CORRELATIONS BETWEEN RELATIVES WHEN INTERMEDIATES ARE FITTEST

By J. W. James*

[*Manuscript received August* 13, 1965]

Summary

It is shown that natural selection for intermediates affects correlations between relatives, the effects being different for different sets of relatives. For additive genes, when heterozygote has superior fitness, half-sib correlations are slightly lower than the parent–offspring correlations. A simple approximate relation between the two is derived, and a similarity to effects of epistasis on selection response is discussed.

I. Introduction

Biometrical genetic theory is now well developed in many respects. In particular, for a population in genetic equilibrium the correlations between relatives can be expressed in terms of additive, dominance, and epistatic components of genetic variance for arbitrary forms of genetic determination of the character concerned (Fisher 1918; Kempthorne 1957). In addition, short-term responses to selection in such populations can be predicted (Griffing 1960; and later). However, although these results are valid for arbitrary gene action on the character analysed, they are subject to the very important qualification that natural selection must be assumed not to effect the genes concerned. This assumption is known to be often false.

Although some effects of natural selection on response to artificial selection have been discussed by Robertson (1956) and others, the influence of natural selection on correlations between relatives appears restricted to work by Penrose (1964) in a context unrelated to artificial selection. In view of the close relationship between selection response and correlations between relatives (Griffing 1960) it seems important that some attention should be given to this point. An attack on the problem with any degree of generality appears very difficult, so a simple situation has been analysed.

II. The Model

The case to be considered is that of a single locus with two alleles, A_1 and A_2, which affect a quantitative character in a population in equilibrium under random mating. The scale of measurement is chosen so that the unit is the phenotypic standard deviation and the homozygotes $A_1\,A_1$ and $A_2\,A_2$ have mean phenotypes u and $-u$ respectively. The mean phenotype of heterozygotes A_1A_2 is au, where a is a dominance parameter, being zero for additive gene action, ± 1 for complete dominance of either allele, and so on.

* School of Wool Technology, University of New South Wales, Kensington, N.S.W.

The frequencies of the alleles A_1 and A_2 are p and q $(= 1-p)$ respectively, so that at fertilization the genotypic frequencies are

$$p^2A_1A_1+2pqA_1A_2+q^2A_2A_2.$$

If no natural selection operated, the additive and dominance components of genetic variance would be

$$\sigma_A^2 = 2pqu^2[1+a(q-p)]^2,$$

$$\sigma_D^2 = 4p^2q^2a^2u^2.$$

For additive genes $(a = 0)$, $\sigma_D^2 = 0$ and $\sigma_A^2 = 2pqu^2$.

It is, however, assumed that the heterozygote is superior in fitness to both homozygotes, and that the fitnesses of A_1A_1 and A_2A_2 relative to that of A_1A_2 are $1-K$ and $1-L$ respectively. Since the population is in equilibrium,

$$p = L/(K+L),$$

and

$$q = K/(K+L),$$

while the mean fitness of the population is $1-S$, where

$$S = KL/(K+L).$$

S is the "segregational genetic load" at this locus.

If natural selection acts through juvenile mortality, and measurements of the character are made on surviving adults, the genotypic proportions among measured individuals are

$$p^2[(1-K)/(1-S)]A_1A_1+2pq/(1-S)A_1A_2+q^2[(1-L)/(1-S)]A_2A_2.$$

For additive gene action the genetic variance is

$$2pqu^2[(1-2S)/(1-S)],$$

so that by lowering the frequencies of extreme genotypes, natural selection reduces the observed genetic variation.

It should be noted that there is a difference in principle as well as notation between this model and that of Penrose (1964). Penrose assumed that frequencies of measured genotypes among progeny were

$$p^2A_1A_1+2pqA_1A_2+q^2A_2A_2,$$

taking differential fertility rather than differential survival as the mode of action of natural selection.

III. Covariances between Relatives

The most widely used methods of estimating genetic variance, at least in animals, are based on half-sib correlations and the regression of offspring on parent. Attention is here confined to the effects of natural selection on estimates derived from these two sources.

It may readily be verified that under the above assumptions the frequencies of observed parent–offspring genotypic combinations are as given in Table 1. This differs from Table 3 of Penrose (1964) for the reason mentioned above. For this model the genotypic distribution is the same for both generations, whereas Penrose's model leads to a lower variance among parents than among progeny.

TABLE 1

FREQUENCIES OF OBSERVED PARENT–OFFSPRING GENOTYPIC COMBINATIONS

Parent	Progeny		
	A_1A_1	A_1A_2	A_2A_2
A_1A_1	$p^3(1-K)^2/(1-S)^2$	$p^2q(1-K)/(1-S)^2$	
A_1A_2	$p^2q(1-K)/(1-S)^2$	$pq/(1-S)^2$	$pq^2(1-L)/(1-S)^2$
A_2A_2		$pq^2(1-L)/(1-S)^2$	$q^3(1-L)^2/(1-S)^2$

It may be confirmed from Table 1 that the mean number of progeny per parent is the same for each genotype and that the mean phenotype for both generations is

$$m = u[(p-q)+2pq/(1-S)a].$$

If we now define the two quantities α and β as

$$\alpha = [u/(1-S)][1-2S+(q-p)a],$$

and

$$\beta = [u/(1-S)]\{1-2S+[(q-p)a/(1-S)]\},$$

it is only a matter of working through the algebra to show that the covariance of half-sibs is given by

$$\text{cov(HS)} = \tfrac{1}{2}pq\alpha^2[(1-2S)/(1-S)],$$

while the covariance of parent and offspring is given by

$$\text{cov(PO)} = pq\alpha\beta.$$

In particular for additive gene action it is readily seen that

$$\text{cov(HS)} = \tfrac{1}{2}pqu^2[(1-2S)/(1-S)]^3,$$

$$\text{cov(PO)} = pqu^2[(1-2S)/(1-S)]^2.$$

Natural selection reduces both covariances below the values they would have in its absence, the effect being greater on cov(HS) than on cov(PO). In the absence of dominance the ratio of covariance is

$$\text{cov(HS)/cov(PO)} = (1-2S)/2(1-S),$$

instead of the value $\frac{1}{2}$ in the absence of natural selection. If S is small this ratio is very nearly $\frac{1}{2}(1-S)$. It may be verified from Table 3 of Penrose (1964) that for his model this ratio is $\frac{1}{2}$, being unaffected by natural selection.

The situation is not quite so simple when there is dominance on the metric trait. For example both cov(HS) and cov(PO) may be zero when $\alpha = 0$. This occurs when

$$a = (1-2S)/(p-q),$$

and since $S \leqslant p \leqslant 1-S$, this requires $|a| > 1$, or that there must be overdominance on the metric trait. If attention is confined to cases where $|a| \leqslant 1$ so that heterozygotes are phenotypically intermediate, $\alpha \neq 0$. It is then still possible that cov(PO) may be negative while cov(HS) is positive, since β may be negative. For this to occur, assuming a is positive, it is necessary that $a > 1-S$ and that p lie in the range $(1-S)(1-\frac{1}{2}S) < p < 1-S$. Thus cov(PO) may be negative if the allele A_1 is nearly dominant on the metric scale, while the genotype A_2A_2 is very nearly lethal.

For an arbitrary degree of dominance the ratio of the covariances is

$$\text{cov(HS)}/\text{cov(PO)} = \tfrac{1}{2}\left([1-2S+(q-p)a]/[1-S+(q-p)a/(1-2S)]\right),$$

so that if a or $(q-p)$ or both factors are small, the ratio will be very nearly $\frac{1}{2}(1-S)$, as for strictly additive gene action.

If there is no epistasis a similar result may be obtained by summing covariances over loci to get

$$\text{cov(HS)} \simeq \tfrac{1}{2}(1-\bar{S})\text{cov(PO)}$$

where $\bar{S}$ is Robertson's (1956) "coefficient of homeostatic strength", defined as the weighted average of S values at the several loci, the weights being the additive genetic variance components at the loci.

For the approximation to be poor at any locus we need both a and $(q-p)$ to be large. But when $(q-p)$ is large either p or q is small, and the locus contributes little to the total variation unless the effect of the locus is disproportionately large. Hence the average result may not be unreasonable.

IV. Response to Selection

The most important practical application of estimates of covariances between relatives is the prediction of response to artificial selection. The question therefore arises as to which covariance gives the more accurate prediction.

For truncation selection with selection differential i it can be shown that the change in frequency of the allele A, among selected parents, is

$$\delta = ipq\{\beta+[2Spau/(1-S)^2]\},$$

and that the genetic gain in the mean of the measured progeny, G, is given by

$$G = 2\alpha\delta,$$

neglecting terms in δ^2, and hence

$$G = 2\,i\,\text{cov(PO)} + 4\,ip^2qau/(1-S)^2.$$

In particular for additive gene action we have

$$G = 2\,i\,\text{cov(PO)}.$$

The error in using $G = 2\,i\,\mathrm{cov(PO)}$ will only be appreciable if there are marked deviations from additivity. This result may also be extended to several loci if there is no epistasis.

In short, then, for additive or nearly additive genes,

$$\mathrm{cov(HS)} \simeq \tfrac{1}{2}(1-\bar{S})\mathrm{cov(PO)},$$

and

$$G \simeq 2\,i\,\mathrm{cov(PO)}.$$

V. Discussion

One striking point about these results is the way in which natural selection for intermediates resembles an important additive × additive component of genetic variance in its effects on these correlations between relatives. Considering only two-locus epistasis in the absence of natural selection,

$$\mathrm{cov(PO)} = \tfrac{1}{2}\sigma_A^2 + \tfrac{1}{4}\sigma_{AA}^2,$$

$$\mathrm{cov(HS)} = \tfrac{1}{4}\sigma_A^2 + \tfrac{1}{16}\sigma_{AA}^2.$$

Thus the effects of a component σ_{AA}^2 of about one-fifth of the genetic variance and a value of $\bar{S}$ of about 0·05 would be very similar with respect to these covariances. It had already been pointed out by Griffing (1960) that the presence of an additive × additive component would give similar results to opposing natural selection following relaxation of artificial selection. The similarity may be extended a little in the light of the present results.

For the additive model with natural selection the result of relaxation of selection is a loss of a fraction $\bar{S}$ of the response obtained. This was shown by Robertson (1956) for the case where measurements are taken before selection acts, and may be obtained for the present model in a similar manner. Thus G' the response after one generation of selection and one of relaxation is

$$G' \simeq 2i\mathrm{cov(PO)}(1-\bar{S})$$

$$\simeq 4i\mathrm{cov(HS)}.$$

It was shown by Griffing (1960) that for two-locus epistasis without natural selection the loss from relaxation was $\frac{1}{2}iy\sigma_{AA}^2$, where y is the recombination fraction. For unlinked loci, $y = \frac{1}{2}$ and for this model also

$$G' = 4i\mathrm{cov(HS)}.$$

This emphasizes further the difficulty of distinguishing between the two models without fitness data.

In a recent study Bradford and Van Vleck (1964) estimated heritability of milk yield in cattle as 0·43±0·03 by daughter–dam regression and 0·25±0·05 by paternal half-sib correlations. They considered a number of factors which could have contributed to this difference, including an important additive × additive component of genetic variance, but not the possibility of greater fitness of intermediates. There

is no evidence suggesting the plausibility of this hypothesis, but it may bear consideration.

Again, Clayton, Morris, and Robertson (1957) estimated heritability of abdominal bristle number in *Drosophila melanogaster* by the two methods as: half-sib correlation, $0{\cdot}48 \pm 0{\cdot}11$; parent–offspring regression, $0{\cdot}51 \pm 0{\cdot}07$. These estimates do not differ significantly, but if the difference were genuine it would correspond in our model to a value of $\bar{S}$ of 0·06. For this population Latter and Robertson (1962) estimated $\bar{S}$ as 0·04.

It has already been seen that the present model gives different results from that of Penrose (1964) for an additive character. If the character considered is fitness itself, it follows from Penrose's model that cov(HS) = cov(PO) = 0. On the other hand, for the model presented here

$$\text{cov(HS)} = [(p-q)^2 S^4(1-2S)]/2pq(1-S)^3,$$

$$\text{cov(PO)} = [(p-q)^2 S^4(3-2S)]/2pq(1-S)^3,$$

and

$$\text{cov(HS)/cov(PO)} = [(1-2S)/(3-2S),$$

or roughly $\frac{1}{3}(1-S)$. This further emphasizes the importance of the difference between the two models.

In general any relationships between fitness and phenotype may be expected to affect correlations between relatives, the nature of the effect being dependent on the connection between phenotype and fitness and also on which relatives are concerned. It will also depend on the stages of the life cycle at which differential fitness occurs and at which measurement is made. These complications appear to make any general treatment of the problem extremely difficult.

VI. References

Bradford, G. E., and Van Vleck, L. D. (1964).—Heritability in relation to selection differential in cattle. *Genetics* **49**: 819–27.

Clayton, G. A., Morris, J. A., and Robertson, A. (1957).—An experimental check on quantitative genetical theory. I. Short-term responses to selection. *J. Genet.* **55**: 131–51.

Fisher, R. A. (1918).—Correlation between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edinb.* **52**: 399.

Griffing, B. (1960).—Theoretical consequences of truncation selection based on the individual phenotype. *Aust. J. Biol. Sci.* **13**: 307–43.

Kempthorne, O. (1957).—"An Introduction to Genetic Statistics." (John Wiley and Sons, Inc.: New York.)

Latter, B. D. H., and Robertson, A. (1962).—The effects of inbreeding and artificial selection on reproductive fitness. *Genet. Res.* **3**: 110–38.

Penrose, L. S. (1964).—Some formal consequences of genes in stable equilibrium. *Ann. Hum. Genet.* **28**: 159–66.

Robertson, A. (1956).—The effect of selection against extreme deviants based on deviation or on homozygosis. *J. Genet.* **54**: 236–48.