Lecture 5. Mutation. Migration and selection.

2.1 Mutation

One locus two alleles model: A wildtype, a mutant
forward mutation rate μ per generation: $A \to a$
backward mutation rate ν: $a \to A$
Typically μ is about
$10^{-4} - 10^{-6}$ mutations per gene per generation

p_t = population frequency of allele A in generation t

Irreversible mutation

If $\nu = 0$, then pure loss of alleles A each generation
$p_t = p_{t-1}(1-\mu) = p_0(1-\mu)^t \approx p_0 e^{-\mu t}$
Fig 5.1, p. 165: $p_t \to 0$ under mutation pressure

$\Delta p = -p\mu$ incremental frequency $\Delta p = p_t - p_{t-1}$

Half-life of an allele

The number of generations
taking to halve the wildtype allele frequency
$t_{0.5} = \frac{\ln 2}{\mu} = \frac{0.693}{\mu}$ solves equation $p_t = 0.5 \cdot p_0$
If $\mu = 10^{-4}$, the half-life is $t_{0.5} = 6,930$ generations
if $\mu = 10^{-6}$, the half-life is $t_{0.5} = 693,000$ generations

Ex 1: mutation rate estimation

Fig 5.3, p. 167: infection resistance gene in $E.coli$
if the cumulative mutation rate μt is small, then
$p_t \approx p_0(1-\mu t)$; if moreover $p_0 \approx 1$, then $q_t \approx q_0 + \mu t$

Ex 2: transposon deletion

D. mauritania, a site with transposon *mariner* insertion
spontaneous deletion at rate $\mu = 0.01$: $A \to a$
If, $D_0 = 1$, find t needed to reach $R_t = 0.05$:
assuming random mating $q_t = \sqrt{R_t} = 0.224$
linear approximation $q_t = 0.01 \cdot t$ gives $t = 23$
exact formula $q_t = 1 - (0.99)^t$ gives $t = 26$ generations

Reversible mutation

If $\mu > 0$ and $\nu > 0$, then
the allele A loss and gain interplay:
$p_t = p_{t-1}(1-\mu) + q_{t-1}\nu = p_{t-1}(1-\mu-\nu) + \nu$

$$\Delta p = -p(\mu + \nu) + \nu$$

Equlibrium frequency $\hat{p} = \frac{\nu}{\mu+\nu}$ solves $\Delta p = 0$

$p_t = \hat{p} + (p_0 - \hat{p})(1 - \mu - \nu)^t$

Fig 5.4, p. 169

$\mu = 10^{-4}$, $\nu = 10^{-5}$, $\hat{p} = 0.091$

Ex 3: intrachromosomal recombination

Salmonella bacterium:

switching between two forms of flagella

due to an intrachromosomal recombination

switching rates are high: $\mu = 8.6 \cdot 10^{-4}$, $\nu = 4.7 \cdot 10^{-3}$

Observed results for two Salmonella cultures

t	0	30	700
p_t	0	0.16	0.85

t	0	388	700
p_t	1	0.88	0.86

Expected frequencies

1: $p_t = 0.845(1 - (0.994)^t)$, $p_{30} = 0.13$, $p_{700} = 0.83$

2: $p_t = 0.845 + 0.155(0.994)^t$, $p_{388} = 0.86$, $p_{700} = 0.85$

Expected equlibrium frequency $\hat{p} = 0.845$

2.2 Migration

Immigration rate m into a subpopulation
= the subpopulation proportion quota for
new immigrants arriving each generation

If $m = 0.05$, then 5% of the subpopulation individuals have immigrated during the last generation period

One-way migration

Fig 5.14, p. 190: mainland to island migration

mainland frequencies are fixed p^*, q^*

Island frequencies change

$p_t = (1 - m)p_{t-1} + mp^*$

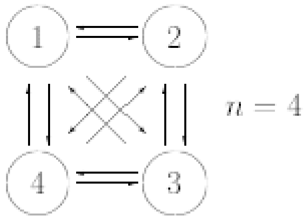
$= \{\text{non-imm. with } A\} + \{\text{immigrants with } A\}$

$$\boxed{\Delta p = -pm + mp^*}$$

Convergence to the mainland frequency

$p_t = p^* + (1 - m)^t(p_0 - p^*)$ so that $p_t \to p^*$

Island model of migration



p_t = allele A frequency
in a certain subpopulation

$\bar{p}$ = allele A frequency
in the metatpopulation
constant over time

The same dynamics as with mainland to island migration

$p_t = (1 - m')p_{t-1} + m'\bar{p}$, where $m' = m \cdot \frac{n}{n-1}$

Gene flow eliminates differences among subpopulations

$p_t = \bar{p} + (1 - m')^t(p_0 - \bar{p})$ so that $p_t \to \bar{p}$

Fig 5.16, p.194

evolution similar to reversible mutation

difference in rates: $m \gg \mu$

2.3 Selection

Haploid selection

Absolute fitnesses W_A, W_a

= offspring numbers for two bacteria strains A and a

Fig 6.1, p. 213: two potential growth rates

Carrying capacity of the habitat is limited

focus on the allele competition within a population

relative fitnesses $w_A : w_a = W_A : W_a$

$w_A = 1$, $w_a = 1 - s$, haploid selection coefficient s

Two potential growth rates

$X_t = X_{t-1}W_A$ number of alleles A in generation t

$Y_t = Y_{t-1}W_a$ number of alleles a in generation t

Allele frequencies

$p_t = \frac{X_t}{X_t+Y_t}$, $q_t = \frac{Y_t}{X_t+Y_t}$ odds ratio $\frac{p_t}{q_t} = \frac{X_t}{Y_t}$

$\frac{p_t}{q_t} = \frac{p_{t-1}}{q_{t-1}}(1 - s)^{-1} = \frac{p_0}{q_0}(1 - s)^{-t}$

$p_t = \frac{p_0}{p_0+q_0(1-s)^t}$

Haploid selection $\Delta p = spq$, if $s \approx 0$

Fixation of the favored allele

$p_t \to 1$ if $s > 0$ and $p_t \to 0$ if $s < 0$ as $t \to \infty$

Estimate s using linear regression

$\ln(\frac{p_t}{q_t}) = \ln(\frac{p_0}{q_0}) - t\ln(1 - s)$ or

$\ln(\frac{p_t}{q_t}) = \ln(\frac{p_0}{q_0}) + st$, if $s \approx 0$

Diploid selection

genotype	AA	Aa	aa
relative fitness	w_{AA}	w_{Aa}	w_{aa}

Three types of the diploid selection

directional selection:

$w_{AA} > w_{Aa} > w_{aa}$ or $w_{AA} < w_{Aa} < w_{aa}$

overdominance: $w_{Aa} > w_{AA}$ and $w_{Aa} > w_{aa}$

stabilizing selection against homozygotes

underdominance: $w_{Aa} < w_{AA}$ and $w_{Aa} < w_{aa}$

disruptional selection against heterozygotes

Biological components of human fitness

survival to maturity, mating success, and fertility

Two stage life history model

adults → random mating → newborns → adults

fitness is proportional to P(Survival to maturity)

Genotype frequencies

in adults (D, H, R) and newborns (d, h, r)

From newborns to adults survival to maturity

$D : H : R = dw_{AA} : hw_{Aa} : rw_{aa}$

From adults to next generation newborns

random mating $d_{\text{next}} = p^2$, $h_{\text{next}} = 2pq$, $r_{\text{next}} = q^2$

Two relations combined

$D_{\text{next}} = p^2\frac{w_{AA}}{\bar{w}}$, $H_{\text{next}} = 2pq\frac{w_{Aa}}{\bar{w}}$, $R_{\text{next}} = q^2\frac{w_{aa}}{\bar{w}}$

Average fitness $\bar{w} = p^2 w_{AA} + 2pqw_{Aa} + q^2 w_{aa}$

is close to one if selection is weak

$$\Delta p = \frac{pq}{\bar{w}}(p(w_{AA} - w_{Aa}) + q(w_{Aa} - w_{aa}))$$

Literature:

1. D.L.Hartl, A.G.Clarc. Principle of population genetics. Sinauer Associates, 2007.

2. R.Nielson, M. Statkin. An introduction to population genetics: theory and applications, Sinauer Associates. 2013.