AN EXTENSION OF THE CONCEPT OF PARTITIONING HEREDITARY VARIANCE FOR ANALYSIS OF COVARIANCES AMONG RELATIVES WHEN EPISTASIS IS PRESENT[1]*

C. CLARK COCKERHAM

North Carolina State College, Raleigh

I T is convenient, for purposes of description and analysis, to consider the phenotypic expression of a characteristic as a sum of an hereditary or genotypic value and of an environmental value. If the actual joint results deviate from this linear description (i.e., if interaction effects exist) the breeder or geneticist must exercise caution in extrapolating from his results because in this case the hereditary and environmental values are defined specifically in terms of each other (NELDER 1950). For some characteristics a transformation of scale may help in coming closer to additivity (WRIGHT 1950).

With this linear description, the total or phenotypic variance may be considered to consist of the hereditary, environmental and interaction variances, and also of covariance terms if the components are correlated in their occurrence. The covariance between heredities and environments is often a troublesome feature in human and livestock populations. For example, where dairy cattle are fed in proportion to their production, the better genotypes are provided better environments. However, in designed experiments correlation in occurrence can, for the most part, be avoided by randomization devices.

FISHER (1918) partitioned the phenotypic variance further by subdividing the hereditary variance into an additive portion resulting from average effects of genes, a portion resulting from dominance effects (allelic interactions) of genes and a portion resulting from epistatic effects (non-allelic interactions) of genes. FISHER showed the distribution of the additive and dominance portions in correlations between various relatives in a randomly mating population. The present paper shows the subdivision of the epistatic variance into components and gives the distribution of these epistatic components in the covariances or correlations between relatives.

PARTITIONING THE HEREDITARY VARIANCE

The partitioning of the hereditary variance of diploid organisms that have no multiple alleles can be illustrated by considering two loci, each with two

alleles (A, a and B, b). If the coupling and repulsion double heterozygotes are identical phenotypically, 9 genetic types, as in table 1, are possible. The Y's and f's in table 1 are the hereditary values and relative frequencies of the indicated types, respectively. The hereditary values are actually phenotypic values averaged over all other loci and environments. The subscripts are related to the loci and to the number of genes present. A dot (.) indicates a marginal frequency or mean.

If frequencies at one locus are uncorrelated with frequencies at another locus (algebraically, all $f_{ij} = f_{i.}f_{.j}$) the total variance among the Y's in table 1 can be partitioned exactly into a marginal variance for the A locus (i.e., variance among the row means in table 1), a marginal variance for the B locus, and a joint or interaction variance. Furthermore, each of the marginal variances can be partitioned into a linear and quadratic variance; and correspondingly, the interaction variances can be broken up into four components which are linear

TABLE 1

Hereditary values and relative frequencies of the nine genetic types for two loci, each with two alleles.

$AABB$	$AABb$	$AAbb$	AA —
Y_{22}	Y_{21}	Y_{20}	$Y_{2.}$
f_{22}	f_{21}	f_{20}	$f_{2.}$
$AaBB$	$AaBb$	$Aabb$	Aa —
Y_{12}	Y_{11}	Y_{10}	$Y_{1.}$
f_{12}	f_{11}	f_{10}	$f_{1.}$
$aaBB$	$aaBb$	$aabb$	aa —
Y_{02}	Y_{01}	Y_{00}	$Y_{0.}$
f_{02}	f_{01}	f_{11}	$f_{0.}$
$-BB$	$-BB$	$-bb$	
$Y_{.2}$	$Y_{.1}$	$Y_{.0}$	$Y_{..}$
$f_{.2}$	$f_{.1}$	$f_{.0}$	$f_{..} = 1$

by linear, linear by quadratic, quadratic by linear, and quadratic by quadratic. This linear and quadratic treatment of a 3×3 factorial, where each of the partitions corresponds to one of the eight degrees of freedom, is found in statistical textbooks. The loci are the factors and each factor has three levels represented by the three combinations that can occur with two alleles. In genetic terminology additive and dominance are used in place of linear and quadratic, respectively.

What is to be shown is equally true in the presence of linkage so long as the frequencies at different loci are uncorrelated. If frequencies are correlated, which they would be under phenotypic assortative mating, the following partitioning does not hold. Other causes of correlated frequencies are discussed by LUSH (1948) under the subject of disequilibrium.

Partitioning the variance will be illustrated by using the eight orthogonal scales in table 2. These orthogonal scales serve the same purpose as orthogonal comparisons or polynomials (SNEDECOR 1946) in computing a portion of the total variance for each degree of freedom. The only new feature here is that the

TABLE 2

Hereditary values, frequencies, and eight orthogonal scales which are utilized in partitioning the hereditary variance.

Scale	Genetic type								
	AABB	AABb	AAbb	AaBB	AaBb	Aabb	aaBB	aaBb	aabb
Y	Y_{22}	Y_{21}	Y_{20}	Y_{12}	Y_{11}	Y_{10}	Y_{02}	Y_{01}	Y_{00}
f	f_{22}	f_{21}	f_{20}	f_{12}	f_{11}	f_{10}	f_{02}	f_{01}	f_{00}
W_1	$2v^*$	$2v$	$2v$	$v-u$	$v-u$	$v-u$	$-2u$	$-2u$	$-2u$
W_2	$1/f_{2.}$	$1/f_{2.}$	$1/f_{2.}$	$-2/f_{1.}$	$-2/f_{1.}$	$-2/f_{1.}$	$1/f_{0.}$	$1/f_{0.}$	$1/f_{0.}$
W_3	$2y$	$y-x$	$-2x$	$2y$	$y-x$	$-2x$	$2y$	$y-x$	$-2x$
W_4	$1/f_{.2}$	$-2/f_{.1}$	$1/f_{.0}$	$1/f_{.2}$	$-2/f_{.1}$	$1/f_{.0}$	$1/f_{.2}$	$-2/f_{.1}$	$1/f_{.0}$
W_5	$4vy$	$2v(y-x)$	$-4vx$	$2y(v-u)$	$(v-u)(y-x)$	$-2x(v-u)$	$-4uy$	$-2u(y-x)$	$4ux$
W_6	$2v/f_{.2}$	$-4v/f_{.1}$	$2v/f_{.0}$	$(v-u)/f_{.2}$	$-2(v-u)/f_{.1}$	$(v-u)/f_{.0}$	$-2u/f_{.2}$	$4u/f_{.1}$	$-2u/f_{.0}$
W_7	$2y/f_{2.}$	$(y-x)/f_{2.}$	$-2x/f_{2.}$	$-4y/f_{1.}$	$-2(y-x)/f_{1.}$	$4x/f_{1.}$	$2y/f_{0.}$	$(y-x)f_{0.}$	$-2x/f_{0.}$
W_8	$1/f_{22}$	$-2/f_{21}$	$1/f_{20}$	$-2/f_{12}$	$4/f_{11}$	$-2/f_{10}$	$1/f_{02}$	$-2/f_{01}$	$1/f_{00}$

*$u = f_{2.} + f_{1.}/2$, $v = 1 - u$, $x = f_{.2} + f_{.1}/2$, $y = 1 - x$.

joint frequencies are not equal to each other but are proportional to the marginal frequencies. The requirements of the orthogonal scales are:

(1) $\sum_{i,j} f_{ij} W_{tij} = 0$,   and   (2) $\sum_{i,j} f_{ij} W_{tij} W_{t'ij} = 0$.

The first requirement insures that deviations around the mean are compared. The second requirement insures that the comparisons are orthogonal, which means simply that they are uncorrelated. The eight scales or partitions of the variance are one for each of the eight separate degrees of freedom in a 3×3 table. The symbols u, v, x and y in table 2 are the frequencies of the genes A, a, B and b, respectively

The partition of the variance, σ_t^2, corresponding to any particular scale, W_t, is found in the following manner:

$$\sigma_t^2 = (\sum_{i,j} f_{ij} Y_{ij} W_{tij})^2 / \sum_{i,j} f_{ij} W_{tij}^2,$$

which, in statistical terminology, is

$$\sigma_t^2 = (Cov\ YW_t)^2 / \sigma_{W_t}^2 = \beta_{YW_t}^2 \sigma_{W_t}^2 = \rho_{YW_t}^2 \sigma_Y^2,$$

where Cov, β and ρ are covariance, regression coefficient and correlation coefficient, respectively. The tth partition of the variance is the variance due to regression on the tth orthogonal scale.

This particular set of scales (among the many others mathematically possible) was chosen for its utility. The scales pertaining to the marginal comparisons of each locus were chosen to separate the marginal variance into the same additive (linear) and dominance (quadratic) portions that were long ago shown to be useful for expressing simply the correlation between parent and offspring and between other relatives. The other four scales, which relate to the interactions among the loci, also permit expressing simply the correlations among the interaction effects of one relative and those of another relative. This is the primary purpose of introducing the orthogonal scales.

The first two scales, W_1 and W_2, are concerned only with the means for the rows in table 1 and thus only with the marginal variance for the A locus. For example, the means and frequencies for the A locus are:

Genetic type	AA	Aa	aa
Mean	$Y_2.$	$Y_1.$	$Y_0.$
Frequency	$f_2.$	$f_1.$	$f_0.$
W_1	$2v$	$v-u$	$-2u$
W_2	$1/f_2.$	$-2/f_1.$	$1/f_0.$

The marginal variance of Y for the A locus is broken into two parts, one part being the variance due to the regression of the marginal means on the linear scale designated as W_1, and the other part being the variance due to deviations from this regression. The variance due to regression,

$$\sigma_1^2 = 2uv \left(\frac{f_{2.}}{u} + \frac{f_{0.}}{v}\right) \left[\frac{\dfrac{f_{2.}}{u}(Y_{2.} - Y_{1.}) + \dfrac{f_{0.}}{v}(Y_{1.} - Y_{0.})}{\dfrac{f_{2.}}{u} + \dfrac{f_{0.}}{v}}\right]^2$$

is the additive variance for the A locus and the variance due to deviations from regression on the linear scale,

$$\sigma_2^2 = \frac{f_{2.}\,f_{1.}\,f_{0.}}{4uv - f_{1.}} \left[(Y_{2.} - Y_{1.}) - (Y_{1.} - Y_{0.})\right]^2,$$

is the dominance variance for the A locus. This latter variance is also the variance due to regression on the scale designated as $W_{2.}$.

The additive variance is proportional to the square of the average effect, β_{YW_1}, of the genes; this average effect being a weighted average of the two effects or differences $Y_{2.} - Y_{1.}$ and $Y_{1.} - Y_{0.}$ corresponding to the comparisons $AA - Aa$ and $Aa - aa$. Each difference represents an effect of replacing a by A, $A - a$, but in each case the effect is measured in the presence of a different allele. If the two differences are not exactly the same, i.e., the effect of replacing a by A changes according to whether a or A is present, then the alleles at this locus interact, and the interaction is reflected in the dominance variance, σ_2^2.

In a similar manner σ_3^2 and σ_4^2 are the additive and dominance variances, respectively, which sum to the marginal variance for locus B; i.e., the variance between the means for the columns in table 1. The partitioning of the variance to this point is the same as that of FISHER (1918) and WRIGHT (1935).

The last four components (σ_5^2 through σ_8^2) account for the remaining or epistatic portion of the variance of Y. The naming of the epistatic components corresponds to the relationships among the orthogonal scales:

$$
\begin{aligned}
W_5 &= W_1 \times W_3 &&\text{(additive} \times \text{additive)}\\
W_6 &= W_1 \times W_4 &&\text{(additive} \times \text{dominance)}\\
W_7 &= W_2 \times W_3 &&\text{(dominance} \times \text{additive)}\\
W_8 &= W_2 \times W_4 &&\text{(dominance} \times \text{dominance)}.
\end{aligned}
$$

The epistatic variance for the case of two loci, therefore, consists of four parts: σ_5^2 is the additive in A by additive in B, σ_6^2 is the additive in A by dominance in B, σ_7^2 is the dominance in A by additive in B and σ_8^2 is the dominance in A by dominance in B. FISHER (1918) and other subsequent workers, in expressing the epistatic variance for two loci, obtained one epistatic component which is actually the sum of the four components indicated above.

The epistatic components arise because the effects of genes at one locus depend on what genes are present at the other locus. For example, the dominance by dominance component is

$$\sigma_\delta^2 = \left(\frac{f_2. f_1. f_0.}{4uv - f_1.}\right) \left(\frac{f_{.2} f_{.1} f_{.0}}{4xy - f_{.1}}\right) (e_{22} - e_{21} - e_{12} + e_{11})^2,$$

where the e's designate the following comparisons among the hereditary values:

$$e_{22} = Y_{22} - Y_{21} - Y_{12} + Y_{11}$$
$$e_{21} = Y_{21} - Y_{20} - Y_{11} + Y_{10}$$
$$e_{12} = Y_{12} - Y_{11} - Y_{02} + Y_{01}$$
$$e_{11} = Y_{11} - Y_{10} - Y_{01} + Y_{00}.$$

All four epistatic components of variance are functions of these four comparisons, and unless each of the comparisons is zero there will be epistatic variance. The comparison e_{22}, for example, corresponds to $(AABB - AABb)$ $- (AaBB - AaBb)$, and will be zero only if the gene effect at the B locus, $BB - Bb$, is the same for each of the phases, AA and Aa, at the A locus. The other comparisons bear similar interpretations and each represents the failure of the effect of a gene replacement at one locus to remain the same when a gene is replaced at the other locus. The epistatic components of variance will be illustrated in more detail later; the essential feature here being that they all arise from non-allelic gene interactions.

The extension to 3 and more loci is apparent. For the case of 3 loci there are 3 additive components, 3 dominance components and 20 epistatic components. The 20 epistatic components are 3 a × a, 6 a × d, 3 d × d, 1 a × a × a, 3 a × a × d, 3 a × d × d and 1 d × d × d (a = additive, d = dominance). For the case of n loci there are $3^n - 1$ components consisting of n components of the additive type, n components of the dominance type and $3^n - 2n - 1$ epistatic components. Of the $3^n - 2n - 1$ epistatic components there are $2n(n-1)$ two-factor components of which one-fourth are of the type a × a, one-half are of the type a × d and one-fourth are of the type d × d; $\frac{4}{3}n(n-1)(n-2)$ three-factor components of which one-eighth are of the type a × a × a, three-eighths are of the type a × a × d, three-eighths are of the type a × d × d and one-eighth are of the type d × d × d; etc. For many purposes the components of the same type may be combined. The types are designated separately because they present different properties in the correlations among relatives which will be discussed later.

Although the method presented here does not include multiple alleles, FISHER (1918) did partition the marginal variance for a locus with any number of alleles in a random mating population into an additive part and a dominance part. Correlations among the additive deviations and dominance deviations of relatives in a randomly mating population were the same as those for two alleles at a locus. It may be possible to partition the epistatic variance for multiple alleles by multiple regression and correlation techniques into components which will bear definitions similar to those where only two alleles are considered. Should the correlations among these epistatic deviations of relatives

be the same as those when only two alleles are considered, the development herein is general for any number of alleles. At present, this must remain as a conjecture.

Explicit expressions (in terms of Y's, gene frequencies and WRIGHT'S inbreeding coefficient, F) are given in table 3 for the eight components of variance for the case of two loci and for any amount of inbreeding. In an inbred population the marginal frequencies for the A locus are $f_{2.} = u^2 + Fuv$, $f_{1.} = 2uv(1 - F)$ and $f_{0.} = v^2 + Fuv$, and those for the B locus are $f_{.2} = x^2 + Fxy$, $f_{.1} = 2xy(1 - F)$ and $f_{.0} = y^2 + Fxy$. When the marginal frequencies are uncorrelated, which is the case under consideration throughout this study, the joint frequencies are found by the simple relationship, $f_{ij} = f_{i.}f_{.j}$.

The partitions for random mating can be found from those in table 3 by letting $F = 0$. It should be recognized that when there is interaction among loci the marginal means, $Y_{i.}$ and $Y_{.j}$, vary with the degree of inbreeding, F. Thus one cannot compare the same marginal partitions for different degrees of inbreeding without writing out the marginal means in more detail. The Y_{ij}'s are also average values, averaged over the various gene combinations at other loci and over the various environmental sources of variation which determine the population of phenotypes. If interactions of 3 or more factors involve these loci, the Y_{ij}'s will also change with inbreeding. Thus the partitions of variance are descriptions of the population under consideration at the moment and will change with inbreeding.

How the inbreeding affects the partitions, or at least the ultimate end of the partitions as inbreeding approaches one, can be seen in table 3. The dominance variance disappears when $F = 1$. The disappearance is not linear with F, however. With some frequencies the dominance variance will actually increase for a period of time with inbreeding (ROBERTSON 1952). The last three of the epistatic partitions in table 3 also disappear as inbreeding becomes complete. Note that these three partitions involve dominance in their nomenclature. Their disappearance is not linear with F and is too complicated to allow any lucid generalizations. In fact, all partitions involving dominance in their nomenclature, and for any number of loci, will disappear as inbreeding becomes complete. If all partitions involving dominance in their nomenclature are zero, the effect of inbreeding on the additive partitions is to increase the one-factor additive partitions by $1 + F$, the two-factor additive partitions by $(1 + F)^2$, the three-factor additive partitions by $(1 + F)^3$, and so on. However, when partitions involving dominance in their nomenclature are not zero, the influence of inbreeding on the additive type of partitions cannot be foretold without specifying the hereditary values and gene frequencies.

CORRELATIONS BETWEEN RELATIVES

The covariances (or correlations) between the phenotypes of relatives can be analyzed in terms of the covariances between their respective components: that is, between the hereditary, environmental, and interaction values of one relative and those of the other relative. Only the covariance between hereditary

C. CLARK COCKERHAM

TABLE 3

Partitions of the hereditary variance and the variances of the orthogonal scales for two loci and any degree of inbreeding (F).

Orthogonal scale	Partitions of the variance	σ_W^2
W_1 a	$\dfrac{2uv}{1+F}\,[(u+Fv)(Y_{2.}-Y_{1.})+(v+Fu)(Y_{1.}-Y_{0.})]^2$	$2uv(1+F)$
W_2 d	$\dfrac{uv(u+Fv)(v+Fu)(1-F)}{1+F}\,(Y_{2.}-2Y_{1.}+Y_{0.})^2$	$\dfrac{1+F}{uv(1-F)(u+vF)(v+uF)}$
W_3 a	$\dfrac{2xy}{1+F}\,[(x+Fy)(Y_{.2}-Y_{.1})+(y+Fx)(Y_{.1}-Y_{.0})]^2$	$2xy(1+F)$
W_4 d	$\dfrac{xy(x+Fy)(y+Fx)(1-F)}{1+F}\,(Y_{.2}-2Y_{.1}+Y_{.0})^2$	$\dfrac{1+F}{xy(1-F)(x+Fy)(y+Fx)}$
W_5 a×a	$\dfrac{4uvxy}{(1+F)^2}\,[(u+Fv)(x+Fy)(y+Fx)e_{22}^*+(u+Fv)(y+Fx)e_{21}+(v+Fu)(x+Fy)e_{12}+(v+Fu)(y+Fx)e_{11}]^2$	$4uvxy(1+F)^2$
W_6 a×d	$\dfrac{2uvxy(1-F)(x+Fy)(y+Fx)}{(1+F)^2}\,[(u+Fv)(e_{22}-e_{21})+(v+Fu)(e_{12}-e_{11})]^2$	$\dfrac{2uv(1+F)^2}{xy(1-F)(x+Fy)(y+Fx)}$
W_7 d×a	$\dfrac{2xyuv(1-F)(u+Fv)(v+Fu)}{(1+F)^2}\,[(x+Fy)(e_{22}-e_{12})+(y+Fx)(e_{21}-e_{11})]^2$	$\dfrac{2xy(1+F)^2}{uv(1-F)(u+Fv)(v+Fu)}$
W_8 d×d	$\dfrac{uvxy(1-F)^2}{(1+F)^2}\,(x+Fy)(y+Fx)(u+Fv)(v+Fu)(e_{22}-e_{21}-e_{12}+e_{11})^2$	$\dfrac{(1+F)^2}{uvxy(1-F)^2(x+Fy)(y+Fx)(u+Fv)(v+Fu)}$

*The e's indicate the following comparisons among the hereditary values:

$$e_{22}=Y_{22}-Y_{21}-Y_{12}+Y_{11} \qquad e_{21}=Y_{21}-Y_{20}-Y_{11}+Y_{10}$$

$$e_{12}=Y_{12}-Y_{11}-Y_{02}+Y_{01} \qquad e_{11}=Y_{11}-Y_{10}-Y_{01}+Y_{00}$$

values will be considered here. The other covariances can sometimes be assumed to be zero, but this will depend on the organism and the circumstances. In human and livestock populations these correlations are often a troublesome feature. For example, in multiparous species, the environments of littermates are correlated. It is always necessary to be wary about assuming that the only covariances between relatives are those between their heredities.

The hereditary deviations of individuals are products of the regression coefficients, β_{YW}'s, and the appropriate scale values, W's. For example, the additive genetic deviation of an individual for the A locus is $\beta_{YW_1}W_1$, where W_1 is 2v, v-u or -2u corresponding to whether the individual is AA, Aa or aa, respectively. WRIGHT's (1935, 1950) additive estimates, G's in his notation, are the additive genetic deviations plus the mean, $\overline{Y}$. His dominance deviations, $D = H - G$ in his notation, are the same as the present dominance deviations; for example, $\beta_{YW_2}W_2$'s for the A locus. An individual's hereditary value is a sum of the mean and its hereditary deviations,

$$Y = \overline{Y} + \sum_t \beta_{YW_t} W_t.$$

The correlations between the additive genetic and between the dominance deviations of relatives in a randomly mating population were given by FISHER (1918). He also gave the covariance between epistatic deviations of relatives for the case of two loci. WRIGHT (1922), with his coefficient of relationship, extended the correlations between additive genetic deviations of relatives to include inbreeding. WRIGHT's coefficient of relationship is the same as the absolute value of the correlation between additive genetic deviations of relatives as described in this paper. COCKERHAM (1952) gave the correlations between epistatic deviations of relatives in a randomly mating population.

The orthogonal scales may again be invoked to obtain these relationships for a randomly mating population or an inbreeding population where gene frequency is not changing, and to obtain the correlations among relatives for the epistatic deviations. Although one cannot find a joint frequency distribution of relatives which is general for all systems of inbreeding they can be computed for any particular system of consanguine mating. Again, the case of two loci with two genes each will be used for illustrative purposes; however, the extension to any number of loci will be apparent.

Let Y and Y' be the hereditary values for two relatives (parent and offspring, full-sibs, etc.) and let W_t and W_t' be the orthogonal scales of their respective generations. The scale values (terms of the scale) vary with the genotypes of course, but the entire scale may vary with the generation (e.g., when there is inbreeding). In a randomly mating population where gene frequencies do not change, relatives with the same genotype have the same hereditary deviations and the same scale values. The different notation (prime) simply designates whether the hereditary deviation or scale value is used for one relative or the other relative. In any system of inbreeding relatives with

the same genotype will have the same deviations and the same scale values provided they are in the same generation but can have different deviations and different scale values if they are in different generations, even though gene frequency does not change.

The covariance between Y and Y' may be expressed as

$$\text{Cov } YY' = \sum_{t,t'} \text{Cov}(\beta_{YW_t} W_t)(\beta_{Y'W'_{t'}} W'_{t'}), \quad (t, t' = 1 \ldots 8)$$

which is equivalent to

$$\sum_{t,t'} \rho(\beta_{YW_t} W_t)(\beta_{Y'W'_{t'}} W'_{t'}) \sigma_t \sigma'_{t'}$$

because $\sigma_t^2 = \beta_{YW_t}^2 \sigma_{W_t}^2$. Now, $\rho(\beta_{YW_t} W_t)(\beta_{Y'W'_{t'}} W'_{t'}) = \pm \rho_{W_t W'_{t'}}$

because the regression coefficients are constants and do not affect the absolute value of the correlation coefficients. The sign (+ or −) will be the same as the sign of the product of the two regression coefficients, and will be considered in more detail later. It is necessary then to determine the correlation between the eight orthogonal scales of one relative and the eight orthogonal scales of the other relative. This is done by constructing a nine by nine joint distribution table for the two relatives. Although the results are simple, the development is involved and is given in an Appendix. If the genotypic frequencies for all relatives are proportional to the marginal frequencies of the loci, if the genes recombine independently, and if gene frequencies do not change, the correlations between the epistatic scales of two relatives are products of the one-factor correlations as given in table 4.

TABLE 4

Correlations between the eight scales of one relative
and the eight scales of another relative.

	W'_1	W'_2	W'_3	W'_4	W'_5	W'_6	W'_7	W'_8
W_1	$\rho_{w_1 w'_1}$	$\rho_{w_1 w'_2}$	0	0	0	0	0	0
W_2	$\rho_{w_2 w'_1}$	$\rho_{w_2 w'_2}$	0	0	0	0	0	0
W_3	0	0	$\rho_{w_3 w'_3}$	$\rho_{w_3 w'_4}$	0	0	0	0
W_4	0	0	$\rho_{w_4 w'_3}$	$\rho_{w_4 w'_4}$	0	0	0	0
W_5	0	0	0	0	$\rho_{w_1 w'_1} \rho_{w_3 w'_3}$	$\rho_{w_1 w'_1} \rho_{w_3 w'_4}$	$\rho_{w_1 w'_2} \rho_{w_3 w'_3}$	$\rho_{w_1 w'_2} \rho_{w_3 w'_4}$
W_6	0	0	0	0	$\rho_{w_1 w'_1} \rho_{w_4 w'_3}$	$\rho_{w_1 w'_1} \rho_{w_4 w'_4}$	$\rho_{w_1 w'_2} \rho_{w_4 w'_3}$	$\rho_{w_1 w'_2} \rho_{w_4 w'_4}$
W_7	0	0	0	0	$\rho_{w_2 w'_1} \rho_{w_3 w'_3}$	$\rho_{w_2 w'_1} \rho_{w_3 w'_4}$	$\rho_{w_2 w'_2} \rho_{w_3 w'_3}$	$\rho_{w_2 w'_2} \rho_{w_3 w'_4}$
W_8	0	0	0	0	$\rho_{w_2 w'_1} \rho_{w_4 w'_3}$	$\rho_{w_2 w'_1} \rho_{w_4 w'_4}$	$\rho_{w_2 w'_2} \rho_{w_4 w'_3}$	$\rho_{w_2 w'_2} \rho_{w_4 w'_4}$

Two simple generalizations are apparent from table 4. (1) One-factor deviations (additive or dominance) of one relative are correlated, if at all, only with one-factor deviations of the other relative, and also the deviations must pertain to the same locus. This is actually a particular case of the broader generalization that n-factor deviations of one relative are correlated, if at all, only with n-factor deviations of the other relative, and the deviations for both

relatives must involve the same loci. (2) The correlations between the epistatic or multi-factor deviations are functions of the correlations between one-factor deviations.

Relatives in a randomly mating population

If the one-factor additive deviations of either relative are uncorrelated with the one-factor dominance deviations of the other relative, i.e., $\rho_{w_1 w_2'}$ = $\rho_{w_2 w_1'}$ = 0 for all loci, then all the correlations in table 4 are zero, except for the diagonal ones. This means that a deviation in one relative is correlated, if at all, only with the same deviation in the other relative. This condition is fulfilled in a randomly breeding population. Since the one-factor additive correlations are the same and the one-factor dominance correlations are the same for all loci, a simple rule can be used for computing the various correlations. Let p be the correlation between one-factor additive deviations of the relatives, and let q be the correlation between one-factor dominance deviations of the relatives. The correlation between any type of deviation of the relatives is

$$(p)^A(q)^D,$$

where A is the number of factors (loci) entering into the deviation with additive nomenclature and D is the number of factors entering into the deviation with dominance nomenclature. The deviation is an $(A + D)$-factor deviation. When the two relatives are parent and offspring, p = ½ and q = 0. No deviations involving dominance in their nomenclature are correlated between parent and offspring. The correlations between the indicated types of deviations are:

a	a × a	a × a × a
$\left(\dfrac{1}{2}\right)^1 (0)^0 = \dfrac{1}{2}$	$\left(\dfrac{1}{2}\right)^2 (0)^0 = \dfrac{1}{4}$	$\left(\dfrac{1}{2}\right)^3 (0)^0 = \dfrac{1}{8}$

Dominance is illustrated by considering the relationships between full sibs; i.e. $p = \dfrac{1}{2}$ and $q = \dfrac{1}{4}$. The correlations between the epistatic deviations are

a × a	a × d	d × d
$\left(\dfrac{1}{2}\right)^2 \left(\dfrac{1}{4}\right)^0 = \dfrac{1}{4}$	$\left(\dfrac{1}{2}\right)^1 \left(\dfrac{1}{4}\right)^1 = \dfrac{1}{8}$	$\left(\dfrac{1}{2}\right)^0 \left(\dfrac{1}{4}\right)^2 = \dfrac{1}{16}$, and so on.

The correlations between the various deviations of relatives in a randomly mating population are all positive, since each one involves the same deviation and the two regression coefficients, β_{YW} and $\beta_{Y'W'}$, are identical in each case.

Since the partitions of the variance are the same for each relative in a randomly mating population the covariance between the hereditary values of two relatives can be given in terms of partitions of the hereditary variance. Also since the correlations among hereditary deviations are the same for all

those of the same type it is convenient to designate the sum of the components of variance of the same type by a single term. Let σ_{AD}^2 be the sum of all the components in whose nomenclature additive appears A times and dominance appears D times. For example, $\sigma_{10}^2 (A = 1, D = 0)$ is the additive genetic variance which is the sum of all the one-factor additive components (one from each locus), σ_{01}^2 is the dominance variance, σ_{11}^2 is the a × d epistatic variance, and so on. Symbolically, the total hereditary variance is

$$\sigma_Y^2 = \sum_{\substack{A,D=0 \\ 1 \leq A+D \leq n}}^{n} \sigma_{AD}^2,$$

where n is the number of loci. The summation is over both A and D where each varies from 0 to n but subject to the limitations that for any term A and D cannot both be zero nor can their sum be greater than n. The covariance between relatives can be given in terms of these partitions,

$$\text{Cov } YY' = \sum_{\substack{A,D=0 \\ 1 \leq A+D \leq n}}^{n} p^A q^D \sigma_{AD}^2.$$

When the two relatives are parent and offspring the covariance reduces to

$$\sum^{n} \left(\frac{1}{2}\right)^A \sigma_{AO}^2,$$

which is one-half the additive genetic variance, plus one-fourth the a × a variance plus one-eighth the a × a × a variance, and so on.

If only the heredities of relatives are correlated, the covariance between their phenotypes is the same as the covariance between their heredities. In this case the phenotypic correlation or the phenotypic regression of one relative on the other is the ratio of the covariance between the heredities of the relatives to the total phenotypic variance.

Relatives in a self-fertilizing population

Although the correlations among epistatic deviations of relatives in an inbred population are functions of the correlations for one-factor deviations, several new considerations are involved. The absolute values of the correlations between one-factor additive deviations are the same as WRIGHT's (1922) coefficient of relationship (the correlations can be negative when the relatives are in different generations depending on dominance and epistatic effects of the genes). This means that the absolute values of the correlations between the epistatic deviations of relatives involving only additive in their nomenclature can be found in the same manner as when mating is random. The correlations between

one-factor dominance deviations of relatives must be computed for each system of inbreeding and involve considerable labor. Also, one-factor additive deviations of one relative are correlated with one-factor dominance deviations of the other relative. This means that many more of the epistatic deviations of one relative are correlated with those of the other relative than in a randomly breeding population.

Self-fertilization was chosen as the system of inbreeding to illustrate the correlations among relatives in an inbred population for three main reasons. First, it is the most extreme form of inbreeding and relationships in a milder system of inbreeding should fall between these for selfing and those indicated earlier for random mating; second, considerable work is being done in plant breeding with normally selfed organisms or by selfing organisms; and third, it is by far the easiest system to work with computationally. The correlations between hereditary deviations of parent and of offspring will be considered at first.

The correlation between one-factor additive scales of the parent and one-factor additive scales of the offspring is

$$\frac{1 + F' + 2F}{2\sqrt{1 + F}\,\sqrt{1 + F'}}, \qquad \text{(Lush 1948)}$$

where F' and F are the inbreeding coefficients of the parent and the offspring, respectively. When selfing

$$F = \frac{1 + F'}{2},$$

so that the correlation becomes

$$\sqrt{\frac{2F}{1 + F}}.$$

The correlation between the one-factor dominance scales of parent and of offspring in a self-fertilizing population is

$$\frac{1}{2}\sqrt{\frac{(1 + F)(u - v + 2Fv)(v - u + 2Fu)}{F(u + Fv)(v + Fu)}},$$

which involves gene frequency, u and v at the A locus in this example. Also in a self-fertilizing population, the one-factor additive scales of the parent may be correlated with the one-factor dominance scales of the offspring. This correlation is

$$\frac{(u - v)(1 - F)}{2}\sqrt{\frac{(1 - F)}{F(1 + F)(u + Fv)(v + Fu)}},$$

which is zero when gene frequency is one-half. The dominance scales of the parent and the additive scales of the offspring are not correlated.

The correlations among the epistatic scales can be found from the one-factor correlations by substitution in table 4. The simple rule used for computing the correlation between any type of deviation in a randomly mating population cannot be used here because (1) the one-factor dominance correlations involve gene frequencies and can differ from locus to locus, (2) different types of epistatic deviations may be correlated, and (3) in terms of hereditary deviations the correlations may be negative.

The signs (+ or --) of the correlations vary with the dominance and epistatic effects of the genes. For example, even in the absence of epistasis, the correlation between the additive deviations of non-inbred parents and first generation selfed offspring is negative for an over-dominant locus when the frequency of the favorable gene is slightly higher than the frequency that would make the mean of the parents a maximum. For any given situation the sign can be found and is the same as the sign of the product of the two regression coefficients, β_{YW} and $\beta_{Y'W'}$, corresponding to the deviations under question. When the relatives are in the same generation the correlation between the same type of deviation will always be positive because the two regression coefficients are identical, but even here the signs of the correlations between different types of deviations have to be determined and also all of the correlations involve gene frequency. The knowledge required to determine the correlations, gene frequencies and regression coefficients, precludes any use of this type of biometrical analysis, with one exception, and this is when all gene frequencies are one-half. In this case a deviation of one relative is correlated, if at all, only with the same deviation of the other relative, as is true for relatives in a randomly mating population.

The frequencies of segregating genes can be reasonably assumed to be one-half in subsequent generations of a cross between two homozygous lines. However, the applicability of the following results is limited somewhat because of the assumption, necessary in this analysis, that the genotypic frequencies are proportional to the marginal frequencies of the loci. Even with random mating following the cross it is several generations before the genotypic frequencies can be reasonably assumed to be proportional to the marginal frequencies of those loci which are closely linked. When generations subsequent to the cross are selfed or inbred rapidly the genotypic frequencies cannot ever be reasonably assumed to be proportional to the marginal frequencies of the linked loci because the parental types are fixed much more rapidly than the recombinants. In this case the partitions of the variance and the following correlations among relatives apply strictly only to those combinations of loci whose genes recombine independently.

Where each generation after the first generation of a cross between two homozygous parental lines is obtained by self-fertilizing the previous generation, the coefficient of inbreeding in the gth generation is

$$\frac{2^{g-2} - 1}{2^{g-2}}$$

The correlation between the parent in the $(g-1)$th generation and the offspring in the gth generation is then:

$$\sqrt{\frac{2(2^{g-2}-1)}{2^{g-1}-1}}$$

for the additive scales, and

$$\sqrt{\frac{2^{g-2}-1}{2^{g-1}-1}}$$

for the dominance scales. These correlations have meaning only when g is greater than two.

By using the correlations between parent and offspring a more general scheme can be developed for indicating the correlations between any two relatives. Let $\rho_{a(k,g_1,g_2)}$ be the correlation between the additive scales of one relative, O_{g_1}, in the g_1th generation and of another relative, P_{g_2}, in the g_2th generation both of whom descended from the last common parent, C_k, in the kth generation. Let $\rho_{d(k,g_1,g_2)}$ be the correlation between the dominance scales of the same two relatives. The following path coefficient diagram shows the relationships between these relatives.
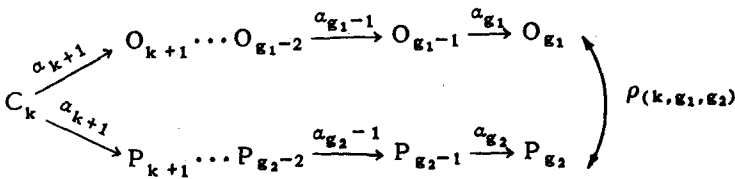
P's are used to indicate the parents of P_{g_2} and O's are used to indicate the parents of O_{g_1} in their descent from the last common parent in the kth generation. The single-headed arrows are path coefficients (WRIGHT 1934), and the double headed arrow indicates a correlation coefficient. The only connection between the genotypes of the two relatives is through the chain of parents via the common parent. Since the additive scales of one individual are correlated only with the additive scales (and not the dominance scales) of the preceding parent (or succeeding offspring), the path coefficient, a, for the additive scales is actually the correlation coefficient between the additive scales of the parent and of the offspring. The correlation between the additive scales of O_{g_1} and P_{g_2} is computed by multiplying together all the path coefficients:

$$\rho_{a(k,g_1,g_2)} = \alpha_{g_1} \alpha_{g_1-1} \cdots \alpha^2_{k+1} \cdots \alpha_{g_2-1} \alpha_{g_2}.$$

Substituting

$$\alpha_g = \sqrt{\frac{2(2^{g-2}-1)}{2^{g-1}-1}},$$

the correlation between the additive scales of the two relatives is

$$\rho_{a(k,g_1,g_2)} = \frac{2^{(g_1+g_2-2k)/2}(2^{k-1}-1)}{\sqrt{(2^{g_1-1}-1)(2^{g_2-1}-1)}}.$$

The correlation between the dominance scales of O_{g_1} and P_{g_2} are found in the same manner, except of course, the path coefficients are correlations between the dominance scales. Substituting

$$\alpha_g = \sqrt{\frac{2^{g-2}-1}{2^{g-1}-1}},$$

the correlation between the dominance scales of the two relatives is

$$\rho_{d(k,g_1,g_2)} = \frac{2^{k-1}-1}{\sqrt{(2^{g_1-1}-1)(2^{g_2-1}-1)}}.$$

The correlation between the various epistatic scales can be found by the familiar formula:

$$\left[\rho_{a(k,g_1,g_2)}\right]^A \left[\rho_{d(k,g_1g_2)}\right]^D$$

When the relatives are in the same generation the covariance between their hereditary values can be given in terms of partitions of the hereditary variance,

$$\text{Cov } YY' = \sum_{\substack{A,D=0 \\ 1 \le A+D \le n}}^{n} \rho^{A}_{a(k,g,g)} \rho^{D}_{d(k,g,g)} \sigma^2_{AD},$$

as was done for the randomly mating population. Remember that these partitions of variance are for the inbred populations and change with the degree of inbreeding. As an example of the procedure, consider a population of F_4 individuals which are classified according to their F_3 parents and F_2 grandparents. The total hereditary variation of these F_4 individuals is distributed in the following analysis of variance as

Source	Variance
Between F_2 progeny means	$\sum_{A,D} \rho^{A}_{a(2,4,4)} \rho^{D}_{d(2,4,4)} \sigma^2_{AD}$
Between F_3 progeny means within F_2 parents	$\sum_{A,D} \left[\rho^{A}_{a(3,4,4)} \rho^{D}_{d(3,4,4)} - \rho^{A}_{a(2,4,4)} \rho^{D}_{d(2,4,4)}\right] \sigma^2_{AD}$
Between F_4 individuals with the same F_3 parent	$\sum_{A,D} \left[1 - \rho^{A}_{a(3,4,4)} \rho^{D}_{d(3,4,4)}\right] \sigma^2_{AD},$
where	$\sum_{A,D} \rho^{A}_{a(2,4,4)} \rho^{D}_{d(2,4,4)} \sigma^2_{AD}$

is the covariance between the hereditary values of F_4 individuals which descended from the same F_2 grandparents but different F_3 parents, and

$$\sum_{A,D} \rho^A_{a(3,4,4)} \, \rho^D_{d(3,4,4)} \, \sigma^2_{AD}$$

is the covariance between the hereditary values of F_4 individuals which descended from the same F_3 parent. The one-factor correlations are $\rho_{a(2,4,4)} = 4/7$, $\rho_{a(3,4,4)} = 6/7$, $\rho_{d(2,4,4)} = 1/7$ and $\rho_{d(3,4,4)} = 3/7$.

Relatives from randomly mated inbred parents

The final situation to be considered is one in which several inbred lines are made from a randomly mating population. The inbreeding program is interrupted at any desired stage of inbreeding to cross the inbred individuals at random. Although the parents are inbred, the offspring will not be. Barring selection and linkage, the offspring will reconstitute the original randomly breeding population from which the inbred individuals were obtained. Whether one can actually obtain unselected inbred lines or individuals is questionable. If the genotypic frequencies are proportional to the marginal frequencies of the loci and if gene frequencies do not change (no selection), the gametic arrays from any generation of inbreeding will be the same as those in the randomly mating population. Thus the offspring from the randomly mated inbred parents will be a reconstitution of the original population. However, the genotypic frequencies among the inbred parents, at least when the parents are the result of self-fertilization, will not be proportional to the marginal frequencies of those loci which have recombination values less than one-half. This will affect the joint distribution of full sib offspring and of half sib offspring. In a few examples considered, the effect on the joint distribution is small. Nevertheless, if the inbreeding is between zero and one ($0 < F < 1$), the results are accurate only for those combinations of loci which have recombination frequencies of one-half. The frequencies of genotypes among parents inbred to homozygosity ($F = 1$) are proportional to the marginal frequencies of the loci, regardless of linkage relationships, if the genotypic frequencies in the original population were proportional to the marginal frequencies of the loci and if there has been no selection or change in gene frequencies.

The covariance between offspring with both parents in common, full sibs, and the covariance between offspring with one parent in common, half sibs, will be considered. The joint distribution of full sibs from randomly mated inbred parents for the A locus is given in table 5. From this table the following correlations are found: $\rho_{w_1 w_1'} = (1 + F)/2$, $\rho_{w_1 w_2'} = \rho_{w_2 w_1'} = 0$, $\rho_{w_2 w_2'} = [(1 + F)/2]^2$, i.e., the correlation between the additive deviations of full sibs is $(1 + F)/2$, the additive deviations of one full sib are not correlated with the dominance deviations of the other full sib, and the correlation between the dominance deviations is the square of the correlation between the additive deviations.

In an analogous manner it is found that of the one-factor deviations only the additive deviations of half sibs are correlated, and this correlation is $(1 + F)/4$.

TABLE 5

Joint frequency of full sibs from randomly mated inbred parents (A locus).

Genetic type	f	W_1	W_2
AA	u^2	$2v$	$1/u^2$
Aa	$2uv$	$v-u$	$-1/uv$
aa	v^2	$-2u$	$1/v^2$

Full sib	Full sib' AA	Aa	aa
AA	$u^2[u + Fv + v^2(1-F)^2/4]$	$u^2v(1-F)(1+u+Fv)/2$	$u^2v^2(1-F)^2/4$
Aa	$u^2v(1-F)(1+u+Fv)/2$	$uv[1+F+uv(1-F)^2]$	$uv^2(1-F)(1+v+Fu)/2$
aa	$u^2v^2(1-F)^2/4$	$uv^2(1-F)(1+v+Fu)/2$	$v^2[v+Fu+u^2(1-F)^2/4]$

Genetic Type	AA	Aa	aa
f'	u^2	$2uv$	v^2
W_1'	$2v$	$v-u$	$-2u$
W_2'	$1/u^2$	$-1/uv$	$1/v^2$

The correlations between the epistatic deviations are found in the manner given previously: $(p)^A(q)^D$. The correlations are all positive.

When the parents are inbred to homozygosity, $F = 1$, the correlations between hereditary deviations of half sibs are the same as those between parent and offspring in the original population. This is a reasonable result when one considers that the relationship, in both cases, is between two individuals in a randomly mating population with a gamete in common.

Again, the covariance between the hereditary values of these relatives, which is the same as the covariance between their phenotypes if only their hereditary values are correlated, can be expressed in the familiar form,

$$\text{Cov } YY' = \sum_{\substack{A,D = 0 \\ 1 \leq A + D \leq n}}^{n} p^A q^D \sigma^2_{AD}.$$

The partitions of variance are the same as those in the original randomly mating population.

<div style="text-align:center">SUMMARY</div>

The hereditary variance of diploid populations, whose genotypic frequencies are proportional to the marginal frequencies of the loci, was partitioned into $3^n - 1$ partitions analogous to the linear and quadratic analysis of a 3^n factorial representation, where n is the number of loci (factors) each with three genetic phases or levels (no multiple alleles). These partitions are grouped into types of which there are $s + 1$ types of s-factor partitions. The one-factor types are (1) additive (linear) variance, or additive genetic variance, which is the sum of the n partitions, one for each locus, resulting from additive effects of genes and (2) dominance (quadratic) variance which is the sum of the n partitions resulting from dominance or allelic interactions of genes. The remaining two and more factor partitions result from non-allelic interactions of genes and constitute the epistatic variance. The three types of two-factor epistatic variances are (1) additive by additive variance which is the sum of the $n(n-1)/2$ partitions, one for each combination of two loci, resulting from joint additive by additive effects of non-allelic genes, (2) additive by dominance variance which is the sum of the $n(n-1)$ partitions resulting from joint additive by dominance effects of non-allelic genes and (3) dominance by dominance variance which is the sum of the $n(n-1)/2$ partitions resulting from joint dominance by dominance effects of non-allelic genes. The three (or more) factor types of epistatic variance bear similar interpretations.

Correlations between the hereditary deviations (a, d, a × a, a × d, and so on; a = additive, d = dominance) of relatives were considered for the following situations: (1) relatives in a randomly mating population, (2) relatives in a self-fertilizing population and (3) relatives which are offspring from randomly mated inbred parents. For situations 1 and 3, and for situation 2 when gene frequencies are one-half and the relatives are in the same generation, the corre-

lations between the epistatic deviations of the relatives are products of the one-factor additive and dominance correlations. For each of these situations only the same types of deviations involving the same locus or loci are correlated and the correlations are the same for deviations of the same type. Therefore, the correlations between the hereditary deviations of the relatives may be generally expressed as

$$(p)^A(q)^D,$$

where additive appears A times and dominance appears D times in the nomenclature of the deviation, and where p and q are the correlations between the additive and dominance deviations, respectively. Thus the covariance between hereditary values of the two relatives is

$$\sum_{\substack{A,D=0 \\ 1 \le A + D \le n}}^{n} p^A q^D \sigma^2_{AD}$$

For example, where the two relatives are parent and offspring in a randomly mating population, $p = 1/2$ and $q = 0$. The covariance between the hereditary values of parent and offspring is $\sum_{A=1}^{n} (1/2)^A \sigma^2_{AO}$, which is $1/2$ the additive genetic variance, plus $1/4$ the a × a variance, plus $1/8$ the a × a × a variance, and so on.

LITERATURE CITED

COCKERHAM, C. CLARK, 1952 Genetic covariation among characteristics of swine. Unpublished Ph.D. Thesis. Iowa State College Library. Ames, Iowa.

FISHER, R. A., 1918 The correlation between relatives on the supposition of Mendelian inheritance. Trans. Roy. Soc. Edinburgh 52: 399–433.

LUSH, J. L., 1948 The genetics of populations. Ames, Iowa. Mimeo.

ROBERTSON, ALAN, 1952 The effect of inbreeding on the variation due to recessive genes. Genetics 37: 189–207.

NELDER, J. A., 1953 Statistical models in biometrical genetics. Heredity 7: 111–119.

SNEDECOR, G. W., 1946 Statistical methods, 4th edition. Iowa State College Press. Ames, Iowa.

WRIGHT, SEWALL, 1922 Coefficients of inbreeding and relationship. Am. Nat. 56: 330–338.
1934 The method of path coefficients. Ann. Math. Stat. 5: 161–215.
1935 The analysis of variance and the correlations between relatives with respect to deviations from an optimum. J. Genet. 30: 243–256.
1950 The genetics of quantitative variability. Colloquium at the Institute of Animal Genetics of the University of Edinburgh, 1949–1950.

APPENDIX

GENERAL DERIVATION OF THE CORRELATIONS BETWEEN SCALES OF RELATIVES

Consider at first the joint distribution of two relatives for the A locus averaged over all other loci, table 6. The h's are the joint frequencies of the two relatives; for example,

TABLE 6

Joint frequency of two relatives (A locus).

	Relative′			Genetic type	f	W_1	W_2
Relative	h_{22}	h_{21}	h_{20}	AA	$f_{2.}$	2v	$1/f_{2.}$
	h_{12}	h_{11}	h_{10}	Aa	$f_{1.}$	v-u	$-2/f_{1.}$
	h_{02}	h_{01}	h_{00}	aa	$f_{0.}$	-2u	$1/f_{0.}$
Genetic type	AA	Aa	aa				
f′	$f'_{2.}$	$f'_{1.}$	$f'_{0.}$				
W'_1	2v	v-u	-2u				
W'_2	$1/f'_{2.}$	$-2/f'_{1.}$	$1/f'_{0.}$				

h_{22} is the frequency with which the AA relative appears with the AA relative′. The row and column frequencies add to the marginal frequencies of the relatives

$$f'_{i.} = \sum_{i} h_{ii'}, \qquad f_{i.} = \sum_{i} h_{ii'},$$

which are also the marginal frequencies for the A locus in each of the relative's generation. The correlation between a one-factor scale for the A locus of one relative and a one-factor scale for the A locus of the other relative is

$$\rho_{WW'} = \frac{\sum_{i,i'} h_{ii'} W_i W'_{i'}}{\sqrt{\sum_{i} f_{i.} W_i^2 \sum_{i} f'_{i.} (W'_{i'})^2}} = \frac{\text{Cov } W W'}{\sigma_W \sigma_{W'}}. \tag{1}$$

As an example let the primed relative be the offspring and the unprimed relative be the parent in a randomly mating population where the gene frequency is the same in the two generations. The joint frequencies are: $h_{22} = u^3$, $h_{21} = u^2v$, $h_{20} = 0$, $h_{12} = u^2v$, $h_{11} = uv$, $h_{10} = uv^2$, $h_{02} = 0$, $h_{01} = uv^2$, $h_{00} = v^3$. The correlations are from (1): $\rho_{W_1' W_1} = \frac{1}{2}$, $\rho_{W_1' W_2} = \rho_{W_2' W_1} = \rho_{W_2' W_2} = 0$.

The joint frequencies, $k_{jj'}$, of the two relatives for the B locus are shown similarly in table 7.

When the joint frequencies of relatives for two loci are proportional to the joint frequencies of the relatives for each locus, the joint frequencies for both loci (table 8) can be found by multiplying together the joint frequencies for each locus (tables 6 and 7). They are proportional in randomly mating populations and in inbreeding populations if gene frequencies are not changing, if genotypic frequencies in each generation are proportional to the marginal frequencies of loci and if loci are not linked. The frequency with which the $AABB$ relative appears with the $AABB$ relative′ is $h_{22} k_{22}$, and so on for the other genotypes. The frequency of the $AABB$ relative is $f_{2.} f_{.2} = f_{22}$ and the frequency of the $AABB$ relative′ is $f_{2.}' f_{.2}' = f_{22}'$. The joint frequencies for locus A

C. CLARK COCKERHAM

TABLE 7

Joint frequency of two relatives (B locus).

	Relative$'$			Genetic type	f	W_3	W_4
Relative	k_{22}	k_{21}	k_{20}	BB	$f_{.2}$	$2y$	$1/f_{.2}$
	k_{12}	k_{11}	k_{10}	Bb	$f_{.1}$	$y-x$	$-2/f_{.1}$
	k_{02}	k_{01}	k_{00}	bb	$f_{.0}$	$-2x$	$1/f_{.0}$
Genetic type	BB	Bb	bb				
f'	$f'_{.2}$	$f'_{.1}$	$f'_{.0}$				
W'_3	$2y$	$y-x$	$-2x$				
W'_4	$1/f'_{.2}$	$-2/f'_{.1}$	$1/f'_{.0}$				

are on the left side of the block of joint frequencies for locus B in table 8. To complete the table, the frequency for A must multiply each term in the block.

The covariance between a scale of one relative and that of the other relative is

$$\text{Cov } WW' = \sum_{i,i',j,j'} h_{ii'} k_{jj'} W_{ij} W'_{i'j'}. \tag{2}$$

These covariances simplify when we consider the relationships among the orthogonal scales. Only the first four scales are given for each relative in table 8 since these can be used to construct the remaining four which pertain to epistasis. Note in table 8 (and in table 3) that the individual terms of the scales have the following relationships:

$$W_{ij} = W_{i.}, \text{ or is constant for each i, for scales } W_1 \text{ and } W_2$$
$$\text{relating to the } A \text{ locus}$$
$$= W_{.j}, \text{ or is constant for each j, for scales } W_3 \text{ and } W_4$$
$$\text{relating to the } B \text{ locus,}$$

and remember that

$$W_{ij} = W_{i.} \cdot W_{.j} \text{ for } W_5, W_6, W_7 \text{ and } W_8 \text{ or the epistatic scales.}$$

(The subscripts i and j designate terms of a particular scale.) The same relationships hold among the W's. Thus, $W_{i.} \cdot W_{.j}$ may be substituted for W_{ij} and $W_{i.}' \cdot W_{.j}'$ may be substituted for $W_{i'j'}$ in (2) for the epistatic scales. With these substitutions the covariance in (2) becomes

$$\text{Cov } WW' = \sum_{i,i'} h_{ii'} W_{i.} W'_{i'.} \cdot \sum_{j,j'} k_{jj'} W_{.j} W'_{.j'}. \tag{3}$$

But $\sum_{i,i'} h_{ii'} W_{i.} W'_{i'.}$ is the covariance pertaining to the one-factor scales for locus A in (1), and $\sum_{j,j'} k_{jj'} W_{.j} W'_{.j'}$ is the covariance pertaining to the one-factor scales for locus B, so that the covariances between the epistatic scales can be evaluated in terms of the covariances of the one-factor scales for each locus. For example,

$$\text{Cov } W_5 W'_5 = \sum_{i,i'} h_{ii'} W_{1i.} W'_{1i'} \sum_{j,j'} k_{jj'} W_{3.j} W'_{3.j}$$

$$= \text{Cov. } W_1 W'_1 \text{ Cov } W_3 W'_3.$$

TABLE 8

Joint frequency of two relatives for two loci.

				Relative'									Genetic type	f	W_1	W_2	W_3	W_4
Relative													$AABB$	f_{22}	$2v$	$1/f_{.2}$	$2y$	$1/f_{.2}$
	h_{22}	k_{22}	k_{21}	k_{20}	k_{22}	k_{21}	k_{20}	h_{20}	k_{22}	k_{21}	k_{20}		$AABb$	f_{21}	$2v$	$1/f_{.2}$	$y-x$	$-2/f_{.1}$
		k_{12}	k_{11}	k_{10}	k_{12}	k_{11}	k_{10}		k_{12}	k_{11}	k_{10}		$AAbb$	f_{20}	$2v$	$1/f_{.2}$	$-2x$	$1/f_{.0}$
		k_{02}	k_{01}	k_{00}	k_{02}	k_{01}	k_{00}		k_{02}	k_{01}	k_{00}		$AaBB$	f_{12}	$v-u$	$-2/f_{.1}$	$2y$	$1/f_{.2}$
	h_{12}	k_{22}	k_{21}	k_{20}	k_{22}	k_{21}	k_{20}	h_{10}	k_{22}	k_{21}	k_{20}		$AaBb$	f_{11}	$v-u$	$-2/f_{.1}$	$y-x$	$-2/f_{.1}$
		k_{12}	k_{11}	k_{10}	k_{12}	k_{11}	k_{10}		k_{12}	k_{11}	k_{10}		$Aabb$	f_{10}	$v-u$	$-2/f_{.1}$	$-2x$	$1/f_{.0}$
		k_{02}	k_{01}	k_{00}	k_{02}	k_{01}	k_{00}		k_{02}	k_{01}	k_{00}		$aaBB$	f_{02}	$-2u$	$1/f_{.0}$	$2y$	$1/f_{.2}$
	h_{02}	k_{22}	k_{21}	k_{20}	k_{22}	k_{21}	k_{20}	h_{00}	k_{22}	k_{21}	k_{20}		$aaBb$	f_{01}	$-2u$	$1/f_{.0}$	$y-x$	$-2/f_{.1}$
		k_{12}	k_{11}	k_{10}	k_{12}	k_{11}	k_{10}		k_{12}	k_{11}	k_{10}		$aabb$	f_{00}	$-2u$	$1/f_{.0}$	$-2x$	$1/f_{.0}$
		k_{02}	k_{01}	k_{00}	k_{02}	k_{01}	k_{00}		k_{02}	k_{01}	k_{00}							

Genetic type	AA BB	AA Bb	AA bb	Aa BB	Aa Bb	Aa bb	aa BB	aa Bb	aa bb
f'	$f_{.2}$	$f_{.1}$	$f_{.0}$	$f_{.2}$	$f_{.1}$	$f_{.0}$	$f_{.2}$	$f_{.1}$	$f_{.0}$
W_1'	$2v$	$2v$	$2v$	$v-u$	$v-u$	$v-u$	$-2u$	$-2u$	$-2u$
W_2'	$1/f_{.2}$	$1/f_{.2}$	$1/f_{.2}$	$-2/f_{.1}$	$-2/f_{.1}$	$-2/f_{.1}$	$1/f_{.0}$	$1/f_{.0}$	$1/f_{.0}$
W_3'	$2y$	$y-x$	$-2x$	$2y$	$y-x$	$-2x$	$2y$	$y-x$	$-2x$
W_4'	$1/f_{.2}$	$-1/f_{.1}$	$1/f_{.0}$	$1/f_{.2}$	$-2/f_{.1}$	$1/f_{.0}$	$1/f_{.2}$	$-2/f_{.1}$	$1/f_{.0}$

The covariance between the additive by additive epistatic scales of one relative and the additive by additive epistatic scales of the other relative is the product of the covariance between the additive scales of the two relatives for one locus and of the covariance between the additive scales of the two relatives for the other locus. The same relationship exists between the correlations as between covariances,

$$\rho_{w_5 \ w_5'} = \rho_{w_1 \ w_1'} \, \rho_{w_8 \ w_8'},$$

because $\sigma_{w_5} = \sigma_{w_1} \sigma_{w_8}$ and $\sigma_{w_5'} = \sigma_{w_1'} \sigma_{w_8'}$. The correlations between the eight scales of one relative and the eight scales of the other relative are given in table 4.