NON-ALLELIC INTERACTION IN CONTINUOUS VARIATION OF RANDOMLY BREEDING POPULATIONS

KENNETH MATHER

Department of Genetics, University of Birmingham

SUMMARY

The consequences are considered of digenic interactions for the variances and covariances commonly estimated from randomly breeding populations. The interaction affects the constitutions of D_R and H_R as well as introducing new quadratic terms by which, in principle, its presence can be detected. In practice the chief consequence of interaction is likely to be to alter the apparent values of D_R and H_R as estimated from the variances and covariances.

1. COMPONENTS OF VARIATION

In the descendants of the cross between true breeding lines, or its equivalent, it has been shown that non-allelic interaction has a two-fold effect on the genetical components of continuous variation (see Mather and Jinks, 1971, whose notation is used here). In the first place the terms in $D = S(d^2)$ and $H = S(h^2)$, measuring variation due to additive and dominance effects of the genes, have added to them terms in $I = S(i^2)$, $J = S(j^2)$ and $L = S(l^2)$, which depend on hom × hom, hom × het and het × het interactions respectively. In first rank variances and covariances these terms appear with coefficients which are the products of the coefficients of the relevant main terms, *i.e.* I takes a coefficient which is the square of that of D, J one which is the product of the D and H coefficients and L one which is the square of that of H. Thus the genetical component of variation in the F_2 is

$$V_{1F2} = \tfrac{1}{2}D + \tfrac{1}{4}H + \tfrac{1}{4}I + \tfrac{1}{8}J + \tfrac{1}{16}L$$

the variance of F_3 means is

$$V_{1F3} = \tfrac{1}{2}D + \tfrac{1}{16}H + \tfrac{1}{4}I + \tfrac{1}{32}J + \tfrac{1}{256}L$$

and the covariance of F_2 parent and F_3 mean is

$$W_{1F23} = \tfrac{1}{2}D + \tfrac{1}{8}H + \tfrac{1}{4}I + \tfrac{1}{16}J + \tfrac{1}{64}L.$$

This relation, however, no longer holds with variances and covariances of higher rank, the mean variance of F_3 families, for example, being

$$V_{2F3} = \tfrac{1}{4}D + \tfrac{1}{8}H + \tfrac{5}{16}I + \tfrac{7}{64}J + \tfrac{1}{32}L.$$

The second effect of this interaction is to change the definitions of D and H. In the F_2 they become $D = S_a (d_a + \tfrac{1}{2} S_b j_{ab})^2$ and $H = S_a (h_a + \tfrac{1}{2} S_b l_{ab})^2$ where $S_a j_{ab}$ indicates the sum of $j_{ab}, j_{ac} \cdots j_{ak}$ and $S_a l_{ab}$ is similarly $l_{ab} + l_{ac} \cdots + l_{ak}$ where there are k genes in the system. These definitions are not, however, invariable: they change with the generation, becoming

for example $D = S_a (d_a + \frac{1}{4} S_b j_{ab})^2$ and $H = S_a (h_a + \frac{1}{4} S_b l_{ab})^2$ in both the variances of F_3. Thus the contribution of the various genes to the additive and dominance variation are no longer absolute, in the forms of d^2 and h^2, but have become conditional on the interactions they display, the degree to which they are affected by these interactions depending on the genetical structure of the population and in particular on the proportion of homozygotes and heterozygotes it carries.

Even in the absence of interaction, the additive and dominance components of variation, D_R and H_R, are more complex in randomly breeding populations where the gene frequencies are in general unequal (Mather and Jinks, *loc. cit.*) Not only do their values depend on the gene frequencies, but the so-called additive variation includes a component from dominance, with

$$D_R = S_a \{4u_a v_a [d_a - (u_a - v_a)h_a]^2\} \text{ and } H_R = S_a [16u_a^2 v_a^2 h_a^2]$$

which, of course, reduce to the standard forms $D = S_a (d_a^2)$ and $H = S_a (h_a^2)$ when $u = v$ for all genes, as in F_2 and the generations derived from it.

Non-allelic interaction again has a two-fold effect on the components of variation in a randomly breeding population, though not surprisingly these are more complex than in an F_2 and its descendants. These effects, whose derivation is algebraically tedious rather than difficult, are set out in table 1, which gives the compositions in terms of D_R, H_R, I_R, J_R and L_R of different variances and covariances commonly calculated from observations in populations, and table 2 which gives the definitions of D_R, H_R, I_R, J_R and L_R. The two-gene case was worked out and gave as its components

$$D_R = 4u_a v_a\{[d_a + 2u_b v_b j_{ab} + (u_b - v_b)i_{ab}] - (u_a - v_a)[h_a + (u_b - v_b)j_{ba} + 2u_b v_b l_{ab}]\}^2$$
$$+ 4u_b v_b\{[d_b + 2u_a v_a j_{ba} + (u_a - v_a)i_{ab}] - (u_b - v_b)[h_b + (u_a - v_a)j_{ab} + 2u_a v_a l_{ab}]\}^2$$

$$H_R = 16u_a^2 v_a^2 [h_a + (u_b - v_b)j_{ba} + 2u_b v_b l_{ab}]^2 + 16u_b^2 v_b^2[h_b + (u_a - v_a)j_{ab} + 2u_a v_a l_{ab}]^2$$

$$I_R = 16u_a v_a u_b v_b [i_{ab} - (u_b - v_b)j_{ab} - (u_a - v_a)j_{ba} + (u_a - v_a)(u_b - v_b)l_{ab}]^2$$

$$J_R = 64u_a v_a u_b^2 v_b^2 [j_{ab} - (u_a - v_a)l_{ab}]^2 + 64u_a^2 v_a^2 u_b v_b [j_{ba} - (u_b - v_b)l_{ab}]^2$$

$$L_R = 256u_a^2 v_a^2 u_b^2 v_b^2 l_{ab}^2.$$

These were then generalised to give the definitions shown in table 2. This table also sets out the components in F_2 (and in the S_3 generation derived from it by random pair matings) which is the special case where all gene frequencies are equal. These F_2 components agree of course with those

TABLE 1

The structure of variances and covariances in randomly breeding populations

Statistic	Structure
Variance of population	$V_R = \frac{1}{2}D_R + \frac{1}{4}H_R + \frac{1}{4}I_R + \frac{1}{8}J_R + \frac{1}{16}L_R + E$
Covariance of parent and offspring	$W_{PO} = \frac{1}{4}D_R + \frac{1}{16}I_R$
Covariance of full-sibs	$W_{FS} = \frac{1}{4}D_R + \frac{1}{16}H_R + \frac{1}{16}I_R + \frac{1}{64}J_R + \frac{1}{256}L_R$
Covariance of half-sibs	$W_{HS} = \frac{1}{8}D_R + \frac{1}{64}I_R$
Variance of sibship means	$V_{1SR} = \frac{1}{4}D_R + \frac{1}{16}H_R + \frac{1}{16}I_R + \frac{1}{64}J_R + \frac{1}{256}L_R + E_b$
Mean variance of sibships	$V_{2SR} = \frac{1}{4}D_R + \frac{3}{16}H_R + \frac{3}{16}I_R + \frac{7}{64}J_R + \frac{15}{256}L_R + E_w$

The contributions from non-heritable variation are denoted by E, E_w and E_b. E_w is the non-heritable variance within a sibship family while E_b reflects the effects of any common factors in the environments of members of the same family (the common family environment). E will be compounded of E_w and E_b, its composition depending on the average size of the sibships of which the population is made up.

derived directly from the F_2 as listed by Mather and Jinks. It should be noted too that the components as they appear in a randomly breeding population are the same as those Mather and Jinks (Table 96) record for the two-gene case in diallel crosses.

The basic structures of the variances and covariances set out in table 1 are the same as for those derivable from an F_2 and its S_3 generation, as indeed is to be expected since the latter are but the special case of the former where all u = all v = $\frac{1}{2}$. The relations of the coefficients of I, J and L to those of D and H which we have already noted in first rank variances and covariances from F_2 and its derivatives also hold with randomly breeding populations; and again as before they are lost in variances and covariances of higher rank. We may note too that any variance or covariance, of what-

TABLE 2

Definitions of D_R, H_R, I_R, J_R and L_R in a randomly breeding population. The special case of F_2, where all gene frequencies are equal, is included for comparison

Component	Randomly breeding population	F_2
D_R	$S_a 4u_a v_a \{[d_a + 2S_b(u_b v_b j_{ab}) + S_b(\overline{u_b - v_b i_{ab}})] - (u_a - v_a)$ $[h_a + S_b(u_b - v_b j_{ba}) + 2S_b(u_b v_b l_{ab})]\}^2$	$S_a(d_a + \frac{1}{2}S_a j_{ab})^2$
H_R	$S_a 16u_a^2 v_a^2[h_a + S_b(\overline{u_b - v_b j_{ba}}) + 2S_b(u_b v_b l_{ab})]^2$	$S_a(h_a + \frac{1}{2}S_b l_{ab})^2$
I_R	$S_{ab} 16u_a v_a u_b v_b[i_{ab} - (u_b - v_b)j_{ab} - (u_a - v_a)j_{ba} + (u_a - v_a)(u_b - v_b)l_{ab}]^2$	$S_{ab}(i_{ab}^2)$
J_R	$S_{ab} 64\{u_a v_a u_b^2 v_b^2[j_{ab} - (u_a - v_a)l_{ab}]^2 + u_a^2 v_a^2 u_b v_b[j_{ba} - (u_b - v_b)l_{ab}]^2\}$	$S_{ab}(j_{ab}^2 + j_{ba}^2)$
L_R	$S_{ab} 256u_a^2 v_a^2 u_b^2 v_b^2 l_{ab}^2$	$S_{ab}(l_{ab}^2)$

S_a indicates summation over all genes.
S_b indicates summation over all genes interacting with A-a.
S_{ab} indicates summation over all pairs of interacting genes.

ever rank, which carries a term in D_R also carries one in I_R, while J_R and L_R also appear wherever H_R enters in. It is thus not easy to separate I_R from D_R or J_R from H_R. In principle the difference between the parent-offspring covariance and twice that of half-sibs offers a means of estimating I_R, since

$$W_{PO} - 2W_{HS} = (\tfrac{1}{4}D_R + \tfrac{1}{16}I_R) - 2(\tfrac{1}{8}D_R + \tfrac{1}{64}I_R) = \tfrac{1}{32}I_R;$$

but this is hardly likely to be a sensitive test unless the two covariances are known with some precision and there is no comparable disturbance from such other factors as maternal effects, assortative mating or selection.

$H_R + \tfrac{1}{4}J_R + \tfrac{1}{16}L_R$ can be found as the difference between W_{PO} and W_{FS}, the covariance of full-sibs, but this comparison does not separate J_R and L_R from H_R. Again in principle this might be approached by finding

$$V_{2SR} - V_{1SR} - 2W_{FS} - 2W_{PO} + 8W_{HS} = \tfrac{1}{16}J_R + \tfrac{3}{64}L_R + (E_w - E_b).$$

Then if we have an external means of evaluating $E_w - E_b$ (such as could be provided by a suitably designed experiment) an estimate of $J_R + \tfrac{3}{4}L_R$ could be obtained. $V_R - 2V_{1SR}$ or $2V_{2SR} - V_R$ could be used in place of $V_{2SR} - V_{1SR}$ or W_{FS} could be substituted for V_{1SR}; but the difficulties would not be overcome as there would always remain the need to estimate an appropriate

non-heritable component before $\jmath_R + \frac{3}{4}L_R$ could be estimated. Furthermore as with the estimate of I_R, any of these comparisons would be uninformative unless the values of the variances and covariances were known with some 'precision and there were no sources of comparable disturbance.

Where the population can be investigated experimentally there is the prospect of coping with the problems raised by non-heritable variation. At the same time the situation may not be eased materially in other respects. Thus the North Carolina designs I and II yield statistics which raise the same difficulties in separating the interactive from the main components of variation. Kearsey and Jinks (1968) triple test cross can, however, give more direct information about the interactions, since if the appropriate lines are available for making the test crosses, in the two gene case the mean of their $\bar{L}_1 + \bar{L}_2 - 2\bar{L}_3$ comparisons yields an estimate of

$$\tfrac{1}{2} S_{ab} \left[i_{ab} - (u_b - v_b) j_{ab} - (u_a - v_a) j_{ba} + (u_a - v_a)(u_b - v_b) l_{ab} \right]$$

and the variance of these comparisons gives an estimate of

$$\tfrac{1}{2} S_{ab} \left\{ u_b v_b \left[j_{ab} - (u_a - v_a) l_{ab} \right]^2 + u_a v_a \left[j_{ba} - (u_b - v_b) l_{ab} \right]^2 + u_a v_a u_b v_b l^2_{ab} \right\}.$$

If experimental inbreeding and subsequent crossing can be practised further possibilities are of course opened up, including the use of comparisons among means in addition to a wider range of comparisons among second degree statistics.

2. EFFECTS OF SELECTION

Since I_R appears with D_R even though not always with the same balance of coefficients, and $\jmath_R$ and L_R similarly appear with H_R, their effects will be to alter the estimates of D_R and H_R obtained from the data if these do not allow isolation of the interactions. The interactions will also, as we have seen in table 2, affect the values of D_R and H_R by internal modification of these components. The precise alterations they will make in this way will depend on the magnitudes and signs of the various i, j and l, and also on the relevant gene frequencies. Two particular cases are worthy of mention. It has been argued by Mather (1960) and by Breese and Mather (1960) that stabilising selection for a character will tend to produce ambidirectional dominance and interaction, if it produces interaction at all, while directional selection will tend to produce unidirectional dominance and duplicate type interaction (see Mather, 1973). With a character under stabilising selection, measured on an appropriate scale, we would expect $\bar{h}, i, j$ and $\bar{l}$, averaged over all genes in the system to tend towards 0; but with a character under directional selection $\bar{h}$ would tend to be positive and i, j and l all negative (Mather and Jinks, 1971). Furthermore since under stabilising selection neither the $+$ allele nor the $-$ allele at any locus has an unconditional advantage over the other, we might expect the average value of $u - v$ taken over all genes to tend to 0 also. With directional selections where the $+$ allele has a general advantage over its $-$ fellow, we might expect the average of $u - v$ taken over all genes to be positive.

Applying these expectations to the composition of D_R and H_R set out in table 2, stabilising selection would give a situation where the contribution of h, i, j and l to D_R would tend to vanish and leave D_R as depending chiefly, indeed almost entirely, on $S_a(d_a^2)$ though reduced somewhat in overall value by the fact that even where, with unequal u and v, $\bar{u}$ and $\bar{v}$ tend to $\frac{1}{2}$, the

average of $4uv$ will still be less than 1. In the same way H_R would be largely a measure of S_a (h^2) though biassed downwards by the mean of $16u^2v^2$ being less than 1.

With duplicate interaction on the other hand, the terms in S_a $(\overline{u-vj})$ and 2 S_b (uvl) would be preponderantly negative and so reduce the value of H_R while 2 S_b (uvj) and S_b $(\overline{u-vi})$ would similarly tend to reduce the value of D_R. True the effect on the dominance portion included in D_R would tend to be positive and so to raise the value of D_R itself, but with a coefficient of $(u-v)$ is hardly likely to offset the reduction springing from the d part of D_R. While the ascertainable values of D_R and H_R would thus tend to be lowered, I_R and J_R would tend to be raised and insofar as they were confounded with D_R and H_R respectively in estimation they would tend to offset the reductions in D_R and H_R themselves; but the coefficients of I_R and J_R are so much lower than those of D_R and H_R in the various statistics (table 1) that the overall effect should still be a reduction in the apparent values of D_R and H_R. We should note too that the interactions included in D_R and H_R are summed over all interacting genes in their effects on the contributions of any pair of alleles to the additive and dominance variation, with the consequence that the more interacting pairs of genes there are in the system the less the magnitude of the interaction between any pair need be to achieve a given reduction. In I_R and J_R the interactions do not sum in the same way.

With duplicate interaction reducing the apparent additive and dominance components in this way, we would expect that, other things being equal, the genetical contribution to the variance of the character in the population would be lower and its heritability to be consequently smaller. Overall therefore, so-called fitness characters, which must be under predominantly directional selection, should have lower heritabilities than non-fitness characters, which must experience a corresponding predominance of stabilising selection. These expectations accord with the heritabilities recorded by Falconer (1960, table 10.1) for various characters in a number of domestic and laboratory animals.

3. References

BREESE, E. L., AND MATHER, K. 1960. The organisation of polygenic activity within a chromosome in *Drosophila*. II. Viability. *Heredity*, *14*, 375-400.

FALCONER, D. S. 1960. *Introduction to Quantitative Genetics*. Oliver and Boyd, Edinburgh.

KEARSEY, M. J., AND JINKS, J. L. 1968. A general method of detecting additive, dominance and epistatic variation for metrical traits. I. Theory. *Heredity*, *23*, 403-409.

MATHER, K. 1960. Evolution in polygenic systems. *Int. Colloquium on Evolution and Genetics*, pp. 131-152. Acad. Naz. dei Lincei, Rome.

MATHER, K. 1973. *Genetical Structure of Populations*. Chapman and Hall, London.

MATHER, K., AND JINKS, J. L. 1971. *Biometrical Genetics*, Second ed. Chapman and Hall, London.