

Partial Tetra-Allel Crosses^{*1}

KLAUS HINKELMANN

Iowa State University and Virginia Polytechnic Institute

Summary. The construction of partial tetra-allel crosses (PTAC) is considered using BIB and PBIB designs with blocks of size four. It is shown how this can lead to certain types of balanced designs. An explicit procedure is given for constructing circulant PTAC's. The analysis of PTAC's — estimation of general effects of the lines involved and analysis of variance — is illustrated in terms of an example.

1. Introduction

Double cross hybrids are frequently used in plant breeding programs. A practical problem may arise if one wants to determine for such a program the best double crosses or the best performing lines in double crosses from a set of n lines, since the number of possible double crosses is $n(n-1)(n-2)(n-3)/8$. This is in general a prohibitively large number, and therefore, various prediction formulae for the performance of double crosses from single or three-way crosses have been used (e.g., EBERHARDT, RUSSEL, and PENNY, 1964). Another way of dealing with this problem is to consider only a sample of all possible double crosses and obtain information about the „general combining ability“ in double crosses for each line. Designs of obtaining suitable samples will be called partial tetra-allel crosses (PTAC). The construction of such designs as considered in this paper is based on balanced incomplete block designs (BIB) and partially balanced incomplete block designs (PBIB). Different types of PTAC's and their analysis are discussed.

2. Parametrization of Double Crosses

Denote a double cross by $(ij)(kl)$, where i, j, k, l are four different lines out of the n lines and where (ij) and (kl) are the two unrelated single crosses forming the double cross. For the genotypic value $y_{(ij)(kl)}$ we assume the following model based on general genetic considerations:

$$y_{(ij)(kl)} = \mu + h_i + h_j + h_k + h_l + d_{ij} + d_{kl} + s_{i,k} + s_{i,l} + s_{j,k} + s_{j,l} + t_{(ij)k} + t_{(ij)l} + t_{(kl)i} + t_{(kl)j} + w_{(ij)(kl)} \quad (1)$$

where μ is the population mean, h_i is the general effect of line i (we shall see later — cf. equation 4a — that h_i is different from the general combining ability for single crosses), d_{ij} is the interaction effect between lines i and j appearing in the same single cross, $s_{i,k}$ is the interaction effect between lines i and k appearing in different single crosses, $t_{(ij)k}$ is the interaction effect between single cross (ij) and line k , and $w_{(ij)(kl)}$ is the interaction effect between the single crosses (ij) and (kl) .

This model is different from the one used by RAWLINGS and COCKERHAM (1962) in that it is not orthogonal; i.e., if the effects are assumed to be random variables with mean zero, there exist covariances between certain parameters. These covariances along with the variances of the parameters in the model, we denote as follows:

$$\left. \begin{aligned} \text{Var } (h_i) &= \sigma_h^2, \text{ Var } (d_{ij}) = \sigma_d^2, \text{ Var } (s_{i,k}) = \sigma_s^2, \\ \text{Var } (t_{(ij)k}) &= \sigma_t^2, \text{ Var } (w_{(ij)(kl)}) = \sigma_w^2, \\ \text{Cov } (d_{ij}, s_{i,j}) &= \sigma_{ds}, \text{ Cov } (t_{(ij)k}, t_{(kl)j}) = \sigma_{tt}, \\ \text{Cov } (w_{(ij)(kl)}, w_{(ik)(jl)}) &= \sigma_{ww} \end{aligned} \right\} \quad (2)$$

There exist eight covariances between double cross hybrids each of which can be expressed in terms of the parametric variance and covariance components defined in (2). Using KEMPTHORNE's (1957) notation, the covariance between the genotypic values y_X and y_Z of two collateral relatives X and Z from a Hardy-Weinberg equilibrium population with no linkage can also be written in terms of genetic variance components as follows:

$$\text{Cov } (y_X y_Z) = \sum_{\alpha, \delta} [2 r_{XZ}(1+F)]^\alpha [u_{XZ}(1+F)^2]^\delta \sigma_{A\alpha D\delta}^2 \quad (3)$$

where F is the inbreeding coefficient for the parents, and r_{XZ} and u_{XZ} are the probabilities of alikeness by descent for the types of relatives when the parents are non-inbred.

The eight covariances are given in Table 1 together with the corresponding values $2 r_{XZ}$, u_{XZ} of (3), and the coefficients for the parametric variance and covariance components. From Table 1 we obtain, for example,

$$\left. \begin{aligned} \sigma_h^2 &= \sum_{\alpha} \left[\frac{1+F}{16} \right]^\alpha \sigma_{A\alpha}^2 & (a) \\ \sigma_d^2 &= \sum_{\alpha} \left[\left(\frac{1+F}{8} \right)^\alpha - 2 \left(\frac{1+F}{16} \right)^\alpha \right] \sigma_{A\alpha}^2 & (b) \\ \sigma_{ds} &= \sigma_d^2 & (c) \\ \sigma_s^2 &= \sum_{\alpha, \delta} \left(\frac{1+F}{8} \right)^\alpha \left(\frac{1+F}{64} \right)^\delta \sigma_{A\alpha D\delta}^2 - 2 \sum_{\alpha} \left(\frac{1+F}{16} \right)^\alpha \sigma_{A\alpha}^2 & (d) \end{aligned} \right\} \quad (4)$$

It follows from (4) that h accounts for all the additive effects as far as they are not contained in the error due to variation within crosses. We also see from Table 1 and (4) that the order of lines is important

* Part of this work was supported by National Science Foundation Grant NSF-16491.

¹ Journal Paper No. J-5831 of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa. Project No. 890.

Table 1. Covariances Between Double Cross Hybrids

Covariances ^a	$2r_{XZ}$	u_{XZ}	σ_h^2	σ_d^2	σ_s^2	σ_{ds}	σ_t^2	σ_{tt}	σ_w^2	σ_{ww}
$\text{Cov}_1 = \text{Cov}(y_{(i)j(k)l}, y_{(i)j(k)l})$	1/4	1/16	4	2	4		4		1	
$\text{Cov}_2 = \text{Cov}(y_{(i)j(k)l}, y_{(i)k(j)l})$	1/4	1/32	4		2	4		4		1
$\text{Cov}_3 = \text{Cov}(y_{(i)j(k)l}, y_{(i)j(k)l'})$	3/16	1/32	3	1	2		1			
$\text{Cov}_4 = \text{Cov}(y_{(i)j(k)l}, y_{(i)k(j)l'})$	3/16	1/64	3		1	2		1		
$\text{Cov}_5 = \text{Cov}(y_{(i)j(k)l}, y_{(i)j(k'l)})$	1/8	1/64	2		1					
$\text{Cov}_6 = \text{Cov}(y_{(i)j(k)l}, y_{(i)j(k'l)})$	1/8	0	2	1						
$\text{Cov}_7 = \text{Cov}(y_{(i)j(k)l}, y_{(i)k(j'l)})$	1/8	0	2			1				
$\text{Cov}_8 = \text{Cov}(y_{(i)j(k)l}, y_{(i)j(k'l)})$	1/16	0	1							

^a Primed subscripts are different from any subscript already appearing in a term.

only if there are dominance effects and dominance type epistatic effects; *i.e.*, in the absence of dominance effects and dominance type epistatic effects we have

$$\begin{aligned}\sigma_d^2 &= \sigma_{ds} = \sigma_s^2, \\ \sigma_{tt} &= \sigma_t^2, \\ \sigma_{ww} &= \sigma_w^2.\end{aligned}$$

For a first approximation we are thus led to adopt the model

$$y_{(ij)(kl)} = \mu + h_i + h_j + h_k + h_l + e_{(ij)(kl)} \quad (5)$$

where the $e_{(ij)(kl)}$'s are assumed to be independent random variables with mean zero and variance σ^2 .

3. Construction of PTAC's

Before we discuss the construction of PTAC's, we first have to define this term properly. A set of matings is said to be a PTAC, if the following conditions are satisfied: (i) every line occurs exactly r times in the set; (ii) every double cross occurs at most once.

This definition does not exclude the possibility that a particular single cross occurs several times, whereas other single crosses do not occur at all. From a purely practical point of view this may be an advantage, but for reasons concerning the inferences to be drawn from a PTAC this is undesirable; for example, this might cause a bias in the estimates of the genetic variance components. We shall, therefore, also consider the following type of PTAC's: A set of matings is said to be a PTAC in the strict sense, if it satisfies conditions (i) and (ii) as given above and (iii) every single cross occurs at most once. Naturally, such PTAC's constitute a particular subclass within the class of all PTAC's, and, hence, are obtained by choosing appropriately the parameters of a design that can be used for constructing PTAC's.

The most general way of constructing PTAC's is that by means of incomplete block designs. Following the same argument as for single and three-way crosses (e.g. CURNOW, 1963; HINKELMANN, 1965), we can set up a correspondence between a double cross and blocks of size four. Since by adopting (5) we do not differentiate between the order of the lines the double cross, a procedure to be followed then is to construct incomplete block designs with blocks of size four and in some way assign the elements to two pairs, each pair constituting a single cross. The totality of all such blocks with every element occurring r times is then a PTAC. The total number of double crosses in a PTAC is $nr/4$, where n is the number of lines. For a complete tetra-allele cross we have

$r = (n-1)(n-2)(n-3)/2$. Hence, the degree of fractionation for a PTAC is $2r/[(n-1)(n-2)(n-3)]$.

Suitable designs for constructing PTAC's are m -associate class PBIB designs with blocks of size four. For $m = 2$, appropriate designs are listed by BOSE, CLATWORTHY and SHRIKHANDE (1954). The definition and examples of some m -associate PBIB designs are given by VARTAK (1959), RAGHAVARAO (1960), THARTHARE (1963), and HINKELMANN (1964). If the λ -parameters of these designs take on only the values zero or one, we obtain a PTAC in the strict sense. But this is not the only way of obtaining PTAC's in the strict sense. If the design is given in replication groups orthogonal to blocks, one could use the elements of the first two replication groups within each block to make up the first single cross and the last two replication groups to make up the other single cross, rather than to arrange the elements randomly into single crosses. In many cases this will lead to PTAC's in the strict sense.

As an illustration consider the following example. Suppose $n = 10$, $r = 4$. From the two-associate class design T_{12} in BOSE, CLATWORTHY, and SHRIKHANDE (1954) we obtain the following PTAC in the strict sense: (2,10) (6,7), (10,1) (2,5), (7,3) (8,2), (6,2) (9,4), (1,9) (10,8), (5,4) (3,10), (8,7) (4,1), (3,5) (7,9), (9,6) (1,3), (4,8) (5,6).

BIB designs can be used to construct balanced PTAC's; *i.e.*, PTAC's in which every line occurs equally often with every other line in a double cross. Under certain circumstances it is even possible to achieve the balance in such a way that also every line occurs equally often with every other line in the same single cross. The following PTAC for $n = 8$, $r = 7$ is such a balanced PTAC in the strict sense obtained from a BIB association scheme with $\lambda = 3$ (cf. COCHRAN and COX, 1957, p. 473): (1,2) (3,4), (5,6) (7,8), (1,3) (6,8), (2,4) (5,7), (1,4) (6,7), (2,3) (5,8), (1,5) (3,7), (2,6) (4,8), (1,6) (2,5), (3,8) (4,7), (1,7) (2,8), (3,5) (4,6), (1,8) (4,5), (2,7) (3,6).

4. Circulant PTAC's

The number of suitable PBIB's is not very large. We shall, therefore, describe now a more systematic way of constructing PTAC's which is a special case of of the procedure given in the previous section. Since these designs are closely related to circulant PBIB's as defined by KEMPTHORNE (1953), they will be called circulant PTAC's.

Denote the following set of single crosses (i,j) , $(i+1, j+1)$, $(i+2, j+2)$, ..., $(i+n-1, j+n-1)$ by $S(i,j)$, where every number is reduced mod n (*i.e.*, replaced by the remainder after dividing by n) with

the convention that 0 is replaced by n . We then adjoin two such sets, $S(i, j)$ and $S(k, l)$, say, with i, j, k, l all different, to obtain the following array: $(i, j) (k, l), (i+1, j+1) (k+1, l+1), \dots, (i+n-1, j+n-1) (k+n-1, l+n-1)$, which we denote by $T(i, j; k, l)$. Each element of $T(i, j; k, l)$ represents a four-way cross, and since every line occurs exactly twice in the sets $S(i, j)$ and $S(k, l)$, it follows that every $T(i, j; k, l)$ is a PTAC with $r = 4$ and of size n , i.e., a PTAC with n double crosses. A circulant PTAC of size cn and $r = 4c$ is obtained by considering c different $T(i, j; k, l)$.

Suppose now without loss of generality that $i < j$ and $k < l$, and denote $j-i$ by d_1 and $l-k$ by d_2 . $T(i, j; k, l)$ is then a PTAC in the strict sense if and only if one of the following two conditions is satisfied:

- (a) $d_1 \neq d_2$ if $d_1, d_2 < n/2$ or $d_1, d_2 > n/2$,
 (b) $d_1 \neq n - d_2$ if $d_1 < n/2, d_2 > n/2$ or $d_1 > n/2, d_2 < n/2$.

As an example, consider the case $n = 10$ and the PTAC $T(1, 2; 3, 5)$ which is given by the following four way crosses: (1,2) (3,5), (2,3) (4,6), (3,4) (5,7), (4,5) (6,8) (5,6) (7,9), (6,7) (8,10), (7,8) (9,1), (8,9) (10,2), (9,10) (1,3), (10,1) (2,4). Since $d_1 = 1$ and $d_2 = 2$, condition (a) is satisfied, and hence, this is a PTAC in the strict sense.

5. Analysis of PTAC's

The analysis of PTAC's relies very heavily on the properties and parameters of PBIB designs and can be given in general terms (HINKELMANN, 1963). Rather than dealing with the general case we shall illustrate the method by considering the analysis of a fictitious numerical example.

Suppose we have $n = 6$ lines, each line being used $r = 8$ times, and suppose we have obtained the following PTAC from design R4 by BOSE et al. (1954) (note that this is one of the many possible PTAC's that can be obtained from this design) as given in Table 2.

Table 2. Double Crosses and Yields

$(i, j) (k, l)$	$y(ij)(kl)$	$(i, j) (k, l)$	$y(ij)(kl)$
(3,4) (5,6)	2	(5,1) (4,3)	7
(5,4) (2,6)	1	(6,1) (4,2)	12
(2,3) (6,5)	5	(4,5) (1,2)	3
(2,3) (6,4)	6	(6,2) (3,1)	6
(1,6) (5,4)	10	(1,5) (2,3)	6
(3,6) (1,5)	4	(4,2) (3,1)	10
Mean (\bar{y})		6	

Table 3. Ω_i and \hat{h}_i

i	Ω_i	\hat{h}_i
1	10	3.50
2	1	-.88
3	-2	-.38
4	3	.88
5	-10	-3.63
6	-2	.50
Sum	0	0

For each line we evaluate the quantity Ω_i (Table 3), where

$$\Omega_i = \sum_{j, k, l} a_{(ij)(kl)} (y_{(ij)(kl)} - \bar{y})$$

with $a_{(ij)(kl)} = 1$ if $(ij)(kl)$ is contained in the PTAC = 0 otherwise.

Estimates of the h_i are obtained by solving the following system of equations:

$$B \begin{bmatrix} \hat{h}_1 \\ \hat{h}_2 \\ \hat{h}_3 \\ \hat{h}_4 \\ \hat{h}_5 \\ \hat{h}_6 \end{bmatrix} = \begin{bmatrix} \Omega_1 \\ \Omega_2 \\ \Omega_3 \\ \Omega_4 \\ \Omega_5 \\ \Omega_6 \end{bmatrix}$$

where $B = (b_{ij})$ is a 6×6 matrix, the elements of which are determined by the parameters of the underlying PBIB design as follows: $b_{ii} = r$ for all i , $b_{ij} = \lambda_1$ if i and j are first associates, $b_{ij} = \lambda_2$ if i and j are second associates (the parameters λ_1 and λ_2 indicate how often a particular line occurs together with its first and second associates, resp., in a double cross). For this particular design, $\lambda_1 = 4, \lambda_2 = 5$ with the association scheme as given in Table 4.

Table 4. Association Scheme

line	1. associates	2. associates
1	6	2,3,4,5
2	5	1,3,4,6
3	4	1,2,5,6
4	3	1,2,5,6
5	2	1,3,4,6
6	1	2,3,4,5

Hence B is of the form

$$B = \begin{bmatrix} 8 & 5 & 5 & 5 & 5 & 4 \\ 5 & 8 & 5 & 5 & 4 & 5 \\ 5 & 5 & 8 & 4 & 5 & 5 \\ 5 & 5 & 4 & 8 & 5 & 5 \\ 5 & 4 & 5 & 5 & 8 & 5 \\ 4 & 5 & 5 & 5 & 5 & 8 \end{bmatrix}$$

One could now invert this matrix if such an inverse exists, but because of the special structure of B , this problem can be reduced to inverting a 3×3 matrix since the inverse B^{-1} will have the same structure as B with elements b^0 along the diagonal, b^1 in positions corresponding to λ_1 , and b^2 in positions corresponding to λ_2 . To set up a system of equations for b^0, b^1, b^2 , one has to use the parameters of the second kind, p_{ij}^k , which are given with the design. These are (BOSE et al., 1954, p. 13) in matrix notation

$$P_1 = (p_{ij}^1) = \begin{pmatrix} 0 & 0 \\ 0 & 4 \end{pmatrix}$$

$$P_2 = (p_{ij}^2) = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$$

Table 5 and Table 6 are not the same.) The comparison of the (Interaction) M. S. with the (Error) M. S. provides a goodness-of-fit criterion for the model (5). If the test is significant, this would indicate that dominance and/or dominance-type epistatic effects are also present.

Zusammenfassung

In dieser Arbeit wird die Konstruktion von unvollständigen Tetra-Allelen (PTAC) mit Hilfe von unvollständigen Blockversuchsplänen (BIB- und PBIB-Plänen) betrachtet. Es wird gezeigt, wie dies zu gewissen ausgewogenen Plänen führen kann. Eine Methode zur Konstruktion von zirkularen PTAC's wird explizit angegeben. Die Analyse von PTAC's, d.h. das Schätzen von allgemeinen Effekten der betrachteten Linien und die Varianzanalyse, wird an einem Beispiel erläutert.

References

1. BOSE, R. C., W. H. CLATWORTHY, and S. S. SHRIKHANDE: Tables of partially balanced designs with two

associate classes. North Carolina Agricultural Experiment Station Technical Bulletin 107 (1954). — 2. COCHRAN, W. G., and G. M. COX: Experimental designs. New York: Wiley 1957. — 3. CURNOW, R. N.: Sampling the diallel cross. *Biometrics* **19**, 287–306 (1963). — 4. EBERHART, S. A., W. A. RUSSEL, and L. H. PENNY: Double cross hybrid prediction in maize when epistasis is present. *Crop Science* **4**, 363–366 (1964). — 5. HINKELMANN, K.: Design and analysis of multi-way genetic cross experiments. Unpubl. Ph. D. Thesis Iowa State University (1963). — 6. HINKELMANN, K.: Extended group divisible partially balanced incomplete block designs. *Ann. Math. Statist.* **35**, 681–695 (1964). — 7. HINKELMANN, K.: Partial triallel crosses. *Sankhyā* **27**, 173–196 (1965). — 8. KEMPTHORNE, O.: A class of experimental designs using blocks of two plots. *Ann. Math. Statist.* **24**, 78–84 (1953). — 9. KEMPTHORNE, O.: An introduction to genetic statistics. New York: Wiley 1957. — 10. KEMPTHORNE O., and R. N. CURNOW: The partial diallel cross. *Biometrics* **17**, 229–250 (1961). — 11. RAGHAVARAO, D.: A generalization of group divisible designs. *Ann. Math. Statist.* **31**, 756–771 (1960). — 12. RAWLINGS, J. O., and C. C. COCKERHAM: Analysis of double cross hybrid populations. *Biometrics* **18**, 229–244 (1962). — 13. THARTHARE, S. K.: Right angular designs. *Ann. Math. Statist.* **34**, 1057 bis 1067 (1963). — 14. VARTAK, M. N.: The non-existence of certain PBIB designs. *Ann. Math. Statist.* **30**, 1051 bis 1062 (1959).