4 Mutation Selection Balance (very brief notes)

Don't worry too much about all of the algebra in these notes, but do understand the general logic and how that leads to eq. 4.4

In previous chapters we saw that natural selection can be a strong force that will eventually lead to the fixation of advantageous alleles (or to the loss of deleterious alleles). Nevertheless, many apparently maladaptive traits can be found in populations. For example, how does a fatal genetic disease such as Tay Sachs disease persist in a population? It would seem that selection would have removed those alleles long ago. One possibility is that there is a balance between the production of new alleles by mutation and the removal of those alleles by natural selection.

Here we will start with a general algebraic model of selection to find an expression for the rate of change in allele frequency, Δp. Then we will try to find conditions where the loss of alleles by selection are exactly balanced by the gain through mutation and see if those conditions are consistent with what we know about mutation rates disease frequency.

4.1 A general model of selection

Let's look at a very general case where we let the fitnesses be

Genotype	AA	Aa	aa
Relative fitness	1	1-*hs*	1-*s*

Here s is the selection coefficient against the aa homozygote. If s=1 the fitness of aa is zero and a is a homozygous lethal allele. If s=0 then alleles A and a are equal in terms of fitness. The coefficient "h" is the dominance coefficient. If h=0 then allele A is dominant. If h=1 then allele a is dominant. And if h=0.5 then the fitness of the heterozygote is exactly intermediate between the two homozygotes. If h is negative the heterozygote will actually have higher fitness than either homozygote.

Now, from equations 2.1 - 2.3

$$p' = \frac{p^2 \cdot 1 + pq(1-hs)}{\bar{w}}$$

$$\bar{w} = p^2 \cdot 1 + 2pq(1-hs) + q^2(1-s)$$

$$\Delta p = p' - p$$

Δp is the expression we want to find. After about a full page of algebra (which is at the end of the document, if you really want to see it), that equation can be simplified to

$$\Delta p = \frac{pqs\left[ph + q(1-h)\right]}{\bar{w}} \qquad \text{eq. 4.1}$$

It is still a bit complicated…

4.2 **Mutation-selection balance for a completely recessive mutant:**

To make more progress lets look at the special case of a completely recessive mutant. In that case *h*=0.

Then $\bar{w} = p^2 + 2pq + q^2(1-s)$. Using the fact that $p^2 + 2pq + q^2 = 1$, equation 4.1 simplifies to

$$\Delta p = \frac{spq^2}{1 - sq^2} \qquad \text{eq. 4.2a}$$

That one looks a bit easier!

As one allele increases in frequency the other must decrease by an equal amount so we can also write

$$\Delta q = \frac{-spq^2}{1 - sq^2} \qquad \text{eq. 4.2b}$$

We will use that equation to follow the fate of the recessive allele.

Now here is a trick: Because we expect the deleterious allele to remain quite rare in the population q^2 will be close to zero and the denominator of 4.2b will be close to 1.0. Also the frequency of the normal allele will be close to p=1. Therefore we can make a good approximation for the change in allele frequency as

$$\Delta q \approx -sq^2 \qquad \text{eq. 4.3}$$

Selection against a deleterious recessive allele will decrease the frequency of that allele by that amount each generation.

At the same time mutations are always occurring. New alleles will be formed by mutation at some rate m per generation. $\Delta q = +\mu$. Eventually there will be an equilibrium where the loss of alleles due to selection against the deleterious recessive is exactly balanced by the gain of alleles through new mutations.

At that equilibrium, . $\mu = sq^2$. Solving for q, we find that

$$q^2 = \frac{\mu}{s} \qquad \text{eq. 4.4a}$$

or

$$\hat{q} = \sqrt{\frac{\mu}{s}} \qquad \text{eq. 4.4b}$$

4.3 **An example: cystic fibrosis.**

Cystic fibrosis is a genetic disease caused by a defect in a particular protein, the CFTR ion transporter. One of the main symptoms of the disease is that patients have excess fluid in their lungs because of thick mucus secretions. Patients have difficulty breathing and are

very susceptible to respiratory infections. Until very recently CF was a fatal disease and the life expectancy of patients with CF is still only about 30 years.

Despite its severity, Cystic Fibrosis is one of the more common genetic diseases. In the US the frequency of cystic fibrosis is about 1/3000 births.

If the disease occurs in 1/3000 births, what is the allele frequency for the CF allele?

q=_____

One might ask why a genetic disease that is so severe persists in the population at all. Selection should be a potent force removing disease alleles from the human population. Can it be explained by a balance between selection and mutation?

For a recessive lethal s=1.

Using equation 4.4, what would the mutation rate have to be to produce the observed frequency of CF alleles in the US?

μ=_____

Observed mutation rates for CF are on the order of 6×10^{-7}. Is mutation selection balance a reasonable explanation for the frequency of CF?

4.4 **(optional) Heterozygote advantage**

If mutation cannot explain the high frequency of the CF allele, then what other explanations are possible?

One possibility is that there is that the mutation may have a fitness advantage in heterozygotes. If the heterozygote has the highest fitness then neither allele will go to complete fixation. The allele frequency will equilibrate at an intermediate frequency.

We can determine that equilibrium from equation 0.1. To find that equilibrium set $\Delta p = 0$. There will be an equilibrium when either p=0, q=0, s=0, or $\left[ph + q(1-h)\right] = 0$ and it is the last possibility that is the interesting situation. In that case $ph = -q(1-h)$.

Letting p=1-q and simplifying, you can get

$h - qh = qh - q$

$h = 2qh - q$

$h = q(2h - 1)$

and finally

$$\hat{q} = \frac{h}{2h-1} \qquad \text{eq. 4.5}$$

You can also solve for h to get

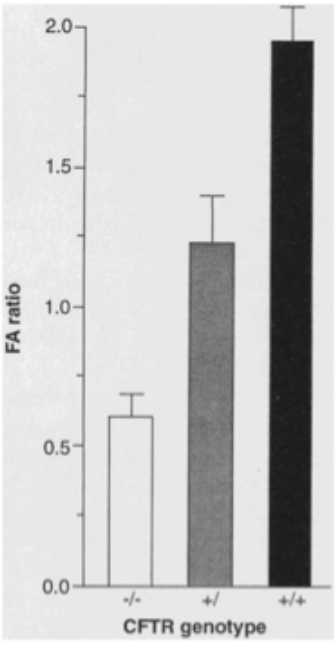
$$h = \frac{\hat{q}}{2\hat{q} - 1} \qquad \text{eq. 4.6}$$

Now if the observed allele frequency of the CF allele is q=0.018, how big does the heterozygote advantage have to be in order to produce that equilibrium frequency? (Look back at the way we defined fitnesses in section 4.1 and notice that h will be negative if the heterozygote has higher fitness than the normal homozygote).

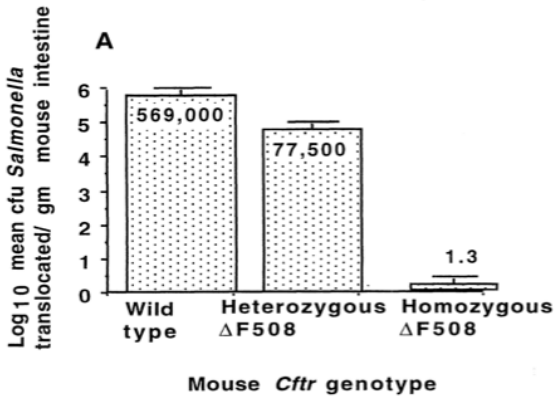
h=_____

A very slight fitness advantage of heterozygotes, on the order of 2%, would be enough to keep the CF allele in the population at the observed frequency.

There has been some speculation about possible advantages of the CF allele. There is some evidence that the CF allele may in fact provide partial protection against diarrheal diseases such as cholera (Gabriel et al. 1994) or typhoid fever (Pier 2000). That evidence comes from studies of mice. There is still no direct evidence for a heterozygote advantage in humans.



Gabriel et al (1994, Science 266:107-109) showed that mice that were heterozygous for the CFTR mutation accumulated less fluid in their intestine after they were inject with the cholera toxin compared to the normal +/+ mice.



Pier (2000, PNAS 97:8822-8828) showed that mice heterozygous for the CFTR mutation had fewer cells of the typhoid fever bacterium than wild type mice.

4.5 **Your turn**

Nick Waser and Mary Price (1981. Evolution35:376-90) examined selection against albino forms of the Rocky Mountain wildflower *Delphinium nelsonii*. They showed that the white- flowered albino plants received fewer visits from pollinators and produced fewer seeds per flower than the normal blue-flowered plants. Nevertheless, the rare white flowered plants were present in most populations. They suggested that those variants might be maintained by mutation selection balance.

Here are some of their data:

	Frequency of white-flowered plants
site1	0.00096
site2	0.00054
site3	0.00103
site4	0.00042
Average	**0.00074**

	Average seed set (# seeds per flower)	Relative fitness
Blue flowered plants	19.7	1.0
White flowered plants	15.8	0.80

- If the relative fitness of albinos is 0.8, what is their value for s?
- Assuming HWE and that the albino allele is completely recessive, what is the observed allele frequency, q?
- What mutation rate would be required for the white allele to be maintained by mutation/selection balance?

4.6 **EXTRA**

Here is the algebra we skipped, in case you are curious. Some of us enjoy this kind of puzzle!

Start with $$\Delta p = \frac{p^2 + pq(1-hs)}{\bar{w}} - p$$

To simplify this equation, we'll multiply p by $\frac{\bar{w}}{\bar{w}}$ so we can have a common dominator, then expand $\bar{w}$ and start canceling terms.

$$\Delta p = \frac{p^2 + pq(1-hs) - p\bar{w}}{\bar{w}}$$

It will first get uglier before we can simplify it:

$$\Delta p = \frac{p^2 + pq(1-hs) - p\left[p^2 + 2pq(1-hs) + q^2(1-s)\right]}{\bar{w}}$$

Next expand the terms in parentheses:

$$\Delta p = \frac{p^2 + pq - pqhs - p\left[p^2 + 2pq - 2pqhs + q^2 - q^2 s)\right]}{\bar{w}}$$

Notice that inside the [ ] there is $p^2 + 2pq + q^2$ which we know equals 1.

$$\Delta p = \frac{p^2 + pq - pqhs - p\left[1 - 2pqhs - q^2 s)\right]}{\bar{w}}$$

We can also factor a p out of everything to get

$$\Delta p = \frac{p\left[p + q - qhs - \left[1 - 2pqhs - q^2 s)\right]\right]}{\bar{w}}$$

Now remove the brackets, changing signs as needed. Also note that p+q=1 so

$$\Delta p = \frac{p\left[1 - qhs - 1 + 2pqhs + q^2 s\right]}{\bar{w}}$$

Here the +1 and -1 cancel. Also there is a +2pqhs and a –qhs so we get

$$\Delta p = \frac{p\left[(2pq - q)hs + q^2 s\right]}{\bar{w}}$$

Move the q and s out of the brackets to get

$$\Delta p = \frac{pqs\left[(2p - 1)h + q\right]}{\bar{w}}$$

The term in parentheses, (2p-1), is equal to p+p-1 or p+(1-q)-1 or finally just (p-q) so

$$\Delta p = \frac{pqs\left[ph + q(1-h)\right]}{\bar{w}} \qquad \text{eq. 4.1}$$

Answers:

p3.
Observed q=0.018
Mutation rate would have to be 0.00032
That is three orders of magnitude (1000x) higher than other estimates of mutation rate, so it does not appear to be a reasonable explanation for the high frequency of CF alleles.

p4
For heterozygote advantage, if q=0.018 then h would have to be -0.0186

P5
From p1, the fitness of the recessive homozygote is defined as w=1-s , so if w=0.8, then s=0.2
The expected frequency of white flowered plants is q^2, so q=sqrt(0.00074) = 0.0272
From eq 4.4, the mutation rate would have to be 0.00015