Theoretical and Applied Genetics 42, 371—377 (1972)

A Comparison of Four Experimental Designs for the Estimation of Heritability

D. G. PEDERSON

Faculty of Agriculture, University of Sydney, Sydney, New South Wales (Australia)

Summary. The partial diallel cross, the complete diallel cross, and the designs known as North Carolina Experiments 1 and 2 are compared for their usefulness in estimating heritability. It is first shown that reliable values for the sampling mean and variance of heritability estimates are obtained from approximate expressions based on the moments of the chi-square distribution. These expressions are then applied to determine the optimum experimental designs for a range of situations.

The main basis for discrimination is the amount of information per unit, defined as

$$i = 1/(N \operatorname{var}(\hat{h}^2)),$$

where $\hat{h}^2$ is the estimate of the heritability h^2 and N is the number of units in the experiment, either individuals or families.

The two parameters considered were the heritability of individuals and the heritability of full-sib families, and for each of these the partial diallel cross was the most preferred, followed in decreasing order of preference by design $NC2$, the complete diallel, and design $NC1$.

It is first shown that there is no optimum number of parents for a partial diallel cross or male parents for designs $NC1$ and $NC2$. The number of crosses per parent for a partial diallel or dams per sire for designs $NC1$ and $NC2$ should generally be six or less. Any expansion should be in the direction of using more parents in the case of the partial diallel, or more male parents in the case of designs $NC1$ and $NC2$. For the two heritability parameters considered in this study it is inefficient to increase the number of replicates beyond two.

1. Introduction

The efficiency of heritability estimates from a diallel cross has been investigated by Pederson (1971) and recommendations were made for both the number of parents and the number of individuals per cross to be used in the experimental design. Three other mating designs are now considered, namely the partial diallel cross and the designs commonly referred to as North Carolina Experiment 1 (*NC* 1) and North Carolina Experiment 2 (*NC* 2) (Comstock and Robinson 1952).

A general evaluation of these three designs and of the half diallel cross has been presented by Kearsey (1965), using data for flowering time in a population of *Papaver dubium*. The usefulness of estimates obtainable from each design was discussed, along with the effects of failure of assumptions on which the genetic model was based. The half diallel cross was found to be the most useful design, followed by the two North Carolina designs. Marquez-Sanchez and Hallauer (1970 a, b) have studied the estimation of genetic components of variance for ten characters in corn using design *NC* 1. The variance of estimates was shown to decrease with increasing sample size, but only slowly as the number of females per male was increased much beyond four.

Since heritabilities and other functions of the components of variance are generally of greater interest than the components themselves the present paper deals with the efficient estimation of heritability, in particular the heritability applicable to individual selection and the heritability applicable to the selection of full-sibs based on family means. In the simulation method used previously (Pederson loc. cit.) the distribution of a heritability estimate was built up by the generation of random values from various chi-square distributions. To apply the same technique to the present three designs would have required a considerable amount of computer time, and an approximate but quicker method has therefore been used.

2. Definitions and Methods

The analysis of variance tables shown in Table 1 are from Kempthorne and Curnow (1961). For the partial diallel cross it is assumed that p parents are chosen at random from the population under study and that there are s crosses per parent. For $NC2$ there are m male parents and f female parents crossed in all possible combinations, and for $NC1$ there are m male parents each crossed to a different set of f females. It is assumed that there are r replicates with n individuals per full-sib family in each case.

In the expectations of mean squares the components σ_G^2 and σ_S^2 are the general combining ability (g.c.a.) and specific combining ability (s. c. a.) variances respectively, and if the contributing loci are assumed to be non-interacting then the interpretations in terms of additive genetic variance (σ_A^2) and dominance variance (σ_D^2) are

$$\sigma_G^2 = 1/4\,(1 + F)\,\sigma_A^2,$$

Table 1. *Analysis of variance tables for the partial diallel cross, North Carolina Experiment 1 and North Carolina Experiment 2 (df = degrees of freedom)*

Source	df	Expected mean square
		a) partial diallel cross
g. c. a.	$p-1$	$\sigma_E^2+n\,\sigma_P^2+n\,r\,\sigma_S^2+\frac{n\,r\,s\,(p-2)}{p-1}\,\sigma_G^2$
s. c. a.	$p(s/2-1)$	$\sigma_E^2+n\,\sigma_P^2+n\,r\,\sigma_S^2$
plots	$(r-1)$ $(p\,s/2-1)$	$\sigma_E^2+n\,\sigma_P^2$
residual	$p\,r\,s\,(n-1)/2$	σ_E^2
		b)North Carolina Experiment 1
sires	$m-1$	$\sigma_E^2+n\,\sigma_P^2+n\,r\,\sigma_S^2+n\,r(f+1)\,\sigma_G^2$
dams in sires	$m(f-1)$	$\sigma_E^2+n\,\sigma_P^2+n\,r\,\sigma_S^2+n\,r\,\sigma_G^2$
plots	$(r-1)\,(mf-1)$	$\sigma_E^2+n\,\sigma_P^2$
residual	$mfr\,(n-1)$	σ_E^2
		c) North Carolina Experiment 2
sires	$m-1$	$\sigma_E^2+n\,\sigma_P^2+n\,r\,\sigma_S^2+n\,r\,f\,\sigma_G^2$
dams	$f-1$	$\sigma_E^2+n\,\sigma_P^2+n\,r\,\sigma_S^2+n\,r\,m\,\sigma_G^2$
sires × dams	$(m-1)(f-1)$	$\sigma_E^2+n\,\sigma_P^2+n\,r\,\sigma_S^2$
plots	$(r-1)(mf-1)$	$\sigma_E^2+n\,\sigma_P^2$
residual	$mfr(n-1)$	σ_E^2

and

$$\sigma_S^2 = 1/4\,(1+F)^2\,\sigma_D^2\,,$$

where F is the degree of inbreeding of the sample relative to the population of interest. The variance component σ_P^2 is due to plot differences and the variance within families (σ_E^2) is considered to have genetic and environmental components

viz. $\sigma_E^2 = \sigma_W^2 + 1/2\,(1-F)\,\sigma_A^2 + 1/4\,(3+F)\,(1-F)\,\sigma_D^2$.

True genotype × environment interaction is assumed to be absent.

For $NC2$ it is necessary to pool the sires and dams mean squares:

pooled mean square =

$$\frac{(m-1)\text{ sires mean square} + (f-1)\text{ dams mean sq.}}{m+f-2}.$$

The expected value of the pooled mean square is then

$$\sigma_E^2 + n\,\sigma_P^2 + n\,r\,\sigma_S^2 + \frac{n\,r\,[f\,(m-1)+m\,(f-1)]}{m+f-2}\,\sigma_G^2\,.$$

A heritability estimate from a particular analysis is a function of the observed mean squares and is always of the form

$$\hat{h}^2 = x/y\,,$$

where x and y are linear functions of components of variance. For example, we may estimate the heritability

$$\hat{h}^2_{individual} = \frac{\hat{\sigma}_A^2}{\hat{\sigma}_A^2+\hat{\sigma}_D^2+\hat{\sigma}_W^2} \qquad (1)$$

This is the heritability appropriate to a scheme in which mass selection is carried out within many populations of size n, such that each population occupies a single plot, and the progeny of selected parents are distributed at random among plots.

Let the expected values of x and y, which are both random variables, be μ_x and μ_y respectively. Then we can write

$$\hat{h}^2 = \frac{\mu_x+e_x}{\mu_y+e_y}\,,$$

where e_x and e_y are random deviations from μ_x and μ_y. Thus

$$\hat{h}^2 = \frac{\mu_x}{\mu_y}\left(1+\frac{e_x}{\mu_x}\right)\left(1+\frac{e_y}{\mu_y}\right)^{-1}.$$

On expanding the final term as a power series, convergent for $e_y/\mu_y < 1$, we obtain

$$\hat{h}^2 = \frac{\mu_x}{\mu_y}\left(1+\frac{e_x}{\mu_x}\right)\left(1-\frac{e_y}{\mu_y}+\frac{e_y^2}{\mu_y^2}-\frac{e_y^3}{\mu^3}+\dots\right).$$

It follows that

$$E[\hat{h}^2] = \frac{\mu_x}{\mu_y}\cdot E\left[1+\frac{e_x}{\mu_x}-\frac{e_y}{\mu_y}-\frac{e_x\,e_y}{\mu_x\,\mu_y}+\dots\right]$$
$$= \frac{\mu_x}{\mu_y}\left\{1-\frac{E[e_x\,e_y]}{\mu_x\,\mu_y}+\frac{E[e_y^2]}{\mu_y^2}+\dots\right\}. \qquad (2)$$

Similarly we can write

$$var\,(\hat{h}^2) = E\,[(\hat{h}^2 - E[\hat{h}^2])^2]$$

in the form

$$\left(\frac{\mu_x}{\mu_y}\right)^2\left\{\frac{E[e_x^2]}{\mu_x^2}-\frac{2\,E[e_x\,e_y]}{\mu_x\,\mu_y}+\frac{E[e_y^2]}{\mu_y^2}+\frac{4\,E[e_x\,e_y^2]}{\mu_x\,\mu_y^2}-\frac{2\,E[e_x^2\,e_y]}{\mu_x^2\,\mu_y}+\dots\right\}. \qquad (3)$$

Now consider a sample from a normally distributed population, giving rise to a mean square s^2 with expectation σ^2 and degrees of freedom m. In successive samples s^2 is distributed as $\sigma^2\,\chi_m^2/m$ and we can therefore determine the moments of the distribution of s^2 from the moments of the chi-square distribution. The expectations in (2) and (3) may therefore be expressed in terms of the population variances and design parameters, since both x and y are linear functions of mean squares.

The question arises as to how many terms should be included in (2) and (3) in order to give reliable results. The simulation method described by Pederson (loc. cit.) was therefore used to determine the distribution of $\hat{h}^2_{individual}$, as determined from a Method 4 diallel analysis (Griffing 1956), for the five populations given in Table 2. The dominance variance and degree of inbreeding were assumed to be zero and the number of individuals per cross was set at two, with two replicates in the experimental design. The mean and variance of the heritability estimates from this simulation procedure were compared with the values given by expressions (2) and (3).

There were problems with convergence of both (2) and (3) when fewer than 10 parents were included in the sample, but satisfactory results for 10 or more parents were obtained by including only the variance and covariance terms.

$$i.e.\ E[\hat{h}^2] = \frac{\mu_x}{\mu_y}\left\{1-\frac{E[e_x\,e_y]}{\mu_x\,\mu_y}+\frac{E[e_y^2]}{\mu_y^2}\right\}, \qquad (4)$$

and

$$var\,(\hat{h}^2) = \left(\frac{\mu_x}{\mu_y}\right)^2\left\{\frac{E[e_x^2]}{\mu_x^2}-\frac{2\,E[e_x\,e_y]}{\mu_x\,\mu_y}+\frac{E[e_y^2]}{\mu_y^2}\right\}. \qquad (5)$$

Table 2. *Parameters of five populations for which an approximate and an exact simulation method were compared*

Population	σ_A^2	σ_W^2	$h^2_{individual}$
1	1	9	0.1
2	3	7	0.3
3	1	1	0.5
4	7	3	0.7
5	9	1	0.9

Table 3. *True values (upper figures) from a simulation procedure and approximate values (lower figures) from expressions given in the text for the expectation (E) and variance (V) of individual heritability estimates from a diallel cross. (Values are × 10^4.) The population parameters are specified in Table 2*

Population	Parameter	Number of parents					
		5	6	7	8	9	10
1	E	798	926	937	966	976	978
		793	918	957	973	982	986
	V	2078	700	533	203	142	107
		1215	556	318	206	146	109
2	E	2576	2779	2855	2868	2893	2904
		2610	2775	2842	2877	2898	2913
	V	4537	1499	575	425	333	273
		1674	941	629	463	362	295
3	E	4414	4561	4695	4720	4764	4788
		4336	4553	4657	4718	4760	4790
	V	2047	1691	819	641	525	446
		2059	1292	929	720	585	491
4	E	6157	6370	6506	6549	6614	6653
		6027	6296	6441	6533	6597	6645
	V	2198	1366	1009	811	679	585
		2270	1521	1138	906	751	641
5	E	7830	8155	8317	8380	8465	8519
		7738	8052	8233	8353	8439	8504
	V	5364	1471	1211	912	771	669
		2256	1576	1209	979	822	707

The latter expression is commonly given, for example by Kempthorne (1957). From Table 3, which shows the "true" values from the simulation procedure and the values obtained from (4) and (5), it is seen that the reliability of each expression improves as the sample size increases.

Expressions (4) and (5) were applied to a range of situations for the three designs under study. In fact the variance of an estimated heritability is not a good indicator of the worth of a design since the variance would be expected to decrease as the sample size increases. The amount of information per individual, defined as

$$i = 1/(N\, var\, (\hat{h}^2))\,,$$

was therefore calculated in each case, N being the total number of individuals scored. The efficiency of a design is directly related to the value of i.

The most efficient method of estimating individual heritability was first considered. In practice there are numerous systems of selection and to give the present results some degree of generality the additional system has therefore been considered in which there are r groups of unreplicated full-sib families, corresponding to the r replicates of the design used for parameter estimation, and selection on the basis of full-sib performance is carried out independently within each group. The heritability applicable to this form of selection is (Falconer 1960)

$$h^2_{family} = \frac{1/2\, \sigma^2_A}{\sigma^2_E/n + \sigma^2_P + cov\, (FS)}\,, \tag{6}$$

where $cov\ (FS)$ is the covariance of full-sibs.

An estimate of h^2_{family} is most conveniently obtained from an analysis of family means, in which case the expectations of mean squares given in Table 1 are divided through by n and the residual mean square is not obtained. There is the assumption that n is the same in the estimation and selection phases. For an analysis based on family means there are complications in applying the estimated variance components to the selection phase unless the degree of inbreeding is zero throughout. Only this case was therefore considered for h^2_{family}.

Further, since family means may often be obtained by treating the family as a unit, the amount of labour involved is effectively independent of n. The parameter i was therefore calculated as

$$i = 1/(N\, var\, (\hat{h}^2_{family}))\,,$$

where N is the total number of family-units scored, and n was treated as a fixed quantity rather than as a variable for which some optimum value is to be determined.

3. Results and Discussion

a) *Individual Heritability*

Values of i were initially determined for the four populations with parameters as shown in Table 4.

Table 4. *Parameters of four populations for which the four mating systems were compared*

Population	σ^2_A	σ^2_D	σ^2_W	$h^2_{individual}$	dd
1	1	0	4	0.2	0
2	2	1	7	0.2	1
3	3	0	2	0.6	0
4	6	3	1	0.6	1

The parameter "*dd*" is the degree of dominance, defined as $(2\,\sigma^2_D/\sigma^2_A)^{1/2}$ (Comstock and Robinson, 1948). Both the number of replicates and the number of individuals per cross were set at two since these values were found to maximise i for $h^2_{individual}$ estimated from a complete diallel cross (Pederson loc. cit.). For inbreeding levels (F) of 0 and 1 the maximum values of i for various numbers of parents in the partial diallel and for various numbers of sires in designs *NC* 1 and *NC* 2 are given in Table 5.

Completely inbred parents are seen to give two or three times as much information per individual as non-inbred parents, and this factor is most evident for the higher heritability in the absence of dominance. However, a more significant result is that i passes through no maximum value as the number of parents or sires is increased but tends towards a limiting value for each population and level of inbreeding. For each design there is no apparent optimum number of parents or sires for the estimation of individual heritability. This result is in contrast to the situation for the complete diallel cross, for which i was always found to pass through a maximum value as the number of parents sampled was increased (Pederson loc.

Table 5. *Estimation of individual heritability; maximum values of amount of information per individual (× 10³) for varying numbers of a) parents in a partial diallel design b) sires in design NC1 and c) sires in design NC2. Two replicates and two individuals per cross are grown; the population parameters are specified in Table 4*

Population	F	Number of parents or sires					
		10	30	50	70	90	100
		a) partial diallel					
1	0	287	327	335	339	340	341
	1	676	754	770	776	779	782
2	0	267	305	313	316	318	319
	1	528	596	609	615	618	620
3	0	131	145	147	149	149	150
	1	700	750	760	764	766	767
4	0	100	125	128	129	129	130
	1	350	380	386	389	390	391
		b) design *NC*1					
1	0	130	139	141	141	142	142
	1	247	263	266	267	268	268
2	0	122	131	132	133	133	134
	1	202	216	218	220	220	220
3	0	40	42	43	43	43	43
	1	82	86	87	88	88	88
4	0	35	37	38	38	38	38
	1	55	58	58	59	59	59
		c) design *NC*2					
1	0	307	309	309	309	309	309
	1	669	685	687	688	688	689
2	0	289	291	291	291	292	292
	1	543	549	549	549	549	550
3	0	130	133	134	134	134	135
	1	643	654	656	657	657	658
4	0	113	114	114	114	114	114
	1	349	356	357	357	357	357

cit.). For the populations specified in Table 4 these maximum values were:

Population	F	maximum i
1	0	.283
	1	.606
2	0	.267
	1	.496
3	0	.117
	1	.568
4	0	.100
	1	.319

If the values in the last column of Table 5 are taken as the limits it is concluded that a partial diallel design potentially gives the greatest amount of information per individual, followed in decreasing order by design *NC* 2, the complete diallel, and design *NC* 1. The differences between the first three are not marked if relatively few parents are sampled.

The number of parents to be sampled for a partial diallel cross or the number of sires for designs *NC* 1 and *NC* 2 is therefore largely a matter of individual choice, although large numbers are preferred. But once this number has been nominated there will be an optimum number of crosses per parent in the case of the partial diallel and an optimum number of dams per sire in the case of designs *NC* 1 and *NC* 2. Table 6 shows these optimum numbers for a range of situations.

The dominant factor is the heritability, with fewer crosses per parent or dams per sire being required when the heritability is high. The optimum number tends to decrease as the level of inbreeding of the parents is increased from 0 to 1 but there is very little dependence on the number of parents or sires, or on the degree of dominance of the character.

There is the further point that a design which maximises i will not be particularly useful if the estimator of heritability so obtained is biassed. The expression (4) was therefore used to determine the expectation of $\hat{h}^2_{individual}$ for each of the designs specified in Table 6.

Table 6. *Estimation of individual heritability; optimum numbers of a) crosses per parent for a partial diallel design, and b) and c) dams per sire for designs NC1 and NC2 (* = greater than 100)*

Number of parents or sires	F	dd	Heritability				
			0.1	0.3	0.5	0.7	0.9
			a) partial diallel				
10	0	0	13	5	5	3	3
		1	13	7	5	—	—
	1	0	7	3	3	3	3
		1	9	5	5	—	—
50	0	0	13	5	5	3	3
		1	13	5	5	—	—
	1	0	7	3	3	3	3
		1	7	5	5	—	—
			b) design *NC*1				
5	0	0	12	5	4	3	3
		1	12	5	4	—	—
	1	0	7	3	3	3	3
		1	7	4	4	—	—
10	0	0	12	5	4	3	3
		1	12	5	4	—	—
	1	0	7	4	3	3	3
		1	7	4	4	—	—
50	0	0	12	5	4	3	3
		1	12	6	4	—	—
	1	0	7	4	3	3	3
		1	8	4	4	—	—
			c) design *NC*2				
5	0	0	*	8	4	3	2
		1	*	11	5	—	—
	1	0	38	3	2	2	2
		1	*	5	4	—	—
10	0	0	21	4	3	2	2
		1	23	5	3	—	—
	1	0	7	3	2	2	2
		1	8	4	3	—	—
50	0	0	8	4	3	2	2
		1	8	4	3	—	—
	1	0	5	2	2	2	2
		1	5	3	3	—	—

The results (Table 7) indicate that the general negative bias is greater when the heritability is high than when it is low, but that none of the designs gives a consistently smaller bias than the others. It would appear that five sires are too few for designs *NC* 1 and *NC* 2, and that even ten may be too few if a possible negative bias of about five per cent is unacceptable.

Table 7. *Estimation of individual heritability; expected values of the heritability estimate (× 10^3) for designs which optimise i (as shown in Table 4)*

Number of parents or sires	F	dd	Heritability (× 10^3)				
			100	300	500	700	900
			a) partial diallel				
10	0	0	99	291	478	662	849
		1	99	290	473	—	—
	1	0	98	285	472	667	882
		1	97	279	461	—	—
50	0	0	100	299	497	694	892
		1	100	298	496	—	—
	1	0	100	298	495	694	897
		1	100	297	493	—	—
			b) design *NC* 1				
5	0	0	98	284	457	613	769
		1	98	282	450	—	—
	1	0	97	267	423	575	735
		1	96	262	415	—	—
10	0	0	99	294	483	664	845
		1	99	293	480	—	—
	1	0	99	288	467	646	829
		1	98	284	464	—	—
50	0	0	100	299	497	694	890
		1	100	299	497	—	—
	1	0	100	298	494	690	887
		1	100	297	494	—	—
			c) design *NC* 2				
5	0	0	*	293	472	643	799
		1	*	291	471	—	—
	1	0	100	277	445	639	867
		1	*	278	450	—	—
10	0	0	100	293	482	663	851
		1	100	293	478	—	—
	1	0	99	288	473	670	884
		1	99	286	467	—	—
50	0	0	100	299	496	693	891
		1	100	298	495	—	—
	1	0	100	297	495	694	897
		1	100	296	493	—	—

Finally, it will be remembered that in each of the designs specified in Table 6 there are two individuals per family, this being the number which will always maximise *i*. But a family size of two may be impractical, particularly for plant material for which border plants are grown but ignored in analyses. Table 8 has therefore been prepared, showing the effect of an increase in family size on the optimum number of crosses per parent for the partial diallel, or dams per sire for designs *NC* 1 and *NC* 2. The general trend is for fewer crosses or dams to be required as more individuals are grown per family, and there is an accompanying decrease in the extent to which the optimum number is a function of heritability.

Table 8. *Estimation of individual heritability; effect of increasing the number of individuals per cross on the optimum number of a) crosses per parent for a partial diallel design, and b), c) dams per sire for designs NC 1 and NC 2 (* ≐ greater than 100)*

Number of parents or sires	F	dd	Heritability					
			0.2			0.6		
			n			n		
			2	5	20	2	5	20
			a) partial diallel					
10	0	0	7	5	3	3	3	3
		1	7	5	3	5	3	3
	1	0	5	3	3	3	3	3
		1	5	5	3	5	5	5
50	0	0	7	5	3	3	3	3
		1	7	5	3	5	3	3
	1	0	5	3	3	3	3	3
		1	5	5	3	5	5	5
			b) design *NC* 1					
5	0	0	7	4	3	3	3	2
		1	7	4	3	4	3	3
	1	0	4	3	2	3	3	2
		1	5	4	3	4	4	4
10	0	0	7	4	3	4	3	2
		1	7	4	3	4	3	3
	1	0	4	3	2	3	3	2
		1	5	4	3	4	4	4
50	0	0	7	4	3	4	3	2
		1	7	4	3	4	3	3
	1	0	4	3	2	3	3	3
		1	5	4	3	4	4	4
			c) design *NC* 2					
5	0	0	38	4	2	3	2	2
		1	*	5	3	4	3	2
	1	0	5	2	2	2	2	2
		1	9	4	3	5	4	3
10	0	0	7	3	2	3	2	2
		1	7	4	2	3	2	2
	1	0	3	2	2	2	2	2
		1	4	3	3	3	3	3
50	0	0	5	3	2	2	2	2
		1	5	3	2	3	2	2
	1	0	3	2	2	2	2	2
		1	4	3	2	3	2	2

b) Family Heritability

Values of *i* were calculated for the same parameter values as were used to obtain Table 5, except that F was set at zero and n was set at both 2 and 20. From the results (Table 9) it is seen that *i* increases as the number of parents or sires is increased and there is again no tendency for a maximum value to be attained along a row. The corresponding maximum values of i for the complete diallel are:

Population	n	Maximum i
1	2	.814
	20	1.232
2	2	.794
	20	.706
3	2	.641
	20	8.218
4	2	.559
	20	1.180

A comparison with the final column of Table 9 indicates that the partial diallel design gives the most efficient estimator of h^2_{family}, followed by *NC* 2, the complete diallel, and *NC* 1.

It is concluded that there is no optimum number of parents or sires, although the chosen number should not be less than about 10, but once the choice has been made then there will be an optimum number of crosses per parent or dams per sire. Typical values are shown in Table 10 and these may be compared with the values for the estimation of $h^2_{individual}$ which have been given in Table 8.

Table 9. *Estimation of full sib family heritability; maximum values of amount of information per full-sib family* ($\times$ *10*^{3}) *for varying numbers of a) parents in a partial diallel design, and b), c) sires in designs NC 1 and NC 2. There are 2 replicates, with either 2 or 20 individuals per cross; the population parameters are specified in Table 4*

Population	Number of parents or sires n	10	30	50	70	90	110
		a) partial diallel					
1	2	813	926	948	957	962	966
	20	1404	1509	1529	1538	1543	1545
2	2	788	894	914	923	928	931
	20	779	846	859	864	867	869
3	2	716	785	798	804	807	809
	20	9614	10161	10266	10311	10336	10352
4	2	618	681	694	699	702	704
	20	1280	1370	1388	1395	1399	1402
		b) design *NC* 1					
1	2	338	361	365	367	368	369
	20	126	133	134	135	135	135
2	2	331	353	357	359	360	361
	20	119	126	127	128	128	129
3	2	158	167	169	170	171	171
	20	117	123	124	125	125	125
4	2	149	158	160	161	161	162
	20	104	110	111	111	111	112
		c) design *NC* 2					
1	2	878	884	886	886	887	887
	20	1389	1437	1446	1450	1452	1453
2	2	854	859	862	863	863	864
	20	778	784	785	785	785	785
3	2	713	720	720	720	720	720
	20	9256	9459	9497	9513	9522	9528
4	2	613	628	630	631	632	632
	20	1283	1288	1288	1293	1296	1298

Table 10. *Estimation of full-sib family heritability; optimum number of a) crosses per parent for a partial diallel design, and b), c) dams per sire for designs NC 1 and NC 2* (* = *greater than 100*)

Number of parents or sires		$h^2_{individual}$ 0.2 n			0.6 n		
	dd	2	5	20	2	5	20
		a) partial diallel					
10	0	7	5	3	5	3	3
	1	9	5	5	5	5	5
50	0	7	5	3	5	3	3
	1	9	5	5	5	5	5
		b) design *NC* 1					
5	0	7	4	3	4	4	3
	1	7	5	4	4	4	3
10	0	7	4	3	4	3	3
	1	7	5	4	4	4	3
50	0	7	4	3	4	3	3
	1	7	5	4	4	4	4
		c) design *NC* 2					
5	0	*	6	3	4	3	2
	1	*	8	4	5	4	4
10	0	7	4	2	3	2	2
	1	8	4	3	3	3	3
50	0	5	3	2	3	2	2
	1	5	3	3	3	3	2

There is a definite similarity of pattern, and the slight tendency for the values of Table 10 to be greater is not important since the efficiency of estimation of both heritabilities is affected very little by small deviations from the optimum with respect to the number of crosses per parent or dams per sire.

The conclusions regarding estimation of $h^2_{individual}$ from a complete diallel cross have been found to be very little affected by a non-zero value of σ^2_P (Pederson loc. cit.), which is probably to be expected since $h^2_{individual}$ is defined to be independent of σ^2_P. However, the latter is not the case for h^2_{family}. Smith (1938) has shown that a relationship of the form

$$\frac{\sigma^2_W}{n} + \sigma^2_P = \frac{\sigma^2_W}{n^b},$$

where the coefficient b is usually between 0.2 and 0.8, has general validity for genetically uniform plant material and this expression may therefore be used to determine realistic values for σ^2_P. If b is set at 0.5 and n at 2 then $\sigma^2_P = 0.83$ for population 1 of Table 4, while for $n = 20$ the value is $\sigma^2_P = 0.69$. The procedure which gave rise to Table 9 was therefore repeated but with $\sigma^2_P = 1$, and also with $\sigma^2_P = 4$ as an extreme value. Only population 1 was considered and the number of replicates was varied between two and five.

The results of this further study will only be stated qualitatively since previous conclusions are very little affected. First, it was found in all cases that an increase in the number of replicates beyond two resulted in a loss of the efficiency of estimation of h^2_{family}. Secondly, the partial diallel cross was always the most preferred method of estimation, followed by design $NC2$, the complete diallel, and design $NC1$. Finally, for efficient estimation of h^2_{family} from a partial diallel cross the required number of crosses per parent was found to increase with σ^2_P, and a similar result was obtained for designs $NC1$ and $NC2$. For example, from Table 10 the average optimum number of crosses per parent or dams per sire is approximately three when $h^2_{individual} = 0.2$, $n = 20$, and $\sigma^2_P = 0$, and this number increases to approximately six when $\sigma^2_P = 1$ and twelve when $\sigma^2_P = 4$.

IV. Conclusions

The choice of an experimental design for the estimation of heritability may be based on several factors, and in this paper the emphasis has been on the amount of information per unit of measurement as a discriminatory measure. The partial diallel design was found to be the most preferred, followed in decreasing order of preference by design $NC2$, the complete diallel, and design $NC1$. This was the case for each of the two heritabilities considered and is probably a general result when a heritability is the parameter under consideration.

The first three designs are closely related in that each male parent is crossed to more than one female and each female is crossed to more than one male. However, for design $NC1$ each male is crossed to a different set of females and this results in an estimate of additive genetic variance with approximately twice the sampling variance of the corresponding estimate from the other three designs. In addition, the estimation of dominance variance using design $NC1$ is a relatively inefficient procedure since a subtraction of the estimated additive genetic variance is involved.

There may be other factors involved in the choice of a mating design. As regards practicability, with animal material there is very little choice but to use a design such as $NC1$ since the partial diallel, design $NC2$, and the complete diallel are only suitable for multi-flowered plants. Kearsey (1965) has stated that if epistatic effects are present then an inflated estimate of dominance variance will result from designs $NC1$ and $NC2$, while a test for the presence of non-allelic effects is possible for the data from a complete diallel. He further considers that the partial diallel is no more informative than $NC1$ or $NC2$ and involves both an awkward crossing programme and a lengthy analysis.

When all of these factors are taken into account it appears that the complete diallel is the best of the four designs considered for the estimation of heritability. The "multiple diallel" outlined by Pederson (1971) is probably the most efficient procedure.

On the question of the number of families to be grown, the indication are that with a family size (n) of two or three there should be approximately six crosses per parent for the partial diallel, or six females per male for designs $NC1$ and $NC2$. The optimum number is generally fewer than six when the value of $h^2_{individual}$ is 0.5 or greater, but is only slightly changed by variation in either the degree of dominance or the level of inbreeding of the parents. If the family size is increased to twenty then the optimum number of crosses per parent or dams per sire decreases to about three, although in the case of h^2_{family} this number would be greater than three if the between plot component of variance were appreciably greater than zero. It follows that any expansion of a partial diallel design should be in the direction of using more parents while maintaining the number of crosses per parent constant, and any expansion of designs $NC1$ and $NC2$ should involve the use of more male parents. For the heritability parameters considered in the present paper the number of replicates should be maintained at two.

Literature

1. Comstock, R. E., Robinson, H. F.: The components of genetic variance in populations of biparental progenies and their use in estimating the average degree of dominance. Biometrics **4**, 254–266 (1948). — 2. Comstock, R. E., Robinson, H. F.: Estimation of average dominance of genes. In: Heterosis pp. 494–516. Ames: Iowa State College 1952. — 3. Falconer, D. S.: Introduction to Quantitative Genetics. New York: Ronald 1960. — 4. Griffing, B.: Concept of general and specific combining ability in relation to diallel crossing systems. Aust. J. biol. Sci. **9**, 463–493 (1956). — 5. Kearsey, M. J.: Biometrical analysis of a random mating population: a comparison of five experimental designs. Heredity **20**, 205–235 (1965). — 6. Kempthorne, O.: An Introduction to Genetic Statistics. New York: Wiley 1957. — 7. Kempthorne, O., Curnow, R. N.: The partial diallel cross. Biometrics **17**, 229–250 (1961). — 8. Marquez-Sanchez, F., Hallauer, A. R.: Influence of sample size on the estimation of genetic variances in a synthetic variety of maize. I. Grain yield. Crop Sci. **10**, 357–361 (1970a). — 9. Marquez-Sanchez, F., Hallauer, A. R.: Influence of sample size on the estimation of genetic variances in a synthetic variety of maize II. Plant and ear characters. Iowa State Jour. Sci. **44**, 423–436 (1970b). — 10. Pederson, D. G.: The estimation of heritability and degree of dominance from a diallel cross. Heredity **27**, 247–264 (1971). — 11. Smith, H. Fairfield: An empirical law describing heterogeneity in the yields of agricultural crops. J. Agric. Sci. **28**, 1–23 (1938).

Received March 30, 1972
Communicated by J. S. F. Barker

Dr. D. G. Pederson
Faculty of Agriculture
University of Sydney
Sydney, New South Wales (Australia) 2006