



Intra-patient evolution of HIV

Richard Neher

Max-Planck-Institute for Developmental Biology

Longitudinal data, envelope gene

- Recombination rate of HIV
- Selection strength in *env*

Drug resistance evolution

- Ultra-deep sequencing data
- Method to estimate selection strength

HIV evolution



MAX-PLANCK-GESELLSCHAFT

Population size:	$N = 10^{10 \pm ?}$
Genome size:	$L = 10^4 bp$
Generation time:	$T = 2d$
Mutation rate:	$\mu = 3 \times 10^{-5} / bp/gen$
Recombination rate??	Selection strength??

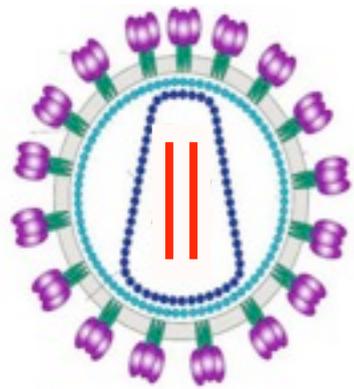
Outline:

1. Estimation of recombination rate in the envelope gene
2. Estimation of the strength of selection in *env*
3. Dynamics of drug resistance evolution, selection in the *pol* gene

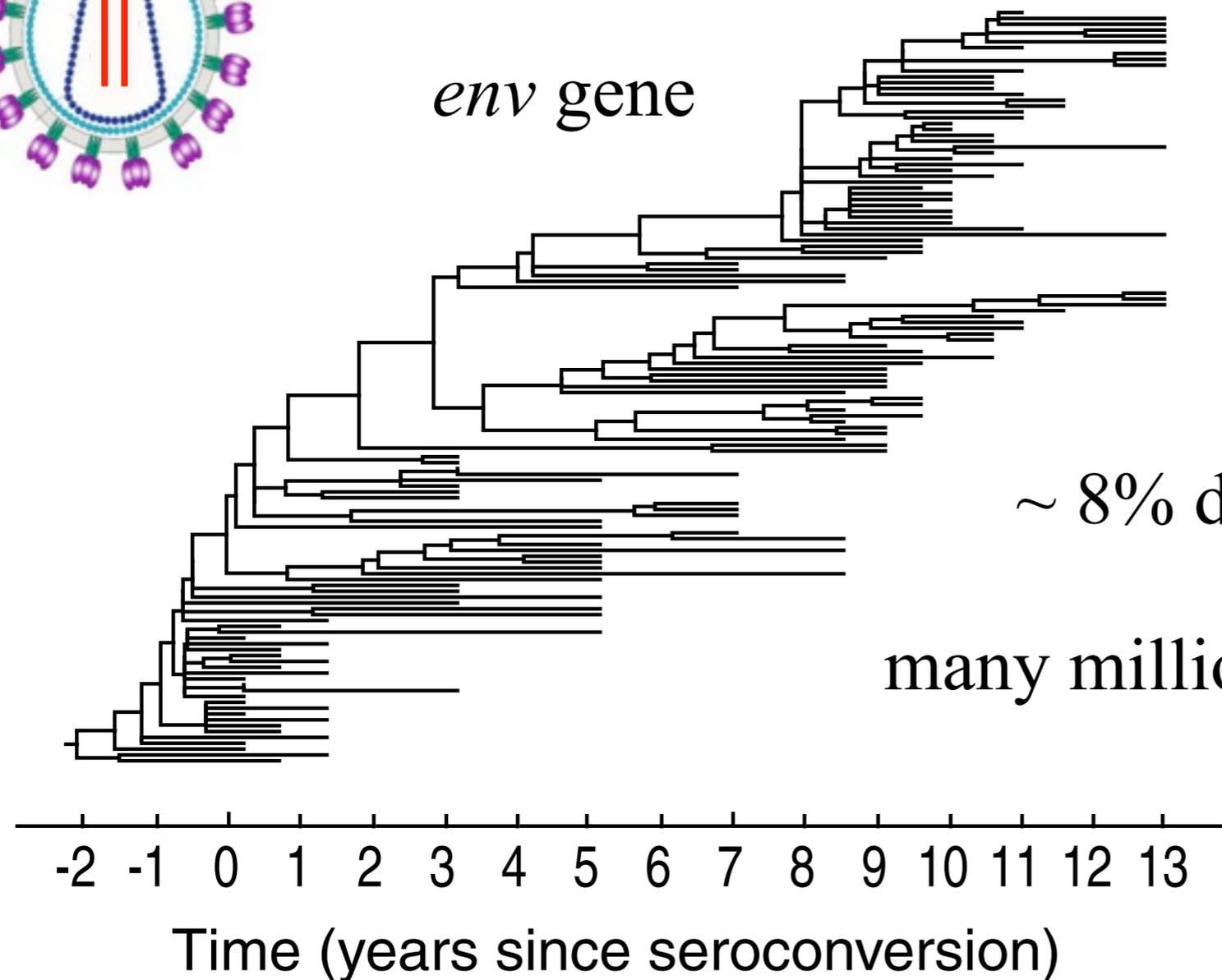
Intra-patient evolution



MAX-PLANCK-GESELLSCHAFT



env gene



~ 8% divergence in 10 years

==

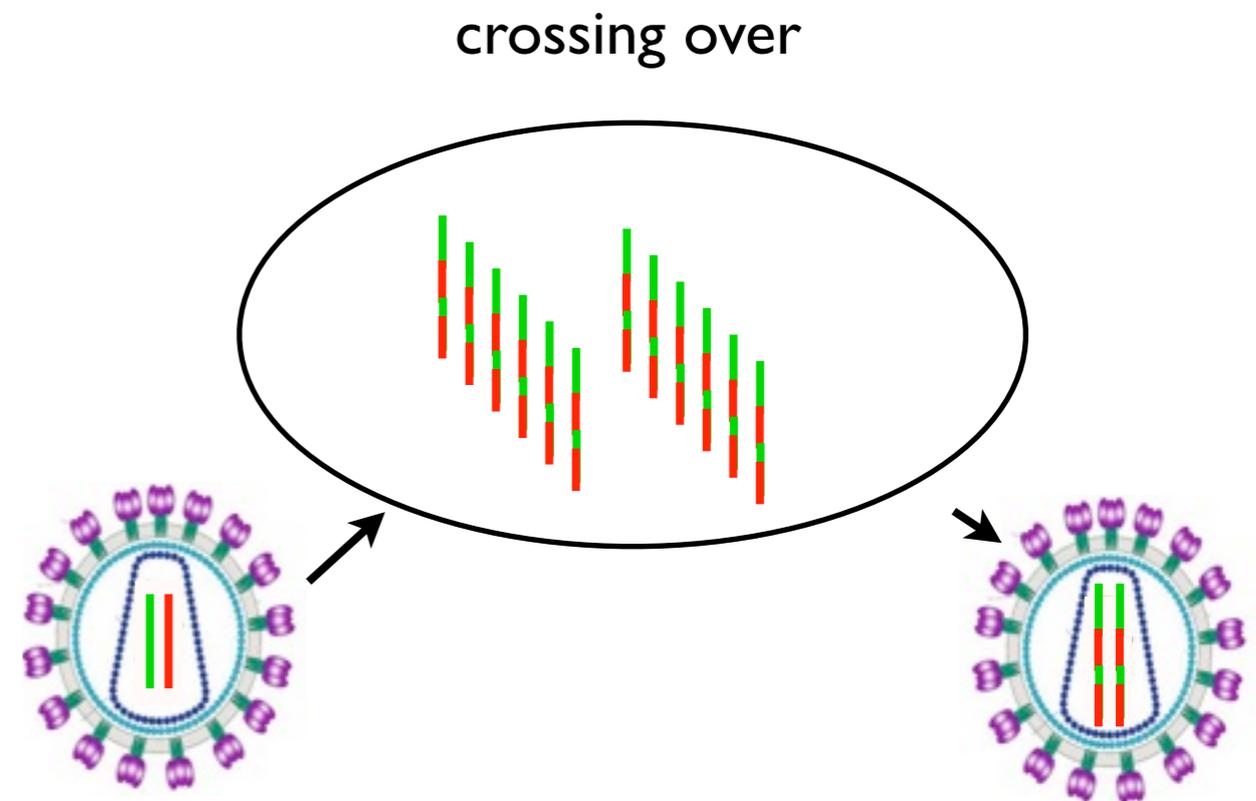
many millions of years in *Drosophila*

Lemey et al., 2006,
Shankarappa et al. 1999

Recombination in HIV



MAX-PLANCK-GESELLSCHAFT

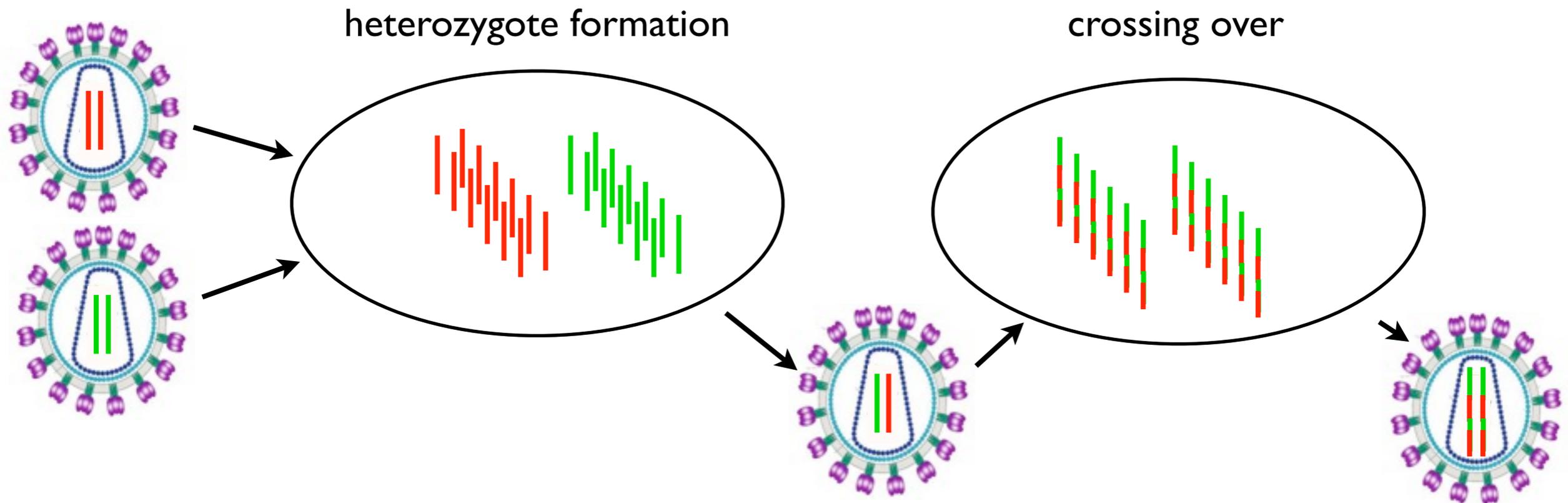


Template switching rate: 3-10 events per replication and genome (Levy et al, 2004)

Recombination in HIV



MAX-PLANCK-GESELLSCHAFT

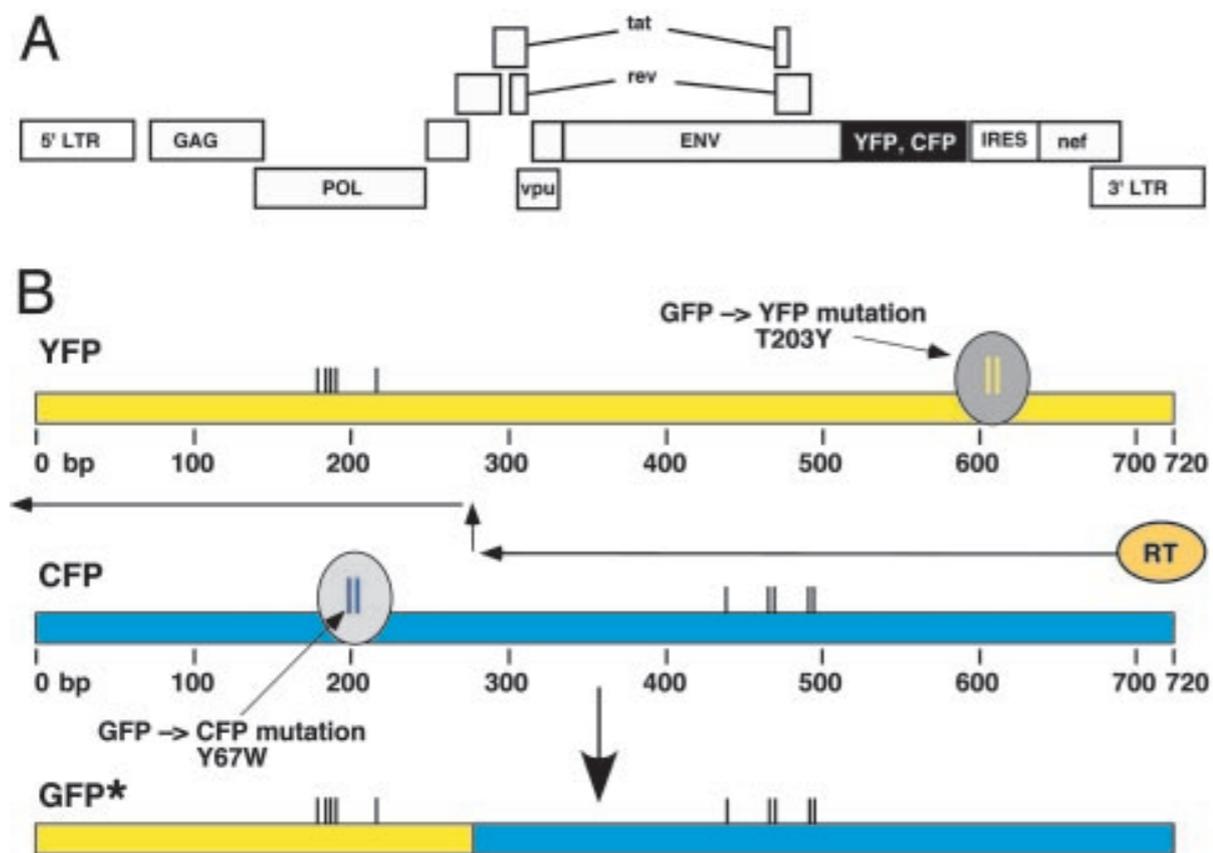


Template switching rate: 3-10 events per replication and genome (Levy et al, 2004)

Template switching – Levy et al.



MAX-PLANCK-GESELLSCHAFT

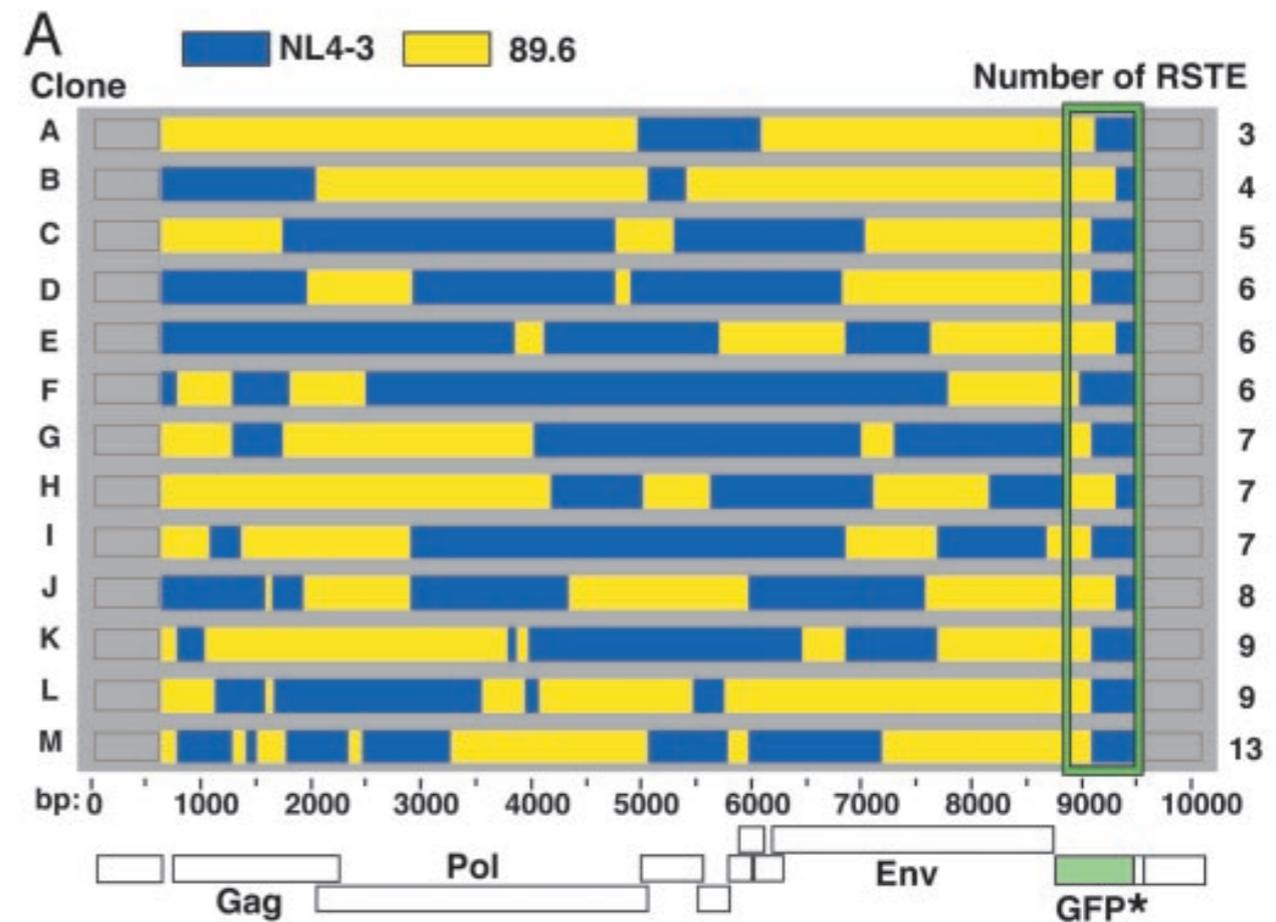
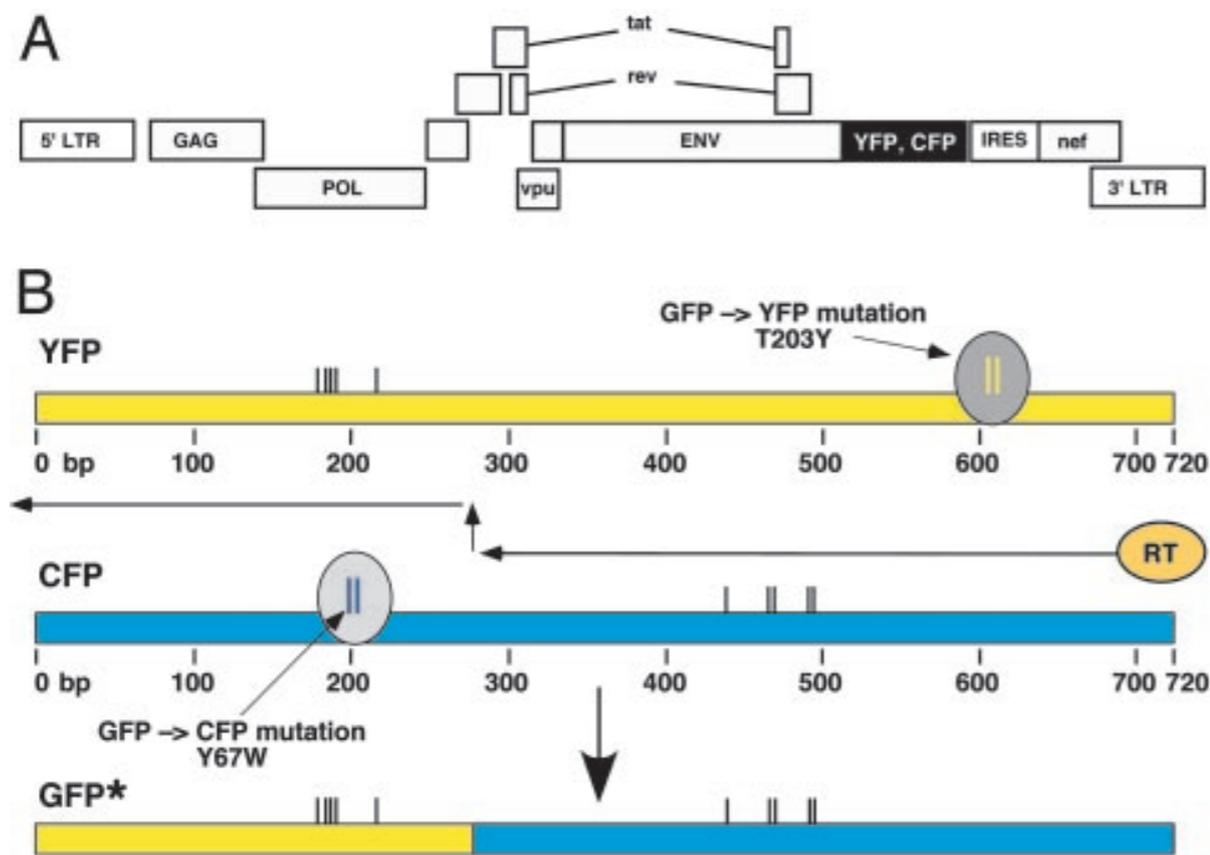


Dynamics of HIV-1 recombination in its natural target cells, Levy et al, PNAS, 2004

Template switching – Levy et al.



MAX-PLANCK-GESELLSCHAFT



Dynamics of HIV-1 recombination in its natural target cells, Levy et al, PNAS, 2004

Recombination in HIV



MAX-PLANCK-GESELLSCHAFT

Relevant effective recombination rate:

Template switching x coinfection frequency

Estimation using population genetic models, Shriner et al. 2004:

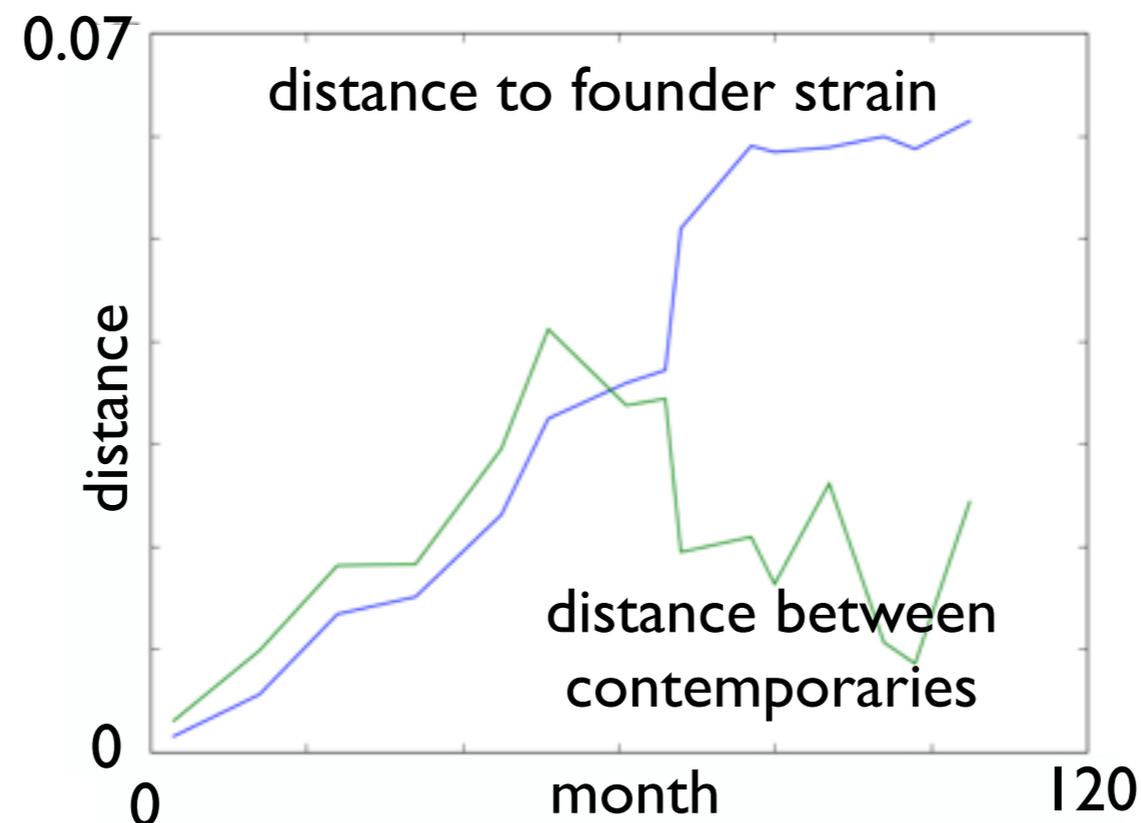
$$\rho = 8 \times 10^{-6} - 1.4 \times 10^{-4} \text{ per site and generation}$$

- implies a very high coinfection frequency
- rests on the assumptions of the neutral coalescent



Data – Shankarappa et al 1999

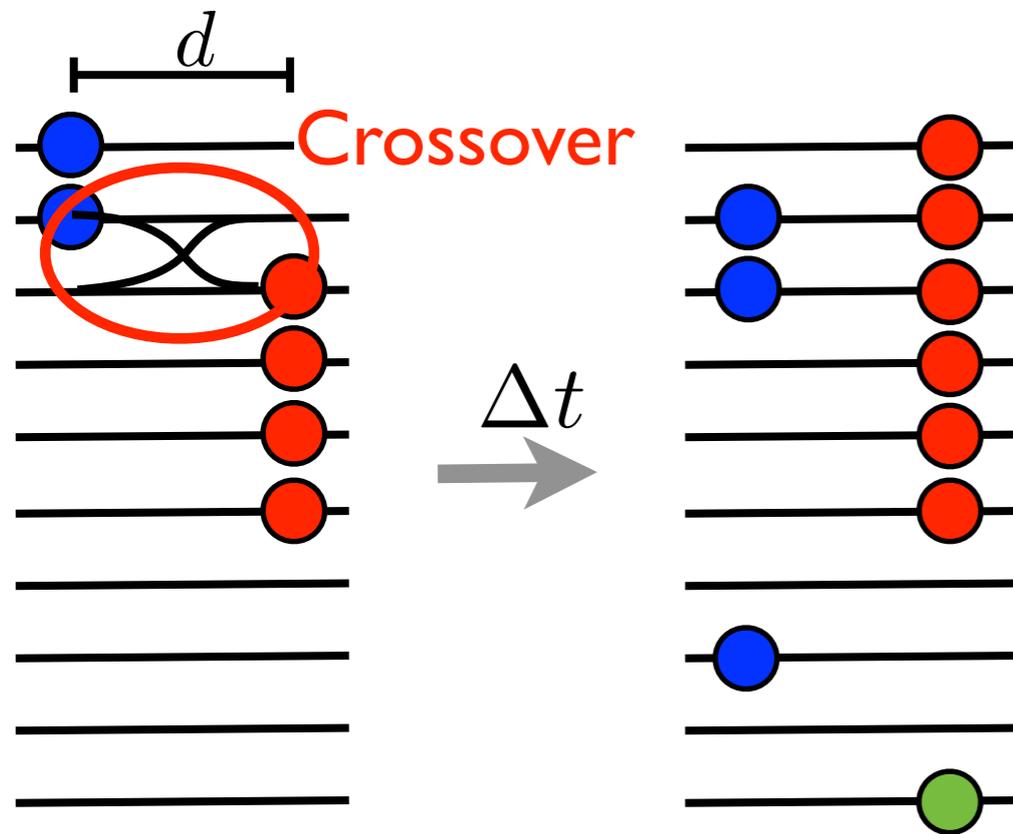
- 11 Patients (part of the MACS cohort)
- 10-20 samples over up to 10 years
- 10-20 partial *env* sequences (700bp, C2V5)
- some received therapy at later time points



Recombination in HIV



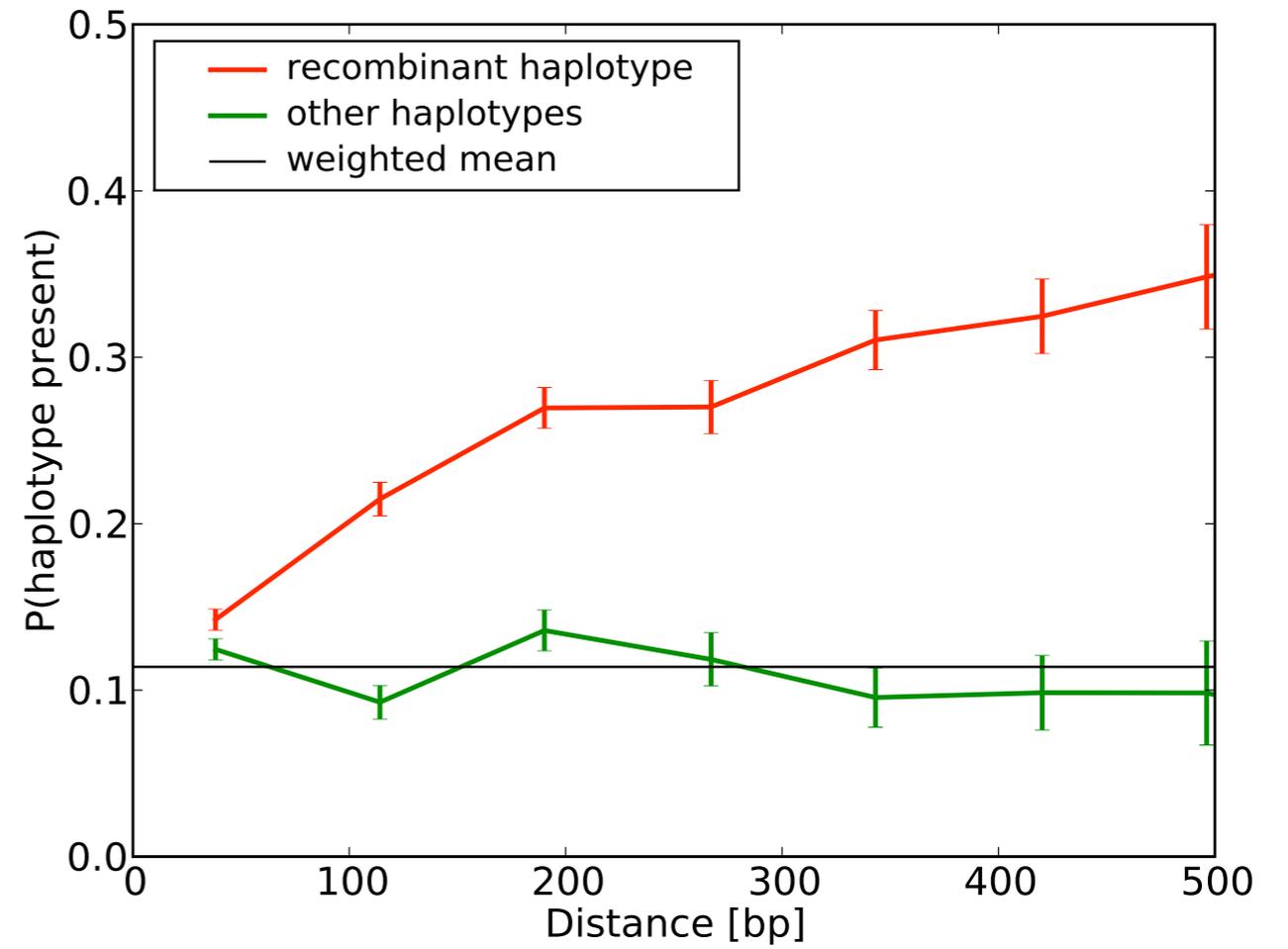
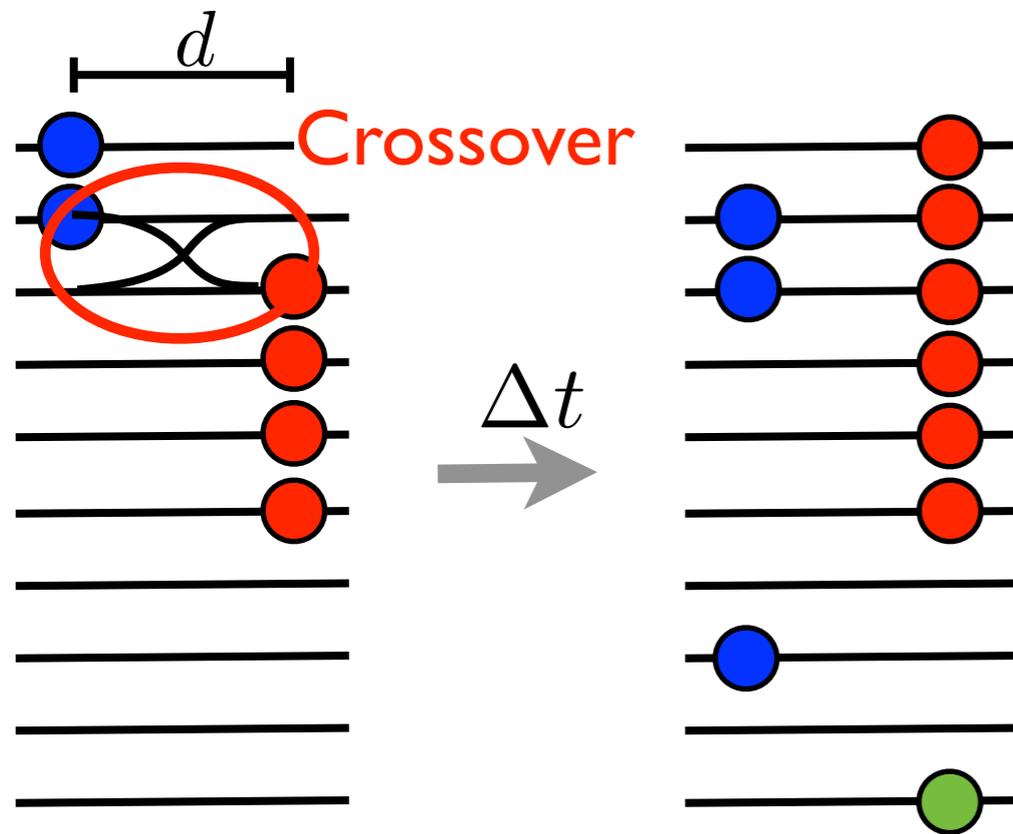
MAX-PLANCK-GESELLSCHAFT



Recombination in HIV



MAX-PLANCK-GESELLSCHAFT

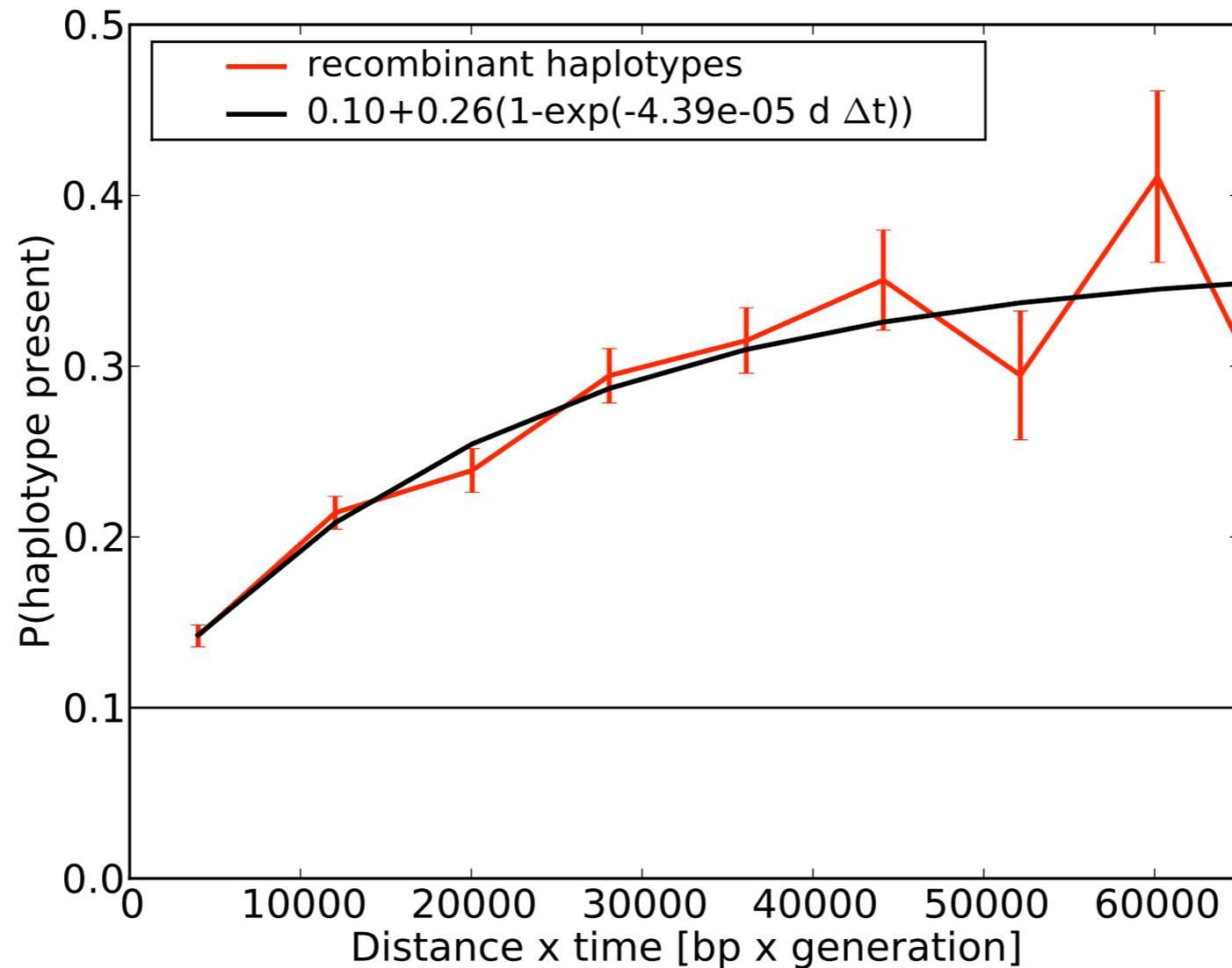


RN and Thomas Leitner, PLoS Comp Bio, 2010

Recombination in HIV



MAX-PLANCK-GESELLSCHAFT

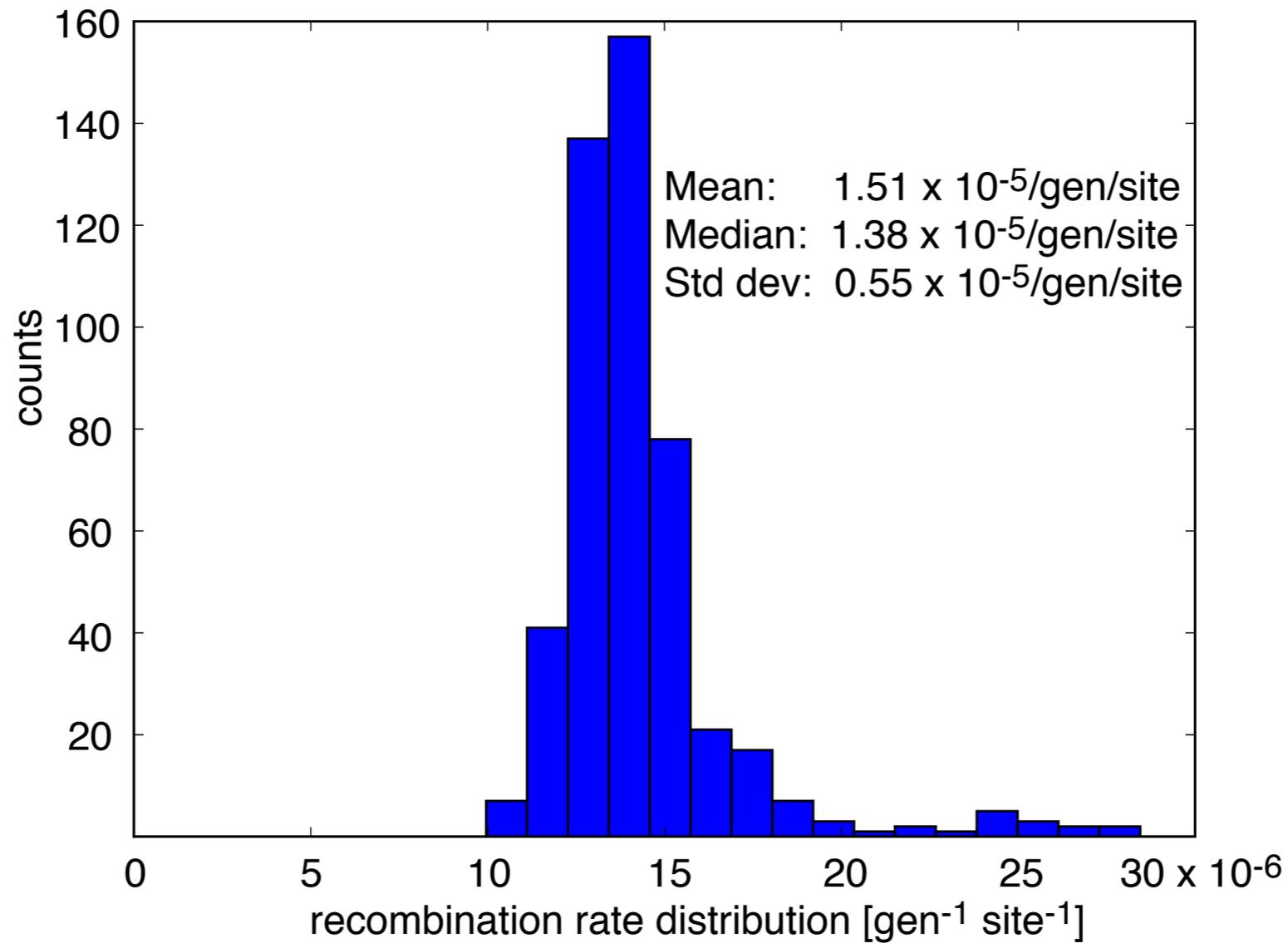


$$p_{AB}(t) = p_{APB} + (p_{AB}(t_0) - p_{APB})e^{-\rho d(t-t_0)} \quad \longrightarrow \quad \rho \approx \frac{\alpha}{\langle Mp_{APB} \rangle}$$

Effective recombination rate



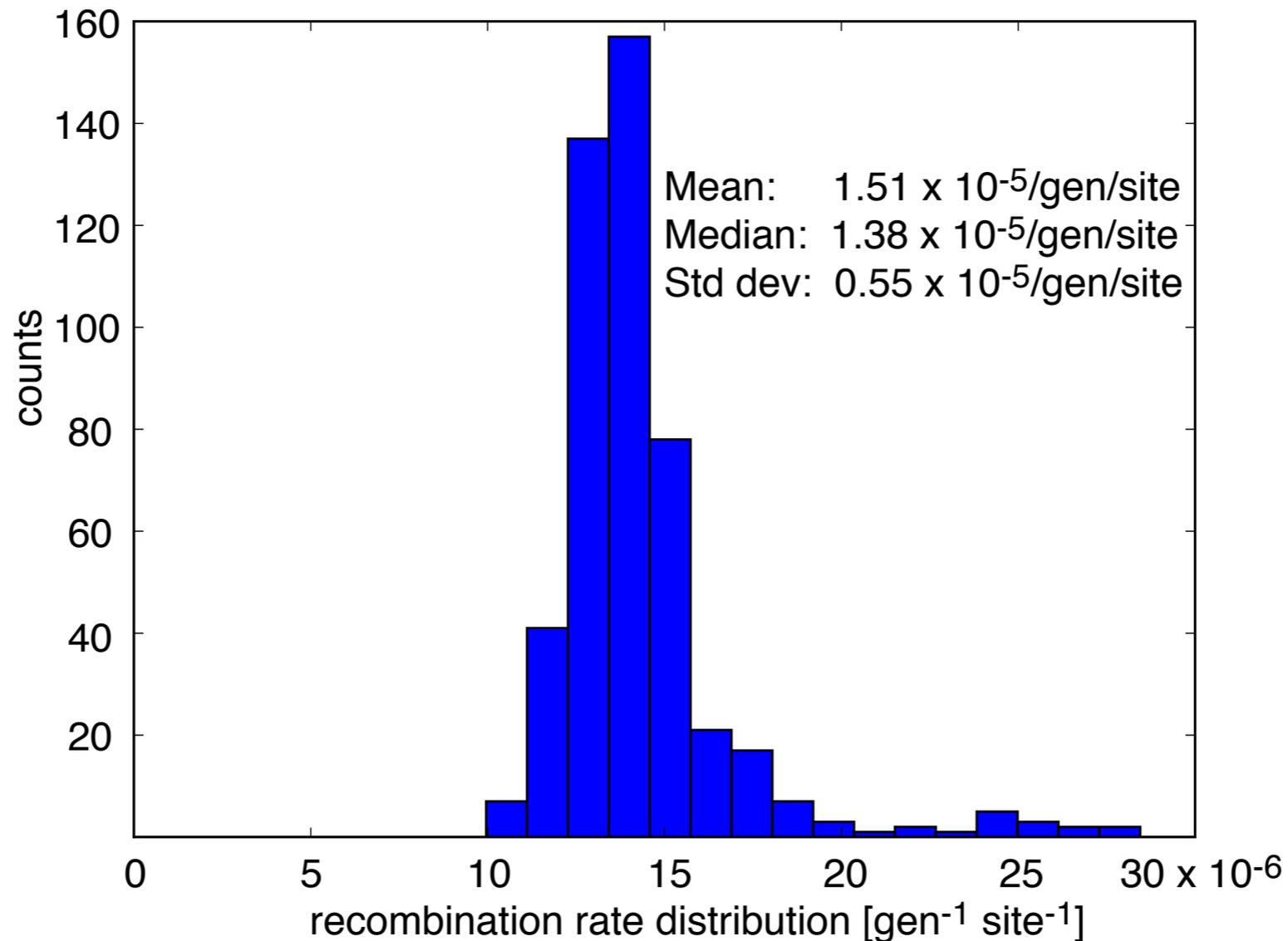
MAX-PLANCK-GESELLSCHAFT



Effective recombination rate



MAX-PLANCK-GESELLSCHAFT



About a factor 50 lower than the template switching rate.

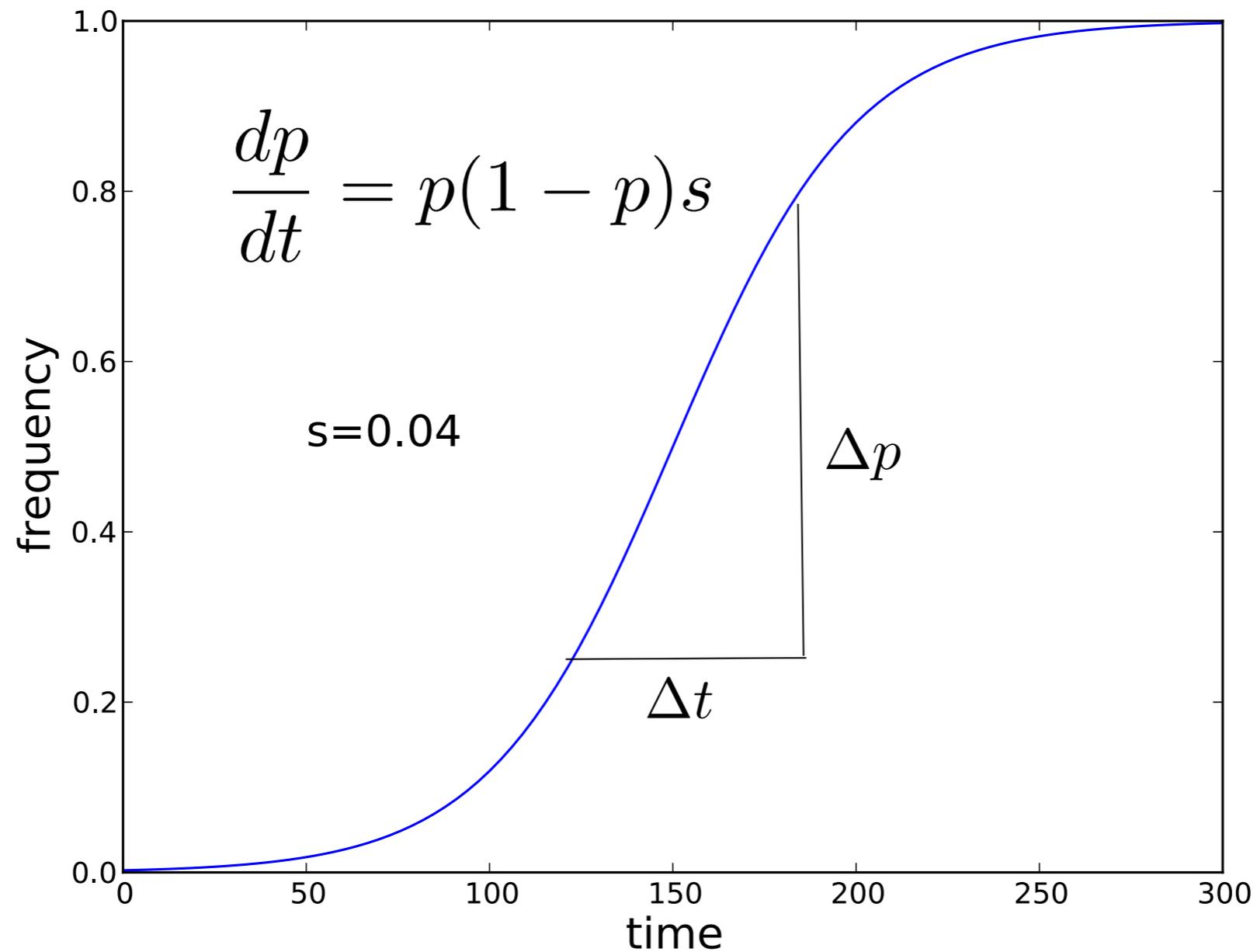


Coinfection $< 5\%$

Selection in HIV evolution



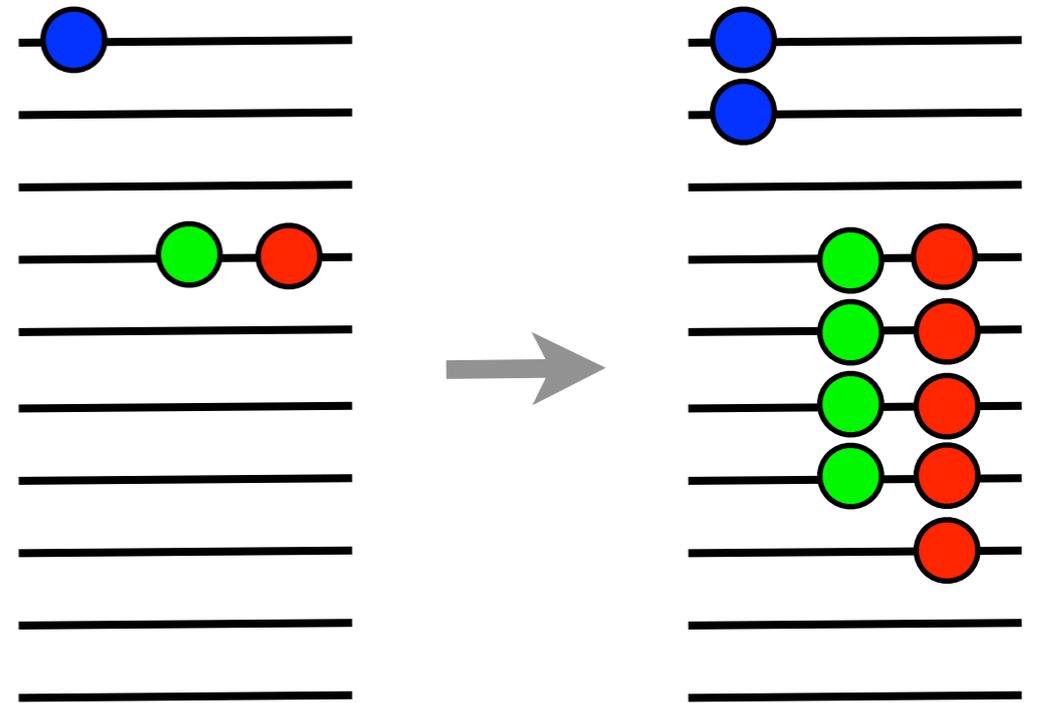
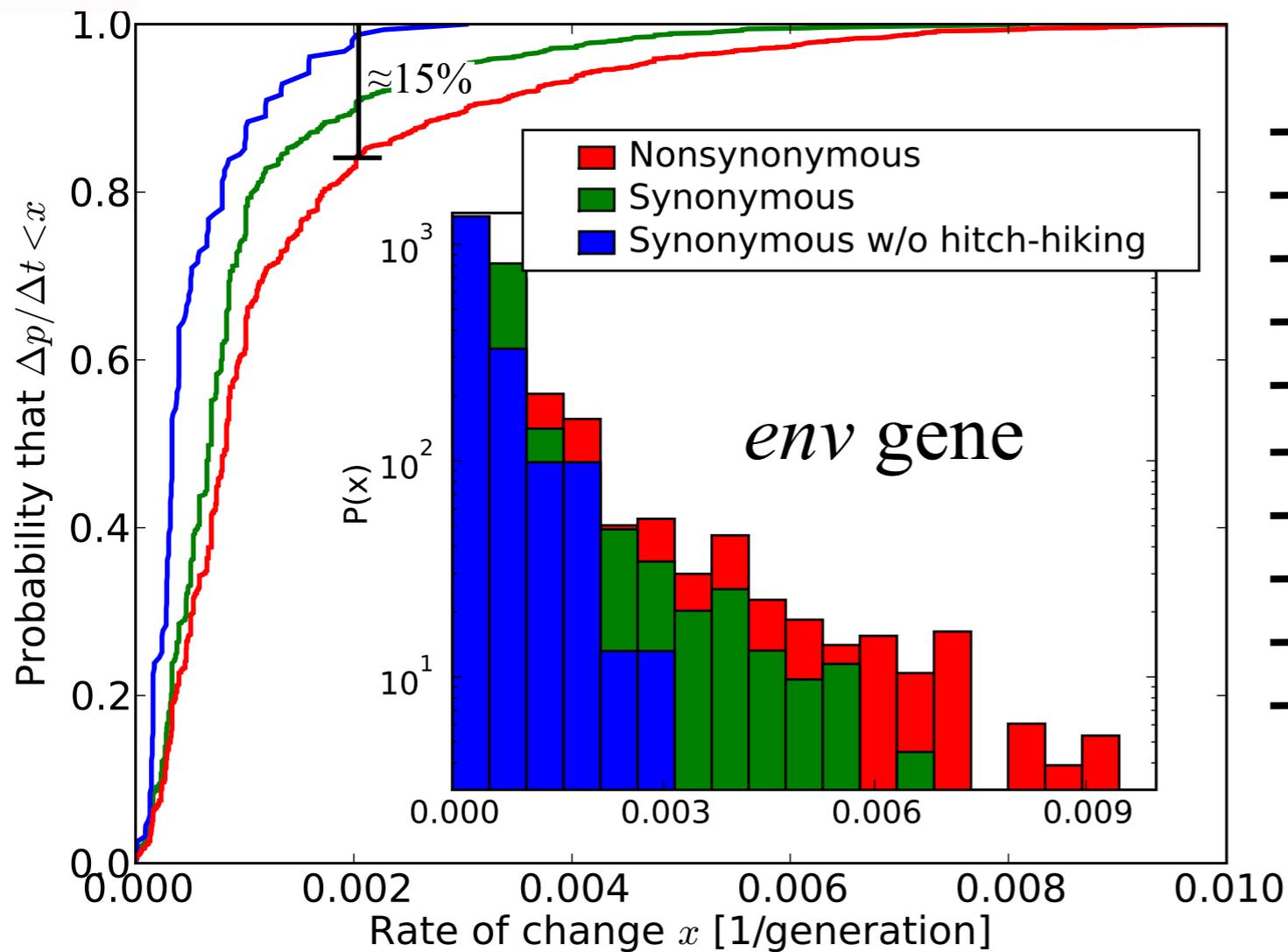
MAX-PLANCK-GESELLSCHAFT



Selection in HIV evolution



MAX-PLANCK-GESELLSCHAFT



Selection for red will affect green

Selection strength: $s = 1-2\%$

Turnover in 50-100 generations

RN and Thomas Leitner, PLoS Comp Bio, 2010

Drug Resistance Evolution



MAX-PLANCK-GESELLSCHAFT

Anti-retroviral drugs:

- Protease inhibitors (PI)
- Nucleoside-Analog Reverse transcriptase inhibitors (NRTI)
- Non-Nucleoside-Analog RT inhibitors (NNRTI)
- Integrase inhibitors

Drug resistance can emerge fast (weeks) or take decades to set in.

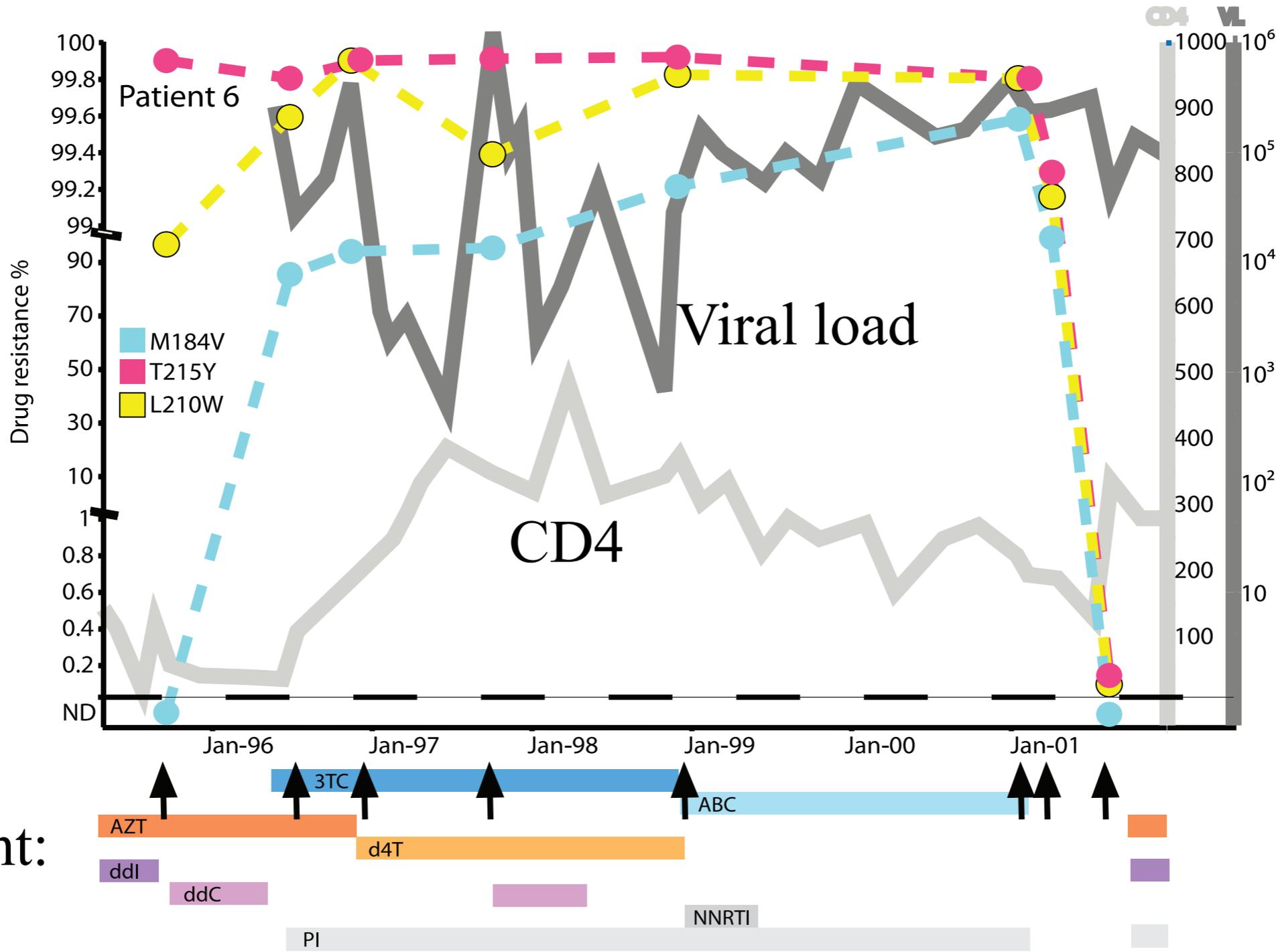
Collab: **Jan Albert** at the Karolinska Institute, Stockholm

- 10000 sequences from each sample
- 180-200 aa of the reverse transcriptase

Drug Resistance Evolution



MAX-PLANCK-GESELLSCHAFT



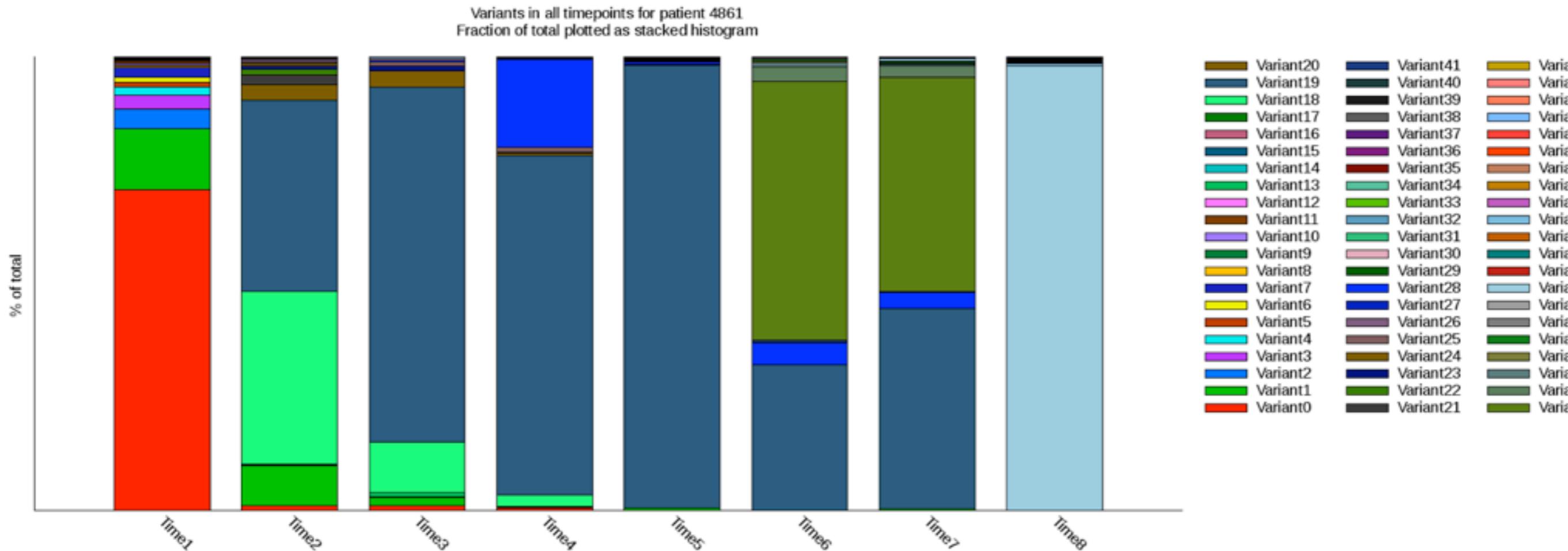
Hedskog, et al, PLoS One, 2010

Drug resistance evolution



MAX-PLANCK-GESELLSCHAFT

- Tremendous diversity, lots of rare variants (~50 in 120bp)
- rapid shifts in the population, latency
- Which variant is replacing others, how fast, and why?



The signature of selection



MAX-PLANCK-GESELLSCHAFT

- Observation: Each amino-acid sequence comes in multiple variants
- The abundances of these variants often show a regular pattern

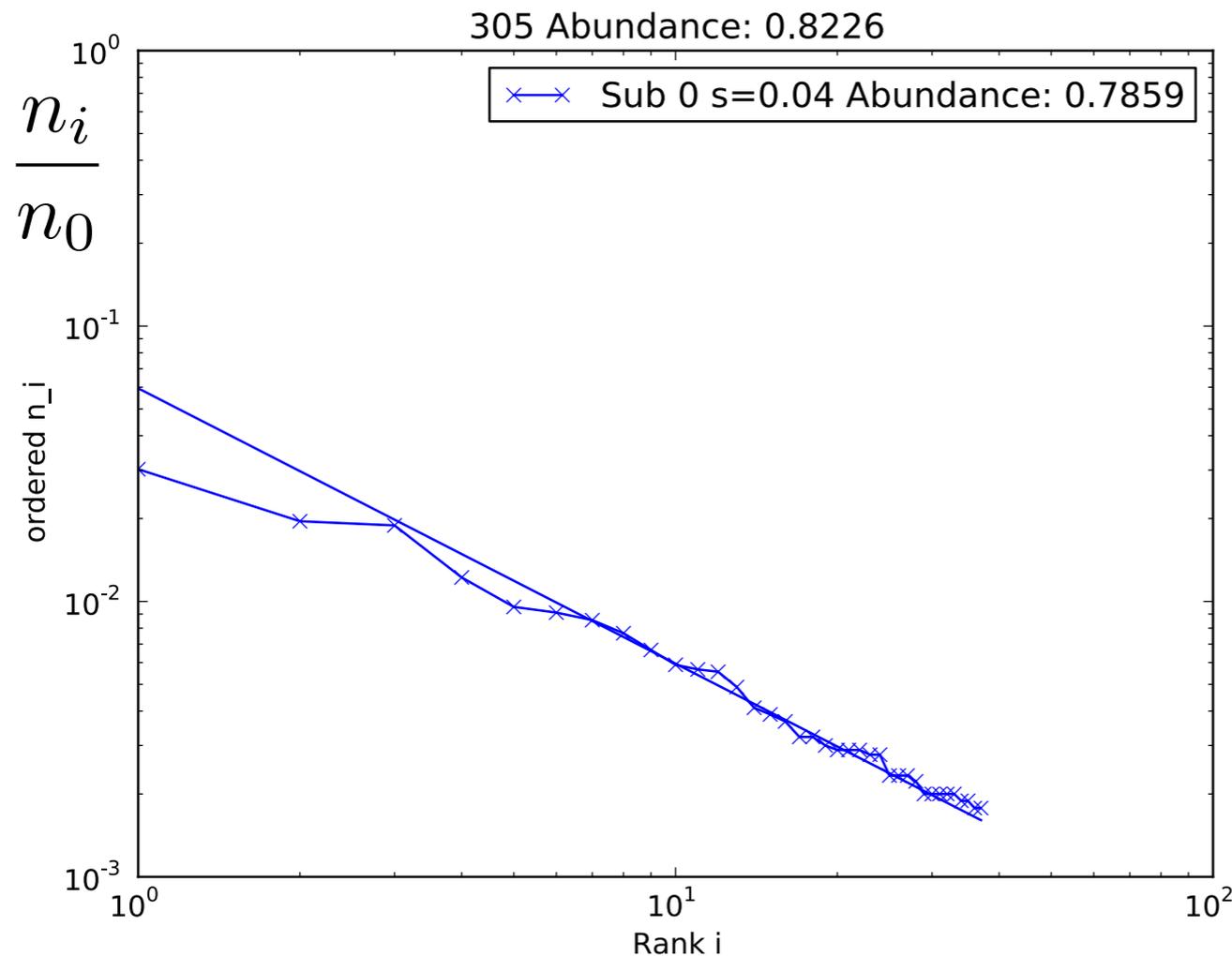
$$n_0 > n_1 > n_2 \dots$$

The signature of selection



- Observation: Each amino-acid sequence comes in multiple variants
- The abundances of these variants often show a regular pattern

$$n_0 > n_1 > n_2 \dots$$



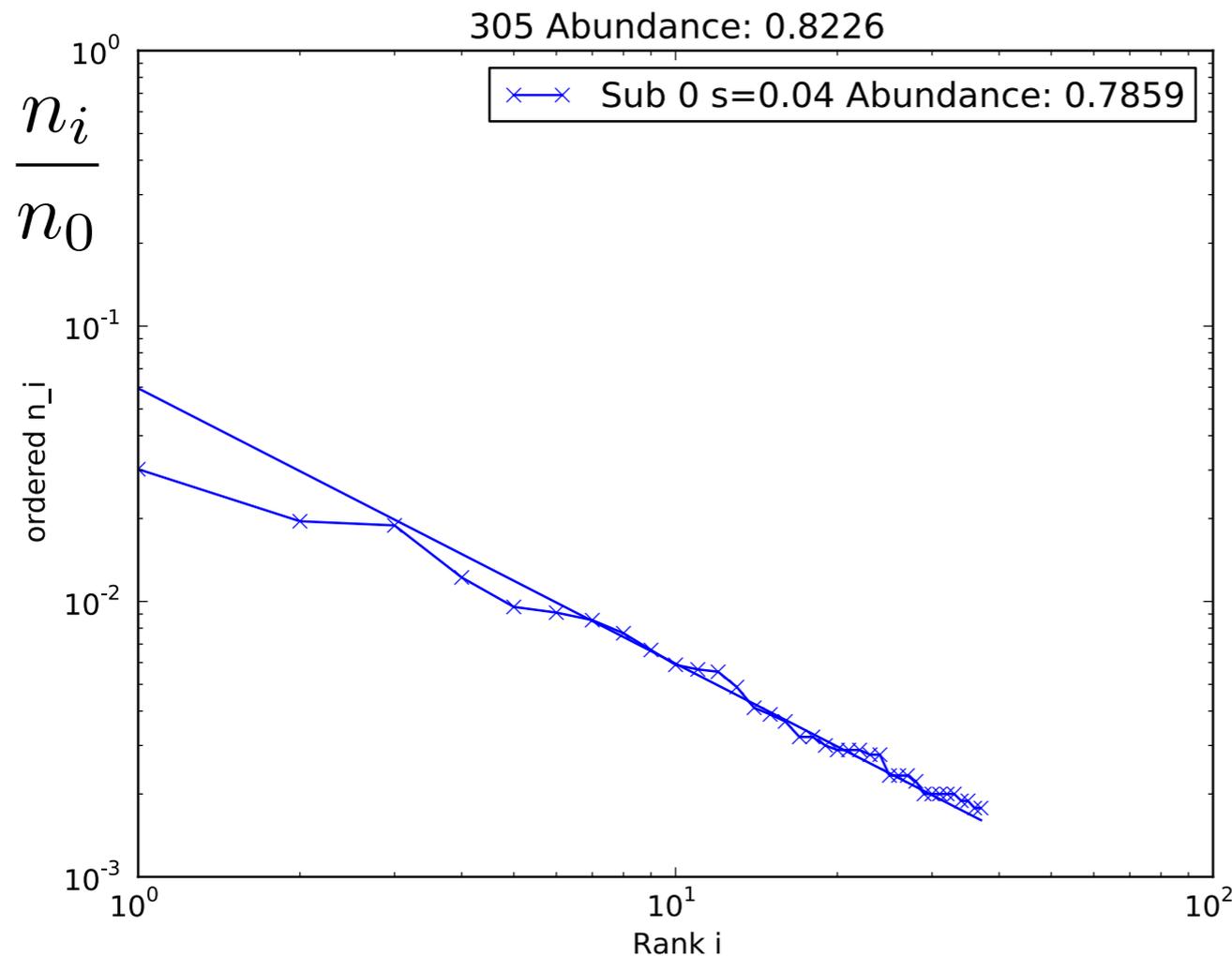
The signature of selection



MAX-PLANCK-GESELLSCHAFT

- Observation: Each amino-acid sequence comes in multiple variants
- The abundances of these variants often show a regular pattern

$$n_0 > n_1 > n_2 \dots$$



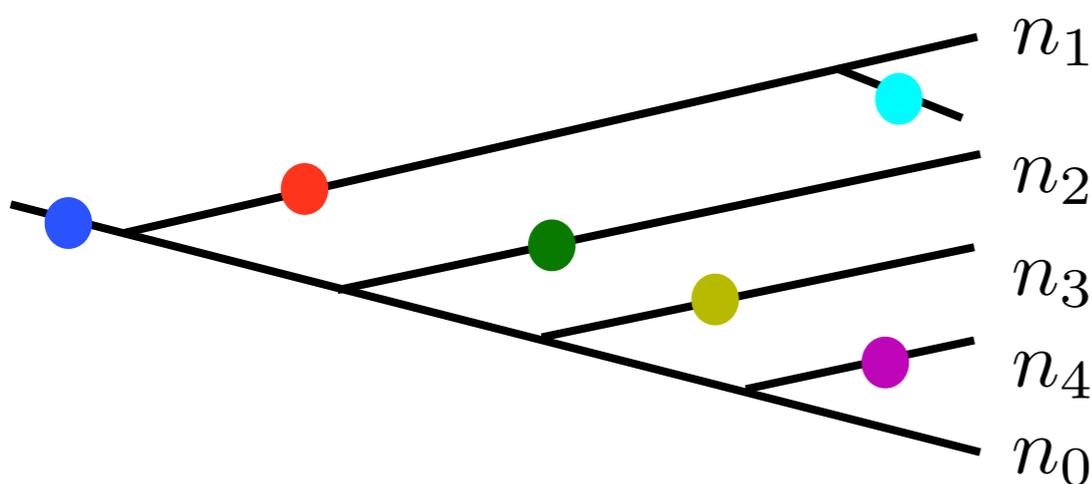
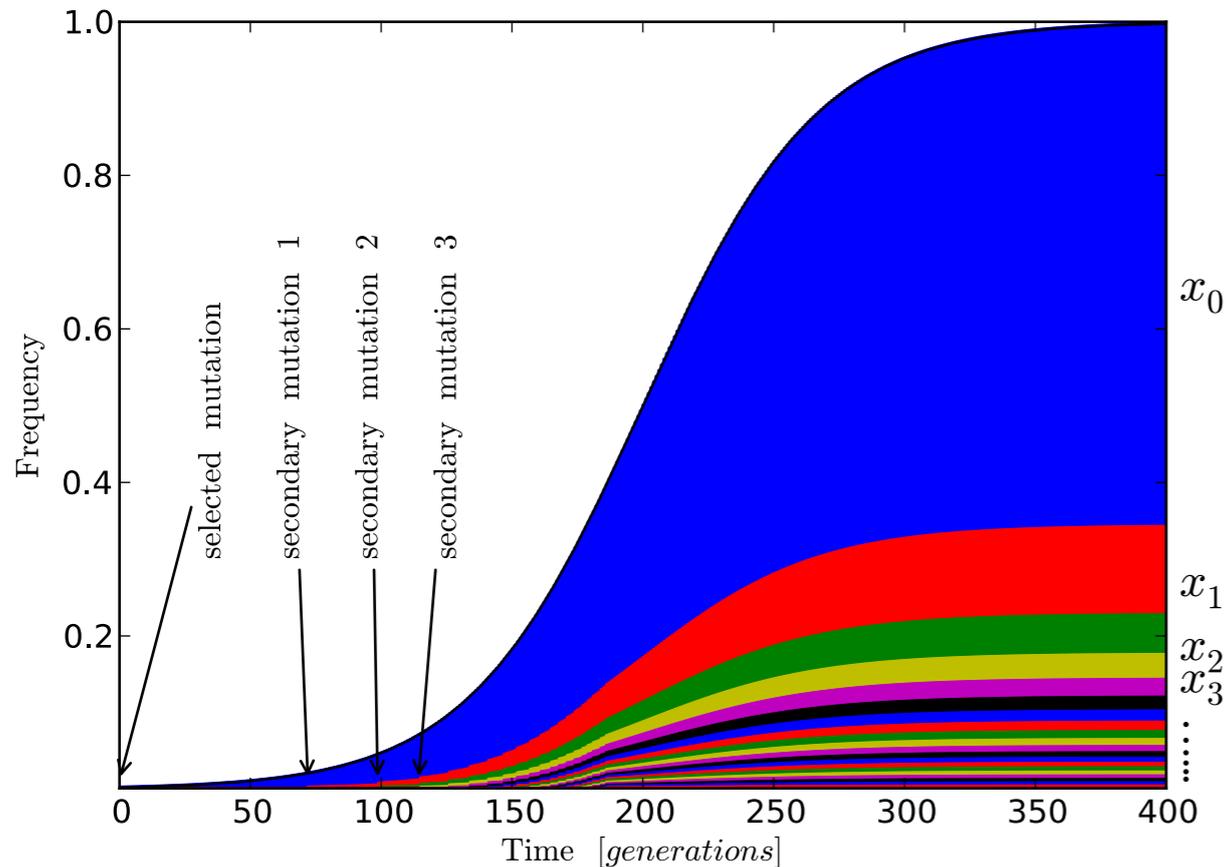
$$\frac{n_i}{n_0} \sim \frac{1}{i}$$

$$\frac{n_i}{n_0} = \frac{\mu}{i^s}$$

Signature of a sweep



MAX-PLANCK-GESELLSCHAFT



$$n_0(t) = e^{(s-\mu)t}$$

$$\int_0^t dt \mu e^{(s-\mu)t'} \approx \frac{\mu e^{(s-\mu)t}}{s-\mu} = 1$$

$$t_i = \frac{1}{s-\mu} \log \frac{i(s-\mu)}{\mu}$$

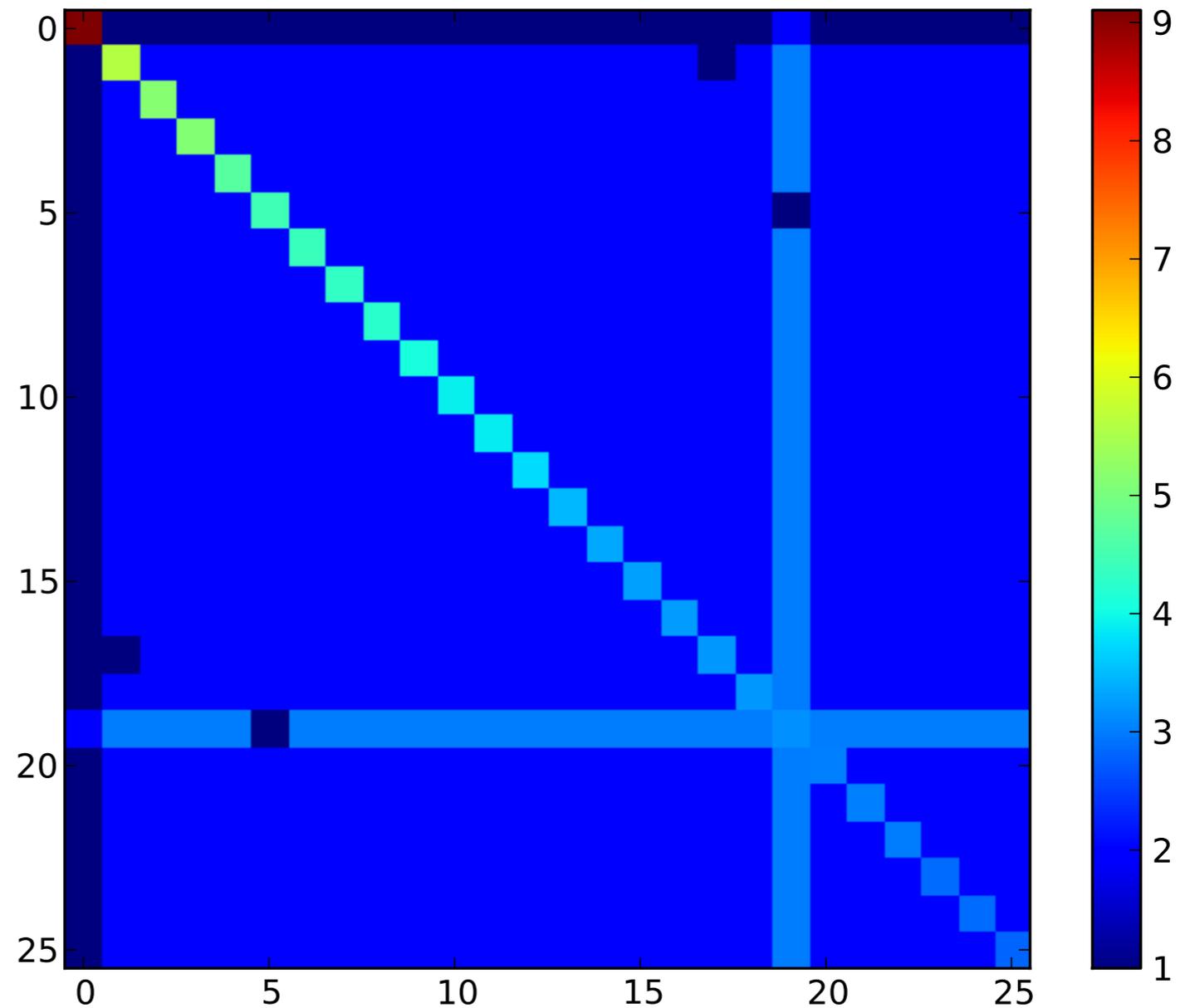
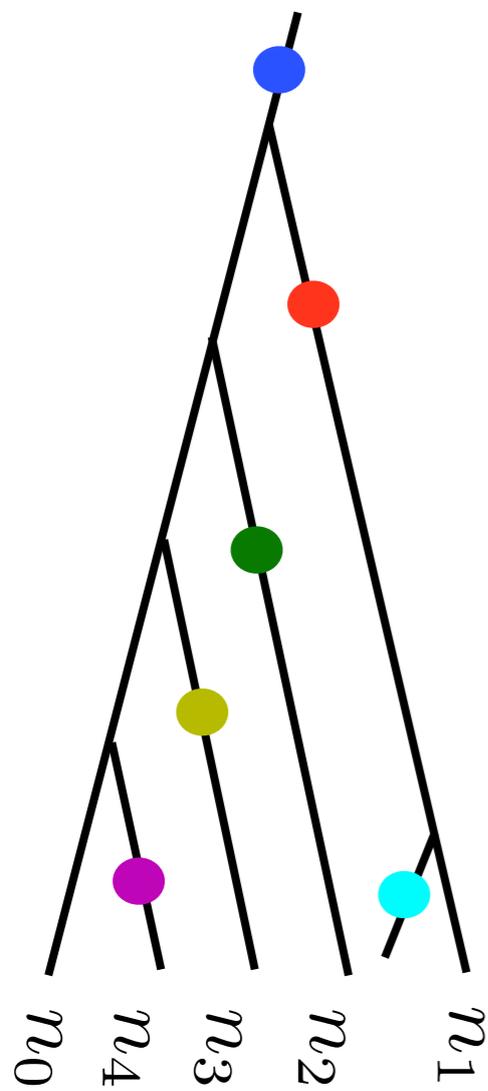
$$\frac{n_i(T)}{n_0(T)} = \frac{e^{(s-\mu)(T-t_i)}}{e^{(s-\mu)T}} = \frac{\mu}{i(s-\mu)}$$

Strength of selection: 5-10%

Star-phylogeny



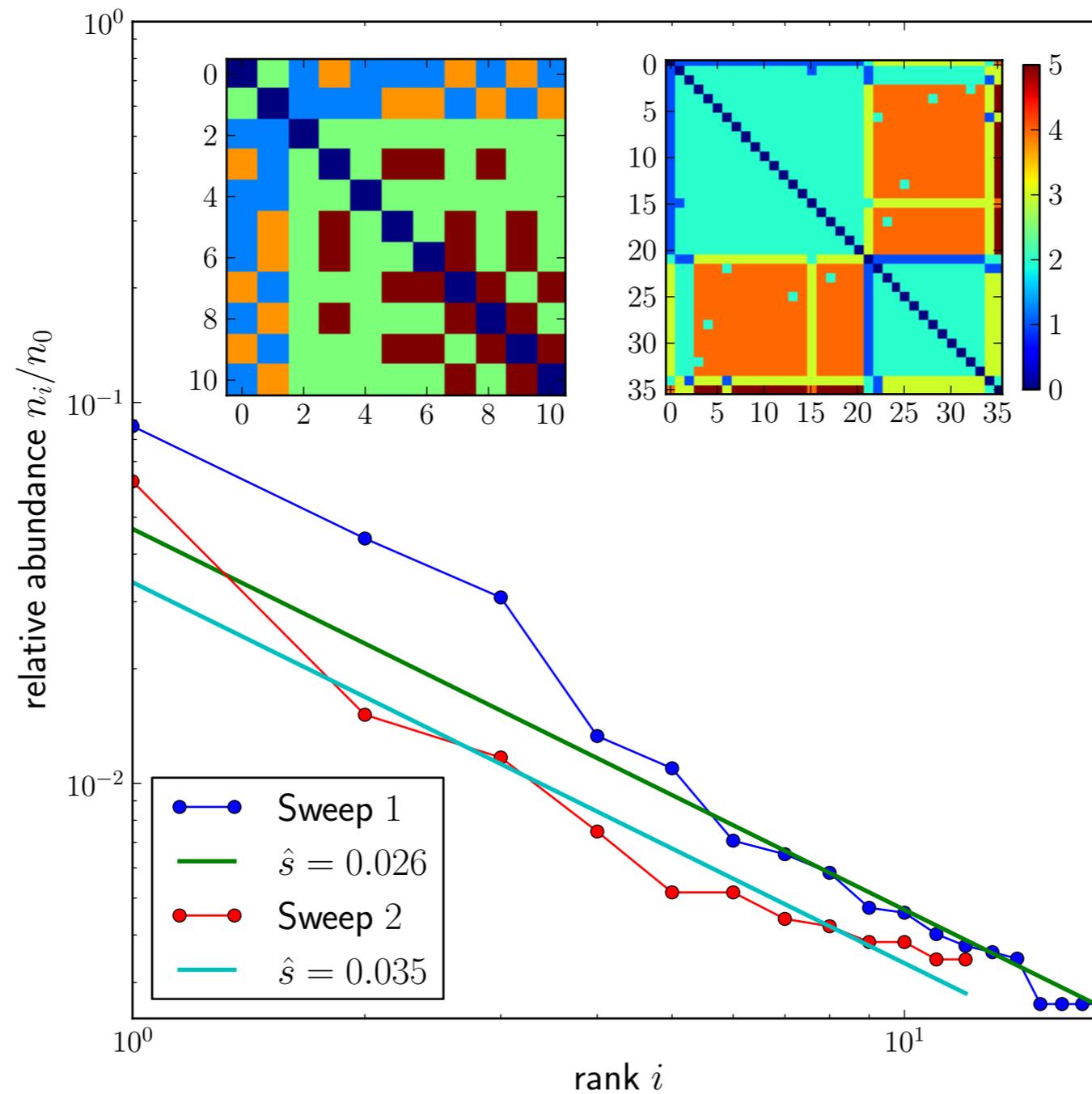
MAX-PLANCK-GESELLSCHAFT



Soft sweeps



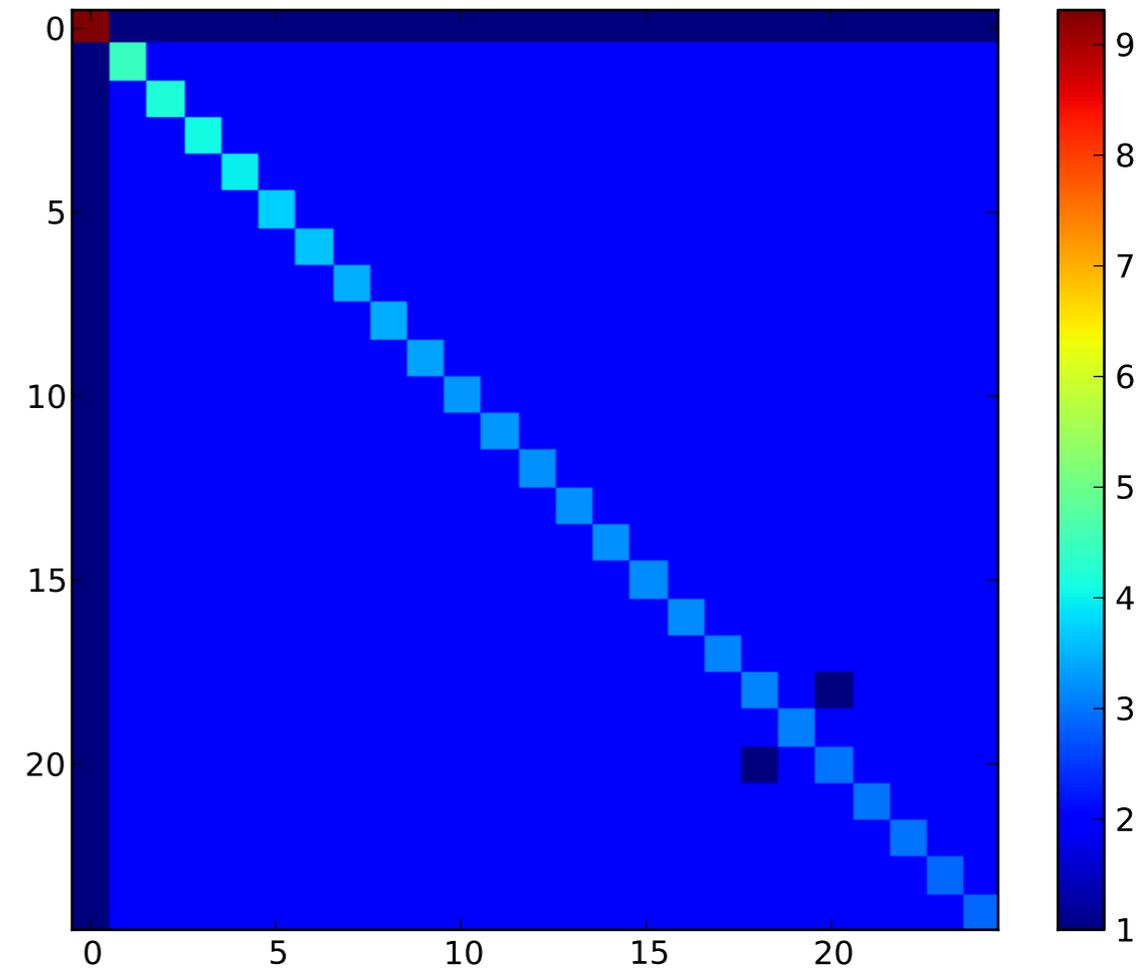
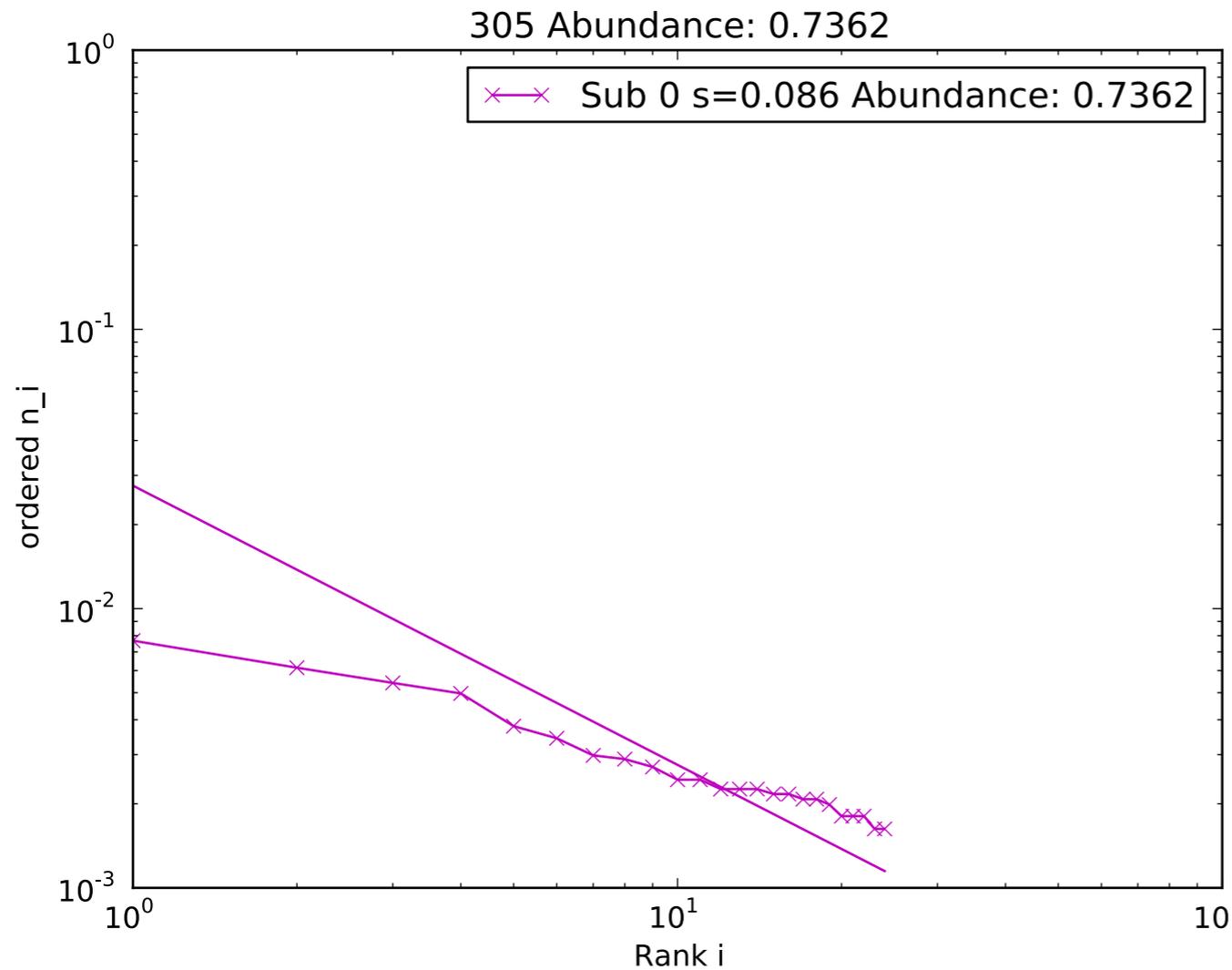
MAX-PLANCK-GESELLSCHAFT



Fast sweep



MAX-PLANCK-GESELLSCHAFT

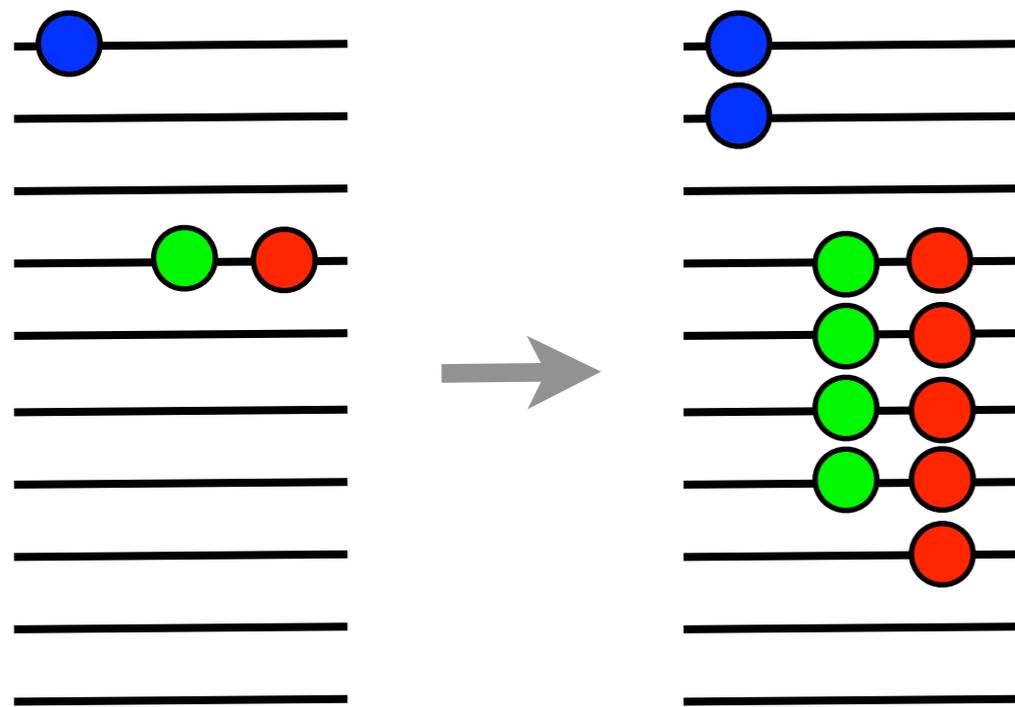


Rare variants consistent with $\nu = \mu t$

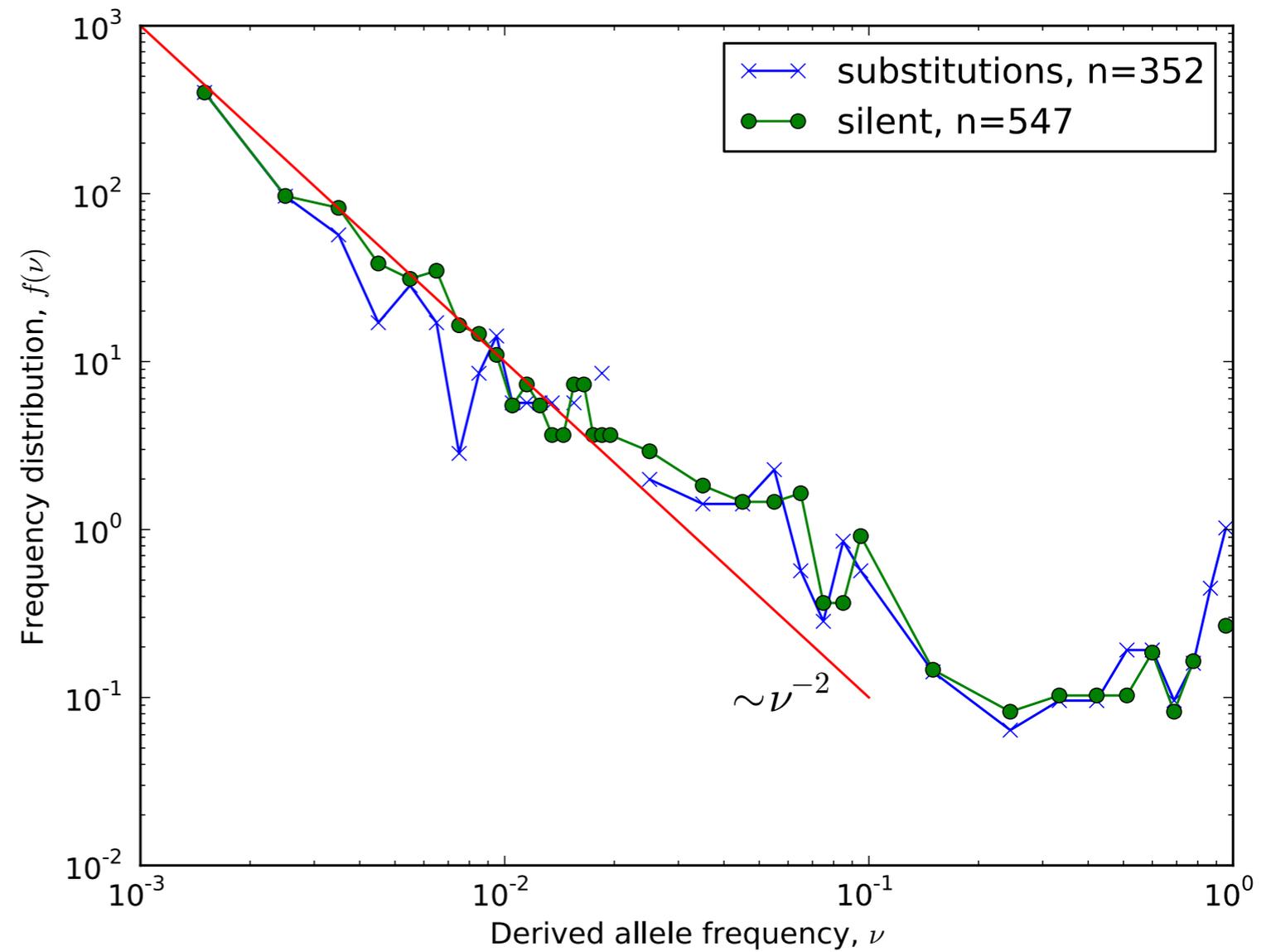
Genetic hitch-hiking



MAX-PLANCK-GESELLSCHAFT



Selection for red will affect green



Acknowledgements



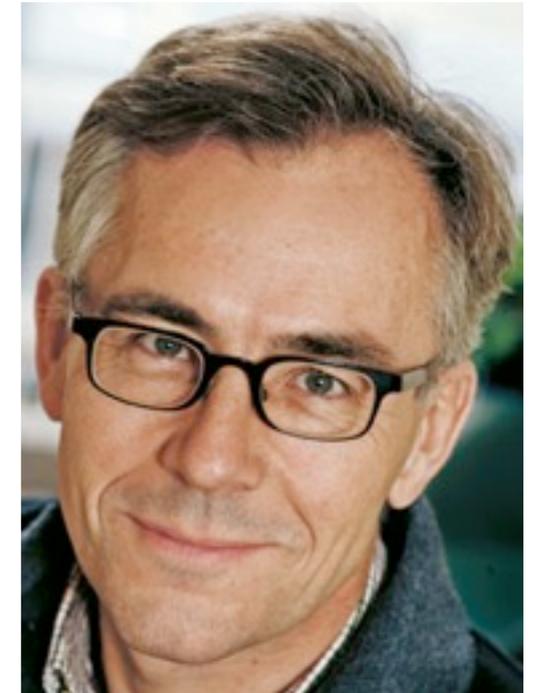
MAX-PLANCK-GESELLSCHAFT



Boris Shraiman, KITP



Thomas Leitner, LANL



Jan Albert, KI



Philipp Messer, Stanford

Johanna Jernberg, KI



Harvey Karp Discovery Award

Richard Neher

KITP, 2011



MAX-PLANCK-GESELLSCHAFT

