Start w/ 1 mutant in a population of wild types

mutant fitness: f

wild type fitness: g

$$s = \frac{f}{g} - 1$$

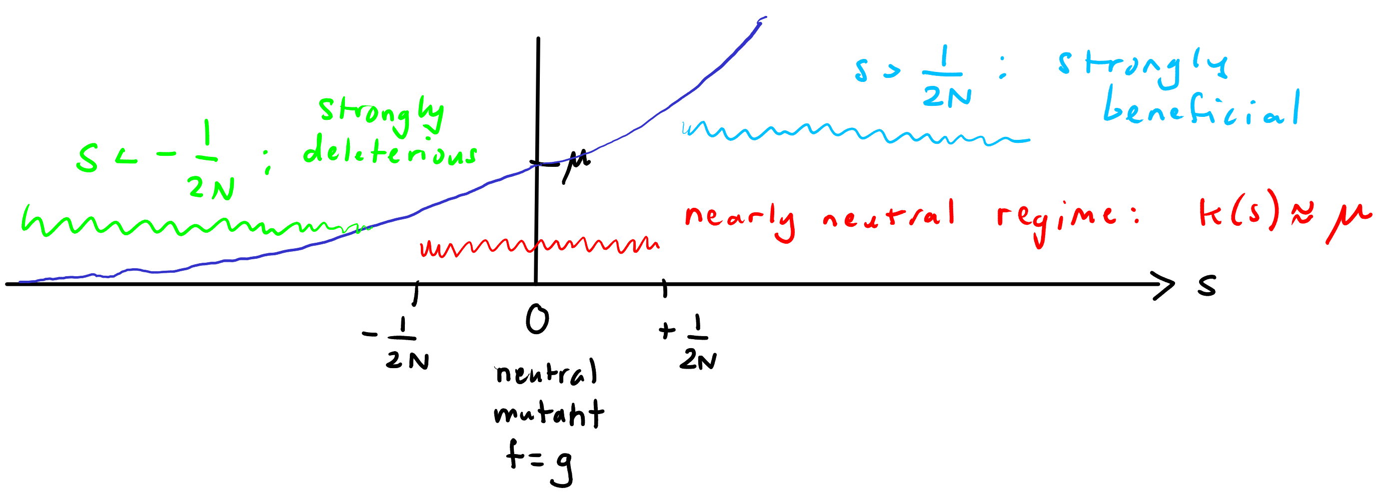
prob. for mutant to take over: (fixation)

$$\pi_{fix} \approx \frac{1 - e^{-s}}{1 - e^{-2Ns}}$$

N = effective population

substitution rate = rate at which new mutations appear & take over at a given location

$$k(s) = \underbrace{2N\mu}_{\text{rate at which new mutations appear}} \frac{1 - e^{-s}}{1 - e^{-2Ns}}$$



$|s| \ll \frac{1}{2N}$:
$$k \approx 2N\mu \frac{1 - (1 - s)}{1 - (1 - 2Ns)} = \mu \text{ when } s = 0$$

nearly const. subs. rate for nearly neutral mutants, indep. of N

"molecular clock" (Kimura, Ohta 1971)

strongly deleterious: $s \ll -\frac{1}{2N}$

$$\pi_{fix} \approx |s| e^{-2N|s|}$$

$$k \approx 2N\mu |s| e^{-2N|s|} \ll \mu$$

strongly beneficial: $s \gg \frac{1}{2N}$

$$\pi_{fix} \approx s$$

$$k = 2N\mu s \gg \mu$$

Generalize this model to multiple possible mutants at a given location
Sella + Hirsh PNAS (2005)

Zoom out

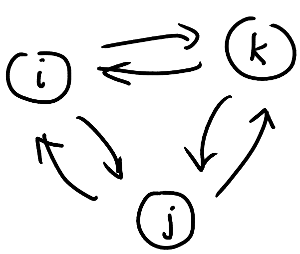
genetic variant i

(i = C, T, A, G for one base pair)

State i:
whole population is same genetic variant w/ fitness f_i

could be very slow

state j:
" "
"
w/ fitness f_j



Ω_{ij} = prob. to reach state i, given start at j per unit time

$$s_{ij} \equiv \frac{f_i}{f_j} - 1 \qquad = 2N\mu \left[\frac{1 - e^{-s_{ij}}}{1 - e^{-2Ns_{ij}}} \right]$$

description is valid in small μ limit $\mu \ll \frac{1}{N}$ (typical for biology)

assume $|s_{ij}| \ll 1$

$$e^{-s_{ij}} = \frac{1}{e^{s_{ij}}} \approx \frac{1}{1 + s_{ij}} = \frac{f_j}{f_i}$$

$$\Omega_{ij} \approx 2N\mu \left[\frac{1 - \frac{f_j}{f_i}}{1 - \left(\frac{f_j}{f_i}\right)^{2N}} \right]$$

$$\frac{\Omega_{ji}}{\Omega_{ij}} = \left(\frac{f_j}{f_i}\right)^{2N-1} = e^{-\left[(2N-1)\left(-\ln f_j - (-\ln f_i)\right)\right]}$$

compare $= e^{-\beta(E_j - E_i)}$

$$\Rightarrow \quad \beta = \frac{1}{k_B T} = 2N - 1 \qquad E_i = -\ln f_i$$
$$E_j = -\ln f_j$$

as $t \to \infty$ system reaches equilibrium:

$$p_i^{eq} = \frac{e^{-\beta E_j}}{Z}$$

$$Z = \sum_i e^{-\beta E_i}$$

low "energy" (high fitness) states preferred, but....

pref. stronger $\Rightarrow T$ is lower

only if $N \to \infty$
does the fittest
actually totally dominate

β is higher
N is larger