ANALYSIS OF GENOTYPE × ENVIRONMENT INTERACTION IN TRIPLE TEST CROSS DATA

JEAN M. PERKINS and J. L. JINKS
Department of Genetics, University of Birmingham

Received 10.v.70

1. Introduction

The triple test cross breeding programme is an efficient method of detecting and partitioning an epistatic component of variation within a population and for estimating the additive and dominance components in the absence of epistasis (Kearsey and Jinks, 1968; Jinks, Perkins and Breese, 1969; Jinks and Perkins, 1970; Perkins and Jinks, 1970). In the present paper an analytical procedure for detecting interactions between the progeny genotypes of triple test crosses and micro-environmental effects is described which permits a partial separation of the interactions into those involving the additive effects of the genes and those involving non-additive effects. An extension of the experimental design and analysis is also described which permits the detection of interaction between the additive, dominance and epistatic effects of the genes and macro-environmental differences. Both are illustrated by reference to triple test crossing programmes carried out on the F_2's of crosses between inbred varieties of *Nicotiana rustica*.

2. Interactions with the micro-environment

The contributions of a single gene difference A, a and a single micro-environmental difference to the means and variances of the L_1 and L_2 progeny families of a triple test cross produced by crossing individuals of an F_2 population to inbred lines (P_1 and P_2) are summarised in table 1. The symbols are those used by Kearsey and Jinks (1968) and Perkins and Jinks (1970).

Two methods are used to detect interactions between genotypes and micro-environmental differences. One is to find heterogeneity of the variance within families where in the absence of interactions the variances are expected to be homogeneous; the other is to find a correlation between the means and variances of families where in the absence of interaction no correlation is expected (Mather, 1949; Perkins and Jinks, 1970; Jinks and Fulker, 1970). Neither of these tests is applicable to the L_1 and L_2 families of a triple test cross. Thus examination of the expectations in table 1 shows that the variances within the families could be heterogeneous and the means and variances of families could be correlated in the absence of interactions, that is, when $g_d = g_h = 0$. If, however, we combine the L_1 and L_2 family means and variances which have the same F_2 parent to give half the sums and differences for the means and variances ($\frac{1}{2}(\bar{L}_{1i}+\bar{L}_{2i})$, $\frac{1}{2}(\bar{L}_{1i}-\bar{L}_{2i})$, $\frac{1}{2}(\sigma^2_{L1i}+\sigma^2_{L2i})$ and $\frac{1}{2}(\sigma^2_{L1i}-\sigma^2_{L2i})$ in table 1) this is no longer the case. The second degree statistics, $\frac{1}{2}(\sigma^2_{L1i}+\sigma^2_{L2i})$ and $\frac{1}{2}(\sigma^2_{L1i}-\sigma^2_{L2i})$, are still expected to be heterogeneous in the absence of interactions but there is now no correlation between corresponding first and second degree statistics, that is, between

$\frac{1}{2}(\bar{L}_{1i}+\bar{L}_{2i})$ and $\frac{1}{2}(\sigma^2_{L1i}+\sigma^2_{L2i})$ and between $\frac{1}{2}(\bar{L}_{1i}-\bar{L}_{2i})$ and $\frac{1}{2}(\sigma^2_{L1i}-\sigma^2_{L2i})$ in the absence of interactions. In the presence of interactions the covariances between these statistics have the expectations:

$$\text{cov. } \tfrac{1}{2}(\bar{L}_{1i}+\bar{L}_{2i}) \,.\, \tfrac{1}{2}(\sigma^2_{L1i}+\sigma^2_{L2i}) \;=\; \tfrac{1}{4}dg_d e$$

$$\text{cov. } \tfrac{1}{2}(L_{1i}-L_{2i}) \,.\, \tfrac{1}{2}(\sigma_{L1i}-\sigma^2_{L2i}) \;=\; -\tfrac{1}{8}hg^2_d+\tfrac{1}{8}hg^2_h+\tfrac{1}{4}hg_h e$$

which are zero in the absence of interactions, *i.e.* when $g_d = g_h = 0$.

TABLE 1

A. *Contributions of single gene difference and single micro-environmental factor to the means of the L_1 and L_2 types of families ($\bar{L}_{1i}$ and $\bar{L}_{2i}$)*

		F_2 population			
Parents	Geno-types	Genotypes *AA* Frequencies $\frac{1}{4}$	*Aa* $\frac{1}{2}$	*aa* $\frac{1}{4}$	
$\times P_1$	*AA*	d	$\frac{1}{2}d+\frac{1}{2}h$	h	L_1 families
$\times P_2$	*aa*	h	$-\frac{1}{2}d+\frac{1}{2}h$	$-d$	L_2 families
$\frac{1}{2}(\bar{L}_{1i}+\bar{L}_{2i})$		$\frac{1}{2}d+\frac{1}{2}h$	$\frac{1}{2}h$	$-\frac{1}{2}d+\frac{1}{2}h$	
$\frac{1}{2}(\bar{L}_{1i}-\bar{L}_{2i})$		$\frac{1}{2}d-\frac{1}{2}h$	$\frac{1}{2}d$	$\frac{1}{2}d+\frac{1}{2}h$	

B. *Contributions to variances (σ^2_{L1i} and σ^2_{L2i})*

$\times P_1$	*AA*	$e^2+g^2_d+2g_d e$	$\frac{1}{4}d^2+\frac{1}{4}h^2-\frac{1}{2}dh+e^2$ $+\frac{1}{2}g^2_d+\frac{1}{2}g^2_h+g_d e+g_h e$	$e^2+g^2_h+2g_h e$	L_1 families
$\times P_2$	*aa*	$e^2+g^2_h+2g_h e$	$\frac{1}{4}d^2+\frac{1}{4}h^2+\frac{1}{2}dh+e^2$ $+\frac{1}{2}g^2_d+\frac{1}{2}g^2_h-g_d e+g_h e$	$e^2+g^2_d-2g_d e$	L_2 families
$\frac{1}{2}(\sigma^2_{L1i}+\sigma^2_{L2i})$		$e^2+\frac{1}{2}g^2_d+\frac{1}{2}g^2_h$ $+g_d e+g_h e$	$e^2+\frac{1}{4}d^2+\frac{1}{4}h^2+\frac{1}{2}g^2_h$ $+g_h e$	$e^2+\frac{1}{2}g^2_d+\frac{1}{2}g^2_h$ $-g_d e+g_h e$	
$\frac{1}{2}(\sigma^2_{L1i}-\sigma^2_{L2i})$		$\frac{1}{2}g^2_d-\frac{1}{2}g^2_h$ $+g_d e-g_h e$	$-\frac{1}{2}dh+g_d e$	$-\frac{1}{2}g^2_d+\frac{1}{2}g^2_h$ $+g_d e+g_h e$	

On extending these expressions to many genes new terms appear in the expectations (*see* Perkins and Jinks, 1970). Thus for a pair of independent genes, $A-a$ and $B-b$ the two covariances become:

$$\tfrac{1}{4}d_a g_{da} e+\tfrac{1}{4}d_b g_{db} e+\tfrac{1}{8}(d_a-d_b)(g_{da}g_{hb}-g_{ha}g_{db})$$

and

$$\tfrac{1}{8}h_a(g^2_{ha}-g^2_{da}-g_{ha}g_{db}+2g_{ha}e)+\tfrac{1}{8}h_b(g^2_{hb}-g^2_{hb}-g_{da}g_{hb}+2g_{hb}e),$$

respectively.

This increase in complexity, however, does not alter the conclusion that providing the genes are independent, significance of either covariance is indicative of the presence of interactions with micro-environmental differences. Since the covariance based on the average family mean and variance depends on g_d it specifically detects interactions with the additive effects of the genes. Similarly, the covariance based on differences largely, but not exclusively, detects interactions with the dominance effects of the genes, g_h.

We may illustrate the analysis with the data from a triple test cross based on the F_2 of a cross between inbred varieties 1 and 5 of *Nicotiana rustica*

grown in 1965. Forty plants (i = 1 to 40) were crossed to parents 1 and 5 (P_2 and P_1, respectively) and to their F_1. The 120 progeny families were grown in two replicate blocks with individually randomised plants of each

TABLE 2

The values of $\frac{1}{2}(L_{1i}+L_{2i})$ *and* $\frac{1}{2}(L_{1i}-L_{2i})$ *for the means and variances of* 40 L_1 *and* L_2 *families of a triple test cross for the character final height*

Family	$\frac{1}{2}(L_{1i}+L_{2i})$		$\frac{1}{2}(L_{1i}-L_{2i})$	
i	mean	variance	mean	variance
1	60·70	26·600	3·10	−24·450
2	62·95	9·225	2·05	4·725
3	57·65	9·350	3·65	1·900
4	58·60	29·150	5·50	−5·600
5	61·25	19·775	5·35	2·775
6	61·40	16·550	1·90	−4·650
7	59·45	17·325	0·65	−6·075
8	61·15	15·250	2·05	4·750
9	59·90	13·500	4·70	4·350
10	58·05	25·350	2·65	−5·200
11	61·55	20·200	4·45	1·600
12	63·25	35·400	1·85	−5·400
13	58·75	15·950	4·55	−8·050
14	58·60	35·675	4·50	9·575
15	60·05	24·775	1·85	5·575
16	60·25	32·325	2·75	0·375
17	59·85	22·125	1·95	−2·375
18	58·75	24·450	−0·45	3·200
19	60·05	20·075	4·05	−5·475
20	58·45	25·075	3·85	−7·175
21	57·15	11·750	2·65	−0·250
22	58·35	23·750	2·55	8·850
23	57·40	53·900	0·40	0·600
24	54·45	33·675	5·75	21·775
25	59·55	38·750	3·35	21·250
26	59·55	25·175	4·35	−11·325
27	57·15	27·225	−0·25	−0·375
28	59·50	21·425	6·10	−4·075
29	59·50	19·900	−0·70	4·000
30	58·00	19·950	5·20	−10·800
31	60·00	16·275	3·10	−0·525
32	58·32	39·715	4·19	11·085
33	58·35	37·875	3·55	−31·525
34	58·95	20·875	5·35	9·025
35	58·25	43·950	4·05	16·050
36	60·75	21·275	2·85	18·525
37	59·37	17·390	1·17	4·240
38	57·35	26·625	−0·95	6·125
39	55·25	23·050	3·45	−9·600
40	56·60	27·875	2·70	11·325
Covariance		−176·09		−18·65
Correlation (38 d.f.)		−0·259		−0·024

family in each block. (For further details *see* Jinks and Perkins (1970) and Perkins and Jinks (1970).) The final heights of these progenies will be analysed as this character satisfies the criterion of independence of the genes in these data (Jinks and Perkins, 1970). The values of $\frac{1}{2}(L_{1i}+L_{2i})$ and $\frac{1}{2}(L_{1i}-L_{2i})$ for the means and variances of the 40 pairs of L_1 and L_2 families averaged over blocks are given in table 2.

Plots of $\frac{1}{2}(\sigma^2_{L1i}+\sigma^2_{L2i})$ against $\frac{1}{2}(\bar{L}_{1i}+\bar{L}_{2i})$ and of $\frac{1}{2}(\sigma^2_{L1i}-\sigma^2_{L2i})$ against $\frac{1}{2}(\bar{L}_{1i}-\bar{L}_{2i})$ reveal no obvious linear or curvilinear relationships. The covariances given at the foot of table 2 and tests of significance based on the corresponding correlations confirm the absence of any relationship. There is, therefore, no evidence of interactions with the micro-environment.

This conclusion is at variance with the previous demonstration of interaction of varieties 1 and 5 and their F_1 with the micro-environment for the same character in the same experiment from which the triple test cross data were taken (Perkins and Jinks, 1970). Parental and F_1 data, however, provide the most sensitive test for genotype × environment interaction that is currently available, namely, the test of homogeneity of their variances within families. But there is no simple relationship between the means and variances of these three kinds of families and it is on the existence of such a relationship that the only test for interactions among triple test cross families depends. This suggests that a relationship between means and variances of the kind sought in the triple test cross data may be a less efficient method of detecting interactions in the families derived from the 1 × 5 cross. It could equally be argued, however, that because the genotypes of varieties 1 and 5 and their F_1 interact with the micro-environment it does not necessarily follow that the wider range of genotypes resulting from their recombination and reassortment in the progenies of the triple test cross will also do so. In which case the difference between the two analyses is not just a reflection of a difference in their efficiencies in detecting genotype × environment interactions.

3. INTERACTIONS WITH THE MACRO-ENVIRONMENT

If r individuals of each of the L_{1i}, L_{2i} and L_{3i} progeny families of a triple test cross, where $i = 1$ to n are raised in each of s environments we can extract the following items from an analysis of variance (*see* Kearsey and Jinks, 1968; Jinks and Perkins, 1970).

Item	d.f.	e.m.s.
Sums $(\bar{L}_{1i}+\bar{L}_{2i}+\bar{L}_{3i})$	$n-1$	$\sigma^2_{e1}+3r\sigma^2_{sm}+3rs\sigma^2_{m}$
Sums × Environments	$(s-1)(n-1)$	$\sigma^2_{e1}+3r\sigma^2_{sm}$
Error	$3ns(r-1)$	σ^2_{e1}
Differences $(\bar{L}_{1i}-\bar{L}_{2i})$	$n-1$	$\sigma^2_{e2}+2r\sigma^2_{sml}+2rs\sigma^2_{ml}$
Differences × Environments	$(s-1)(n-1)$	$\sigma^2_{e2}+2r\sigma^2_{sml}$
Error	$2ns(r-1)$	σ^2_{e2}

There are, of course, other items in the complete analysis of variance, but the above are the only ones of direct interest to a biometrical genetical analysis. In the absence of epistasis (see later) significant sums × environments and differences × environments items show the presence of interactions between the environments and the additive and dominance effects of the genes, respectively. Estimates of σ^2_{sm} and σ^2_{sml} provide measures of these interactions which are directly comparable with the estimates of the additive and dominance components of variation, σ^2_m and σ^2_{ml}. The specification of these σ^2's in terms of the parameters of the biometrical genetical model will depend on the nature of the parental population used in the triple test cross

(Kearsey and Jinks, 1968). For example, for the simplest situation, an F_2 population with no linkage:

$$\sigma^2_m = \tfrac{1}{8}D$$

$$\sigma^2_{ml} = \tfrac{1}{8}H$$

$$\sigma^2_{sm} = \tfrac{1}{8}G_{2D}$$

$$\sigma^2_{sml} = \tfrac{1}{8}G_{2H}$$

Where D and H are as originally defined by Mather (1949) and $G_{2D} = \Sigma g^2_{dij}$ and $G_{2H} = \Sigma g^2_{hij}$ are the macro-environmental equivalents of the micro-environmental interaction components defined by Perkins and Jinks (1970).

In the presence of epistasis the additive and dominance components and their interactions with the environments will be confounded with contributions from the epistatic effects of the genes. However, in these circumstances we can still unambiguously detect the presence of an epistatic component of variation and its interaction with the environment. This requires only that we obtain the epistatic items of the analyses of variance described by either Kearsey and Jinks (1968) or Jinks and Perkins (1970) and the interaction of these items with the environment.

These analyses may be illustrated by a triple test cross on the F_2 of a cross between inbred varieties 2 and 12 of *Nicotiana rustica*. A sample of 18 F_2 individuals ($i = 1$ to 18, $n = 18$) were crossed to 2, 12 and their F_1 (2×12) and five individuals of each progeny were grown in 1969 as individually randomised plants in each of two replicate blocks in each of two environments ($s = 2$). The two environments were two sowing dates, 23rd April and 21st May, which are the earliest and latest that these genotypes can be successfully grown in the Birmingham area. The characters to be analysed are final height and the number of days from sowing to flowering.

The relevant items from the complete analysis of variance of these data are given in table 3. The within families within blocks items (4, 8 and 15) have been used to test, as χ^2's, the significance of the other items except where one or more of the block or interaction with environment items are significant. In the latter cases the interaction item has been tested as a variance ratio against the appropriate significant block item and the main effects against the appropriate significant interaction item. Reference to table 3 shows that there is a significant epistatic component of variation for both characters. For flowering time this is mainly an interaction between homozygous combinations of genes (item 9) while for final height it is mainly an interaction involving heterozygotes (item 10). In the presence of epistasis the sums and differences items will detect and measure the variation due to additive and dominance effects of the genes confounded by epistatic variation. While this can lead to difficulties in interpretation, in the present context it is of little concern since the principal evidence of interaction with the two environments implicates only the epistatic component of variation for final height (item 11). This result is compatible with an earlier finding that the epistatic component of variation for height is significant for the second sowing but not for the first.

It would appear from this analysis that the epistatic action of the genes controlling final height are more sensitive to the environmental difference

than their additive or dominance action. This agrees with earlier analyses of the same character in a number of varieties of *N. rustica* in which the relative sensitivities of the additive, dominance and epistatic action of the genes to seasonal differences were compared (Jinks and Stevens, 1959; Perkins, 1970). In the present case it would appear to be the epistatic action involving homozygous combination of genes (item 11) that is sensitive to the environmental difference rather than that involving heterozygous combinations (item 12).

TABLE 3

Analysis of triple test cross grown in two environments

Item	Flowering time d.f.	Flowering time M.S.	Flowering time $\chi^2(4)$	Flowering time VR(7)	Height M.S.	Height $\chi^2(4)$	Height VR(7)
1 Sums	17	115·04	***		921·32	***	
2 Sums × Environments	17	9·86	NS		63·78	5%	
3 Blocks within Environments	34	8·81	NS		53·24	NS	
4 Within families within blocks	1924†	6·26	—		39·59	—	
			$\chi^2(8)$	VR(7)		$\chi^2(8)$	VR(7)
5 Differences	17	45·37	—	***	490·81	—	***
6 Differences × Environments	17	16·28	—	NS	68·36	—	NS
7 Blocks within Environments	34	10·93	**	—	68·46	**	—
8 Within families within blocks	1277†	6·42	—	—	36·07	—	—
			$\chi^2(15)$			$\chi^2(15)$	VR(11)
9 Overall epistasis	1	93·77	***		155·35	—	NS
10 Epistasis	17	16·62	*		80·43	**	—
11 Overall epistasis × Environments	1	3·06	NS		321·82	**	—
12 Epistasis × Environments	17	6·26	NS		50·93	NS	—
13 Overall blocks within Environments	2	6·71	NS		54·20	NS	—
14 Blocks within Environments	34	8·63	NS		47·51	NS	—
15 Within families within blocks	1924†	6·26	—		39·59	—	—

† The degree of freedom for the within family items are one greater for height.
NS $P > 0{\cdot}05$. * $P = 0{\cdot}05 - 0{\cdot}01$. ** $P = 0{\cdot}01 - 0{\cdot}001$. *** $P < 0{\cdot}001$.

4. SUMMARY

1. An analysis of the means and variances of the families of a triple test cross breeding programme is described which tests for interactions of genotypes with micro-environmental effects.

2. This analysis partially partitions these interactions into those involving the additive action of the genes and those involving the dominance action.

3. An extension of the triple test cross experimental design to provide tests for interactions of the genotypes with macro-environmental effects is also described.

4. This extension allows the interactions between the environment and the additive, dominance and epistatic effects of the genes to be independently detected and measured.

5. Both analyses are illustrated by data from triple test crosses involving F_2 populations derived from inbred varieties of *Nicotiana rustica*.

5. References

JINKS, J. L., AND FULKER, D. W. 1970. Comparison of the biometrical genetical, Mava and classical approaches to the analysis of human behaviour. *Psychological Bull.*, *73*, 311-349.

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 PERKINS, JEAN M. 1970. A general method of detecting additive, dominance and epistatic components of variation III F_2 and backcross populations. *Heredity*, *25*, 419-429.

JINKS, J. L., PERKINS, JEAN M., AND BREESE, E. L. 1969. A general method of detecting additive, dominance and epistatic components of variation for metrical traits. II. Application to inbred lines. *Heredity*, *24*, 45-57.

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*, *23*, 403-409.

MATHER, K. 1949. *Biometrical Genetics. The Study of Continuous Variation.* Methuen, London.

PERKINS, JEAN M. 1970. Environmental and genotype-environmental components of variability. VI. Diallel sets of crosses. *Heredity*, *25*, 29-40.

PERKINS, JEAN M., AND JINKS, J. L. 1970. Detection and estimation of genotype-environmental, linkage and epistatic components of variation for a metrical trait. *Heredity*, *25*, 157-177.