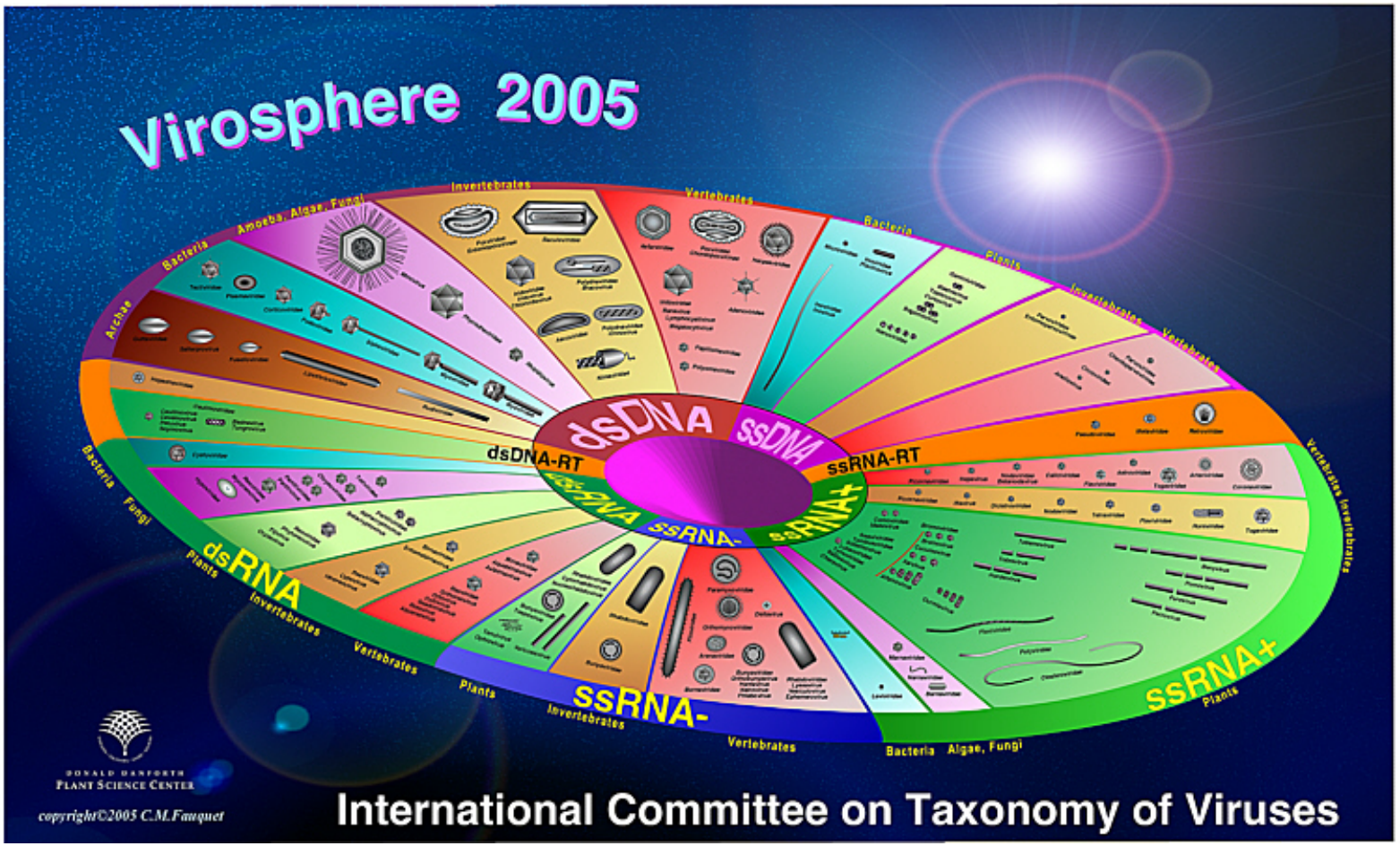
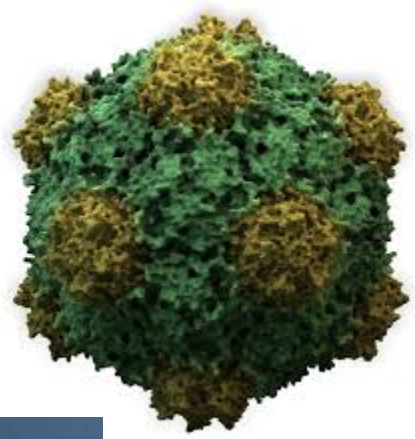
Virology in a nutshell, quasispecies and experimental virus evolution

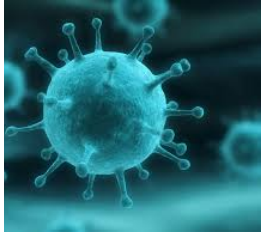
Santiago F. Elena

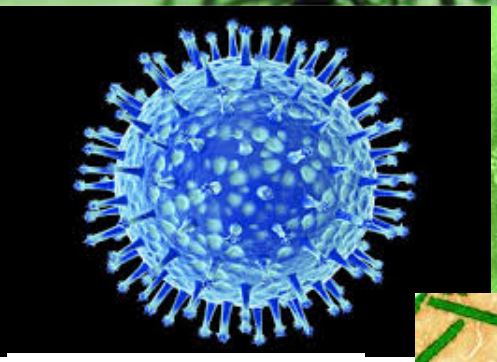
Evolutionary Systems Virology Group

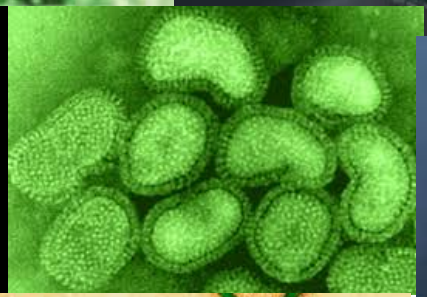
The virosphere

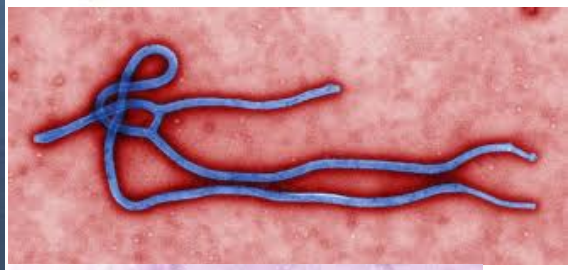


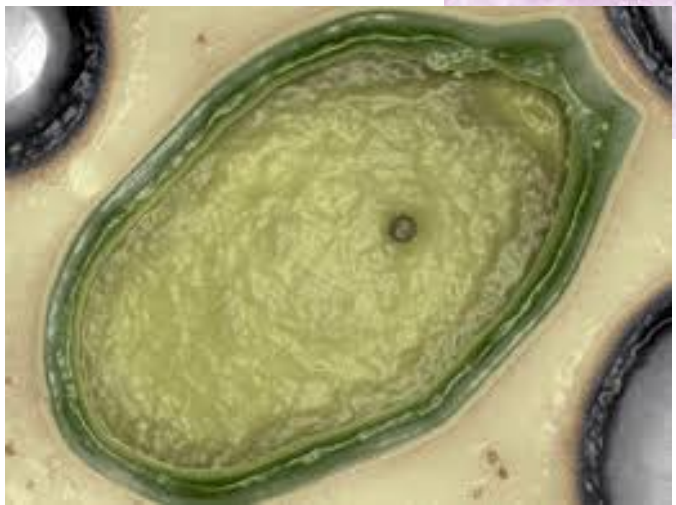




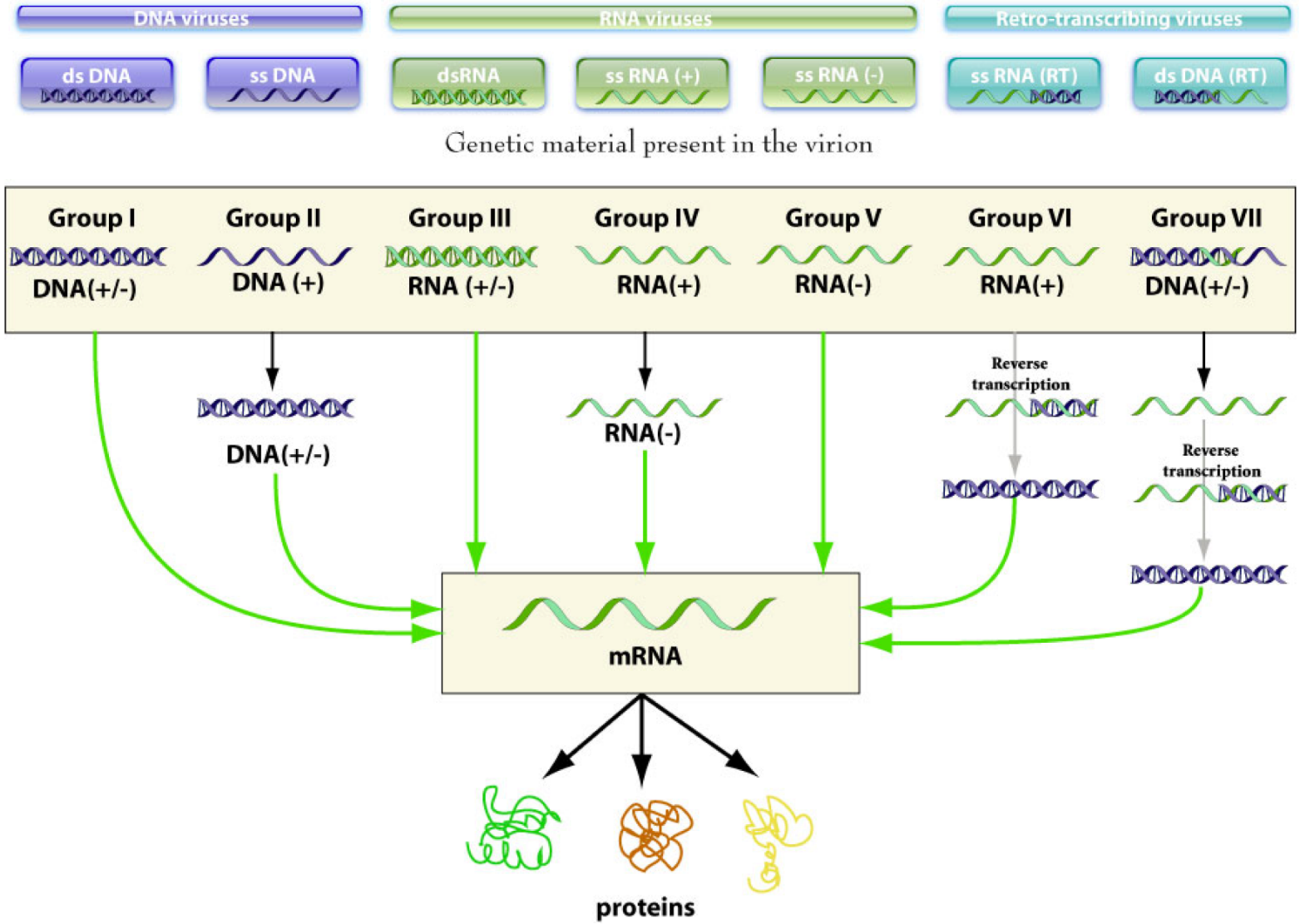








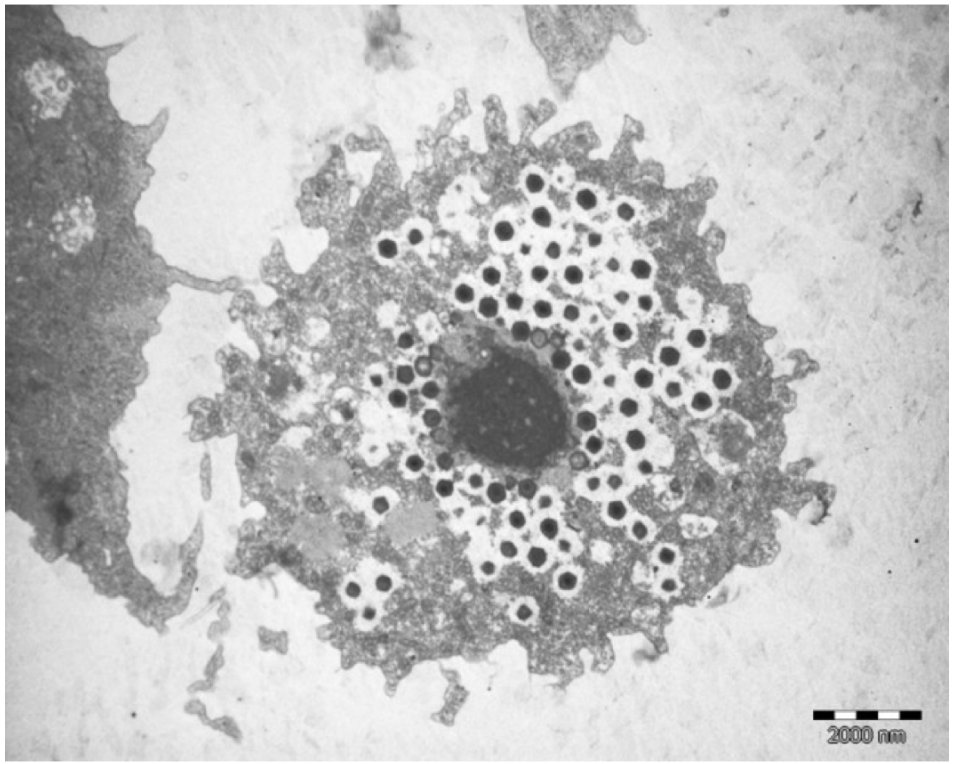
The Baltimore's classification



Are viruses alive?

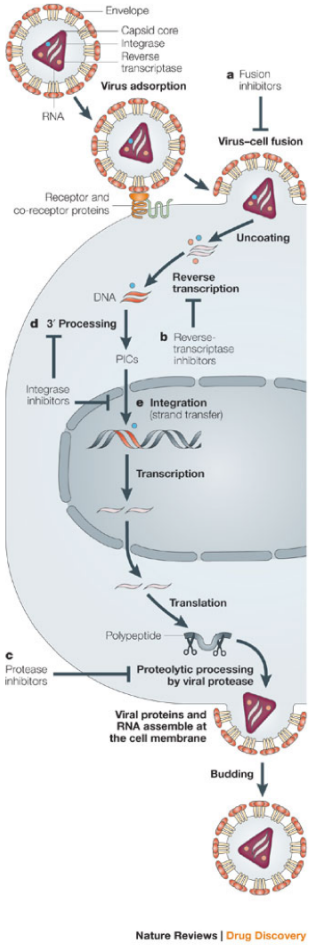
The viral particle (virion) is not, it is just a transmission inert stage. Likewise a dormant plant seed.

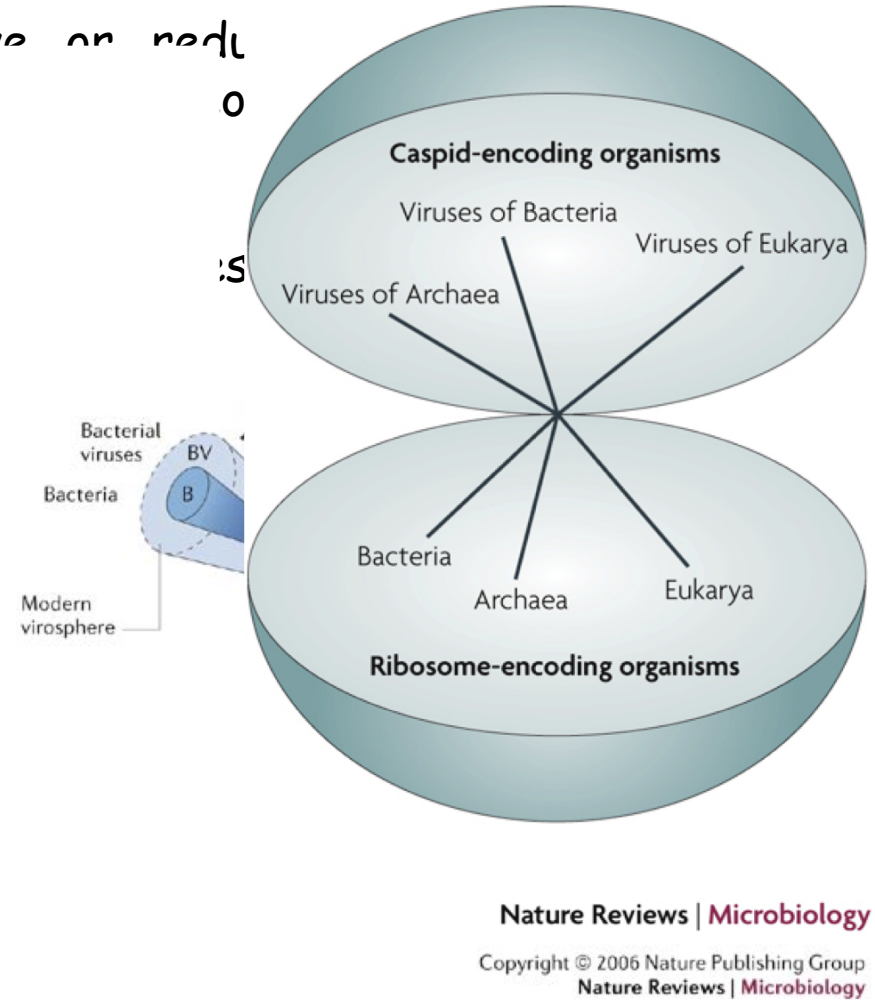
Once in their right environment (*e.g.*, the cell nucleus or the cytoplasm) they are alive: the **virocell** (*sensu* Forterre).



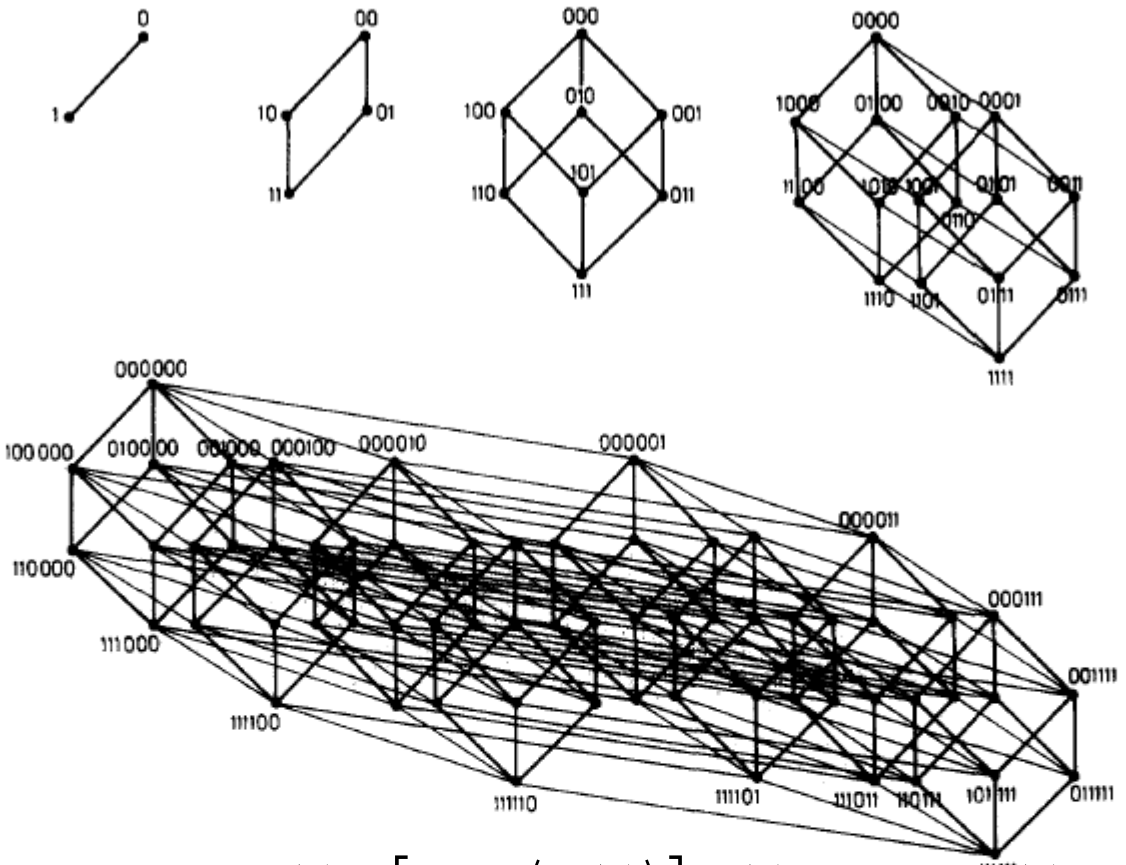
Virus origins

- ✓ The progressive or scape hypothesis: viruses arose from genetic elements that gained the ability to move between cells (*e.g.*, retroelements and retroviruses).
- ✓ The regressive or redu… are remnants of cellular o… viruses as poxvirus or
- ✓ …s … evolved with their current





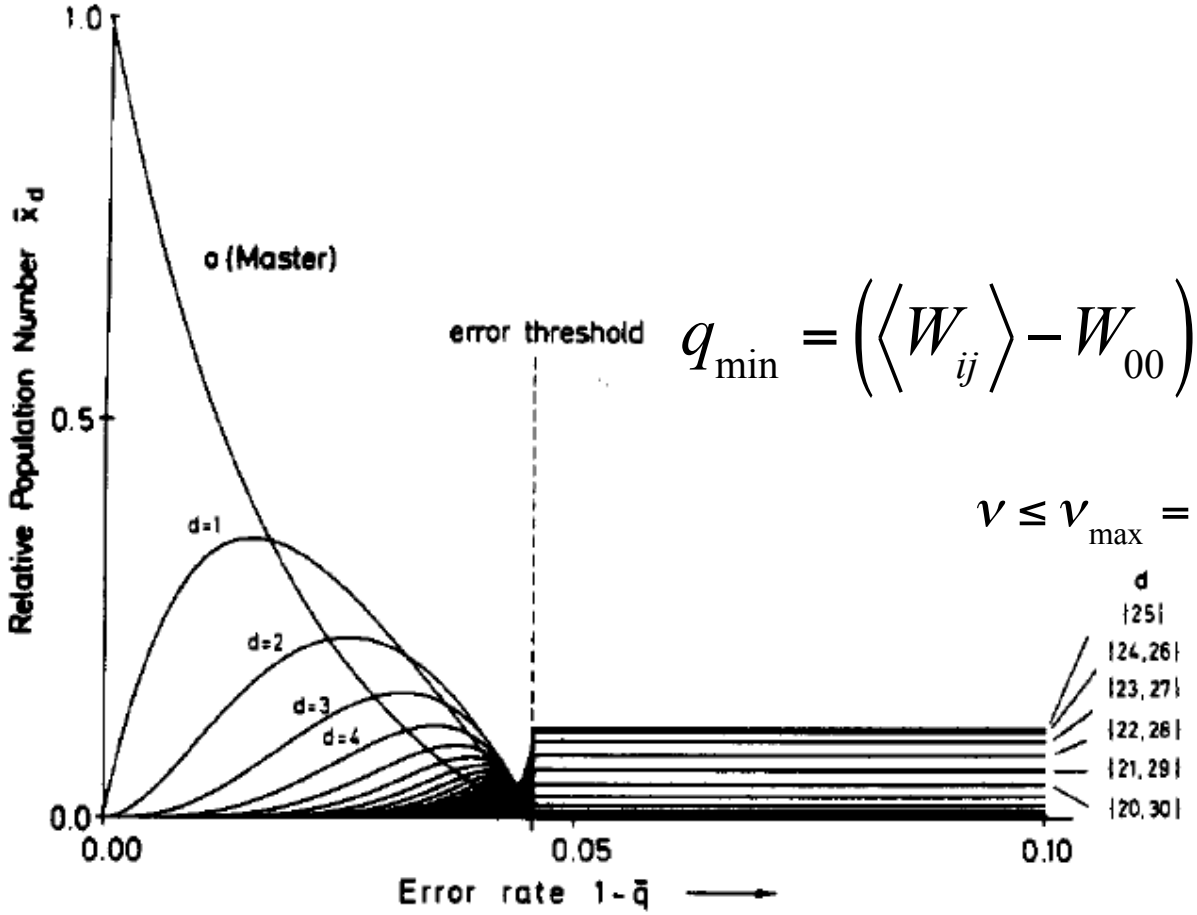
The Quasispecies model of virus evolution



$$\dot{x}_i(t) = \left[W_{ii} - \left\langle E(t) \right\rangle\right] x_i(t) + \sum_{j \neq i} W_{ij} x_j(t)$$

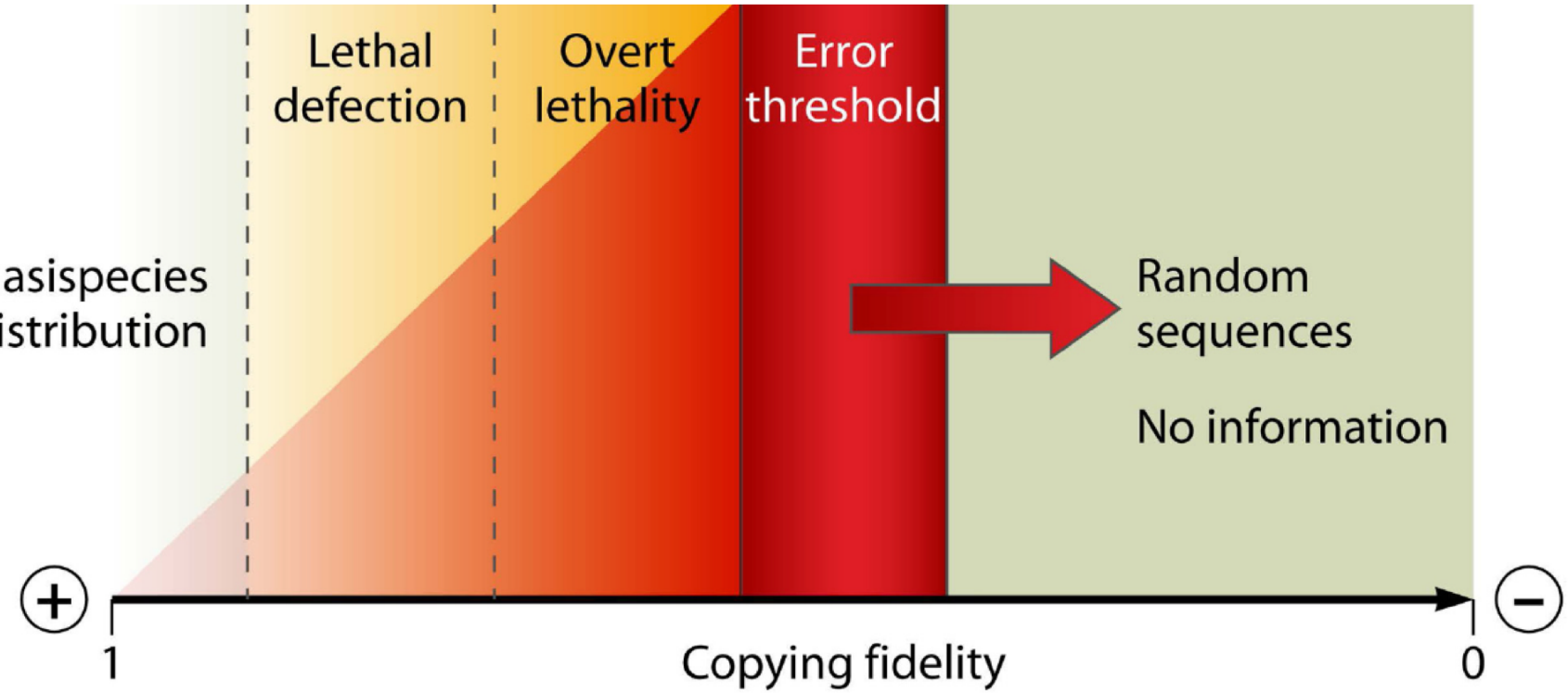
$$W_{ij} = A_i q_{ij}$$

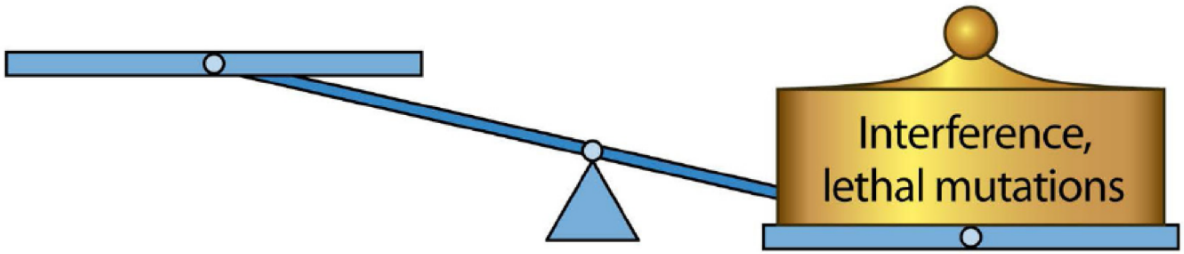
The Quasispecies model of virus evolution

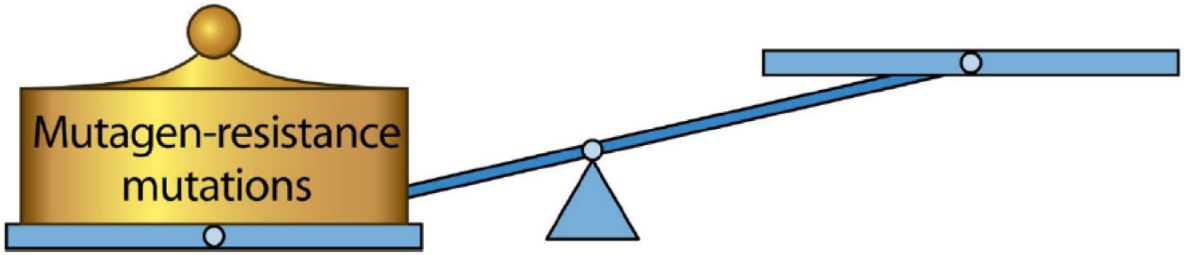


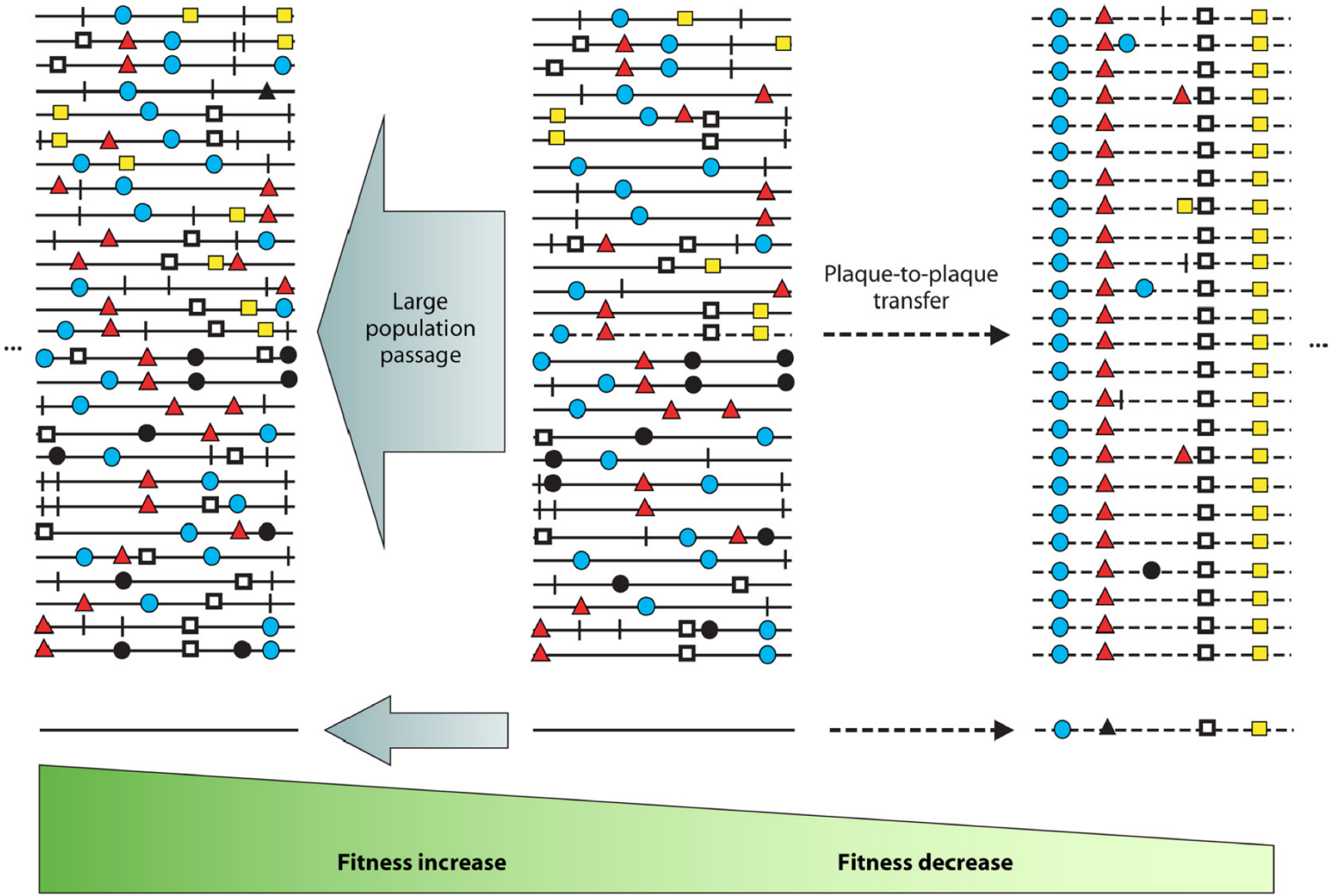
$$q_{\min} = \left(\langle W_{ij} \rangle - W_{00} \right)^{1/\nu}$$

$$\nu \leq \nu_{\max} = \frac{\ln\left(\langle W_{ij} \rangle - W_{00} \right)}{1 - \bar{q}}$$

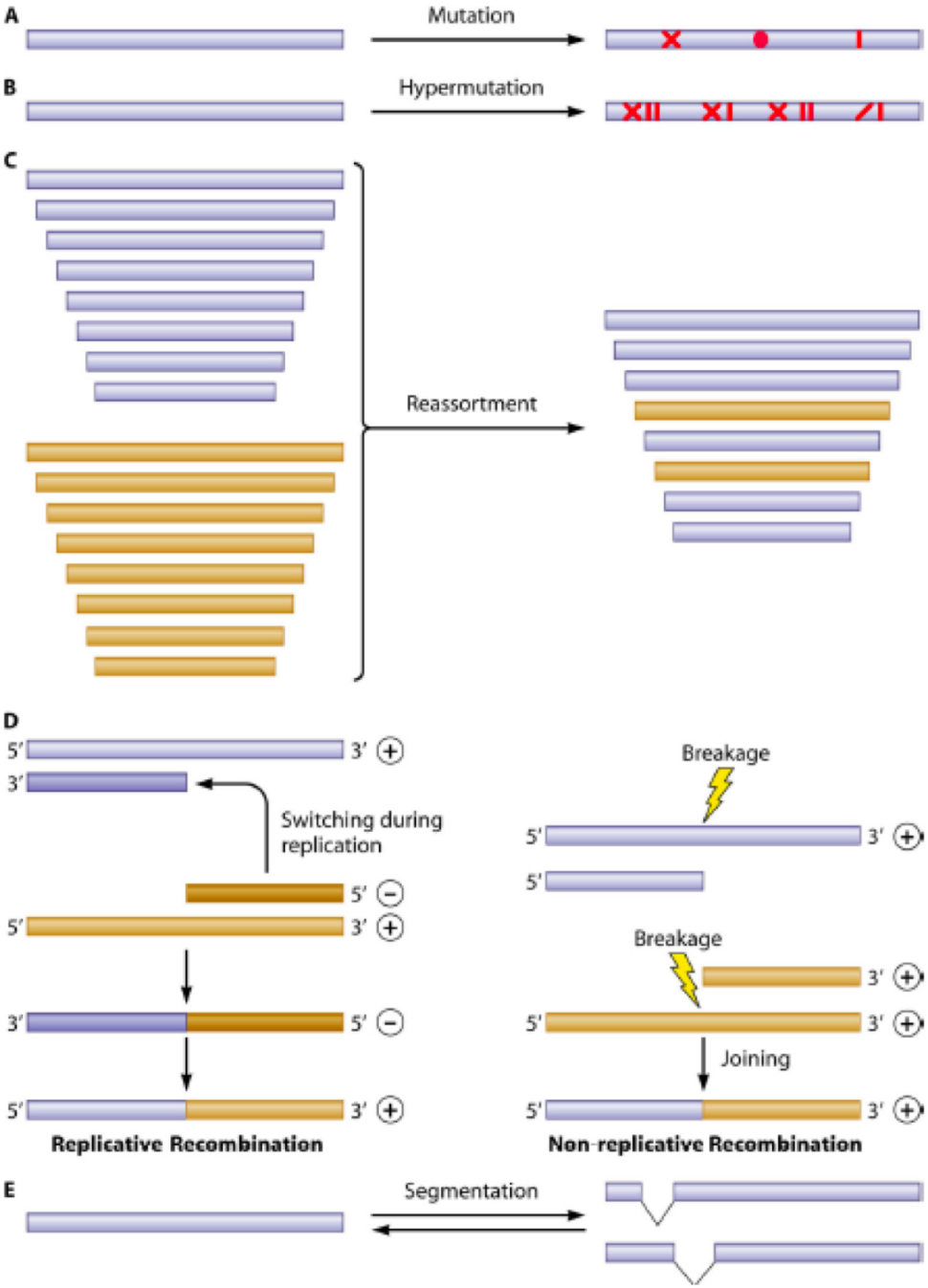

Lethal defection
Overt lethality
Error threshold
Quasispecies distribution
Random sequences
No information
+
−
1
Copying fidelity
0


Interference, lethal mutations


Mutagen-resistance mutations


Large
population
passage
Plaque-to-plaque
transfer
Fitness increase
Fitness decrease

Viral sex

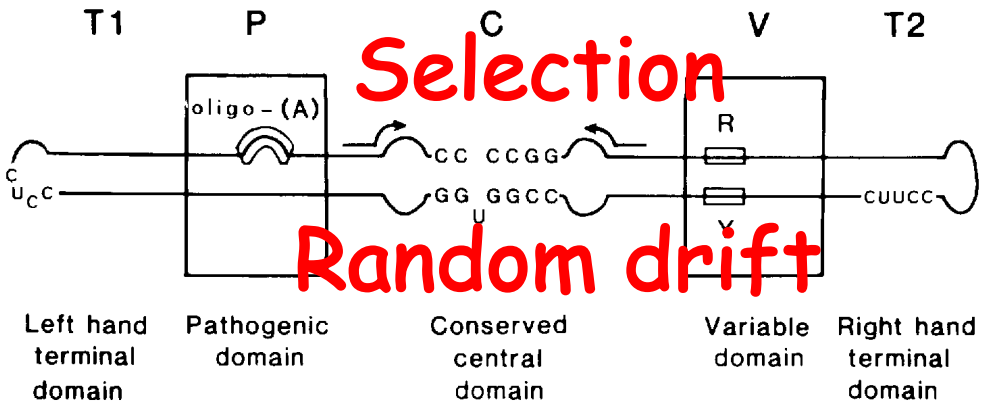

A
Mutation
B
Hypermutation
C
Reassortment
D
5′
3′
Switching during replication
Replicative Recombination
Breakage
Breakage
Joining
Non-replicative Recombination
E
Segmentation

Mutation

Selection

Random drift

Contingency



Advantages of microorganisms for evolution experiments

✓ They are easy to propagate and enumerate.

✓ They reproduce quickly, which allows experiments to run for many generations.

✓ They allow large populations in small spaces, which facilitates experimental replication.

✓ They can be stored in suspended animation and later revived, which allows the direct comparison of ancestral and evolved types.

✓ Many microbes reproduce asexually and the resulting clonality enhances the precision of experimental replication.

✓ Asexuality also maintains linkage between a genetic marker and the genomic background into which it is placed, which facilitates fitness measurements.

✓ It is easy to manipulate environmental variables, such as resources, as well as the genetic composition of founding populations.

✓ There are abundant molecular and genomic data for many species, as well as techniques for their precise genetic analysis and manipulation.

Peculiarities of RNA viruses

✓ High genetic variability. Orders of magnitude greater than for DNA-based organisms.

✓ High mutation rates: 2.5×10^{-4} s/s/r for VSV, 5×10^{-5} for TEV and 2.5×10^{-3} for CChMVd. Such mutation rates are consequence of the lack of proofreading mechanisms in viral RdRp.

✓ Compacted genome: 11162 nts for VSV and 9494 for TEV.

✓ Huge numbers of generations per time unit: $\sim10^{3}$ PFU/cell in 6 - 8 hpi for VSV or $\sim10^{6}$ LFU/g 5 dpi for TEV.

✓ The variability is a key factor for pathogenicity.

✓ It is impossible talking about a single defined entity. Instead we shall talk on a distribution of genomes centered around a more frequent one: *Quasispecies*.

✓ Relatively easy to map genotypes into phenotypic space.

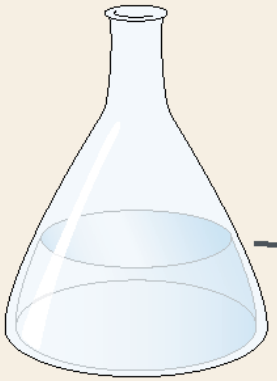
✓ Viral infectious diseases represent the most important threat to animal and plant health.

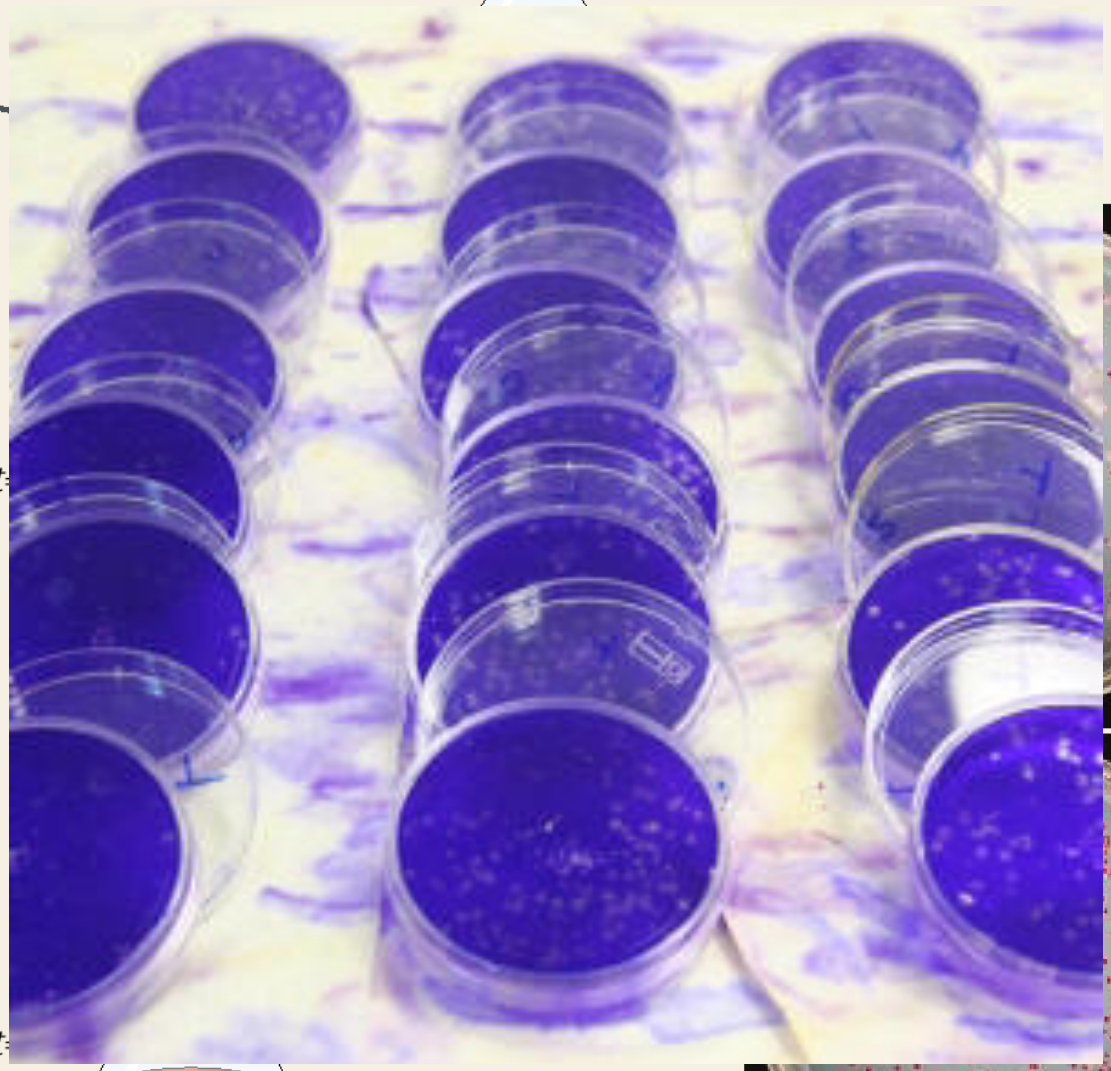
Quantifying the degree of adaptation

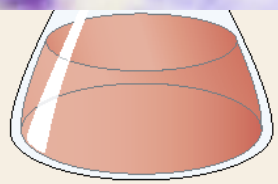
- ✓ Relative Darwinian fitness: Reproductive ability of a given viral strain in a defined environment. This is a macroscopic property that includes components such as replication, transcription and encapsidation rates as well as virion stability in the environment, resistance to antiviral responses and transmission or adsorption rates.
- ✓ Competition experiments between ancestral and evolved strains.

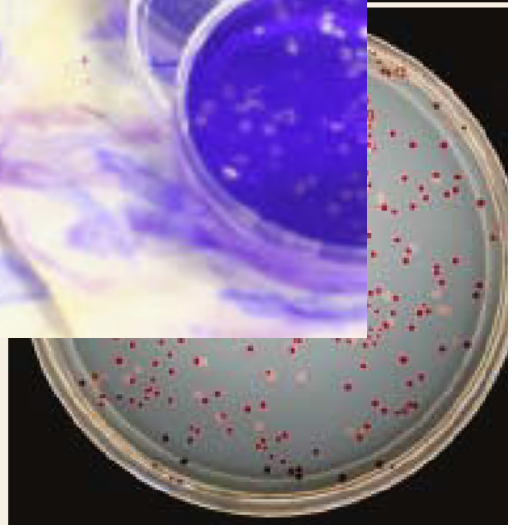
Ancestral

Evolved

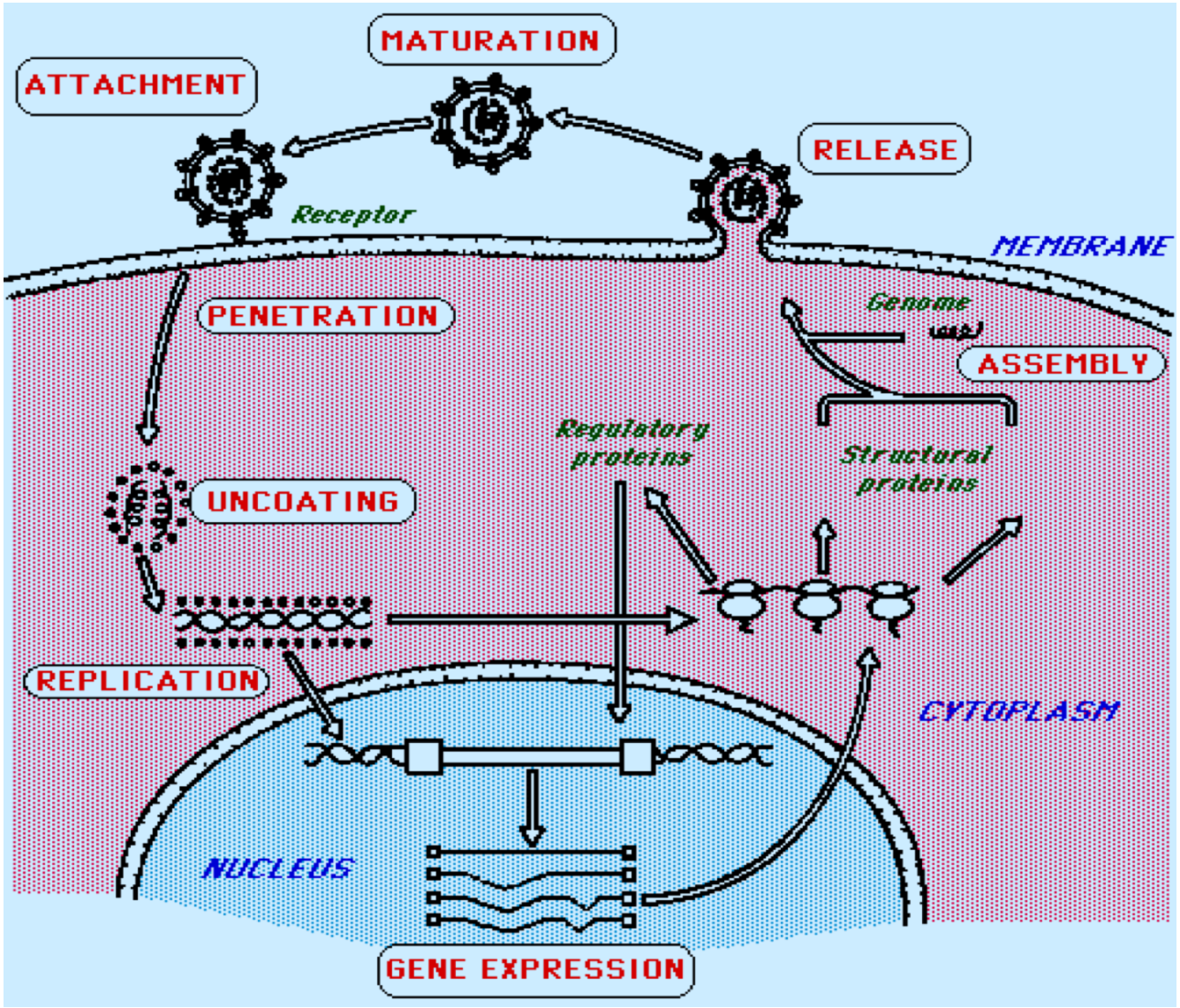








✓ Within-cell fitness components for a lytic virus.



✓ Among-cells fitness components:

Virion stability in intercellular space
Diffusion rate
Escape from cell-mediated and humoral immune responses
Tissue tropism
Abundance of receptors and binding affinity
…

✓ Among-individuals R_0 fitness components:

Virion long-term stability
Transmission rates and routes (horizontal, vertical)
Transmission scheme (early, late)
Virulence (as far as it is related to replication rate)
Virus induced mortality
Virus induced morbidity
Vector specificity and abundance
Host range and host abundance
…

✓ Considerations:

Frequency-dependent fitness

Elena *et al.* (1997) *Evolution* **51**, 984.
Yuste *et al.* (2002) *J. Gen. Virol.* **83**, 103.

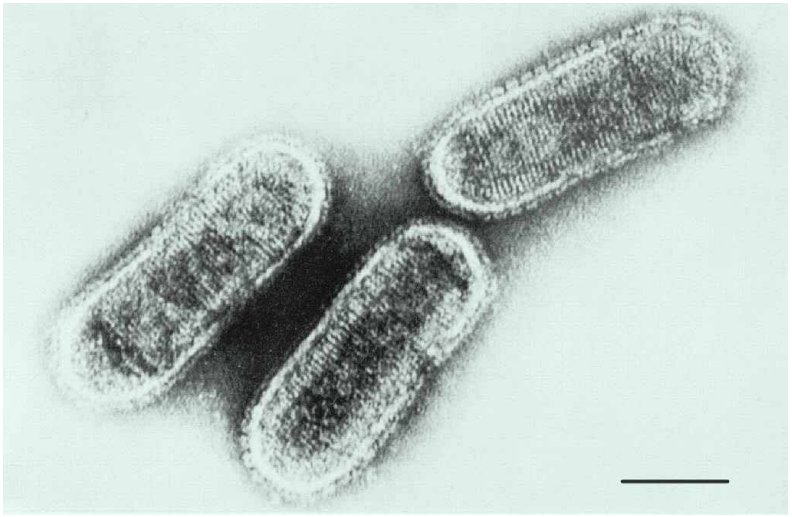
Density-dependent fitness

Bordería & Elena. (2002) *Infect. Genet. Evol.* **2**, 137.
Novella *et al.* (2004) *J. Virol.* **78**, 5799.
Sevilla *et al.* (1998) *J. Gen. Virol.* **79**, 2971.

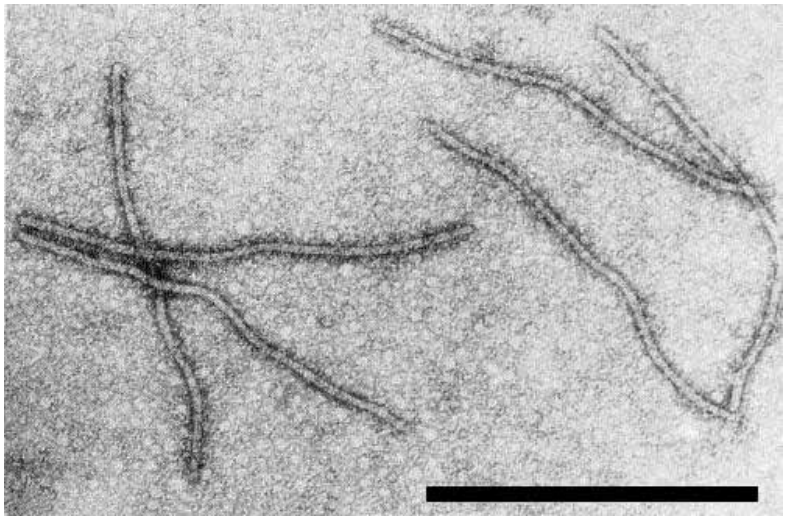
Cooperation and defection

Turner & Chao (1999) *Nature* **398**, 441.
Turner & Chao (2003) *Am. Nat.* **161**, 497.
Chao & Elena (2017) *Proc. R. Soc. B* **284**, 20170228.

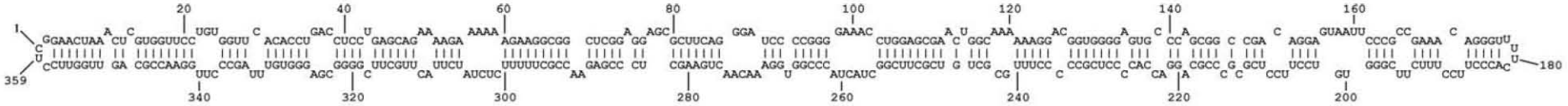
Vesicular stomatitis *rhabdovirus*



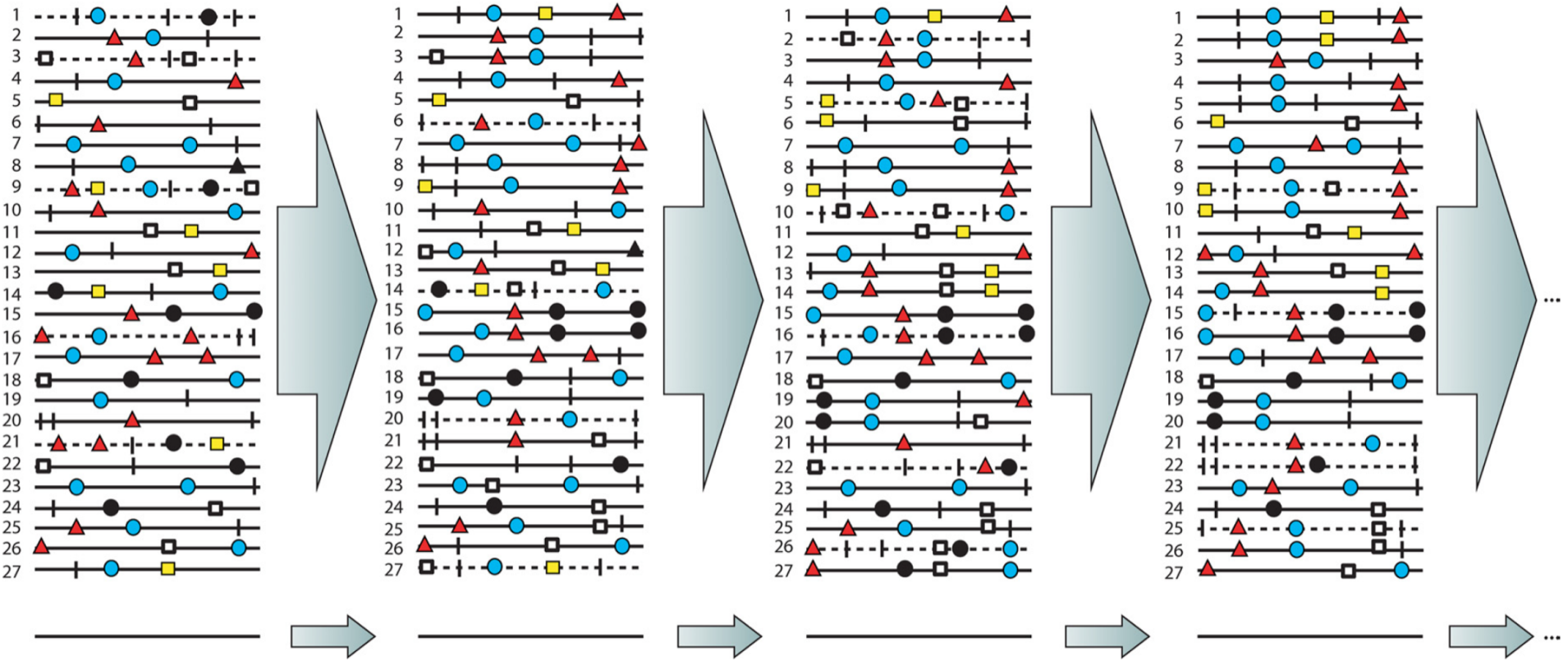
Tobacco etch *potyvirus*



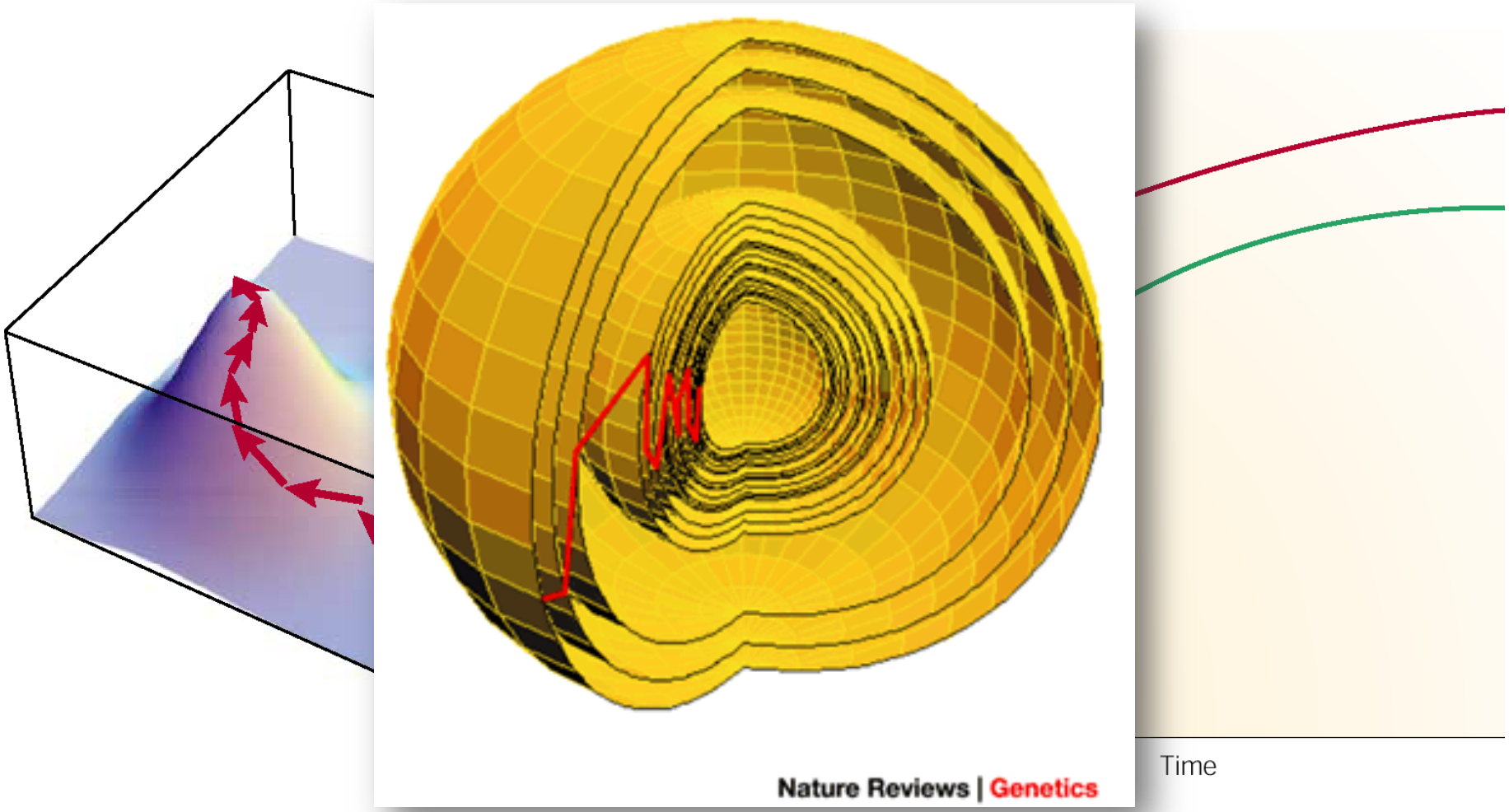
The viroids



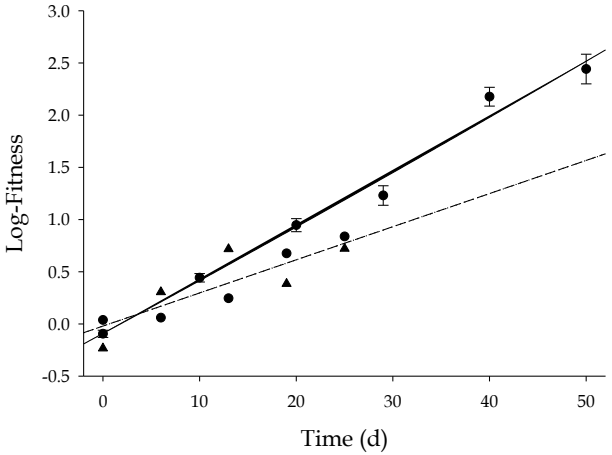
The dynamics of evolutionary adaptation


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
...

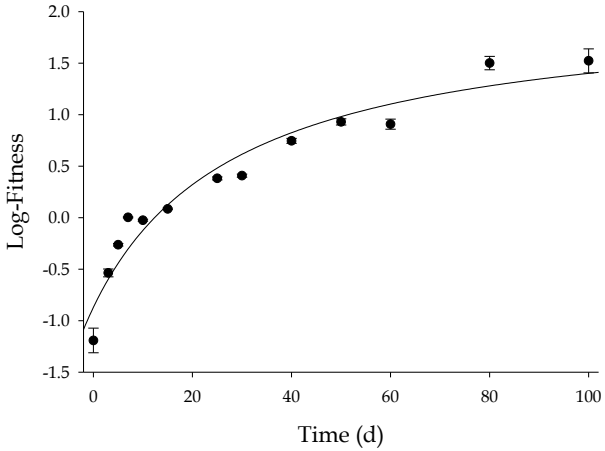
Walking throughout Wright's adaptive landscapes



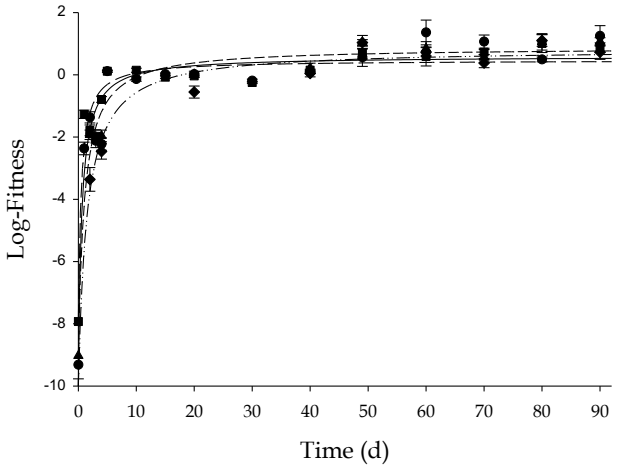
MARM *C*



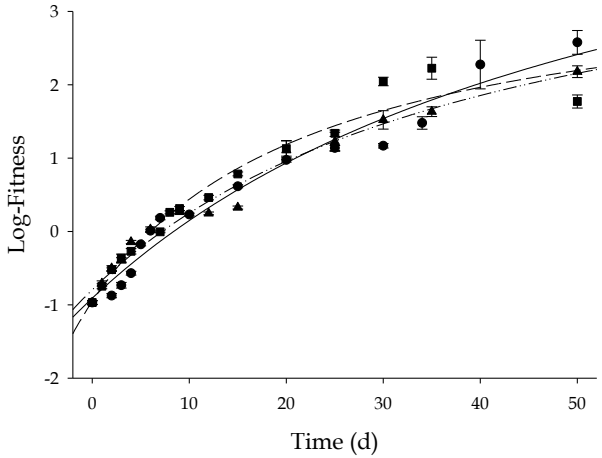
MARM *D*



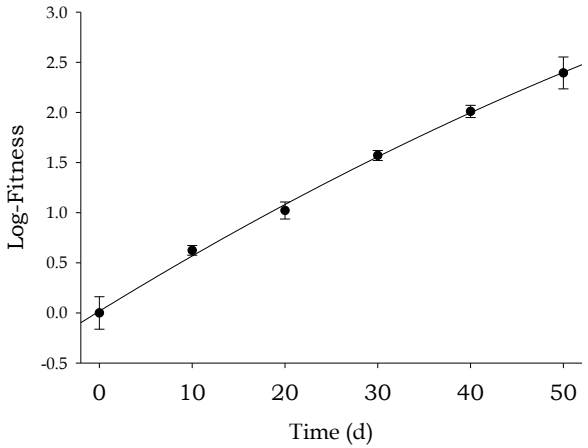
MARM *F*



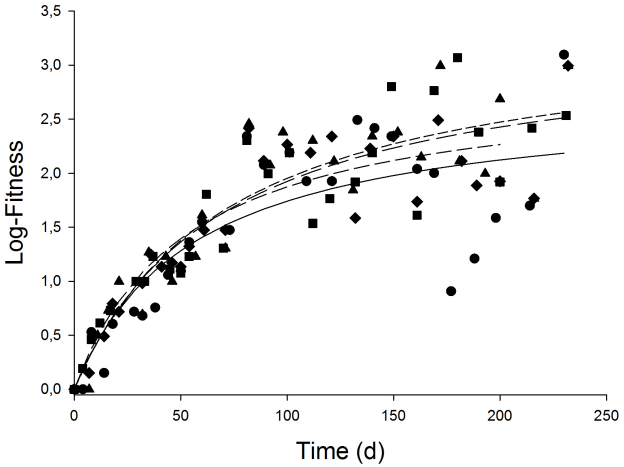
MARM *N*



MARM *U*



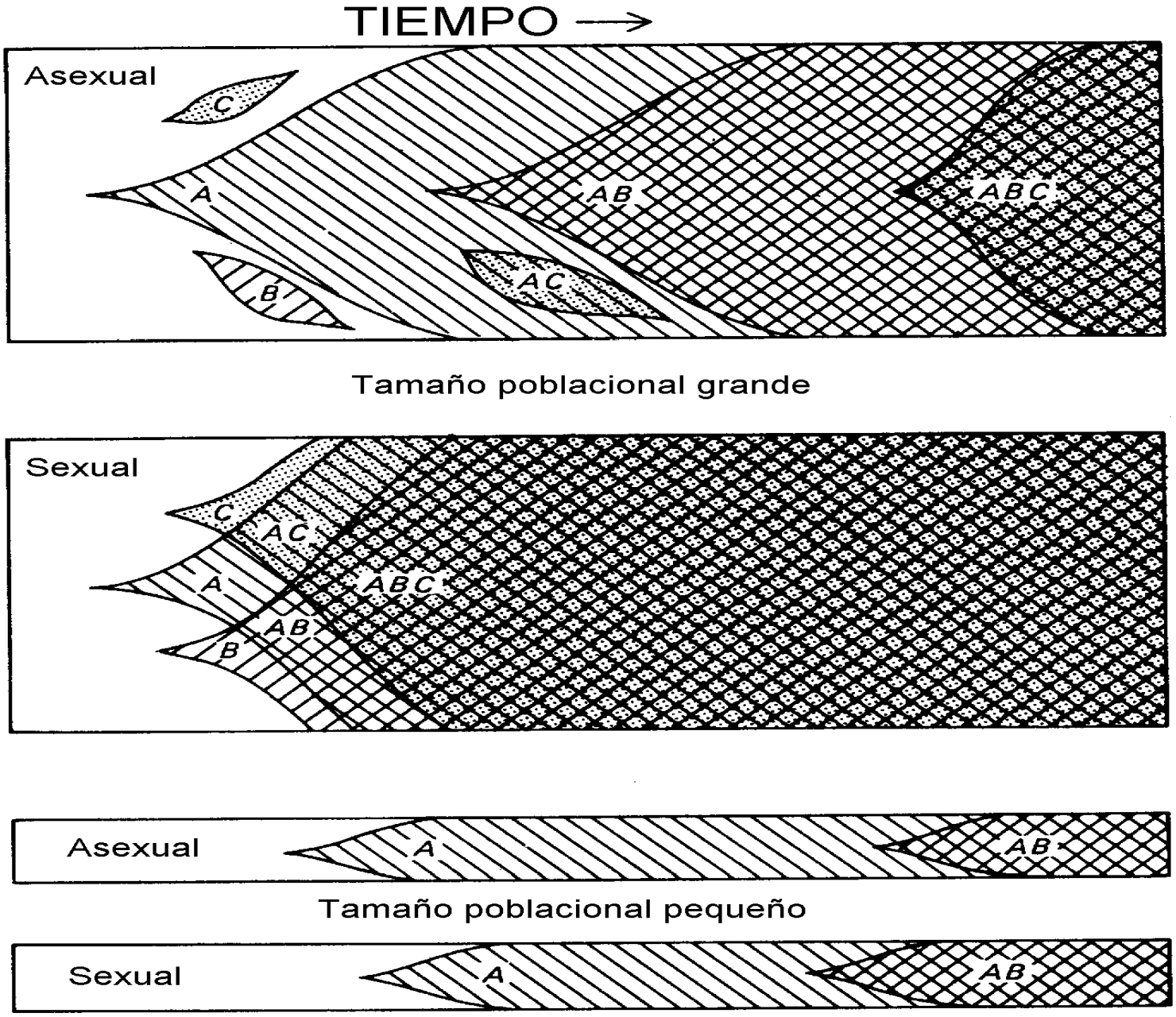
WT



Novella *et al.* (1995) *Proc. Natl. Acad. Sci. USA* **92**, 5841
Elena *et al.* (1998) *Evolution* **52**, 309
Novella *et al.* (1999) *J. Virol.* **73**, 1668

Limits to viral evolution?

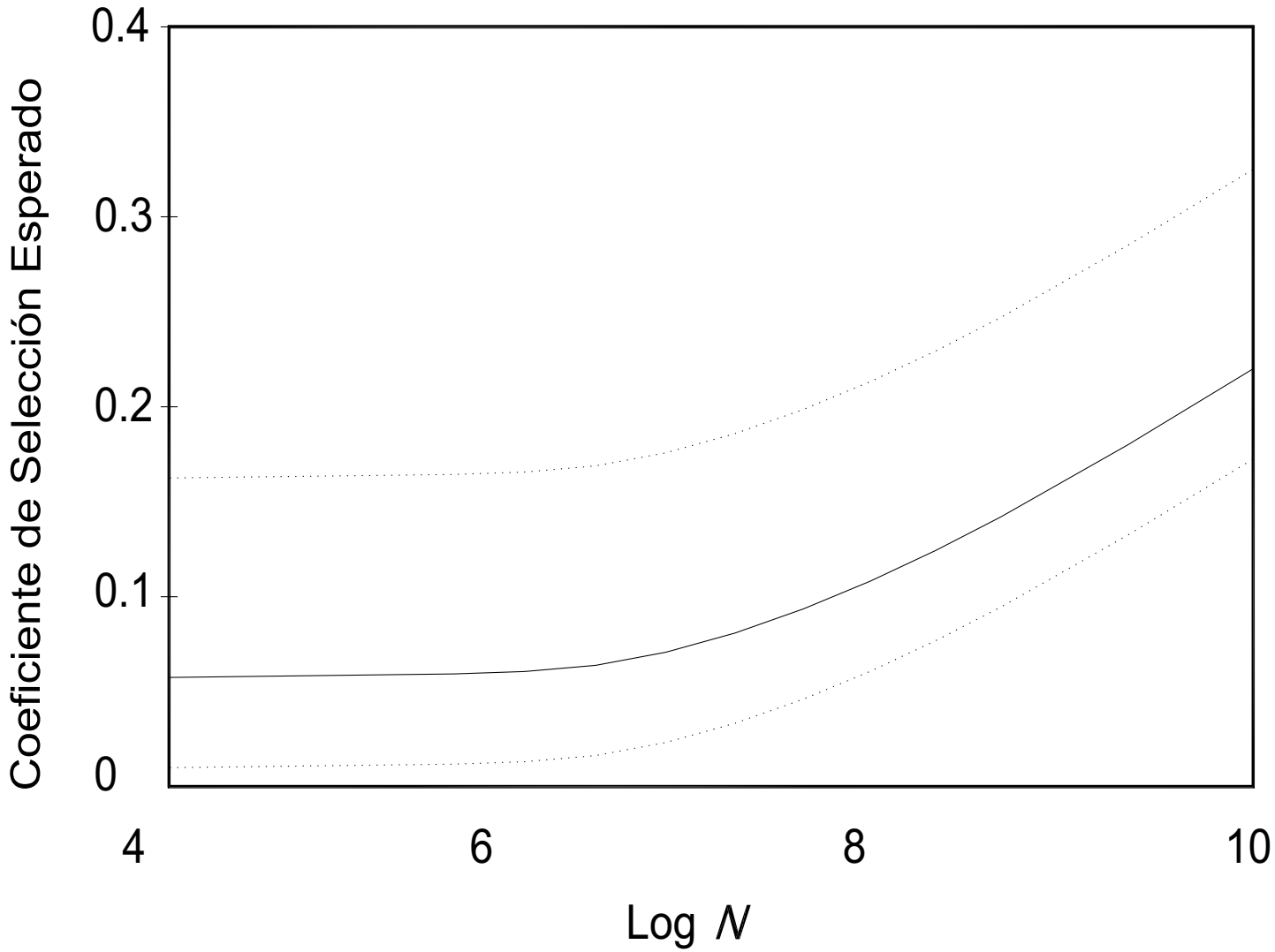
Clonal Interference



✓ Predictions of Gerrish & Lenski (1998) *Genetica* **102**, 127:

1. The probability of fixation of a given beneficial mutation decreases both with population size and mutation rate.

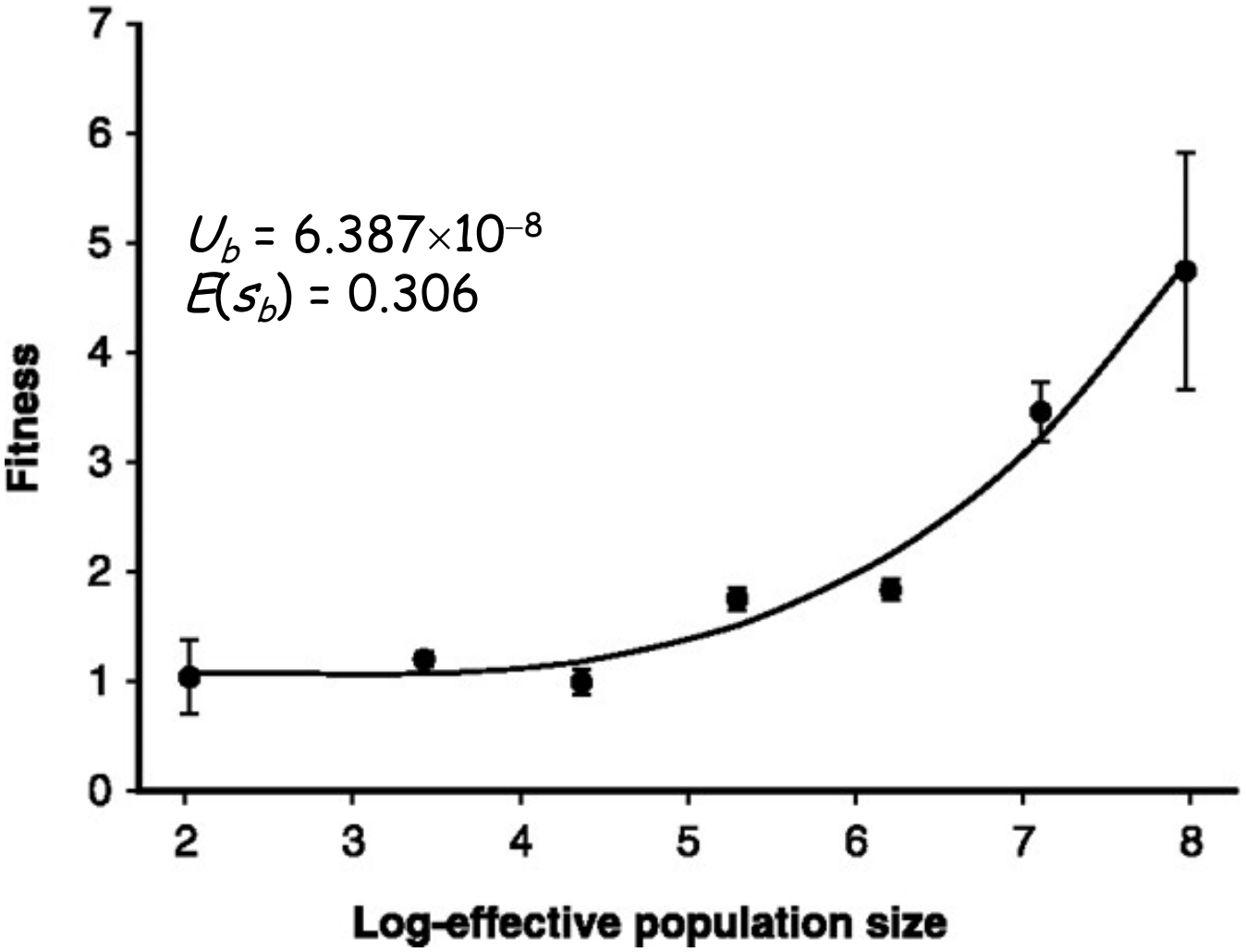
2. As population size or mutation rate increase, adaptive substitutions result in larger fitness increases.



Gerrish & Lenski (1998) *Genetica* **102-103**, 127

Fixed beneficial effects increase with N_e

$r = 0.893$, 5 *df*, 1-tailed $P = 0.003$



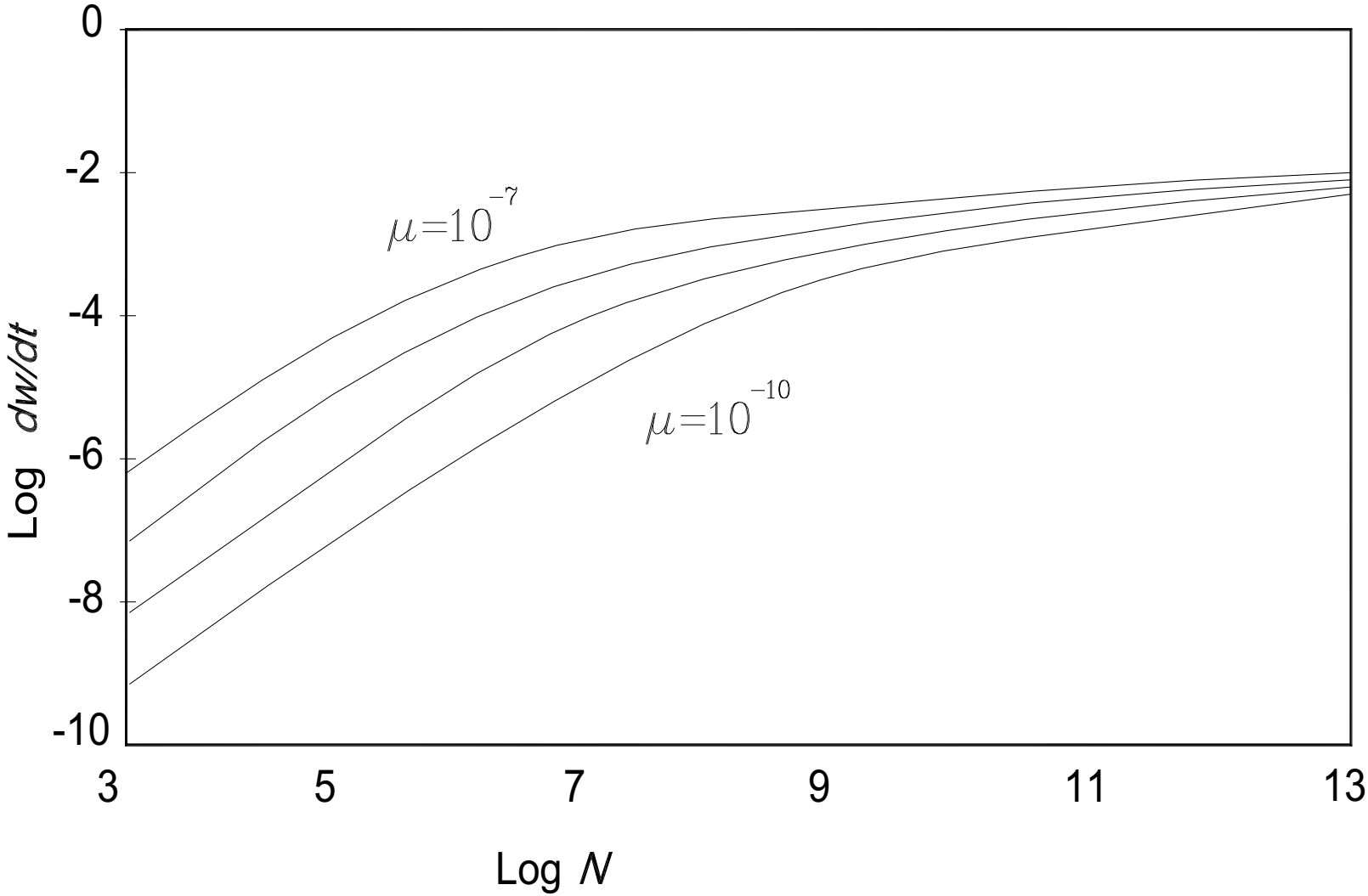
Miralles *et al.* 1999. *Science* **285**, 1745

✓ Predictions of Gerrish & Lenski. 1998. *Genetica* **102**, 127:

1. The probability of fixation of a given beneficial mutation decreases both with population size and mutation rate.

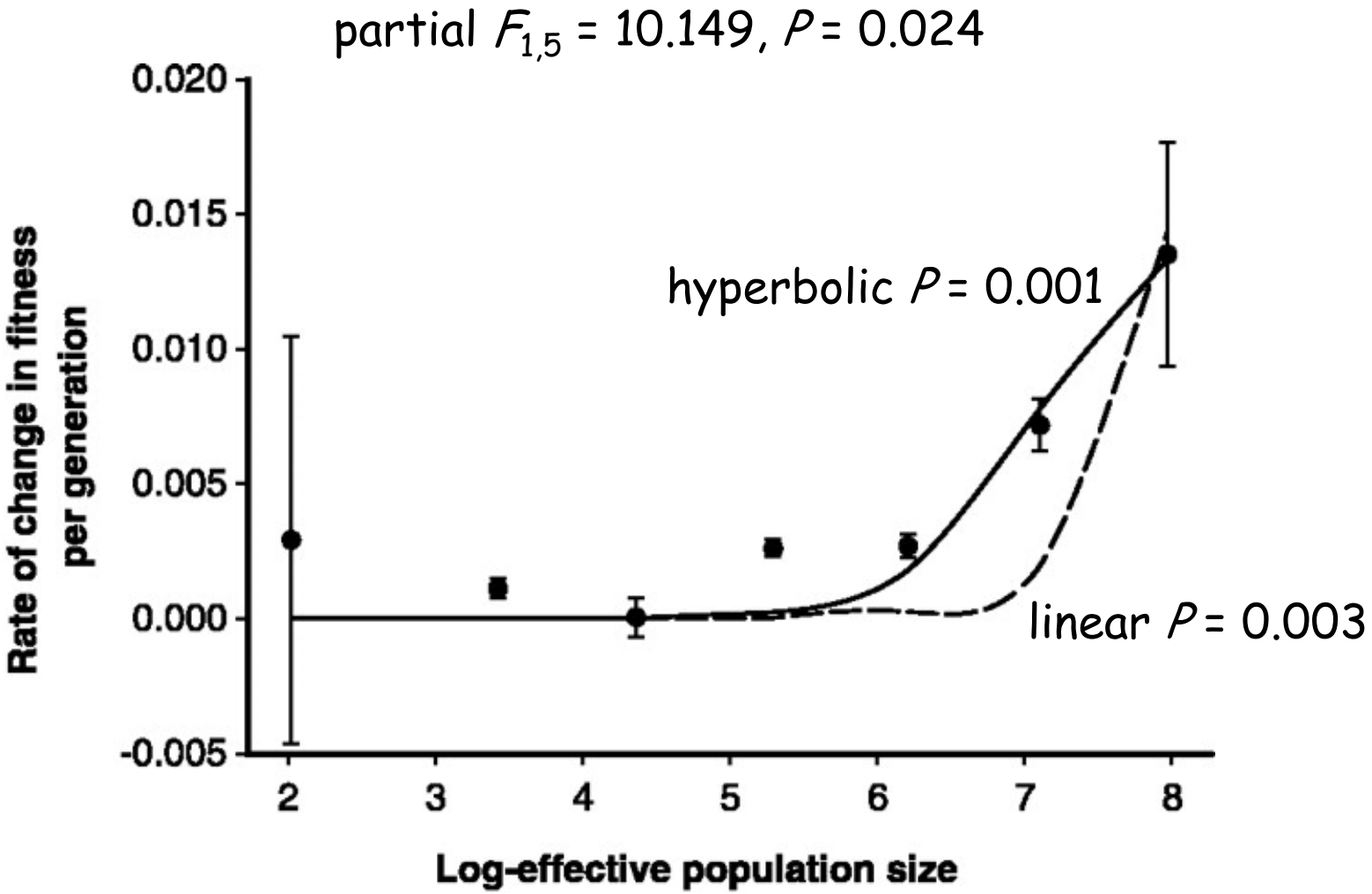
2. As population size or mutation rate increase, adaptive substitutions result in larger fitness increases.

3. The rate of adaptation is an increasing, but decelerating, function of both population size and mutation rate.



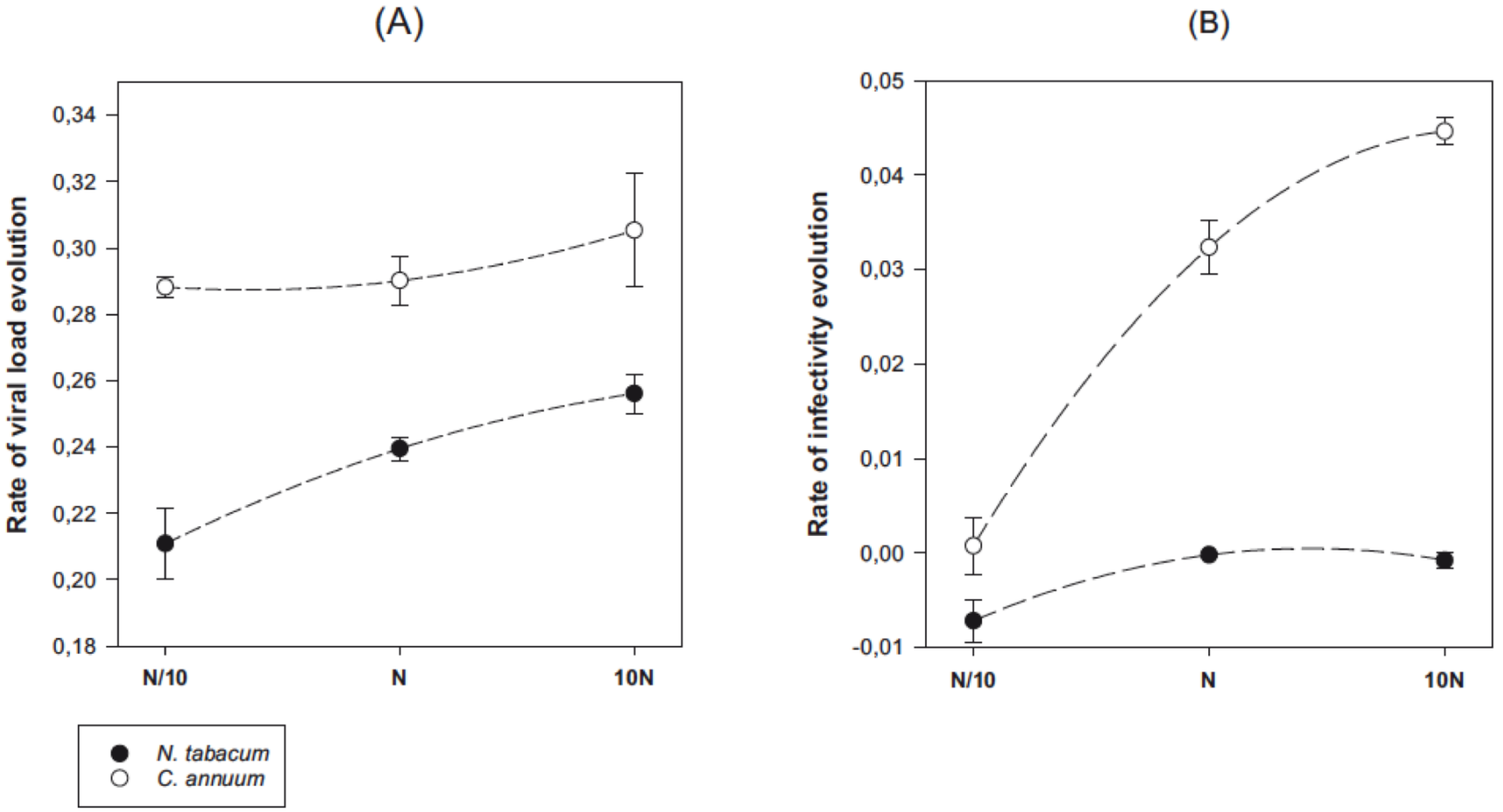
Gerrish & Lenski (1998) *Genetica* **102-103**, 127

Effect of N_e in the rate of adaptation



Miralles *et al.* 1999. *Science* **285**, 1745

Effect of N_e in the rate of adaptation



Navarro *et al.* (2017) *Europhys. Lett.* **120**, 38001

✓ Predictions of Gerrish & Lenski. 1998. *Genetica* **102**, 127:

1. The probability of fixation of a given beneficial mutation decreases both with population size and mutation rate.

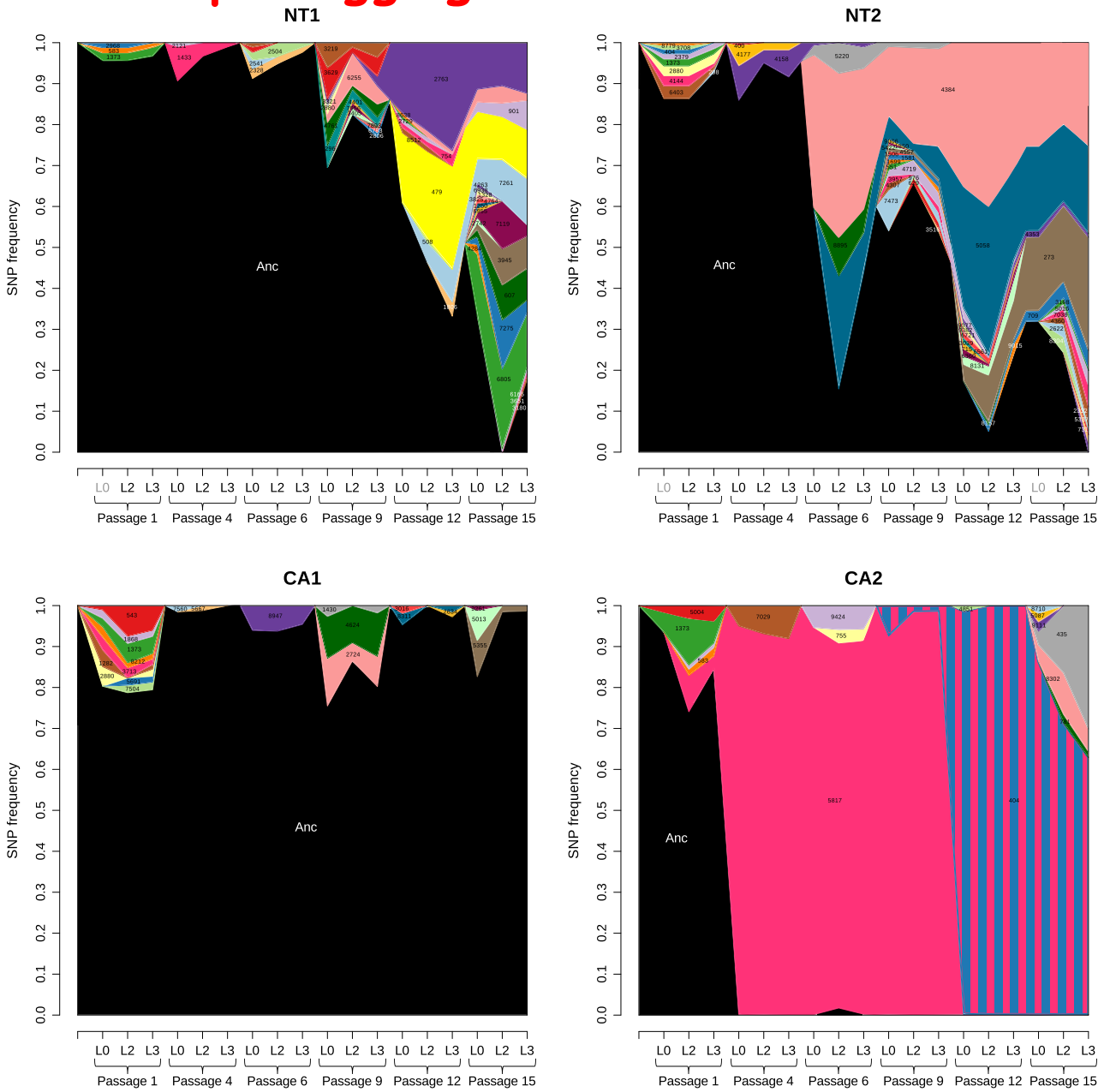
2. As population size or mutation rate increase, adaptive substitutions result in larger fitness increases.

3. The rate of adaptation is an increasing, but decelerating, function of both population size and mutation rate.

4. Beneficial mutations that become transiently common but do not achieve fixation due to interfering beneficial mutations are relatively abundant.

5. Transient polymorphisms may give rise to a "leapfrog" effect, where the most common genotype at a given moment might be less closely related to the immediately preceeding one than with an earlier genotype.

The leap-frogging of beneficial mutations



Cuevas *et al.* (2015) *Mol. Biol. Evol.* **32**, 1132

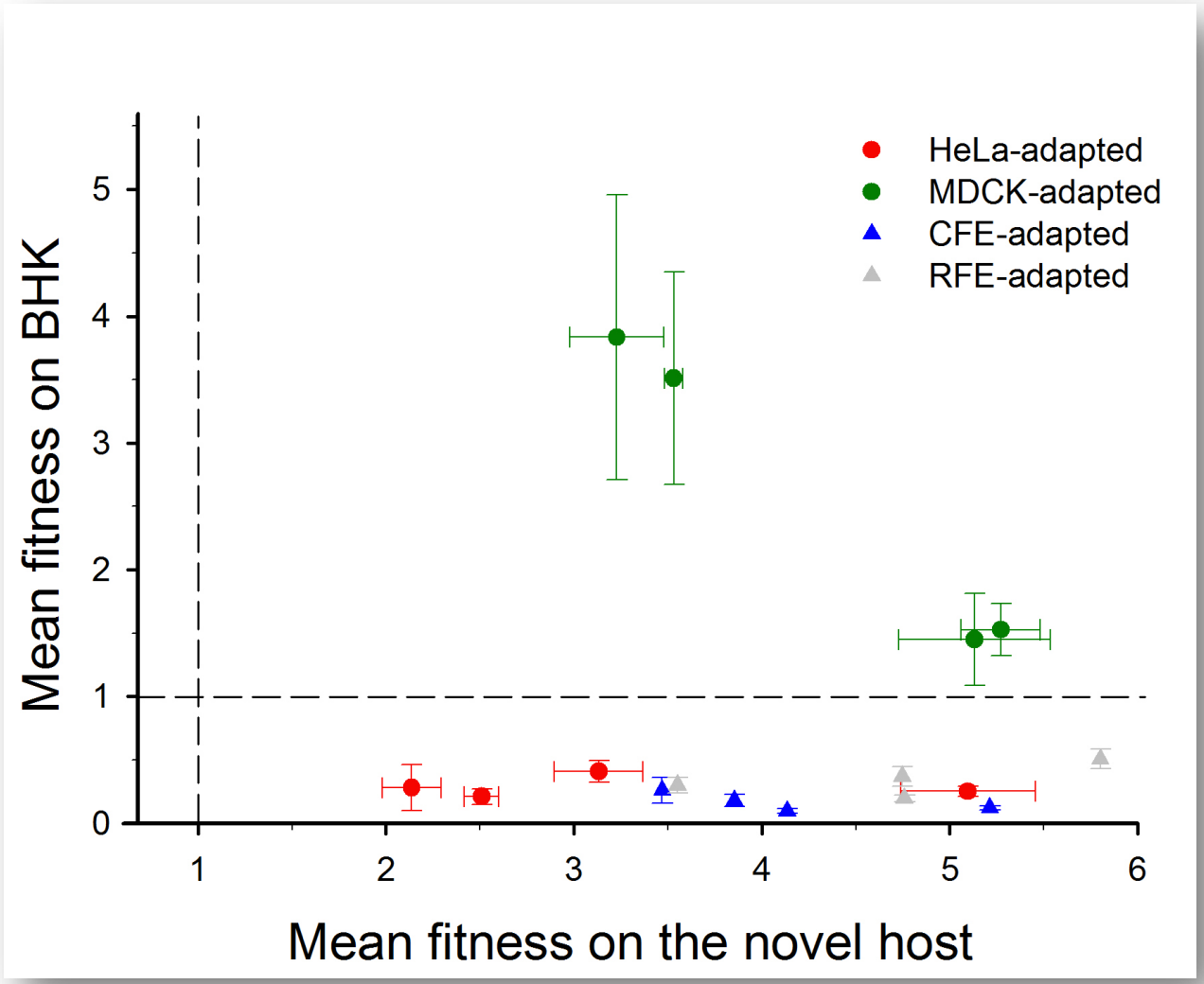
✓ Three important conclusions can be drawn from the clonal interference model:

1. Adaptive substitutions appear as discrete events. They do not occur simply as the result of a single mutational event but instead represent the best possible candidate

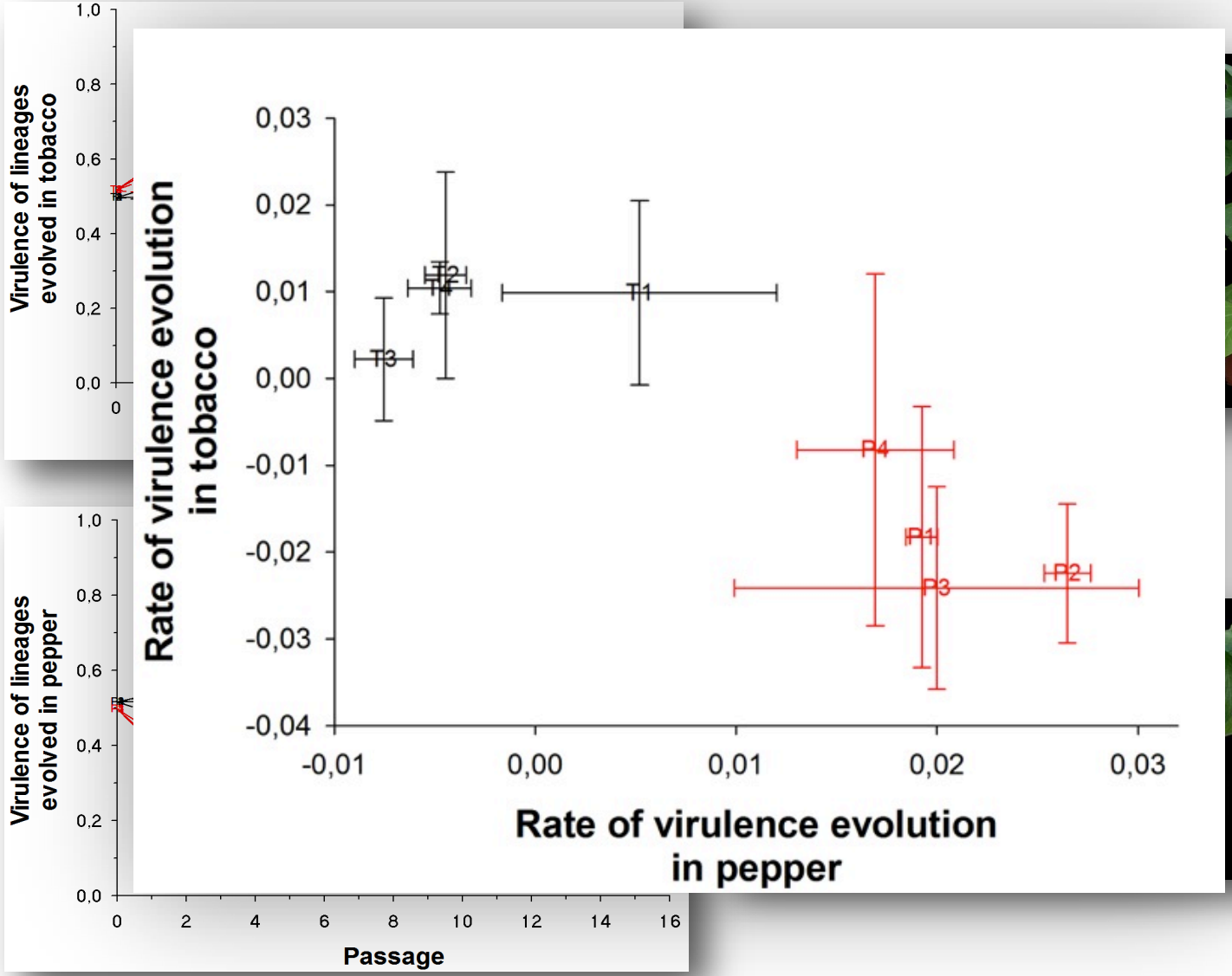
2. Because the rate of adaptation is not positively affected by increases in mutation availability, it is questionable whether the high mutation rate shown by RNA viruses has evolved because of the adaptive capacity it confers. Instead, a decrease in mutation rate would benefit the population by slowing the accumulation of deleterious mutations.

3. Consequently, high mutation rates are the result of a trade-off between keeping a compacted genome and the costs of maintaining the enzymatic system required for error detection and correction.

Evolution in a new single host promotes specialization and pays the cost of host-range expansion.

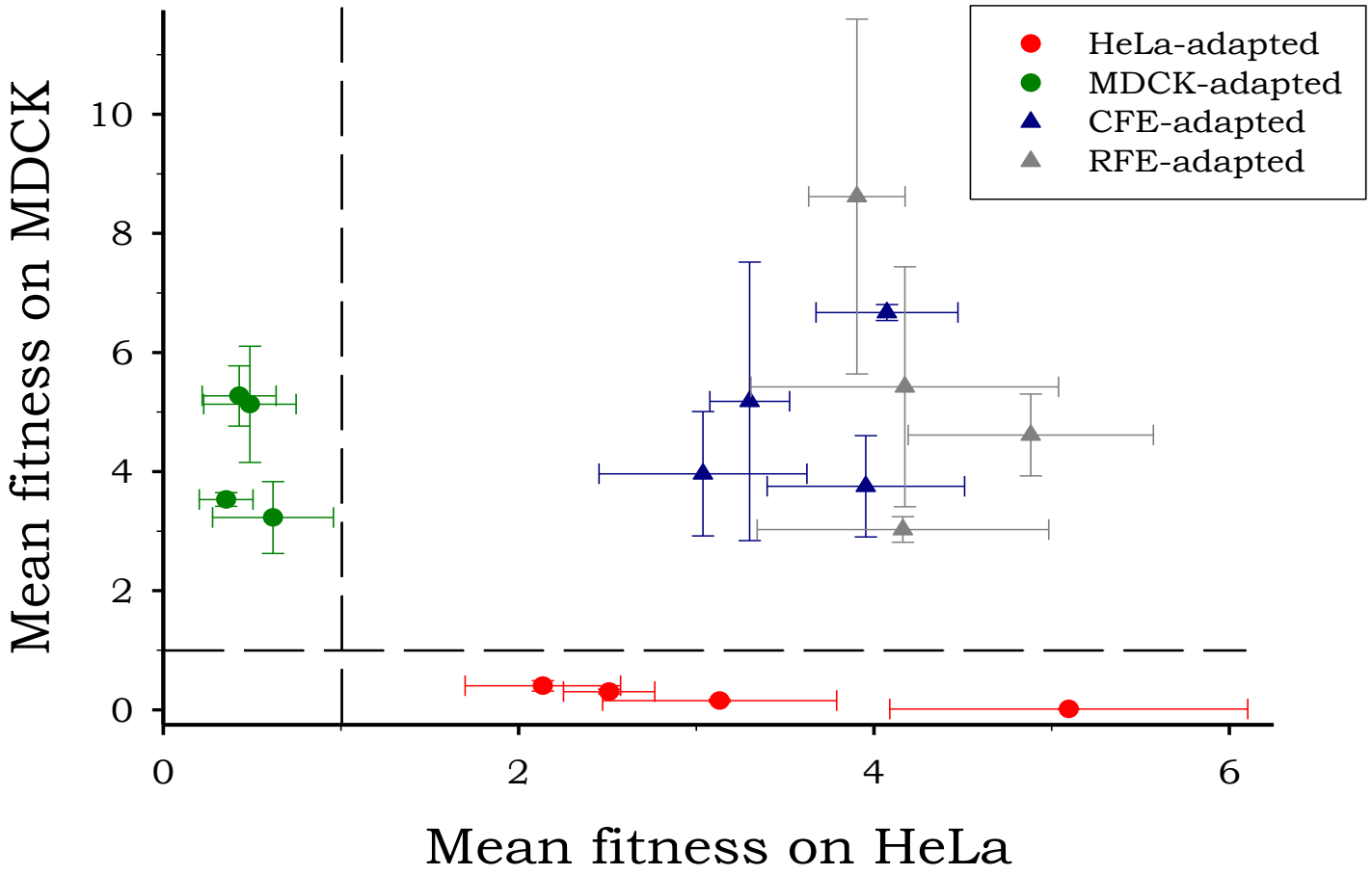


Turner & Elena (2000) *Genetics* **156**, 1465

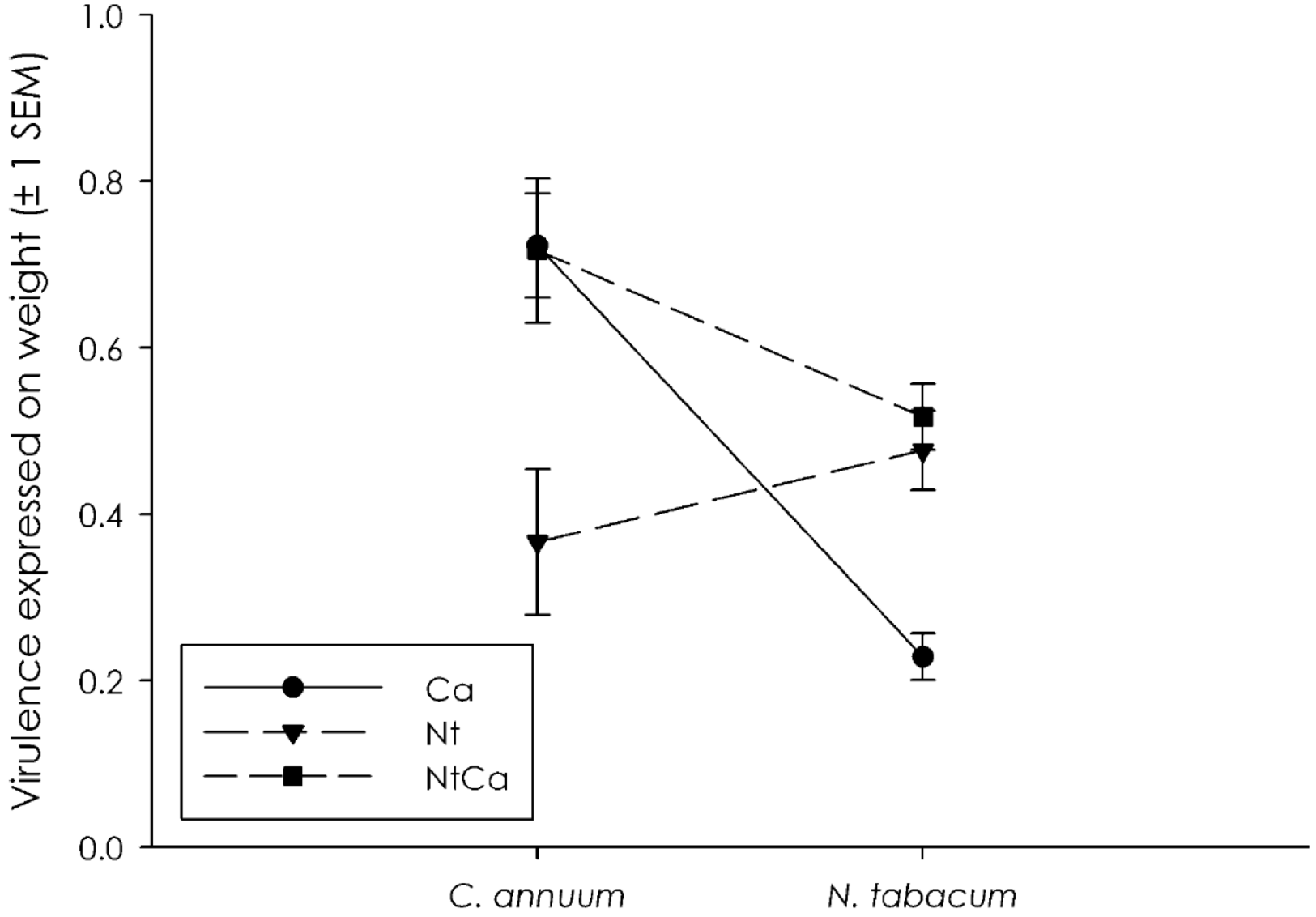


Agudelo-Romero *et al.* (2008) *Infect. Genet. Evol.* **6**, 806
Bedhomme *et al.* (2012) *Mol. Biol. Evol.* **29**, 1481

No-cost generalists evolve under host-switching regimes

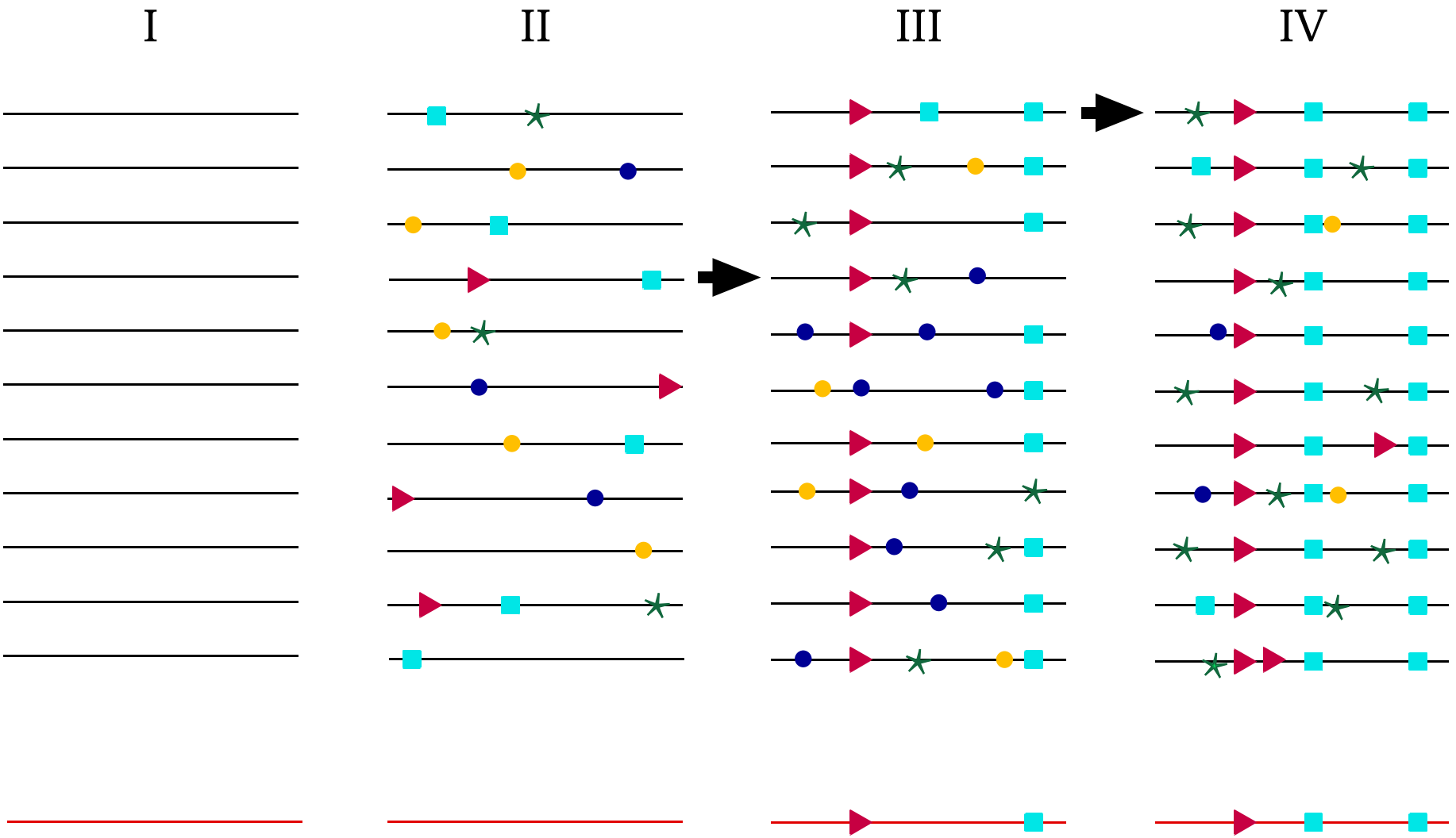


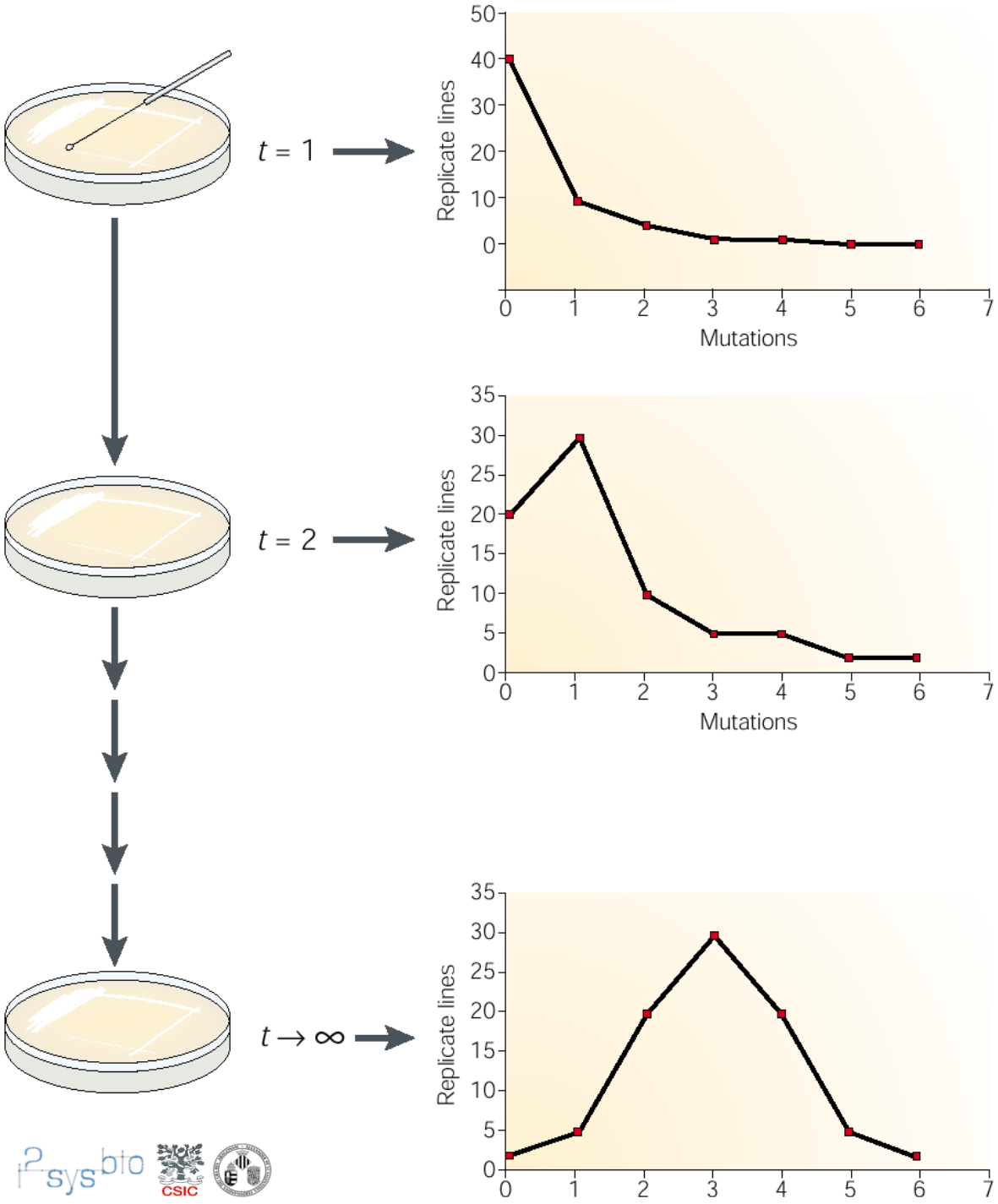
Turner & Elena (2000) *Genetics* **156**, 1465



S. Bedhomme *et al.* (2012) *Mol. Biol. Evol.* **29**, 1481-92

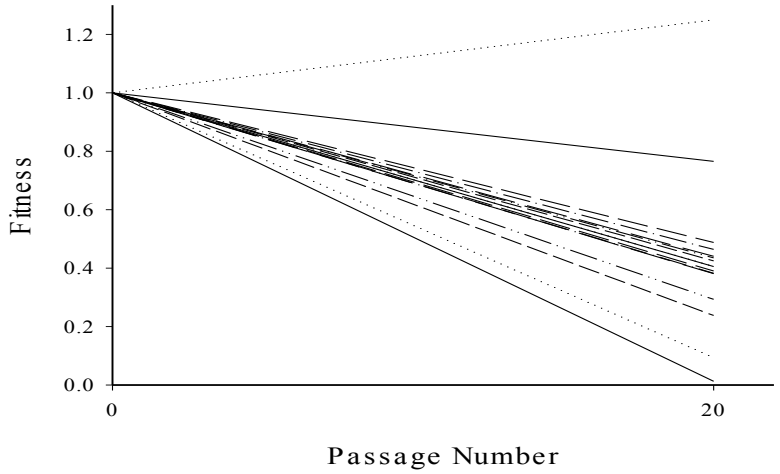
Drift and decay in very small populations


I
II
III
IV

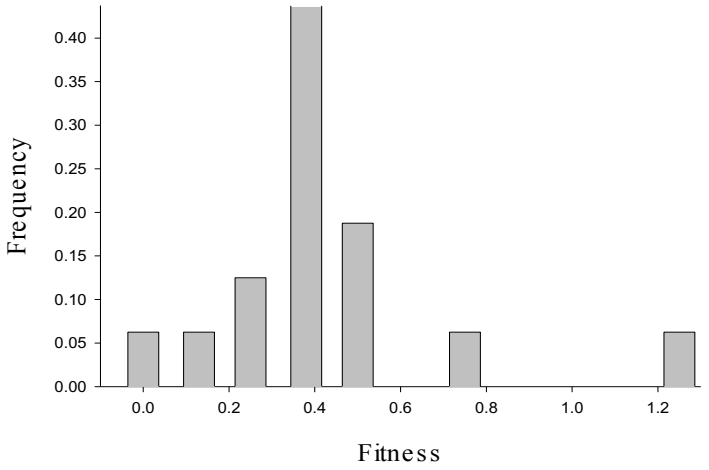

t = 1
t = 2
t → ∞
Replicate lines
Mutations
50
40
30
20
10
0
35
25
15
5
0
1
2
3
4
5
6
7

MARM X

$U_d E(s_d) = -0.012 \pm 0.001\ d^{-1}$

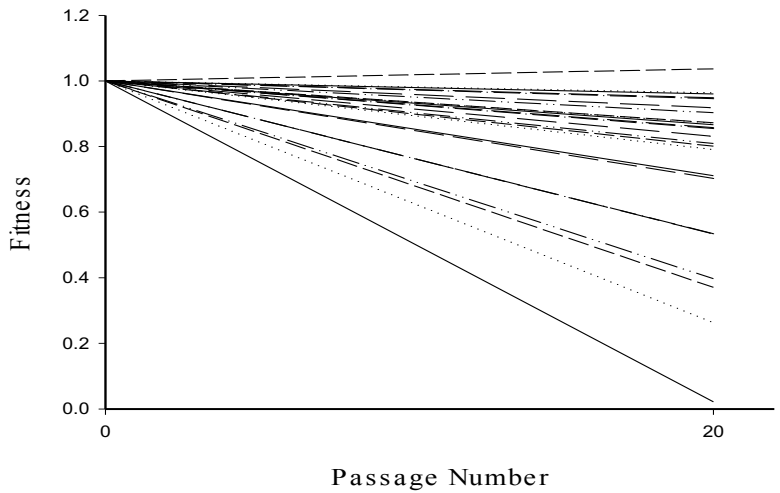


MARM X

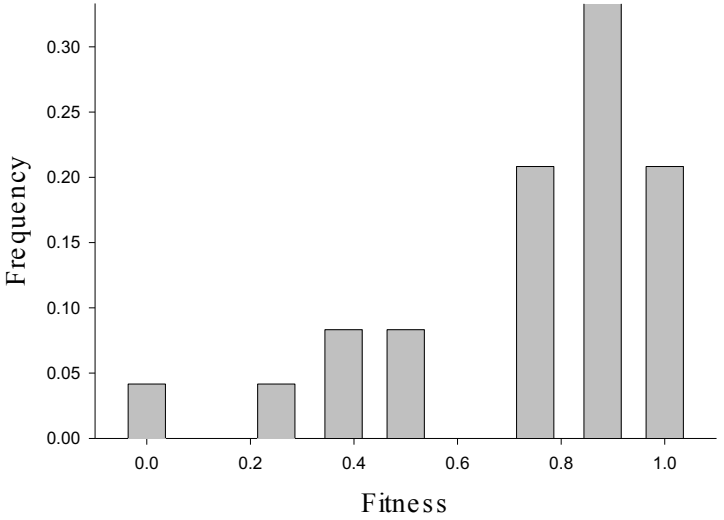


MARM C

$U_d E(s_d) = -0.008 \pm 0.0001\ d^{-1}$

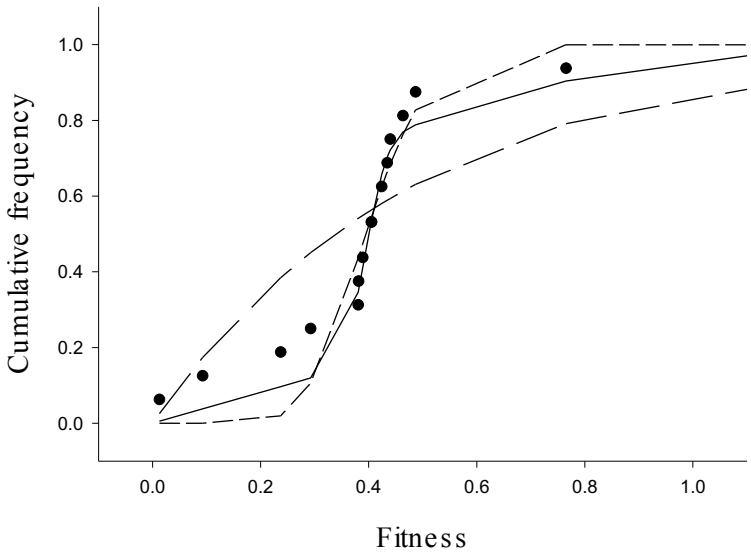


MARM C

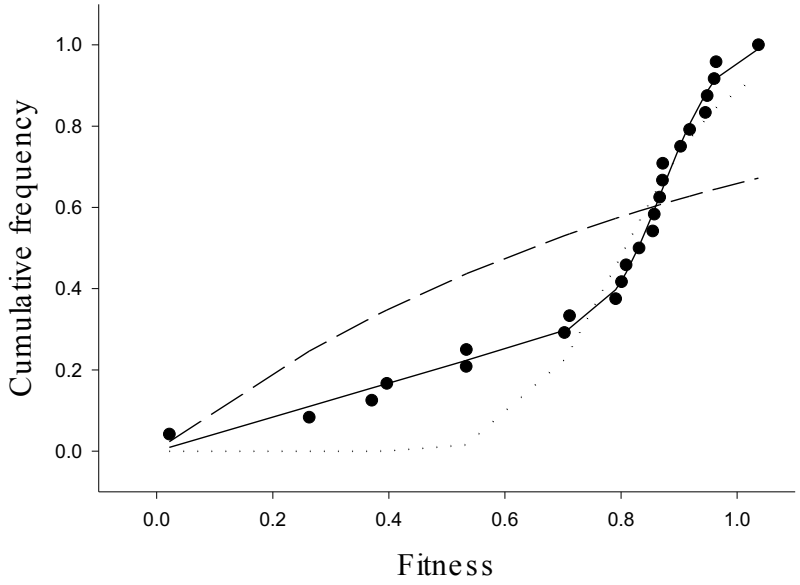


Duarte *et al.* (1992) *Proc. Natl. Acad. Sci. USA* **89**, 6015
Clarke *et al.* (1993) *J. Virol.* **67**, 222
Elena & Moya (1999) *J. Evol. Biol.* **12**, 1078

MARM X



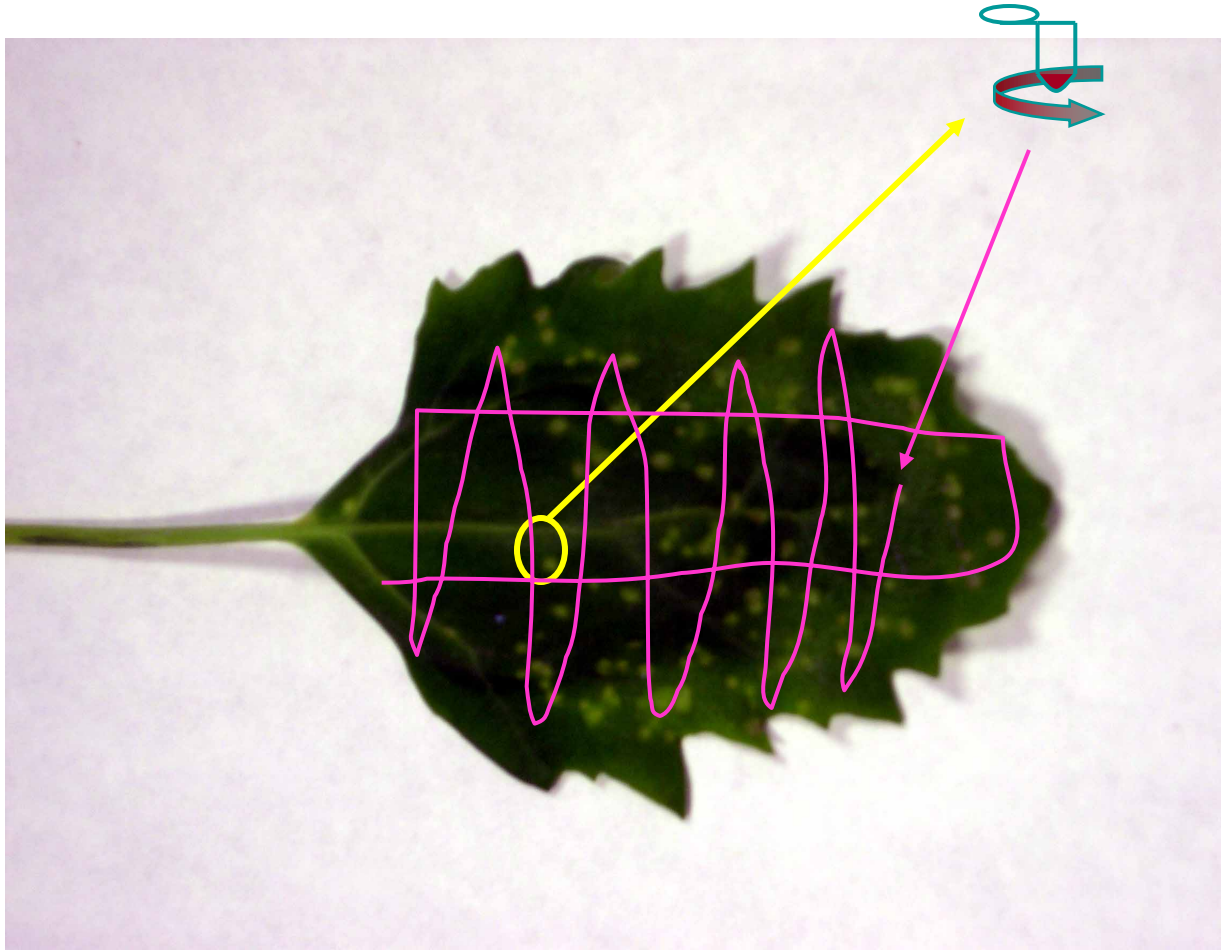
MARM C

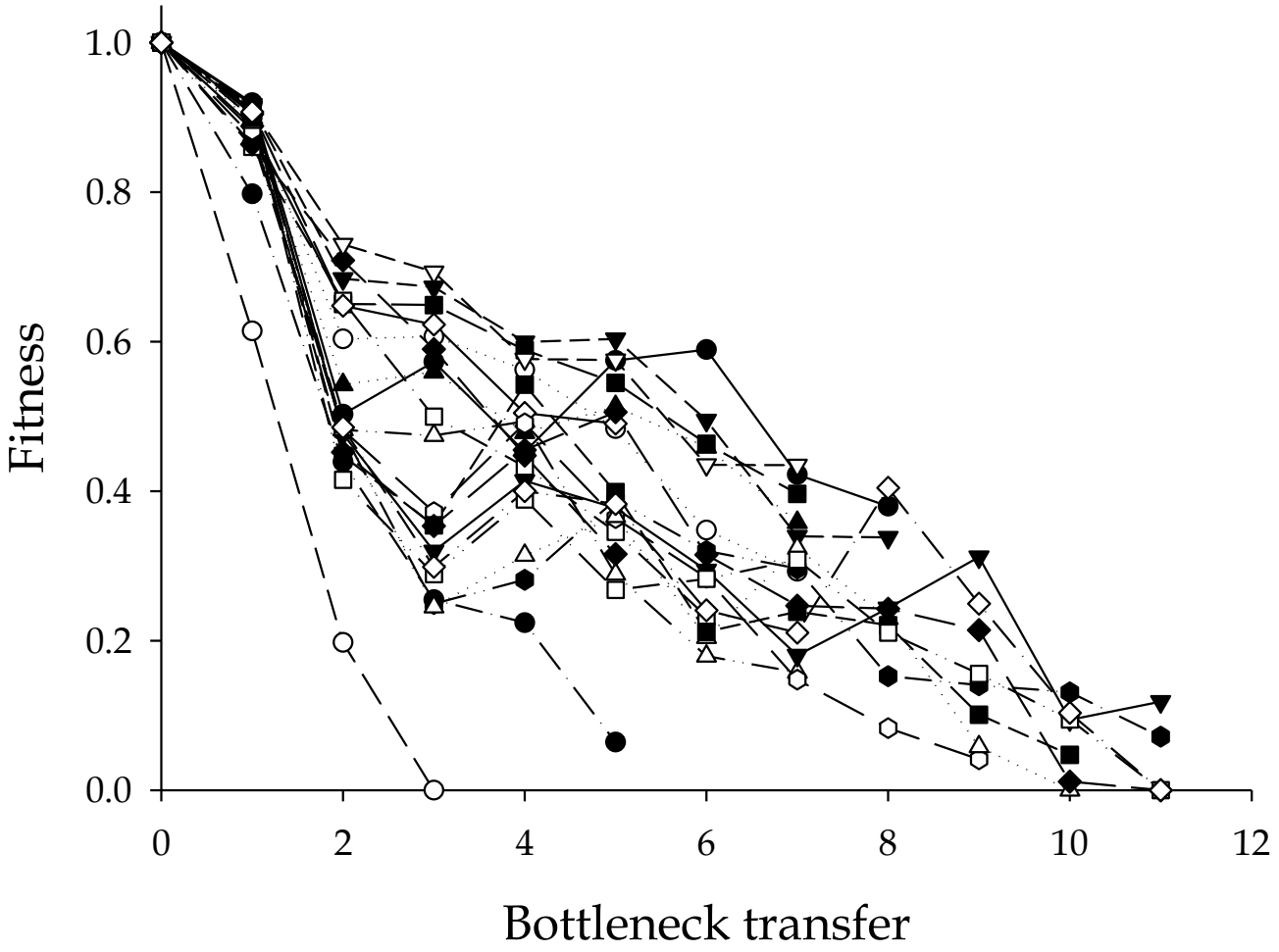


Gamma + uniform R^2 = 0.962
U_d = 3.090
$E(s_d)$ = -0.100

Gamma + uniform R^2 = 0.992
U_d = 1.768
$E(s_d)$ = -0.150

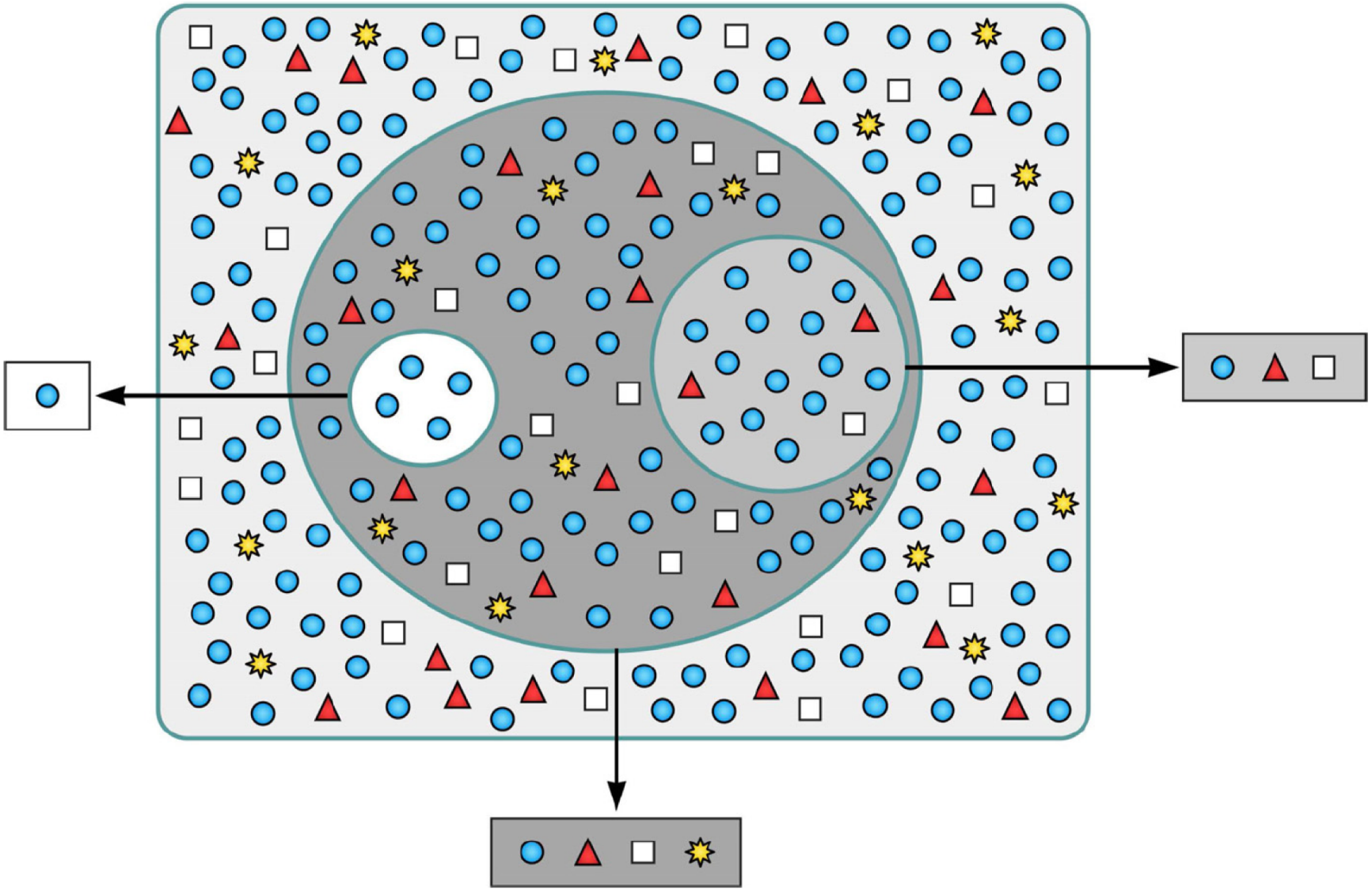
Elena & Moya (1999) *J. Evol. Biol.* **12**, 1078





$UE(s) = -0.121 \pm 0.005$, $t_{19} = 25.374$, 1-tailed $P < 0.001$

Bottleneck size and genetic diversity



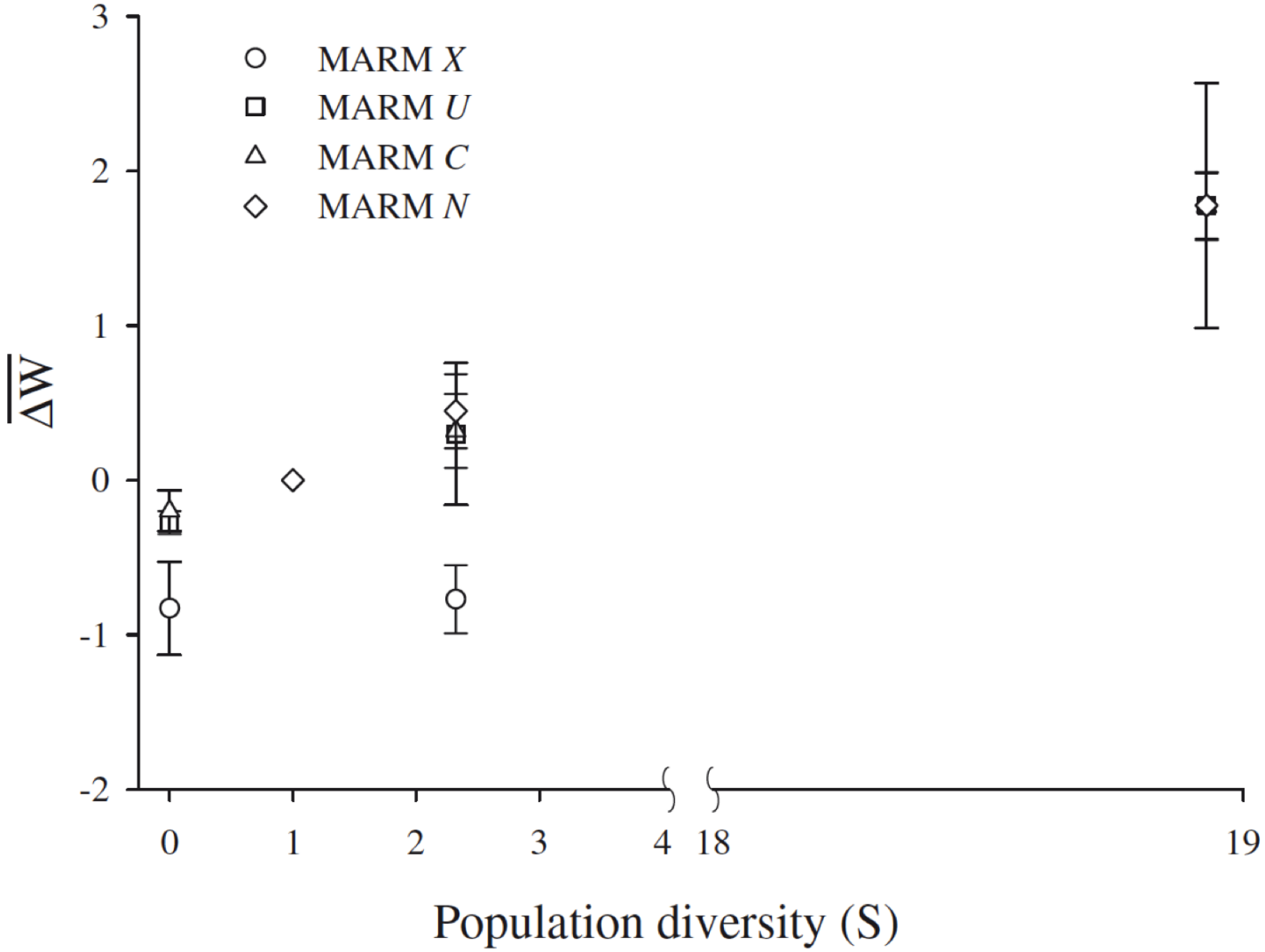
Effect of bottleneck size and genotype in the outcome of Muller's ratchet. Dynamic makes reference to the number of clones isolated and pooled at every of 20 consecutive infectious passages of *n* plaques-to-*n* plaques.

MARM	Dynamic	Fitness	t_5	1-tailed *P*
X	5-to-5	1.7±0.2	5.725	0.001
	30-to-30	3.0±0.4	0.161	0.439
U	5-to-5	1.3±0.2	0.813	0.226
C	5-to-5	1.2±0.2	1.292	0.126
N	5-to-5	0.55±0.01	3.048	0.986
	2-to-2	0.38±0.01	0.271	0.398

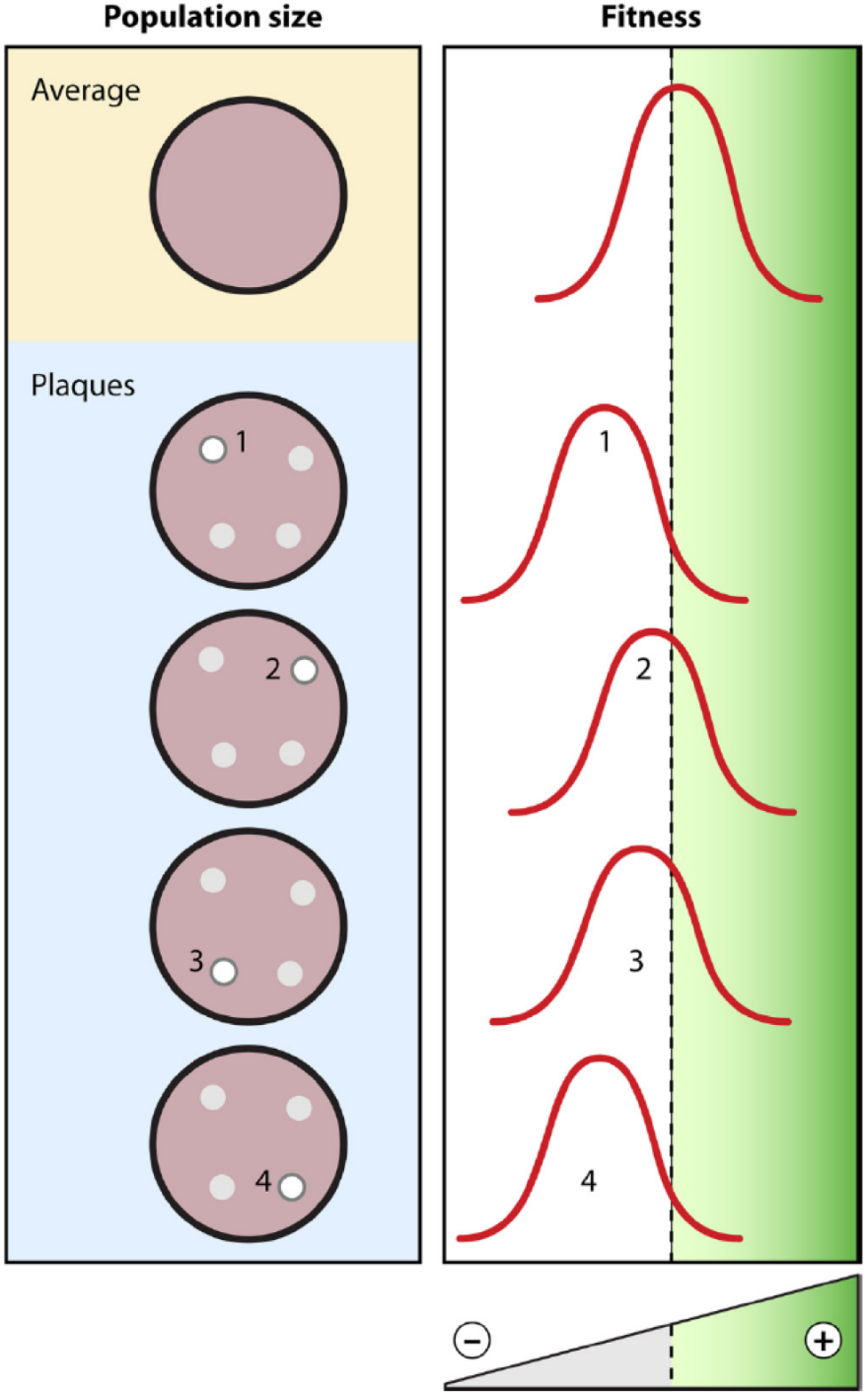
The initial fitnesses were 3.05±0.03, 1.0±0.2, 0.91±0.03 and, 0.38±0.01 respectively. In all cases, experiments were 6-fold replicated.

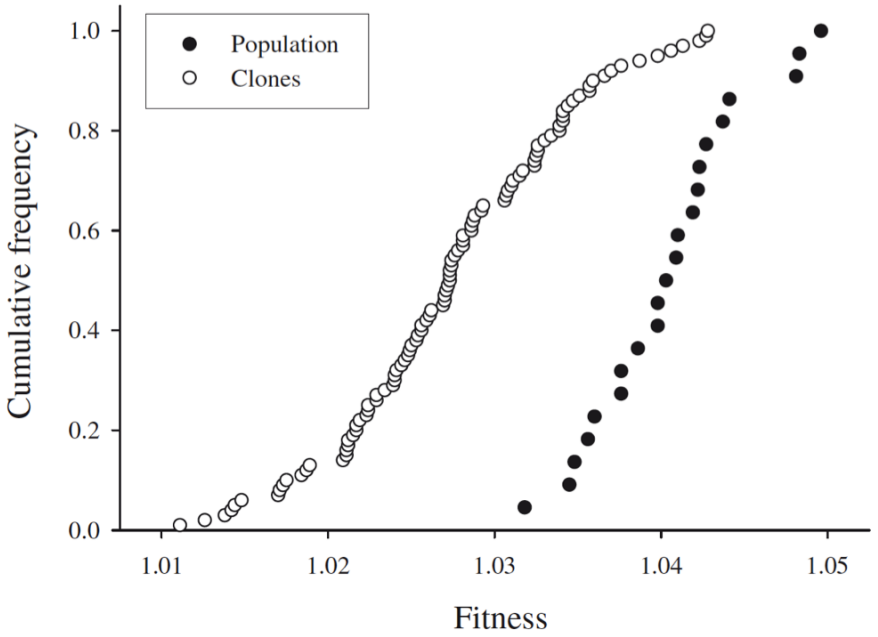
Novella *et al.* (1995) *J. Virol.* **69**, 2869
Novella *et al.* (1996) *Mol. Gen. Genet.* **252**, 733

$r = 0.957$, 9 d.f., 1-tail $P < 0.001$



$$S = -\sum_{i=1}^{n} p_i \log_2 p_i \approx \log_2 N$$

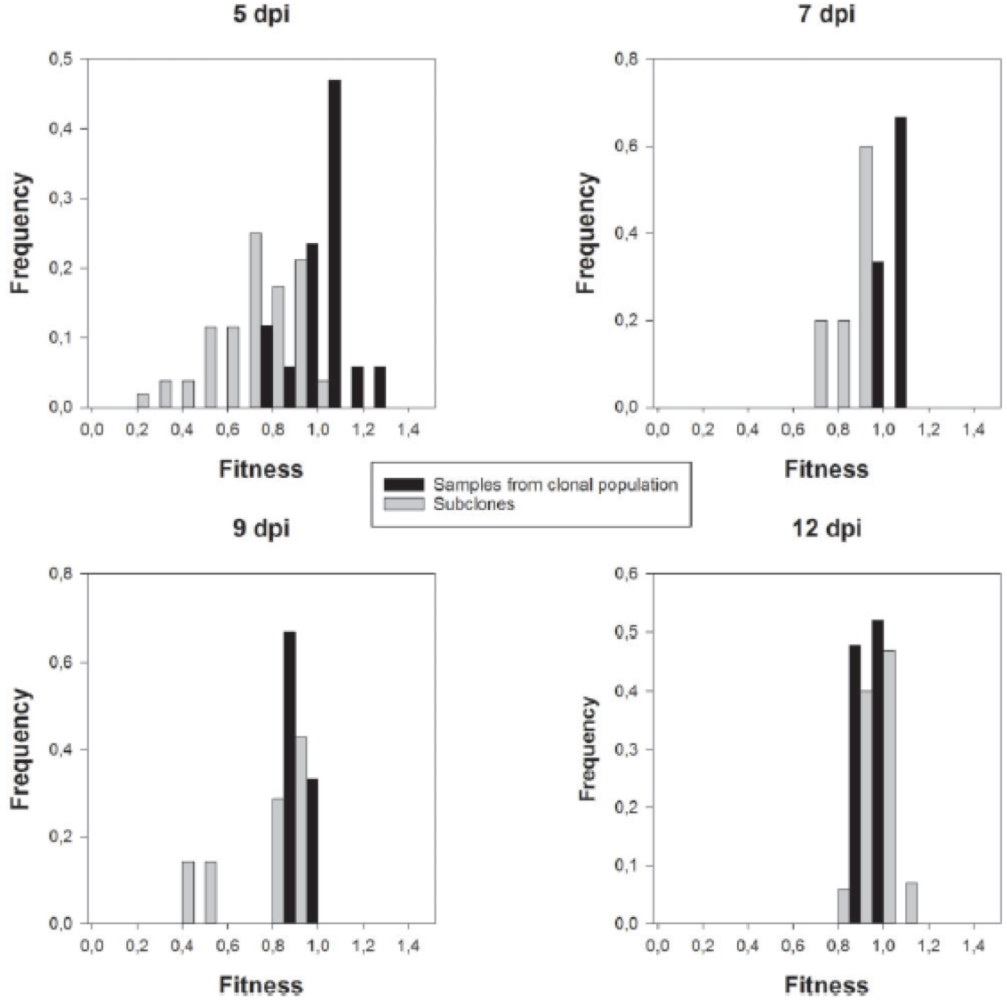


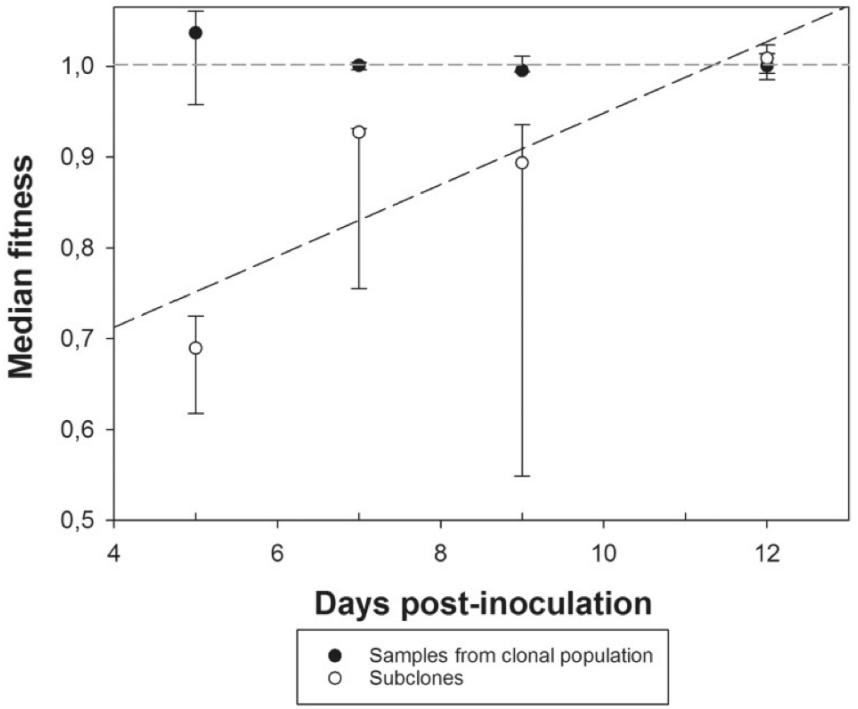


Mann-Whitney $U = 136$, $P < 0.001$

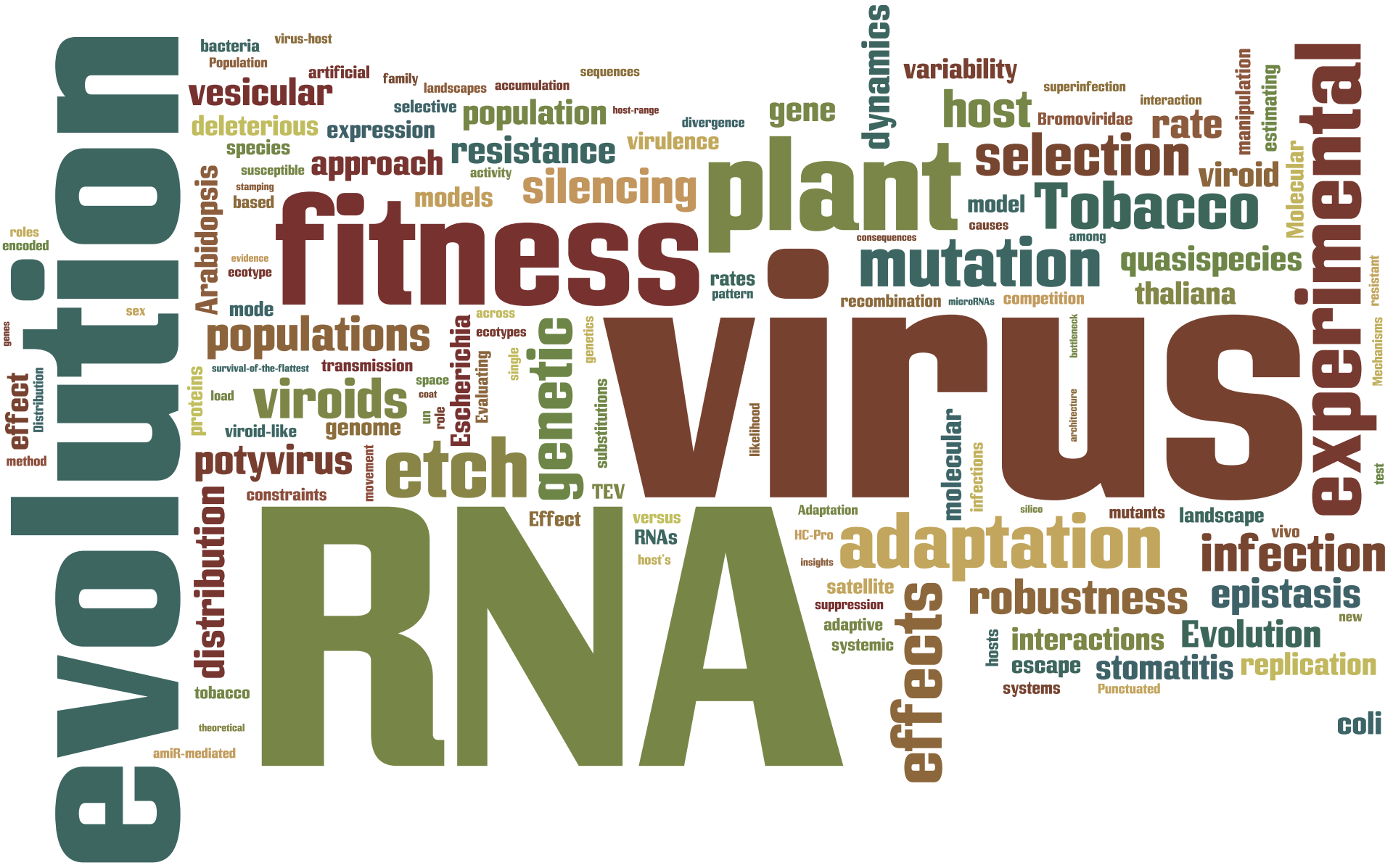
Duarte *et al.* (1994) *J. Virol.* **68**, 4295
Elena *et al.* (2003) *Biol. J. Linn. Soc.* **79**, 17

SANTA FE INSTITUTE





Cervera & Elena (2016) *Virus Evol.* **2**, vew006


evolution
RNA
virus
fitness
plant
experimental
adaptation
infection
mutation
Tobacco
host
selection
resistance
silencing
populations
viroids
genetic
etch
potyvirus
distribution
effects
robustness
epistasis
Evolution
interactions
stomatitis
replication
quasispecies
thaliana
Arabidopsis
vesicular
deleterious
expression
population
approach
models
based
dynamics
variability
rate
viroid
Molecular
Escherichia
genome
transmission
molecular
RNAs
versus
Effect
TEV
constraints
movement
substitutions
likelihood
infections
satellite
suppression
adaptive
systemic
hosts
escape
systems
Punctuated
tobacco
theoretical
amiR-mediated
coli
bacteria
virus-host
Population
artificial
family
landscapes
accumulation
sequences
selective
host-range
divergence
gene
virulence
species
susceptible
stamping
activity
evidence
ecotype
rates
pattern
recombination
microRNAs
competition
consequences
among
model
causes
superinfection
interaction
Bromoviridae
manipulation
estimating
roles
encoded
sex
mode
across
ecotypes
survival-of-the-flattest
space
coat
un
role
Evaluating
single
genetics
bottleneck
architecture
proteins
load
viroid-like
effect
Distribution
method
genes
Mechanisms
resistant
test
Adaptation
HC-Pro
insights
silico
mutants
landscape
vivo
new
coli