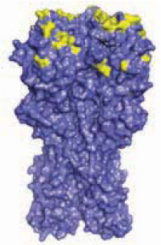
The role of deleterious mutations in influenza's adaptive (antigenic) evolution

Katia Koelle, David Rasmussen

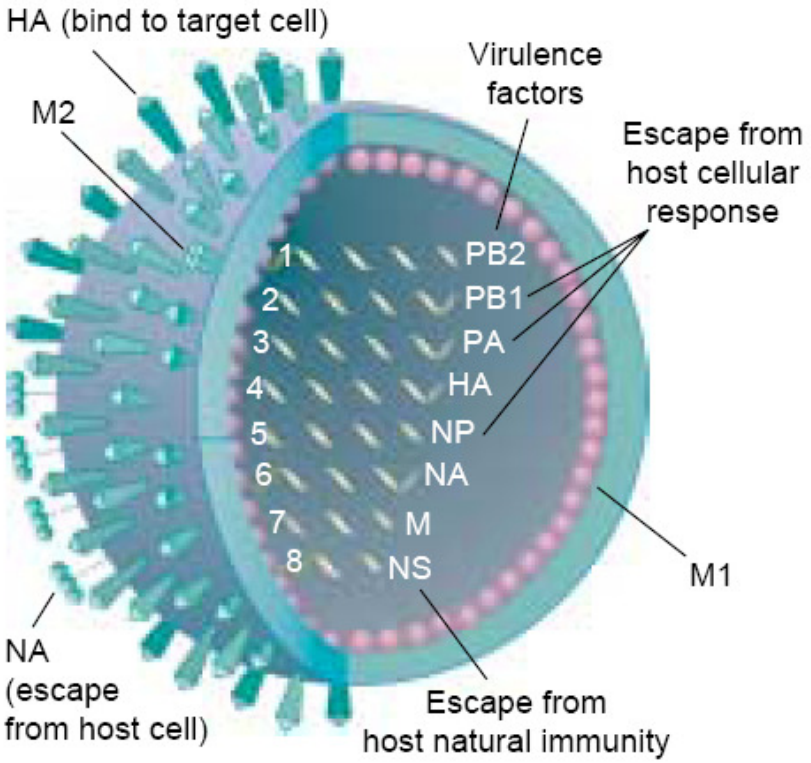
Department of Biology
Duke University

August 19, 2014

The virus: Influenza A/H3N2



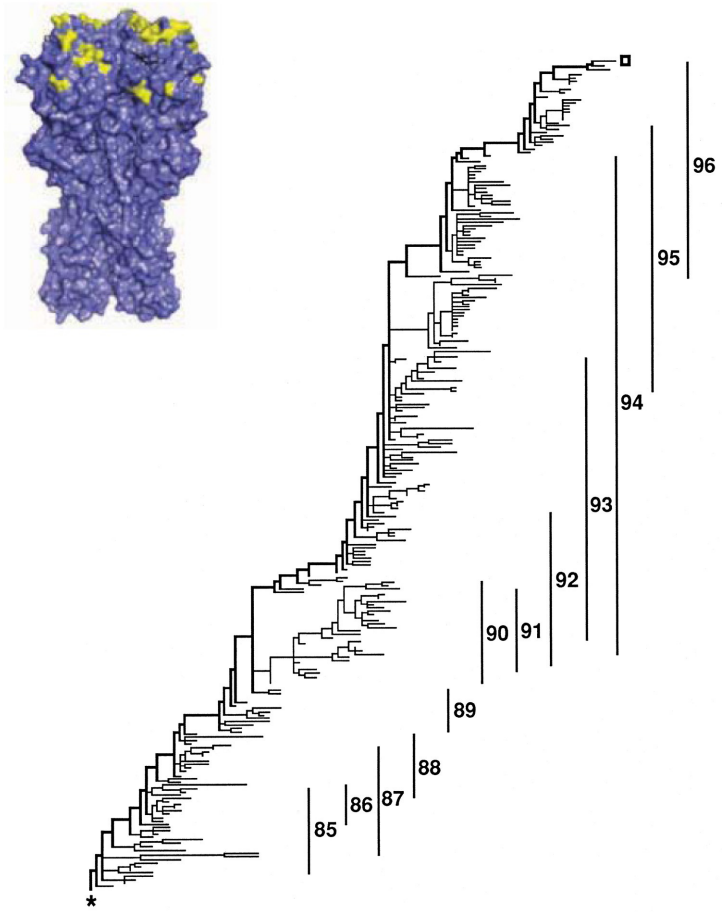
- Seasonal flu in humans caused by: **influenza A/H3N2**, A/H1N1, B



Influenza A/H3N2:

- Present in humans since 1968
- Segmented RNA virus (8 segments)
- Hemagglutinin HA = 'H' of H3N2
- Importance of HA for antigenic evolution

‘Spindly’/‘ladderlike’ HA phylogeny

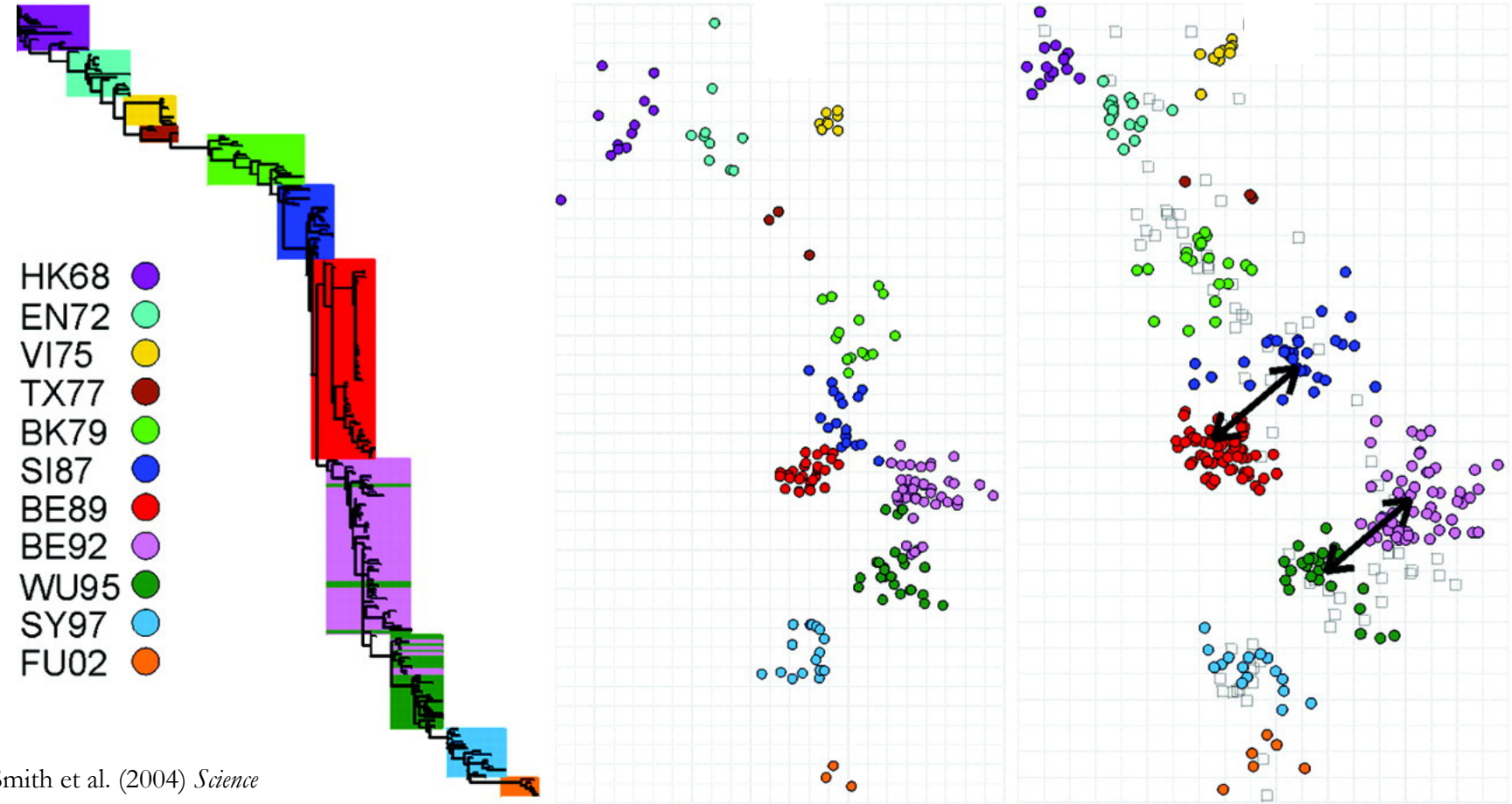


Fitch et al. (1997) *PNAS*

- Rapid turnover of viral population
- Low genetic diversity at any point in time

Evolutionary patterns of H3N2

Antigenic 'clusters' **Genetic map** **Antigenic map**



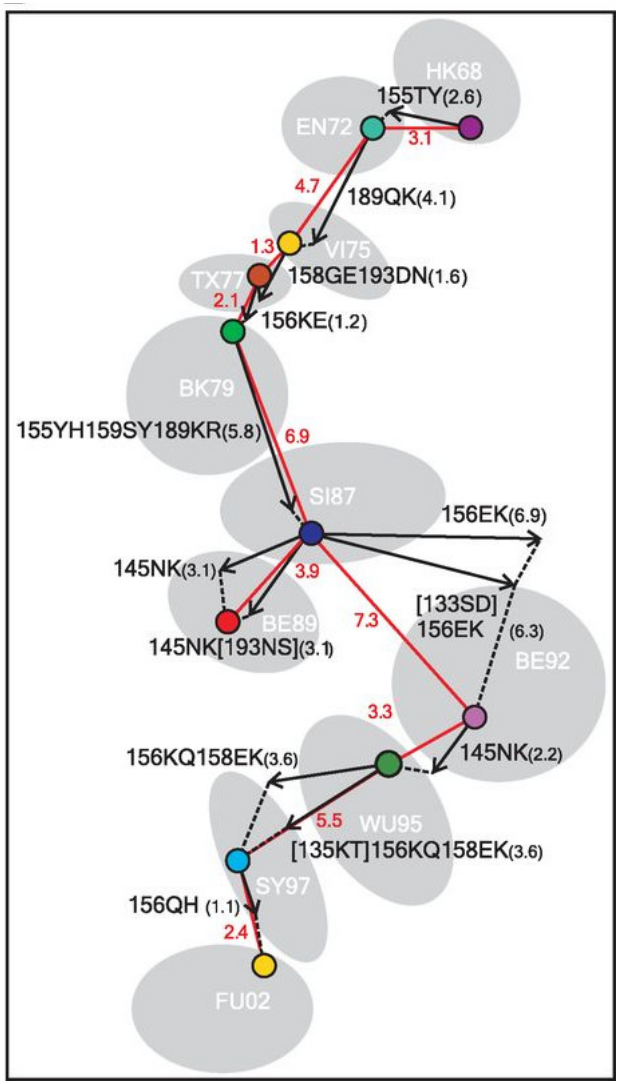
Smith et al. (2004) *Science*

- Gradual genetic change
- Punctuated antigenic change

- Multiple antigenic variants co-circulate

Strelkowa and Lässig (2012) *Genetics*

Cluster transitions precipitated by very few amino acid changes

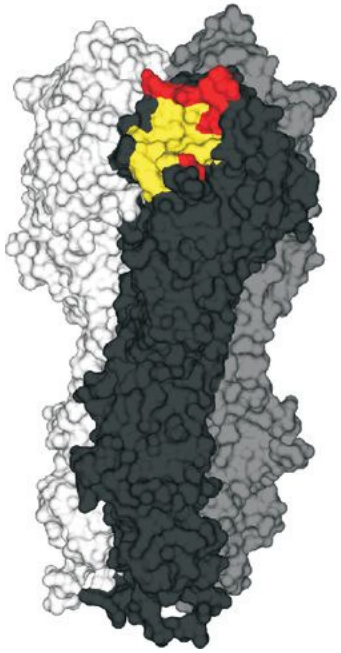


Koel et al. (2013) *Science*

Cluster transitions caused by:

- a single amino acid change (7 out of 10 instances)
- two amino acid changes (2 out of 10 instances)
- three amino acid changes (1 out of 10 instances)

In the latter cases, majority of antigenic change comes from single amino acid change



Non-temperate (Asian) source of global genetic diversity

- Source of antigenic clusters and smaller variants: China and SE Asia

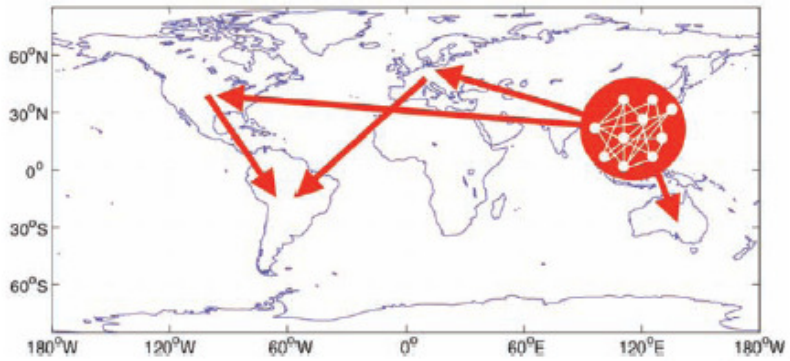
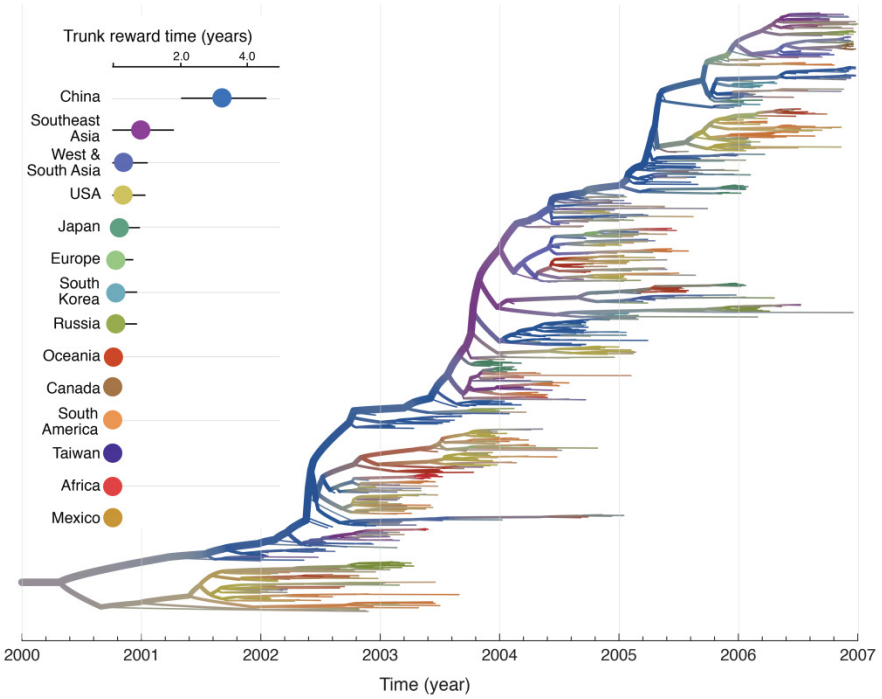


Fig. 5. Schematic of the dominant seeding hierarchy of seasonal influenza A (H3N2) viruses. The structure of the network within E-SE Asia is unknown.

Russell et al. (2008) *Science*



Lemey et al. (2014) *PLoS Pathogens*

- See also:

  Rambaut et al. (2008) *Nature*
  Bedford et al. (2010) *PLoS Pathogens*
  Bahl et al. (2011) *PNAS*

Models/hypotheses for influenza's antigenic evolution

Limited number of antigenic strains
Recker et al. (2007) *PNAS*

Antigenic drift as a side effect of binding avidity changes
Hensley et al. (2009) *Science*

Neutral networks leading to epochal evolution
Koelle et al. (2006) *Science*

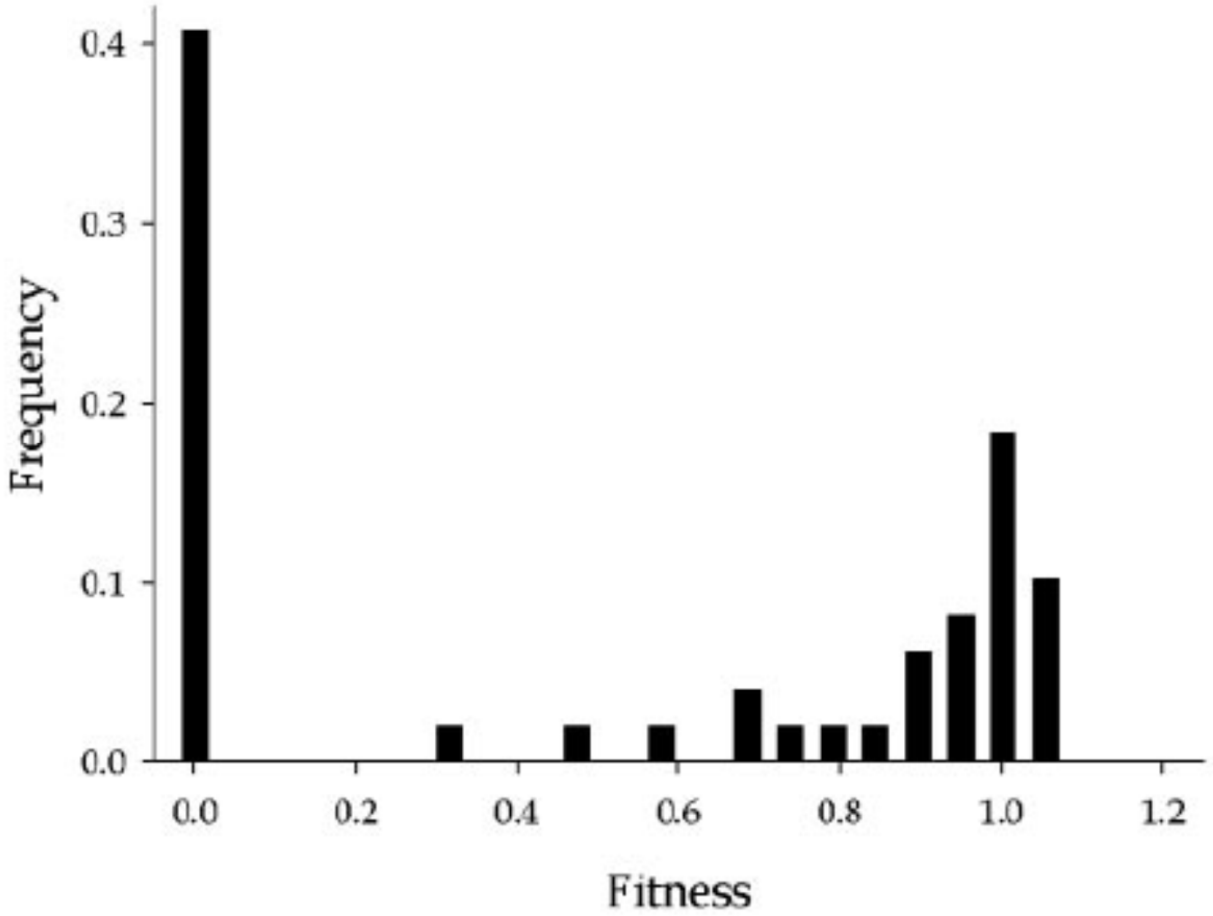
Clonal interference
Strelkowa and Lässig (2012) *Genetics*

Canalization
Bedford et al. (2012) *BMC Biology*

Generalized cross immunity
Ferguson et al.(2003) *Nature*

Deleterious mutations: they occur

Site-directed mutagenesis study of vesicular stomatitis virus



Sanjuán et al. (2004) *PNAS*

Deleterious mutations: they transiently circulate

Phylogenetic Evidence for Deleterious Mutation Load in RNA Viruses and Its Contribution to Viral Evolution

Oliver G. Pybus, Andrew Rambaut,* Robert Belshaw,* Robert P. Freckleton,* Alexei J. Drummond,*† and Edward C. Holmes‡§*

*Department of Zoology, University of Oxford, Oxford, United Kingdom; †Department of Computer Science, University of Auckland, Auckland, New Zealand; ‡Center for Infectious Disease Dynamics, Department of Biology, The Pennsylvania State University; and §Fogarty International Center, National Institutes of Health, Bethesda, Maryland

Populations of RNA viruses are often characterized by abundant genetic variation. However, the relative fitness of these mutations is largely unknown, although this information is central to our understanding of viral emergence, immune evasion, and drug resistance. Here we develop a phylogenetic method, based on the distribution of nonsynonymous and synonymous changes, to assess the relative fitness of polymorphisms in the structural genes of 143 RNA viruses. This reveals that a substantial proportion of the amino acid variation observed in natural populations of RNA viruses comprises transient deleterious mutations that are later purged by purifying selection, potentially limiting virus adaptability. We also demonstrate, for the first time, the existence of a relationship between amino acid variability and the phylogenetic distribution of polymorphisms. From this relationship, we propose an empirical threshold for the maximum viable deleterious mutation load in RNA viruses.

Pybus et al. (2007) *MBE*

Strong genetic linkage in influenza, and importance of accounting for deleterious mutations in successfully predicting short-term HA evolution

Strelkowa and Lässig (2012) *Genetics*

Łuksza and Lässig (2014) *Nature*

Simple model: Viral population subject to only deleterious mutations

λ = per-genome per-transmission deleterious mutation rate

s_d = fitness effect of deleterious mutations

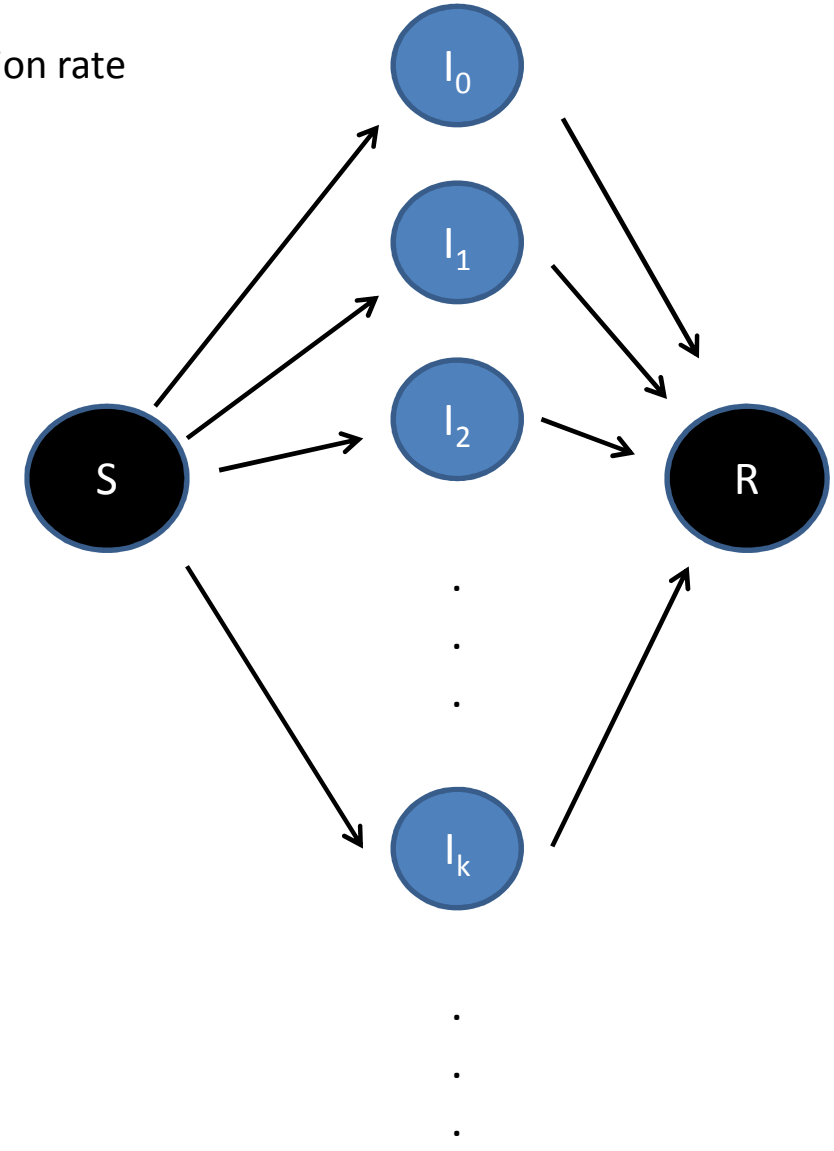
$$\beta_i = \beta_0(1 - s_d)^i$$

ρ_L = probability that a mutation is lethal

$$\frac{dS}{dt} = \mu N - \mu S - \beta_0 e^{-\lambda p_L} \frac{S}{N} \sum_{k=0}^{\infty} (1 - s_d)^k I_k$$

$$\frac{dI_0}{dt} = \beta_0 e^{-\lambda} \frac{S}{N} I_0 - (\mu + \gamma) I_0$$

$$\frac{dI_k}{dt} = \beta_0 I_{tot} e^{-\lambda} \frac{S}{N} \sum_{j=0}^{k} \left((1 - s_d)^{k-j} \frac{(\lambda(1 - p_L))^j}{j!} p_{k-j} \right) - (\mu + \gamma) I_k$$



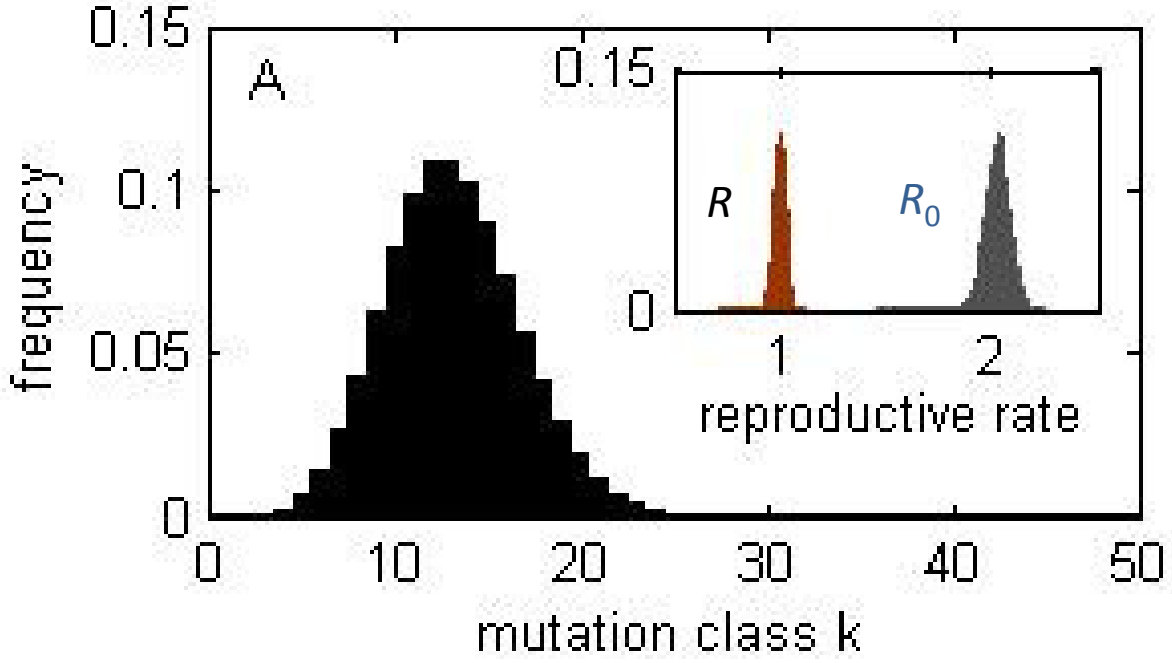
Mutation-selection balance

$$p_k = e^{-\theta}\frac{\theta^k}{k!}$$

$$\theta = \lambda(1 - p_L)/s_d$$

In comparison, Haigh (1978) with constant population size with discrete generations; no lethal mutations considered:

$$\theta = \lambda/s_d$$



Parameterization

$$p_k = e^{-\theta}\frac{\theta^k}{k!}$$

$$\theta = \lambda(1 - p_L)/s_d$$

λ = genome size × # substitutions that occur per nucleotide per cell infection

= 12741 nucleotides × 2.3×10^{-5} substitutions per nucleotide (for influenza A)

= 0.293 per genome per transmission

Roughly a third of mutations are synonymous;

Roughly 25% of non-syn. mutations are phenotypically neutral

= **0.15**

Sanjuán et al. (2010) J. Virology

ρ_L = 0.20-0.41 for five RNA viruses studied.

Assume ρ_L = 0.3.

Sanjuan (2010) *Phil. Trans.*

average s_d for five RNA viruses studied: 0.103, 0.107, 0.112, 0.126, 0.132 (remarkably narrow range).

Sanjuan (2010) *Phil. Trans.*

But these are *in vitro* fitness costs associated with viral growth in cells.

Translating *in vivo* into transmission fitness costs

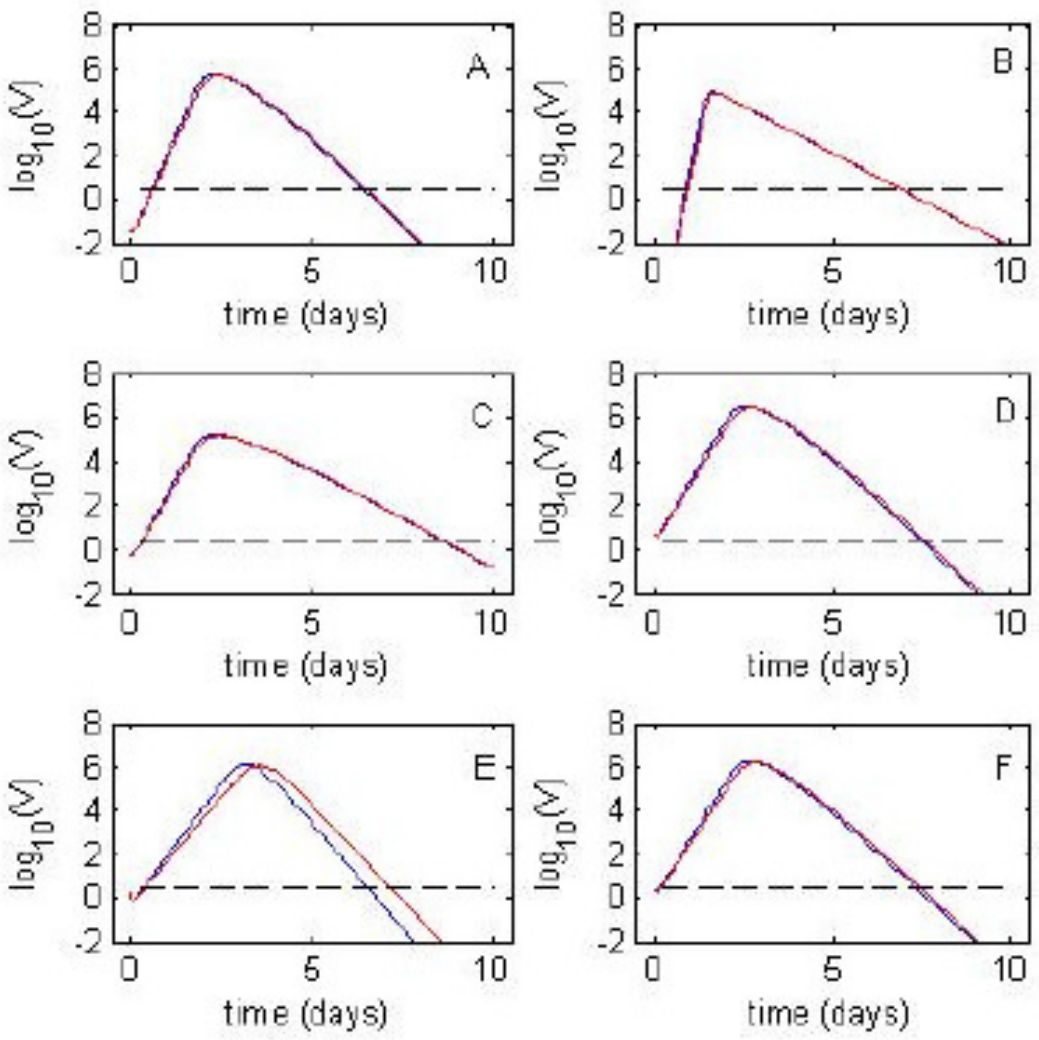
$$\frac{dT_w}{d\tau} = -\beta_w T_w V_w$$

$$\frac{dE_w}{d\tau} = \beta_w T_w V_w - k_w E_w$$

$$\frac{dI_w}{d\tau} = k_w E_w - \delta_w I_w$$

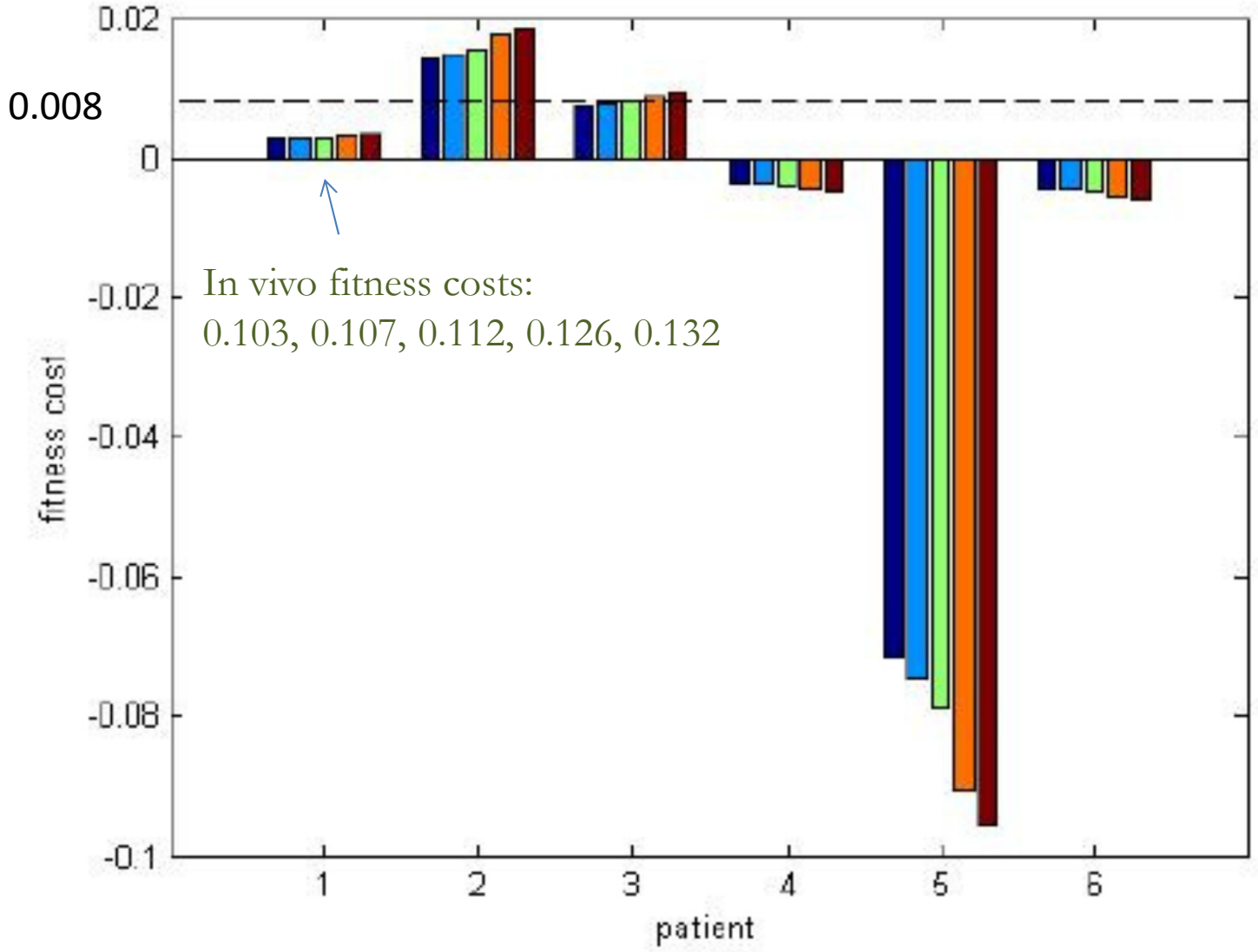
$$\frac{dV_w}{d\tau} = p_w I_w - c_w V_w$$

Baccam et al. (2006)

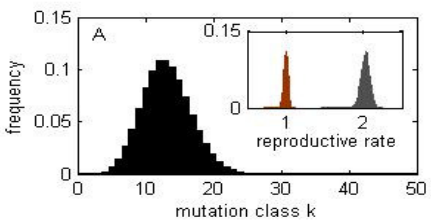


wild-type, mutant (fitness cost in p_w = 0.112)

Transmission fitness costs



Fates of antigenic mutants



Antigenic mutations occur in a certain genetic background with i deleterious mutations

Antigenic mutations are of a certain size σ, which quantifies the degree of immune escape

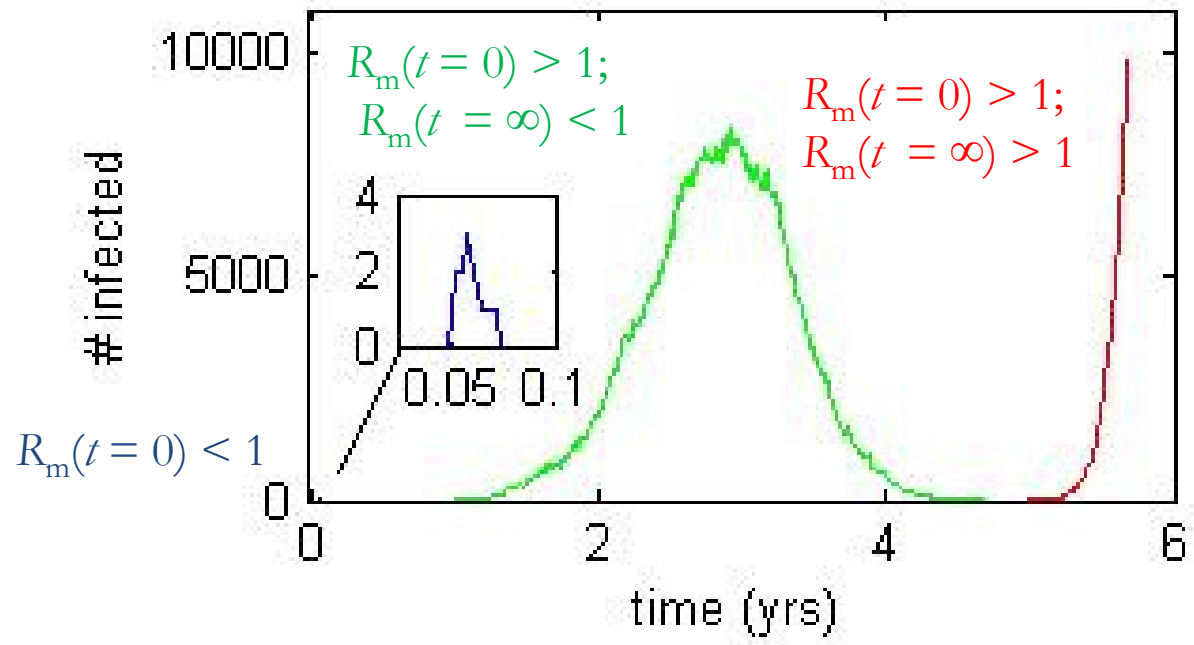
Calculate reproductive rate R of invading mutant strain: $R_m(t = 0)$ and $R_m(t = \infty)$

Calculation of $R_m(t = \infty)$ assumes no change in susceptible population, but allows for mutant strain to reach its mutation-selection balance

$R_m(t = 0)$ and $R_m(t = \infty)$ depend on i and σ

THREE POSSIBLE FATES

à la Peck (1994) *Genetics*

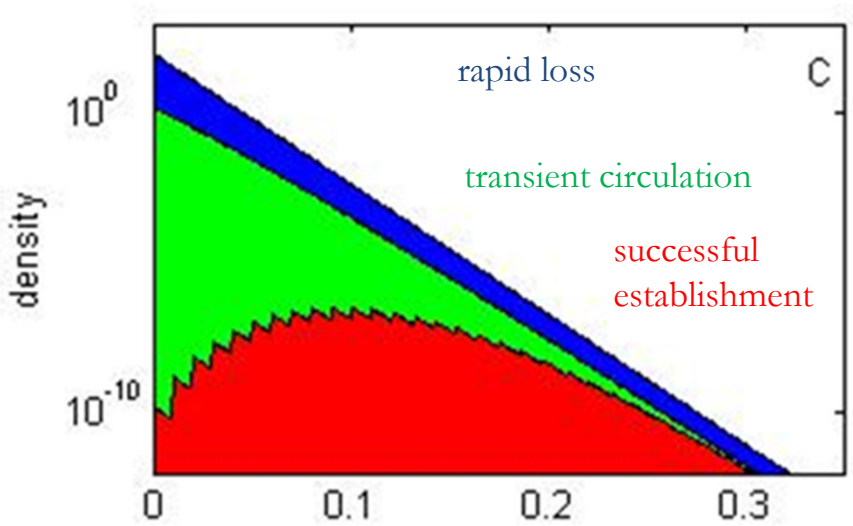


Probabilities of fates

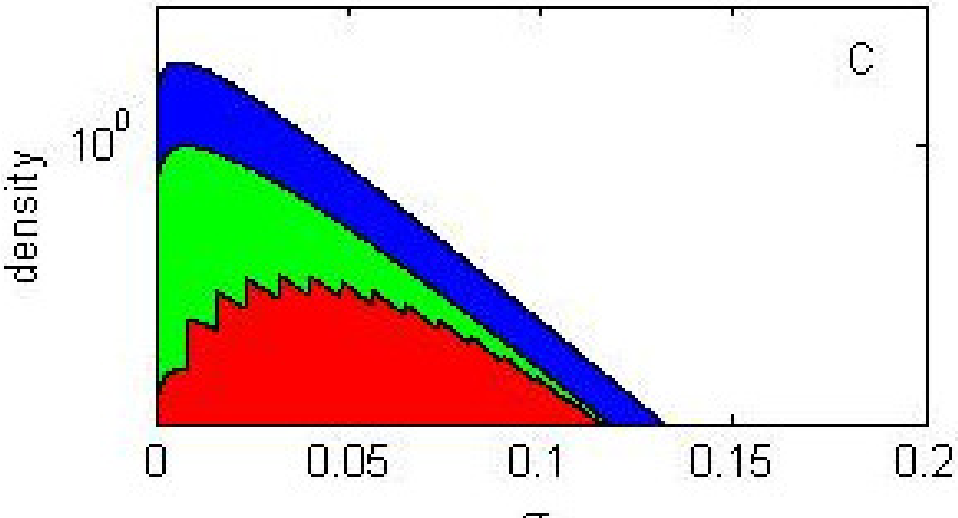
Which fate occurs depends on:

- the antigenic mutant's deleterious mutation load:
  lower i → higher $R_m(t = 0)$ and $R_m(t = \infty)$

- the antigenic mutant's degree of immune escape σ:
  higher σ → higher $R_m(t = 0)$ and $R_m(t = \infty)$

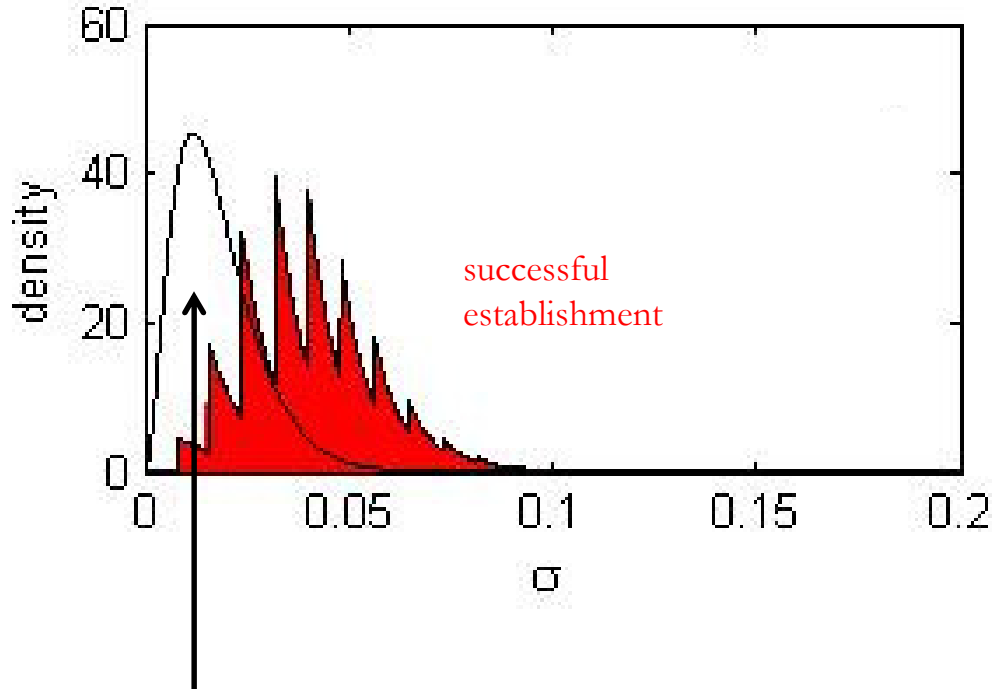


(assuming σ is exponentially distributed)

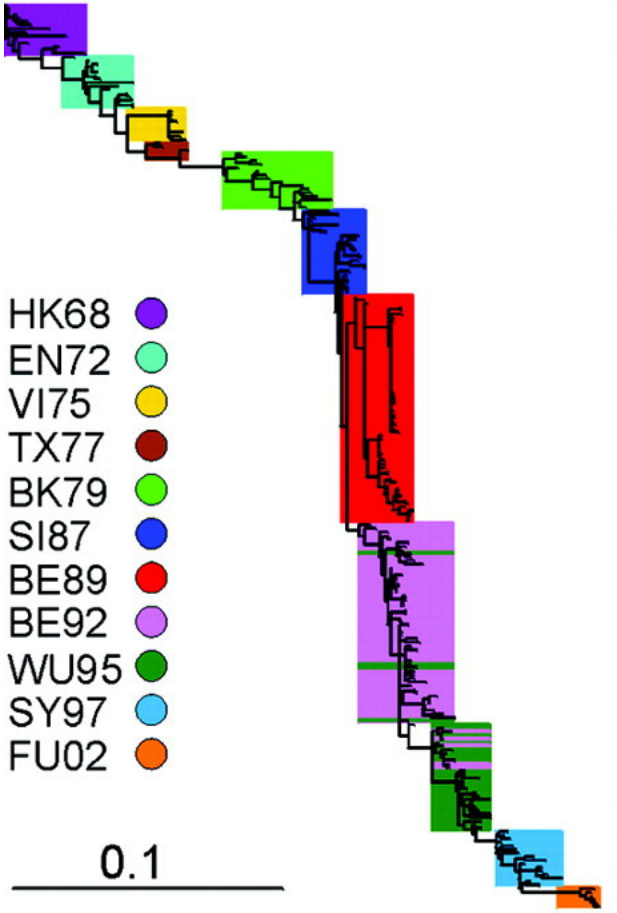


(assuming σ is gamma distributed)

Size distribution of antigenic mutants that successfully establish



no deleterious mutations scenario



Expectation of punctuated antigenic evolution

Smith et al. (2004) *Science*

Simple model w/explicit epidemiological dynamics

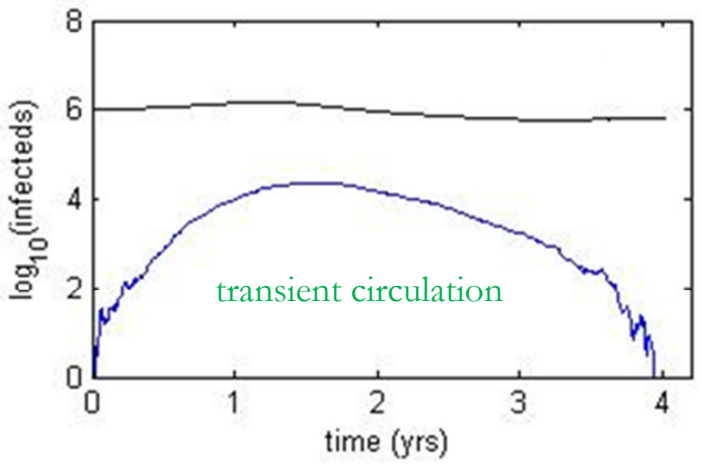
Model allows number of susceptible hosts to change dynamically

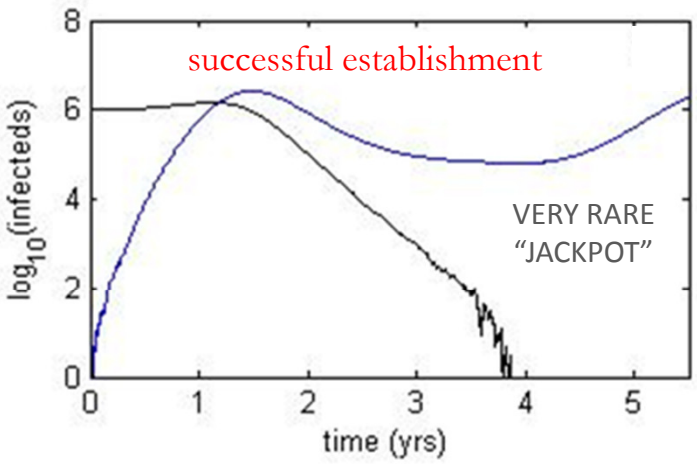
low immune escape σ

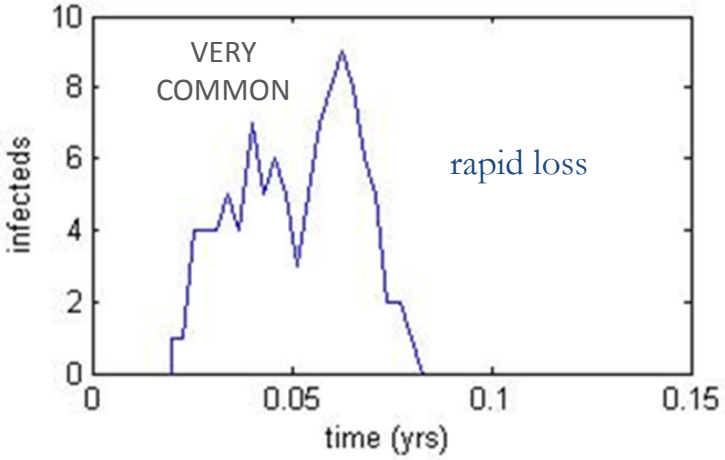
high immune escape σ

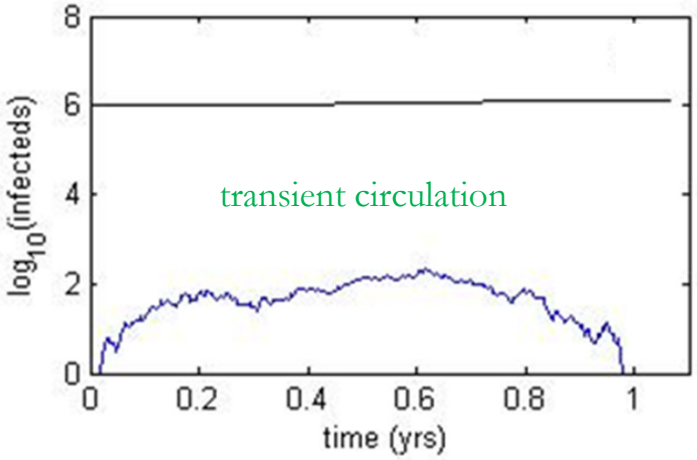
low load *i*

avg load *i*

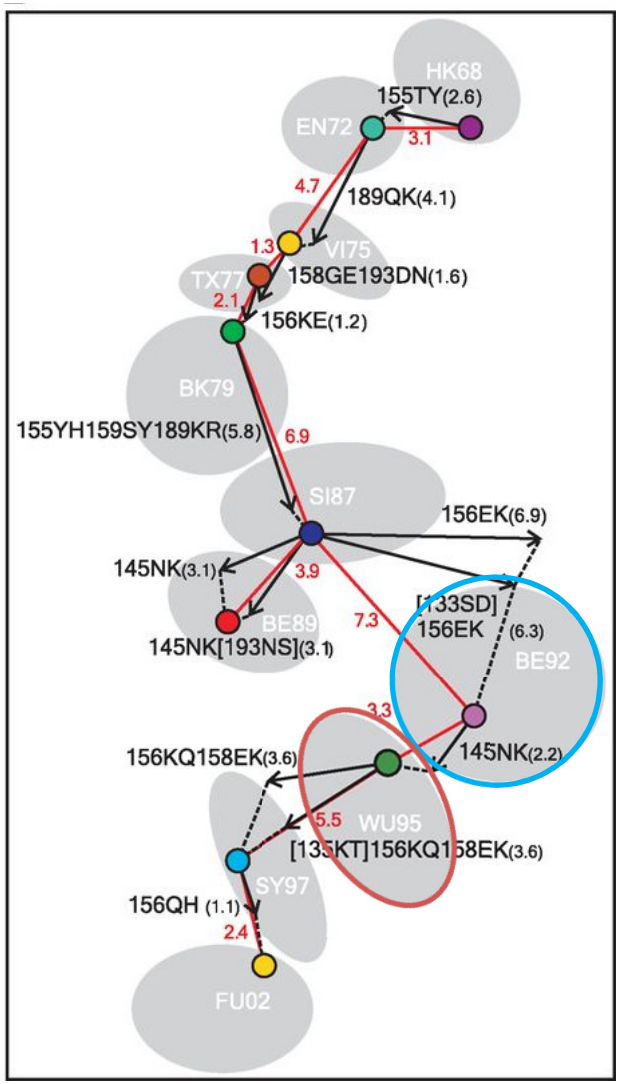




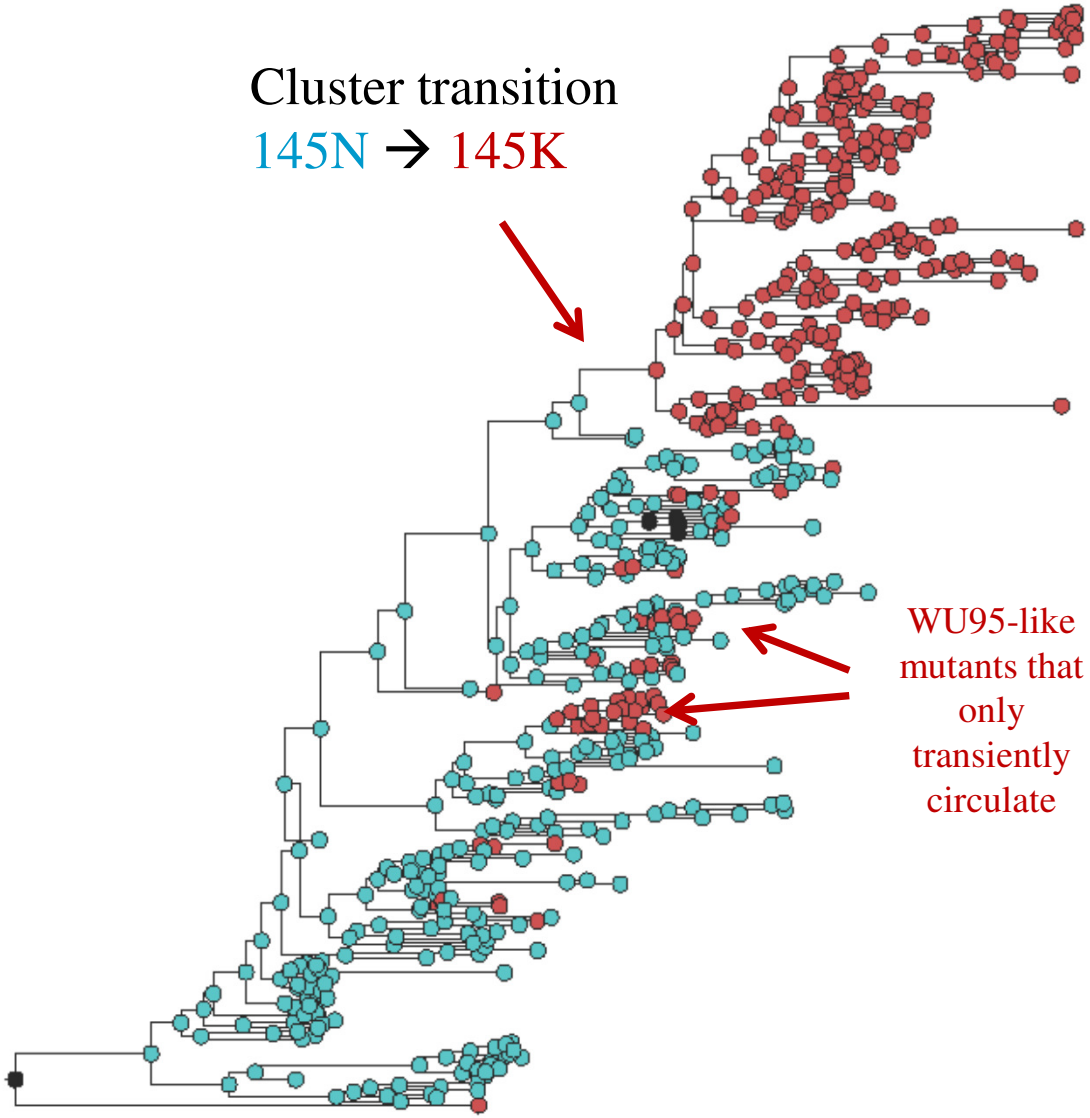




Empirical support for the 'jackpot' strategy



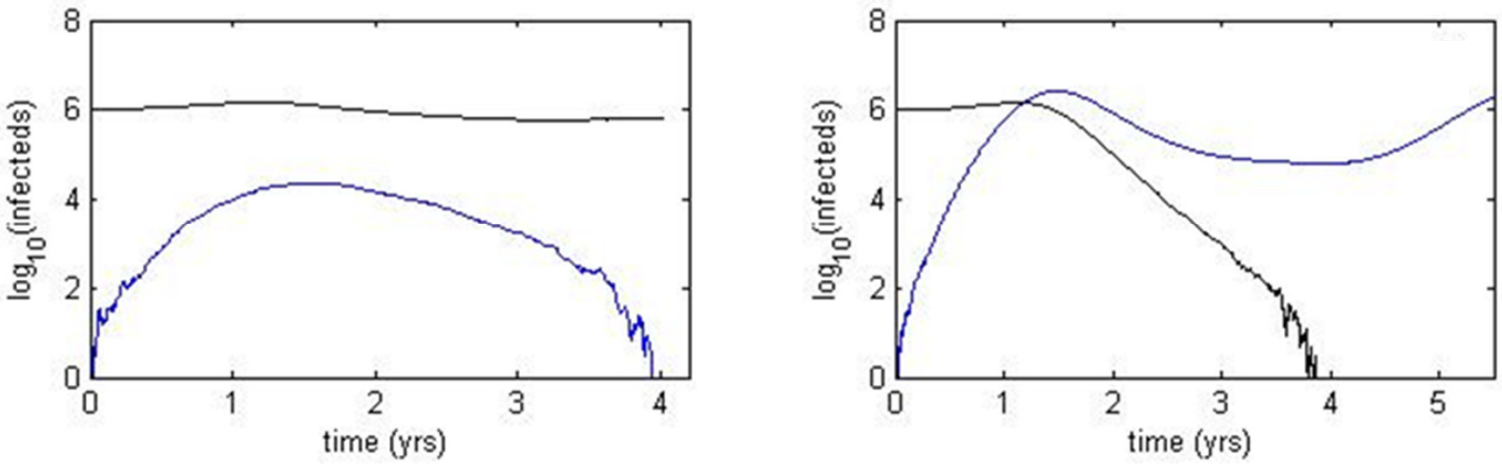
Koel et al. (2013) *Science*



Alternative strategies for hitting the jackpot

low σ high σ

low *i*

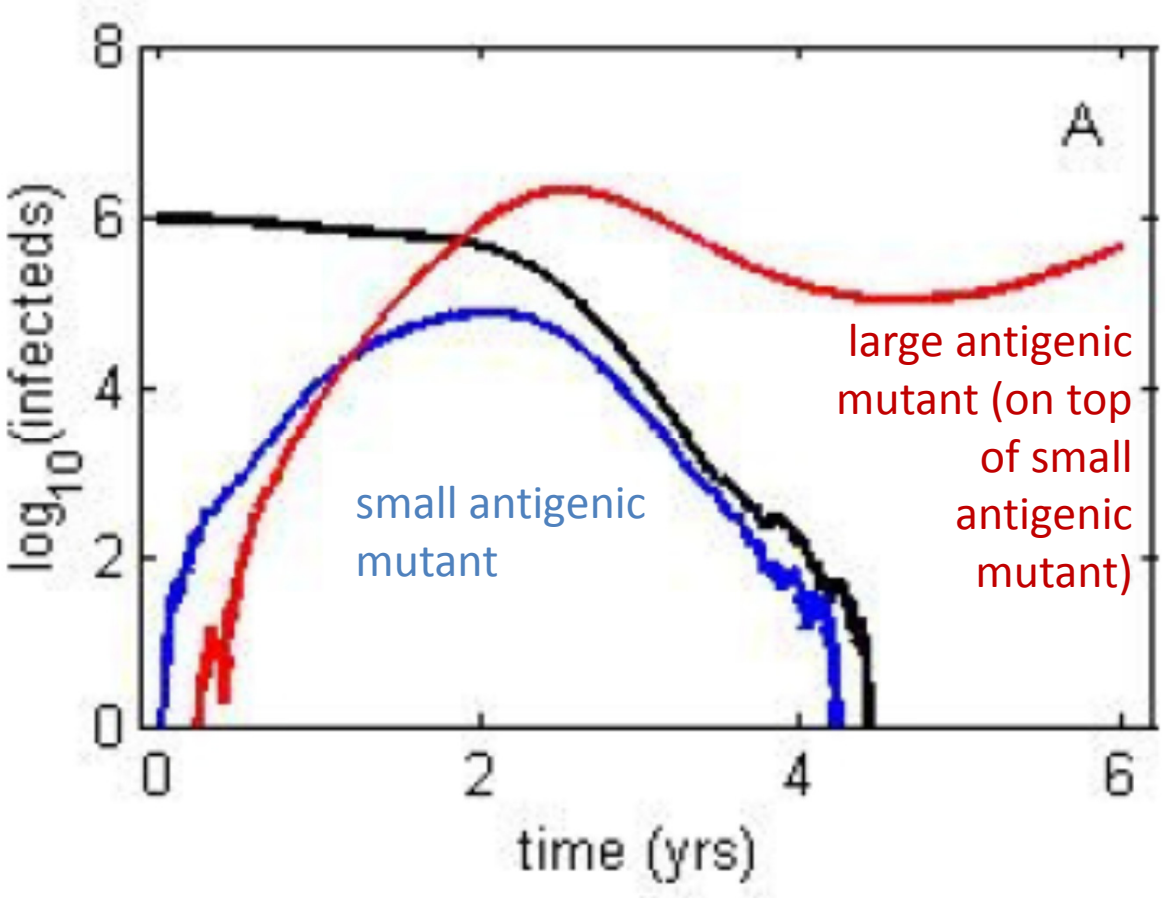


Step 1: More likely small- to moderate-sized mutation occurs first in 'good' background, transiently rises, thereby "inflating" low-*i* viral counts

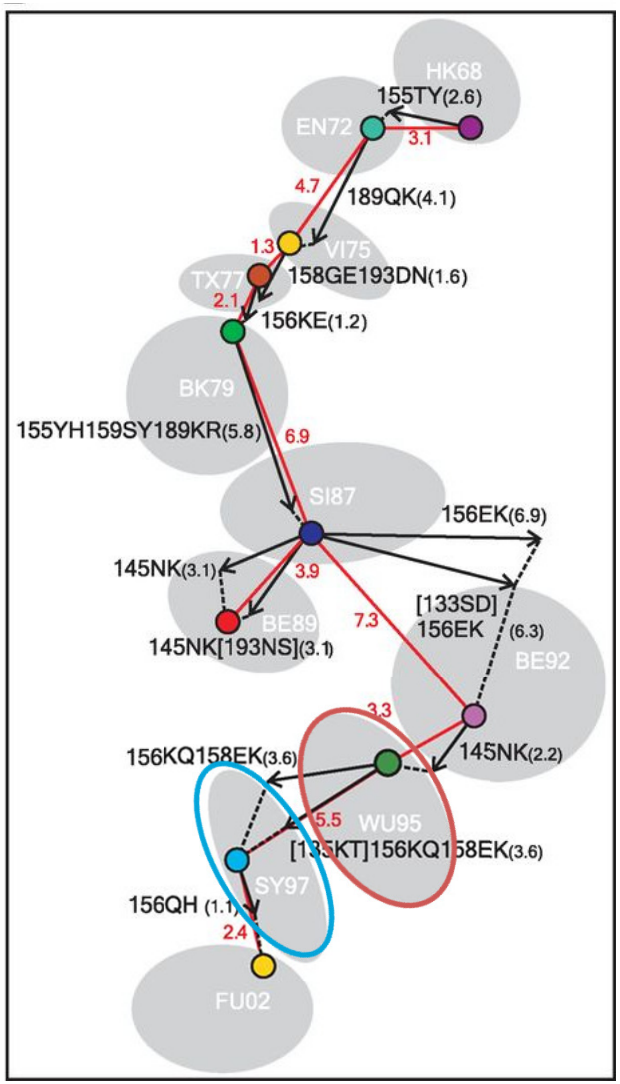
Step 2: Less likely large antigenic mutation occurs shortly thereafter, resulting in an antigenic cluster transition

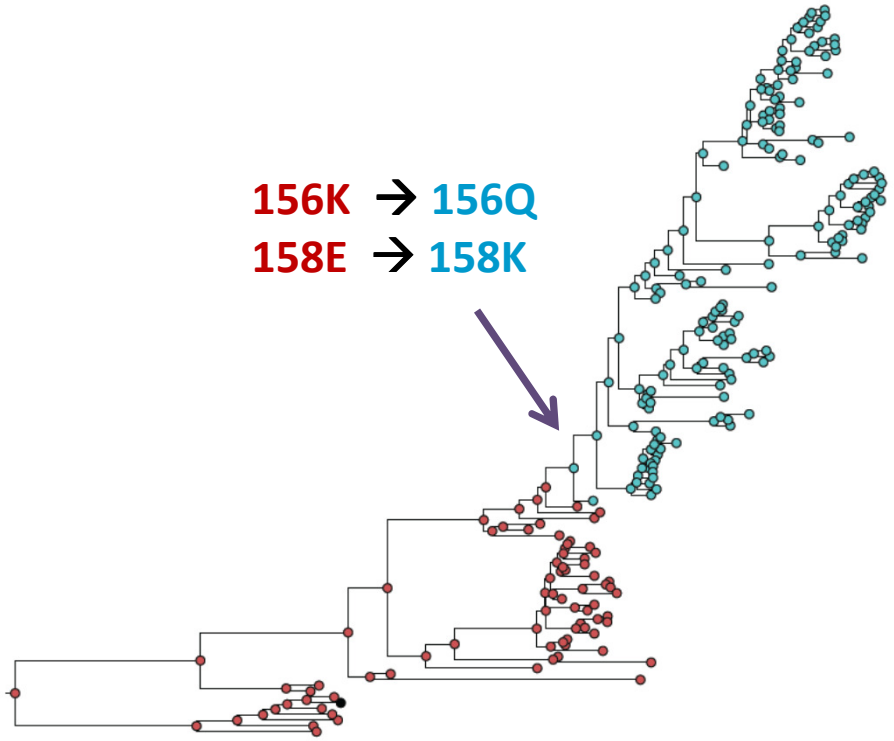
Will effectively appear as two simultaneous antigenic amino acid changes in a viral phylogeny

Alternative strategies for hitting the jackpot



Empirical support for this two-step antigenic change strategy





Koel et al. (2013) *Science*

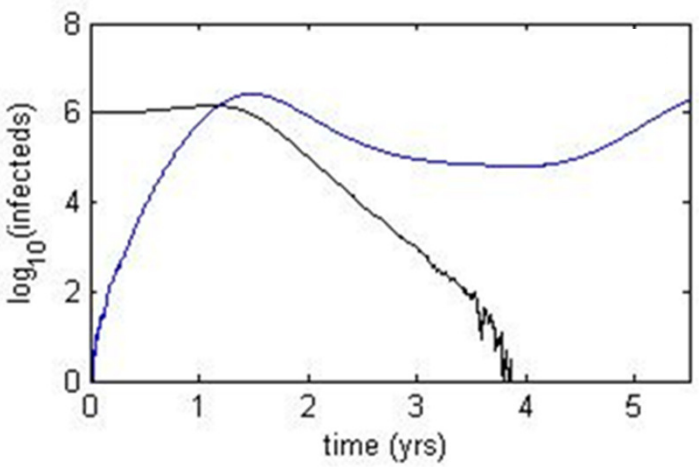
Alternative strategies for hitting the jackpot

Step 1: Large antigenic mutation occurs in average genetic background, and transiently rises.

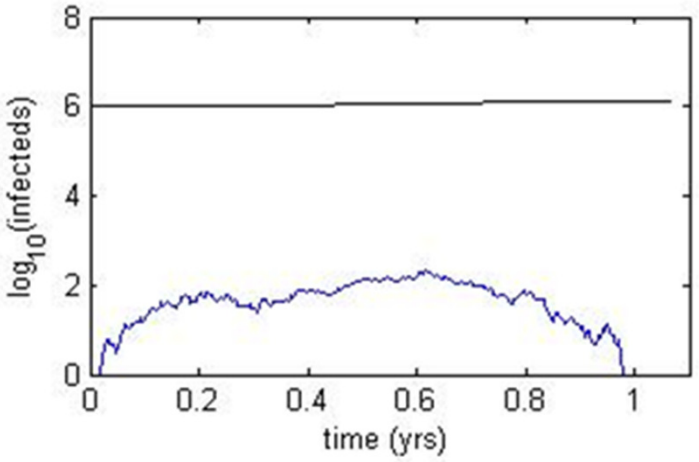
Step 2: Co-infection occurs (with resident strain, also likely with average genetic background). Reassortment leads to purging of deleterious mutations, and therewith an antigenic cluster transition.

high σ

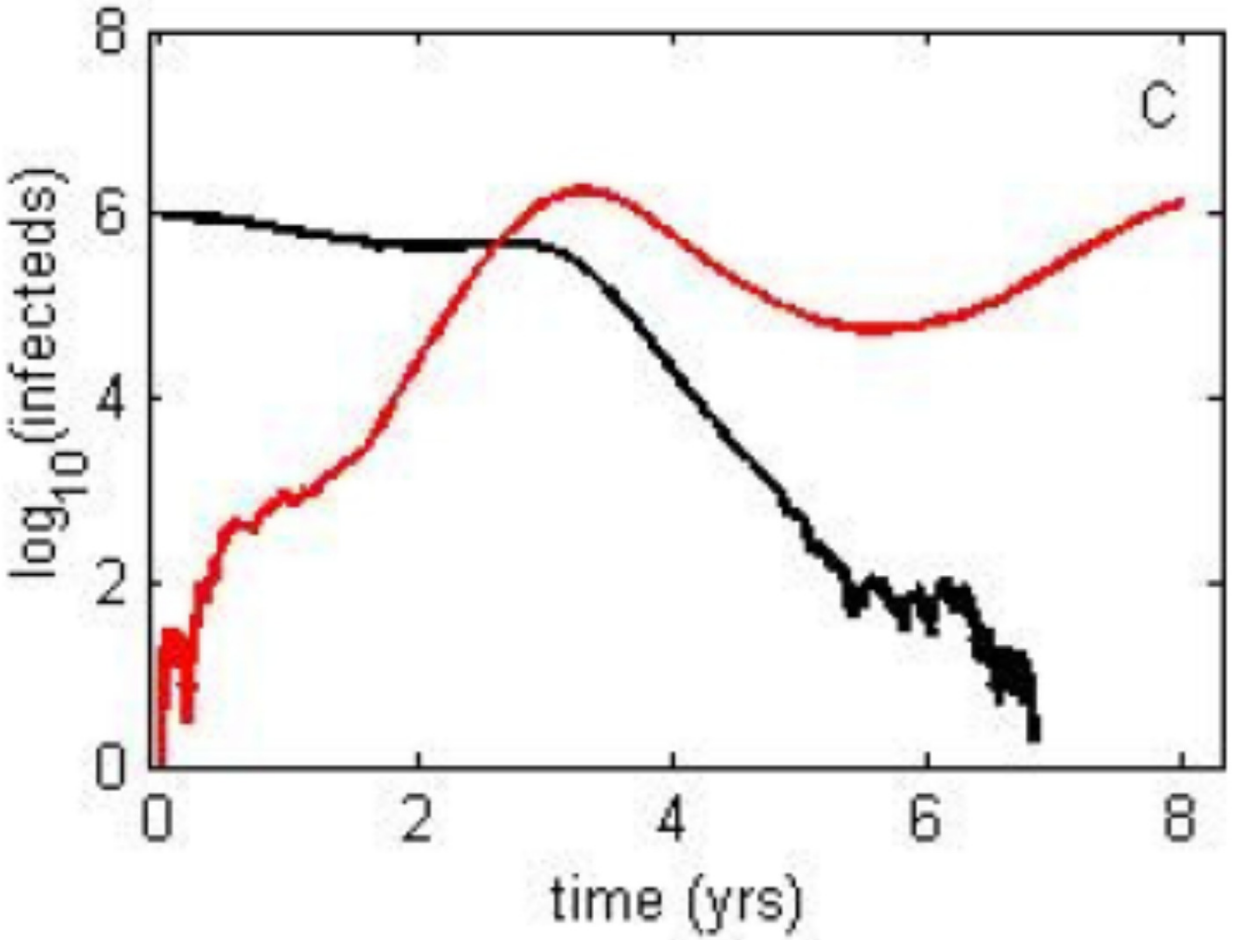
low i



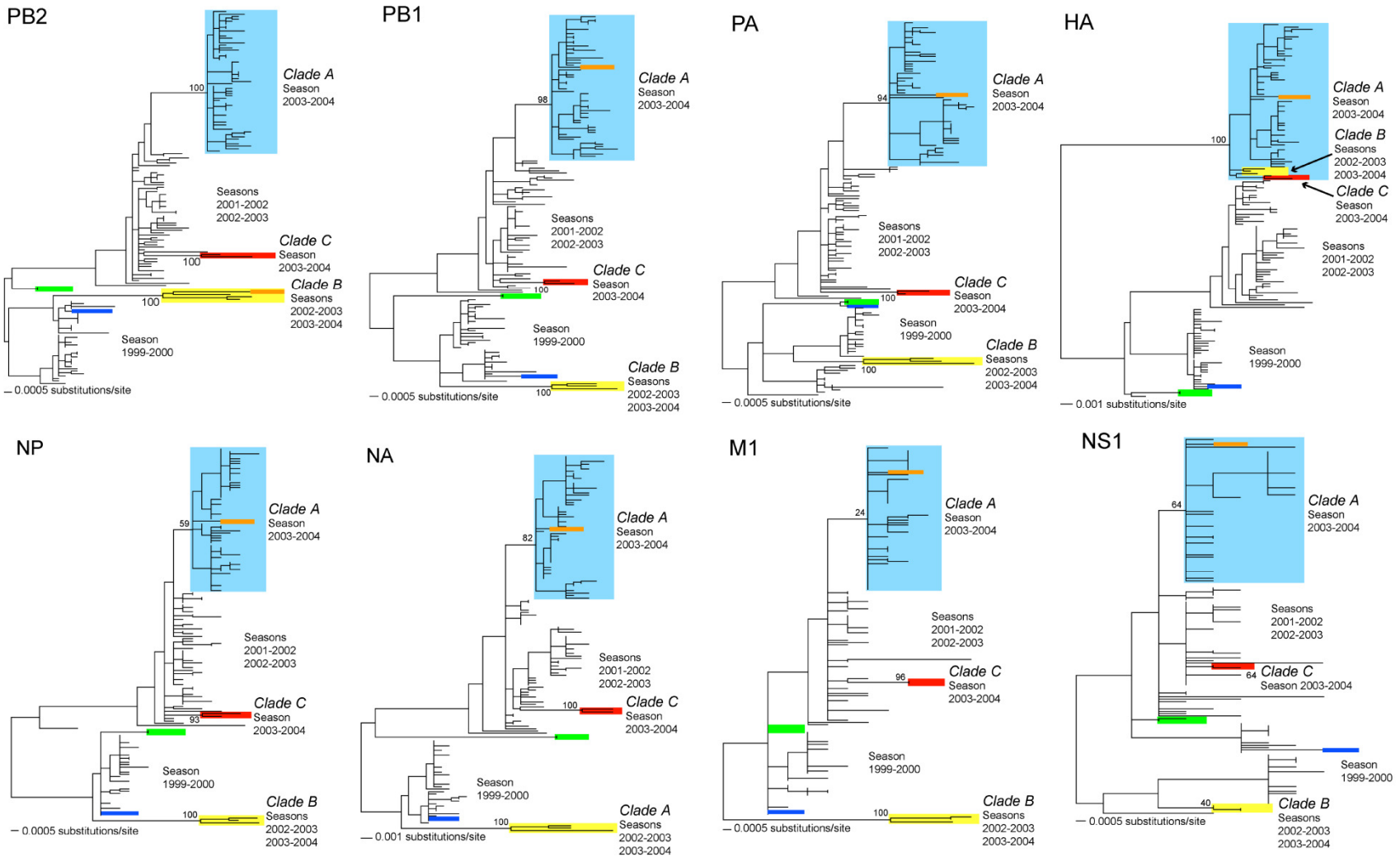
avg i



Alternative strategies for hitting the jackpot



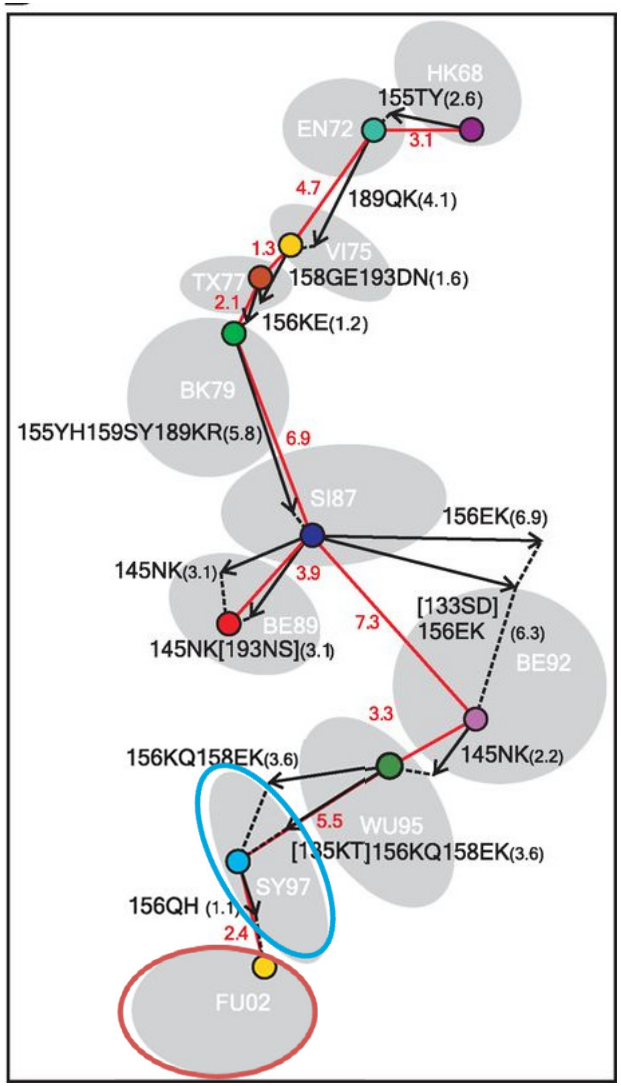
Evidence in support of this alternative strategy

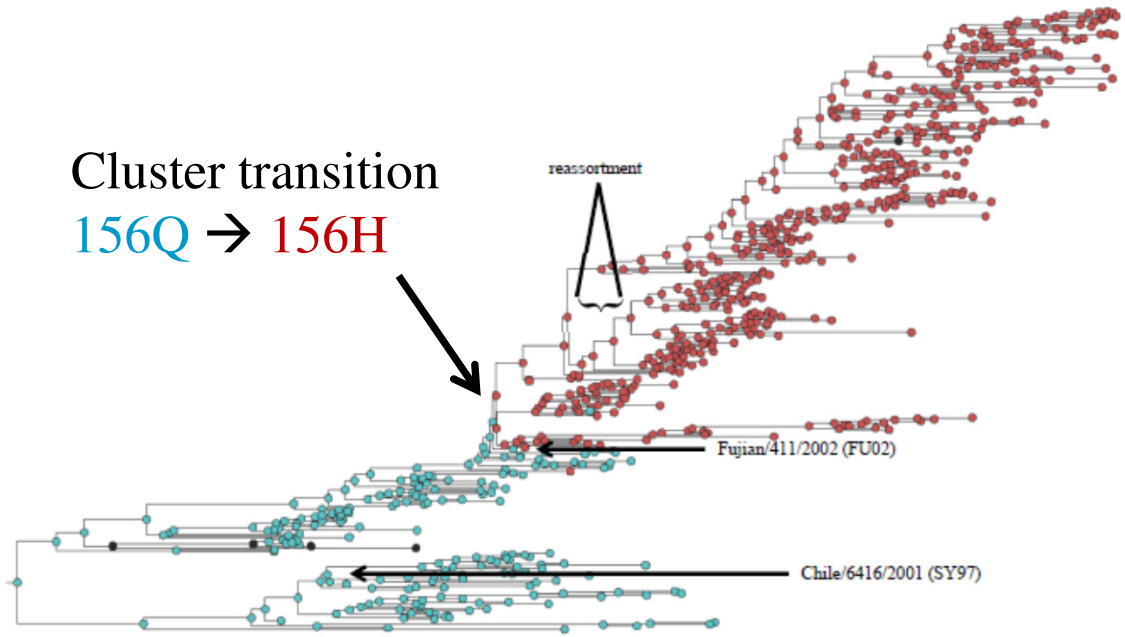


FU02: antigenic mutant was circulating for several years, reassortment occurred, then antigenic cluster transition

"Functional compatibility", "effective gene constellation" Holmes et al. (2005) *PLoS Biology*

Empirical support for this two-step reassortment strategy



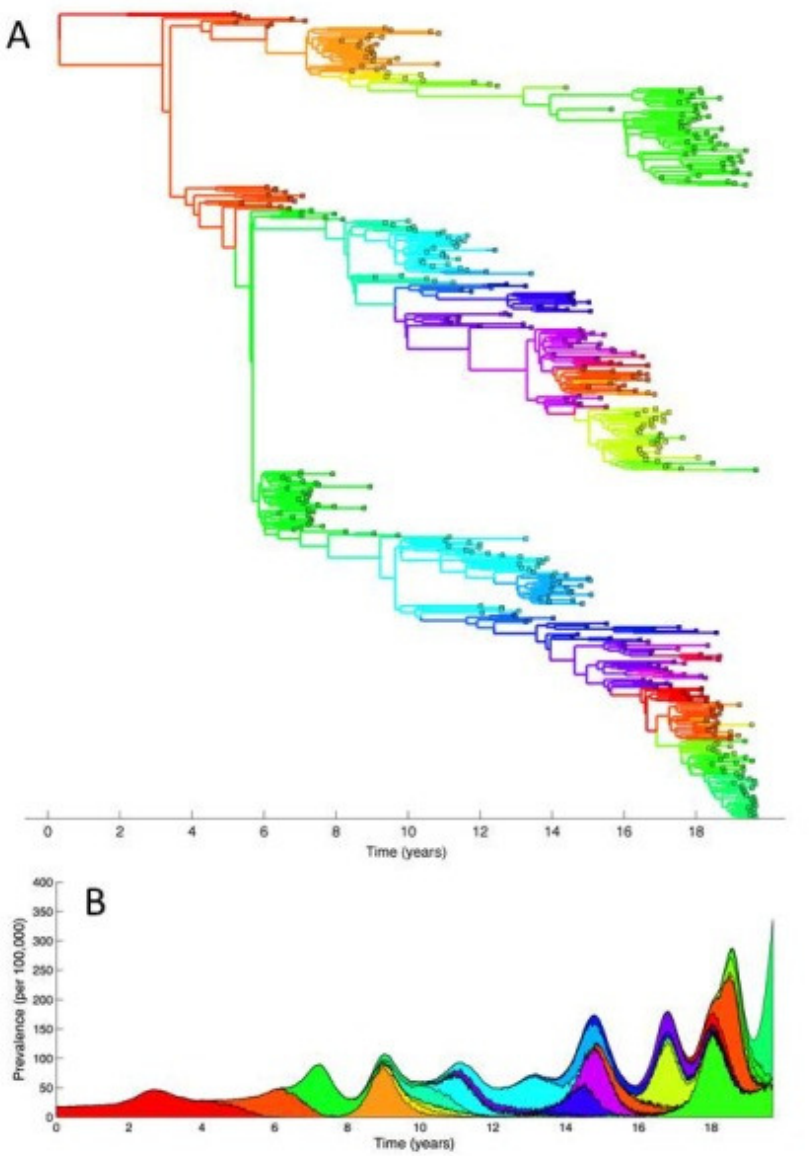


Koel et al. (2013) *Science*

Full ‘phylodynamic’ simulations

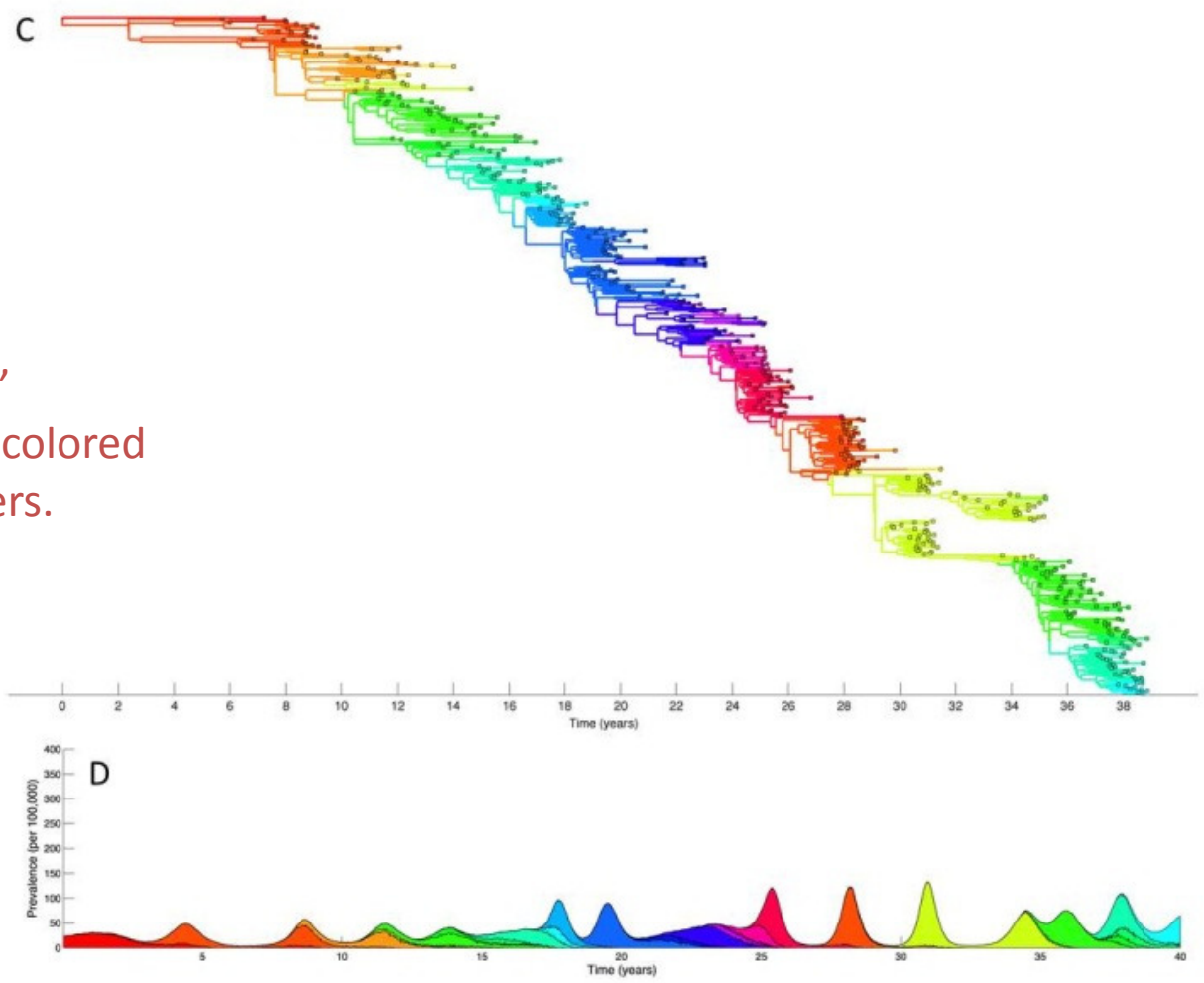
No load simulation
(10 yrs)

Explosive genetic &
antigenic diversity



Full ‘phylodynamic’ simulations

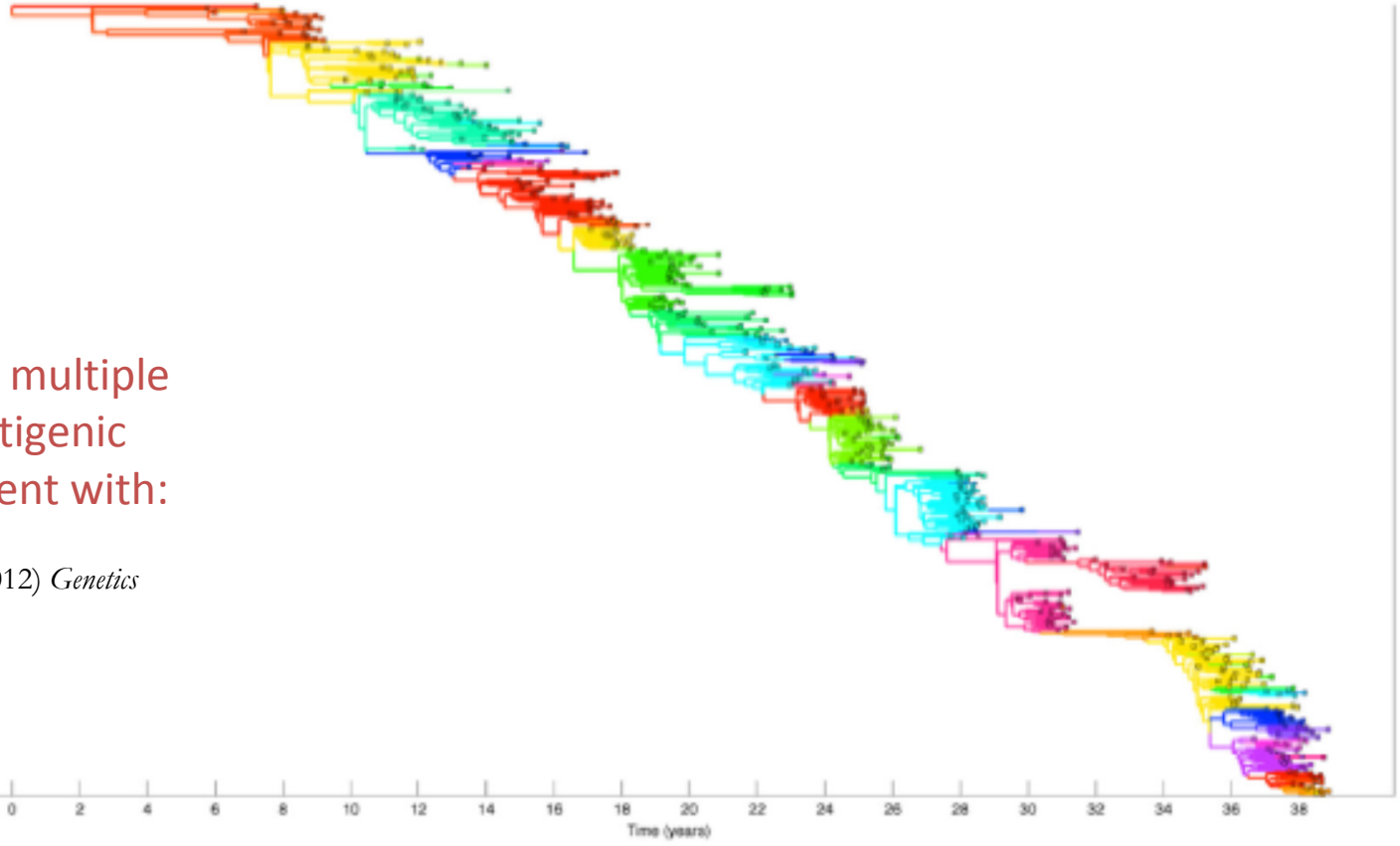
Load simulations
(40 yrs)



Reproduction of ‘spindly’
HA phylogeny. Lineages colored by major antigenic clusters.

Full ‘phylodynamic’ simulations

Load simulations
(40 yrs)



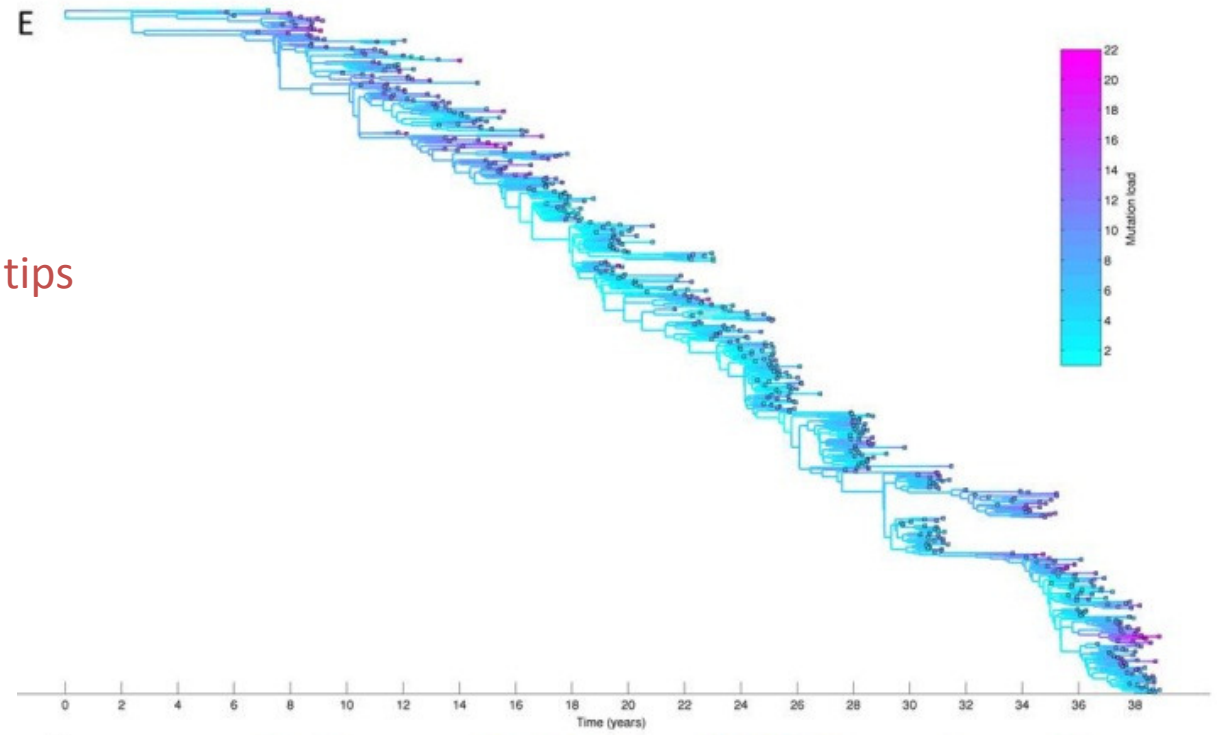
Co-circulation of multiple (more minor) antigenic variants. Consistent with:

Strelkowa and Lässig (2012) *Genetics*

Full ‘phylodynamic’ simulations

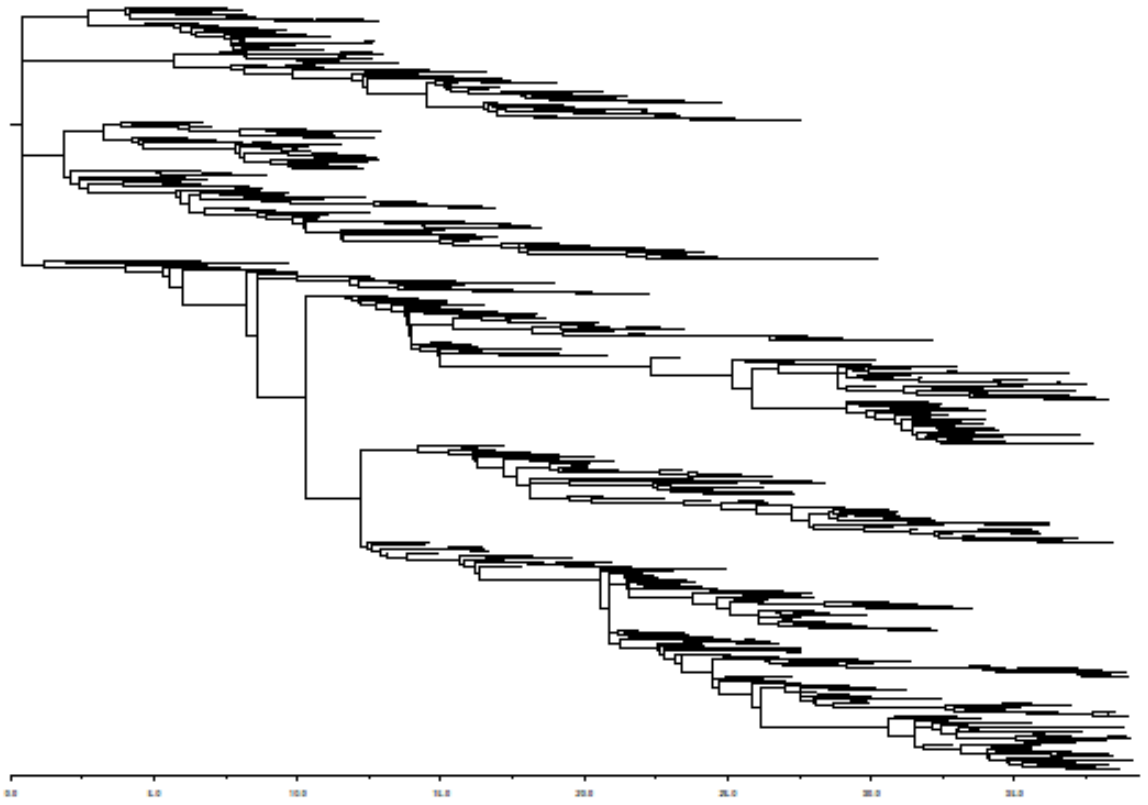
Load simulations
(40 yrs)

Low-load trunk, high-load tips



Full ‘phylodynamic’ simulations

Load simulations, no antigenic evolution
(40 yrs)

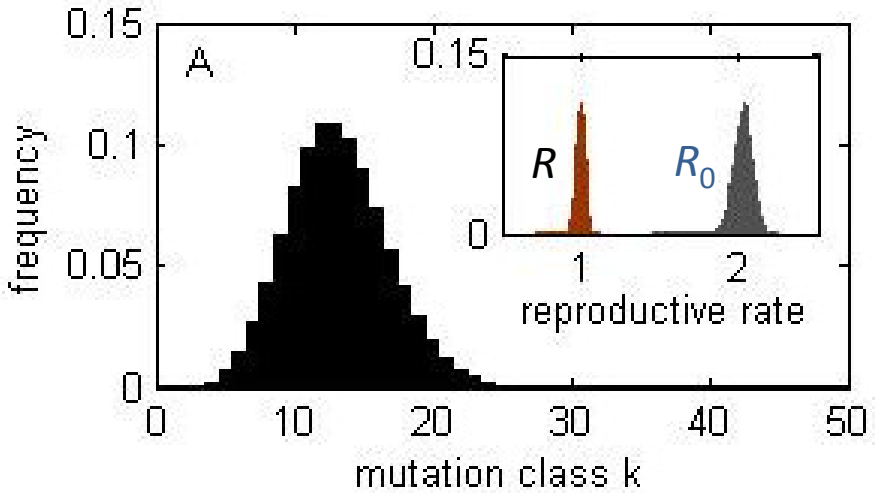


Influenza's phylogeography

Phylogenetic evidence that seasonal epidemics in temperate regions (e.g. US) are seeded from the outside

Nelson et al. (2006) *PLoS Pathogens*

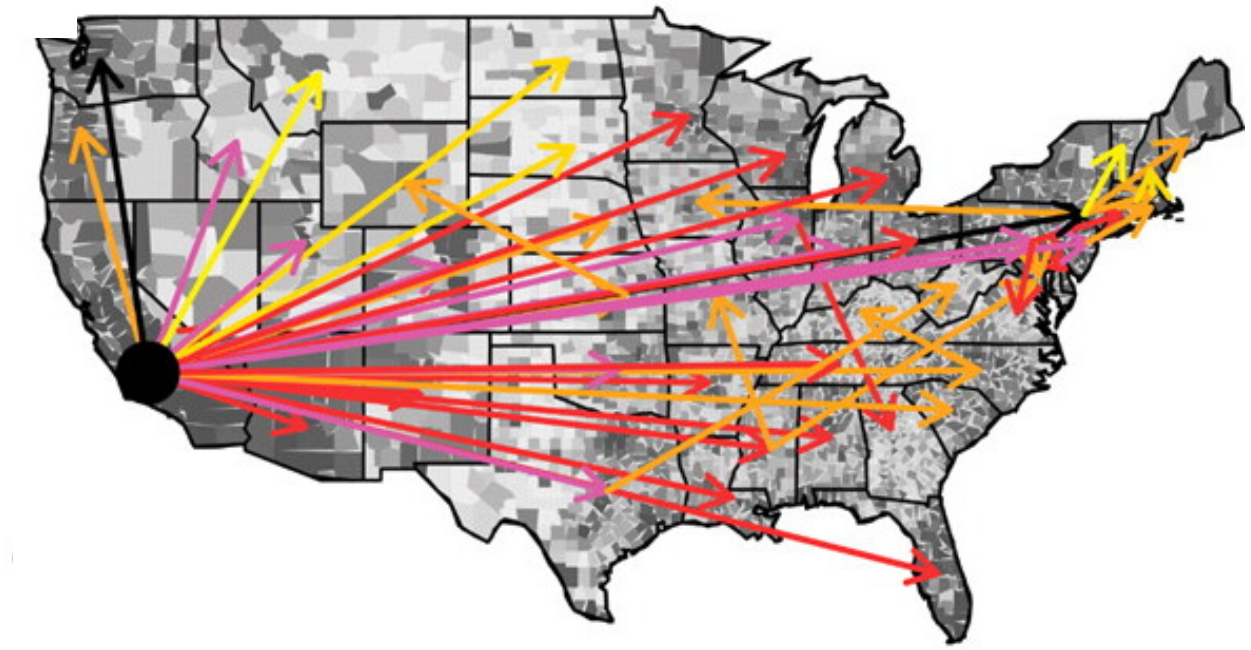
Outside world



Temperate region

Extension of Viboud et al. (2006) *Science* for spread within US

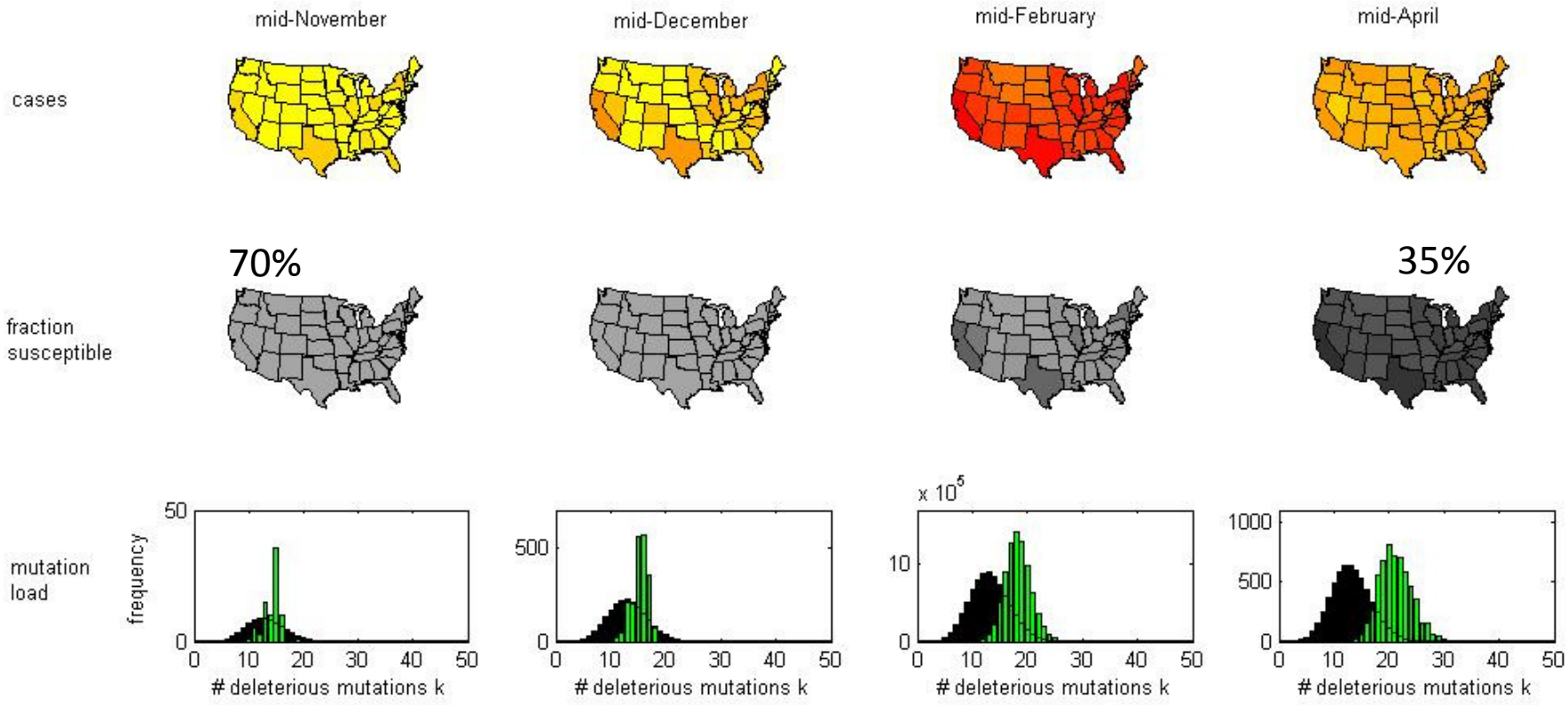
Spatial spread within US



- State-to-state workflow data
- Gravity model: transmission between states depends on their population sizes and distances between one another
- Seasonal spread of flu through US – takes ~3-5 weeks to reach last states

Viboud et al. (2006) *Science*

Modeling the spatiotemporal spread of H3N2 in the US



Very low number of viruses that have 'good' genetic backgrounds. 'Jackpot', or even transient circulation, is therefore very difficult to attain. Only way in which 'jackpot' can be hit is via reassortment. Temperate regions therefore unlikely to be responsible for major drift variants, and therefore will not populate the trunk.

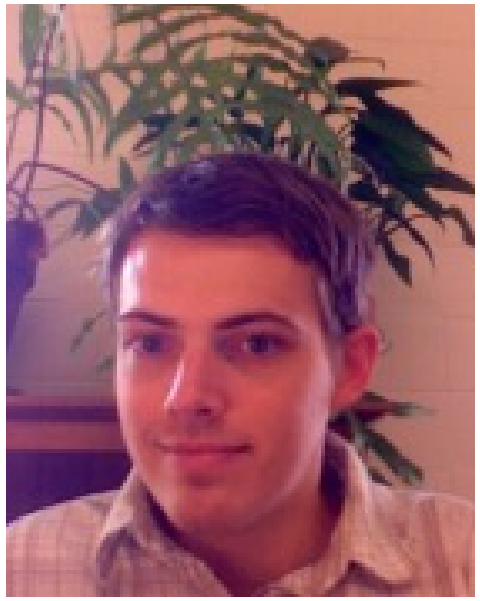
Why? Surfing the expanding wavefront… Travis et al. (2007) *MBE*, Hallatschek and Nelson (2010) *Evolution*


Ruby in the rubbish

Acknowledgments

James S. McDonnell Foundation

David Rasmussen



Koelle lab