Theor Appl Genet (1990) 79:219–224

Effects of epistasis on tests for linkage in self-pollinated species

R.-C. Yang and R. J. Baker

Department of Crop Science and Plant Ecology, University of Saskatchewan, Saskatoon, Saskatchewan S7N OWO, Canada

Received June 5, 1989; Accepted October 20, 1989
Communicated by A. R. Hallauer

Summary. Tests for linkage based on covariances among relatives in self-pollinated species are usually based upon an assumption that epistasis is not important. This study was conducted to determine the impact of epistasis on, and to investigate the sensitivity of, such tests. Thirty covariances were calculated for each of ten non-epistatic and ten epistatic genetic models with varying probabilities of recombination between two coupling or repulsion loci. Each set of covariances was tested for linkage by comparing covariances calculated for the model with those expected for an additive-dominance model with no linkage. Results showed that the test for linkage is quite insensitive to the effects of linkage due to the disproportionate influence of inbreeding. Repulsion linkages should be easier to detect than coupling linkages for all models. Epistasis was found to mimic or counteract the effects of linkage. Tests for linkage based on covariances within a hierarchical mating design appear to be insensitive to linkage and may confuse the effects of linkage and epistasis.

Key words: Linkage – Epistasis – Genetic covariances – Inbred species

Introduction

Linkages among genes at different loci can alter genetic variances of quantitative traits and can affect correlations between traits. Theoretically, coupling linkages among genes that affect a single quantitative trait will result in increased genetic variance, while repulsion linkages will result in decreased genetic variance. Genetic correlations will be positive with coupling linkages between desirable alleles for each of two traits, and negative with repulsion linkages. In general, coupling linkages between desirable alleles for one or more traits will increase response to selection, while repulsion linkages will restrict response.

Linkage disequilibrium must be anticipated in self-pollinated crops where selected populations in most breeding programs are developed by inbreeding after a cross of two (single cross or backcross), three (three-way cross) or four (double cross) inbred lines. Whether or not linkages among genes are primarily in the coupling or repulsion phase will depend upon the nature of association among genes in the inbred parents. A major issue in crop improvement concerns whether or not intermating should be used to help break linkage groups (Hanson 1959; Pederson 1974). Proper resolution of this issue could be facilitated if one could estimate the extent and phase of linkages among genes that affect quantitative traits.

Two basic approaches have been used to assess the importance of linkage in self-pollinated species. One approach is based on theoretical considerations of the impact of linkage on genetic variances and correlations, and consists of comparing estimates of these parameters in inbred and random-mated populations. Miller and Rawlings (1967) showed that natural intercrossing after a cross of two contrasting inbred lines of cotton (*Gossypium hirsutum* L.) helped reduce linkage disequilibrium to the extent that genetic variance was decreased in six traits and increased in one. Genetic correlations in all cases approached what would be expected in an equilibrium population. Meredith and Bridge (1971) used manual intermating in cotton and found that genetic correlations were generally reduced in magnitude.

While comparisons of inbred and intermating populations within the same cross provide a sound basis for detecting linkages among genes, the method is generally too tedious in self-pollinated species, where intermating has to be facilitated through hand crossing.

The second method for detecting linkage is through comparison of covariances between relatives in a hierarchical set of inbred progeny from a cross of two inbred lines. Mather and Jinks (1971, pp. 187–201) discussed a test for presence of linkage among genes controlling ear conformation in barley (*Hordeum vulgare* L.). The analysis was based on a consideration of variation among F_2 plants and among F_3 plants within F_2 families, as well as on covariances between F_2 plants and the means of their F_3 progenies.

Gates et al. (1960) proposed to test for the presence of linkage by comparing estimates of numerous covariances within a hierarchical set with those expected in the absence of linkage. The method is equivalent to that of Mather and Jinks (1971), but utilizes a larger set of covariance estimates. Gates et al. (1960) used their own test to show that repulsion linkages were important among genes controlling height and yield in a soybean (*Glycine max* L.) cross, while coupling linkages were important among genes controlling flowering time. The method has since been used by Croissant and Torrie (1971) to investigate linkage in soybean, and by VandeLogt et al. (1984) to study two barley crosses.

This method of detecting linkage relies on a comparison of estimated covariances between relatives with those expected in the absence of linkage and epistasis. Mather and Jinks (1971) and Opsahl (1956) have shown that it is not possible to clearly differentiate between the effects of linkage and epistasis on different genetic covariances. There is a possibility that certain types of epistasis may affect genetic variances and genetic correlations in the same way as does linkage, while the effects of other types of epistasis may counteract those of linkage. For this reason, the purpose of this research was to investigate the extent to which epistasis can mimic or counteract linkage in its contribution to covariances among relatives in self-pollinated species.

Theory

Let f_i be the frequency of genotype *i* in the F_k generation after selfing a cross between two inbred lines. Further, let X_i be the average genotypic value of all progeny of genotype *i* following a further *n-k* generations of selfing, and let Y_i be the average genotypic value of all progeny of genotype *i* after a further *n′-k* generations of selfing. Then, X_i represents the mean genotypic value of an F_k-derived F_n line, and and Y_i represents the mean genotypic value of the related F_k-derived $F_{n'}$ line. The covariance between pairs of lines, each developed from the same plant in F_k, can be written as $\mathrm{Cov}(k, 1; n, n')$ (Gates et al. 1960). By definition.

$$\mathrm{Cov}\,(k, 1; n, n') = \sum f_i X_i Y_i - (\sum f_i X_i)(\sum f_i Y_i),$$

where $\sum$ indicates summation over all possible genotypes in F_k.

Table 1. Expected frequencies of ten genotypes after *s* generations of selfing of a coupling ($\delta = 1$) or repulsion ($\delta = -1$) double heterozygote

Genotype	Expected frequency after *s* generations of selfing
AB/AB	$(1+\delta)\,a/4 + (1-\delta)\,pa/2 - b/2 - \delta c + d/4$
AB/Ab	$b/2 \quad - d/2$
Ab/Ab	$(1-\delta)\,a/4 + (1+\delta)\,pa/2 - b/2 + \delta c + d/4$
AB/Ab	$b/2 \quad - d/2$
AB/aB	$d/2 + \delta e/2$
Ab/aB	$d/2 - \delta e/2$
Ab/ab	$b/2 \quad - d/2$
aB/aB	$(1-\delta)\,a/4 + (1+\delta)\,pa/2 - b/2 + \delta c + d/4$
aB/ab	$b/2 \quad - d/2$
ab/ab	$(1+\delta)\,a/4 + (1-\delta)\,pa/2 - b/2 - \delta c + d/4$

where p = probability of recombination between A-a and B-b,
$a = 1/[1+2p]$,
$b = 2^{-s}$,
$c = (1+2p)\,[(1-2p)/2]^s/4\,(1+2p)$,
$d = [(1-2p+2p^2)/2]^s$, and
$e = (0.5-p)^s$

In this expression, the second argument (1 in this case) provides facility for subdividing the covariance into portions related to various subpopulations of F_k, i.e., $\mathrm{Cov}\,(k, 1; n, n') = \sum_k \mathrm{Cov}\,(k, k-1; n, n')$ for $k > 1$ (Gates et al. 1960). If $n' = n$, then the expression represents the genetic variance among F_k-derived F_n lines. Furthermore, if $k = n = n'$, then the expression becomes the total genetic variance in the F_k generation. Terminology used by Mather and Jinks (1971) is less general than that used by Gates et al. (1960). $\mathrm{V}_{(k-1)Fn}$ of Mather and Jinks can be written as $\mathrm{Cov}\,(k, k-1; n, n)$ and $\mathrm{W}_{(k-1)Fnn'}$ as $\mathrm{Cov}\,(k, k-1; n, n')$. Mather and Jinks (1971) refer to $k-1$ as the rank of the corresponding variance or covariance, and *n or n′* as its generation. Rank reflects the number of opportunities for recombination through gametogenesis.

It is clear that epistasis cannot affect of genotypic frequencies, f_i, and can therefore have an effect on covariances only through its impact on genotypic means X_i and Y_i. Linkage, on the other hand, can affect all three components of covariances, and is expected to have a major impact on genotypic frequencis. Mather and Jinks (1971) indicated that the test for linkage is primarily a test for heterogeneity of estimates of additive and dominance variance from covariances of different rank.

We proceed now to develop a general expression for $\mathrm{Cov}\,(k, 1; n, n')$, the covariance between progenies of a coss between two homozygous parents, for a two-locus genetic model with arbitrary linkage and epistasis. The frequencies of ten possible genotypes for two linked loci in the F_k generation after selfing a cross between two inbred lines were given by Nelder (1952) and are presented in a modified form in Table 1.

With nine distinct genotypes, genotypic values can be expressed in terms of nine parameters. The choice of parameters is somewhat arbitrary. In this research, we use the F_∞-metric with values of each genotype represented as linear functions of m, d_a, d_b, h_a, h_b, i_{ab}, j_{ab}, j_{ba} and l_{ab} as described by Mather an Jinks (1971, p. 83).

Mean genotypic values of families produced by *s* generations of selfing individuals of each of ten different genotypes are given in Table 2. These expectations allow for arbitrary levels of linkage, arbitrary types of two-locus epistasis, and an arbitrary number of generations of selfing.

Table 2. Mean genotypic values of progeny after s generations of selfing for each of ten genotypes

Genotype of parent	Mean of progeny after s generations of selfing
AB/AB	$m+d_a+d_b+i_{ab}$
AB/Ab	$m+d_a+0.5^s h_b+0.5^s j_{ab}$
Ab/Ab	$m+d_a-d_b-i_{ab}$
AB/aB	$m+0.5^s h_a+d_b+0.5^s j_{ba}$
AB/ab	$m+0.5^s h_a+0.5^s h_b+K_1 i_{ab}+K_2 l_{ab}$
Ab/aB	$m+0.5^s h_a+0.5^s h_b-K_1 i_{ab}+K_2 l_{ab}$
Ab/ab	$m+0.5^s h_a-d_b-0.5^s j_{ba}$
aB/aB	$m-d_a+d_b-i_{ab}$
aB/ab	$m-d_a+0.5^s h_b-0.5^s j_{ab}$
ab/ab	$m-d_a-d_b+i_{ab}$

where $K_1 = (1-2p)\,[1-(0.5-p)^s]/(1+2p)$, and
$K_2 = (0.5-p+p^2)^s$

For an arbitrary two-locus genetic model expressed in terms of the F_∞-metric (Mather and Jinks 1971), the information in Tables 1 and 2 can be used to calculate expected genotypic components of covariances between relatives. For example, to calculate Cov $(k, 1; n, n')$, replace s in Table 1 with $s=k-1$ to calculate the expected frequencies of ten genotypes in the F_k generation after selfing a cross of two inbred parents in coupling (AB/AB × ab/ab; $\delta=1$) or in repulsion (Ab/Ab × AB/aB; $\delta=-1$). Linkage can be specified by the probability of recombination between two loci ($p=0.5$ for no linkage, p near zero for tight linkage). Once the genotypic model has been expressed in terms of the F_∞-metric (Mather and Jinks 1971), the results in Table 2 can be used to calculate X_i, the genotypic means of F_k-derived F_n lines, and Y_i, the genotypic means of related F_k-derived $F_{n'}$ lines. In using Table 2, s, the number of generations of selfing, is taken as $s=n-k$ for X_i and $s=n'-k$ for Y_i. Use of the definition of a covariance leads directly to the required expression.

It is possible to combine the algebra of Tables 1 and 2 to derive a single expression for the genetic composition of Cov $(k, 1; n, n')$ that would be suitable for any linkage value and any arbitrary two-locus genetic model. The complexity of the resulting expression suggests that such an approach is of little practical value. Because of the unwieldly expressions obtained for variances and covariances under a completely general genetic model, Weir and Cockerham (1977) suggested use of simplifying assumptions for approximation in specific cases. Gallais (1974), in deriving a general expression for covariances between inbred relatives, restricted consideration to additive types of epistasis with arbitrary linkage. In this research, we pursued numerical evaluation of a specific set of covariances.

Use of Table 1 allows one to assess the impact of linkage value and phase on the frequencies of genotypes in generation F_k after k-1 generations of selfing from a cross between two homozygous lines. Table 2 then allows one to incorporate the effects of epistasis, as well as linkage, on the genotypic means of F_k-derived F_n lines. Any covariance between progenies of a single cross can be calculated from these two sets of results.

In some cases, it may be desirable to calculate covariances within subpopulation, and to then average those values over all such subpopulations. Gates et al. (1960) used the term Cov $(k, k'; n, n')$ to indicate the average of covariances between F_k-derived F_n and F_k-derived $F_{n'}$ lines within subpopulations derived from individuals in generations $F_{k'}$ $(1 \le k' < k \le n \le n')$. The hierarchical nature of the mating scheme leads to the relationships Cov $(k, k'; n, n')=$ Cov $(k, 1; n, n')-$ Cov $(k', 1; n, n')$. This relationship provides a basis for calculating any covariance from considerations given above.

The test for linkage described by Gates et al. (1960) is based upon a comparison of estimated Cov $(k, 1; n, n')$, with approximate expectations derived under the assumption of no epistasis. Briefly, the method consists of using estimated covariances from a hierarchical mating structure as the response variable (Y) in a multiple regression analysis with six predictor variables. The six predictor variables (Gates et al. 1960) are:

$X_1=2-2^{2-k}$,
$X_2=(2^{k-1}-1)/(2^{n+n'-4})$,
$X_3=k/(2^{k-3})-4$,
$X_4=-[2+2^{k-1}(k-3)]/(2^{n+n'-4})$,
$X_5=8-(k_2-k+2)/(2^{k-3})$, and
$X_6=[2^{k-2}(k^2-7k+14]/(2^{n+n'-4})$.

Provided that estimates of at least seven covariances are available for analysis, the sum of squares for deviations from regression of covariances y on X_1-X_6 provides an estimate of error variation that can be used to test for linkage. The difference in the residual sum of squares for regression on X_1 and X_2 and the residual sum of squares for regression on X_1-X_2 provides a sum of squares for linkage with four degrees of freedom. All regressions are calculated with an intercept of zero (Gates et al. 1960). In the absence of epistasis, significant regression on predictors X_3, X_4, X_5 or X_6 indicates linkage (Gates et al. 1960). The signs and magnitudes of significant regressions on these predictors can provide evidence as to whether linkages are predominantly in the coupling or repulsion phases.

Mather and Jinks (1971) used a similar approach to test for linkage. They used separate estimates of environmental variance, rather than deviations from regression of X_1-X_6, as the denominator for testing significance of linkage. Furthermore, they restricted consideration to covariances in early generations (Mather and Jinks 1971, pp. 187–201), while the method of Gates et al. (1960) can be generalized to any generation. The linkage tests of Gates et al. (1960) and Mather and Jinks (1971) are both based upon analysis of deviations of covariances from what is expected in the absence of linkage and epistasis.

In the absence of linkage and epistasis,

$$\mathrm{Cov}\,(k, 1; n, n')=X_1\,\sigma_A^2+X_2\,\sigma_D^2,$$

where σ_A^2 is the additive genetic variance and σ_D^2 is the dominance genetic variance (Gates et al. 1960). In the absence of environmental variation, the test for linkage requires that the sum of squares due to regression covariances on X_1 and X_2 be significantly less than the uncorrected sum of squares of covariance.

Computations

Twenty genetic models (Table 3), with varying levels of additive, dominance, and epistatic effects, were evaluated at different levels of linkage in both coupling and repulsion crosses, to determine the impact of epistasis on the test for linkage and to assess the sensitivity of the test. The relationships shown in Table 3 were derived from the F_∞-metric (Mather and Jinks 1971). The relative importance of epistasis in each of these 20 models was evaluated by considering the contribution to total variation in an

Table 3. Relationships among gene effects, F_2 segregation ratios, and proportions of total variation in an equilibrium F_2 population accounted for by additive, dominance, and epistatic variation for each of 20 genetic models

Model no.	Relationships	Segregation ratio	Proportion of variation in F_2		
			Additive	Dominance	Epistatic
1	$d_a=d_b;\ h_a=h_b=i_{ab}=j_{ab}=j_{ba}=l_{ab}=0$	1:4:6:4:1	1.00	0.00	0.00
2	$d_a=d_b=h_a=h_b;\ i_{ab}=j_{ab}=j_{ba}=l_{ab}=0$	9:6:1	0.67	0.33	0.00
3	$d_a=d_b=-h_a=-h_b;\ i_{ab}=j_{ab}=j_{ba}=l_{ab}=0$	1:6:9	0.67	0.33	0.00
4	$2d_a=2d_b=h_a=h_b;\ i_{ab}=j_{ab}=j_{ba}=1_{ab}=0$	4:4:1:4:2:1	0.33	0.67	0.00
5	$2d_a=2d_b=-h_a=-h_b;\ i_{ab}=j_{ab}=j_{ba}=l_{ab}=0$	1:2:4:1:4:4	0.33	0.67	0.00
6	$2d_a=d_b\ ;\ h_a=h_b=i_{ab}=j_{ab}=j_{ba}=l_{ab}=0$	1:2:3:4:3:2:1	1.00	0.00	0.00
7	$2d_a=d_b=2h_a=h_b;\ i_{ab}=j_{ab}=j_{ba}=l_{ab}=0$	9:3:3:1	0.70	0.30	0.00
8	$2d_a=d_b=-2h_a=-h_b;\ i_{ab}=j_{ab}=j_{ba}=l_{ab}=0$	1:3:3:9	0.70	0.30	0.00
9	$4d_a=2d_b=2h_a=h_b;\ i_{ab}=j_{ab}=j_{ba}=l_{ab}=0$	1:2:2:4:1:2:1:2:1	0.30	0.70	0.00
10	$4d_a=2d_b=-2h_a=-h_b;\ i_{ab}=j_{ab}=j_{ba}=l_{ab}=0$	1:2:1:2:1:4:2:2:1	0.30	0.70	0.00
11	$d_a=d_b=h_a=h_b=i_{ab}=j_{ab}=j_{ba}=l_{ab}$	9:7	0.57	0.28	0.15
12	$d_a=d_b=i_{ab}=l_{ab};\ h_a=h_b=j_{ab}=j_{ba}=0$	1:8:6	0.70	0.09	0.21
13	$d_a=d_b=h_a=h_b=-i_{ab}=-j_{ab}=-j_{ba}=-l_{ab}$	15:1	0.26	0.13	0.61
14	$-d_a=d_b=-h_a=h_b=i_{ab}=j_{ab}=j_{ba}=l_{ab}$	13:3	0.49	0.26	0.25
15	$d_b=h_b=i_{ab}=j_{ab}=j_{ba}=l_{ab};\ d_a=h_a=0$	9:3:4	0.51	0.26	0.23
16	$d_a/3=h_a=d_b/3=h_b=i_{ab}=j_{ab}=j_{ba}=l_{ab}$	12:3:1	0.60	0.30	0.10
17	$-d_a=d_b=-h_a=i_{ab}=j_{ba};\ h_b=j_{ab}=l_{ab}=0$	7:6:3	0.72	0.11	0.17
18	$d_a=d_b/3=h_a=h_b/5=i_{ab}=j_{ab}/3=j_{ba}=l_{ab}/3$	6:3:3:4	0.39	0.52	0.09
19	$d_a=h_a=h_b=i_{ab}=-j_{ab}/4=j_{ba}=-l_{ab}/4;\ d_b=0$	7:4:3:2	0.14	0.11	0.75
20	$d_a=d_b=h_a=h_b=i_{ab}/2=j_{ab}/2=j_{ba}/2=l_{ab}/2$	9:1:6	0.48	0.24	0.28

equilibrium F_2 population due to additive, dominance, and epistatic effects (Table 3). Models 1 through 10 were non-epistatic and differed only in the levels of dominance and the size of effects at the two loci. The remaining ten models (Models 11–20) included various levels of epistasis and were chose to represent known models of epistasis. Model 13 (duplicate epistasis) and Model 19 were the most epistatic of the models considered, with epistasis accounting for more than 60% of the equilibrium F_2 genetic variance in both cases.

For each of the 20 genetic models, 30 Cov (k, k'; n, n') were calculated for 15 levels of coupling and repulsion linkages varying from $p=0.005$ to $p=0.50$. Covariances were those used by Gates et al. (1960) in testing for linkage in soybean, and consisted of partial combinations of $k=2$ to 6, $n=k$ to 7 and $n'=n$ to 7. For each set of 30 covariances, multiple regression (with intercept = 0) of covariances on predictors X_1 and X_2 was calculated. The sum of squares due to regression was expressed as a proportion of the total uncorrected sum of squares of covariances.

Results and discussion

Proportions of the total uncorrected sums of squares of covariances attributable to regression on X_1 and X_2 were generally close to unity for all models and all probabilities of recombination between the two loci (Table 4). Results can be adequately represented by presentation of 7 of the 15 cases that were evaluated. For models without epistasis (Models 1 through 10), additive and dominance effects accounted for all of the sums of squares when there was no linkage. Comparison of the left side of Table 4 with its right side showed that repulsion linkages generally led to greater deviations from values expected in the absence of linkage than did coupling linkages.

Except for Model 1, all non-epistatic models showed greatest deviations when probabilities of recombination were at intermediate levels. In these cases, tight linkages make the two-locus model appear similar to a single locus, while probabilities of recombination close to 0.5 resembled two unlinked loci. Model 1 consists of two loci with equal and additive effects. With this model, discrepancies due to linkage were large with tight repulsion linkages than with intermediate linkages This difference from the typical pattern for non-epistatic models probably reflects the fact that inbreeding will have less of an impact on the relationship between covariances for an additive model than for any model that includes dominance genetic effects.

For Models 1–10, the greatest discrepancy from an additive-dominance model was observed with tight repulsion linkages for Model 1 (proportion = 0.968 for

Table 4. Proportion of the total uncorrected sum of squares of 30 covariances that can be attributed to additive and dominance effects for 20 models and various probabilities of recombination

Model[a]	Probability of recombination						
	Repulsion				Coupling		
	0.025	0.10	0.30	0.50	0.30	0.10	0.025
1	0.9680	0.9776	0.9946	1.0000	0.9986	0.9995	1.0000
2	0.9927	0.9845	0.9952	1.0000	0.9986	0.9995	1.0000
3	0.9927	0.9845	0.9952	1.0000	0.9986	0.9995	1.0000
4	0.9993	0.9951	0.9966	1.0000	0.9987	0.9995	1.0000
5	0.9993	0.9951	0.9966	1.0000	0.9987	0.9995	1.0000
6	0.9985	0.9942	0.9971	1.0000	0.9990	0.9996	1.0000
7	0.9990	0.9954	0.9974	1.0000	0.9990	0.9996	1.0000
8	0.9990	0.9954	0.9974	1.0000	0.9990	0.9996	1.0000
9	0.9997	0.9979	0.9981	1.0000	0.9991	0.9996	1.0000
10	0.9997	0.9979	0.9981	1.0000	0.9991	0.9996	1.0000
11	0.9987	0.9955	0.9965	0.9985	0.9996	1.0000	1.0000
12	0.9814	0.9744	0.9871	0.9965	0.9998	1.0000	1.0000
13	0.8701	0.8891	0.9376	0.9797	0.9977	1.0000	1.0000
14	1.0000	1.0000	0.9999	0.9968	0.9864	0.9724	0.9663
15	0.9998	0.9991	0.9981	0.9976	0.9966	0.9975	0.9995
16	1.0000	0.9997	0.9991	0.9994	0.9997	0.9998	1.0000
17	1.0000	1.0000	0.9999	0.9989	0.9950	0.9886	0.9910
18	0.9999	0.9992	0.9983	0.9984	0.9991	0.9998	1.0000
19	0.9981	0.9822	0.9145	0.8623	0.8719	0.9536	0.9953
20	0.9982	0.9939	0.9942	0.9957	0.9969	0.9984	0.9997

[a] Genetic models from Table 3

$p = 0.025$, Table 4). With such small deviations from the non-linkage model (3.2% of the total sum of squares of covariances), there is a question as to whether or not the test is sufficiently sensitive to detect linkage. In the study by Gates et al. (1960), linkage was detected for flowering time, height, and yield. For flowering time, the estimated ratio of sum of squares due to regression on X_1 and X_2 to the sum of squares due to regression on $X_1 - X_6$ is $(6*1{,}013.51 - 4*64.844)/(6*1{,}013.51) = 5{,}818.624/6{,}081.060 = 0.957$. Corresponding ratios were 0.948 for height and 0.880 for yield. While not strictly equivalent to values reported in Table 4, the similarity in magnitude serves to emphasize the observation that, for two-locus models, linkage in the absence of epistasis has rather little impact on the relationships among covariances in a hierarchical mating design.

In this case, inbreeding may overpower the impact of linkage on covariances. Heterozygosity decreases one-half in each successive generation of selfing. The effect of linkage can be expressed only when the double heterozygotes are present. Thus, the impact of linkage on the covariances decreases as the generation number increases. In the absence of epistasis ($i_{ab} = j_{ab} = j_{ba} = l_{ab} = 0$), the general expression for Cov $(k, k-1; n, n')$ is found, by using the results of Tables 1 and 2, to be

$$\mathrm{Cov}\,(k, k-1; n, n') = (0.5)^{k-1}\, D_{k-1} + (0.5)^{n+n'-k}\, H_{k-2},$$

where $k-1$ is the rank of the covariance (Mather and Jinks 1971),

$$D_{k-1} = (d_a^2 + d_b^2) \pm 2\,(1-2p)^{k-1}\, d_a d_b, \text{ and}$$
$$H_{k-2} = (h_a^2 + h_b^2) + 2\,(1-2p)^2\,(1-2p+2p^2)^{k-2}\, h_a h_b.$$

The ± sign in D_{k-1} should be + for coupling and − for repulsion linkages. From this expression, one can see that cross products of additive and dominance effects at different loci occur only in the presence of linkage, and that these cross products decrease geometrically as the rank $(k-1)$ of the covariances increase arithmetically.

Comparisons of discrepancies due to epistasis in the absence of linkage (Models 11–20 with probability of recombination = 0.50) with discrepancies due to linkage in the absence of epistasis (Models 1–10) show that certain types of epistasis can mimic the effects of linkage on the proportions of the total sum of squares attributable to an additive-dominance model. With model 19, e.g., in the absence of linkage, only 86.23% of the total sum of squares could be explained by an additive-dominance model. With the linkage test of Gates et al. (1960), such a discrepancy (13.77%) would be incorrectly attributed to linkage of non-epistatic genes rather than to epistatic interaction among unlinked genes. In this extreme case, epistasis of unlinked genes produced much larger effects on the magnitude of covariances than did linkage of genes in all non-epistatic models. Duplicate epistasis (Model 13) can also cause discrepancies in the absence of linkage that are as large as those due to linkage in non-epistatic models. This finding raises considerable doubt as to whether purported detection of linkage is, in fact, due to linkage or to epistasis. These results suggest that analysis of covariances between inbred relatives is not a reliable test for linkage.

It is interesting to note that the effects of epistasis can sometimes cancel the effects of linkage. With Models 14 and 17, for example, the proportions of total sum of squares attributable to an additive-dominance model approached unity as the level of repulsion linkage increased (Table 4). This is yet another weakness of the test for linkage.

This research has concerned a general treatment of only a two-locus model. It is felt that similar trends would be observed with more complex genetic models. With several loci, the frequency of repulsion linkages would be reduced, so that discrepancies due to linkage would likely be less than calculated for the two-locus model. It is unlikely that the test for linkage would be more sensitive for linkages among several loci than it appears to be for linkage between two loci.

It is more difficult to assess the probable effects of multi-locus epistatic models on tests for linkage. There is a great difficulty in trying to define what would be reasonable epistatic interaction for a multi-locus model. However, Crow and Kimura (1970) argued that three-locus and higher order interactions probably have smaller impact on genetic variances than interactions between pairs of loci.

In this research, it was assumed that environment had no effect on covariances between relatives. Even if this were true, it would be difficult to detect linkage in any of the genetic models considered. With environmental variation and genotype-environmental interaction, it would be even more difficult to identify patterns in estimates of genetic covariances that would clearly indicate the effects of linkage.

This research has shown that tests for linkage based on analysis of covariances among relatives in a hierarchical mating design are rather insensitive to linkage. This insensitivity is due in part to the overpowering effect of inbreeding on the relationships among the different covariances that might be studied. Moreover, epistasis can mimic or counteract the effects of linkage, even when genes are not linked. For these reasons, these types of tests hold little promise for clearly identifying cases in which linkage is important. One should proceed to test for linkage only after showing that epistasies is unimportant. Unfortunately, unequivocal tests for epistasis, such as that decribed by Kearsey and Jinks (1968), require considerable effort and are only suited to a few crops. Perhaps the approach of comparisons among inbred and random-mated populations continues to be the best method for detecting linkage.

Acknowledgements. This research was supported in part by an Operating Grant from the Natural Sciences and Engineering Research Council of Canada and in part by funding from the Agricultural Development Fund of the Government of Saskatchewam.

References

Croissant GL, Torrie JH (1971) Evidence of nonadditive effects and linkage in two hybrid populations of soybeans. Crop Sci 11: 675–677

Grow JF, Kimura M (1970) An introduction to population genetics theory. Harper and Row, New York

Gallais A (1974) Covariances between arbitrary relatives with linkage and epistasis in the case of linkage disequilibrium. Biometrics 30: 429–446

Gates CE, Weber CR, Horner TW (1960) A linkage study of quantitative characters in a soybean cross. Agron J 52: 45–49

Hanson WD (1959) The breakup of initial linkage blocks under selected mating systems. Genetics 44: 857–868

Kearsey MJ, Jinks JL (1968) A general method of detecting additive, dominance and epistatic variation for metrical traits. I. Theory. Heredity 23: 403–409

Mather K, Jinks JL (1971) Biometrical genetics. Chapman and Hall, London

Meredith WR Jr, Bridge RR (1971) Breakup of linkage blocks in cotton, *Gossypium hirsutum* L. Crop Sci 11: 695–698

Miller PA, Rawlings JO (1967) Breakup of initital linkage blocks through intermating in a cotton breeding population. Crop Sci. 7: 199–204

Nelder JA (1952) Some genotypic frequencies and variance components occurring in biometrical genetics. Heredity 6: 387–394

Opsahl B (1956) The discrimination of interaction and linkage in continuous variation. Biometrics 12: 415–432

Pederson DG (1974) Arguments against intermating before selection in self-fertilizing species. Theor Appl Genet 45: 157–162

VandeLogt MB, Brinkman MA, Forsberg RA (1984) Variance components and linkage effects for quantitatively inherited traits in two barley crosses. Crop Sci 24: 323–327

Weir BS, Cockerham CC (1977) Two-locus theory in quantitative genetics. In: Pollack E, Kempthorne O, Bailey TB Jr (eds) Proc Int Conf on Quantit Genet. Iowa State University Press, Ames/IA, pp 247–269