EXPERIMENTAL SIZES FOR DETECTING DOMINANCE VARIATION

M. J. KEARSEY

Department of Genetics, University of Birmingham, England

1. INTRODUCTION

THE techniques of biometrical genetics provide many potentially useful approaches to the study of the genetical architecture of metrical traits. However, the more detailed and specialised information one requires the more restricted is the range of material that may be investigated. For example the work of Jinks and Perkins (1969) and Perkins and Jinks (1970) on *Nicotiana rustica* has provided information not only on the additive and dominance properties of genes, but also allowed epistasis, linkage and genotype-environment interaction to be investigated in crosses between a few inbred lines.

The population geneticist and breeder alike, on the other hand, wish to obtain information about their populations, but such material poses many difficulties. Nevertheless, by means of various multiple mating schemes, it is, at least theoretically, possible to partition the variation within and between full and half-sib families into additive genetic, non-additive genetic and environmental components.

The present paper is concerned with the practical utility of such designs simply for the purpose of detecting dominance variation, assuming that the additive variance is easily estimated.

This problem was originally considered by Comstock and Robinson (1952) in connection with overdominance. Single locus overdominance is no longer widely held to be an important feature of non-additive variance, and in fact it appears that for characters for which epistasis is not a significant feature of the genetic architecture, dominance ratios appear to be considerably less than unity (Kearsey and Kojima, 1967; Mather, 1960). Since non-additive components can be unambiguously ascribed to dominance only in the absence of epistasis, restricting our attention to such characters is not too unrealistic.

There are three questions we might ask about the dominance properties of a character. Firstly, is dominance present? Secondly, what is the average magnitude of the dominance ratio? And thirdly, if dominance is a feature of the genetic control, is there a directional element to it? In this paper attention will be restricted to the first and last points above, the possibility of answering the second being very dependent on the other two.

Pursuing the first point, then, we will consider just how large an experiment it is necessary to raise to detect dominance by means of an F test of appropriate mean squares, and see to what extent we might reduce the total experimental size by inbreeding prior to the experiment. Obviously

the necessary experimental size will depend on the magnitude of dominance, and a range of dominance ratios up to 1·0 are considered.

In the limiting case of the additive effects (d_i) the absolute values of the dominance effects $(|h_i|)$ and allele frequencies (u_i) being constants (d, h, u) at all the loci $(i = l - k)$ controlling the character the population can be completely defined in terms of three parameters. These are the dominance ratio $b\left(= \left|\dfrac{h}{d}\right|\right)$; the narrow heritability, $h_n^2\left(= \dfrac{V_A}{V_P}\right.$ where V_A, V_P are the additive genetic and total phenotypic variances respectively$\Big)$ and the frequency of the increasing allele (u).

The narrow heritability of the original population can be determined with some accuracy and thus is a useful standard. It is then possible, as will be shown, to determine the minimum experimental size needed to detect dominance for a given value of u and to show how this size varies with allele frequency.

2. EXPERIMENTAL DESIGNS

There are basically four crossing schemes which allow direct tests for non-additive variation for a metrical trait in a randomly mating population. In the absence of epistasis, non-nuclear effects and genotype-environment interaction this variation is due to dominance irrespective of the magnitude of linkage disequilibrium. It will be assumed for the purposes of illustration that the progeny are all individually randomised in a single block, with r individuals in every family. Obviously such a design restricts the type of organisms for which it is suitable, but the randomised plot design will be considered in the discussion.

The four crossing schemes, together with a summary of their appropriate analysis of variance, are as follows.

(i) The Experiment II of Comstock and Robinson (1952) in which n_1 males and n_2 females are chosen at random from the population and crossed in all combinations. This design will be referred to as N.C. Exp. II. Provided $n_1 = n_2 = n$ the analysis takes the following form:

Source	d.f.	M.S.	e.m.s.
Between male parents	$n-1$	MS_4	—
Between female parents	$n-1$	MS_3	—
Males × Females	$(n-1)^2$	MS_2	$\sigma_1^2 + r\sigma_2^2$
Within families	$n^2(r-1)$	MS_1	σ_1^2

Here, as in the other designs described below, only the expected mean squares (e.m.s.) of those items relevant to the detection of dominance are indicated. The genetical and environmental components of σ_1^2, σ_2^2 are shown for all designs in table 1 in terms of the degree of inbreeding, f (Jinks and Broadhurst, 1965).

(ii) The Experiment III of Comstock and Robinson (1952)—N.C. Exp. III. As originally described, this involved sampling n males from a randomly mated derivative of a cross between two inbred lines (i.e. an F_2 or later generation). Each male was then crossed to females of both inbred lines. When the population is of unknown ancestry the inbred lines can be replaced

by extreme selection lines from the population (Kearsey and Jinks, 1968). The analysis of such data is as shown below.

Source	d.f.	M.S.	e.m.s.
Sums (male groups)	$n-1$	MS_3	—
Differences (males $\times$ fenales)	$n-1$	MS_2	$\sigma_1^2 + 2r\sigma_2^2$
Within families	$2n(r-1)$	MS_1	σ_1^2

(iii) The diallel cross (Hayman, 1954) in which n individuals are sampled and crossed both as male and female parents.

TABLE 1

Genetic and environmental components of expected mean squares for the four experimental designs. (For explanation, see text)

Genetic components

Design	e.m.s.	$\frac{1}{4}D_1$ Σuvd^2	$\frac{1}{4}H_1$ Σuvh^2	$\frac{1}{16}H_2$ $\Sigma u^2v^2h^2$	$\frac{1}{8}F_1$ $\Sigma uv(u-v)dh$	E e^2
N.C. Exp. II	σ_2^2	—	—	$(1+f)^2$	—	—
and Diallel	σ_1^2	$(1-f)$	$(1-f)$	$-(1-f)^2$	$-(1-f)$	1
N.C. Exp. III	σ_2^2	—	—	$(1+f)/2$	—	—
	σ_1^2	$(1-f)/2$	$(1-f)/2$	—	—	1
A.B.I.P.'s	σ_2^2	—	$(1+7f)/16$	$[1-f(10-f)]/8$	—	—
	σ_1^2	$(1-f)$	$3(1-f)/4$	$-(1-f)^2/2$	$-(1-f)$	1

In a complete diallel all n^2 matings are constructed, but selfs and/or one set of reciprocal crosses may be omitted (Griffing, 1956; Jones, 1965). Although the parents selfed do not contribute to the mean square for dominance variance (b_3 in Hayman's notation) we shall, however, consider the complete diallel cross. The relevant items in the analysis of variance are:

Source	d.f.	M.S.	e.m.s.
b_3	$\dfrac{n(n-3)}{2}$	MS_2	$\sigma_1^2 + 2r\sigma_2^2$
Within families	$n^2(r-1)$	MS_1	σ_1^2

The b_3 MS in Hayman's notation is identical to Griffings specific combining ability (Griffing, 1956).

(iv) This is an extension of the biparental mating scheme (B.I.P.'s) described by Mather (1949) to include families derived from selfing the parents and will be referred to as Augmented B.I.P.'s (A.B.I.P.'s).

Since the analysis and interpretation of this has not been described previously a brief account will be presented here. A sample of $2n$ individuals are taken and paired off at random. Call the parents of the jth pair P_{j1}, P_{j2}, where $j = 1, 2, \ldots n$. Four full-sib families can be produced from every such pair by selfing each parent ($F_{j1\cdot1}$, $F_{j2\cdot2}$) and by reciprocally crossing

$(F_{j1\cdot2}, F_{j2\cdot1})$. The following comparisons can then usefully be made between the four families.

$F_{j1\cdot1}$	$F_{j1\cdot2}$	$F_{j2\cdot1}$	$F_{j2\cdot2}$	Comparison
1	0	0	-1	C_{j1}
-1	1	1	-1	C_{j2}
0	1	-1	0	C_{j3}

C_{j3} is a measure of reciprocal differences and in the absence of such effects should be zero for all sets. Thus $\sum_{j=1}^{n} \frac{C_{j3}^2}{2r}$ will be a sum of squares (S.S.) with n d.f. testing for reciprocal differences. Similarly the S.S. for C_{j1} (again with n d.f.) measures the variation of selfed families (see Dickinson and Jinks, 1956).

Interest in this design centres around C_{j2}. Consider a set of parents derived, from a population in linkage equilibrium, by inbreeding, without selection, to some arbitary degree f. Then the frequencies of the three genotypes at some locus with alleles A, a are:

$$\begin{array}{ccc} AA & Aa & aa \\ u^2+uvf & 2uv(1-f) & v^2+uvf \\ \alpha & \beta & \gamma \end{array}$$

Taking pairs of individuals at random results in the matings shown in table 2. Also shown in this table are the genetic values of the selfed and crossed family

TABLE 2

Parental genotypes and the genetic values of the progeny for the A.B.I.P.'s design. (For explanation, see text.)

Frequency	Parental genotypes P_{j1} P_{j2}	Genetic values of progeny family means $F_{j1\cdot1}$ $F_{j2\cdot2}$ $F_{j1\cdot2}$ $F_{j2\cdot1}$	$C_{j2}/4$ $(F_{j1\cdot2}+F_{j2\cdot1}-F_{j1\cdot1}-F_{j2\cdot2})/4$
α^2	$AA \times AA$	d d d	0
$2\alpha\beta$	$AA \times Aa$	d $\frac12 h$ $\frac12(d+h)$	$\frac14 h$
$2\alpha\gamma$	$AA \times aa$	d $-d$ h	$\frac12 h$
β^2	$Aa \times Aa$	$\frac12 h$ $\frac12 h$ $\frac12 h$	0
$2\beta\gamma$	$Aa \times aa$	$\frac12 h$ $-d$ $\frac12(-d+h)$	$\frac14 h$
γ^2	$aa \times aa$	$-d$ $-d$ $-d$	0
		mean	$\frac14(1+f)uvh$

means and of the comparison $C_{j2}/4$. Since C_{j2} is always a function of h alone, $\sum_{j=1}^{n} \frac{C_{j2}^2}{4r}$ is a S.S. with n d.f. measuring dominance. However, it is convenient to partition this S.S. into that S.S. for deviations of the C_{j2} around their own mean and the correction factor. The former can be readily expressed in conventional dominance components (table 1) while the correction factor is a measure of directional dominance analagous to the b_1 item in Hayman's (1954) analysis of diallel tables. The analysis of variance takes the form:

Source	d.f.	M.S.	e.m.s.
Directional dominance	1	—	—
Dominance variation (C_2)	$n-1$	MS_2	$\sigma_1^2+4r\sigma_2^2$
Within families	$4n(r-1)$	MS_1	σ_1^2

The variance components are defined in the terminology of Dickinson and Jinks (1956).

3. METHOD OF ESTIMATING EXPERIMENTAL SIZES

(a) *Biometrical*

It can be seen from table 1 that the genetic components of variance from all designs can be expressed using four components (D_1, H_1, H_2, F_1). Provided that at the k loci controlling the character u_t, d_t and h_t are constants (u, d, h) and that the dominance ratio $(b) = h/d$, then

$$D_1 = 4kuvd^2$$
$$H_1 = 4kuvd^2b^2$$
$$H_2 = 16ku^2v^2d^2b^2$$
$$F_1 = 8kuv(u-v)d^2b.$$

Further, since we are interested only in the relative values of the mean squares and not their absolute values, we can put the total phenotypic variance equal to unity. Thus the narrow heritability equals the additive genetic variance $VA(= \tfrac{1}{2}D_R)$.

$$i.e. \quad h_n^2 = \tfrac{1}{2}D_R = 2kuvd^2 [1+b(v-u)]^2.$$

From which it follows that

$$kuvd^2 = h_n^2/2 [1+b(v-u)]^2.$$

The right-hand side of this equation contains all the variables we are going to consider, and thus we can calculate D_1, H_1, H_2, and F_1. The environmental variance is calculated as:

$$E = 1 - h_n^2 - \tfrac{1}{4}H_2.$$

The parameters computed as above are then combined to give values for the e.m.s. according to the relationships set out in table 1. Finally the expected values of MS_1, MS_2 are calculated for each design for a given family size, r.

The only unknown in this sytem is n. The problem is first to find that value of n, for a given family size and genetic situation, which will result in MS_2 being significant at some level (α_1) on a certain proportion of occasions (α_2). And secondly, to find that combination of n and r which minimises the total experimental size.

Although the approach to the problem is described below for general values of α_1, α_2 we have used the values 0·05 and 0·95 respectively. That is, we want to find the values of n for which MS_2 is significant at the 0·05 level on 95 per cent. of occasions.

(b) *Statistical*

Let the degrees of freedom of MS_1, MS_2 be df_1, df_2 respectively and further let

$$\frac{\epsilon MS_2}{\epsilon MS_1} = k; \quad \frac{MS_2}{MS_1} \text{ (observed)} = k^\circ.$$

We require that k° be significant at some level (α_1) on a certain proportion of occasions (α_2).

2 M

The tables of F give $F\alpha_i(1)$, the value of F corresponding to some probability α_i, calculated on the assumption that $k = 1$. If, in fact, k is greater than unity, the value of F corresponding to α_i will be obtained by multiplying the tabulated value by k. Call this value $F_{\alpha_i}(k)$.

The null hypothesis of no dominance is rejected if

$$F \geqslant F_{\alpha_i}(1).$$

We wish this to happen on α_2 of occasions. If we write $F(k)$ for the probability density of F, given a value of k greater than unity, we want F such that

$$\int_F^\infty F(k)dF(k) = \alpha_2.$$

We may call this F value $F_{\alpha_2}(k)$. Thus we require

$$F_{\alpha_1}(1) = F_{\alpha_2}(k).$$

But as explained above

$$F_{\alpha_2}(k) = kF_{\alpha_2}(1)$$

and hence we require

$$F_{\alpha_1}(1) = kF_{\alpha_2}(1)$$

$$\text{or } k = \frac{F_{\alpha_1}(1)}{F_{\alpha_2}(1)}. \tag{1}$$

For large df_1, F can be replaced by χ^2 such that

$$k = \frac{\chi^2_{\alpha 1}}{\chi^2_{\alpha 2}}. \tag{2}$$

In the present paper α_1, α_2 are set at 0·05, 0·95 respectively, *i.e.* we require k° to be significant at P = 0·05 on 95 per cent. of occasions.

The problem is thus to find the degrees of freedom (df_1, df_2) that satisfy equations (1) and (2). In all the designs (see section 2),

$$k = \frac{\sigma_1^2 + rt\sigma_2^2}{\sigma_1^2} = 1 + \frac{rt\sigma_2^2}{\sigma_1^2}. \tag{3}$$

(Where $t = 1$ for N.C. Exp. II; $= 2$ for diallel and N.C. Exp. III; $= 4$ for A.B.I.P.'s; and df_1, df_2 are functions of n and r).

To calculate n for a given value of r the approximate method suggested by Comstock and Robinson (1952) was used. That is for large df_1, df_2 we can use the relationship

$$\sigma_z^2 = \tfrac{1}{2}\left(\frac{1}{df_1} + \frac{1}{df_2}\right). \tag{4}$$

Now
$$k = \frac{F_{0·05}(1)}{F_{0·95}(1)} = \frac{e_{0·05}^{2z}}{e_{0·95}^{2z}}.$$

For large df_1, df_2,

$$z_{0.95} \simeq -z_{0.05} \tag{5}$$

and hence

$$\log_e k = 4z_{0.05}$$

i.e.
$$z_{0.05} = \log_e k/4. \tag{6}$$

Putting

$$\frac{z_{0.05}}{\sigma_z} = c_{0.1} \text{ (Normal deviate)} = 1.644854. \tag{7}$$

We obtain from (6) and (7) that

$$\sigma_z^2 = \left(\frac{\log_e k}{4 \times 1.644854}\right)^2. \tag{8}$$

Now df_1, df_2 are functions of n and r only. Thus for a given r, n can in all cases be calculated to a close approximation.

For example, consider a N.C. Exp. II. We have (see section 2 (i))

$$df_1 = n^2(r-1)$$

$$df_2 = (n-1)^2$$

For large n we may approximate $(n-1)^2$ by n^2 so that

$$2\sigma_z^2 \simeq \frac{1}{n^2(r-1)} + \frac{1}{n^2} = \frac{r}{n^2(r-1)} \tag{9}$$

giving
$$n \simeq \sqrt{\frac{r}{2\sigma_z^2(r-1)}} \tag{10}$$

where σ_z^2 is given by equation (8).

This approximate solution which need not of course be an integer, will be slightly too small. Hence the integer immediately above this approximate n was substituted in

$$\frac{1}{n^2(r-1)} + \frac{1}{(n-1)^2} \tag{11}$$

and this value compared with σ_z^2 as given by (8). If expression (11) was greater than σ_z^2, n was incremented by units of 1 until such time as the expression was less than σ_z^2. This final value of n was the value actually used.

Obviously n is inversely related to r and thus with large values of r, df_2 may be sufficiently small for the validity of the approximation in equation (5) above to be in doubt. In such cases n was adjusted to satisfy the criteria of equations (1) or (2).

All the calculations were carried out by computer which enabled n to be calculated for a range of family sizes ($r = 5$ to 1000), and the minimal experimental sizes (*e.g.* n^2r for N.C. Exp. II) obtained numerically.

4. RESULTS

The minimum experimental sizes necessary with small family sizes ($r = 5$ and 10) are shown for the N.C. Exp. II in table 3. In view of the

generally large number of individuals necessitated by such small families, r was then incremented as described in section 3 to determine that combination of n and r that minimised the total experimental size.

TABLE 3

Minimum total experimental sizes for N.C. Exp. II design for small family sizes ($r = 5$, 10), f $= 0$

Narrow heritability	Family size	0·2	0·4	0·6	0·8	1·0
0·25	5	2,671,805	171,125	35,280	11,520	5,120
	10	1,204,090	79,210	16,810	6,250	2,890
0·50	5	496,125	32,805	6,845	2,420	1,125
	10	225,000	16,000	4,000	1,440	810
0·75	5	154,880	10,580	2,420	980	—
	10	72,250	5,760	1,440	640	—

Preliminary runs indicated that, for a given mating system, the value of n yielding the minimum total experimental size varied little with any of the different combinations of dominance, heritability, gene frequency or degree of inbreeding considered. What variation existed, did not appear to follow any simple pattern and was due to some extent to the approximations employed and also to the minimum size being in some cases associated with values of r in excess of 1000. It should be noted that since n and r are both integers the relation between them is not a smooth curve but proceeds in steps, hence the need for an iterative numerical procedure to obtain the minimum size.

The modal values of n obtained for each design were as follows

N.C. Exp. II	5
Diallel	8
N.C. Exp. III	12
A.B.I.P.'s	12

For these values of n, r was generally large and hence the χ^2 method (section 3 (*b*) equation (2)) was employed.

Since $k = 1 + \dfrac{rt\sigma_2^2}{\sigma_1^2}$ (equation (3) section 3 (*b*))

$$\frac{\chi_{0 \cdot 05}^2}{\chi_{0 \cdot 95}^2} = 1 + \frac{rt\sigma_2^2}{\sigma_1^2}$$

and

$$r = \left[\frac{\chi_{0 \cdot 05}^2}{\chi_{0 \cdot 95}^2} - 1\right] \frac{\sigma_2^2}{t\sigma_1^2}. \tag{12}$$

Using the values of n above to determine df_2 for each design, r was estimated using the tables of χ^2. The minimum experimental sizes so obtained are shown in table 4, 5, 6 for $f = 0 \cdot 0$, $0 \cdot 5$, $1 \cdot 0$ respectively. Again for small df_1, r has been adjusted to satisfy equation (1). (*N.B.* The value of $F_{0 \cdot 95}$ for df_1, df_2 degrees of freedom is equal to $1/F_{0 \cdot 05}$ for df_2, df_1 degrees of freedom).

5. Discussion

It is apparent from the figures in tables 3 to 6 that for moderate dominance values and heritabilities all experiments need to be carried out on a very large scale, particularly at extreme gene frequencies.

TABLE 4

Minimum experimental sizes for detecting dominance by the four designs when the population sample is not inbred (f = 0) (a, b and c represent frequencies of the dominant allele of 0·5, 0·1, 0·9 respectively)

Heritability	Design	Dominance ratio				
		0·2	0·4	0·6	0·8	1·0
a 0·25	N.C. Exp. II	40,250	10,025	4,425	2,475	1,575
	Diallel	42,432	10,560	4,672	2,624	1,664
	A.B.I.P.'s	73,776	18,336	8,064	4,464	2,832
	N.C. Exp. III	12,840	3,192	1,392	768	480
0·50	N.C. Exp. II	17,225	4,275	1,875	1,025	650
	Diallel	18,176	4,544	1,984	1,088	704
	A.B.I.P.'s	31,536	7,776	3,408	1,824	1,152
	N.C. Exp. III	4,920	1,200	528	288	168
0·75	N.C. Exp. II	9,550	2,350	1,025	550	—
	Diallel	10,048	2,496	1,088	576	—
	A.B.I.P.'s	17,472	4,272	1,824	960	—
	N.C. Exp. III	2,280	552	240	120	—
b 0·25	N.C. Exp. II	150,575	48,725	27,200	18,775	14,450
	Diallel	158,592	51,328	28,672	19,776	15,232
	A.B.I.P.'s	124,176	39,744	22,032	15,120	11,616
	N.C. Exp. III	16,992	5,448	3,024	2,088	1,608
0·50	N.C. Exp. II	64,500	20,850	11,625	8,025	6,172
	Diallel	67,968	21,952	12,288	8,448	6,528
	A.B.I.P.'s	51,936	16,320	8,976	6,096	4,656
	N.C. Exp. III	6,336	1,992	1,104	744	576
0·75	N.C. Exp. II	35,825	11,575	6,450	4,425	3,400
	Diallel	37,760	12,224	6,784	4,672	3,584
	A.B.I.P.'s	27,840	8,544	4,608	3,072	2,352
	N.C. Exp. III	2,784	840	456	312	240
c 0·25	N.C. Exp. II	78,925	12,900	3,325	850	—
	Diallel	83,136	13,568	3,520	896	—
	A.B.I.P.'s	68,304	11,808	3,408	1,152	—
	N.C. Exp. III	9,408	1,656	480	192	—
0·50	N.C. Exp. II	33,800	5,500	1,400	350	—
	Diallel	35,648	5,824	1,472	384	—
	A.B.I.P.'s	30,432	5,616	1,776	720	—
	N.C. Exp. III	3,816	720	264	120	—
0·75	N.C. Exp. II	18,750	3,025	750	—	—
	Diallel	19,776	3,200	832	—	—
	A.B.I.P.'s	17,808	3,504	1,248	—	—
	N.C. Exp. III	1,944	432	168	—	—

2 M 2

Table 5

Minimum experimental sizes for detecting dominance by the four designs when the population sampled is partly inbred (f = 0·5). The heritabilities refer to the base, non-inbred population (a, b and c represent frequencies of the dominant allele of 0·5, 0·1, 0·9 respectively)

		Dominance ratio				
Heritability	Design	0·2	0·4	0·6	0·8	1·0
a 0·25	N.C. Exp. II	16,575	4,125	1,800	1,000	625
	Diallel	17,472	4,352	1,920	1,088	640
	A.B.I.P.'s	39,072	9,696	4,224	2,304	1,440
	N.C. Exp. III	8,232	2,040	888	480	312
0·50	N.C. Exp. II	6,350	1,550	675	350	200
	Diallel	6,720	1,664	704	384	320
	A.B.I.P.'s	14,976	3,648	1,536	816	480
	N.C. Exp. III	2,952	720	312	168	120
0·75	N.C. Exp. II	2,950	700	275	200	—
	Diallel	3,136	768	320	320	—
	A.B.I.P.'s	6,912	1,632	624	288	—
	N.C. Exp. III	1,176	264	120	120	—
b 0·25	N.C. Exp. II	62,125	20,075	11,200	7,725	5,950
	Diallel	65,408	21,184	11,840	8,128	6,272
	A.B.I.P.'s	35,856	11,520	6,432	4,416	3,408
	N.C. Exp. III	10,992	3,528	1,968	1,368	1,056
0·50	N.C. Exp. II	23,875	7,700	4,275	2,950	2,250
	Diallel	25,152	8,128	4,544	3,136	2,368
	A.B.I.P.'s	13,584	4,320	2,400	1,632	1,248
	N.C. Exp. III	3,888	1,248	696	480	360
0·75	N.C. Exp. II	11,100	3,575	1,975	1,350	1,025
	Diallel	11,712	3,776	2,112	1,408	1,088
	A.B.I.P.'s	6,144	1,920	1,056	720	528
	N.C. Exp. III	1,512	480	264	168	144
c 0·25	N.C. Exp. II	32,550	5,300	1,350	325	—
	Diallel	34,304	5,568	1,408	384	—
	A.B.I.P.'s	19,296	3,264	912	288	—
	N.C. Exp. III	5,928	1,008	288	120	—
0·50	N.C. Exp. II	12,500	2,000	500	150	—
	Diallel	13,184	2,112	512	320	—
	A.B.I.P.'s	7,632	1,344	384	240	—
	N.C. Exp. III	2,208	384	120	120	—
0·75	N.C. Exp. II	5,800	925	200	—	—
	Diallel	6,144	960	320	—	—
	A.B.I.P.'s	3,696	672	240	—	—
	N.C. Exp. III	960	192	120	—	—

The N.C. Exp. III is invariably the most efficient design particularly at extreme gene frequencies and low *f* values. Only at very low dominance levels and heritabilities does it cease to be practical, and furthermore it is less sensitive to changes in gene frequency and inbreeding than are the other designs.

TABLE 6

Minimum experimental sizes for detecting dominance by the four designs when the population sampled is completely inbred (f = 1·0). The heritabilities refer to the base, non-inbred population (a, b and c represent frequencies of the dominant allele of 0·5, 0·1, 0·9 respectively)

		Dominance ratio				
Heritability	Design	0·2	0·4	0·6	0·8	1·0
a 0·25	N.C. Exp. II	8,600	2,125	925	500	300
	Diallel	9,088	2,240	960	512	320
	A.B.I.P.'s	23,616	5,808	2,496	1,344	816
	N.C. Exp. III	5,904	1,464	624	336	216
0·50	N.C. Exp. II	2,825	675	275	150	150
	Diallel	3,008	704	320	320	320
	A.B.I.P.'s	7,776	1,824	768	384	240
	N.C. Exp. III	1,944	456	192	120	120
0·75	N.C. Exp. II	925	200	150	150	—
	Diallel	960	320	320	320	—
	A.B.I.P.'s	2,496	528	144	120	—
	N.C. Exp. III	624	144	120	120	—
b 0·25	N.C. Exp. II	32,225	10,400	5,800	3,975	3,075
	Diallel	33,984	10,944	6,144	4,224	3,264
	A.B.I.P.'s	19,488	6,288	3,504	2,448	1,872
	N.C. Exp. III	7,992	2,592	1,440	1,008	768
0·50	N.C. Exp. II	10,725	3,450	1,900	1,300	1,000
	Diallel	11,328	3,648	2,048	1,408	1,088
	A.B.I.P.'s	6,480	2,112	1,152	816	624
	N.C. Exp. III	2,664	864	480	336	264
0·75	N.C. Exp. II	3,550	1,125	600	400	300
	Diallel	3,776	1,216	640	448	320
	A.B.I.P.'s	2,160	672	384	240	240
	N.C. Exp. III	888	288	168	120	120
c 0·25	N.C. Exp. II	16,875	2,725	675	150	—
	Diallel	17,792	2,880	704	320	—
	A.B.I.P.'s	10,224	1,680	432	240	—
	N.C. Exp. III	4,200	696	168	120	—
0·50	N.C. Exp. II	5,600	875	200	150	—
	Diallel	5,888	960	320	320	—
	A.B.I.P.'s	3,408	528	240	240	—
	N.C. Exp. III	1,392	216	120	120	—
0·75	N.C. Exp. II	1,825	275	200	—	—
	Diallel	1,920	320	320	—	—
	A.B.I.P.'s	1,104	240	240	—	—
	N.C. Exp. III	456	120	120	—	—

The diallel and N.C. Exp. II designs differ little as would be expected from their e.m.s.'s, and they are both very sensitive to changes in gene frequency and inbreeding.

Augmented B.I.P.'s require a much larger experimental size than the other designs at gene frequencies of 0·5 but generally require less individuals at extreme gene frequencies.

Inbreeding markedly decreases the total experimental sizes. For example, one generation of selfing (*i.e.* $f = 0.5$) results in the size of all designs except the N.C. Exp. III being reduced by more than a half.

These results can be simply extended to randomised plot designs, the narrow heritability now being that appropriate to plot means. Consider the progeny from a random mating population raised at random in r plots of size m. If the phenotypic variance of a population within plots is

$$V_P = \tfrac{1}{2}D_R + \tfrac{1}{4}H_R + E,$$

then the variance of plot means $(V\bar{x})$ is

$$V\bar{x} = \frac{V_P}{m} + E_2 = \frac{\tfrac{1}{2}D_R + \tfrac{1}{4}H_R + E_1 + mE_2}{m},$$

where E_1 is the environmental variance within plots and E_2 is the environmental variance of plot means.

The narrow heritability of plot means (*not* family means) thus becomes

$$\frac{\tfrac{1}{2}D_R}{m} \cdot \frac{1}{V\bar{x}} = \frac{\tfrac{1}{2}D_R}{\tfrac{1}{2}D_R + \tfrac{1}{4}H_R + E_1 + mE_2}.$$

The narrow heritability is unchanged, while E (in σ_1^2) becomes $E_1 + mE_2$. Thus MS_1, MS_2 based on plot means have the following expectations

$$MS_2 = \frac{\sigma_1^2}{m} + rt\sigma_2^2$$

$$MS_1 = \frac{\sigma_1^2}{m}$$

Thus
$$k = 1 + \frac{rtm\sigma_2^2}{\sigma_1^2} \quad \text{(cf. equation (3))}$$

and
$$rm = \left[\frac{\chi_{0.05}^2}{\chi_{0.95}^2} - 1 \right] \frac{\sigma_1^2}{t\sigma_2^2} \quad \text{(cf. equation (12))}.$$

Since the number of families for any design is fixed, the total experimental sizes remain as shown in tables 3 to 6.

The experimental sizes shown in tables 4, 5 and 6 all involve experiments with few families of large size, indeed in many cases of a size too large to be practical except perhaps for certain plant species. Because of this small family structure one might well heistate to extrapolate any conclusions drawn from them to the population at large. However, they do represent the *minimum* experimental sizes required to satisfy the criterion that dominance if present should be detected on 95 per cent. of occasions. It is possible to increase the number of families somewhat without excessively increasing the number of individuals but a comparison of table 4 for the N.C. Exp. II with those of table 3 indicates what happens in the limiting case of many families of small size.

Total experimental sizes of up to 5000 randomised individuals are not necessarily excessive, but the experimenter is frequently interested not just in a single population but in comparing the genetic architecture of several populations. With very few exceptions this is obviously impracticable.

Furthermore, we have considered only the simplifying case of many loci, with additive, dominance effects and gene frequencies constant at all loci and that these loci are in linkage equilibrium with only two alleles per locus. With the possible exception of the last one, all these assumptions are likely to be unrealistic and in general will underestimate the true size required.

Thus in many real experiments there is likely to be a very high chance of failure to detect non-additive variance when it is present, $i.e.$ the null hypothesis of no-dominance is accepted when it is false.

It follows, therefore, that since the sensitivity of these variance approaches to detecting dominance is often low, the estimates of dominance obtained from the expected mean squares will be of low precision also. It would appear then that these methods are not generally very useful for detecting and estimating dominance in randomly mating populations. As is shown in tables 4, 5 and 6, inbreeding prior to crossing improved their sensitivity markedly, but it is possible that even if such inbreeding is feasible, natural selection during the inbreeding process may so distort the genotypic frequencies as to make variance estimates so obtained irrelevant to the base population.

Two of the designs, the diallel and augmented B.I.P.'s, allow tests for directional dominance in the comparison of selfed $versus$ outcrossed progeny. Since all the data are used in obtaining the means of selfed and outcrossed progeny, such a comparison will be more sensitive than those previously described. The answer obtained from such a comparison, however, is a qualitative one in that it will indicate whether or not the genes display directional dominance and will not give a quantitative measure of the dominance contribution to the phenotypic variance.

If the tester lines used in the N.C. Exp. III are extreme selection lines from the population, then directional dominance may be detected by the covariance of sums on differences (Jinks, Perkins and Breese, 1968).

Since in all these designs MS_2 (or particularly σ_2^2) is a measure of dominance irrespective of the degree of disequilibrium, one can increase the sensitivity of the test by deliberately choosing extreme parents in the mating scheme. In that such a system will lead to matings either between like or unlike extreme individuals it will tend to mimic selfing and crossing with the same advantages and disadvantages as described above.

The difficulties involved in detecting and estimating dominance components suggest that it is too ambitious to attempt to investigate these components in terms of population parameters. Moreover, even if these difficulties did not exist, the utility of the parameters for predictive purposes or for providing some understanding of the genetic control of the characters is limited.

In view of these difficulties, it is likely to prove more useful to take advantage of the more sensitive approaches available with restricted non-random samples. For example, generations ($e.g.$ F_1's, F_2's backcrosses, etc.) derived from crosses between extreme groups or individuals in the population, or lines derived by selection.

6. Summary

1. The efficiency of four methods designed to detect the presence of non-additive variation for metrical traits in a randomly mating population are compared theoretically.

2. These four methods are the Experiments II and III of Comstock and Robinson (1952), the diallel cross, and an extension of the bi-parental mating scheme of Mather (1949) and termed A.B.I.P.'s.

3. On the assumption that the non-additive variation is due to dominance alone, the methods are compared on the basis of the minimum experimental sizes required to detect significant dominance (P = 0·05) with 95 per cent. certainty.

4. These sizes are estimated for populations with all combinations of the following properties: dominance ratios (0·2, 0·4, 0·6, 0·8, 1·0); narrow heritabilities (0·25, 0·50, 0·75); gene frequencies (0·1, 0·5, 0·9). Various degrees of inbreeding prior to crossing ($f = 0·0, 0·5, 1·0$) are also included.

5. The Experiment II and diallel are similar in efficiency and require a considerably smaller experimental size than does the A.B.I.P.'s at gene frequencies of a half. At more extreme gene frequencies, the latter design is generally superior. The Experiment III invariably requires fewer individuals than all the other designs, and this size varies least with changes in the genetic architecture.

6. Over most of the situations considered, the minimum experimental size of all designs is large, usually too large to make them practical for comparing different populations.

7. It is suggested that in view of their low efficiency in general, it might be better to use methods involving restricted non-random samples, and to take advantage of the greater precision of such methods even though the parameters detected and estimated are not population parameters.

7. REFERENCES

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

DICKINSON, A. G., AND JINKS, J. L. 1956. A generalised analysis of diallel crosses. *Genetics, 41*, 65-78.

GRIFFING, B. 1956. Concept of general and specific combining ability in relation to diallel crossing systems. *Australian J. Biol. Sci., 9*, 463-493.

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

JINKS, J. L., AND BROADHURST, P. L. 1965. The detection and estimation of heritable differences in behaviour among individuals. *Heredity, 20*, 97-115.

JINKS, J. L., AND PERKINS, J. M. 1969. The detection of linked epistatic genes for a metrical trait. *Heredity, 24*, 465-575.

JINKS, J. L., PERKINS, J. M., AND BREESE, E. L. 1968. A general method of detecting additive, dominance and epistatic variation for metrical traits. II. Application to inbred lines. *Heredity, 24*, 45-57.

JONES, R. MORLEY. 1965. Analysis of variance of the half diallel table. *Heredity, 20*, 117-121.

KEARSEY, M. J., AND JINKS, J. L. 1968. A general method of detecting additive, dominance and epistatic variation for metrical traits 1. Theory. *Heredity, 23*, 403-409.

KEARSEY, M. J., AND KOJIMA, K. 1967. The genetic architecture of body weight and egg hatchability in *Drosophila melanogaster*. *Genetics, 56*, 23-27.

MATHER, K. 1949. *Biometrical Genetics*. Methuen, London.

MATHER, K. 1960. Evolution in polygenic systems. *Int. Colloquium on Evolution and Genetics*, 131-152. Rome, Acad. Maz. dei Lincei.

PERKINS, J. 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.