
DETECTION AND ESTIMATION OF GENOTYPE-ENVIRONMENTAL, LINKAGE AND EPISTATIC COMPONENTS OF VARIATION FOR A METRICAL TRAIT

JEAN M. PERKINS and J. L. JINKS

Department of Genetics, University of Birmingham

Received 5.iii.69

1. CURRENT PROCEDURES

WHEN, in an adequately designed experiment, a simple model which allows for additive and dominance genetic and additive environmental effects fails to account for the observed variation in the generations that can be derived from a cross between two inbred lines, there are only three possible causes, namely, the presence of genotype-environmental interactions, linkage or epistasis. To proceed beyond the initial stage of recognising that the simple model is inadequate, it is necessary to detect unambiguously the particular causes of failure so that where possible an extended model which recognises their presence can be fitted.

A test for the presence of a genotype-environmental interaction component of variance based on a comparison of the within family variance of inbred lines and their F_1 's has been available for some time (Mather, 1949). Similarly, an epistatic component of variance can be detected unambiguously by the failure of a model which allows for linkage between genes with additive and dominance effects in the absence of genotype-environmental interactions (Mather and Vines, 1952). In the absence of a significant epistatic component the same model fitting procedure provides an unambiguous test for linkage (Mather, 1949). However, in the presence of epistasis this approach provides an insensitive and often unreliable test for linkage (Mather and Vines, 1952; Jinks, 1956; Opsahl, 1956).

Unambiguous tests for linkage based on comparisons between the observed variances of pairs of specific types of families have been described by Van der Veen (1959) and equally unambiguous tests for an epistatic component of variation based on a triple test cross have been described by Kearsy and Jinks (1968). Neither of these tests has yet been used for the analysis of breeding data.

Among these methods of detecting causes of failure of the simple model we can recognise two basically different approaches. There is the approach exemplified by the methods of Mather (1949), Van der Veen (1959) and Kearsy and Jinks (1968) in which genotype-environmental interactions, linkage and epistasis may be detected and often classified by raising particular combinations of generation. The approach is essentially that of a "scaling test" providing statistically simple comparisons which give unambiguous qualitative answers.

The alternative approach initiated by Mather and Vines (1952) and elaborated by Jinks (1956), Opsahl (1956), Cooke and Mather (1962) and Hill (1966) consists of fitting models of increasing complexity, starting with the simplest model, until an adequate model is obtained. This approach has

the advantage in that it is prospectively capable of determining which sources of variation are present and at the same time of providing estimates of the parameters in the model. It is, however, a less sensitive means of deciding which sources of variation are present and often leads to ambiguous conclusions when many different sources are present simultaneously.

In the present paper both approaches will be illustrated and compared by the analysis of the variation in plant height and time of flowering for the generations which can be derived from an initial cross between inbred varieties 1 and 5 of *Nicotiana rustica*. It will be shown that application of the qualitative scaling tests followed by the estimation of the parameters of the appropriate model is the most efficient approach to the analysis of data in which genotype-environmental interactions, linkage and epistasis occur together.

2. EXPERIMENTAL DESIGN

Features of the design of the *N. rustica* experiment relevant to the analysis of their generation means have been described by Jinks and Perkins (1969), these are:

1. The maximum number of different generations that could be produced in one season were obtained by making all possible selfings, sib-matings and backcrosses among the six generations P_1 (= variety 5), P_2 (= variety 1), F_1 , F_2 , B_1 and B_2 .
2. The relative numbers of individuals that were grown from each of the 21 generations so produced were determined as the reciprocal of the within generation variances for final height and flowering time that had been observed in, or predicted from, earlier experiments with the same varieties.
3. The experiment was grown in two replicate blocks, the plants in each block being individually randomised.

The aim of this design was to obtain the same amount of information about the means of as many generations as possible within the practical limitations of space and time and to obtain reliable estimates of the amount of information for each generation. Further features of the experimental design which have not previously been described relate specifically to the analysis of the variances within the families and generations. Of the 21 generations, 15 consist of a number of families, each family being the progeny of a single mating involving one or a pair of individuals chosen at random from among the F_2 , B_1 and B_2 parental generations. The procedure used for determining the relative numbers of individuals in each generation was used in these 15 generations to determine the relative numbers of families and the number of individuals per family. The relative numbers were converted into absolute numbers of plants and rounded off to the nearest five plants per family as described earlier (Jinks and Perkins, 1969). The final design arrived at in this way is summarised in table 1.

In so far as the design succeeds in its objective, least squares procedures should lead to maximum likelihood estimates of the components of the generation means. It also goes a long way towards equalising the variances of the within and between family variances although, of course, a design which equalises the variances of the generation means will not also equalise the variances of the variances.

TABLE 1
The design of the *N. rustica* experiment

	Generation	No. of families	No. of individuals	
			Per family per block	Total
1.	P ₁	1	80	160
2.	P ₂	1	50	100
3.	F ₁	1	70	140
4.	F ₂	1	150	300
5.	B ₁	1	80	160
6.	B ₂	1	80	160
7.	F ₃	40	5	400
8.	F ₂ × P ₁	40	5	400
9.	F ₂ × P ₂	40	5	400
10.	F ₂ × F ₁	40	5	400
11.	F ₂ bip*	20	10	400
12.	B ₁₁	14	5	140
13.	B ₁₂	14	5	140
14.	B ₂₂	14	5	140
15.	B ₂₁	14	5	140
16.	B ₁ × F ₁	14	10	280
17.	B ₂ × F ₁	14	10	280
18.	B ₁ bip	7	15	210
19.	B ₂ bip	7	15	210
20.	B ₁ S	14	10	280
21.	B ₂ S	14	10	280

Total = 5120

* Usually referred to as F₃bip but this terminology cannot be extended to the backcross series. Hence for the bip and S series we have specified the parental generation.

3. DERIVATION OF EXPECTED VARIANCES

The principles underlying the derivation of the expected variances within the generations that can be obtained from an initial cross between two inbred lines when the individual plants are randomised over the whole environment and when there are interactions between the genotypes and the environment have been described by Mather and Jones (1958) and Jones and Mather (1958). These have been used not only to derive some of the expectations they discussed in a form more appropriate for the present analyses, but also to derive expectations for further generations.

In the simple model only four parameters are required to specify the additive-dominance genetic and the additive environmental contributions to the 36 variances which can be derived from the 21 generations of the *N. rustica* experiment, these are, the three genetic components, *D*, *H* and *F* and the environmental component, *E*₁ (Mather, 1949; Opsahl, 1956; Hill, 1966). To allow for the contributions of genotype-environmental interactions to these variances in the absence of linkage, a further seven parameters are required. These are:

$G_D = \sum g_{atj}^2$, which is the contribution of the interaction between the additive genetic effect at the *i*th locus and the *j*th environment summed over all loci and environments.

$G_H = \sum g_{htj}^2$, which is the corresponding component for the dominance genetic effect.

$Wg_{aik}g_{akj}$,	which is the product of the interactions of the additive genetic effects at the i th and k th loci and the j th environment summed over all pairs of loci and environments.
$Wg_{nij}g_{nkj}$,	which is the corresponding component for the dominance genetic effect.
$Wg_{aij}g_{nij}$,	which is the corresponding component for the interaction between the j th environment and the additive and dominance effects at the i th locus summed over all loci and environments.
We_jg_{aij} ,	which is the product of the additive environmental effect of the j th environment and the interaction of the latter with the additive genetic effect at the i th locus summed over all loci and environments.

And

We_jg_{nij} ,	which is the corresponding component for the dominance genetic effect.
-----------------	--

Similar parameters have previously been described by Bucio Alanis, Perkins and Jinks (1969) for the analysis of genotype-environmental interactions between the means of generations and macroenvironmental effects such as seasons and locations.

Examination of the contributions which these parameters make to the variances which can be obtained from the *N. rustica* experiment, shows that five of these parameters always appear associated in the following combinations:

$$\begin{aligned} E_1 + G_D + 2We_jg_{aij} \\ E_1 + G_D - 2We_jg_{aij} \\ E_1 + G_H + 2We_jg_{nij} \end{aligned}$$

which we will designate GE_1 , GE_2 and GE_3 , respectively. We have, therefore, effectively only three equations with which to solve for these five parameters. Hence, we can only solve for three combinations of these parameters. One possible set of three is GE_1 , GE_2 and GE_3 , the other is $(E_1 + G_D)$, We_jg_{aij} and GE_3 . Since the former set are, in fact, the major components of the variances of P_1 , P_2 and F_1 families, respectively, this set is the obvious choice. While, therefore, we required 11 parameters to specify the expectations, we can estimate only nine. The expectations of the 36 variances of the *N. rustica* experiment are given in terms of these nine parameters in table 2. The expectation on the simple model (D , H , F and E_1) can be derived from these by putting $G_D = G_H = g_a = g_n = 0$, in which case $GE_1 = GE_2 = GE_3 = E_1$ so that E_1 appears with a coefficient of 1 in the σ_w^2 's of all statistics and all the W terms = 0.

If we now relax the assumption of no linkage, 15 parameters are required to specify the expected variances but 11 of these are compound like GE_1 , etc., containing parameters which can be independently specified but which cannot be independently estimated. This is a considerable reduction compared with the model described by Van der Veen (1959) in which 15 parameters are required to describe the contributions of the genetic components alone in the presence of linkage compared with the eight required on the present system. This must mean that there are hidden correlations among the parameters of Van der Veen's model. Some of the parameters of the present model, which is given in table 3, are directly comparable

TABLE 2

Model specifying genetic, environmental and genotype-environmental interaction contributions to the within family ($V_{FAM} = \sigma_w^2$) and between family ($V_{FAM} = 1/nc_w + \sigma_b^2$) variances. In the absence of genotype-environmental interactions $GE_1 + GE_2 + GE_3 = E_1$ and all $W_{i,j} = 0$

N	V_{FAM}	n^*	D	H	F	GE_1	GE_2	GE_3	$W_{adj}^{G_A B_1}$	$W_{adj}^{G_A B_2}$	$W_{adj}^{G_A B_3}$
1.	$V_{FAM} P_1$	—	—	—	—	1	—	—	2	—	—
2.	F_2	—	—	—	—	—	1	—	2	—	—
3.	F_1	—	—	—	—	—	—	1	—	—	—
4.	B_1	—	1/2	1/4	—	1/4	1/4	1/2	1/2	1/2	—
5.	B_2	—	1/4	1/4	-1/2	1/2	1/2	1/2	1/2	1/2	1/2
6.	B_3	—	1/4	1/4	1/2	—	—	—	1/2	—	-1/2
7.	$F_2 \times P_1$	—	1/4	1/8	—	3/8	3/8	1/4	1/8	1/8	—
8.	$F_2 \times P_2$	—	1/8	1/8	-1/4	1/2	1/2	1/2	1/2	1/2	1/2
9.	$F_2 \times F_1$	—	3/8	1/4	1/4	1/4	1/4	1/2	1/2	1/2	—
10.	$F_2 \text{ bip}$	—	1/4	3/16	—	1/4	1/4	1/2	—	—	—
11.	B_{21}	—	1/8	1/8	—	3/4	—	1/2	—	—	—
12.	B_{12}	—	1/8	1/8	-1/4	—	—	1/4	9/8	1/8	3/8
13.	B_{22}	—	1/8	1/8	1/4	—	—	3/4	1/8	1/8	-3/8
14.	B_{23}	—	1/8	1/8	1/4	—	—	1/4	9/8	1/8	-3/8
15.	$B_{21} \times F_1$	—	1/8	1/8	-1/4	1/4	1/8	1/2	1/8	1/8	1/4
16.	$B_2 \times F_1$	—	3/8	1/4	—	1/4	1/8	1/2	1/2	1/2	—
17.	$B_2 \times F_2$	—	1/4	3/16	1/4	3/8	3/8	1/2	1/2	1/2	—
18.	$B_1 \text{ bip}$	—	1/4	3/16	—	1/16	1/16	3/8	9/32	9/32	3/8
19.	$B_2 \text{ bip}$	—	1/4	3/16	1/4	1/16	9/16	3/8	1/2	1/2	—
20.	$B_1 S$	—	1/4	1/8	—	5/8	1/8	1/4	1/2	1/2	1/4
21.	$B_2 S$	—	1/4	1/8	—	1/8	5/8	1/4	1/2	1/2	-1/4
22.	$V_{FAM} F_2$	5	1/4n+1/2	1/8n+1/16	—	3/8n	3/8n	1/4n	1/2n	1/2n	—
23.	$F_2 \times P_1$	5	1/8n+1/8	1/8n+1/8	-1/4n-1/4	1/2n	1/2n	1/4n	1/2n	1/2n	1/2n
24.	$F_2 \times P_2$	5	1/8n+1/8	1/8n+1/8	1/4n+1/4	—	—	1/2n	1/2n	1/2n	-1/2n
25.	$F_2 \times F_1$	5	3/8n+1/8	1/4n	—	1/4n	1/4n	1/2n	1/2n	1/2n	—
26.	$F_2 \text{ bip}$	10	1/4n+1/4	3/16n+1/16	—	1/4n	1/4n	1/4n	—	—	—
27.	B_{11}	5	1/8n+1/16	1/8n+1/16	-1/4n-1/8	—	—	1/4n	9/8n	1/8n	3/8n
28.	B_{12}	5	1/8n+1/16	1/8n+1/16	1/4n+1/8	—	—	3/4n	1/8n	1/8n	-3/8n
29.	B_{22}	5	1/8n+1/16	1/8n+1/16	1/4n+1/8	—	—	1/4n	9/8n	1/8n	3/8n
30.	B_{23}	5	1/8n+1/16	1/8n+1/16	-1/4n-1/8	—	—	3/4n	1/8n	1/8n	-3/8n
31.	$B_1 \times F_1$	10	3/8n+1/16	1/4n	—	1/4n	—	1/2n	1/2n	1/2n	—
32.	$B_2 \times F_1$	10	3/8n+1/16	1/4n	-1/4n	1/8n	1/8n	1/2n	1/2n	1/2n	1/4n
33.	$B_1 \text{ bip}$	15	1/4n+1/8	3/16n+3/64	-1/4n-1/8	3/8n	3/8n	1/2n	1/2n	1/2n	-1/4n
34.	$B_2 \text{ bip}$	15	1/4n+1/8	3/16n+3/64	1/4n+1/8	9/16n	9/16n	3/8n	9/32n	9/32n	3/8n
35.	$B_1 S$	10	1/4n+1/4	1/8n+1/16	-1/4	1/16n	9/16n	3/8n	1/2n	1/2n	-3/8n
36.	$B_2 S$	10	1/4n+1/4	1/8n+1/16	1/4	5/8n	5/8n	1/4n	1/2n	1/2n	1/4n

* n = Number of individuals per family.

with those given by Jones and Mather (1958) for linked genes which interact with the environment.

In the parameters of the present model p is the recombination frequency between the pairs of genes specified and the linkage parameters have been defined as ascending powers of p (Jinks and Perkins, 1969). For example, the parameter $Wg_{aijg_{akj}}$ of the no linkage model now appears in the additional forms $Wp_{ik}g_{aijg_{akj}}$ and $Wp_{ik}^2g_{aijg_{akj}}$ and similarly for the $Wg_{aijg_{hij}}$ and $Wg_{hijg_{hkj}}$ terms. Three further groups of parameters occur in the presence of linkage, namely, the products of the additive and dominance effects of the linked pairs of genes, that is Wd_id_k , Wd_ih_k and Wh_ih_k and the products of these with ascending powers of p . Appropriate combinations of the latter parameters with D , H , and F give the various modified forms of D , H and F

TABLE 3

The definitions of the genetic and genotype-environmental interaction parameters in the presence of linkage. For simplicity p_{ik} has been written as p throughout

Parameter	Definition
1. D_1	$D + 2Wd_id_k - 4Wpd_id_k$
2. D_2	$D + 2Wd_id_k - 8Wp^2d_id_k - 8Wp^2g_{aijg_{akj}}$
3. H_1	$H + 2Wh_ih_k - 4Wph_ih_k$
4. H_2	$H + 2Wh_ih_k - 8Wp^2h_ih_k$
5. H_3	$H + 2Wh_ih_k - 16Wp^3h_ih_k - 16Wp^3g_{hijg_{hkj}}$
6. H_4	$H + 2Wh_ih_k - 32Wp^4h_ih_k - 32Wp^4g_{hijg_{hkj}}$
7. F_1	$F + Wd_ih_k - 2Wpd_ih_k$
8. F_2	$F + Wd_ih_k - 4Wp^2d_ih_k - 4Wp^2g_{aijg_{hkj}}$
9. GE_{L1}	$E_1 + C_D - 2We_jg_{aij} + 2Wg_{aijg_{akj}}$
10. GE_{L2}	$E_1 + C_D - 2We_jg_{aij} + 2Wg_{aijg_{akj}}$
11. GE_{L3}	$E_1 + C_H + 2We_jg_{hij} + 2Wg_{hijg_{hkj}}$
12. $Wp^0g_{aijg_{akj}}$	$Wp^0g_{aijg_{akj}}$
13. $Wp^0g_{hijg_{hkj}}$	$Wp^0g_{hijg_{hkj}}$
14. $Wp^2g_{hijg_{hkj}}$	$Wp^2g_{hijg_{hkj}}$
15. $Wp^2g_{aijg_{hkj}}$	$Wp^2g_{aijg_{hkj}}$

used by Mather (1949) and others to specify the contributions of linked genes to the variances of different rank. However, by breaking down these modified forms into the original D , H and F and the cross product terms (W 's) fewer parameters are required to completely specify the contributions of linkage to the 36 variances of the present study.

The cross product parameters based on d_id_k will be positive if the breakable linkages are predominantly in the repulsion phase. Similarly, the parameters based on h_ih_k will be positive if the breakable linkages are predominantly between genes with reinforcing dominance, *i.e.* with h 's of same sign, and negative if they are predominantly between genes with opposing dominance, *i.e.* with h 's of opposite signs.

The expectations of the 36 variances in terms of the 15 parameters are given in tables 4 and 5. By summing over the coefficients of D_1 and D_2 in any expectation the coefficient of D for the corresponding expectation in the absence of linkage is obtained. This relationship also holds for the sums of the coefficients of H_1 , H_2 , H_3 and H_4 and F_1 and F_2 and the corresponding coefficients of H and F , respectively. Hence the expectations in the absence of linkage (table 2) may be readily derived from those in its presence (tables 4 and 5). Similarly, the expectations for the presence of linkage in the absence of genotype-environmental interactions can be derived by putting

TABLE 4

Model specifying the genetic contributions to the within family ($V_{FAM} = c_w^2$) and between family ($V_{FAM} = 1/n\sigma_w^2 + \sigma_b^2$) variances in the presence of linkage. For definitions of the parameters, see table 3

Statistic	n*	Parameters									
		D ₁	D ₂	H ₁	H ₂	H ₃	H ₄	F ₁	F ₂		
1. $V_{FAM} P_1$	—	—	—	—	—	—	—	—	—	—	—
2. F ₁	—	—	—	—	—	—	—	—	—	—	—
3. F ₂	—	—	—	—	—	—	—	—	—	—	—
4. F ₃	—	—	—	—	—	—	—	—	—	—	—
5. B ₁	—	—	—	—	—	—	—	—	—	—	—
6. B ₂	—	—	—	—	—	—	—	—	—	—	—
7. F ₃	—	—	—	—	—	—	—	—	—	—	—
8. F ₂ × P ₁	—	—	—	—	—	—	—	—	—	—	—
9. F ₂ × P ₂	—	—	—	—	—	—	—	—	—	—	—
10. F ₂ × F ₁	—	—	—	—	—	—	—	—	—	—	—
11. F ₂ bip	—	—	—	—	—	—	—	—	—	—	—
12. B ₁₁	—	—	—	—	—	—	—	—	—	—	—
13. B ₁₂	—	—	—	—	—	—	—	—	—	—	—
14. B ₂₁	—	—	—	—	—	—	—	—	—	—	—
15. B ₂₂	—	—	—	—	—	—	—	—	—	—	—
16. B ₂ × F ₁	—	—	—	—	—	—	—	—	—	—	—
17. B ₂ × F ₂	—	—	—	—	—	—	—	—	—	—	—
18. B ₂ bip	—	—	—	—	—	—	—	—	—	—	—
19. B ₁ S	—	—	—	—	—	—	—	—	—	—	—
20. B ₂ S	—	—	—	—	—	—	—	—	—	—	—
21. $V_{FAM} F_3$	—	—	—	—	—	—	—	—	—	—	—
22. F ₂ × P ₁	5	1/2n + 1/2	—	3/8n + 1/8	—	1/4n	—	—	—	—	—
23. F ₂ × P ₂	5	1/4n + 1/8	—	1/4n + 1/8	—	1/4n + 1/8	—	—	—	—	—
24. F ₂ × F ₁	5	1/2n + 1/4	—	5/8n	—	1/2n	—	—	—	—	—
25. F ₂ bip	10	3/16n + 1/16	—	3/16n + 1/16	—	3/16n + 1/16	—	—	—	—	—
26. B ₁₁	5	3/16n + 1/16	—	3/16n + 1/16	—	3/16n + 1/16	—	—	—	—	—
27. B ₁₂	5	3/16n + 1/16	—	3/16n + 1/16	—	3/16n + 1/16	—	—	—	—	—
28. B ₂₁	5	3/16n + 1/16	—	3/16n + 1/16	—	3/16n + 1/16	—	—	—	—	—
29. B ₂₂	5	3/16n + 1/16	—	3/16n + 1/16	—	3/16n + 1/16	—	—	—	—	—
30. B ₂ × F ₁	10	7/16n + 1/16	—	1/2n	—	1/2n	—	—	—	—	—
31. B ₂ × F ₂	10	7/16n + 1/16	—	1/2n	—	1/2n	—	—	—	—	—
32. B ₁ bip	15	3/8n + 1/8	—	7/16n + 1/16	—	7/16n + 1/16	—	—	—	—	—
33. B ₂ bip	15	3/8n + 1/8	—	7/16n + 1/16	—	7/16n + 1/16	—	—	—	—	—
34. B ₁ S	10	3/8n + 1/4	—	5/16n + 1/16	—	5/16n + 1/16	—	—	—	—	—
35. B ₂ S	10	3/8n + 1/4	—	5/16n + 1/16	—	5/16n + 1/16	—	—	—	—	—

* n = Number of individuals per family.

$G_D = G_H = g_a = g_b = 0$ in the expectations in tables 4 and 5, which leaves only one parameter for environmental effects E_1 which has a coefficient of 1 in all the σ_w^2 's of all statistics.

TABLE 5

Model specifying the environmental and genotype-environmental interaction contributions to the within ($V_{FAM} = \sigma_w^2$) and between ($V_{FAM} = 1/n\sigma_w^2 + \sigma_b^2$) family variances in the presence of linkage. For definitions of the parameters see table 3. In the absence of genotype-environmental interactions $GE_{L_1} + GE_{L_2} + GE_{L_3} = E_1$ and all W 's = 0.

Statistic	n*	Parameters						
		GE_{L_1}	GE_{L_2}	GE_{L_3}	$Wpg_{a1j}g_{akj}$	$Wpg_{b1j}g_{bkj}$	$Wp^2g_{a1j}g_{bkj}$	$Wpg_{a1j}g_{bkj}$
1. $\sqrt{V_{FAM}}$ P ₁	—	1	—	—	—	—	—	—
2. P ₂	—	—	1	—	—	—	—	—
3. F ₁	—	—	—	1	—	—	—	—
4. F ₂	—	1/4	1/4	1/2	-2	-2	2	—
5. B ₁	—	1/2	—	1/2	-1	-1	—	1
6. B ₂	—	—	1/2	1/2	-1	-1	—	-1
7. F ₃	—	3/8	3/8	1/4	-4	-2	4	—
8. F ₂ × P ₁	—	1/2	—	1/2	-3/2	-3/2	1	3/2
9. F ₂ × P ₂	—	—	1/2	1/2	-3/2	-3/2	1	-3/2
10. F ₂ × F ₁	—	1/4	1/4	1/2	-5/2	-5/2	4	—
11. F ₂ bip	—	1/4	1/4	1/2	-3	-3	13/2	—
12. B ₁₁	—	3/4	—	1/4	-1	-1	1/2	1
13. B ₁₂	—	—	1/4	3/4	-1	-1	1/2	-1
14. B ₂₂	—	—	3/4	1/4	-1	-1	1/2	-1
15. B ₂₁	—	1/4	—	3/4	-1	-1	1/2	1
16. B ₁ × F ₁	—	3/8	1/8	1/2	-2	-2	5/2	1/2
17. B ₂ × F ₁	—	1/8	3/8	1/2	-2	-2	5/2	-1/2
18. B ₁ bip	—	9/16	1/16	3/8	-2	-2	3	1
19. B ₂ bip	—	1/16	9/16	3/8	-2	-2	3	-1
20. B ₁ S	—	5/8	1/8	1/4	-5/2	-3/2	2	1/2
21. B ₂ S	—	1/8	5/8	1/4	-5/2	-3/2	2	-1/2
22. $\sqrt{V_{FAM}}$ F ₃	5	3/8n	3/8n	1/4n	-4/n	-2/n	4/n	—
23. F ₂ × P ₁	5	1/2n	—	1/2n	-3/2n	-3/2n	1/n	3/2n
24. F ₂ × P ₂	5	—	1/2n	1/2n	-3/2n	-3/2n	1/n	-3/2n
25. F ₂ × F ₁	5	1/4n	1/4n	1/2n	-5/2n	-5/2n	4/n	—
26. F ₂ bip	10	1/4n	1/4n	1/2n	-3/n	-3/n	13/2n	—
27. B ₁₁	5	3/4n	—	1/4n	-1/n	-1/n	1/2n	1/n
28. B ₁₂	5	—	1/4n	3/4n	-1/n	-1/n	1/2n	-1/n
29. B ₂₂	5	—	3/4n	1/4n	-1/n	-1/n	1/2n	-1/n
30. B ₂₁	5	1/4n	—	3/4n	-1/n	-1/n	1/2n	1/n
31. B ₁ × F ₁	10	3/8n	1/8n	1/2n	-2/n	-2/n	5/2n	1/2n
32. B ₂ × F ₁	10	1/8n	3/8n	1/2n	-2/n	-2/n	5/2n	-1/2n
33. B ₁ bip	15	9/16n	1/16n	3/8n	-2/n	-2/n	3/n	1/n
34. B ₂ bip	15	1/16n	9/16n	3/8n	-2/n	-2/n	3/n	-1/n
35. B ₁ S	10	5/8n	1/8n	1/4n	-5/2n	-3/2n	2/n	1/2n
36. B ₂ S	10	1/8n	5/8n	1/4n	-5/2n	-3/2n	2/n	-1/2n

* n = number of individuals per family.

4. SCALING TESTS FOR VARIANCE COMPONENTS

Examination of the expectations of the 36 variances in tables 2 and 5 shows that the within family variances of the P₁, P₂ and F₁ families are unique in differing only in the presence of genotype-environmental interactions. This is true irrespective of the presence of linkage and as the following consideration will show, irrespective of the presence of epistasis. Thus since the epistatic components of inbred lines and their F₁'s are completely

confounded with the additive and dominance components, respectively, their interactions with the environment are similarly confounded. Hence, what is true for an additive-dominance model is still true when epistasis is present.

The genotype-environmental interaction components of the variances of the parental and F_1 families differ more than those of the variances of any other families in the experiment. The parental and F_1 families, therefore, provide the most sensitive comparisons for the detection of genotype-environmental interactions.

Three of the generations included in the experiment have identical expectations for the total σ^2 in the absence of linkage irrespective of the presence of genotype-environmental interactions. These are the F_2 , $F_2 \times F_1$ and F_2 bip generations for all of which

$$\sigma_w^2 + \sigma_b^2 = \frac{1}{2}D + \frac{1}{4}H + \frac{1}{4}GE_1 + \frac{1}{4}GE_2 + \frac{1}{2}GE_3 + \frac{1}{2}Wg_{hij}g_{hkf} \quad (\text{see table 2}).$$

Furthermore, since in the absence of linkage genotype frequencies in respect of all gene differences will be identical in the three types of families, this equality of the total σ^2 's will be independent of the presence of epistasis. Hence, inequality of the total σ^2 's is diagnostic of the presence of linkage.

Reference to table 5 shows that in the presence of linkage there is a greater difference between the expectations of the F_2 and F_2 bip families due to linkage than between any other pair of families. Furthermore, the expectation of the $F_2 \times F_1$ families is intermediate. Hence, the total σ^2 's of the F_2 and F_2 bip families provide the most sensitive test for linkage available in the *N. rustica* experiment.

If the breakable linkages are predominantly in the coupling phase and the dominance is reinforcing the total σ^2 for the F_2 will be the largest, if they are predominantly repulsion and opposing, respectively, the F_2 value will be the smallest.

Similar considerations show that inequality of the total σ^2 's of the $F_2 \times P_1$ and B_1 families and of the $F_2 \times P_2$ and B_2 families are also diagnostic of linkage. Again if the B_1 and B_2 values are the larger the breakable linkages are in the coupling and reinforcing phases, etc. (Van der Veen, 1959).

Kearsey and Jinks (1968) have shown that an epistatic component of variation within any population can be unambiguously detected by an appropriate analysis of variance of the families produced by crossing individuals in the populations to each of two inbred lines and to the F_1 produced by crossing these lines. The most efficient test is provided where the population is an F_2 or a combination of the two first backcrosses ($B_1 + B_2$) crossed to the inbred parents (P_1 and P_2) and the F_1 from which the F_2 and backcrosses were derived, *i.e.* the $F_2 \times P_1$, $F_2 \times P_2$ and $F_2 \times F_1$ and $B_1 \times P_1$, $B_2 \times P_1$, $B_1 \times P_2$, $B_2 \times P_2$, $B_1 \times F_1$, and $B_2 \times F_1$ generations. An extension of this analysis, described in section 6 (i) allows the epistatic component to be partitioned and classified.

Thus by including families of the P_1 , P_2 , F_1 , F_2 , F_2 bip, $F_2 \times P_1$, $F_2 \times P_2$ and $F_2 \times F_1$ generations in our experiment we are able to carry out the most sensitive and unambiguous tests for genotype-environmental interactions, linkage and epistatic components of variation that are readily available.

5. MODEL FITTING PROCEDURES

The aim of model fitting in biometrical genetics is to determine the simplest model of gene and environmental action and interaction that will account for the observed variances. We begin, therefore, by fitting the simplest model, namely, one in which there are additive and dominance genetic effects (D , H and F) and additive environmental effects (E_1) and proceed progressively to models which allow for genotype-environmental interactions, linkage and epistasis only where the simpler models fail.

If the simplest model fails all three possible causes may be responsible. Unfortunately we cannot fit a model which allows for epistasis because more parameters are required than we have statistics for their solution. The next step, therefore, is to fit the models which allow for genotype-environmental interactions only and linkage only (tables 2, 4 and 5). If both of these models prove to be inadequate we must then fit the model which allows for the presence of both (tables 4 and 5). The latter model can be inadequate for one reason only, namely, the presence of epistasis. Hence the cause of failure of this model is the only one which is unambiguous. The problem of determining whether or not genotype-environmental interactions and linkage are present when the model which includes both fails, *i.e.* epistasis is present, will be discussed in relation to the analysis of the *N. rustica* data.

6. RESULTS

Estimates of the 21 within family variances (V_{FAM}) and the 15 between family variances (V_{FAM}) of the *N. rustica* experiment are given in table 6 for plant height and time of flowering.

(i) *Scaling tests*

Genotype-environmental interactions. The variances of the P_1 , P_2 and F_1 generations along with their generation means are given in table 7. Bartlett's test for both characters shows that the three variances are heterogeneous ($P < 0.001$) hence genotype-environmental interactions are present. Where there is a simple relationship between family means and variances such interactions may be removed by rescaling the data (Mather, 1949). Reference to table 7 reveals no simple relationship between the means and variances for either character, because the largest variances are associated with means that are intermediate in value. For final height the F_1 is taller than the taller parent (P_1) but it has a smaller variance than the more stable parent (P_2). For flowering time the F_1 flowers earlier than the earlier flowering parent (P_2) but it has a variance which equals that of the more stable parent (P_1). Thus the mean expression of the characters and their response to environmental variation appear to be independently determined; a conclusion which is in agreement with more extensive investigations on varieties of *N. rustica* (Jinks and Mather, 1955; Perkins and Jinks, 1968).

Linkage. Three comparisons can be made between the total σ^2 's of generations which specifically test for linkage. These are:

1. Total $\sigma^2(F_2 \times P_1) = \text{Total } \sigma^2(B_1)$
2. Total $\sigma^2(F_2 \times P_2) = \text{Total } \sigma^2(B_2)$
3. Total $\sigma^2(F_2 \text{ bip}) = \text{Total } \sigma^2(F_2)$

In addition we have the expectation that in the presence of linkage the total σ^2 of $F_2 \times F_1$ will be intermediate between those of the F_2 and F_2 bip families. The estimates of the total σ^2 's of these generations obtained from the values of \hat{V}_{FAM} and $V_{\overline{FAM}}$ in table 6 are given in table 8. The expected equalities have been tested as variance ratios using the smaller estimate as

TABLE 6

Estimates of the within family, \hat{V}_{FAM} , and between family, $V_{\overline{FAM}}$ variances for the two characters final height and flowering time in each of two replicate blocks

		Final height			Flowering time		
		Block I	Block II	I + II	Block I	Block II	I + II
1.	$\hat{V}_{FAM} P_1$	21.1074	35.9702	57.0776	10.1209	7.6157	17.7366
2.	P_2	20.1216	12.4490	32.5706	22.3376	13.1547	35.4923
3.	F_1	18.0433	13.0435	31.0868	9.8418	7.4586	17.3004
4.	F_2	36.8519	31.5832	68.4351	27.5929	19.7587	47.3516
5.	B_1	32.9816	23.4034	56.3850	13.6202	11.3999	25.0201
6.	B_2	24.8036	23.0404	47.8440	21.4025	9.6479	31.0504
7.	F_3	38.0000	26.6850	64.6850	17.7642	13.8458	31.6100
8.	$F_2 \times P_1$	23.1454	27.6933	50.8387	15.7412	9.8675	25.6087
9.	$F_2 \times P_2$	22.6688	25.2950	47.9638	13.3179	10.9821	24.3000
10.	$F_2 \times F_1$	35.5225	30.2848	65.8073	20.6600	14.0388	34.6988
11.	F_2 bip	33.2522	29.8344	63.0866	20.3453	14.7913	35.1366
12.	B_{11}	34.1643	24.0071	58.1714	12.1429	8.4369	20.5798
13.	B_{12}	23.1000	38.6643	61.7643	15.4143	10.6429	26.0572
14.	B_{22}	22.6429	17.9905	40.6334	15.5893	7.9726	23.5619
15.	B_{21}	33.1786	22.9071	56.0857	12.3286	6.8500	19.1786
16.	$B_1 \times F_1$	35.1587	33.1175	68.2762	21.3936	16.6869	38.0805
17.	$B_2 \times F_1$	42.8956	38.8984	81.7940	22.2259	14.9056	37.1315
18.	B_1 bip	28.5388	29.4912	58.0300	21.9510	13.8139	35.7649
19.	B_2 bip	37.0067	32.6599	69.6666	17.1090	14.3633	31.4723
20.	B_1 S	26.9603	26.2222	53.1825	16.4365	10.1667	26.6032
21.	B_2 S	28.4889	24.8155	53.3044	15.6222	11.1710	26.7932
22.	$V_{\overline{FAM}} F_2$	30.5013	26.0798	56.5811	12.3196	10.8596	23.1792
23.	$F_2 \times P_1$	9.1674	8.4810	17.6484	9.1118	5.6524	14.7642
24.	$F_2 \times P_2$	7.7707	12.7600	20.5307	5.9959	7.8590	13.8549
25.	$F_2 \times F_1$	12.7717	9.7670	22.5387	11.1137	5.1999	16.3136
26.	F_2 bip	16.4148	18.8099	35.2247	9.7756	6.1869	15.9625
27.	B_{11}	4.9442	5.7273	10.6715	3.8596	4.3240	8.1836
28.	B_{12}	7.9950	6.2031	14.1981	10.0565	4.5629	14.6194
29.	B_{22}	10.0668	8.9594	19.0262	4.6748	2.2271	6.9019
30.	B_{21}	13.8075	7.2380	21.0455	7.8182	4.1213	11.9395
31.	$B_1 \times F_1$	9.2596	6.2782	15.5378	4.1791	3.5259	7.7050
32.	$B_2 \times F_1$	8.7757	6.5488	15.3245	5.4979	1.9823	7.4802
33.	B_1 bip	1.5843	1.5442	3.1285	2.2333	0.9727	3.2060
34.	B_2 bip	12.5830	9.6798	22.2628	15.8527	4.1828	20.0355
35.	B_1 S	11.5046	12.1348	23.6394	17.0811	11.4547	28.5358
36.	B_2 S	22.0044	26.3108	48.3152	9.1598	9.0884	18.2482

the denominator and doubling the probability. This is an approximate test since we are comparing total σ^2 's and not mean squares, the two being identical only for families of size one. However, in view of the large numbers of degrees of freedom and the low probabilities of the significant comparisons the results are probably reliable.

For final height all three tests in table 8 are consistent with repulsion linkage although the first test is not significant. Furthermore the value of the $F_2 \times F_1$ statistic lies between the F_2 bip and F_2 statistics as expected (see

section 4). For flowering time all three tests in table 8 are also consistent with repulsion linkage but for this character only the first test is significant but its significance is high. There is, therefore, evidence of breakable linkages for the genes controlling height and flowering time.

TABLE 7

Means and variances of the P₁, P₂ and F₁ generations of varieties 1 and 5 of N. rustica for final height and flowering time (P₁ = variety 5) averaged over blocks

Generation	Final height		Flowering time	
	Mean	Variance	Mean	Variance
P ₁	60.46	28.5388	17.80	8.8683
P ₂	49.92	16.2853	15.62	17.7462
F ₁	61.30	15.5434	13.52	8.6502

Epistasis. The test for an epistatic component of variations is a modification of that proposed by Kearsey and Jinks (1968). If for the *i*th individual of the F₂ or combined B₁ and B₂ families we write

L_{1t} for the family produced by crossing the *i*th individual to P₁,
 L_{2t} for the family produced by crossing the *i*th individual to P₂, and
 L_{3t} for the family produced by crossing the *i*th individual to the
 $F_1(P_1 \times P_2)$,

TABLE 8

Specific tests for linkage based on comparisons of the total σ^2 's between the generations indicated

Test	Generation	Total σ^2	d.f.	P
Height				
1	F ₂ × P ₁	29.16	398	N.S.
	B ₁	28.19	158	
2	F ₂ × P ₂	29.46	398	0.02-0.002
	B ₂	23.92	158	
3	F ₂ bip	46.00	398	< 0.002
	F ₂	34.22	298	
	F ₂ × F ₁	37.59	398	
Flowering time				
1	F ₂ × P ₁	17.63	398	< 0.002
	B ₁	12.51	158	
2	F ₂ × P ₂	16.65	398	N.S.
	B ₂	15.53	158	
3	F ₂ bip	23.80	398	N.S.
	F ₂	23.68	298	
	F ₂ × F ₁	22.04	398	

then $L_{1t} + L_{2t} - 2L_{3t} = 0$ in the absence of epistasis for $i = 1$ to n where L is the progeny family mean and n is the number of individuals of the F₂ or B₁ + B₂ families sampled. Deviation from this expectation may be tested within each set of three families arising from the *i*th individual, as a *t* test using a standard error derived from the variances of the family means. Alternatively, we may compute the mean squared deviations from zero as:

$$\sum_{i=1}^n (L_{1t} + L_{2t} - 2L_{3t})^2 / 6nr,$$

for n degrees of freedom, where r is the number of replicate families (2 in the present experiment). In the absence of epistasis this statistic is expected to be non-significant when tested against its own replicate error mean square. In the presence of epistasis between pairs of unlinked genes this will test the significance of

$$\frac{1}{32}\bar{J} + \frac{1}{64}L + \frac{1}{16}[i]^2 \text{ for the } F_2 \text{ sample,}$$

and

$$\frac{1}{32}\bar{J} + \frac{1}{32}\sum^i j_{ik}j_{kl} + \frac{3}{256}L + \frac{1}{16}[i]^2 + \frac{1}{256}[i]^2 + \frac{1}{32}[i][l] \text{ for the } B_1 + B_2 \text{ sample.}$$

Using the symbols of Hayman and Mather (1955) in which i denotes homozygote \times homozygote interactions, j , heterozygote \times homozygote interactions and l the heterozygote \times heterozygote interactions and the definitions of i , j and l using the F_∞ notations (Van der Veen, 1959), $\bar{J} = \sum j^2$, $L = \sum l^2$ and $[i]$ and $[l]$ are the components of generation means for the i and l type interactions (Jinks and Jones, 1958; Jinks and Stevens, 1959).

We can similarly derive the squared mean deviation as

$$\left[\sum_{i=1}^n (L_{1t} + L_{2t} - 2L_{3t}) \right]^2 / 6nr$$

for 1 degree of freedom, which is also expected to be non-significant when tested against its own replicate error mean square in the absence of epistasis, but in its presence it tests the significance of

$$\frac{1}{18}[i]^2 \text{ for the } F_2 \text{ sample,}$$

and

$$\frac{1}{18}[i]^2 + \frac{1}{256}[l]^2 + \frac{1}{32}[i][l] \text{ for the } B_1 + B_2 \text{ sample.}$$

By correcting the mean squared deviation for the squared mean deviation we obtain a mean square which measures the variation of the deviations around their own mean, that is, epistasis \times crosses for $n-1$ degrees of freedom, which provides a test of significance, against its own replicate error, of:

$$\frac{1}{32}\bar{J} + \frac{1}{62}L \text{ for the } F_2 \text{ sample}$$

and

$$\frac{1}{32}\bar{J} + \frac{3}{256}L + \frac{1}{32}\sum^i j_{ik}j_{kl} \text{ for the } B_1 + B_2 \text{ sample.}$$

Hence, we can partition the epistatic component of variation for the F_2 sample into that due to i type interactions and that due to j and l type interactions. The analyses of variance for testing these portions of the epistatic variance are given in table 9.

For flowering time there is clear evidence for an epistatic component of variance of both kinds over the two samples. For height on the other hand, there is little evidence of epistasis other than a suggestion of an interaction between epistasis and the block environments in the F_2 sample.

(ii) *Model fitting*

Four models have been fitted to the variances in table 6. These are the four-parameter additive-dominance model, the nine-parameter model that also includes linkage, the genotype-environmental model which assumes

no linkage and the fifteen-parameter model which includes both linkage and genotype-environmental interactions (table 2, 4 and 5). The remainder mean squares after fitting each of these models (the observed minus expected mean squares) are given in table 10 along with the tests of significance of

TABLE 9

Analysis of variance of the triple test crosses to test for epistasis in the F_2 and $B_1 + B_2$ generations

Item	Flowering time			Height	
	d.f.	M.S.	χ^2	M.S.	χ^2
F_2 Sample					
1. Epistasis	1	28.0333	9.87***	1.6922	— N.S.
2. Epistasis \times crosses	39	3.8769	53.26*	6.6095	45.66 N.S.
1 \times blocks	1	0.5603	— N.S.	21.4630	3.80*
2 \times blocks	39	2.4378	33.49 N.S.	6.4040	44.24 N.S.
Replicates within blocks	951	2.8389	—	5.6454	—
$B_1 + B_2$ Sample					
1. Epistasis	1	4.0305	1.89 N.S.	3.0858	—
2. Epistasis \times crosses	27	4.4250	56.03****	6.2738	34.77 N.S.
1 \times blocks	1	2.3333	1.09 N.S.	0.0171	—
2 \times blocks	27	1.5980	20.23 N.S.	2.9437	16.32 N.S.
Replicates within blocks	946†	2.1323	—	4.8712	—

**** $P < 0.001$; * $P \approx 0.05$; N.S. $P < 0.05$.

† 948 degrees of freedom for height.

these mean squares against their replicate error mean squares. It is clear from these tests that all four models are inadequate for final height, therefore the only conclusion that we can safely draw from these tests is that an epistatic component of variance is present. The comparable analyses of time

TABLE 10

Variance ratio tests of the goodness of fit of the models specified

Model	No. of parameters	Degrees of freedom†	V.R.	P
Final height				
Additive dominance	4	32	2.1997	0.01-0.05
Additive dominance + Linkage	9	27	2.1605	0.01-0.05
Additive dominance + $G \times E$	9	27	2.7378	0.001-0.01
Additive dominance + Linkage + $G \times E$	15	21	2.5543	0.01-0.05
Flowering time				
Additive dominance	4	32	3.2230	0.001-0.01
Additive dominance + Linkage	9	27	3.3550	0.001-0.01
Additive dominance + $G \times E$	9	27	2.8803	0.001-0.01
Additive dominance + Linkage + $G \times E$	15	21	2.0511	≈ 0.05

† Degrees of freedom for the remainder mean square (numerator of the V.R.) and for its interaction with blocks (denominator of V.R.).

of flowering are similar to those for height except that there is no clear evidence of an epistatic component of variation since the remainder mean square after fitting the joint linkage, genotype-environmental interaction model has only borderline significance ($P = 0.05$).

Another way of looking at the results of the successive model fittings is presented in tables 11 and 12. These are the standard analyses of variance for partitioning the improvement in fit of successive models among the additional parameters included in these models. The results of two alternative sequences of model fitting are given for each character, one in which the inclusion of linkage precedes that of genotype-environmental interactions and the other in which the sequence is reversed. Because of the overall negative correlation between the linkage and interaction parameters, we detect linkage as the improvement in the goodness of fit of the model resulting from the inclusion of the linkage parameters after first allowing for the genotype-environmental interactions, that is, the improvement of fit of the model allowing for linkage and interactions over that which allows only for interactions. The reverse sequence provides the means of detecting genotype-environmental interactions. However, where the model which allows for both is inadequate due to the presence of epistasis any correlations between the epistatic contributions to the statistics and those due to linkage and genotype-environmental interactions will bias the mean squares attributable to the former in the first analysis, and the latter in the second analysis. In the only case where this correlation has been estimated, it was found to be positive (Opsahl, 1956). Since, therefore, the mean squares attributed to linkage and genotype-environmental interactions are probably inflated in the presence of epistasis, we cannot safely infer their presence unless their mean squares are significant against the significant mean squares attributed to epistasis. Even so, this will overestimate the significances of the linkage and genotype-environmental interaction components if their correlations with the epistatic component are greater than a half, but equally it will underestimate their significances if they are less than a half.

Bearing these points in mind, there is evidence of a significant epistatic component for final height and when tested against the latter the genotype-environmental interaction and linkage components are both found to be non-significant (table 11). For flowering time there is clear evidence of a genotype-environmental interaction component and borderline evidence for epistasis. If we assume no epistasis, the test for a linkage component is significant, if we accept the suggestion of its presence there is no evidence of linkage (table 12).

Since the scaling tests (section 4) gave clear evidence of the presence of repulsion linkages and genotype-environmental interactions for final height and flowering time, the estimates of the parameters of the model which includes both linkage and genotype-environmental interactions would seem to be more appropriate for further examination than those from the model which exclude both. Since there is some evidence of the presence of epistasis for both characters the estimates of these parameters will be biased to an unknown extent. However, because of this, the standard errors of the estimates have been computed from the remainder mean squares (epistasis mean squares of tables 11 and 12) after fitting the joint linkage, genotype-environmental interaction model. The estimates of the parameters and their standard errors are given in table 13.

Few of the estimates are significantly greater than zero ($P < 0.05$) although the overall picture is that expected from the results of the scaling tests (section 6 (i)). Thus while the estimates of D_1 are not significantly greater than those of D_2 (table 13) they are consistent with negative values of

TABLE 11
Analysis of variance of the improvement in goodness of fit of the models in table 10 partitioned among the components of the models for final height

		G × E interaction before linkage			Sequence of model fitting			Linkage before G × E interaction		
Item	d.f.	M.S.	V.R.	P.	Item	d.f.	M.S.	V.R.	P.	
1. Additive-dominance	4	9810·12	—	—	1. Additive-dominance	4	9810·12	—	—	
2. G × E interaction	5	49·47	—	—	2. Linkage	5	15·95	—	—	
3. Linkage	6	24·12	4·31(7)†	0·01-0·05	3. G × E interaction	6	52·05	1·59(7)†	N.S.	
			0·69(4)	N.S.				1·49(4)	N.S.	
4. Epistasis	21	34·95	2·55(8)	0·01-0·05	4. Epistasis	21	34·95	2·55(8)	0·01-0·05	
5. 1 × blocks	4	34·25	—	—	5. 1 × blocks	4	34·25	—	—	
6. 2 × blocks	5	38·19	—	—	6. 2 × blocks	5	5·53	—	—	
7. 3 × blocks	6	5·60	—	—	7. 3 × blocks	6	32·82	—	—	
8. 4 × blocks	21	13·68	—	—	8. 4 × blocks	21	13·68	—	—	

† The figure in brackets is the item used as the denominator in the V.R.

TABLE 12
Analysis of variance of the improvement in goodness of fit of the models in table 10 partitioned among the components of the models for flowering time

		Sequence of model fitting							
		G × E interaction before linkage			Linkage before G × E interaction				
Item	d.f.	M.S.	V.R.	P.	Item	d.f.	M.S.	V.R.	P.
1. Additive-dominance	4	2700.36	—	—	1. Additive-dominance	4	2700.36	—	—
2. G × E interaction	5	21.42	—	—	2. Linkage	5	3.99	—	—
3. Linkage	6	21.00	12.87(7)†	0.001-0.01	3. G × E interaction	6	35.52	12.05(7)†	0.001-0.01
			1.82(4)	N.S.				3.08(4)	0.01-0.05
4. Epistasis	21	11.52	2.05(8)	≅ 0.05	4. Epistasis	21	11.52	2.05(8)	≅ 0.05
5. 1 × blocks	4	95.85	—	—	5. 1 × blocks	4	95.85	—	—
6. 2 × blocks	5	3.93	—	—	6. 2 × blocks	5	2.35	—	—
7. 3 × blocks	6	1.63	—	—	7. 3 × blocks	6	2.95	—	—
8. 4 × blocks	21	5.62	—	—	8. 4 × blocks	21	5.62	—	—

† The figure in brackets is the item used as the denominator in the V.R.

Wd_1d_k , Wpd_1d_k and $Wp^2d_1d_k$ and hence with repulsion linkages. Similarly while the estimates of GE_{L1} , GE_{L2} , and GE_{L3} do not differ significantly their relative magnitudes are consistent with the interpretation of the genotype-environmental interactions derived from the scaling tests (section 4).

TABLE 13

The estimates of the genetic and genotype-environmental interaction parameters in the presence of linkage for the two characters final height and flowering time for the 21 generations derived from a cross between varieties 1 and 5 of Nicotiana rustica

Parameter	Final height	Flowering time
1. D_1	51.84 ± 12.32****	31.58 ± 7.07****
2. D_2	2.38 ± 32.20	10.65 ± 18.48
3. H_1	-4.39 ± 24.58	14.20 ± 14.11
4. H_2	17.33 ± 71.06	94.93 ± 40.80*
5. H_3	-21.72 ± 70.43	-63.08 ± 40.44
6. H_4	-56.46 ± 193.66	-94.61 ± 111.18
7. F_1	9.37 ± 6.46	0.81 ± 3.71
8. F_2	5.50 ± 18.72	0.85 ± 10.75
9. GE_{L1}	26.84 ± 3.87****	9.20 ± 2.22****
10. GE_{L2}	15.64 ± 3.87****	16.05 ± 2.22****
11. GE_{L3}	19.47 ± 3.67****	9.34 ± 2.11****
12. $WpG_{a1j}G_{akj}$	4.55 ± 2.84	3.66 ± 1.63
13. $WpG_{h1j}G_{hkj}$	-0.10 ± 6.24	4.97 ± 3.58
14. $Wp^2G_{h1j}G_{hkj}$	3.35 ± 9.62	15.32 ± 5.52**
15. $WpG_{a1j}G_{hkj}$	0.55 ± 4.08	1.45 ± 2.34

**** $P < 0.001$; ** $P = 0.01-0.02$; * $P = 0.02-0.05$.

7. CONCLUSIONS

We are now able to compare the advantages and disadvantages of scaling tests and sequential model fitting as methods of detecting genotype-environmental, linkage and epistatic components of variance in terms of the outcome of the analyses of the *N. rustica* experiment. For the detection of genotype-environmental interactions the two methods appear to be equally sensitive for flowering time but the scaling test is the more sensitive for final height (tables 7, 11 and 12). This, in general, is to be expected, because whereas the scaling test detects the presence of these interactions independently of the presence of epistasis, the sensitivity of the model fitting method depends on the magnitude of the uncorrelated epistatic component of variance; the larger the latter the less sensitive the test for the interactions becomes (tables 11 and 12).

While the scaling tests unambiguously detect the presence of breakable linkages among the genes controlling final height and give a clear indication for flowering time, model fitting provides no acceptable evidence of its presence. The reason for this insensitivity once again can be found in the correlation with the epistatic component. The effect of this correlation can be seen by comparing the significance levels of the linkage items in tables 11 and 12 when tested against their own block interactions (7) and when tested against the epistatic item (4). For both characters the former test, which is unacceptable in the presence of epistasis, is significant, and the latter test non-significant.

For detecting epistasis the two methods do not appear to differ. Thus the scaling tests gives the higher significance level for flowering time and

model fitting the higher significance for height. In general, however, we might expect model fitting to be less sensitive, even though both methods are equally unambiguous, because it only attributes to epistasis that part of the variance, due to this cause, that is not correlated with the additive-dominance effects, the genotype-environmental interactions or linkage.

As a means of determining the type of linkage, genotype-environmental interaction and epistasis the scaling tests invariably have the advantage over model fitting when all models fail. Thus in the *N. rustica* data we have been unable to establish the linkage phase, the kind of epistasis or the relative susceptibilities of different kinds of gene action to interaction with the environment, by model fitting. On the other hand, the scaling tests provide some information on all these points. However, the estimates of the parameters from the best fitting model (table 13) are consistent with the phase of linkage and kind of genotype-environmental interactions deduced from the scaling tests.

The power of the combined use of a novel experimental design (section 2), scaling tests, and a reversed sequence of model fitting can be assessed by comparing the results with those from previous investigations of the same pair of varieties. In none of the previous experiments and analyses have linkage, genotype-environmental interactions and epistatic components of variance been detected and estimated simultaneously, let alone classified as to type, although occasional evidence of linkage or epistasis has been obtained in a minority of experiments (Mather and Vines, 1952; Breese, personal communication; Opsahl, 1956; Hill, 1966).

Earlier analyses of the means of the 21 generations of the *N. rustica* experiment have shown that additive, dominance and epistatic gene action and linkage contribute to the differences among the means (Jinks and Perkins, 1969). Digenic interactions of a predominantly duplicate kind between pairs of genes, probably linked in repulsion, were detected and estimated for both characters, however, for flowering time there was evidence of more complex interactions involving more than two genes at a time. The analysis of means, of course, provides no information about within family environmental variation and its interaction with the genetic effects. It does, however, permit a more detailed analysis of the epistasis into digenic, trigenic, etc., components. But while an analysis of means can detect linkage between genes showing epistatic interactions it cannot reveal the presence of linkages between genes which have only additive and dominance action.

Furthermore, the action of genes which are dispersed between the inbred parents or whose contributions to the mean phenotype are ambidirectional cannot be detected by an analysis of generation means. Thus while there is a certain amount of agreement between the results of the analyses of the means and of the variances of the *N. rustica* experiment there are important differences in the kind of information they yield; the two analyses are complementary rather than alternatives. Hence, we can never substitute the statistically more satisfactory analysis of means for the analysis of the variances, for if we do we are in danger of overemphasising the importance of certain kinds of gene action and interaction for the variation among individuals within generations missing altogether other important sources of variation.

8. SUMMARY

1. Biometrical genetical models, scaling tests and sequential model fitting procedures are described for detecting, classifying and estimating the contributions of genotype-environmental interactions, linkage and epistasis when a simple model, which assumes their absence, fails to account for all the observed variation within the generations which can be derived from an initial cross between a pair of inbred strains.

2. The use of these procedures when all three causes of failure of the simple model are present simultaneously is illustrated by the analysis of 36 variances, for each of two characters, obtained from 21 generations of a cross between varieties 1 and 5 of *Nicotiana rustica*.

3. The results of the analyses show that the scaling tests are in general more sensitive than sequential model fitting procedures for detecting and classifying linkage, genotype-environmental interactions and epistasis when all models fail, although estimates of their contributions to the variances can only come from model fitting.

4. Comparison of these results with those of earlier investigations of the same cross leave no doubt of the greater sensitivity of both the analyses and the experimental design used here.

5. The advantages and disadvantages of biometrical genetical analyses of generation means and within generation variances are discussed and illustrated from the present analyses and an earlier analysis of the means of the same experiment. It is stressed that the two are complementary and not alternatives and extrapolation from the control of a character revealed at one level to that at the other can be misleading.

Acknowledgments.—The sequential model fitting was carried out on the KDF9 computer of the University of Birmingham.

9. REFERENCES

- BUCIO-ALANIS, L., PERKINS, JEAN M., AND JINKS, J. L. 1969. Environmental and genotype-environmental components of variability. V. Segregating generations. *Heredity*, *24*, 115-127.
- COOKE, PATRICIA, AND MATHER, K. 1962. Estimating the components of continuous variation. II. Genetical. *Heredity*, *17*, 211-236.
- HAYMAN, B. I., AND MATHER, K. 1955. The description of genic interactions in continuous variation. *Biometrics*, *2*, 69-82.
- HILL, J. 1966. Recurrent backcrossing in the study of quantitative inheritance. *Heredity*, *21*, 85-120.
- JINKS, J. L. 1956. The F_2 and backcross generations from a diallel set of 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 MATHER, K. 1955. Stability in development of heterozygotes and homozygotes. *Proc. Roy. Soc., B*, *143*, 561-578.
- JINKS, J. L., AND PERKINS, JEAN M. 1969. The detection of linked epistatic genes for a metrical trait. *Heredity*, *24*, 465-475.
- JINKS, J. L., AND STEVENS, J. M. 1959. The components of variation among family means in diallel crosses. *Genetics*, *44*, 297-308.
- JONES, R. MORLEY, AND MATHER, K. 1958. Interaction of genotype and environment in continuous variation. II. Analysis. *Biometrics*, *14*, 489-498.
- KEARSEY, M. J., AND JINKS, J. L. 1968. A general method of detecting additive, dominance and epistatic variation for metrical traits. I. Theory. *Heredity*, *23*, 403-409.
- MATHER, K. 1949. *Biometrical Genetics*. Methuen, London.

- MATHER, K., AND JONES, R. MORLEY. 1958. Interaction of genotype and environment in continuous variation. I. Description. *Biometrics*, 14, 343-359.
- MATHER, K., AND VINES, A. 1952. The inheritance of height and flowering time in a cross of *Nicotiana rustica* in *Quantitative Inheritance*, Ed. E. C. Reeve and C. H. Waddington, 49-79. H.M.S.O., London.
- OPSAHL, B. 1956. The discrimination of interactions and linkage in continuous variation. *Biometrics*, 10, 415-432.
- PERKINS, JEAN M., AND JINKS, J. L. 1968. Environmental and genotype-environmental components of variability. III. Multiple lines and crosses. *Heredity*, 23, 339-356.
- VAN DER VEEN, J. H. 1959. Tests of non-allelic interaction and linkage for quantitative characters in generations derived from two diploid pure lines. *Genetics*, 30, 201-232.

