Recovery of genetic variation in HIV

Pleuni Pennings, Sergey Kryazhimskiy, John Wakeley
Harvard University

Marseille CIRM June 2012

Goal of my work

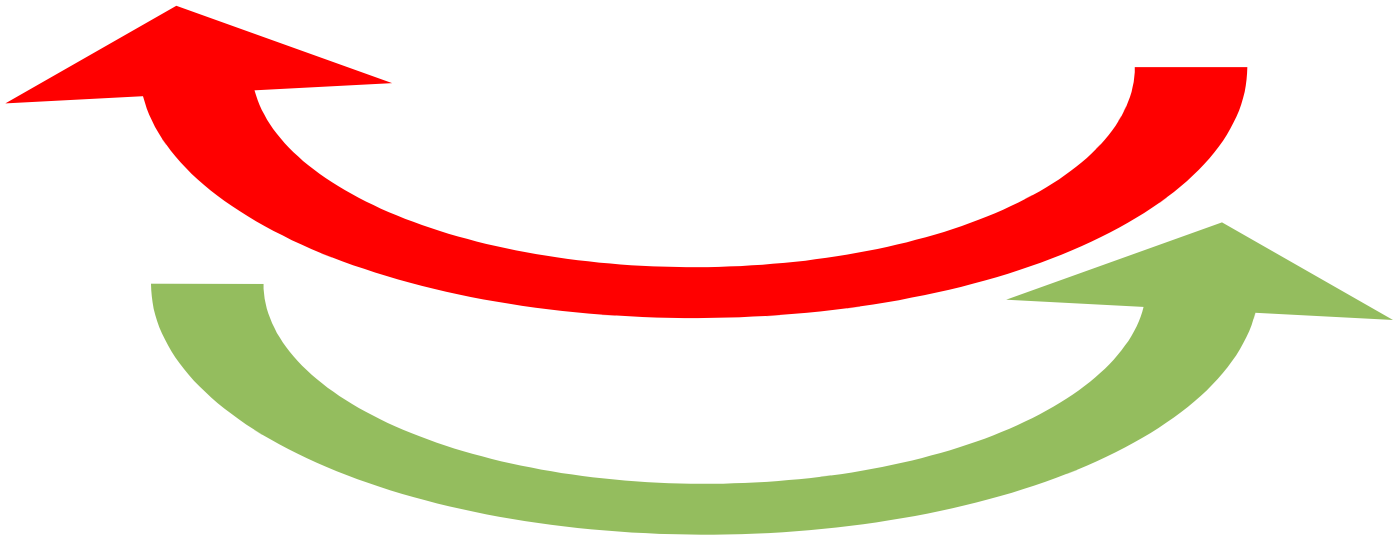
Evolutionary biology

Population genetics

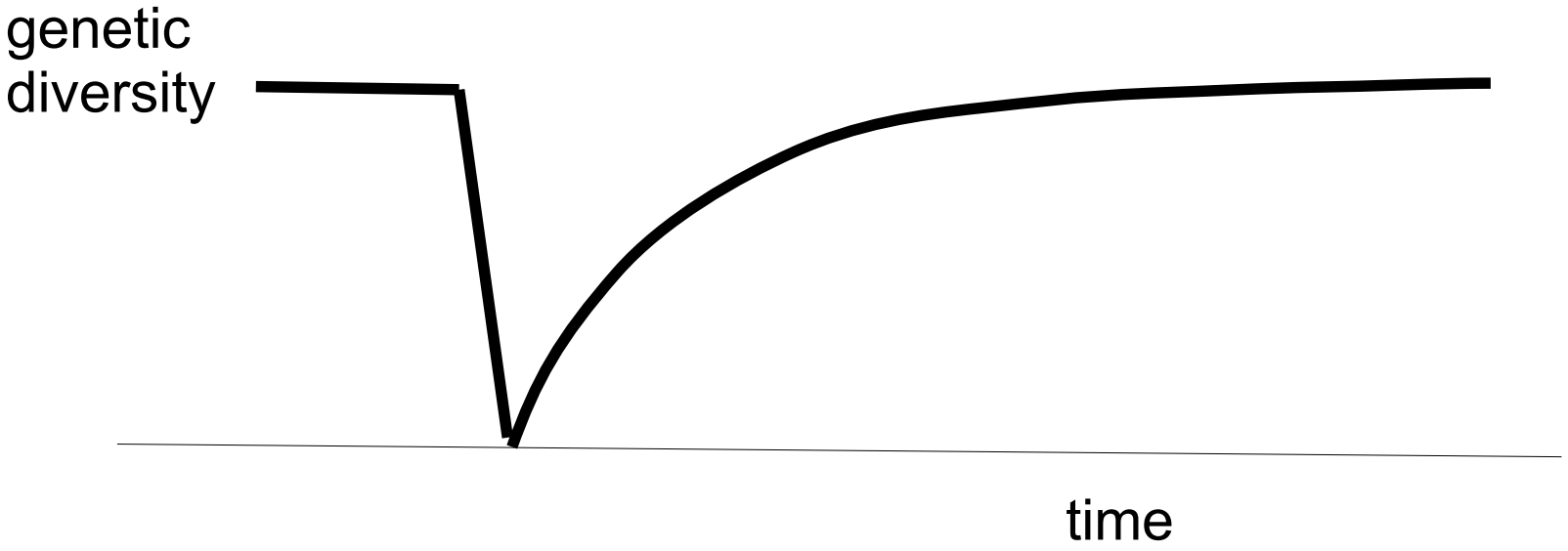
HIV within host

Drug resistance

Data & Questions

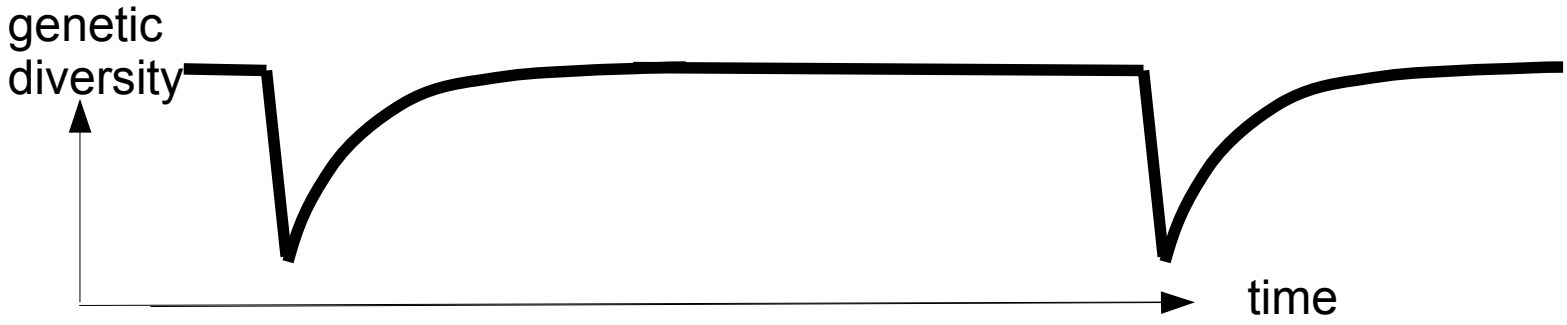


Why look at recovery of diversity?



Why look at recovery of diversity?

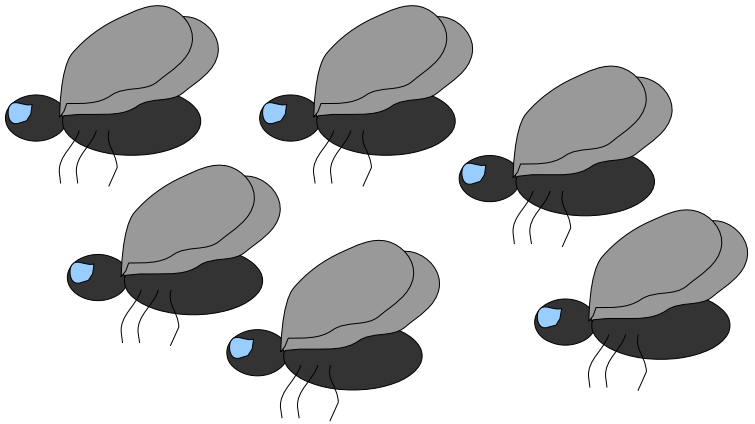
1. If recovery fast, then equilibrium assumptions justified

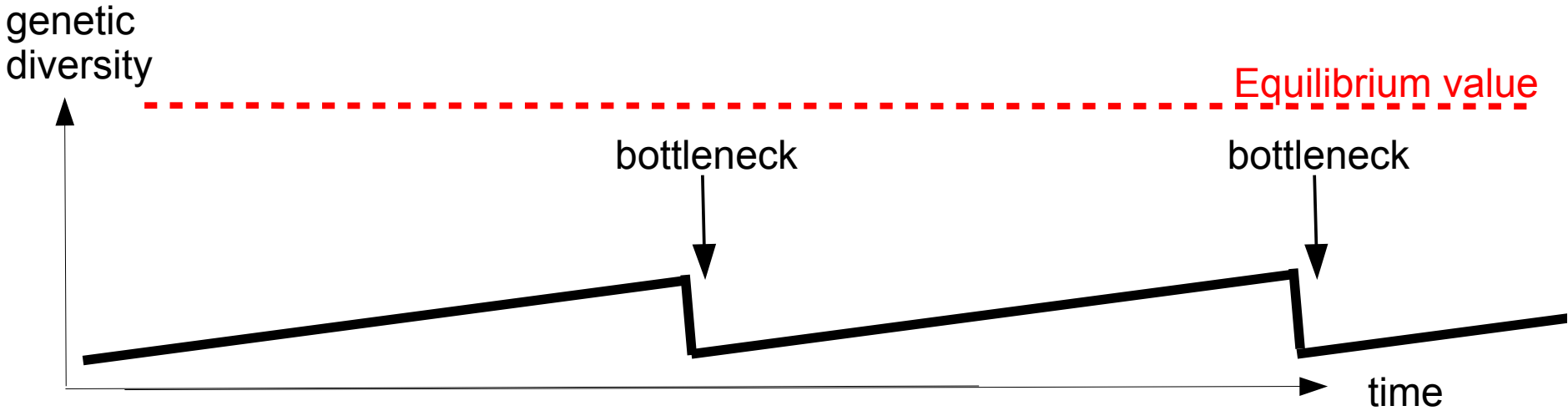


2. Recovery tells us about evolutionary parameters

Examples

Drosophila melanogaster
Karasov, Messer, Petrov 2010

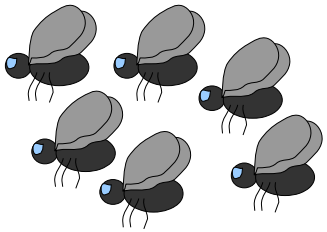




Genetic diversity never reaches expected equilibrium

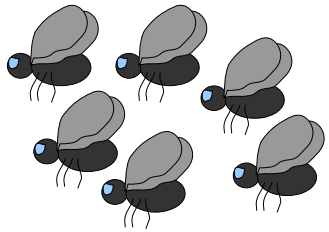
Examples

Drosophila melanogaster
never recovers
Karasov, Messer, Petrov 2010

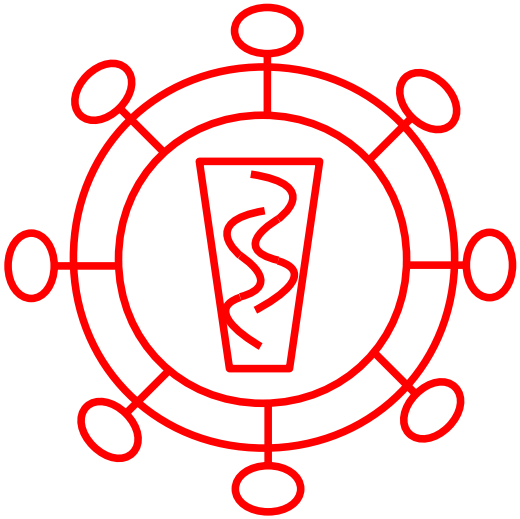


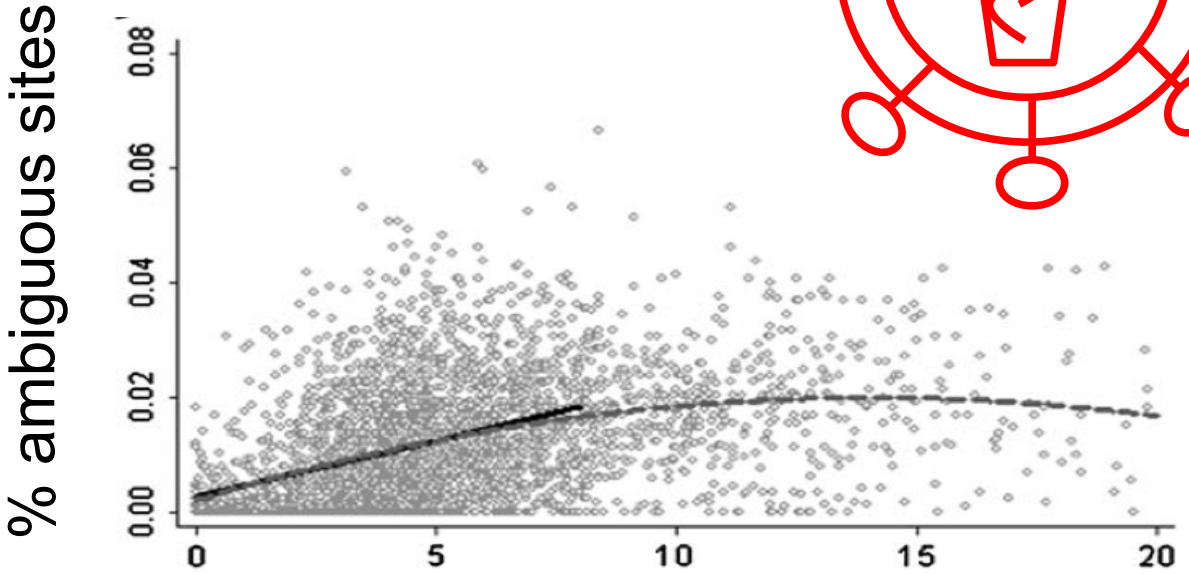
Examples

Drosophila melanogaster
never recovers
Karasov, Messer, Petrov 2010



HIV
Kouyos et al 2011

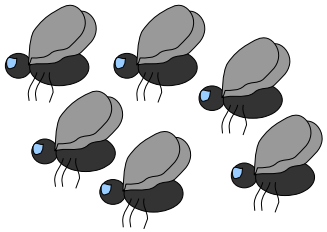




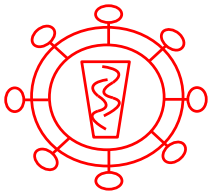
Takes few years to recover

Examples

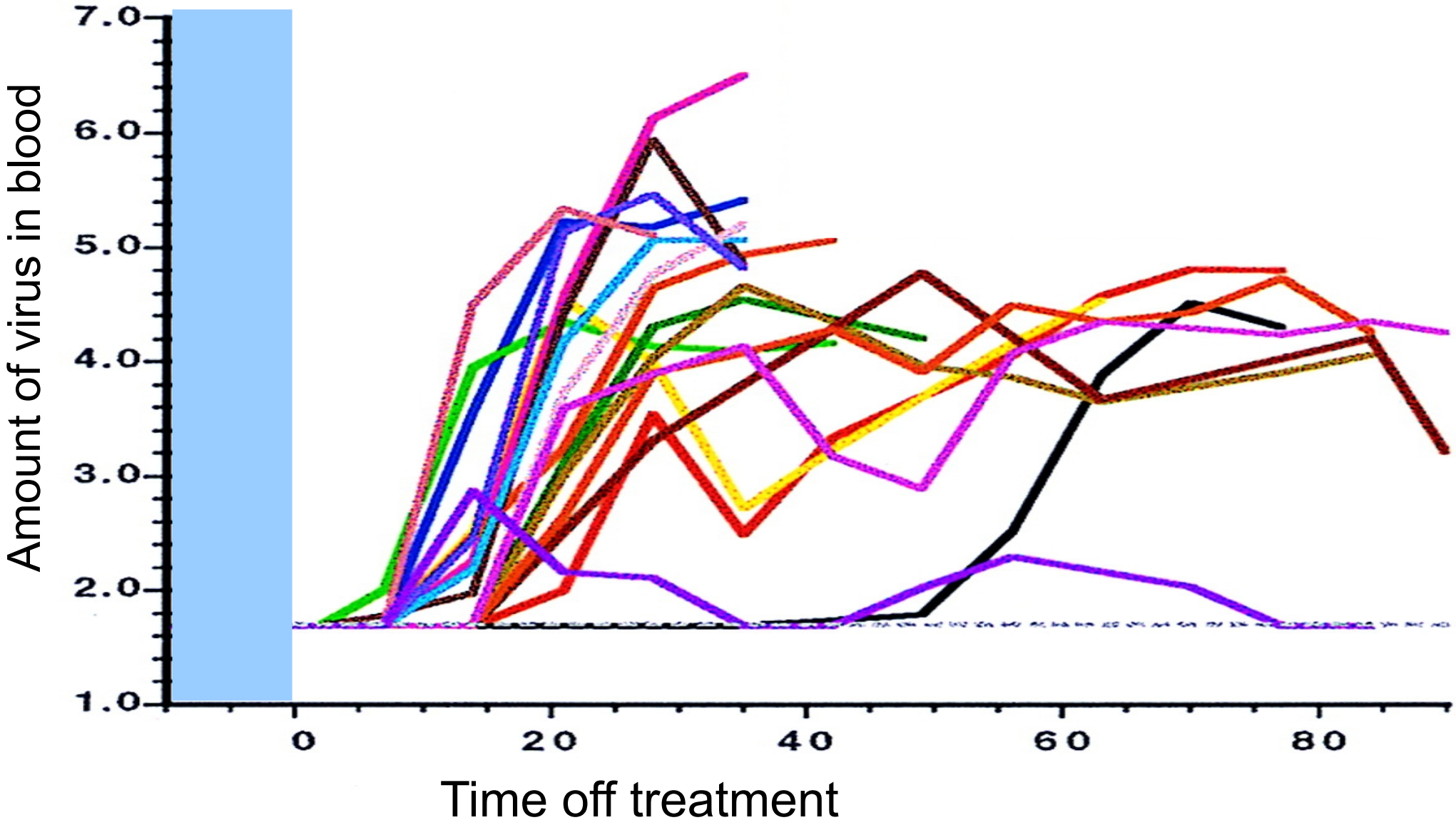
Drosophila melanogaster
never recovers
Karasov, Messer, Petrov 2010



HIV
few years to recover
Kouyos et al 2011

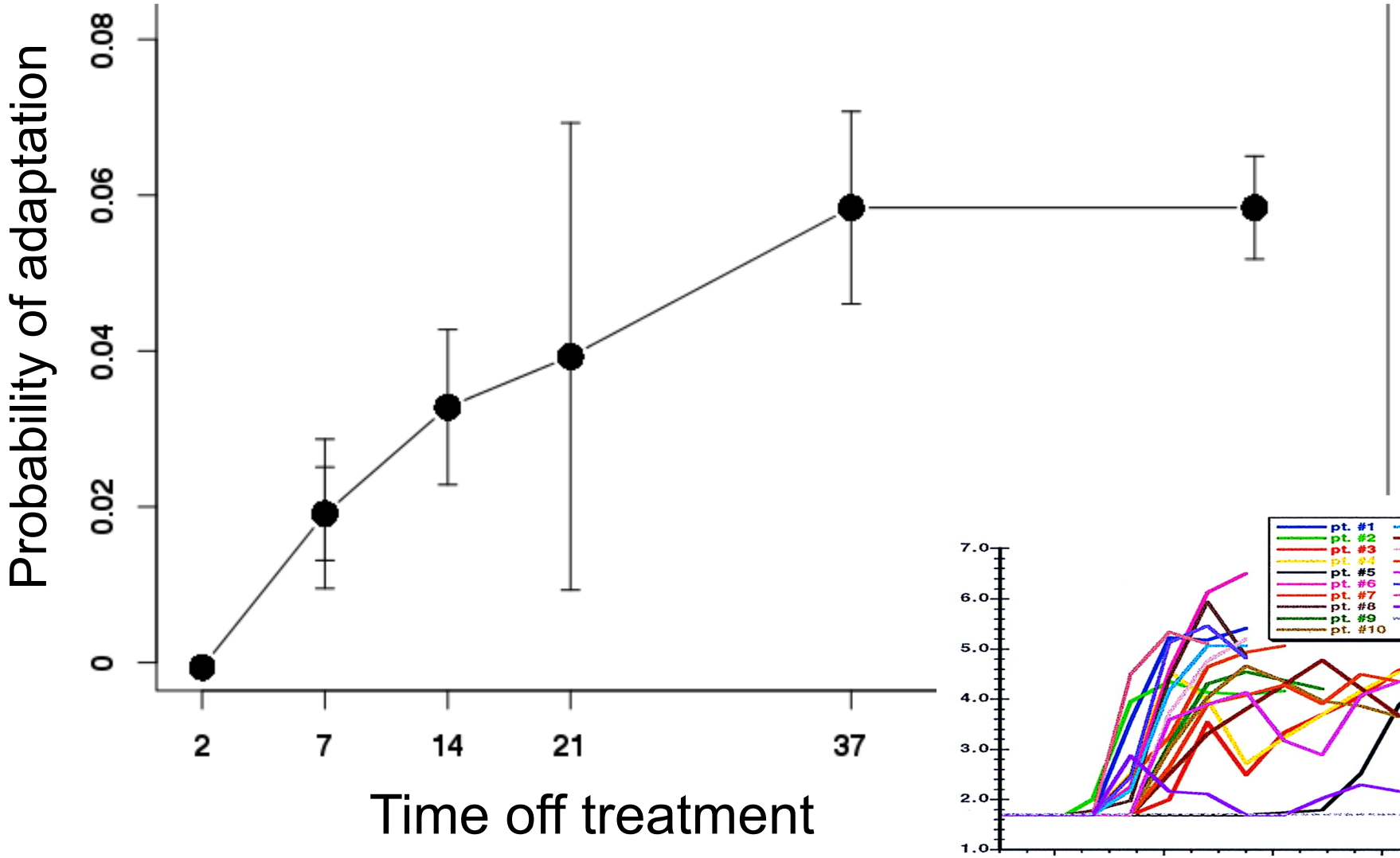


Examples


Amount of virus in blood
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0
20
40
60
80
Time off treatment

Examples

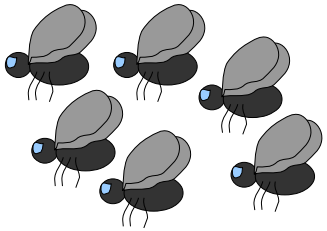
Adaptive potential during treatment interruption



Pennings, 2012 PLoS Comp Bio

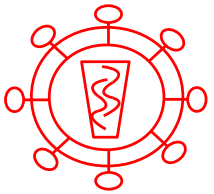
Examples

Drosophila melanogaster
never recovers
Karasov, Messer, Petrov 2010



HIV
few years to recover
Kouyos et al 2011

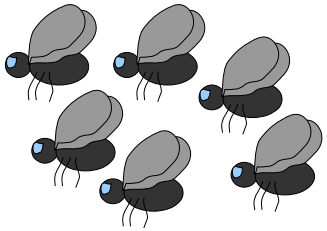
HIV
days - weeks to recover
Pennings 2012



Bottleneck size 1
Neutral diversity

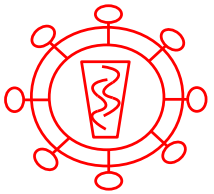
Larger bottleneck
Resistance muts

Drosophila melanogaster
never recovers
Karasov, Messer, Petrov 2010



HIV
few years to recover
Kouyos et al 2011

HIV
days - weeks to recover
Pennings 2012

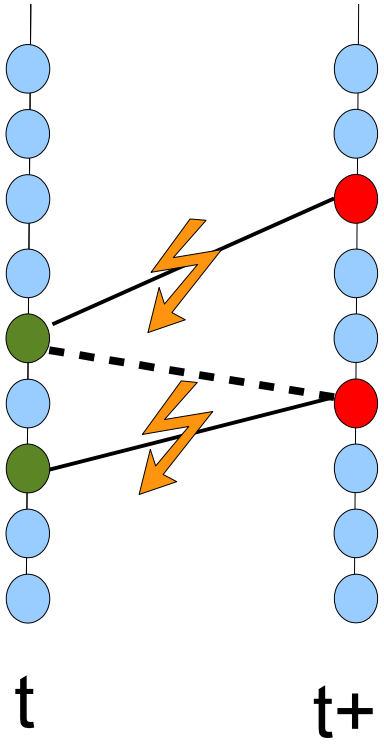


Theoretical expectations

Neutrality

Theoretical expectations

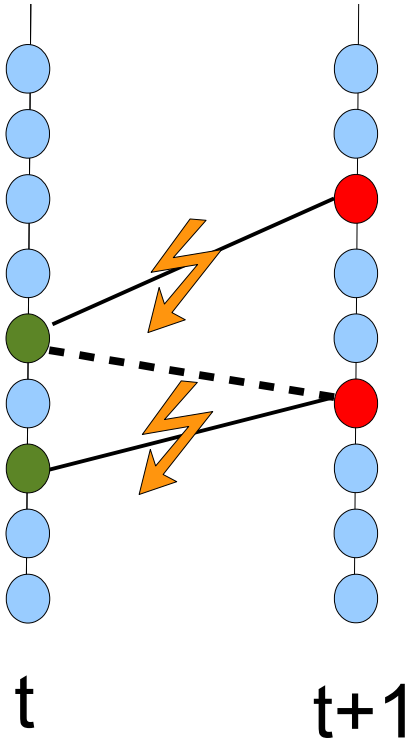
Neutrality



$$H_{t+1} = H_t\left(1 - \frac{1}{N}\right) + 2u$$

Theoretical expectations

Neutrality



$$H_{t+1} = H_t\left(1 - \frac{1}{N}\right) + 2\mathrm{u}$$

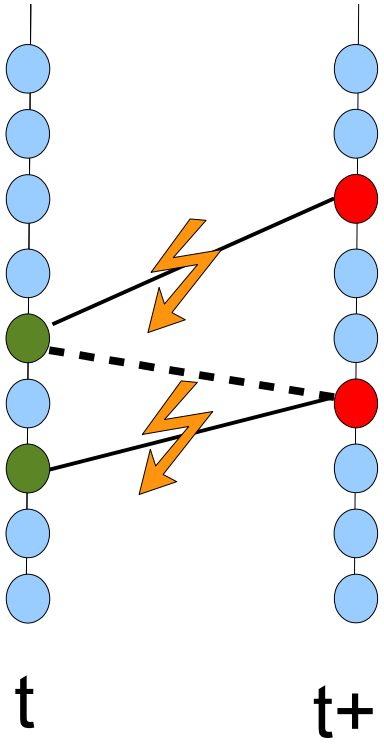
$$\frac{dH}{dt} = -\left(\frac{1}{N}\right)H + 2\mathrm{u}$$

$$H_0 = 0$$

$$H_t = 2\mathrm{Nu}\left(1 - e^{-\left(\frac{1}{N}t\right)}\right)$$

Theoretical expectations

Neutrality

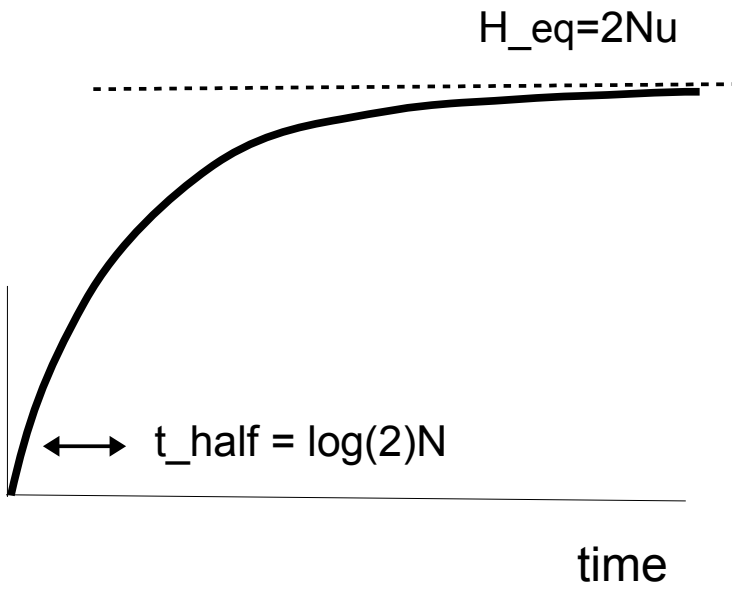


$$H_{t+1} = H_t\left(1 - \frac{1}{N}\right) + 2u$$

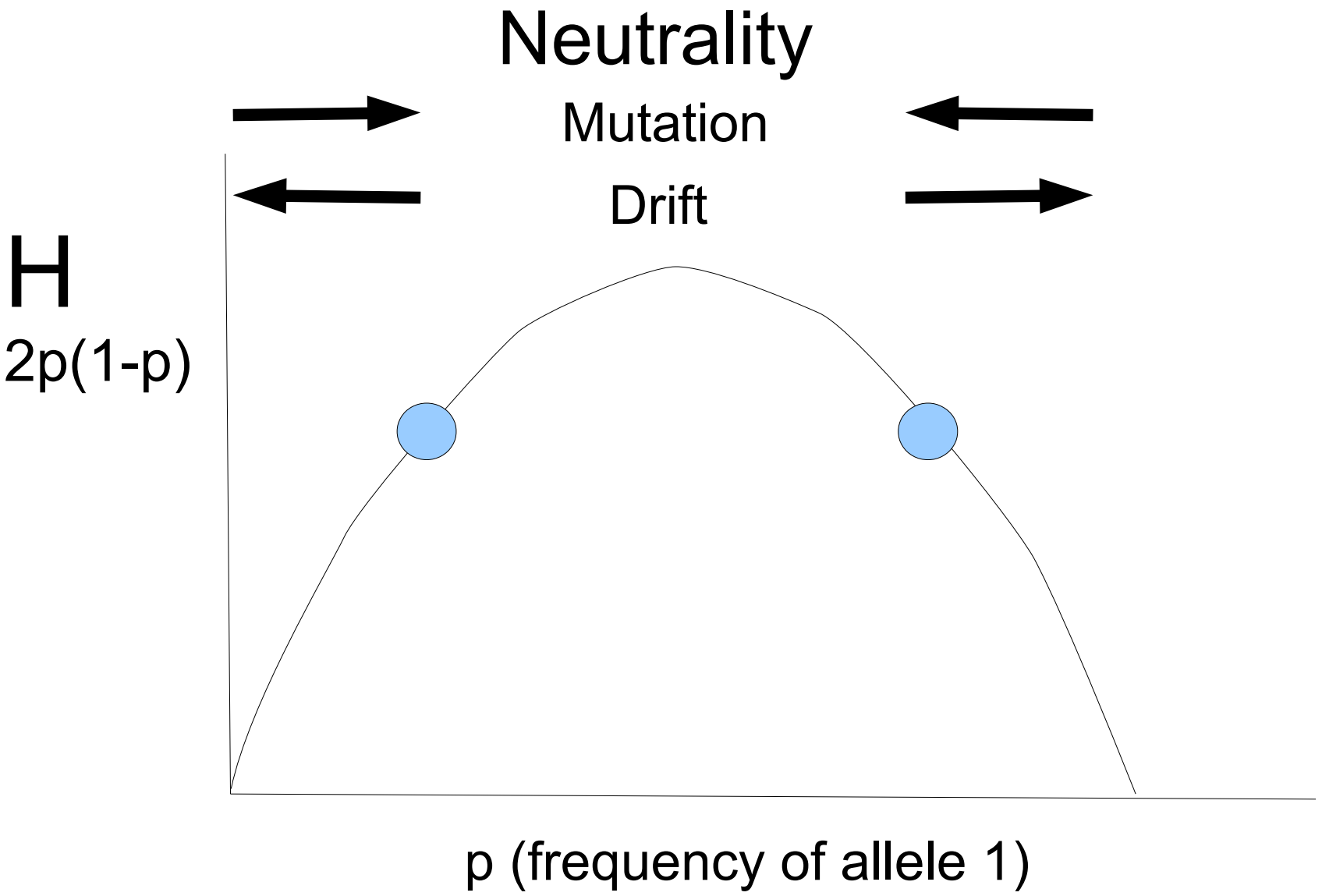
$$\frac{dH}{dt} = -\left(\frac{1}{N}\right)H + 2u$$

$$H_0 = 0$$

$$H_t = 2Nu\left(1 - e^{-\left(\frac{1}{N}t\right)}\right)$$

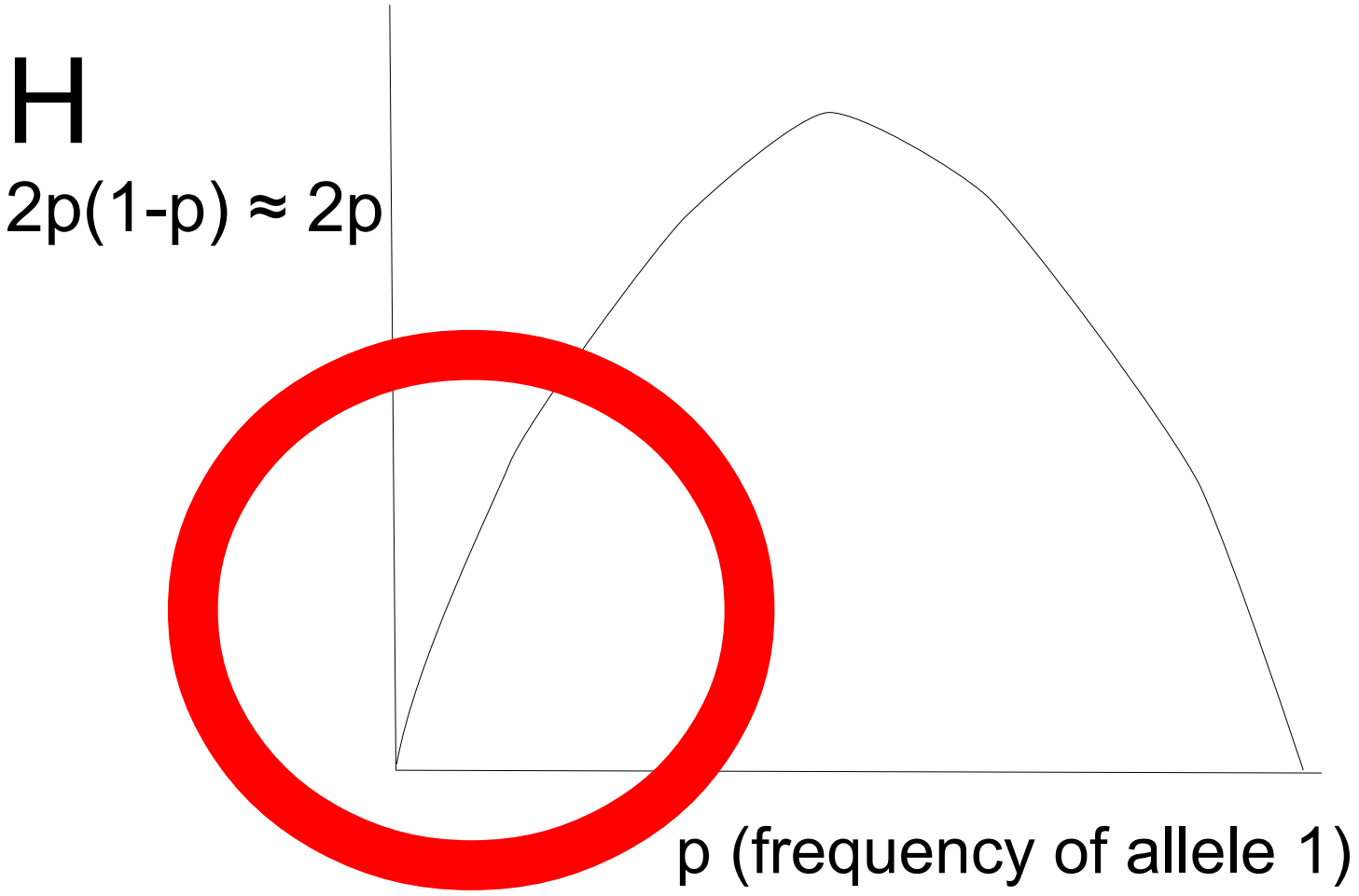


Theoretical expectations



Theoretical expectations

Strong selection



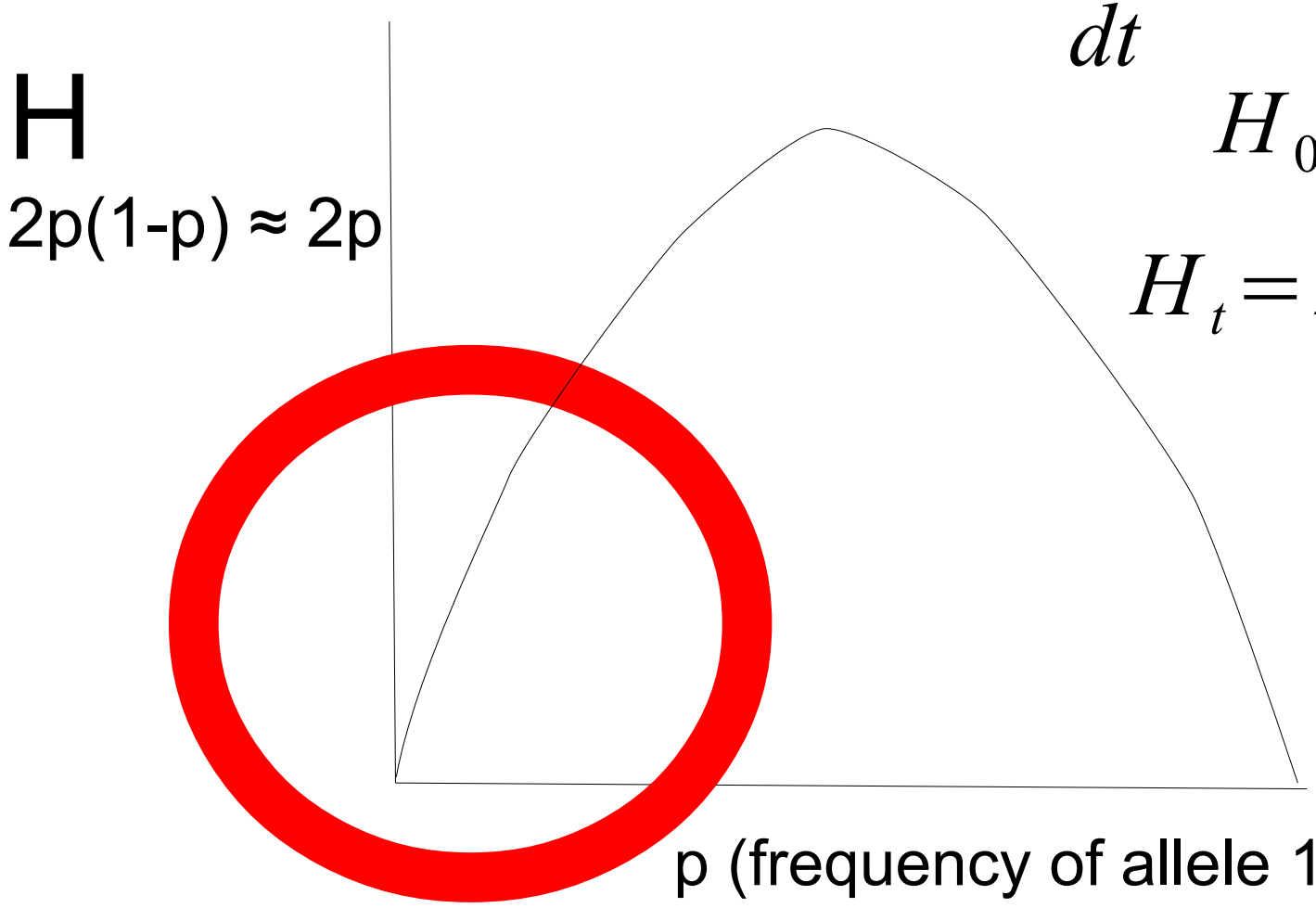
Theoretical expectations

Strong selection

$$\frac{dH}{dt} = -sH + 2u$$

$$H_0 = 0$$

$$H_t = 2u/s\left(1 - e^{-(st)}\right)$$

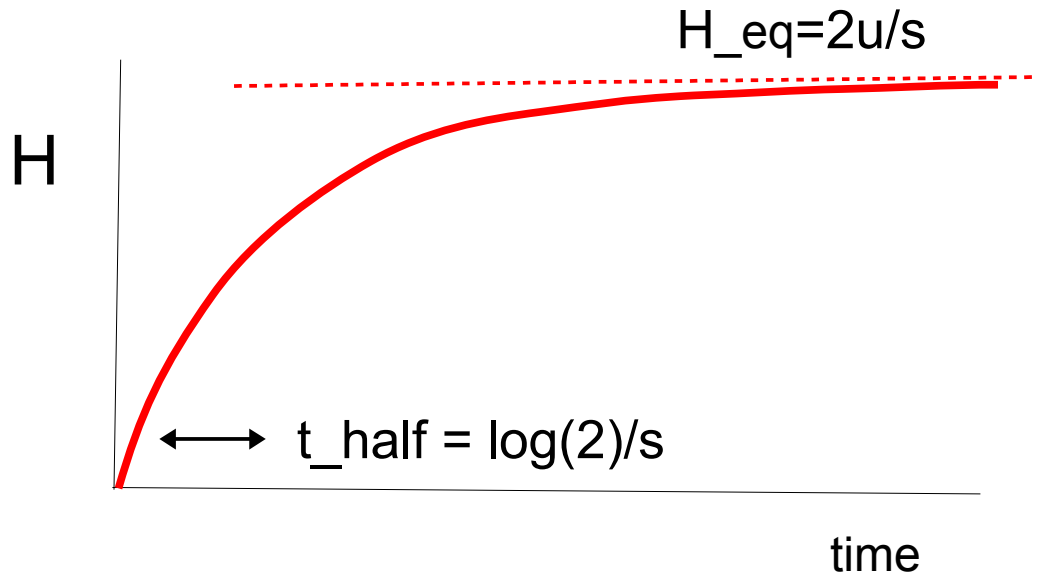


Theoretical expectations

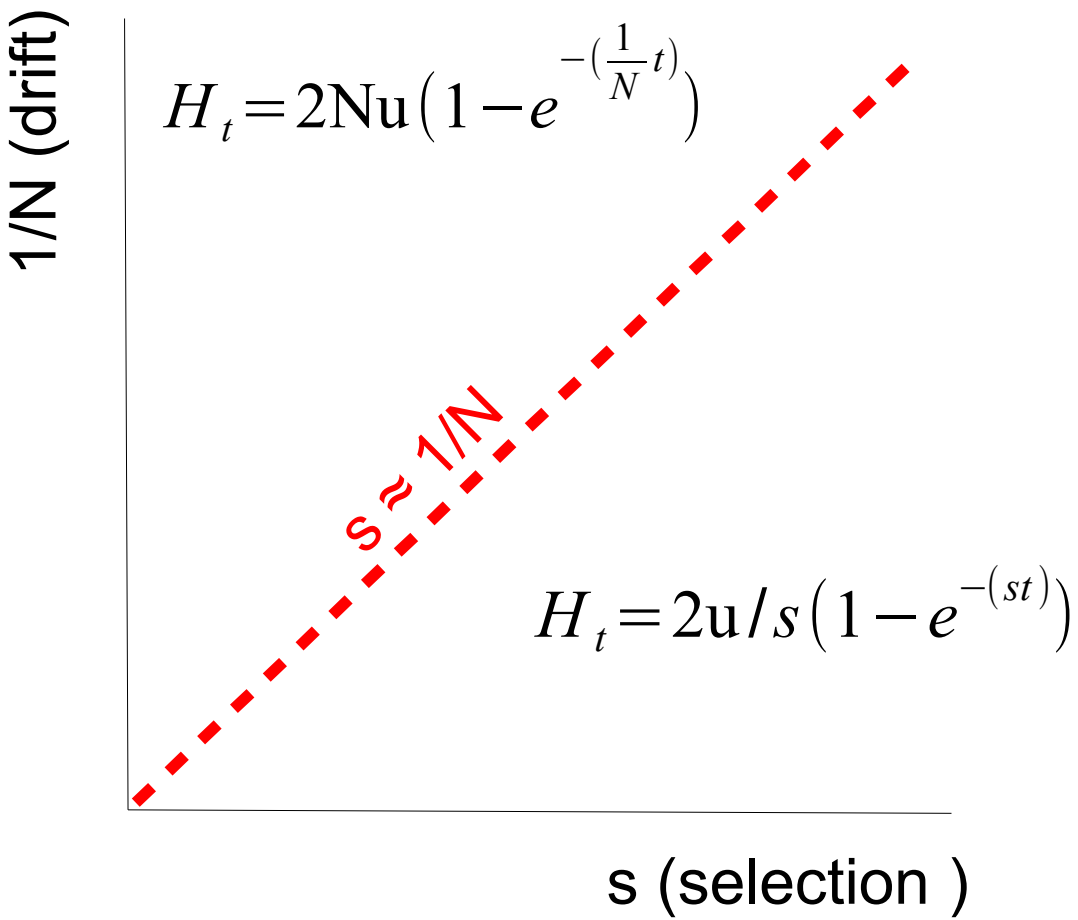
$$\frac{dH}{dt} = -sH + 2\mathrm{u}$$

$$H_0 = 0$$

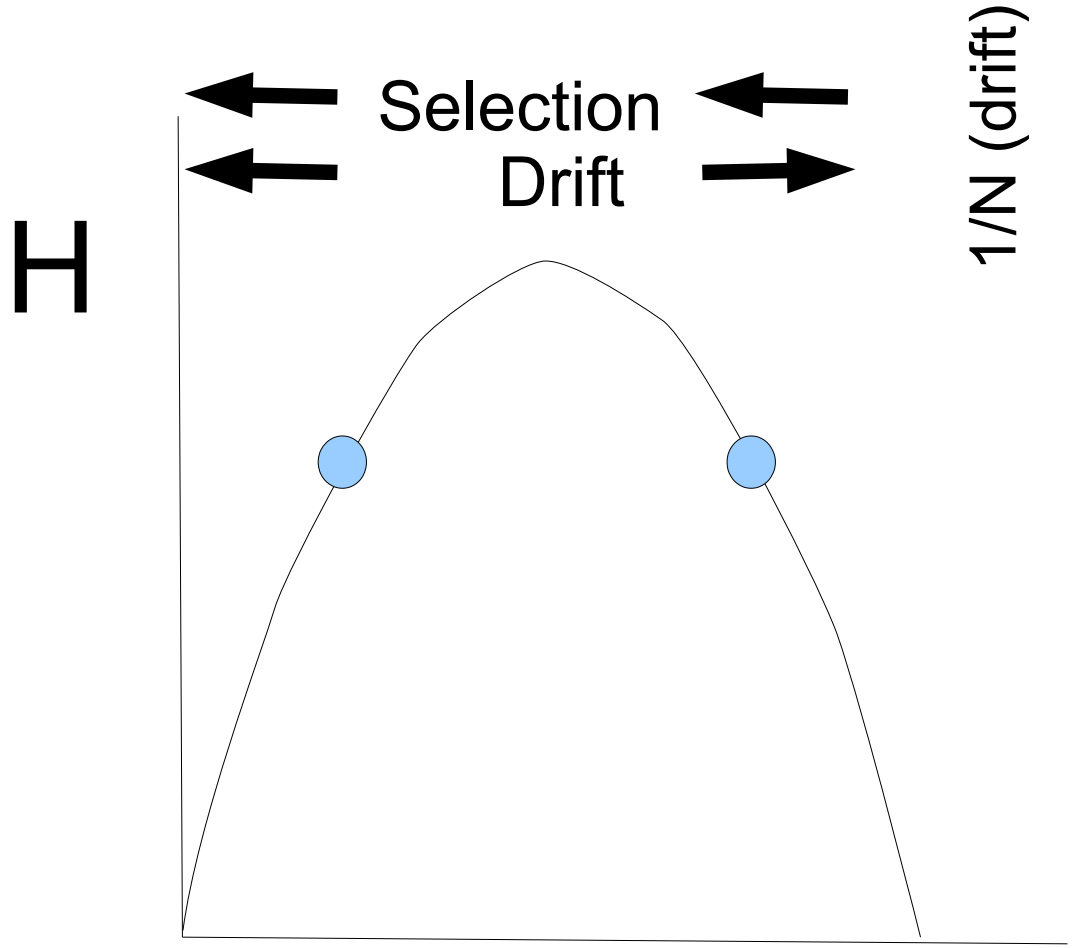
$$H_t = 2\mathrm{u}/s\left(1 - e^{-(st)}\right)$$

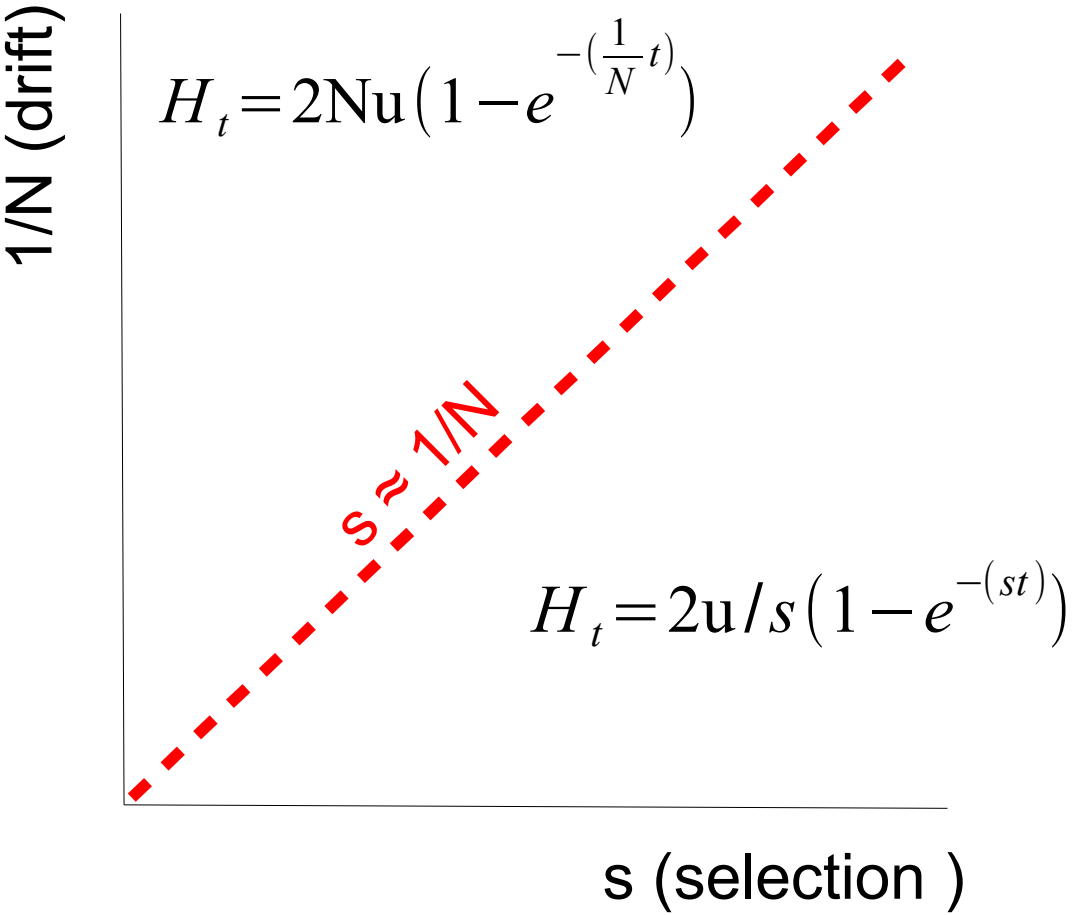


Theoretical expectations

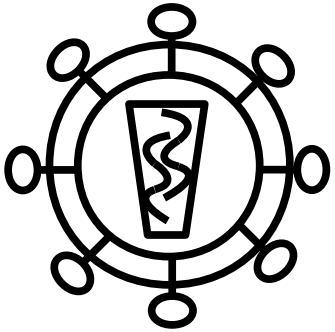


Theoretical expectations

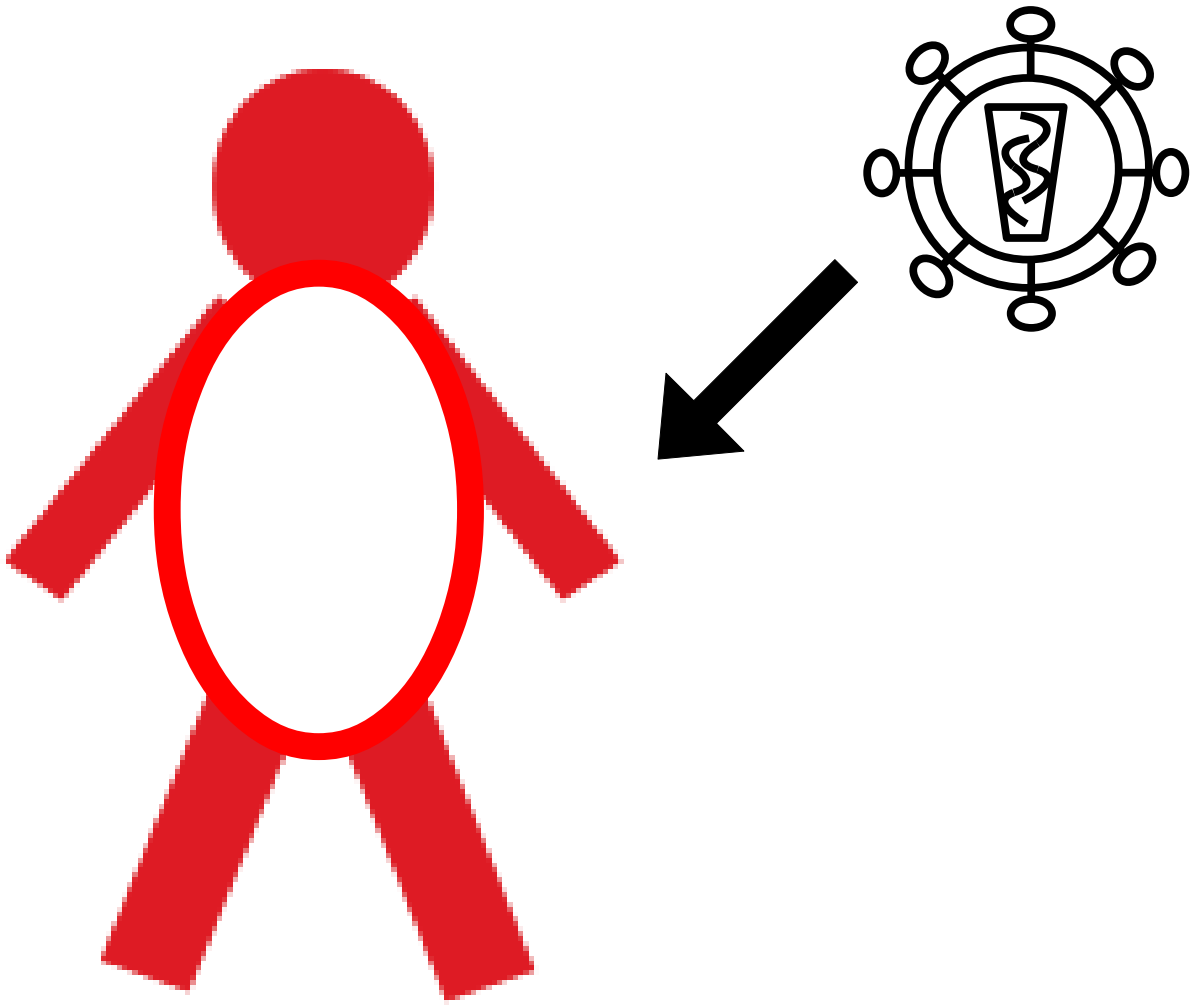




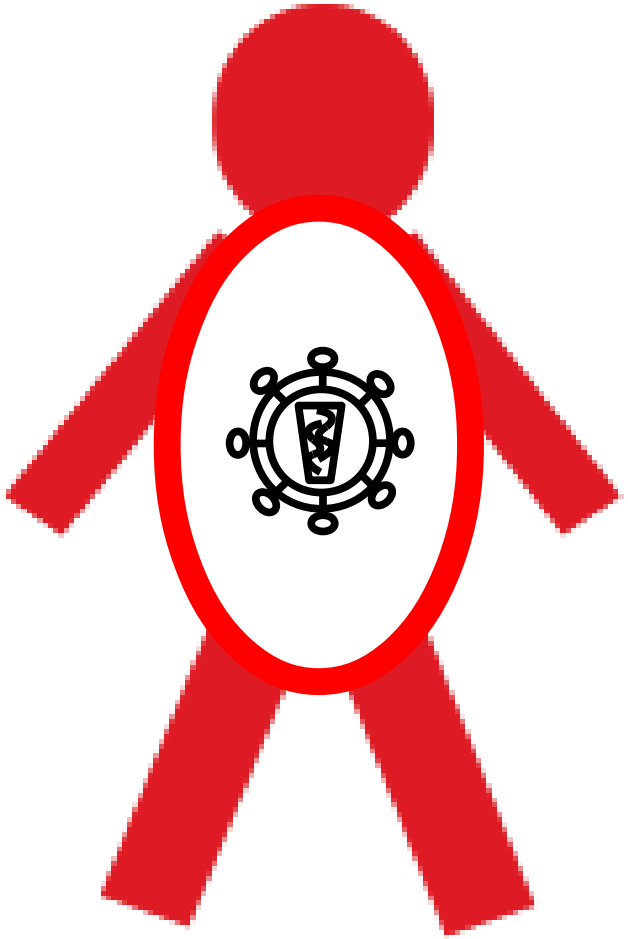
DATA: HIV



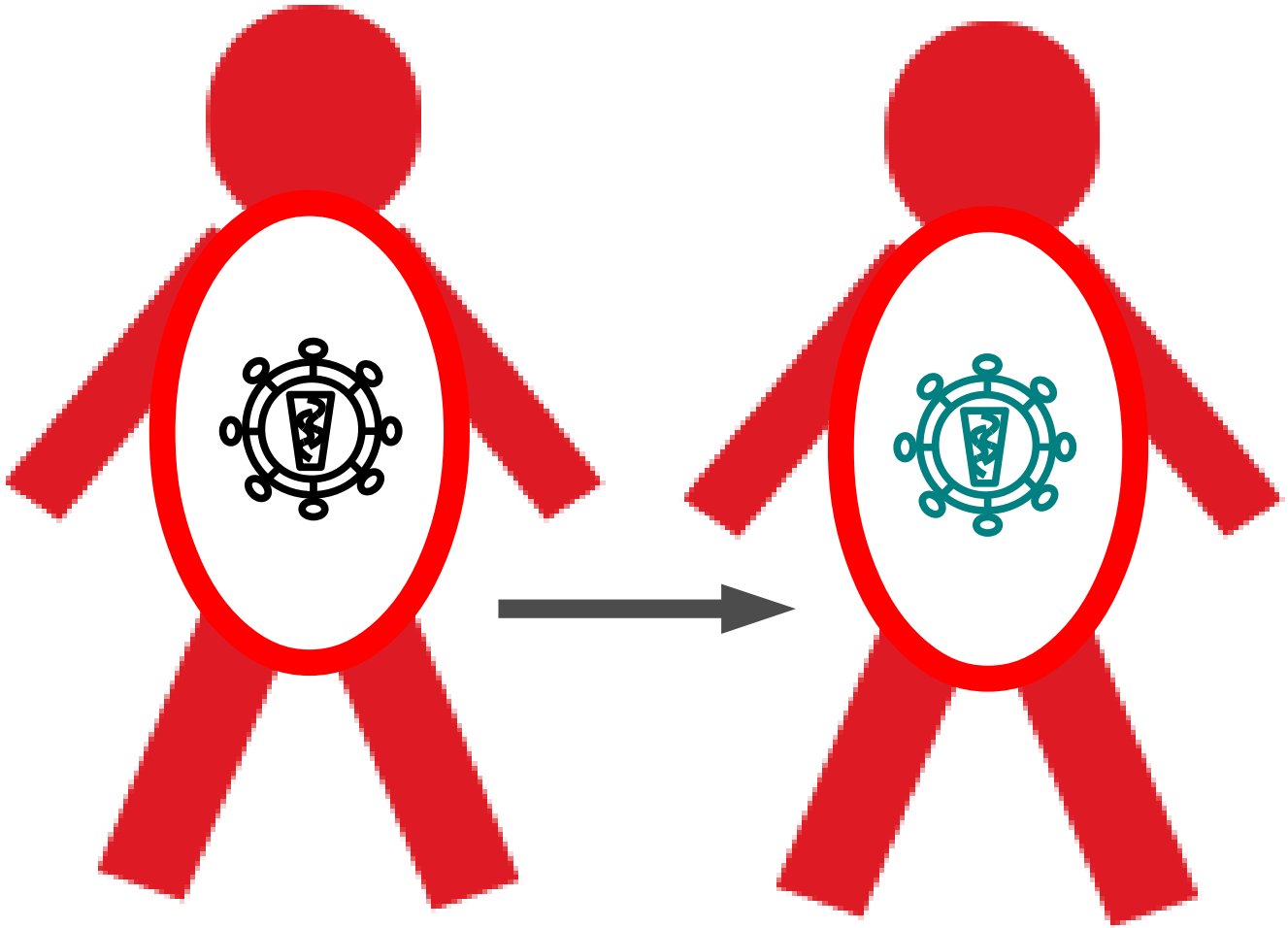
DATA: HIV



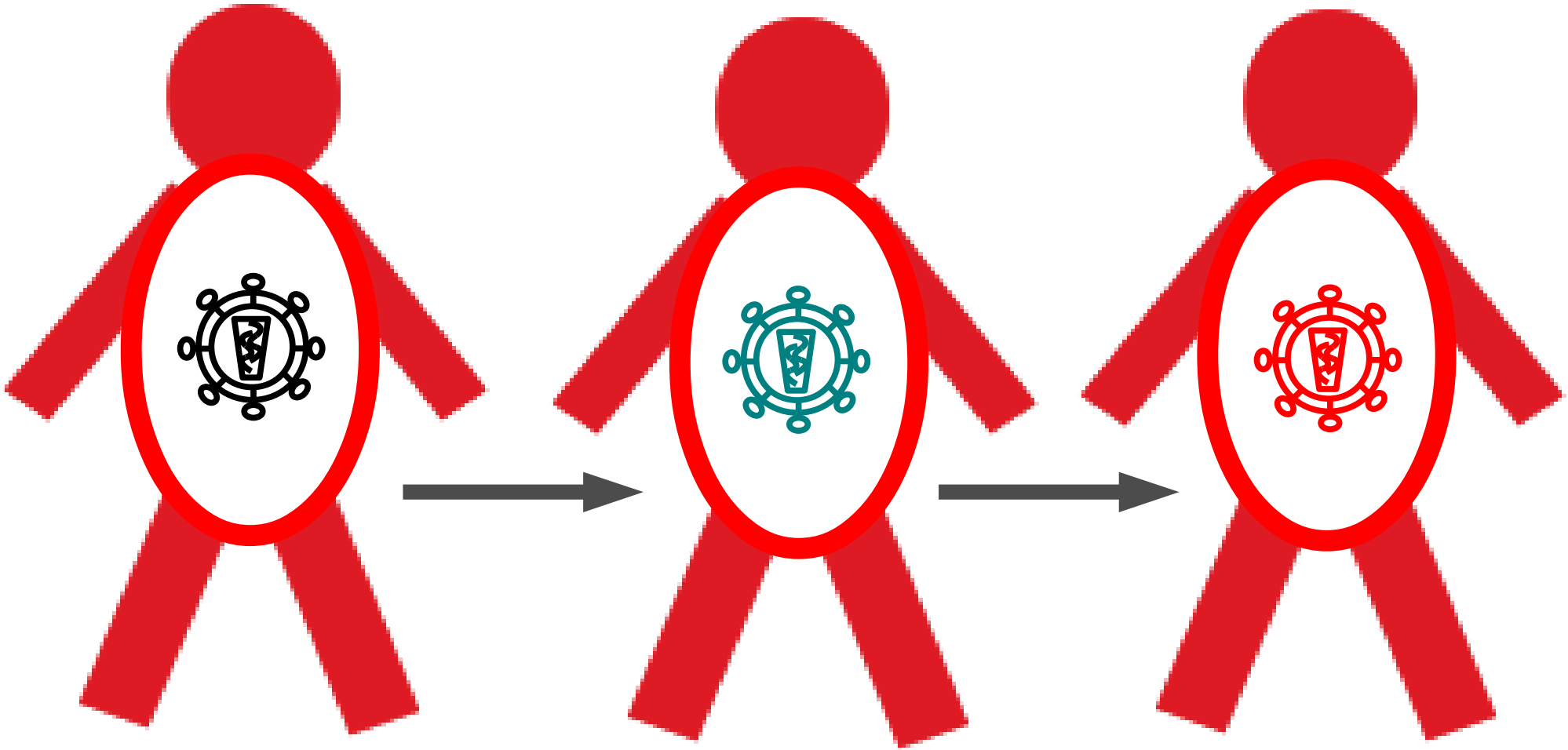
DATA: HIV



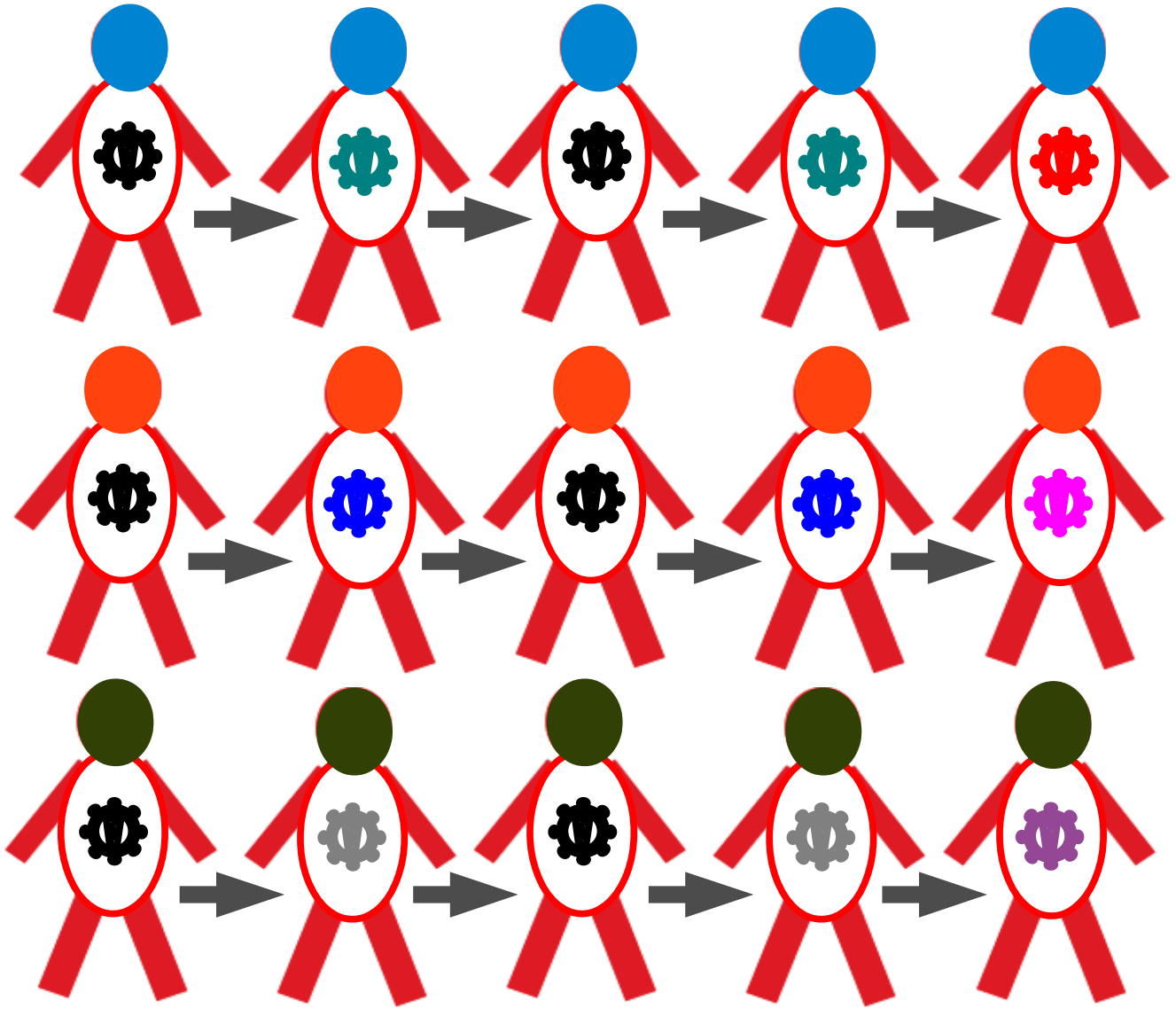
DATA: HIV



DATA: HIV



DATA: HIV



DATA: HIV

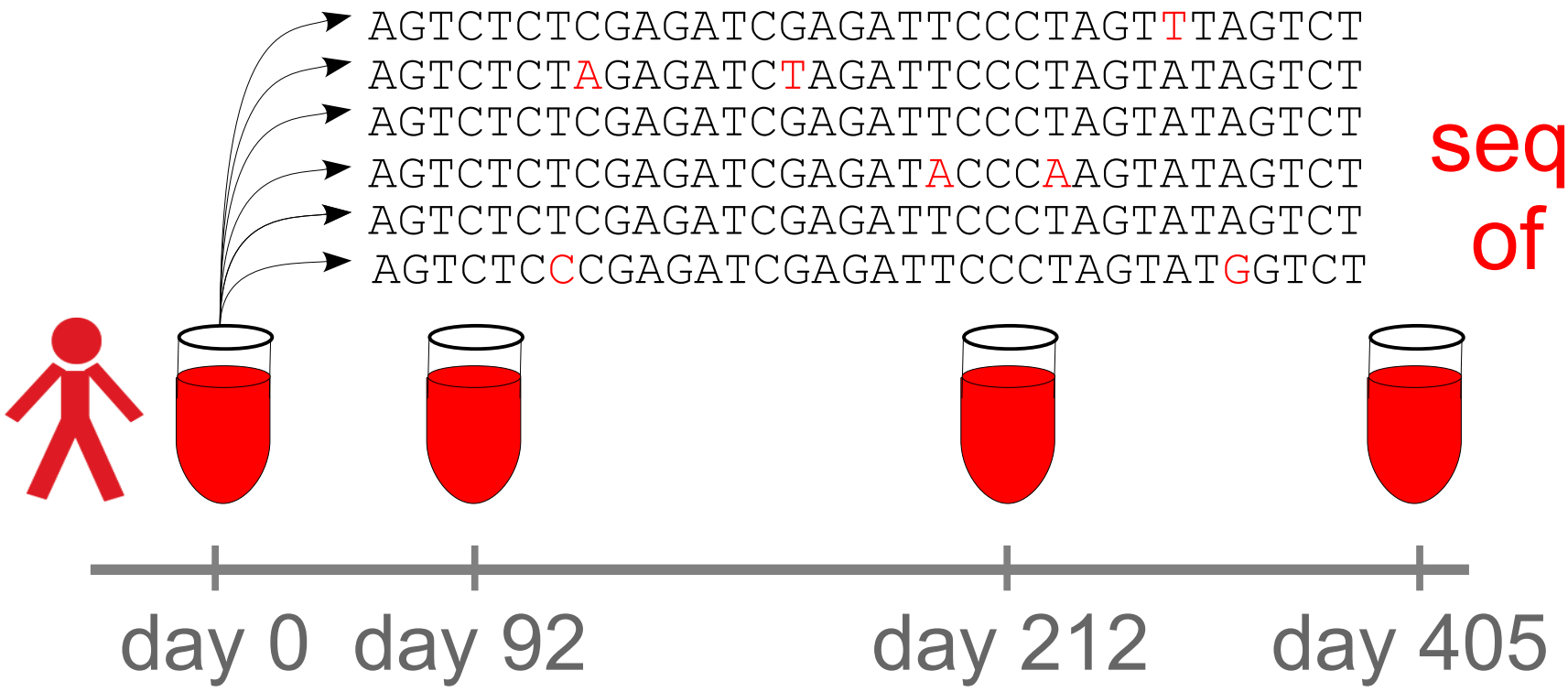
TCCCTAGT**A**TAGTCTCT susceptible virus

TCCCTAGT**T**TAGTCTCT resistant against drug X

HIV: Sampling is easy

Bacheler 2000 dataset

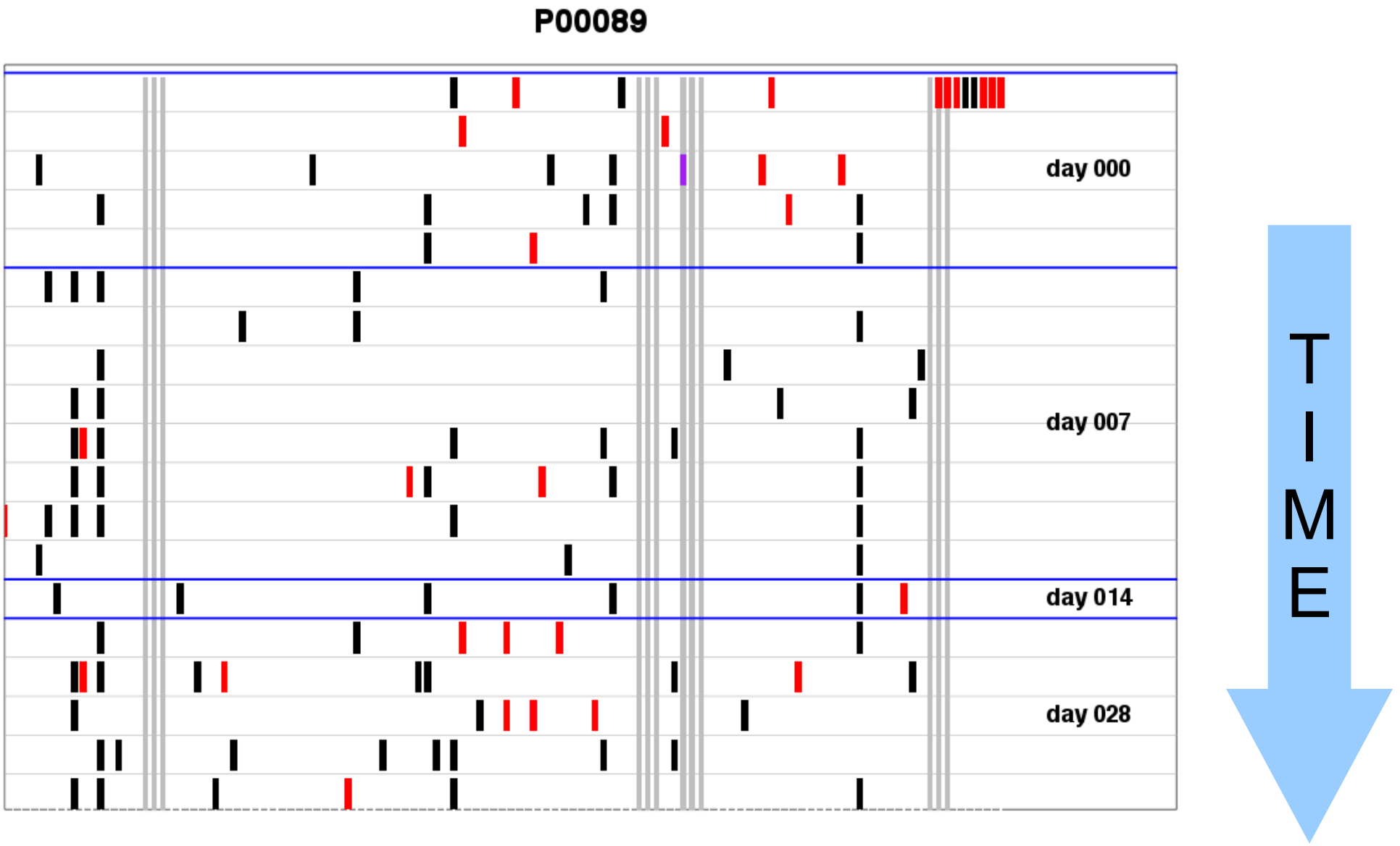
118 treated patients



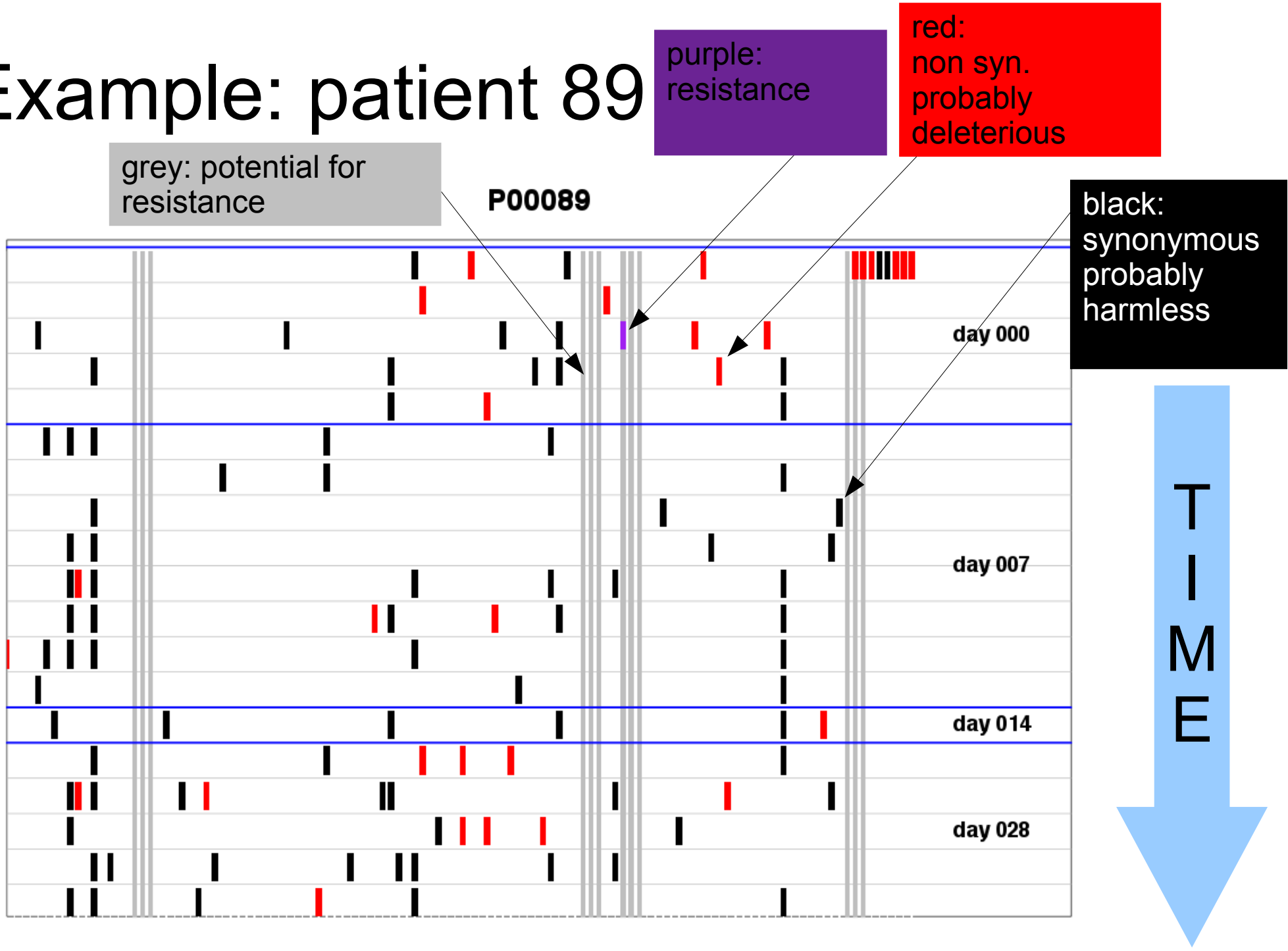
~ 6 sequences of 984 bp

~ 4 time points/patient

Example: patient 89

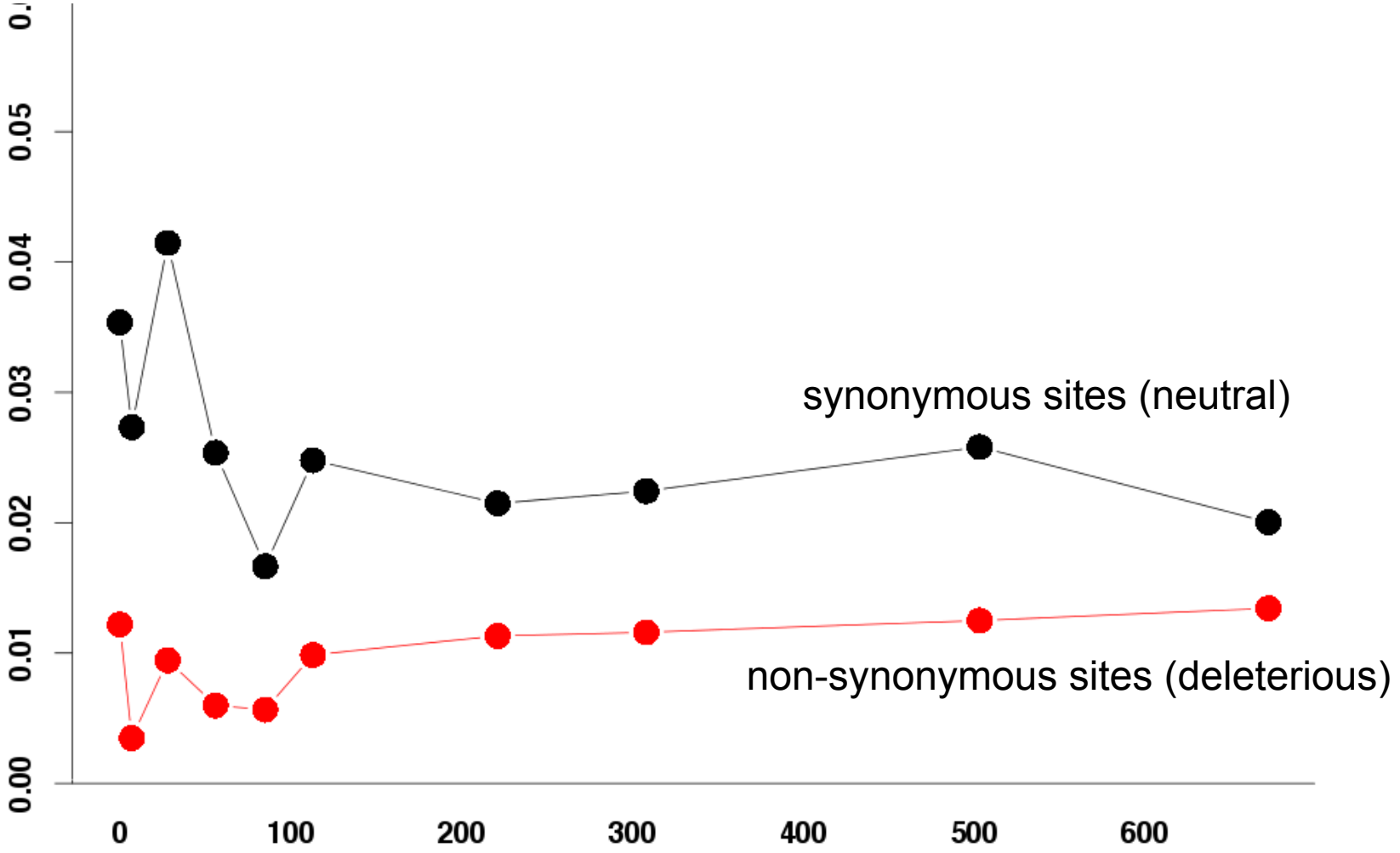


Example: patient 89



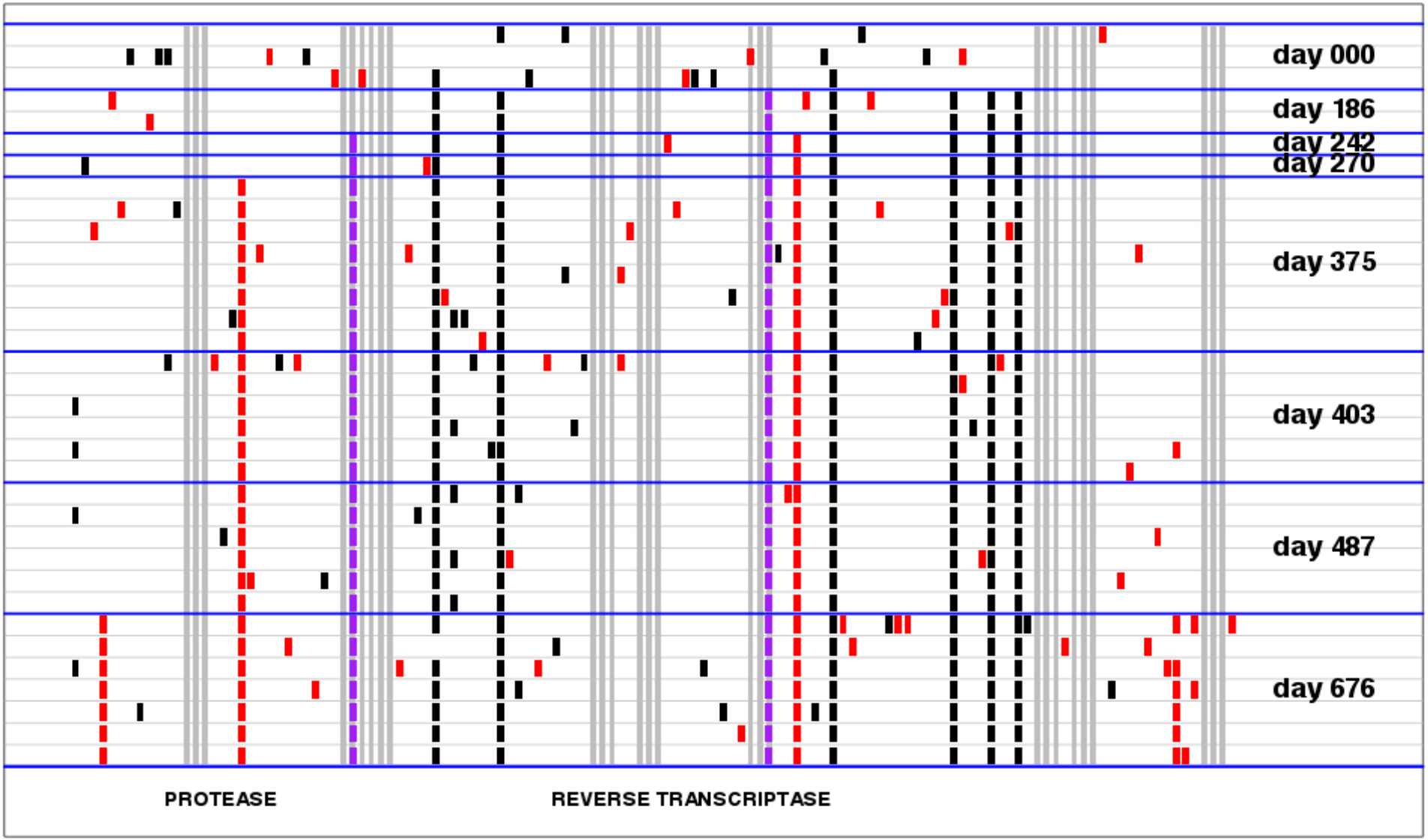
Example: patient 89

PI



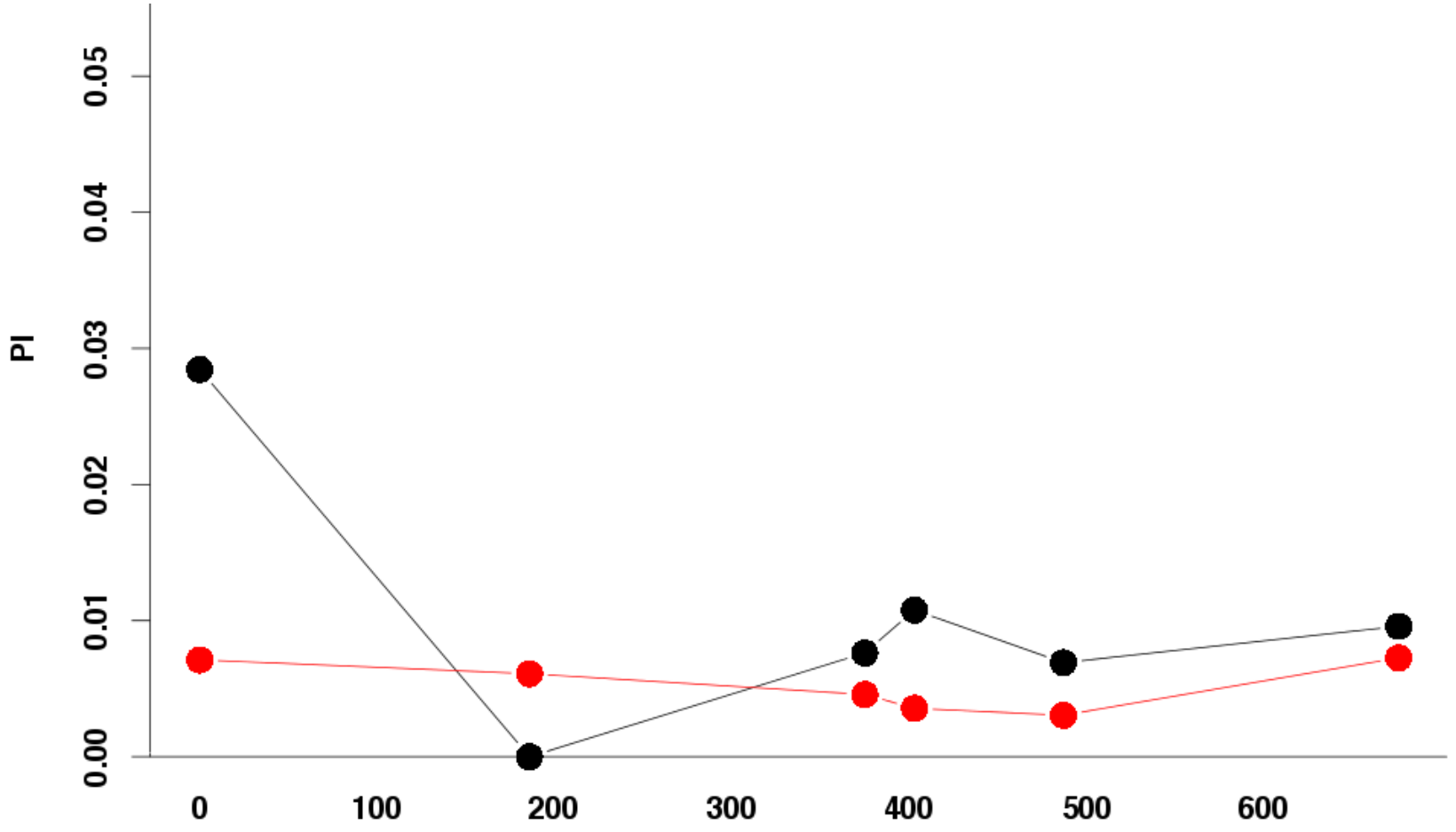
TIME

Example: patient 94



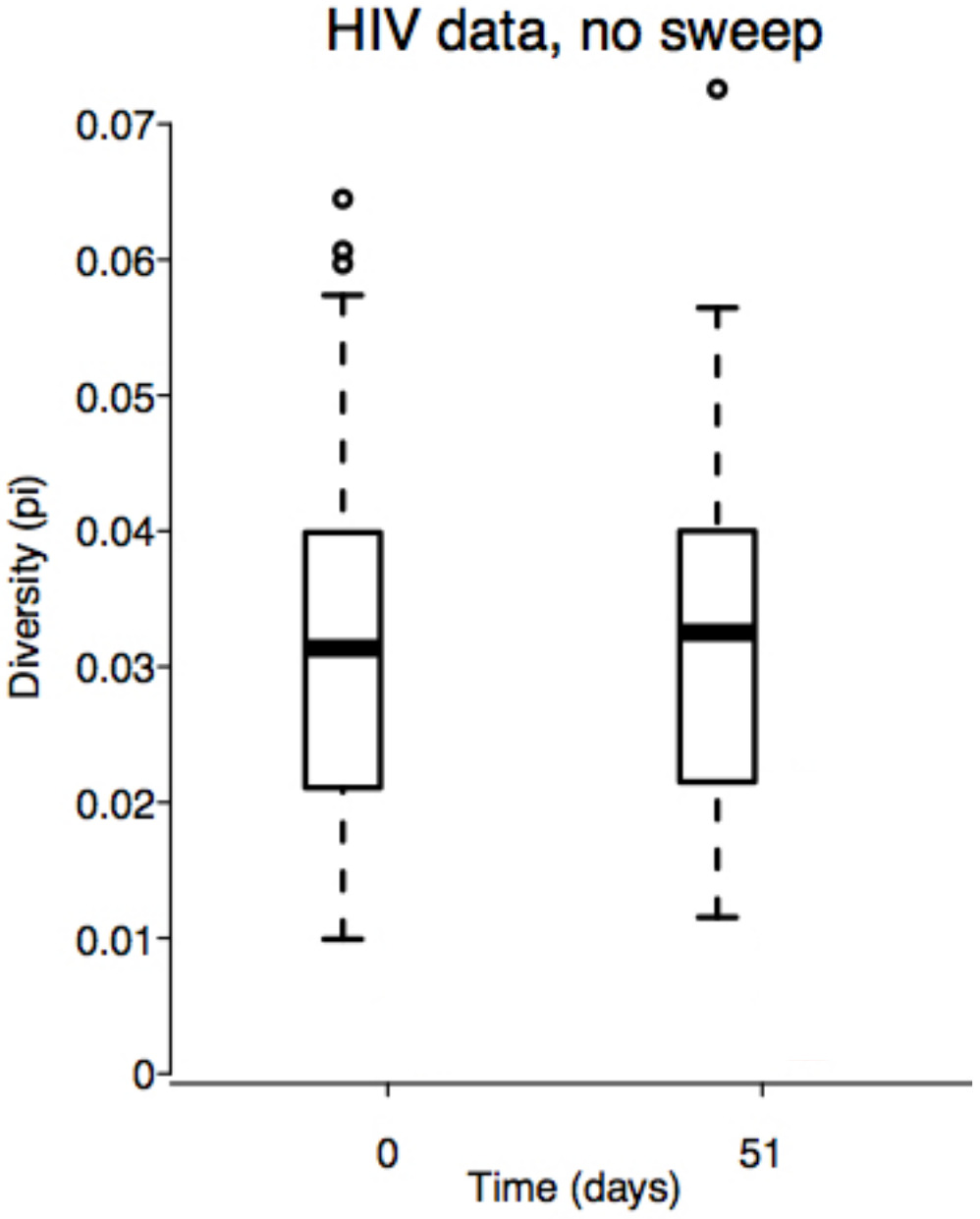
Example: patient 94

PI

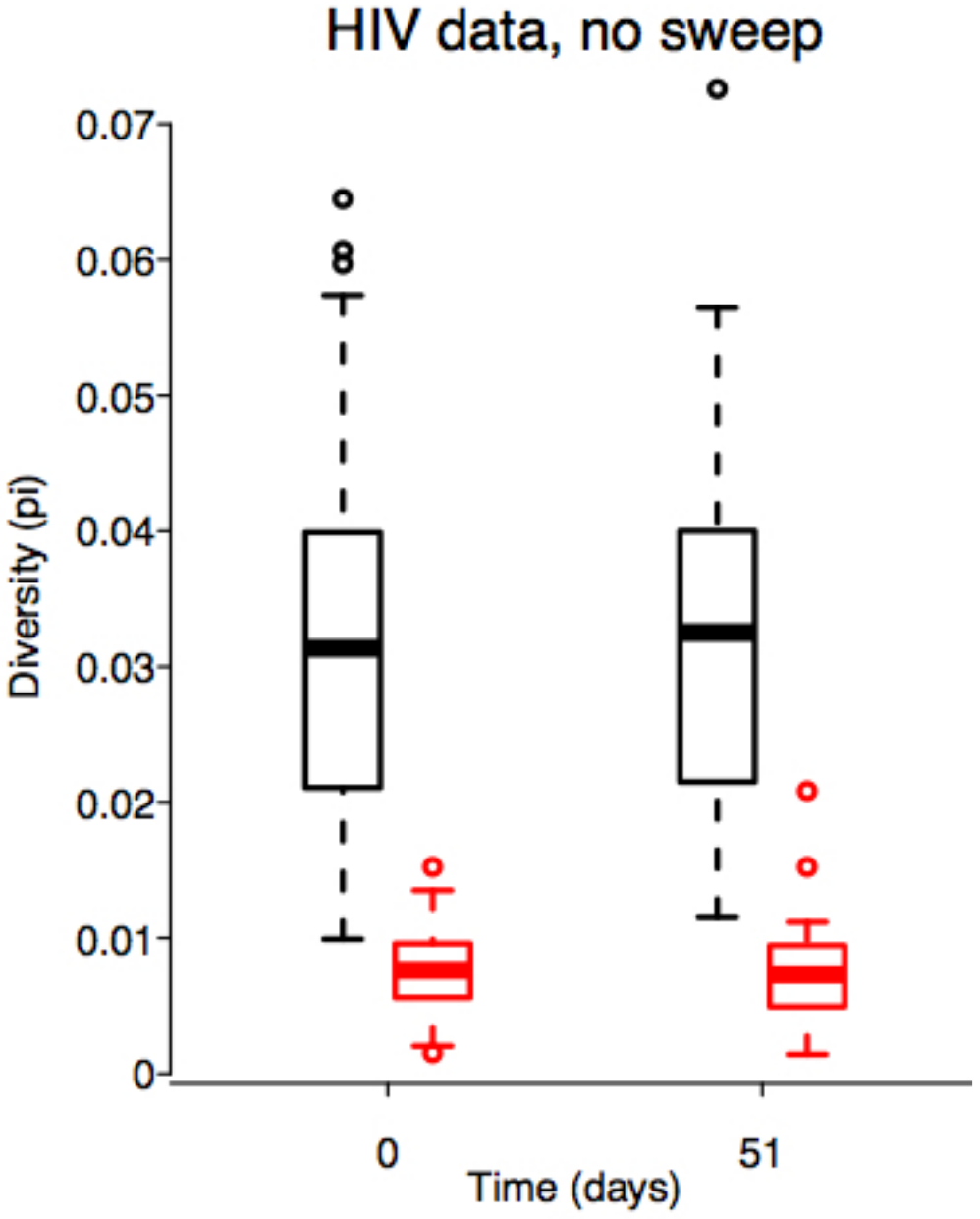


TIME

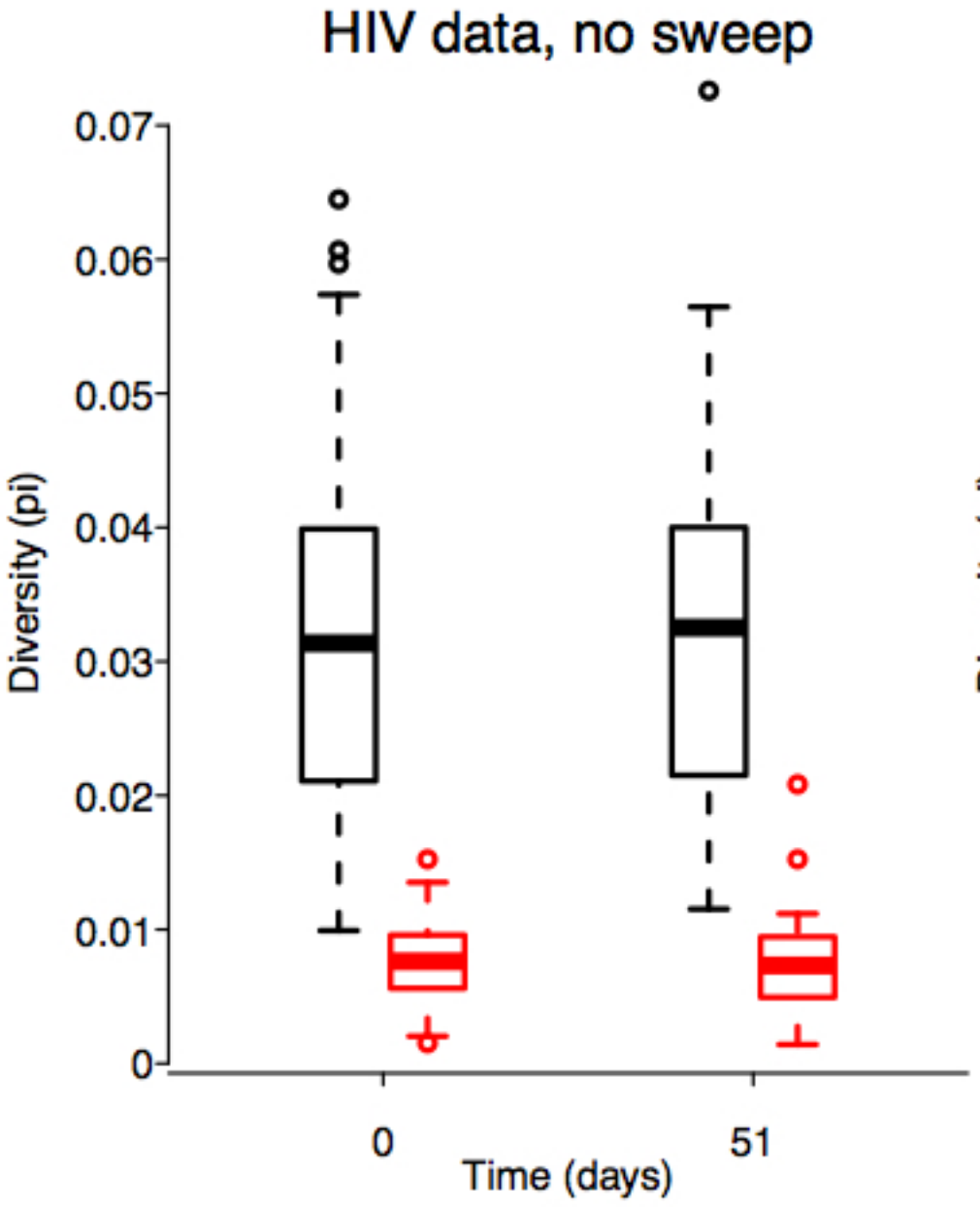
Effect of fixation on pi

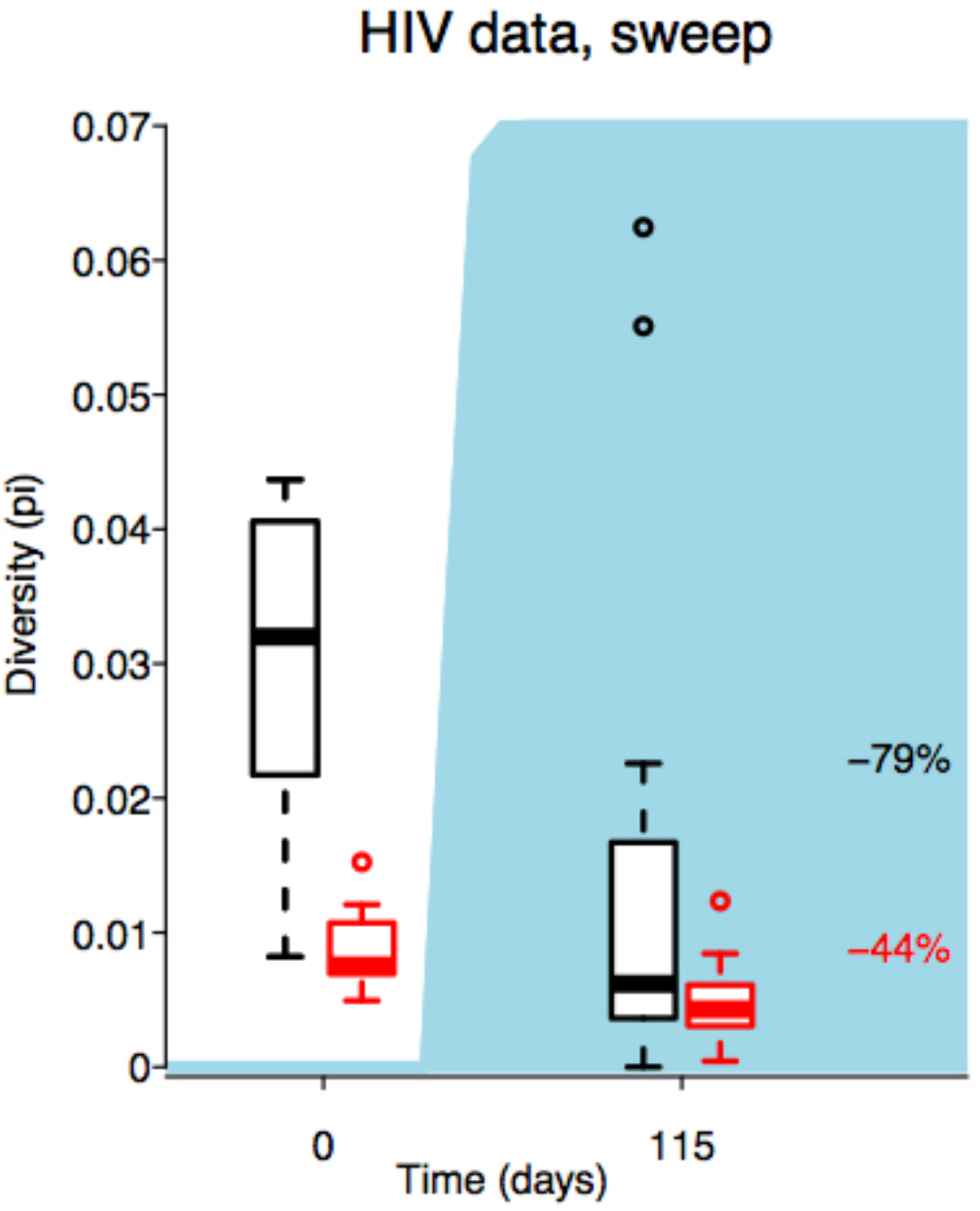


Effect of fixation on pi



Effect of fixation on pi





Fixation of resistance mutation reduces surrounding genetic divesity.

How does it recover?

Fixation of resistance mutation reduces surrounding genetic divesity.

How does it recover?

Only use patients with complete, hard sweep

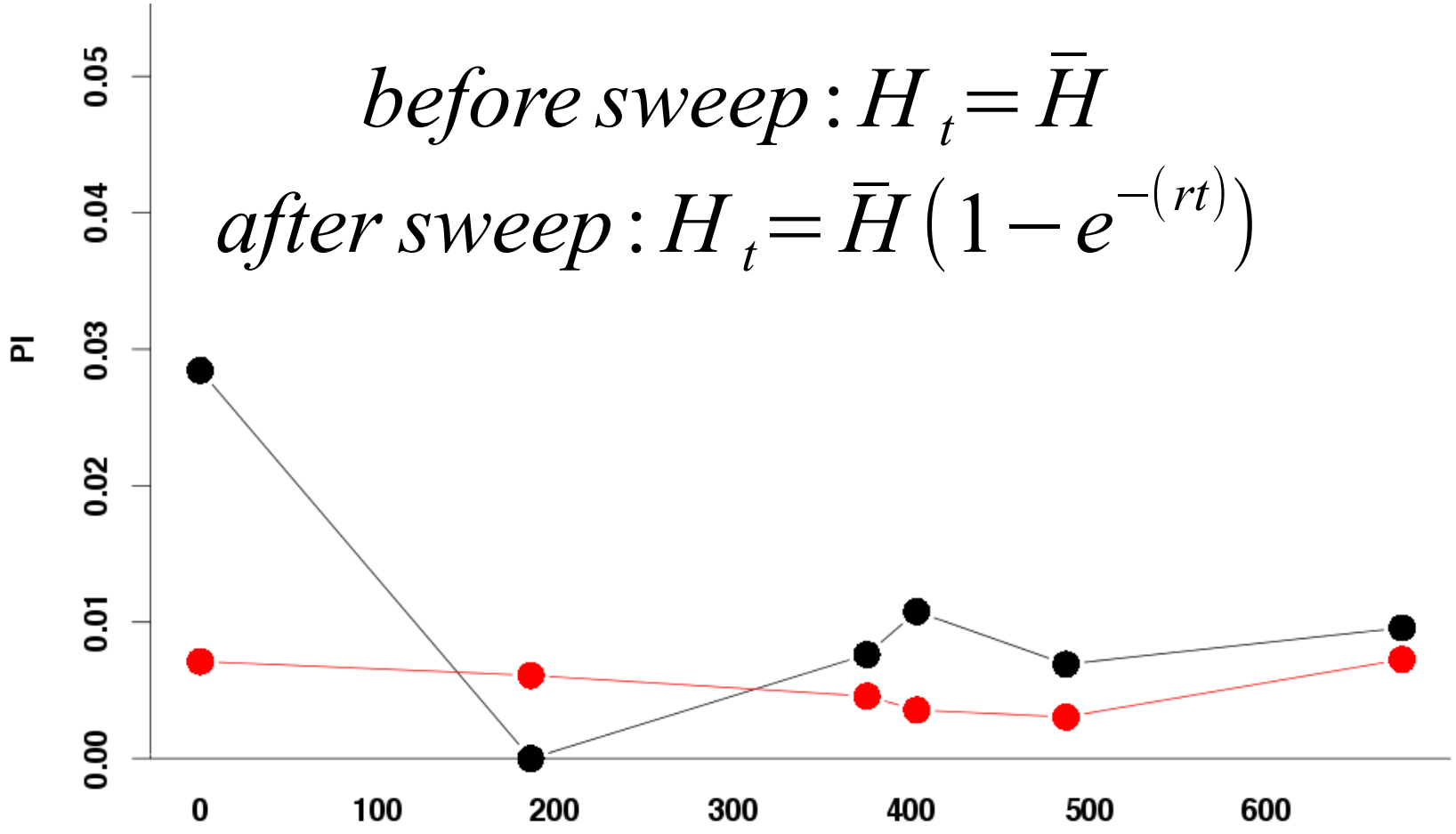
Fixation of resistance mutation reduces surrounding genetic divesity.

How does it recover?

PI

$$before\ sweep: H_t = \bar{H}$$

$$after\ sweep: H_t = \bar{H}\left(1 - e^{-(rt)}\right)$$



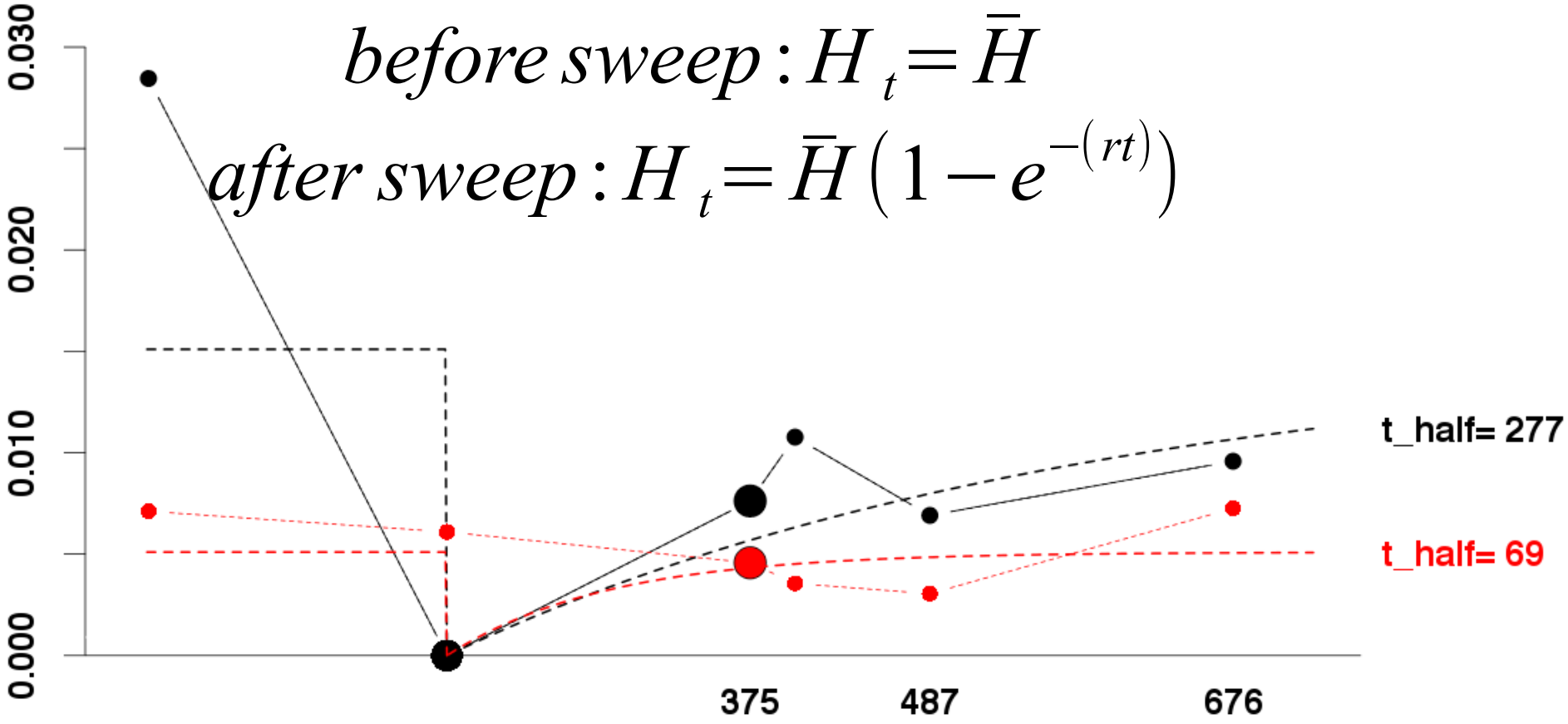
TIME

Fixation of resistance mutation reduces surrounding genetic divesity.

How does it recover?

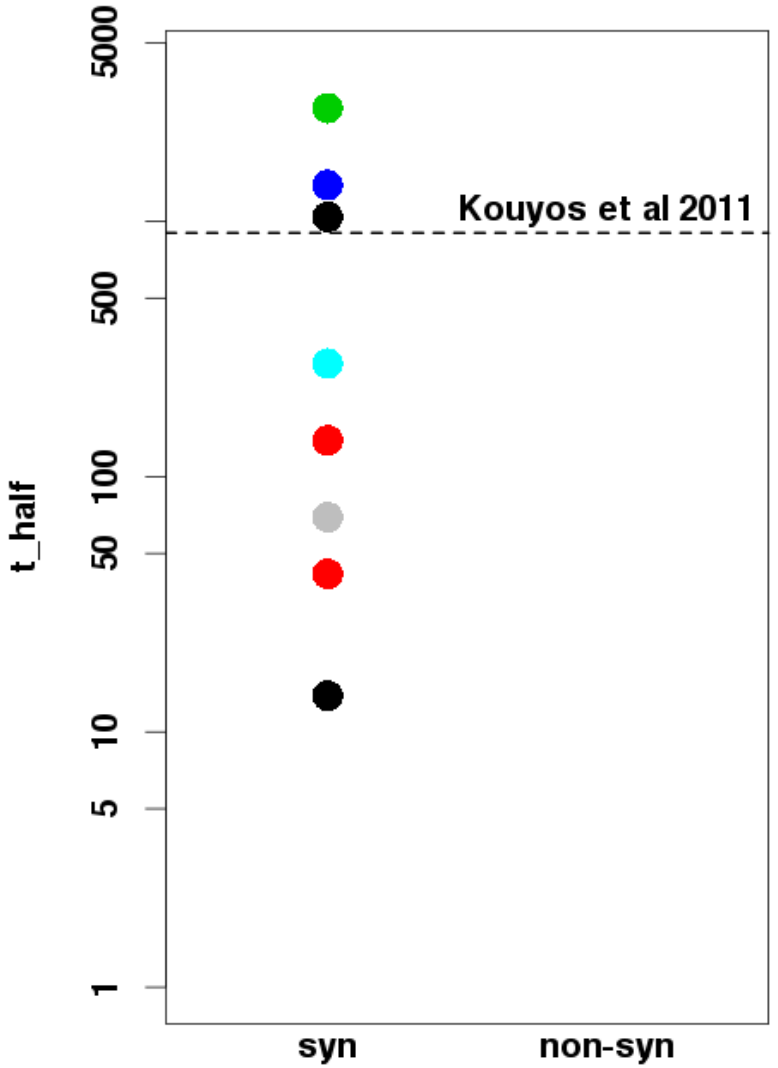
$$before\ sweep: H_t = \bar{H}$$

$$after\ sweep: H_t = \bar{H}\left(1 - e^{-(rt)}\right)$$



Estimated t_half values

Recovery is relatively fast



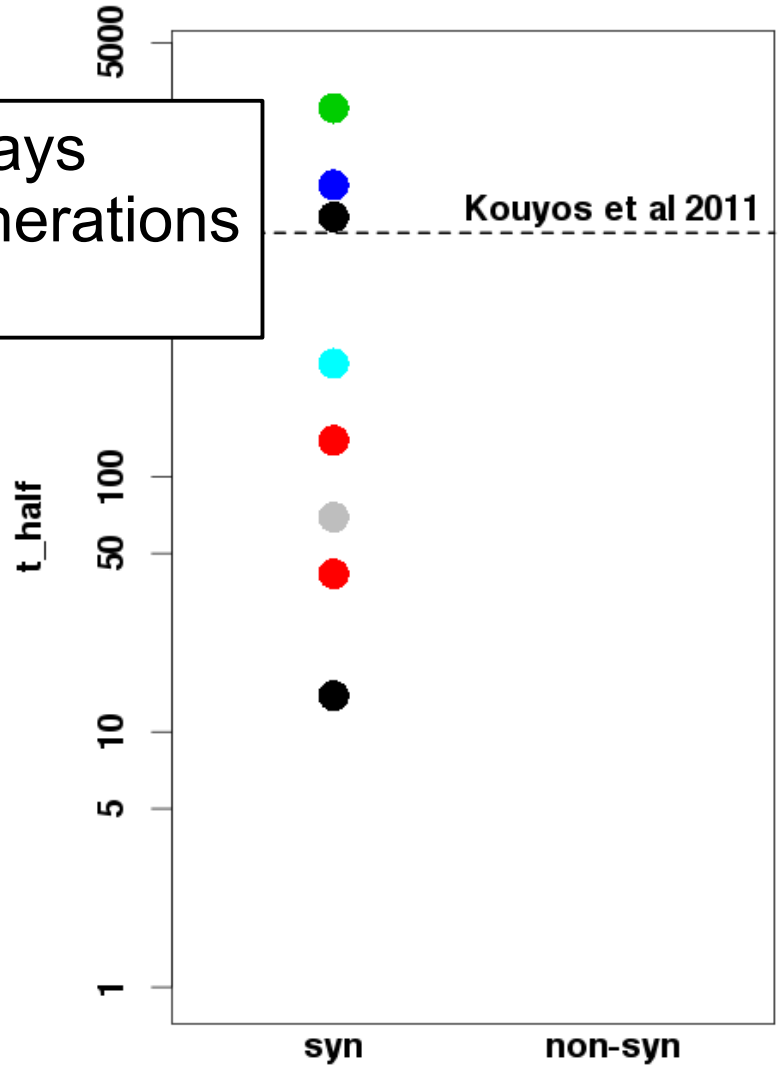
Estimated t_half values

Recovery is relatively fast

median t_half = 92 days
≈ 60 generations
→ Ne ≈ 90!!

H_pred ≈ 0.005

H_obs ≈ 0.025



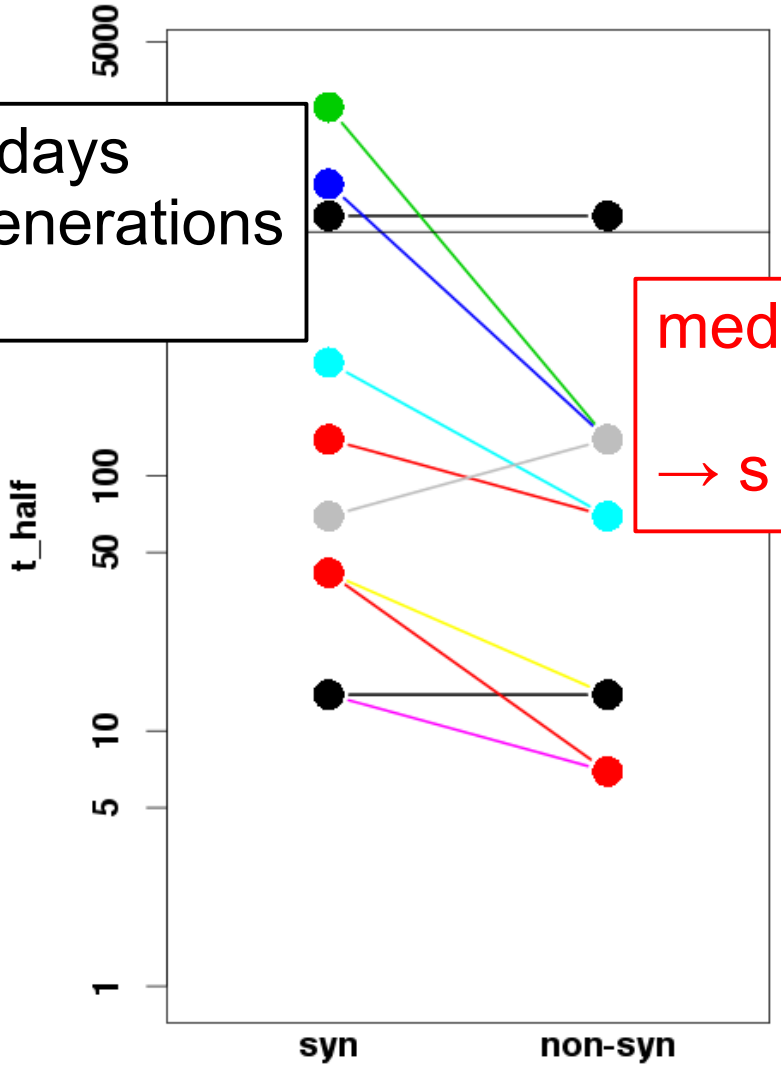
Estimated t_half values

Recovery is relatively fast

median t_half = 92 days
≈ 60 generations
→ Ne ≈ 90!!

H_pred ≈ 0.005

H_obs ≈ 0.025



Recovery is faster for non-syn mutations

median t_half = 70 days
≈ 45 generations
→ s ≈ 1.5%

H_pred ≈ 0.003

H_obs ≈ 0.01

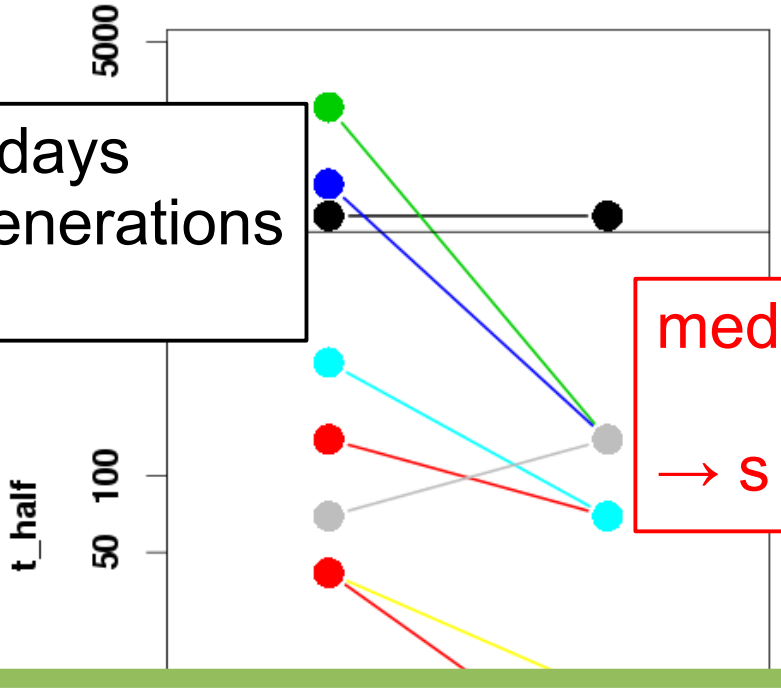
Estimated t_half values

Recovery is relatively fast

median t_half = 92 days
≈ 60 generations
→ Ne ≈ 90!!

H_pred ≈ 0.005

H_obs ≈ 0.025



Recovery is faster for non-syn mutations

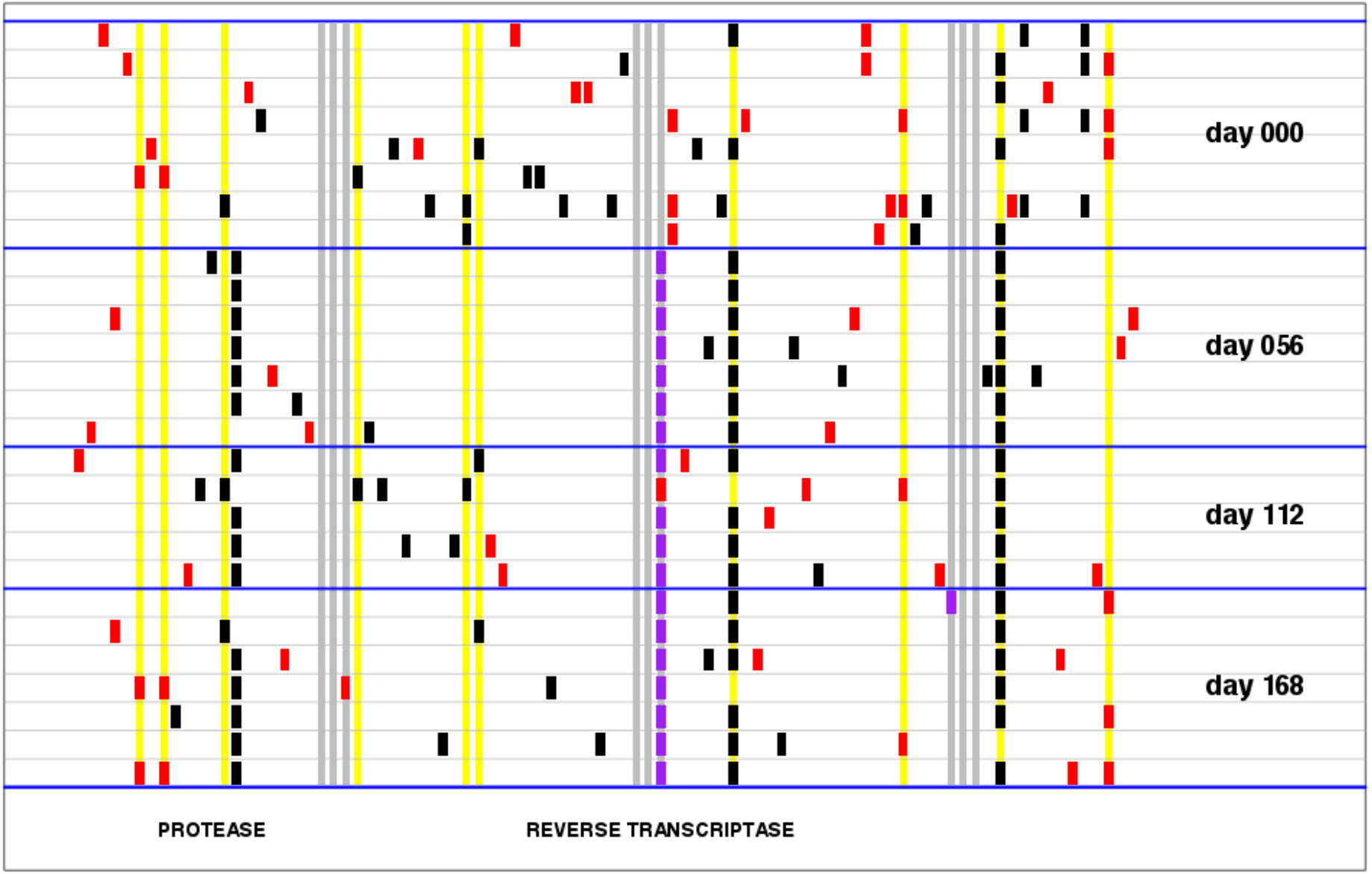
median t_half = 70 days
≈ 45 generations
→ s ≈ 1.5%

H_pred ≈ 0.003

H_obs ≈ 0.01

Times too fast & N estimates too low?
Mutation rate higher?
Generation time faster?
Migration?

Migration?


day 000
day 056
day 112
day 168
PROTEASE
REVERSE TRANSCRIPTASE

Conclusions

- Recovery can be predicted & observed

Conclusions

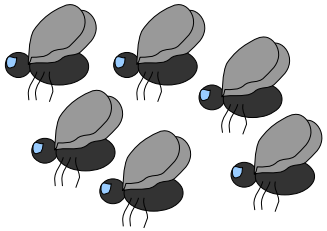
- Recovery can be predicted & observed
- Predictions and observation: faster for non-syn sites
  - ratio nonsyn:syn polymorphism will be high at first and go down during recovery
  - this can explain why European humans have high ratio nonsyn:syn polymorphism (Lohmueller et al 2008)

Conclusions

- Recovery can be predicted & observed
- Predictions and observation: faster for non-syn sites
  - ratio nonsyn:syn polymorphism will be high at first and go down during recovery
  - this can explain why European humans have high ratio nonsyn:syn polymorphism (Lohmueller et al 2008)
- Recovery in HIV data is variable, and mostly fast
- Parameter estimates inconsistent w. steady state values
- Due to wrong estimates/assumptions for mutation rate, generation time, migration?

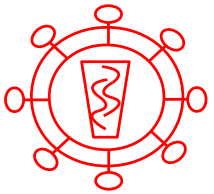
Examples

Drosophila melanogaster
never recovers
Karasov, Messer, Petrov 2010



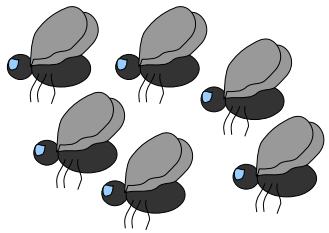
HIV
few years to recover
Kouyos et al 2011

HIV
days - weeks to recover
Pennings 2012



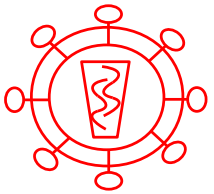
Examples

Drosophila melanogaster
never recovers
Karasov, Messer, Petrov 2010



HIV
few years to recover
Kouyos et al 2011

HIV
days - weeks to recover
Pennings 2012

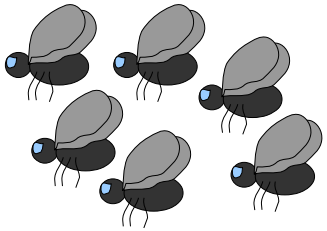


HIV
in treated patients
after sweep
few months to recover
This talk

bottleneck of size 1,
large N
neutral sites

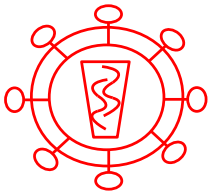
larger bottleneck,
resistance mutations

Drosophila melanogaster
never recovers
Karasov, Messer, Petrov 2010

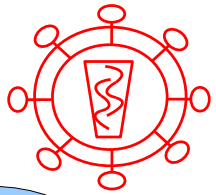


HIV
few years to recover
Kouyos et al 2011

HIV
days - weeks to recover
Pennings 2012



HIV
in treated patients
after sweep
few months to recover
This talk

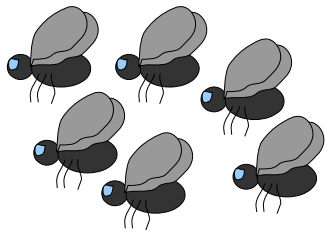


smaller N
neutral and delet. muts
better data

bottleneck of size 1,
large N

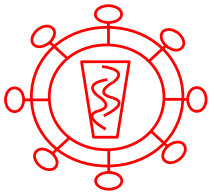
larger bottleneck,
deleterious mutations

Drosophila melanogaster
never recovers
Karasov, Messer, Petrov 2010

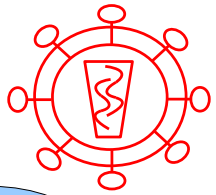


HIV
few years to recover
Kouyos et al 2011

HIV
days - weeks to recover
Pennings 2012



HIV
in treated patients
after sweep
few months to recover
This talk



smaller N?
better data

Questions?
Suggestions?
Thank you!

funded by HFSP