BIOL2007 EVOLUTION OF GENETIC DIVERSITY

SO FAR:

a) Evolution is change in gene frequencies.
b) Selection can lead to **fixation**.

IN THE NEXT LECTURE:

Kevin will discuss mutation: new raw material for evolution.

HOWEVER:

If alleles always evolved until they become **fixed** (invariant), or lost...
Most of the time, populations would rarely be under selection, and there would be little standing variation. But, in nature things are very different ...

TODAY:

We'll explore how variation can be maintained by natural selection
We'll test for deviation from Hardy-Weinberg, & use it to estimate selection

PLENTY OF GENETIC VARIATION EXISTS!

e.g. snail shell colour/banding, human eye/hair colour, protein variation in just about everything, DNA variation in absolutely everything

Understanding this genetic diversity is a major goal.

Possible explanations:

1) **Selection** on its own.

a) **heterozygous advantage** - selection for heterozygotes

b) **diversifying frequency-dependent selection** - selection for rare forms when their frequencies are low**:**

morphs of *Acleris cristana*

2) Different forces may result in **equilibrium**:

a) **Mutation/selection balance**

b) **Drift/mutation balance** - drift can cause fixation or loss of mutants, mutation introduces them. The **neutral theory of evolution**.
Drift is very slow in large populations, so polymorphisms often result.

c) Migration/selection balance, or spatial variation in selection
e.g. the peppered moth.

HERE WE WILL DEAL ONLY WITH:

(1a) **heterozygote advantage**

HETEROZYGOUS ADVANTAGE

Heterozygous advantage at a locus with 2 alleles leads to polymorphism:

AA Aa aa
-------> !!FITTER!! <-------

Why?

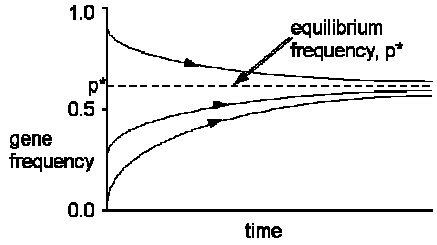
Suppose a population of AA, + a few Aa, so that p_A is almost 1
Aa will do better than AA, a will increase.

Population fixed for A is at an **unstable equilibrium**; can be invaded by **a**.

Similarly, when aa common, rare Aa will do better.
A will this time increase.

p_A=0 is also an unstable equilibrium

Between the unstable equilibria p=0 and p=1, there must be a **stable equilibrium**.



Evolution under heterozygous advantage

The stable equilibrium p^* is the gene frequency at which no change results. You can play with this on your own using **natural selection simulation programs**.

BUT we can go further than this simple verbal argument!

We can calculate this equilibrium (i.e. when there is no further evolutionary change); using the symbols as before:

$\Delta p = 0$ when $p' - p = 0$

(change in p) = 0 when (new p) - (old p) = 0

Genotype	AA	Aa	aa
Frequency	p^2	$2pq$	q^2
Fitnesses	$1-s$	1	$1-t$

(t and s are selection coefficients)

In the next generation, once again we count up the **A** and **a** alleles to get the frequency after selection,

$$p' = (\text{frequency AA}) + \frac{(\text{frequency Aa})}{2}$$

$$= \frac{p^2(1-s) + \frac{1}{2}2pq}{p^2(1-s) + 2pq + q^2(1-t)}$$

note: we divide by the total (or mean fitness) to get frequencies, as before

So:

$$\Delta p = 0 \text{ when } p' - p = \frac{p[p(1-s)+q]}{p^2(1-s) + 2pq + q^2(1-t)} - p = 0$$

at equilibrium

After a little manipulation†, this nasty looking piece of work can be shown to be equivalent to a much simpler form:

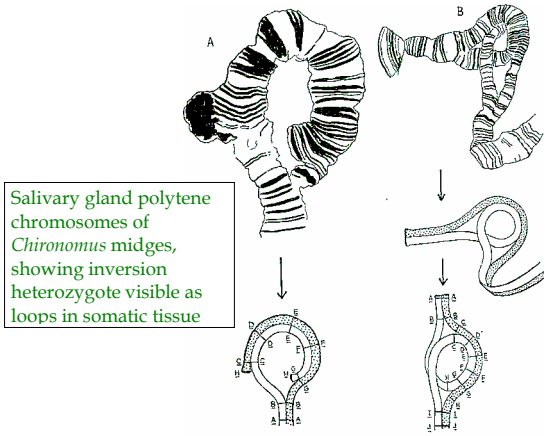
$$pq(-sp + tq) = 0$$

The analytical solution for the equilibrium is:

$$p^* = \frac{t}{s+t}$$

So, if we know the selection s, t, we can estimate the equilibrium frequency; or ... if we know an equilibrium frequency of a gene, we can estimate the relative strengths of selection coefficients.

† See: http://www.ucl.ac.uk/~ucbhdjm/courses/b242/MaintVar/proof.html



Salivary gland polytene chromosomes of *Chironomus* midges, showing inversion heterozygote visible as loops in somatic tissue

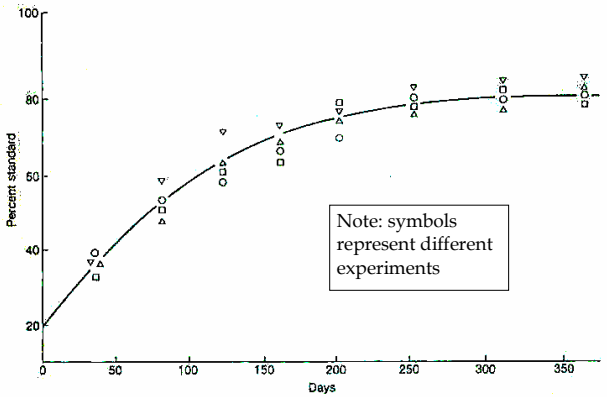


FIG. 4.6. Frequency of STANDARD chromosomes in a population cage experiment with Drosophila pseudoobscura. Data from four replicates are shown. The populations were started with 20 per cent of STANDARD and 80 per cent of CHIRICAHUA chromosomes. (After Dobzhansky 1951.)

What happens when the selection pressures s, and t are *negative*? In other words, if the homozygotes are **fitter** than heterozygotes

(i.e. ***heterozygote DISadvantage***).

AA Aa aa
<------ !!LESS FIT!! ------>

Here when we introduce a few **Aa** to a pure population of **AA**, the **AA** do better, so the **a** alleles are lost.

A population fixed for **A** is at a stable equilibrium; similarly, a population fixed for **a** is also stable.

Genotype	AA	Aa	aa
Frequency	p^2	$2pq$	q^2
Fitnesses	$1-s$ *more fit*	1 *less fit*	$1-t$ *more fit*

Is there an internal equilibrium? Yes, our formula for the equilibrium p^*

$$p^* = \frac{t}{s+t}$$

….. shows it is again in the possible range 0 to 1.

But the equilibrium only remains if the population starts at **exactly** that frequency.

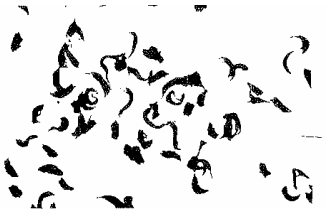
The **equilibrium** is ***unstable***, in which case the only ***stable equilibria*** are p=0 and p=1. e.g. ***hybrid inviability*** between species.

Now let's try a real example, in humans: malaria resistance.

ESTIMATING SELECTION

Based on observations of beta-haemoglobin **S** (sickle-cell) and **A** (normal) genotypes in a malaria-infested region of West Africa.

Blood phenotype of SS genotype:



Geographic distribution of Hb_β^S allele and of malaria

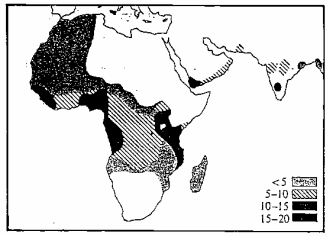


FIGURE 4.8
Frequency of the sickle-cell gene Hb_β^S in various parts of the Old World. The key in the right-hand corner shows the percentages of the populations that bear the gene. (After Allison, 1961.)

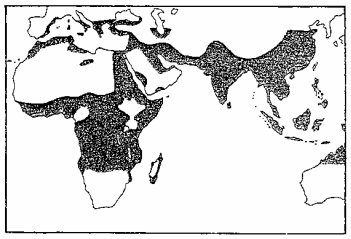


FIGURE 4.9
Distribution of falciparum malaria before 1930. (From M. F. Boyd's *Malariology*.)

Evidence for selection: correlation of geographic distribution of disease with distribution of S allele.

An assumption of Hardy-Weinberg equilibrium is no selection. If we can rule out inbreeding, migration or other factors, selection may be the cause.

From samples in SW Nigeria:

(Gilles et al. 1967)		Hb type	
	Sample size	A	AS
Children with severe *falciparum* malaria	100	96%	4%
Children with no malaria (control)	200	82%	18%

It is advantageous to have the S allele in SW Nigeria!

Here we use a χ^2 test to find whether there is evidence for deviation from Hardy-Weinberg in the population.

Testing whether data conforms to Hardy-Weinberg ratios:

Table of genotype frequency calculations:

a useful check!

Genotype	AA	AS	SS	TOTAL
O(bserved number)	25,374	5,482	67	30,923
Expected fraction	0.8266 ‡	0.1651 ‡	0.0082 ‡	0.9999*
E(xpected number)	(25,561.98)†	(5,106.03)†	(254.98)†	30,922.99*

Calculations of gene frequency:

$$p(A) = \frac{25374 + \frac{1}{2}5482}{30923} = 0.9092; \quad q(S) = 1 - p = 0.0908$$

... assuming Hardy-Weinberg.

‡ p_A^2 = 0.9092 ² and so on...
† NOT whole numbers!
* Calculator rounding errors

Statistical (χ^2) test:

Chi-square test: $\chi^2 = \sum \frac{(O-E)^2}{E}$

$$\chi^2 = \frac{(25374 - 25561.98)^2}{25561.98} + \frac{(5482 - 5106.03)^2}{5106.03} + \frac{(67 - 254.98)^2}{254.98}$$

χ^2 = 1.4 + 27.7 + 138.6 = 167.7

Degrees of freedom:

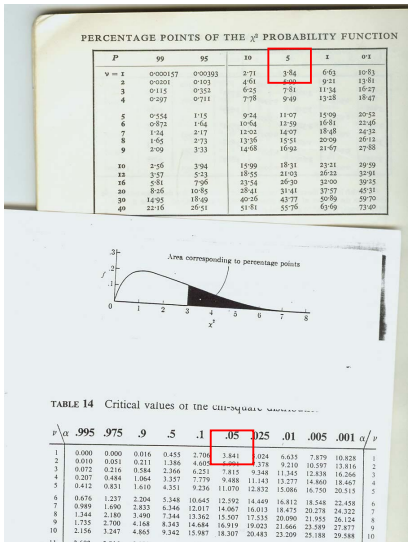
There are **3** classes of genotypes, and we lose degrees of freedom for any estimates we make from the data

We lose **1** because we obtain the total from the data
We lose **1** because we also estimate ***p*** from the data
... leaving **3** – **2** = **1** degree of freedom

Look for this value of χ^2 in your χ^2 tables, under **1** degree of freedom. You find that χ^2=167.7 greatly exceeds the value for *P*=0.001, which is χ^2=10.83.

Probability of getting a χ^2 this big in a large number of trials under the "null hypothesis" (i.e. Hardy-Weinberg ratios) is much less than one in a thousand.

Chi-square tables showing significance levels


PERCENTAGE POINTS OF THE χ^2 PROBABILITY FUNCTION

P	99	95	10	5	1	0·1
ν = 1	0·000157	0·00393	2·71	3·84	6·63	10·83
2	0·0201	0·103	4·61	5·99	9·21	13·81
3	0·115	0·352	6·25	7·81	11·34	16·27
4	0·297	0·711	7·78	9·49	13·28	18·47
5	0·554	1·15	9·24	11·07	15·09	20·52
6	0·872	1·64	10·64	12·59	16·81	22·46
7	1·24	2·17	12·02	14·07	18·48	24·32
8	1·65	2·73	13·36	15·51	20·09	26·12
9	2·09	3·33	14·68	16·92	21·67	27·88
10	2·56	3·94	15·99	18·31	23·21	29·59
12	3·57	5·23	18·55	21·03	26·22	32·91
16	5·81	7·96	23·54	26·30	32·00	39·25
20	8·26	10·85	28·41	31·41	37·57	45·31
30	14·95	18·49	40·26	43·77	50·89	59·70
40	22·16	26·51	51·81	55·76	63·69	73·40

TABLE 14 Critical values of the chi-square distribution

P \ α	.995	.975	.9	.5	.1	.05	.025	.01	.005	.001	α / P
1	0.000	0.000	0.016	0.455	2.706	3.841	5.024	6.635	7.879	10.828	1
2	0.010	0.051	0.211	1.386	4.605	5.991	7.378	9.210	10.597	13.816	2
3	0.072	0.216	0.584	2.366	6.251	7.815	9.348	11.345	12.838	16.266	3
4	0.207	0.484	1.064	3.357	7.779	9.488	11.143	13.277	14.860	18.467	4
5	0.412	0.831	1.610	4.351	9.236	11.070	12.832	15.086	16.750	20.515	5
6	0.676	1.237	2.204	5.348	10.645	12.592	14.449	16.812	18.548	22.458	6
7	0.989	1.690	2.833	6.346	12.017	14.067	16.013	18.475	20.278	24.322	7
8	1.344	2.180	3.490	7.344	13.362	15.507	17.535	20.090	21.955	26.124	8
9	1.735	2.700	4.168	8.343	14.684	16.919	19.023	21.666	23.589	27.877	9
10	2.156	3.247	4.865	9.342	15.987	18.307	20.483	23.209	25.188	29.588	10

Estimating selection coefficients:

If we know that the population is not inbreeding, and we suspect selection against homozygotes; and if we can also assume that the selection has caused all of the deviation from Hardy Weinberg; then...

We can estimate the selection based on the data just analysed:

Genotype	AA	AS	SS
Fitnesses, O/E	$\frac{25374}{25561.98}$	$\frac{5482}{5106.03}$	$\frac{67}{254.98}$
	0.99	1.07	0.26

now divide these through by by 1.07 to obtain ..

Relative fitnesses	W_{AA}	W_{AS}	W_{SS}
Values of fitnesses Relative to to AS:	0.92	1.00	0.24
Our selection model	1 - *s*	1	1 - *t*

Therefore estimated selection coefficients: s = 0.08, t = 0.76.

TAKE HOME POINTS

- Polymorphisms are common, and can be explained in a variety of ways:
  - Natural selection directly favours polymorphism
    - heterozygote advantage (done today)
    - frequency-dependent selection (Kevin Fowler's lecture)
  - A balance of evolutionary forces results in polymorphic equilibrium
    - equilibrium between mutation and selection
    - equilibrium between mutation and genetic drift (neutral evolution)
    - equilibrium between spatially varying selection and migration

- Fitnesses for heterozygote advantage at a single locus with two alleles ⇒ we can predict the polymorphic equilibrium frequency p^*
- If heterozygote advantage is suspected ⇒ can use deviations from Hardy-Weinberg to estimate selection

FURTHER READING

FUTUYMA, DJ 2005. Evolution. Chapter 9; Chapter 12 (pp. 274-285).