The reduction in fitness from genetic drift at heterotic loci in small populations

By ALAN ROBERTSON

Institute of Animal Genetics, Edinburgh

SUMMARY

In finite populations, loci maintained segregating by heterozygote superiority will be disturbed from their equilibrium positions by genetic sampling and the mean fitness of individuals will consequently be reduced. A general expression for this reduction is obtained for the segregation of two alleles. If the probability of continued segregation at the locus is high, the reduction tends to $1/4N$, irrespective of the strength of selection, where N is the effective population size. This will always be much less than the segregation load. If n alleles are segregating, so that all heterozygotes have the same fitness, the reduction tends to $(n-1)/4N$.

If, at a locus with two alleles, the heterozygote is fitter than either homozygote, segregation will be maintained in an infinite population. If the relative fitness of the A_1A_1, A_1A_2 and A_2A_2 genotypes are $1-s_1$, 1 and $1-s_2$ respectively, then at equilibrium the frequency $\bar{q}$ of the A_1 allele is given by $s_2/(s_1+s_2)$ and the mean fitness of individuals, $\overline{W}$, equals $1-s_1s_2/(s_1+s_2)$. The latter term has been called the 'segregation load' (Crow, 1958).

In a finite population, the gene frequency will not be at its equilibrium value because of the sampling which takes place each generation. If the population is very small this may lead to fixation, but the effect of sampling will otherwise be balanced by selection tending to return the frequency to its equilibrium value. As a consequence, the gene frequency will vary about this value over the generations and the average fitness will be rather less than the above expression. It will be shown in this note that if the chance of the locus being fixed is small, the expected reduction in fitness is close to $1/4N$, where N is the effective population size.

If the gene frequency is not equal to $\bar{q}$, the mean fitness of individuals in the population is given by

$$\overline{W} = 1 - s_1 s_2/(s_1+s_2) - (s_1+s_2)(q-\bar{q})^2.$$

The expected change in gene frequency in the next generation due to selection is given by

$$\Delta q = -(s_1+s_2)\bar{q}(1-\bar{q})(q-\bar{q}),$$

ignoring terms in $(q-\bar{q})^2$, etc.

Wright (1943) has shown that in such populations of size N, the gene frequency is of the form

$$\phi(q) \sim q^{4Nm\bar{q}-1}(1-q)^{4Nm(1-\bar{q})-1},$$

where $m = (s_1+s_2)\bar{q}(1-\bar{q})$. Such a distribution has mean $\bar{q}$ and variance

$$\bar{q}(1-\bar{q})/(4N(s_1+s_2)\bar{q}(1-\bar{q})+1).$$

The expected reduction in fitness due to variations in gene frequency about the equilibrium value is

$$\frac{(s_1+s_2)\,\bar{q}(1-\bar{q})}{4N(s_1+s_2)\,\bar{q}(1-\bar{q})+1} = \frac{L_0}{4NL_0+1}.$$

If $4NL_0$ is small, the reduction due to drift equals L_0, the infinite population segregation load. In fact, all populations will then be fixed. As $4NL_0$ increases, the reduction falls to a limiting value of $1/4N$, irrespective of the strength of the selection pressure. The variance in gene frequency equals $1/4N(s_1+s_2)$. At a given value of N, an increase in the selection pressure merely concentrates the gene frequencies around $\bar{q}$ but leaves the mean fitness unaffected.

What values of $4NL_0$ are to be expected for loci which remain segregating in finite populations because of heterozygote superiority? Robertson (1962) showed that, given s_1+s_2, the effect of selection in maintaining segregation depends critically on $\bar{q}$, being greatest when this is close to 0·5. From his Fig. 1. we see that, if $N = 1000$ and $s_1 = s_2 = s$, then for less than 10^{-6} of such loci to be fixed each generation s must be greater than 0·01. Since $L_0 = \frac{1}{2}s$, this critical value of $4NL_0$ is 20 and the standard deviation of gene frequency over populations is 0·11.

If $s_2 = 2s$, so that $\bar{q} = \frac{1}{3}$, s_1 must be greater than 0·02 for a fixation rate greater than 10^{-6} with $N = 1000$. At this value, $4NL_0$ is 52 and the standard deviation of gene frequency is 0·06. It therefore follows that if we are concerned with loci which remain segregating over long periods, $4NL_0$ will be much greater than unity and the reduction in fitness will be close to $1/4N$, irrespective of selection pressures. The reduction must then be expected always to be much smaller than the segregation load, since the ratio of the two is $1/(4NL_0+1)$.

A more general expression has been obtained for the situation with n alleles, when all heterozygotes have unit fitness and the homozygote for the mth allele has fitness $1-s_m$. At equilibrium in an infinite population, the frequency of the mth allele, $\bar{q}_m$, is equal to K/s_m where

$$1/K = \sum_{m=1}^{n} 1\bigg/s_m$$

and the mean fitness of a population not at equilibrium by

$$\overline{W} = 1-K- \sum_{m=1}^{n} s_m(q_m-\bar{q}_m)^2.$$

Selection then increases $\overline{W}$ by causing the gene frequencies to return towards their equilibrium values. Fisher's fundamental theorem of natural selection states that the increase in fitness each generation is V_a, the additive genetic variance in fitness, which here equals

$$2\left[K \sum_{m=1}^{n} s_m\,(q_m-\bar{q}_m)^2 + \sum_{m=1}^{n} s_m^2(q_m-\bar{q}_m)^3 - \left(\sum_{m=1}^{n} s_m(q_m-\bar{q}_m)^2\right)^2\right].$$

The final terms, being in higher powers of the deviations from equilibrium values, will be smaller than the first if the population is close to equilibrium and can be ignored. The gain by selection each generation is then a fraction $2K$ of the reduction due to deviations from equilibrium.

The expected decrease in fitness each generation due to sampling drift from the equilibrium point can be calculated since the expected increase in $(q_m-\bar{q}_m)^2$ is equal to

$$\bar{q}_m(1-\bar{q}_m)/2N.$$

The overall decrease in fitness each generation due to sampling is

$$\sum_{m=1}^{n} s_m \bar{q}_m \frac{1-\bar{q}_m}{2N} = \sum_{m=1}^{n} \frac{K(1-q_m)}{2N}, \quad \text{since } s_m \bar{q}_m = K$$
$$= K(n-1)/2N.$$

Equating the expectation of gain with that of loss, we find the expected value of the reduction due to drift is $(n-1)/4N$.

I am indebted to Dr B. Charlesworth for his comments on an earlier version of this paper.

REFERENCES

CROW, J. F. (1958). Some possibilities for measuring selection intensities in Man. *Hum. Biol.* **30**, 1–13.

ROBERTSON, A. (1962). Selection for heterozygotes in small populations. *Genetics* **47**, 1291–1300.

WRIGHT, S. (1943). Isolation by distance. *Genetics* **28**, 114–138.