THEORETICAL ANALYSES OF DOUBLE-CROSS AND THREE-WAY HYBRID POPULATIONS

by

JOHN OREN RAWLINGS

&

C. CLARK COCKERHAM

NORTH CAROLINA STATE COLLEGE

Institute of Statistics
Mimeo Series No. 248
March, 1960

1.0 TABLE OF CONTENTS

	Page
2.0 LIST OF TABLES	vi
3.0 INTRODUCTION	1
4.0 REVIEW OF LITERATURE	2
4.1 Correlations between Relatives	2
4.2 Analysis of Single-Cross Hybrids	6
5.0 ANALYSIS OF VARIANCE FOR DOUBLE-CROSS HYBRIDS	10
5.1 Introduction	10
5.2 Expectations of Quadratic Forms in Terms of Covariances	11
5.3 Genetic Expectations of Covariances	13
5.4 Partitioning of Sums of Squares	14
5.5 Expectations of Sums of Squares in Terms of Covariances	19
5.6 Model of Linear Effects	25
5.7 Covariances as Functions of the Components of Variance	28
5.8 Expectations of Mean Squares in Terms of Components of Variance and Estimation of Components of Variance	28
5.9 Components of Variance as Functions of Covariances between Hybrid Relatives	32
5.10 Genetic Interpretations of the Components of Variance and Tests of Significance	35
5.11 Estimation of Genetic Parameters	42
5.12 Discussion of the Analysis of Double-Crosses	45
6.0 ANALYSIS OF VARIANCE FOR THREE-WAY HYBRIDS	48
6.1 Introduction	48
6.2 Expectation of Quadratic Forms in Terms of Covariances	48
6.3 Genetic Expectations of Covariances	51
6.4 Partitioning of Sums of Squares	52
6.5 Expectations of Sums of Squares in Terms of Covariances	57
6.6 Model of Linear Effects	57
6.7 Covariances as Functions of the Components of Variance	63
6.8 Expectations of Mean Squares in Terms of Components of Variance and Estimation of Components of Variance	64
6.9 Components of Variance as Functions of Covariances between Hybrid Relatives	66
6.10 Genetic Interpretations of the Components of Variance and Tests of Significance	67
6.11 Estimation of Genetic Parameters	73
6.12 Discussion of the Analysis of Three-Way Hybrids	76

1.0 TABLE OF CONTENTS(continued)

Page

7.0 GENERAL DISCUSSION 79

8.0 SUMMARY AND CONCLUSIONS. 82

9.0 LIST OF REFERENCES 84

2.0 LIST OF TABLES

Page

1. Correlations between additive deviations, p, and dominance deviations, q, of double-cross hybrid relatives. 13

2. Description of sums pertinent to the analysis of double-cross hybrids. 16

3. Analysis of variance of double-cross hybrids 17

4. Expectations of uncorrected sums of squares in terms of the covariances between double-cross hybrid relatives. . . 20

5. Expectations of mean squares in terms of the covariances between hybrid relatives for the analysis of double-cross hybrids. 24

6. Covariances between double-cross hybrid relatives expressed as functions of the components of variance 28

7. Expectations of mean squares with Model A for the analysis of double-crosses. 29

8. Expectations of mean squares with Model B for the analysis of double-crosses. 30

9. Components of variance as functions of the covariances between double-cross hybrid relatives for Model A 33

10. Components of variance as functions of the covariances between double-cross hybrid relatives for Model B 34

11. Components of variance in the analysis of double-crosses as functions of the components of genetic variance up to and including the three-factor epistatic components . . 36

12. F-tests of hypotheses that various components of variance are zero in the analysis of double-crosses 39

13. Correlations between additive deviations, p, and dominance deviations, q, of three-way hybrid relatives. 50

14. Description of sums pertinent to the analysis of three-way hybrids 54

15. Analysis of variance of three-way hybrids 55

16. Expectations of uncorrected sums of squares in terms of the covariances between three-way hybrid relatives 58

2.0 LIST OF TABLES (continued)

Page

17. Expectations of mean squares in terms of the covariances between hybrid relatives for the analysis of three-way hybrids 60

18. Covariances between three-way hybrid relatives expressed as functions of the components of variance 63

19. Expectations of mean squares for the analysis of three-way hybrids 65

20. Components of variance as functions of the covariances between three-way hybrid relatives 68

21. Components of variance in the analysis of three-way crosses as functions of the components of genetic variance up to and including the three-factor epistatic components . . . 69

22. F-tests of hypotheses that various components of variance are zero in the analysis of three-way crosses 71

3.0 INTRODUCTION

The use of statistical concepts to characterize the variation in quantitative characteristics has become widely accepted as an approach to studying the inheritance of such characters. The nature of the variation in quantitative characters, that is, the absence of discrete classes, precludes the use of classical genetic methods of investigating the inheritance. A knowledge of the mode of inheritance of characters is essential to the development of efficient breeding programs for improvement of organisms. As a rule, however, precise estimation of the genetic parameters which prove most useful is difficult and generally requires very elaborate and extensive experiments. It would be desirable if these estimates could be obtained as a matter of course in a breeding program without extensive special effort.

The evaluation of potential hybrids constitutes a large part of most corn breeding programs. The analysis of variance of all possible single-cross hybrids from a set of parents has been shown to provide estimates of some of the genetic parameters. It is the purpose of this dissertation to develop analyses of variance for all possible double-cross hybrids and three-way hybrids from a set of p lines and to determine the genetic information available from such analyses.

4.0 REVIEW OF LITERATURE

4.1 Correlations between Relatives

The first general treatment of the theoretical genetic correlations between relatives was given by Fisher (1918). The heritable variation for a locus was partitioned into additive variance and dominance variance. Non-additivity of the effects between loci was defined as epistacy. The correlations for parent-offspring, full-sib, uncle-nephew, cousin, and double-first-cousins were treated for random mating populations with any degree of dominance but no epistacy. The general solution considered an arbitrary number of alleles. The effects of two-locus epistatic deviations on the covariances of the above relatives were presented. Assortative mating and coupling effects were studied.

Wright (1921) proposed the use of the method of path coefficients for the study of the genetic correlations between relatives in the case of no dominance or epistasis. Dominance was considered for some special cases but, primarily, only the additive effects of the genes were considered. The method of path coefficients did not provide a general treatment of non-additive gene effects so that the covariances between dominance deviations were not encountered by this method. The parent-offspring and full-sib correlations were given by Wright (1935) for a special model in which the value of a secondary character is dependent upon the deviation from an optimum of a primary character. The cases of either no dominance or complete dominance at all loci were considered.

Malécot[1] (Kempthorne, 1954, 1957) gave a general treatment of the covariance between relatives for the case of random mating for a single locus with arbitrary dominance among an arbitrary number of alleles in 1948. The only restriction applied by Malécot was that of no inbreeding. The covariance between two individuals was given as

$$\left(\frac{\phi + \phi'}{2}\right)\sigma_A^2 + \phi\phi'\sigma_D^2$$

where σ_A^2 is the additive genetic variance, σ_D^2 is the dominance variance and ϕ and ϕ' are measures of the relationship between the two individuals for the locus under consideration. ϕ and ϕ' can be used as above only in the particular case when x_s, y_s can be identified as genes contributed to individuals X and Y, respectively, on one chromosome and x_d, y_d as genes contributed to X and Y on the other chromosome with no relationship between the two chromosomes received by X or by Y, i.e., when the sire genes, x_s, and the dam genes, x_d, are distinct sets. Then $\phi = P(x_s = y_s)$ and $\phi' = P(x_d = y_d)$, where $P(x_s = y_s)$ is the probability that genes x_s and y_s are identical by descent. The more general result for the covariance between $X = (A_{x_s}\ A_{x_d})$ and $Y = (A_{y_s}\ A_{y_d})$ was given as

$$[P(x_s = y_s) + P(x_d = y_s) + P(x_s = y_d) + P(x_d = y_d)]\frac{\sigma_A^2}{2}$$
$$+ [P(x_s = y_s,\ x_d = y_d) + P(x_s = y_d,\ x_d = y_s)]\sigma_D^2 \ . \qquad (4.1.1)$$

Anderson and Kempthorne (1954) presented a factorial gene model which included epistatic parameters for an arbitrary number of loci with gene frequency of one-half. The model was applied to populations

---

[1] Malécot, G., 1948. Les mathematiques de l'heredite. Paris, Masson et Cie.

obtained by successive selfing of a heterozygous individual and to populations arising from crossing two inbred lines with subsequent crossing and selfing.

Cockerham (1954) partitioned the hereditary variance of diploid populations into additive, dominance, additive x additive, additive x dominance, dominance x dominance and higher order epistatic components for the case of two alleles at each of an arbitrary number of loci with no linkage. Correlations between the hereditary deviations were considered for (1) relatives from random mating populations, (2) relatives from a self-fertilizing population and (3) offspring from randomly mated inbred parents. The correlations between the epistatic deviations are products of the one-factor additive and one-factor dominance correlations for cases one and three above and for case two when gene frequency is one-half and the relatives are in the same generation. Hence, the covariance between the hereditary values of two relatives can be expressed for the above situations as

$$\sum_{\substack{A,D=0 \\ 1\leq A+D\leq n}}^{n} p^A q^D \sigma^2_{AD} \qquad (4.1.2)$$

where p and q are the correlations between additive and dominance deviations, respectively, A is the number of times additive appears in the component, D is the number of times dominance appears and n is the number of loci.

Kempthorne (1954, 1955a, 1957) presented the same type of analysis but extended it to include an arbitrary number of alleles for a random mating population but with no inbreeding. The covariance of any two individuals was expressed by an extension of Malécot's notation as

$$\left(\frac{\phi + \phi'}{2}\right)^r (\phi\phi')^s \sigma^2_{A_r D_s} \qquad (4.1.3)$$

where A appears r times and D appears s times. ϕ and ϕ' are as defined previously if the sire and dam genes are distinct sets. If sire and dam genes are not distinct sets, $\frac{\phi + \phi'}{2}$ and $\phi\phi'$ are defined as the coefficients of σ^2_A and σ^2_D, respectively, as given in expression (4.1.1).

Horner and Kempthorne (1955) determined the correlation between relatives in random mating populations for which the gene effects were the same for all loci assuming only two alleles per locus and gene frequency of one-half with no linkage. Some of the classical epistatic models were considered.

Kempthorne (1955b) considered the correlations between relatives derived from a random mating population by some regular system of inbreeding. The covariance between full-sibs after n generations of inbreeding and the covariance between a parent in the nth generation and offspring in the (n + 1)th generation were derived for the particular case of full-sib inbreeding.

The effects of linkage in studies of correlations between relatives have generally been assumed absent. Cockerham (1956) studied the effects of linkage on the covariances between relatives in random mating populations which are in linkage equilibrium assuming absence of position effects. Ancestral covariances were found to be unaffected by linkage, whereas, the covariances not involving ancestral relationships were affected. The effect of linkage in nonancestral relationships was to increase the covariance but only in the coefficients of the epistatic components and not in the coefficients of the additive and dominance components.

4.2 Analysis of Single-Cross Hybrids

The analysis of diallel crosses or the analysis of all possible single-crosses from a set of lines has been used to obtain information concerning quantitative inheritance. The methods of analysis have been varied.

Hull (1946) discussed the interpretation of the constant parent regression analysis of all possible single-crosses from a set of homozygous lines. Epistacy was assumed to be absent. The regression of the constant parent regression coefficients on the phenotypic values of the constant parent, b_2, was interpreted as an indication of the degree of dominance.

Griffing (1950) described in detail the constant parent regression analysis of quantitative genetic data. With no epistacy, b_2 varied inversely with h, the degree of dominance, i.e., $b_2 = -\frac{h}{4d}$. However, with dominance and a multiplicative type epistacy at two loci, the direction of slope and degree of curvilinearity of b_2 was shown to be dependent upon the particular combination of degree of dominance, h, and the degree of epistacy, e. For $h = 0$, i.e., no dominance, and $e > 0$, the trend of the constant parent regression values are opposite to that for dominance alone.

Burdick (1957) presented a constant parent regression model for multiplicative type epistasis with the assumption of no dominance. With this model, b_2 varied directly with e, a measure of the degree of contribution of multiplicative effects, in contrast to the dominance model with no epistacy in which b_2 varies inversely with the degree of dominance. The sign of e with this model may be either positive or negative. Hence, if multiplicative type epistasis were the cause of heterosis, b_2 could be either positive or negative depending on the sign of e.

Sprague and Tatum (1942) presented an application of the analysis of variance technique to the analysis of the n(n-1)/2 single-crosses from a set of n lines for the estimation of variance due to general combining ability and variance due to specific combining ability. The expectation of the mean square for within line groups involved only variance due to specific combining ability and an error variance.

Yates (1947) presented an analysis of variance of all possible reciprocal crosses from a set of lines for the following cases: (1) self-sterility, (2) no self-sterility and (3) self-sterility with incompatibility within groups of lines. Additivity of parental effects was assumed but a test for the departure from this additivity was available.

Griffing (1950) presented a partitioning of the variance among F_1's within a constant parent group into a basic gene variance, a dominance variance and a covariance between the basic gene effect and the dominance effect. It was assumed that epistasis was either absent or that the proper transformation had been applied to the data to make between locus effects and environmental effects additive.

Hayman (1954a, 1954b) included dominance in the model for the analysis of the diallel set of all n^2 matings with the following assumptions:

(1) Diploid segregation

(2) No difference between reciprocals

(3) Independent action of non-allelic genes

(4) No multiple allelism

(5) Homozygous parents

(6) Genes independently distributed between the parents. Methods for detecting epistacy and for estimating the degree of dominance were presented.

Dickinson and Jinks (1956) extended this method to include heterozygous parental lines. A method was presented for estimating the degree of heterozygosity or the inbreeding coefficient and whether dominant or recessive genes were in excess.

Griffing (1956) suggested the use of the term "modified diallel" to designate the diallel set where the parents are not included. The analysis of the modified diallel from homozygous parents with and without reciprocal crosses was generalized to include an arbitrary number of alleles at an arbitrary number of loci with any magnitude of additive genetic, dominance and epistatic effects. The assumptions and conditions necessary for valid inductive inferences about the parental population were presented as follows:

(1) The set of inbreds is a random sample from a population of inbred lines derived from the parent population by means of an inbreeding system free from forces which change gene frequency.

(2) The parent population is under equilibrium conditions with random mating.

(3) It is necessary to use the "modified diallel" to obtain unbiased estimates of population parameters.

The variances of general and specific combining ability were given explicitly in terms of additive, dominance and epistatic components of variance. The variance of general combining ability was shown to be equal to the parent-offspring covariance in a random mating population in equilibrium.

Kempthorne (1956) removed the assumptions of no epistacy and no multiple alleles as given by Hayman (1954b) and presented the generalized expectations of the sums of squares of the analyses given by Yates (1947) and Hayman (1954a). The parents were assumed to be homozygous lines obtained without selection from a random mating population at equilibrium. It was illustrated that only in the absence of epistacy could information be obtained from the diallel analysis about the original population. Arguments were presented to show that previous analyses in which the parents were not a random sample from a random mating population at equilibrium are not valid.

Matzinger and Kempthorne (1956) extended the analysis of the "modified diallel" to include an arbitrary degree of inbreeding for the parents of the diallel and to obtain estimates of the interactions of the genotypic components of variance and environments.

Kempthore (1957) extended the diallel analysis presented earlier (Kempthorne, 1956) to include arbitrary inbreeding of the parental lines.

5.0 ANALYSIS OF VARIANCE FOR DOUBLE-CROSS HYBRIDS

5.1 Introduction

The term double-cross hybrid is used to denote the progeny of the cross of two F_1 hybrids in which none of the four parents or lines are the same and may be symbolized by (A x B)(C x D). Omitting reciprocal crosses, there are 3 ${}_pC_4$ different double-cross hybrids possible from a set of p grandparents or lines, where ${}_pC_n = \frac{p!}{n!\, p_n!}$; $p! = 1.2...p_2.p_1.p$. The notation p_i is used herein to denote $(p - i)$, e.g., $p_1 = (p - 1)$. There are ${}_pC_4$ different combinations of four grandparents, each combination giving rise to three different double-crosses by changing the pairing, which will be referred to as ordering, of the grandparents in the single-cross parents. For example, the three orders or double-crosses for four lines A, B, C, and D are (A x B)(C x D), (A x C)(B x D), and (A x D)(B x C).

The analysis of variance presented is of the complete set of 3 ${}_pC_4$ double-cross hybrids from a set of p lines. It is assumed that the p lines constitute a random sample of all possible lines produced from some random mating population by some system of inbreeding to an arbitrary degree without selection of any form. The only restrictions on the system of inbreeding are that the inbreeding coefficient can be determined at any particular stage and that all p lines are equally inbred.

5.2 Expectations of Quadratic Forms in Terms of Covariances

The phenotype of the double-cross (A x B)(C x D) grown in the mth replication will be denoted in the following derivations by $Y_{(AB)(CD)m}$ or, in general, $Y_{(ij)(kl)m}$. The paired subscripts within the parentheses indicate the parents of each of the parental single-crosses. The model for the analysis of variance in such an experiment is

$$Y_{(ij)(kl)m} = \mu + r_m + G_{(ij)(kl)} + e_{(ij)(kl)m} \qquad (5.2.1)$$

where

μ = a contribution common to all entries,

r_m = a contribution particular to replication m,

$G_{(ij)(kl)}$ = a contribution particular to double-cross (i x j)(k x l), and

$e_{(ij)(kl)m}$ = a random deviation of the observation on the double-cross (i x j)(k x l) grown in replication m.

It is assumed that $e_{(ij)(kl)m}$'s, r_m's and $G_{(ij)(kl)}$'s are independent random variables with zero means and variances σ_e^2, σ_r^2 and σ_y^2, respectively, and that the e's, r's and G's are uncorrelated with each other. A constant number of individuals in each double-cross progeny in each replication is assumed.

The expectations of the quadratic forms which will be encountered can be defined in terms of μ^2, σ_e^2, σ_r^2, and the covariances between the effects of the particular double-cross hybrids, i.e., the $G_{(ij)(kl)}$'s. The expectation of the square of the phenotype of a double-cross grown in replication m will contain $\mu^2 + \sigma_r^2 + \sigma_e^2$ plus the covariance of the effect of a double-cross with itself, i.e., $\mathrm{Cov}(G_{(ij)(kl)}\ G_{(ij)kl)})$. The expectation of the product of the phenotypes of two different double-crosses in the same replication will contain $\mu^2 + \sigma_r^2$ plus the covariance of the effects of the two double-crosses, i.e., $Cov(G_{(ij)(kl)}\ G_{(mn)(op)})$. The expectation of the product of the phenotypes of the same or different double-crosses grown in different replications will contain μ^2 plus the covariance of the effects of the two double-crosses involved.

The covariances between the effects of double-crosses are defined as listed in Table 1. It will be noted that:

Cov_1 = the covariance of the effect of a double-cross with itself,

Cov_2 = the covariance between the effects of two double-cross hybrids with all four lines common but in different orders,

Cov_3 = the covariance between the effects of two double-cross hybrids with only three lines common and in the same order,

Cov_4 = the covariance between the effects of two double-cross hybrids with only three lines common but in different orders,

Cov_5 = the covariance between the effects of two double-cross hybrids with two lines common and with the two common lines in different parental single-crosses in each case,

Cov_6 = the covariance between the effects of two double-cross hybrids with two lines common and with both of the common lines in the same parental single-cross in each case,

Cov_7 = the covariance between the effects of two double-cross hybrids with two lines common and with the common lines in the same parental single-cross in one double-cross and in different parental single-crosses in the other double-cross and

Cov_8 = the covariance between the effects of two double-cross hybrids with only one line common.

The covariance between the effects of two double-cross hybrids with no lines common is zero.

Table 1. Correlations between additive deviations, p, and dominance deviations, q, of double-cross hybrid relatives[a]

Covariance	Number of lines in common	Correlations Inbreeding = F p	q	F = 1 p	q
$Cov_1 = Cov\ [G_{(ij)(kl)}G_{(ij)(kl)}]$	4	F/2	$F^2/4$	1/2	1/4
$Cov_2 = Cov\ [G_{(ij)(kl)}G_{(ik)(jl)}]$	4	F/2	$F^2/8$	1/2	1/8
$Cov_3 = Cov\ [G_{(ij)(kl)}G_{(ij)(km)}]$	3	3F/8	$F^2/8$	3/8	1/8
$Cov_4 = Cov\ [G_{(ij)(kl)}G_{(ik)(jm)}]$	3	3F/8	$F^2/16$	3/8	1/16
$Cov_5 = Cov\ [G_{(ij)(kl)}G_{(im)(kn)}]$	2	F/4	$F^2/16$	1/4	1/16
$Cov_6 = Cov\ [G_{(ij)(kl)}G_{(ij)(mn)}]$	2	F/4	0	1/4	0
$Cov_7 = Cov\ [G_{(ij)(kl)}G_{(ik)(mn)}]$	2	F/4	0	1/4	0
$Cov_8 = Cov\ [G_{(ij)(kl)}G_{(im)(no)}]$	1	F/8	0	1/8	0
$Cov\ [G_{(ij)(kl)}G_{(mn)(op)}]$	0	0	0	0	0

[a]From Cockerham, C. C. (1957) (see footnote 2, Section 5.3).

5.3 Genetic Expectations of Covariances

In order to express the covariances between effects of relatives as defined in Section 5.2 in terms of the genetic parameters as defined by Cockerham (1954) or Kempthorne (1954), it is necessary to impose the assumption that the p parents are a random sample of lines which would result by inbreeding some equilibrium random mating population without any type of selection. In addition it is necessary to assume no linkage and equal inbreeding for all parents. Cockerham (1954) indicates that the assumption of no linkage is necessary only if the parents are partially inbred ($0<F<1$). For completely inbred parents, the assumption of linkage equilibrium in lieu of no linkage is adequate. With these assumptions, the covariances between the phenotypes of relatives, if the environments are not correlated, can be expressed as

$$\sum_{\substack{A,D=0 \\ 1\le A+D\le n}}^{n} p^A q^D \sigma^2_{AD} \text{ (Cockerham, 1954).} \tag{5.3.1}$$

Cockerham[2] derived the correlations between the additive and dominance deviations, p and q in expression (5.3.1), for various hybrid relatives. The correlations for double-cross hybrid relatives are given in Table 1.

5.4 Partitioning of Sums of Squares

The following notation will be used to indicate the individual observations and sums which are pertinent to the partitioning of the sums of squares:

$Y_{(ij)(kl)m}$ = the phenotypic value of double-cross $(i \times j)(k \times l)$ in the mth replication; $m = 1,2,\ldots,r$; $i,j,k,l = 1,2,\ldots,p$ where $i \neq j \neq k \neq l$.

$Y_{(ij)(kl).} = \sum_{m=1}^{r} Y_{(ij)(kl)m}$ = sum of double-cross $(i \times j)(k \times l)$ over r replications.

$Y_{ijkl.} = \sum_{o}^{3} Y_{(ij)(kl).}$ = sum of the three orders or double-crosses involving the combination of four lines i,j,k and l.

---

[2] Cockerham, C. C. 1957, Coefficients of the additive and dominance components of variance between hybrid relatives. (unpublished).

$$Y_{(ij)(k.).} = \sum_{l\neq i,j,k}^{p_3} Y_{(ij)(kl).}$$ = sum of all double-crosses with parental single-cross (i x j) and grandparent k.

$$Y_{ijk..} = \sum_{o}^{3} Y_{(ij)(k.).}$$ = sum of all double-crosses with lines i,j and k as grandparents.

$$Y_{(ij)(..).} = \frac{1}{2}\sum_{k\neq i,j}^{p_2} Y_{(ij)(k.).} = \sum_{\substack{k,l\neq i,j\\ k<l}}^{p_2 C_2} Y_{(ij)(kl).}$$ = sum of all double-crosses with parental single-cross (i x j).

$$Y_{(i.)(j.).} = \sum_{k\neq i,j}^{p_2} Y_{(ik)(j.).} = \sum_{\substack{k,l\neq i,j\\ k<l}}^{p_2 C_2} \sum_{o}^{2} Y_{(ik)(jl).}$$ = sum of all double-crosses with grandparental lines i and j in different parental single-crosses.

$$Y_{ij...} = Y_{(ij)(..).} + Y_{(i.)(j.).} = \frac{1}{2}\sum_{k\neq i,j}^{p_2} Y_{ijk..}$$

$$= \sum_{\substack{k,l\neq i,j\\ k<l}}^{p_2 C_2} \sum_{o}^{3} Y_{(ij)(kl).}$$ = sum of all double-crosses with line i and j as grandparents

$$Y_{i....} = \frac{1}{3}\sum_{j\neq i}^{p_1} Y_{ij...} = \sum_{\substack{j,k,l\neq i\\ j<k<l}}^{p_1 C_3} \sum_{o}^{3} Y_{(ij)(kl).}$$ = sum of all double-crosses with line i as one grandparent.

$$Y_{.....} = \frac{1}{4}\sum_{i}^{p} Y_{i....} = \sum_{\substack{i,j,k,l \\ i<j<k<l}}^{{}_pC_4} \sum_{o}^{3} Y_{(ij)(kl).} = \text{grand total.}$$

Other pertinent facts regarding these summations are presented in Table 2.

Table 2. Description of sums pertinent to the analysis of double-cross hybrids

Sum	Number of observations/sum	Number of such sums	Total of all such sums equals
$Y_{(ij)(kl)m}$	1	$\frac{rpp_1p_2p_3}{8}$	$Y_{.....}$
$Y_{(ij)(kl).}$	r	$\frac{pp_1p_2p_3}{8}$	$Y_{.....}$
$Y_{ijkl.}$	3r	$\frac{pp_1p_2p_3}{24}$	$Y_{.....}$
$Y_{(ij)(k.).}$	rp_3	$\frac{pp_1p_2}{2}$	$4Y_{.....}$
$Y_{ijk..}$	$3rp_3$	$\frac{pp_1p_2}{6}$	$4Y_{.....}$
$Y_{(ij)(..).}$	$\frac{rp_2p_3}{2}$	$\frac{pp_1}{2}$	$2Y_{.....}$
$Y_{(i.)(j.).}$	rp_2p_3	$\frac{pp_1}{2}$	$4Y_{.....}$
$Y_{ij...}$	$\frac{3rp_2p_3}{2}$	$\frac{pp_1}{2}$	$6Y_{.....}$
$Y_{i....}$	$\frac{rp_1p_2p_3}{2}$	p	$4Y_{.....}$
$Y_{.....}$	$\frac{rpp_1p_2p_3}{8}$	1	$Y_{.....}$

One possible method of partitioning the sums of squares from all possible double-cross hybrids grown in r replications in one environment is presented in Table 3. The partitions of the sums of squares into

Table 3. Analysis of variance of double-cross hybrids

Source	d.f.	Sums of squares	M.S.
Correction factor	1	$C = \dfrac{8Y^2_{\cdots\cdots}}{rpp_1p_2p_3}$	C
Replications	$(r-1)$	$R = \dfrac{8\sum_{m=1}^{r} Y^2_{\cdots\cdots m}}{pp_1p_2p_3} - C$	R*
Hybrids	$3\,{}_pC_4 - 1$	$H = \dfrac{\sum_{i<j<k<l}^{{}_pC_4} \sum_{o}^{3} Y^2_{(ij)(kl).}}{r} - C$	H*
1-line general	p_1	$G = \dfrac{2\sum_{i}^{p} Y^2_{i\cdots\cdots}}{rp_2p_3p_4} - \dfrac{4p_1}{p_4} C$	G*
2-line general	$pp_3/2$	$S = \dfrac{2\sum_{i<j}^{{}_pC_2} Y^2_{ij\cdots}}{3rp_4p_5} - \dfrac{6p_2p_3}{p_4p_5} C - \dfrac{3p_3}{p_5} G$	S*
3-line general	$pp_1p_5/6$	$U = \dfrac{\sum_{i<j<k}^{{}_pC_3} Y^2_{ijk\cdot\cdot}}{3rp_6} - \dfrac{4p_3}{p_6} C - \dfrac{3p_4}{p_6} G - \dfrac{2p_5}{p_6} S$	U*

Table 3(continued)

Source	d.f.	Sums of squares	M.S.
4-line general	$pp_1p_2p_7/24$	$W = \dfrac{\sum_{i<j<k<l}^{{}_pC_4} Y^2_{ijkl.}}{3r} - C - G - S - U$	W*
2-line order	$pp_3/2$	$T = \dfrac{2\sum_{i<j}^{{}_pC_2} Y^2_{(ij)(..).}}{rp_1p_2} + \dfrac{\sum_{i<j}^{{}_pC_2} Y^2_{(i.)(j.).}}{rp_1p_2} - \dfrac{2\sum_{i<j}^{{}_pC_2} Y^2_{ij...}}{3rp_1p_2}$	T*
3-line order	$pp_2p_4/3$	$V = \dfrac{\sum_{i<j<k}^{{}_pC_3} \sum_{o}^{3} Y^2_{(ij)(k.).}}{rp_3} - \dfrac{\sum_{i<j<k}^{{}_pC_3} Y^2_{ijk..}}{3rp_3} - \dfrac{2p_2}{p_3} T$	V*
4-line order	$pp_1p_4p_5/12$	$X = \dfrac{\sum_{i<j<k<l}^{{}_pC_4} \sum_{o}^{3} Y^2_{(ij)(kl).}}{r} - \dfrac{\sum_{i<j<k<l}^{{}_pC_4} Y^2_{ijkl.}}{3r} - T - V$	X*
Error	$(r-1)(3\ {}_pC_4 - 1)$	$E = M - R - H$	E*
Total	$3r\ {}_pC_4 - 1$	$M = \sum_{m}^{r} \sum_{i<j<k<l}^{{}_pC_4} \sum_{o}^{3} Y^2_{(ij)(kl)m} - C$	

variance due to replications, hybrids, and error are the least squares partitions using Model (5.2.1) and are orthogonal. The partitioning of the hybrid sums of squares was arrived at by largely a trial and error procedure. It is, however, an orthogonal partitioning. The proof of orthogonality is long and involved and is not presented in this thesis. Mean squares are indicated by an asterisk on the letter symbolizing the corresponding sum of squares.

5.5 Expectations of Sums of Squares in Terms of Covariances

The expectations of the sums of squares can be derived in terms of μ^2, σ_r^2, σ_e^2 and the covariances between the effects of relatives as defined in Section 5.2. These expectations are presented in Table 4 for the uncorrected sums of squares. The derivation of the expectation of the one-line general sum of squares,

$$E\left[\frac{2\sum_i^p Y_{i....}^2}{rp_1p_2p_3}\right],$$

will be presented for illustration. The sum $Y_{i....}$ contains $rp_1p_2p_3/2$ terms of $Y_{(ij)(kl)m}$ so that $Y_{i....}^2$ will contain, altogether, $r^2p_1^2p_2^2p_3^2/4$ squares or products of these terms. Of these, $rp_1p_2p_3/2$ will be squares and have expectation $\mu^2 + \sigma_e^2 + \sigma_r^2 + \text{Cov}_1$. Each of the $rp_1p_2p_3/2$ terms will be involved in products with (r-1) like terms in other replications so that $r(r-1)p_1p_2p_3/2$ products will have expectation $\mu^2 + \text{Cov}_1$. There will be $r^2p_1p_2p_3$ products between double-crosses with four lines common but in different order. Of these, $rp_1p_2p_3$ of them will be between double-crosses in the same replication with expectation $\mu^2 + \sigma_r^2 + \text{Cov}_2$ and $r(r-1)p_1p_2p_3$ of them will be between double-crosses in different replications with expectation $\mu^2 + \text{Cov}_2$.

Table 4. Expectations of uncorrected sums of squares in terms of the covariances between double-cross hybrid relatives

Sums of squares	Coefficients of covariances									
	$(r\mu^2 + \sigma_r^2)$	σ_e^2	Cov_1	Cov_2	Cov_3	Cov_4	Cov_5	Cov_6	Cov_7	Cov_8
$\frac{Y^2_{.....}}{3r\ {}_pC_4}$	$3\ {}_pC_4$	1	r	2r	$4rp_4$	$8rp_4$	$4rp_4p_5$	rp_4p_5	$4rp_4p_5$	$2rp_4p_5p_6$
$\frac{2\sum_i^p Y^2_{i....}}{rp_1p_2p_3}$	$12\ {}_pC_4$	p	rp	2rp	$3rpp_4$	$6rpp_4$	$2rpp_4p_5$	$\frac{rpp_4p_5}{2}$	$2rpp_4p_5$	$\frac{rpp_4p_5p_6}{2}$
$\frac{\sum_{i<j}^{{}_pC_2} Y^2_{ij...}}{3rp_2p_3}$	$18\ {}_pC_4$	${}_pC_2$	$\frac{rpp_1}{2}$	rpp_1	rpp_1p_4	$2rpp_1p_4$	$\frac{rpp_1p_4p_5}{3}$	$\frac{rpp_1p_4p_5}{12}$	$\frac{rpp_1p_4p_5}{3}$	0
$\frac{\sum_{i<j<k}^{{}_pC_3} Y^2_{ijk..}}{3rp_3}$	$12\ {}_pC_4$	${}_pC_3$	$r\ {}_pC_3$	$2r\ {}_pC_3$	$rp_4\ {}_pC_3$	$2rp_4\ {}_pC_3$	0	0	0	0
$\frac{\sum_{i<j<k<l}^{{}_pC_4} Y^2_{ijkl.}}{3\ r}$	$3\ {}_pC_4$	${}_pC_4$	$r\ {}_pC_4$	$2r\ {}_pC_4$	0	0	0	0	0	0

Table 4(continued)

Sums of squares	Coefficients of covariances									
	$(r\mu^2 + \sigma_r^2)$	σ_e^2	Cov_1	Cov_2	Cov_3	Cov_4	Cov_5	Cov_6	Cov_7	Cov_8
$\frac{2\sum_{i<j}^{{}_pC_2} Y^2_{(ij)(..).}}{rp_2p_3}$	$6\ {}_pC_4$	${}_pC_2$	$r\ {}_pC_2$	0	$2rp_4\ {}_pC_2$	0	0	$\frac{rp_4p_5\ {}_pC_2}{2}$	0	0
$\frac{\sum_{i<j}^{{}_pC_2} Y^2_{(i.)(j.).}}{rp_2p_3}$	$12\ {}_pC_4$	${}_pC_2$	$r\ {}_pC_2$	$r\ {}_pC_2$	$2rp_4\ {}_pC_2$	$2rp_4\ {}_pC_2$	$rp_4p_5\ {}_pC_2$	0	0	0
$\frac{\sum_{i<j<k}^{{}_pC_3} \sum_{o}^{3} Y^2_{(ij)(k.).}}{rp3}$	$12\ {}_pC_4$	$3\ {}_pC_3$	$3r\ {}_pC_3$	0	$3rp_4\ {}_pC_3$	0	0	0	0	0
$\frac{\sum_{i<j<k<l}^{{}_pC_4} \sum_{o}^{3} Y^2_{(ij)(kl).}}{r}$	$3\ {}_pC_4$	$3\ {}_pC_4$	$3r\ {}_pC_4$	0	0	0	0	0	0	0

Of the $3r^2p_1p_2p_3p_4/2$ products between double-crosses with three lines common and in the same order, $3rp_1p_2p_3p_4/2$ will be between double-crosses in the same replication with expectation $\mu^2 + \sigma_r^2 + Cov_3$ and $3r(r-1)p_1p_2p_3p_4/2$ will be between double-crosses in different replications with expectation $\mu^2 + Cov_3$. There will be twice as many products involving hybrids with three lines common but in different order. Hence, $3rp_1p_2p_3p_4$ of the products will have expectation $\mu^2 + \sigma_r^2 + Cov_4$ and $3r(r-1)p_1p_2p_3p_4$ of them will have expectation $\mu^2 + Cov_4$.

Of the $r^2p_1p_2p_3p_4p_5$ products between double-crosses which have two lines common but in different parental single-crosses in both hybrids, $rp_1p_2p_3p_4p_5$ will be between hybrids in the same replication with the expectation $\mu^2 + \sigma_r^2 + Cov_5$ and $r(r-1)p_1p_2p_3p_4p_5$ will be between hybrids in different replications with expectation $\mu^2 + Cov_5$. There will be one-fourth as many products between hybrids with two lines common but in the same parental single-cross; hence, $rp_1p_2p_3p_4p_5/4$ will have expectation $\mu^2 + \sigma_r^2 + Cov_6$ and $r(r-1)p_1p_2p_3p_4p_5/4$ will have expectation $\mu^2 + Cov_6$. The number of products between hybrids with two lines common but in different order will equal the number with the two common lines in different parental single-crosses; hence, $rp_1p_2p_3p_4p_5$ products will have expectation $\mu^2 + \sigma_r^2 + Cov_7$ and $r(r-1)p_1p_2p_3p_4p_5$ will have expectation $\mu^2 + Cov_7$.

Of the $r^2p_1p_2p_3p_4p_5p_6/4$ products between double-crosses with only one line common, $rp_1p_2p_3p_4p_5p_6/4$ will be between double-crosses in the same replication and $r(r-1)p_1p_2p_3p_4p_5p_6/4$ will be between double-crosses in different replication with expectations $\mu^2 + \sigma_r^2 + Cov_8$ and $\mu^2 + Cov_8$, respectively.

There will be no products with expectation μ^2 or $\mu^2 + \sigma_r^2$ since all double-crosses in the sum $Y_{i....}$ have at least one line common.

In summary,

$$E\left[\frac{2\sum\limits_{i}^{p} Y_{i....}^2}{rp_1p_2p_3}\right] = \frac{pp_1p_2p_3}{2}(r\,\mu^2 + \sigma_r^2) + p\sigma_e^2 + rp(\mathrm{Cov}_1 + 2\mathrm{Cov}_2)$$
$$+ 3rpp_4(\mathrm{Cov}_3 + 2\mathrm{Cov}_4) + \frac{rpp_4p_5}{2}(4\mathrm{Cov}_5 + \mathrm{Cov}_6 + 4\mathrm{Cov}_7)$$
$$+ \frac{rpp_4p_5p_6}{2}\mathrm{Cov}_8.$$

The expectations of the sums of squares as partitioned in Table 3 may be derived as simple linear functions of the expectations presented in Table 4, i.e., $E(\Sigma\, a_i x_i) = \Sigma\, a_i\, E(x_i)$.

If we define $\underline{S}$ as the (9 x 1) vector of sum of squares as listed in Table 4, M as the (9 x 10) matrix of coefficients in Table 4, and $\underline{C}' = (r\mu^2 + \sigma_r^2, \sigma_e^2, \mathrm{Cov}_1, \mathrm{Cov}_2, \mathrm{Cov}_3, \mathrm{Cov}_4, \mathrm{Cov}_5, \mathrm{Cov}_6, \mathrm{Cov}_7, \mathrm{Cov}_8)$, then

$$E(\underline{S}) = M\underline{C}.$$

Further, if we define $\underline{S}_p' = (G, S, U, W, T, V, X)$ from Table 3 and N equal to the (7 x 9) transformation matrix such that $\underline{S}_p = N\underline{S}$ (i.e., N is the matrix of coefficients which transforms the uncorrected sums of squares, $\underline{S}$, into the partitioned sums of squares, $\underline{S}_p$), then

$$E(\underline{S}_p) = NE(\underline{S}) = (NM)\underline{C}$$

Each expectation will contain, in addition to the covariances, the degrees of freedom times σ_e^2. The terms μ^2 and σ_r^2 drop out of the expectations of all the hybrid sums of squares. The expected mean squares are then found by dividing the expected sums of squares by the degrees of freedom. The mean square expectations are shown in Table 5 in terms of σ_e^2 and eight functions of the covariances.

Table 5. Expectations of mean squares in terms of the covariances between hybrid relatives for the analysis of double-cross hybrids

M.S.	Coefficients of covariances							
	σ_e^2	$Cov_1 + 2Cov_2$	$Cov_1 - Cov_2$	$Cov_3 + 2Cov_4$	$Cov_3 - Cov_4$	$4Cov_5 + Cov_6 + 4Cov_7$	$Cov_5 + Cov_6 - 2Cov_7$	Cov_8
G*	1	r	0	r(3p-16)	0	$rp_5p_8/2$	0	$rp_5p_6p_{16}/2$
S*	1	r	0	$2rp_7$	0	$r(p^2-21p+92)/6$	0	$-3rp_6p_9$
U*	1	r	0	rp_{10}	0	$-rp_8$	0	3r(3p-22)
W*	1	r	0	-4r	0	2r	0	-12r
T*	1	0	r	0	$2rp_4$	0	$rp_4p_5/3$	0
V*	1	0	r	0	rp_7	0	$-rp_5$	0
X*	1	0	r	0	-4r	0	2r	0
E*	1	0	0	0	0	0	0	0

5.6 Model of Linear Effects

The development of a model of linear effects is not essential for the estimation of genetic parameters. The results of Section 5.3 (the genetic expectations of the covariances) and Section 5.5 (expectations of the sums of squares in terms of covariances) can be used directly to estimate the components of genetic variance. The development of a model of linear effects is necessary, however, for the development of tests of significance of various genetic hypotheses.

The effect model presented in this section is the join result of intuitive reasoning as to what would constitute a logical model and an attempt to make the linear model compatible with the previously defined covariances at least in their more apparent features. It is known, for example, on the basis of genetic expectations of the covariances that $Cov_6 = Cov_7 > Cov_8$ (Table 1). It would be most desirable that these covariances expressed in terms of the components of variance would follow the same pattern. With these and other considerations, the following is proposed as one possible model of linear effects:

$$Y_{(ij)(kl)m} = \mu + r_m + G_{(ij)(kl)} + e_{(ij)(kl)m} \tag{5.6.1}$$

$$\begin{aligned} &= \mu + r_m + g_i + g_j + g_k + g_l + s_{ij} + s_{ik} + s_{il} \\ &\quad + s_{jk} + s_{jl} + s_{kl} + t_{ik} + t_{il} + t_{jk} + t_{jl} + u_{ijk} \\ &\quad + u_{ijl} + u_{ikl} + u_{jkl} + v_{(ij)k} + v_{(ij)l} + v_{(kl)i} \\ &\quad + v_{(kl)j} + w_{ijkl} + x_{(ij)(kl)} + e_{(ij)(kl)m} \end{aligned} \tag{5.6.2}$$

where $Y_{(ij)(kl)m}$, μ, r_m, $G_{(ij)(kl)}$ and $e_{(ij)(kl)m}$ are as defined in Section 5.2 and where

g_i = a contribution particular to line i,

s_{ij} = a contribution particular to line i and line j appearing together as a single-cross parent,

$s_{ij} + t_{ij}$ = a contribution particular to line i and line j appearing in different parental single-crosses,

u_{ijk} = a contribution particular to lines i, j and k appearing together in any order.

$v_{(ij)k}$ = a contribution particular to lines i, j and k appearing in the order (ixj)k,

w_{ijkl} = a contribution particular to lines i, j, k and l appearing as the four grandparents in any order.

$x_{(ij)(kl)}$ = a contribution particular to lines i, j, k and l appearing together in the order (ixj)(kxl).

The terms g_i, s_{ij}, u_{ijk}, and w_{ijkl} will be referred to as one-, two-, three-, and four-line general effects, respectively, and t_{ij}, $v_{(ij)k}$, and $x_{(ij)(kl)}$ as two-, three-, and four-line order effects, respectively. It is assumed that effects g_i, s_{ij}, t_{ij}, u_{ijk}, and w_{ijkl} are independent random variables with means zero and variances σ_g^2, σ_s^2, σ_t^2, σ_u^2 and σ_w^2, respectively, and that they are uncorrelated with each other.

The effects $v_{(ij)k}$ and $x_{(ij)(kl)}$ may be treated under two different assumptions. In Model A, the effects $v_{(ij)k}$ and $x_{(ij)(kl)}$ are assumed to be random variables with means zero and variances σ_v^2 and σ_x^2, respectively. In Model B, the three 3-line order effects involving the same set of three lines and the three 4-line order effects involving the same four lines are assumed to sum to zero within each set;

that is,

$$v_{(ij)k} + v_{(ik)j} + v_{(jk)i} = 0 \text{ for all combinations of three lines}$$

and

$$x_{(ij)(kl)} + x_{(ik)(jl)} + x_{(il)(jk)} = 0 \text{ for all combinations of four lines.}$$

The variances of the effects $v_{(ij)k}$ and $x_{(ij)(kl)}$ are defined for Model B as

$$1/2\ E(v^2_{(ij)k} + v^2_{(ik)j} + v^2_{(jk)i}) = \sigma^2_{v_{ijk}} = \sigma^2_v$$

and

$$1/2\ E(x^2_{(ij)(kl)} + x^2_{(ik)(jl)} + x^2_{(il)(jk)}) = \sigma^2_{x_{ijkl}} = \sigma^2_x$$

or

$$E(v^2_{(ij)k}) = 2/3\ \sigma^2_v,\ E(v_{(ij)k}\ v_{(ik)j}) = -1/3\ \sigma^2_v,$$

$$E(x^2_{(ij)(kl)}) = 2/3\ \sigma^2_x, \text{ and } E(x_{(ij)(kl)}\ x_{(ik)(jl)}) = -1/3\ \sigma^2_x.$$

The covariances between three-line order effects not involving the same combination of three lines and the covariances between four-line order effects not involving the same combination of four lines are all assumed to be zero.

It should be noted that the two-line order effect, t_{ij}, could be handled under similar alternative models by defining two effects, $t_{(ij)}$ and $t_{(i)(j)}$, for example, such that $t_{(ij)} + 2t_{(i)(j)} = 0$. No arrangement of this system, however, was found which was compatible with the genetic expectations of the covariances.

5.7 Covariances as Functions of the Components of Variance

The covariances as defined in Section 5.2 may be expressed in terms of the components of variance as defined in Section 5.6 by taking the expectations with respect to the model of linear effects. These results are presented in Table 6 for Model A. $Cov_6 = Cov_7 > Cov_8$ in both Table 1 and Table 6. Hence, at least in this major respect, the linear model of effects is compatible with the genetic expectations of the covariances.

Table 6. Covariances between double-cross hybrid relatives expressed as functions of the components of variance

Covariance	Coefficients of components of variance						
	σ_g^2	σ_s^2	σ_t^2	σ_u^2	σ_v^2	σ_w^2	σ_x^2
Cov_1	4	6	4	4	4	1	1
Cov_2	4	6	2	4	0	1	0
Cov_3	3	3	2	1	1	0	0
Cov_4	3	3	1	1	0	0	0
Cov_5	2	1	1	0	0	0	0
Cov_6	2	1	0	0	0	0	0
Cov_7	2	1	0	0	0	0	0
Cov_8	1	0	0	0	0	0	0

5.8 Expectations of Mean Squares in Terms of Components of Variance and Estimation of Components of Variance

The expectations of the mean squares in terms of the components of variance are presented for Model A and Model B in Table 7 and Table 8, respectively. The derivation of the expectations of the mean squares are straight forward and are not presented here. For the completely

Table 7. Expectations of mean squares with Model A for the analysis of double-crosses

M.S.	Coefficients of components of variance								
	σ_e^2	σ_x^2	σ_v^2	σ_t^2	σ_w^2	σ_u^2	σ_s^2	σ_g^2	σ_r^2
R*	1	0	0	0	0	0	0	0	$3\ {}_pC_4$
G*	1	r	$3rp_4$	$2rp_3p_4$	3r	$9rp_4$	$9rp_3p_4/2$	$rp_2p_3p_4/2$	0
S*	1	r	$2rp_5$	$2rp_4p_5/3$	3r	$6rp_5$	$3rp_4p_5/2$	0	0
U*	1	r	rp_6	0	3r	$3rp_6$	0	0	0
W*	1	r	0	0	3r	0	0	0	0
T*	1	r	$2rp_2$	$rp_1p_2/3$	0	0	0	0	0
V*	1	r	rp_3	0	0	0	0	0	0
X*	1	r	0	0	0	0	0	0	0
E*	1	0	0	0	0	0	0	0	0

random model, Model A, the expectation of the mean squares can be derived directly by simple substitution of the results in Table 6, where the covariances are expressed as functions of the components, in the results in Table 5, where the expectations of mean squares are given in terms of covariances.

It is apparent from the expectations of the mean squares that exact tests of significance are available only for the hypotheses $\sigma_x^2 = 0$, $\sigma_w^2 = 0$ and $\sigma_v^2 = 0$ for Model A and $\sigma_x^{2'} = 0$, $\sigma_w^{2'} = 0$, $\sigma_v^{2'} = 0$ and $\sigma_u^{2'} = 0$ in Model B. Tests of significance of the remaining components are complex and are approximations. The appropriate F-tests and their interpretations are presented in Section 5.10.

Table 8. Expectations of mean squares with Model B for the analysis of double-crosses

M.S.	Coefficients of components of variance								
	$\sigma_e^{2'}$	$\sigma_x^{2'}$	$\sigma_v^{2'}$	$\sigma_t^{2'}$	$\sigma_w^{2'}$	$\sigma_u^{2'}$	$\sigma_s^{2'}$	$\sigma_g^{2'}$	$\sigma_r^{2'}$
R*	1	0	0	0	0	0	0	0	$3\ {}_pC_4$
G*	1	0	0	$2rp_3p_4$	3r	$9rp_4$	$9rp_3p_4/2$	$rp_2p_3p_4/2$	0
S*	1	0	0	$2rp_4p_5/3$	3r	$6rp_5$	$3rp_4p_5/2$	0	0
U*	1	0	0	0	3r	$3rp_6$	0	0	0
W*	1	0	0	0	3r	0	0	0	0
T*	1	r	$2rp_2$	$rp_1p_2/3$	0	0	0	0	0
V*	1	r	rp3	0	0	0	0	0	0
X*	1	r	0	0	0	0	0	0	0
E*	1	0	0	0	0	0	0	0	0

The formulae for the estimation of the various components of variance for the two models are as follows where the estimate of σ^2 is denoted by s^2. A prime is used on the component to indicate an estimate for Model B.

Model A:

$$s_x^2 = (X^* - E^*)/r$$

$$s_w^2 = (W^* - X^*)/3r$$

$$s_v^2 = (V^* - X^*)/rp_3$$

$$s_u^2 = [U^* - W^* - \frac{p_6}{p_3}(V^* - X^*)]/3rp_6$$

$$s_t^2 = 3[T^* - \frac{2p_2}{p_3}V^* + \frac{p_1}{p_3} X^*]/rp_1p_2$$

$$s_s^2 = 2[S^* - \frac{2p_5}{p_6} U^* + \frac{p_4}{p_6} W^* - \frac{2p_4p_5}{p_1p_2} T^* + \frac{4p_4p_5}{p_1p_3} V^*$$

$$- \frac{2p_4p_5}{p_2p_3} X^*]/3rp_4p_5$$

$$s_g^2 = 2[G^* - \frac{3p_3}{p_5} S^* + \frac{3p_2}{p_6} U^* - \frac{p_2p_3}{p_5p_6} W^*]/rp_2p_3p_4$$

Model B:

$$s_x^{2'} = s_x^2$$

$$s_w^{2'} = (W^* - E^*)/3r$$

$$s_v^{2'} = s_v^2$$

$$s_u^{2'} = (U^* - W^*)/3rp_6$$

$$s_t^{2'} = s_t^2$$

$$s_s^{2'} = s_s^2$$

$$s_g^{2'} = s_g^2$$

Only σ_w^2 and σ_u^2 require different estimators in the two models. Further discussion of the above components will be delayed until after the expectations of these components are given in terms of genetic parameters.

Since the analysis as presented is an orthogonal partitioning of the sums of squares, the mean squares are uncorrelated. If the observations are normally distributed, then each mean square is a sum of squared linear funtions of NID variates and is distributed as $\chi^2\sigma_i^2$ with f_i degrees of freedom. Hence, the variances of the mean squares are equal to $\frac{2\sigma_i^4}{f_i}$ and are estimated by $\left[\frac{2(MS_i)^2}{f_i + 2}\right]$ (Anderson and Bancroft, 1952) where

f_i indicates degrees of freedom for the ith mean square. The variance of the estimated components of variance may then be computed by applying the rules for finding the variance of a linear function. That is, the

$$V_{(\Sigma a_i X_i)} = \Sigma a_i^2 V_{(X_i)} + 2 \sum_{i<j}\sum a_i a_j \operatorname{Cov}(X_i X_j)$$

but since the mean squares are uncorrelated,

$$V_{(\Sigma a_i MS_i)} = \Sigma a_i^2 V_{(MS_i)} = 2 \Sigma \left[\frac{a_i^2 (MS_i)^2}{f_i + 2}\right].$$

For example,

$$V_{(s_x^2)} = \frac{1}{r^2}\left[\frac{2X*^2}{f_x + 2} + \frac{2E*^2}{f_e + 2}\right]$$

and

$$V_{(s_u^2)} = \frac{2}{9r^2 p_6^2}\left[\frac{U*^2}{f_u + 2} + \frac{W*^2}{f_w + 2} + \frac{p_6^2}{p_3^2}\left(\frac{V*^2}{f_v + 2} + \frac{X*^2}{f_x + 2}\right)\right]$$

5.9 Components of Variance as Functions of Covariances between Hybrid Relatives

The expectations of the mean squares have been derived in terms of the components of variance for Models A and B, Table 7 and Table 8, respectively, and in terms of the covariances between hybrid relatives, Table 5. From these results, the components of variance for the two models may be expressed as functions of the covariances between hybrid relatives.

Model A:

Let $\underline{S}_e' = (G* - E*, S* - E*, U* - E*, W* - E*, T* - E*, V* - E*, X* - E*)$,

$\underline{P}' = (\sigma_x^2, \sigma_v^2, \sigma_t^2, \sigma_w^2, \sigma_u^2, \sigma_s^2, \sigma_g^2)$, A = the 7 x 7 matrix of coefficients of the components of variance in the expected mean squares of Table 7, excluding R* and E* mean squares and the columns for σ_e^2 and σ_r^2,

B = the 7 x 7 matrix of the coefficients of the functions of covariances given in Table 5, excluding E* mean square and the column for σ_e^2, and

$\underline{C}_f' = (\text{Cov}_1 + 2\ \text{Cov}_2,\ \text{Cov}_1 - \text{Cov}_2,\ \text{Cov}_3 + 2\ \text{Cov}_4,\ \text{Cov}_3 - \text{Cov}_4,$
$4\ \text{Cov}_5 + \text{Cov}_6 + 4\ \text{Cov}_7,\ \text{Cov}_5 + \text{Cov}_6 - 2\ \text{Cov}_7,\ \text{Cov}_8)$.

Then, from Table 7, $E(\underline{S}_e) = A\underline{P}$ and, from Table 5, $E(\underline{S}_e) = B\underline{C}_f$. Hence, $A\underline{P} = B\underline{C}_f$ or $\underline{P} = A^{-1}B\underline{C}_f$. This solution expresses the components of variance in Model A as functions of the covariances between relatives (Table 9).

Table 9. Components of variance as functions of the covariances between double-cross hybrid relatives for Model A

Component of variance	Coefficients of covariances							
	Cov_1	Cov_2	Cov_3	Cov_4	Cov_5	Cov_6	Cov_7	Cov_8
σ_g^2	0	0	0	0	0	0	0	1
σ_s^2	0	0	0	0	0	-1/3	4/3	-2
σ_t^2	0	0	0	0	1	1	-2	0
σ_u^2	0	0	0	1	-1	0	-2	3
σ_v^2	0	0	1	-1	-1	-1	2	0
σ_w^2	0	1	0	-4	2	0	4	-4
σ_x^2	1	-1	-4	4	2	2	-4	0

Model B:

Let $\underset{\sim}{Q}' = (\sigma_x^{2'}, \sigma_v^{2'}, \sigma_t^{2'}, \sigma_w^{2'}, \sigma_u^{2'}, \sigma_s^{2'}, \sigma_g^{2'})$ and

D = the 7 x 7 matrix of coefficients of the components of variance in the expected mean squares in Table 8, excluding R* and E* mean squares and the columns for σ_e^2 and σ_r^2. Then, $E(\underset{\sim}{S}_e) = D\underset{\sim}{Q}$.

Hence, $D\underset{\sim}{Q} = B\underset{\sim}{C}_f$ or $\underset{\sim}{Q} = D^{-1}B\underset{\sim}{C}_f$. This solution is given in Table 10.

It is apparent from these results that $\sigma_g^2 = \sigma_g^{2'}$, $\sigma_s^2 = \sigma_s^{2'}$, $\sigma_t^2 = \sigma_t^{2'}$, $\sigma_v^2 = \sigma_v^{2'}$ and $\sigma_x^2 = \sigma_x^{2'}$ but $\sigma_u^2 \neq \sigma_u^{2'}$ and $\sigma_w^2 \neq \sigma_w^{2'}$. Hence, changing the definition of the three- and four-line order effects from random to fixed, i.e., Model A to Model B, changes the interpretations of only σ_u^2 and σ_w^2.

Table 10. Components of variance as functions of the covariances between double-cross hybrid relatives for Model B

Component of variance	Coefficients of covariances							
	Cov_1	Cov_2	Cov_3	Cov_4	Cov_5	Cov_6	Cov_7	Cov_8
$\sigma_g^{2'}$	0	0	0	0	0	0	0	1
$\sigma_s^{2'}$	0	0	0	0	0	-1/3	4/3	-2
$\sigma_t^{2'}$	0	0	0	0	1	1	-2	0
$\sigma_u^{2'}$	0	0	1/3	2/3	-4/3	-1/3	-4/3	3
$\sigma_v^{2'}$	0	0	1	-1	-1	-1	2	0
$\sigma_w^{2'}$	1/3	2/3	-4/3	-8/3	8/3	2/3	8/3	-4
$\sigma_x^{2'}$	1	-1	-4	4	2	2	-4	0

5.10 Genetic Interpretations of the Components of Variance and Tests of Significance

The covariances between double-cross hybrid relatives can be expressed in terms of the genetic components of variance by function (5.3.1) where the values of p and q in the function are given in Table 1 for each of the covariances. The genetic interpretations of the components of variance can then be derived by applying (5.3.1) to the results of Section 5.5 (Table 9, Table 10). For example,

$$\sigma_g^2 = \sigma_g^{2'} = \mathrm{Cov}_8 = \sum_{\substack{A,D=0 \\ 1 \le A+D \le n}}^{n} \left(\frac{F}{8}\right)^A (0)^D \sigma_{AD}^2$$

$$= \frac{F}{8}\sigma_{10}^2 + \frac{F^2}{64}\sigma_{20}^2 + \frac{F^3}{512}\sigma_{30}^2 + \ldots$$

and

$$\sigma_t^2 = \sigma_t^{2'} = \mathrm{Cov}_5 + \mathrm{Cov}_6 - 2\,\mathrm{Cov}_7 = \mathrm{Cov}_5 - \mathrm{Cov}_7$$

$$= \Sigma\left(\frac{F}{4}\right)^A\left(\frac{F}{16}\right)^D \sigma_{AD}^2 - \Sigma\left(\frac{F}{4}\right)^A (0)^D \sigma_{AD}^2 = \sum_{D\neq 0} \left(\frac{F}{4}\right)^A\left(\frac{F}{16}\right)^D \sigma_{AD}^2$$

$$= \frac{F^2}{16}\sigma_{01}^2 + \frac{F^3}{64}\sigma_{11}^2 + \frac{F^4}{256}\sigma_{02}^2 + \ldots .$$

The expressions of the components of variance in terms of one-, two-, and three-factor genetic components of variance are presented in Table 11.

The genetic interpretations of the effect components of variance are apparent from Table 11. The genetic expectations of σ_s^2, σ_u^2, σ_v^2, σ_w^2 and σ_x^2 involve only epistatic components of genetic variance. $\sigma_x^2 = \sigma_x^{2'}$ is a function of the dominance x dominance epistatic component of genetic variance and all three-factor or higher epistatic components except the all-additive types. Hence, a test of the hypothesis that

Table 11. **Components of variance in the analysis of double-crosses as functions of the components of genetic variance up to and including the three-factor epistatic components**

Component	Coefficients of components of genetic variance								
	σ^2_{10}	σ^2_{01}	σ^2_{20}	σ^2_{11}	σ^2_{02}	σ^2_{30}	σ^2_{21}	σ^2_{12}	σ^2_{03}
$\sigma^2_g = \sigma^{2'}_g$	$F/8$	0	$F^2/64$	0	0	$F^3/512$	0	0	0
$\sigma^2_s = \sigma^{2'}_s$	0	0	$F^2/64$	0	0	$3F^3/256$	0	0	0
$\sigma^2_t = \sigma^{2'}_t$	0	$F^2/16$	0	$F^3/64$	$F^4/256$	0	$F^4/256$	$F^5/1024$	$F^6/4096$
σ^2_u	0	0	0	$F^3/128$	0	$3F^3/256$	$5F^4/1024$	$F^5/2048$	0
$\sigma^{2'}_u$	0	0	0	$F^3/96$	$F^4/384$	$3F^3/256$	$5F^4/768$	$5F^5/3072$	$F^6/2048$
$\sigma^2_v = \sigma^{2'}_v$	0	0	0	$F^3/128$	$F^4/128$	0	$5F^4/1024$	$7F^5/2048$	$3F^6/2048$
σ^2_w	0	0	0	0	$F^4/128$	0	$F^4/256$	$F^5/256$	$3F^6/2048$
$\sigma^{2'}_w$	0	0	0	0	$F^4/96$	0	$F^4/192$	$5F^5/768$	$F^6/256$
$\sigma^2_x = \sigma^{2'}_x$	0	0	0	0	$F^4/128$	0	$F^4/256$	$F^5/128$	$15F^6/2048$

$\sigma_x^2 = 0$ is a test of the hypothesis that this particular function of the epistatic variances equals zero. Similarly, the tests that $\sigma_w^2 = 0$ and $\sigma_w^{2'} = 0$ are tests that a slightly different function of the same epistatic variances equal zero. The functions for σ_w^2 and $\sigma_w^{2'}$ place relatively less emphasis on the higher-factor components than the function for σ_x^2.

The test of the hypothesis $\sigma_v^2 = \sigma_v^{2'} = 0$ is a test that a function involving additive x dominance, dominance x dominance, and all higher-factor epistatic variances except the all-additive types is equal to zero. $\sigma_u^{2'}$ is a function of all epistatic components of variance except additive x additive epistatic variance and a test of $\sigma_u^{2'} = 0$ is a test of these components. The test of $\sigma_u^2 = 0$ is a test of additive x dominance epistatic variance and all higher-factor epistatic variances except the all-dominance types. The test of $\sigma_s^2 = \sigma_s^{2'} = 0$ is a test of only the all-additive types of epistasis. The test of $\sigma_t^2 = \sigma_t^{2'} = 0$ is a test of dominance variance and all epistatic variances involving dominance in their nomenclature and the test of $\sigma_g^2 = \sigma_g^{2'} = 0$ is a test of additive genetic variance plus the all-additive types of epistatic variance.

Thus, testing the hypotheses that the various effect components of variance are zero will provide considerable information on the gene action involved. However, as indicated in Section 5.2, simple F-tests are possible only for the hypotheses concerning σ_x^2, σ_v^2, and σ_w^2 in Model A and $\sigma_x^{2'}$, $\sigma_v^{2'}$, $\sigma_w^{2'}$, and $\sigma_u^{2'}$ in Model B unless some of these components can be assumed equal to zero, either on the basis of the F-test or from other genetic information. If some can be assumed equal to zero, then exact F-tests will be available for more of the remaining components of variance.

Satterthwaite (1946) suggested that a linear function of mean squares, $\Sigma(a_i MS_i)$, is approximately distributed as $\chi^2 \sigma^2/f'$ with f' degrees of freedom where

$$f' = (\Sigma a_i MS_i)^2/\Sigma(a_i^2 MS_i^2/f_i) \ . \tag{5.10.1}$$

The degrees of freedom for mean square MS_i is denoted by f_i. With the use of Satterthwaite's approximation, two different approximate tests have been proposed for the cases where no simple error term is available. One procedure is to construct an error term with the proper expectation as linear function of mean squares and assume it is distributed as $\chi^2\sigma^2/f'$ with f' degrees of freedom. Satterthwaite, however, has indicated that caution should be exercised when using this approximation if the linear function of mean squares involves some negative a_i. Cochran (1951) suggested an F-test whereby negative a_i can be avoided by adding all negative terms to both numerator and denominator of the F-ratio involving the constructed error term. This ratio is then assumed to follow the F-distribution with f_1' and f_2' degrees of freedom. Anderson and Bancroft (1952) indicated that Cochran's test is more powerful than the first test but whether or not it is enough better to offset the need for computing two estimated degrees of freedom instead of one is doubtful. The exact and approximate F-tests are presented in Table 12 for the hypotheses that the various components of variance equal zero. $\bar{F}$ is used to indicate an F-test to avoid confusion with F used to indicate the inbreeding coefficient. $\bar{F}'$ indicates an approximate test formed by constructing an appropriate error term and $\bar{F}''$ indicates an approximate test using Cochran's suggestion.

A much simpler system of tests, which are valid F-tests, are available but they will not generally yield as much information concerning

Table 12. F-tests of hypotheses that various components of variance are zero in the analysis of double-crosses

Component	F-test	Degrees of freedom[a]
σ_x^2, $\sigma_x^{2'}$	$\bar{F} = X^*/E^*$	$pp_1p_4p_5/12$, $(r-1)(3\ {}_pC_4 - 1)$
σ_w^2	$\bar{F} = W^*/X^*$	$pp_1p_2p_7/24$, $pp_1p_4p_5/12$
$\sigma_w^{2'}$	$\bar{F} = W^*/E^*$	$pp_1p_2p_7/24$, $(r-1)(3\ {}_pC_4 - 1)$
σ_v^2, $\sigma_v^{2'}$	$\bar{F} = V^*/X^*$	$pp_2p_4/3$, $pp_1p_4p_5/12$
σ_u^2	$\bar{F}' = \dfrac{U^*}{W^* + \dfrac{p_6}{p_3}(V^* - X^*)}$	$pp_1p_5/6$, f'
	$\bar{F}'' = \dfrac{U^* + \dfrac{p_6}{p_3}X^*}{W^* + \dfrac{p_6}{p_3}V^*}$	f_1'', f_2''
$\sigma_u^{2'}$	$\bar{F} = U^*/W^*$	$pp_1p_5/6$, $pp_1p_2p_7/24$
σ_t^2, $\sigma_t^{2'}$	$\bar{F}' = \dfrac{T^*}{\dfrac{2p_2}{p_3}V^* - \dfrac{p_1}{p_3}X^*}$	$pp_3/2$, f'

Table 12(continued)

Component	F-test	Degrees of freedom[a]
$\sigma_t^2,\ \sigma_t^{2'}$	$\bar{F}''= \dfrac{T^* + \frac{p_1}{p_3} X^*}{\frac{2p_2}{p_3} V^*}$	$f''_1,\ f''_2$
$\sigma_s^2,\ \sigma_s^{2'}$	$\bar{F}'= \dfrac{S^*}{\frac{2p_5}{p_6} U^* - \frac{p_4}{p_6} W^* + \frac{2p_4p_5}{p_1p_2} T^* - \frac{4p_4p_5}{p_1p_3} V^* + \frac{2p_4p_5}{p_2p_3} X^*}$	$pp_3/2,\ f'$
	$\bar{F}''= \dfrac{S^* + \frac{p_4}{p_6} W^* + \frac{4p_4p_5}{p_1p_3} V^*}{\frac{2p_5}{p_6} U^* + \frac{2p_4p_5}{p_1p_2} T^* + \frac{2p_4p_5}{p_2p_3} X^*}$	$f''_1,\ f''_2$
$\sigma_g^2,\ \sigma_g^{2'}$	$\bar{F}'= \dfrac{G^*}{\frac{3p_3}{p_5} S^* - \frac{3p_2}{p_6} U^* + \frac{p_2p_3}{p_5p_6} W^*}$	$p_1,\ f'$
	$\bar{F}''= \dfrac{G^* + \frac{3p_2}{p_6} U^*}{\frac{3p_3}{p_5} S^* + \frac{p_2p_3}{p_5p_6} W^*}$	$f''_1,\ f''_2$

[a] f' and f" indicate approximate degrees of freedom as suggested by Satterthwaite (1946)(see equation 5.10.1)

gene action as the tests indicated in Table 12. The error mean square for the experiment is a valid error term for testing each of the entry mean squares. The hypotheses being tested, however, are the composite hypotheses that all genetic components of variance appearing in the expectation of the respective mean squares are equal to zero. Since the genetic expectations of the mean squares are independent of the effect model, the interpretations of the F-tests using the common error term are the same for both models. The tests X*/E*, W*/E*, V*/E* and T*/E* have essentially the same interpretation as the F-tests of the corresponding components of variance σ_x^2, σ_w^2, σ_v^2, and σ_t^2, respectively. $\bar{\bar{F}}$ = U*/E* is, however, a test of all the components of epistatic variance except the additive x additive component. $\bar{\bar{F}}$ = S*/E* is a test of dominance variance and all epistatic components and $\bar{\bar{F}}$ = G*/E* is a test of all components of genetic variance. The ability to test only the all-additive components of epistatic variance and to test the additive genetic variance plus only the all-additive epistatic components is lost with this procedure of testing.

The most powerful test of dominance and epistasis is

$$\bar{\bar{F}} = \frac{8(H - G)/p(p_1p_2p_3 - 8)}{E^*}$$

with $p(p_1p_2p_3 - 8)$ and $(r-1)(3\ {}_pC_4 - 1)$ degrees of freedom. The simplest overall test of epistasis is

$$\bar{\bar{F}} = \frac{8(U + W + V + X)/pp_2p_5(p + 1)}{E^*}$$

with $pp_2p_5(p + 1)/8$ and $(r-1)(3\ {}_pC_4 - 1)$ degrees of freedom. This test, however, excludes the effects of two-factor additive x additive epistasis. In order to include these effects in an all-epistatic test, the test would have to be a complex test.

5.11 Estimation of Genetic Parameters

Equating the estimates of the components of variance to their expectations in terms of components of genetic variance gives a set of seven equations in $\frac{n(n+3)}{2}$ unknowns, σ^2_{10}, σ^2_{01}, σ^2_{20}, ..., where n is the number of loci involved (Table 9). It is necessary to assume a restricted genetic model in order that the number of unknowns be reduced to an estimable number. We may assume (1) additive and dominance effects but complete independence between loci, i.e., $\sigma^2_G = \sigma^2_{10} + \sigma^2_{01}$, where σ^2_G = total genetic variance, (2) additive and dominance effects and only dual-factor epistatic effects, i.e., $\sigma^2_G = \sigma^2_{10} + \sigma^2_{01} + \sigma^2_{20} + \sigma^2_{11} + \sigma^2_{02}$, or (3) additive and dominance effects, dual-factor epistatic effects and the 3-factor additive x additive x additive epistatic effects, i.e., $\sigma^2_G = \sigma^2_{10} + \sigma^2_{01} + \sigma^2_{20} + \sigma^2_{11} + \sigma^2_{02} + \sigma^2_{30}$. Unless our assumed genetic model includes σ^2_{12}, σ^2_{03} or 4-factor or higher epistatic parameters, there are only six independent equations since $E(\sigma^2_w) = E(\sigma^2_x) = kE(\sigma^{2'}_w)$ for (1), (2), and (3) above. Since the expectations of the mean squares are independent of the effect model, the same estimates will be obtained using either effect Model A or B. The results from Model A will be used to derive the estimators for the six genetic variances

$$\sigma^2_G = \sigma^2_{10} + \sigma^2_{01} + \sigma^2_{20} + \sigma^2_{11} + \sigma^2_{02} + \sigma^2_{30},$$

which assumes no epistatic effects involving more than two loci other than additive x additive x additive effects. The seven equations in six unknowns to be solved are shown in Table 11 by deleting the last three columns and substituting estimates for the components and for the genetic parameters.

Two estimates of σ^2_{02} are available from

$$s^2_w = \frac{F^4}{128} s^2_{02} \text{ and } s^2_x = \frac{F^4}{128} s^2_{02}$$

Hence, σ^2_{02} may be estimated by using either a weighted or unweighted mean of the two components. Using an unweighted mean, $\underline{P}_1 = F\underline{G}$ or $G = F^{-1}\underline{P}_1$, where

$$\underline{G}' = (s^2_{10}, s^2_{01}, s^2_{20}, s^2_{11}, s^2_{02}, s^2_{30}),$$

$$\underline{P}'_1 = [s^2_g, s^2_s, s^2_t, s^2_u, s^2_v, (s^2_w + s^2_x)/2],$$

and

$$F = \begin{bmatrix} F/8 & 0 & F^2/64 & 0 & 0 & F^3/512 \\ 0 & 0 & F^2/64 & 0 & 0 & 3F^3/152 \\ 0 & F^2/16 & 0 & F^3/64 & F^4/256 & 0 \\ 0 & 0 & 0 & F^3/128 & 0 & 3F^3/256 \\ 0 & 0 & 0 & F^3/128 & F^4/128 & 0 \\ 0 & 0 & 0 & 0 & F^4/128 & 0 \end{bmatrix}$$

from Table 9. Then

$$F^{-1} = \begin{bmatrix} 8/F & -8/F & 0 & 8/3F & -8/3F & 8/3F \\ 0 & 0 & 16/F^2 & 0 & -32/F^2 & 24/F^2 \\ 0 & 64/F^2 & 0 & -32/F^2 & 32/F^2 & -32/F^2 \\ 0 & 0 & 0 & 0 & 128/F^3 & -128/F^3 \\ 0 & 0 & 0 & 0 & 0 & 128/F^4 \\ 0 & 0 & 0 & 256/3F^3 & -256/3F^3 & 256/3F^3 \end{bmatrix}$$

and the estimation equations are as follows:

$$s^2_{10} = \frac{8}{3F} [3s^2_g - 3s^2_s + s^2_u - s^2_v + (\frac{s^2_w + s^2_x}{2})]$$

$$s_{01}^2 = \frac{8}{F^2}\left[2s_t^2 - 4s_v^2 + 3\left(\frac{s_w^2 + s_s^2}{2}\right)\right]$$

$$s_{20}^2 = \frac{32}{F^2}\left[2s_s^2 - s_u^2 + s_v^2 - \left(\frac{s_w^2 + s_x^2}{2}\right)\right]$$

$$s_{11}^2 = \frac{128}{F^3}\left[s_v^2 - \left(\frac{s_w^2 + s_x^2}{2}\right)\right]$$

$$s_{02}^2 = \frac{128}{F^4}\left[\left(\frac{s_w^2 + s_x^2}{2}\right)\right]$$

$$s_{30}^2 = \frac{256}{3F^3}\left[s_u^2 - s_v^2 + \left(\frac{s_w^2 + s_x^2}{2}\right)\right]$$

Using a weighted mean of s_w^2 and s_x^2 yields a similar set of equations with the weighted mean substitued for $(s_w^2 + s_x^2)/2$. The variances for the estimates of genetic variances may be found by applying the rules for variances of linear functions, i.e.,

$$V(\Sigma a_i x_i) = \Sigma a_i^2\, V(x_i) + 2\sum_{i<j}\Sigma a_i a_j\, \mathrm{Cov}(x_i x_j)$$

For example,

$$V(s_{02}^2) = \frac{(128)^2}{4F^8}\left[V_{(s_w^2)} + V_{(s_x^2)} + 2\mathrm{Cov}_{(s_w^2 \cdot s_x^2)}\right]$$

It must be remembered that although the individual mean squares are uncorrelated, many of the functions of mean squares estimating the components of variance are correlated. Hence, the covariances between estimates of variance components must be considered in deriving the variances of the estimates of the genetic components of variance.

5.12 Discussion of the Analysis of Double-Crosses

The analysis of variance for double-crosses is parallel in many respects to the analysis of a p^4 factorial experiment. The four factors in this case are the four positions a line may take in the pedigree of a double-cross and the p "levels" are the p parental lines. This degenerates into a partial factorial because of the restriction that a particular line can occur only once in a particular double-cross. There are only 3 ${}_pC_4$ combinations instead of p^4 as there would be in a complete factorial. The effects of the factors are not distinguishable in the double-cross analysis so that the one-line general sum of squares is comparable to the sum of squares due to the pooled main effects in a factorial. Similarly, the two-, three- and four-line general sums of squares are comparable to the sums of squares due to the pooled two-, three- and four-factor interaction effects. The order sums of squares do not have parallels in a factorial analysis. They are new features which depend upon the order in which the four lines are combined.

There is lack of complete balance in the double-cross analysis in that a particular line cannot appear more than once in a particular pedigree. There is, however, balance in the sense that every line appears an equal number of times with every other line. There is also a certain symmetry about the way the lines are combined in the hybrid. This balance and symmetry have simplified the development of a model of linear effects which is compatible with the genetic expectations of the covariances between hybrid relatives.

Considerable information concerning gene action can be obtained from the analysis of double-cross hybrids under the assumptions as stated. Direct F-tests are available for different fractions of the genetic variance including several all epistatic fractions. A qualitative comparison of the relative importance of the different types of genetic variance may be obtained from these tests. F-tests are available for the detection of dominance plus epistatic variance and for the detection of all components of epistatic variance except the additive x additive component. A complex function is necessary in order to test all components of epistatic variance with a single F-test. A quantitative estimate of additive genetic variance, dominance variance and four epistatic components of genetic variance can be obtained as shown in Section 5.11 if all other epistatic components are assumed to be equal to zero. The bias, if this assumption is not valid, is not likely to be large, if the four epistatic parameters estimated are the three two-factor components and the three-factor additive x additive x additive component, since the coefficients of the higher order components of epistatic variance in most cases decrease in size rapidly as order increases. The relatively small size of the coefficients of the genetic variances in Table 9 may cause some difficulty in obtaining enough precision to detect two-factor epistasis, even if it is an important part of total genetic variance.

The variances due to the interaction of the various hybrid effects with replications have been assumed to be equal in this analysis and have been pooled into a single error sum of squares. It is possible and may be desirable, if there are reasons to believe inequality exists, to partition the error sum of squares into interaction sums of squares corresponding to the partitioning of the hybrid sums of squares.

The analysis of variance for double-cross hybrids as presented was for r replications in one environment. This can readily be extended to include tests over years and locations with a more complex model. It would be desirable in the extended analysis to partition the sums of squares due to the interaction of hybrids with years, locations and years x locations into the sums of squares due to the interaction of each of the separate effects defined in the linear model with years, locations and years x locations. These interaction components may be different. If so, they will have different biological interpretations as they do in the analysis of single-cross hybrids (Matzinger, et al., 1959).

A serious limitation on the practical usefulness of this analysis is the rapidity with which the number of double-crosses increases as p increases. The degrees of freedom indicate that $p = 8$ is a minimum number of parents which can be used for a complete analysis since there will not be any degrees of freedom with which to estimate W* if $p \leq 7$. Thus, 210 entries, $p = 8$, are required for a complete analysis. If $p = 9$, there are 378 double-crosses and if $p = 10$, there are 630 double-crosses. From a practical standpoint, $p = 8$ or $p = 9$ are probably upper bounds on the number of parents for the complete analysis of double-crosses.

6.0 ANALYSIS OF VARIANCE FOR THREE-WAY HYBRIDS

6.1 Introduction

The term three-way hybrid is used herein to denote the hybrid progeny resulting from the cross of a line with the F_1 of a cross of two other lines and may be symbolized by A(B x C). Omitting reciprocal crosses, there are 3 ${}_pC_3$ different three-way hybrids possible from a set of p lines. Any one of the three parents in a particular combination of three lines may appear as the single parent in the hybrid so there are three order hybrids for each combination of lines, For example, the three orders or three-way hybrids for the lines A, B, and C are A(B x C), B(A x C) and C(A x B).

The analysis of variance presented is of the complete set of all three-way hybrids from a set of p random lines. The p lines are, as in the double-cross analysis, assumed to be a random sample of lines derived from a random mating population. It is assumed that all lines have an equal but arbitrary degree of inbreeding.

6.2 Expectation of Quadratic Forms in Terms of Covariances

The phenotype of the three-way hybrid A(B x C) grown in the mth replication in one environment will be denoted in the following derivations by $Y_{A(BC)m}$ or, in general, $Y_{i(jk)m}$. The subscripts within the parentheses denote the lines which form the single-cross parent and the single subscript before the parentheses denotes the single parent in the three-way cross.

The model for the analysis of variance is

$$Y_{i(jk)m} = \mu + r_m + G_{i(jk)} + e_{i(jk)m} \qquad (6.2.1)$$

where

μ = a contribution common to all entries,

r_m = a contribution particular to replication m,

$G_{i(jk)}$ = a contribution particular to three-way cross i(j x k),

and $e_{i(jk)m}$ = a random deviation of the observation on the three-way cross i(j x k) grown in replication m.

It is assumed that the $e_{i(jk)m}$'s, r_m's and $G_{i(jk)}$'s are independent random variables with zero means and variances σ_e^2, σ_r^2 and σ_y^2, respectively, and that the e's, r's and G's are not correlated with each other. A constant number of individuals in each hybrid progeny in each replication is assumed.

The expectations of the quadratic forms which will be encountered can be defined in terms of μ^2, σ_e^2, σ_r^2 and the covariances between the effects of the three-way hybrids, i.e., the $G_{i(jk)}$'s. The expectation of the square of the phenotype of a hybrid grown in replication m will contain $\mu^2 + \sigma_r^2 + \sigma_e^2$ plus the covariance of the effect of the three-way hybrid with itself. The expectation of the product of the phenotypes of two different hybrids grown in the same replication will contain $\mu^2 + \sigma_r^2$ plus the covariance of the effects of the two three-way hybrids. The expectation of the product of the phenotypes of the same or different hybrids grown in different replications will contain μ^2 plus the covariance of the effects of the two hybrids involved.

The covariances between the effects of three-way crosses are defined as listed in Table 13. It will be noted that:

Cov_1 = the covariance of the effect of a three-way cross with itself,

Table 13. Correlations between additive deviations, p, and dominance deviations, q, of three-way hybrid relatives[a]

Covariance	Number of lines in common	Correlations Inbreeding = F p	q	F = 1 p	q
$Cov_1 = Cov\ [G_{i(jk)}G_{i(jk)}]$	3	3F/4	$F^2/2$	3/4	1/2
$Cov_2 = Cov\ [G_{i(jk)}G_{j(ik)}]$	3	5F/8	$F^2/4$	5/8	1/4
$Cov_3 = Cov\ [G_{i(jk)}G_{i(jl)}]$	2	5F/8	$F^2/4$	5/8	1/4
$Cov_4 = Cov\ [G_{i(jk)}G_{j(il)}]$	2	F/2	$F^2/4$	1/2	1/4
$Cov_5 = Cov\ [G_{i(jk)}G_{l(ij)}]$	2	3F/8	0	3/8	0
$Cov_6 = Cov\ [G_{i(jk)}G_{l(jk)}]$	2	F/4	0	1/4	0
$Cov_7 = Cov\ [G_{i(jk)}G_{i(lm)}]$	1	F/2	0	1/2	0
$Cov_8 = Cov\ [G_{i(jk)}G_{j(lm)}]$	1	F/4	0	1/4	0
$Cov_9 = Cov\ [G_{i(jk)}G_{l(jm)}]$	1	F/8	0	1/8	0
$Cov\ [G_{i(jk)}G_{l(mn)}]$	0	0	0	0	0

[a]From Cockerham, C. C. (1957) (see footnote 2, Section 5.3).

Cov_2 = the covariance between the effects of two three-way hybrids with all three lines common but in different orders,

Cov_3 = the covariance between the effects of two three-way hybrids with two lines common and with the same line as the single parent in both hybrids,

Cov_4 = the covariance between the effects of two three-way hybrids with two lines common and with a different line as the single parent in the two hybrids,

Cov_5 = the covariance between the effects of two three-way hybrids with two lines common and with one of the lines the single parent in only one of the hybrids,

Cov_6 = the covariance between the effects of two three-way hybrids with two lines common and with both the lines as a single-cross in both hybrids,

Cov_7 = the covariance between the effects of two three-way hybrids with one line common as the single parent in both hybrids,

Cov_8 = the covariance between the effects of two three-way hybrids with one line common but with the common line as the single parent in only one hybrid and

Cov_9 = the covariance between the effects of two three-way hybrids with one line common in the single-cross parent in both hybrids.

The covariance between the effects of two three-way hybrids with no lines common is zero.

6.3 Genetic Expectations of Covariances

In order to express the covariances between three-way hybrid relatives in terms of the genetic parameters as defined by Cockerham (1954), or Kempthorne (1954), it is necessary to impose the same assumptions as given in Section 5.3 for the covariances of related double-cross hybrids. With these assumptions, the covariances between the effects of relatives, if the environments are not correlated, can be expressed by equation 5.3.1. The correlations between the additive and dominance deviations, p and q, respectively, in equation 5.3.1, were derived by Cockerham (see footnote 2, Section 5.3) for various hybrid relatives and are presented in Table 13 for three-way hybrid relatives.

6.4 Partitioning of Sums of Squares

The notation for the observations and various sums which will be encountered in partitioning the sums of squares are defined as follows:

$Y_{i(jk)m}$ = the phenotypic observation on three-way cross $i(j \times k)$ grown in replication m. $m = 1,2,\ldots,r$; $i, j, k = 1,2,\ldots, p$; $i \neq j \neq k$.

$Y_{i(jk).} = \sum_{m=1}^{r} Y_{i(jk)m}$ = sum over replications of the three-way hybrid $i(j \times k)$.

$Y_{ijk.} = \sum_{o}^{3} Y_{i(jk).}$ = sum of the three orders of hybrids involving the combination of lines i, j, and k.

$Y_{i(j.).} = \sum_{k\neq i,j}^{p_2} Y_{i(jk).}$ = sum of all hybrids in which line i appears as the single parent and line j appears in the single-cross parent.

$Y_{.(ij).} = \sum_{k\neq i,j}^{p_2} Y_{k(ij).}$ = sum of all hybrids in which lines i and j appear in the single-cross parent.

$Y_{ij..} = Y_{i(j.).} + Y_{j(i.).} + Y_{.(ij).} = \sum_{k\neq i,j}^{p_2} Y_{ijk.}$ = sum of all hybrids in which lines i and j appear in any order.

$$Y_{i(..).} = \frac{1}{2}\sum_{j\neq i}^{p_1} Y_{i(j.).} = \sum_{\substack{j,k\neq i\\ j<k}}^{p_1 C_2} Y_{i(jk).} = \text{sum of all hybrids in}$$

which line i appears as the single parent.

$$Y_{.(i.).} = \sum_{j\neq i}^{p_1} Y_{.(ij).} = \sum_{\substack{j,k\neq i\\ j<k}}^{p_1 C_2} \sum_{o}^{2} Y_{j(ik).} = \text{sum of all hybrids}$$

in which line i appears in the single-cross parent.

$$Y_{i...} = Y_{i(..).} + Y_{.(i.).} = \frac{1}{2}\sum_{j\neq i}^{p_1} Y_{ij..} = \sum_{\substack{j,k\neq i\\ j<k}}^{p_1 C_2} Y_{ijk.}$$

= sum of all hybrids in which line i appears.

$$Y_{....} = \frac{1}{3}\sum_{i}^{p} Y_{i...} = \sum_{i<j<k}^{p C_3} Y_{ijk.} = \text{grand total of all}$$

observations.

Other pertinent facts regarding these sums are given in Table 14.

One possible partitioning of the sums of squares from all possible three-way hybrids from a set of p parents grown in r replications in one environment is shown in Table 15. The sums of squares for replications, hybrids, and error are the ordinary least squares partitions for these effects using Model (6.2.1) and are orthogonal. The partitioning of the hybrid sum of squares is the result of largely trial and error procedures. The partitions are, however, an orthogonal set. The proof of orthogonality is long and involved and is not presented in this thesis.

This system of partitioning appears logical and closely parallels the double-cross partitioning with two exceptions. First, it is possible

Table 14. Description of sums pertinent to the analysis of three-way hybrids

Sum	Number of observations/sum	Number of such sums	Total of all such sums equals
$Y_{i(jk)m}$	1	$\frac{rpp_1p_2}{2}$	$Y_{....}$
$Y_{i(jk).}$	r	$\frac{pp_1p_2}{2}$	$Y_{....}$
$Y_{ijk.}$	3r	$\frac{pp_1p_2}{6}$	$Y_{....}$
$Y_{i(j.).}$	rp_2	pp_1	$2Y_{....}$
$Y_{.(ij).}$	rp_2	$\frac{pp_1}{2}$	$Y_{....}$
$Y_{ij..}$	$3rp_2$	$\frac{pp_1}{2}$	$3Y_{....}$
$Y_{i(..).}$	$\frac{rp_1p_2}{2}$	p	$Y_{....}$
$Y_{.(i.).}$	rp_1p_2	p	$2Y_{....}$
$Y_{i...}$	$\frac{3rp_1p_2}{2}$	p	$3Y_{....}$
$Y_{....}$	$\frac{rpp_1p_2}{2}$	1	$Y_{....}$

in the three-way partitioning to identify two different types of one-line sums, i.e., $Y_{i(..).}$ and $Y_{.(i.).}$, giving rise to one-line general and one-line order sums of squares. This was not possible in the double-cross analysis. Secondly, in the three-way analysis the two-line order sums $Y_{i(j.).}$ and $Y_{j(i.).}$ are distinct sums and give rise to an

Table 15. Analysis of variance of three-way hybrids

Source	d.f.	Sums of squares	M.S.
Correction factor	1	$C = \frac{2Y^2_{....}}{rpp_1p_2}$	C
Replications	$(r-1)$	$R = \frac{2\sum_{m=1}^{r} Y^2_{...m}}{pp_1p_2} - C$	R*
Hybrids	$(3\ {}_pC_3 - 1)$	$A = \frac{\sum_{i<j<k}^{{}_pC_3} \sum_{o}^{3} Y^2_{i(jk).}}{r} - C$	A*
1-line general	p_1	$G = \frac{2\sum_{i}^{p} Y^2_{i...}}{3rp_2p_3} - \frac{3p_1}{p_3} C$	G*
2-line general	$pp_3/2$	$S = \frac{\sum_{i<j}^{{}_pC_2} Y^2_{ij..}}{3rp_4} - \frac{3p_2}{p_4} C - \frac{2p_3}{p_4} G$	S*
3-line general	$pp_1p_5/6$	$U = \frac{\sum_{i<j<k}^{{}_pC_3} Y^2_{ijk.}}{3r} - C - G - S$	U*

Table 15(continued)

Source	d.f.	Sums of squares	M.S.
1-line order	p_1	$H = \dfrac{2\sum\limits_{i}^{p} Y^2_{i(..).} + \sum\limits_{i}^{p} Y^2_{.(i.)}}{rpp_2} - \dfrac{2\sum\limits_{i}^{p} Y^2_{i...}}{3rpp_2}$	H*
2-line order			
$Y_{.(ij).}$ vs. $Y_{i(j.).} + Y_{j(i.).}$	$pp_3/2$	$T = \dfrac{\sum\limits_{i<j}^{{}_pC_2} [Y_{i(j.).} + Y_{j(i.)}]^2}{2rp_1} + \dfrac{\sum\limits_{i<j}^{{}_pC_2} Y^2_{.(ij)}}{rp_1} - \dfrac{\sum\limits_{i<j}^{{}_pC_2} Y^2_{ij..}}{3rp_1} - \dfrac{p}{2p_1} H$	T*
$Y_{i(j.).}$ vs. $Y_{j(i.).}$	$p_1p_2/2$	$D = \dfrac{\sum\limits_{i<j}^{{}_pC_2} Y^2_{i(j.).} + \sum\limits_{i<j}^{{}_pC_2} Y^2_{j(i.).}}{rp_3} - \dfrac{\sum\limits_{i<j}^{{}_pC_2} [Y_{i(j.).} + Y_{j(i.).}]^2}{2rp_3} - \dfrac{3p_2}{2p_3} H$	D*
3-line order	$pp_2p_4/3$	$V = \dfrac{\sum\limits_{i<j<k}^{{}_pC_3} \sum\limits_{o}^{3} Y^2_{i(jk).}}{r} - \dfrac{\sum\limits_{i<j<k}^{{}_pC_3} Y^2_{ijk.}}{3r} - H - T - D$	V*
Error	$(r-1)(3\ {}_pC_3-1)$	$E = M - R - A$	E*
Total	$(3r\ {}_pC_3-1)$	$M = \sum\limits_{i<j<k}^{{}_pC_3} \sum\limits_{o}^{3} \sum\limits_{m=1}^{r} \Sigma^2_{i(jk)m} - C$	

additional source of variance, the variance of the difference between these two sums. In the double-cross partitioning, $Y_{(i.)(j.).} = Y(j.)(i.).$ so that this source of variance does not exist.

6.5 Expectations of Sums of Squares in Terms of Covariances

The expectations of the sums of squares in Table 15 can be derived in terms of μ^2, σ_r^2, σ_e^2 and the covariances between the effects of relatives as defined in Section 6.2 in the same manner as illustrated in the analysis of double-crosses. The expectations for the uncorrected sums of squares are presented in Table 16.

The expectations of sums of squares as partitioned in Section 6.4 may be derived from the results presented in Table 16 by applying the rules for the expectation of linear functions. If we let $\underline{S}$ be the (10 x 1) column vector of uncorrected sums of squares in Table 16, $\underline{C}$ be the (11 x 1) column vector of $(r\mu^2 + \sigma_r^2$, σ_e^2 and the nine covariances, and M be the (10 x 11) matrix of coefficients in Table 16, then $E(\underline{S}) = M\underline{C}$. Letting $\underline{S}_p$ be the (7 x 1) column vector of partitioned hybrid sums of squares, i.e., $\underline{S}_p' = (G, S, U, H, T, D, V)$, and N be the transformation matrix such that $\underline{S}_p = N\underline{S}$, then $E(\underline{S}_p) = NE(\underline{S}) = NM\underline{C}$. The coefficient of σ_e^2 in the expectation of each of the sums of squares is the degrees of freedom for the sum of squares. The expectations of the mean squares are then found by dividing by degrees of freedom. The expectations of the mean squares are presented in Table 17 in terms of σ_e^2 and seven functions of the covariances.

6.6 Model of Linear Effects

It is important to recognize that the development of a model of linear effects is not essential for the estimation of genetic parameters.

Table 16. Expectations of uncorrected sums of squares in terms of the covariances between three-way hybrid relatives

Sums of Squares	Coefficients of covariances $(r\mu^2 + \sigma_r^2)$	σ_e^2	Cov_1	Cov_2	Cov_3	Cov_4	Cov_5	Cov_6	Cov_7	Cov_8	Cov_9
$\frac{2Y^2_{....}}{rpp_1p_2}$	$\frac{pp_1p_2}{2}$	1	r	2r	$2rp_3$	$2rp_3$	$4rp_3$	rp_3	$\frac{rp_3p_4}{2}$	$2rp_3p_4$	$2rp_3p_4$
$\frac{2\sum_i^p Y^2_{i...}}{3rp_1p_2}$	$\frac{3pp_1p_2}{2}$	p	rp	2rp	$\frac{4rpp_3}{3}$	$\frac{4rpp_3}{3}$	$\frac{8rpp_3}{3}$	$\frac{2rpp_3}{3}$	$\frac{rpp_3p_4}{6}$	$\frac{2rpp_3p_4}{3}$	$\frac{2rpp_3p_4}{3}$
$\frac{\sum_{i<j}^{{}_pC_2} Y^2_{ij..}}{3rp_2}$	$\frac{3pp_1p_2}{2}$	$\frac{pp_1}{2}$	$\frac{rpp_1}{2}$	rpp_1	$\frac{rpp_1p_3}{3}$	$\frac{rpp_1p_3}{3}$	$\frac{2rpp_1p_3}{3}$	$\frac{rpp_1p_3}{6}$	0	0	0
$\frac{\sum_{i<j<k}^{{}_pC_3} Y^2_{ijk.}}{3r}$	$\frac{pp_1p_2}{2}$	$\frac{pp_1p_2}{6}$	$\frac{rpp_1p_2}{6}$	$\frac{rpp_1p_2}{3}$	0	0	0	0	0	0	0
$\frac{2\sum_i^p Y^2_{i(..).}}{rp_1p_2}$	$\frac{pp_1p_2}{2}$	p	rp	0	$2rpp_3$	0	0	0	$\frac{rpp_3p_4}{2}$	0	0
$\frac{\sum_i^p Y^2_{.(i.).}}{rp_1p_2}$	pp_1p_2	p	rp	rp	rpp_3	0	$2rpp_3$	rpp_3	0	0	rpp_3p_4

Table 16(continued)

Sums of squares	Coefficients of covariances										
	$(r\mu^2 + \sigma_r^2)$	σ_e^2	Cov_1	Cov_2	Cov_3	Cov_4	Cov_5	Cov_6	Cov_7	Cov_8	Cov_9
$\dfrac{\sum_{i<j}^{{}_pC_2} Y^2_{\cdot(ij)\cdot}}{rp_2}$	$\frac{pp_1p_2}{2}$	$\frac{pp_1}{2}$	$\frac{rpp_1}{2}$	0	0	0	0	$\frac{rpp_1p_3}{2}$	0	0	0
$\dfrac{\sum_{i<j}^{{}_pC_2} [Y^2_{i(j.)\cdot} + Y^2_{j(i.)\cdot}]}{rp_2}$	pp_1p_2	pp_1	rpp_1	0	rpp_1p_3	0	0	0	0	0	0
$\dfrac{\sum_{i<j}^{{}_pC_2} [Y_{i(j.)\cdot} + Y_{j(i.)\cdot}]^2}{2rp_2}$	pp_1p_2	$\frac{pp_1}{2}$	$\frac{rpp_1}{2}$	$\frac{rpp_1}{2}$	$\frac{rpp_1p_3}{2}$	$\frac{rpp_1p_3}{2}$	0	0	0	0	0
$\dfrac{\sum_{i<j<k}^{{}_pC_3} \sum_{o}^{3} Y^2_{i(jk)\cdot}}{r}$	$\frac{pp_1p_2}{2}$	$\frac{pp_1p_2}{2}$	$\frac{rpp_1p_2}{2}$	0	0	0	0	0	0	0	0

Table 17. Expectations of mean squares in terms of the covariances between hybrid relatives for the analysis of three-way hybrids.

M.S.	Coefficients of covariances							
	σ_e^2	$Cov_1 + 2Cov_2$	$Cov_1 - Cov_2$	$2Cov_3 + 2Cov_4 + 4Cov_5 + Cov_6$	$Cov_3 - Cov_4$	$Cov_3 + Cov_4 - 4Cov_5 + 2Cov_6$	$Cov_7 + 4Cov_8 + 4Cov_9$	$Cov_7 - 2Cov_8 + Cov_9$
G*	1	r	0	$r(2p-9)/3$	0	0	$rp_4p_9/6$	0
S*	1	r	0	$rp_7/3$	0	0	$-r(2p-11)/3$	0
U*	1	r	0	-r	0	0	r	0
H*	1	0	r	0	$3rp_3/2$	$rp_3/6$	0	$rp_3p_4/3$
T*	1	0	r	0	$-3r/2$	$r(2p-5)/6$	0	$-rp_4/3$
D*	1	0	r	0	$r(2p-9)/2$	$-r/2$	0	$-rp_4$
V*	1	0	r	0	$-3r/2$	$-r/2$	0	r
E*	1	0	0	0	0	0	0	0

The results of the previous five sections can be used directly to estimate the components of genetic variance. A model of linear effects is necessary, however, in order to develop tests of significance of various genetic hypotheses. As in the double-cross analysis, the model of linear effects which is presented is only one of the possible models which could be used. The model as presented closely parallels the model of linear effects for the double-crosses (5.6.2) with few exceptions. The effect g_i as defined here, has approximately the same genetic interpretation as the effect g_i for the double-cross model but does not have the same statistical definition. Additional effects, h_i and $d_{i(j)}$, have been added to the model of linear effects for three-way crosses to account for the extra sources of variation described in Section 6.4.

The following model of linear effects appears reasonable for the analysis of three-way hybrids:

$$Y_{i(jk)m} = \mu + r_m + G_{i(jk)} + e_{i(jk)m} \qquad (6.6.1)$$

$$= \mu + r_m + 2g_i + g_j + g_k + h_i + s_{ij} + s_{ik} + s_{jk}$$

$$+ t_{ij} + t_{ik} + d_{i(j)} + d_{i(k)} + u_{ijk} + v_{i(jk)}$$

$$+ e_{i(jk)m} \qquad (6.6.2)$$

where $Y_{i(jk)m}$, μ, r_m, $G_{i(jk)}$ and $e_{i(jk)m}$ are defined in Section 6.2 and where

g_i = a basic contribution particular to line i appearing as any one of the parental lines,

$g_i + h_i$ = an additional contribution particular to line i appearing as the single parent,

s_{ij} = a contribution particular to lines i and j appearing together in the three-way hybrid,

t_{ij} = an additional contribution particular to line i and line j appearing together with one or the other as the single parent,

$d_{i(j)}$ = an additional contribution particular to line i and line j appearing together with line i as the single parent,

u_{ijk} = a contribution particular to lines i, j, and k appearing together regardless of the order,

and $v_{i(jk)}$ = an additional contribution particular to lines i, j, and k appearing in the order i(j x k).

It is assumed that the $e_{i(jk)m}$'s are uncorrelated with mean zero and variance σ_e^2 and that there is a constant number of individuals in each plot. The terms g_i, h_i, s_{ij}, t_{ij}, and u_{ijk} are assumed to be random variables with zero means and variances σ_g^2, σ_h^2, σ_s^2, σ_t^2 and σ_u^2, respectively. The use of $g_i + h_i$ to represent the additional effect due to line i being the single parent is synonymous with defining an effect h_i^* to represent this contribution where $h_i^* = g_i + h_i$. However, h_i^* and g_i are then correlated with Cov (h_i^*, g_i) = Cov $(g_i + h_i, g_i) = \sigma_g^2$ and $\text{Var}(h_i^*)$ = Var $(g_i + h_i) = \sigma_g^2 + \sigma_h^2$. Hence, the use of $g_i + h_i$ gives effects which are uncorrelated. The terms $d_{i(j)}$ and $v_{i(jk)}$ may be treated either as completely random variables with means zero and variances σ_d^2 and σ_v^2, respectively, or with the restrictions that

$d_{i(j)} = - d_{j(i)}$ for all ij combinations and

$v_{i(jk)} + v_{j(ik)} + v_{k(ij)} = 0$ for all ijk combinations, respectively.

It was illustrated in the double-cross analysis that the choice of the model, in so far as random or fixed effects were concerned, had no affect on the information obtained from the analysis concerning genetic parameters. Thus, for the analysis of three-way hybrids, only one model will be used; that model assuming all effects to be completely random and independent variables.

6.7 Covariances as Functions of the Components of Variance

The covariances as defined in Section 6.2 can be expressed in terms of the effect components of variance (Section 6.6) by taking the expectations of the respective quadratic forms in terms of the model of linear effects. These results are presented in Table 18.

Table 18. Covariances between three-way hybrid relatives expressed as functions of the components of variance

Covariance	Coefficients of components of variance						
	σ_g^2	σ_h^2	σ_s^2	σ_t^2	σ_d^2	σ_u^2	σ_v^2
Cov_1	6	1	3	2	2	1	1
Cov_2	5	0	3	1	0	1	0
Cov_3	5	1	1	1	1	0	0
Cov_4	4	0	1	1	0	0	0
Cov_5	3	0	1	0	0	0	0
Cov_6	2	0	1	0	0	0	0
Cov_7	4	1	0	0	0	0	0
Cov_8	2	0	0	0	0	0	0
Cov_9	1	0	0	0	0	0	0

It would be desirable for the linear effects model to be compatible with the genetic expectations of the covariances as defined in Section 6.2. However, by the genetic expectations, $Cov_6 = Cov_8$ and $Cov_2 = Cov_3$. These equalities are not satisfied by the model of linear effects. No way was found to reconcile these differences with a linear model. It appears that a non-linear model would be necessary to completely describe the situation in the analysis of three-way hybrids.

6.8 Expectations of Mean Squares in Terms of Components of Variance and Estimation of Components of Variance

The expectations of the mean squares in terms of the effect components of variance are presented in Table 19. The detailed derivations of the expectations of the mean squares are not presented here. They may be derived either by the usual method of taking expectations of quadratic forms or by substituting the results of Table 18 in Table 17. In most of the following derivations, including the results presented in Table 19, the components of variance σ_g^2 and σ_h^2 are grouped so as to present a variance due to average one-line effect, σ_g^2, and a variance due to one-line order effect, $\sigma_h^{2*} = \sigma_h^2 + \sigma_g^2$. This makes the derivations simpler, since these variances coincide with particular sources of variation in the analysis of variance.

The estimators for the effect components of variance can be determined by inspection from Table 19. They are as follows with the estimate of σ^2 denoted by s^2:

$$s_v^2 = (V^* - E^*)/r$$

$$s_u^2 = (U^* - V^*)/3r$$

$$s_d^2 = (D^* - V^*)/rp_3$$

$$s_t^2 = 3(T^* - \frac{p_1}{3p_3} D^* - \frac{2p_4}{3p_3} V^*)/2rp_1$$

$$s_s^2 = [S^* - U^* - \frac{2p_4}{p_1}(T^* - V^*)]/3rp_4$$

$$s_{h^*}^2 = s_h^2 + s_g^2 = 3(H^* - \frac{p}{2p_1} T^* - \frac{3p_2}{2p_3} D^* + \frac{pp_2}{p_1p_3} V^*)/rpp_2 \text{ and}$$

$$s_g^2 = 2(G^* - \frac{p_3}{2p} H^* - \frac{2p_3}{p_4} S^* + \frac{p_3}{4p_1} T^* + \frac{3p_2}{4p} D^* + \frac{p_2}{p_4} U^* - \frac{p_2}{2p_1} V^*)/5rp_2p_3$$

Table 19. Expectations of mean squares for the analysis of three-way hybrids

M.S.	Coefficients of components of variance σ_e^2	σ_v^2	σ_d^2	σ_t^2	$\sigma_{h^*}^2 = \sigma_h^2 + \sigma_g^2$	σ_u^2	σ_s^2	σ_g^2	σ_r^2
R*	1	0	0	0	0	0	0	0	$3\ {}_pC_3$
G*	1	r	$4rp_3/3$	$8rp_3/3$	$rp_2p_3/6$	3r	$6rp_3$	$5rp_2p_3/2$	0
S*	1	r	$2rp_4/3$	$4rp_4/3$	0	3r	$3rp_4$	0	0
U*	1	r	0	0	0	3r	0	0	0
H*	1	r	$r(5p-9)/3$	$rp/3$	$rpp_2/3$	0	0	0	0
T*	1	r	$rp_1/3$	$2rp_1/3$	0	0	0	0	0
D*	1	r	rp_3	0	0	0	0	0	0
V*	1	r	0	0	0	0	0	0	0
E*	1	0	0	0	0	0	0	0	0

In lieu of estimating σ_g^2, which has an extraordinarily complex estimator, we may estimate $\sigma_{h*}^2 + 15\sigma_g^2 = \sigma_h^2 + 16\sigma_g^2$ as follows:

$$s_g^2 + 16s_g^2 = 6(G* - \frac{2p_3}{p_4} S* + \frac{p_2}{p_4} U*)/rp_2p_3$$

This is a much simpler estimator and, consequently, has a smaller variance.

The variances of these estimators can be approximated as shown in Section 5.8 in the double-cross analysis.

6.9 Components of Variance as Functions of Covariances between Hybrid Relatives

The components of variance can be expressed as functions of the covariances between hybrid relatives by use of the expectations in Table 17 and Table 19. In matrix notation, $E(\underline{S}_e) = A\underline{P}$, where

$\underline{S}_e' = (G* - E*, S* - E*, U* - E*, H* - E*, T* - E*, D* - E*, V* - E*)$,

A = the (7 x 7) matrix of coefficients of the components of variance for the expectations of the mean squares in Table 19 omitting mean squares R* and E* and the columns for σ_e^2 and σ_r^2 and

$\underline{P}' = (\sigma_v^2, \sigma_d^2, \sigma_t^2, \sigma_h^2*, \sigma_u^2, \sigma_s^2, \sigma_g^2)$. Also, $E(\underline{S}_e) = B\underline{C}_f$ where

B = the (7 x 7) matrix of coefficients of the functions of covariances in the expectations of the mean squares in Table 17 omitting mean square E* and the column for σ_e^2 and

$$\underline{C}_f' = (Cov_1 + 2Cov_2,\ Cov_1 - Cov_2,\ 2Cov_3 + 2Cov_4 + 4Cov_5 + Cov_6,$$
$$Cov_3 - Cov_4,\ Cov_3 + Cov_4 - 4Cov_5 + 2Cov_6,\ Cov_7 + 4Cov_8$$
$$+ 4Cov_9,\ Cov_7 - 2Cov_8 + Cov_9).$$

Then $A\underline{P} = B\underline{C}_f$ or $\underline{P} = A^{-1}B\underline{C}_f$.

Note that $\underline{P} = A^{-1}E(\underline{S}_e)$ or A^{-1} is the transformation matrix giving the estimators of the components of variance. Hence, A^{-1} may be written from the estimation equations in Section 6.8. Thus,

$$A^{-1} = \frac{1}{r}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1/p_3 & -1/p_3 \\ 0 & 0 & 0 & 0 & 3/2p_1 & -1/2p_3 & -p_4/p_1p_3 \\ 0 & 0 & 0 & 3/pp_2 & -3/2p_1p_2 & -9/2pp_3 & 3/p_1p_3 \\ 0 & 0 & 1/3 & 0 & 0 & 0 & -1/3 \\ 0 & 1/3p_4 & -1/3p_4 & 0 & -2/3p_1 & 0 & 2/3p_1 \\ 2/5p_2p_3 & -4/5p_2p_4 & 2/5p_3p_4 & -1/5pp_2 & 1/10p_1p_2 & 3/10pp_3 & -1/5p_1p_3 \end{bmatrix}$$

A rearrangement of the product matrix $(A^{-1}B)$ is given in Table 20.

6.10 Genetic Interpretations of the Components of Variance and Tests of Significance

The covariances between three-way hybrid relatives can be expressed in terms of genetic components of variance from Table 13 by expression (5.3.1). The genetic interpretations can then be placed on the components of variance from the relationships shown in Table 20. These are presented in Table 21 up to and including the three-factor epistatic components of variance.

The interpretations for various tests of significance are apparent from Table 21. The component σ_v^2 is a function of additive x dominance epistatic variance and all higher order epistatic components except the

Table 20. Components of variance as functions of the covariances between three-way hybrid relatives

Component of variance	Coefficients of covariances								
	Cov_1	Cov_2	Cov_3	Cov_4	Cov_5	Cov_6	Cov_7	Cov_8	Cov_9
σ_g^2	0	0	0	0	0	0	0	2/5	1/5
σ_{h*}^2	0	0	0	0	0	0	1	-2	1
σ_s^2	0	0	0	0	4/3	-1/3	0	-4/3	-2/3
σ_t^2	0	0	0	1	-2	1	0	0	0
σ_d^2	0	0	1	-1	0	0	-1	2	-1
σ_u^2	0	1	0	-1	-2	0	0	2	1
σ_v^2	1	-1	-2	1	2	-1	1	-2	1

all-additive types. Hence, the test of the hypothesis that $\sigma_x^2 = 0$ is a test that this particular combination of epistatic components of variance is equal to zero. The tests of the hypotheses $\sigma_u^2 = 0$ and $\sigma_d^2 = 0$ are tests that the additive x dominance and all higher epistatic components of variance except the all-dominance types are equal to zero. The test of $\sigma_u^2 = 0$ gives relatively more weight to the three-factor all-additive epistatic component than the test of $\sigma_d^2 = 0$.

The test of the hypothesis $\sigma_t^2 = 0$ is a test of all components of variance except additive genetic variance. The test of $\sigma_s^2 = 0$ is a test of the all-additive types epistatic variances.

The tests of the hypotheses that $\sigma_{h*}^2 = 0$ and $\sigma_g^2 = 0$ are tests of additive genetic variance and the all-additive epistatic components of variance with the latter giving considerably more weight, relatively, to the additive component of genetic variance.

Table 21. Components of variance in the analysis of three-way crosses as functions of the components of genetic variance up to and including the three-factor epistatic components

Component	Coefficients of components of genetic variance								
	σ^2_{10}	σ^2_{01}	σ^2_{20}	σ^2_{11}	σ^2_{02}	σ^2_{30}	σ^2_{21}	σ^2_{12}	σ^2_{03}
σ^2_g	$F/8$	0	$9F^2/320$	0	0	$17F^3/2560$	0	0	0
σ^2_{h*}	$F/8$	0	$9F^2/64$	0	0	$49F^3/512$	0	0	0
σ^2_s	0	0	$7F^2/96$	0	0	$11F^3/256$	0	0	0
σ^2_t	0	$F^2/4$	$F^2/32$	$F^3/8$	$F^4/16$	$9F^3/256$	$F^4/16$	$F^5/32$	$F^6/64$
σ^2_d	0	0	0	$F^3/32$	0	$3F^3/128$	$9F^4/256$	$F^5/128$	0
σ^2_u	0	0	0	$F^3/32$	0	$5F^3/64$	$9F^4/256$	$F^5/128$	0
σ^2_v	0	0	0	$F^3/32$	$F^4/8$	0	$13F^4/256$	$13F^5/128$	$3F^6/32$

Hence, considerable information can be obtained concerning gene action from the tests of the hypotheses that the various components of variance equal zero. It is evident from Table 19 that simple F-tests are available only for the hypotheses that $\sigma_v^2 = 0$, $\sigma_u^2 = 0$ and $\sigma_d^2 = 0$. The remaining tests make use of Satterthwaite's approximation as explained in Section 5.10. The exact and approximate F-tests for the various hypotheses are given in Table 22. The difference between the tests $\bar{\bar{F}}'$ and $\bar{\bar{F}}''$ is explained in Section 5.10. $\bar{\bar{F}}$ is used to indicate an F-test in order to avoid confusion with the F used to indicate the degree of inbreeding.

Another system of F-tests is available in which all the tests are exact tests but which will not generally yield as much information concerning gene action as the tests indicated in Table 22. As in the double-cross analysis, the error mean square is a valid error term for testing all of the hybrid mean squares. The hypotheses being tested, however, are the composite hypotheses that all genetic components of variance appearing in the respective mean squares are equal to zero. As a result, the interpretations of the tests will be considerably different from those in Table 22. The test $\bar{\bar{F}} = V^*/E^*$ is the same test as for the hypothesis $\sigma_v^2 = 0$. $\bar{\bar{F}} = U^*/E^*$ and $\bar{\bar{F}} = D^*/E^*$ are, however, tests of all epistatic components of variance except the two-factor additive x additive component. $\bar{\bar{F}} = T^*/E^*$ is a test of all components of genetic variance except additive genetic variance. $\bar{\bar{F}} = S^*/E^*$ is a test of the same hypothesis with relatively more weight given to the all additive types of epistasis. $\bar{\bar{F}} = H^*/E^*$ and $\bar{\bar{F}} = G^*/E^*$ are both tests of all components of genetic variance with the latter giving relatively more weight to the additive genetic variance component. It is apparent

Table 22. F-tests of hypotheses that various components of variance are zero in the analysis of three-way crosses

Component	F-test	Degrees of freedom[a]
σ_v^2	$\bar{F} = V^*/E^*$	$pp_2p_4/3$, $(r-1)(3\ {}_pC_3 - 1)$
σ_u^2	$\bar{F} = U^*/V^*$	$pp_1p_5/6$, $pp_2p_4/3$
σ_d^2	$\bar{F} = D^*/V^*$	$p_1p_2/2$, $pp_2p_4/3$
σ_t^2	$\bar{F}' = \bar{F}'' = \dfrac{T^*}{\dfrac{p_1}{3p_3} D^* + \dfrac{2p_4}{3p_3} V^*}$	$pp_3/2, f'$
σ_s^2	$\bar{F} = \dfrac{S^*}{U^* + \dfrac{2p_4}{p_1}(T^* - V^*)}$	$pp_3/2$, f'
	$\bar{F}'' = \dfrac{S^* + \dfrac{2p_4}{p_1} V^*}{U^* + \dfrac{2p_4}{p_1} T^*}$	f''_1, f''_2

Table 22(continued)

Component	F-test	Degrees of freedom[a]
$\sigma^2_{h*} = \sigma^2_h + \sigma^2_g$	$\bar{F}' = \dfrac{H*}{\frac{p}{2p} T* + \frac{3p_2}{2p_3} D* - \frac{pp_2}{p_1 p_3} V*}$	p_1, f'
	$\bar{F}'' = \dfrac{H* + \frac{pp_2}{p_1 p_3} V*}{\frac{p}{2p_1} T* + \frac{3p_2}{2p_3} D*}$	f''_1, f''_2
σ^2_g	$\bar{F}' = \dfrac{G*}{\frac{p_3}{2p} H* + \frac{2p_3}{p_4} S* - \frac{p_3}{4p_1} T* - \frac{3p_2}{4p} D* - \frac{p_2}{p_4} U* + \frac{p_2}{2p_1} V*}$	p_1, f'
	$\bar{F}'' = \dfrac{G* + \frac{p_3}{4p_1} T* + \frac{3p_2}{4p} D* + \frac{p_2}{p_4} U*}{\frac{p_3}{2p} H* + \frac{2p_3}{p_4} S* + \frac{p_2}{2p_1} V*}$	f''_1, f''_2

[a] f' and f'' indicate the approximate degrees of freedom as suggested by Satterthwaite (1946) (see equation 5.10.1)

that this system will not generally yield as much information. For example, the ability to test only the all-additive types of epistasis and the ability to test additive genetic variance plus only the all-additive types epistasis are lost.

The most powerful test of dominance and epistasis is

$$\bar{F} = \frac{2(S + T + D + U + V)/(p + 1)(p^2 - 4p + 2)}{E^*}$$

with $(p + 1)(p^2 - 4p + 2)/2$ and $(r - 1)(3\ {}_pC_3 - 1)$ degrees of freedom. The simplest most powerful test for epistasis is

$$\bar{F} = \frac{2(D + U + V)/(p^3 - 5p^2 + 4p + 2)}{E^*}$$

with $(p^3 - 5p^2 + 4p + 2)/2$ and $(r-1)(3\ {}_pC_3 - 1)$ degrees of freedom. This test, however, does not include the two-factor additive x additive epistatic deviations. In order to include these deviations in a single all-epistatic test, the test must be complex and an approximation.

6.11 Estimation of Genetic Parameters

There are two approaches available for estimation of the genetic components of variance. One approach is to fit by least squares procedures progressively more complex models to the estimated components of variance until no significant deviation from the model is obtained. The components of genetic variance are then estimated as the regression coefficients in the particular models.

Another method of estimating the genetic parameters is to equate the estimates of the components of variance to their expectations in terms of the genetic parameters and solve the set of equations. However, since the number of genetic parameters involved in the expectations is $n(n + 3)/2$, where n is the number of loci, a restricted genetic model must be assumed in order to reduce the number of parameters to an estimable number. In the three-way analysis there are seven equations but a dependency exists among them such that there are only six independent equations. Hence, only six genetic parameters may be estimated. An independent set of six equations can be formed by adding s_d^2 and s_u^2. This set will allow estimation of the parameters for the model

$$\sigma_G^2 = \sigma_{10}^2 + \sigma_{01}^2 + \sigma_{20}^2 + \sigma_{11}^2 + \sigma_{02}^2 + \sigma_{30}^2,$$

where σ_G^2 is total genetic variance, which assumes no epistatic effects involving more than two loci other than additive x additive x additive effects. For this genetic model, $\underline{P} = F\underline{G}$ or $\underline{G} = F^{-1}\underline{P}$, where

$$\underline{P}' = (s_{h*}^2 + 15s_g^2,\ s_{h*}^2,\ s_s^2,\ s_t^2,\ s_d^2 + s_u^2,\ s_v^2),$$

$$\underline{G}' = (s_{10}^2,\ s_{01}^2,\ s_{20}^2,\ s_{11}^2,\ s_{02}^2,\ s_{30}^2),$$

and

$$F = \begin{bmatrix} 2F & 0 & 9F^2/16 & 0 & 0 & 25F^3/128 \\ F/8 & 0 & 9F^2/64 & 0 & 0 & 49F^3/512 \\ 0 & 0 & 7F^2/96 & 0 & 0 & 11F^3/256 \\ 0 & F^2/4 & F^2/32 & F^3/8 & F^4/16 & 9F^3/256 \\ 0 & 0 & 0 & F^3/16 & 0 & 13F^3/128 \\ 0 & 0 & 0 & F^3/32 & F^4/8 & 0 \end{bmatrix}$$

from Table 21.

Then

$$F^{-1} = \begin{bmatrix} 0.30/F & 3.19/F & -8.47/F & 0 & 0 & 0 \\ -1.89/F^2 & 30.17/F^2 & -45.35/F^2 & 4/F^2 & -7/F^2 & -2/F^2 \\ 1.73/F^2 & -27.61/F^2 & 53.65/F^2 & 0 & 0 & 0 \\ 4.76/F^3 & -76.13/F^3 & 110.12/F^3 & 0 & 16/F^3 & 0 \\ -1.19/F^4 & 19.03/F^4 & -27.53/F^4 & 0 & -4/F^4 & 8/F^4 \\ -2.93/F^3 & 46.85/F^3 & -67.76/F^3 & 0 & 0 & 0 \end{bmatrix}$$

and the estimation equations are as follows:

$$s_{10}^2 = [0.30(s_{h*}^2 + 15s_g^2) + 3.19s_{h*}^2 - 8.47s_s^2]/F$$

$$s_{01}^2 = [-1.89(s_{h*}^2 + 15s_g^2) + 30.17s_{h*}^2 - 45.35s_s^2 + 4s_t^2 - 7(s_d^2 + s_u^2) - 2s_v^2]/F^2$$

$$s_{20}^2 = [1.73(s_{h*}^2 + 15s_g^2) - 27.61s_{h*}^2 + 53.65s_s^2]/F^2$$

$$s_{11}^2 = [4.76(s_{h*}^2 + 15s_g^2) - 76.13s_{h*}^2 + 110.12s_s^2 + 16(s_d^2 + s_u^2)]/F^3$$

$$s_{02}^2 = [-1.19(s_{h*}^2 + 15s_g^2) + 19.03s_{h*}^2 - 27.53s_s^2 - 4(s_d^2 + s_u^2) + 8s_v^2]/F^4$$

$$s_{30}^2 = [-2.93(s_{h*}^2 + 15s_g^2) + 46.85s_{h*}^2 - 67.76s_s^2]/F^3$$

For these estimation equations, the simpler estimator, $s_{h*}^2 + 15s_g^2$, is used in lieu of s_g^2. The alternative set of equations can be derived directly from these by rearranging the coefficients of s_{h*}^2 and s_g^2. The variances of these estimators may be approximated as shown in Section 5.11.

6.12 Discussion of the Analysis of Three-Way Crosses

The analysis of variance for three-way hybrids may be compared to the analysis of a p^3 factorial experiment. The three factors are the three positions a line may take in the pedigree of the hybrid and the p levels are the parental lines. This is not a complete factorial because of the restriction that a particular line cannot appear more than once in a particular hybrid. There are only 3 ${}_pC_3$ combinations instead of p^3 as there would be in a complete factorial. The effects of the factors, i.e., the positions of the lines in the pedigrees, are not distinguishable so that the one-line general sum of squares is comparable to the sum of squares due to the pooled main effects of the factors in a factorial. Similarly, the two- and three-line general sums of squares are comparable to the sums of squares due to the pooled two- and three-factor interaction effects, respectively. The order sums of squares are new features which depend upon the order in which the four lines are combined.

There is lack of complete balance in the three-way analysis, since a particular line may not appear in a pedigree with itself. There is, however, balance in the sense that each line appears an equal number of times with every other line. Because of the asymmetry of the three-way hybrids no linear model has been found as yet which is compatible with the covariances. The inability to reconcile the model and the covariances is apparently due to the inability to find a linear function the square of which will reconcile itself with all possible quadratic forms.

Qualitative estimation of the importance of various functions of the components of genetic variance, including several all-epistatic functions, is available from the F-tests in the three-way analysis of variance. Tests for deviations from an additive model and from an additive plus dominance plus additive x additive model are available. No simple test for deviations from an additive plus dominance model was found. Quantitative estimates of additive genetic variance, dominance variance, the two-factor epistatic variances and the three-factor all-additive epistatic variance component are obtainable if all other components of epistatic variance are assumed equal to zero. There will be some bias in the estimates if this assumption is not valid. This bias may under some conditions be quite large, if epistatic variance is important, since some of the coefficients of the higher epistatic components in the genetic expectations of the effect components of variance are of about the same magnitude as the coefficients of the estimated components.

The variances due to the interaction of the various hybrid effects with replications have been assumed to be equal in the three-way analysis of variance and have been pooled into a single error sum of squares. It may be desirable under some conditions to partition the error sum of squares into the individual interaction sums of squares.

The analysis as presented is for r replications grown in only one environment. This can be extended to include tests over years and locations by incorporating these effects in the model of linear effects. It would be desirable in the extended analysis to partition the hybrids x environments sum of squares into the sums of squares due to the interactions of each of the hybrid effects defined in the linear model with environments.

The degrees of freedom in the analysis of variance indicate that the complete analysis can be computed only if $p \geq 6$. For $p = 5$, there are no degrees of freedom available for estimation of the three-line general mean square. Hence, the complete three-way hybrid analysis can be computed with only sixty hybrids. This size analysis, however, will not have enough degrees of freedom to give good estimates of the parameters. For example, the largest number of degrees of freedom among the hybrid mean squares will be ten. Considering the size of the variance generally associated with estimates of components of variance, it would be desirable to use some larger number of lines if the purpose of the study is primarily that of estimation of the components of genetic variance.

7.0 General Discussion

The primary purpose of developing the analyses of variance for double-crosses and three-way crosses was to determine what genetic information could be obtained from some of the designs commonly used in practical breeding programs. Evaluation of hybrids is a routine part of corn breeding programs. It is not likely that the hybrids in any routine evaluation program would satisfy all of the assumptions necessary for the genetic analysis, particularly the assumption regarding the source of the lines. The analyses have shown, however, that it is possible to obtain considerable genetic information from double-crosses and three-way crosses and possibly, with some preliminary planning, as a by-product of a practical breeding program.

Both of the analyses presented allow estimation of six components of genetic variance. One of the limitations to the use of double-cross and three-way cross analyses for estimation is the rapidity with which the number of hybrids increases as the number of lines increases. The sample of lines from the population must be kept small in order to handle the resulting hybrids. The use of the complete double-cross and three-way analyses probably is not the most efficient method of estimation. There is some indication that a systematic subset of the crosses in each case, with an increase in the number of parental lines, may give more precision to some of the estimates than the complete analyses. More precision may also be obtained by using an incomplete analysis on several sets of hybrids from less than the minimum number of lines required for the complete analysis. For example, the use of five different sets of all possible double-crosses from six lines gives a total number of hybrids only slightly more than would one set of $p = 8$ lines necessary if the complete analysis is to be used. The five sets of $p = 6$ lines would provide a better sample of the original population than eight lines and, at the same time, the smaller blocks, 45 entries versus 210, will allow more control of environmental variation.

Total information might also be increased by including single-crosses and computing the three covariance analyses between the three sets of hybrids, that is, between the single-crosses, the three-way crosses and the double-crosses. The information contained in the covariance analyses remains to be elaborated.

The theoretical analyses of variance for double-cross hybrids and three-way hybrids in themselves enhance the understanding of the genetics involved. It is apparent from the double-cross analysis that, if the additive effects are the only genetic effects of importance, the order in which four lines are combined has no effect on the performance of the hybrid, since the one-line general sum of squares is the only one which involves additive genetic variance. With three-way hybrids, ordering is important even if all the genetic variance is due to additive effects as is indicated by the presence of additive genetic variance in the two-line order sum of squares. Further, it is evident from the analyses that the generally accepted method of predicting hybrid performance, i.e., using the mean of the non-parental single-crosses, is the appropriate method for both double-cross hybrids and three-way hybrids if epistasis is not important. These, of course, are not new concepts but they are exemplified by the analyses.

A problem which warrants investigation in connection with the double-cross and three-way cross analyses is the development of prediction formulae for use in selection of the better hybrids from three-way or double-cross hybrid evaluation trials. The usual method of selecting in an evaluation study is to simply pick the hybrids on the basis of their own performance. This, however, does not make use of all the information contained in the analysis. For example, it does not take into consideration the information contained in the rest of the test concerning the average performance of the lines in any particular hybrid. The same applies to all of the other effects defined in the model. A better selection scheme would be to base selection on a predicted phenotypic performance using a weighting function of the estimates of the effects obtained from the entire analysis.

8.0 Summary and Conclusions

Statistical concepts provide methods for the estimation of genetic parameters which are essential for the designing of efficient breeding programs. It would be desirable to be able to estimate these parameters from designs commonly used in breeding programs.

Analyses of variance are developed for all-possible double-cross hybrids and for all-possible three-way hybrids from a set of lines which have arbitrary but equal inbreeding to determine what genetic information is available from such analyses. The set of lines are assumed to be a random sample derived from a random mating population.

The hybrid sum of squares is partitioned into seven orthogonal sums of squares in both analyses. The expectations of the mean squares are presented in terms of the covariances between hybrid relatives and in terms of the effects defined in the models of linear effects for double-cross hybrids and for three-way hybrids.

The genetic expectations of the components of variance are derived from the covariances between hybrid relatives by means of which the genetic interpretations of various F-tests are determined. Several F-tests are presented in each analysis which provide information relative to various genetic hypotheses, many of which involve only epistatic components of genetic variance. Relatively powerful F-tests are presented in each analysis for dominance plus epistatic variance and for deviations from an additive plus dominance plus additive x additive model.

Each analysis provides information which allows the estimation of additive genetic variance, dominance variance, the three two-factor epistatic components of variance and the three-factor all-additive epistatic component. The estimators for these components are presented for each analysis and the method of computing their variances is indicated.

Limitations on the usefulness of these designs for estimation purposes are discussed and some possible modifications to improve efficiency are indicated. It is suggested that an analysis of a systematic subset of hybrids may provide more efficient estimation.

It is concluded that considerable information relative to gene action can be obtained from the analysis of variance of double-cross and three-way cross hybrids through inferences from the F-tests and from the estimates of the genetic components of variance.