The immune system is a distributed complex system

Rob J de Boer, Utrecht University

Explain to you how the adaptive immune system works:

to highlight its complexity features.

Whenever possible I will use equations.

Ask the question how one should model such a system.

Show an example of host-pathogen co-evolution

(that is typically studied with ODEs).

For that I even need some epidemiology.

Utrecht Center for Quantitative Immunology

Lymphocyte dynamics (modeling deuterium labeling)
life spans of naive and memory T cells

Lymphocyte migration (quantifying 2PM videos)
http://2ptrack.net/: Motilitylab.

Epitope identification (NetMHCpan)
predict pMHC complexes of HIV and cancers

T cell repertoire sequencing (diversity): RTCR
bioinformatic pipeline with superior recall and precision

Universiteit Utrecht

http://tbb.bio.uu.nl/ucqi

Universitair Medisch Centrum
Utrecht

The immune system is a distributed complex system taking decisions on how to respond to molecular patterns in its environment (mango vs cholera).

It memorizes these decisions specifically.

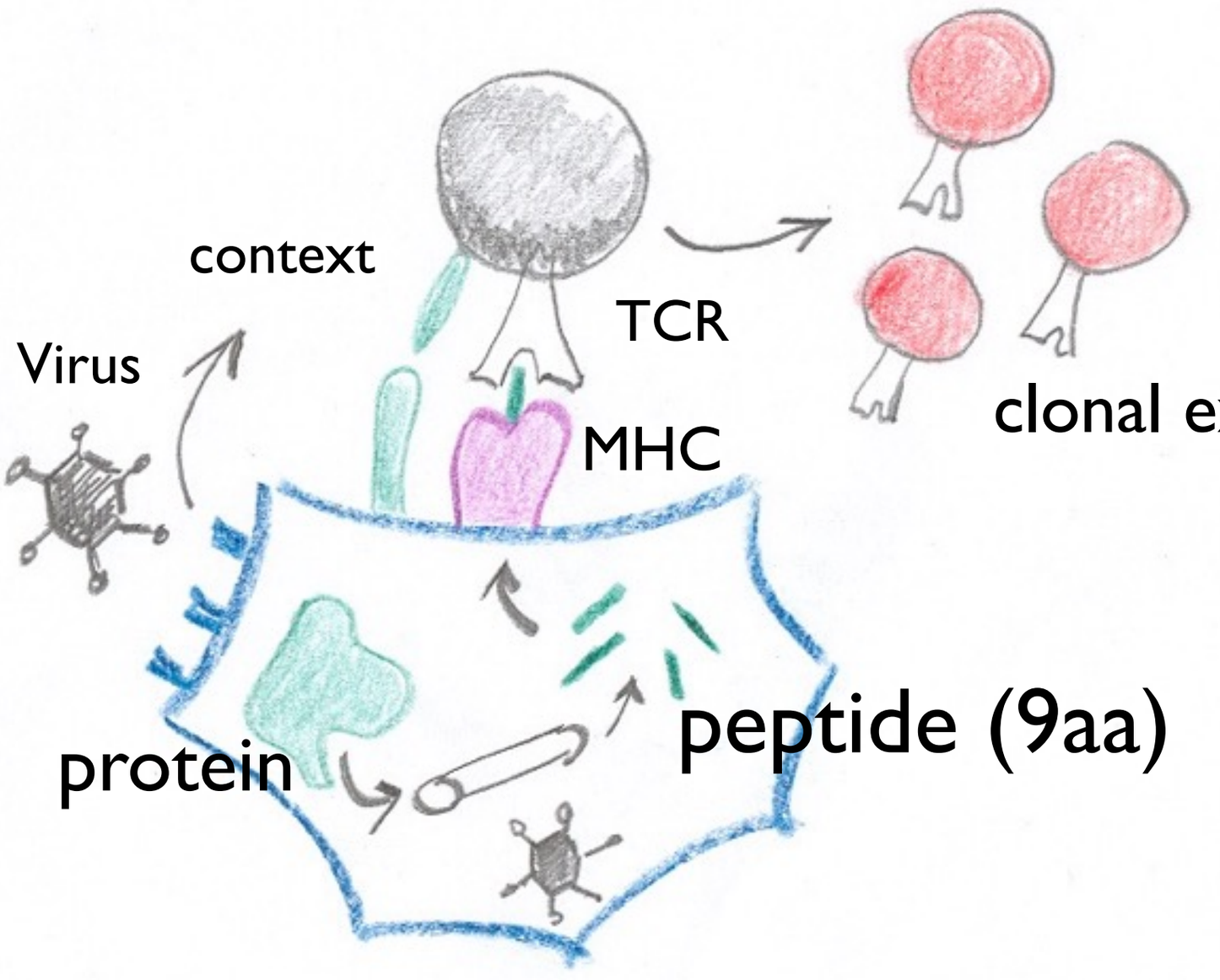
It uses a large array of random detectors.

It protects its host to rapidly evolving pathogens.

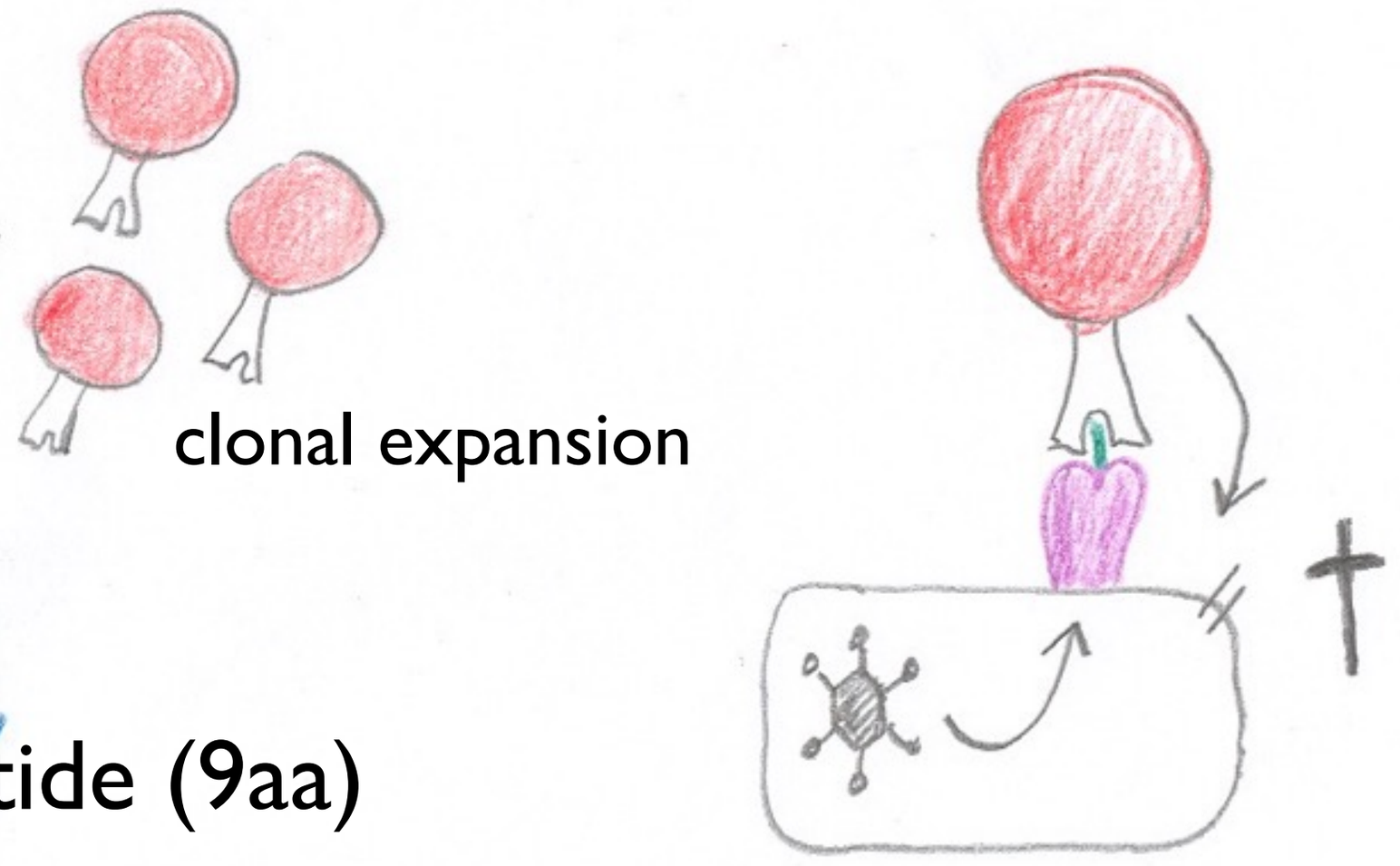
During immune responses we see rapid biological evolution on a time scale of weeks

Immune systems samples a few peptides and stores contextual information in memory cells

naive T cell



educated T cells



kill target using the info from one peptide

Diversity of TCR repertoire and polymorphism of MHC

Lymphocyte receptors are made by randomly assorting gene segments and deliberate random mutations:

diversity of potential receptors (10^{20}) >> number of genes (10^{4})

Each circulating T cell is a random detector expressing a unique receptor.

After stimulation lymphocytes adopt a particular memory phenotype.

MHC molecules presenting the peptides to these receptors are **polymorphic**:

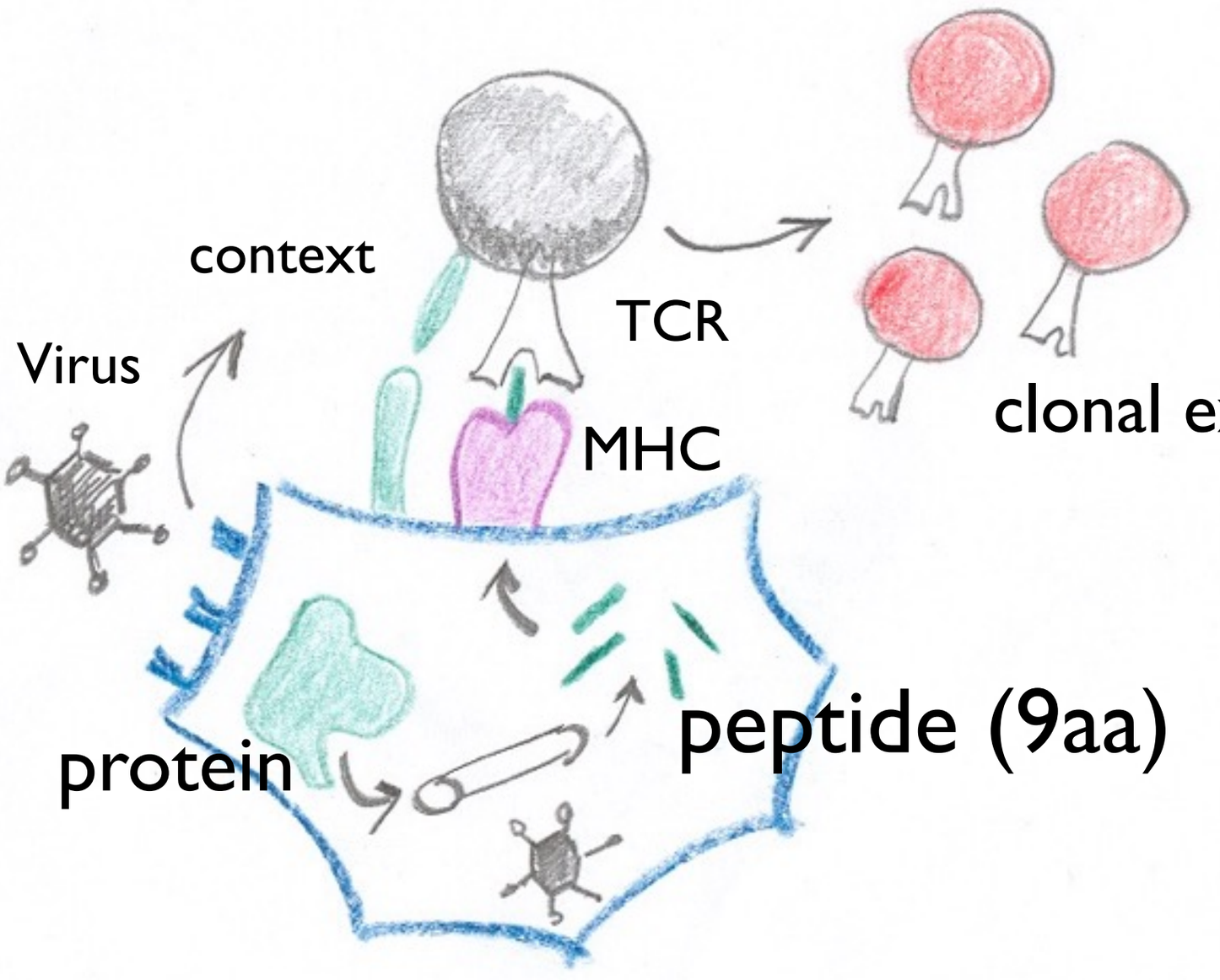
thousands of alleles in the population.

We all sample different peptides from the same pathogen.

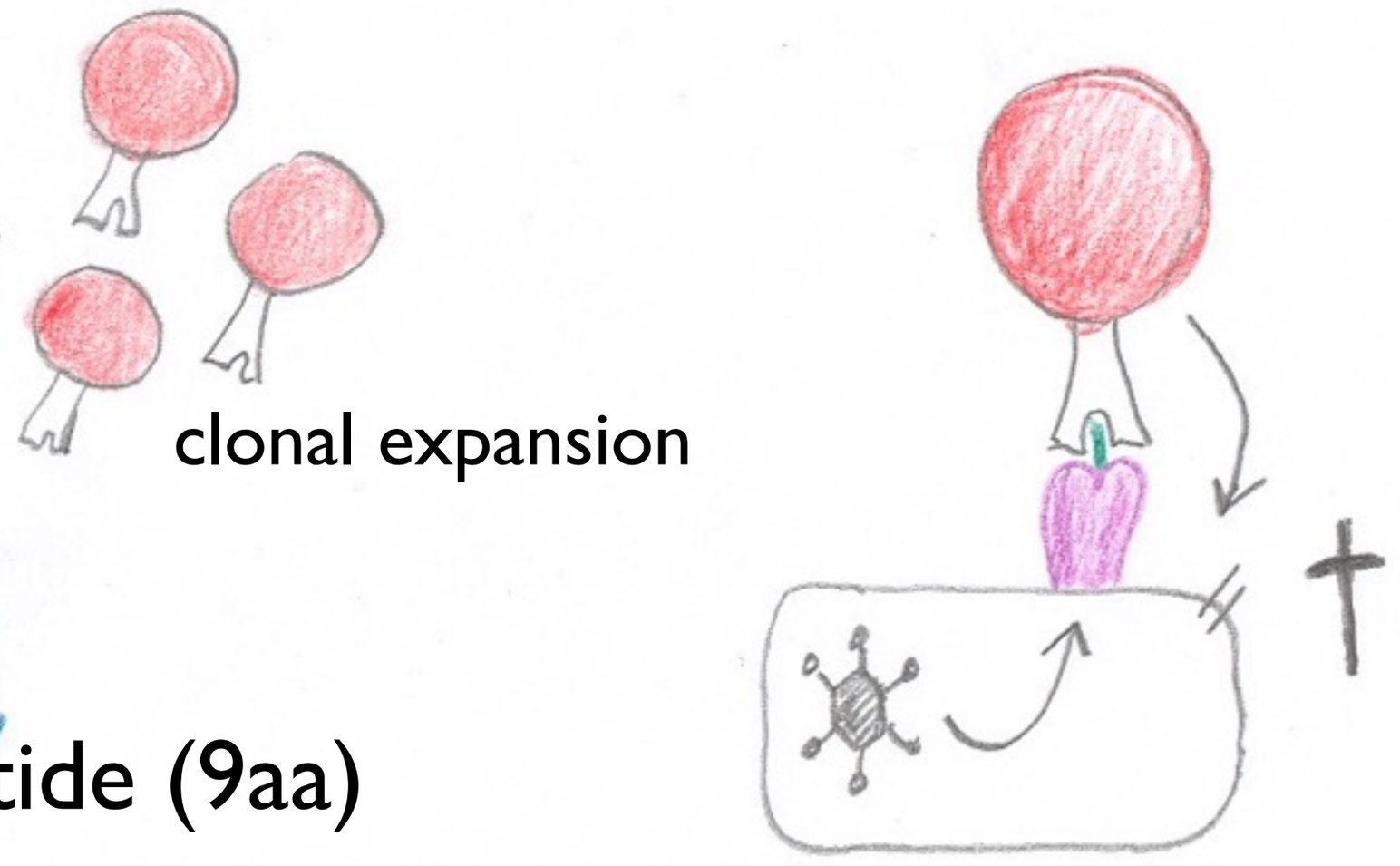
Due to a rare allele advantage: its is good to be different

Immune systems samples a few peptides and stores contextual information in memory cells

naive T cell



educated T cells



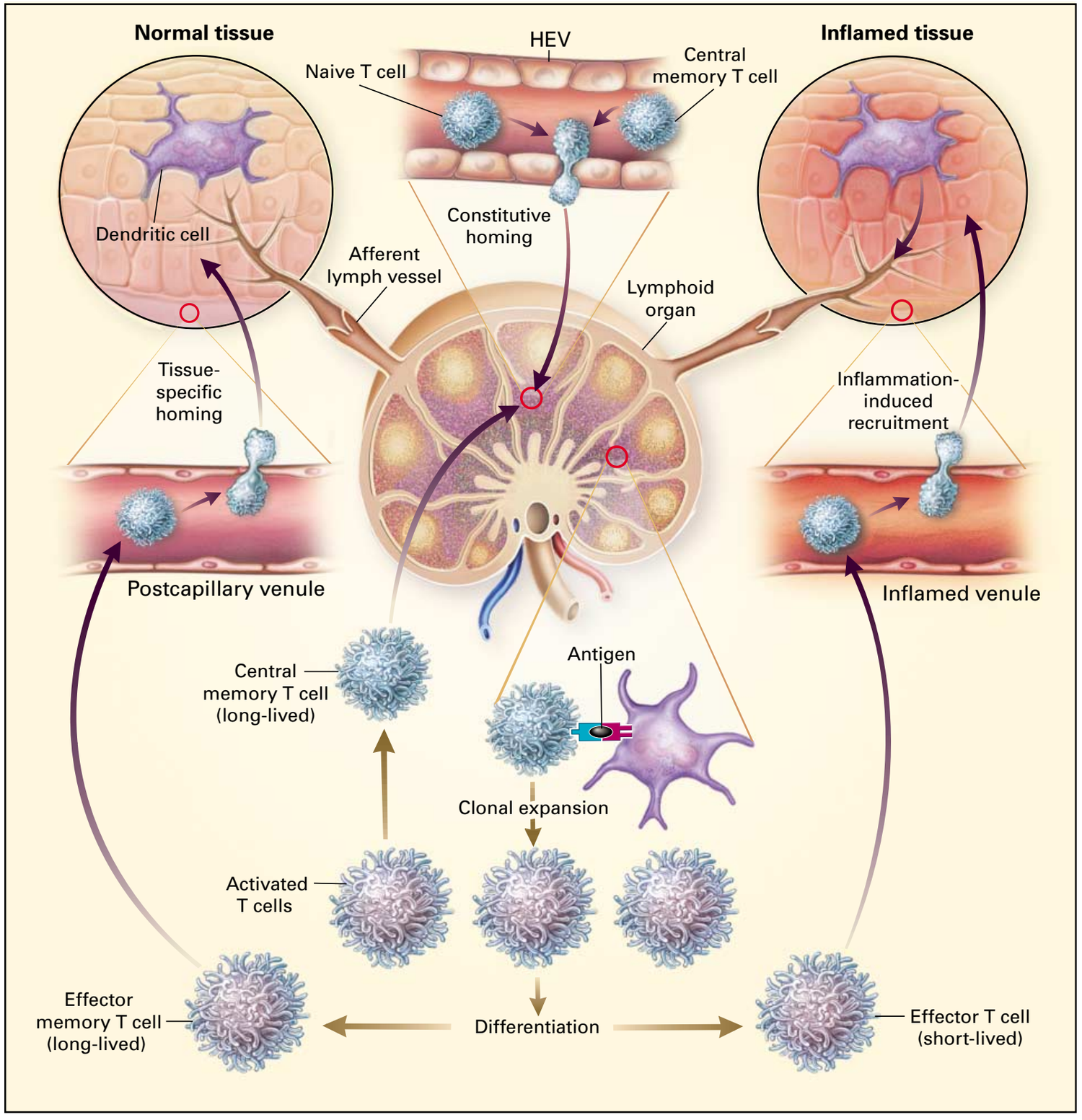
kill target using the info from one peptide

Immune responses develop in draining lymph nodes

Dendritic cells (DC) scan peripheral tissues, and migrate to draining lymph nodes to present their antigens.

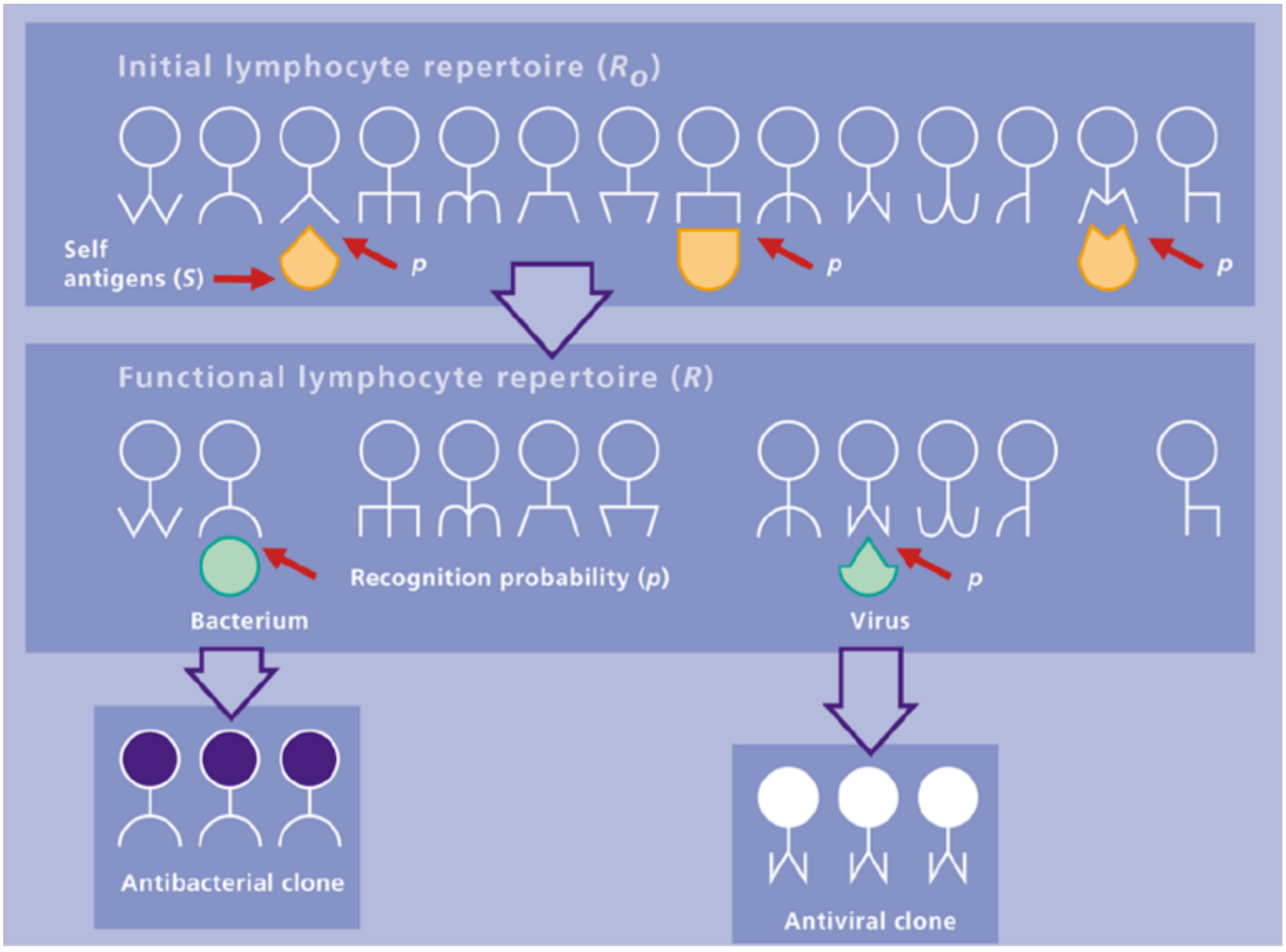
Millions of different naive T cells migrate through lymph nodes, and bind these DC.

Only 1:100000 T cells will become activated, expand, and emigrate as effector cells that move back to the inflamed tissues.



Von Andrian & Mackay, NEJM, 2000

Lymphocyte receptors are made randomly by (VDJ) recombination and mutation



$$R_0 = 10^9$$

$$S = 10^5$$

10^7 9-mers,
1% on MHC

$$R = R_0(1-p)^S$$

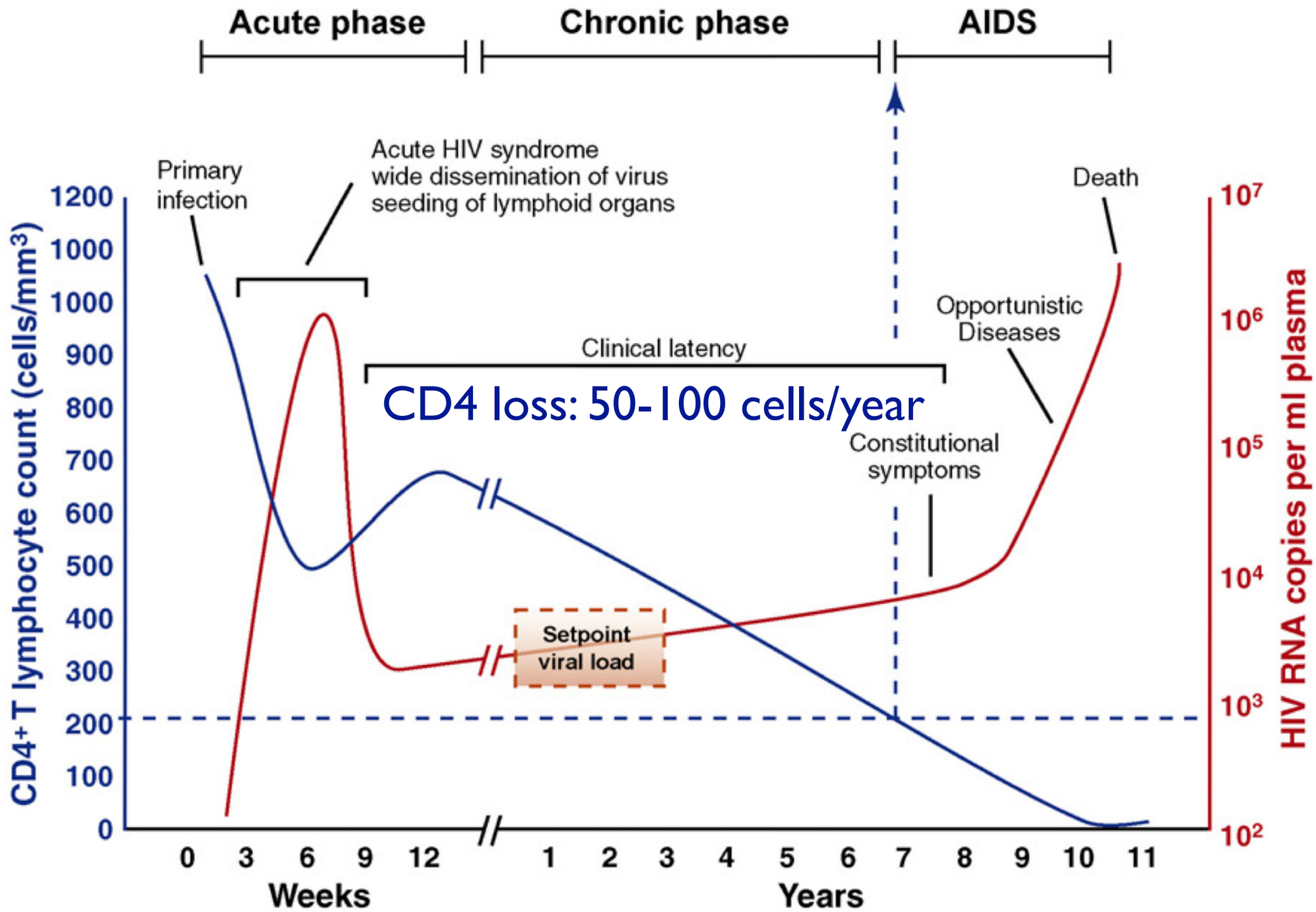
$$P_i = 1 - (1-p)^R$$

Optimize (Pi'=0):

$$p = \frac{1}{S} \simeq 10^{-5}$$

Receptors have to specific to avoid massive deletion [De Boer, Perelson, Borghans, 1993, 1999]

Let me introduce the pathogen: HIV

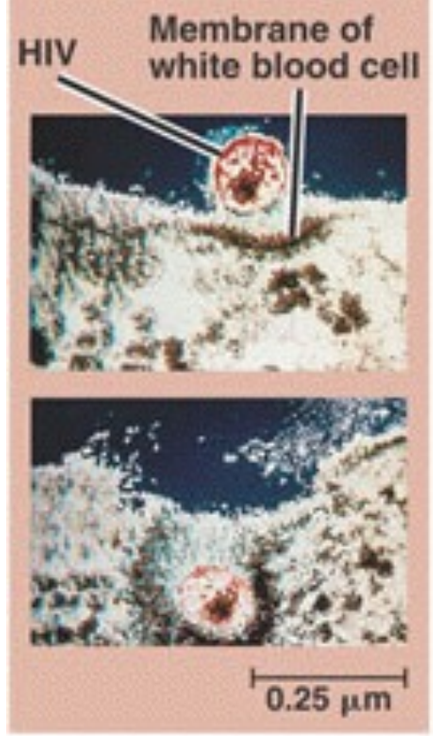


O'Brien & Hendrickson, Genome Biol. 2013

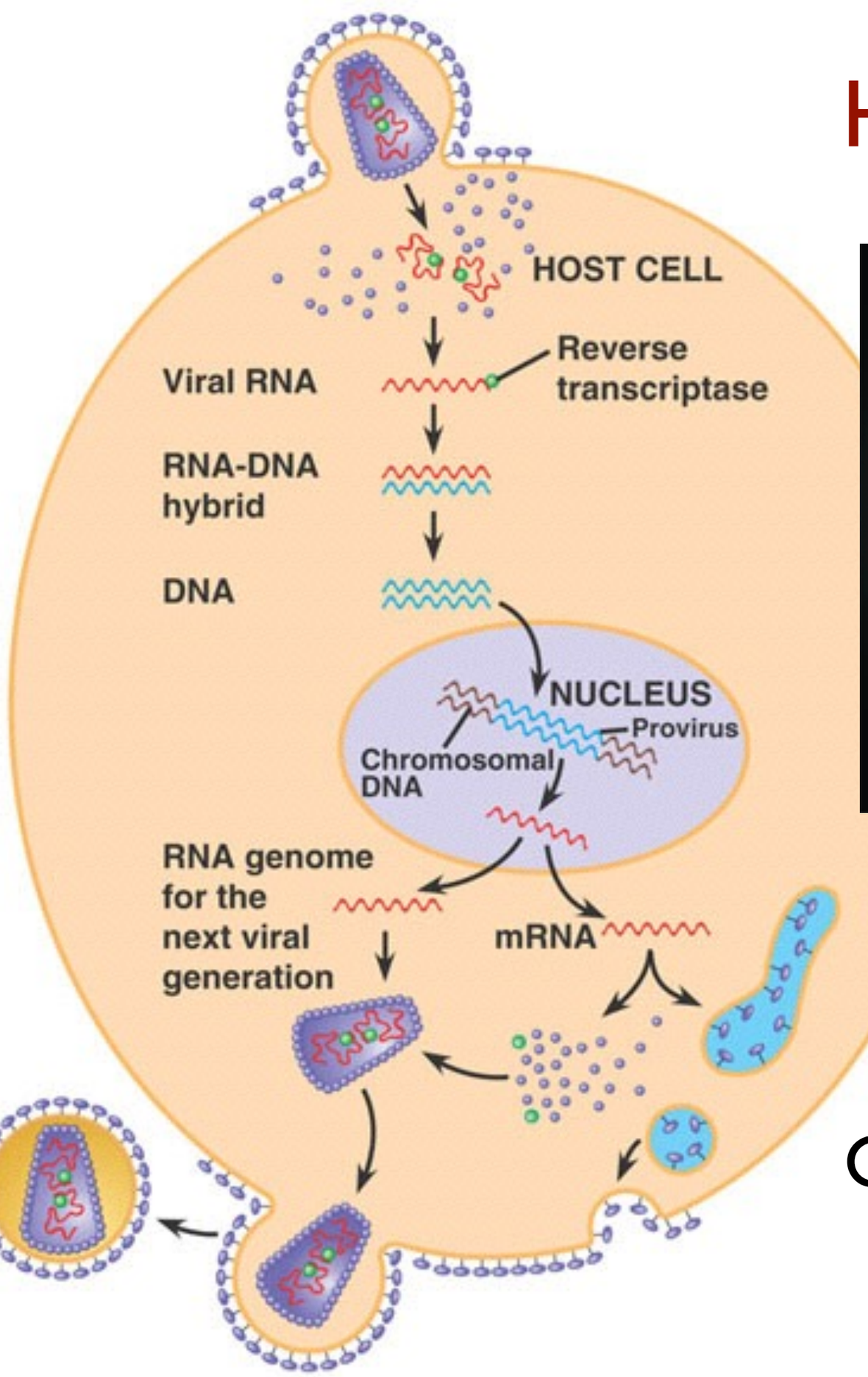
Slow decline of $CD4^+$ T cells: AIDS due to loss of immunity

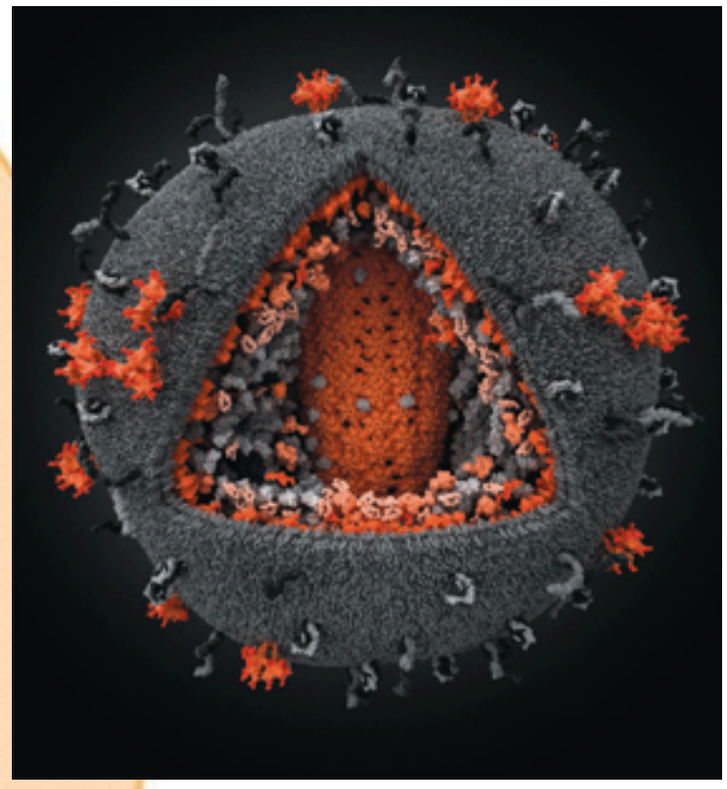
Fairly stable viral setpoint for many years: time to AIDS

HIV life cycle



HIV entering a cell





Konstantinov Science 2011

New HIV leaving a cell

CD4$^+$ T cell

10^8 in 5×10^{11}

Viral dynamics during chronic phase inferred by modeling

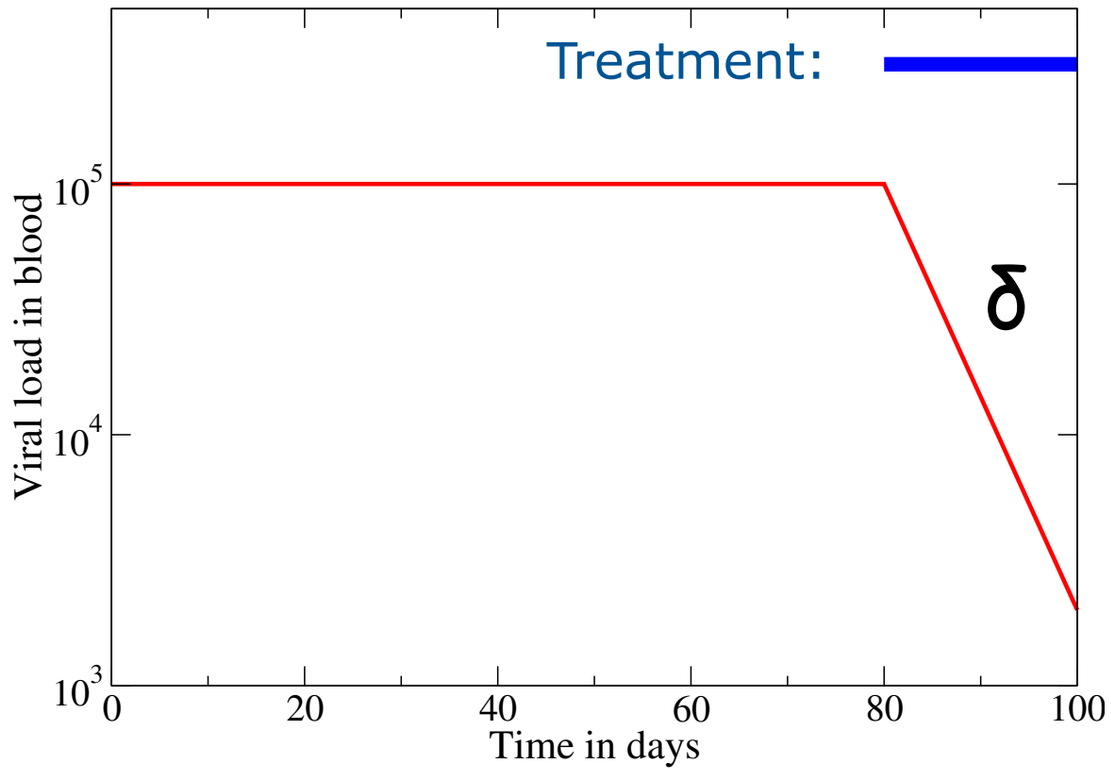
ARTICLES

Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection

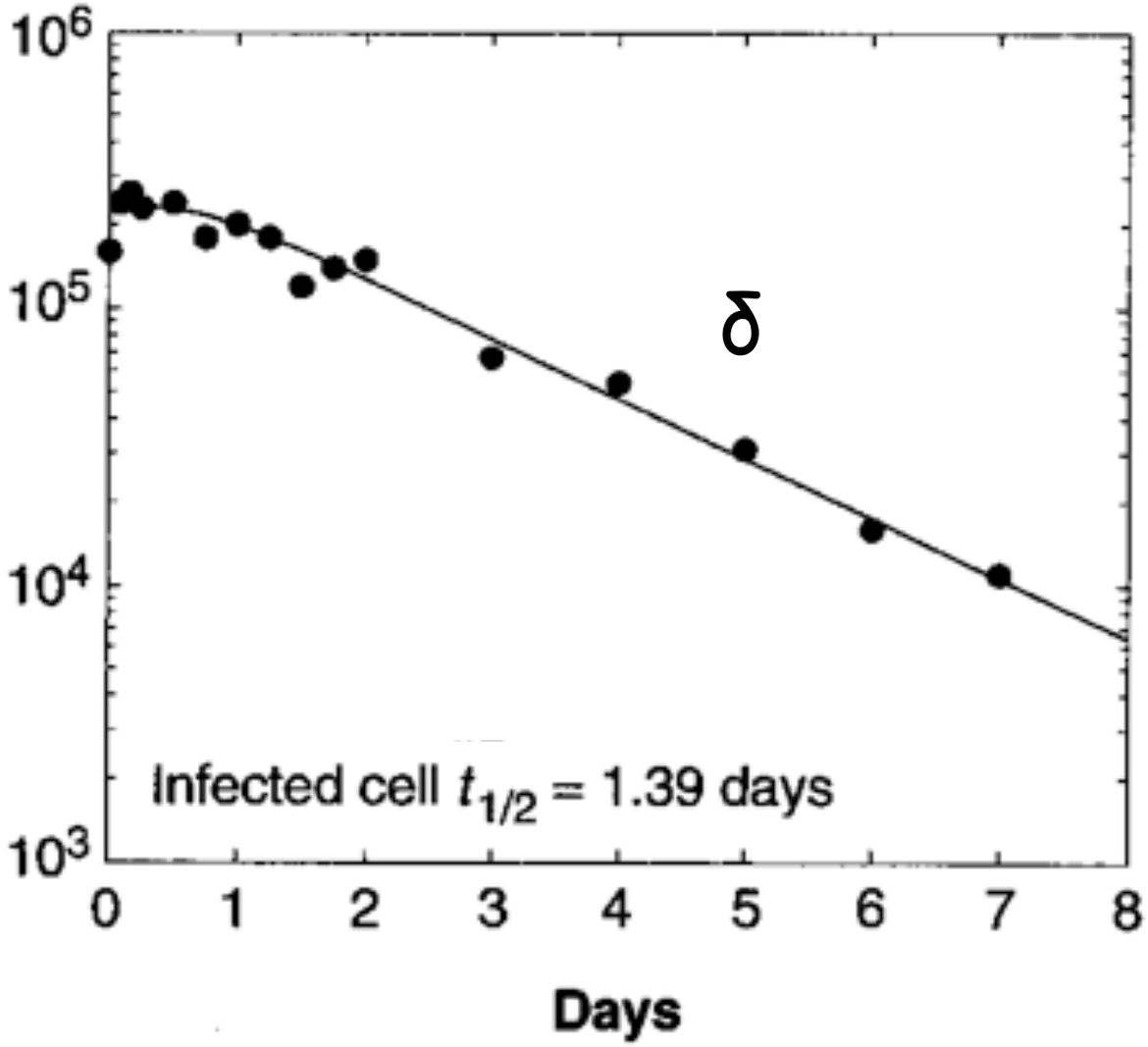
David D. Ho, Avidan U. Neumann*†, Alan S. Perelson†, Wen Chen, John M. Leonard‡ & Martin Markowitz

Nature 1995

Determine life span of infected cells from treatment data



Ho and Perelson, Nature 1995, Science 1998



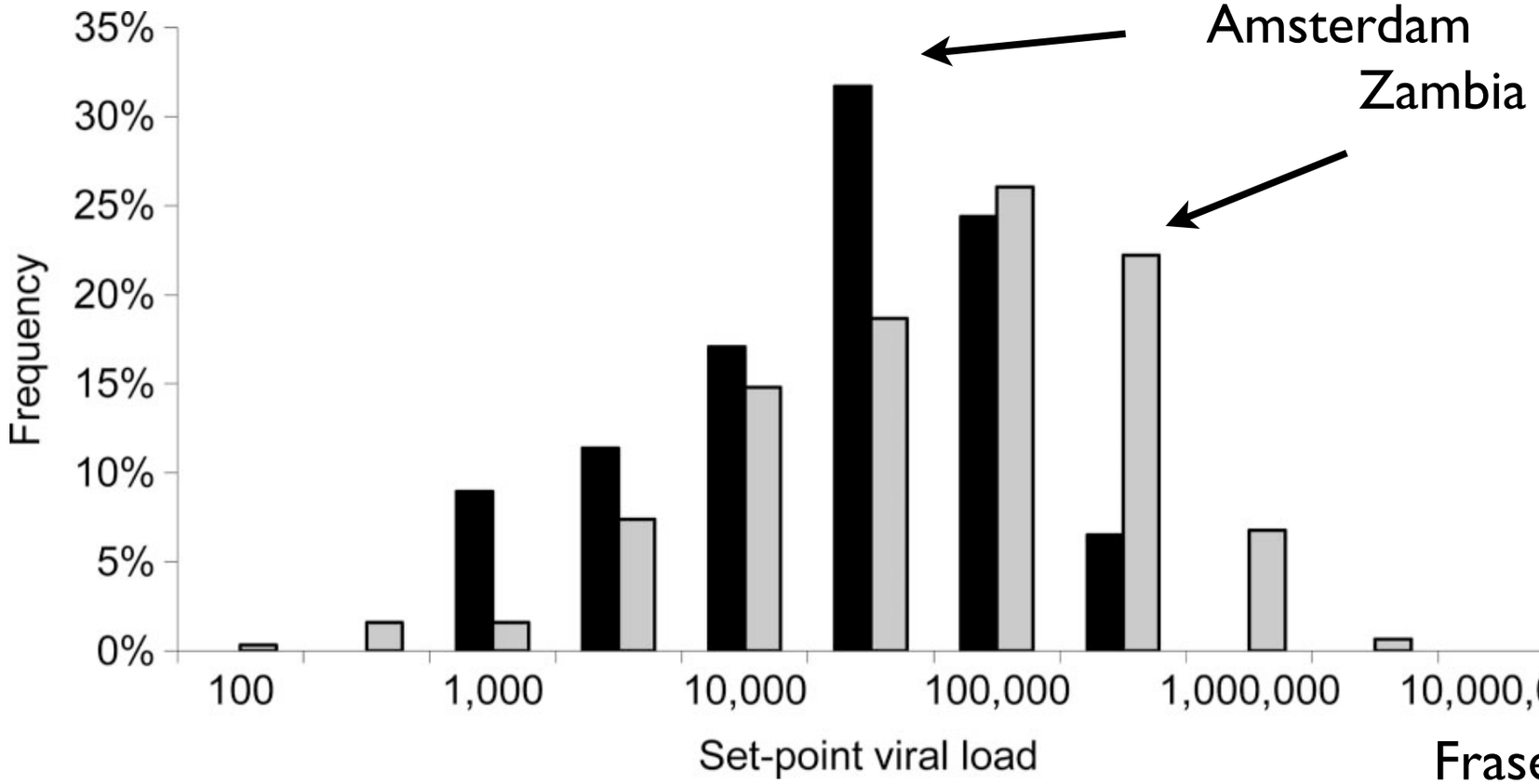
$$\frac{\mathrm{d}T}{\mathrm{d}t} = \sigma - \delta_T T - \beta' T I$$

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \beta' T I - \delta I$$

Death rate δ is estimated from down-slope during therapy

Generation time of 1-2 days!

Epidemiology: HIV-1 set-points vary orders of magnitude

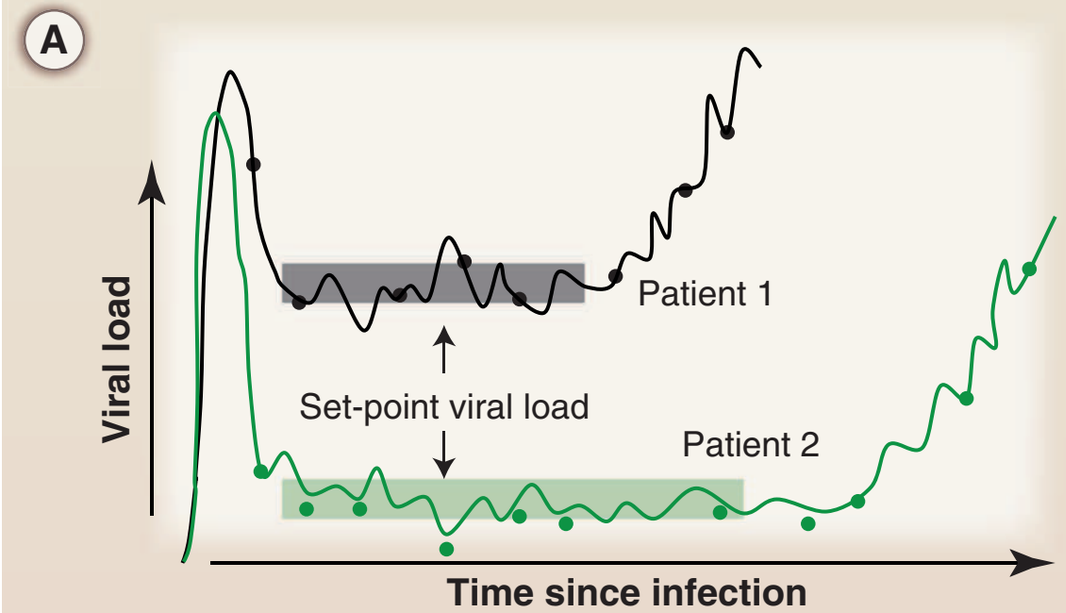


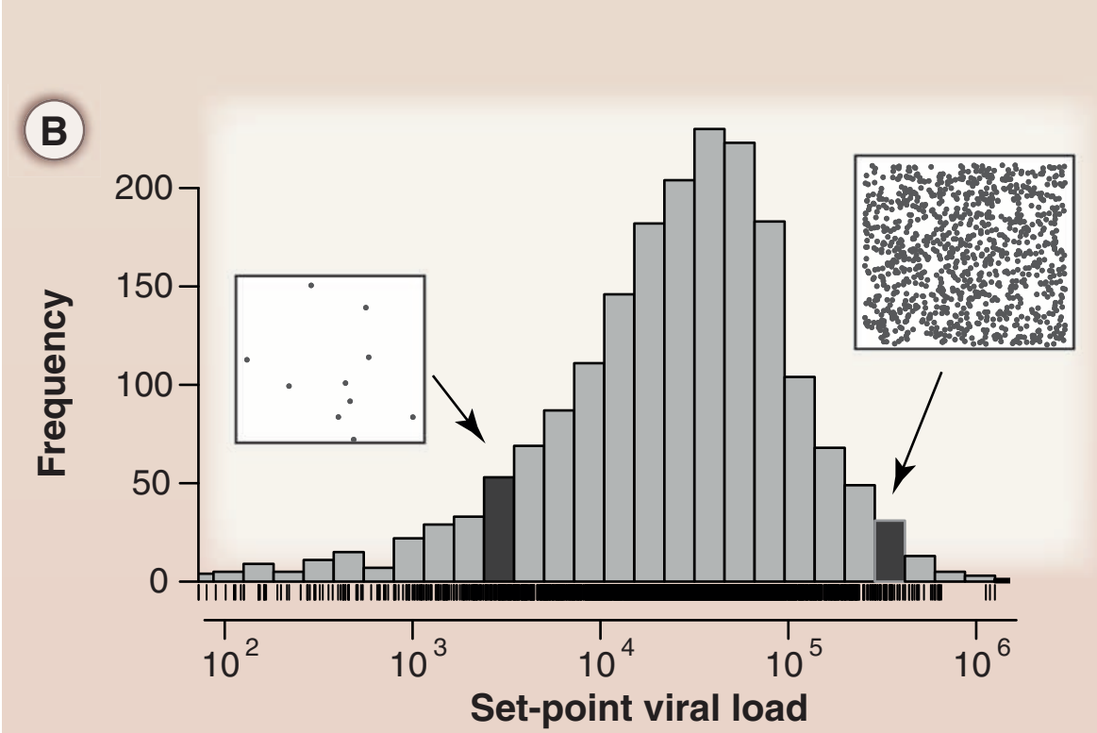
Fraser et al. PNAS 2007

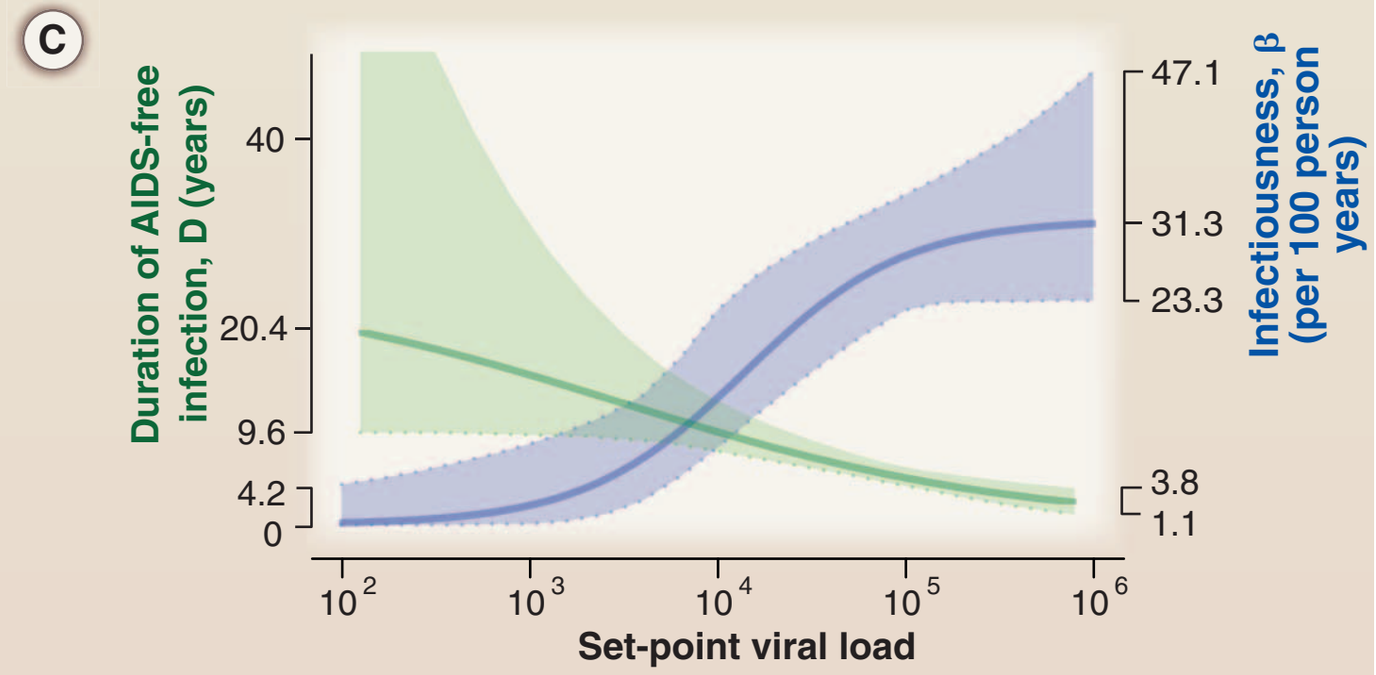
Partly due to host factors (MHC) & Δ-32 deletion in CCR5
Partly due viral factors: heritability & crippling mutations

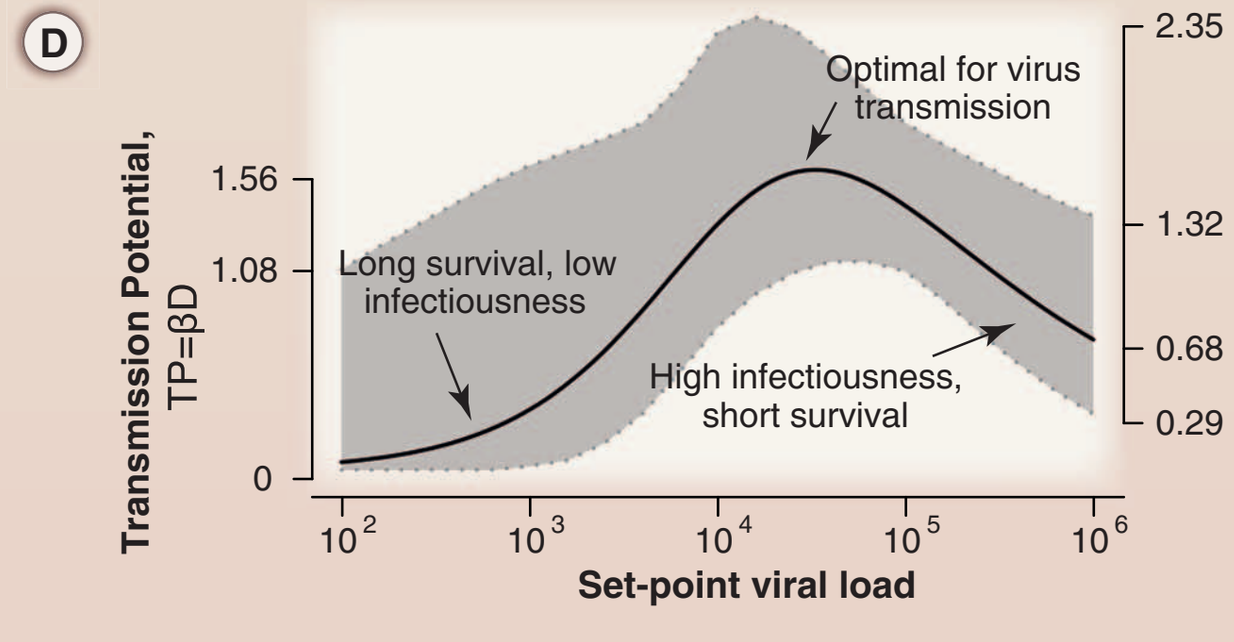
Major unresolved problem in HIV research

Is the HIV-1 set-point optimized for transmission?





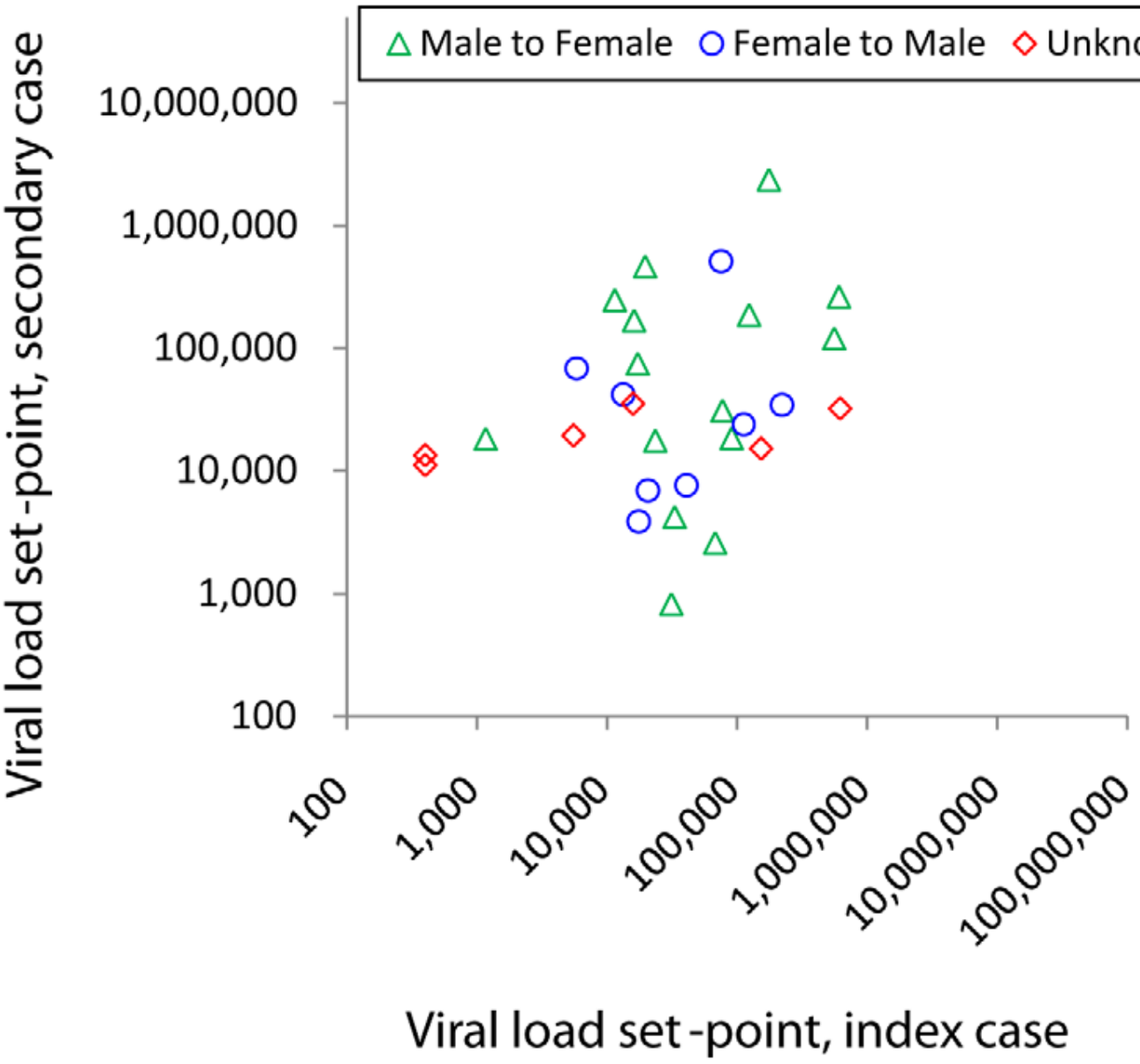




Fraser et al. PNAS 2007, Science 2014

Classical trade-off between virulence and transmission

Set-point viral load is heritable



Hollingsworth et al, PLoS Path. 2010

N couples	Heritability	Study (reference)
97	36% (6 to 66%)	Hollingsworth *et al.* (*17*)
141	44% (19 to 69%)	Lingappa *et al.* (*18*)
195	26% (8 to 44%)	Yue *et al.* (*19*)
433	33% (20 to 46%)	Overall summary estimate (weighted by standard error)

Ample heritability to allow for viral evolution

Fraser et al, Science, 2014

Viral factors accounting for variation never identified

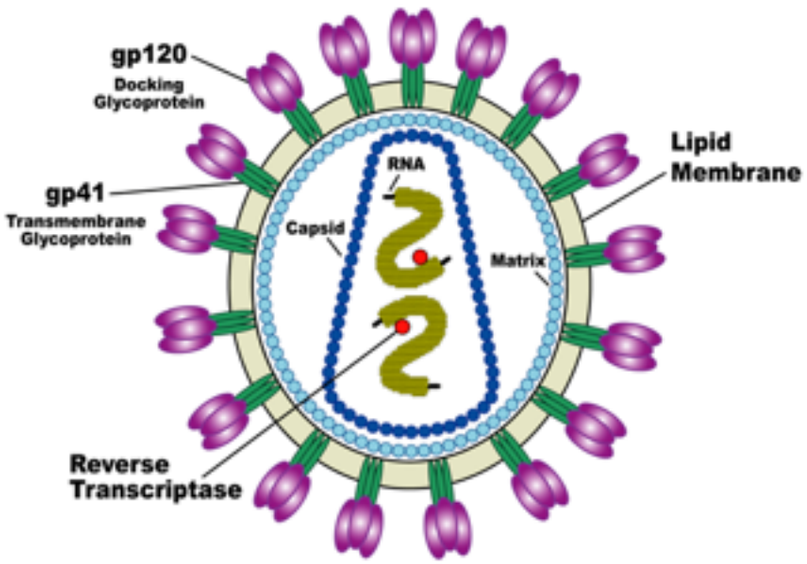
Adaptation at the epidemic level is unexpected

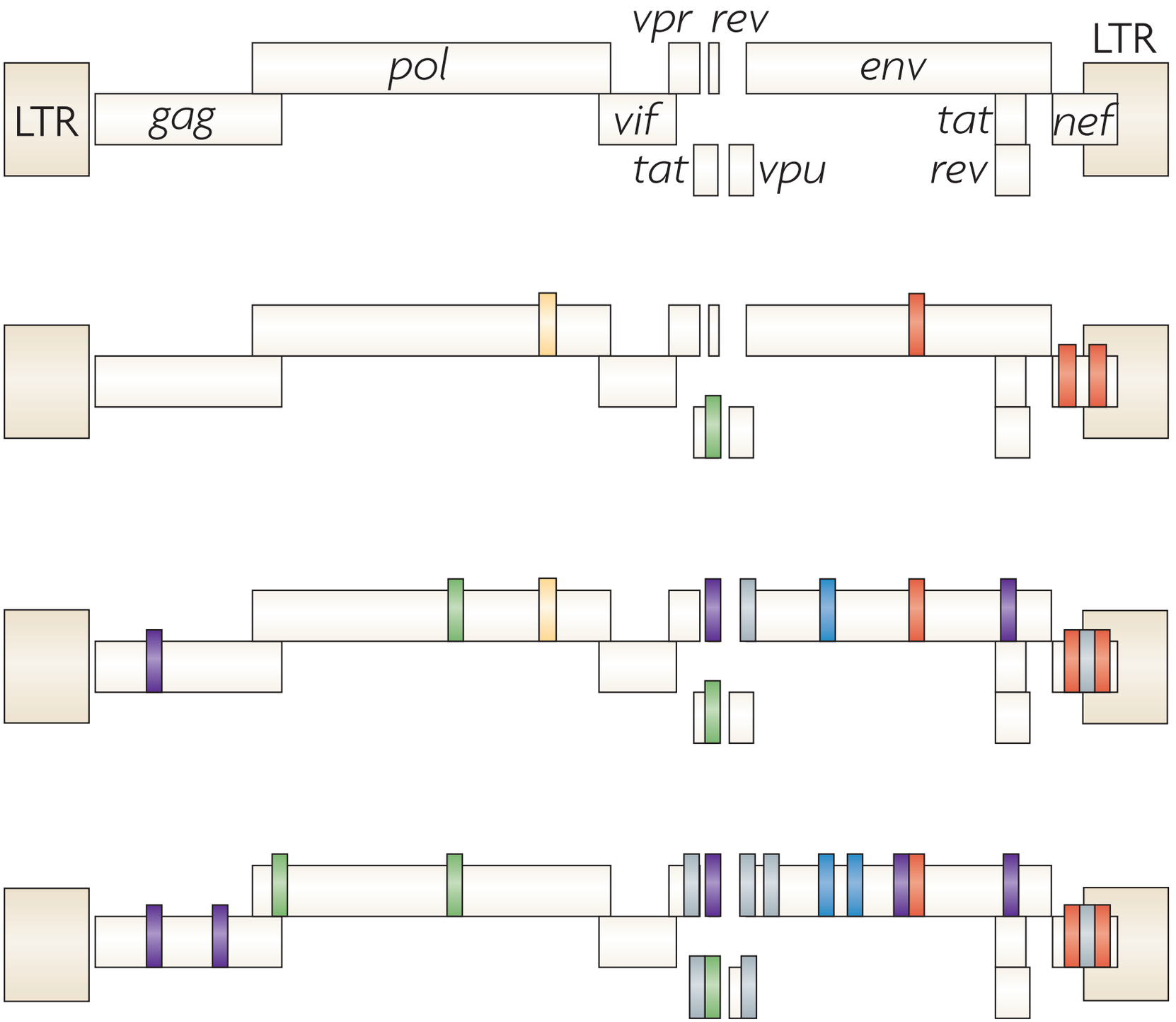
Transmission events are separated by hundreds of generations where HIV-1 is evolving within a single host.

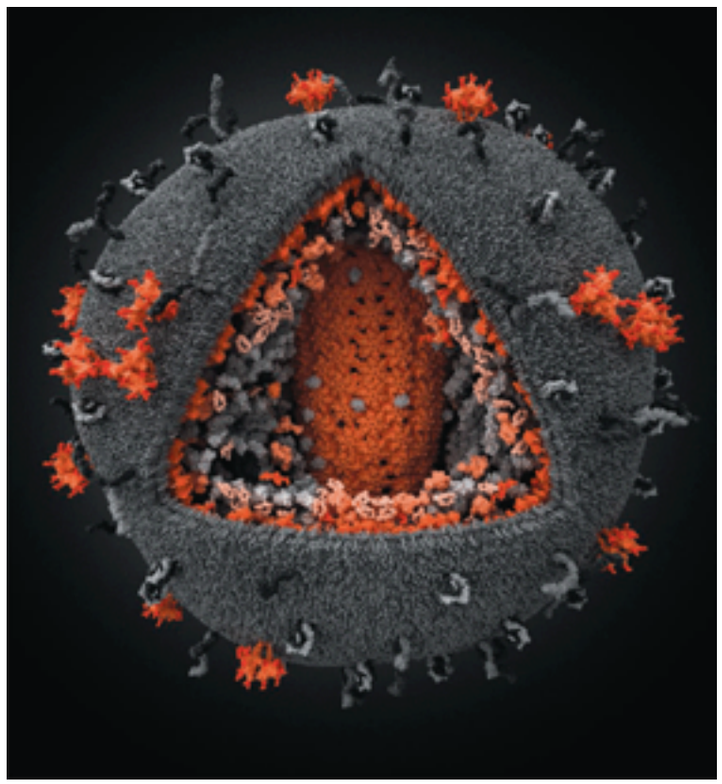
Within each host strains compete and evolve

Indeed HIV transmits quite poorly

Immunology: HIV rapidly accumulates immune escape mutations







Konstantinov Science 2011

McMichael Nature Rev Immunol 2010

Immunology brings **host heterogeneity** to epidemiology

Not only **phenotypic** heterogeneity (immunity)
but importantly also **genetic** heterogeneity

Major factor is the extremely polymorphic MHC (HLA)

1. we are all different,
2. almost everyone is heterozygous,
3. most SNP correlations with disease in MHC region.

Major strategy of the host is to be different

Epidemiology: the SIR model
for Susceptible, Infected and Recovered

$$\frac{\mathrm{d}S}{\mathrm{d}t} = rN(1-N/k) - dS - \beta SI$$

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \beta SI - (d+\delta+\gamma)I$$

$$\frac{\mathrm{d}R}{\mathrm{d}t} = \gamma I - dR$$

$$N = S + I + R$$

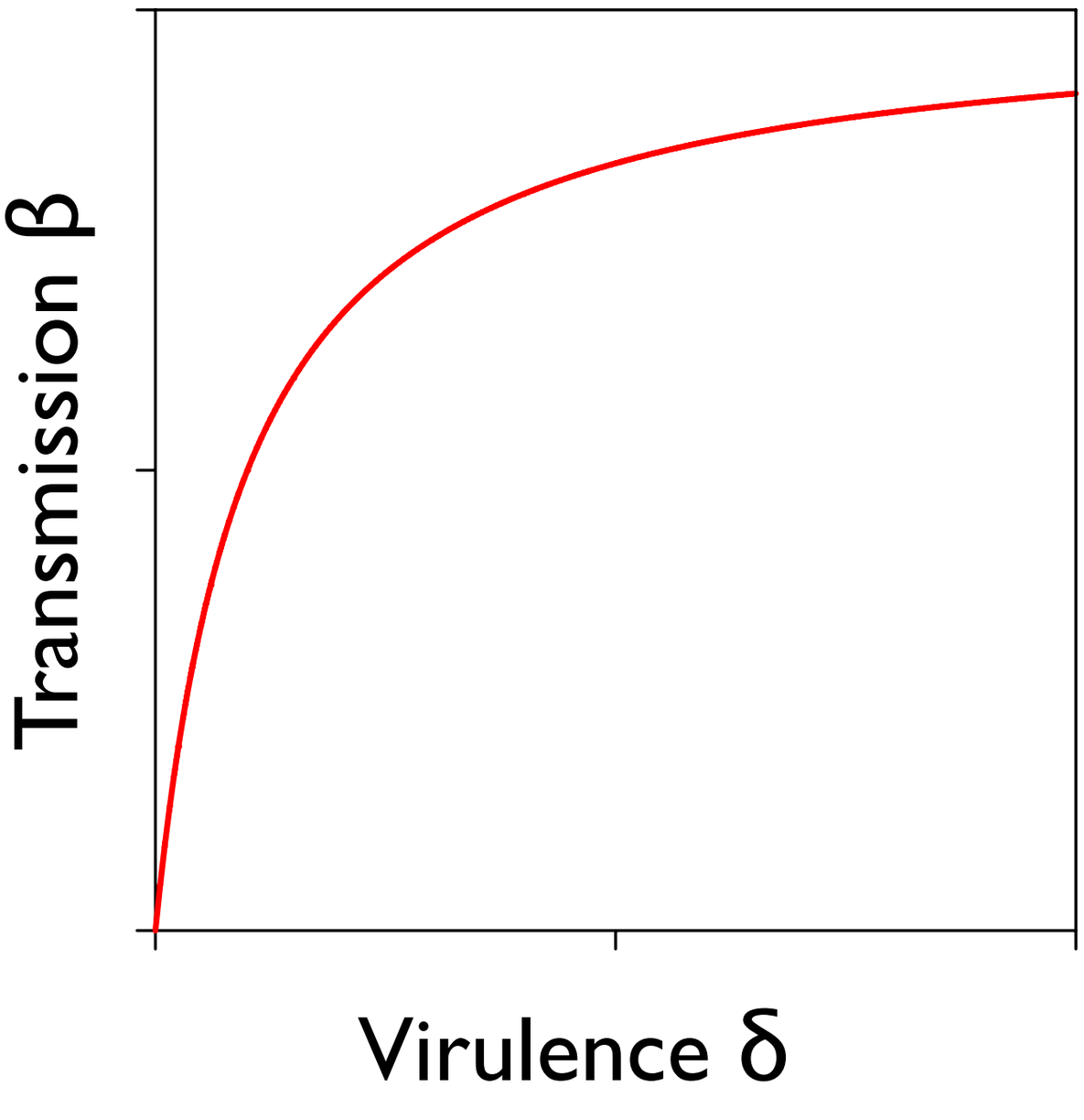
$$R_0 = \frac{\beta K}{d+\delta+\gamma}$$

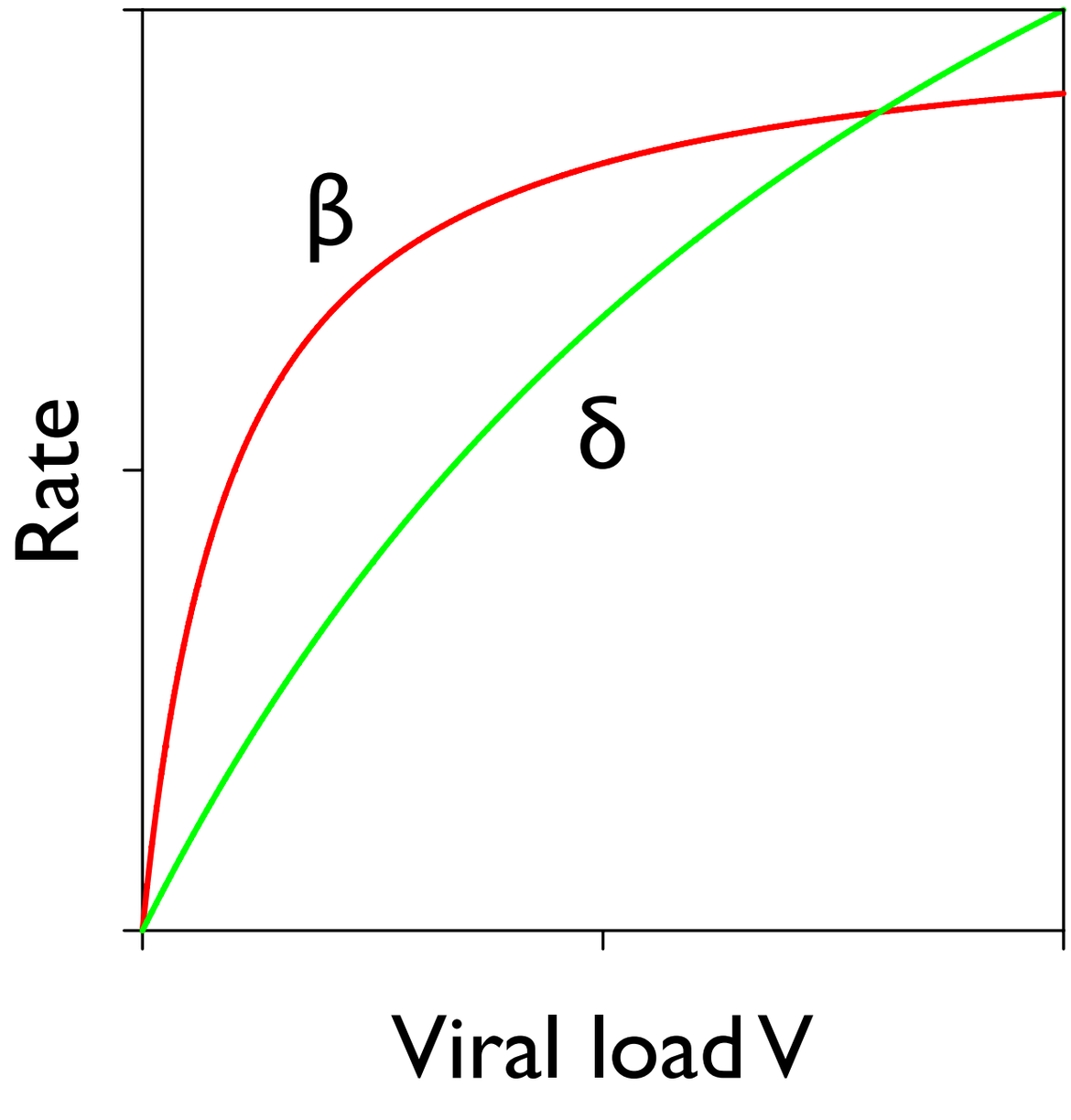
Pathogens are expected to optimize R_0:
trade-off between infectiousness and virulence

(K is steady state of N in the absence of infection)

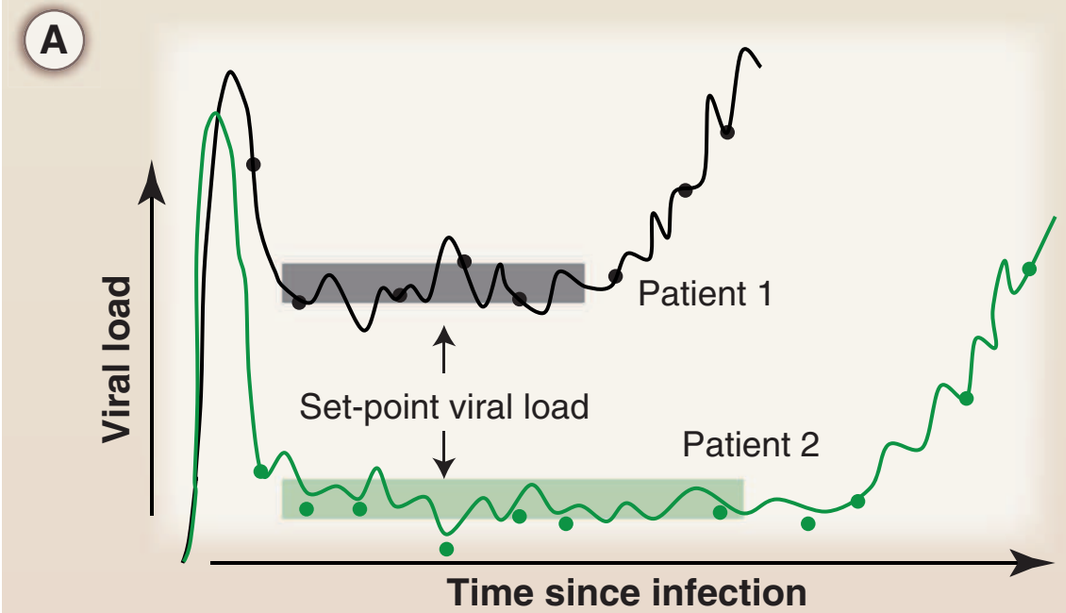
Trade-off between transmission and virulence leads to an optimum R_0

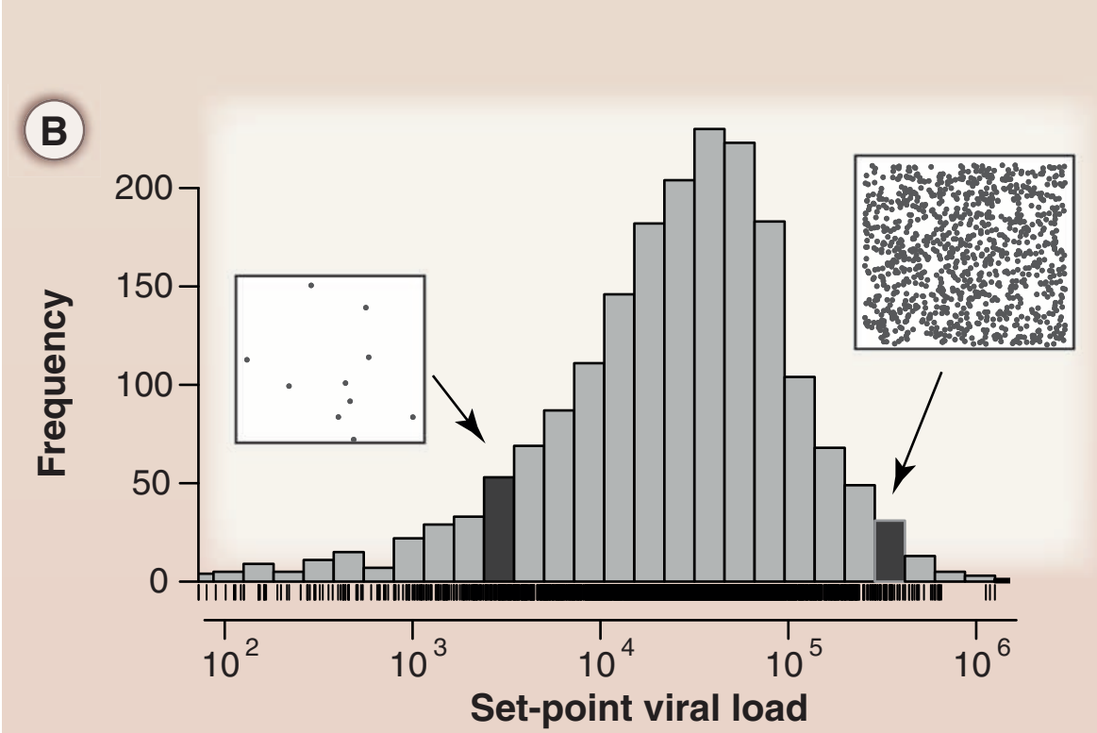
$$R_0 = \frac{\beta K}{d + \delta + \gamma}$$

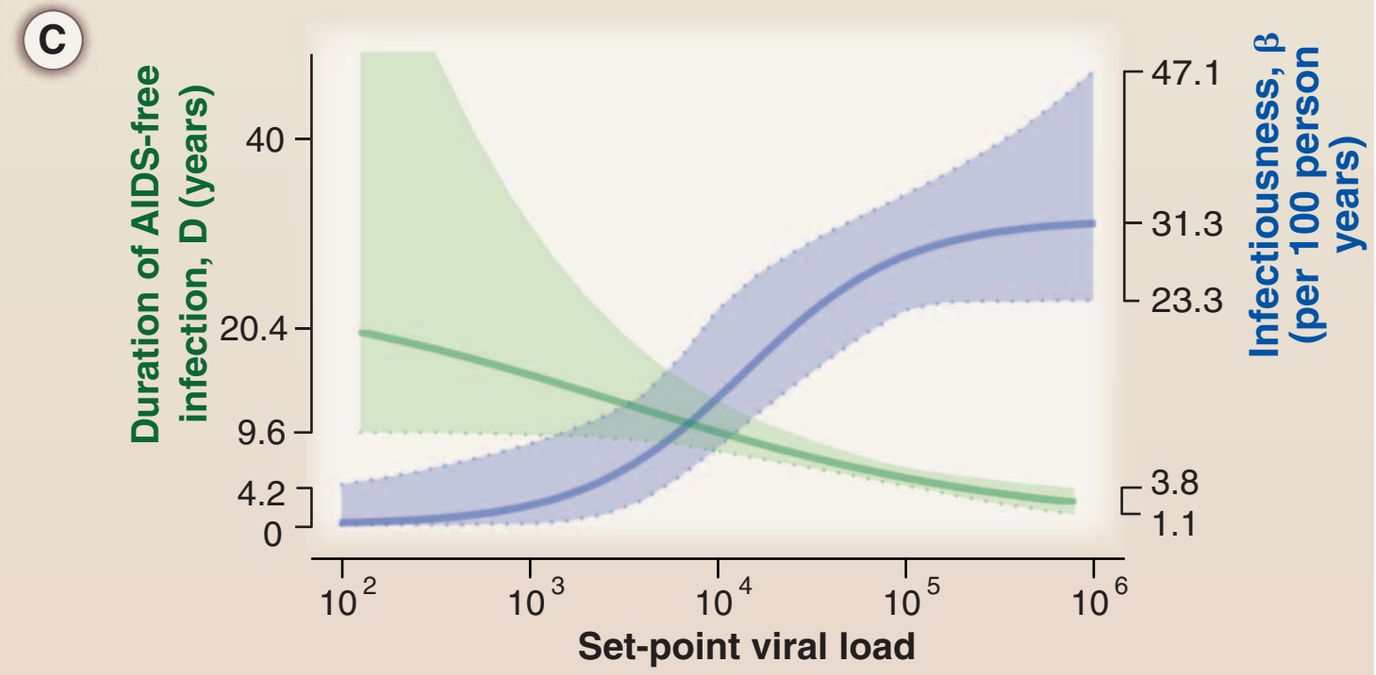


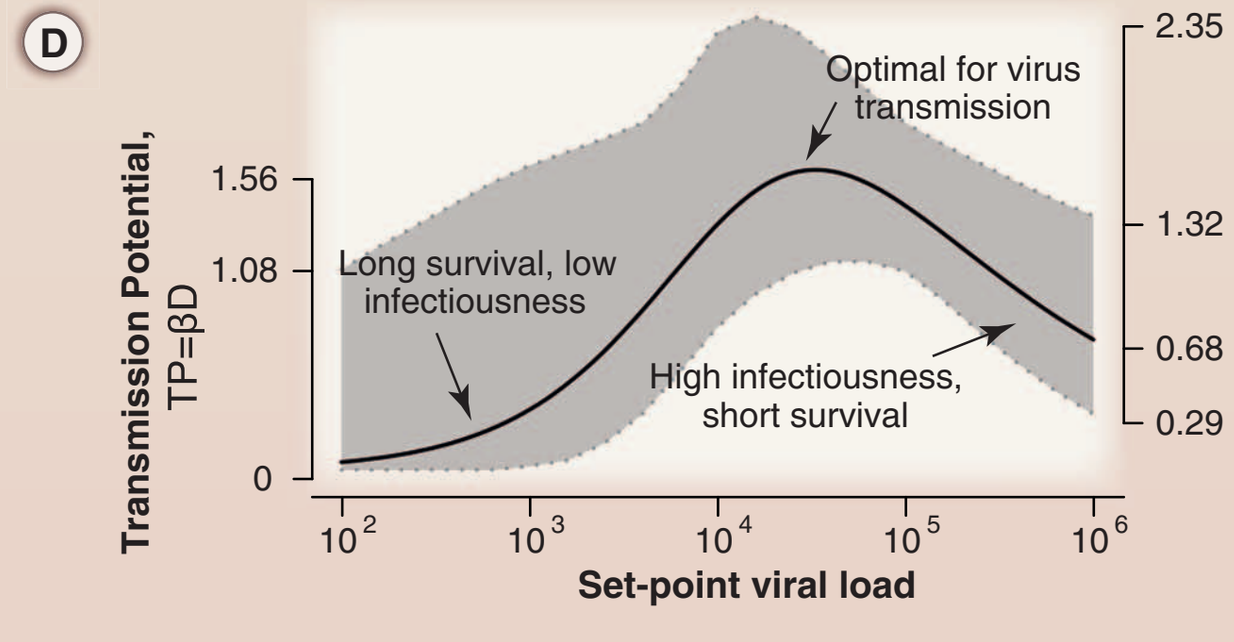


Is the HIV-1 set-point optimized for transmission?









Fraser et al. PNAS 2007, Science 2014

Classical trade-off between virulence and transmission

Complexity features

Two levels of evolution: within and between hosts

Hundreds of generations within a host separate the transmission events.

Host mounts unique immune responses

HIV rapidly evolves in every new host: escapes & reversions

Can virus evolve optimal transmission if that selection occurs just once every hundreds of generations?

So what next: how do we model such a system combining immunology, epidemiology, host heterogeneity, and virus evolution?

Agent based model with two levels of selection

Virus contains n=300 potential epitopes (bitstring), which can be wild-type or mutated (2^{300} viruses).

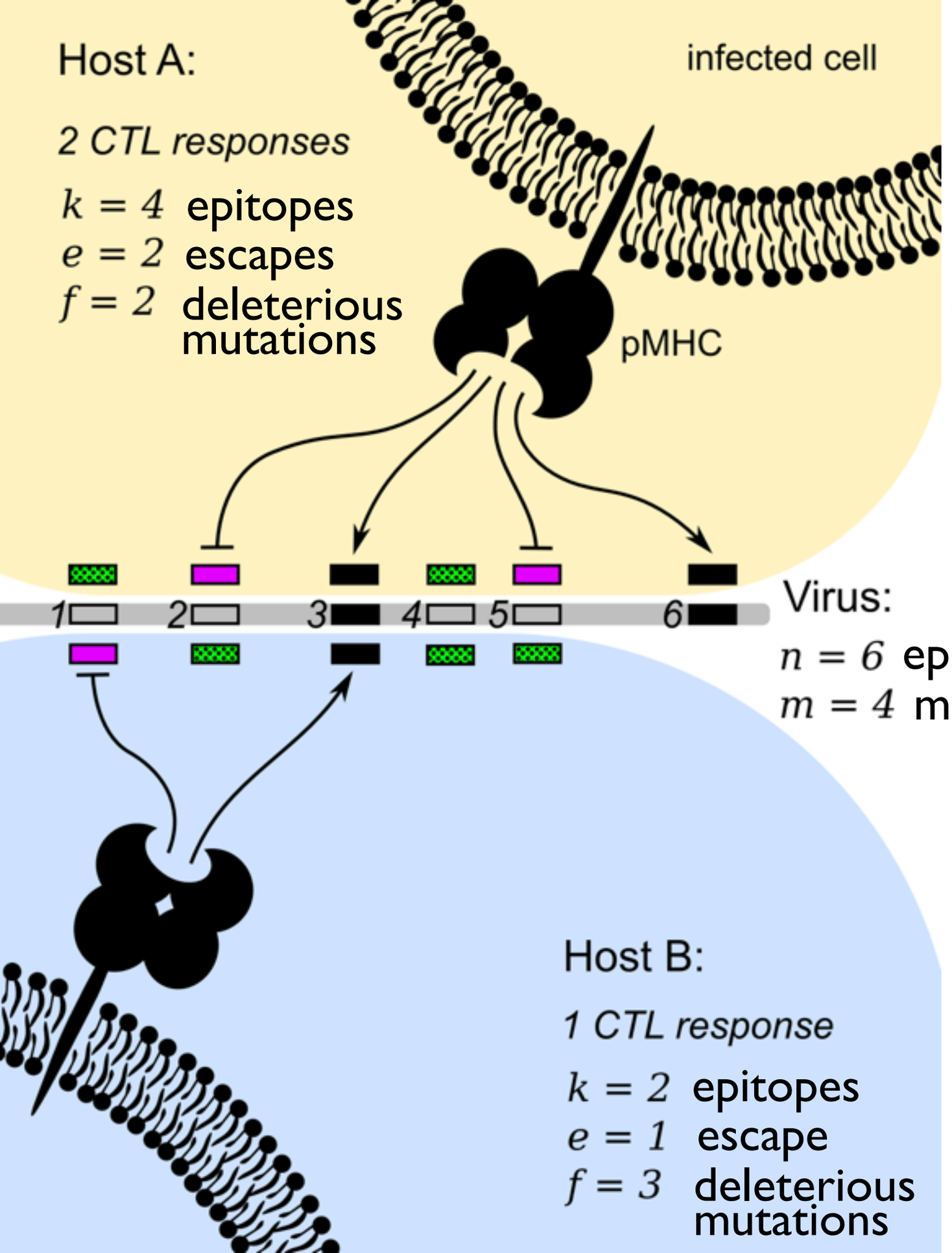
Agents randomly select approximately k=15 epitopes. With e escapes they will k-e immune responses.

Mutated sites that are not selected do confer a fitness cost and are "deleterious mutations" (f).

Within the host the virus stochastically mutates epitopes to escape immune responses and reverts deleterious mutations to increase its fitness:

$$e \xrightarrow{\lambda_{\text{esc}} \cdot (k-e)} e+1, \qquad f \xrightarrow{\lambda_{\text{rev}} \cdot f} f-1,$$

Van Dorp, Van Boven & De Boer, PLoS Comp Biol, 2014.



The virus load in a host is determined by the number of remaining immune responses (k-e), and by the total fitness cost of all mutations (e+f)

$$V = V_{max} - \sigma(k-e) - \varphi(e+f)$$

wild-type

mutated

immune escape

deleterious mutation

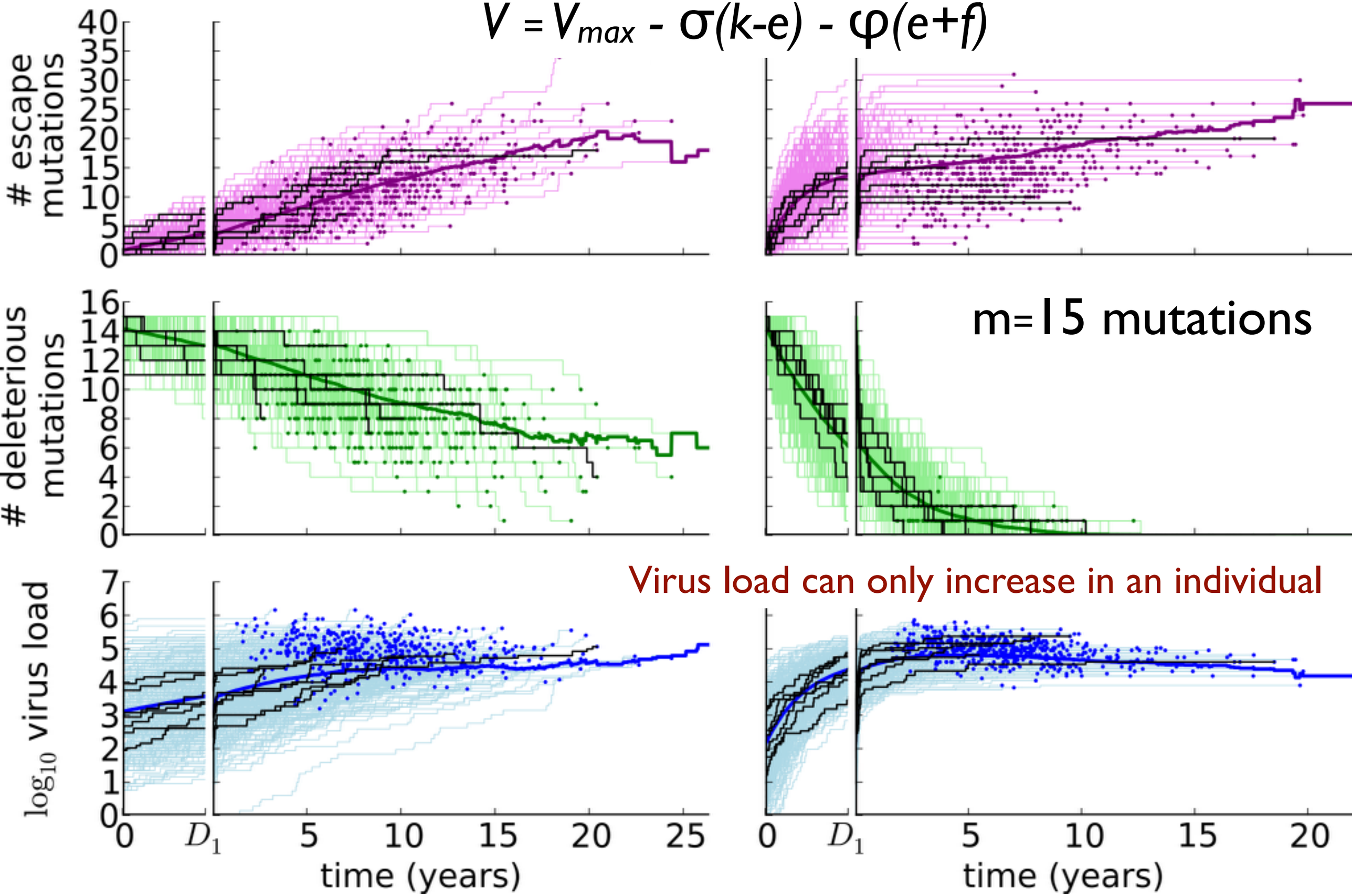
V_{max}=7, escape rate λ=1 y^{-1}

V_{max}=6, escape rate λ=10 y^{-1}

$V = V_{max} - \sigma(k\text{-}e) - \varphi(e+f)$

m=15 mutations

Virus load can only increase in an individual



Modeling the epidemic level

Simulate a population of infected individuals (*S*+*I*=*25000*)

Each transmission event happens with a virus having accumulated *e* escapes and *f* deleterious mutations: a virus with *e*+*f* = *m* mutations in the *n* sites.

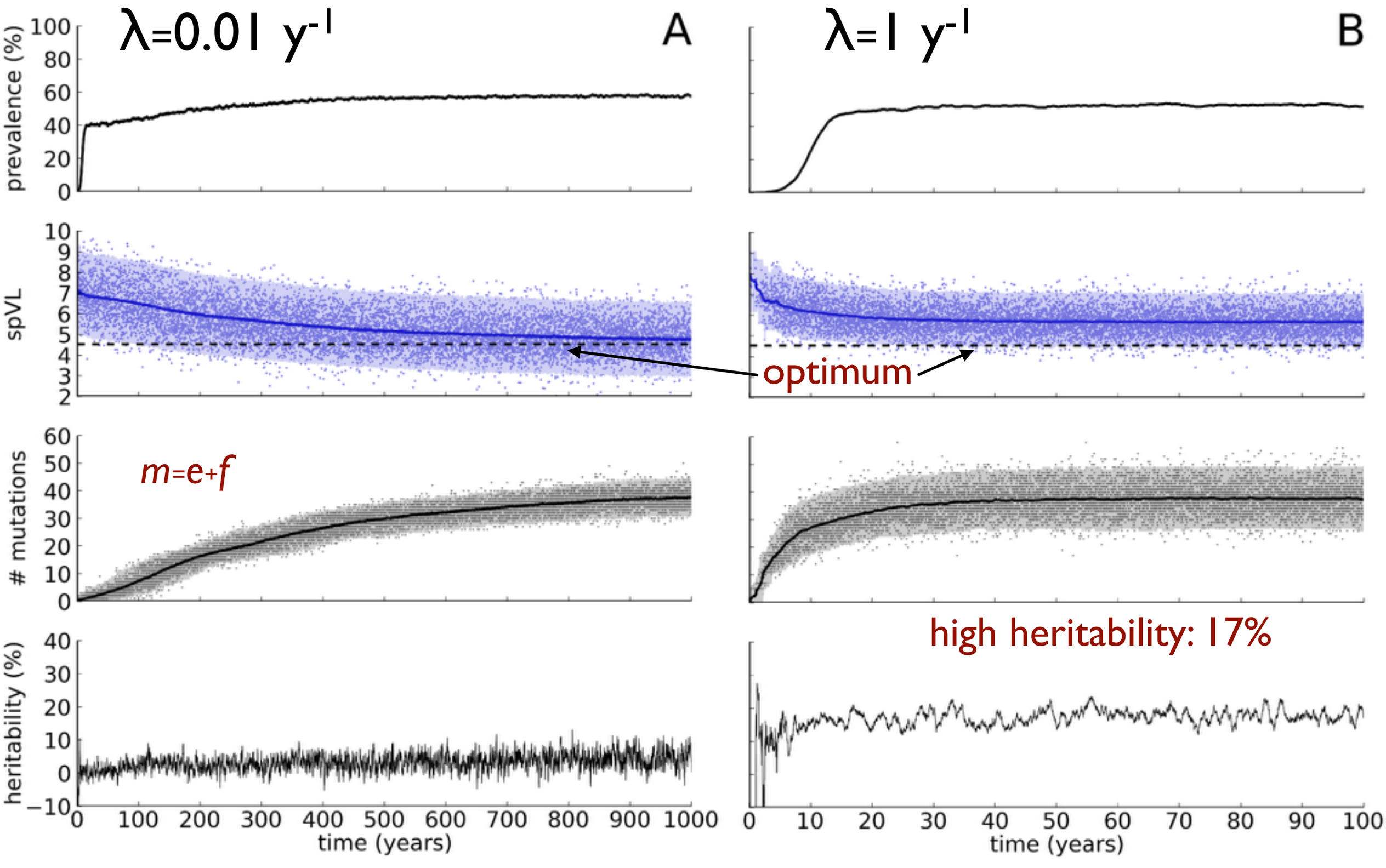
In the next random host we draw a new random binding repertoire of *k* immune responses, and draw the expected number of escaped epitopes *e'*.

The remaining *m*-*e'* form the new *f'* deleterious mutations.

Parameters of Fraser *et al.* for infection and death: $V^*=10^{4.5}$

Van Dorp, Van Boven & De Boer, PLoS Comp Biol, 2014.

Epidemics with two different mutation rates do not approach same set-point

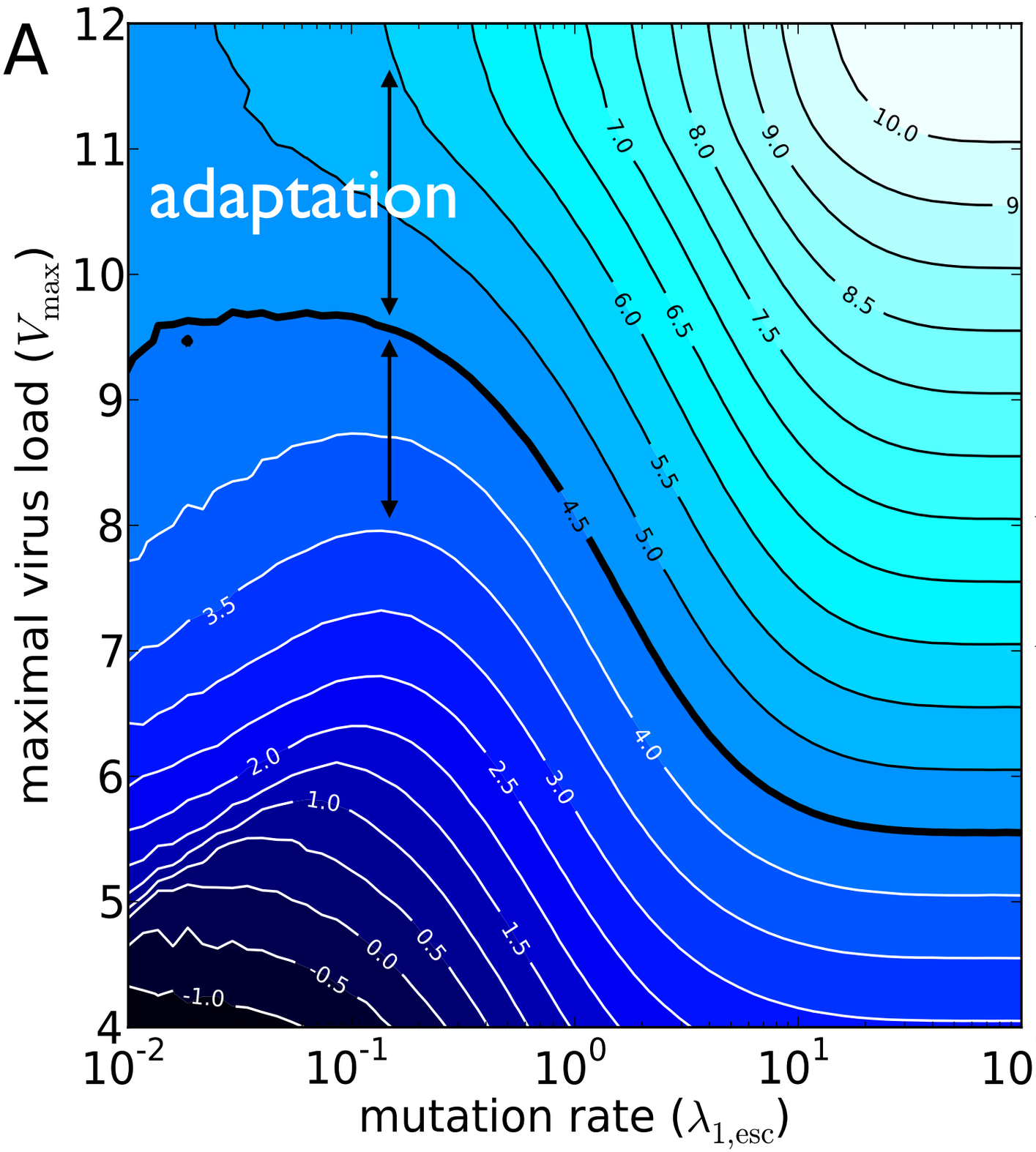


Evolved mean set-point virus load

V=4.5 is the "optimal" viral load (heavy line)

realistic range

A



At high mutation rates all escapes and reversions happen early. Hence

$V=V_{max} - k\varphi$

$V=V_{max} - k\varphi$

$V=V_{max} - k\varphi$

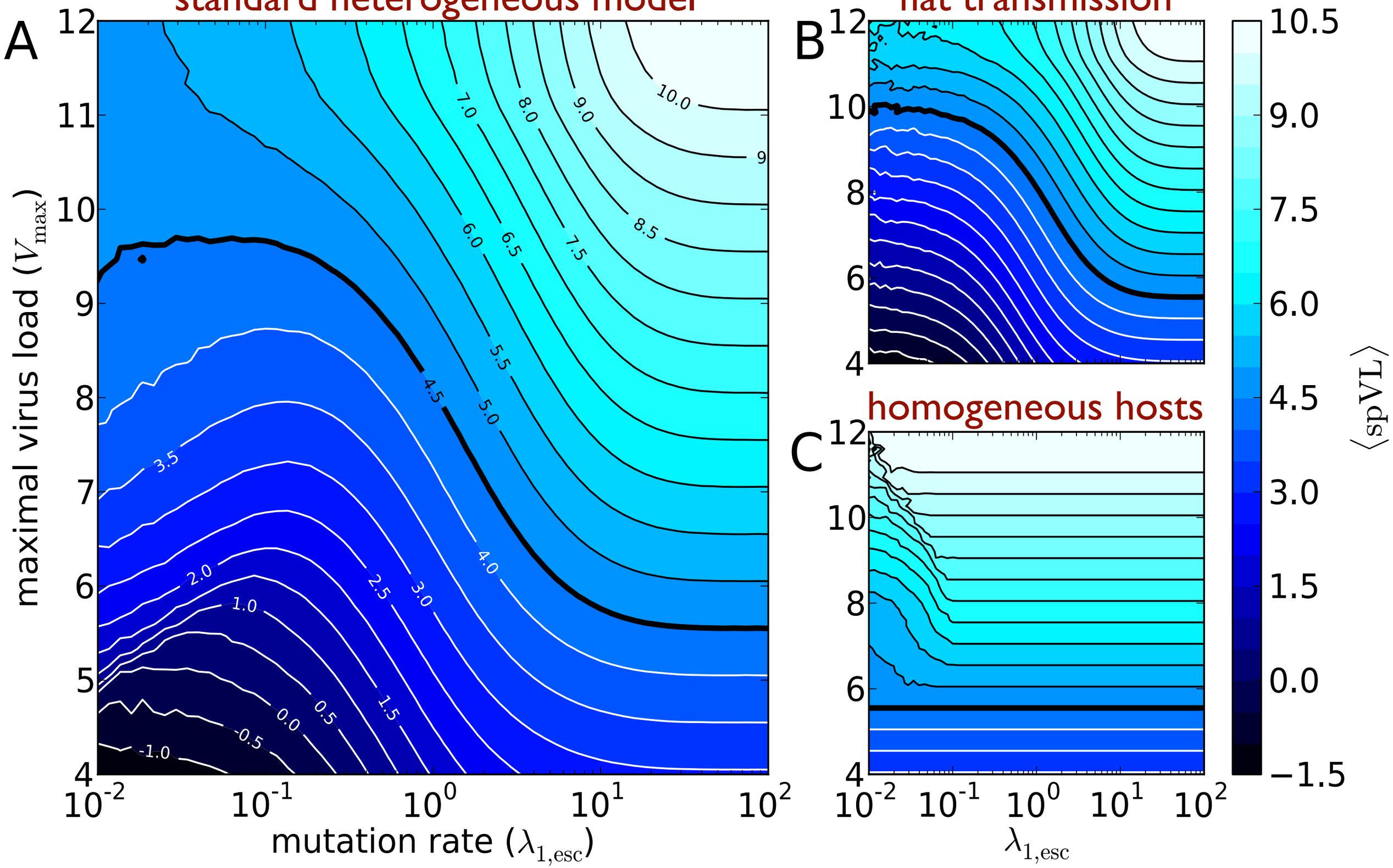
effect of increasing V_{max}

effect of decreasing V_{max}

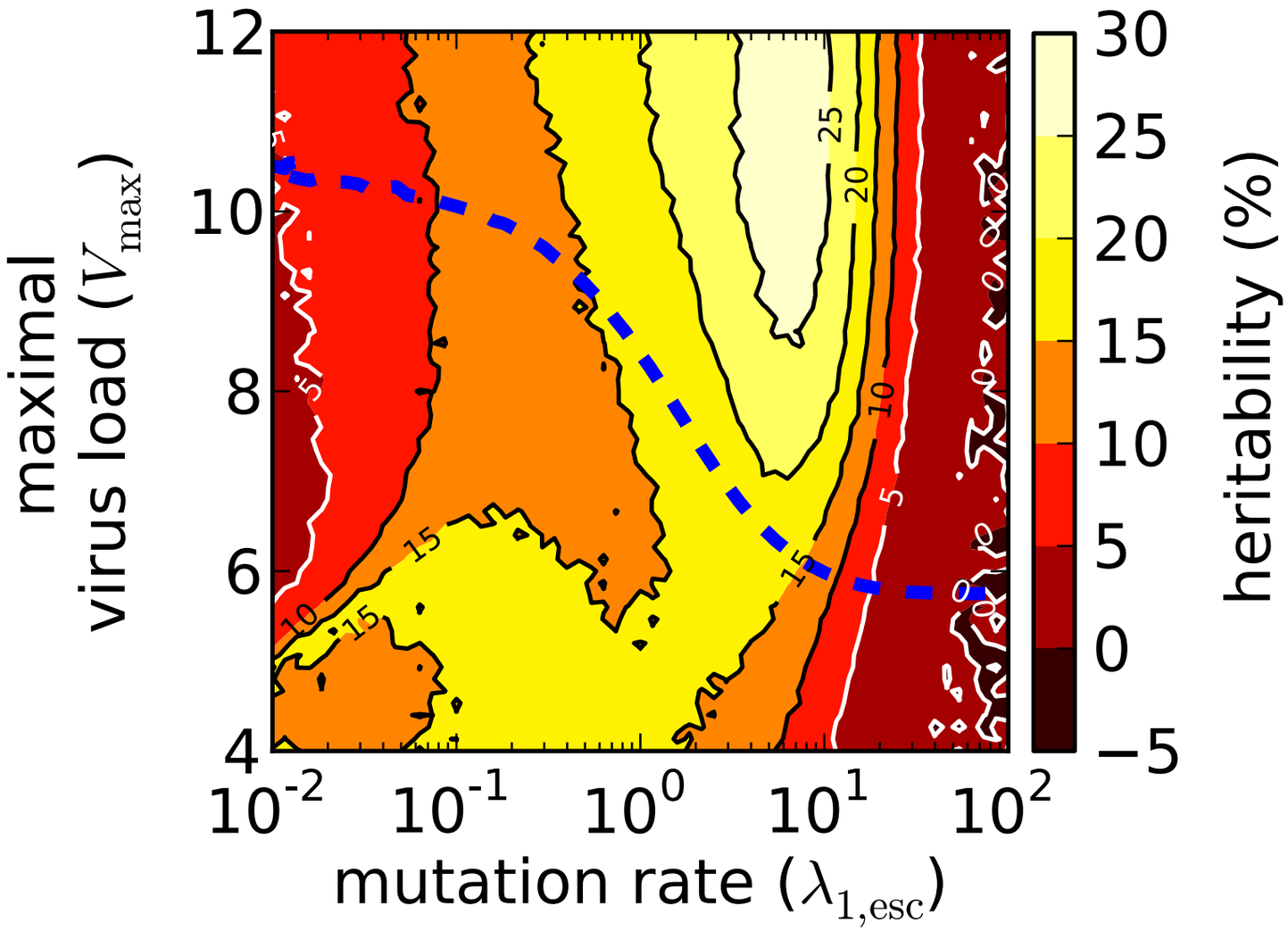
$V = V_{max} - \sigma(k\text{-}e) - \varphi(e+f)$

Evolved mean set-point virus load

V=4.5 is the "optimal" viral load (heavy line)



Measure heritability during simulations

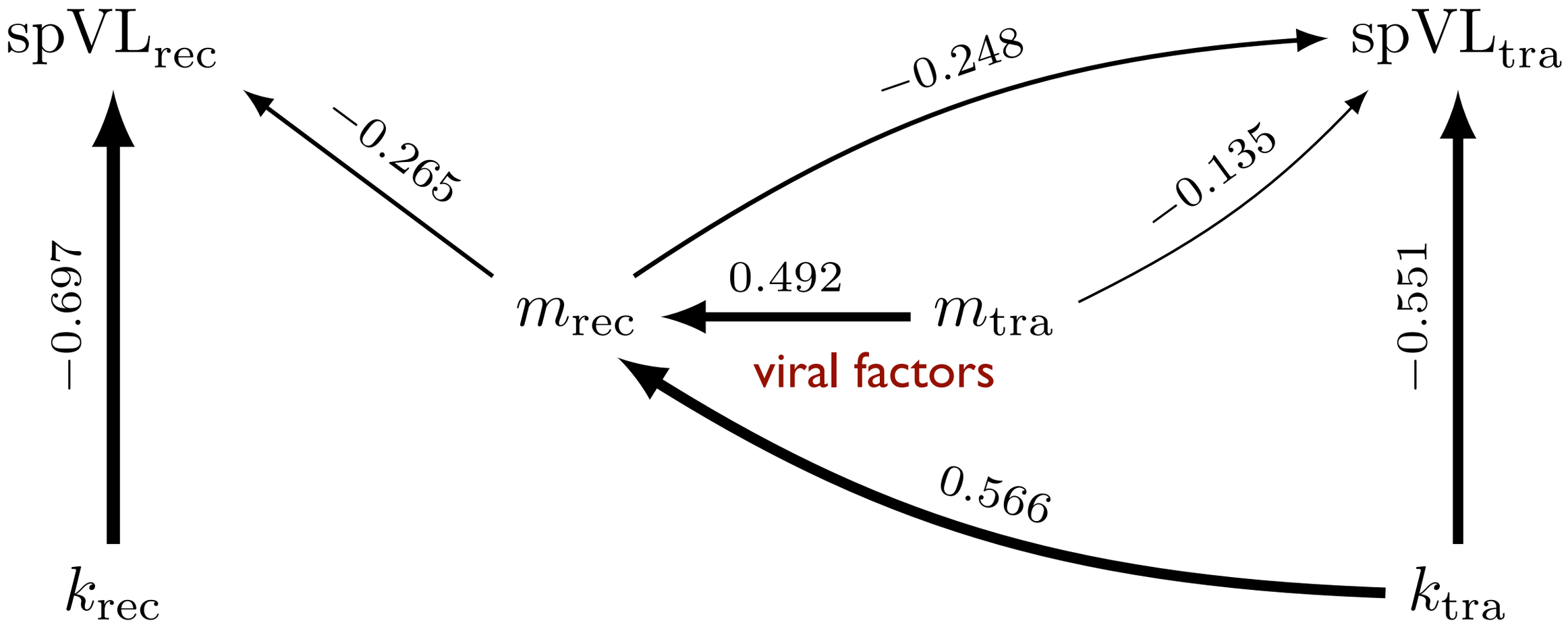


Low mutation rate: viruses with many deleterious mutations are crippled in both donor and recipient (viral factor)

High mutation rate: hosts with high *k* control well and transmit crippled virus, leading to low viral load in both:

Immunological footprint

Structural equation model



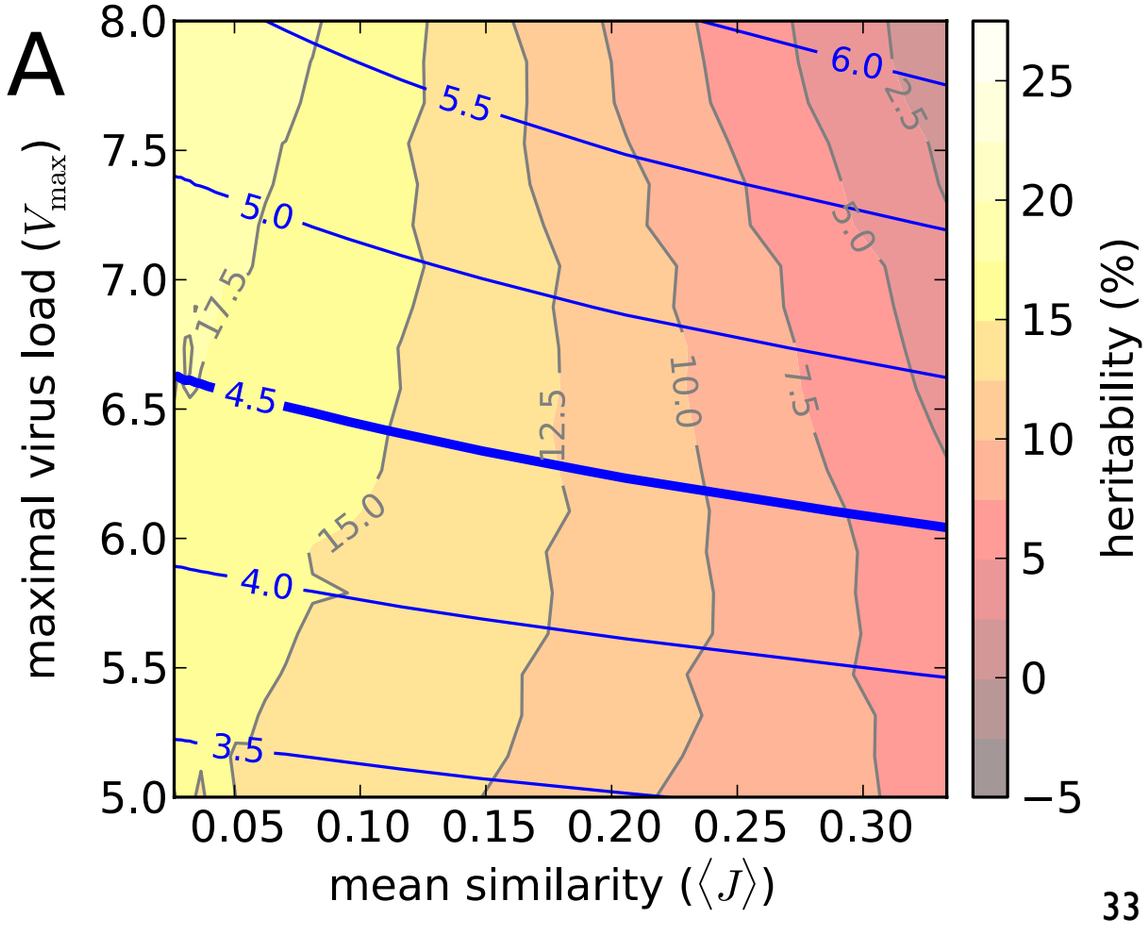
Heritability of spVL is the sum of viral fitness and the breadth of the immune response of the transmitting host

For realistic parameter values, half of the observed heritability is due to the footprint effect

$$P_1 \quad := \quad \mathrm{spVL_{tra}} \leftarrow m_{\mathrm{rec}} \rightarrow \mathrm{spVL_{rec}}\,,$$

$$P_2 \quad := \quad \mathrm{spVL_{tra}} \leftarrow m_{\mathrm{tra}} \rightarrow m_{\mathrm{rec}} \rightarrow \mathrm{spVL_{rec}} \quad \text{and,}$$

$$P_3 \quad := \quad \mathrm{spVL_{tra}} \leftarrow k_{\mathrm{tra}} \rightarrow m_{\mathrm{rec}} \rightarrow \mathrm{spVL_{rec}}\,,$$



Heritability due to HLA heterogeneity

Decrease heterogeneity by decreasing the number of epitopes *n*

Measure similarity *J* by Jaccard index (overlapping epitopes)

Conclusions

For realistic mutation rates we do not expect the virus to adapt much on the population level.

Evolution indeed dominated by the hundreds of generations within heterogeneous hosts.

High heritability partly due to immunological footprint in heterogeneous populations.

These variable “viral factors” are not identifiable.

Agent based models required to answer this question

Extensions: immune responses:

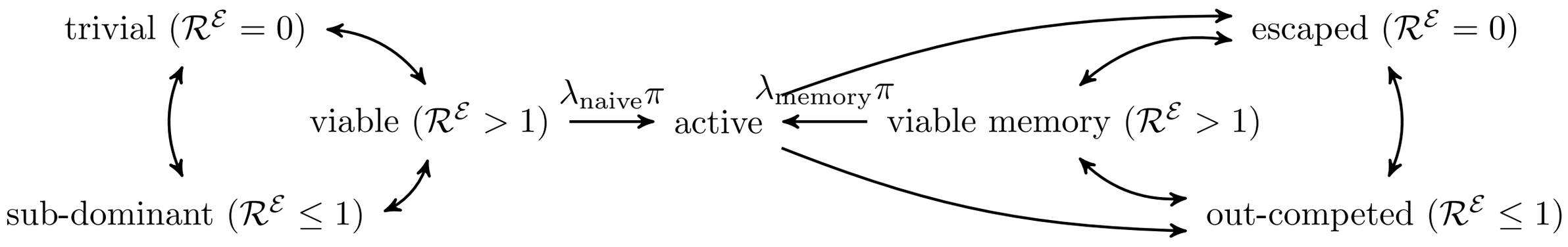
$$\frac{dT}{dt} = s - d_{\mathcal{T}} T - \beta I T\,, \quad \frac{dI}{dt} = \beta c I T - d_{\mathcal{I}} I - nkEI\,, \quad \frac{dE}{dt} = \frac{pEI}{h + I + E} - d_{\mathcal{E}} E\,.$$

h_i

Total number of immune effector cells:

$$E := \sum_{i=1}^{n} E_{\ell_i, a_i} = n \cdot \left(\frac{p - d_{\mathcal{E}}}{d_{\mathcal{E}}} \cdot I - \overline{h} \right)\,,$$

Scheme to allow for immune responses:

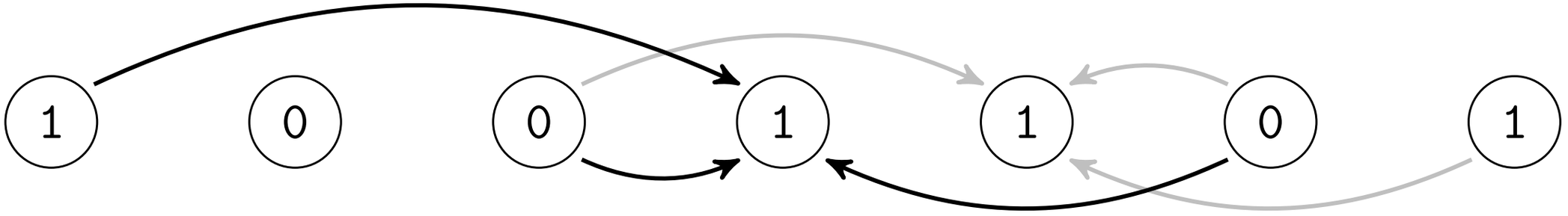


Extensions: rugged fitness landscape

J. theor. Biol. (1989) **141,** 211–245

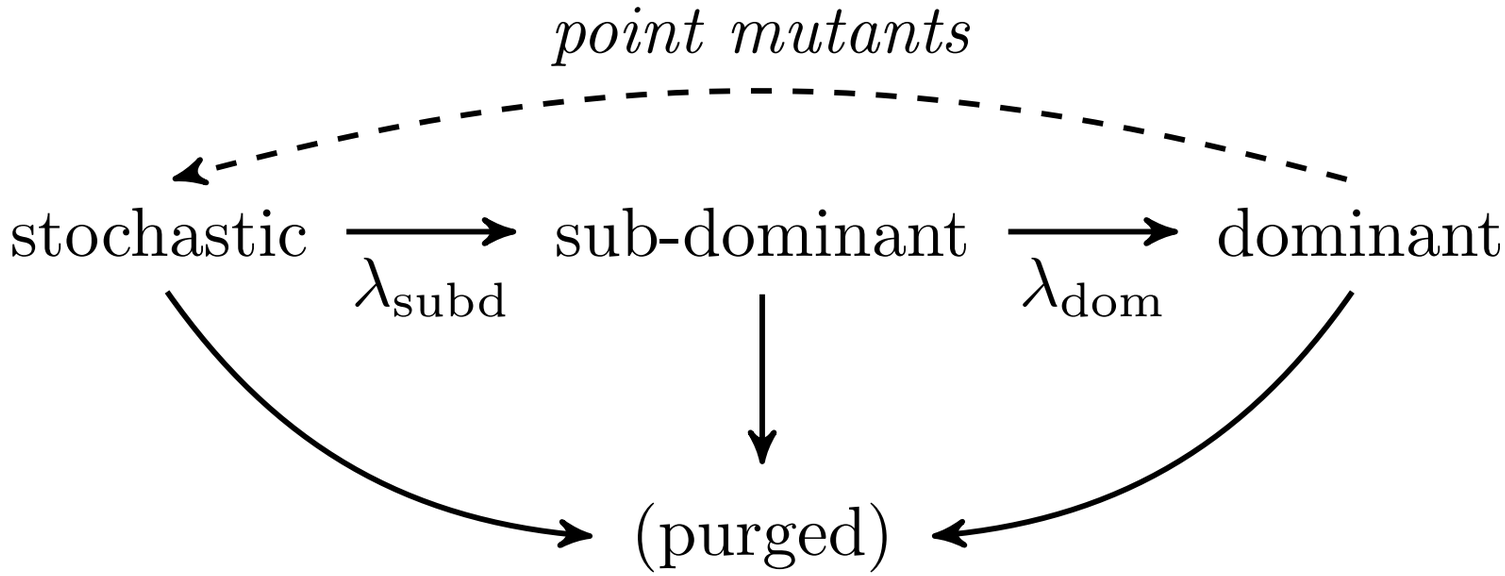
The *NK* Model of Rugged Fitness Landscapes And Its Application to Maturation of the Immune Response

STUART A. KAUFFMAN AND EDWARD D. WEINBERGER



a_4	a_1	a_3	a_6	H_4
0	0	0	0	0.0
⋮	⋮	⋮	⋮	⋮
1	1	0	0	**0.34**
1	1	0	1	0.65

Extensions: viral mutants & interference



Probability that mutant survives initial stochastic phase:

$$1 - 1/\mathcal{R}^{\mathcal{I}}, \text{ where } \mathcal{R}^{\mathcal{I}} = \frac{c\beta' T}{d_{\mathcal{I}} + n' E}$$

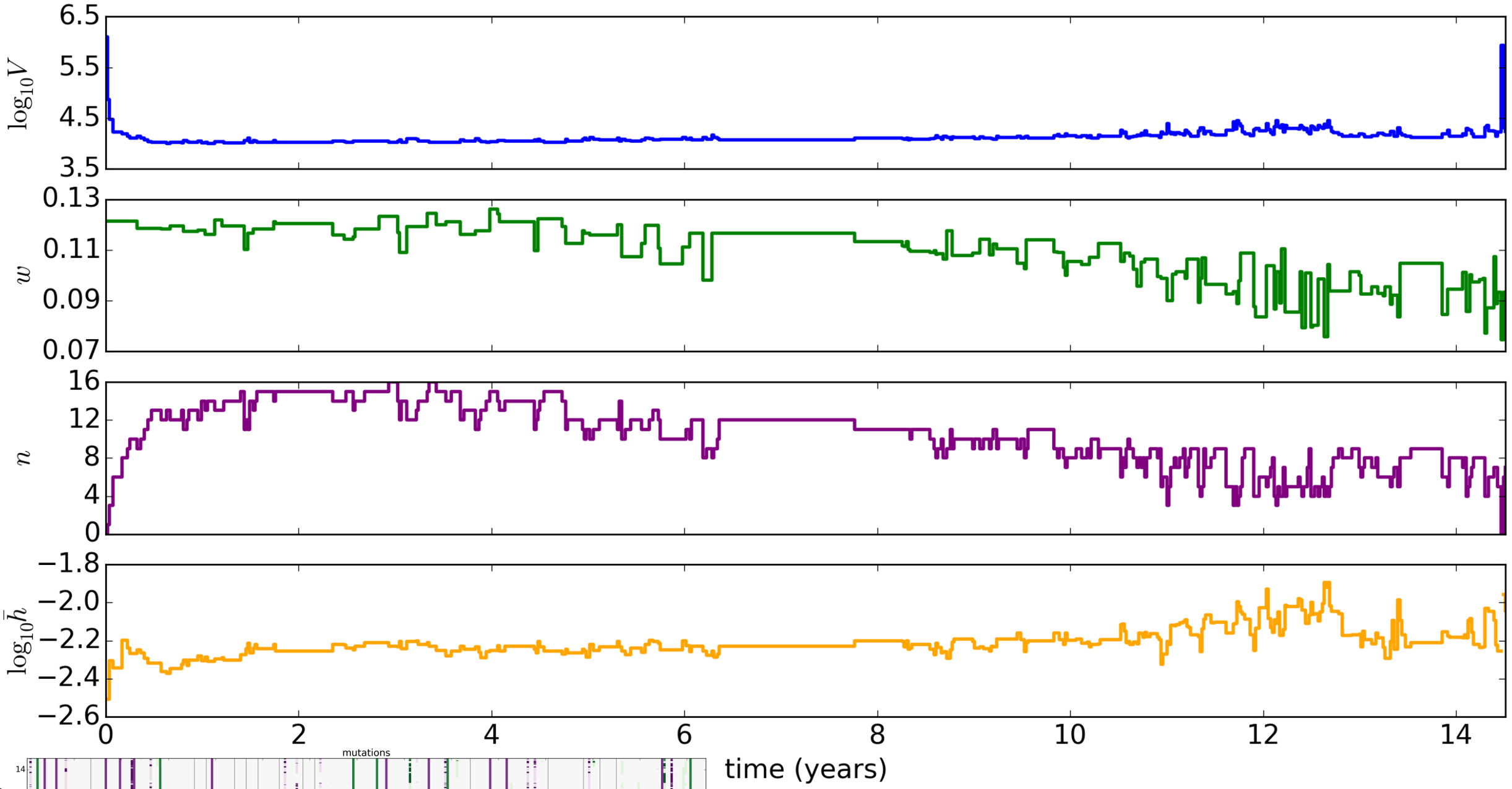
Rate at which they leave this phase:

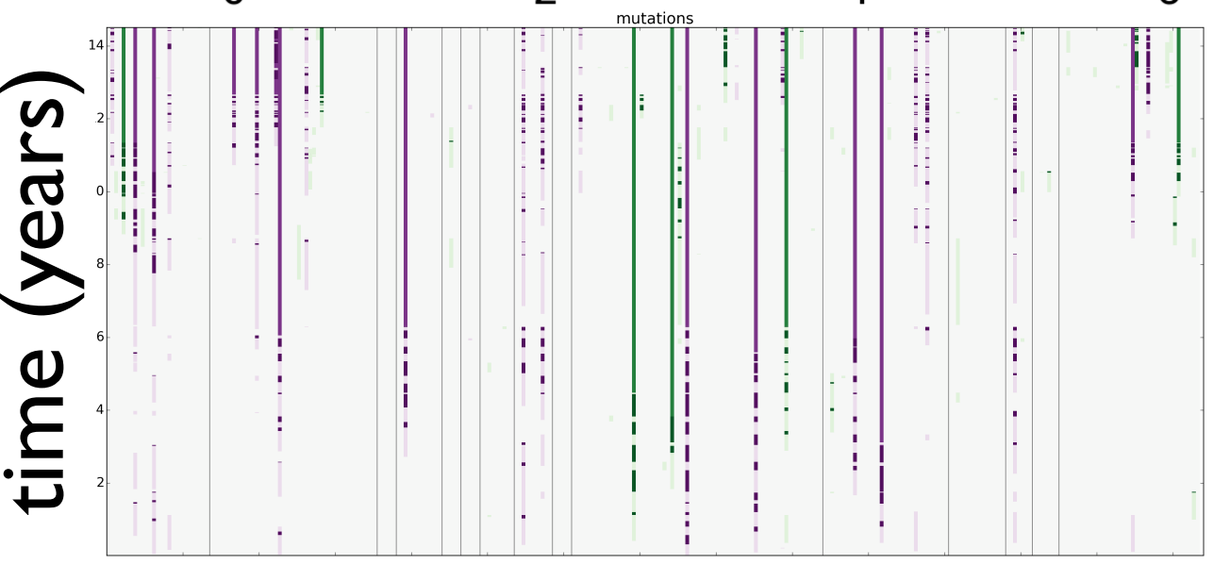
$$\lambda_{\mathrm{subd}} := q\beta c I T (1 - 1/\mathcal{R}^{\mathcal{I}}) \, .$$

Subsequent exponential growth phase:

$$\lambda_{\mathrm{dom}} := c\beta' T - (d_{\mathcal{I}} + n' E)$$

One host





purple: epitope loci
green: other loci

virus genome

One epidemic