8.592 Lecture 4 2/10/2011

Selection, Wrigth's equation, Effective population size

<u>Selection</u> (frequency independent selection)

	A_1A_1	A_1A_2	A_2A_2	
	w_{11}	w_{12}	w_{22}	← fitness
	p^2	$2pq$	q^2	← freq. before selection
	$\frac{w_{11}}{\bar{w}}p^2$	$\frac{w_{12}}{\bar{w}}2pq$	$\frac{w_{22}}{\bar{w}}q^2$	← after selection
	1	$1-hs$	$1-s$	← relative fitness

<u>Where</u> $\bar{w} = p^2 w_{11} + 2p(1-p)w_{12} + (1-p)^2 w_{22}$ ← mean fitness of the population

$\frac{w_{11}}{\bar{w}}$ $\frac{w_{12}}{\bar{w}}$ $\frac{w_{22}}{\bar{w}}$ ← relative fitness

s - selectione coeffeceint (strength of selection), s>0

h determines the type of selection

* <u>Types of selection</u>

h=0 A1 dominant

h=1 A2 dominant

0<h<1 incomplete dominance -> directional selection

h<0 overdominance -> balancing selection

h>1 underdominance -> disruptive selection

* Special type of balancing selection-

- frequency-dependent selection.

* <u>Let's calculate the effect of selection on allele freq. *p*</u>

(in one generation)

1

$$p_{t+1} = p_t^2 \frac{w_{11}}{\bar{w}} + \frac{2(1-p)p}{2} \frac{w_{12}}{\bar{w}}$$

$$p' = p^2 \frac{w_{11}}{\bar{w}} + 2\frac{(1-p)}{2}p \frac{w_{12}}{\bar{w}}$$

$$\bar{w} = p^2 w_{11} + 2p(1-p)\, w_{12} + w_{22}(1-p)^2$$

$$\frac{d\bar{w}}{dp} = 2pw_{11} + 2(1-2p)\, w_{12} - 2w_{22}(1-p)$$

$$\Delta p = \frac{1}{\bar{w}}\left[p^2 w_{11} + (1-p)pw_{12} - p[p^2 w_{11}] - 2p^2(1-p)w_{12} - p(1-p)^2 w_{22}\right.$$

$$= \frac{1}{\bar{w}}\left[w_{11}(1-p)p^2 + w_{12}(1-p)\,p(1-2p) - w_{22}\,p(1-p)^2\right]$$

$$= \frac{p(1-p)}{2\bar{w}}\left[2w_{11}p + 2w_{12}(1-2p) - 2w_{22}(1-p)\right]$$

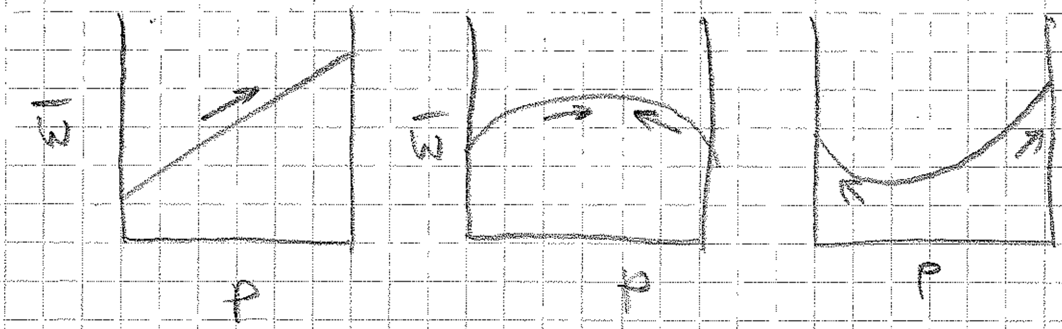
$$\boxed{\Delta p = \frac{p(1-p)}{2\bar{w}} \frac{d\bar{w}}{dp}}$$ Wright Equation

$$= \frac{p(1-p)}{2} \frac{d \log \bar{w}}{dp}$$

1. Direction of evolution is determined by $\frac{d \log \bar{w}}{dp}$ ← adaptive landscape

$$\frac{d \log \bar{w}}{dp} > 0 \Rightarrow \Delta p > 0$$

2. Stable equilibria are points where $\frac{d \log \bar{w}}{dp} = 0$, i.e. local maxima of $\bar{w}$



3. Evolution maximizes mean fitness of the population

* Selection + drift

Probability of fixation

Selection $\langle \Delta p \rangle = \frac{p(1-p)}{2\bar w} \cdot \frac{d\bar W}{dp}$ assuming $h = 1/2$

$\bar W = p^2 + 2p(1-p)(1-hs) + (1-p)(1-s) \doteq 1 + ps - s = 1 - qs$

$\boxed{\langle \Delta p \rangle = \frac{p(1-p)}{2(1+ps-s)} \cdot s \approx \frac{p(1-p)}{2} s}$ assuming $s \ll 1$

$$\underbrace{\Pi'(p)\langle \Delta p \rangle}_{M(p)} + \frac{1}{2}\Pi''(p)\underbrace{\langle \Delta p^2 \rangle}_{V(p) = Var(p) + \langle \Delta p \rangle^2} = 0$$

$$\frac{1}{2}\Pi''(p)V(p) + M\Pi'(p) = 0$$

$M(p) \equiv \langle \Delta p \rangle = \frac{p(1-p)s}{2}$

$V(p) = \frac{p(1-p)}{2N} + \cancel{s^2}$

$$\frac{\Pi''}{\Pi'} = -\frac{2M}{V}$$

$$-\log \Pi' = 2\int_0^p \frac{M(x')}{V(x')} dx' + C$$

$$= 2\int_0^p Ns\, dx + C = 2Nsp + C$$

$$\Pi'(p) = Ce^{-2Nsp} \Rightarrow \Pi(p) = C_1 e^{-2Nsp} + C_2$$

boundary
$\Pi(0) = 0$
$\Pi(1) = 1$

$\Rightarrow \Pi(0) = C_1 + C_2 = 0$ $\quad$ $\Pi(1) = C_1 e^{-2Ns} - C_1 = C_1(e^{-2Ns} - 1) = 1$

$C_2 = -C_1$ $\quad$ $C_1 = 1/(e^{-2Ns} - 1)$

Haldein's equation

$$\boxed{\pi(p) = \frac{1-e^{-2Nsp}}{1-e^{-2Ns}}}$$

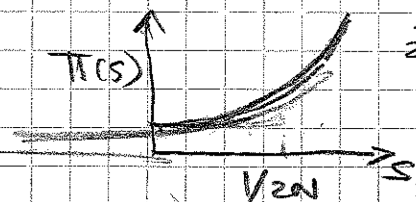
$s>0 \Leftrightarrow A_2$ disadvantageous mutation

1) New mutations

$$p = \frac{1}{2N}$$

$$\pi\left(\frac{1}{2N}\right) = \frac{1-e^{-s}}{1-e^{-2Ns}} \approx \frac{s}{1-e^{-2Ns}} \approx s \quad \text{for } 2Ns \gg 1$$

$$\frac{\cancel{s}}{2N\cancel{s}} = \frac{1}{2N} \quad \text{for } 2Ns < 1$$



2) For $h \neq \frac{1}{2}$

$$\pi\left(\frac{1}{2N}\right) \approx 2\underbrace{(1-h)}_{\text{selective advantage of heterozygot}}s$$

depends on the fitness of heterozygote

3) Deleterious $\pi_1 = \pi$

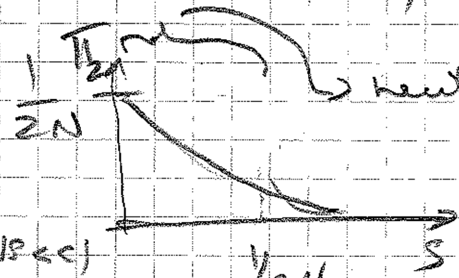
$$\pi_2 = 1-\pi_1(1-q) = \frac{e^{2Nsq}-1}{e^{2Ns}-1}$$

$$\frac{1-e^{-2Ns}-1+e^{-2Ns(1-q)}}{1-e^{-2Ns}} = \frac{-1+e^{2Nsq}}{e^{2Ns}-1} =$$

New mutation
$$q = \frac{1}{2N}$$

$$\pi_2 = \frac{s}{e^{2Ns}-1}$$

$$\approx \frac{1}{2N} \quad \text{for } 2Ns \ll 1$$



$$k = \pi\left(\tfrac{1}{2N}\right) 2N\mu F \quad \frac{2N\mu s}{e^{2Ns}-1}$$

rate of fixation

rate of a.a. subst = rate of neutral subst.

2. Effective population size

Recall

H - heterozygosity (fraction of different nucleotides)

$$H_{t+1} = H_t \left(1 - \frac{1}{2N}\right)$$

if N_t, then

$$H_{t+1} = H_t \left(1 - \frac{1}{2N_t}\right)$$

$$H_t = H_0 \prod_{i=0}^{t-1} \left(1 - \frac{1}{2N_i}\right) \quad \text{compare to} \quad H_t = H_0 \left(1 - \frac{1}{2N_e}\right)^t$$

effective pop. size

then

$$\left(1 - \frac{1}{2N_e}\right)^t = \prod_{i=0}^{t-1} \left(1 - \frac{1}{2N_i}\right)$$

$$\prod_{i=1}^{n} (1+x_i) = e^{\sum \log(1+x_i)}$$

$x \ll 1$

using $\log(1+x) \approx x - \frac{x^2}{2} \approx x$ for $x \sim 0$.

we obtain

$$= e^{\sum x_i} \approx 1 + \sum x_i$$

$$e^{-t/2N_e} = e^{-\sum_{i=0}^{t-1} \frac{1}{2N_i}}$$

$$N_e = \left(\frac{1}{t} \sum_{i=0}^{t-1} \frac{1}{N_i}\right)^{-1}$$

Harmonic mean

- the value is dominated by the smallest population size in history

Sensitive to bottlenecks