INCREASE IN HETEROZYGOTE FREQUENCY WITH DIFFERENTIAL FERTILITY

A. F. PURSER

A.R.C. Animal Breeding Research Organisation, Edinburgh, 9

Received 20.x.65

1. INTRODUCTION

STUDIES of the effects of natural selection at an autosomal locus have usually assumed equal fertilities for all genotypes. In a full treatment of natural selection it appears necessary to distinguish among selection effects operating at different stages of the life cycle. The Hardy-Weinberg law for genotypic frequencies holds when there is selection among parents but not when selection acts at any other stage. Wallace (1958) has pointed out that differences in viability of individuals prior to observation can lead to a deviation from Hardy-Weinberg frequencies, the deviation being in either direction according to the relative viability of each genotype. In this paper it is shown that when differential fertility occurs the deviation is almost always in the same direction, namely, there is an excess of heterozygotes. This effect of differential fertility is related to that for other forms of selection and some of the consequences are discussed.

2. FERTILITY VARIATION IN ONLY ONE SEX

Consider a large random mating population in which the effect of random sampling of gametes is negligible. Let us suppose for a single locus with two alleles A and a that the three genotypes AA , Aa and aa have frequencies D , $2H$ and R respectively. Let us also assume that there is no variation of fertility in one sex (males, say) while the fertilities in the other sex (females) are x , y and z for genotypes AA , Aa and aa respectively.

If p denotes the frequency of gene A among the parents in any generation, the frequency of A gametes in the male gametic output will also be p . The female gametic output will be affected by the differential fertility, the frequency of A gametes being

$$p^* = \frac{x D + y H}{x D + 2y H + z R}$$

$$= p + \Delta, \text{ say.}$$

Writing $q = 1 - p$, the frequency of a gametes from females is correspondingly

$$q^* = \frac{y H + z R}{x D + 2y H + z R}$$

$$= q - \Delta.$$

The frequency of genotype AA in the next generation will be, under random mating,

$$D' = p p^* = p^2 + p\Delta,$$

and for the other genotypes,

$$2H' = pq^* + p^*q = 2pq - (p - q)\Delta \text{ and}$$

$$R' = q q^* = q^2 - q\Delta.$$

The frequency of gene A will be

$$p' = D' + H' = p + \frac{\Delta}{2} = \frac{p + p^*}{2}. \quad (1)$$

Deviations of genotypic frequencies from the Hardy-Weinberg proportions are as follows:

$$\left. \begin{aligned} D' - p'^2 &= p^2 + p\Delta - \left(p + \frac{\Delta}{2}\right)^2 = -\frac{\Delta^2}{4}, \\ 2H' - 2p'q' &= 2pq - (p - q)\Delta - 2\left(p + \frac{\Delta}{2}\right)\left(q - \frac{\Delta}{2}\right) = \frac{\Delta^2}{2}, \text{ and} \\ R' - q'^2 &= q^2 + q\Delta - \left(q - \frac{\Delta}{2}\right)^2 = -\frac{\Delta^2}{4}. \end{aligned} \right\} \quad (2)$$

Hence, there is an excess of heterozygotes to the extent of $\frac{1}{2}\Delta^2 = \frac{1}{2}(p^* - p)^2$, a quantity which cannot be negative, whatever the values of x , y and z , though it may be zero.

The condition that

$$\Delta = p^* - p = 0$$

implies

$$xD + yH = p (xD + 2yH + zR),$$

or,

$$\frac{xD}{z} + \frac{yH}{z} = p \left(\frac{xD}{z} + \frac{2yH}{z} + R \right).$$

If we assume D , H and R and hence p , known and that we are only interested in ratios of fertilities there are still two unknowns and only one equation. There are an infinite number of solutions but all except one (viz. $x = y = z$) are frequency dependent. Thus for any given D , H and R and any x , y and z it will be very rare that $\Delta = 0$.

3. FERTILITY VARIATION IN BOTH SEXES

The general problem involving fertility of each mating varying according to the genotypes mated is not amenable to general solution. The case where fertility of a mating is determined by the sum of the fertility effects from each sex gives virtually the same equations as in the previous case. The only difference is that x , y and z are each defined as the sum of the male and female fertilities of corresponding genotypes.

A slightly different result is obtained if it is assumed that the fertility of a mating is proportional to the product of fertilities for the genotypes in

that mating. Assuming fertilities of x_1, y_1 and z_1 for males and x_2, y_2 and z_2 for females of genotypes AA, Aa and aa respectively, the frequency of A gametes from males is

$$p^*_{1} = \frac{x_1 D + y_1 H}{x_1 D + 2y_1 H + z_1 R} = p + \Delta_1.$$

The frequency of A gametes from females is

$$p^*_{2} = \frac{x_2 D + y_2 H}{x_2 D + 2y_2 H + z_2 R} = p + \Delta_2.$$

Genotypic frequencies in the next generation are calculated as before and the deviations from Hardy-Weinberg equilibrium are then found to be

$$\begin{aligned} D' - p'^2 &= R' - q'^2 = -\frac{1}{4}(\Delta_1 - \Delta_2)^2, \text{ and} \\ 2H' - 2p'q' &= \frac{1}{2}(\Delta_1 - \Delta_2)^2. \end{aligned} \quad (3)$$

These results are in the same direction as before, though now the deviations can be zero if $\Delta_1 = \Delta_2$, *i.e.* if the fertility patterns are the same in both sexes. A corollary is that, for multiplicative effects between sexes, different fertility patterns in each sex leads to an excess of heterozygotes. Robertson (1965) has noted that any difference in gene frequency between males and females will lead to an excess of heterozygotes. However, a sex difference in gene frequency is not a necessary condition. If, for example, the fertility patterns are equal in each sex and fertility of a mating is the sum of the fertilities in each sex the result is the same as that in the previous section.

4. MULTIPLE ALLELES

The extension of these results to multiple alleles is straight forward. The following illustration is for the case of differential fertility in one sex only.

Suppose we have a series of alleles at one locus A_1, A_2, A_3 , etc., with frequency p_i for i^{th} allele A_i , where $\sum p_i = 1$. Suppose differential fertility gives rise to gametic frequencies p_i^* for allele A_i and

$$p_i^* = p_i + \Delta_i,$$

where

$$\sum \Delta_i = 0.$$

The frequency of genotype $A_i A_i$ in the next generation, say,

$$D'_i = p_i (p_i + \Delta_i) = p_i^2 + p_i \Delta_i$$

while the frequency of a heterozygote, say $A_i A_j$ is given by

$$2H'_{ij} = (p_i + \Delta_i)p_j + p_i(p_j + \Delta_j) = 2p_i p_j + \Delta_i p_j + \Delta_j p_i.$$

For the frequency of gene A_i we find, by summation,

$$p'_i = p_i + \frac{\Delta_i^2}{2}.$$

Hence,

$$D'_i - p_i'^2 = \frac{-\Delta_i^2}{4},$$

and similarly for homozygotes of all other alleles. For heterozygotes we have

$$\begin{aligned} 2H'_{ij} - 2p'_i p'_j &= 2p_i p_j + \Delta_i p_j + \Delta_j p_i - 2\left(p_i + \frac{\Delta_i}{2}\right)\left(p_j + \frac{\Delta_j}{2}\right) \\ &= \frac{-\Delta_i \Delta_j}{2}. \end{aligned}$$

Thus for any given heterozygote the deviation from Hardy-Weinberg proportions may be in either direction. However, since $\Sigma \Delta = 0$, the Δ 's are negatively correlated and in general there will be an excess for any given heterozygote. For the sum of all heterozygote frequencies involving one allele, say A_1 , there is always an excess, since,

$$\begin{aligned} \sum_{j, j \neq i} (2H'_{ij} - 2p'_i p'_j) &= \sum \left(-\frac{\Delta_i \Delta_j}{2} \right) = \frac{\Delta_i}{2} (\Delta_i - \Sigma \Delta) \\ &= \frac{\Delta_i^2}{2}. \end{aligned}$$

5. DISCUSSION

The Hardy-Weinberg law for genotypic frequencies within a population is only applicable under certain restrictive conditions. These are that the population is very large (effectively infinite), mating is at random and there are no differences among genotypes in viability or fertility. Here viability is thought of as survival of individuals from zygote formation to time of census. This may be contrasted with selection, in the Haldane (1924) sense, operating between census and mating times in the choice of individuals to be parents of the next generation. All these ideas may be unified by noting that selection can occur at one or more of three stages. These stages are, in biological order:

- (1) Fertility, or gamete production,
- (2) Viability of the zygotes (Wallace's (1958) "fitness"),
- (3) Choice of parents (Haldane's (1924) "selection").

The third form of selection leads to Hardy-Weinberg equilibrium in the next generation. The second form of selection, where there is differential viability or fitness, does lead to a deviation from Hardy-Weinberg frequencies, as pointed out by Wallace (1958). For example, we may redefine the x , y and z in section (2) of this paper as viabilities of corresponding genotypes. We find that the deviation from Hardy-Weinberg is still equal for both homozygotes and in particular

$$2H' - 2p'q' = 2p^2 q^2 (y^2 - xz) / (xp^2 + 2ypq + zq^2)^2.$$

The deviation can be in either direction depending on whether y^2 is greater or less than xz . As Lewontin and Cockerham (1959) have remarked, the deviation is only zero when the viability of the heterozygote is the geometric mean of the viabilities of the two homozygotes.

The first form of selection, differential fertility, as has been shown in this paper, will always lead to an excess of heterozygotes. The three cases may be summarised by saying that deviation from Hardy-Weinberg frequencies is (1) non-existent when selection is amongst individuals to be

parents, (2) is capable of being in either direction according to the relative magnitudes of the coefficients when there is differential survival, and (3) is in one direction only when there is differential fertility, the sign of the effect is unaltered by changing the coefficients.

The results given in this paper for variable fertility have only been proved for very large random mating populations. It is clear, though, that they will also apply to finite populations, not only through genetic differences in fertility but also in random sampling effects. In most natural populations family size varies and so for any finite population the average family size for each genotype in each sex will never be exactly equal to the overall mean. Thus, even if there are no real differences in fertility there will be a deviation from Hardy-Weinberg frequencies automatically in any finite population, provided family size varies. Constant family size may, however, be achieved (mainly in experimental situations) by counting only a fixed number of progeny for each mating. In this case there would be no deviation from Hardy-Weinberg frequencies.

In populations in which inbreeding is also taking place the deficiency in frequency of each homozygote, as a result of variable fertility, is in the opposite direction to that caused by inbreeding. If the level of inbreeding in a population were to be estimated from genotypic frequencies in such a situation then these estimates will be biased. The extent of this bias will depend on the variability of family size, both between and within genotypes, and on population size and gene frequency.

When gene action is considered another effect of variable fertility becomes apparent. There will be a shift in gene frequency from one generation to another as indicated in equation (1) and a corresponding change in mean genetic value. At the same time, for additive gene action, genetic variance will be decreased. Thus, for a population having $p^2 AA$ individuals, $2pq Aa$ individuals and $q^2 aa$ individuals, the additive genetic variance assuming no dominance is $2pqh^2$, where $2h$ is the gene substitution effect $AA-aa$. However, using equations (2) it is found that with differential fertility the genetic variance is $2(pq - \frac{1}{4}\Delta^2)h^2$. When gene action is non-additive it will still be true that the additive portion of genetic variance will be reduced by differential fertility. Furthermore for quantitative traits controlled by many loci the additive variance will be reduced in the same manner.

6. SUMMARY

It is demonstrated that differential fertility among genotypes always leads to a deficiency of homozygotes compared with the expected frequency under the Hardy-Weinberg law. Correspondingly there is an excess of heterozygotes, though for multi-allelic loci this only applies strictly to the total frequency of all heterozygotes. This result is compared to that for other forms of selection where either the deviation from Hardy-Weinberg frequencies can be in either direction or there is no deviation at all.

The effect of differential fertility will occur both as a result of real genetic differences in fertility or family size and also by random sampling variation in small populations.

Estimates of inbreeding level from genotypic frequencies will be biased to give values which are too low. Additive genetic variance will also be reduced from that expected with populations in Hardy-Weinberg equilibrium.

7. REFERENCES

- HALDANE, J. B. S. 1924. A mathematical theory of natural and artificial selection. I. *Trans. Camb. Phil. Soc.*, 23, 19-41.
- LEWONTIN, R. C., AND COCKERHAM, C. C. 1959. The goodness-of-fit test for detecting natural selection in random mating populations. *Evolution*, 13, 561-564.
- ROBERTSON, A. 1965. The interpretation of genotypic ratios in domestic animal populations. *Anim. Prod.*, 7, 319-324.
- WALLACE, B. 1958. The comparison of observed and calculated zygotic distributions. *Evolution*, 12, 113-115.