Mating Designs

By William Beavis, Kendall Lamkey (ISU)

Except otherwise noted, this work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Introduction

Introduction

There are many mating designs developed for purposes of estimating the magnitude of genetic variability in a reference population. This information is most often useful to the plant breeder who is developing a new breeding program in a new crop species or developing a novel germplasm resource for established crop species. For example, large estimates of additive genetic variability and small estimates of genotype by environment variability suggest that rapid progress from selection can be made with minimal allocation of testing resources. While most recently trained plant breeders will assume responsibilities for established plant breeding programs, most established programs began with an evaluation of genetic variability using one of the many mating designs. Thus, we feel it is instructive to understand the genetic basis upon which these programs were established.

The choice of mating designs is based on:

1. Natural mode of reproduction and mating flexibilities of the species.
2. The objective(s) in estimating genetic variances such as:
    - General interest in knowledge of gene action for quantitative characters
    - Making a choice among alternative selection and breeding procedures
    - The prediction of response to selection.
3. Joint purposes such as estimating genetic variances and simultaneously selecting among progenies or evaluating hybrid combinations
4. Precision of the estimates.

Objectives

Students will learn about methods used to evaluate potential for genetic improvement in germplasm with unknown estimates of heritability through the application of the Variance-Covariance principle in various types of mating designs.

Design Setup

Setting up the treatment and experimental designs for mating designs create unique challenges. Several things need to be considered:

- Ease of making crosses in the species
- Inbreeding generation of the parents of the crosses
- The number of parents that will be used (male and females)
- Fixed versus random parents
- The type of mating design to be used
- The type of experimental design to be used
- The environmental design to be used

Diallel Crosses

Diallel Crosses - General Mating Scheme

Diallel matings are used to make inferences regarding the types of gene effects controlling traits. Diallels are particularly important in cross-pollinated crops and for determining the importance of general and specific combining ability. Consider the following general mating scheme. This scheme is very similar in structure to the two-way tables we have seen for studying interactions.

Parents	**P1**	**P2**	**P3**	**P4**	**Pn**	**Totals**
P1	Y_{11}	Y_{12}	Y_{12}	Y_{14}	Y_{1n}	$Y_{1.}$
P2	Y_{21}	Y_{22}	Y_{23}	Y_{24}	Y_{2n}	$Y_{2.}$
P3	Y_{31}	Y_{32}	Y_{33}	Y_{34}	Y_{3n}	$Y_{3.}$
P3	Y_{41}	Y_{42}	Y_{43}	Y_{44}	Y_{4n}	$Y_{4.}$
Pn	Y_{n1}	Y_{n2}	Y_{n3}	Y_{n4}	Y_{nn}	$Y_{n.}$
Totals	$Y_{.1}$	$Y_{.2}$	$Y_{.3}$	$Y_{.4}$	$Y_{.n}$	$Y_{..}$

Number of Diallel Crosses and Entries

Number of diallel crosses for n parents with and without reciprocal crosses. The number of entries is the number that would have to be evaluated if the parents are included in the experiment.

Without Reciprocals			**With Reciprocals**	
No. of Parents	**No. of Crosses**	**Number of entries**	**No. of Crosses Including Reciprocals**	**Number of entries**
n	$\frac{n(n-1)}{2}$	$\frac{n(n-1)}{2}$	$n(n-1)$	$n(n-1)$
5	10	15	20	20
6	15	21	30	30
7	21	28	42	42
8	28	36	56	56
9	36	45	72	72
10	45	55	90	90
11	55	66	110	110
12	66	78	132	132
13	78	91	156	156
14	91	105	182	182
15	105	120	210	210
20	190	210	380	380
50	1225	1275	2450	2450
100	4950	5050	9900	9900

Types of Diallel Analysis:

Model	**Method**	**Parents Included**	**Crosses**	**Reciprocals**
I (Fixed)	1	Yes	Yes	Yes
I (Fixed)	2	Yes	Yes	No
I (Fixed)	3	No	Yes	Yes
I (Fixed)	4	No	Yes	No

Model	Method	Parents Included	Crosses	Reciprocals
II (Random)	1	Yes	Yes	Yes
II (Random)	2	Yes	Yes	No
II (Random)	3	No	Yes	Yes
II (Random)	4	No	Yes	No

Common Diallel Experiment

The most common diallel experiment is conducted with selected parents, which means a fixed effects analysis where only gene effects and not variance will be estimated. The reason for this is simple: It is very hard to sample a population adequately with a diallel. Diallels are useful mating designs, however, despite this limitation.

Therefore, we will not present any analyses related to estimating variance components — only gene effects. This makes this section somewhat out of place, but it fits in with the other mating designs from the structure point of view. The analyses we will present are a combination of those presented by Griffing (1956) and Gardner and Eberhart (1966).

Methods 2 and 4 are the most common types of diallels. Most scientists grow the parents and the crosses or just the crosses. The method 4 analysis is, however, the most commonly used analysis, because Griffing assigns specific combining ability effects to the parents per se and these are hard to interpret relative to Sprague and Tatum's (1942) definitions of general and specific combining ability.

The general model underlying the diallel can be written as follows:

$$Y_{ijk} = +g_i + g_j + s_{ij} + r_k + e_{ijk}),$$

where g_i is the general combining ability effect (marginal effect) of the i^{th} parent, s_{ij} is the specific combining ability effect (interaction effect) of the i^{th} and j^{th} parents, r_k is the effect if the k^{th} replication, e_{ijk} is the error, and μ is the mean.

ANOVA Table

Source	**df**	**df (n-10)**	**SS**	**MS**	**EMS (Model I - Fixed Effe**
Replications	$r-1$	$r-1$			
Entries	$\frac{n(n-1)}{2}-1$	44	S_2	M_2	Invalid Equation
Among Margins	$n-1$	9	S_{21}	M_{21}	$\sigma_\varepsilon^2+(\frac{n-2}{n-1})\sum_i$
Among cells/Margins	$\frac{n(n-3)}{2}$	35	S_{22}	M_{22}	$\sigma_\varepsilon^2+(\frac{2r}{n(n-3)})$
Error	$(r-1)(\frac{n(n-1)}{2}-1)$	$44(r-1)$	S_1	M_1	σ_ε^2

F-Tests

Model I F-Tests:

Invalid Equation

These F-tests evaluate whether differences among the parents and crosses within parents are significant. Also, it is possible to show that the effects can be estimated as follows:

$$\hat{u} = \frac{2}{n(n-1)} Y_{..}$$
$$\hat{g}_1 = \frac{1}{n(n-2)} (nY_{i.} - 2Y_{..})$$
$$\hat{s}_{ij} = Y_{ij} - \frac{1}{n-2}(Y_{i.} + Y_{.j}) + \frac{2}{(n-1)(n-2)} Y_{..}$$

The variances of the effects can be estimated as:

$$V(\hat{u}) = \frac{2}{n(n-1)} \hat{\sigma}^2_{\bar{Y}}$$
$$V(\hat{g}_i) = \frac{1}{n(n-2)} \hat{\sigma}^2_{\bar{Y}}$$
$$V(\hat{s}_{ij}) = \frac{n-3}{n-1} \hat{\sigma}^2_{\bar{Y}}, (i \neq j)$$
and
$$\hat{\sigma}^2_{\bar{Y}} = \frac{\hat{\sigma}^2_{\varepsilon}}{r} = \frac{M_1}{r}$$

Gardner and Eberhart Diallel Analysis II

Gardner and Eberhart Diallel Analysis II

The Gardner and Eberhart Analysis II for the diallel is a more general analysis designed for the case of when the diallel includes random mating varieties. The model is best laid out by starting with the following single locus theory for the j^{th} variety and i^{th}locus:

Frequency	Genotype	Genotypic value
p_{ji}^2	AA	$\mu' + a_i$
$2p_{ji}(1-p_{ji})$	Aa	$\mu' + \delta_i$
$(1-p_{ji})^2$	aa	$\mu' - a_i$

Where:

$$= (AA + aa)/2$$

The population mean can be written as:

$$\mu' + \sum_i (2p_{ji} - 1)\alpha_i + 2\sum_i (p_{ji} - p_{ji}^2)\delta_i$$

Equations

Let:

$$a'_j = \sum_i (2p_{ji} - 1)\alpha_i$$
$$\bar{a} = \frac{1}{n}\sum_j a'_j$$
$$a_j = a'_j - \bar{a}$$
$$\mu = \mu' + \bar{a}$$

Similarly, let:

$$d_j = 2\sum_i (p_{ji} - p_{ji}^2)\delta_i$$
$$h_{jj'} = \sum_i (p_{ji} - p_{ji}^2)^2\delta_i$$

Then, the population mean can be written as:

$$= \mu' + a'_j + d_j$$
$$= \mu + a_j + d_j$$

and a population cross can be written as:

$$C_{jj'} = \mu + \frac{1}{2}(a_j + a_{j'}) + \frac{1}{2}(d_j + d_{j'} + h_{jj'}$$

If the varieties, varieties selfed, population crosses, population crosses selfed, and population crosses random mated are included in the analysis then all of these genetic effects can be estimated. Usually this is not the case and only varieties and variety crosses are included in the analysis, which are confounded and they have to be estimated together. We can then define the following parameters:

$$\mu_v = \mu + \frac{1}{n}\sum_j d_j$$
$$= \mu + \bar{d}$$

$=$ the mean of all parental varieties included in the analysis

and

$$v_j = a_j + (d_j + \bar{d})$$

=the variety effect when parents are included in the analysis

Models

We can then fit the following four models to the data:

Invalid Equation

where:

$$y = \begin{Bmatrix} 0 \; when \; j = j' \\ 1 \; when \; j \neq j' \end{Bmatrix}$$

ANOVA Table

The following ANOVA table can be written:

Source	df	Sum of squares
Populations	$[n(n-1)/2]-1$	S'
Varieties (v_j)	$n-1$	Invalid Equation
Heterosis ($h_{jj'}$)	$n(n-1)/2$	Invalid Equation
Average ($\bar{h}$)	1	Invalid Equation
Variety (h_j)	$n-1$	Invalid Equation
Specific ($s_{jj'}$)	$n(n-3)/2$	$S'_{23} = (B'G)_4 - (B'G)_3$

Equivalent Analysis

An equivalent analysis can be made with just the crosses as follows:

$$Let: (\mu_c = \mu_v + \bar{h})(= \mu + \bar{d} + \bar{h}) = themeanofcrossesinthediallel$$

Invalid Equation

$$then(C_{jj'} = \mu_c + g_j + g'_j + s_{jj'}$$

$$and(\sum_{j} g_j = 0)(\sum_{j \neq j'} s_{jj'} = 0$$

Analysis III of Gardner and Eberhart

Then the following ANOVA can be written (Analysis III of Gardner and Eberhart)

Source	$n-1\,(v_j)\,S'_{23}=S''_{32}$ **d.f.**	**Sum of squares**
Varieties		S''_1
Varieties vs. Crosses	1	S''_2
Crosses	$[n(n-1)/2]-1$	S''_3
GCA(g_j)	$n-1$	S''_{31}
SCA ($s_{jj'}$)	$n(n-3)/2$	S''_{32}

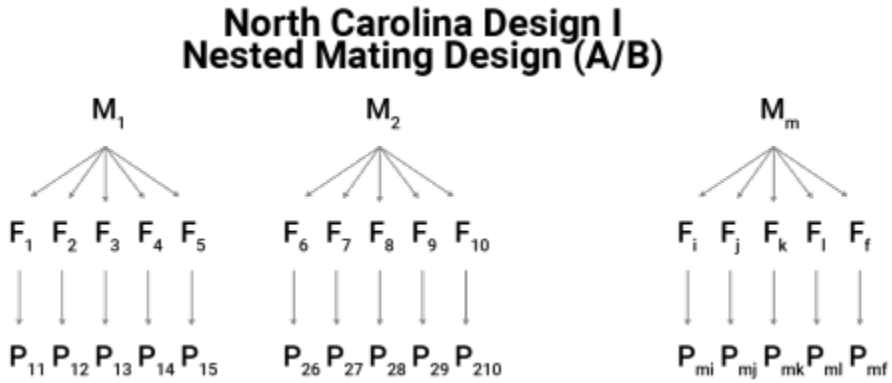
The analysis of Crosses, GCA, and SCA is all that can be done if only the crosses are included in the analysis. This analysis is equivalent to the Model 4 analysis of Griffing. If varieties or parents are also included, then the analysis Varieties and Varieties vs. Crosses can also be calculated.

The Analysis III is related to the Analysis II in the following ways that the ($s_{jj'}$) are the same in the two analyses $S'_{21} = S''_2$, meaning that average heterosis is simply a contrast of the mean of the varieties with the mean of the crosses.

$$S' + S'_{22} = S'_{31}), since(g_j = \frac{1}{2}v_j + h_j$$

North Carolina Design I

North Carolina Design I



- Consider m male plants,
- each of which is mated to f female plants
- to produce n full-sib families within each male,
- for a total of mf half-sib families.
- There are a total of m half-sib families.
- Different female plants are used to cross with each male.
- The progeny P are grown in a replicated experiment design.

$$Y_{ijk} = \mu + m_i + f_{ij} + r_k + e_{ijk}$$

where,

μ = mean

m_i = the effect of male i

f_{ij} = the effect of female j when crossed to male i

r_k = replication effect

e_{ijk} = the residual

ANOVA Table

Then the ANOVA can be written as:

Source of Variation	**d.f.**	**MS**	**EMS**
Replications	$r-1$		
Males	$m-1$	M4	$\sigma + r\sigma^2_{f(m)} + rf\sigma^2_m$
Females/Males	$m(f-1)$	M3	$\sigma^2 + r\sigma^2_{f(m)}$
Error	$(mf-1)(r-1)$	M2	σ^2
Total	$rmf-1$		
Within	$rmf(k-1)$	M1	σ^2_W

The table can be rewritten in terms of the covariance of relatives as follows:

Source of Variation	**d.f.**	**MS**	**EMS**
Replications	$r-1$		
Males	$m-1$	M4	$\sigma^2 + r[Cov(FS) - Cov(HS)] + rfCov(HS$
Females/Males	$m(f-1)$	M3	$\sigma^2 + r[Cov(FS) - Cov(HS)]$
Error	$(mf-1)(r-1)$	M2	σ^2
Total	$rmf-1$		
Within	$rmf(k-1)$	M1	σ^2_W

$$And(\sigma^2_u = \frac{\sigma^2_{\mu\varepsilon} + [\sigma^2_G - Cov(FS)]}{k}$$

Variance Estimates

Estimation:

Invalid Equation

So (ignoring epistasis),

Invalid Equation

Consider the case when $F_m = F_f = 0$, i.e., all the parents are noninbred.

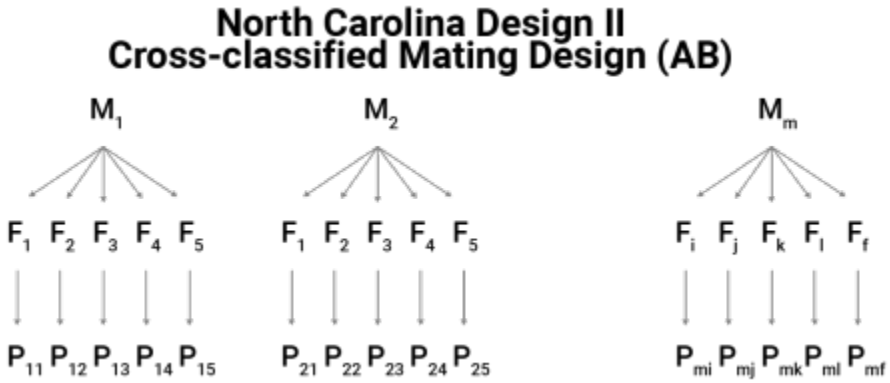
Invalid Equation

When $F_m = F_f = 1$, both the male and female parents are inbred, then

Invalid Equation

North Carolina Design II

North Carolina Design II



- Consider m male plants,
- each of which is mated to f female plants
- to produce f full-sib families within each male,
- for a total of mf half-sib families.
- There are a total of m + f half-sib families.
- The same female plants are crossed with each male.
- The progeny P are grown in a replicated experiment design.

The design is related to the diallel and another more simple way to represent the design is:

Parents	**M1**	**M2**	**M3**	**M4**	**Totals**
F5	Y_{15}	Y_{25}	Y_{35}	Y_{45}	$Y_{.5}$
F6	Y_{16}	Y_{26}	Y_{36}	Y_{46}	$Y_{.6}$
F7	Y_{17}	Y_{27}	Y_{37}	Y_{47}	$Y_{.7}$
F8	Y_{18}	Y_{28}	Y_{38}	Y_{48}	$Y_{.8}$
Totals	$Y_{1.}$	$Y_{2.}$	$Y_{3.}$	$Y_{4.}$	$Y_{..}$

Model

$$Y_{ijk} = +m_i + f_{ij} + mf_{ij} + r_k + e_{ijk}$$

where,

μ = mean

m_i = the effect of male *i*

f_j= the effect of female *j*

mf_{ij} = the interaction effect of female *j* when crossed to male *i*

r_k = replication effect

e_{ijk} = the residual

Source of Variation	**d.f.**	**MS**	**EMS**
Replications	*r* - 1		
Males (M)	*m* - 1	M5	$\sigma^2 + r\sigma^2_{mf} + r\sigma^2_m$
Females (F)	*f* - 1	M4	$\sigma^2 + r\sigma^2_{mf} + rm\sigma^2_f$
M x F	(*m* - 1)(*f* - 1)	M3	$\sigma^2 + r\sigma^2_{mf}$
Error	*mf* - 1)(*r* - 1)	M2	σ^2
Total	*rmf* - 1		
Within	*rmf*(*k* - 1)	M1	σ^2_w

Covariance of Relatives

The table can be rewritten in terms of the covariance of relatives as follows:

Table 1 A general ANOVA table for any type of related progeny.

Source of Variation	d.f.	MS	EMS
Replications	$r-1$		
Males (M)	$m-1$	M5	$\sigma^2+r[Cov(FS)-Cov(HS_m)-Cov(HS_f)]+rfCov(H$
Females (F)	$f-1$	M4	$\sigma^2+r[Cov(FS)-Cov(HS_m)-Cov(HS_f)]+rmCov(H$
M x F	$(m-1)(f-1)$	M3	$\sigma^2+r[Cov(FS)-Cov(HS_m)-Cov(HS_f)]$
Error	$(mf-1)(r-1)$	M2	σ^2
Total	$rmf-1$		
Within	$rmf(k-1)$	M1	σ_w^2

$And(\sigma_u^2=\frac{\sigma_{w\varepsilon}^2+[\sigma_G^2-Cov(FS)]}{k}$

Estimation:

Invalid Equation

Variance Estimates

Ignoring epistasis:

Invalid Equation

Consider the case when $F_m = F_f = 0$, i.e., all the parents are noninbred.

Invalid Equation

When $F_m = F_f = 1$, both the male and female parents are inbred, then

Invalid Equation

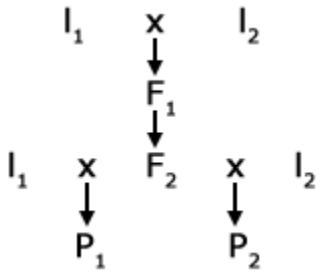
Equation 1

North Carolina Design III

North Carolina Design III

The main use of the Design III is for estimating the average degree of dominance.

North Carolina Design III Backcross Design



This design involves crossing two inbred lines and obtaining the F_1 and F_2 generations. An individual F_2 plant is then backcrossed to each of the inbred parents generating a pair of progeny using the F_2 plants and pollen parents. Then for n F_2 plants there are 2n progenies produced.

$$Y_{ijk} = +l_i + m_j + ml_{ij} + r_k + e_{ijk}$$

where,
μ = mean
l_i = contrast of the two inbred parents (i = 1,2)
m_j = the effect of F_2 parent *j*
ml_{ij}= the interaction effect of inbred line *i* and F_2 parent *j*
r_k = replication effect
e_{ijk} = the residual

ANOVA Table

Source of Variation	**d.f.**	**MS**	**EMS**
Replications	*r* - 1		
Inbred Lines	1		
F_2 parents	*n* - 1	M3	$\sigma^2 + 2r\sigma^2_m$
F_2 parent x inbred line	*n* - 1	M2	$\sigma^2 + r\sigma^2_{ml}$
Error	(2*n* - 1)(*n* - 1)	M1	σ^2
Total	*rmf* - 1		

Estimation:

Invalid Equation

F-Tests

Remember that in an F_2 population:

Invalid Equation

so that,

Invalid Equation

Note that this design is very specialized for the specific case of F_2 populations when p = q = 0.5. This design provides exact F-tests of two important hypotheses:

1. The null hypothesis of no dominance. This is tested by:

$$(F = \frac{M2}{M1}$$

if this F-test is significant then it means that $d \succ |0|$ and there is no dominance.

2. The null hypothesis that dominance is complete. If there is complete dominance, then the ratio:

$$\frac{M3}{M2} = 1$$

A significant departure of this ratio from one indicates that d departs significantly from 1.

Acknowledgements

This module was developed as part of the Bill & Melinda Gates Foundation Contract No. 24576 for Plant Breeding E-Learning in Africa.

Quantitative Genetics Mating Designs Author: William Beavis, and Kendall Lamkey (ISU)

Multimedia Developers: Gretchen Anderson, Todd Hartnell, and Andy Rohrback (ISU)

How to cite this module: Beavis, W. and K. Lamkey. 2016. Mating Designs. *In* Quantitative Genetics, interactive e-learning courseware. Plant Breeding E-Learning in Africa. Retrieved from **https://pbea.agron.iastate.edu**.

Source URL: https://pbea.agron.iastate.edu/course-materials/quantitative-genetics/mating-designs-0?cover=1