A GENERAL METHOD OF DETECTING ADDITIVE, DOMINANCE AND EPISTATIC VARIATION FOR METRICAL TRAITS

II. APPLICATION TO INBRED LINES

J. L. JINKS, JEAN M. PERKINS and E. L. BREESE

Dept. of Genetics, University of Birmingham, and Welsh Plant Breeding Station

1. INTRODUCTION

KEARSEY and Jinks (1968) have described an extension of the design III experiment of Comstock and Robinson (1952) that is applicable to any population irrespective of its mating system and its gene and genotype frequencies. The extended analysis provides tests for the presence of epistatic variation in addition to estimates of the additive and dominance components of variation when epistasis is absent. In the present paper we shall apply this analysis to a population of inbred lines. A special test for detecting epistasis under these circumstances will be developed which is only applicable to inbred lines. The analysis will be illustrated by reference to two experiments on inbred lines of *Nicotiana rustica* and the results will be compared with those obtained from earlier investigations of these lines using alternative analyses.

2. THEORY

The experimental design consists of crossing each of n inbred lines to two inbred tester lines, L_1 and L_2, chosen because they are the opposite extreme phenotypes available for the character under investigation to yield $2n$ progeny families each of size r. If each cross is carried out reciprocally the number of families is doubled. The progeny families along with representatives of the n inbred lines are grown in a replicated block design.

The analysis falls into two parts: (i) the test for epistasis, (ii) the estimation of the additive and dominance components if epistasis is absent.

(i) *Test for epistasis*

The general test for epistasis which has been described by Kearsey and Jinks (1968) and which can be used with any kind of population requires a more complex breeding programme than that just described. In the present paper we will use a simpler, albeit, a less efficient test that is applicable only to inbred lines and depends on the fact that for the ith inbred line (where $i = 1$ to n)

$$\bar{L}_{1i} + \bar{L}_{2i} - \bar{P}_i = \mu + \sum_j h_j$$

where $(\mu + \sum_j h_j)$ is a constant for all n lines in the absence of epistasis, j being the number of loci at which L_1 and L_2 differ. $\bar{L}_{1i}$ is the mean of the progeny of the cross between the ith inbred line and the tester L_1, $\bar{L}_{2i}$, is the mean of the progeny of the cross between the ith inbred line and the tester L_2 and $\bar{P}$ is the mean of the ith inbred line.

It follows, therefore, that the variance of $(\bar{L}_{1i} + \bar{L}_{2i} - \bar{P}_i)$ over the $i = 1$ to n sets of progenies around the overall mean $(\mu + \sum_j h_j)$ is zero in the absence of epistasis irrespective of the gene and genotype frequencies among the inbred lines. Hence, if this variance is significantly greater than the error variance (σ^2) we have unambiguous evidence of the presence of non-allelic interactions.

The contribution that non-allelic interactions make to the variance of $(\bar{L}_{1i} + \bar{L}_{2i} - \bar{P}_i)$ will be illustrated for the two gene case with arbitrary gene frequencies u and v and a random distribution of the genes among the inbred lines. The relevant statistics are summarised in table 1.

TABLE 1

The expected contributions of non-allelic interactions to $(\bar{L}_{1i} + \bar{L}_{2i} - \bar{P}_i)$
in the two gene case with random gene distributions, where the genotypes
of L_1 and L_2 are AABB and aabb, respectively

Genotype of inbred Lines	Frequency	$\bar{L}_{1i} + \bar{L}_{2i} - \bar{P}_i$
AABB	$u_a u_b$	l_{ab}
AAbb	$u_a v_b$	$j_{ab} - j_{ba} - i_{ab}$
aaBB	$v_a u_b$	$-j_{ab} + j_{ba} + i_{ab}$
aabb	$v_a v_b$	l_{ab}

The variance of $\bar{L}_{1i} + \bar{L}_{2i} - \bar{P}_i$ is then

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

This variance is independent of additive and dominance effects and depends solely on the presence of epistatic effects. It should be noted, however, that in the special case where $i_{ab} = l_{ab}$ and $j_{ab} = j_{ba}$ this variance reduces to zero even though epistasis is present. It follows, therefore, that when these relationships hold for all pairs of interacting genes this test for epistasis will fail to detect its presence; a weakness it shares with most tests for epistasis.

A further weakness of this test is that if, in spite of L_1 and L_2 being the most extreme phenotypes available, they are homozygous for the same allele at loci for which the lines under test carry different alleles, then $\bar{L}_{1i} + \bar{L}_{2i} - \bar{P}_i$ is no longer constant even in the absence of non-allelic interactions. Hence, under these circumstances it is no longer an unambiguous method for detecting non-allelic interactions. These circumstances, however, do not upset the more general test of Kearsey and Jinks (1968). Where L_1 and L_2 are chosen from a collection of inbred lines with diverse origins the risk that they will not differ at all loci at which the lines under test differ is greater than where they are extreme selections derived from the same initial population as the lines under test (see, for example, section 3 experiments A and B).

(ii) *Estimation of Additive and Dominance effects*

When epistasis is absent the analysis follows the procedure described by Kearsey and Jinks. The nature of the expectations, however, depends on the composition of the inbred population. If the genes are distributed at random then

$$\sigma_m^2 = \sum_j uv d_j^2 = \tfrac{1}{4}D$$

and
$$\sigma^2_{ml} = \sum_j uvh^2_j = \tfrac{1}{4}H_1$$

if the genes are not distributed at random

$$\sigma^2_m = \sum_j uvd^2_j + 2\sum_{jk} D_{jk}\, d_jd_k = \tfrac{1}{4}D'$$

and
$$\sigma^2_{ml} = \sum_j uvh^2_j \pm 2\sum_{jk} D_{jk}\, h_jh_k = \tfrac{1}{4}H'_1$$

Where D_{jk} is the linkage disequilibrium between the jth and kth loci in the gametes which produce the population.

Since L_1 and L_2 will have been chosen so that the maximum number of genes will be in the association phase the sign of $2\sum_{jk} D_{jk}\, h_j h_k$ will be positive. Hence in both situations σ^2_{ml}/σ^2_m will measure the dominance ratio. In all cases σ^2 will be an estimate of E_2.

One further useful statistic can be obtained that was not discussed by Kearsey and Jinks, this is the covariance of $\bar{L}_{1i}+\bar{L}_{2i}$ on $\bar{L}_{1i}-\bar{L}_{2i}$ for all values of i. In the absence of epistasis and correlated gene distributions this covariance has the expectation:

$$\text{cov}_{\text{sum/diff}} = -\sum_j uvd_jh_j = -\tfrac{1}{4}F$$

F, therefore, has the same coefficient as D and H_1 but it measures the sum of the products of the d and h terms. Both the magnitude and the sign of the covariance provide information about the magnitude and direction of dominance which supplements that obtained from σ^2_{ml}. To determine whether or not the covariance is significant it can be converted into a correlation coefficient with $(n-3)$ degrees of freedom.

$$r_{\text{sum/diff}} = \frac{\text{Cov}_{\text{sum/diff}}}{\sqrt{V_{\text{sum}} \times V_{\text{diff}}}}$$

which in terms of our model equals

$$\frac{\sum uvd_jh_j}{(\sum_j uvd^2_j + \tfrac{1}{2}r\sigma^2)(\sum_j uvh^2_j + \tfrac{1}{2}r\sigma^2)}$$

A number of situations can occur in practise each of which has its own unique interpretation. These are:

(1) *σ^2_{ml} is significant and $r_{\text{sum/diff}}$ is significant*

This means that there is a dominance contribution to the variation and the dominance is predominantly in one direction. The nearer that $r_{\text{sum/diff}}$ approaches a value of one the greater the directional element. By examining the sign of F (which is the opposite of the sign of the covariance) the predominant direction of the dominance effects can be determined. If F is positive then the increasing allele is dominant more often than the decreasing allele; if F is negative, the reverse is true.

(2) *σ^2_{ml} is significant and $r_{\text{sum/diff}}$ is non-significant*

This means that there is a dominance contribution to the variation but the dominance is ambidirectional, increasing and decreasing alleles being dominant and recessive to the same extent.

(3) σ^2_{ml} *is non-significant and* $r_{\mathrm{sum/diff}}$ *is non-significant*

This means that there is no evidence of a dominance contribution to the variation. The only remaining possibility, where σ^2_{ml} is non-significant and $r_{\mathrm{sum/diff}}$ significant, is trivial and could only arise as a result of sampling error.

3. ANALYSIS

A. *Experiment* 1. Population of 20 inbred lines with equal gene frequencies.

The population of inbred lines consisted of 20 F_9 families each derived from an independent plant of an F_2 of a cross between varieties 1 and 5 of *Nicotiana rustica* (Mather and Vines, 1952). These 20 lines were crossed in all possible combinations, including reciprocal crosses, to give a diallel set. Our immediate interest, however, is with the crosses between the 20 lines and the two extreme lines for the characters final height and flowering time. These extreme lines proved to be the same for both characters, namely, lines 5 and 17. The crosses along with the 20 inbred lines, which were now F_{10}'s, were grown in two replicate randomised blocks, each block containing one plot of five plants for each of the reciprocal crosses and two plots of five plants for each of the inbred lines.

Because all the plants in any one family had correlated environments during the seedling stage prior to planting in the experimental field (Jinks, 1956) the difference between reciprocal crosses within blocks, which were independently randomised at all stages, will be used as the source of the error variance (σ^2).

(i) *Test for epistasis*

The variances of $(\bar{L}_{1i}+\bar{L}_{2i}-\bar{P}_i)$ for $i = 1$ to 20 for final height and for flowering time were non-significant when tested against the appropriate error variance (table 2). There is, therefore, no reason to believe that an additive-dominance model is inadequate for either character.

TABLE 2

Test for epistasis

	Final height	Flowering time
Expt. 1	N.S.	N.S.
Expt. 2		
Season 1	*	+
2	*	N.S.
3	*	N.S.

significant epistasis, * = $P < 0.1\%$; + = P near 1%.

(ii) *Fitting an additive-dominance model*

The analysis of variance of sums $(L_{1i}+L_{2i})$ and differences $(L_{1i}-L_{2i})$ for the two characters are summarised in table 3. The analyses have been carried out on plot means hence r is effectively 1. Since each cross is represented by four plots (two blocks, two reciprocals per block) the coefficient of both σ^2_m and σ^2_{ml} is 8.

The sum and difference mean squares are highly significant when tested against the error mean squares for both characters. Hence, there is a

significant additive and dominance component of variation for both characters. The estimate of the σ^2's and the components derived from them, namely, D, H_1, and E_2 are given in tables 3 and 4 respectively. The estimates of F obtained from the covariance of sums and differences are also given in table 4. For final height the correlation between sums and differences is non-significant (($r = -0.156$, $P > 0.20$) but for flowering time it is significant

TABLE 3

Experiment 1. Analysis of variance for sums and differences

| | Final height | | Flowering time | | |
	df	MS	df	MS	E.M.S.
Sums	19	77·59*	19	25·18*	$\sigma^2 + 8\sigma_m^2$
Differences	19	16·06*	19	10·20*	$\sigma^2 + 8\sigma_{ml}^2$
Blocks	1	68·38†	1	62·13*	$\sigma^2 + 80\sigma_B^2$
Error	115	9·46	111	3·58	σ^2
$\hat{\sigma}_m^2$		8·52		2·70	
$\hat{\sigma}_{ml}^2$		0·83		0·83	
$\hat{\sigma}^2$		9·46		3·58	

*, P = <0·1%; †, P = 0·1−1·0%.

TABLE 4

Experiment 1. A comparison between the estimates of the additive, dominance and environmental components of variation obtained in the new analysis and those previously obtained from the analysis of the diallel

| | Final height | | Flowering time | |
Component	Diallel	New analysis	Diallel	New analysis
D	36·49	34·06	14·20	10·80
H_1	3·60	3·30	1·91	3·31
H_2	0·28	—	1·29	—
F*	12·20	2·76	6·07	−5·21
E_2	9·70	9·46	2·98	3·58

* F has a different expectation in the diallel and in the new analysis.

($r = 0.650$, $P = 0.01$-0.001). Since for height, F is not significantly different from zero there is no evidence of directional dominance. There is, however, clear evidence of a low level of dominance for height (dominance ratio $\sqrt{\dfrac{H_1}{D}} = 0.31$) hence the dominance must be ambidirectional.

For flowering time F is significant and negative. Hence the low level of dominance (dominance ratio $\sqrt{\dfrac{H_1}{D}} = 0.55$) must result from a preponderance of dominant decreasing genes, *i.e.* there is a low level of dominance for the early flowering genes.

D

B. *Experiment* 2. Population of eight inbred lines of diverse origins.

The eight inbred lines were a representative sample from a world col-
lection (over 60) of inbred lines of *Nicotiana rustica* maintained at Birmingham.
Each of the eight lines were crossed to the two extremes which were lines 4
and 7 for final height and lines 3 and 7 for flowering time.   All crosses were
made reciprocally and were grown as part of a much larger crossing pro-
gramme in three successive seasons.   Each cross was represented by a single
plot of five plants in each of two replicate randomised blocks.   Each inbred
line was also represented by one plot of five plants per block in the first two
seasons and by two plots of five plants per block in the third.   As in experi-
ment 1, the difference between reciprocal crosses was used to estimate the
error variance to overcome the effects of correlated seedling environments on
the replicate block differences.

(i) *Test for epistasis*

The results of the test for the presence of epistasis for both characters in
each of three seasons are summarised in table 2.   Final height consistently
shows significant variation for $\bar{L}_{1i} + \bar{L}_{2i} - \bar{P}_i$ over the eight lines in all three
seasons while flowering time shows significant variation in the first season
only.   If we are correct in assuming that the varieties chosen to be L_1 and L_2
differ at all loci, this result indicates the presence of epistasis for height in all
three seasons and for flowering time in the first season.   On the other hand, a
part or the whole of the significant variation could, in the present instance,
be due to the failure of this assumption (see section 2 (i)).   It is extremely
unlikely, however, that such a failure could account for all or even most of
the variation because of its magnitude.   Thus for final height this variation
accounts for over a quarter of the variation ascribable to the additive and
dominance components of variation and this would require that the tester
lines had identical alleles at half the loci for which the lines under test differ.
We can conclude, therefore, that the significant variation for $\bar{L}_{1i} + \bar{L}_{2i} - \bar{P}_i$
must largely be due to epistasis.   There would appear, therefore, to be
little point in fitting an additive and dominance model to final height.

For flowering time there is evidence of epistasis in the first but not in the
last two seasons.   However, the evidence from other analyses of flowering
time presented in section 4 (ii) is inconsistent, some tests showing the presence
and others the absence of epistasis in each of the three seasons.   It seems
likely, therefore, that the level of epistasis is low in that it only reaches
significance in certain tests in certain seasons.   We shall, therefore, estimate
the additive and dominance components for all of them.

(ii) *Fitting an Additive-dominance model*

The analysis of variance of sums and differences for flowering time in the
three seasons are summarised in table 5.   Both the sum and difference mean
squares are significant when tested against the error mean squares in all
seasons.   Hence, there are significant additive and dominance components
of variation in these seasons.

The estimates of the σ^2's and the estimates of D, H_1, and E_2 derived from
them are given in tables 5 and 6 respectively.   Also included in table 6 are
the estimates of F obtained from the covariance of sums and differences.
The correlation coefficients testing the significance of F are significant in all

seasons ($r = 0.67$, 0.93 and 0.91 respectively). Hence, F is consistently negative and significantly different from zero. There is, therefore, incomplete dominance ($\sqrt{\dfrac{H_1}{D}}$, the dominance ratio, equals 0.47, 0.50 and 0.74)

TABLE 5

Experiment 2. Analysis of variance of sums and differences

	Flowering time						
	1951		1952		1953		E.M.S.
	df	MS	df	MS	df	MS	
Sums	7	282·42*	7	84·29*	7	324·24*	$\sigma^2 + 8\sigma^2_m$
Differences	7	66·42*	7	29·37‡	7	186·46*	$\sigma^2 + 8\sigma^2_{ml}$
Error	42	5·78	42	10·71	44	15·33	σ^2
$\hat{\sigma}^2_m$		34·58		9·20		33·61	
$\hat{\sigma}^2_{ml}$		7·58		2·33		21·39	
$\hat{\sigma}^2$		5·78		10·71		15·33	

*, P = <0·1%; ‡, P = 1-5%

TABLE 6

Experiment 2. A comparison between the estimates of the additive, dominance and environmental components of variation for flowering time obtained in the new analysis and those previously obtained from the analysis of the diallel

	Season									
	1		2				3			
				Diallel				Diallel		
Component	Diallel	New analysis	Perfect fit	Least squares	New analysis		Perfect fit	Least squares	New analysis	
D	102·57	138·32	16·59	20·03 ± 3·16	36·79		147·95	131·98 ± 3·93	154·45	
H_1	51·73	30·32	8·40	4·43 ± 20·07	9·33		88·02	75·08 ± 21·84	85·56	
H_2	35·95	—	7·74	2·16 ± 20·36	—		69·30	65·55 ± 25·27	—	
F*	16·57	−45·96	−0·40	5·29 ± 14·45	−23·04		40·47	31·08 ± 17·00	−111·58	
E_2	2·56	5·78	3·37	3·37†	10·71		5·11	5·11†	15·33	

* F has a different expectation in the diallel and new analysis.

† The perfect fit solution of E_2 was used to obtain the least squares estimates of the other parameters (Jinks, 1956).

in the direction of early flowering. This conclusion agrees with that from the analysis of the 20 inbred lines (section 3, expt. 1 (ii)). It should, however, be borne in mind that the additive and dominance components in the 8×8 diallel will be biassed to some extent by the low level of epistasis that is detected by one or more of the tests in every season.

4. COMPARISON WITH ALTERNATIVE ANALYSES

The population of 20 inbred lines of experiment 1 and the eight inbred lines of experiment 2 have been extensively investigated by alternative and

more complex breeding programmes than those described in the present paper. We can, therefore, compare the outcome of the past and present analyses.

(i) *Experiment* 1

The 20 inbred lines have been investigated by a complete 20×20 diallel set of crosses, including reciprocals, grown in two replicated blocks. The Hayman (1954) analysis of variance of the data for final height and flowering time showed the presence of significant additive and non-additive components of variation for both characters. The Wr/Vr analysis (Jinks, 1954) showed that the non-additive variation for height could be traced to a low level of ambidirectional dominance, there being no evidence of non-allelic interactions. For flowering time, on the other hand, the Wr/Vr analysis indicates a low level of dominance in the direction of early flowering but in addition there is some suggestion of the presence of non-allelic interactions or of correlated gene distributions (P = 0·05); Wr having a curvilinear relationship to Vr of a kind that is consistent with the presence of duplicate gene interactions.

The estimates of the components of variation (Jinks, 1954), assuming an additive-dominance model, are given for both flowering time and height in table 4. Three of these diallel components, namely, D, H_1 and E_2 have the same expectations as the D, H_1 and E_2 of the new analysis (table 4) hence they can be compared directly. The two F's, however, have different expectations and are therefore not directly comparable, the new F being $\Sigma \ 4uvd_jh_j$ and the diallel F, $\Sigma \ 8uv(u-v)d_jh_j$. Irrespective of the direction of dominance, the two F's will always have the same sign if u on average is greater than v and the opposite signs if u on average is smaller than v. As u approaches v in value *i.e.* equal gene frequencies, the new F will be larger than the diallel F. Whereas, when u or v approaches unity the F of the new analysis will be the smaller.

In the present instance the new F is not significantly different from zero for height, indicating ambidirectional dominance. For flowering time the F of the new analysis is significant and negative while that for the diallel is the same order of magnitude but positive. This result is expected when the decreasing genes, which for flowering time are the dominant genes, are somewhat more frequent than their increasing alleles.

The D, H_1 and E_2 estimates for height and flowering time for the 20×20 diallel set of crosses and for the 2×20 set of the new analysis (table 4) are virtually identical. Hence we can conclude that the new analysis of the much smaller experiment has yielded the same information as the more extensive diallel analysis.

(ii) *Experiment* 2

The eight inbred lines of the second experiment have also been investigated by a complete 8×8 diallel set of crosses grown in three successive seasons (Jinks, 1954, 1956). The results of the analyses have been reported in a number of papers (Jinks, 1954, 1956; Jinks and Jones, 1958; Jinks and Stevens, 1959; Hayman, 1954, 1957) and only the main conclusions will be summarised here.

There is consistent evidence from four sources that there are non-allelic interactions among the genes controlling height in each of the three seasons.

1. The linear relationship between Wr and Vr, which is expected if non-allelic interactions are absent, was not obtained in any season. (Jinks, 1954, 1956).

2. The individual and joint scaling tests, which uniquely detect the presence of non-allelic interactions, indicated its presence in 11 out of 28 crosses and 14 out of 28 crosses of the diallel set in the two seasons in which they could be carried out (Jinks, 1956; Jinks and Jones, 1958).

3. Estimates of the non-allelic interaction components of the generation means were significantly different from zero for the same crosses of the diallel set that showed the presence of interactions on the scaling tests (Jinks and Jones, 1958; Jinks and Stevens, 1959).

4. A model which assumed the absence of a non-allelic interaction component of the variation within segregating generations was inadequate in the two seasons in which it could be tested. Since all other factors which could adversely affect the fit of the model, including linkage, were allowed for its failure is unambiguous evidence of the presence of non-allelic interactions (Jinks, 1956).

The new analysis agrees with all the previous evidence in showing non-allelic interactions for height. Because of the presence of these interactions there is little point in estimating the additive and dominance components of variation either by the new analysis or by any other analysis. There is, however, one direction in which the analyses can be pursued further, namely, in locating the particular pairwise crosses between the inbred lines which are responsible for the non-allelic interactions. These were previously located both by the Wr/Vr analysis and by the scaling tests (Jinks, 1954, 1956). In the new analysis we can attempt to locate the inbred lines responsible for the interactions by examining the magnitude of $L_{1i} + L_{2i} - \bar{P}_i$ for each line. This is constant over all lines and equal to $(\mu + \sum_j h_j)$ only in the absence of inter-actions. Lines that interact with the testers L_1 and L_2 might be expected, therefore, to show the greatest deviations in the value of $L_{1i} + L_{2i} - \bar{P}_i$ from its mean value over all lines. Earlier analyses (Jinks, 1954, 1956) have shown that the crosses between the testers and lines 1, 3, 6 and 8 in season 2 and 1, 3, 5, 6 and 8 in season 3 exhibit epistasis. Examination of the eight values of $L_{1i} + L_{2i} - \bar{P}_i$ show that those for $i = 1$ and 3 in season 2 and $i = 1$, 6 and 8 in season 3 have the largest deviations from the mean value. Thus the new analysis allows us to locate the parental lines responsible for the epistasis in five out of nine of the known cases.

The various methods of detecting non-allelic interactions are not consistent in indicating the presence or absence of interaction for flowering time in any one season.

1. The Wr/Vr relationship did not indicate the presence of significant non-allelic interactions in any of the three seasons although the Wr/Vr relationship is consistent with the presence of duplicate gene interactions in all seasons (Jinks, 1954, 1956).

2. The scaling tests which could only be applied to the last two of the three seasons, indicated the presence of non-allelic interactions only in the third season (Jinks, 1956).

D 2

3. A model which assumed that there was no non-allelic interaction component of the variation within segregating generations was inadequate in the two seasons in which it could be tested, namely the second and third seasons.

Thus we have some evidence for the presence of non-allelic interactions in every season but consistent evidence over all tests for the third season only. This contrasts with the outcome of the new analysis which indicates the presence of non-allelic interactions in the first season only. This inconsistency which is not surprising in view of the presence of genotype-environmental interaction that produces a tenfold difference in total variation between seasons, suggests that epistasis is not a major source of variation for flowering time in most seasons. We can, therefore, usefully compare the estimates of the additive and dominance components of the variation in each of the three seasons for the new analysis and the diallel analysis obtained by assuming no epistasis (table 6). Both within seasons and between seasons the relative magnitudes of the estimates of D and H_1 are remarkably consistent between the new analysis and the diallel analysis. The F's differ as expected. The F from the new analysis is larger in magnitude and opposite in sign to the diallel F in every season. This suggests that on average v must be somewhat larger than u, i.e. there is an excess of decreasing genes which for flowering time, are the dominant alleles. Once again the new analysis based on 2×8 crosses has yielded comparable information to the more extensive analysis based on 8×8 sets of diallel crosses carried to the F_2 and first backcross generations.

5. DISCUSSION

In the last section we have shown that the same quantitative and quali- tative conclusions can often be drawn from the new analysis involving few crosses $(2 \times n)$ as from the complete diallel set (n^2), even when the latter is supplemented by information from the F_2 and first backcross generations. Thus the test for epistasis in the two analyses gave identical answers for height in both experiments. For flowering time in experiment 2, on the other hand, there were inconsistencies between the new analysis and the diallel analysis but there were equally marked discrepancies in the diallel analysis between seasons and between the alternative tests. Similarly, in experiment 1 there is a suggestion of duplicate gene interaction (P = 0·05) from the Wr/Vr test in the diallel analysis of flowering time which is not detected by the new analysis carried out on data extracted from this diallel. Unfortunately, there are no further tests that can be applied to this data to confirm either the presence or absence of interactions. There is, however, evidence of duplicate gene interaction in other generations derived from the same initial cross when grown in some seasons, but not in others (Opsahl, 1956; Mather and Vines, 1952). This situation is similar to that found for the same character in experiment 2 where there were indications of the presence of duplicate gene interactions in different tests in different seasons. There are, therefore, clear indications of genotype-environmental interactions over seasons for flowering time which affect the genic interactions as well as possible differences in the sensitivities and reliabilities of alternative methods of detecting their presence.

In all cases, where comparisons have been made, the estimates of the

additive and dominance components of variation obtained from the new analysis are remarkably similar to those obtained from the full diallel analysis. Furthermore, the new analysis has the advantage of providing tests of significance for the two components prior to their estimation and the estimates are uncorrelated.

The diallel analysis provides estimates of two parameters, F and H_2, which are not obtainable from the new analysis. These allow us to determine the relative frequencies of dominant to recessive genes and of increasing to decreasing genes in the parental populations. The new analysis, however, provides an estimate of a parameter similar to F, (and designated F) which allows us to distinguish between directional and ambidirectional dominance and to determine the direction of the dominance, *i.e.* positive or negative.

An important point which is illustrated by the analyses described in this paper is the danger of fitting an additive-dominance model before first determining whether or not such a model is adequate. The failure of the approaches typified by the three designs of Comstock and Robinson (1952) and the designs commonly used by animal breeders (see Falconer, 1964 for details) to provide tests of the adequacy of their models means that the estimates of additive and dominance components of variation they provide are frequently biassed to an unknown extent by epistatic variation. Equally the failure to provide tests for the presence of epistasis, let alone methods for estimating its magnitude and nature, has led to serious underestimation of the importance of the epistatic component of variation and a failure to recognise its contribution to heterosis. It is no longer possible to justify the use of a biometrical genetical analysis which does not have a built-in test for epistasis now that its widespread occurence, particularly for characters which are directly related to fitness, has been established.

A number of alternative methods for detecting epistasis are now available covering almost every kind of biological material and breeding system. These methods include scaling tests, (Mather, 1949; Cavalli, 1952; Jinks, 1956; Jinks and Jones, 1958), the Wr/Vr relationship (Jinks, 1954), the triple test-cross (Kearsey and Jinks, 1968) and the test described in the present paper.

The scaling tests in their original and variously modified forms provide unambiguous methods for detecting the presence of epistasis in the generations derivable from an initial cross between a pair of inbred lines. For these tests to be effective, the F_2 and first backcross generations are the minimum that must be raised. The Wr/Vr relationship, similarly, provides a test for epistasis in multiline crosses between inbred lines which is applicable as early as the F_1 generation. However, because this test depends on the fact that $Wr - Vr$ is constant over arrays in the absence of epistasis, it lacks the statistical precision of the scaling tests since there is no entirely satisfactory way of testing this expectation. In general a combination of a joint regression analysis of Wr on Vr and an analysis of variance of $(Wr - Vr)$, both of which use an empirical error derived from differences between replicate blocks, have been used to detect deviations from the expectation that $(Wr - Vr)$ is a constant. The Wr/Vr test is also ambiguous in that it can fail because of correlated gene distributions (Jinks, 1956), although no failure has been traced to this cause except where correlations were deliberately introduced (Hill, 1964).

The triple test-cross described by Kearsey and Jinks (1968) is applicable

to the first generation crosses between the members of any population and two inbred lines and their F_1. It, therefore, takes no more generations to carry out than the Wr/Vr test, but it has the advantage of being unambiguous, statistically reliable and universal in its applicability.

The new test described in the present paper retains many of these advantages but its use is confined to a population of inbred lines. It is however less efficient in that there are two sets of conditions in which it could give misleading results. Thus it may wrongly indicate the presence of epistasis if the tester lines do not differ at all the loci for which the lines under test differ and it fails to indicate its presence in the unlikely event that for all pairs of interacting genes $i_{ab} = l_{ab}$ and $j_{ab} = j_{ba}$. The latter type of failure, which results from the balancing out of the interactions, occurs for particular sets of relative values and signs of the interaction parameters in almost every test that has been devised. The other cause of failure is unique to this test. However, this cause of failure is also unlikely to be important except when the testers have been derived independently of one another and of the population being tested. And even here it is unlikely that the testers, chosen because their phenotypes are opposite extremes, would have a sufficiently high proportion of genes in common as to produce a misleading result.

One danger of the smaller $2 \times n$ crossing programme compared with the full diallel of n^2 crosses is that if the non-allelic interactions are confined to crosses between pairs of lines which do not include one or the other of the tester lines, L_1 and L_2, they will not be detected by the former design. This could well have occurred in the second experiment, the 8×8 diallel between inbred lines of diverse origins, where the non-allelic interactions are known to be confined to specific crosses. However, L_1 was a line which showed marked interaction with the other lines for final height. Hence, the presence of epistasis for final height was detected by the smaller design. Nevertheless, the possibility of not detecting an epistatic component of the variation although it is an important part of the genetical architecture, should be borne in mind whenever the tester lines are extreme phenotypes chosen from unrelated inbred lines.

6. Summary

1. A new analysis for detecting additive, dominance and epistatic variation applicable to populations of inbred lines is described.

2. The analysis is illustrated by data on final height and flowering time taken from a 20×20 and an 8×8 diallel of *Nicotiana rustica* lines.

3. The results are compared with those obtained from the original diallel analyses and scaling tests.

4. The conclusions drawn from the new analysis show remarkably good agreement with those from the more extensive diallel analyses.

5. Hence the new analysis which requires only $2 \times n$ crosses out of the diallel set of n^2 crosses yields the same information as the full diallel analysis

7. References

CAVALLI-SFORZA, L. I. 1952. An analysis of linkage in Quantitative Inheritance. " Quantitative Inheritance ", Ed. E. C. Reeve and C. H. Waddington. H.M.S.O., pp. 135-144. London.

COMSTOCK, R. E., AND ROBINSON, H. F. 1952. Estimation of average dominance of genes. *Heterosis*, Ch. 30. *Iowa State College Press*.

FALCONER, D. S. 1964. *Introduction to Quantitative Genetics*. Oliver and Boyd, Edinburgh and London.

HAYMAN, B. I. 1954. The analysis of variance of diallel tables. *Biometrics 10*, 235-244.

HAYMAN, B. I. 1957. Interaction, heterosis and diallel crosses. *Genetics, 42*, 336-355.

HILL, J. 1964. Effects of correlated gene distributions in the analysis of diallel crosses. *Heredity 19*, 27-46.

JINKS, J. L. 1954. The analysis of continuous variation in a diallel cross of *Nicotiana rustica* varieties. *Genetics, 39*, 767-788.

JINKS, J. L. 1956. The F_2 and backcross generations from a set of diallel crosses. *Heredity, 10*, 1-30.

JINKS, J. L., AND JONES, R. MORLEY. 1958. Estimation of the components of Heterosis. *Genetics, 43*, 223-234.

JINKS, J. L., AND STEVENS, J. M. 1959. The components of variation among family means in diallel crosses. *Genetics, 44*, 297-308.

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

MATHER, K. 1949. Biometrical Genetics. The Study of Continuous Variation. *Dover Publications, Inc.*

MATHER, K. AND VINES, A. 1952. The inheritance of height and flowering time in a cross of *Nicotiana rustica*. " Quantitative Inheritance," Ed. E. C. Reeve and C. H. Waddington. H.M.S.O., pp. 135-144. London.

OPSAHL, B. 1956. The discrimination of Interactions and linkage in continuous variation. *Biometrics, 10*, 415-432.