

effective theory for immune-pathogen interaction

Armita Nourmohammad

Princeton University

February 9th, 2016

acknowledgements

Jakub Otwinowski (UPenn)

Josh Plotkin (UPenn)

Luca Peliti (IAS)

Bill Bialek (Princeton)

Alan Perelson (LANL)

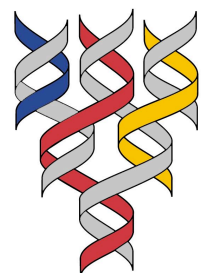
Marta Łuksza (IAS)



Michael Lässig (Cologne)

Stephan Schiffels (Sanger)

Torsten Held (Cologne)

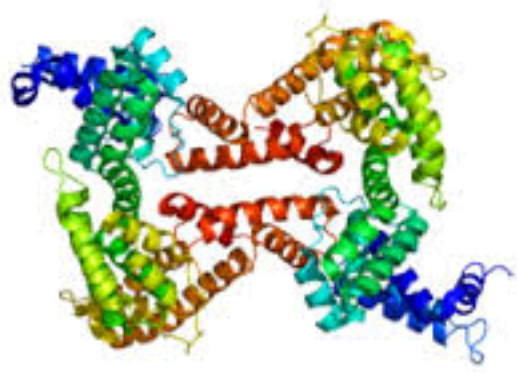


SFB 680
Molecular Basis of
Evolutionary Innovations



James S. McDonnell Foundation



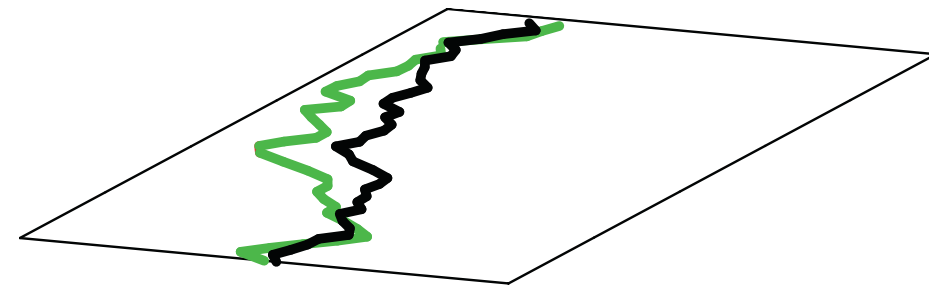
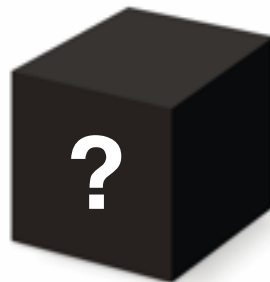


evolution of molecular phenotypes
in **time-dependent** environments

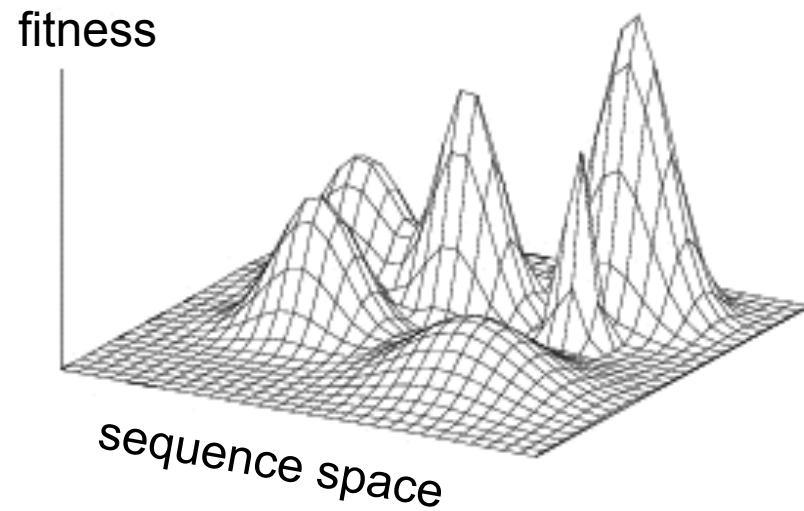
environment



population

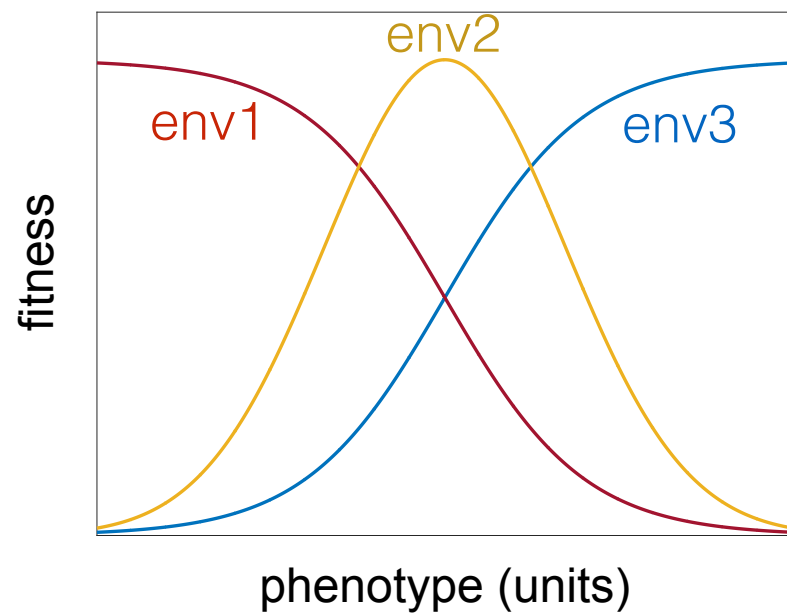


genotypes vs. phenotypes



→ genotypes

- high dimensional, sparsely sampled
- low predictability (degenerate)



→ phenotypes

- closer relation to fitness
- lower dimensional theories (ignoring degeneracy)
- difficult to identify

$$\frac{\partial}{\partial t} P[\rho(E), t] = ?$$

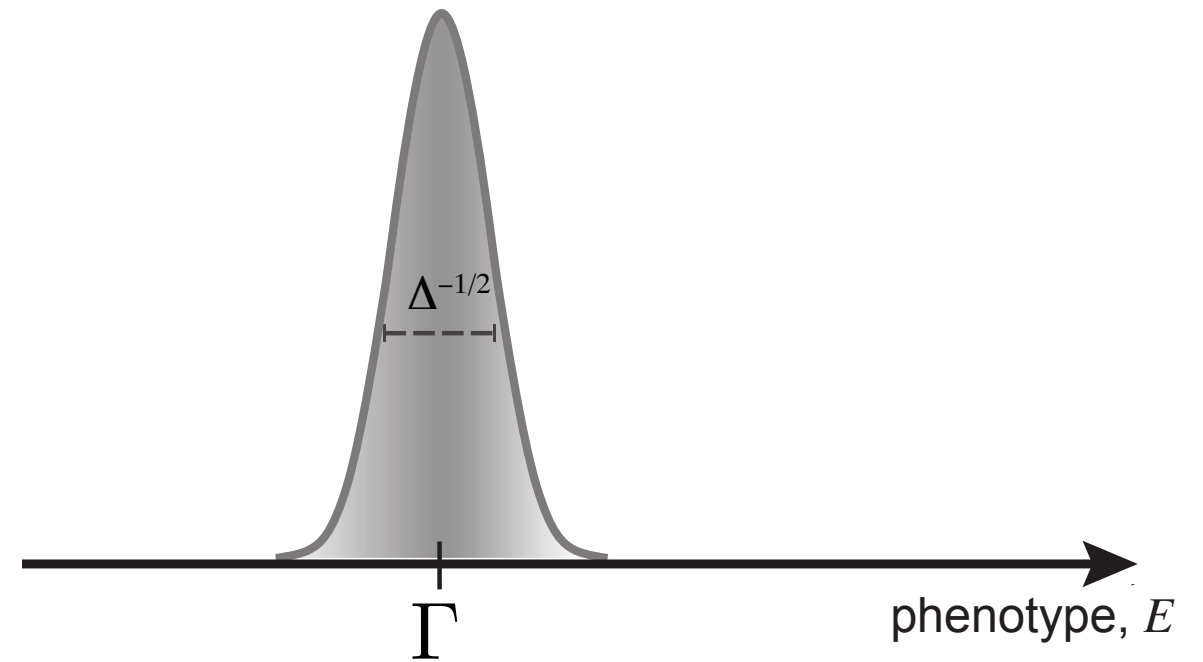
statistical physics of phenotype evolution (equilibrium)

- ▶ characterizing the **intra-population** phenotype distribution by its moments

trait mean: $\Gamma = \overline{E}$

trait diversity: $\Delta = \overline{(E - \Gamma)^2}$

inter-population (ensemble) statistics: $Q(\Gamma, \Delta, t)$



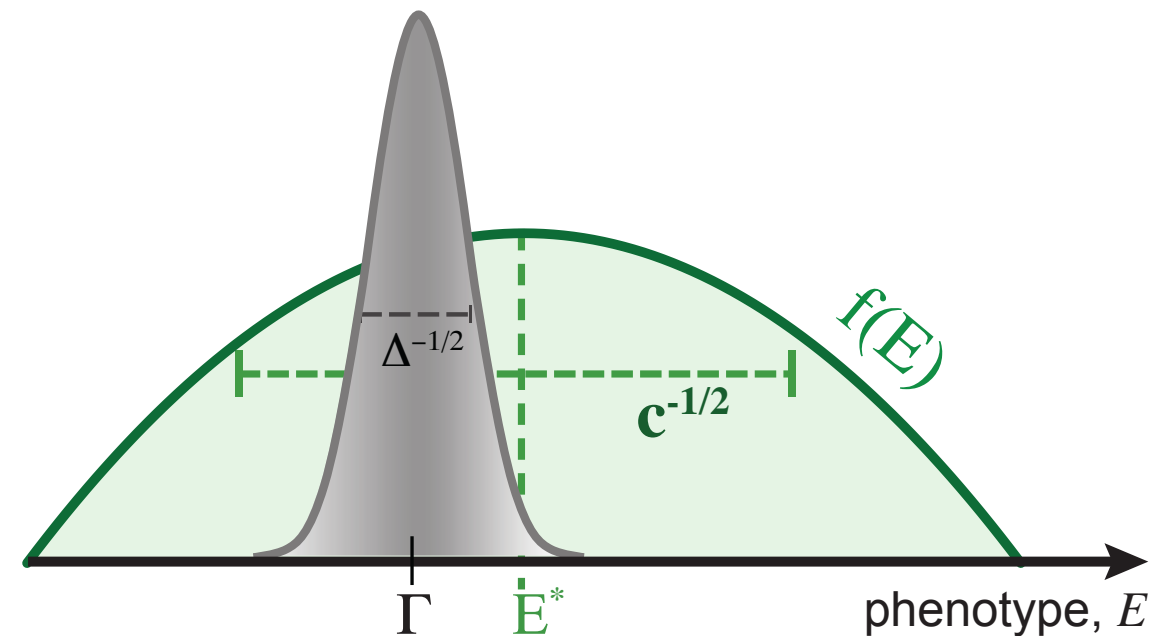
statistical physics of phenotype evolution (equilibrium)

- ▶ characterizing the **intra-population** phenotype distribution by its moments

trait mean: $\Gamma = \overline{E}$

trait diversity: $\Delta = \overline{(E - \Gamma)^2}$

inter-population (ensemble) statistics: $Q(\Gamma, \Delta, t)$



- ➔ **stochastic evolution of trait statistics**

$$\frac{d}{dt}\Gamma = \underbrace{-2\mu(\Gamma - \Gamma_0)}_{\text{mutation}} + \underbrace{\Delta \partial_{\Gamma} F(\Gamma)}_{\text{selection}} + \underbrace{\sqrt{\frac{\Delta}{N}} \chi_{\Gamma}}_{\text{stochasticity genetic drift}}$$

similar for $Q(\Delta, t)$

- ➔ **Boltzmann-type equilibrium distributions**

$$Q_{\text{eq}}(\Gamma) = \frac{1}{Z} \tilde{Q}_0(\Gamma) e^{2N\tilde{F}(\Gamma)}$$

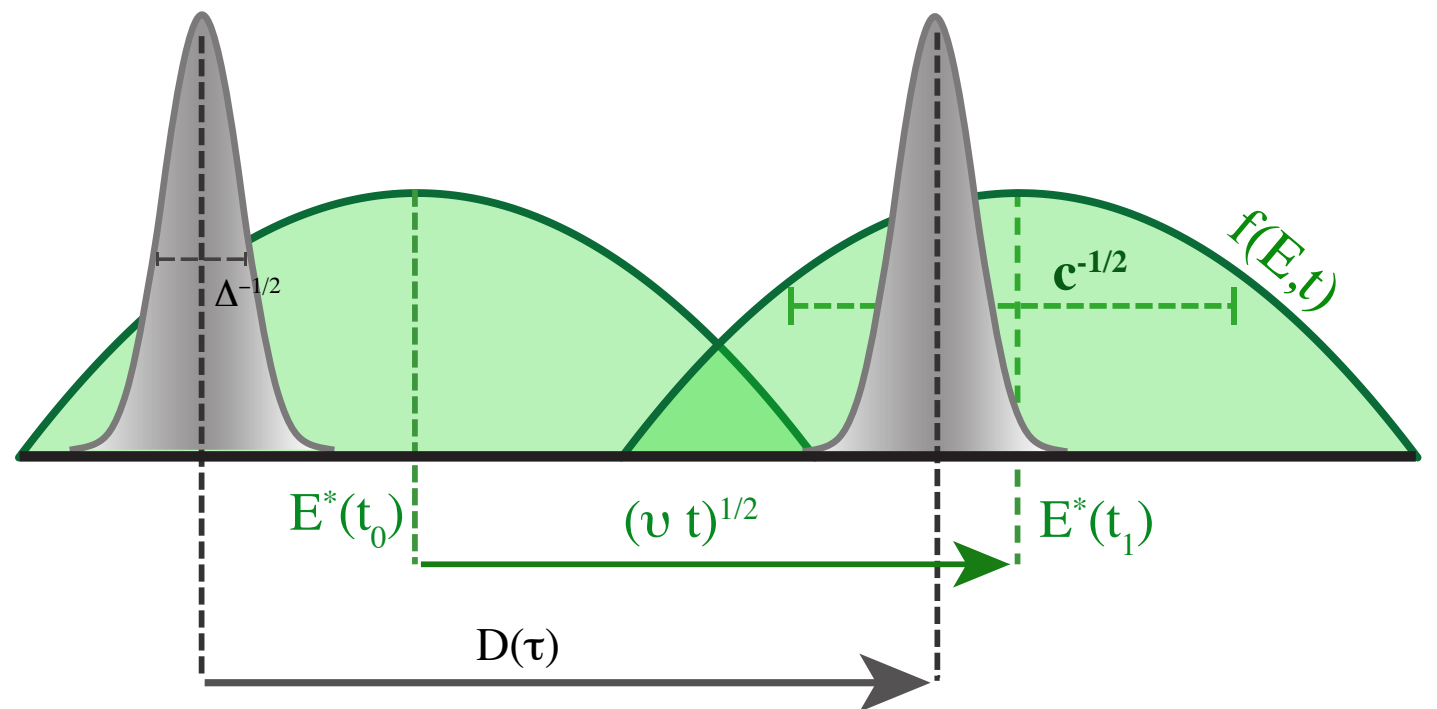
molecular traits in time-dependent environments

▶ adaptive evolution with stabilizing selection (non-equilibrium)

▶ time-dependent trait divergence:

$$D(\tau) = \langle (\Gamma_\tau - \Gamma_0)^2 \rangle$$

- environmental driving rate: \mathcal{U}
- stabilizing strength: \mathcal{C}



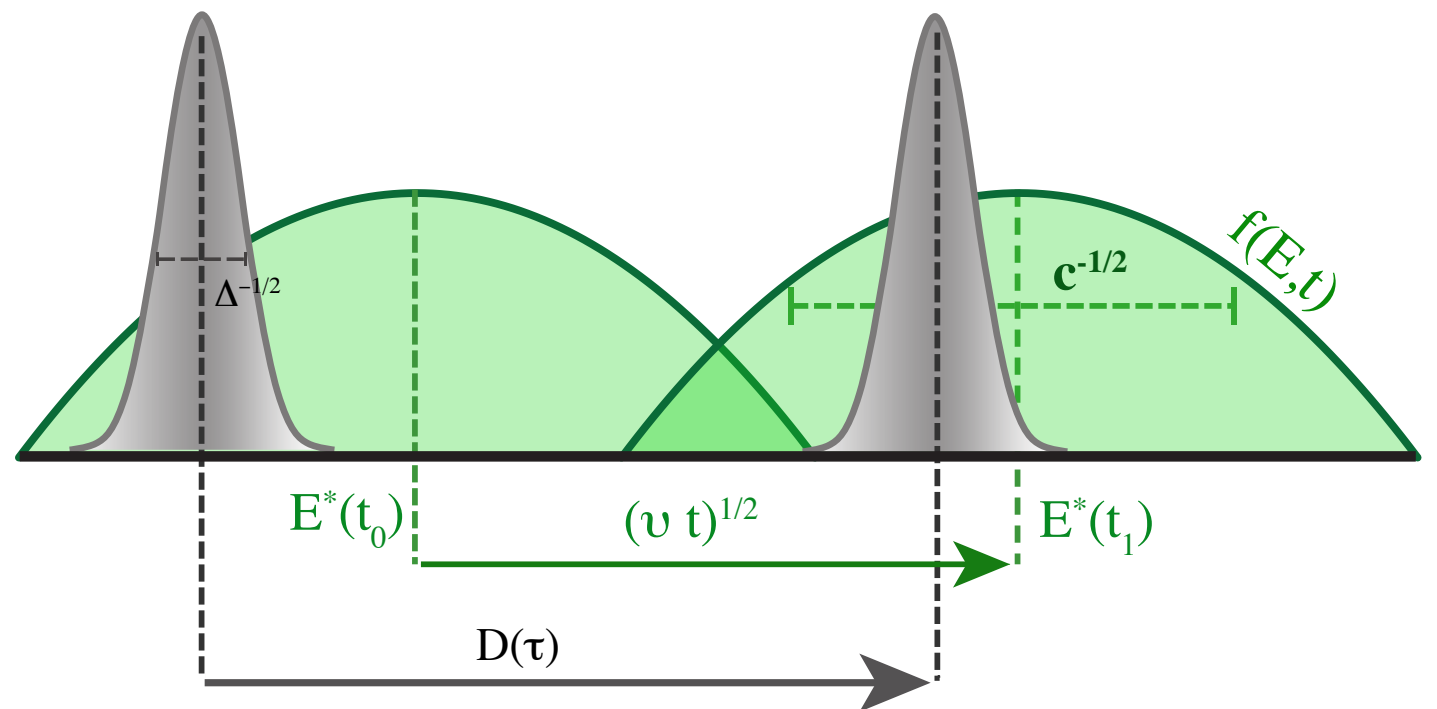
molecular traits in time-dependent environments

▶ adaptive evolution with stabilizing selection (non-equilibrium)

▶ time-dependent trait divergence:

$$D(\tau) = \langle (\Gamma_\tau - \Gamma_0)^2 \rangle$$

- environmental driving rate: \mathcal{U}
- stabilizing strength: \mathcal{C}



▶ fitness peak follows an Ornstein-Uhlenbeck process on macroevolutionary time-scales

➔ **joint** stochastic dynamics of fitness and population $Q(\Gamma, E^*, t)$

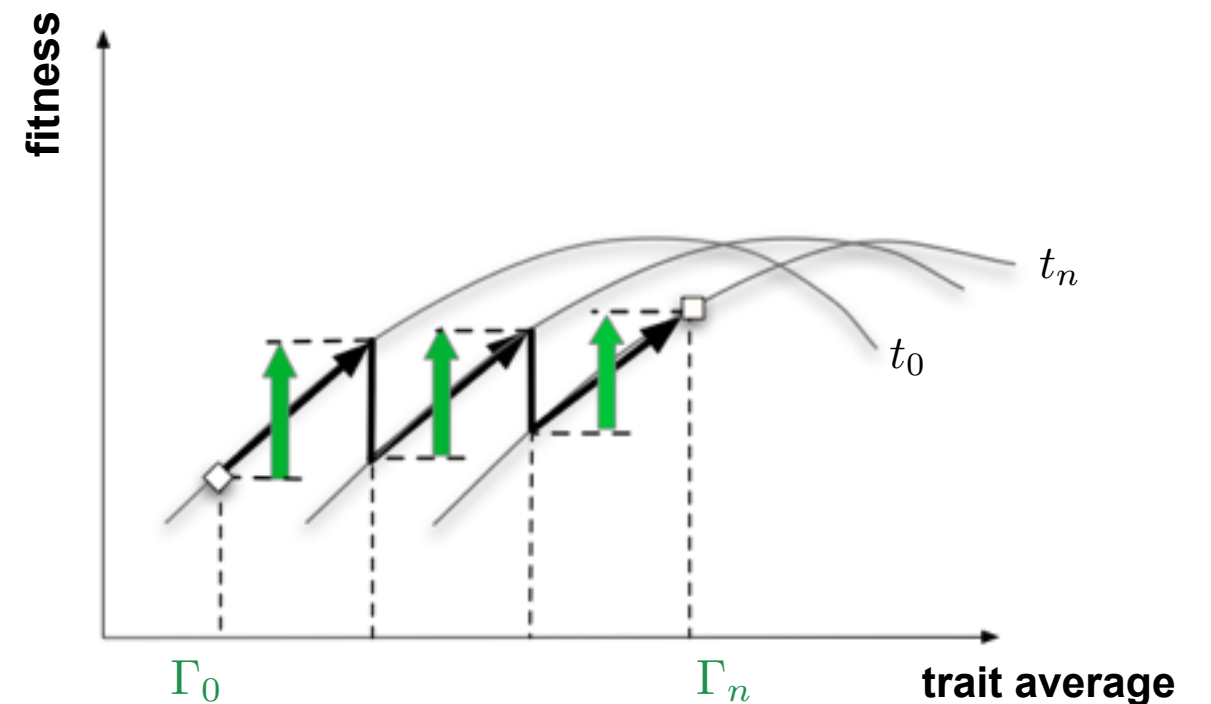
fitness flux: measure of adaptation

- ▶ **population history** is a sequence of trait measurements

$$(\Gamma_0, \dots, \Gamma_n) \quad \text{measured over time: } (t_0, \dots, t_n)$$

- ▶ **fitness flux** of a population history is the cumulative **selective effect of trait changes**.

$$\Phi_n = \sum_{i=1}^n \delta\Gamma_i \nabla F(\Gamma_i, t_i) \neq F(\Gamma_n, t_n) - F(\Gamma_0, t_0)$$



fitness flux: measure of adaptation

- ▶ **population history** is a sequence of trait measurements

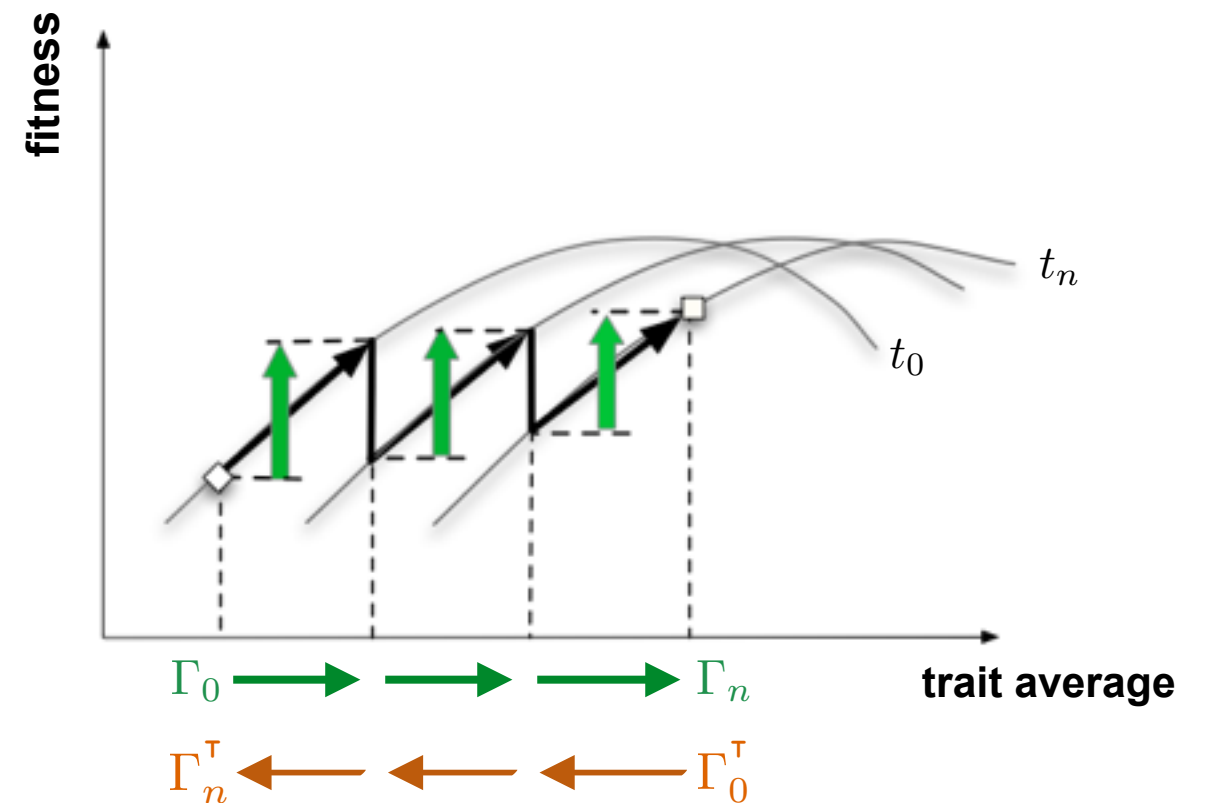
$$(\Gamma_0, \dots, \Gamma_n) \quad \text{measured over time: } (t_0, \dots, t_n)$$

- ▶ **fitness flux** of a population history is the cumulative **selective effect of trait changes**.

$$\Phi_n = \sum_{i=1}^n \delta\Gamma_i \nabla F(\Gamma_i, t_i) \neq F(\Gamma_n, t_n) - F(\Gamma_0, t_0)$$

- ▶ fitness flux measures the **deviation from equilibrium**

$$\langle 2N\Phi \rangle = D_{KL}(\mathcal{P} || \mathcal{P}^\top) + \text{boundary terms}$$



fitness flux: measure of adaptation

- ▶ **population history** is a sequence of trait measurements

$$(\Gamma_0, \dots, \Gamma_n) \quad \text{measured over time: } (t_0, \dots, t_n)$$

- ▶ **fitness flux** of a population history is the cumulative **selective effect of trait changes**.

$$\Phi_n = \sum_{i=1}^n \delta\Gamma_i \nabla F(\Gamma_i, t_i) \neq F(\Gamma_n, t_n) - F(\Gamma_0, t_0)$$

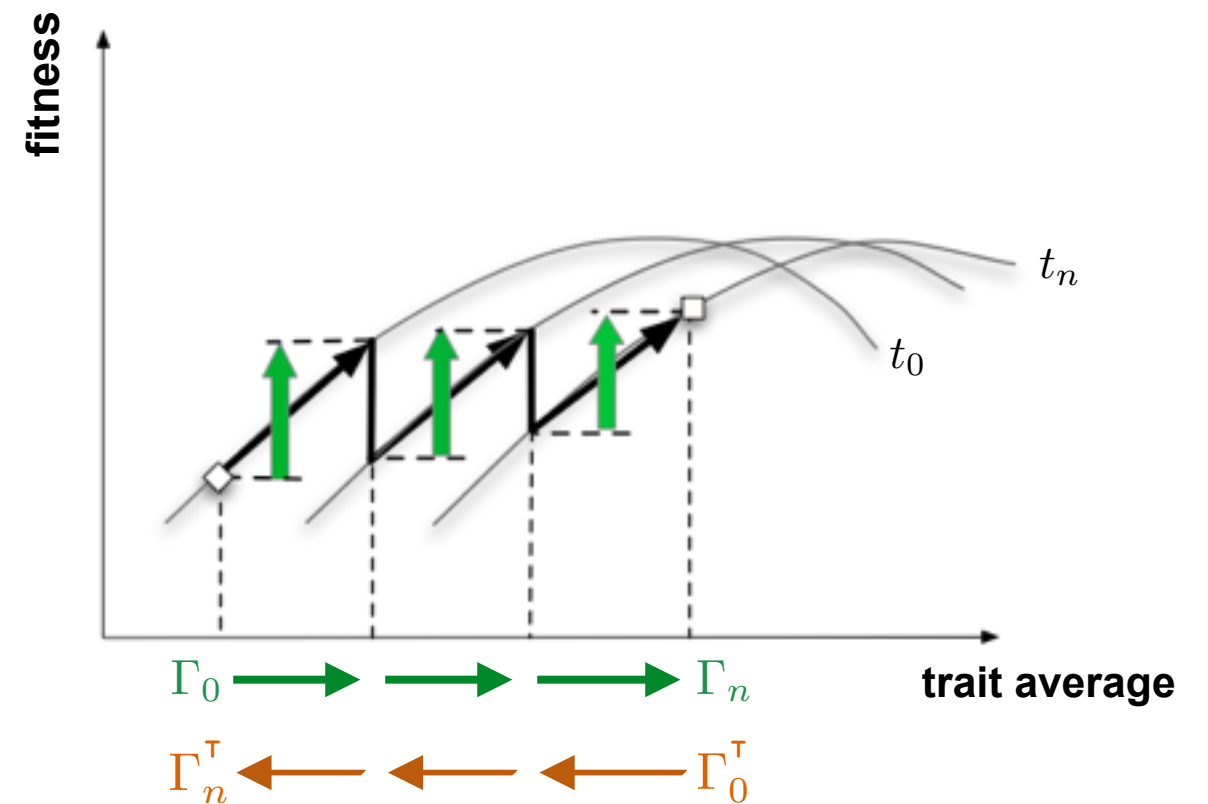
- ▶ fitness flux measures the **deviation from equilibrium**

$$\langle 2N\Phi \rangle = D_{KL}(\mathcal{P} || \mathcal{P}^\top) + \text{boundary terms}$$

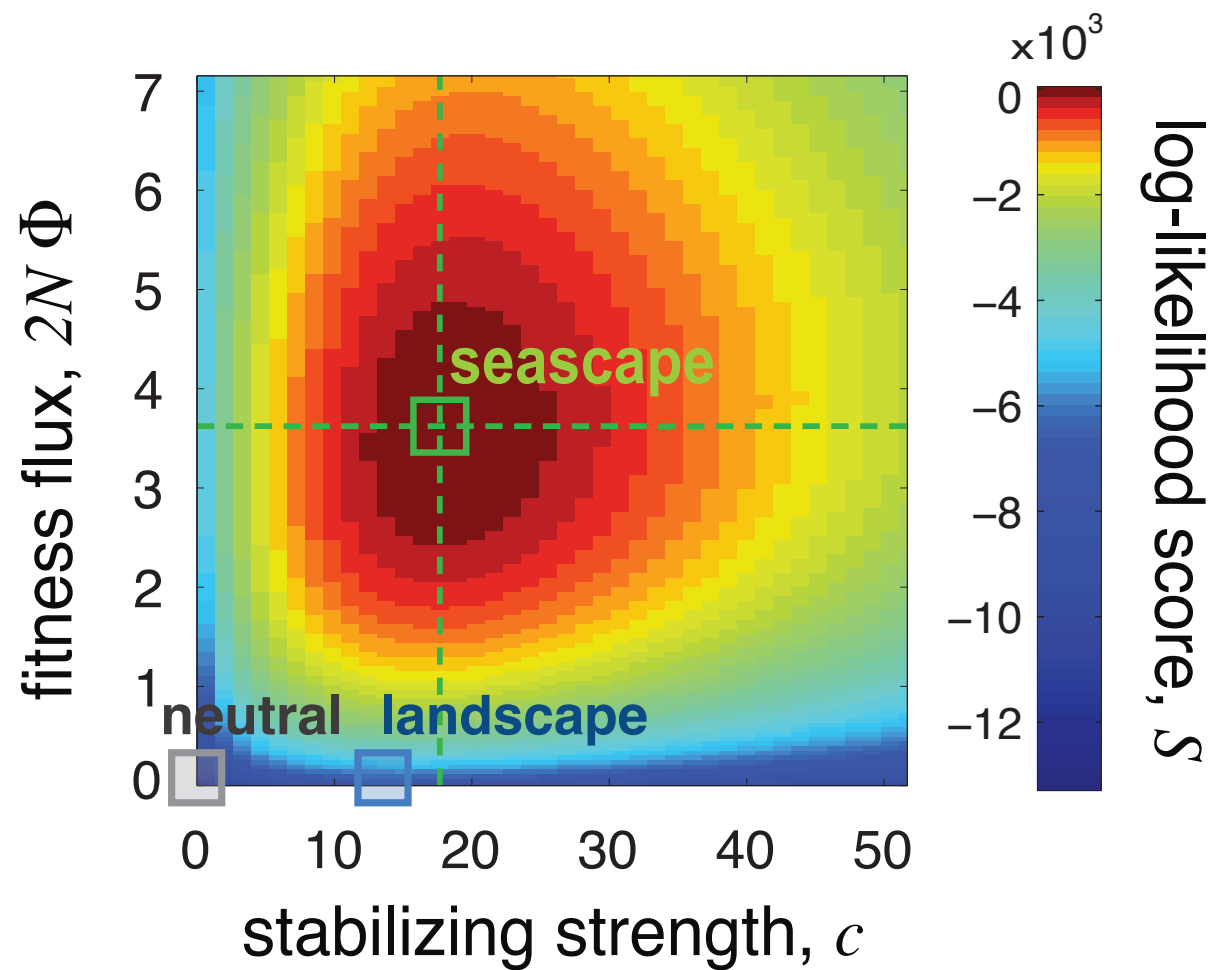
- ▶ fitness flux in stochastic **stabilizing seascape**

fitness flux \sim **curvature** x **driving rate**

$$\langle 2N\Phi(\tau) \rangle \simeq c v \tau$$



gene expression adaptation in *Drosophila* (probabilistic inference)



adaptive classes:

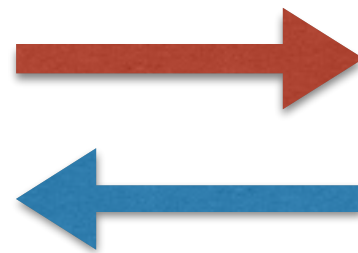
- morphology
- response to stimulus
- cognition
- stress response
- growth regulation
- positive regulation of TFs
- mating behavior

► **Adaptation:** 63% of divergence, and 54% of the genes



co-evolution of molecular phenotypes in interacting populations

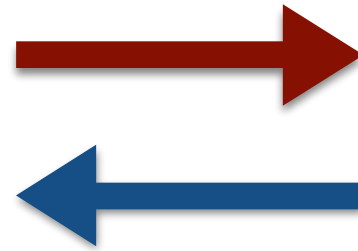
population A



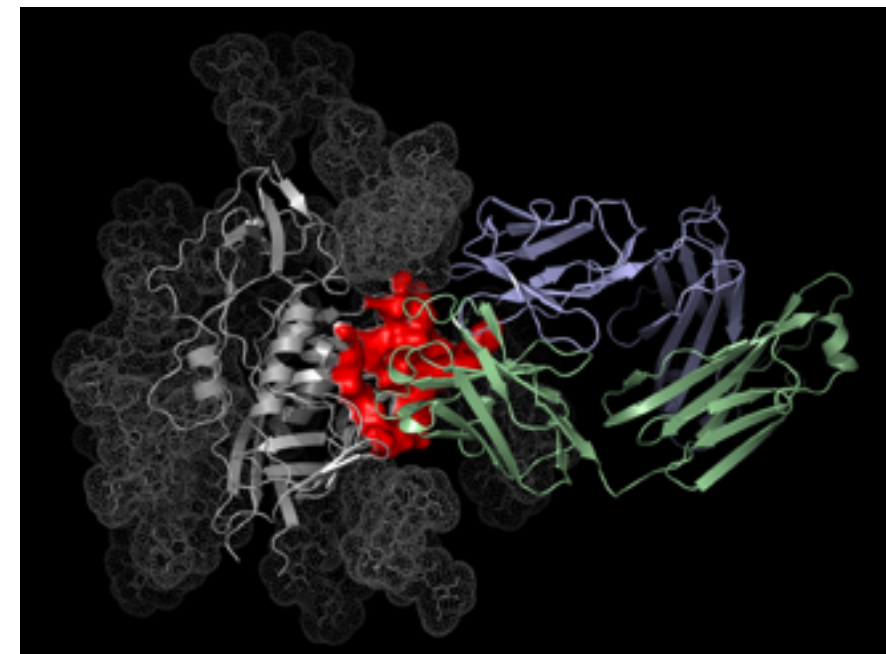
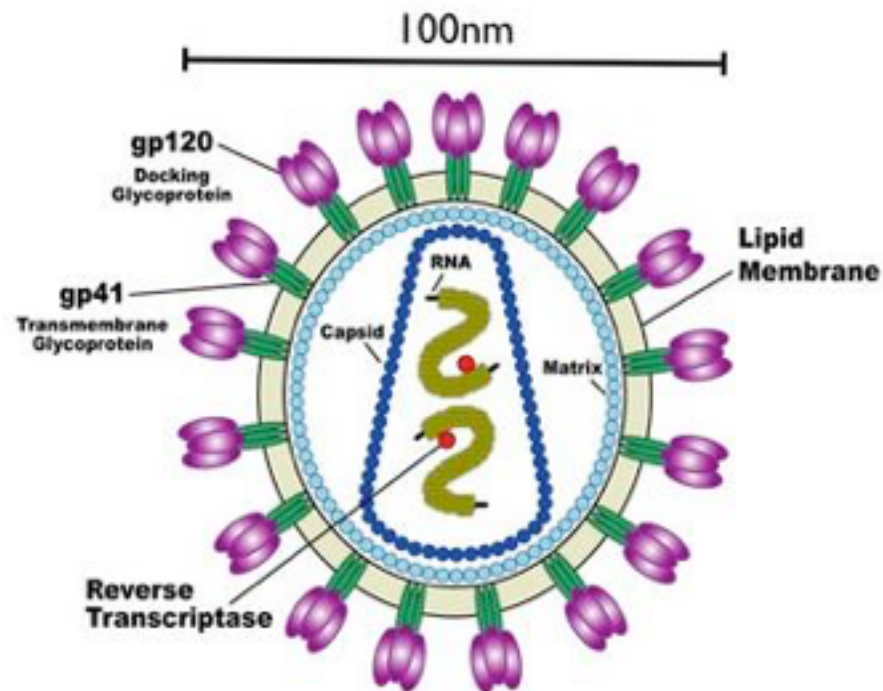
population B

intra-patient progression of chronic infections (HIV)

chronic virus
(HIV)



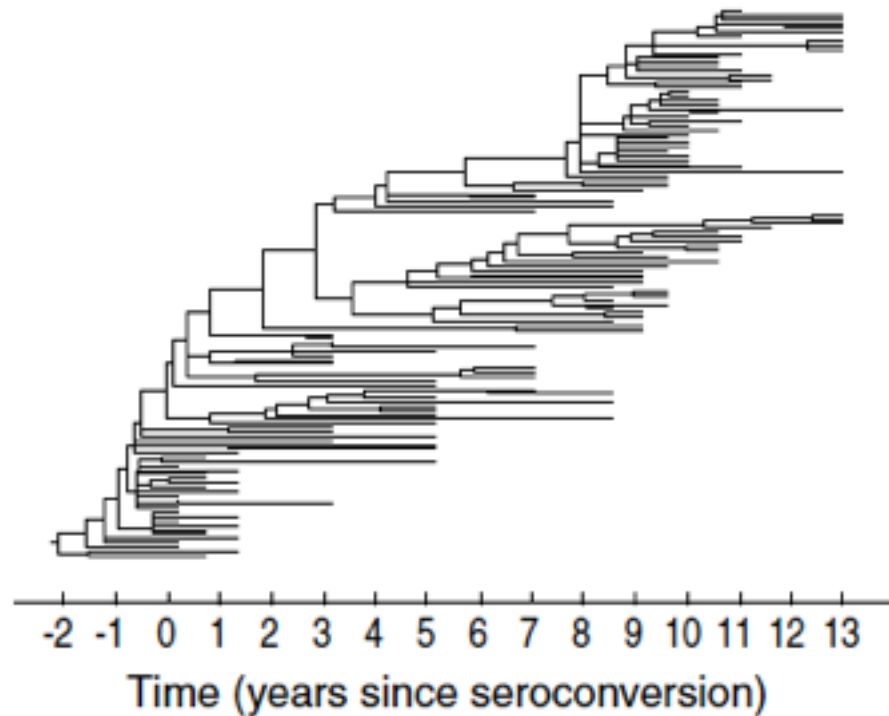
adaptive immune
system



VRC01 antibody (blue and green)
binding to HIV (red and gray)

intra-patient evolution of HIV

HIV evolution: intra-patient

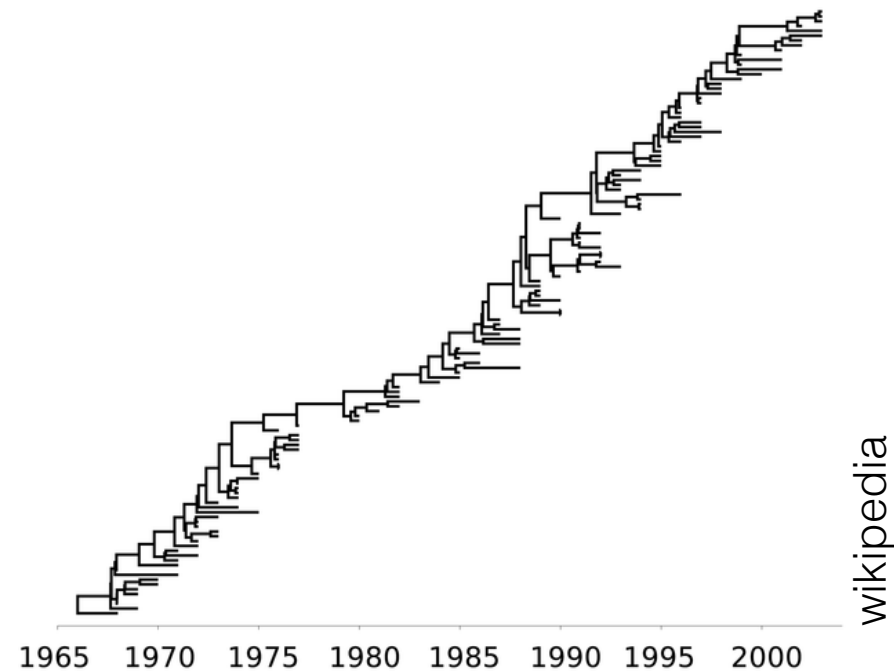


Lemey et al. (2006)

- generation time: ~ 2 days
 - mutation rate: 10^{-5} / bp.gen
 - *env* protein: $\sim 10^3$ bp
- diversity $\sim 10^{-3}$ - 10^{-2} /bp

$$\theta l = 1 - 10$$

influenza world-wide

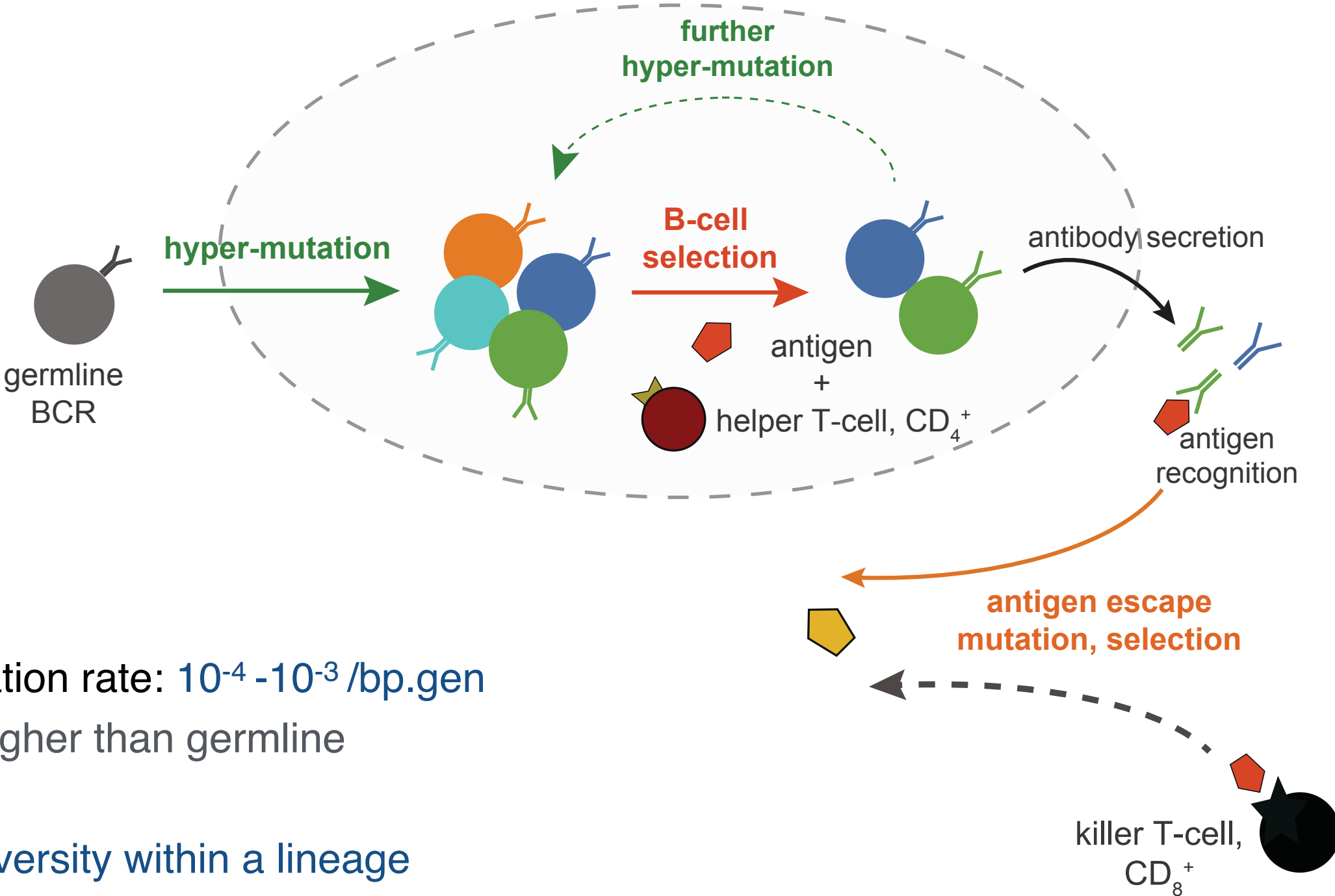


wikipedia

8% divergence in 10 years in a patient
20 yrs in global influenza

env protein: 1-10 aa mutations present

antibody affinity maturation



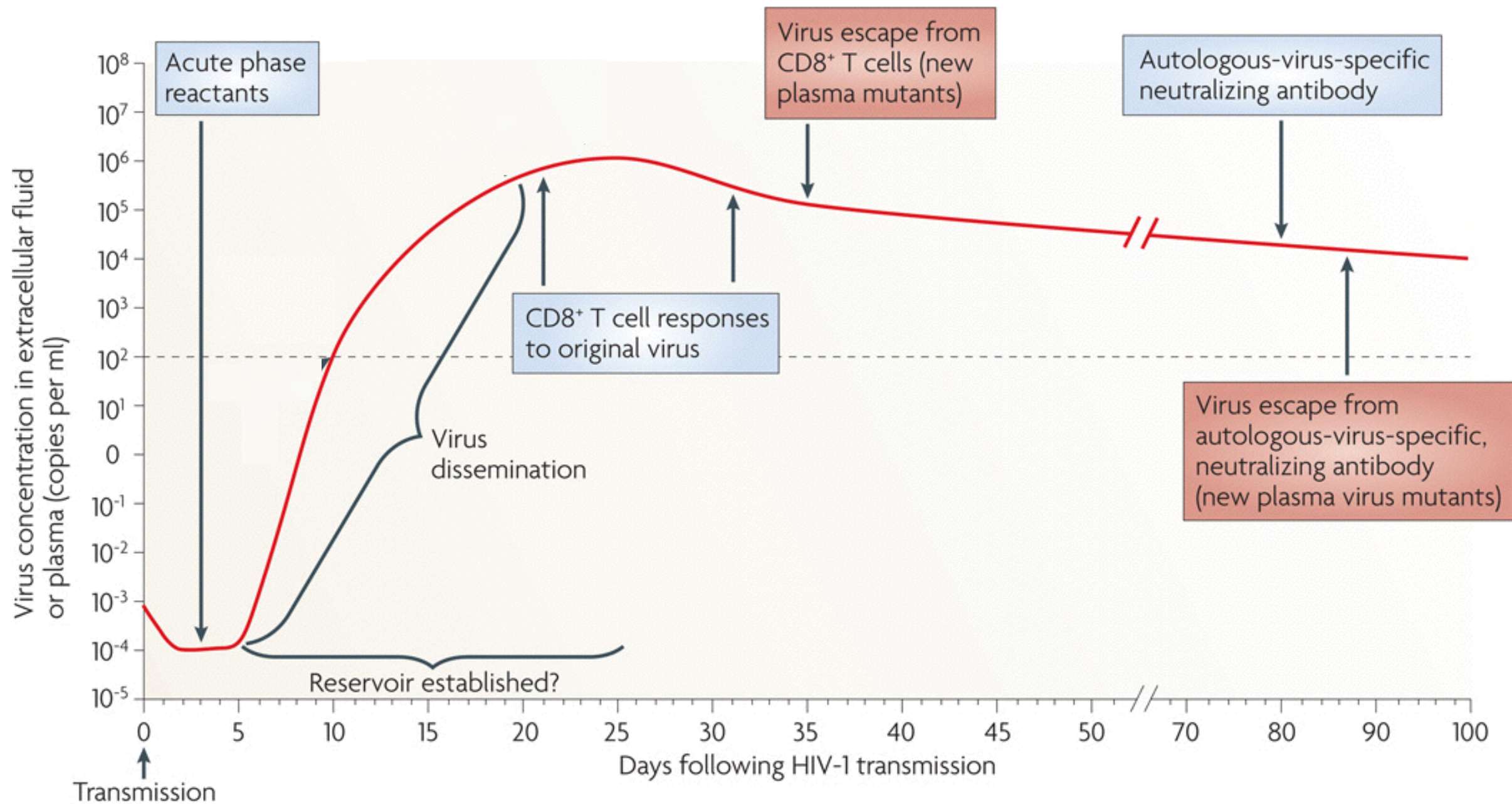
hyper-mutation rate: $10^{-4} - 10^{-3}$ /bp.gen
much higher than germline

~ 5-10% diversity within a lineage

$$\theta l = 1 - 10$$

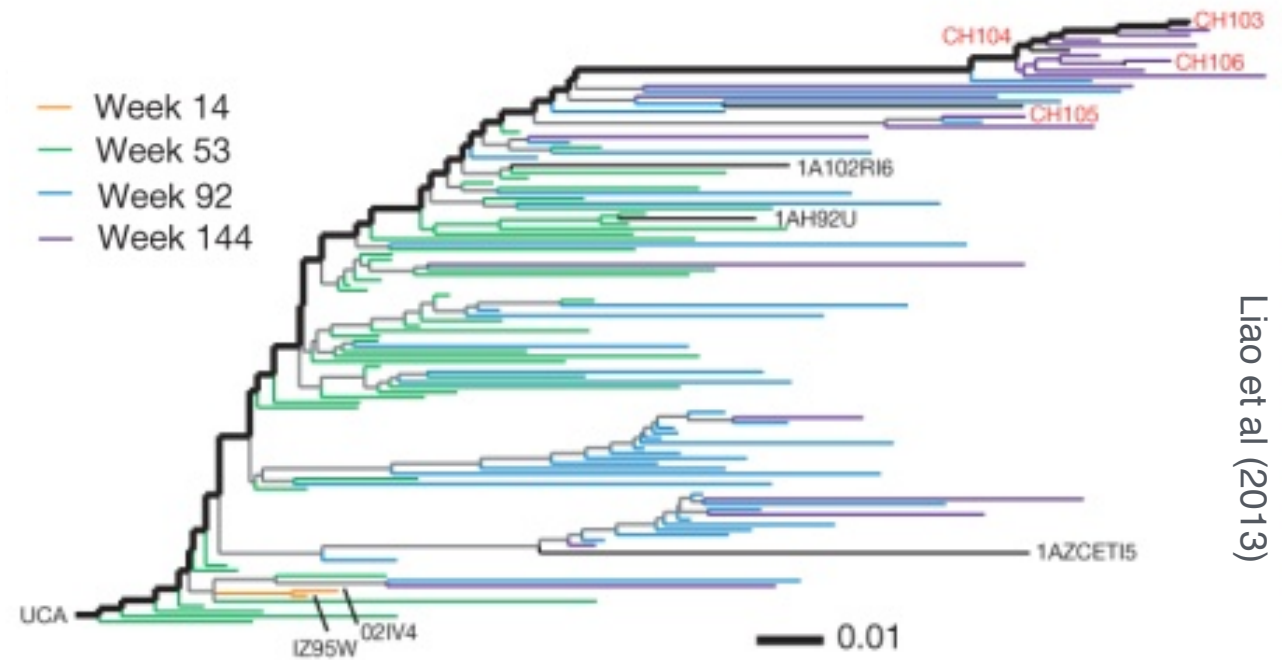
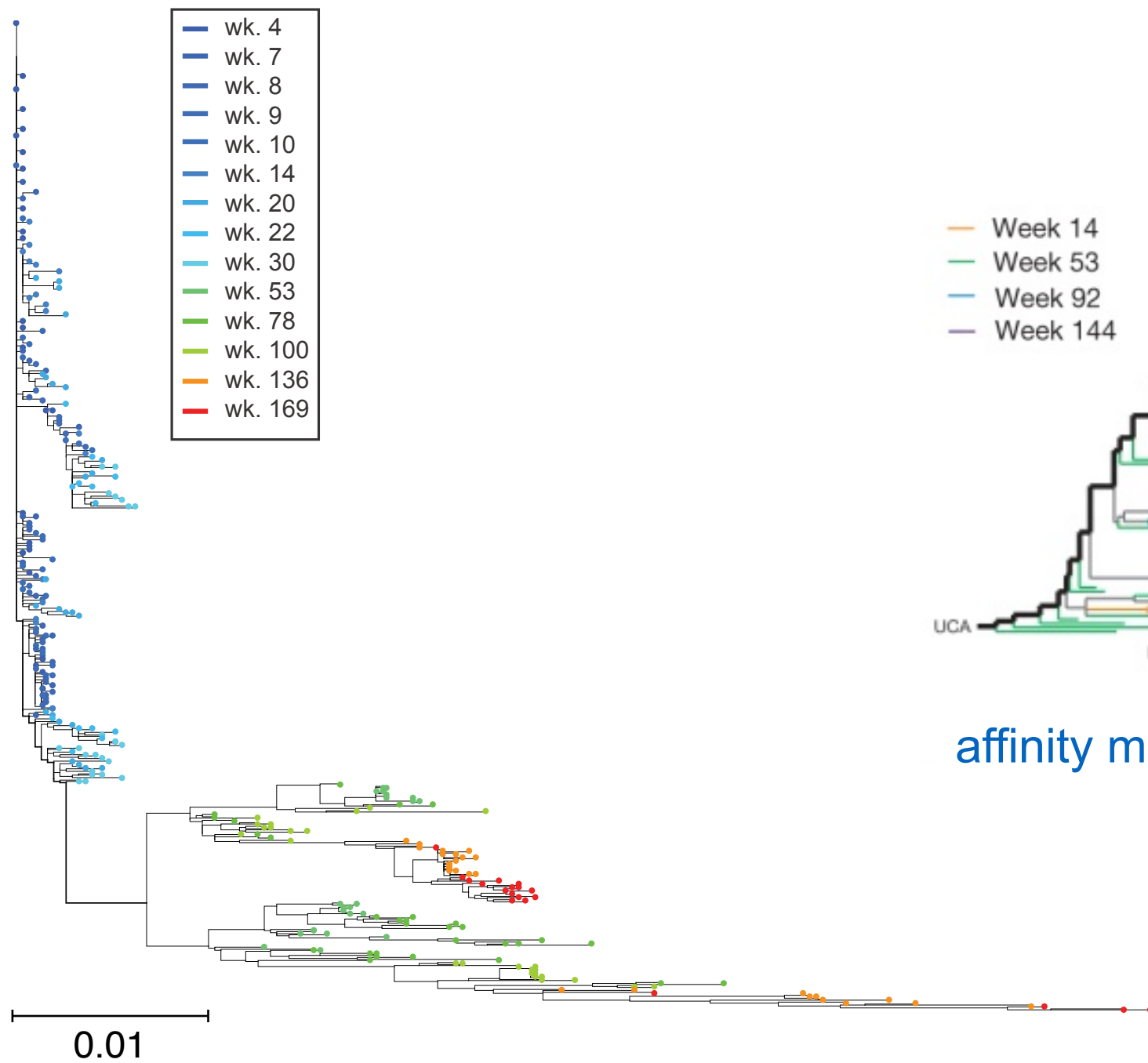
increases binding affinity **10-100 fold**

immune response to HIV



intra-patient coevolution of antibodies and HIV

ch505

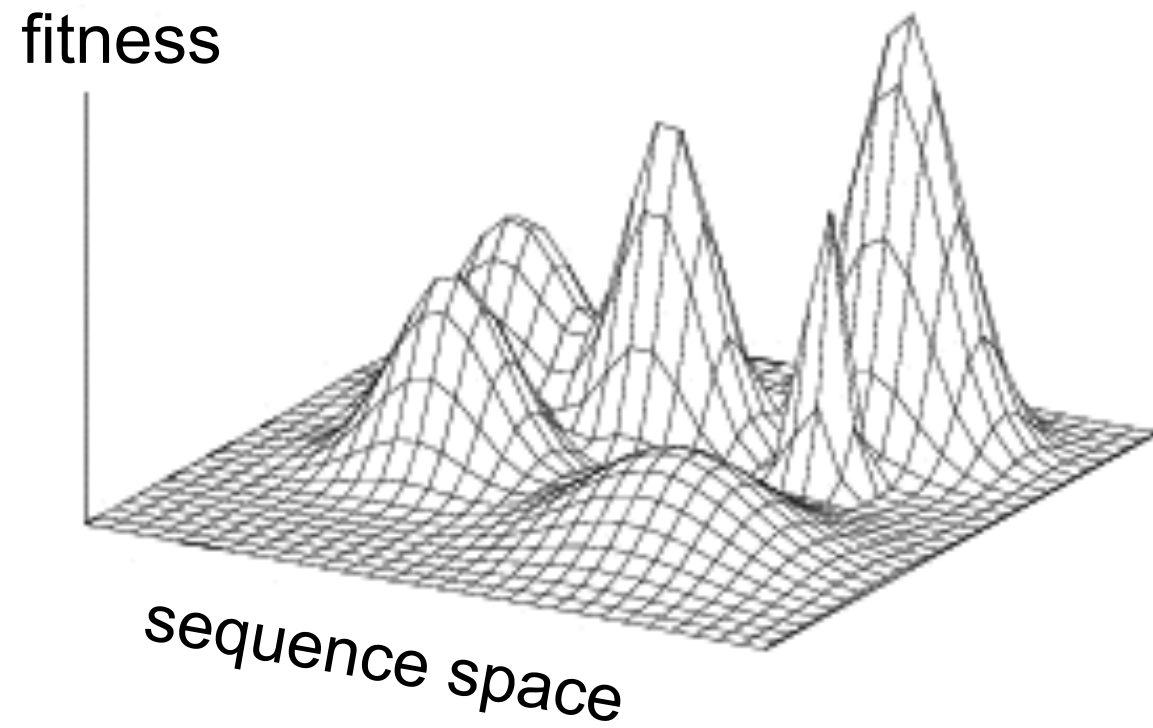


Liao et al (2013)

affinity maturation of an antibody lineage

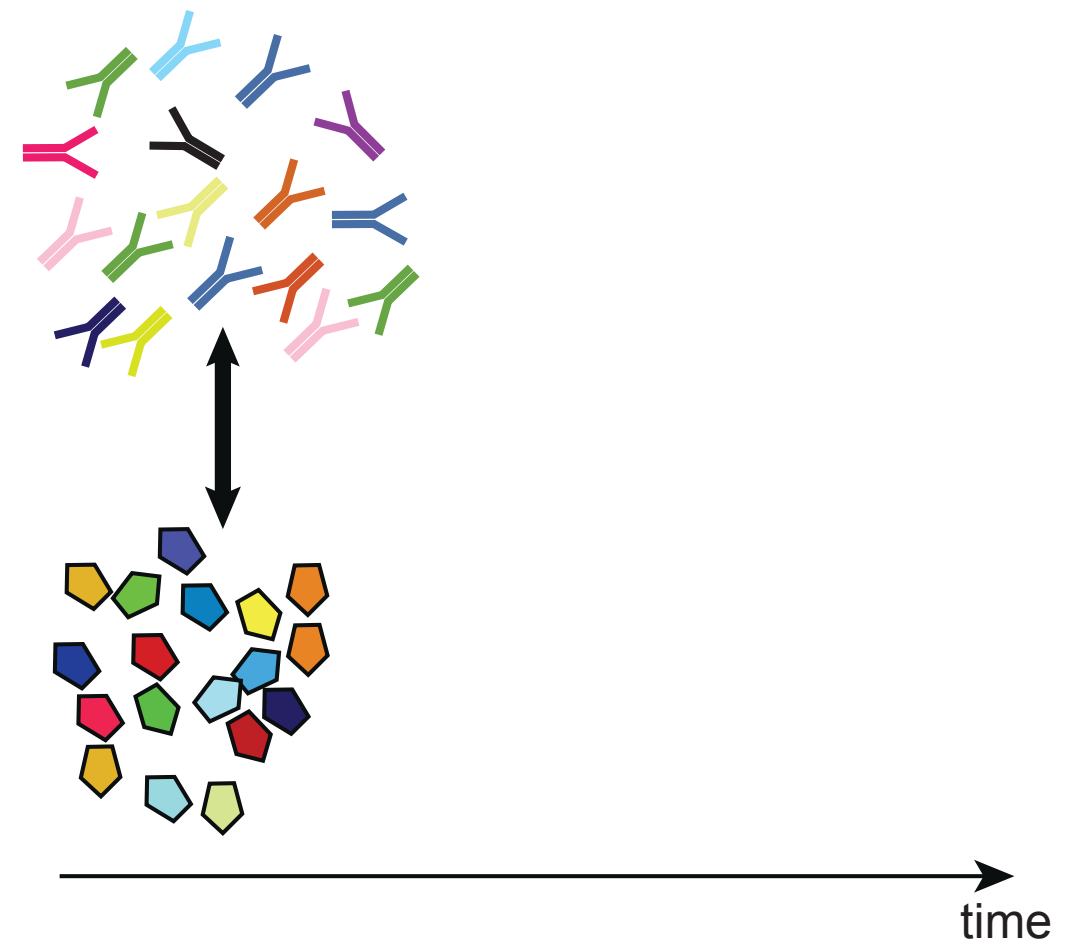
escape of HIV envelope protein (*env*)

co-evolution in the genotype space

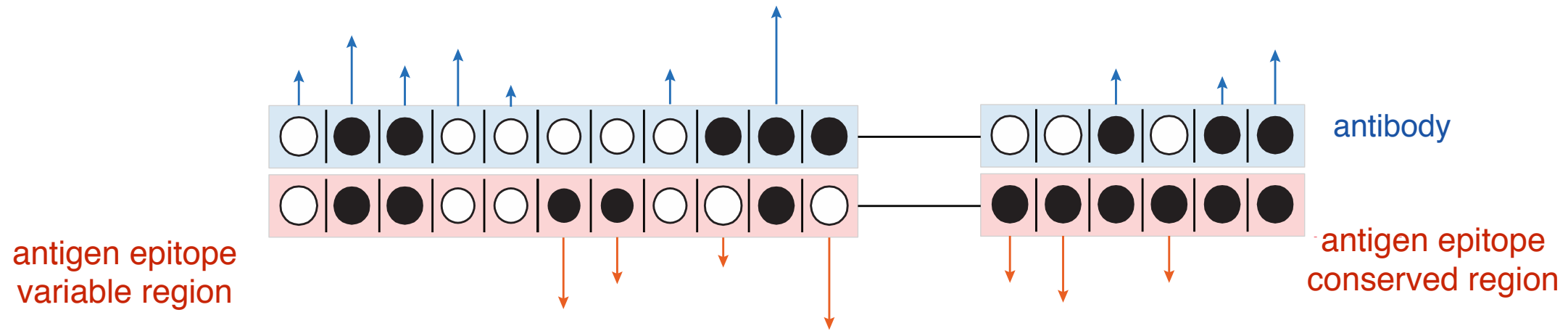


difficulties:

- high dimensional
- highly **under-sampled**
- **cross-reactivity**
- **time-dependence**



co-evolution in the phenotype space



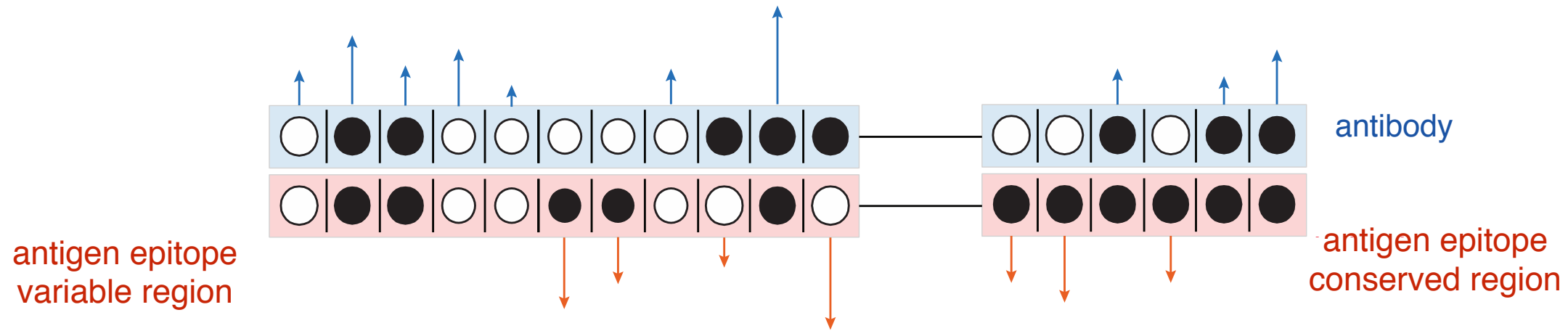
$$E(\mathbf{A}^\alpha, \mathbf{V}^\gamma) = \underbrace{\sum_{i=1}^{\ell} \varepsilon_i A_i^\alpha V_i^\gamma}_{\text{antigen variable region}} + \underbrace{\sum_{i=\ell+1}^{\ell+\tilde{\ell}} \tilde{\varepsilon}_i A_i^\alpha}_{\text{antigen conserved region (stability)}}$$

$$A_i, V_i = \begin{cases} -1, & \text{if } \circ \\ 1, & \text{if } \bullet \end{cases}$$

target for BnAbs

accessibility of an antibody lineage to viral epitopes: $\{\varepsilon_i, \tilde{\varepsilon}_i\}$

co-evolution in the phenotype space



$$E(\mathbf{A}^\alpha, \mathbf{V}^\gamma) = \underbrace{\sum_{i=1}^{\ell} \varepsilon_i A_i^\alpha V_i^\gamma}_{\text{antigen variable region}} + \underbrace{\sum_{i=\ell+1}^{\ell+\tilde{\ell}} \tilde{\varepsilon}_i A_i^\alpha}_{\text{antigen conserved region (stability)}}$$

$$A_i, V_i = \begin{cases} -1, & \text{if } \circ \\ 1, & \text{if } \bullet \end{cases}$$

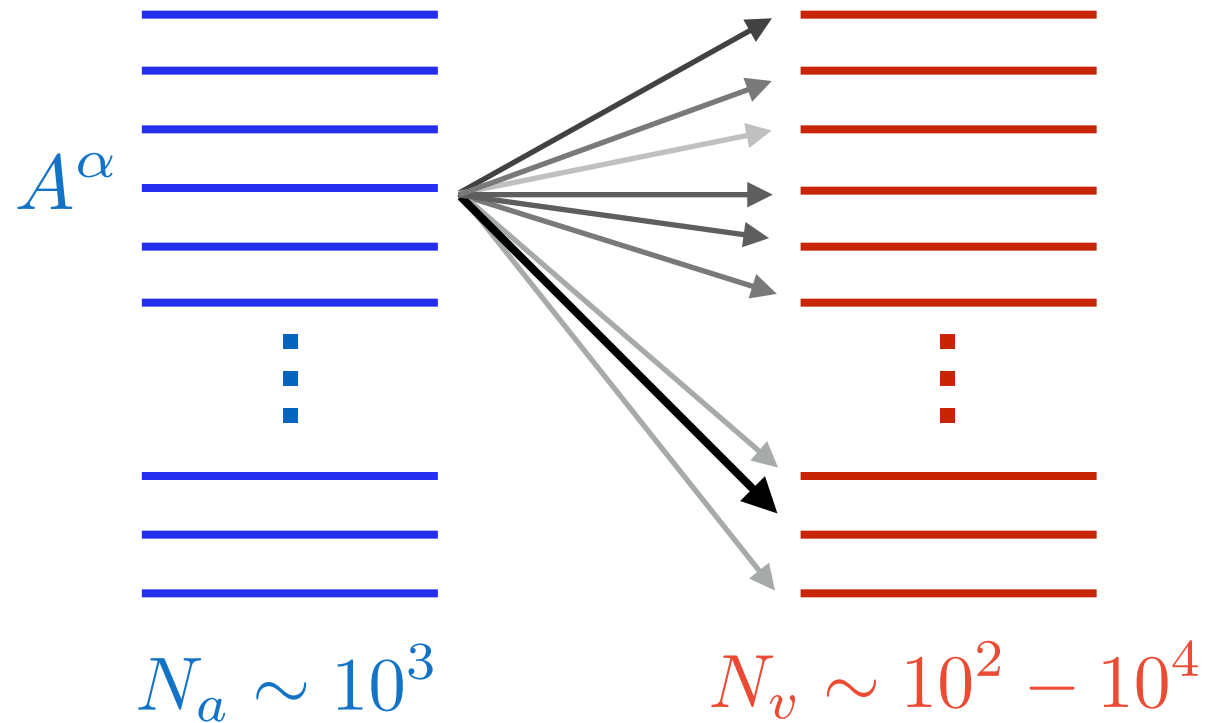
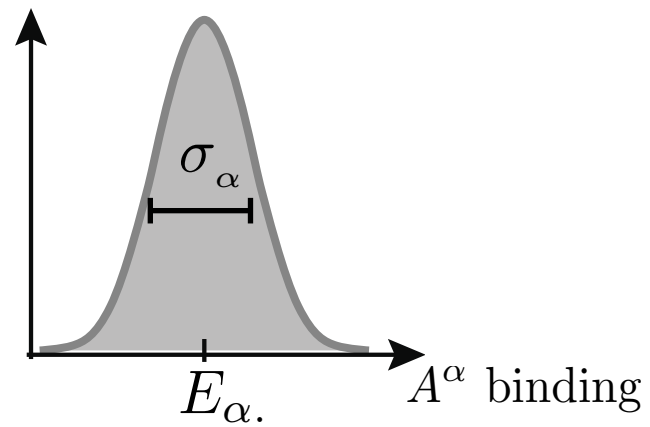
target for BnAbs

accessibility of an antibody lineage to viral epitopes: $\{\varepsilon_i, \tilde{\varepsilon}_i\}$

potential for breadth of a lineage: $\frac{\sum_i \tilde{\varepsilon}_i^2}{\sum_i \varepsilon_i^2 + \tilde{\varepsilon}_i^2}$

phenotype distributions in populations

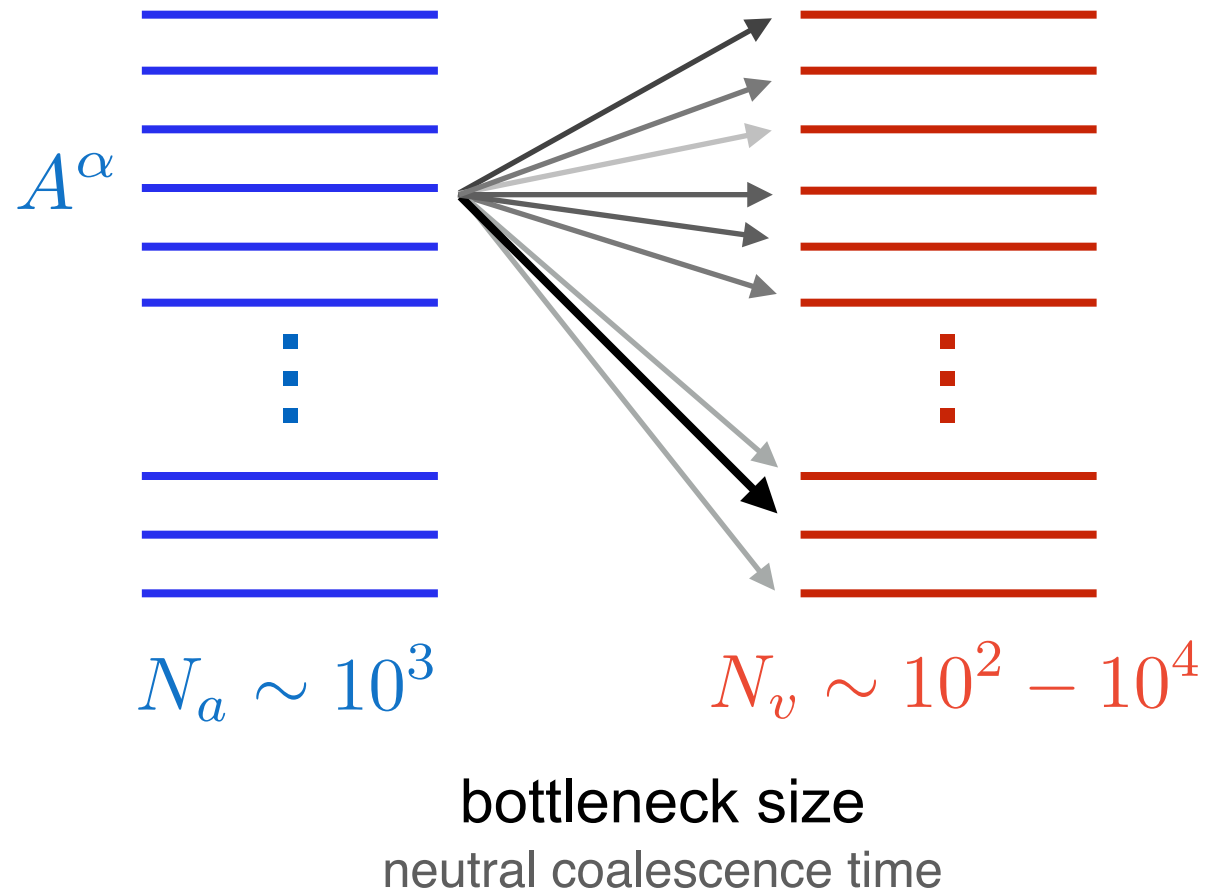
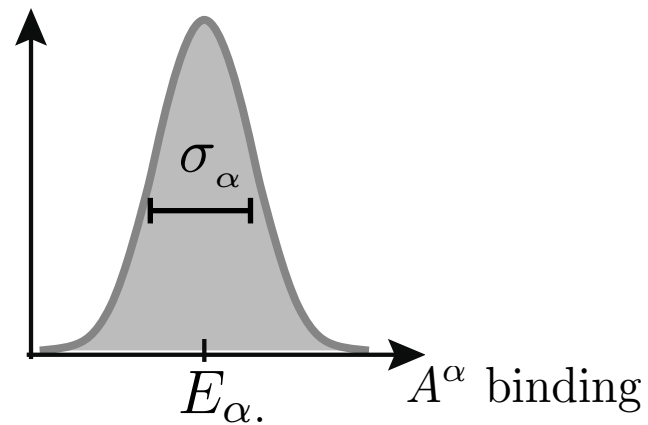
- binding properties of a BCR



bottleneck size
neutral coalescence time

phenotype distributions in populations

- binding properties of a BCR

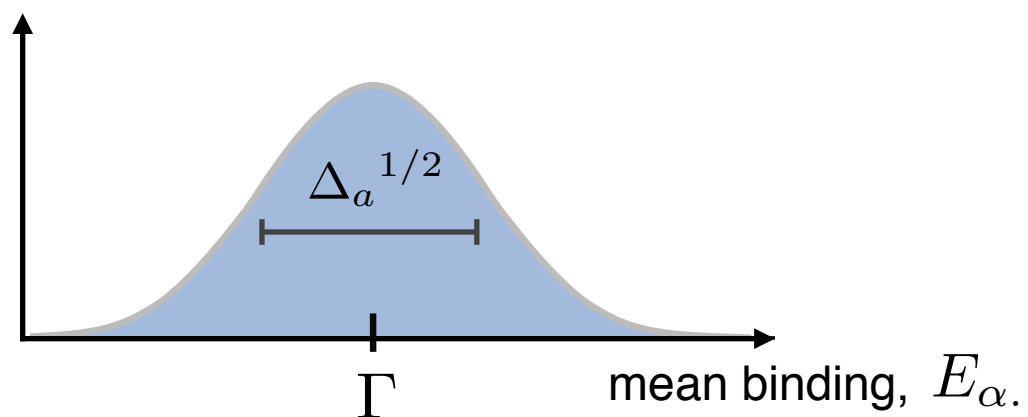


- mean binding affinity

$$E_{\alpha.}(t) = \sum_{\gamma \in \text{vir.}} E(A^\alpha, V^\gamma) y^\gamma(t)$$

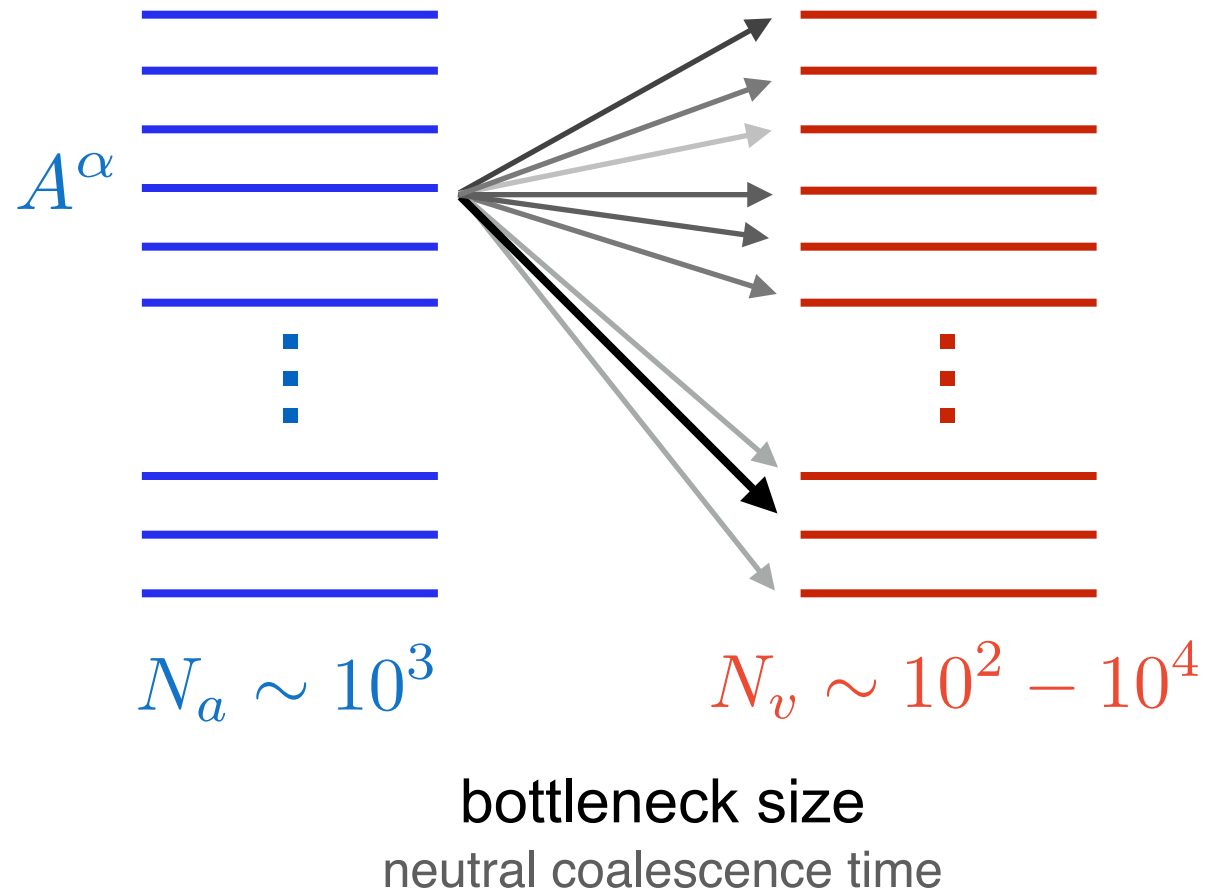
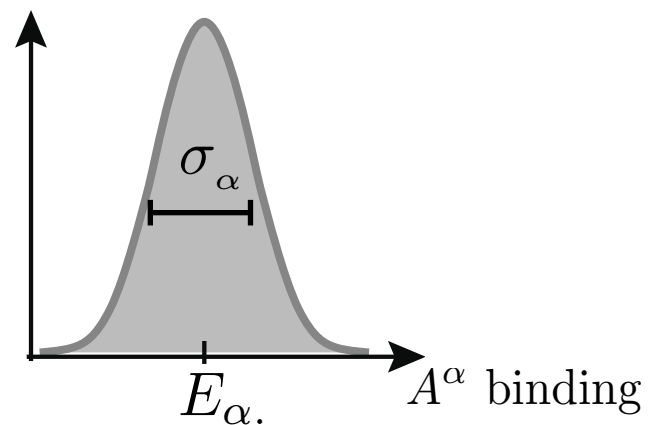
- best binding of R interactions

$$E_{\max_\alpha}(t) \simeq E_{\alpha.}(t) + \sqrt{2 \sigma_\alpha^2 \ln R}$$



phenotype distributions in populations

- binding properties of a BCR



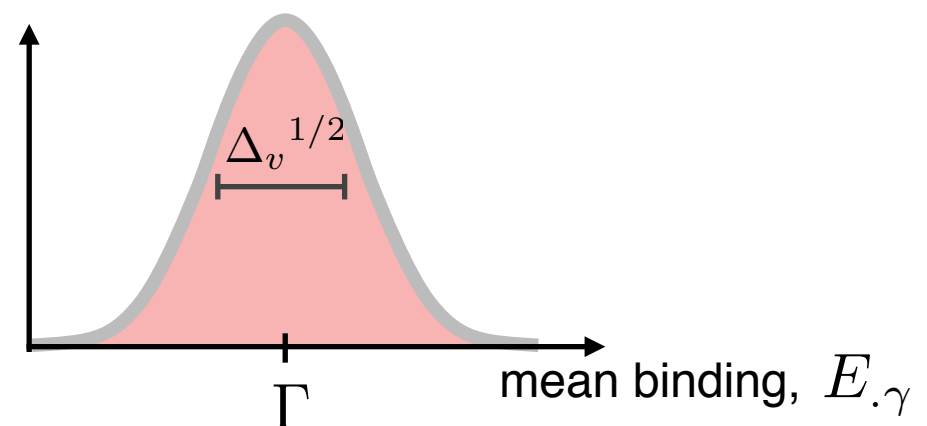
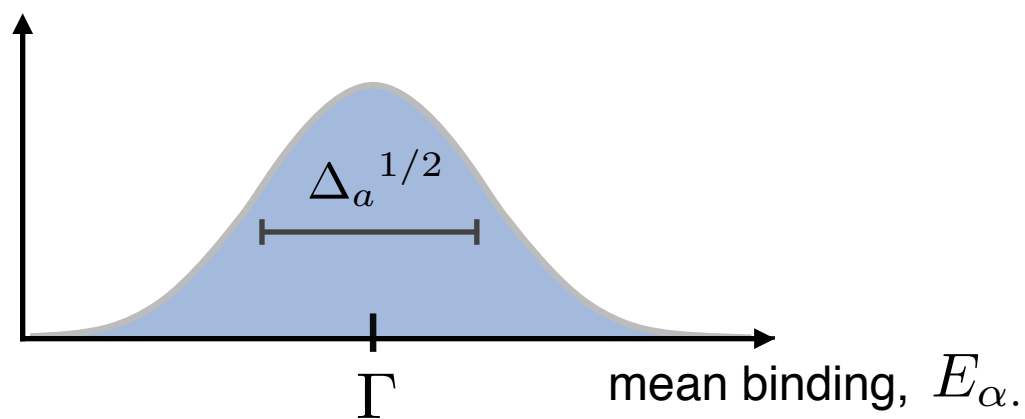
- mean binding affinity

$$E_{\alpha.}(t) = \sum_{\gamma \in \text{vir.}} E(A^\alpha, V^\gamma) y^\gamma(t)$$

- best binding of R interactions

$$E_{\max_\alpha}(t) \simeq E_{\alpha.}(t) + \sqrt{2 \sigma_\alpha^2 \ln R}$$

similar for viruses: $E_{.\gamma}(t)$



stochastic evolution of the phenotypes

→ coupled fitness functions (simplest case)

antibody fitness: $f(A^\alpha | \{V\}) = f_a^* + S_a E_\alpha.$

antigen fitness: $f(V^\gamma | \{A\}) = f_v^* - S_v E_{\cdot\gamma}$

+ higher moments
& non-linearities

stochastic evolution of the phenotypes

→ coupled fitness functions (simplest case)

antibody fitness: $f(A^\alpha | \{V\}) = f_a^* + S_a E_\alpha.$

antigen fitness: $f(V^\gamma | \{A\}) = f_v^* - S_v E_{\cdot\gamma}$

+ higher moments & non-linearities

→ evolution of grand mean binding, Γ :

$$\frac{d}{dt}\Gamma = -2(\mu_v + \mu_a)(\Gamma - \Gamma_0) + S_a \Delta_a(t) - S_v \Delta_v(t) + \sqrt{\frac{\Delta_a}{N_a} + \frac{\Delta_v}{N_v}} \chi_\Gamma$$

genetic drift

mutation

diversity in the host

diversity in the pathogen

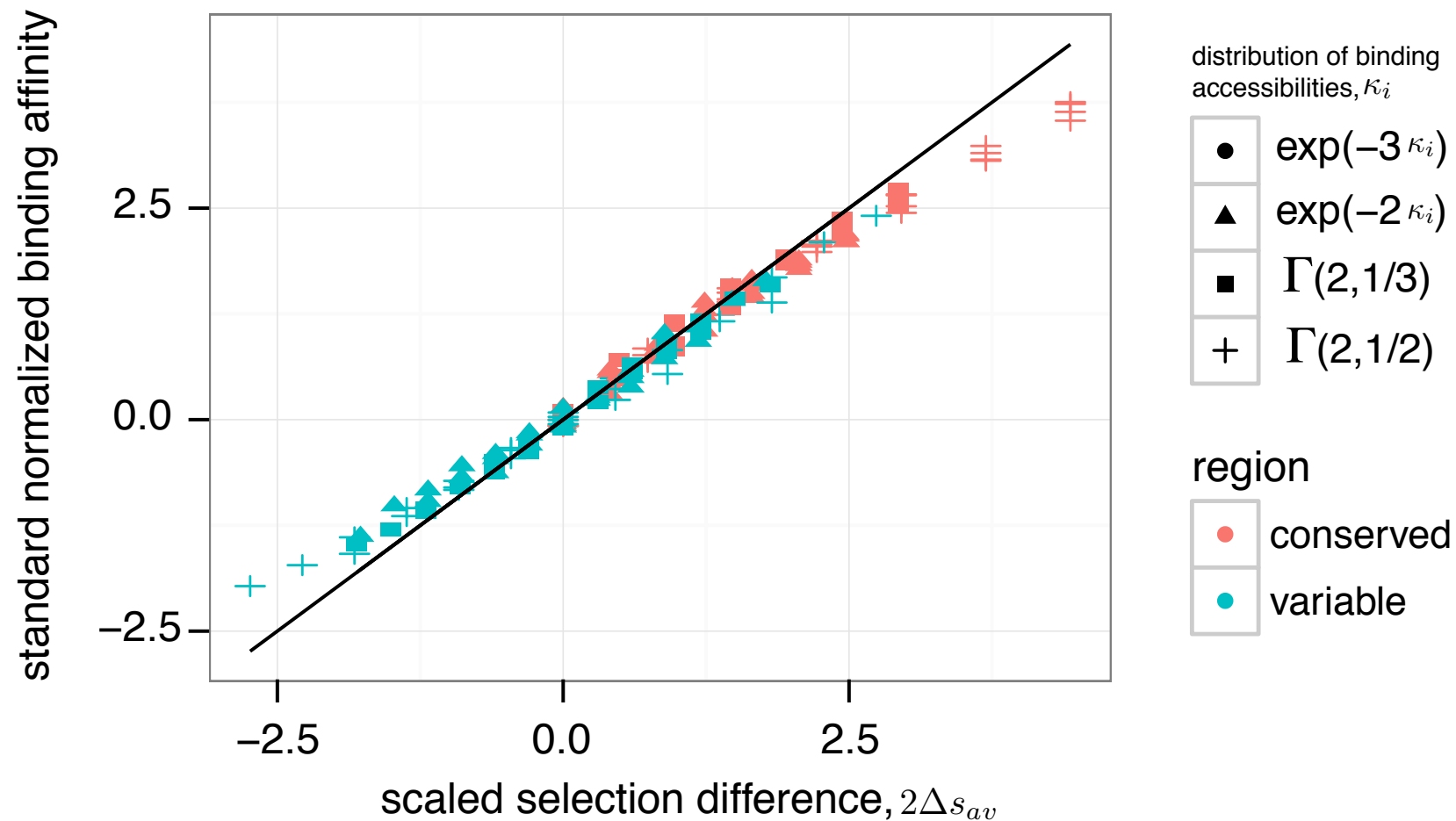
selection

.... and for higher moments

→ antibodies with access to conserved pathogen regions (BnAbs) are not restrained by co-evolution: $(\mu_v \sim 0, \Delta_v \sim 0)$

evolution of the mean binding affinity

- macroscopic statistics (e.g., mean affinity) are insensitive to details of antibody-antigen interactions



scaled selection coef.: $s = NSE_0$

selection difference: $\Delta s_{av} = \frac{s_a \theta_a - s_v \theta_v}{\theta_a + \theta_v}$

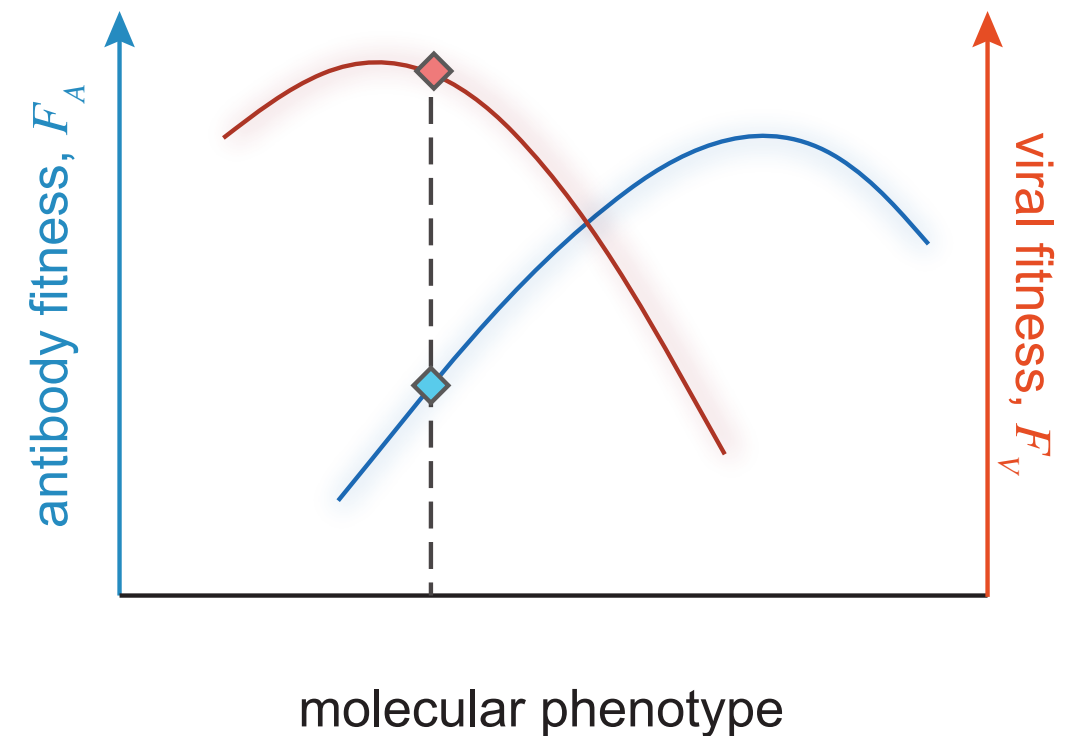
co-evolutionary fluxes

► **population history** is a sequence of joint measurements

$$((\mathbf{x}_0, \mathbf{y}_0, \Gamma_0), \dots, (\mathbf{x}_n, \mathbf{y}_n, \Gamma_n))$$

measured over time: (t_0, \dots, t_n)

selection coefficients: $s_a = \frac{\partial}{\partial \Gamma} F_a$, $s_v = \frac{\partial}{\partial \Gamma} F_v$



co-evolutionary fluxes

► **population history** is a sequence of joint measurements

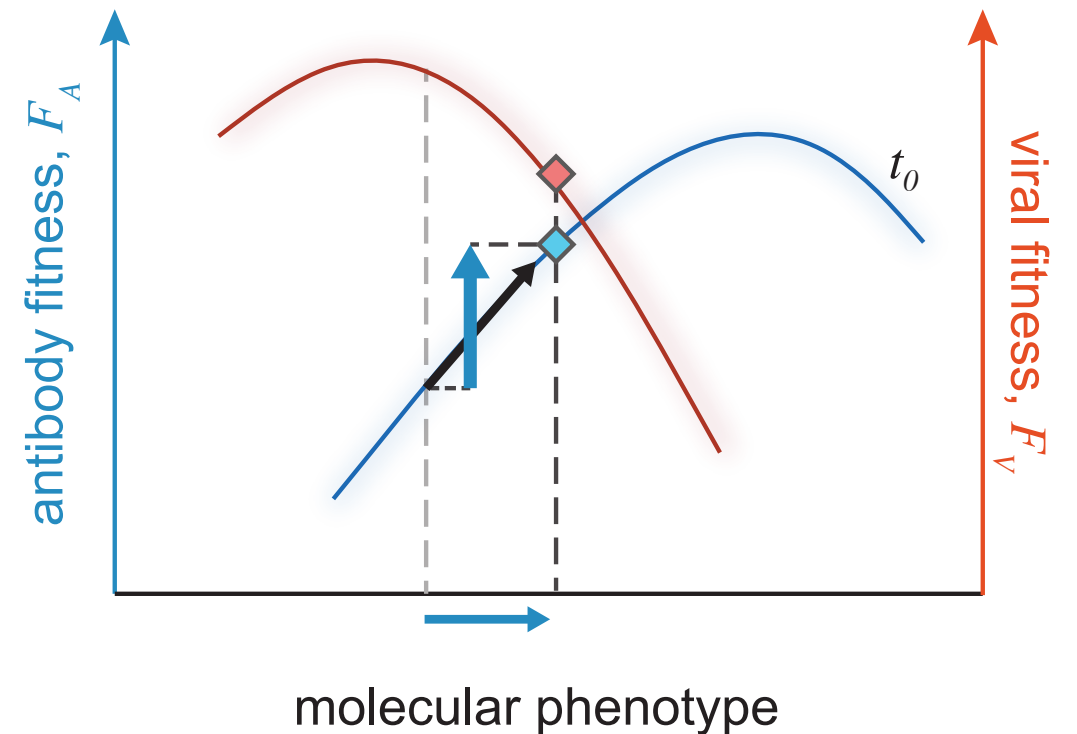
$$((\mathbf{x}_0, \mathbf{y}_0, \Gamma_0), \dots, (\mathbf{x}_n, \mathbf{y}_n, \Gamma_n))$$

measured over time: (t_0, \dots, t_n)

► **fitness flux**, measure of adaptation

$$\phi_A = \sum_{\alpha: \text{antibody}} \delta x^\alpha s_a(\Gamma, t) \partial_\alpha \Gamma$$

selection coefficients: $s_a = \frac{\partial}{\partial \Gamma} F_a$, $s_v = \frac{\partial}{\partial \Gamma} F_v$



co-evolutionary fluxes

- ▶ **population history** is a sequence of joint measurements

$$((\mathbf{x}_0, \mathbf{y}_0, \Gamma_0), \dots, (\mathbf{x}_n, \mathbf{y}_n, \Gamma_n))$$

measured over time: (t_0, \dots, t_n)

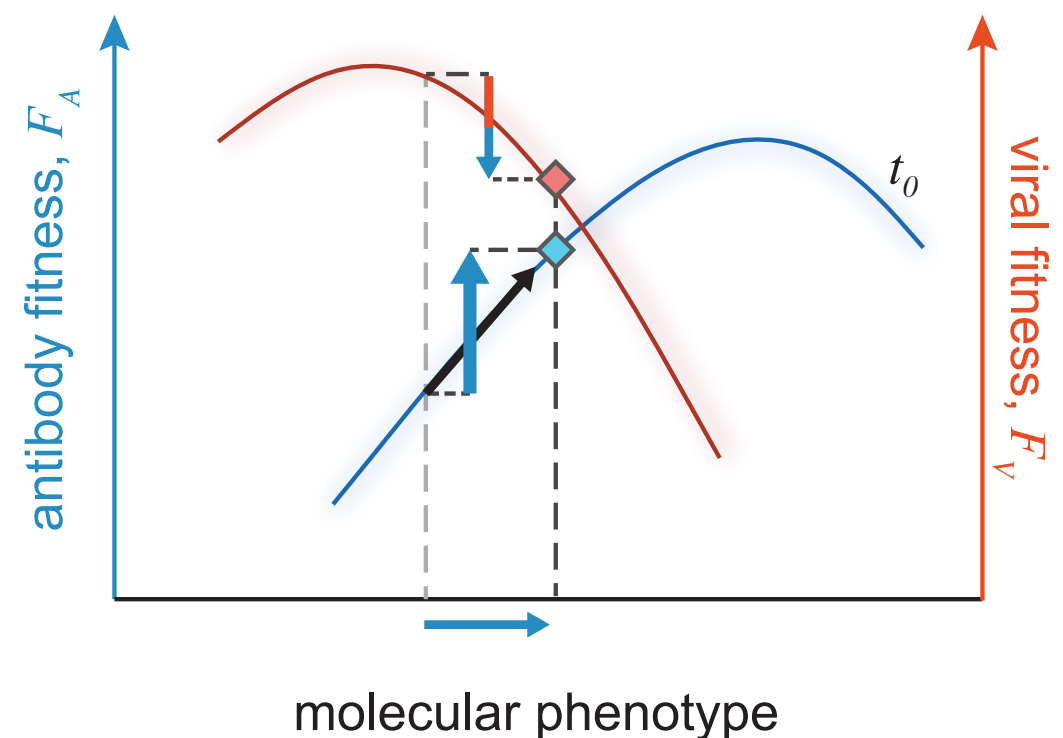
- ▶ **fitness flux**, measure of adaptation

$$\phi_A = \sum_{\alpha: \text{antibody}} \delta x^\alpha s_a(\Gamma, t) \partial_\alpha \Gamma$$

- ▶ **transfer flux**, measure of interaction

$$\mathcal{T}_{A \rightarrow V} = \sum_{\alpha: \text{antibody}} \delta x^\alpha s_v(\Gamma, t) \partial_\alpha \Gamma$$

selection coefficients: $s_a = \frac{\partial}{\partial \Gamma} F_a$, $s_v = \frac{\partial}{\partial \Gamma} F_v$



co-evolutionary fluxes

► **population history** is a sequence of joint measurements

$$((\mathbf{x}_0, \mathbf{y}_0, \Gamma_0), \dots, (\mathbf{x}_n, \mathbf{y}_n, \Gamma_n))$$

measured over time: (t_0, \dots, t_n)

► **fitness flux**, measure of adaptation

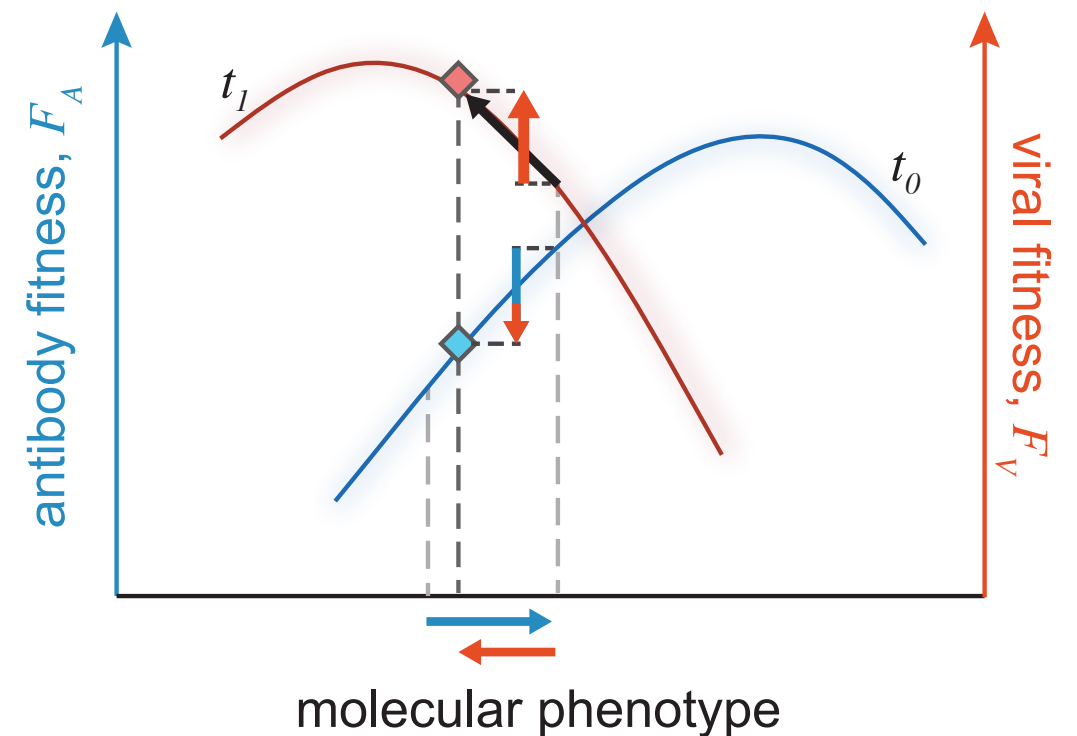
$$\phi_A = \sum_{\alpha: \text{antibody}} \delta x^\alpha s_a(\Gamma, t) \partial_\alpha \Gamma$$

► **transfer flux**, measure of interaction

$$\mathcal{T}_{A \rightarrow V} = \sum_{\alpha: \text{antibody}} \delta x^\alpha s_v(\Gamma, t) \partial_\alpha \Gamma$$

similar for viruses

selection coefficients: $s_a = \frac{\partial}{\partial \Gamma} F_a$, $s_v = \frac{\partial}{\partial \Gamma} F_v$



co-evolutionary fluxes

- ▶ **population history** is a sequence of joint measurements

$$((\mathbf{x}_0, \mathbf{y}_0, \Gamma_0), \dots, (\mathbf{x}_n, \mathbf{y}_n, \Gamma_n))$$

measured over time: (t_0, \dots, t_n)

- ▶ **fitness flux**, measure of adaptation

$$\phi_A = \sum_{\alpha: \text{antibody}} \delta x^\alpha s_a(\Gamma, t) \partial_\alpha \Gamma$$

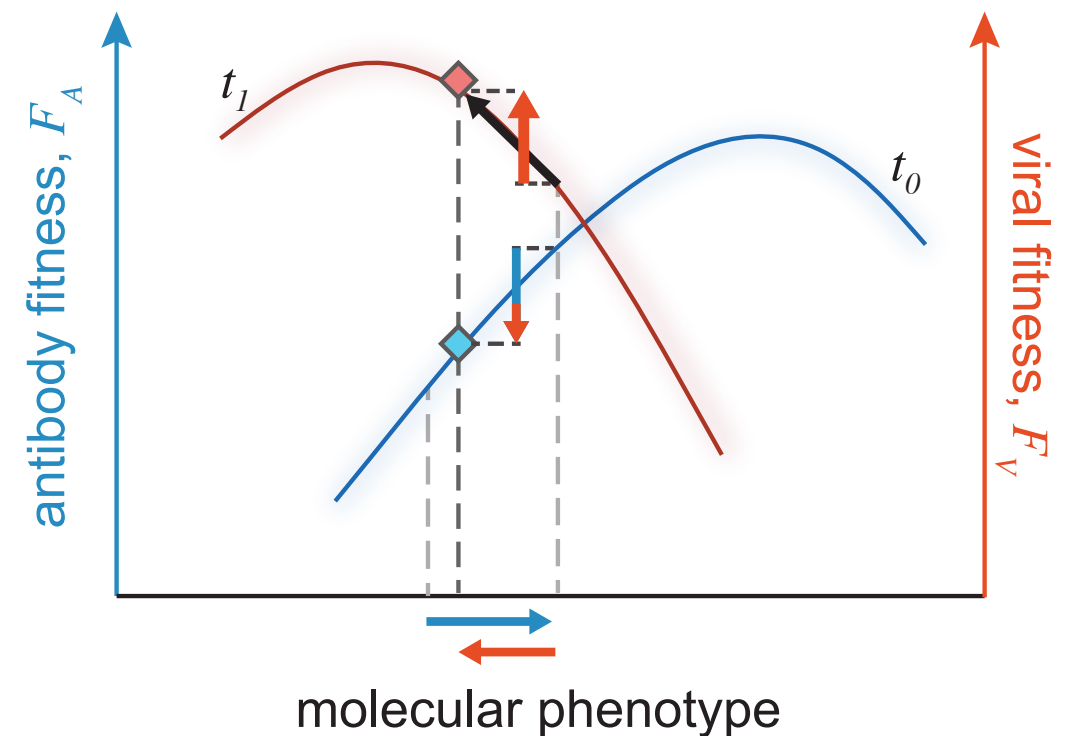
- ▶ **transfer flux**, measure of interaction

$$\mathcal{T}_{A \rightarrow V} = \sum_{\alpha: \text{antibody}} \delta x^\alpha s_v(\Gamma, t) \partial_\alpha \Gamma$$

- ▶ **cumulative stationary fluxes**

$$\langle \Phi_A(\tau) \rangle = \left[-2\theta_a \frac{s_a}{E_0} \langle \Gamma \rangle + \frac{s_a^2}{E_0^2} \langle \Delta_a \rangle \right] \tau$$

selection coefficients: $s_a = \frac{\partial}{\partial \Gamma} F_a$, $s_v = \frac{\partial}{\partial \Gamma} F_v$



$$\langle \mathcal{T}_{A \rightarrow V}(\tau) \rangle = \left[-2\theta_a \frac{s_v}{E_0} \langle \Gamma \rangle - \frac{s_v s_a}{E_0^2} \langle \Delta_a \rangle \right] \tau$$

information theory of fluxes

- ▶ fitness flux: measures the amount of non-equilibrium

$$((\mathbf{x}_0, \mathbf{y}_0), \dots, (\mathbf{x}_n, \mathbf{y}_n)) \text{ vs. } ((\mathbf{x}_n, \mathbf{y}_0), \dots, (\mathbf{x}_0, \mathbf{y}_n))$$

→
←

$$\langle \Phi_A(t_n) \rangle \simeq D_{KL}(\mathcal{P}(x_{0:n}; y_{0:n}) || \mathcal{P}(x_{n:0}; y_{0:n}))$$

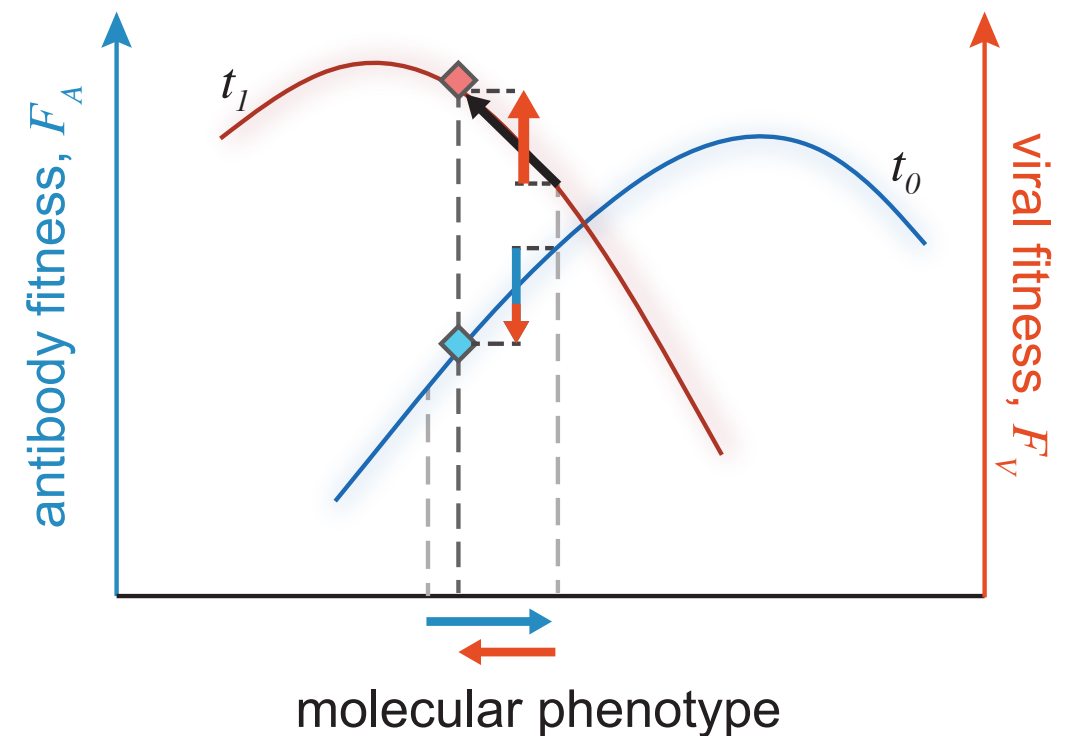
- ▶ transfer flux, measure of information transfer

$$(\mathbf{y}_0, \dots, \mathbf{y}_n) \text{ vs. } ((\mathbf{x}_0, \mathbf{y}_0), \dots, (\mathbf{x}_n, \mathbf{y}_n))$$

→
⇔

$$\langle T_{A \rightarrow V}(t_n) \rangle \simeq H(y_n | y_{0:n-1}) - H(y_n | y_{0:n-1}, x_{0:n-1})$$

time points: (t_0, \dots, t_n)



antibody-antigen adaptation

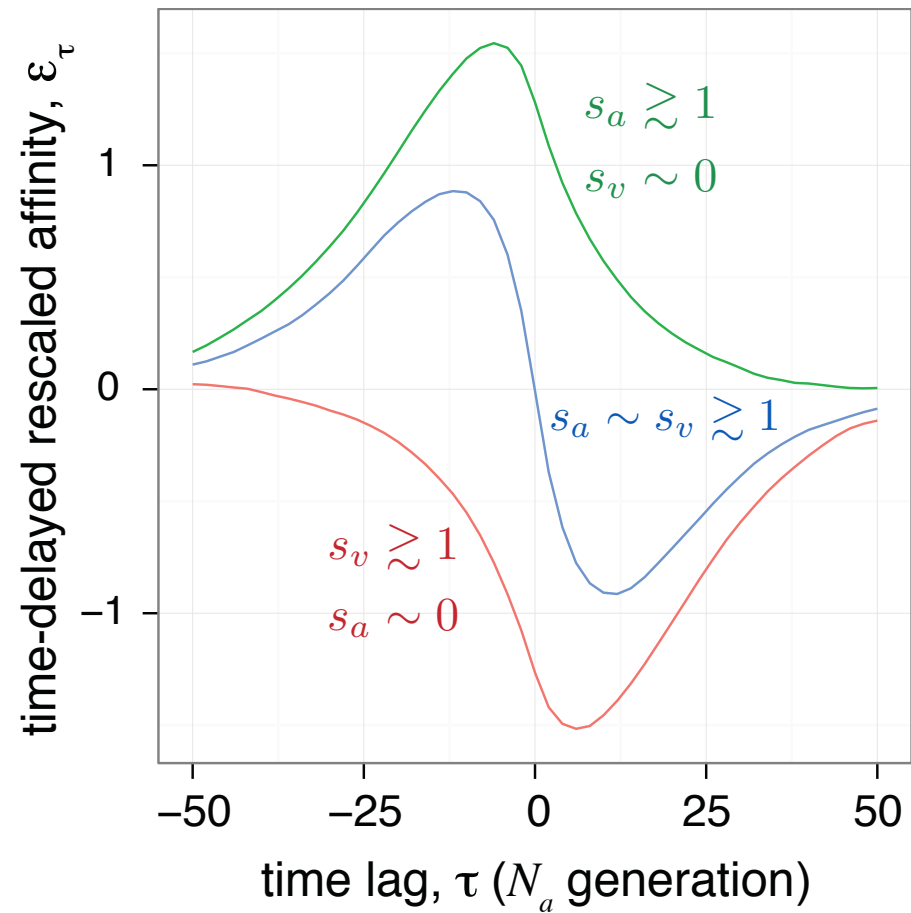


It takes all the running you can do, to keep in the same place. If you want to get somewhere else, you must run at least twice as fast as that!

Through the looking glass, Lewis Carroll

measuring the reciprocity

→ time-dependent observables for the non-equilibrium process

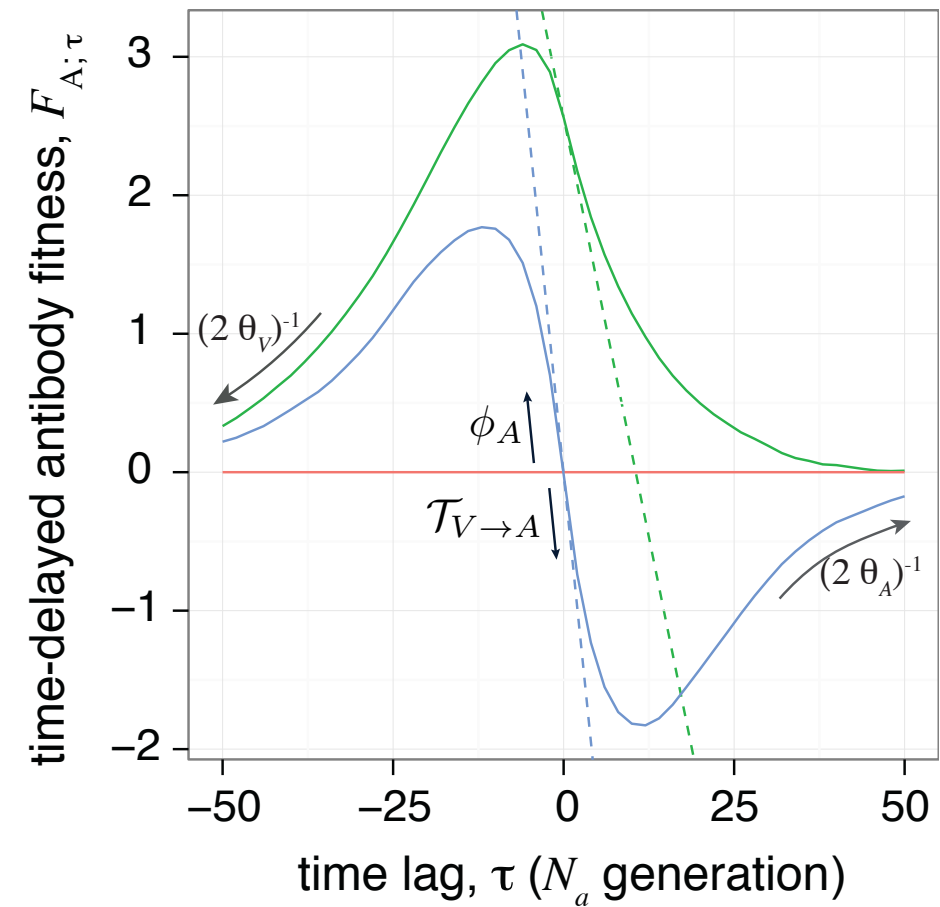
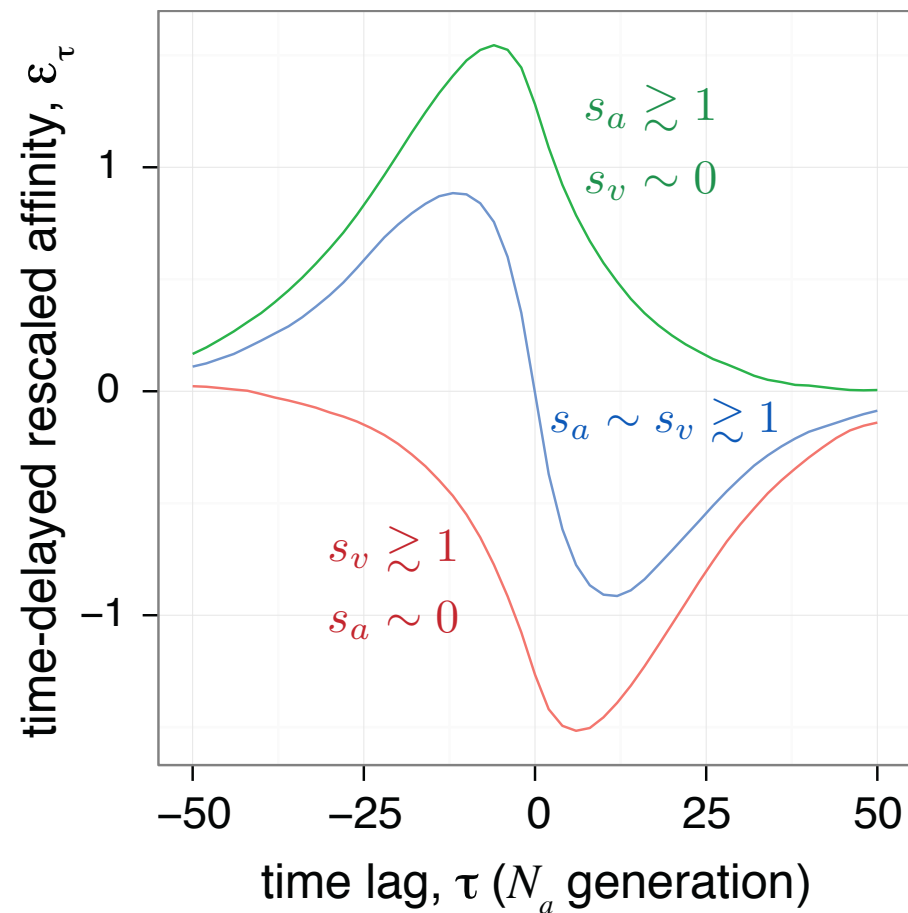


→ time-delayed interaction

$$\varepsilon_\tau(t) = \frac{1}{E_0} \sum_{\alpha, \gamma} E_{\alpha\gamma} x^\alpha(t) y^\gamma(t + \tau)$$

measuring the reciprocity

→ time-dependent observables for the non-equilibrium process



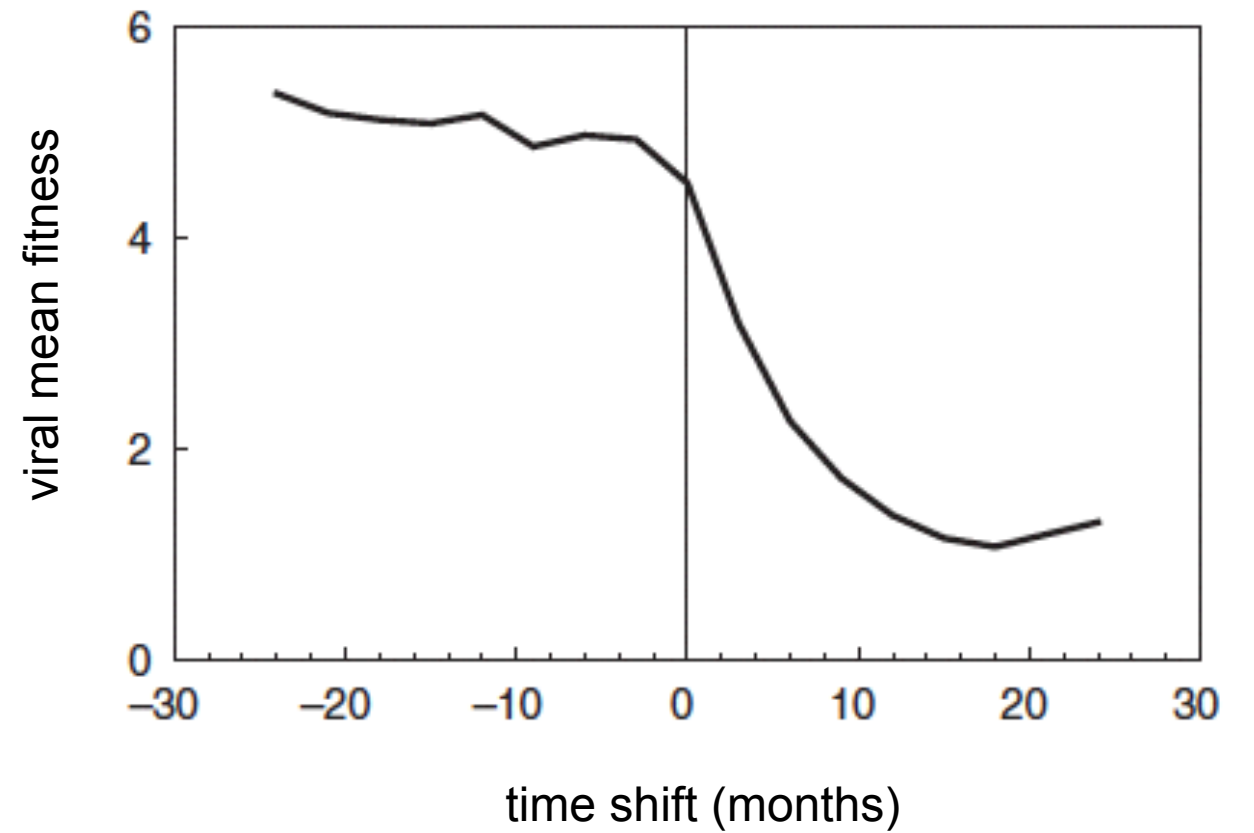
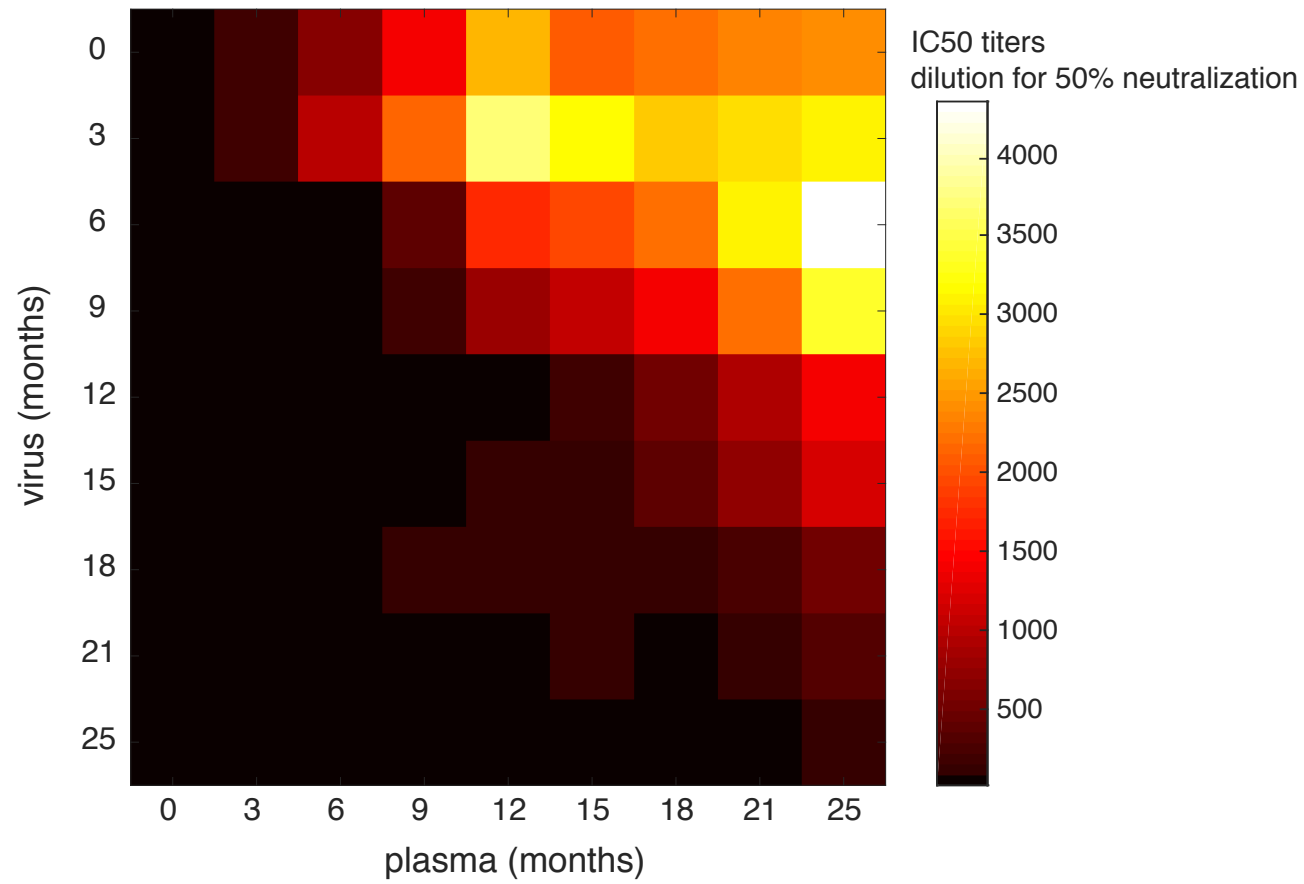
→ time-delayed interaction

$$\varepsilon_\tau(t) = \frac{1}{E_0} \sum_{\alpha, \gamma} E_{\alpha\gamma} x^\alpha(t) y^\gamma(t + \tau)$$

→ time-delayed fitness

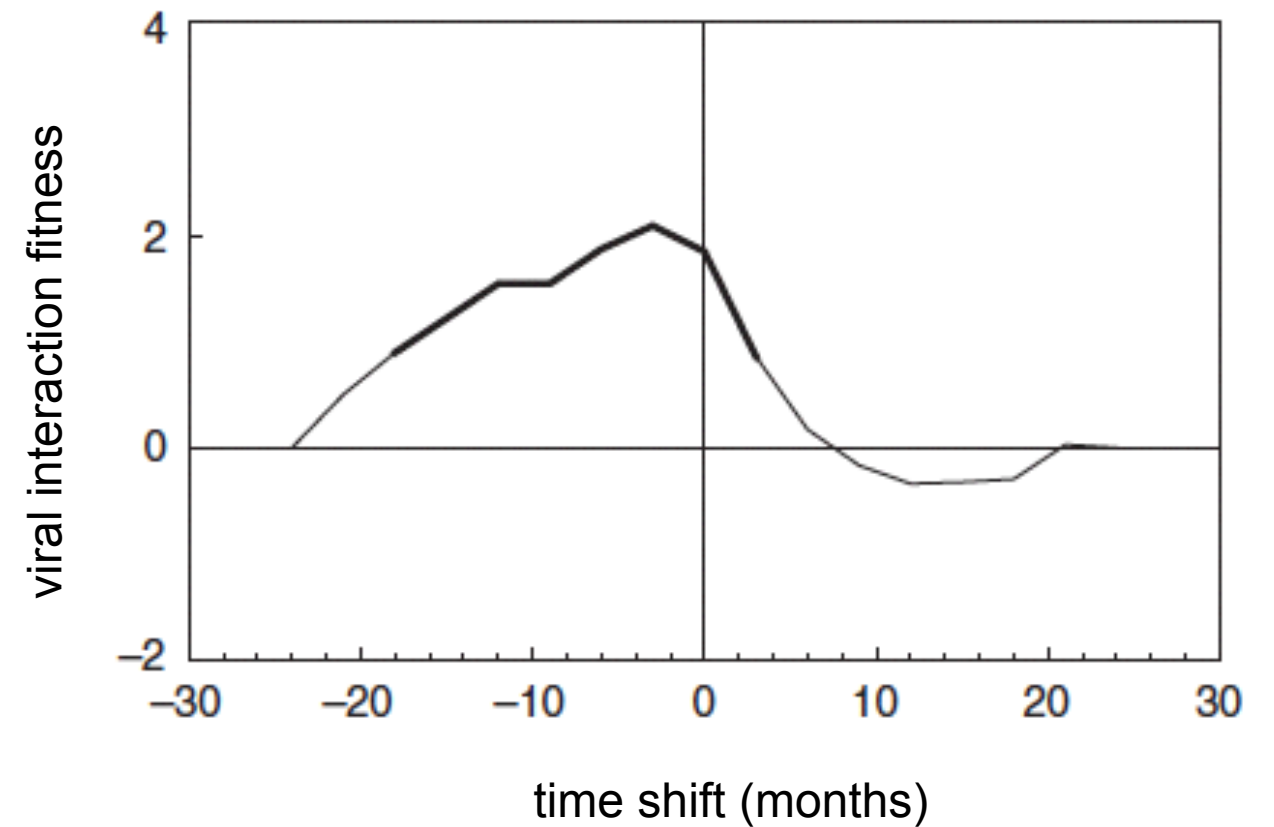
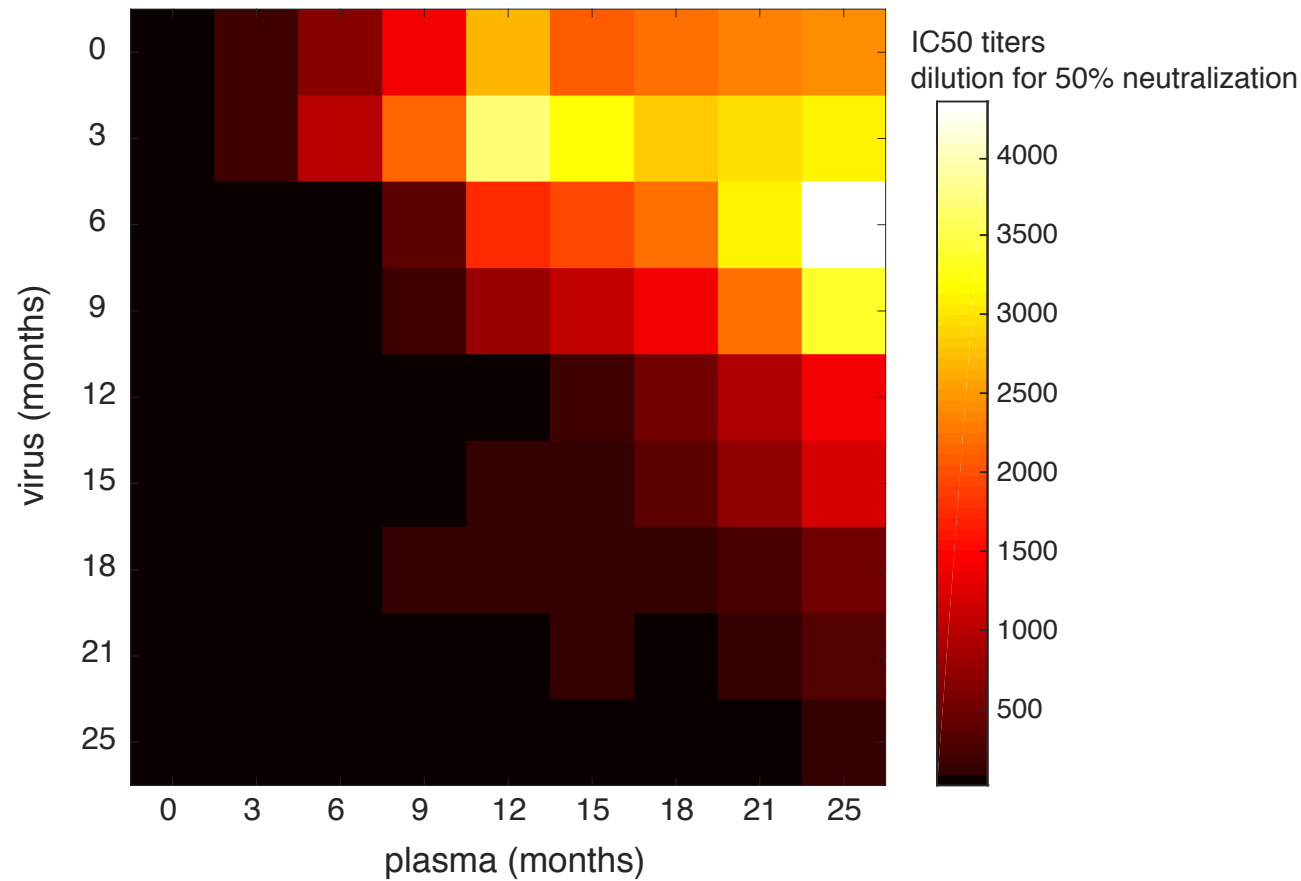
$$F_{A;\tau} = s_a \varepsilon_\tau(t)$$

time-shift experiments in HIV



$$F_v(t) = -\log \text{neut.}(t)$$

time-shift experiments in HIV



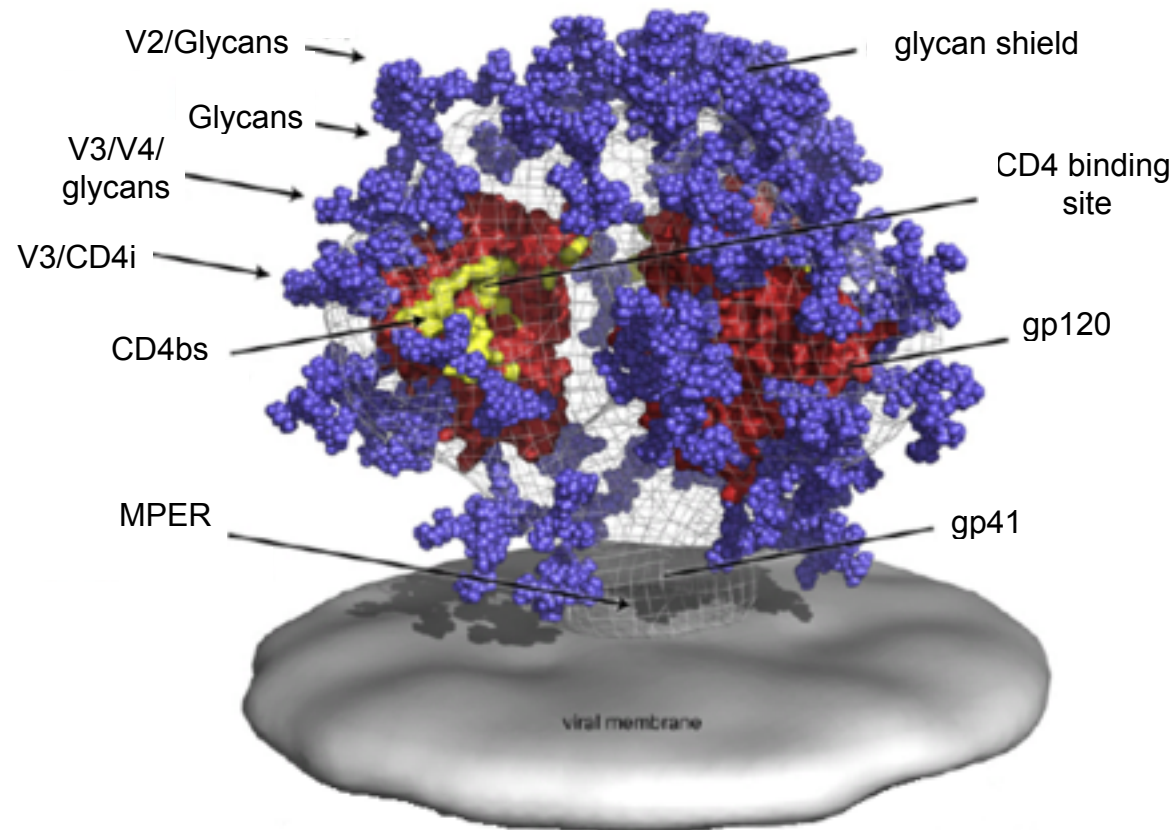
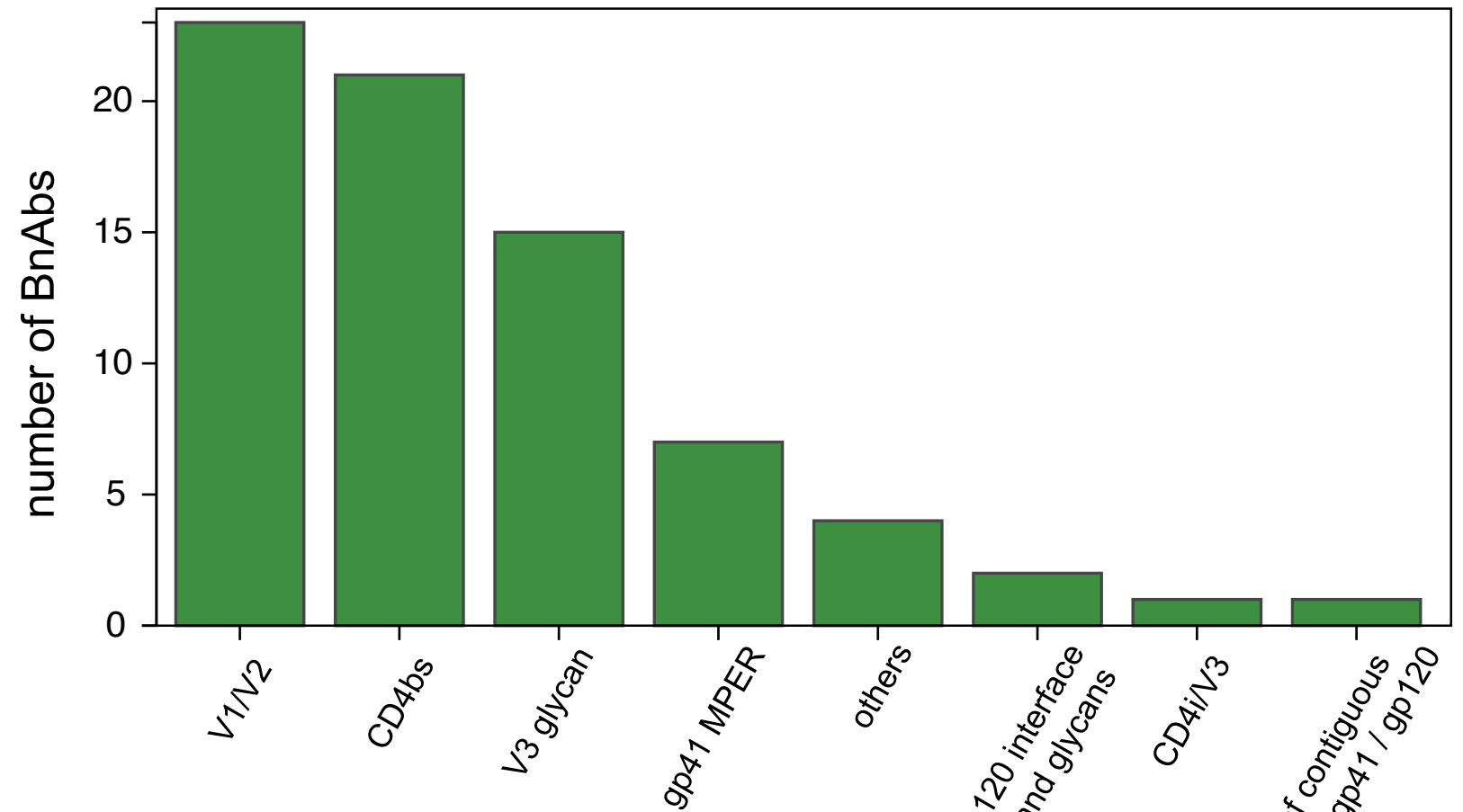
$$F_{v;\tau}(\mathbf{y}_t; \mathbf{x}(t - \tau)) \simeq (F_v(t) + \phi_{v_0} + \dots) + (\phi_{a_0} + \dots) + \underbrace{\sum_{\alpha, \gamma} y^\gamma(t) x^\alpha(t - \tau) \partial_\alpha \partial_\gamma F_v}_{\text{interaction}}$$

fate of an antibody lineage

**fate of a broadly neutralizing
antibody lineage**

BnAb targets

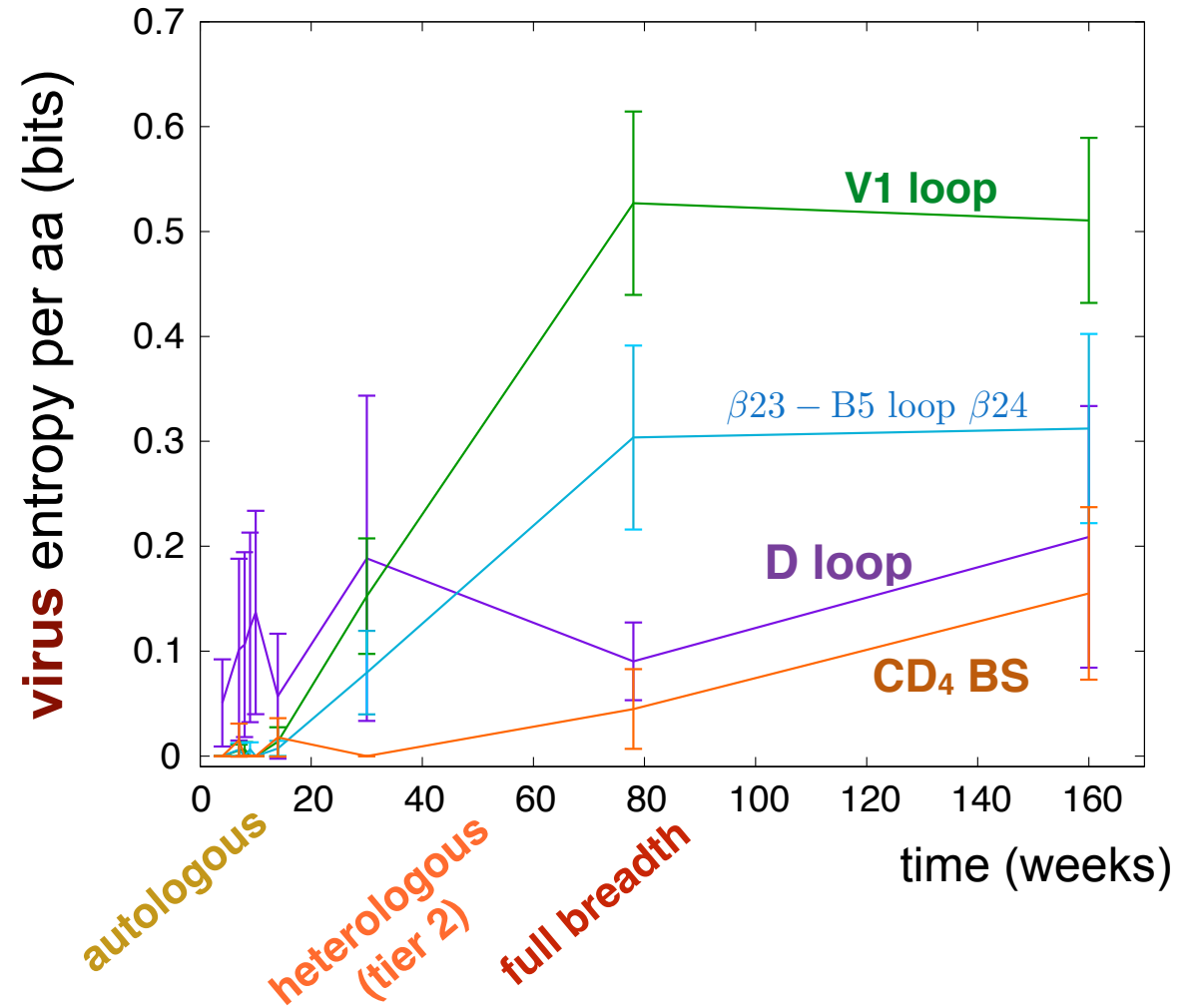
number of unique BnAbs : 87



intra-patient sequence turnover

CH505

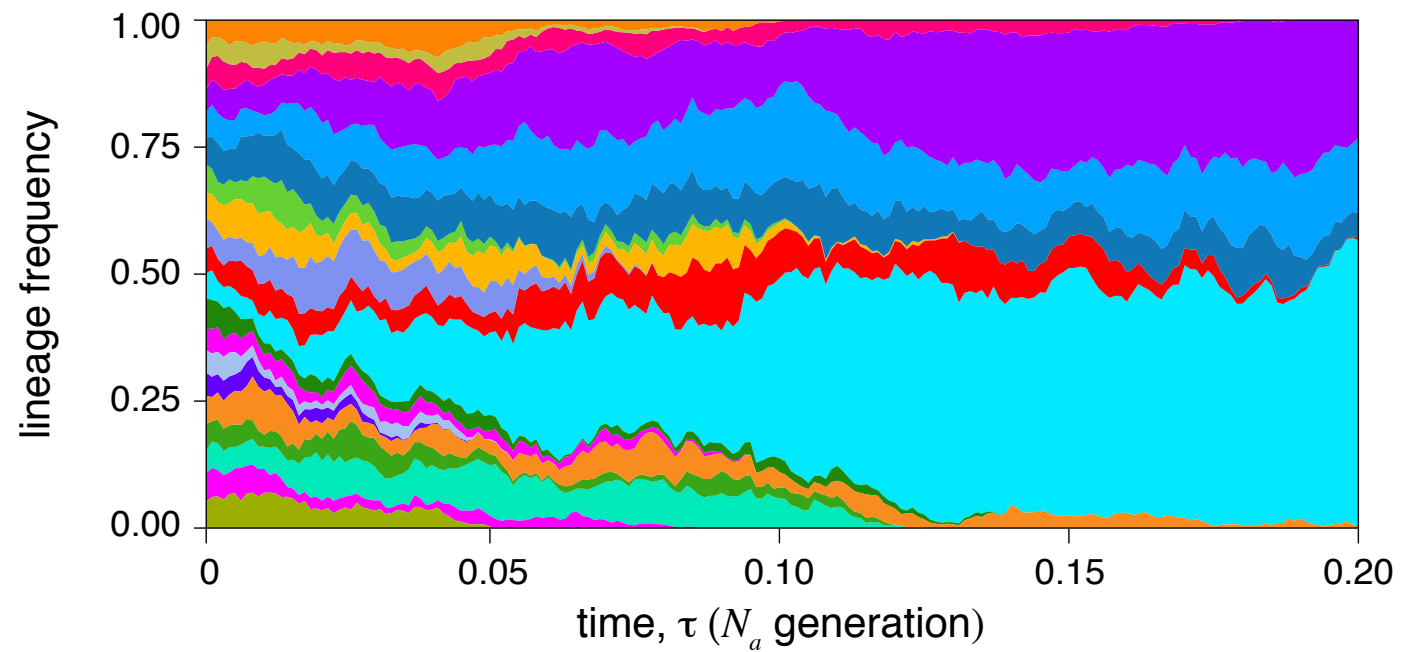
→ restricted escape of CD₄ BS in HIV after full breadth (~week 80)



fate of an antibody lineage

- ➔ within lineage competition
- ➔ between lineage competition

lineage accessibility: $\{\varepsilon_i, \tilde{\varepsilon}_i\}_c$

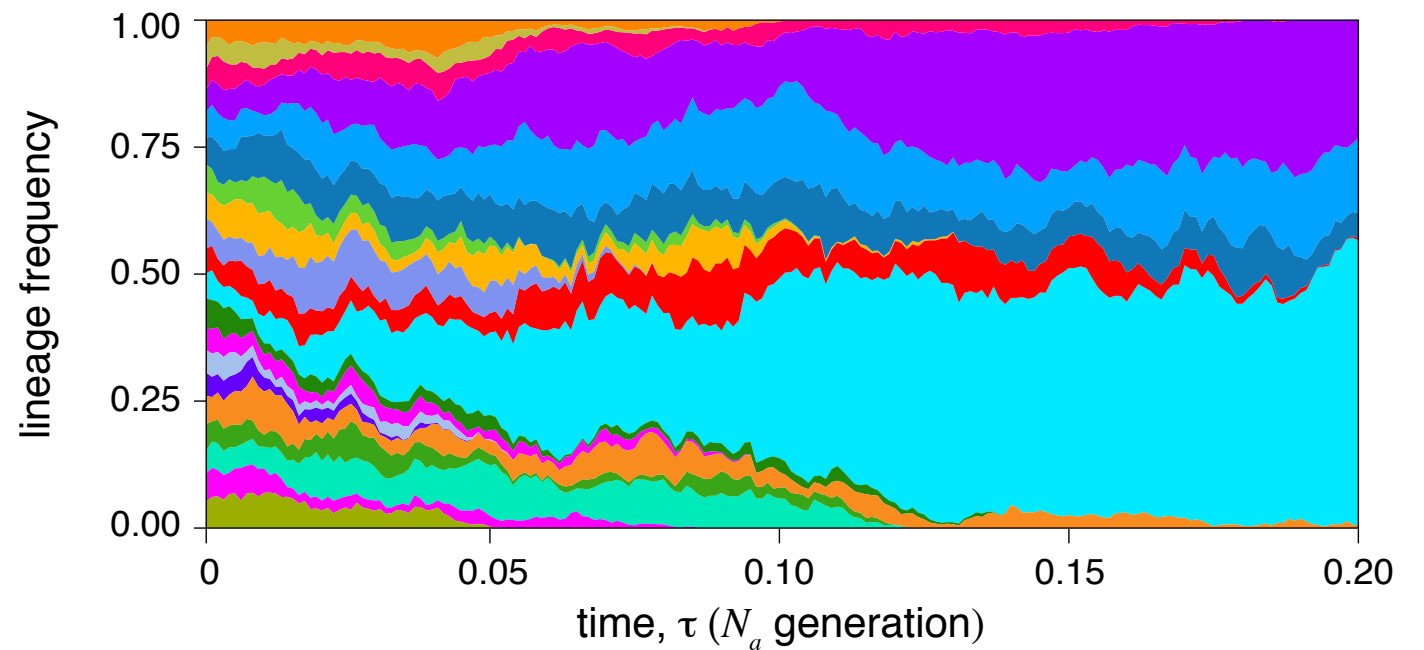


➔ **new lineage ~ clones in an asexual population**

fate of an antibody lineage

- ➔ within lineage competition
- ➔ between lineage competition

lineage accessibility: $\{\varepsilon_i, \tilde{\varepsilon}_i\}_c$



➔ **new lineage ~ clones in an asexual population**

$$\frac{d}{dt}\rho^c = (F_{Ac} - F_A)\rho^c + \sqrt{\frac{\rho^c(1 - \rho^c)}{N_a}}\chi_c$$

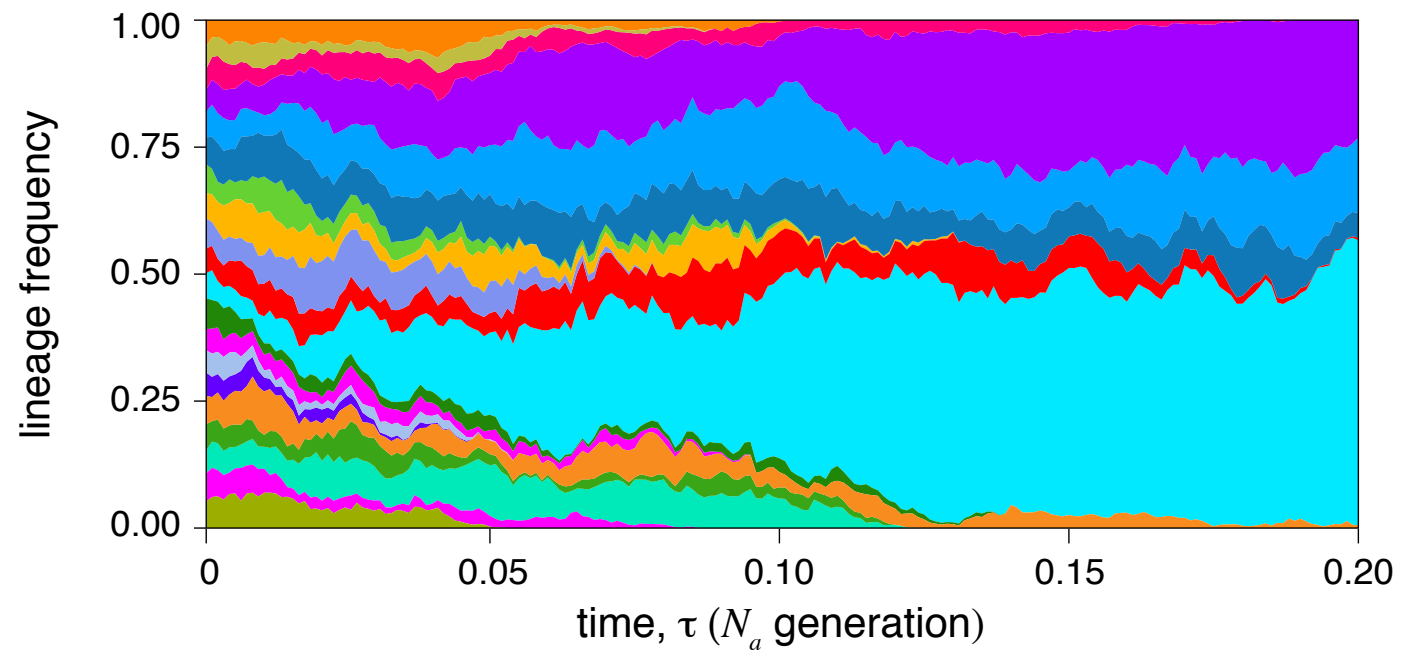
➔ **fixation probability:** asymptotic value of ensemble-averaged lineage frequency

$$P_{\text{fix}}(\mathcal{C}) = \lim_{t \rightarrow \infty} \langle \rho^c(t) \rangle$$

fate of an antibody lineage

- ➔ within lineage competition
- ➔ between lineage competition

lineage accessibility: $\{\varepsilon_i, \tilde{\varepsilon}_i\}_C$



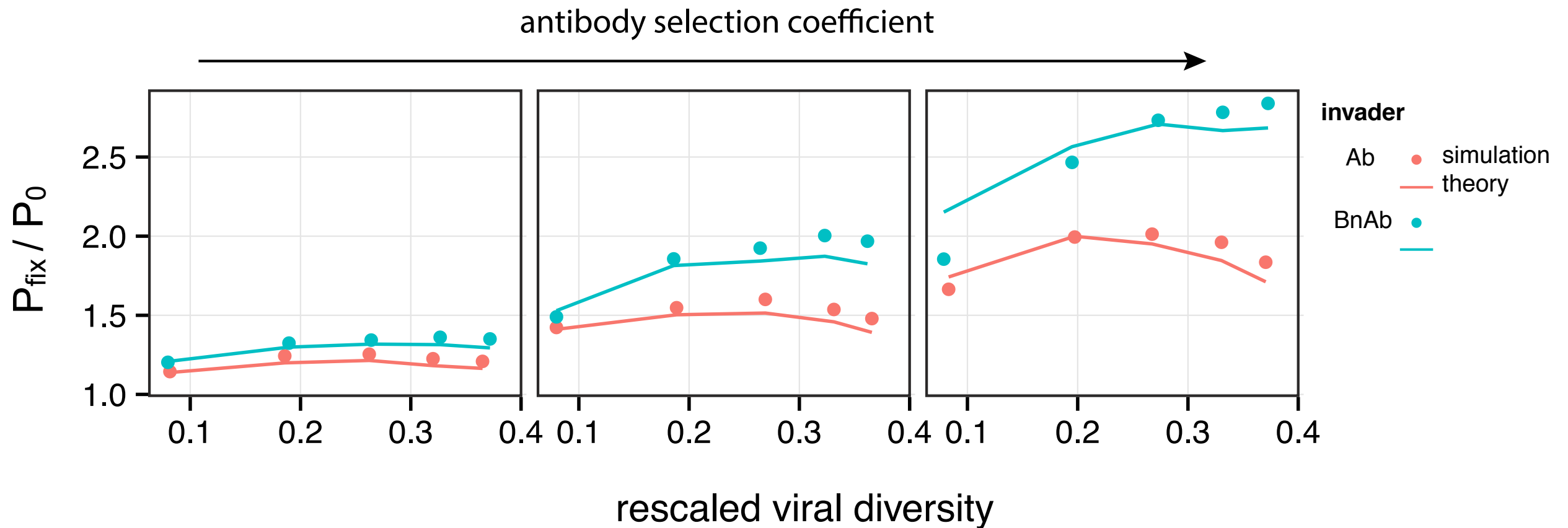
$$\begin{aligned}
 P_{\text{fix}}(\mathcal{C})/P_{0\text{fix}} &\simeq 1 + N_a \underbrace{\langle (F_{Ac}(0) - F_A(0)) \rangle}_{\text{growth rate}} \\
 &+ \frac{N_a^2}{3} \underbrace{\langle \phi_{Ac}(0) - \phi_A(0) \rangle}_{\text{rate of Ab adaptation}} - N_a N_v \underbrace{\langle |\mathcal{T}_{V \rightarrow Ac}(0)| - |\mathcal{T}_{V \rightarrow A}(0)| \rangle}_{\text{rate of viral escape}} + \dots
 \end{aligned}$$

➔ fixation probability depends on:

- efficacy of the new lineage relative to the resident antibodies
- diversity of the viral population

fate of an antibody lineage

➔ higher fixation of **BnAbs** when viral population is diverse (e.g. late in infection)



summary

▶ trait co-evolution in interacting populations

- i. co-evolutionary **observable**: **time-shifted fitness** measurements
- ii. **fitness and transfer flux** characterize coevolution
- iii. **BnAbs** may be induced by controlling the **viral diversity**

