Intra patient evolution of HIV

Richard Neher

Max-Planck-Institute for Developmental Biology

Acknowledgements

erc

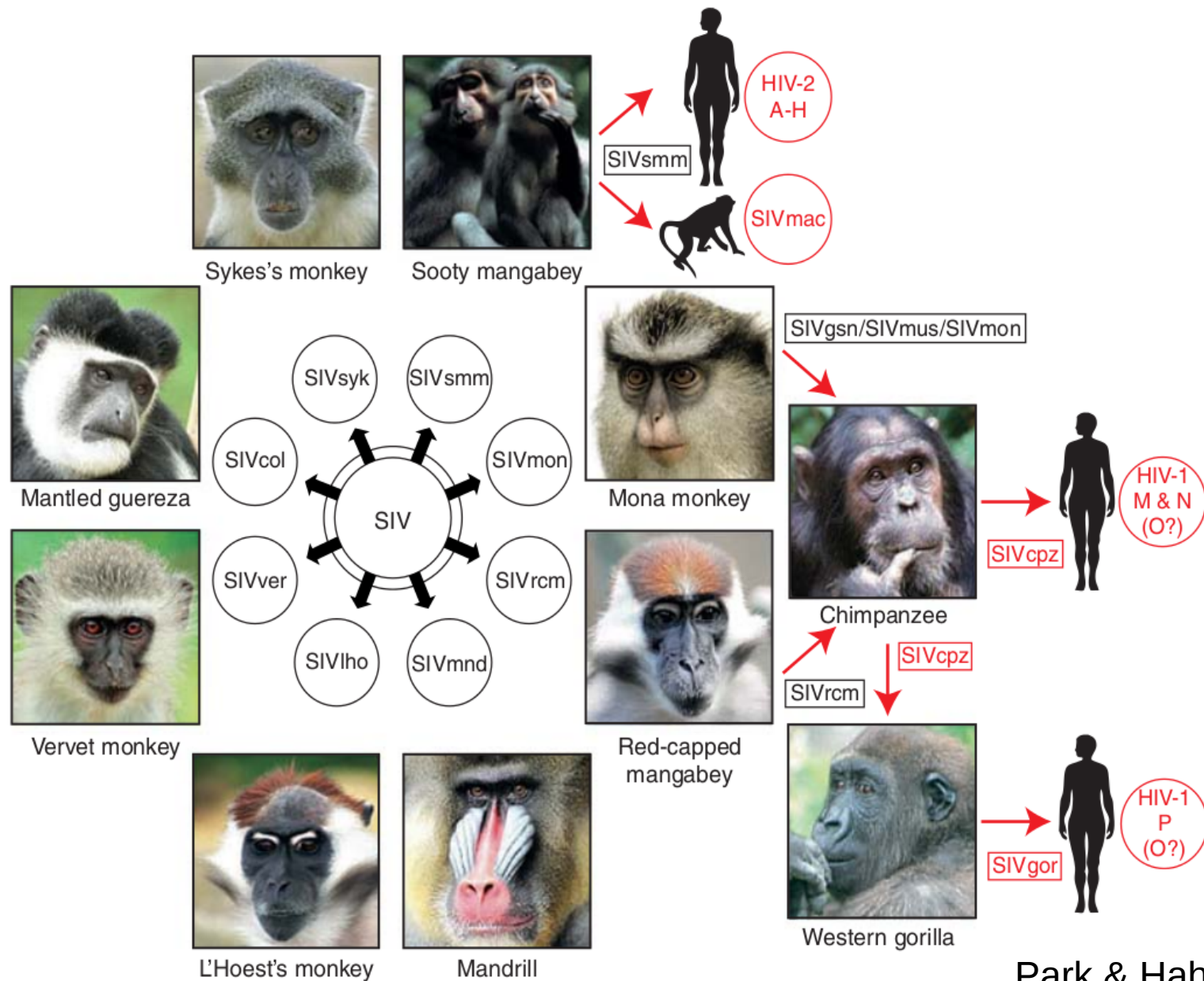# Origins of HIV

Park & Hahn, CSH Pers. Med, 2011

Zoonosis of SIV/HIV-1
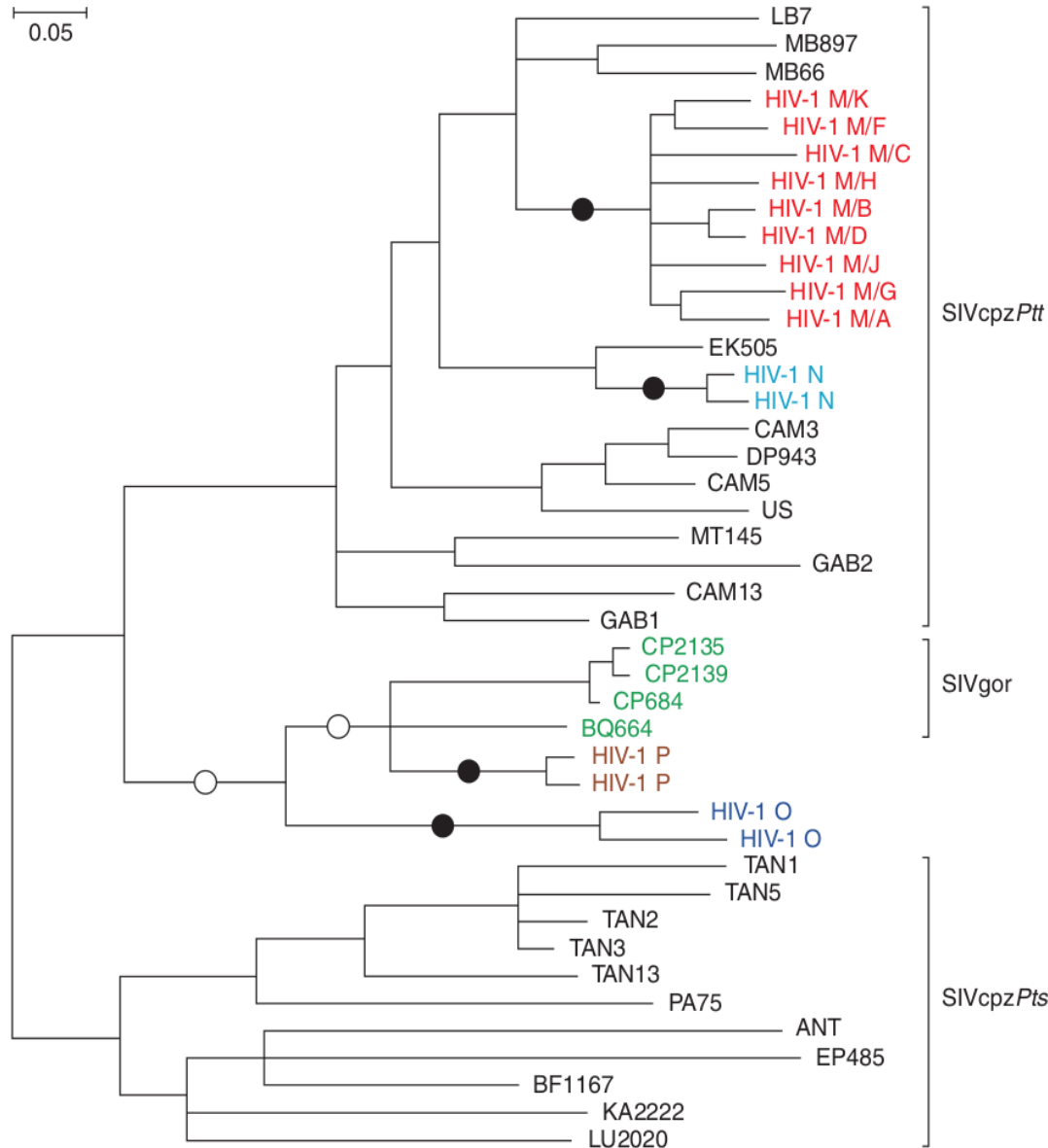
- Several primate-human transmissions
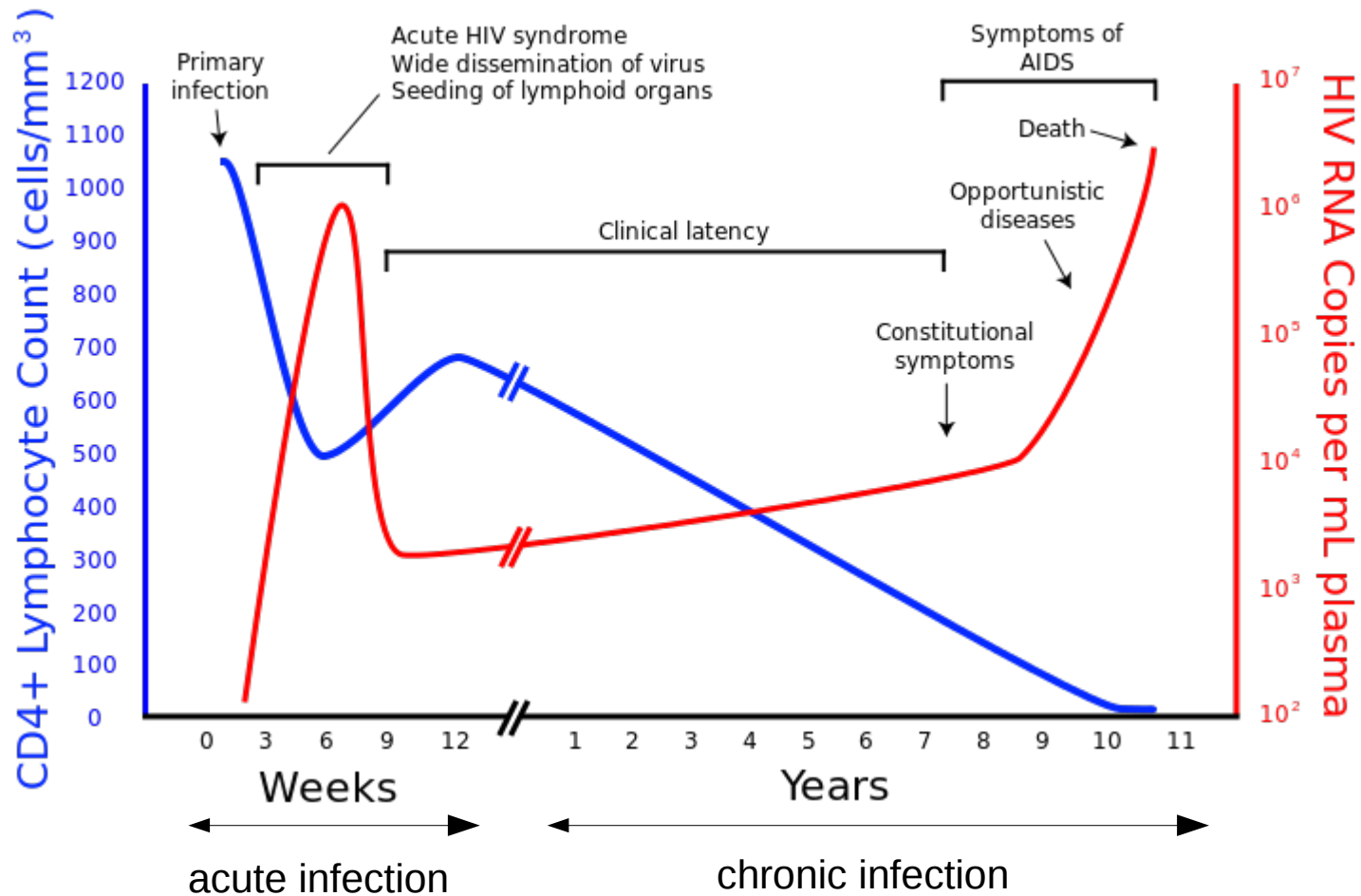- MRCA of HIV-1 M early 1900
- Diversification into subtypes
- North America, western Europe: mainly subtype B

Nucleotide distances
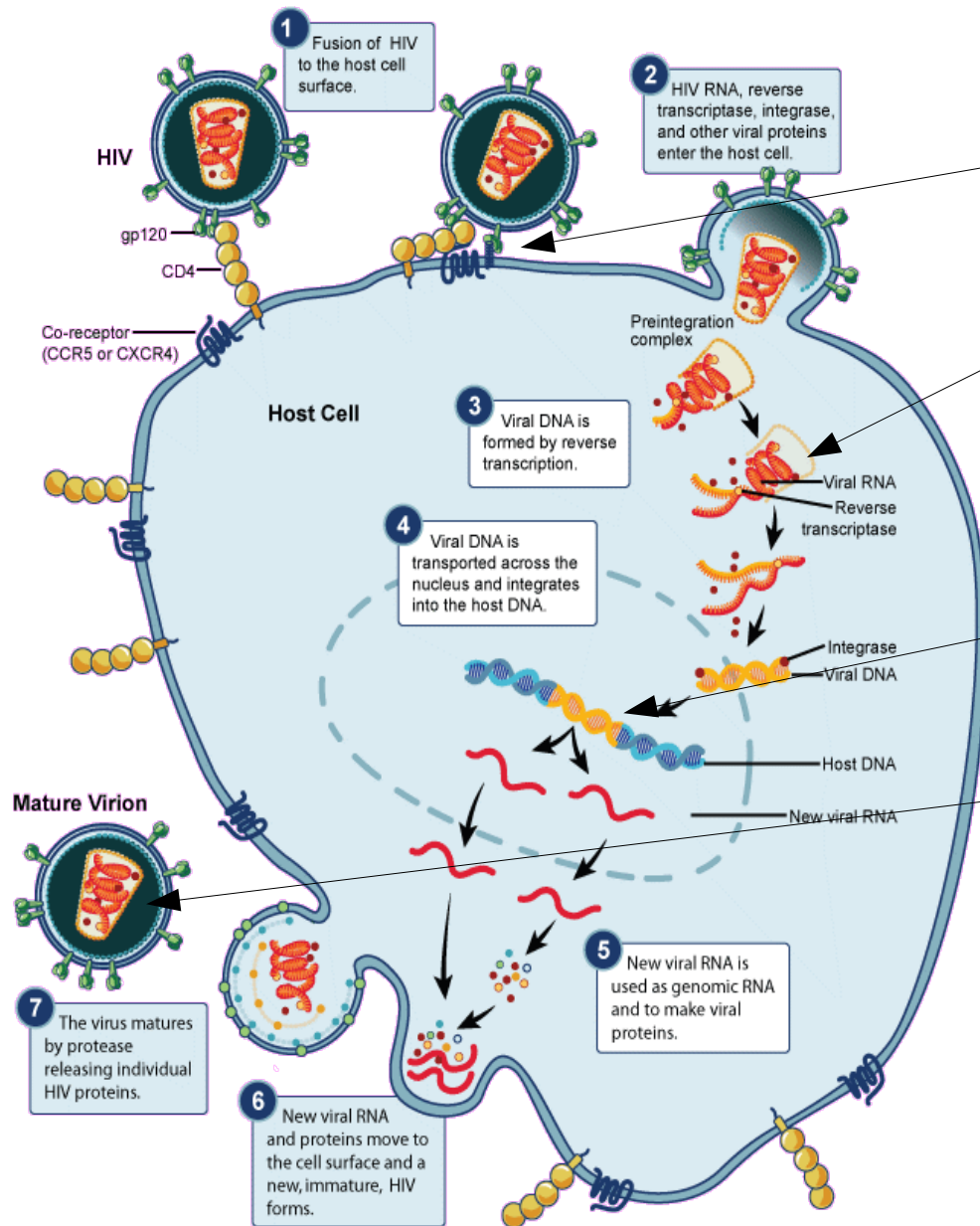
Sharp & Hahn, CSH Pers. Med, 2011

Disease progression

- 10^7 CD4+ cells infected per day
- generation time 2 days
- mutation rate approx 10^{-5}/nuc

→ every mutation produced every day

Image: wikipedia

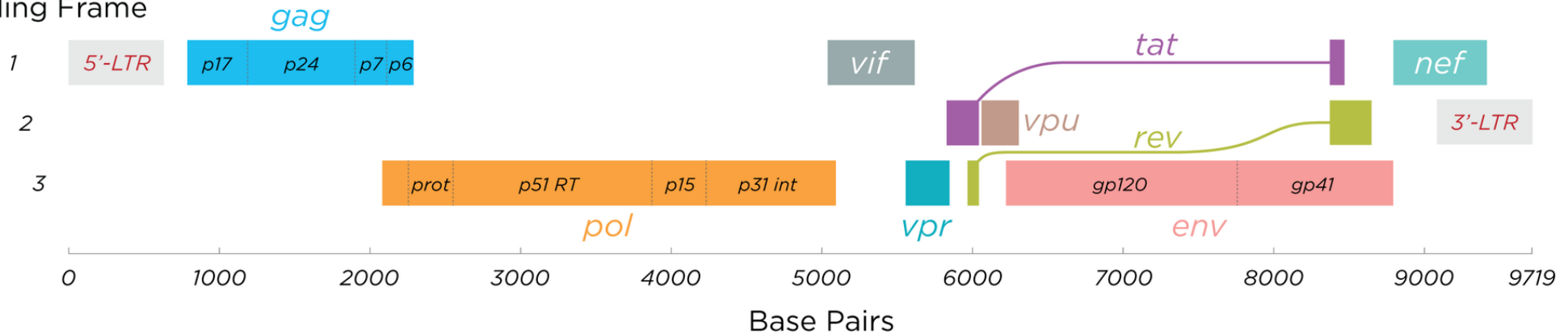# Drug targets

Fusion inhibitors

Reverse transcriptase inhibitors
 - nucleoside analogs
 - non-nucleoside analogs

Integrase inhibitors

Protease inhibitors

Genome and drug resistance

| Subtype[2]	protease		RT		integrase	
	PI-Naive	PI-Treated	RTI-Naive	RTI-Treated	INI-Naive	INI-Treated
B	27,672	9,557	25,469	29,800	3,840	191
A	4,672	262	3,849	1,285	337	1
C	10,391	1,208	8,623	7,774	1,154	2
D	1,751	299	1,319	778	138	6
F	1,105	1,526	723	732	228	4
G	1,570	339	1,577	2,003	152	1
CRF01_AE	6,021	186	5,538	3,852	1,121	1
CRF02_AG	4,217	215	2,906	1,779	548	2
Other	1,617	261	1,957	3,063	237	
Total	59,016	13,853	51,961	51,066	7,755	208

Image: https://en.wikipedia.org/wiki/HIV_genome#mediaviewer/File:HIV-genome.png
Shafer RW(2006). Rationale and Uses of a Public HIV Drug-Resistance Database. Journal of Infectious Diseases 194 Suppl 1:S51-8
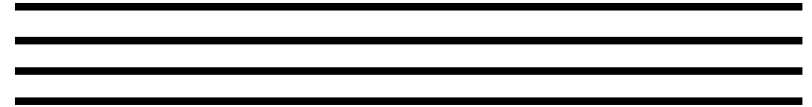
Drug resistance
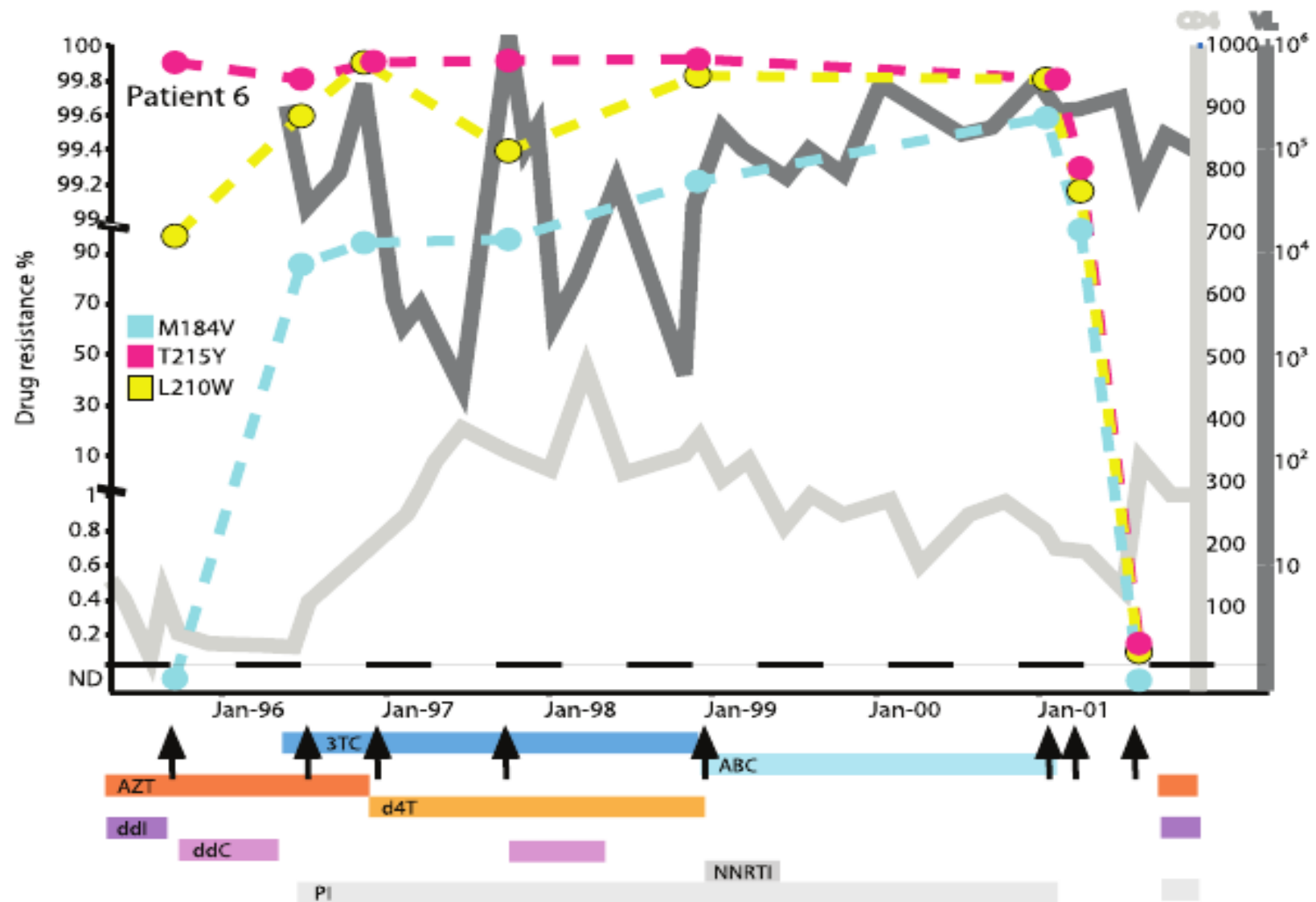
HIV and the immune system

CD4+ (infected cell)

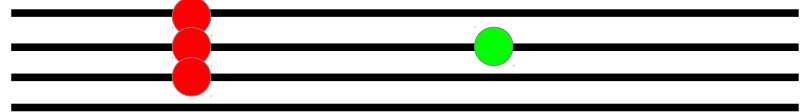CD8+ (killer cell)

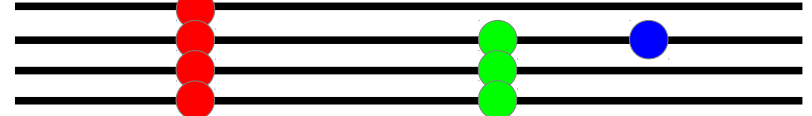MHC I molecule loaded
with viral peptide
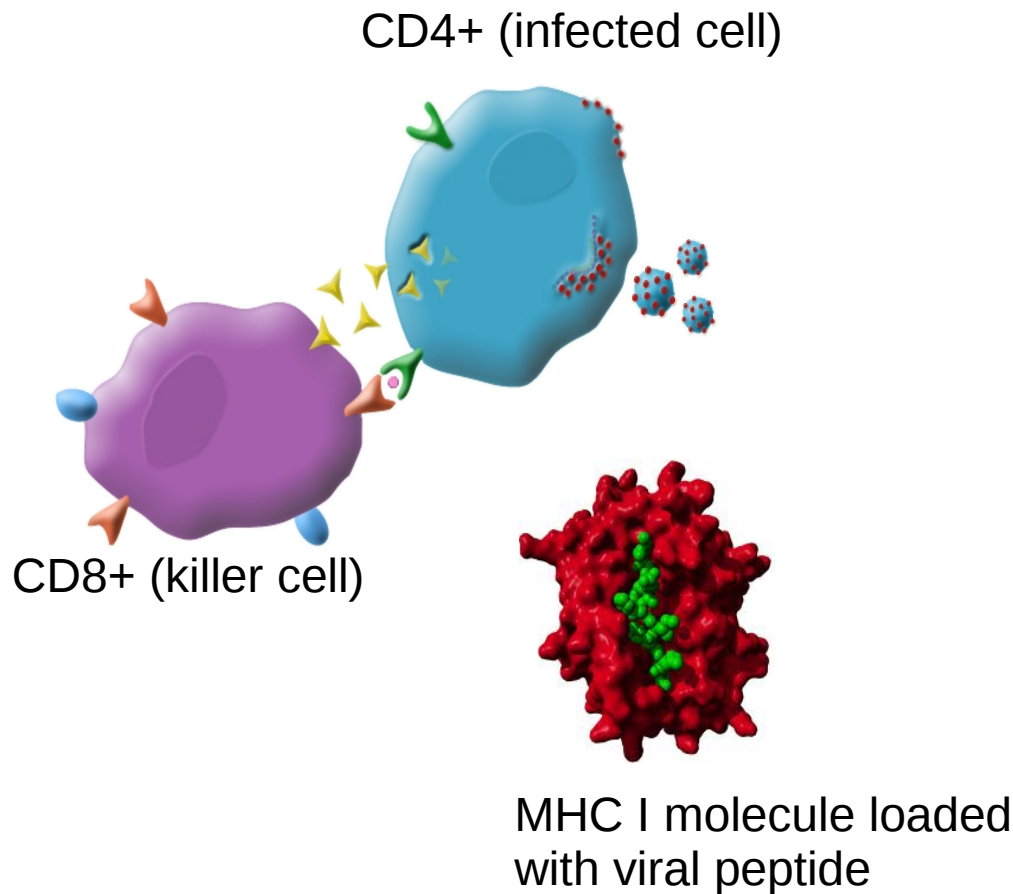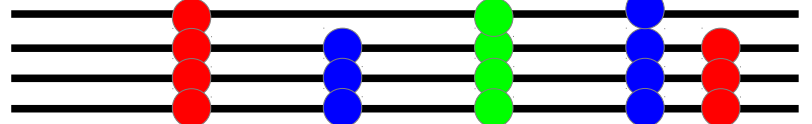
The virus changes
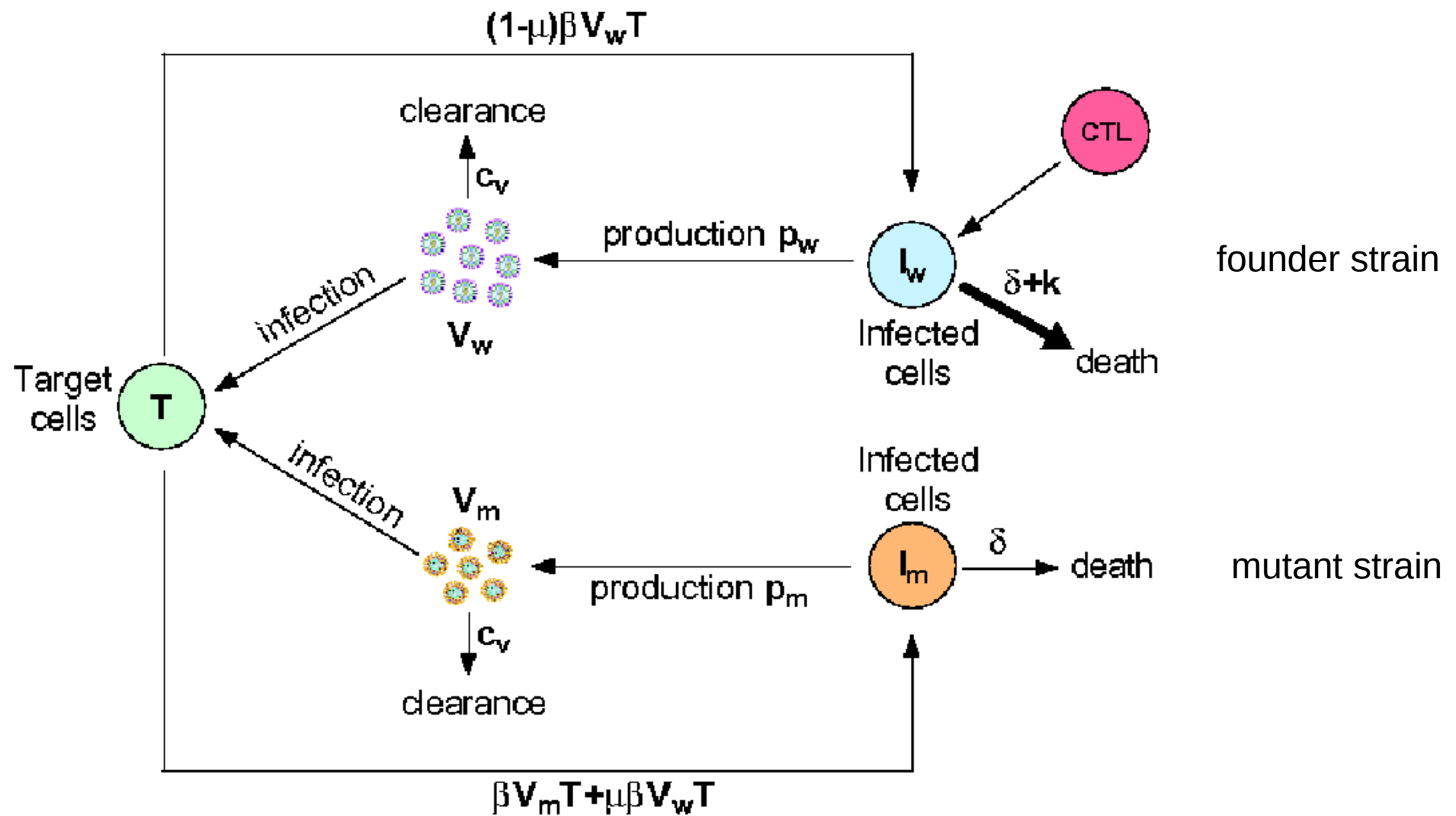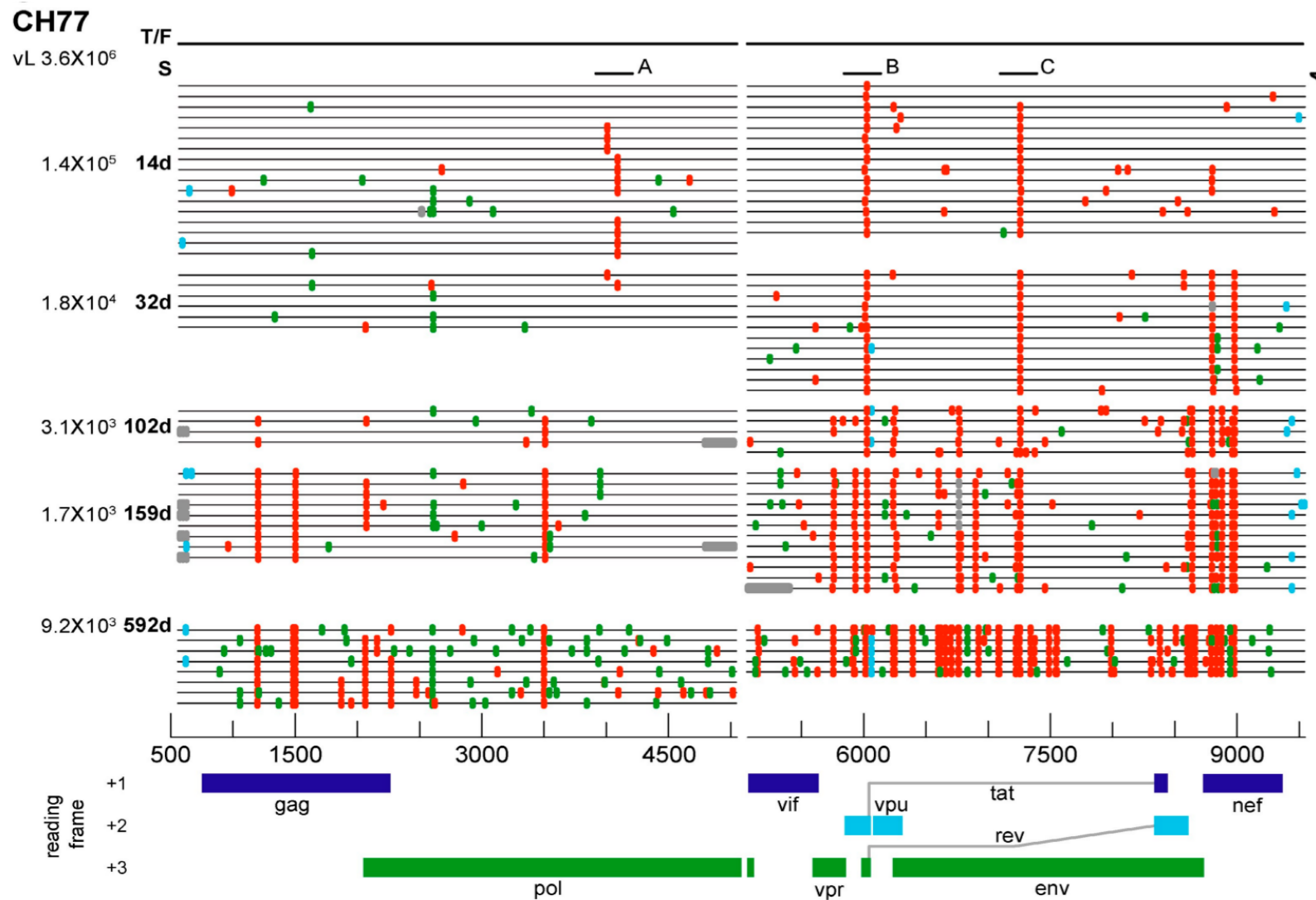
7 days

3 weeks

8 weeks

6 weeks

- MHC molecules present 8-10aa peptide

- Which peptides are presented depends on the HLA genotype

- HIV avoids presentation by adapting to the HLA genotype

Image: pasteur.fr

Modeling of the virus population
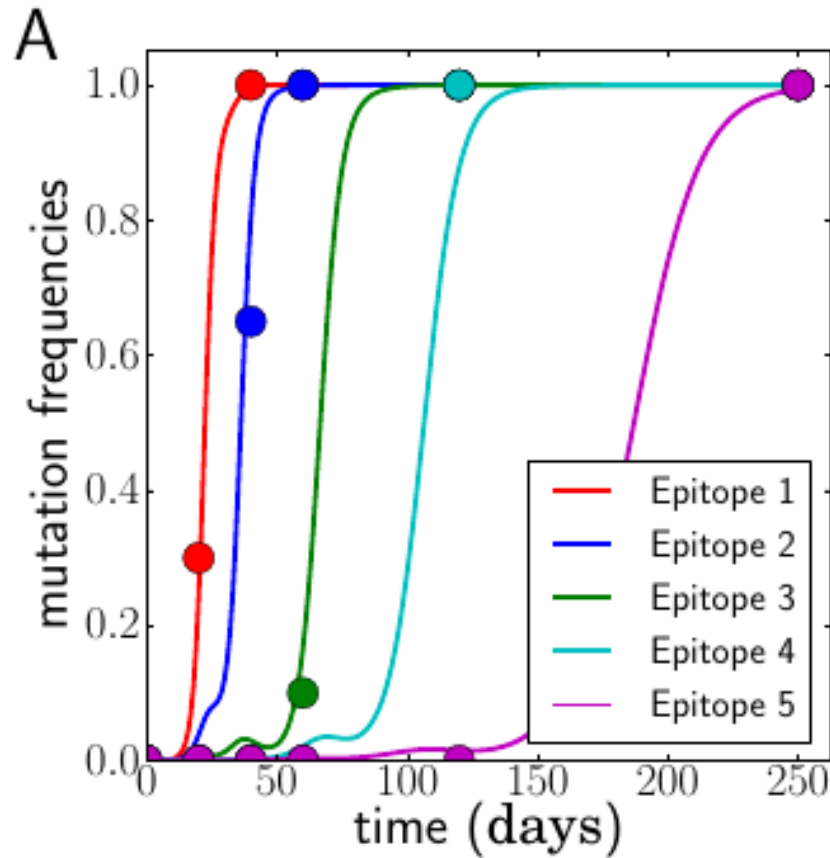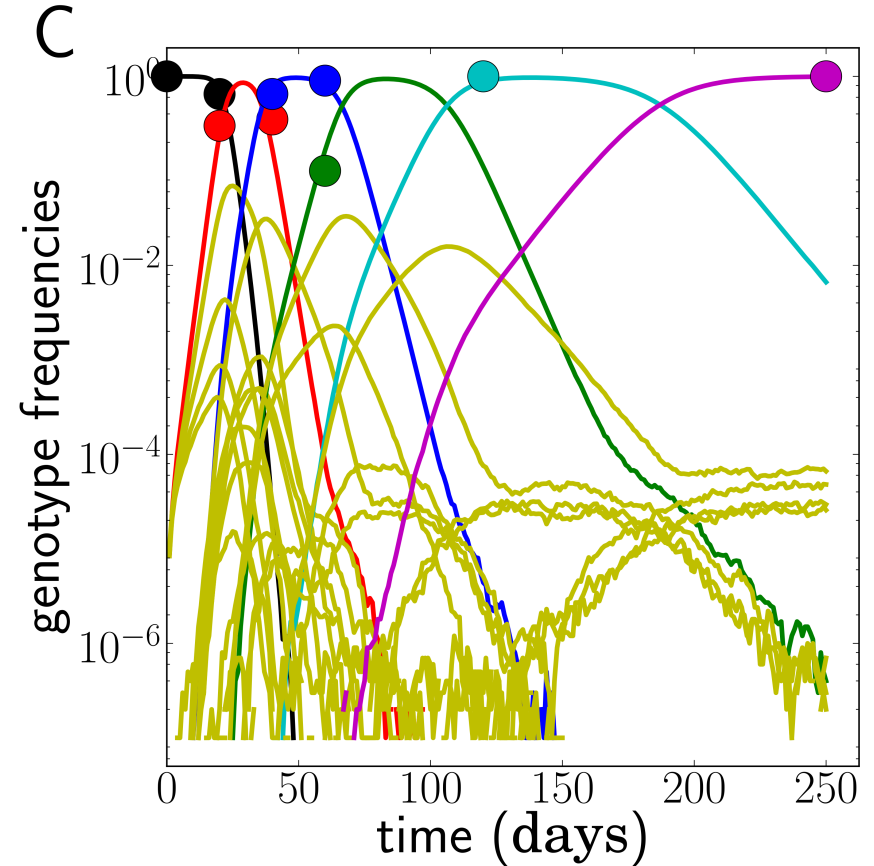

A Perelson

CTL escape

Rapidly spreading mutations that prevent presentation or recognition of CTL epitopes.

Salazar-Gonzales et al, JEM, 2009
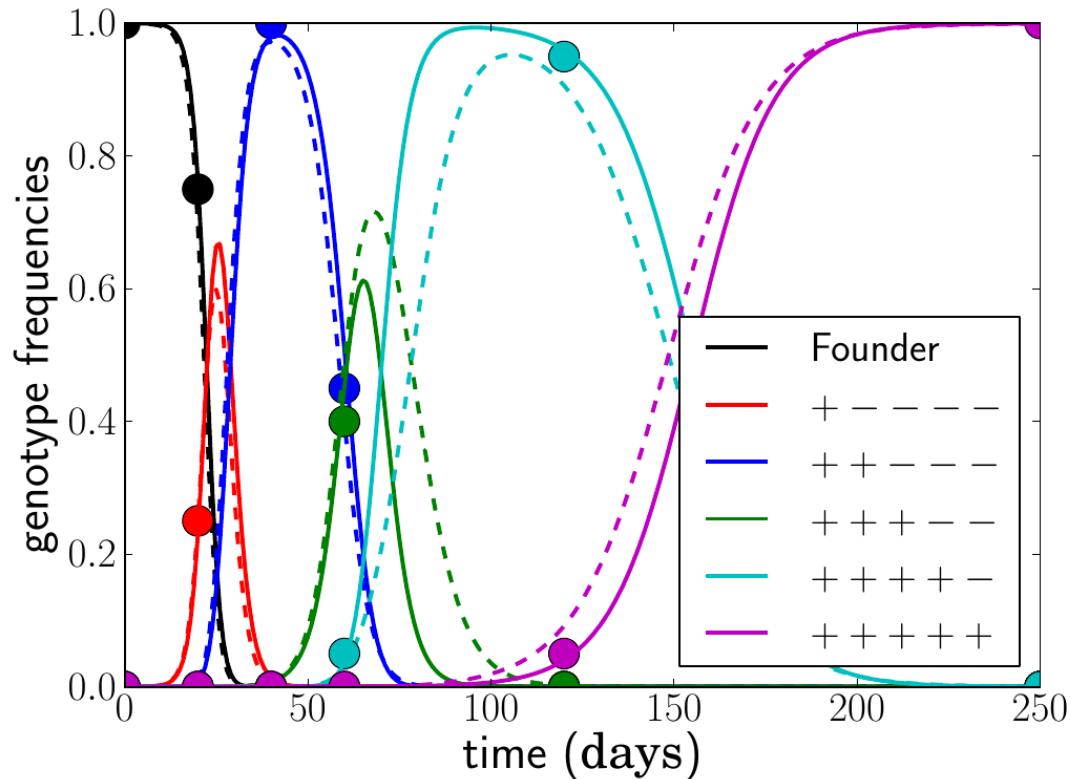
Inferring CTL escape rates

A — single site variant frequencies

C — dominant genotypes

Inferring CTL escape rates

- Dashed: simulation
- Dots: samples used for fitting
- Solid: model fit

Legend:
- Founder (black)
- + − − − − (red)
- + + − − − (blue)
- + + + − − (green)
- + + + + − (cyan)
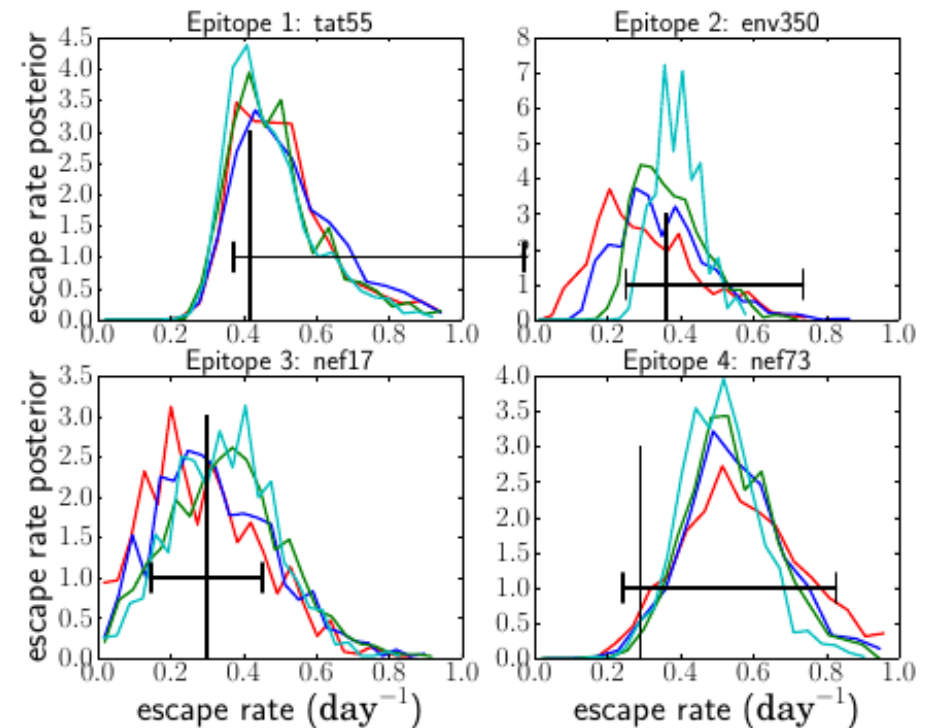- + + + + + (magenta)

- additive model for growth rate
- each locus parametrized by escape rate & establishment time
- maximize the likelihood of the sample

Kessinger, Perelson, RN. Frontiers in Immunology
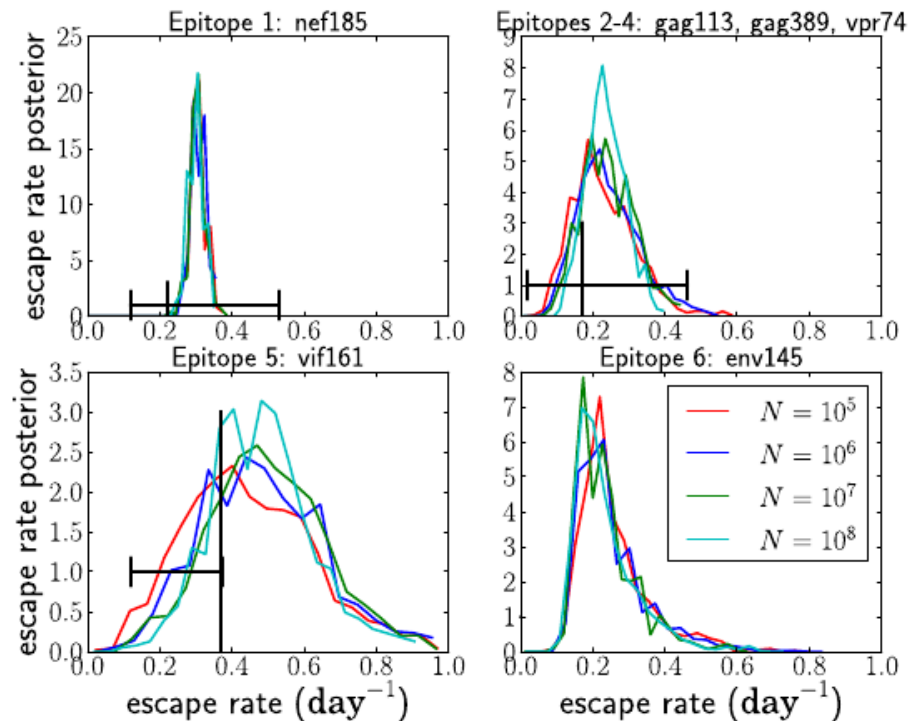
CTL escape rates

Fit results for two patients

- Different colors: N=10^5-10^8

- Escape rates around 0.4

Kessinger, Perelson, RN. Frontiers in Immunology

Estimating CTL escape rates

	control	Day 0	Day 16	Day 45
SLAFRHVAR	99.93	99.14	48.38	0.37 %
--------Q	0	0.02	4.49	86.71 %
----H----	0.01	0	24.07	9.08 %
N---------	0	0.02	1.80	2.51 %
R---------	0	0	3.65	0.11 %
-------M--	0.01	0	5.43	0.05 %
-------E--	0	0	1.86	0.05 %
----C-----	0.01	0.02	5.56	0 %
I---------	0.01	0	1.40	0 %
G---------	0	0.05	0.69	0 %
-------A--	0	0.12	0.48	0 %
---V-----	0	0	0.35	0 %
other	0.04	0.62	1.86	1.12 %
Unique epitope peptides:	9	18	55	20
Unique full-length DNAs:	85	127	455	124
Total sequence reads:	11,318	4,048	7,754	1,873
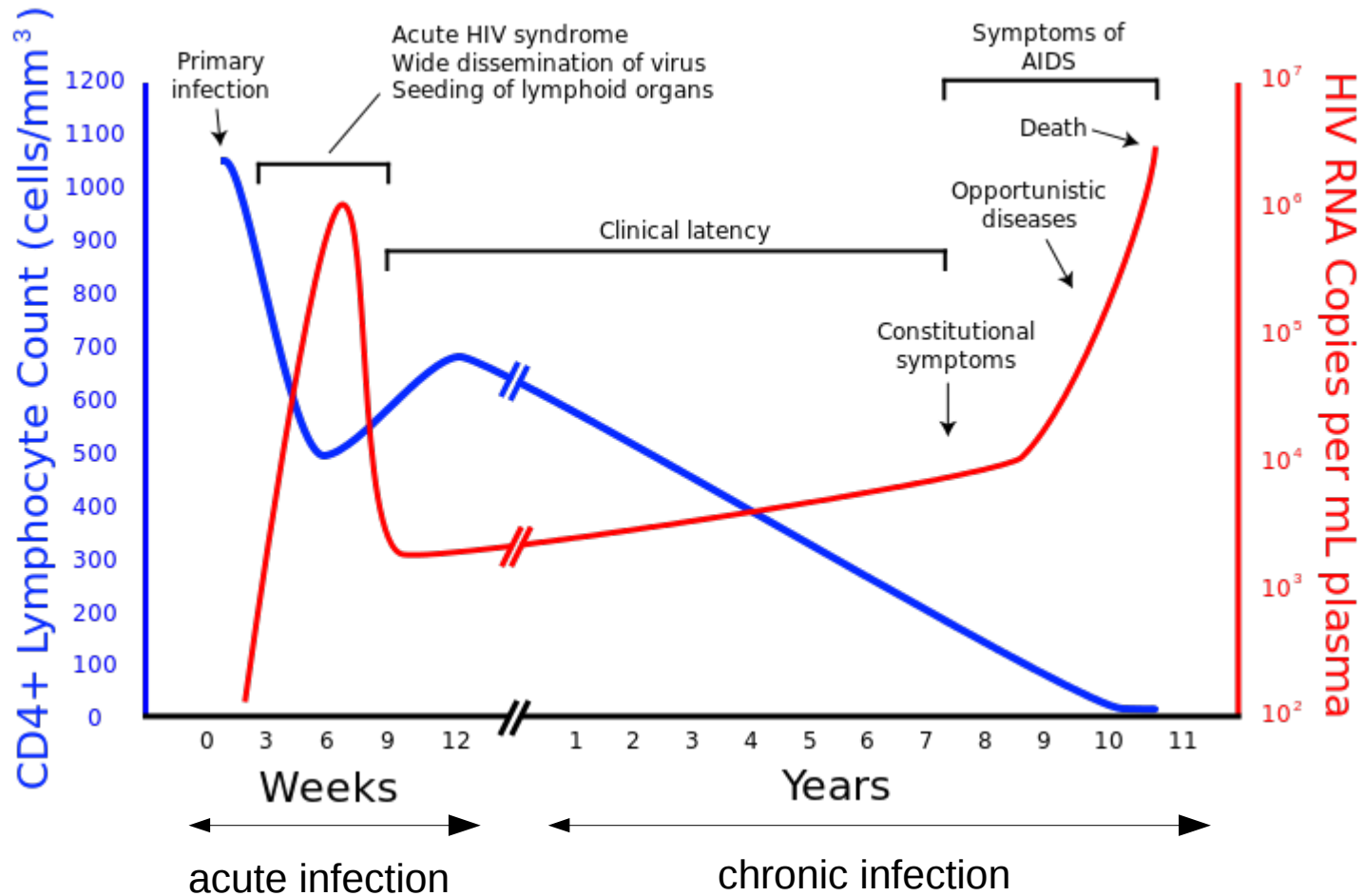
Fischer et al, PLoS One, 2010

HIV infection

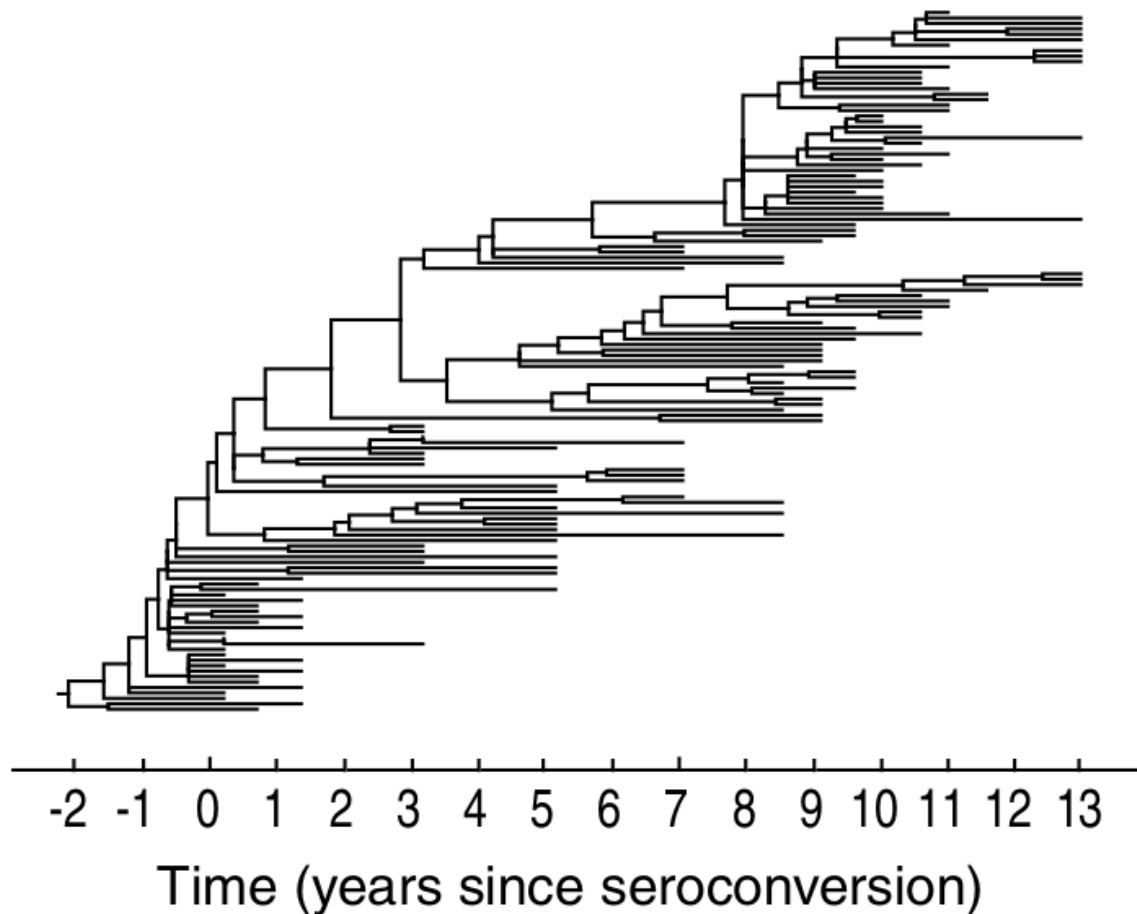- 10^7 CD4+ cells infected per day
- generation time 2 days
- mutation rate approx 10^{-5}/nuc

→ every mutation produced every day

Image: wikipedia
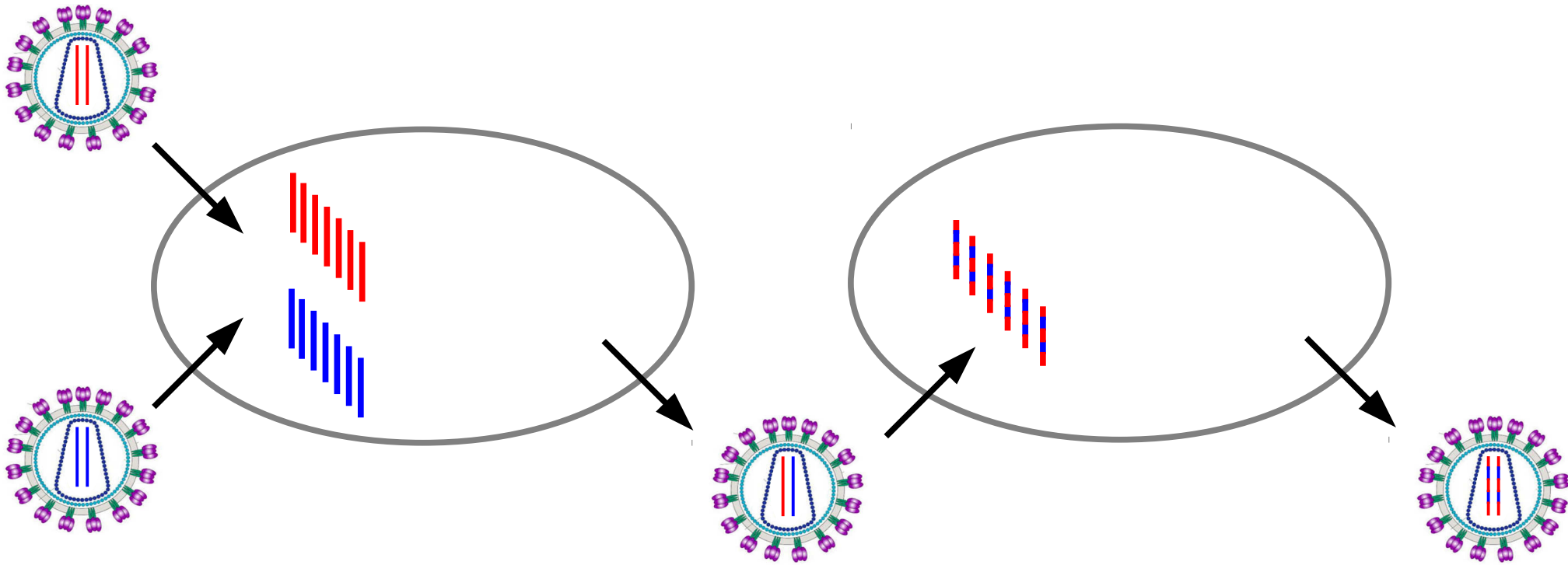
HIV vs S. aureus evolution

patient 2

Time (years since seroconversion)

~ 10 years,
~ 50 mutations in 600 bases
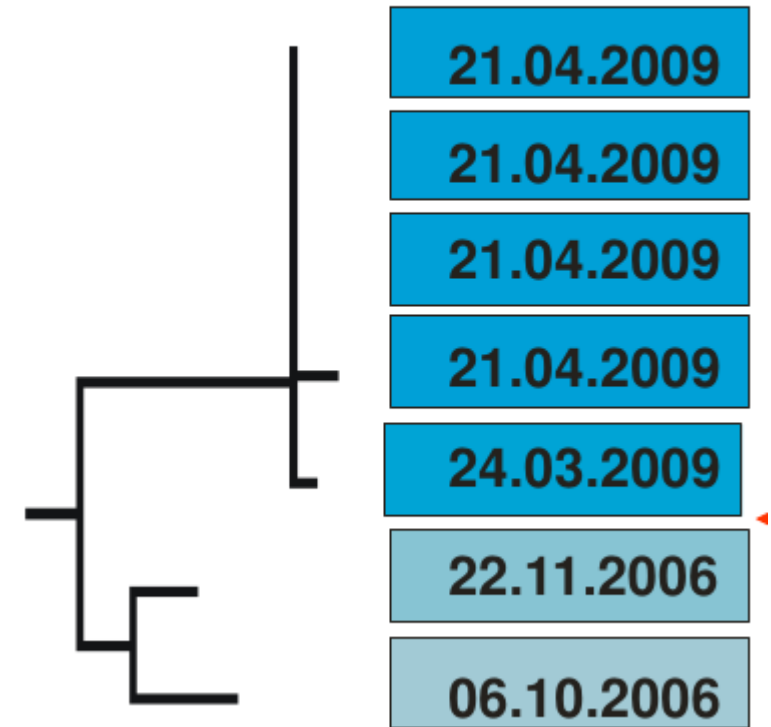
A few mutations in 3e6 bases
→ but duplications and HGT

Figure from Lemey et al, data from Shankarappa et al, 1999   by Ulrich Nuebel

Recombination in HIV
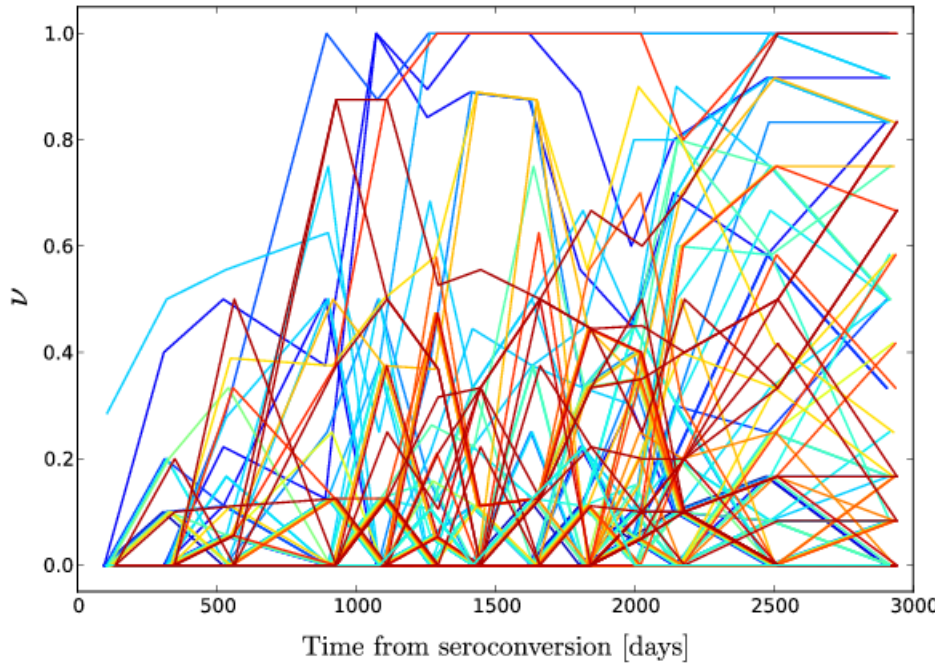

Template switching during reverse transcription: 3-10 times (Levy et al, 2004)

Recombination in HIV
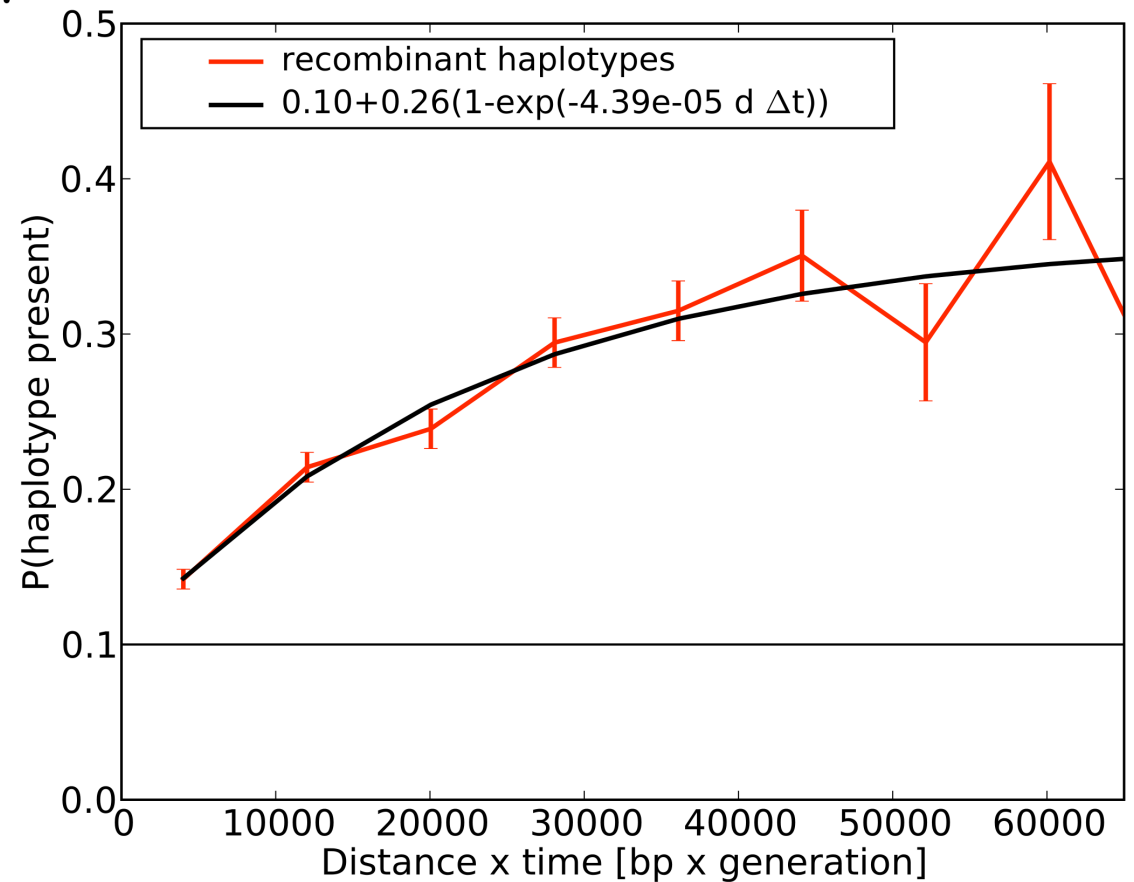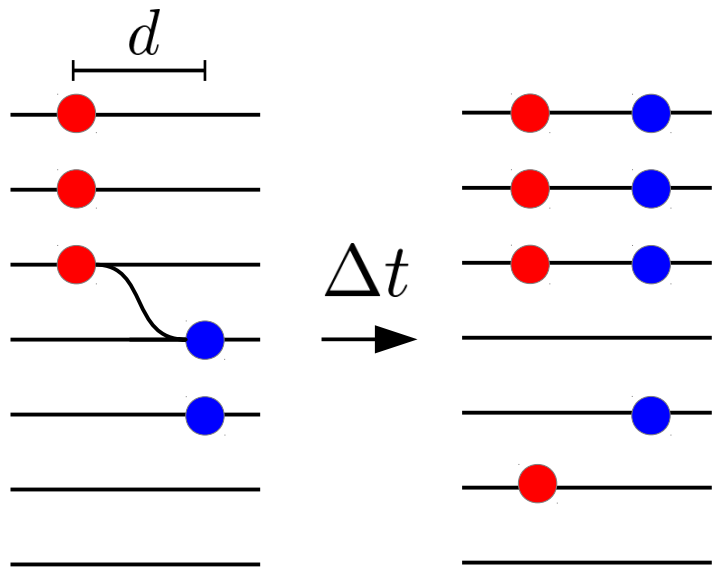
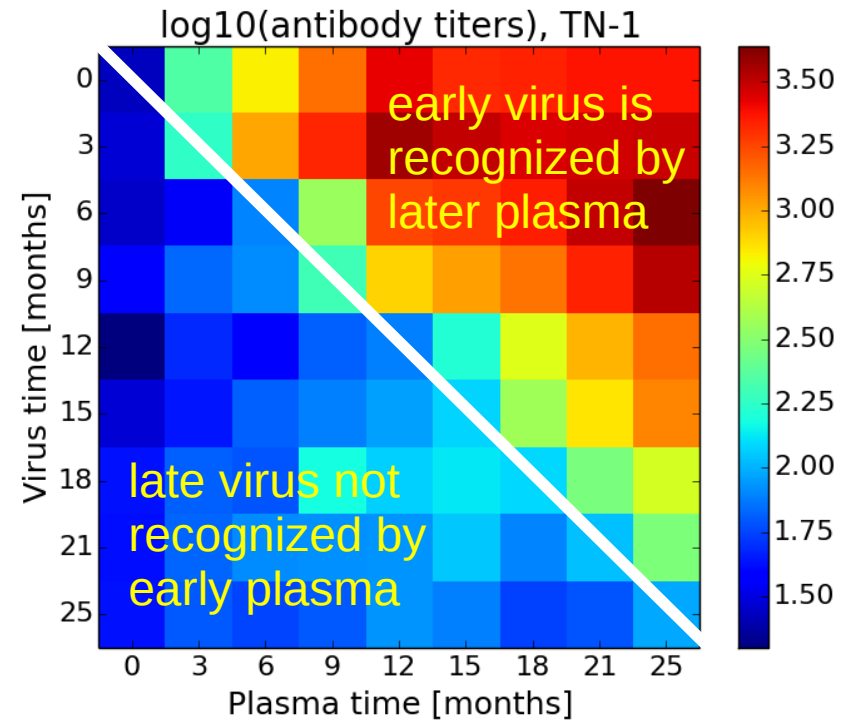$$\rho \approx 1.4 \times 10^{-5}/nuc/generation \qquad \text{implies coinfection <5\%}$$

RN and T Leitner, PLOS Comp Bio 2010   See also Batorsky et al, 2011

Synonymous and nonsynonymous mutations

Shankarrappa et al, 1999

log10(antibody titers), TN-1

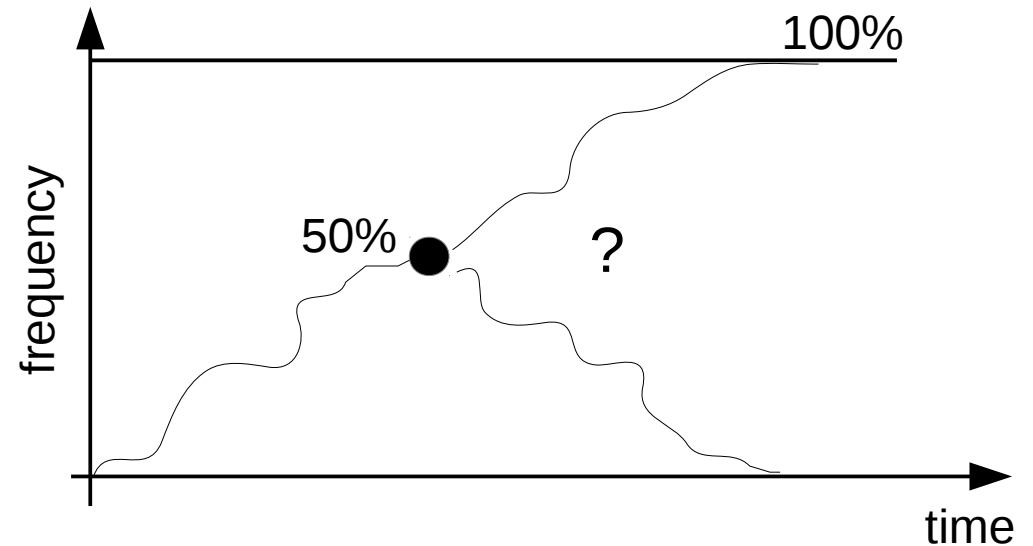early virus is recognized by later plasma
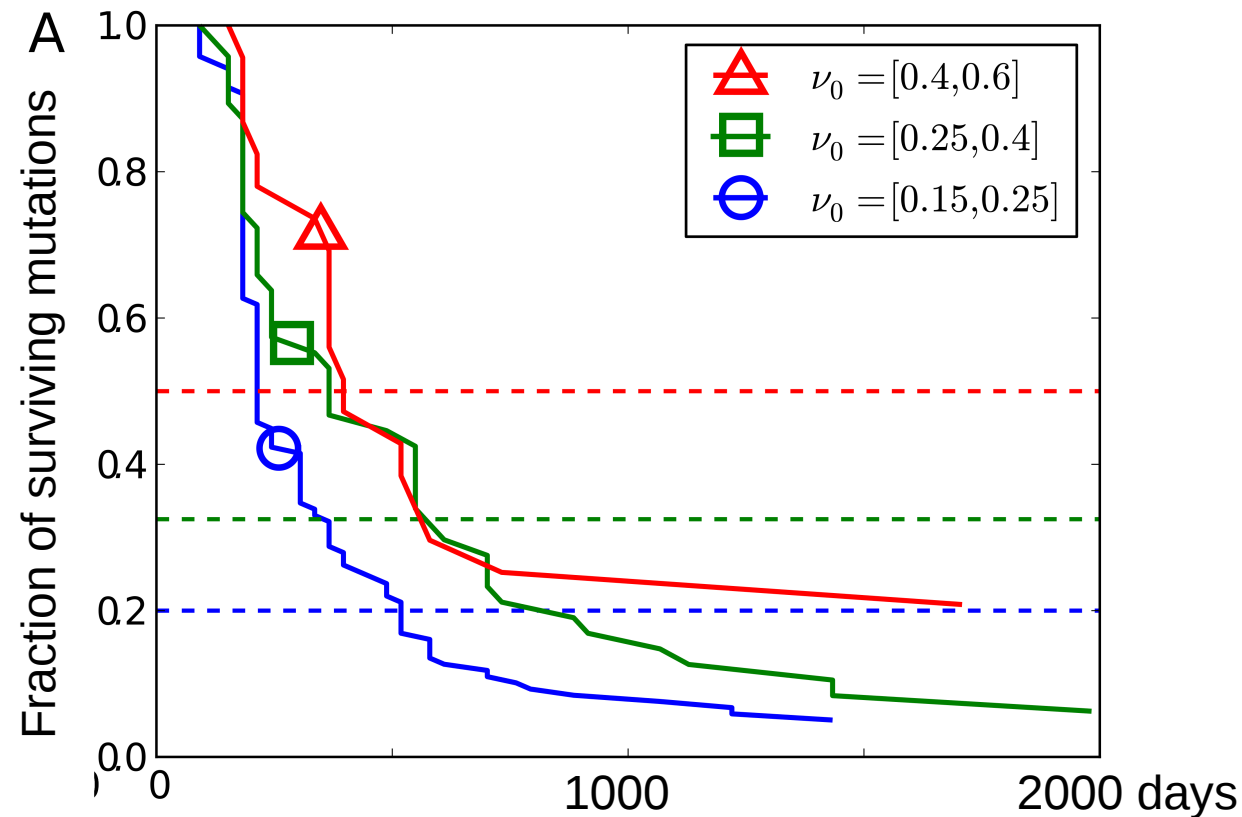
late virus not recognized by early plasma

Richman et al, 2003

- neutral mutations spread by chance:

  fixation probability == frequency

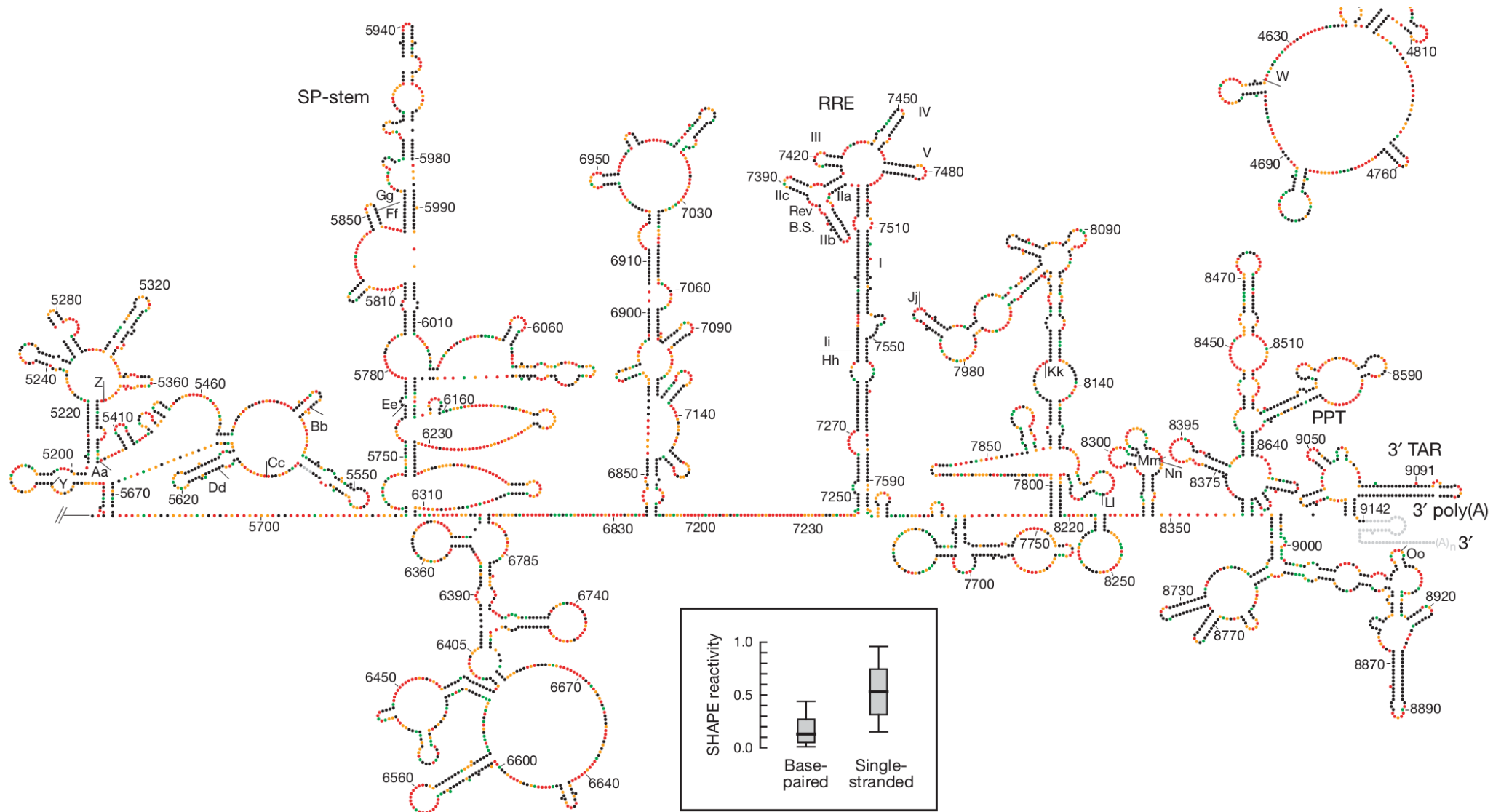- deleterious mutations disappear

- beneficial mutations spread

100%

50% ● ?

frequency

time

Fabio Zanini, RN, J Virology, 2013

Fixation probability

A

Legend:
- $\nu_0 = [0.4, 0.6]$
- $\nu_0 = [0.25, 0.4]$
- $\nu_0 = [0.15, 0.25]$

y-axis: Fraction of surviving mutations

x-axis: days (0, 1000, 2000)

- Synonymous mutations are systematically selected against
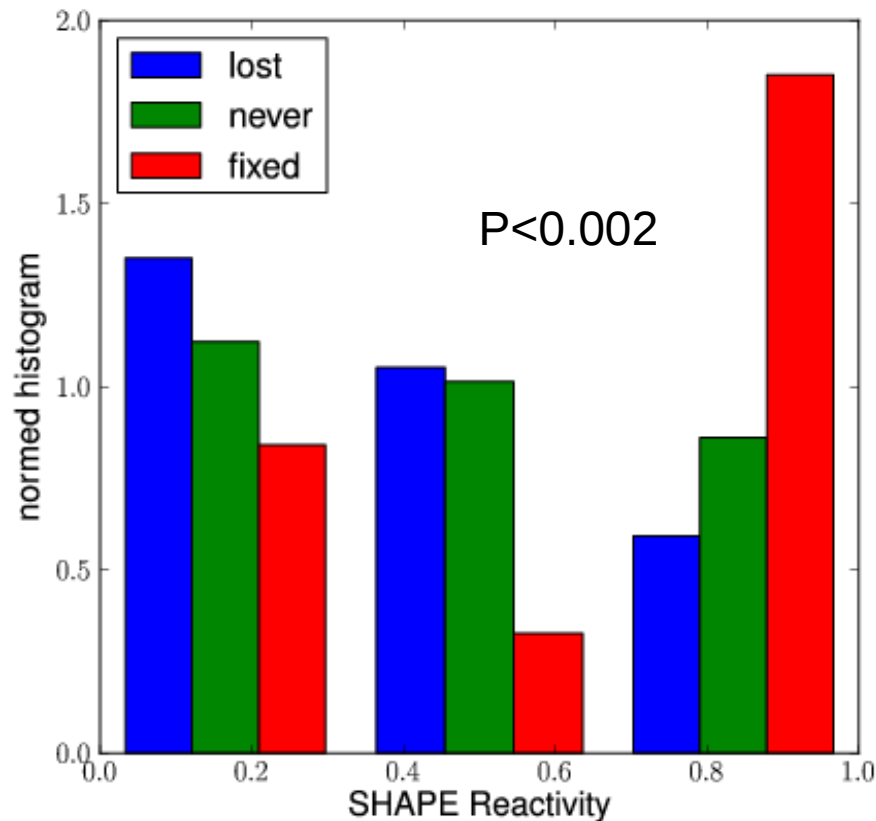- RNA secondary structure??
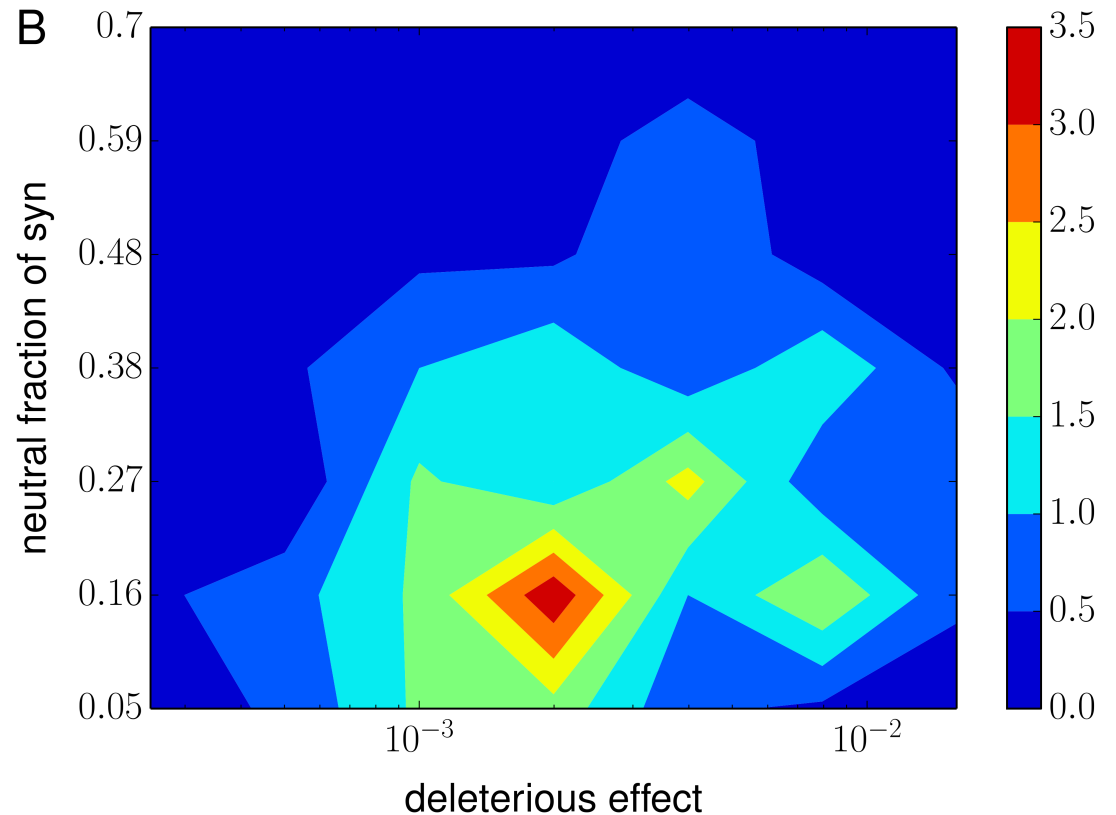
RNA structure

Watts et al, Nature, 2009
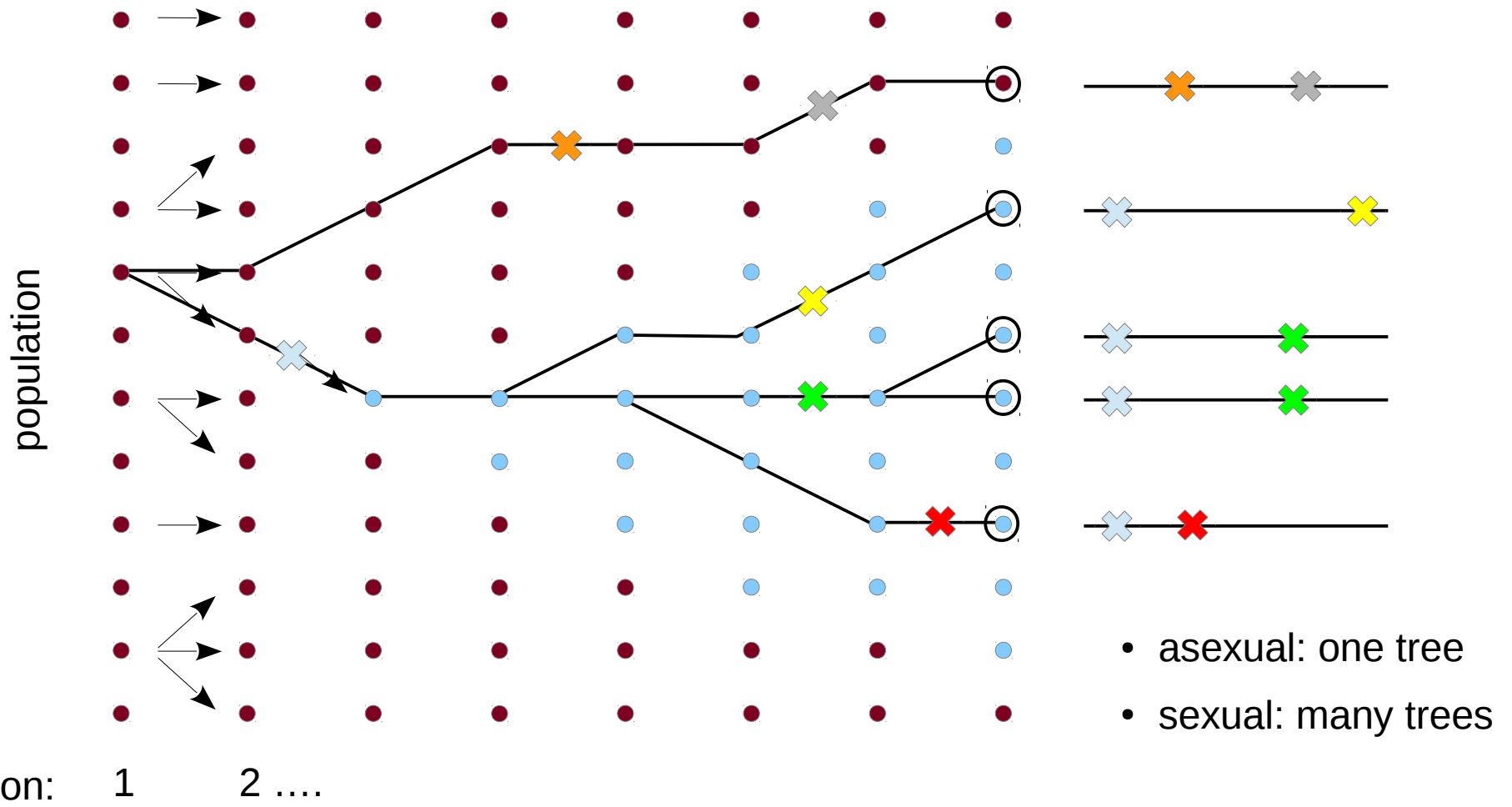
Synonymous mutations break RNA structures

Mutations that fix are not part of RNA structures.

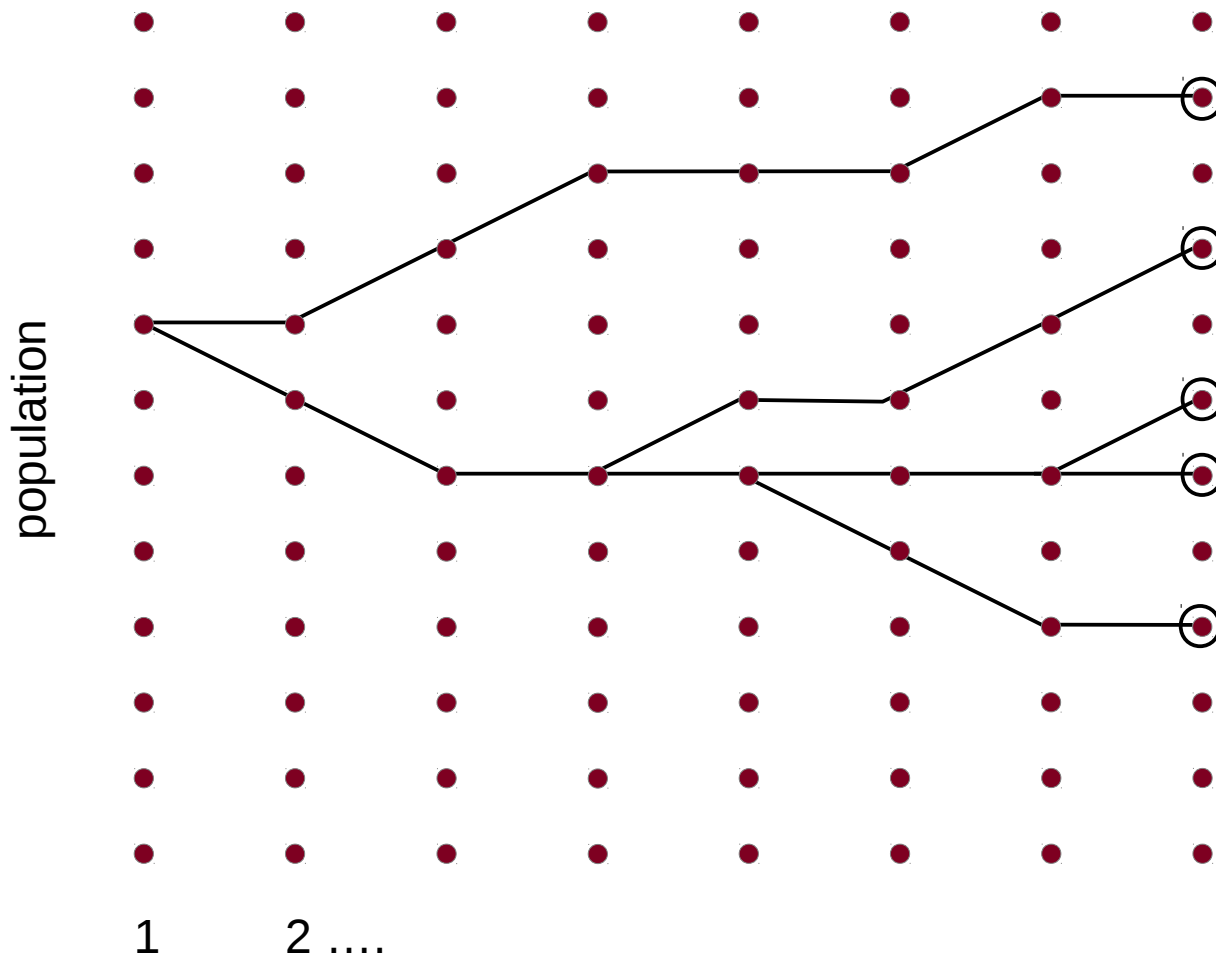Simulations: most synonymous sites are weakly deleterious

Population genetic models

- asexual: one tree
- sexual: many trees

generation: 1   2 ....

evolutionary processes ↔ statistics of trees ↔ patterns of genetic diversity

Neutral coalescence
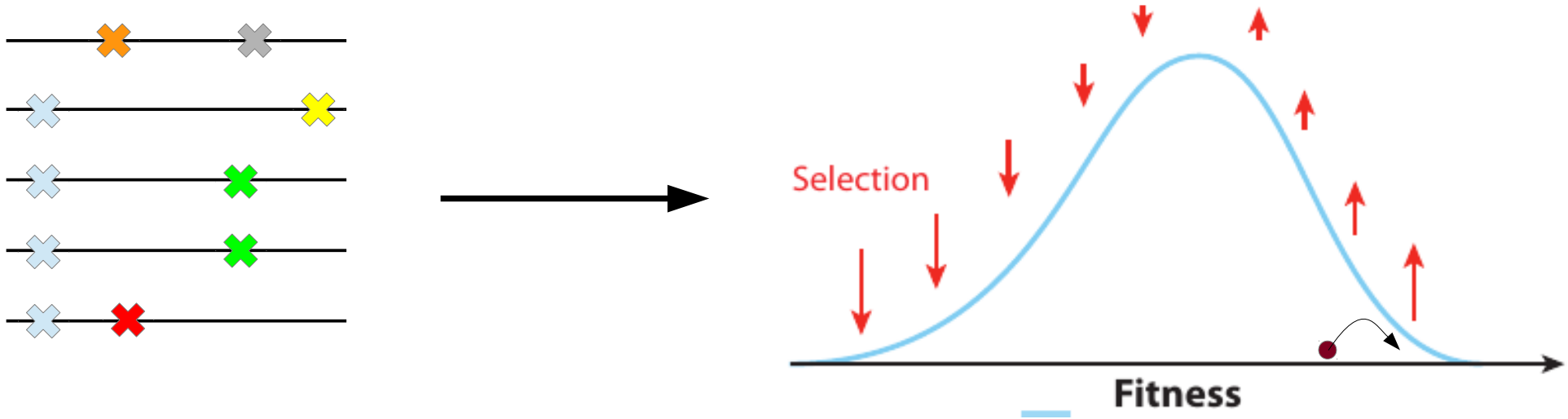
Any pair of lineages merges independently → simply predictions for diversity

$$\text{merging} \sim \frac{1}{N}\binom{k}{2}$$

$$T_{MRCA} \sim N$$

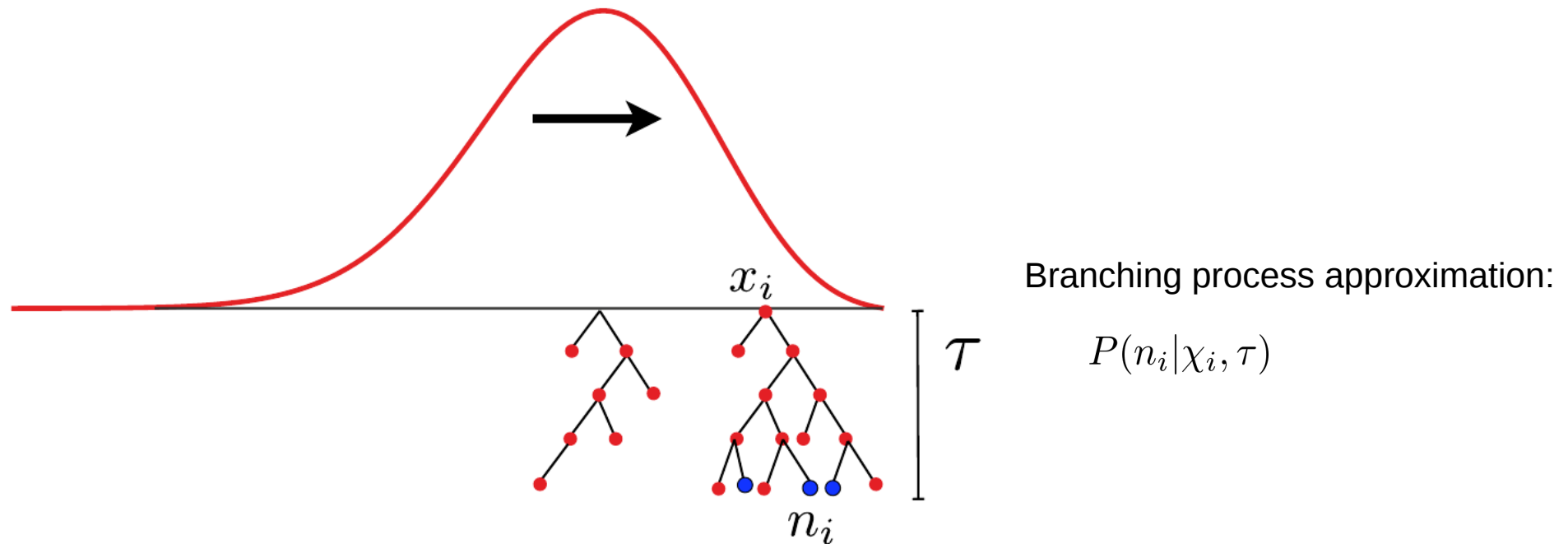# Adapting populations

Deterministic approximation:

$$\frac{\partial}{\partial t} n(x,t) = D \frac{\partial^2}{\partial x^2} n(x,t) + (x - \langle x \rangle) n(x,t)$$

Stochastic at the high fitness tail:

- Similar to stochastic FKPP, see work by Kessler, Levine et al

- But: No deterministic limit. Velocity grows without bounds as *N* grows

$$v \approx 24^{\frac{1}{3}} D^{\frac{2}{3}} (\log N D^{\frac{1}{3}})^{\frac{1}{3}} \qquad\qquad x_c \approx \frac{\sigma^4}{4D} = 3^{\frac{2}{3}} D^{\frac{1}{3}} (\log N D^{\frac{2}{3}})^{\frac{1}{3}}$$

Tsimring et al, Cohen et al, Rouzine et al, Desai et al.... Review: RN, Annual Reviews Evo&Eco, 2013

Coalescence in adapting populations

Branching process approximation:

$$P(n_i | \chi_i, \tau)$$
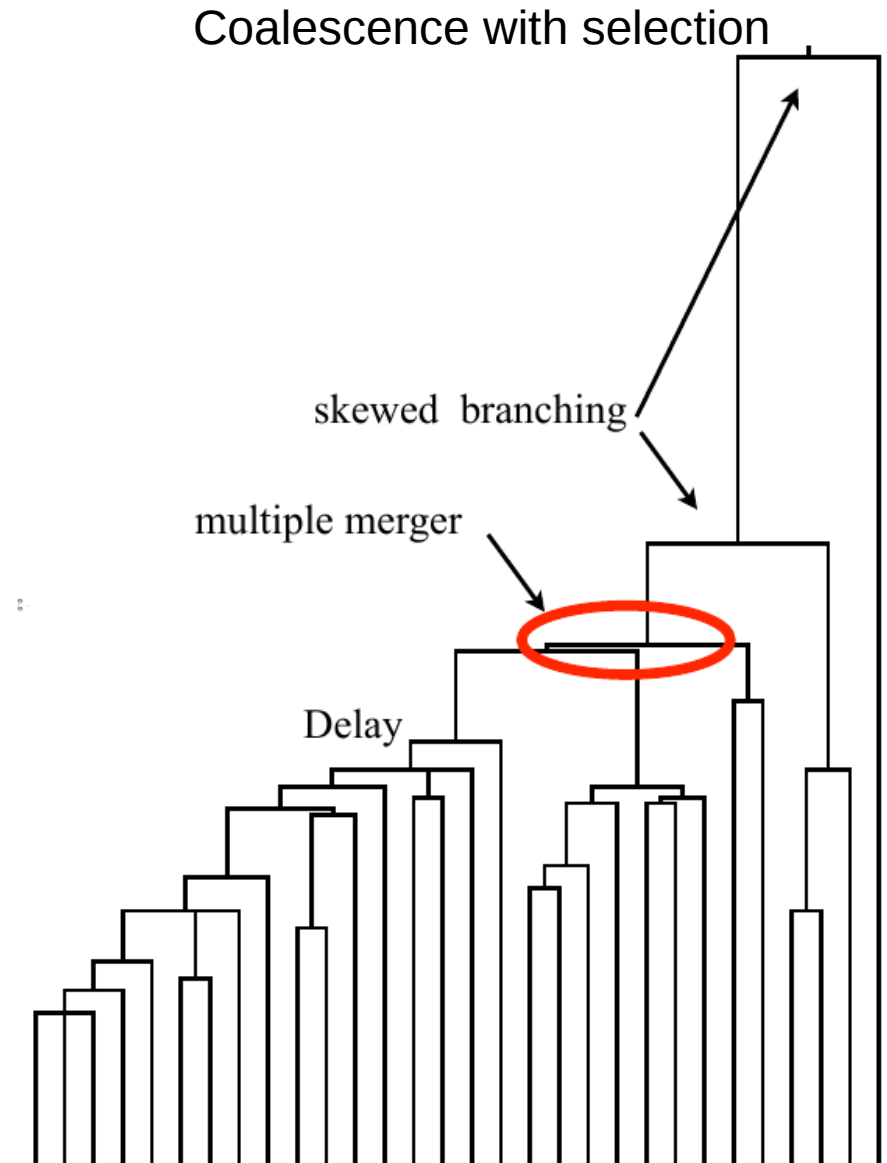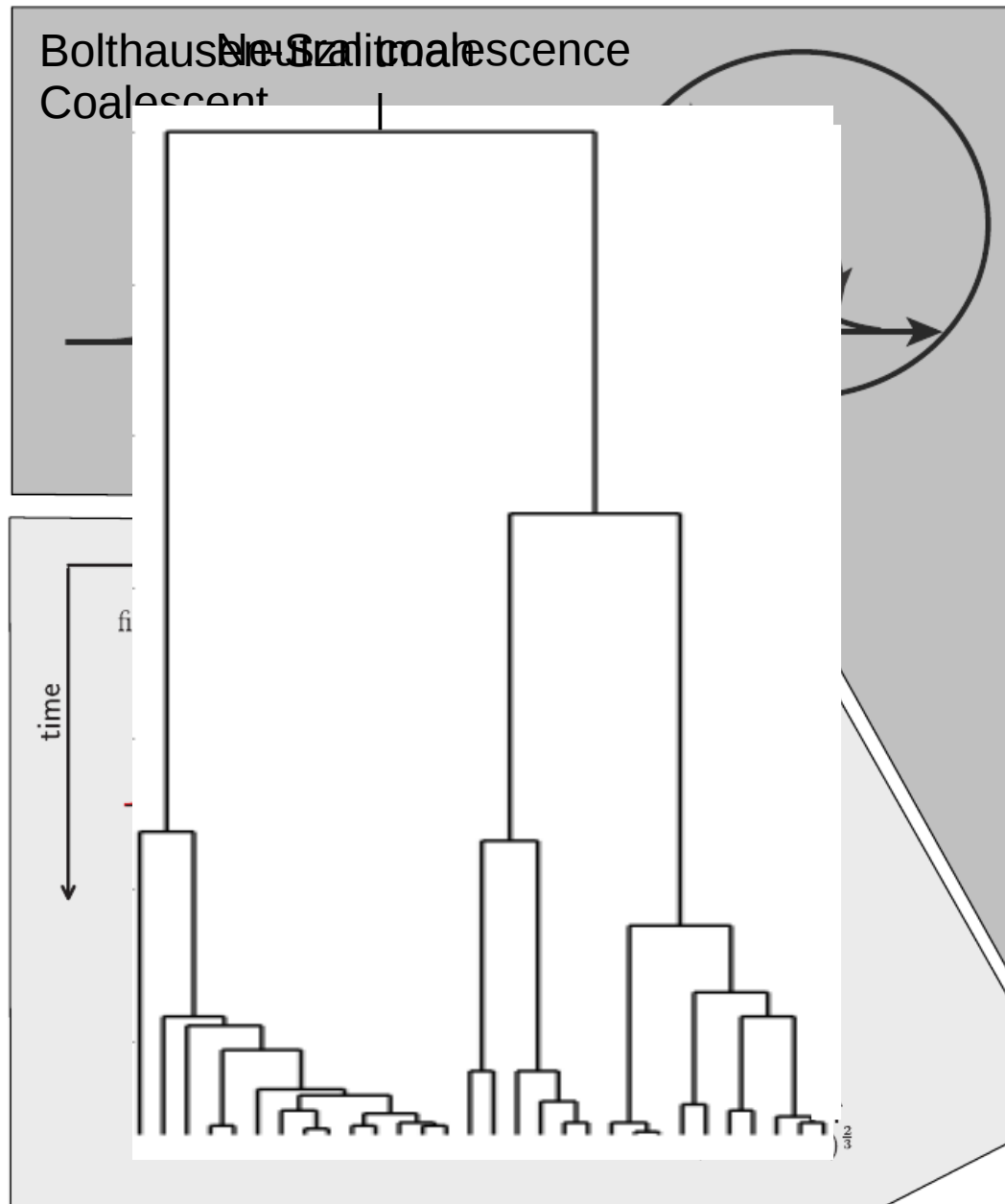
Does a sample (blue dots) have a common ancestor τ generations ago?

$$Q_b = \langle \sum_i \left(\frac{n_i}{\sum_j n_j} \right)^b \rangle = \begin{cases} \mathcal{O}(1/N) & \tau < T_c \\ \frac{\tau - T_c}{T_c(b-1)} & \tau > T_c \end{cases}$$
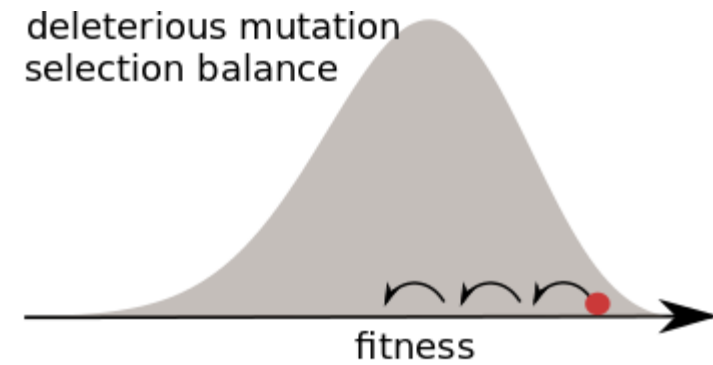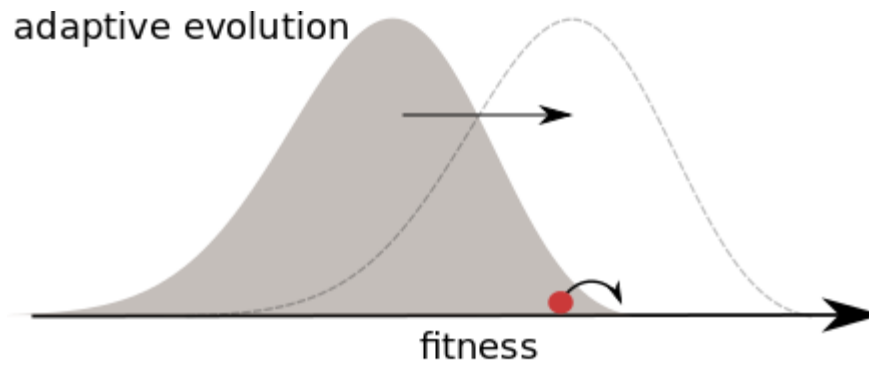
All other merger rates also suggest a Bolthausen-Sznitman coalescent

Brunet et al, PRE, 2007, RN and Hallatschek, 2013

Coalescence in adapting populations

Bolthausen-Sznitman
Coalescent

Neutral coalescence

Coalescence with selection

time

skewed branching

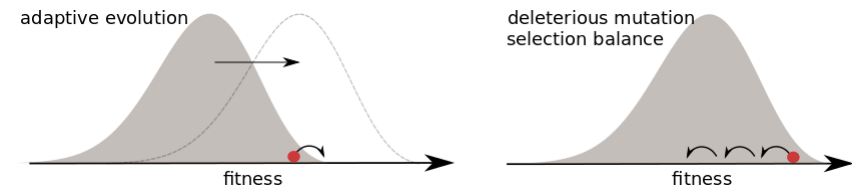multiple merger

Delay

RN and Hallatschek, PNAS, 2013

Coalescence by selection

- universal: many selected mutations → same tree statistics

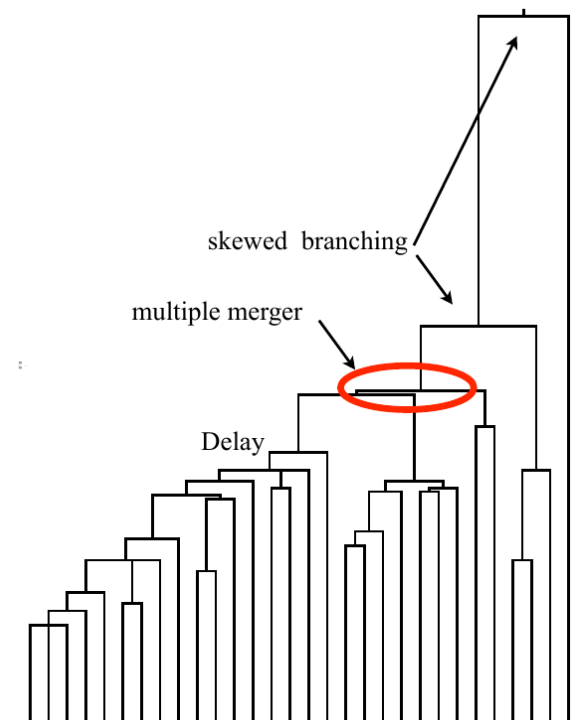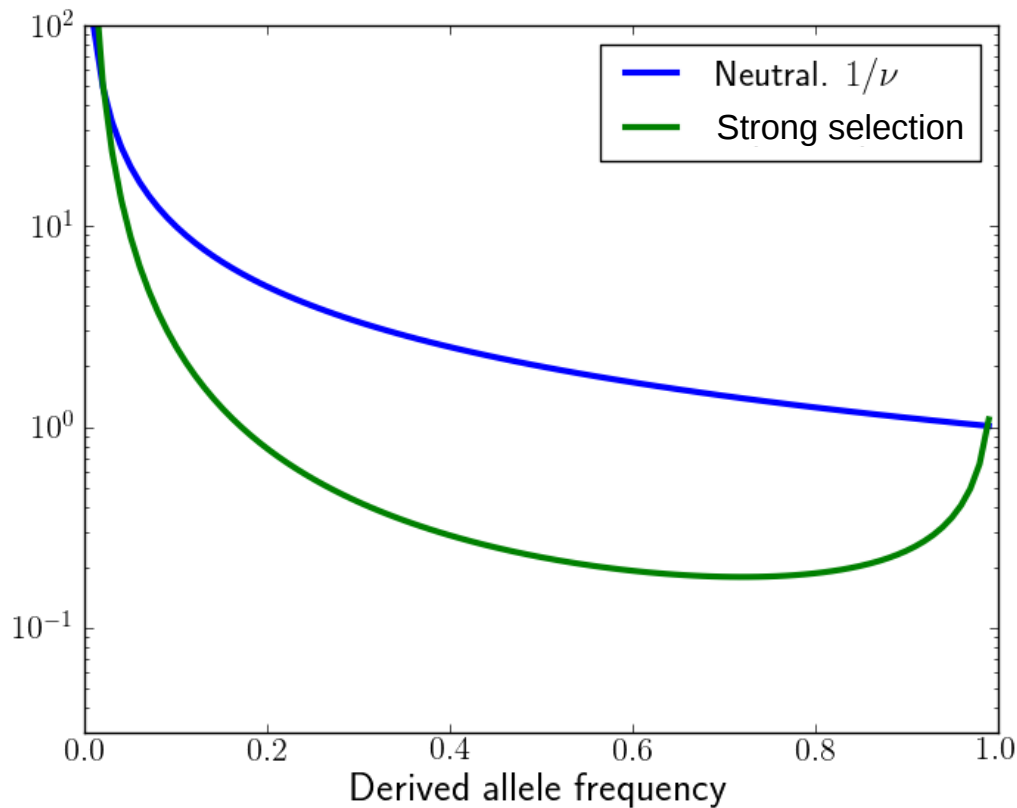# Coalescence by selection

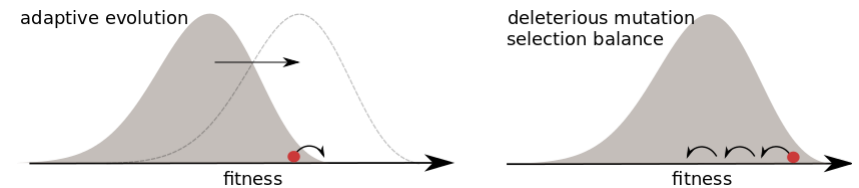- universal: many selected mutations → same tree statistics

- explicit predictions for observable quantities (allele frequency spectra)

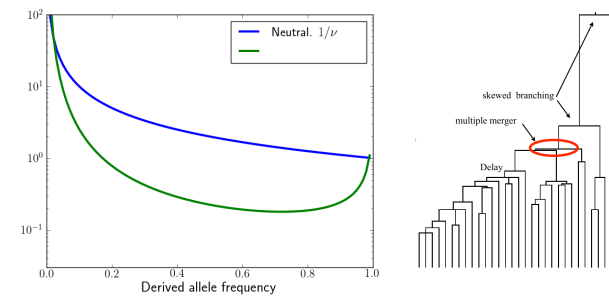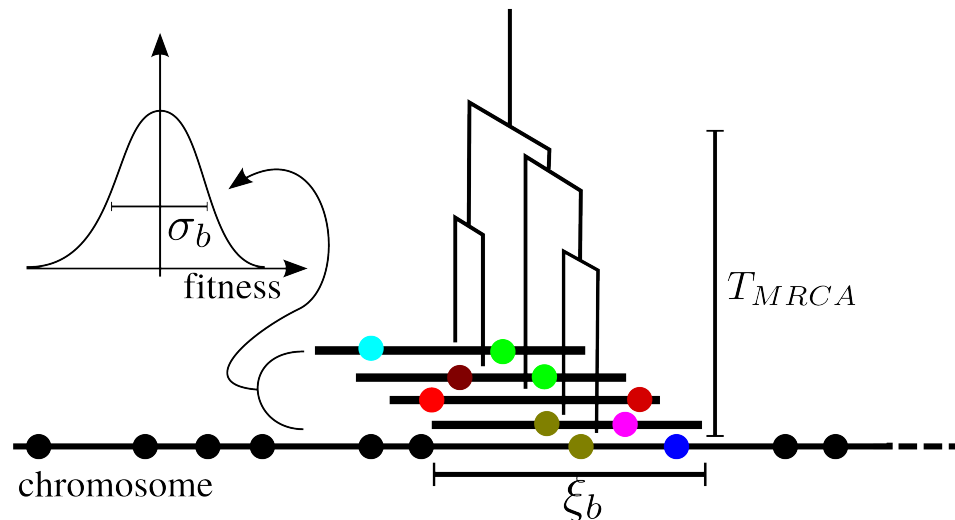# Coalescence by selection

- universal: many selected mutations → same tree statistics

- explicit predictions for observable quantities (allele frequency spectra)
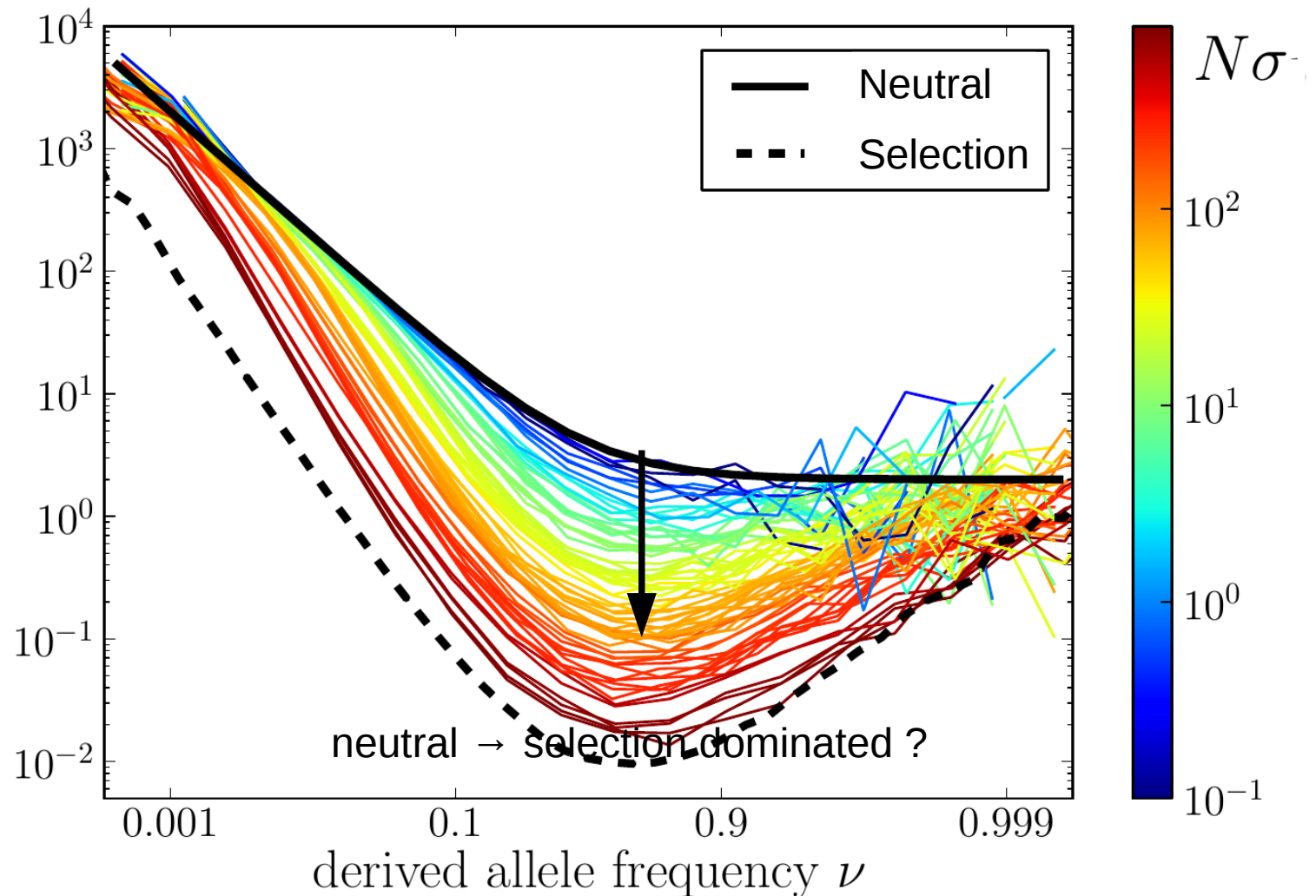

- extension to sexual species: recombination and linkage blocks

RN, Kessinger, Shraiman, PNAS, 2013

From neutral to strong selection

- neutral: *1/N*

- selection: *σ*

- Parameter: *Nσ*

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

Summary

- **RNA viruses sample every single mutation, many double mutations**

- **Coevolution with the immune system**

  - CTL escape: few sites – largely idiosyncratic

  - antibody escape: many mutations – statistical patterns

- **Theory of adaptation**

  - tractable coalescence model → verifiable predictions for diversity

- **Predict influenza evolution → next week**

Acknowledgements

erc