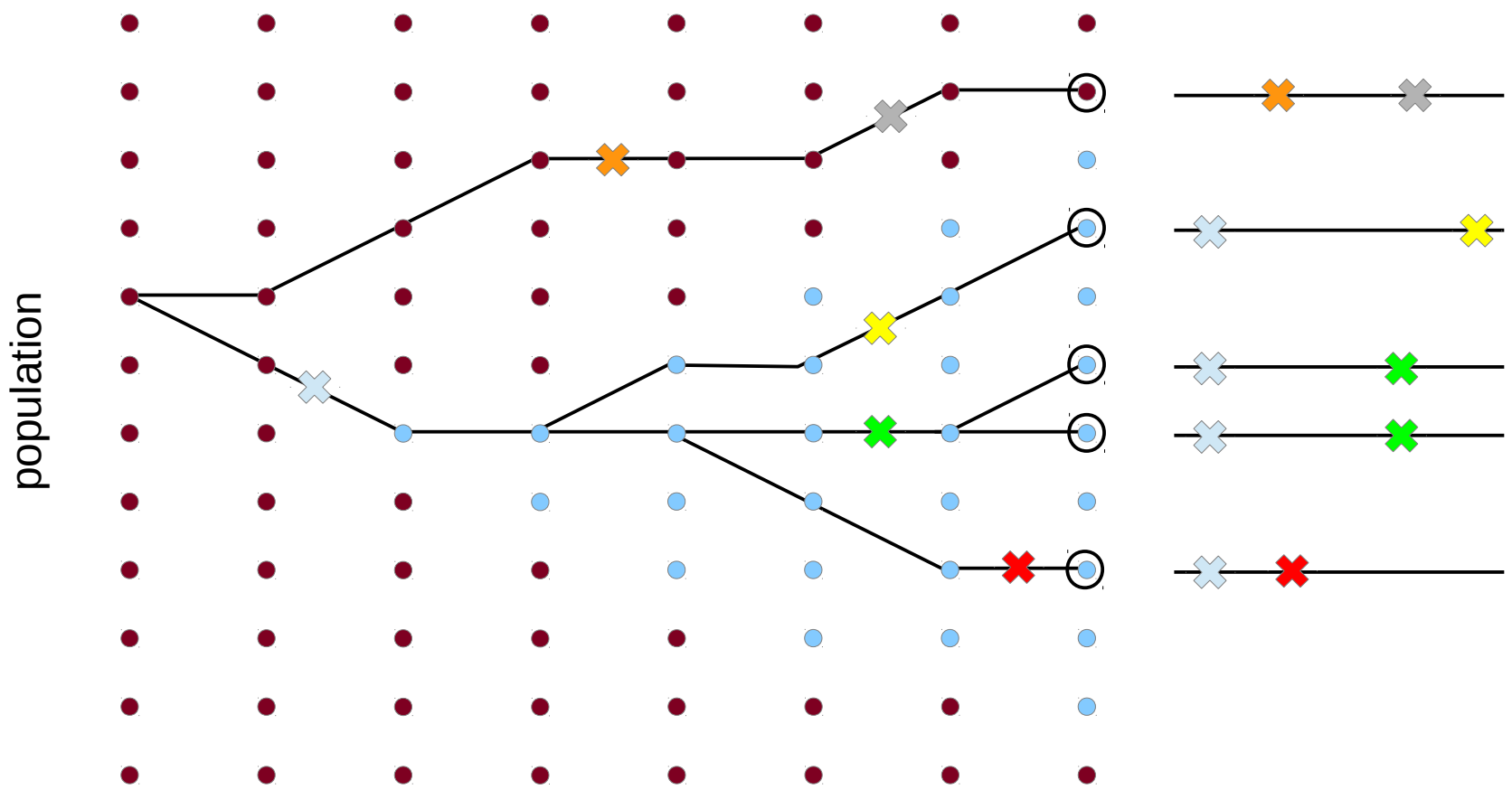


Richard Neher
Max Planck Institute for Developmental Biology

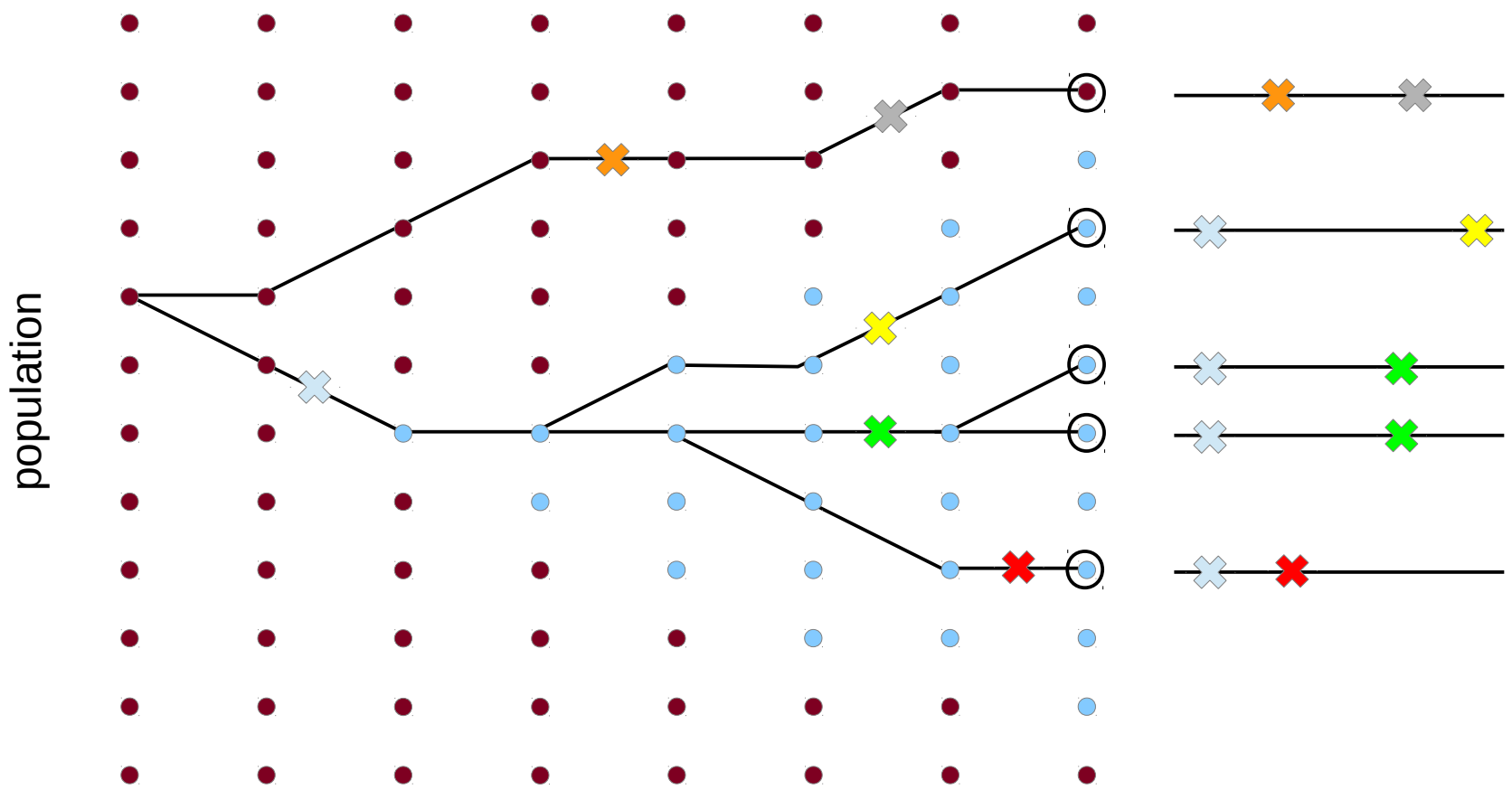
Slowly evolving populations



Slowly evolving populations



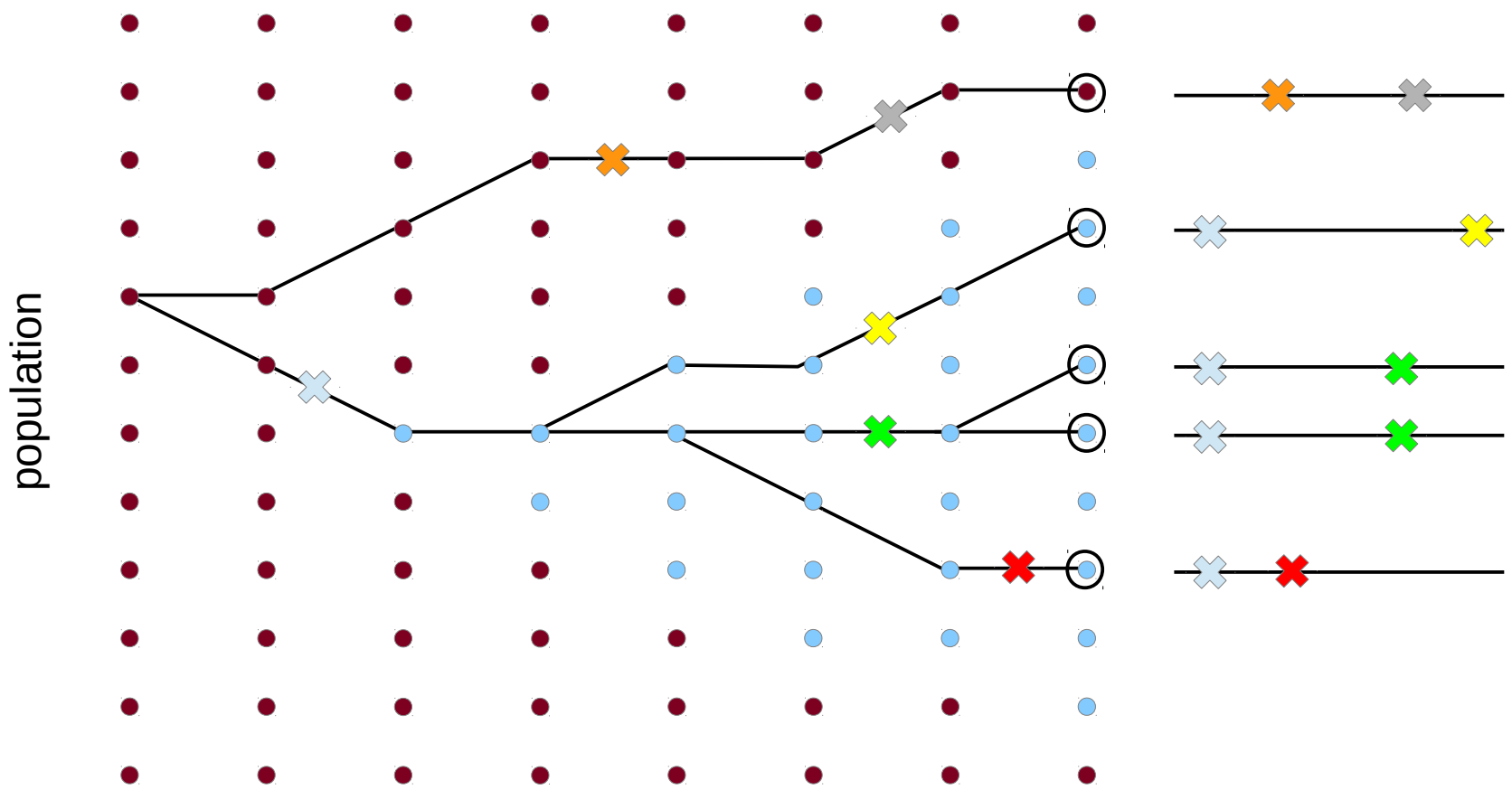
Slowly evolving populations



evolutionary processes ↔ statistics of trees ↔ patterns of genetic diversity

→ time resolved data allows direct observation

Slowly evolving populations

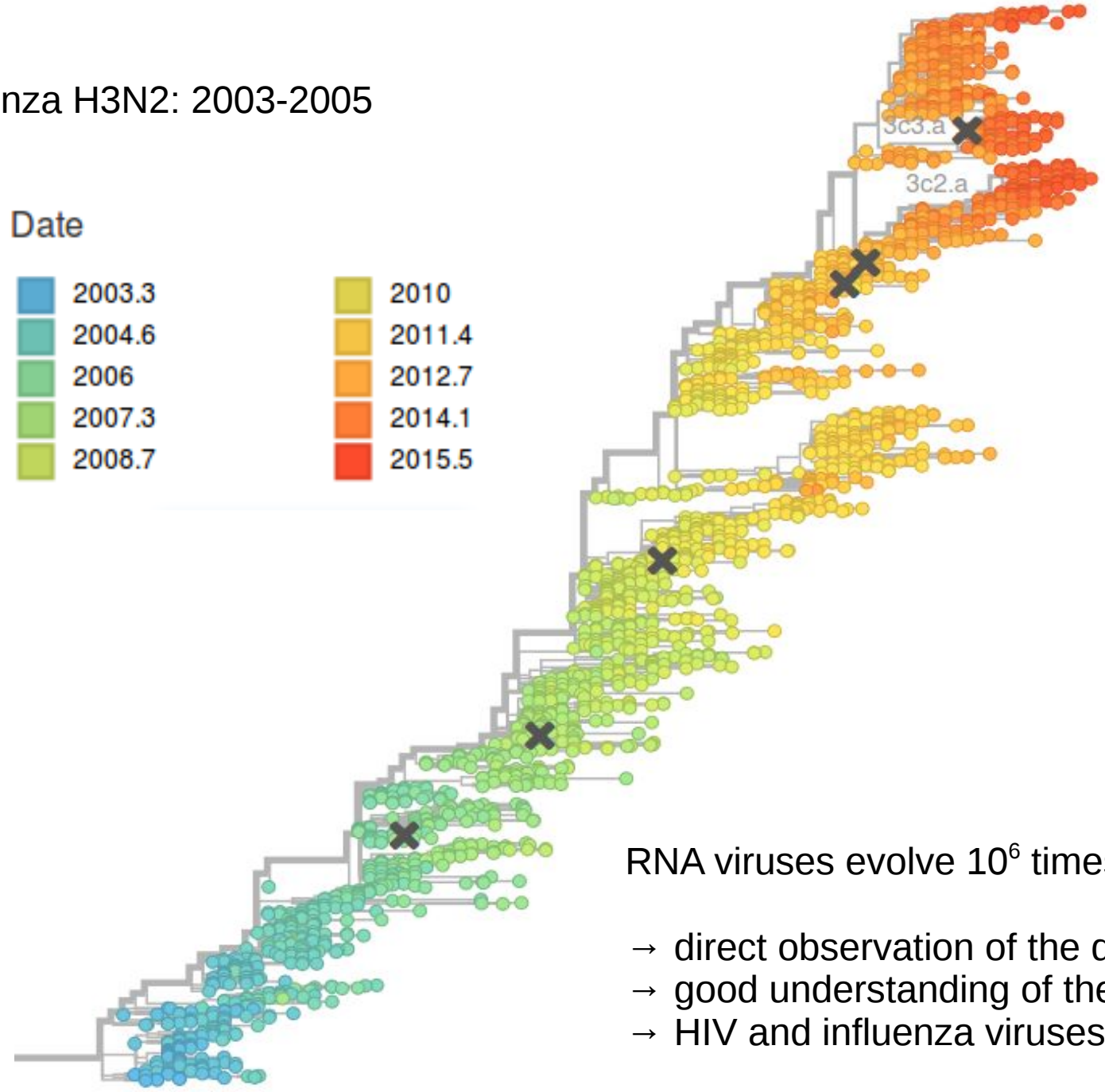


evolutionary processes ↔ statistics of trees ↔ patterns of genetic diversity

→ time resolved data allows direct observation

Rapidly evolving populations

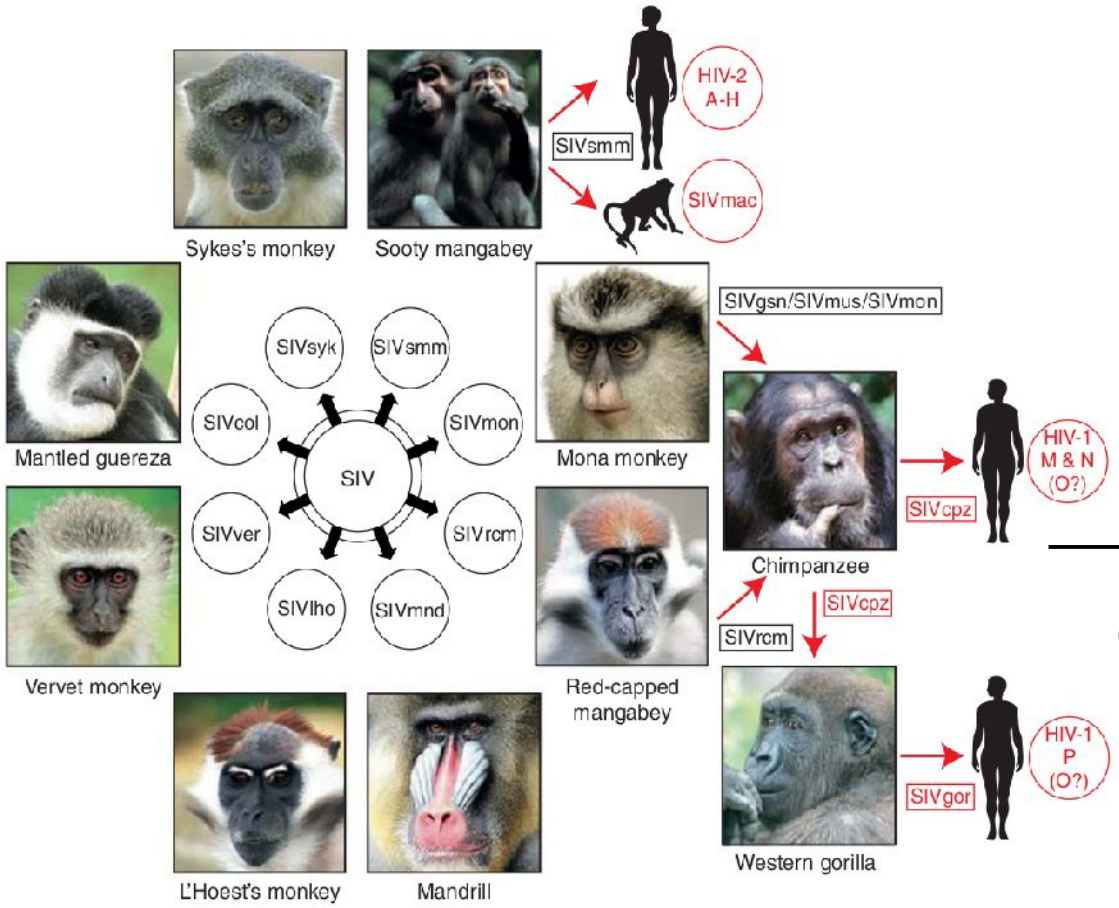
Influenza H3N2: 2003-2005

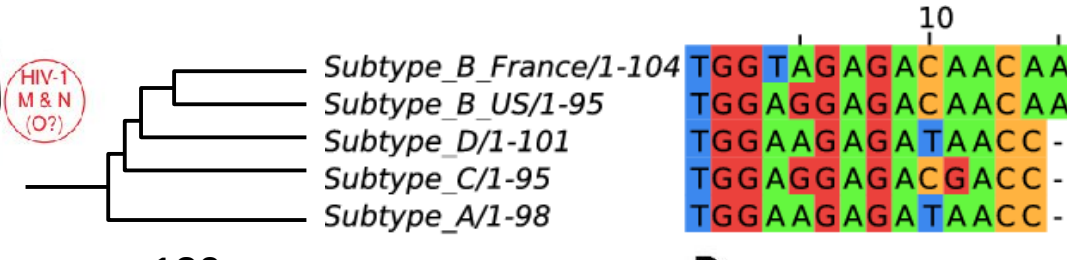


RNA viruses evolve 10^6 times than eukaryotes

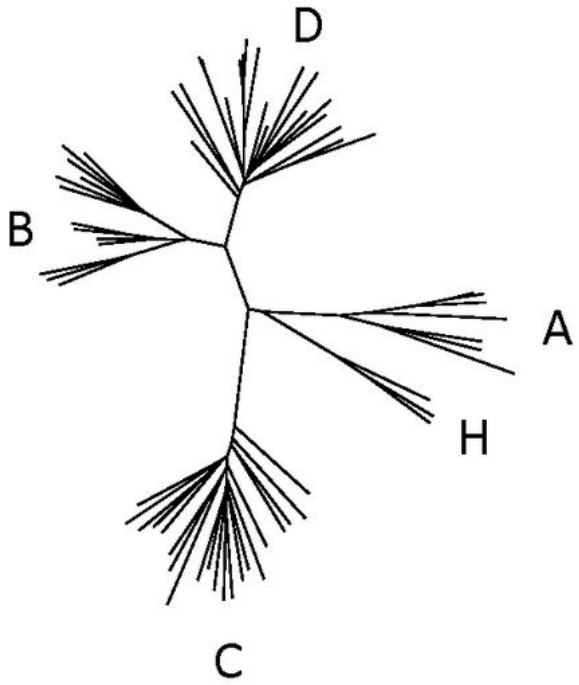
→ direct observation of the dynamics
→ good understanding of the phenotype
→ HIV and influenza viruses

History of HIV





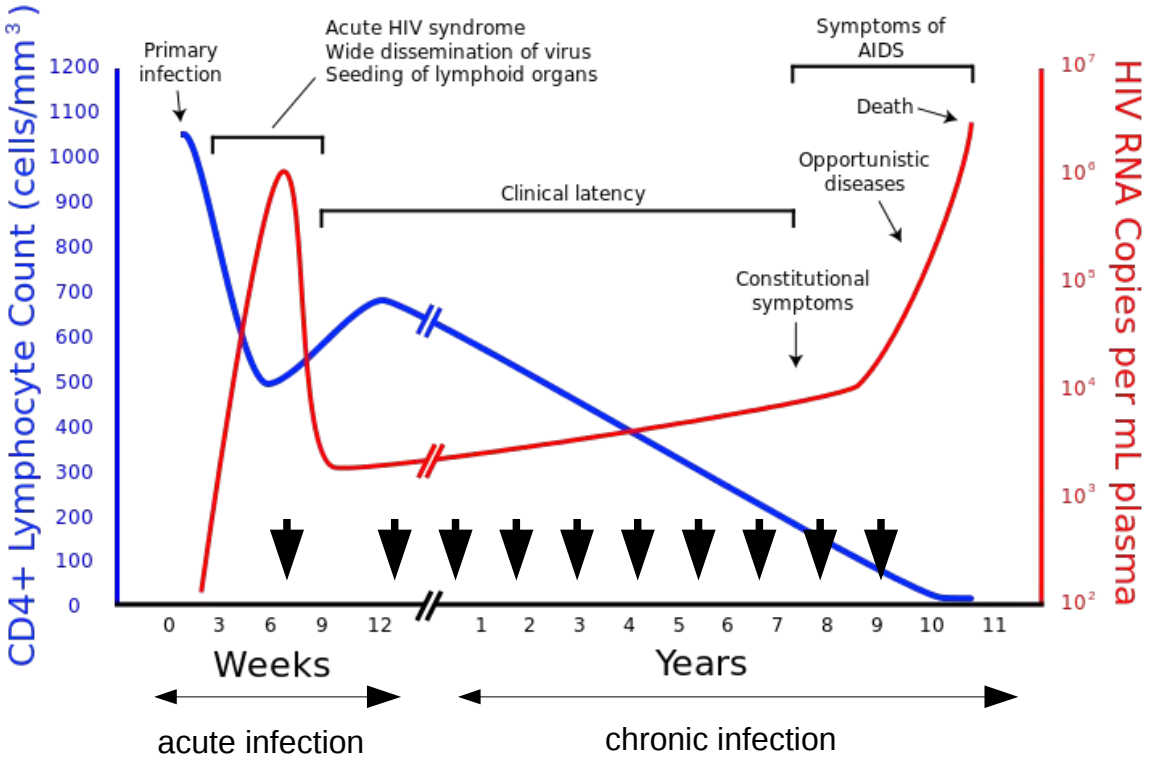
ca 120 years



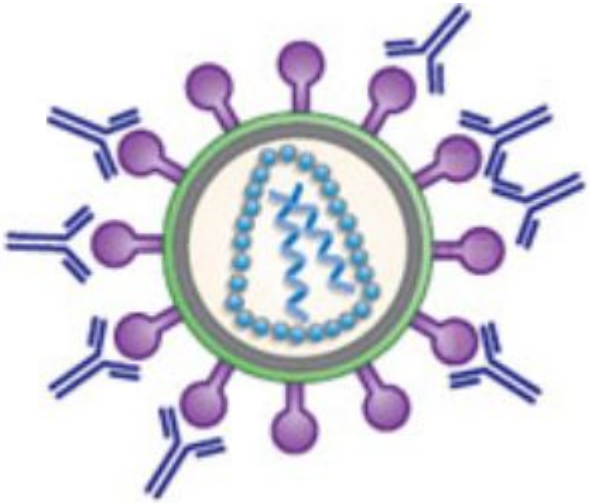
Sharp & Hahn, CSH Pers. Med, 2011

Klenerman, Fleming, Barnes, PLoS Pathogens, 2009

HIV-1 infection and immune selection



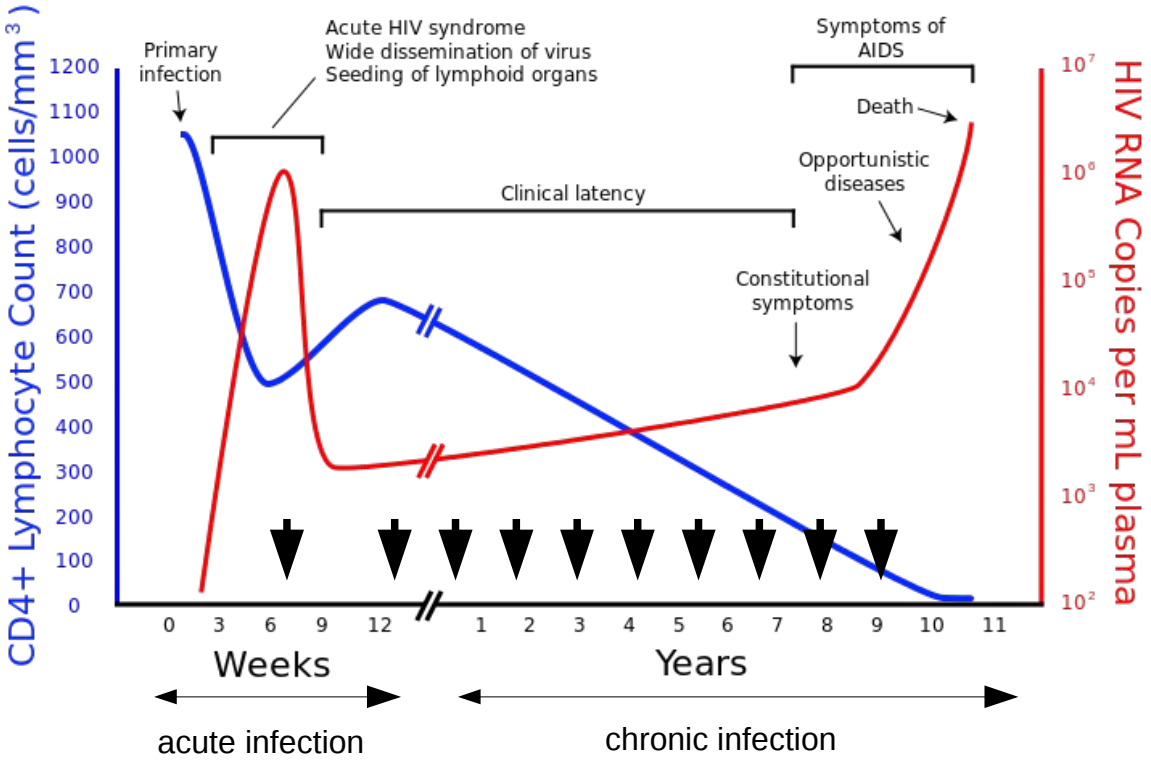
virus-immune system coevolution



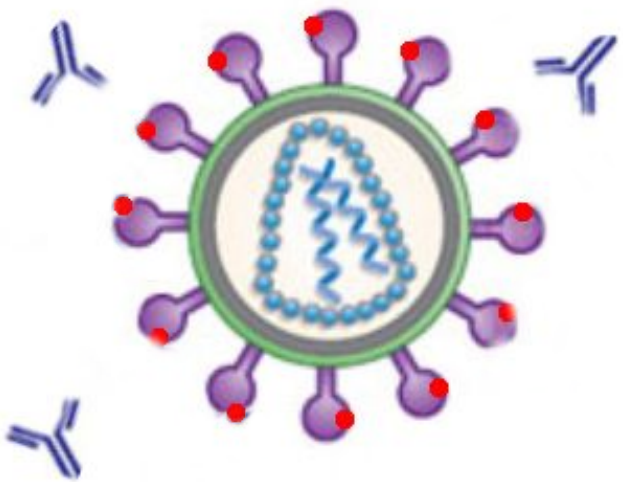
- rapid evolution facilitates chronic infection
- paradigmatic example of host-parasite co-evolution
- can be studied in exquisite detail

Image: wikipedia

HIV-1 infection and immune selection



virus-immune system coevolution

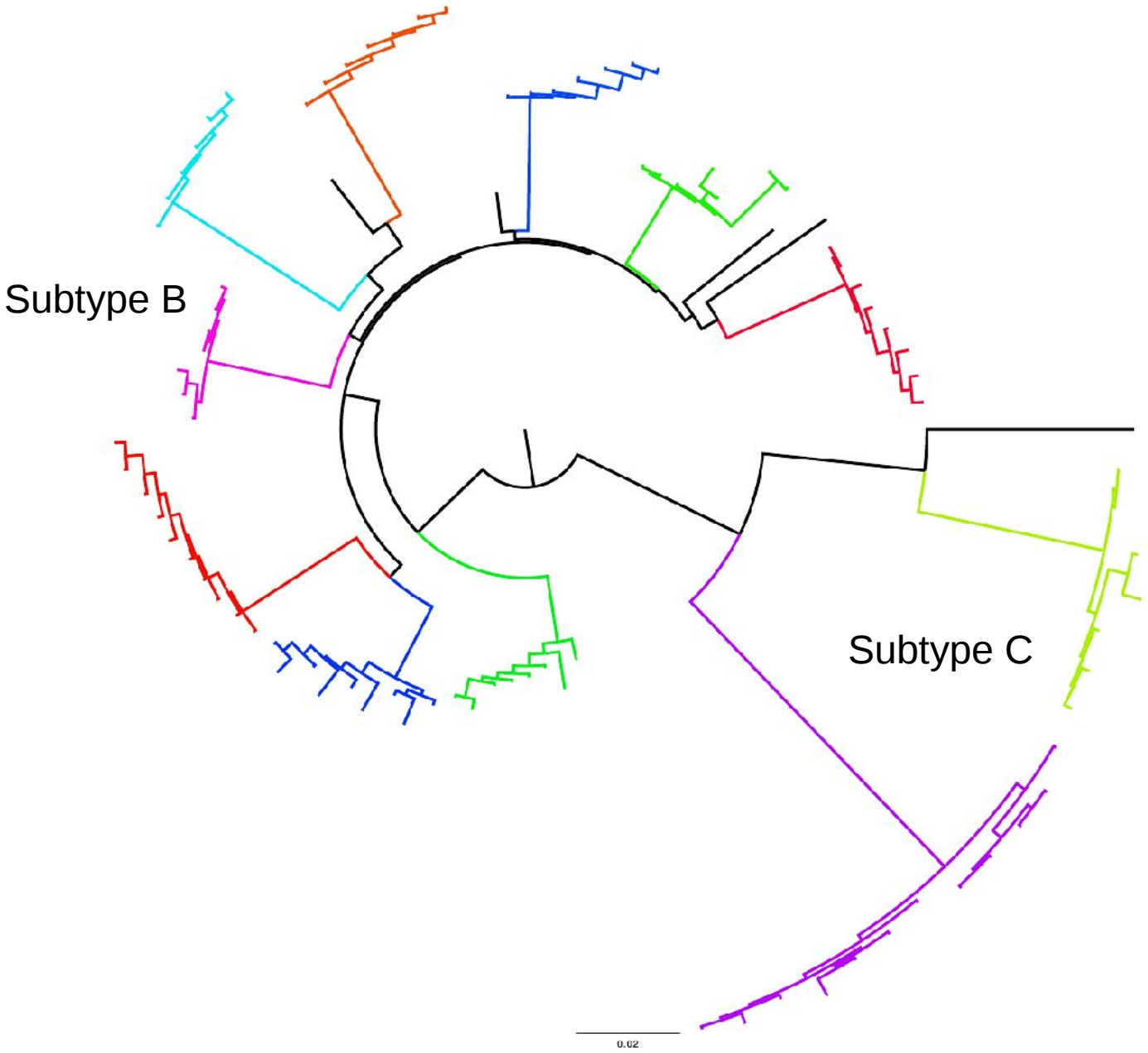


- rapid evolution facilitates chronic infection
- paradigmatic example of host-parasite co-evolution
- can be studied in exquisite detail

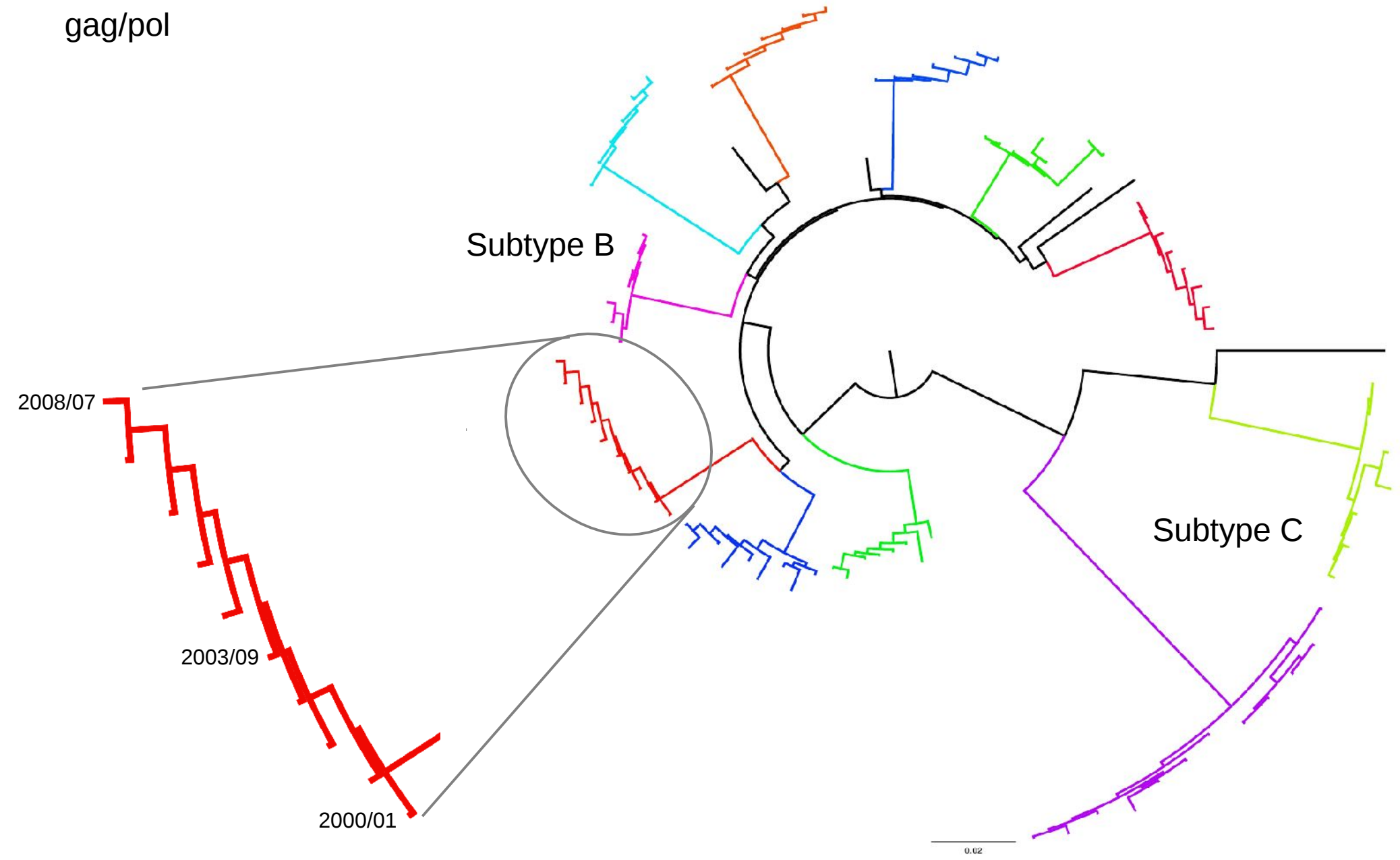
Image: wikipedia

Sequencing of serial HIV-1 RNA samples

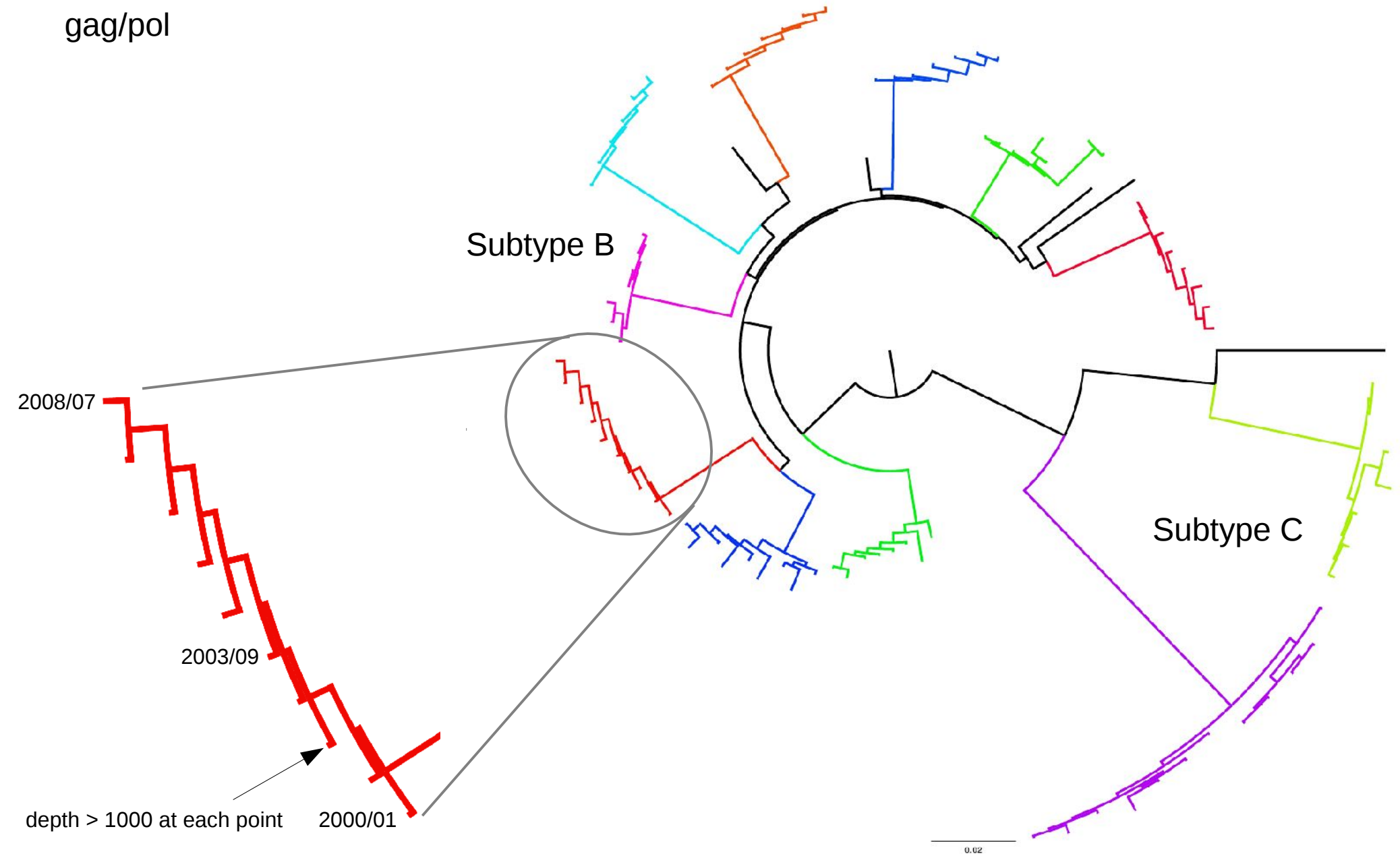
gag/pol



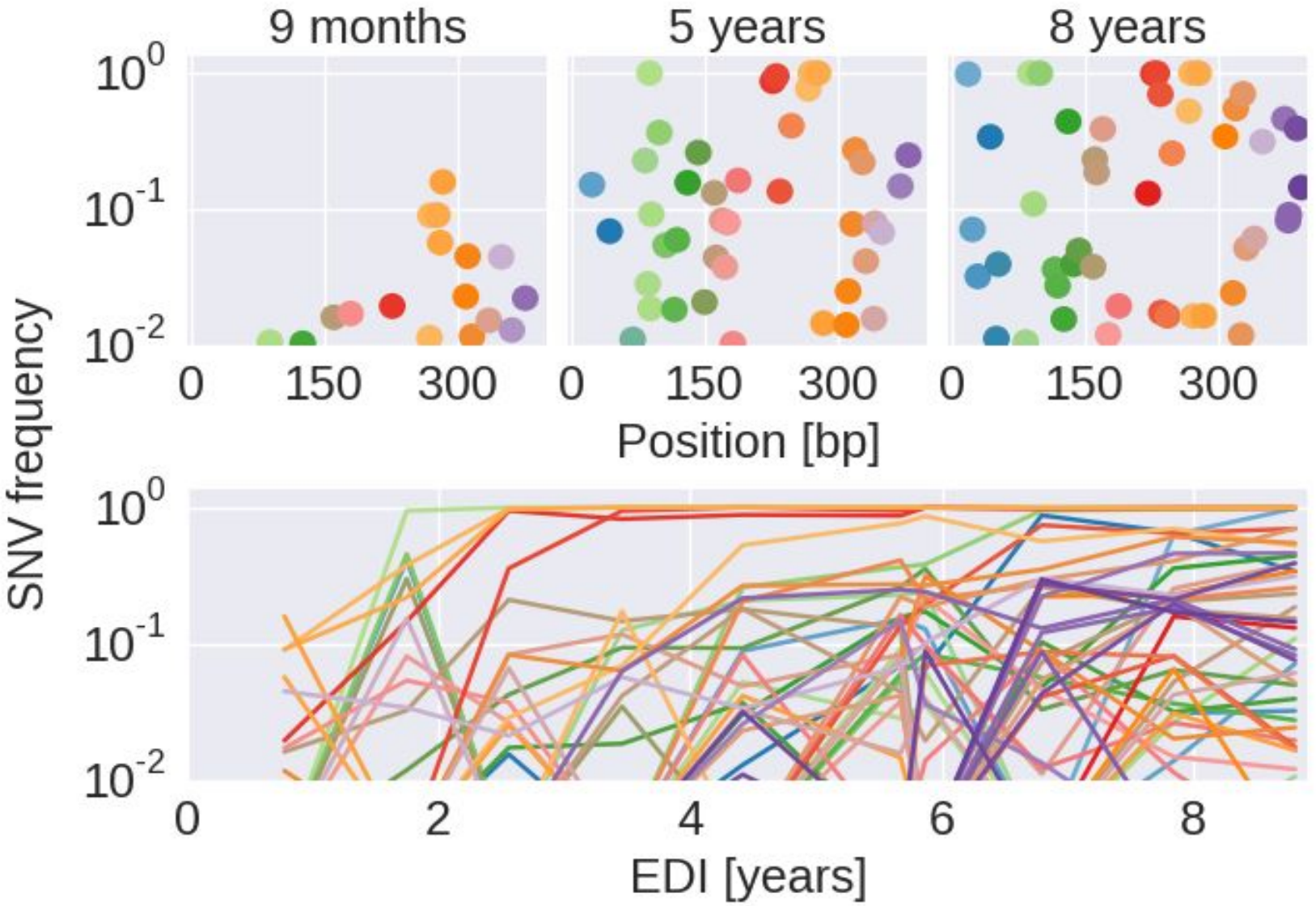
Sequencing of serial HIV-1 RNA samples



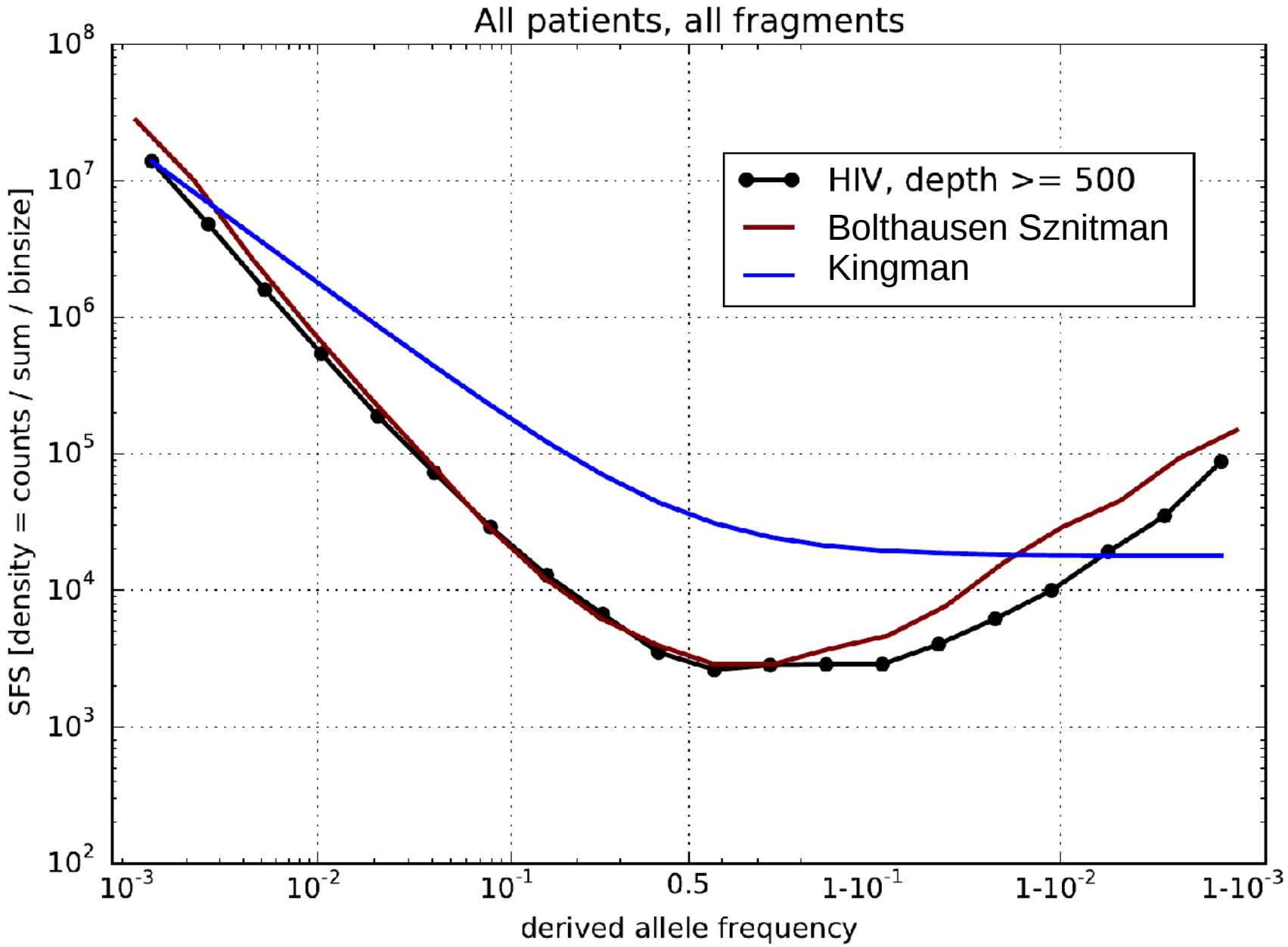
Sequencing of serial HIV-1 RNA samples



The dynamics of mutations



What is an appropriate model?

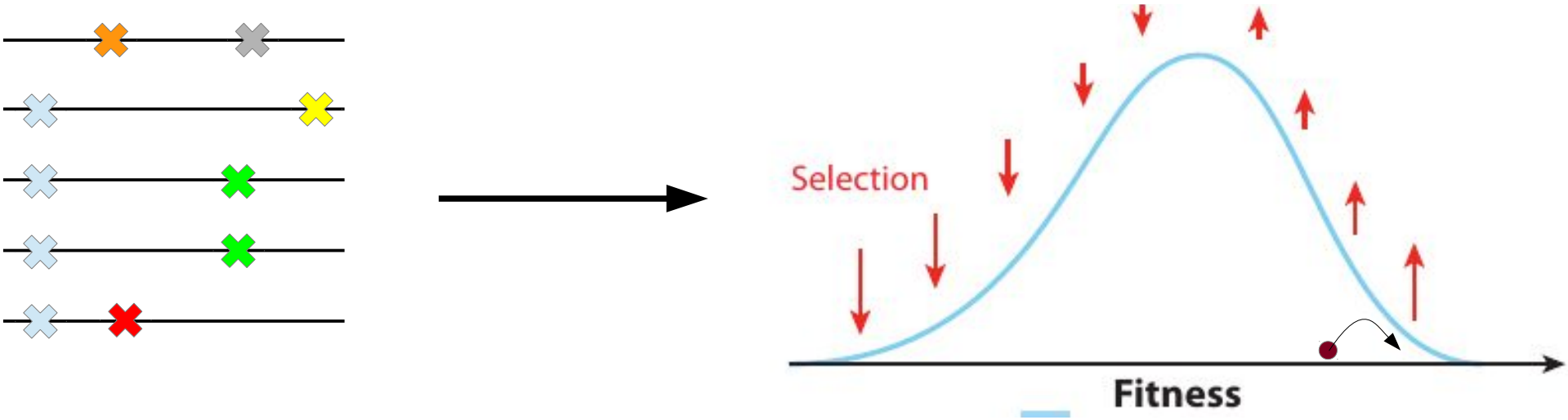


Why does it look like Bolthausen-Sznitman?

A large fraction of genetic variation is under selection

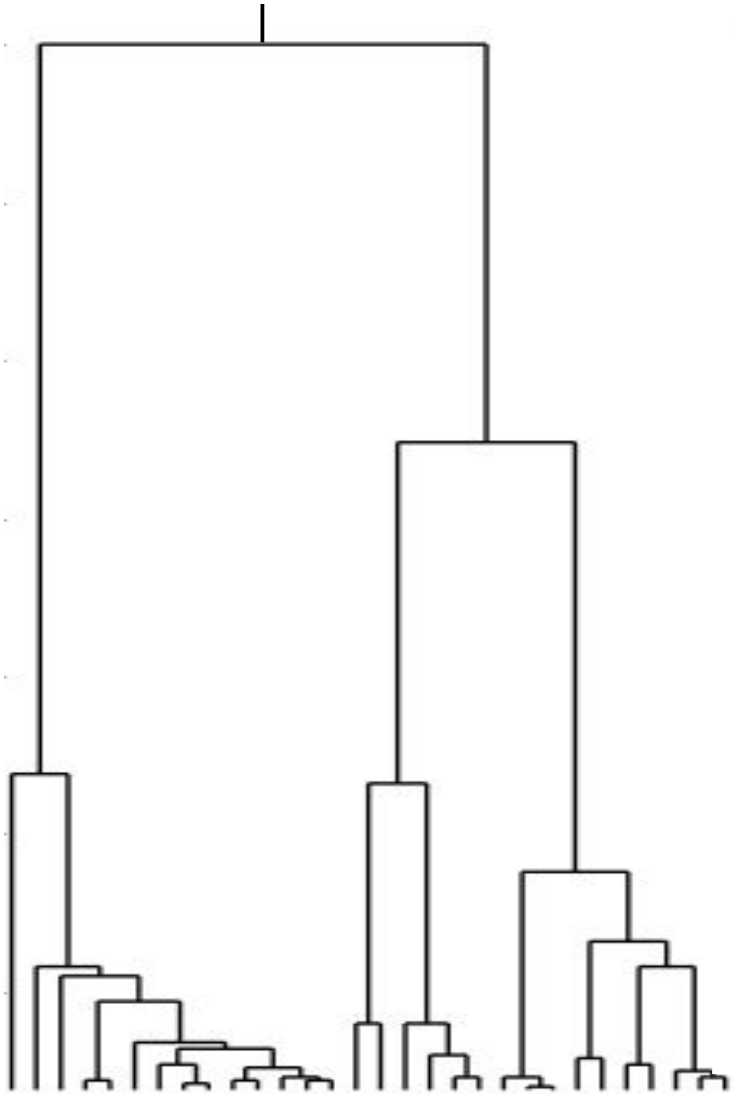
→ there is a broad fitness distribution

→ successful parents have successful offspring

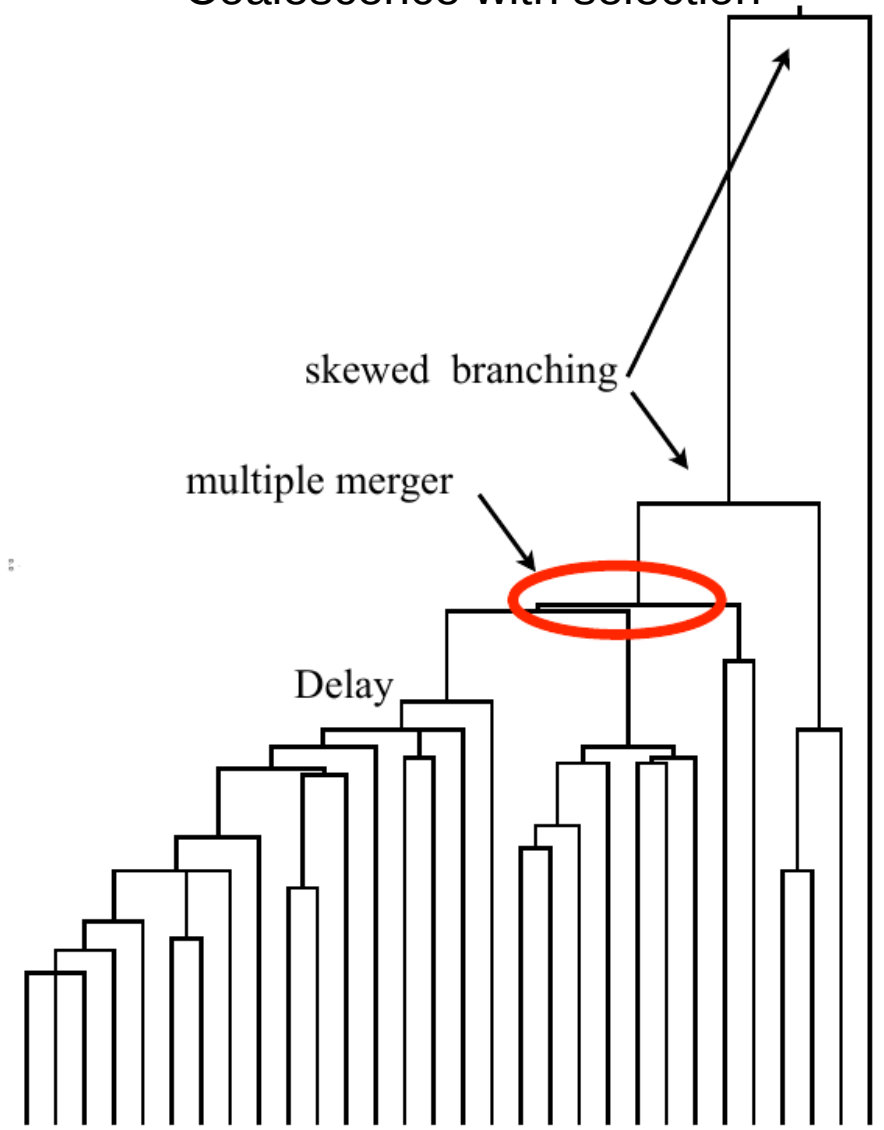


Coalescence with selection

Neutral coalescence

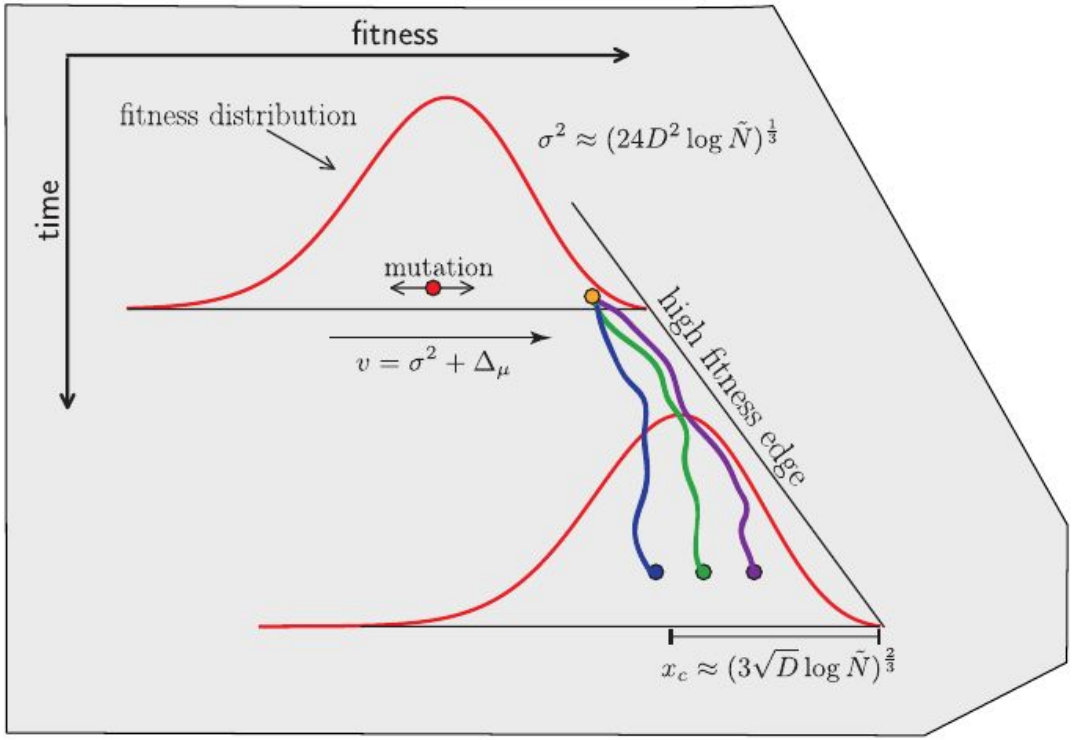


Coalescence with selection

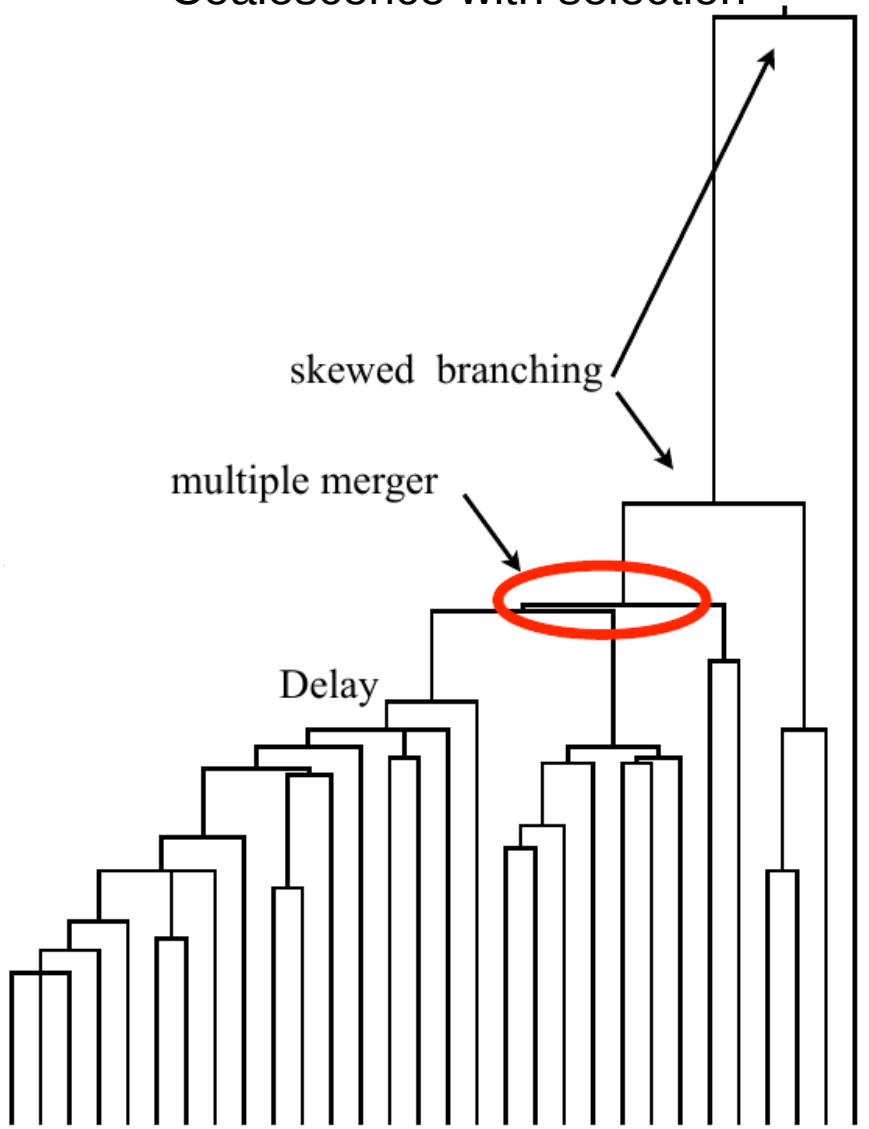


RN and Hallatschek, PNAS, 2013
Desai, Walczak, Fisher, Genetics, 2013
Brunet et al, PRE, 2007

Coalescence with selection

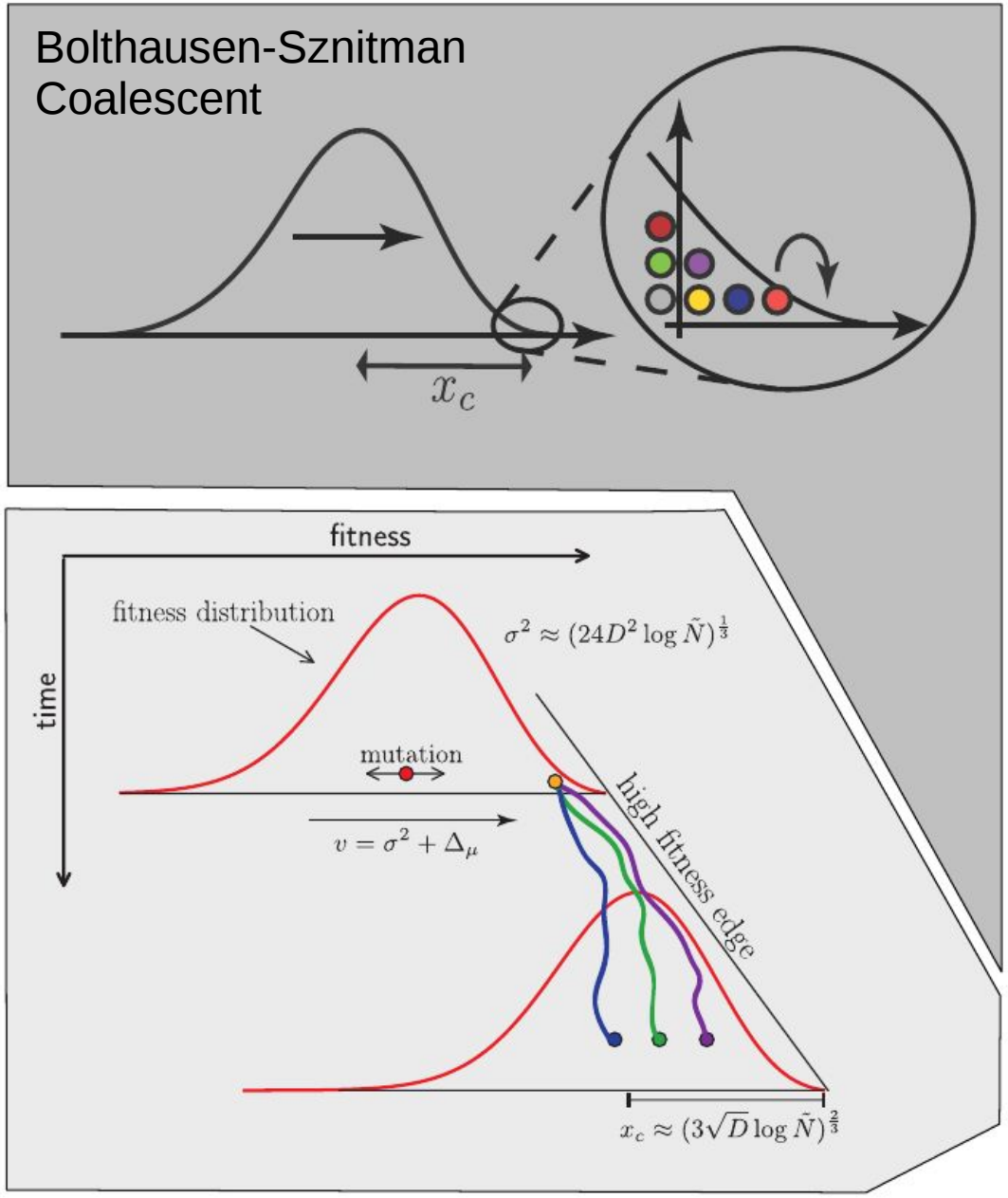


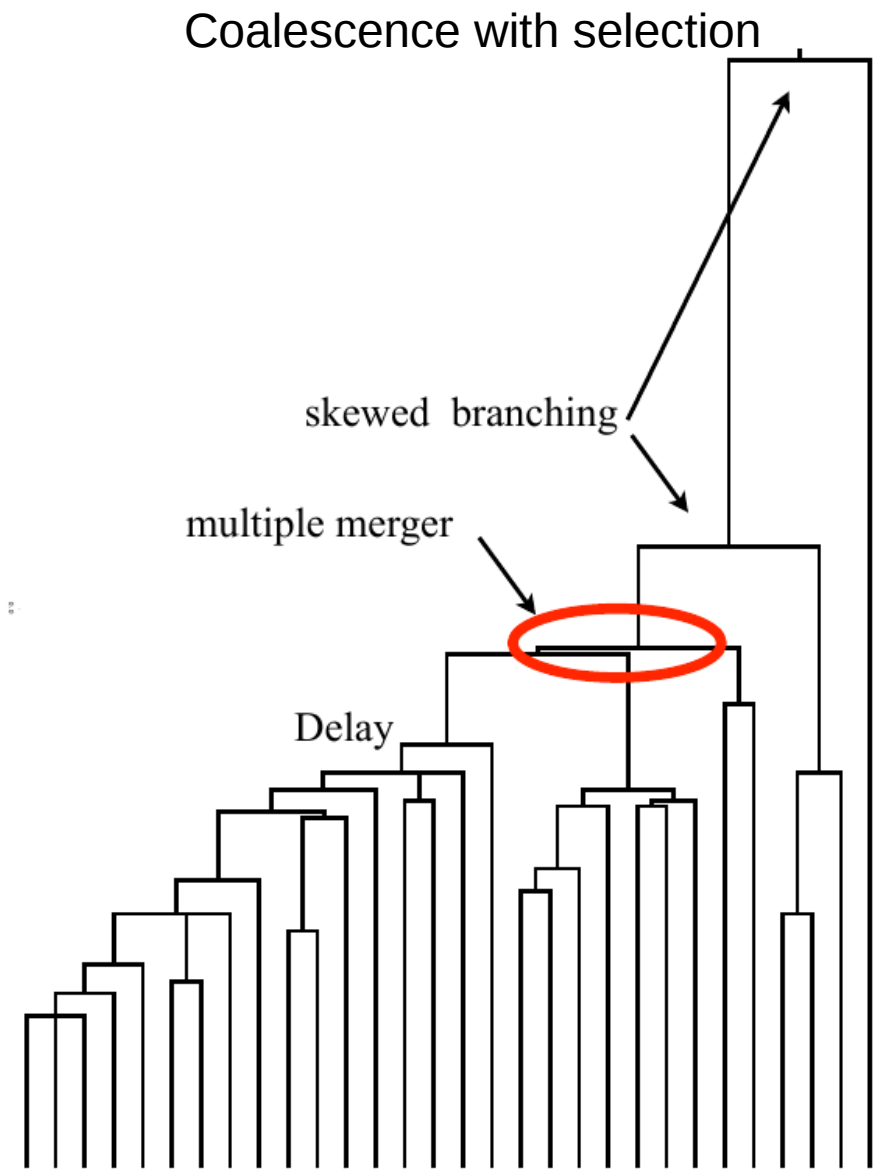
Coalescence with selection



RN and Hallatschek, PNAS, 2013
Desai, Walczak, Fisher, Genetics, 2013
Brunet et al, PRE, 2007

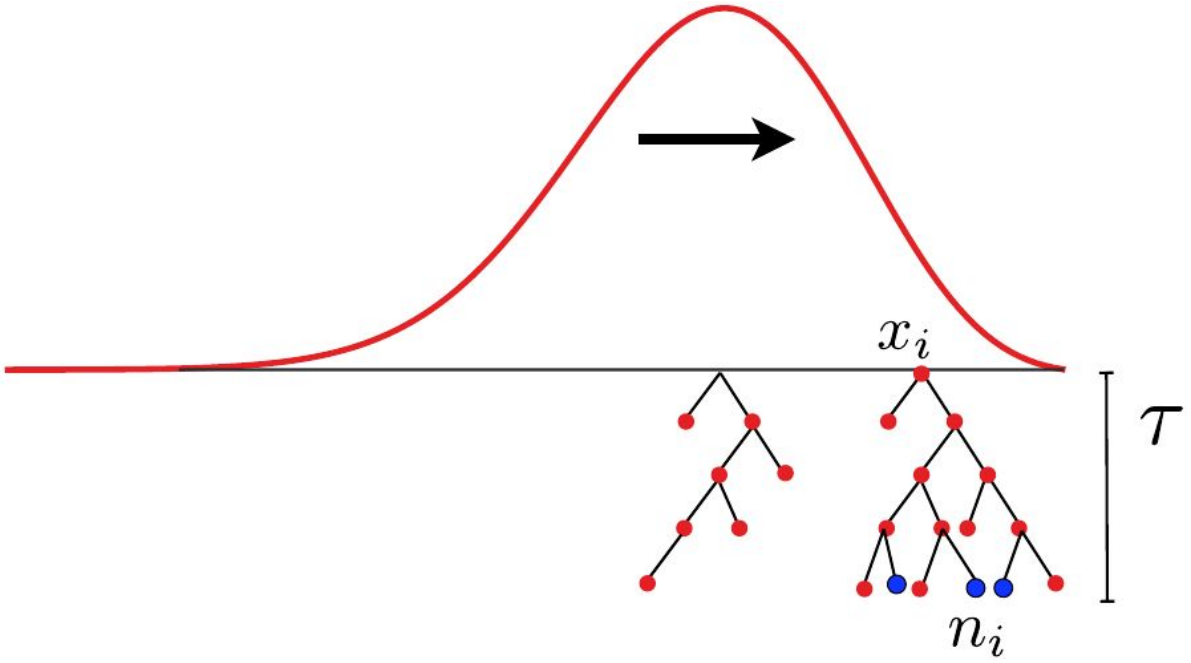
Coalescence with selection





RN and Hallatschek, PNAS, 2013
Desai, Walczak, Fisher, Genetics, 2013
Brunet et al, PRE, 2007

Coalescence with selection

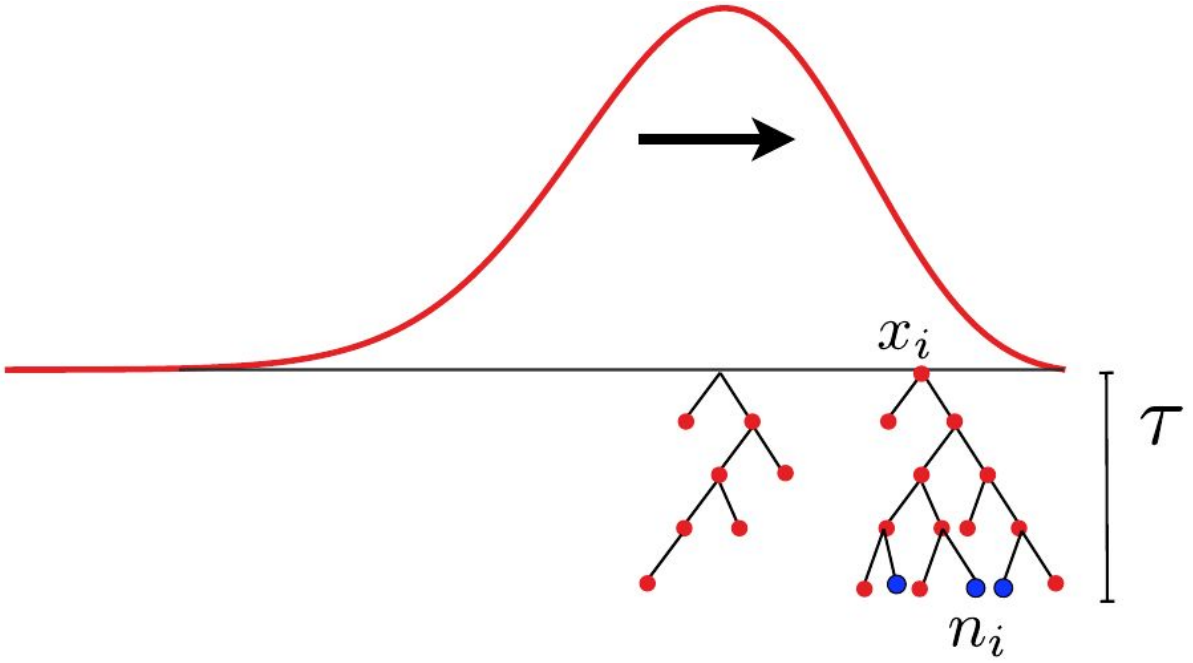


Branching process approximation:

$$P(n_i|\chi_i,\tau)$$

RN and Hallatschek, PNAS, 2013

Coalescence with selection



Branching process approximation:

$$P(n_i|\chi_i,\tau)$$

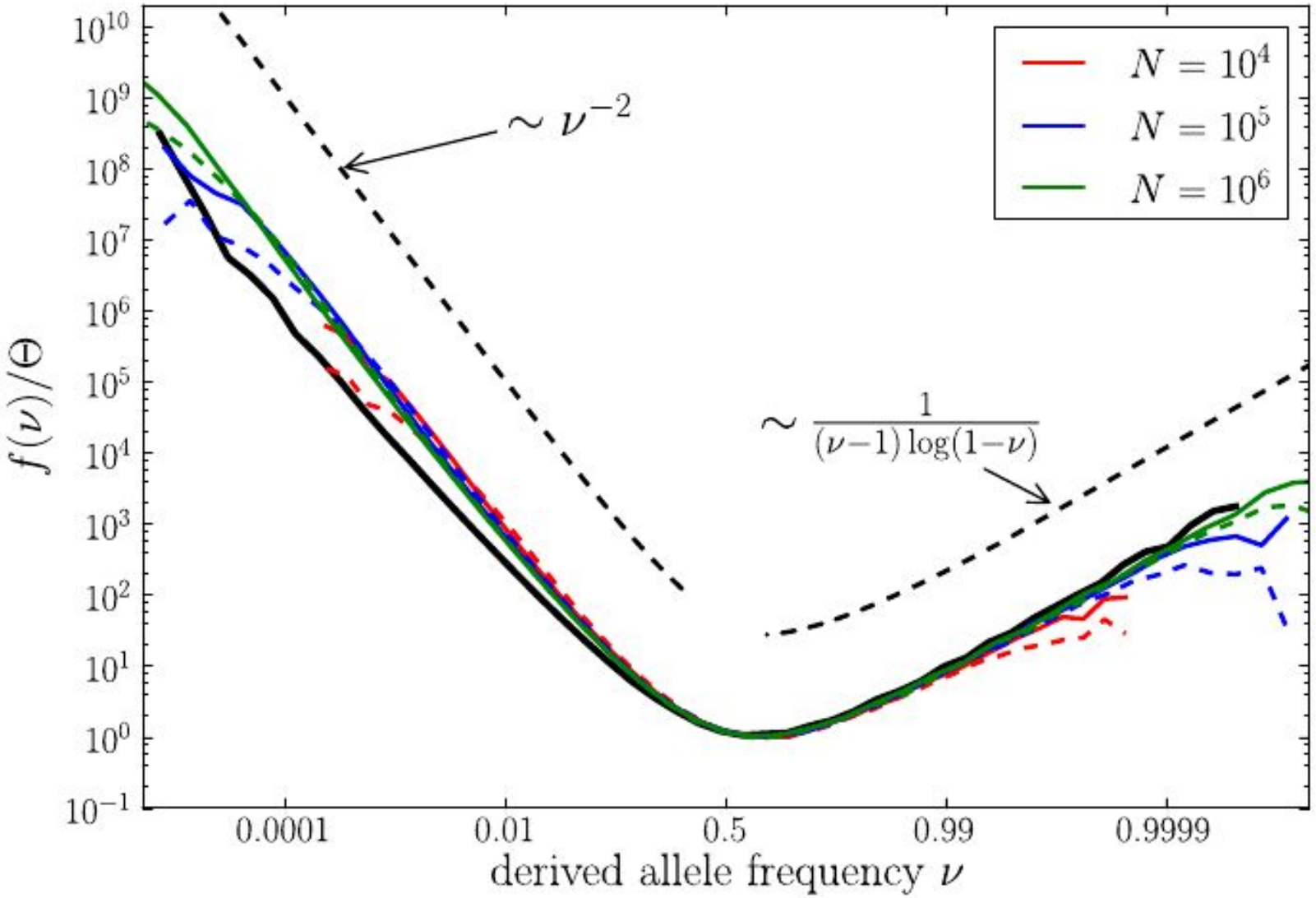
Does a sample (blue dots) have a common ancestor tau generations ago?

$$Q_b = \langle \sum_i \left(\frac{n_i}{\sum_j n_j} \right)^b \rangle = \begin{cases} \mathcal{O}(1/N) & \tau < T_c \\ \frac{\tau - T_c}{T_c(b-1)} & \tau > T_c \end{cases}$$

All other merger rates also suggest a Bolthausen-Sznitman coalescent

RN and Hallatschek, PNAS, 2013

Allele frequency spectra

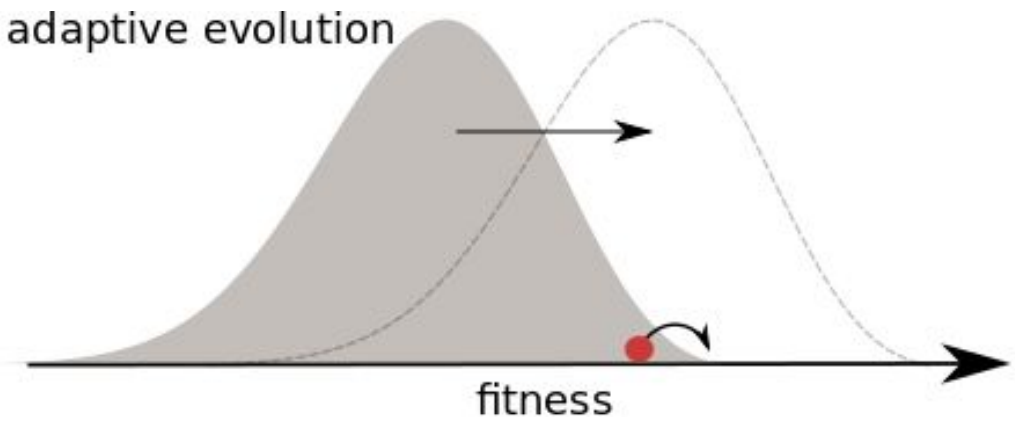


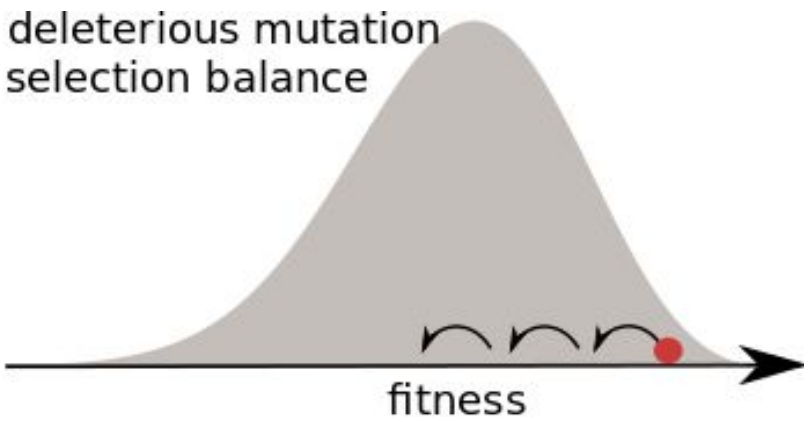
RN and Hallatschek, PNAS, 2013

Genetic diversity in adapting populations

Time scale of coalescence: $T_c \sim \frac{\sqrt{\log N}}{\sigma}$

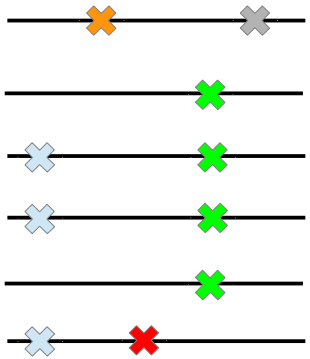
Universal: many selected mutations → same tree statistics

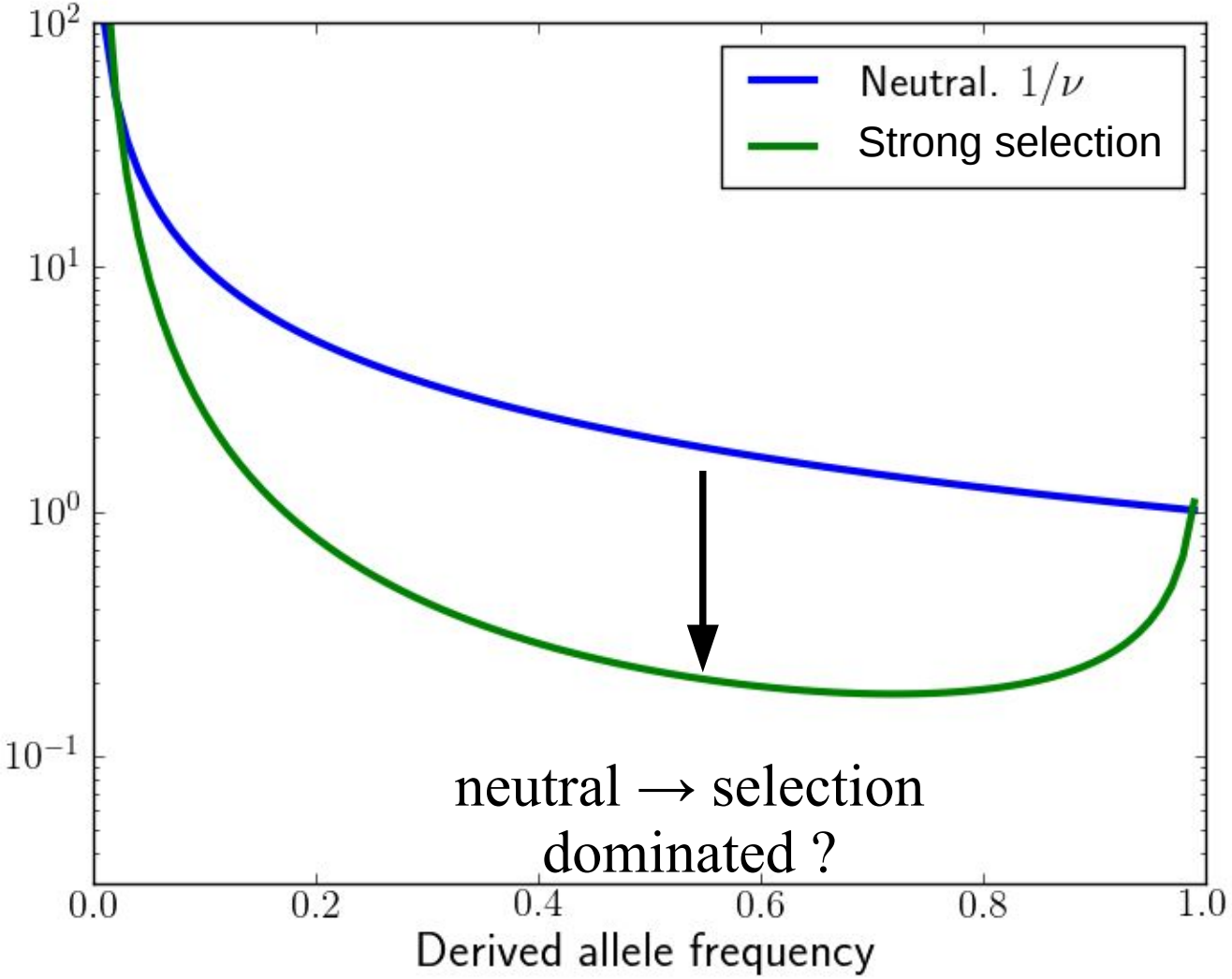




RN and Hallatschek, PNAS, 2013
Desai, Walczak, Fisher, Genetics 2013

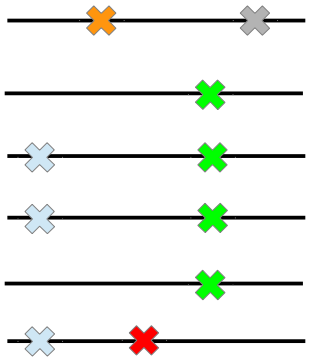
Continuous cross-over from neutral to strongly selected

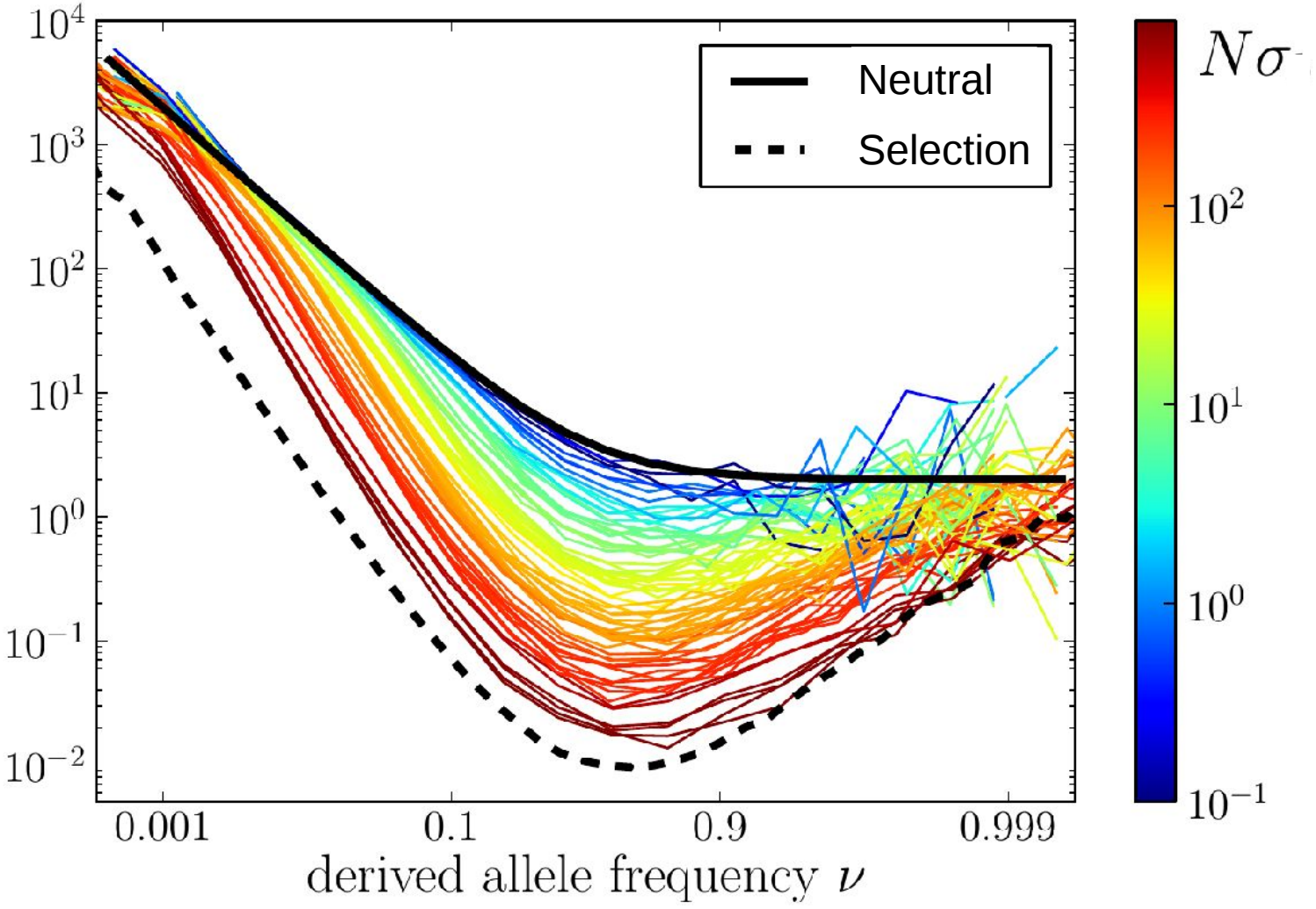




RN, Kessinger, Shraiman. PNAS, 2013
Good et al, PloS Genetics, 2014

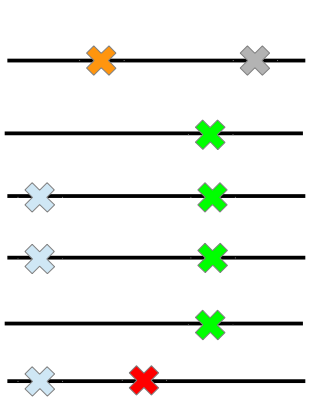
Continuous cross-over from neutral to strongly selected

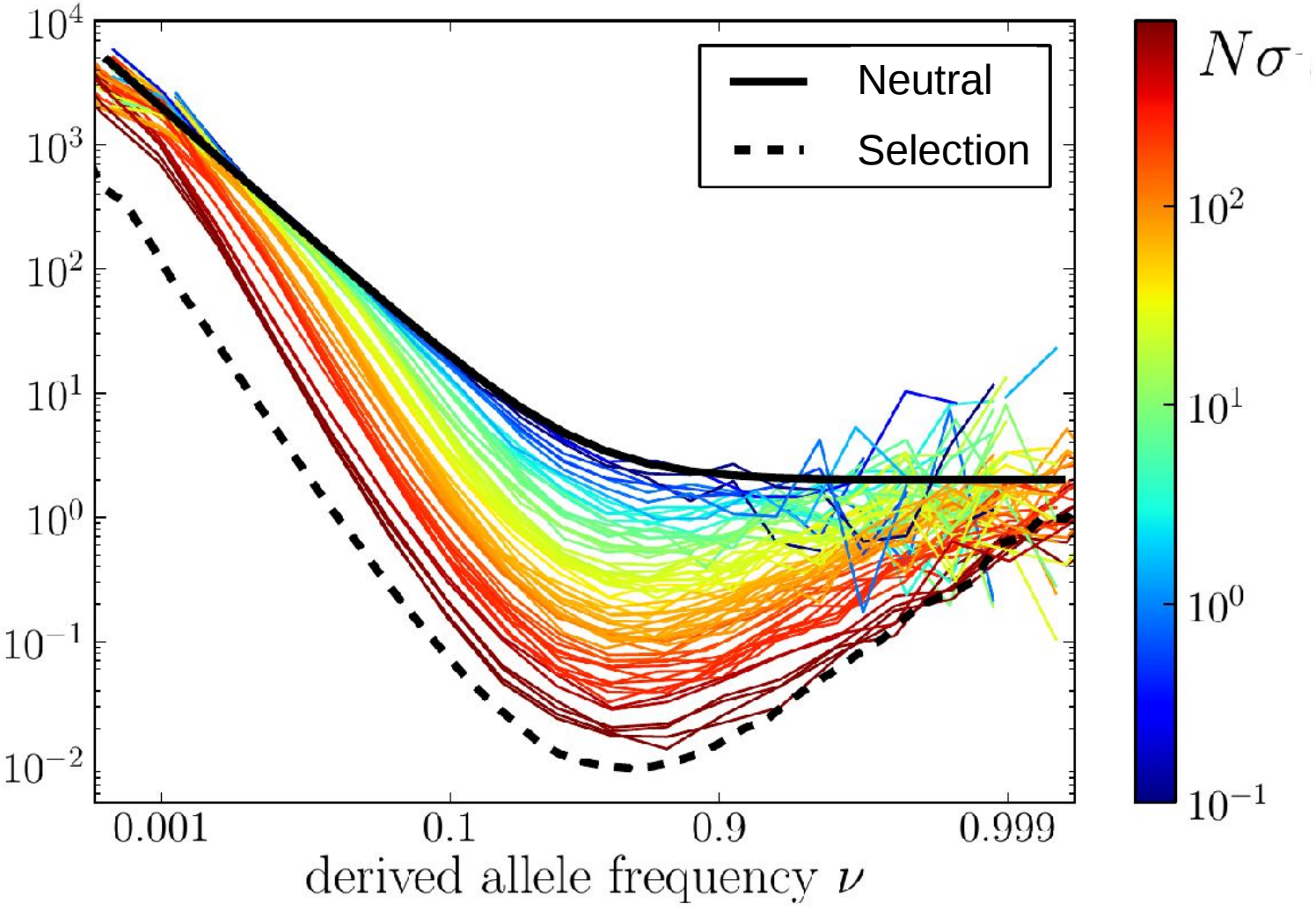


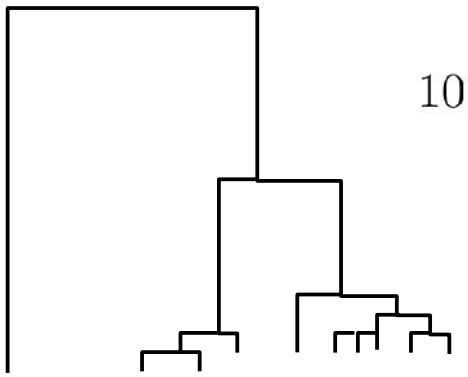


RN, Kessinger, Shraiman. PNAS, 2013
Good et al, PloS Genetics, 2014

Continuous cross-over from neutral to strongly selected

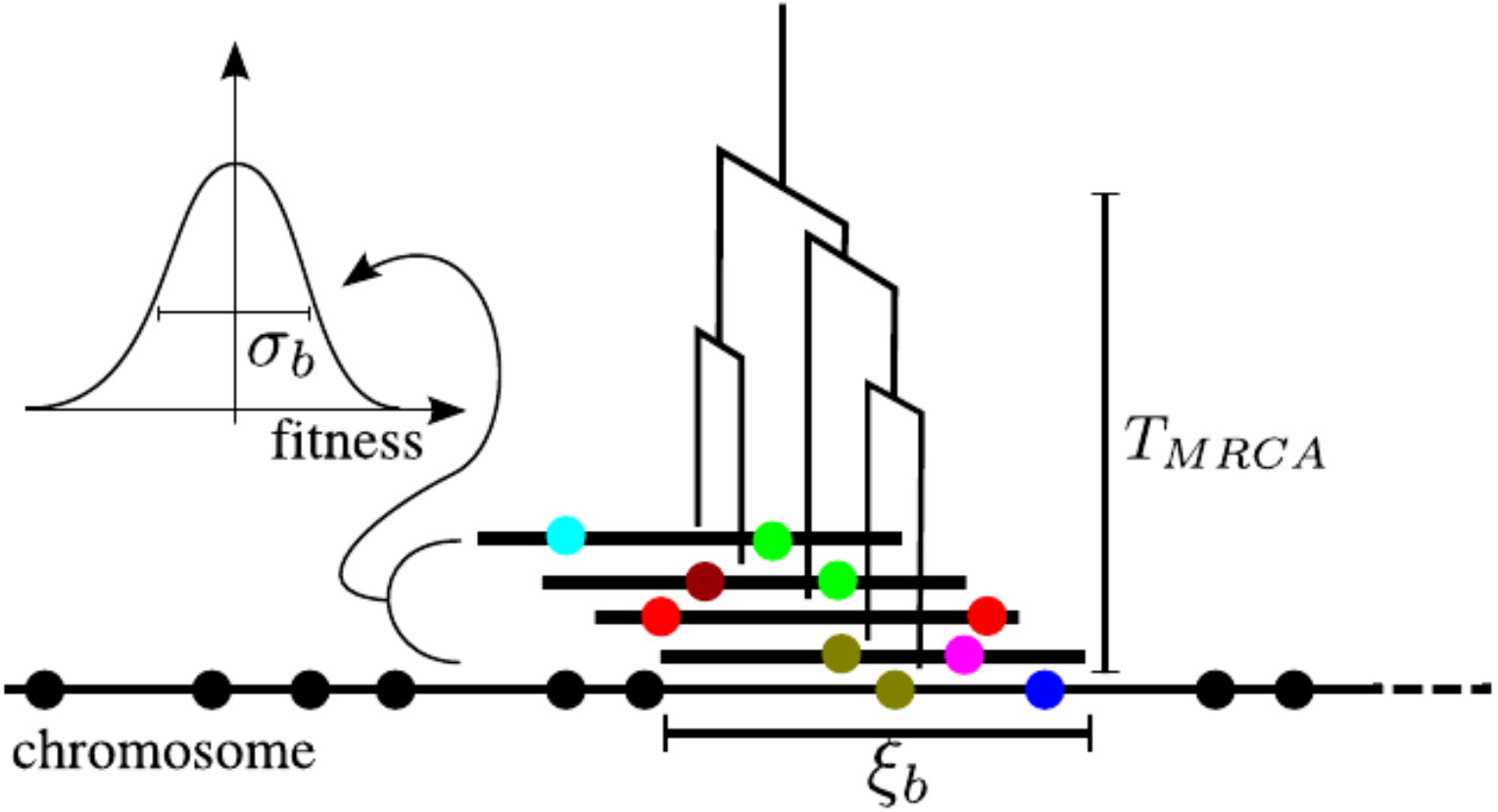






RN, Kessinger, Shraiman. PNAS, 2013
Good et al, PloS Genetics, 2014

Extension to recombining populations



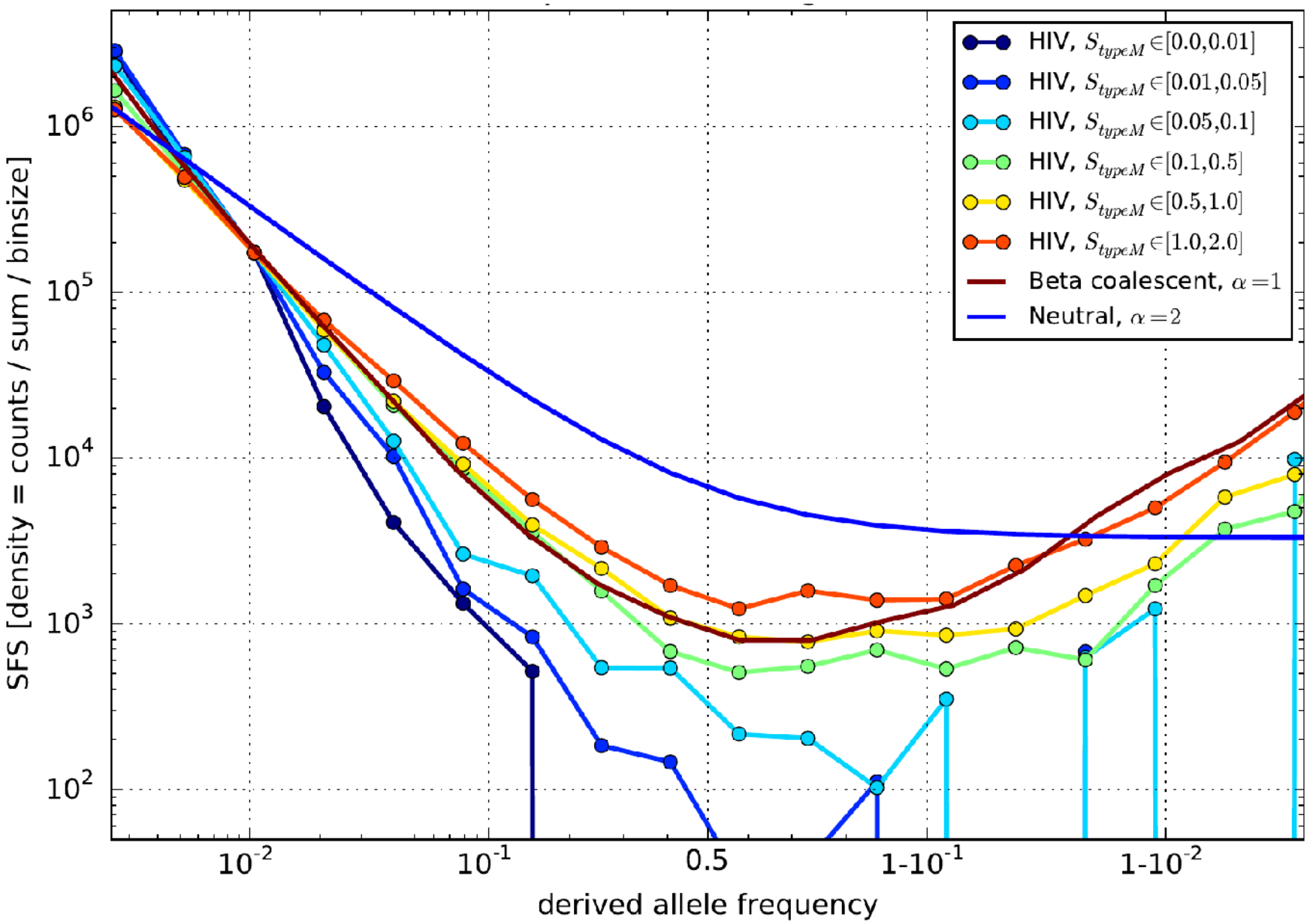
Coalescent time

Fitness diversity σ_b

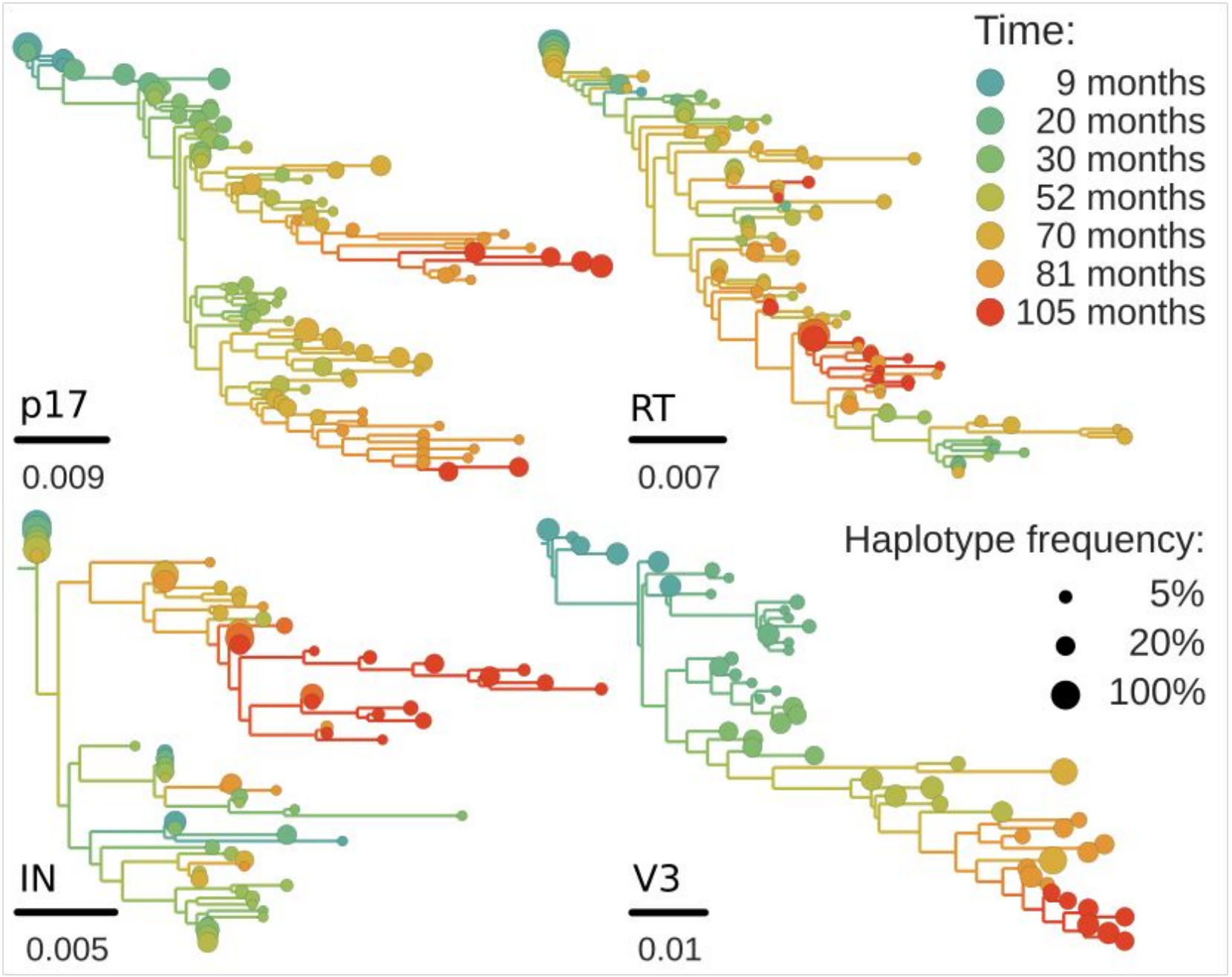
Length of linkage block ξ_b

RN, Kessinger, Shraiman, PNAS, 2013

Site frequency spectra

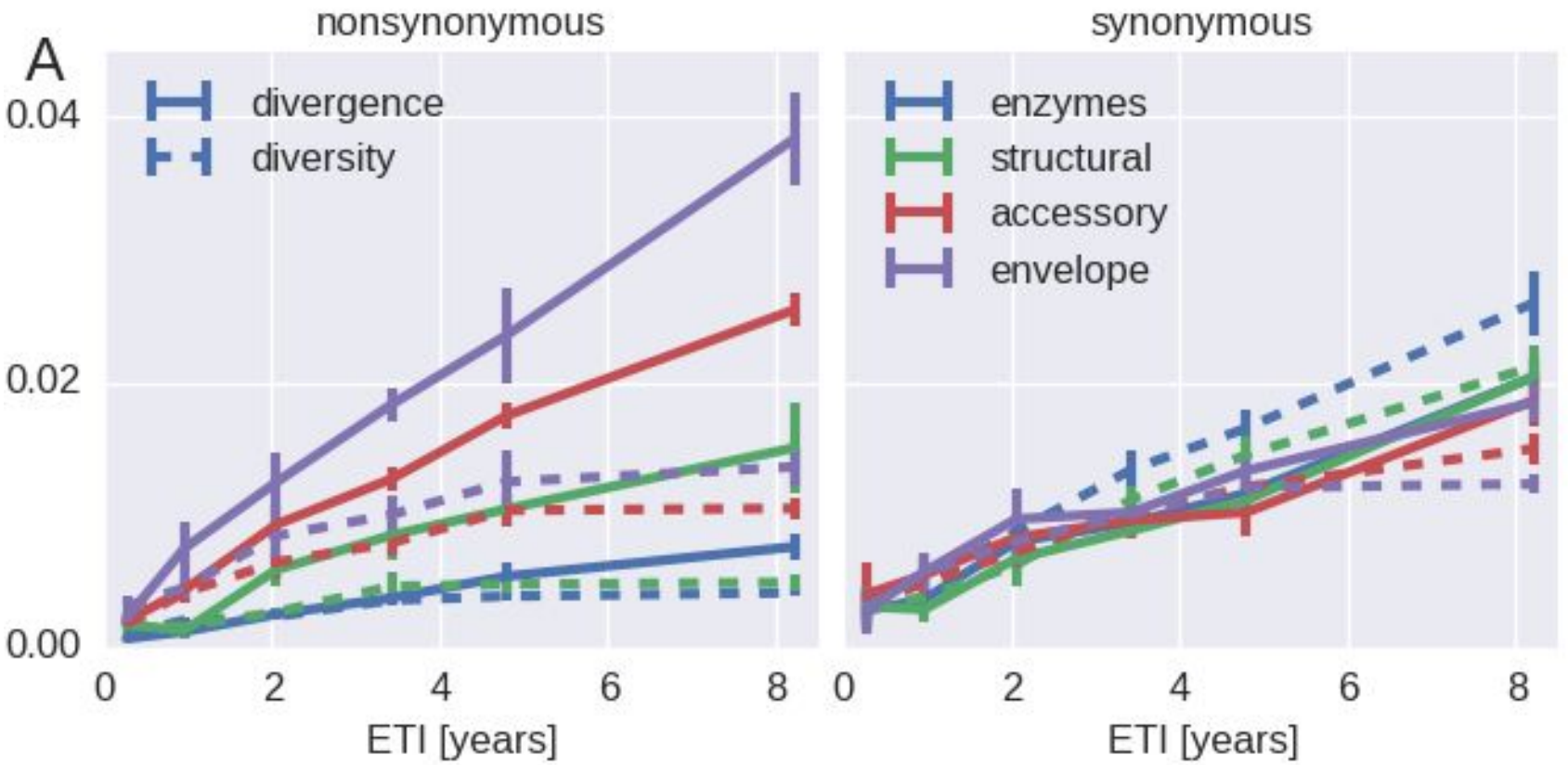


Phylogenetic trees in different regions



Movie

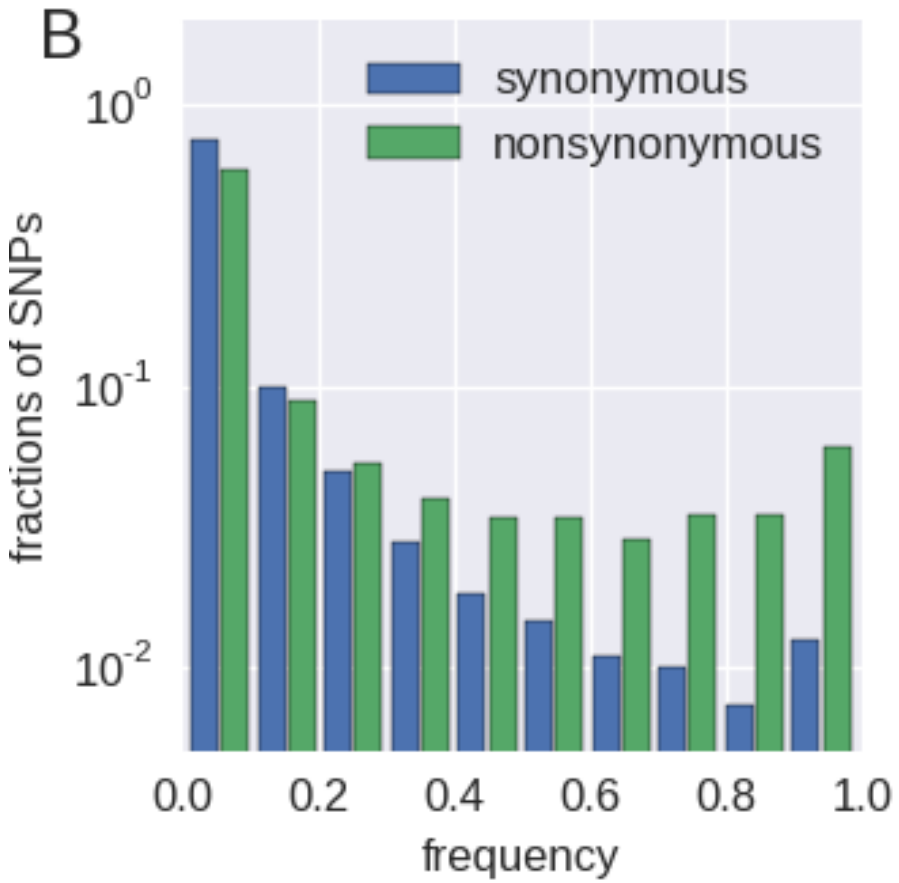
Recombination facilitates adaptation



Constrasting behavior:

- Non-synonymous diversity is low, region specific rate of evolution
- Synonymous diversity keeps increasing, rate is identical across regions
- Synonymous diversity is inversely related to non-synonymous divergence

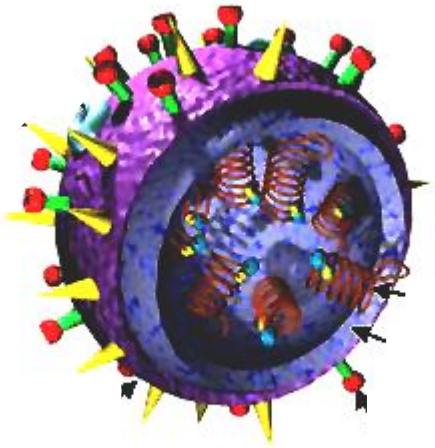
Recombination facilitates adaptation



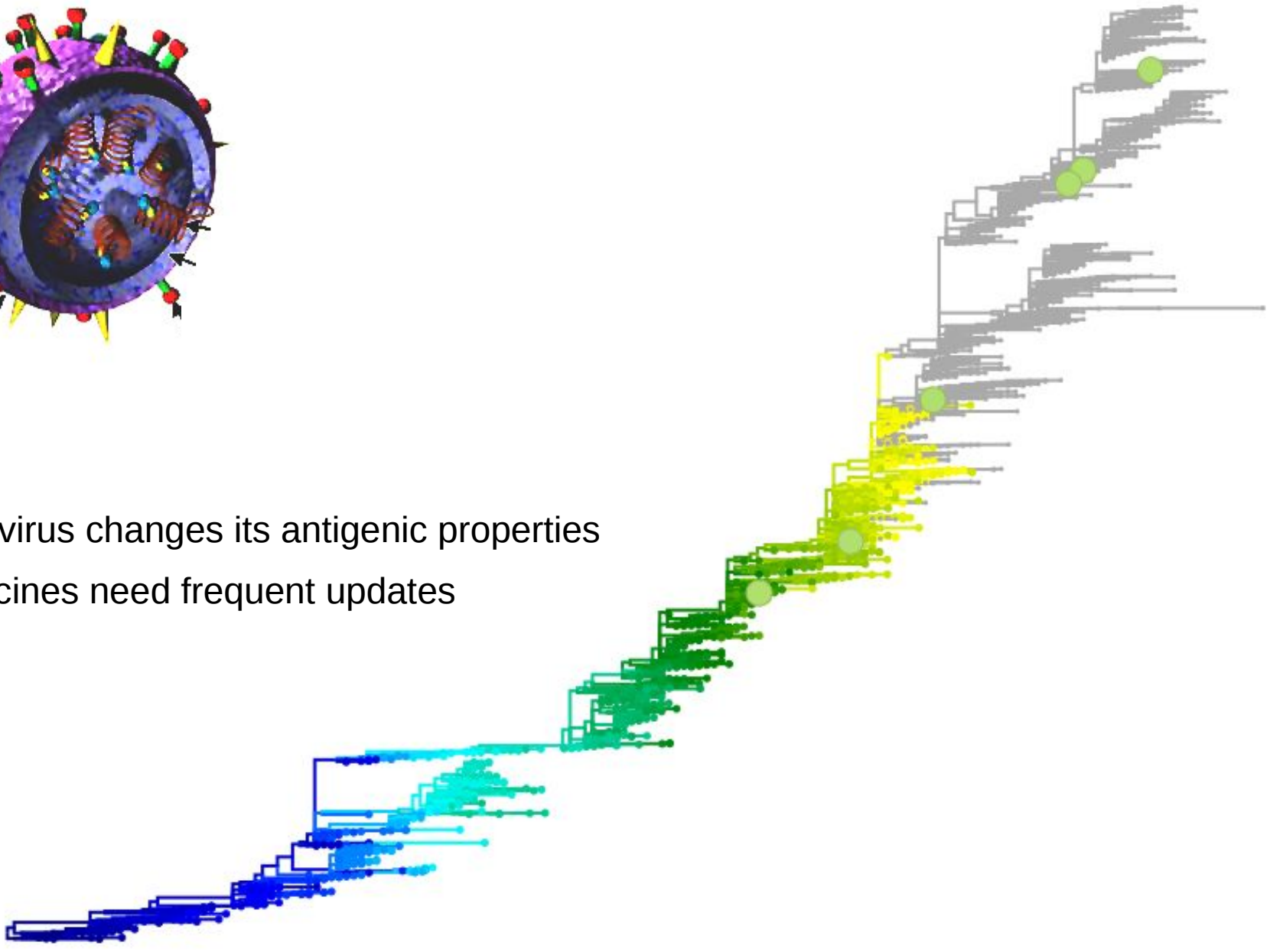
Constrasting behavior:

- Non-synonymous diversity is low, region specific rate of evolution
- Synonymous diversity keeps increasing, rate is identical across regions
- Synonymous diversity is inversely related to non-synonymous divergence

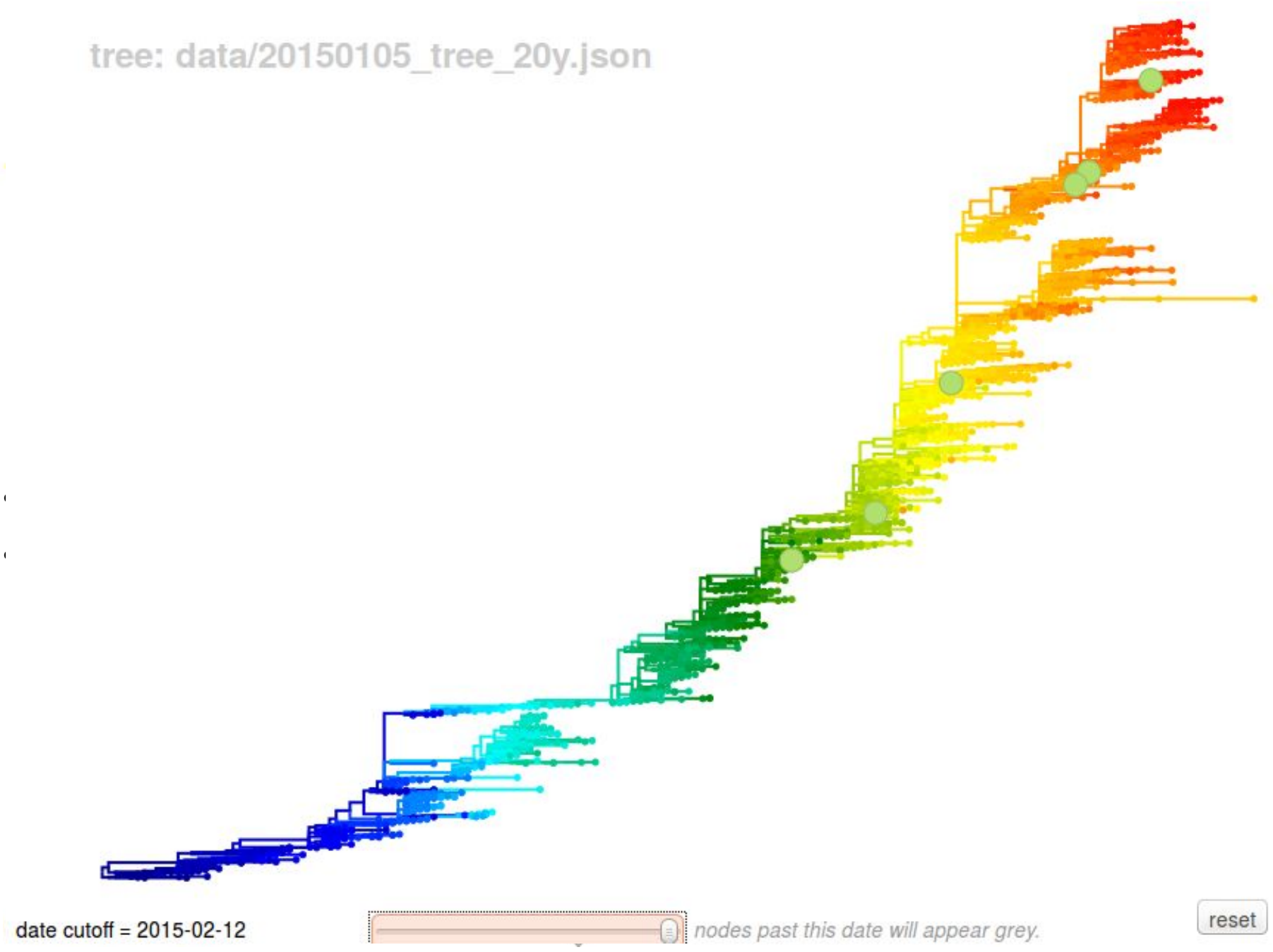
Predicting Influenza evolution?



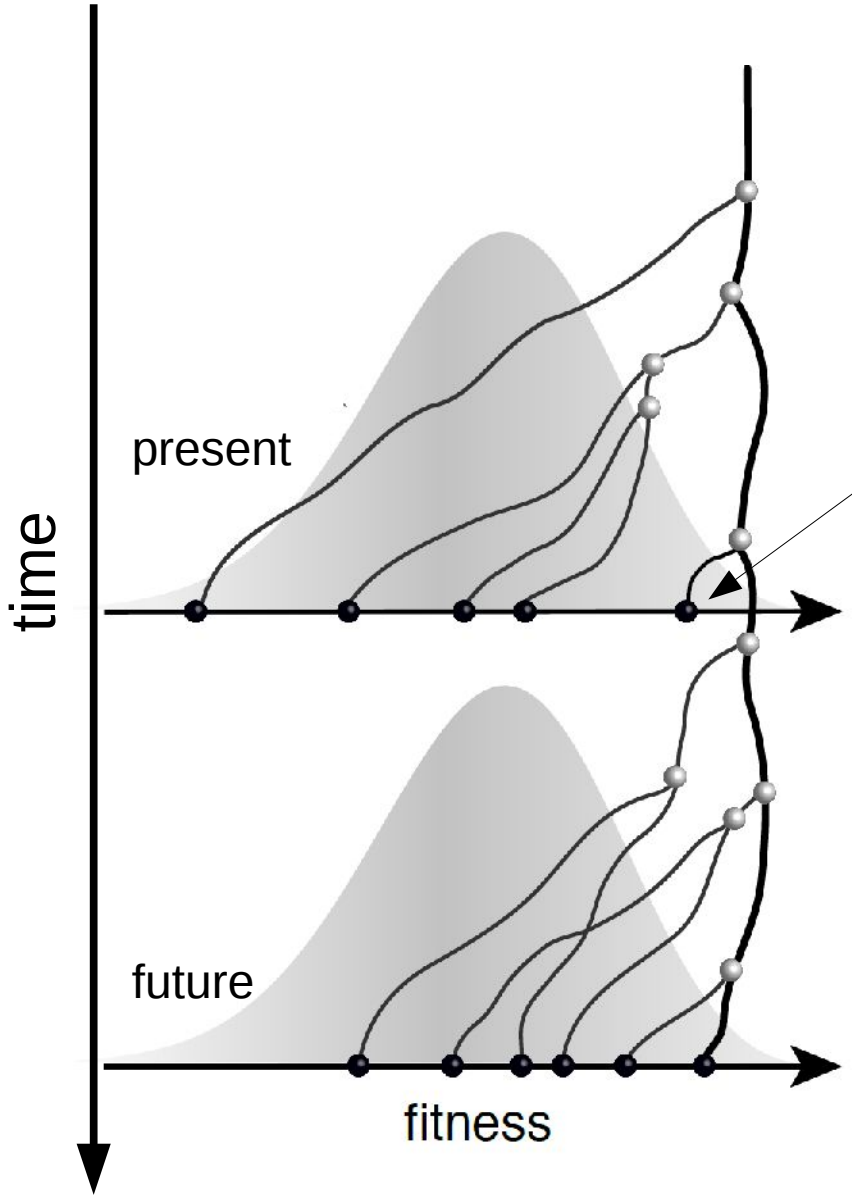
- the virus changes its antigenic properties
- vaccines need frequent updates



Predicting Influenza evolution?



Predicting Influenza evolution?

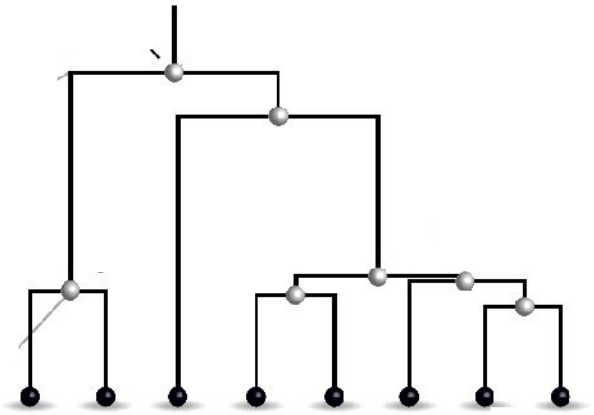


Best pick for new vaccine!

Given the branching pattern, can we

- predict fitness?
- pick the most likely progenitor of the future?

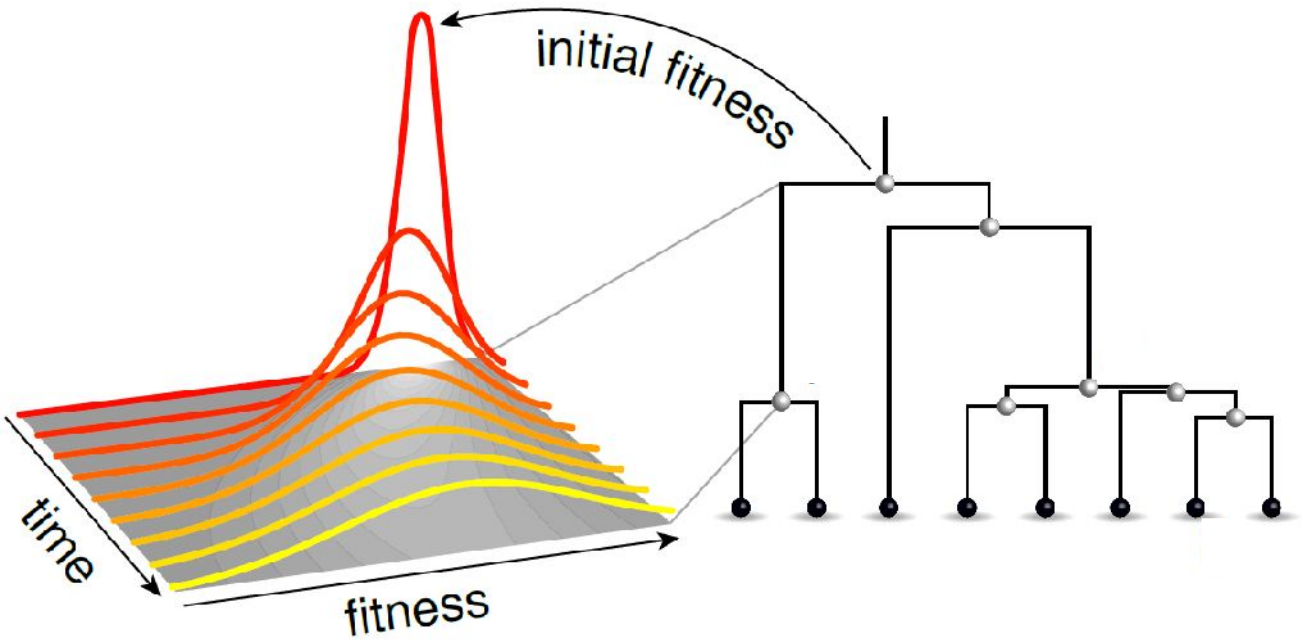
Fitness can be inferred from trees



$$P(\mathbf{x}|T) = \frac{1}{Z(T)} p_0(x_0) \prod_{i=0}^{n_{int}} g(x_{i_1}, t_{i_1}|x_i, t_i) g(x_{i_2}, t_{i_2}|x_i, t_i)$$

RN, Russell, Shraiman. eLife, 2014

Fitness can be inferred from trees

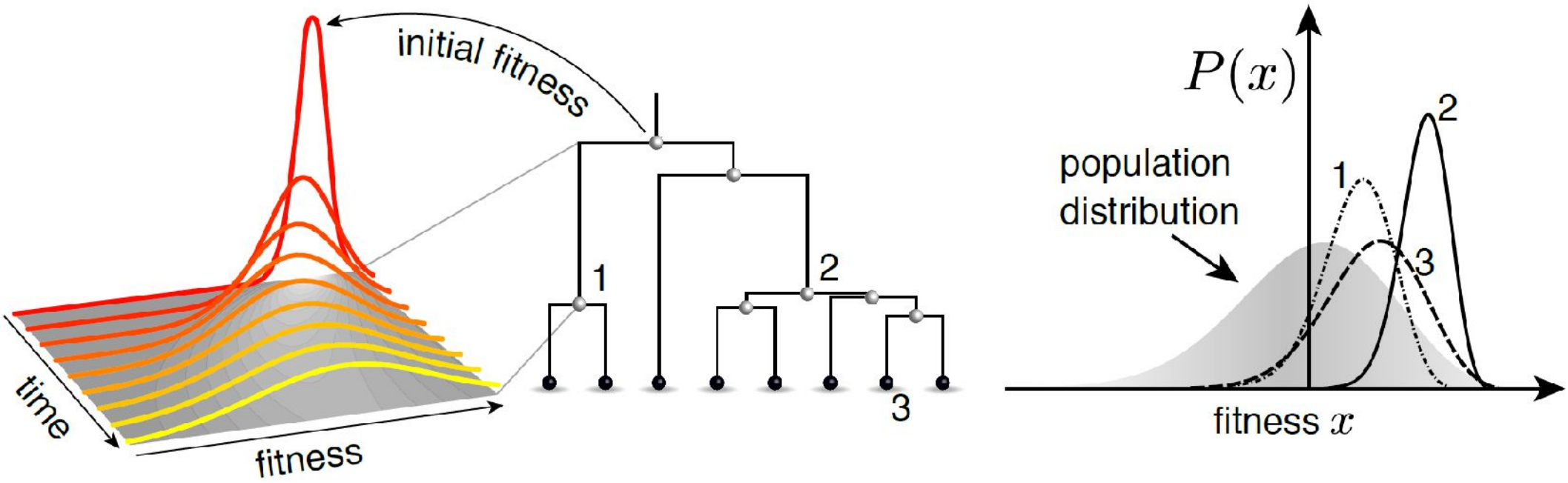


$$P(\mathbf{x}|T) = \frac{1}{Z(T)} p_0(x_0) \prod_{i=0}^{n_{int}} g(x_{i_1}, t_{i_1}|x_i, t_i) g(x_{i_2}, t_{i_2}|x_i, t_i)$$

$$\partial_t g(x, t'|y, t) = [y - 2\phi_\omega(y, t)] g(x, t'|y, t) - \sigma^2 \partial_y g(x, t'|y, t) + D\partial_y^2 g(x, t'|y, t)$$

RN, Russell, Shraiman. eLife, 2014

Fitness can be inferred from trees



$$P(\mathbf{x}|T) = \frac{1}{Z(T)} p_0(x_0) \prod_{i=0}^{n_{int}} g(x_{i_1}, t_{i_1}|x_i, t_i) g(x_{i_2}, t_{i_2}|x_i, t_i)$$

$$\partial_t g(x, t'|y, t) = [y - 2\phi_\omega(y, t)] g(x, t'|y, t) - \sigma^2 \partial_y g(x, t'|y, t) + D\partial_y^2 g(x, t'|y, t)$$

RN, Russell, Shraiman. eLife, 2014

High fitness corresponds to local bursts

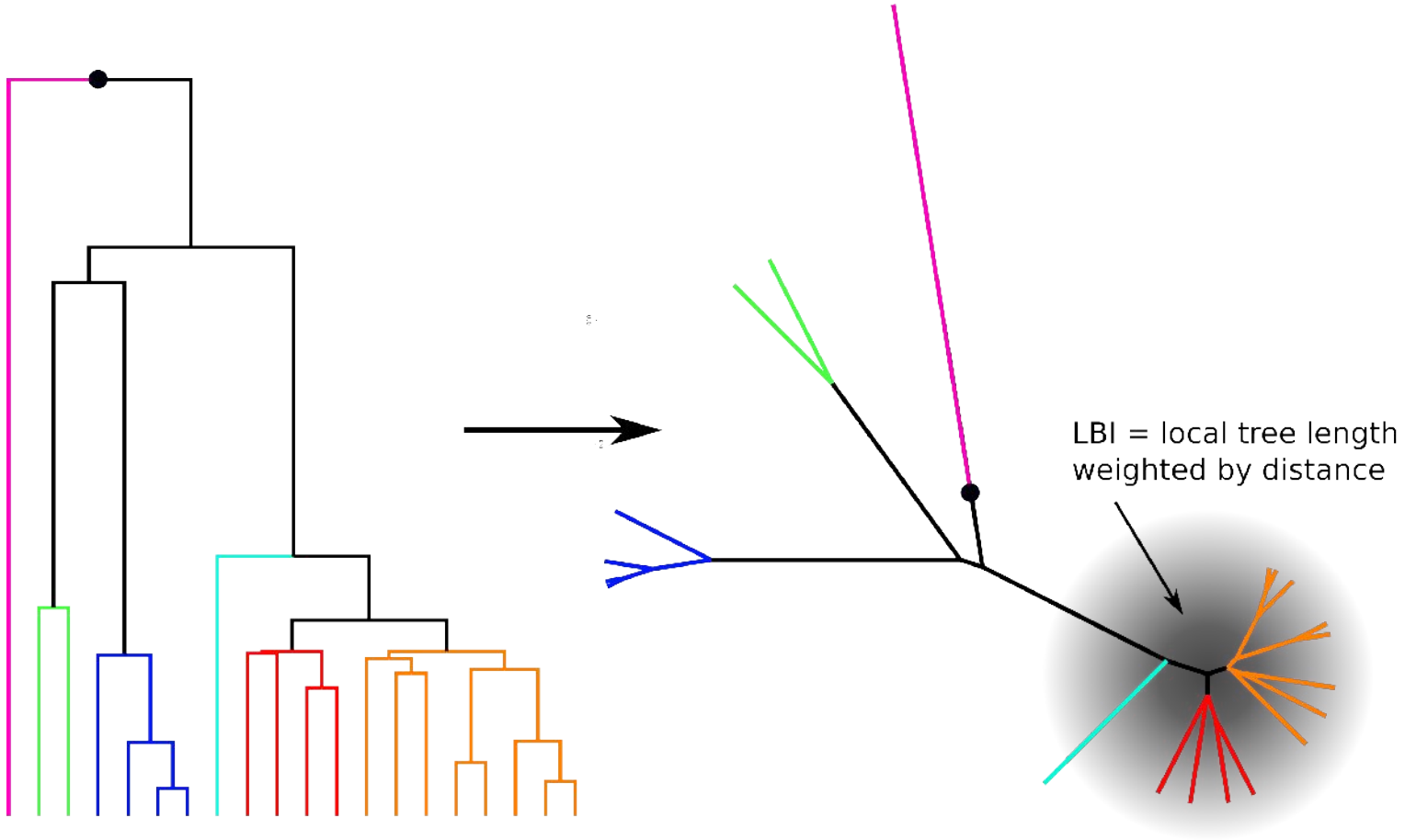
At short times: $\int dx\, g(x,t|y,t') \sim e^{y(t-t')}$

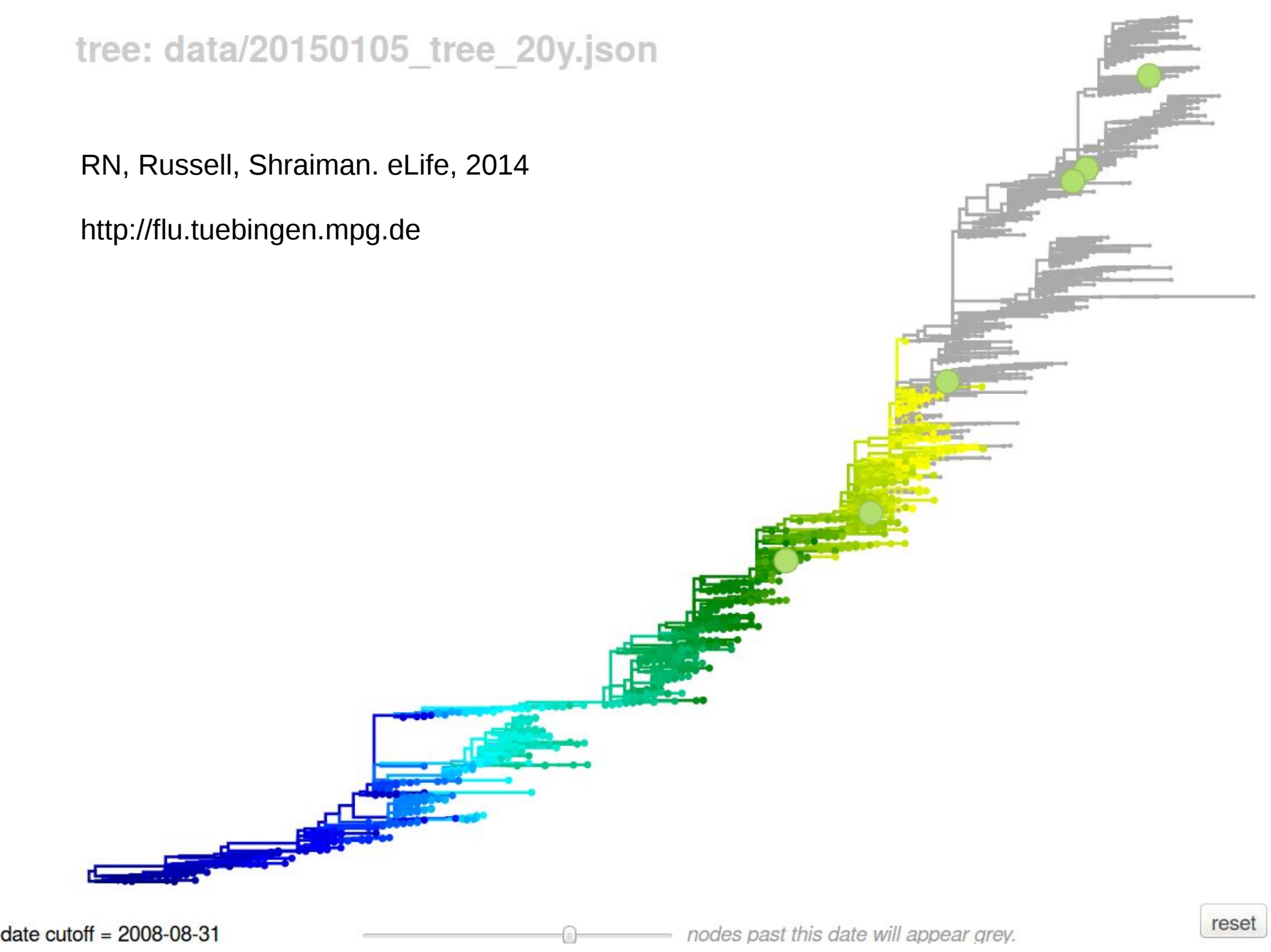
→ the product of propagators and prior: $\sim e^{yT-y^2/2}$

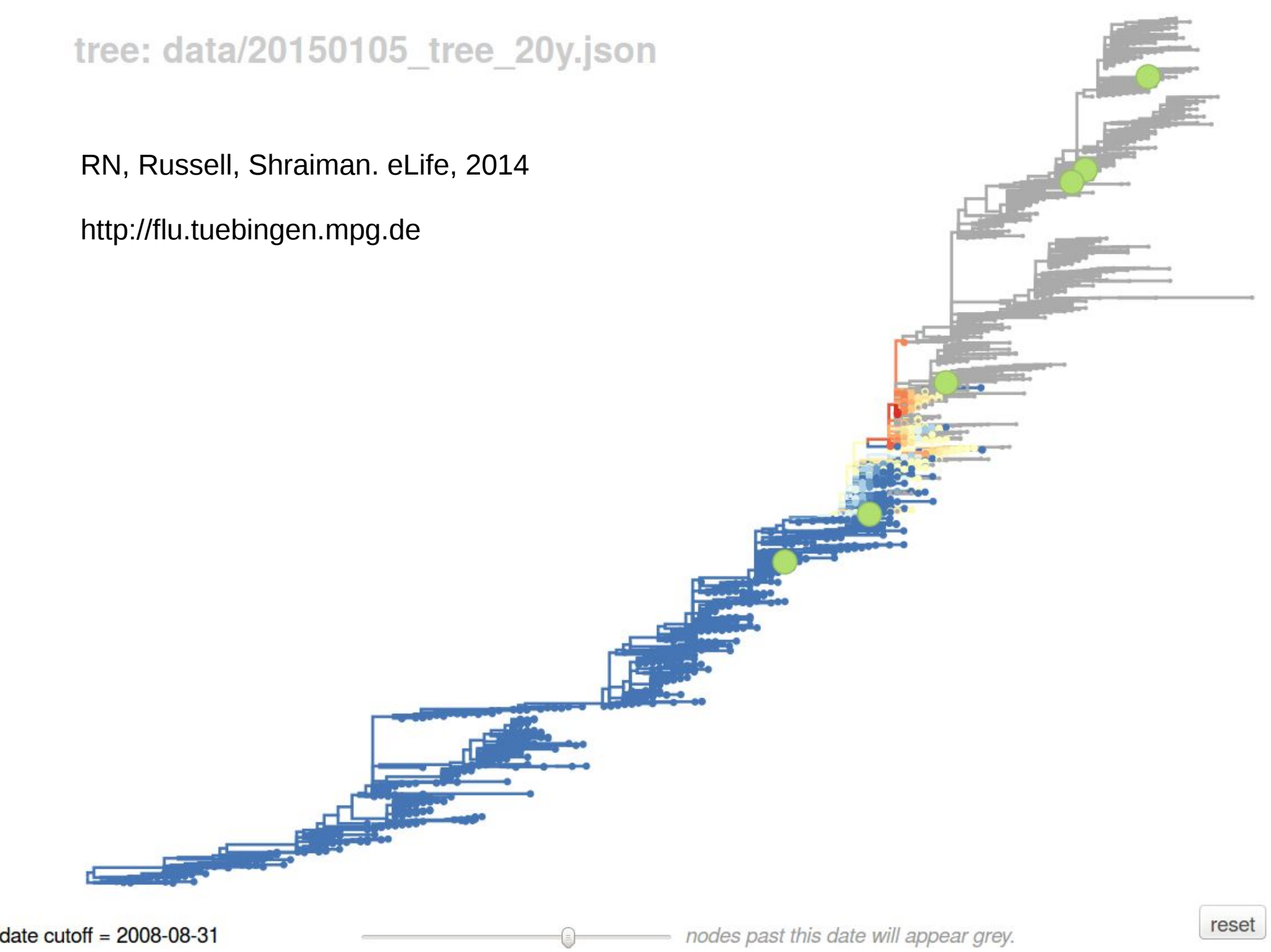
High fitness corresponds to local bursts

At short times: $\int dx\, g(x,t|y,t') \sim e^{y(t-t')}$

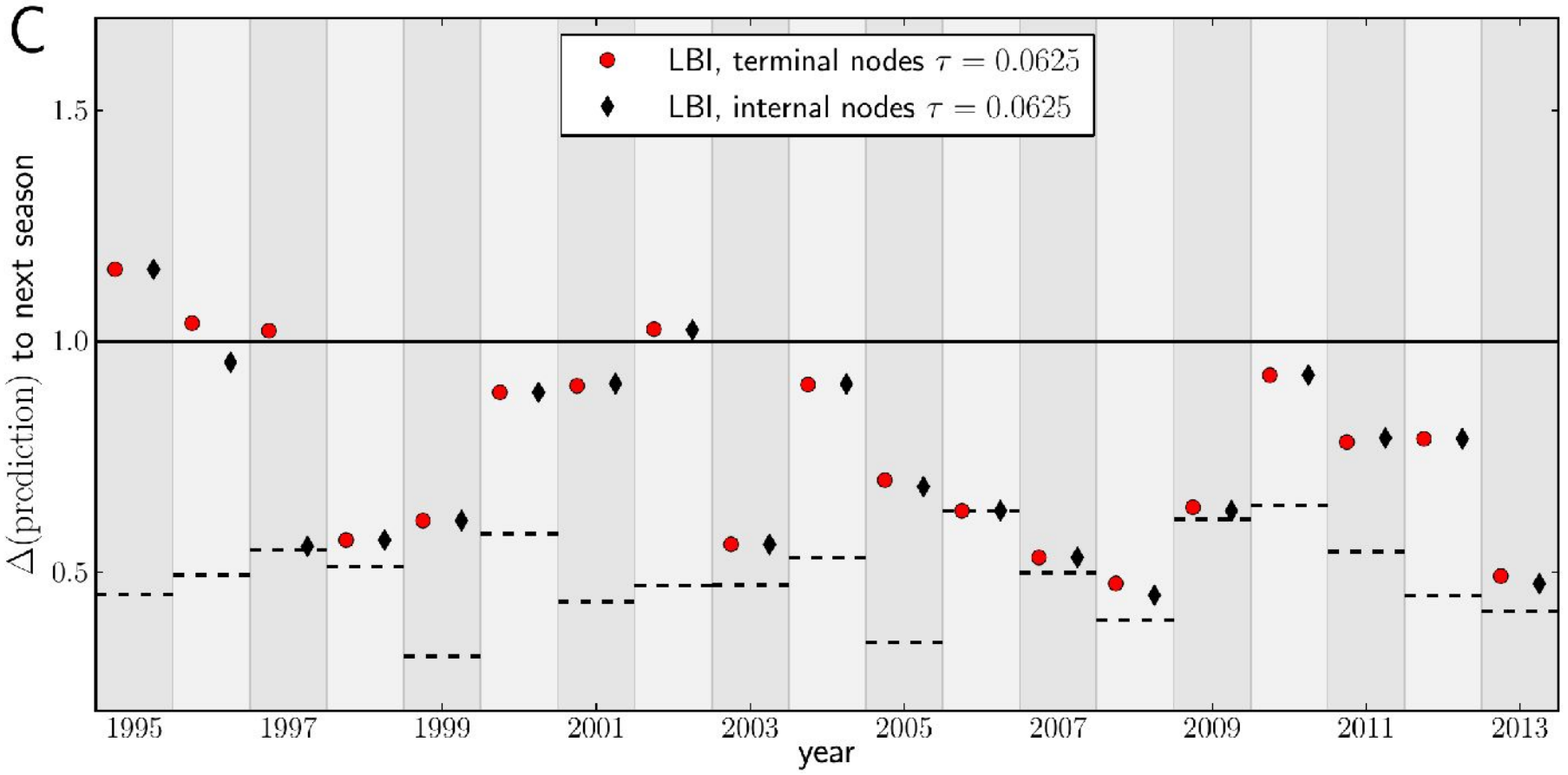
→ the product of propagators and prior: $\sim e^{yT-y^2/2}$



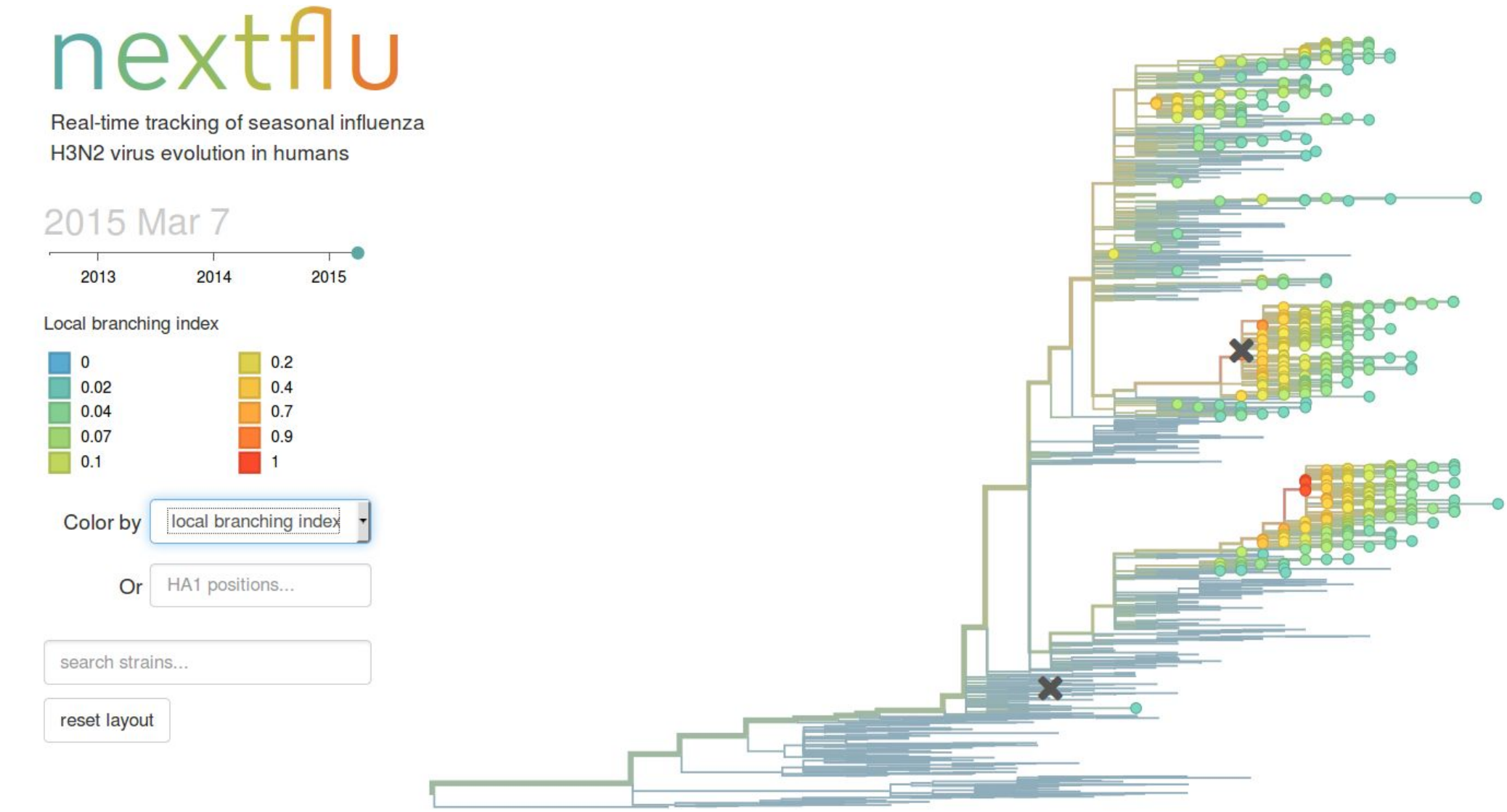

tree: data/20150105_tree_20y.json
RN, Russell, Shraiman. eLife, 2014
http://flu.tuebingen.mpg.de
date cutoff = 2008-08-31
nodes past this date will appear grey.
reset


tree: data/20150105_tree_20y.json
RN, Russell, Shraiman. eLife, 2014
http://flu.tuebingen.mpg.de
date cutoff = 2008-08-31
nodes past this date will appear grey.
reset

Past evolution of influenza A/H3N2 is predictable



- requires only tree as input: generally applicable
  - → emerging diseases
  - → cancer cell populations



with Trevor Bedford, available at nextflu.org

nextflu

Real-time tracking of seasonal influenza
H3N2 virus evolution in humans

2015 Mar 7

2013 2014 2015

Local branching index

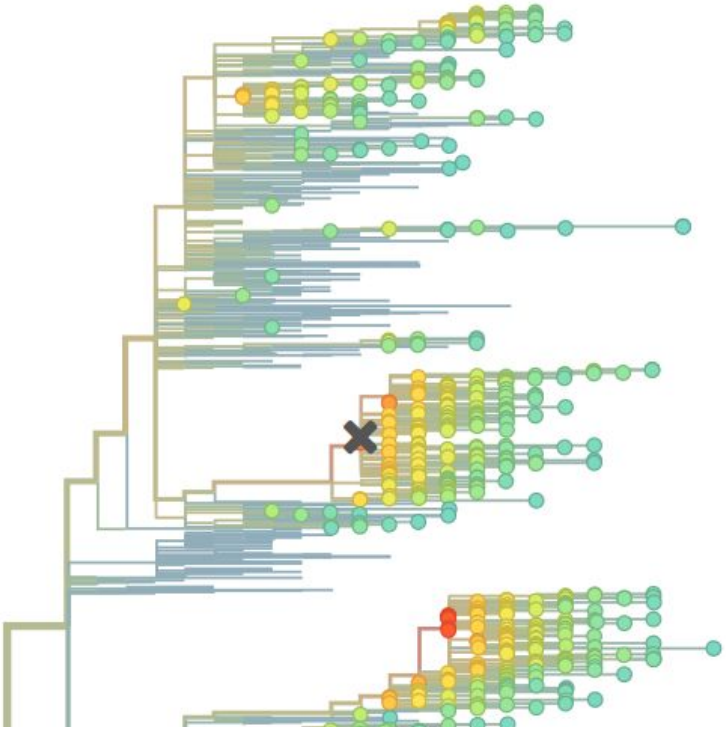
0 0.2
0.02 0.4
0.04 0.7
0.07 0.9
0.1 1

Color by: local branching index

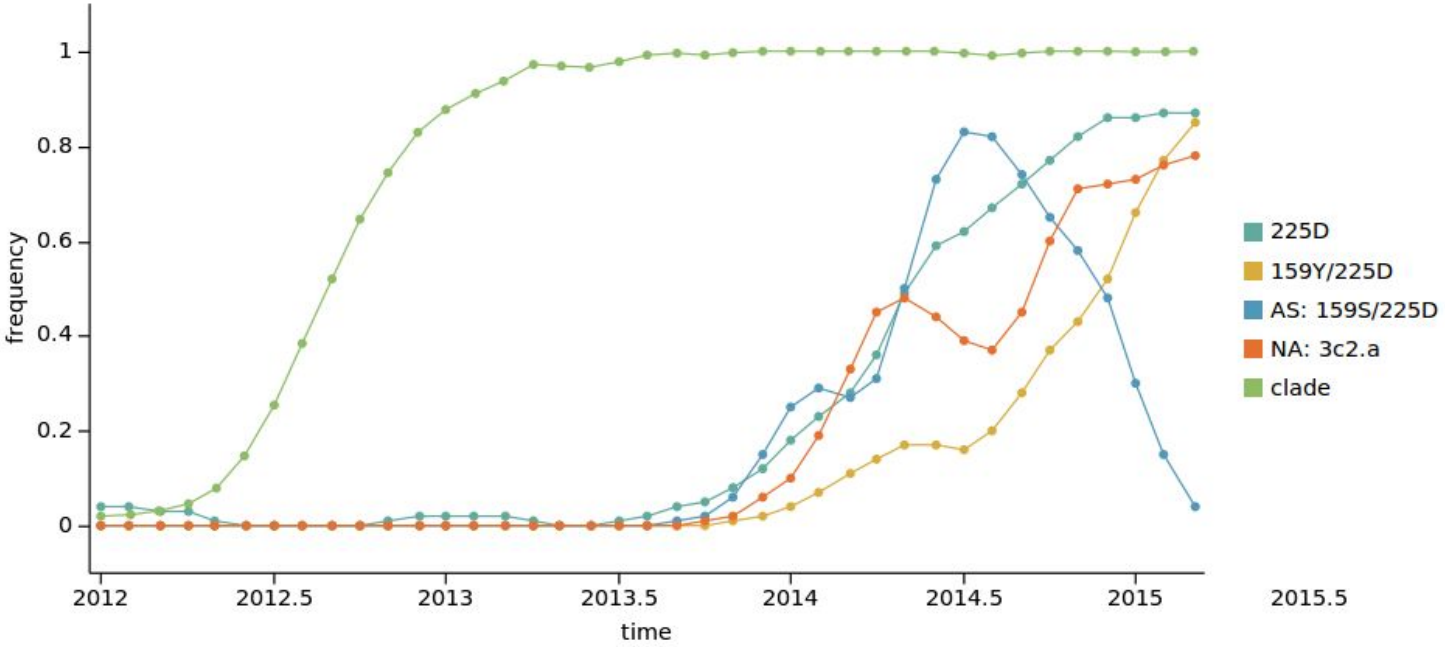
Or: HA1 positions...

search strains...

reset layout

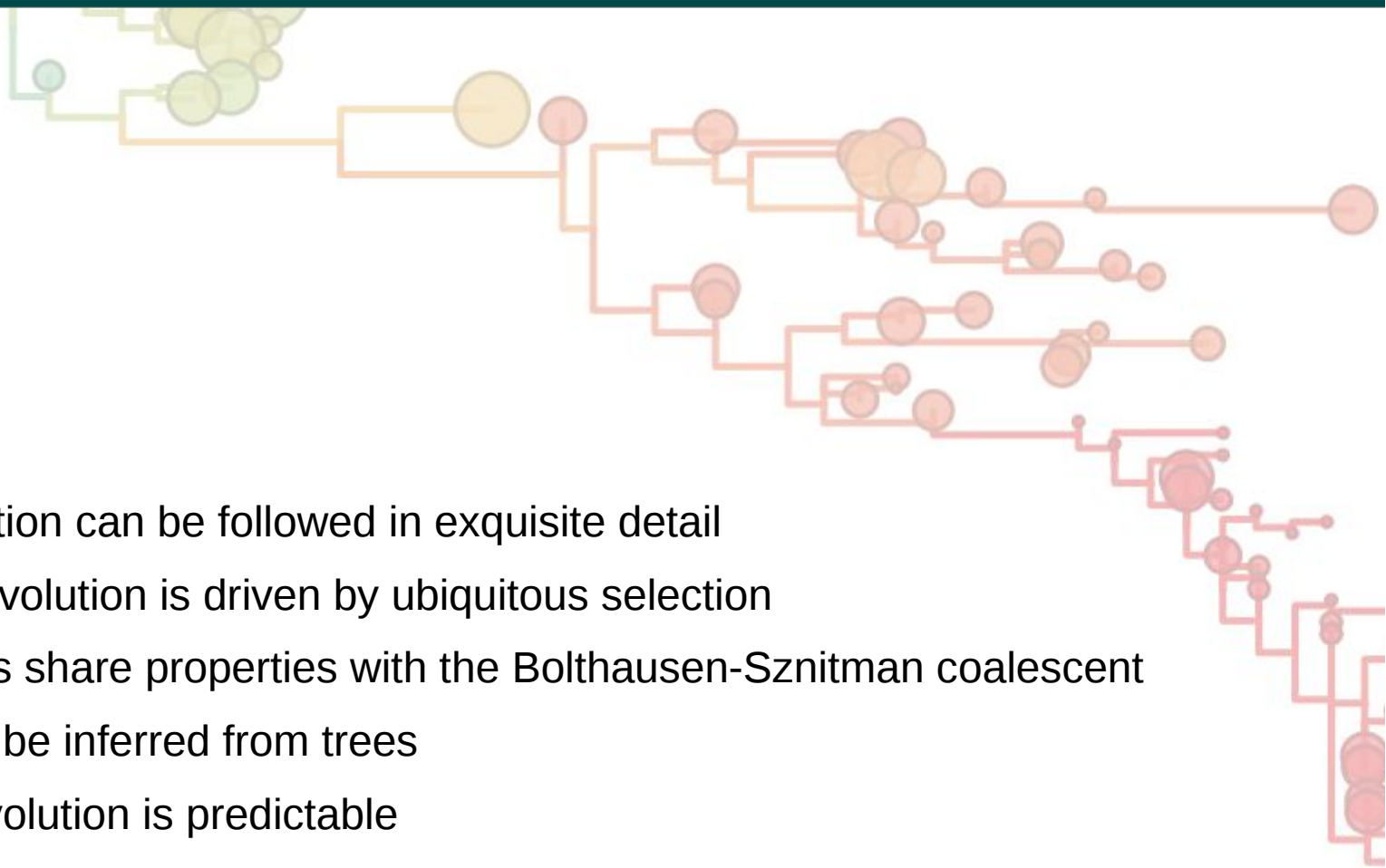


Mutation and clade frequencies



with Trevor Bedford, available at nextflu.org

Summary



- HIV-1 evolution can be followed in exquisite detail
- RNA virus evolution is driven by ubiquitous selection
- Genealogies share properties with the Bolthausen-Sznitman coalescent
- Fitness can be inferred from trees
- Influenza evolution is predictable
- Browser-based tools to integrate information

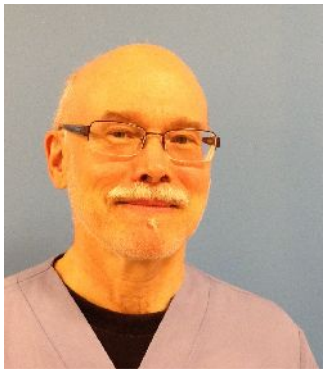
Acknowledgements – HIV

Fabio Zanini
MPI

Jan Albert
KI

Christa Lanz
MPI

Johanna Brodin
KI



Göran Bratt
SSGH

Lina Thebo
KI

Discussions: Thomas Leitner, LANL

Web design: Bianca Regenbogen, Tübingen

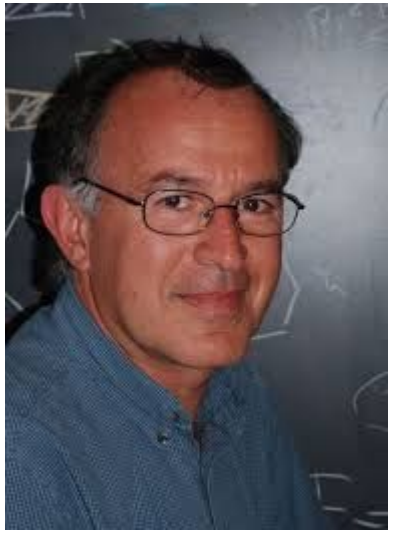
MAX-PLANCK-GESELLSCHAFT

erc

KAROLINSKA INSTITUTET ANNO 1810

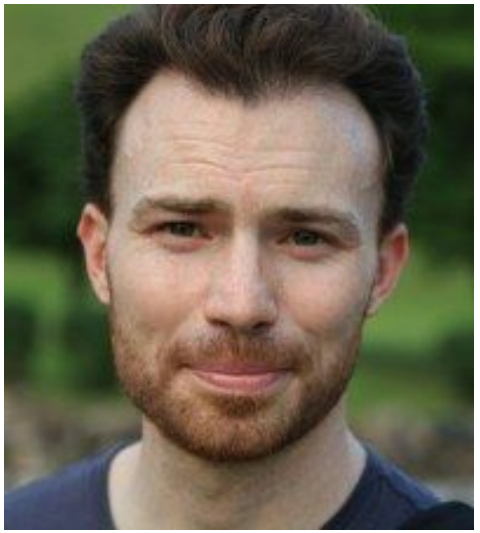
Karolinska Institutet

Acknowledgements – Theory and Influenza



Boris Shraiman

Oskar Hallatschek



Colin Russell

Trevor Bedford

MAX-PLANCK-GESELLSCHAFT

erc