

The rate of adaptation in large sexual populations

Daniel Weissman, Nick Barton

2/14/11

The question

- ▶ How fast can evolution go?
- ▶ What determines the rate?

What are the important parameters?

“Forces”: mutation, recombination, drift, selection

- ▶ Mutation rates: $\sim 10^{-10}$? (*E. coli*), $\sim 10^{-4}$? (poliovirus)
- ▶ Recombination rates: 0? (TB), 10^{-8} – 0.5 (*Drosophila*)
- ▶ Numbers of organisms: $\sim 10^2$ (tiger), $\sim 10^{21}$ (*E. coli* in people)
- ▶ Selective coefficients: ???
- ▶ Genome sizes (L): $\approx 10^4$ (poliovirus), $\approx 3 \times 10^9$ (people)

How do these numbers combine?

What's the speed limit on adaptation?

Best/simplest case:

- ▶ 'big' rate of beneficial mutations U ($\propto L\mu$)
- ▶ all have advantage s , add together
- ▶ Constant population size N

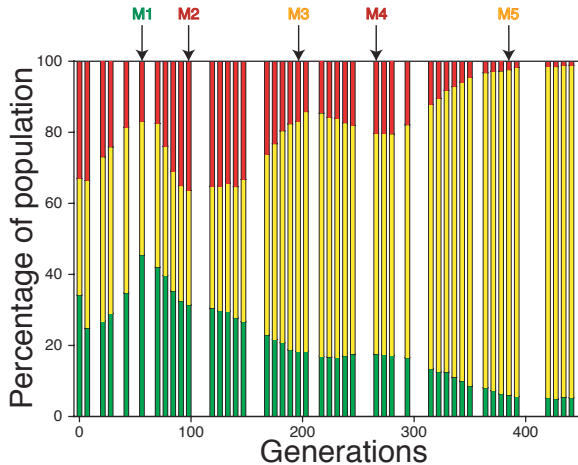
speed of adaptation:

$$v = NU p_{\text{fix}} s$$

- ▶ when mutations are rare:

$$v = 2NU s^2 \equiv v_0$$

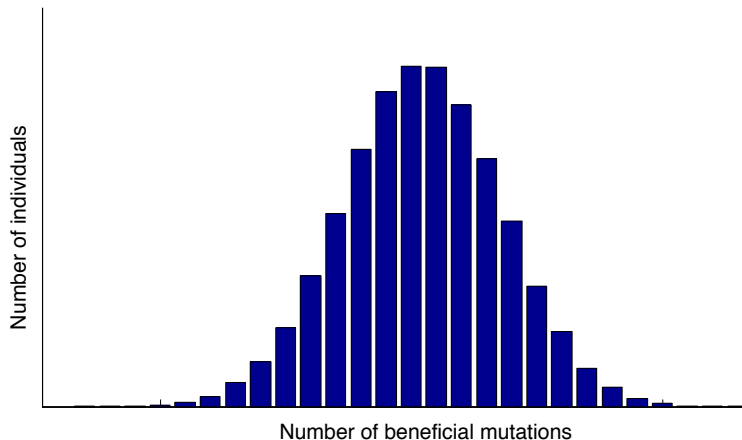
Problem: "clonal interference"



Kao and Sherlock, 2008

$$p_{\text{fix}} \neq 2s$$

Many mutations



Fundamental Theorem: $v =$ variance

Asexual: $v \sim 2s^2 \log(N\sqrt{sU}) / \log^2(s/U)$ (Desai and Fisher, '07; Rouzine et al, '08)

Large recombination limit

(Many) unlinked loci

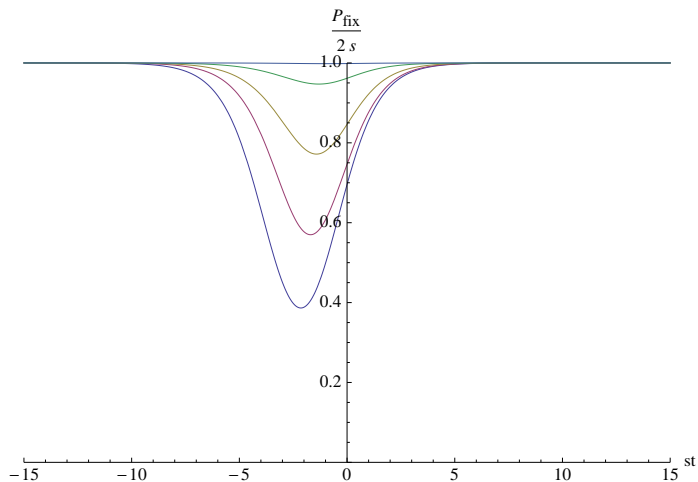
$$p_{\text{fix}} = 2s/\text{variance in reproductive value} = 2se^{-4v}$$

$$\begin{aligned}\Rightarrow v &= \frac{1}{4} \mathfrak{W}(8NUs^2) = \frac{1}{4} \mathfrak{W}(4v_0) \\ &\approx \begin{cases} v_0(1 - 4v_0) & \text{for } v_0 \ll 1 \\ \frac{1}{4} \log(4v_0) & \text{for } v_0 \gg 1 \end{cases}\end{aligned}$$

Tightly linked loci

- ▶ How much does one beneficial mutant interfere with the ones around it?
- ▶ Mutations' fates determined while rare, interfere with others while common
⇒ assume interfering sweep has logistic shape

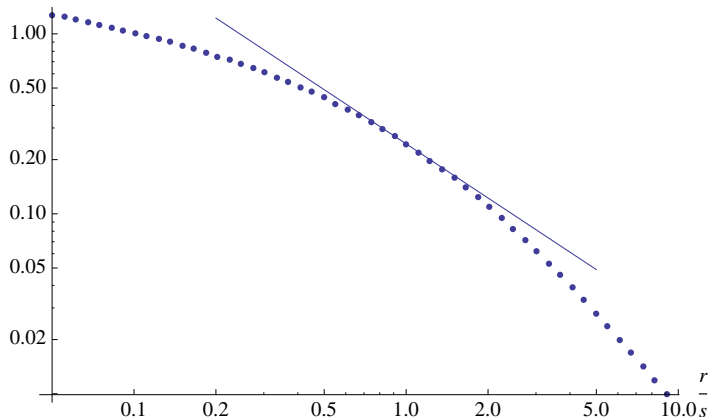
Interference over time, genome



$$r/s = 10, 1, 10^{-1}, 10^{-2}, 10^{-3}$$

Interference over genome integrated over time

$$\int (2s - P_{\text{fix}}(r, t)) dt$$



interference dominated by $r \sim s$

$$\int \int \left(1 - \frac{P_{\text{fix}}}{2s}\right) dt dr = 2$$

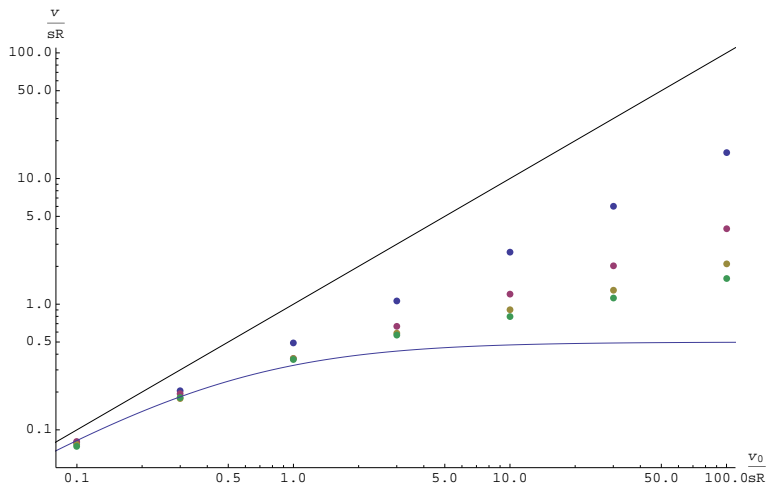
Multiple interfering sweeps?

Guess:

- ▶ don't interfere with each other
- ▶ effects on rare mutants multiply

$$\Rightarrow v = \frac{v_0}{1 + 2v_0/sR}$$

Works for up to moderate interference



$s = .05$, $R = 1$, $N = 10^2, 10^3, 10^4, 10^5$ (moving down)

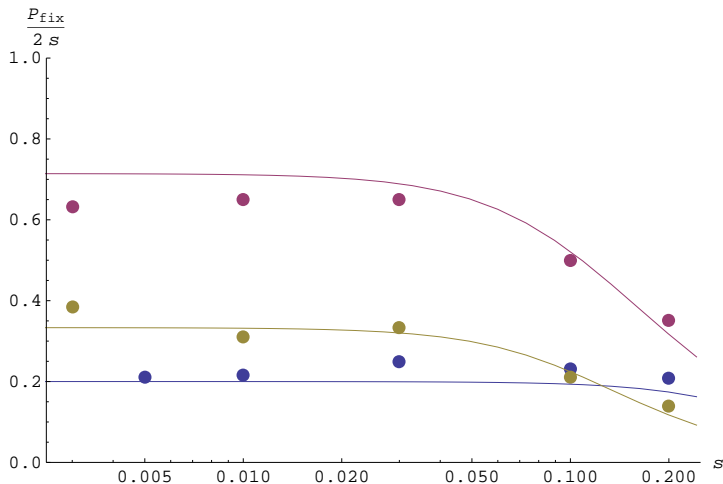
Combining close and far sweeps

$$v = v_0 \left(1 - \frac{2v}{sR} \right) e^{-4v} \text{ for } R \gg s, v_0 \gg sR, 1$$

$$v \approx v_0 \left(1 - \frac{2v_0}{sR} - 4v_0 \right) \text{ for } v_0 \ll sR, 1$$

For $Rs < 1$, interference mostly from few close sweeps, not many far sweeps

Works for up to moderate interference



purple: $N = 10^5, U = 10^{-4}, R/s = 100$

gold: $N = 10^5, U = 5 \times 10^{-4}, R/s = 100$

blue: $N = 10^5, U = 10^{-4}, R/s = 10$

Summary

- ▶ $v \approx v_0 \left(1 - \frac{2v_0}{sR} - 4v_0\right)$, where $v_0 = 2NUs^2$
 - ▶ works until strong interference, then logarithmic
- ▶ interference between closely linked sweeps dominates for $Rs < 1$
- ▶ viral and microbial adaptation may be recombination-limited

Next steps

- ▶ high interference regime
- ▶ distribution of selective coefficients
 - ▶ neutral mutations
- ▶ epistasis
- ▶ population structure