Theor Appl Genet (1992) 85: 21–25

THEORETICAL AND APPLIED GENETICS

© Springer-Verlag 1992

Selecting the size of a diallel cross experiment *

Perry Y. Jui and L. P. Lefkovitch

Statistical Research Section, Research Program Service, Research Branch, Agriculture Canada, Ottawa, K1A OC6, Canada

Received March 5, 1992; Accepted March 24, 1992
Communicated by J. Mac Key

Summary. The size of a diallel cross experiment can be determined by using statistical tables and past diallel cross results; this will provide more flexibility in planning experiments and increase the opportunity to detect real differences.

Key words: Diallel cross – Experimental size – Griffing's models

Introduction

Most statistics texts give a method for determining the number of replications sufficient to detect differences among several populations (Snedecor and Cochran 1980; Steel and Torrie 1980). However, determination of the size of an experiment (i.e., how the number of replications and treatments are related) is not emphasized. In genetic studies in agriculture, the size of the experiment is often the most critical cost factor in choosing a design to achieve the objectives. Usually, the size is determined solely by the genetic material available and by management restrictions, and seldom by statistical considerations. The aim of this investigation is to provide some guidelines for determining the size of a diallel cross experiment based on statistical considerations in addition to any others factors.

Materials and methods

Under the assumption of a simple randomized complete block design, the partitioning of the degrees of freedom in four different methods for diallel cross experiments (Griffing 1956) are given in Table 1. These methods are distinguished by whether they include or exclude parental lines, F_1 or reciprocal crosses. Method 1, which includes the parental lines, is the so-called full diallel, while methods 2, 3 and 4 are sometimes called modified diallel crosses. In Table 1, p denotes the number of parental lines, r the number of replications, g.c.a. the general combining ability, and s.c.a. the specific combining ability; the reciprocal effect is also identified.

In order to provide an index by which the size of an experiment can be chosen, and since the variance components are usually unknown, the ratio of sums of squares (S.S.) is used as an indicator. Assuming a randomized complete-block design, the minimum number of replications needed to achieve significance was calculated for different sizes of diallel crosses and different S.S. ratios. Significance was tested by converting the S.S. to mean squares, and the appropriate ratios referred to the F-distribution. The block, s.c.a., and reciprocal effects were tested by comparison with the error mean square, and the g.c.a. against that for the s.c.a. A 5% or lower type-I error is considered to be significant. The minimal number of replications for a specified number of parental lines required for detecting a significant effect were determined.

In constructing the tables, the procedure adopted can be illustrated as follows. Assume the S.S. ratio between s.c.a. and error is 0.75; divide the sum of squares by its corresponding degrees of freedom; in an 8×8 diallel cross using Method II, the degrees of freedom for three replications are 28 and 70 for s.c.a. and error respectively. Since the variance ratio is 1.875, which exceeds the tabulated F value of 1.65, it follows that three replications should be sufficient to detect differences in the s.c.a. effect. For two replications and the same S.S. ratio, the degrees of freedom are 28 and 35 for s.c.a. and error respectively; the calculated variance ratio of 0.938 is smaller than the tabulated F value of 1.84, so that two replications should not be expected to detect a real differences in the s.c.a. effect.

Results and discussion

Table 2–5 present the expected number of replications required for various diallel crossing methods, tabulated by the number of parental lines and S.S. ratios. The S.S.

* Contribution No. R-015, Research Program Service, Research Branch, Agriculture Canada, Ottawa, Ont. K1A OC6
Correspondence to: P. Y. Jui

Table 1. Partition of degrees of freedom according to Griffing's model II; p is the number of parental lines, r the number of replications

Source of variation	Degrees of freedom			
	Method I P+F1+REC	Method II P+F1	Method III F1+REC	Method IV F1 Only
Block	r−1	r−1	r−1	r−1
Crosses	(p*p)−1	p(p+1)/2−1	p(p−1)−1	p(p−1)/2−1
g.c.a.	p−1	p−1	p−1	p−1
s.c.a.	p(p−1)/2	p(p−1)/2	p(p−3)/2	p(p−3)/2
Reciprocal	p(p−1)/2	–	p(p−1)/2	–
Error	(p*p−1) *(r−1)	[p(p+1)/2−1] *(r−1)	[p(p−1)−1] *(r−1)	[p(p−1)/2−1] *(r−1)

–, not applicable

Table 2. Number of replications needed for significance at 5% level for crosses effect in Griffing's methods I–IV; S.S., sum of squares

Method	S.S. ratio	Number of parental lines												
		4	5	6	7	8	9	10	11	12	13	14	15	16
I	2.4	2												
	2.0	3	2											
	1.75	3	2	2										
	1.5	3	3	3	3	3	3	2						
	1.20	3	3	3	3	3	3	3	3	3	2			
II	3.2	2												
	2.5	3	2											
	2.0	3	3	3	2									
	1.75	3	3	3	3	3	2							
	1.5	3	3	3	3	3	3	3	3	3	2			
	1.25	3												
	1.0	4	4	3										
	0.5	6	5	5	5	5	4							
III	2.2	2												
	1.9	3	3	2										
	1.75	3	3	3	2									
	1.5	3	3	3	3	3	3	2						
	1.3	3	3	3	3	3	3	3	3	3	3	2		
	1.0	4	3	3	3	3	3	3	3	2				
IV	5.0	2												
	3.2	3	2											
	2.5	3	3	2										
	2.0	3	3	3	3	2								
	1.75	3	3	3	3	3	3	2						
	1.5	4	3	3	3	3	3	3	3	3	2			
	1.4	4	3	3	3	3	3	3	3	3	3	3	2	

Blank means the same as the nearest left non-blank number
The ranges of S.S. ratio reflect the nature of the methods

Table 3. Number of replications needed for significance at 5% level for GCA in Griffing's methods I–IV; S.S., sum of squares.

Method	S.S. ratio	Number of parental lines												
		4	5	6	7	8	9	10	11	12	13	14	15	16
I	2.9	2												
	2.0	*	2											
	1.5	*	2											
	1.0	*	*	2										
	0.5	*	*	*	*	*	2							
	0.33	*	*	*	*	*	*	*	*	2				
II	2.9	2												
	2.0	*	2											
	1.5	*	2											
	1.0	*	*	2										
	0.75	*	*	*	2									
	0.5	*	*	*	*	*	2							
	0.38	*	*	*	*	*	*	*	2					
	0.3	*	*	*	*	*	*	*	*	*	2			
III	4.2	2												
	2.0	*	*	2										
	1.5	*	*	*	2									
	1.0	*	*	*	*	2								
	0.5	*	*	*	*	*	*	*	2					
	0.38	*	*	*	*	*	*	*	*	*	2			
	0.33	*	*	*	*	*	*	*	*	*	*	2		
IV	4.2	*	2											
	2.0	*	*	2										
	1.25	*	*	*	2									
	1.0	*	*	*	*	2								
	0.5	*	*	*	*	*	*	*	2					
	0.38	*	*	*	*	*	*	*	*	*	2			
	0.33	*	*	*	*	*	*	*	*	*	*	2		

Blank means the same as the nearest left non-blank number
* implies not possible with more replications
The ranges of S.S. ratio reflect the nature of the methods

ratios in these tables are confined to either g.c.a. against s.c.a. or crosses, and s.c.a. or reciprocal against error; the ratios for the block effect, of lesser interest, are not given here.

The following examples illustrate the use of these tables. In crosses (Table 2) with a S.S. ratio of 2.0, method 1 needs three replications in a 4×4 diallel, but just two for a 5×5 diallel. For method 4, which involves fewer crosses, just two replications are required for eight parental lines, and three for 4–7 parental lines.

For g.c.a. (Table 3), an increase in the number of replications does not increase the power to detect differences.

Table 4. Number of replications needed for significance at 5% level for SCA in Griffing's methods I–IV; S.S., sum of squares.

Method	S.S. ratio	Number of parental lines												
		4	5	6	7	8	9	10	11	12	13	14	15	16
I	1.0	2												
	0.75	3	3	3	3	2								
	0.66	3	3	3	3	3	3	3	2					
	0.5	4	3											
	0.33	5	4	4	4	4	4	4	3					
	0.25	6	5	5	5	4								
	0.13	–	–	–	–	–	–	–	–	6				
II	1.9	2												
	1.7	3	2											
	1.5	3	3	3	2									
	1.25	3	3	3	3	3	3	3	3	3	3	2		
	1.0	3												
	0.75	4	4	3										
	0.67	4	4	4	4	3								
	0.5	5	4											
	0.25	7	7	7	6									
III	0.75	2												
	0.5	3												
	0.4	3												
	0.3	4	4	4	4	4	3							
	0.25	4												
	0.2	5	5	5	5	5	4							
	0.15	5	5	5	4									
	0.1	8	8	8	7									
IV	2.4	2												
	2.0	3	2											
	1.75	3	3	2										
	1.5	3	3	3	3	2								
	1.33	3	3	3	3	3	3	3	2					
	1.0	3												
	0.75	4	4	3										
	0.5	4												
	0.33	6	6	5										
	0.25	7	7	7	6									

Blank means the same as the nearest left non-blank number
–, implies at least more than eight replications needed
The ranges of S.S. ratio reflect the nature of the methods

Even with eight replications and four or five parental lines, the g.c.a. effect is unlikely to show a significant result for small ratios. This can be seen in Table 1, which shows that the degrees of freedom for g.c.a. and s.c.a. involve only the number of parental lines, and so a significance test for g.c.a. cannot be affected by increasing the number of replications. It is also noted that methods 1 and 2, which involve the parental lines, are more sensitive in detecting small S.S. ratios than methods 3 and 4, which do not involve them.

By contrast, in determining the s.c.a. for the same S.S. ratio, an increase in the number of parental lines can compensate for a decrease in the number of replications in detecting differences (Table 4). Since the degrees of freedom for error involve the number of replications, any

Table 5. Number of replications needed for significance at 5% level of reciprocal in Griffing's methods I–IV; S.S., sum of squares.

Method	S.S. ratio	Number of parental lines												
		4	5	6	7	8	9	10	11	12	13	14	15	16
I	1.0	2												
	0.9	3	2											
	0.75	3	3	3	3	2								
	0.5	4	3											
	0.33	5	4	4	4	4	4	4	3					
	0.25	5	5	5	5	4								
	0.166	–	–	6	6	6	6	6	5					
	0.13	–	–	–	–	–	–	–	–	6				
III	1.50	2												
	1.20	3	2											
	1.0	3	3	3	2									
	0.75	3	3	3	3	3	2							
	0.67	4	3	3	3	3	3	3	3	2				
	0.5	4	4	3										
	0.3	5	5	5	4									
	0.25	6	6	5	5	4								

Blank means the same as the nearest left non-blank number
–, implies at least more than eight replications needed
The ranges of S.S. ratio reflect the nature of the methods

test of the effects based on the error mean squares will be affected by the number of replications, and so also applies to the reciprocal effects in Table 5.

Method 3, in which parental lines are not included, can detect differences in s.c.a. for smaller S.S. ratios than can be detected by a full diallel cross; this is consistent with Griffing's statement that method 3 can provide unbiased estimates for specific combining ability.

Even with only a small proportion of the total variation in a 12×12 diallel experiment, it was observed that s.c.a. or reciprocal effects can be detected as being significant with few replications; in experiments with four or five parental lines, more replications are needed to achieve the same level of significance.

A sample of published results of diallel cross experiments has been assembled to illustrate the use of these tables (Liang et al. 1968; Cross 1975; Marin and Lipert 1975; Chiang and Crete 1976; Mason and Zuber 1976; Suh et al. 1976; Miller 1977; Watson et al. 1978; Wilson et al. 1978; Caunter and Gracen 1979; Cisar et al. 1982; Boyle 1987). The published mean squares were first converted into sums of squares, and the S.S. ratios calculated. Table 6 gives the actual size of each experiment, the diallel method used, the ratio of the sum of squares, and the minimal number of replications required as determined by this study. For a few g.c.a. in Table 6, the present study suggests that it is not possible to achieve significance with the number of parental lines used. Increasing the replication here does not improve the ability to detect differences. If a range of ratios had first been examined to

Table 6. Published diallel crosses results reconstructed into sum of squares ratio, number of parental lines (p), the number of replications (r), diallel method and optimal replications in brackets

Crops and Traits	p	r	Method	GCA	SCA	REC
Chili Peppers						
Endocarps	9	3	II	7.21* (2)	1.14* (3)	–
Seed	9	3	II	8.46* (2)	1.26* (3)	–
Stem	9	3	II	1.99* (2)	0.69* (3)	–
Placenta	9	3	II	6.59* (2)	0.42 NS (4)	–
Exocarp	9	3	II	8.39* (2)	1.19* (3)	–
Sorghum						
Stomata density	6	2	II	3.91* (2)	0.54 NS (4)	–
Leaf-blade area	6	2	II	1.41* (2)	2.13* (2)	–
Grain	6	4	I	0.45 NS (–)	2.51* (2)	0.12 NS (–)
Protein %	6	4	I	0.36 NS (–)	4.04* (2)	0.11 NS (–)
Protein yield	6	4	I	0.79* (2)	1.60* (2)	0.09 NS (–)
Grain yield	6	2	I	0.37 NS (–)	2.73* (2)	–
Anthesis time	6	2	I	3.21* (2)	3.13* (2)	–
Protein %	6	2	I	3.23* (2)	0.77 NS (3)	–
Tall fescue						
Perloline	6	2	IV	1.96* (2)	3.51* (2)	–
	6	2	IV	2.21* (2)	0.37 NS (4)	–
Nitrogen	6	2	IV	3.21* (2)	1.48 NS (3)	–
	6	2	IV	5.54* (2)	0.41 NS (4)	–
Corn						
Lesion length	7	3	IV	6.80* (2)	1.98* (2)	–
	7	2	IV	1.67* (2)	5.03* (2)	–
Lesion area	7	3	IV	3.53* (2)	1.53* (2)	–
	7	2	IV	1.91* (2)	5.64* (2)	–
Leaf angle	6	4	IV	0.49 NS (–)	33.50* (2)	–
Mean yield	6	4	IV	7.00* (2)	0.62* (3)	–
Leaf area index	6	4 P1	IV	1.08 NS (–)	0.48* (4)	–
	6	4 P3	IV	0.63 NS (–)	0.64* (3)	–
Grain	6	4 P1	IV	1.09 NS (–)	0.55* (4)	–
	6	4 P3	IV	4.76* (2)	0.26 NS (6)	–
No. of ears	6	4 P1	IV	1.73 NS (–)	0.46* (4)	–
	6	4 P3	IV	0.73 NS (–)	1.42* (3)	–
Filling	7	2	IV	3.80* (2)	2.55* (2)	–
Lag period	7	2	IV	2.45* (2)	0.46 NS (4)	–
Rate of fill	7	2	IV	6.00* (2)	1.02 NS (3)	–
Grain yield	7	2	IV	2.17* (2)	1.90* (2)	–
Sugarcane						
Stalk diam.	5	4	IV	2.28 NS (–)	0.50* (4)	–
Stalk length	5	4	IV	1.50 NS (–)	4.00* (2)	–
Stalk number	5	4	IV	1.25 NS (–)	0.44* (4)	–
Stalk weight	5	4	IV	0.25 NS (–)	2.03* (2)	–
Can yield	5	4	IV	0.54 NS (–)	0.96* (3)	–
Brix	5	4	IV	3.09 NS (–)	1.30* (3)	–
Cabbage						
Infection	4	4	II	16.06* (2)	0.31 NS (6)	–
Spruce						
Height 6 yrs	7	3	I	2.78* (2)	0.67* (3)	0.54* (3)
Height 12 yrs	7	3	I	3.06* (2)	0.71* (3)	0.62* (3)
Wheat						
Yield	12	2	III	1.90* (2)	0.96* (2)	0.64 NS (3)
Yield %C Fall	12	2	III	0.51* (2)	0.74* (2)	0.60 NS (3)
Yield %C Spring	12	2	III	1.63* (2)	0.81* (2)	0.74* (2)

*, Significant at $P<0.05$; NS, Not significant; –, Not applicable; ( ), The number in brackets is the minimal number of replications required; (–), Not significant even with more replications

determine if this effect could potentially have become significant by increasing the number of parental lines then experiments unlikely to be informative on g.c.a. might have been avoided. By contrast, for some of the published experiments to detect s.c.a. and reciprocal effects, an increase in the number of replications may have allowed differences to be detected.

Note that the results in Table 6 are from published papers. It is well known that experiments not showing any significant effects tend not to be reported. If these had been available for inclusion in Table 6, there may well have been more examples showing the need either to increase the number of replications, or the number of parental lines, or both.

Although there are many different diallel analysis methods, each with their own statistical and genetical interpretations, the objective of this paper is not to compare them but to show how previous experiments can be used to plan further experiments having the potential to detect genetic information. If it is known that the g.c.a. or s.c.a. is relatively small for a certain crop, the sensitivity can be increased either by increasing the number of replications or the number of parental lines. By contrast, if the g.c.a. or s.c.a. are relatively large, a lower number of replications may be all that is needed. Expressed in another way, if an additive genetic effect is important, an increase in the number of replications will not improve the capability to detect it. If a non-additive genetic effect is important, adding more replications will improve the precision of the experiment, and hence give a greater chance of revealing it.

The approach recommended in this paper is not restricted to Griffing's random model, since it can also be applied to any fixed model, and so reduce the risk of carrying out an experiment unlikely to produce significant results.

References

Boyle TJB (1987) A diallel cross in black spruce. Genome 29:180–186

Counter IG, Gracen VE (1979) A comparative analysis of nuclear-determined resistance to *Helminthosporium maydis* races T and O in maize. Crop Sci 19:333–336

Chiang MS, Crete R (1976) Diallel analysis of the inheritance of resistance of race 6 of *Plasmodiophora brassicae* in cabbage. Can J Plant Sci 56:865–868

Cisar G, Brown CM, Jedlinski H (1982) Diallel analyses for tolerance in winter wheat to the barley yellow dwarf virus. Crop Sci. 22:328–333

Cross HZ (1975) Diallel analysis of duration and rate of grain filling of seven inbred lines of corn. Crop Sci 15:532–535

Griffing B (1956) Concept of general and specific combining ability in relation to diallel crossing systems. Aust J Biol Sci 9:463–493

Liang GHL, Heyne EG, Chung JH, Koh YO (1968) The analysis of heritable variation for three agronomic traits in a 6-variety diallel of grain sorghum, *Sorghum vulgare* Pers. Can J Genet Cytol 10:460–469

Marin O, Lippert LF (1975) Combining ability analysis of anatomical components of the dry fruit in chili pepper. Crop Sci 15:326–329

Mason L, Zuber MS (1976) Diallel analysis of maize for leaf angle, leaf area, yield, and yield components. Crop Sci 16:693–696

Miller JD (1977) Combining ability and yield component analyses in a five-parent diallel cross in sugarcane. Crop Sci 17:545–547

Snedecor GW, Cochran WG (1980) Statistical Methods, 7th edn. Iowa State University Press, Ames, Iowa

Steel RGD, Torrie JH (1980) Principles and procedures of statistics: a biometrical approach, 2nd edn. McGraw-Hill Book Co. New York

Suh HW, Dayton AD, Casady AJ, Liang GH (1976) Diallel cross analysis of stomatal density and leaf-blade area in grain sorghum, *Sorghum bicolor*. Can J Genet Cytol 18:679–686

Watson CE Jr, Frakes RV, Chilcote DO, Sleper DA, Matches AG (1978) Genetic variation for perloline, N, and digestibility in tall fescue. Crop Sci 18:458–460

Wilson ND, Weibel DE, Mcnew RW (1978) Diallel analysis of grain yield, percent protein, and protein yield in grain sorghum. Crop Sci 18:491–494