Australian Journal of

Crop Science

AJCS 14(12):1855-1869 (2020) doi: 10.21475/ajcs.20.14.12.p2588



Mating designs commonly used in plant breeding: A review

Jane Muthoni^{1*, 2}, Hussein Shimelis²

¹Kenya Agricultural and Livestock Research Organization (KALRO), Kenya ²African Centre for Crop Improvement, University of KwaZulu-Natal, College of Agriculture, Engineering and Science, School of Agricultural, Earth and Environmental Sciences, Private Bag X01, Scottsxille 3209, Pietermaritzburg, South Africa

*Corresponding author: jayney480@yahoo.com

Abstract

Mating design represents "rules" for arranging different controlled crossings; a mating design is a procedure of producing progenies. This review describes mating designs commonly used in plant breeding. Biparental is the simplest design in which a number of *P* plants are paired off to give $\frac{1}{2}P$ families; the parents are mated only once in pairs. The design is simple and it provides information needed to determine if significant genetic variation is present in a population for a long term selection program. A polycross is a mating arrangement for interpollinating a group of cultivars or clones using natural hybridization in an isolated crossing block. The design is often used for generating synthetic cultivars. The North Carolina (NC) mating designs permit determination and/or estimation of variance components (additive and dominance components) by using the information from half-sib (HS) families. The experimental material of North Carolina designs I, II and III is developed from F₂ generation as a base material. The design III (NCIII) involves backcrossing the F₂ plants to the two parental inbred lines from which the F₂ were derived. The NCIII design was extended to include a third tester. This third tester is the F₁ from the two parental inbred lines; in this extended form, this design is useful in deciding the relative ability of a number of female and male inbreds to produce desirable hybrid combinations. When the same parents are used as females and males in breeding, the mating design is called diallel. Parents used range from inbred lines to broad genetic base varieties to clones. The design is the most commonly used in crop plants to estimate GCA and SCA and variances.

Generally, it should be noted that is not the mating design *per se*, but the breeder who breeds a new cultivar. Consequently, proper choice and use of a mating design will provide the most valuable information for breeding.

Keywords: Breeding, Covariance, Designs, GCA, Mating, Plant, SCA, Variances.

Introduction

Mating design represents "rules" for arranging different controlled crossings; a mating design is a procedure of producing progenies (Klein et al., 1973). Selecting the kinds of mating techniques and arrangements depends upon: 1) the predominant type of pollination (self or cross); 2) type of crossing used (artificial or natural); 3) type of pollen dissemination (wind or insect); 4) unique features, such as cytoplasmic or genetic sterility; 5) purpose of project (breeding or genetic); and 6) size of population required (Stuber, 2004). The breeder influences the outcome of a mating by the choice of the parents, the control over the frequency with which each parent is involved in mating, and the number of offspring per mating, among other ways (Acquaah, 2007). Mating designs are used to generate genetic pedigrees, genetic information, and materials that can be used in a breeding program (Jenkins, 1934); they are also used in estimating information on general combining ability (GCA) and specific combining ability (SCA).

In general, the objectives of mating designs are to provide information for evaluating parents (GCA and SCA), provide estimates of genetic parameters, provide estimates of genetic gains, and provide a base population for selection (Klein et al., 1973). General combining ability is the average performance of a parental line in hybrid combinations and SCA is the contribution of an inbred line to hybrid performance in a cross with a specified inbred line, in relation to its contributions in crosses with an array of specified inbred lines (Sleper and Poehlman, 2006)

Mating designs offer different hierarchical structures such as half-, full-sib family, and individuals within family in the progeny population (Lindgren and Matheson, 1986). In evaluating mating designs with respect to estimation, the sampling variance of the genetic variances is used (Nasoetion et al., 1967; Namkoong and Roberds 1974; Pepper 1983). In general, for a given number of parents, the mating design that includes the larger number of crossings will produce the smaller sampling variance. For example, Klein et al. (1973) indicated that it would take at least 400 families to estimate heritability with a standard error less than 0.1.

The mating designs have been classified into one, two, three, and four factor designs according to the number of

factors, parents, and the modalities of combination of these factors (Cockerham, 1963; Table 1).

The simplest are one way mating designs corresponding to independent families, for example half-sibs or S₁ families. They allow for estimation of genetic variance among families, but not the components of genetic variance, unless only additivity is present (Griffings, 1956b). Two way mating designs require related crosses among plants thereby allowing cov HS and cov FS or GCA and SCA variances to be estimated. The two-way mating designs (North Carolina design I, North Carolina Design II and diallel) are based on the development of related single crosses among plants. They differ according to the origin and crossing modalities of plants used as males and females. Two way mating designs will allow estimation of general and specific combing ability and variances and thus of just two covariances, among halfand fullsibs. They are useful for estimating genetic variance components if only two, generally σ_A^2 and σ_D^2 explain a major part of the total genotypic variance.

Three-way mating designs, for example triallel, requiring the development of related three-way crosses, allow for the estimation of at least three parameters. Four-way mating designs, for example quadriallel, which are mating designs based on the development of related double crosses, allow at least four variance components to be estimated (LeClerg. 1966). For effective genetic analyses, components of variance of a design are translated into covariances of relatives, which are readily interpretable into components of genotypic variance (Klein et al., 1973). In most cases, estimates of variance components obtained through translation of covariance of relatives refer to the population from which the mating individuals were sampled. This population is termed as a "reference population", both for the genotype and environments. It is in terms of variances of the reference population that various procedures and designs are interpreted and/or assessed for purposes of decision making (Kang and Namkoong 1988).

In this review, the following mating designs commonly used in plant breeding are described: Biparental mating (paired crossing); Polycross mating; North Carolina Designs I, II, and III; Line x Tester mating design; Topcross and Diallel mating (full, half, smart, partial and disconnected).

Biparental mating (BPIs) (or paired crosses)

It is the simplest design in which a number of *P* plants are paired off at random to give $\frac{1}{2}P$ families. The parents are mated only once in pairs. The *P* parents generate $\frac{1}{2}P$ full-sib families (Mather and Jinks, 1982).

1 x 2 3 x 4 5 x 6 7 x 8 ... x ...

COVFS=1/2 VA+1/4VD+...

The progeny are tested and the observed variation partitioned by a straight forward analysis of variance into between and within families. If m plants per progeny family are evaluated, the variation within (w) and between (b) families may be analysed as shown in table 2.

Where σ_b^2 is the covariance of full-sibs, which equals to $\frac{1}{2}V_A + \frac{1}{4}V_D + V_{EC} = 1/m (MS_1-MS_2)$ while σ_W^2 is equal to $\frac{1}{2}V_A + \frac{3}{4}V_D + V_{EW} = MS_2$. The design is simple and it provides information needed to determine if significant genetic variation is present in a population for a long term selection program. However, no information is available for the type of genetic variation. It is unable to yield sufficient information to estimate all the parameters required by the model. Only two statistics are available for estimating V_A , V_D , V_{EC} and V_{EW} . This is because the progeny from this design are either full-sibs or unrelated.

This design provides good information for full-sib family performance and estimates of some genetic parameters. It provides information on maximum unrelatedness, it is simple to implement, and it is a low cost design. However, it is not good for estimating genetic gains from V_A and it is not optimum for selection (Klein et al., 1973). An F test is used to determine if the variation among crosses is significantly different from zero and an intraclass correlation is can be computed.

Polycross Mating Design

A polycross is a mating arrangement for interpollinating a group of cultivars or clones using natural hybridization in an isolated crossing block (Klein et al., 1973). The term polycross means progeny from a line that was subjected to out-crossing with other selected lines growing within the same nursery. This design is for intermating a group of cultivars by natural crossing in an isolated block. If an isolation block is not available, hand-crossing is required and the entries must be planted to facilitate the required interpollinations. The mating design is often used for generating synthetic cultivars and may be used for recombining selected entries or families in recurrent selection programs. Progeny from each entry have a common parent in the polycross design. Thus, half-sib families are generated and these are frequently used for evaluating general combining abilities. The polycross mating design is most suited to species that are obligate crosspollinators such as forage grasses and legumes, sugarcane, sweet potatoes, and in particular those that can be propagated vegetatively and where the resultant clones can be maintained for several years (Frandsen and Frandsen, 1948). Because the purpose of the polycross is to provide an equal opportunity for each entry to be crossed with every other entry, the field layout is the critical feature of the design. It is critical that the entries be equally represented and randomly arranged in the crossing block. If the entries are 10 or less, the Latin square design is appropriate because it gives chance for an entry to be randomly allocated to each column and each row, thus giving them an equal chance of mating with each other (Figure 1). With more than 10 entries, a randomized complete block design (RCBD), with adequate replications, is usually preferred to the Latin square. When the seed is harvested, equal quantities from each replication of each entry are bulked (Klein et al., 1973; Stuber, 2004). The ideal requirements are hard to meet in practice because of several problems placing the system in jeopardy of deviating from random mating. Problems include entries flowering at different times and non-random pollen dispersal (Acquaah, 2007). Because there is a possibility of entries flowering at different times, it is usually

necessary to stagger planting dates for annual crops so that flowering occurs simultaneously in order to ensure mating at the same time. However, prior knowledge on the flowering behaviour of the entries involved is required. The polycross is often used for generating synthetic cultivars, and may be used for recombining selected entries or families in recurrent selection programs. If s off-springs from each of n mother plants are progeny tested, the resultant variation is partitioned into within and between maternal groups. A family structure can no longer be recognised because each maternal group results from matings with an unspecified and unknown number of different males. Although the ANOVA resembles that for paired crosses, its precise form differs because individuals within maternal groups have one parent in common and are normally half-sibs (Table 3).

North Carolina Mating Designs

The North Carolina mating designs are multipurpose as they permit determination and/or estimation of the variance components (i.e., additive and dominance components) by using the information from half-sib (HS) families i.e. sets of progeny which have either the same father but different mothers or vice versa. The North Carolina Mating Designs comprise of North Carolina Mating Design I, Design II, and Design III (Stuber, 2004; Acquaah, 2007). In all cases, the base population for genetic studies is the F₂ generation from which the parents are assumed to be randomly drawn and are not a selected sample. If the F₂ generation is developed from two true breeding homozygous lines, then the assumption regarding gene frequencies is satisfied. Additionally, it is assumed that the progenies are raised in a completely randomized experimental design, although at times blocking may be necessary for convenient experimental and data management. Furthermore, field arrangement of experimental material is an important aspect because large numbers of progenies are raised for purposes of getting reliable estimates of components of genotypic variances. Major differences between choice of the experimental designs for this kind and other agronomic experiments are that the quantities of interest are variances rather than means. Replications-in-blocks designs ought to be suitable when blocking of genetic material is necessary in order to control the environmental variability. However, the mating design must be such that it permits the sets of progenies to be randomly assigned to the blocks in a way that will allow an independent estimate of variances of interest from analysis of each block (LeClerg, 1966). The assumptions needed for valid genetic interpretation of NC I, II, III experiments are (1) random choice of individuals mated for production of experimental progenies (2) random distribution of genotypes relative to variation in environments (3) no non-genetic maternal effect (4) regular diploid behaviour at meiosis (5) no multiple alleles (6) no linkage except where equilibrium between coupling and repulsion phase exist (7) no epistasis, and (8) for estimating degree of dominance, gene frequencies of 0.5 at all loci when segregation occurs (except for design III)(LeClerg, 1966).

North Carolina Design I (Nested mating or Hierarchical Design)

This is a nested type of mating design in which each member of a group of parents used as males is mated to a different group of parents used as females thus generating full-sib families nested within a half-sib family. No female parent is involved in more than one mating (Klein et al., 1973). As such the females are nested within males and each cross generates one family (Figure 2).

COVHS_M=1/4 VA

V female/male = COVFS -COVHS_M

=1/2 VA+1/4VD -1/4 VA

=1/4 VA +1/4 VD

The females are grown within isolation plots surrounded by the pollinating male. The test materials for North Carolina design I (NC I) originate from matings among F₂ plants from a cross of two inbred lines. The *m* pollen plants (males) are each mated to *n* seed plants (female) with the limitation that a female plant is used in one mating only. There are s sets of matings with a different selection of males in each set. The progeny of the s x m x n matings make up the material for future experimentation. Full-sibs with the same mother and father plants are within families; half-sibs with the same mother or father are across families (Acquaah, 2007).

The design is commonly used to estimate additive and dominance variances as well as for the evaluation of full-sib and half-sib recurrent selection; it provides information on parents and full-sib families (Comstock and Robinson, 1952; Stuber, 2004; Acquaah, 2007). The design requires sufficient seeds for replicated evaluation trials and hence it is not of practical application in breeding species that are not capable of producing large amounts of seeds. It is applicable to both self and cross pollinated species that meet this criterion. The design is low cost although it is not efficient for selection (LeClerg, 1966). It can be used in self-pollinated crops with multiple flowers (Le Clerg, 1966).

Evaluation of Variance Components

The model for this design for one environment is

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

Where μ is the overall mean, m_i is the effect of the ith male, f_{ii} is the effect of the jth female mated to the ith male, r_k is the replication effect, and e_{iik} is the environmental error. Because the mating design is nested, expected mean squares are obtained by a hierarchical type of design. Also, because of the genetic structure of the mating design, expected mean squares can be expressed in the more useful covariance of relatives.

The "F test" is usually used to test the significance of variance components. Skeleton analysis of variance is given in table 4.

 σ^2_{e} is the error variance and comprises of intra plot environmental variance and genetic variance among individuals of same progeny.

is the variance of plot effects

 $\overset{\sigma^2{}_p}{\sigma^2{}_m}$ is the variance of progeny of different males mated to different females.

 σ_{f}^{2} is the variance of progeny families of different females

The model so developed is known as a "random effects model" because it is the population from which individuals are chosen that is considered for inferences. In this model, it is assumed that the sets, replications in sets, males in sets and females in males in sets are randomly distributed with zero mean and variances σ_{s}^{2} , σ_{r}^{2} , σ_{m}^{2} , and σ_{f}^{2} respectively. Errors are also assumed to be distributed with zero mean but variance σ_{e}^2

Number	of	Mating designs		Number of covariances
factors				among relatives
2	(AA) Diallel	(A/B) NCI	(AB) NC II	2
3	(A(AA)triallel	(A/B/C)	(A(BC)	3 to 7
	(A(AB)	(A/B/C)	(A(B/C)	
	(A(BB)	(A/(BC)		
4	(A(A(AA)	(A(A(BC)	(AA)(AA) quadriallel	4 to 15
	(A(A(AA)	(A(B(CD)	(AA)(AB)	
	(A(B(AA)	(A/B/(CD)	(AA)(BB)	
	(B(A(AA)	(A/B(CD)	(AB)(AB)	
	(B(B(AA)	(A(A(B/C)	(AA)(BC)	
	(B(A(BA)	(A(B(C/D)	(AB)(AC)	
	(C(B(AA)	(A(B/C/D)	(AB)(CD)	
	(C(A(BA)	(A/B/C/D)	(AA)(B/C)	
	(A(C(BA)		(AB)(C/D)	
			(A/B)(C/D)	

Table 1. Classification of ma	ating designs. A, B, C, D represent	groups of parents from	n the same populatio	n (same letter) or from
different populations (different	ent letter). A/B means a hierarchica	I (NC I) and (AB) means	a factorial (NC II) cros	ssing.

Source: Cockerham, 1963.

- F										
	А	В	С	D	E	Е	D	С	В	А
	В	С	D	Е	А	D	E	А	С	В
	С	D	E	А	В	С	Α	В	D	Е
	D	E	А	В	С	В	С	Е	А	D
	Е	А	В	С	D	А	В	D	E	С
		a) Re	ep 1			b) Re	p 2			

Figure 1. Polycross field nursery arrangement showing two Latin squares for 5 genotypes. (Comstock and Robinson, 1952).

Table 2. Outline of ANOVA table for a Biparental mating (paired crosses) design.

Source	Degree of freedom	MS	Expected mean squares comp.
			of variance
Between families	(½n) − 1	MS1	$\sigma^2 w + m \sigma^2_b$
Within families	½n(m – 1)	MS2	$\sigma^2 w$
Total	(nm/2)-1		

n and *m* refer to the number of parents sampled and plants within each cross, respectively. Source: Acquaah, 2007.

		5		
		1	2	3
	Α	х		
	В	х		
4	С	х		
	D		Х	
	E		Х	
	F		Х	
	G			Х
	Η			Х
	1			Х

Figure 2. North Carolina mating design I (Comstock and Robinson, 1952)

Table 3. Format ANOVA table for Polycross m	nating design.
---	----------------

Source	df	Expected mean squares		
		Component of variance	Covariance of relatives	
Between maternal groups	n-1	σ^2_{WM} + s σ^2_{BM+} σ^2_{e}	σ ² _{WM} + s(Cov HS)	
Within maternal groups	n(σ-1)	$\sigma^2_{WM} + \sigma^2_e$	σ^2_{EW} + V _G – Cov HS	
Residual		σ_{e}^{2}		

Source: LeClerg, 1966.

			8			
		1	2	3	4	5
	Α	Х	Х	х	Х	Х
	В	Х	х	х	Х	Х
4	С	Х	х	х	Х	Х
	D	Х	х	х	Х	Х
	Е	х	х	Х	Х	х

Figure 3. North Carolina design II (Comstock and Robinson, 1952)

Table 4. Format of the ANOVA table for North Carolina design I.

		0	
Source	df	MS	Expected MS
Sets	(s-1)		
Replications in sets	s(r-1)		
Males in sets	s(m-1)	M_1	$\sigma_{e}^{2} + k\sigma_{p}^{2} + rk\sigma_{f}^{2} + rkf\sigma_{m}^{2}$
Females in males in sets	sm(f-1)	M ₂	$\sigma_{e}^{2} + k\sigma_{p}^{2} + rk\sigma_{f}^{2}$
Reps x Females	s(mf-1)(r-1)	M ₃	$\sigma_{e}^{2} + k\sigma_{p}^{2}$
Residual	smfr(k-1)	M_4	σ_{e}^{2}
Total	smfrk-1		

Source: Comstock and Robinson, 1952.



Figure 4. Mating scheme for top cross mating design (Acquaah, 2007).

Table 5. F	ormat of the	ANOVA ta	able for l	North (Carolina	Design I	١.
------------	--------------	----------	------------	---------	----------	----------	----

Source	Df	Expected MS
Sets	s-1	
Replications in sets	S(r-1)	
Between males	S(m-1)	$\sigma^2_W + r\sigma^2_{mxf} + rf\sigma^2_m$
Between females	S(f-1)	$\sigma_{W}^{2} + r\sigma_{mxf}^{2} + rm\sigma_{f}^{2}$
Males x females	s(m-1) (f-1)	$\sigma^2_W + r\sigma^2_{mxf}$
Plots within replications	S(mf-1)(r-1)	σ^2_W
Total	Srmf-1	

Source: Le Clerg, 1966.

			8			
		1	2	3	4	5
	1	Х	х	х	Х	Х
Ŷ	2	х	х	х	х	х
	3	Х	х	х	Х	Х
	4	Х	х	х	Х	Х
	5	х	х	х	Х	Х

Figure 5. A full diallel (Kempthorne, 1957)

Table 6. Format of the ANOVA table for the North Carolina design III.

Ŷ

ç

Source	Df	MS	Ems
Replications	(r-1)		
Parents/Testers (T)	1	MS _T	$\sigma_{W}^{2} + r\sigma_{Tm}^{2} + mrk_{T}^{2}$
F ₂ (m)	m-1	MS _m	$\sigma^2_{W} + 2r\sigma^2_{m}$
ТхМ	m-1	MS _{Tm}	$\sigma_{W}^{2} + r\sigma_{Tm}^{2}$
Within FS families	(r-1)(2m-1)	MSw	σ^2_W
Total	2mr-1		

Source: Comstock and Robinson, 1952.

	1	2	3	4	5
1		Х	Х	Х	х
2		•	Х	Х	Х
3			•	Х	х
4					х
5					•

Figure 6. A half diallel (Kempthorne, 1957)

Table 7. Orthogonal comparisons in a Triple Testcross.

Comparison	L_{1i}	L _{2i}	L _{3i}	Divisor
(i)	+1	+1	+1	3
(ii)	+1	+1	-2	6
(iii)	+1	-1	0	2

Source: LeClerg, 1966.

			3			_	
	1	2	3	4	5	6	
1	•	х		х		х	
2			Х	•	Х		
3				Х			
4							
5							
6							

Figure 7. A smart diallel (Kempthorne, 1957)

Tab	le 8	3 . 1	Tripl	e t	estc	ross.	De	rivat	ion	of	the	means	s and	vari	ance
-----	------	--------------	-------	-----	------	-------	----	-------	-----	----	-----	-------	-------	------	------

Means					
Genotype		$A^{+}A^{+}$	A ⁺ A ⁻	A ⁻ A ⁻	Mean
Frequency		1/4	1/2	1/4	
L _{1n} – Mean ge	notypic value	А	½(a+d)	d	½(a+d)
L _{2n} – Mean ge	notypic value	D	½(-a+d)	-a	½(-a+d)
L _{3n} – Mean ge	notypic value	½(a+d)	½d	½(-a+d)	½d
Orthogonal co	omparisons				
(i)	$\frac{1}{3} \sum (L_{1i} + L_{2i} + L_{3i})$	½(a+d)	½d	½(-a+d)	½d
(ii)	$[2L_{3i} - (L_{1i} + L_{3i})]$	0	0	0	0
(iii)	$\frac{1}{2} \sum (L_{1i} - L_{2i})$	½(a+d)	½a	½(a+d)	1⁄2a
Variances					
Variance of 1/3	$\sum (L_{1i} + L_{2i} + L_{3i}) = \frac{1}{2} [\frac{1}{2}(a+d)]^2 + \frac{1}{2} (\frac{1}{2}d)^2 + \frac{1}{2} [\frac{1}{2}(-a+d)]^2 + 1$	$d)]^{2} - (\frac{1}{2}d)^{2} = \frac{1}{8}a^{2}$			
Variance of 1/2	$\sum (L_{1i} - L_{2i}) = \frac{1}{2} [\frac{1}{2}(a-d)]^2 + \frac{1}{2} (\frac{1}{2}a)^2 + \frac{1}{4} [\frac{1}{2}(a+d)]^2 - (a+d)^2 - \frac{1}{2} (\frac{1}{2}a+d)^2 - \frac{1}{2} (\frac{1}{2}a+d)^2 + \frac{1}{2} (\frac$	$(\frac{1}{2}a)^2 = \frac{1}{8}d^2$			
Variance with	in L1	0	¼(a-d) ²	0	
Variance with	in L2	0	¼(a+d) ²	0	
Mean varianc	e within L_1 and L_2 families = $\frac{1}{2}[\frac{1}{8}(a-d)^2 + \frac{1}{8}(a+d)^2]$	$^{2}] = \frac{1}{8}a^{2} + \frac{1}{8}d^{2}$			

Source: LeClerg, 1966

				3				
		1	2	3	4	5	•	n
	1		х	х	х	•	•	
	2		•	х	х	х	•	
Ŷ	3	•	•	•	х	х	х	•
	4	•	•	•	•	х	х	х
	5	х	•	•		•	х	х
	•	х	х	•	•	•	•	х
	n	x	x	x	•	•	•	•
	Fig 8. A partial diallel (Kempthorne, 1957).							

Table 9. Format of ANOVA	table for a	triple cross	design
--------------------------	-------------	--------------	--------

Item		Df	Expected mean squares
Sums	$L_{1i} + L_{2i} + L_{3i}$	n - 1	$\sigma_{w}^{2} + 3\sigma_{RS}^{2} + 3mr\sigma_{S}^{2}$
	$L_{1i} + L_{2i}$	n -1	$\sigma_{w}^{2} + 2\sigma_{RS}^{2} + 2mr\sigma_{S}^{2}$
Differences	L _{1i} - L _{2i}	n — 1	$\sigma_{w}^{2} + 2\sigma_{RD}^{2} + 3mr\sigma_{D}^{2}$
Sums x replicates	$L_{1i} + L_{2i} + L_{3i}$	(n - 1) (r - 1)	$\sigma_{w}^{2} + 3\sigma_{RS}^{2}$
	$L_{1i} + L_{2i}$	(n - 1) (r - 1)	$\sigma_{w}^{2} + 2\sigma_{RS}^{2}$
Difference x replicate		(n - 1) (r - 1)	$\sigma_{w}^{2} + 3\sigma_{RD}^{2}$
Within families		2nr(m – 1)	σ_w^2

Source: LeClerg, 1966.

Diallel 1



Figure 9. A disconnected half diallel (Kempthorne, 1957)

Table 10. Format	of ANOVA table for	r the Line x Tester	mating design

Source	df	M.S	Expected mean square	
		•	Random model	Fixed model
Replications	(<i>r</i> -1)			
Lines	(<i>m</i> -1)	M_1	$\sigma_{e}^{2} + rv_{sca} + rfv_{gca(m)}$	$\sigma_e^2 + rf(1/m-1)\Sigma_i g_i^2$
Testers	(f-1)	M_2	$\sigma_{e}^{2} + rv_{sca} + rmv_{gca(f)}$	$\sigma_e^2 + rm(1/f-1)\Sigma_j g_j^2$
Lines x testers	(<i>m</i> -1)(<i>f</i> -1)	M_3	$\sigma_e^2 + rv_{sca}$	$\sigma_{e}^{2} + r\{1/(m-1)(f-1)\} \Sigma_{i} \Sigma_{j} S_{ij}^{2}$
Residual (error)	(r-1)(<i>mf</i> -1)	M_4	σ_{e}^{2}	σ_e^2

Source: Kempthorne, 1957.

Table 11. Format of ANOVA table when diallel progeny (a genotypes) are planted in b blocks having c plants in each of the ab plots.

Source		Df	MS	Expected mean squares	
				Model I (fixed effect)	Model II (random effect)
Replications		(b-1)	M _b	$\sigma_{e}^{2} + ac\{1/(b-1)\}\Sigma_{k}b_{k}^{2}$	$\sigma_{e}^{2} + c\sigma_{bv}^{2} + ac\sigma_{b}^{2}$
Genotypes		(a-1)	Mv	$\sigma_{e}^{2} + bc \{1/(b-1)\}\Sigma_{i}v_{i}^{2}$	$\sigma_{e}^{2} + c\sigma_{bv}^{2} + bc\sigma_{v}^{2}$
Genotype	х	(a-1)(b-1)	M_{bv}	$\sigma_e^2 + c\{1/(a-1)(b-1)\} \Sigma_i \Sigma_j \Sigma_k (bv)_{ijk}^2$	$\sigma_{e}^{2} + c \sigma_{bv}^{2}$
Replications					
Error		ab(c-1)	M _e	σ_{e}^{2}	σ ² _e

Source: LeClerg, 1966.

 Table 12. Analysis of variance of n parents and their n(n-1)/2 variety crosses for variety and heterosis effects (GEAN II) model

 MS

IVIS				
Source	df	SS	Gardner-Eberhart	Diallel
Replications	r-1			
Entries	[n(n+1)/2]-1	S'2	M'2	M ₂
Varieties (vi)	n-1	S' ₂₁ = (B'G) ₁ -CF	M'21	
Heterosis (hij)	n(n-1)/2	S' ₂₂ =(B'G) ₄ -(B'G) ₁	M'22	
Average (\overline{h})	1	S' ₂₂₁ = (B'G) ₂ -(B'G) ₁	M' ₂₂₁	M ₂₂
Variety (hi)	n-1	S' ₂₂₂ = (B'G) ₃ -(B'G) ₂	M' ₂₂₂	
Specific (sij)	n(n-3)/2	S' ₂₂₃ =(B'G) ₄ -(B'G) ₃	M' ₂₂₃	M ₂₃₂
Error	(r-1){[n(n+1)/2]-1}	S'1	M ₁	M_1
Total	[rn(n+1)/2]-1			

Gardner and Eberhart, 1966.

Table 13. Format for ANOVA table for partial diallel analysis

Source	df	Expected MS	
Replication	r-1		
Crosses	(ns/2)-1	$\sigma^{2+}r\sigma_{c}^{2}$	
Gca	n-1	$\sigma_{e}^{2} + r \sigma_{sca}^{2} + [rs(n-2)/(n-1)] \sigma_{gca}^{2}$	
Sca	n[(s/2)-1]	$\sigma_{e+}^2 r \sigma_{sca}^2$	
Error	(r-1)[(ns/2) -1]	σ_{e}^{2}	
Total	(rns/2)-1		

Source: Griffing, 1956a.

Significance of mean square for males suggests significant differences among the progeny families of different males (because of the genetic differences among the males). Because the progeny families of a given male are half-sibs, the mean squares for males measure variance of half-sib progeny families. The variance of half-sibs is also the covariance of half-sibs, this is equal to $\frac{1}{4} \sigma_{A}^{2}$ (Falconer and Mackay, 1996). Therefore, the significance of the mean square for males provides a direct test of significance of additive genetic variance. On the other hand, when mean square for females in males is significant, it implies that there are significant genetic differences among the plants randomly chosen as the seed parents. Because the experimental material produced by the North Carolina design I is a biparental progeny, total heritable variation of the experimental material is $\frac{1}{2}\sigma_{A}^{2} + \frac{1}{4}\sigma_{D}^{2}$, which is the covariance of full-sib when several loci are taken into consideration (Dabholkar, 1992). Therefore, the heritable variation of progenies of the females can be obtained as the difference between the total variation of the experimental material and the expectation of heritable variation of males $(\frac{1}{4}\sigma_{A}^{2})$. Thus,

 $\sigma_{f}^{2} = (\frac{1}{2} \sigma_{A}^{2} + \frac{1}{4} \sigma_{D}^{2}) - \frac{1}{4} \sigma_{A}^{2}$ $= \frac{1}{4} \sigma_{A}^{2} + \frac{1}{4} \sigma_{D}^{2}$

North Carolina design I is most widely used in animal studies. In plants, it has been extensively used in maize breeding for estimating genetic variances (Acquaah, 2007).

North Carolina Design II (Factorial mating design)

This is a modification of NC design I. However, design II has greater precision, it is more applicable to self-pollinated crops, and has a direct estimate of the level of dominance. In NC design II, each member of a group of males is mated to each member of a group of females (Comstock and Robinson, 1952). The number of entries tested is $f \times m$, where f is the number of females and m is the number of males. It is therefore a rectangular mating design, unless f = m. In this design, every progeny family has half-sib relationships through both a common male and a common female (Stuber, 2004; Figure 3).

COVHS_M=1/4 VA

 $COVHS_F = 1/4 VA$ V female x male = $COVFS - COVHS_M - COVHS_F$

= 1/4 VD

Reciprocal crosses may be carried out to analyse maternal effects. The factorial design is the most important in practical plant breeding programs for selection for testcross performance. The factorial design requires that each female must be able to be mated with several males (no

incompatibilities), there should be synchrony in the time of flowering between the male and female group of plants, and that genetically identical (mother) plants are available (clones, inbred lines, double haploid lines) if the number of female inflorescences per plant is limited (LeClerg, 1966; Namkoong and Roberds, 1974). The design provides good information on parents and full-sib families, provides estimates of both additive and dominance effects, and provides estimates of genetic gains from both V_A and V_D. The North Carolina design II is used to evaluate inbred lines for general combining ability and specific combining ability and to estimate genetic variances. The design is suitably employed in the estimation of average degree of dominance of multi-flowered crop species like cotton, tobacco, sorghum, and linseed. The design is well-suited to multipleflowered plants because each plant can be used repeatedly as both male and female. Blocking is used in this design to allow all the mating involving a single group of males to a single group of females to be kept intact as a unit. For example, if 6 plants are chosen as male parents and mated to 6 plants selected as females, there will be 36 biparental progeny families. In contrast to North Carolina Design I, both paternal and maternal half-sibs are produced in North Carolina Design II. The variation is therefore divided into two parts; between full-sib families and within full-sib families. The variation between families is further divided into components due to differences among males, that due to differences among females and that due to male x female interaction. Consequently, the North Carolina Design II is factorial with respect to the male and female effects. As for NCI the experimental materials originates from matings among F₂ plants from a cross of two inbred lines (LeClerg, 1966).

Field arrangement of North Carolina design II is similar to that for North Carolina Design I. Because we have two sets of parents, we have two independent estimates of GCA. The expectations of males and females for design II are equivalent to GCA, and the male x female source is equivalent to SCA of the diallel analysis. Design II is widely used in maize breeding. Appropriate F-tests can be made to test for the difference among males, among females, and for the interaction of males and females. The structure of analysis of variance for k plants of each progeny family (pf) raised in each of the r "plots", with observations recorded on plant basis is given in Table 5.

Where $\sigma_W^2 = \frac{1}{2} V_A + \frac{3}{4} V_D + V_{EW}$ $\sigma_{2mxf}^2 = \frac{1}{4} V_D + V_{EC}, \sigma_M^2 = \frac{1}{4} V_{A}, \sigma_F^2 = \frac{1}{4} V_A$

 $\sigma^2_{m xf}$ is progeny variance arising from interaction of genotypes of male and female parents.

The mean squares for males and females give separate and independent estimates of the additive component of variation. Furthermore the interaction mean square between male and females yields an estimate of the non-additive genetic variance V_D . Should this item be significant then it serves as for testing the difference between both males and females (Acquaah, 2007).

This mating design generates sufficient information to allow all the parameters to be estimated because of the relatedness among the progeny families; thus:

$$V_A = 2(\sigma^2_M + \sigma^2_F)$$

$$V_D = 4(\sigma^2_{MF} - V_{EC})$$

$$V_{EW} = \sigma^2_W - [(\sigma^2_M + \sigma^2_F) + 3(\sigma^2_{MF} - V_{EC})]$$

 $V_{EW} = \sigma^2_{W} - [(\sigma^2_M + \sigma^2_F) + 3(\sigma^2_{MF} - V_{EC})]$ If individual plant randomisation has been used, $V_{EC} = 0$; alternatively it can be estimated directly from a replicated experiment. Because every male is mated to each female, NCII is a factorial design, similar to a two way ANOVA in which the variation can be partitioned into differences between males and females and the interaction between them. Assessment of GCA of the parents can be done from the average performance of the progeny, and measure the SCA of each individual cross (LeClerg, 1966). The genetic model for NCII is as follows: $Y_{ijk} = \mu + m_i + f_j + (mxf)_{ij} + e_{ijk}$ Where,

Where, $Y_{ijk} = k^{th}$ observation on ixjth progeny $\mu =$ the general mean, m_i =the effect of the ith male f_j =effect of the jth female $(mxf)_{ij}$ =is the interaction effect e_{ijk} = the error associated with each observation.

North Carolina Mating Design III

The design III was developed by Comstock and Robinson (1948). This specific design was made with the purpose to estimate the average level of dominance of genes affecting a trait. However, it also provides good estimates of σ_A^2 and σ_D^2 for F₂ populations assuming absence of linkage and epistasis. The advantage of this design over designs I and II is that the estimation of dominance is not subjected to any assumption regarding allele frequencies. The North Carolina Mating Design III (NCIII) involves backcrossing the F₂ plants to the two parental inbred lines from which the F_2 were derived. The F_2 plants are used as male parents while the two parental inbred lines serve as testers, against which the F2 are assessed. However, being the progenitors of the F_2 , the parental inbred lines are very special testers because the F_2 is segregating at all the loci for which the testers differ but for no other loci. The F₂ plants serve as pollen parents, the number of inbred plants crossed to each F2 should be large enough to ensure sufficient seed for field evaluations. Design III is used infrequently and primarily to estimate the average dominance of genes. Estimates of genetic parameters tend to be more biased from epistasis in this design than in Designs I and II. Design III is by far the most powerful of the North Carolina mating designs, with design I being the least powerful (Acquaah, 2007). The NC III was made even more powerful by the modifications made by Kearsey and Jinks (1968) that adds a third tester (not just the two inbreds). The modification is called the triple testcross and is capable of testing for non-allelic (epistatic) interactions, which the other designs cannot, and also capable of estimating additive and dominance variance (LeClerg, 1966).

The ANOVA for NCIII would be similar to that of NCII, if it was not for the two testers that are not a random sample from a population, but are particular lines i.e. the grandparents of the F_2 . Hence the item "testers" in the ANOVA is treated as a statistically fixed effect. An example is an experiment in which $n F_2$ individuals are all crossed to the two testers and that r full sib progeny are raised from the resulting 2n families in a completely randomized design. If these full sibs are scored for some trait, then the mean of r sibs from the cross between the ith F_2 with P_1 and P_2 can be

presented as L_{1i} and L_{2i} respectively. The format for the analysis of variance for the North Carolina Mating Design III is presented in table 6.

Mean square expectations to focus on in design III are the components of variance among males (σ_m^2) and the one for the interaction of males and inbred parents (σ_{tm}^2) . Because the testers are fixed, the interaction item (T x M) and the main effect (M) are both tested against the basic error as the expected mean squares (*ems*) show. The genetic components of the σ_m^2 of the analysis can be derived by recalling that when we have a 2 x m table, the sums of squares for the analysis of variance is computed from the sums and differences of the paired observations. The same approach is used when deriving the expected mean squares (Comstock and Robinson, 1952; LeClerg, 1966). Direct F tests are possible for σ_{Tm}^2 and σ_m^2 with the error term but they do not give us knowledge of the genetic structure of the progenies and how they relate to components of variance.

Triple Testcross

The NCIII design was extended by Kearsey and Jinks (1968) to include a third tester (L₃). This third tester is the F₁ from the original parents L₁ X L₂. In this extended form, this design is known as the triple test cross (TTC) (LeClerg, 1966). The inclusion of the third tester increases the power of this design considerably because it provides a sensitive and unambiguous test for non-allelic interactions. Both in its original form and extended form, NC design III has a general utility for investigating any population, irrespective of gene frequency or mating system. In a triple test cross a random sample of n individuals from the population under investigation is crossed to the same three testers L₁, L₂ and L₃, to give 3n progeny families. as shown in table 7.

The analysis of this design may be divided into two parts; the first part provides a test for epistasis, the second assesses the significance and provides estimates of the additive and dominance components of variation. Using a single gene model A^+ .. A^- , segregating in an F₂ population, the analysis is shown in table 8.

From table 8, the first comparison (if only two parental testers are used) detects additive, the second epistatic and third dominance variations. The difference $[2L_{3i} - (L_{1i} + L_{3i})]$ is computed for all *n* individuals sampled in the population, the additive and dominance components cancel, leaving only epistatic terms as shown in table 8. This applies for number of loci and interactions of any complexity (Kearsey and Jinks, 1968). For any base population $2L_{3i} - L_{1i} - L_{2i}$ is not expected to differ significantly from zero in the absence of epistasis. The variance of this difference over all *n* individuals should not therefore be significantly greater than the experimental error if non-allelic interactions are absent. From table 7, the heritable variance of $\frac{1}{3}\sum(L_{1i} + L_{2i} + L_{3i})$ or $\frac{1}{2}\sum(L_{1i} + L_{2i}) = \frac{1}{8}a^2$, which upon summation over all loci gives ½∑a². Likewise, the heritable variance of $\frac{1}{2}(L_{1i} - L_{2i})$, the last comparison, is $\frac{1}{2}d^2$ after summation. The analysis, with r replicates and m individuals per family per replicate, can be set out as shown in table 9.

Where the variance within families
$$\begin{split} \sigma^2_{\ w} &= \mathscr{Y}_s \sum a^2 + \mathscr{Y}_s \sum d^{2+} V_{EW} \\ \sigma^2_{\ s} &= \mathscr{Y}_s \sum a^2 \\ \sigma^2_{\ D} &= \mathscr{Y}_s \sum d^2 \end{split}$$

 $\sigma_{RS}^{2} = \sigma_{RD}^{2} = V_{EC'}$ from which $V_{EW} = \sigma_{w}^{2} - \sigma_{s}^{2} - \sigma_{D}^{2}$

If the analysis is conducted on progeny means, the expectations of mean squares, together with the sum of squares within families, should be divided by *m*, the number of plants per family per replicate. Otherwise the expectations as shown apply. When extended to populations of unknown allelic frequencies, $\sigma_{\rm S}^2$ and $\sigma_{\rm D}^2$ still yield estimates of the genetical, as opposed to the statistical, components of variation.

The immediate advantages of this design are two:

• It gives a precise test of significance for non-allelic interactions

• In the absence of such interactions both the additive and dominance components of variation are estimated with comparable precision, a property shared with no other mating design.

Line x Tester Mating Design

The Line x Tester mating design is a modification of NC II. The Line x Tester mating design uses inbred lines as the base population. Randomly chosen m inbreds are used as males (referred to as lines) and mated to each of the randomly selected f inbreds, used as females (designated as testers), to generate mf progeny families which are usually F_1 hybrids. This mating design permits mating all the m inbred lines to all the f testers; therefore, it is not restricted to the multiflowered crop species (Comstock and Robinson, 1952; LeClerg, 1966).

This mating design is useful in deciding the relative ability of a number of female and male inbreds to produce desirable hybrid combinations. Consequently, line x tester designs permit evaluation of general combining abilities (GCA) of lines crossed to a set of testers as well as the general combining abilities of the latter; crossing of m lines to ftesters produces mf hybrids, which provide information about the general combining ability effects of m + fgenotypes and specific combining ability effects of mf crosses. Additionally, line x tester design provides information regarding the usefulness of male and female inbreds as parents for hybridization to generate segregating populations which are expected to give prodigious selections. In crops like wheat, maize and cotton, where emasculation and pollination are not difficult, line x tester design could be employed to screen a large number of genotypes for their combining ability (Kempthorne, 1957).

Line x Tester mating design usually assumes a fixed effects model; however, the initial mating design proposed by Kempthorne (1957) is based on random effects model. The field arrangement of experimental material, comprising of mf progeny families, is usually a randomized block design with r replications. The analysis of variance outline for the line x tester mating design is presented in table 10.

Where *r* = number of replications

m = number of male parents (lines)

f = number of females (testers)

 V_{qca} = Covariance half-sibs

V_{sca} = Cov. Full-sibs – 2 Cov. Half-sibs

The various Mean Squares (MS) are calculated as:

Individuals within each of the mf progeny families are fullsibs. Heritable variation of the mf progeny families is, therefore, covariance of full-sibs. i.e. $\frac{1}{2} \sigma_A^2 + \frac{1}{2} \sigma_D^2$, if nonallelic interaction is assumed to be absent. Since each line is crossed to all testers, progeny families belonging to each line are half-sibs. Variation of lines is therefore, covariance of half-sibs, i.e. $\frac{1}{3} \sigma_A^2$. Similarly, variation of progeny families of different testers is covariance of half-sibs, i.e. $\frac{1}{3} \sigma_A^2$. Therefore, the variation attributable to lines x testers is gotten as a difference. Thus,

 $\frac{1}{2}\sigma_{A}^{2} + \frac{1}{4}\sigma_{D}^{2} - \frac{1}{4}\sigma_{A}^{2} - \frac{1}{4}\sigma_{A}^{2} = \frac{1}{4}\sigma_{D}^{2}$

Significance of mean square for line x testers provides a direct test of significance of dominance variance. Significance of σ_A^2 is provided by significance of lines and testers mean squares.

Topcross mating design

The topcross mating scheme involves the crossing of a number of selections, lines, or clones to a common parent (tester) which may be a cultivar, an inbred line, a single cross etc., where the tester is the same for each mating. Because a common tester is used for all crosses, all progeny families produced are half-sibs; therefore, topcross mating design permits the evaluation of GCA for the group of lines, clones, or selections involved in the crosses. The topcross mating design is mainly used in cross-pollinated crops such as maize where it is commonly an inbred-cultivar cross. Additionally, the design is used for initial evaluations of breeding potentials in new maize accessions (Stuber, 20004). The mating scheme is illustrated in figure 4. The tester serves as the male parent while the lines or clones to be tested serve as females. The ratio of number of male rows to female rows varies with species; in maize, a ratio of 1 male row to 2 female rows is adequate for effective topcrossing. Alternating rows may be required in crops such as sugar beet whereas two male rows may be adequate to pollinate 10 to 12 female rows in castor. It may be necessary to delay planting dates if the male and female parents differ in days to flowering. If wind is a factor in pollen dissemination, tester rows should be planted perpendicular to prevailing winds. If isolation is not possible, the tester may be planted in paired rows with the materials to be evaluated. Hand pollinations are made between rows and the tester may be used as either the male or the female parent. Replication and randomization of the female genotypes may help to minimize the effects of inter-crossing among females (Stuber, 2004). If there are wide differences in days to flowering among the materials under study, synchronization of flowering may be difficult to achieve in paired rows. However, pollination may be achieved by planting the materials to be evaluated in separate blocks according to the number of days to flowering (Acquaah, 2007).

Diallel mating design

When the same parents are used as females and males in breeding, the mating design is called diallel (Griffing 1956a). If properly analysed, it is a very powerful design. The full diallel mating design involves crossing of parents in all possible combinations including parents and reciprocals (Darbeshwar, 2000; Stuber, 2004). This is the kind of mating scheme required to achieve Hardy-Weinberg equilibrium in a population (Acquaah, 2007). The parents range from inbred lines to broad genetic base varieties to clones. The diallel mating design has been used and abused more extensively than any other for developing breeding populations for recurrent selection (Jenkin, 1934; Stuber, 2004). The diallel mating design is the most commonly used in crop plants to

estimate GCA, SCA and variances, and it is also widely used for developing breeding populations for recurrent selection (Sprague and Tatum, 1942; Acguaah, 2007). The mating combinations for a full diallel of *n* parents is n^2 crosses which includes $\frac{1}{2}n(n-1) = F_1 + \frac{1}{2}n(n-1)$ reciprocals + n parents. With single crosses (i.e. without parents and reciprocals) the number is n(n-1)/2 crosses. In practice, a diallel with reciprocals is necessary if the character is maternally inherited (Acquaah, 2007). The number of possible crosses increases rapidly with an increase in parents e.g. for n=5, the number of single crosses is 10; for n=15, it is 105; and for n=50, it is 1225 (LeClerg, 1966). Therefore with limited facilities available for testing, a diallel cross may only be possible for a relatively small number of inbred lines. However, if only a small number of lines are included in a full diallel, the estimate of the variance of the general combining abilities in the whole populations of potentially available inbred lines is subject to a large sampling error; in addition, many potentially high yielding inbred lines may be left out completely untested. In mitigation, methods of design and analysis which utilize only part of the possible matings (called the partial diallel cross) has been formulated (Griffing, 1956a).

Nursery arrangements for the application of full and partial diallel are varied. Because large numbers of crosses are made, diallel mating design takes a large amount of space, seed, labour and time to conduct. Because all possible pairs are contained in one half of a symmetric Latin square layout, this layout may be used to address some of the space needs. The diallel design requires that each female must be able to be mated with several males (no incompatibilities and there should be synchrony in the time of flowering), and that genetically identical (mother) plants are available (clones, inbred lines, double haploid lines) if the number of female inflorescences per plant is limited (Comstock and Robinson, 1952). The diallel mating design provides information on the nature and magnitude of genetic parameters, and GCA and SCA of parents and their crosses (Griffing, 1956a).

Cockerham (1963) have discussed analysis of variance with diallel mating design for fixed effects model (Model I) (where the parents are the genotypes under consideration) and for random effects model (Model II) (where the parents are a sample of genotypes from a reference population). For model I estimation of components of variance is not adequate but estimation of GCA and SCA effects is valid. This model is very useful for choice of parents (GCA) and/or hybrids (SCA). Hybrid seed companies are always searching for the best SCA.

1.1 Analysis of diallel mating design

Although several strategies for diallelic analysis have been proposed, few of them are commonly applied: Griffing (1956b) and Gardner and Eberhart (1966) methods are the most frequently applied. The main reasons that justify the widespread use of the Griffing (1956b) method are its generality since the parents can be clones, pure lines, inbred lines, or populations of a self-pollinated, cross-pollinated or intermediate species, and the ease of analysis and interpretation. Although diallel designs were originally intended for crosses between homozygous lines, they may be applied to crosses between heterozygous clones (Kearsey and Pooni, 1996).The linear model for combining abilities for the half and full diallel considering $p(p-1)F_{1\,s}$ is written as:

 $x_{ij} = \mu + gca_i + gca_j + sca_{ij} + r_{ij} + error$ for observations in the cross $i \times j$, where $i = 1, 2, \dots, p$ u = population mean

 gca_i = general combining ability of the i th line or variety sca_{ij} = specific combining ability of the cross $i \times j$ $(sca_{ij} = sca_{ji})$

 r_{ij} = reciprocal effect of the cross $i \times j$ ($r_{ij} = -r_{ii}$)

Griffing Approach

The Griffing (1956b)'s approach considers four methods of diallel analysis that vary in either omission of parents or the reciprocals in the crosses. In each method two steps are involved in the analysis of data: The first step involves analysis of data for testing the null hypothesis that there are no genotypic differences among the F₁'s, parents and reciprocals. Only when the significant differences among these are established, do we proceed to the second step of analysis i.e. the combining ability analysis. In method 1 the number of progeny families (pf) comprises of all the n^2 progeny families, i.e. the *n* parents, $n(n-1)/2 = F_1$ and n(n-1)/2reciprocals; method 2 includes *n* parents and n(n-1)/2 F₁, i.e. n(n+1)/2 families; method 3 uses F₁ and their reciprocals i.e. n(n-1)/2 F₁ and n(n-1)/2 reciprocal to give n(n-1) progeny families while method 4 uses F_1 only i.e. n(n-1)/2 progeny families, to constitute the experimental material. For convenience in data management, a randomized block design is used for planting the experimental material developed from the above-mentioned methods. In using a randomized block design, it is assumed that there are 'a' varieties (i.e. genotypes developed by the diallel crossing method), each of which is assigned at random to each of the b blocks, and that there are c individuals in each of the ab plots. The mathematical model for the *ijkl*th observation is expressed as:

 $Y_{ijkl} = \mu + v_{ij} + b_k + (bv)_{ijk} + e_{ijkl}$

Where μ is the population mean, v_{ij} is the effect of *i* xj^{th} genotype, b_k is the effect of k^{th} block, $(bv)_{ijk}$ is the interaction of ij^{th} genotype with k^{th} block, e_{ijkl} is the environmental effect peculiar to $ijkl^{th}$ observation.

Regarding combining ability analysis, which does not include reciprocals, variety effect v_{ij} is expressed in terms of general and specific combining ability effects as:

 $v_{ij} = g_i + g_j + s_{ij}$

However, when reciprocal cross is included, variety effect is expressed as:

$$r_{ij} = g_i + g_j + s_{ij} + r_{ij}$$

where g_i and g_j are the general combining ability effects of i^{th} and j^{th} lines, respectively, s_{ij} is the specific combining ability effect of ij^{th} cross and r_{ij} is the reciprocal effect associated with ij^{th} cross.

In terms of the covariance among relatives, the diallel family structure is identical to a NCII, with progeny being related either as full or as half-sibs. The main difference between them is that in the diallel, where reciprocals are included (methods 1 and 3) the same set of parents act as males and females (Hill et al., 1998).

Griffing (1956b) suggested two models for combining ability analysis of the experimental progenies developed by the four different methods listed above. In model I, variety and block effects are regarded as constants; while in model II, both variety and block effects are assumed to be random variables. If the 'a' genotype produced by diallel mating according to any of the 4 methods proposed by Griffing (1956b) are planted in a randomized block design with b blocks such that there are c plants per plot of each progeny family within each block, the format for the analysis of variance is presented in table 11.

For model I, the experimental material is the population from which the inferences are to be made. In this model it is possible to compare the combining abilities of parents, when the parents themselves are used as testers, in addition to identifying high performing combinations. Therefore, model I aims at estimating combining ability effects of parents and their crosses and computing appropriate standard errors for differences between the effects. It is assumed that eiikl are identical, independent and normally distributed with mean zero ($\mu_e = 0$) and variance, $\sigma_{e..}^2$ Model I implies that parents were not randomly chosen or that there was bias in choosing parents. Model II on the other hand assumes dealing with random samples from some parent population and inferences are not drawn about individual lines in the sample but about the parameters in the parent population. The objective of model II is to estimate genetic and environmental components of complex population variance. It is assumed that all the effects in this model except μ , are normally, independently distributed with mean zero (i.e. μ_{b} =0, $\mu_g = 0$, $\mu_s = 0$, $\mu_r = 0$) and variance σ_{θ}^2 where θ refers to b or g or s or r. The variance components are estimated by translating the observed mean squares to their expected values. Standard errors for estimates of variance components are calculated from variance of appropriate mean squares (Kempthorne, 1957; LeClerg, 1966).

The Gardner-Eberhart analysis

Gardner and Eberhart (1966) proposed alternative analysis of data from diallel crosses produced from heterogeneous parents/populations ("varieties"). Gardner-Eberhart Analysis (GEAN) I is resource intensive, requiring the evaluation of the *n* parents, n(n-1)/2 F1 crosses, and inbred progeny of parents and crosses. It provides information on additive and dominance gene action, heterosis, and inbreeding depression (Gardner and Eberhart, 1966). The GEAN II is useful in evaluating n populations (varieties) and their n(n-1)1)/2 F1 crosses; variation among populations is partitioned into populations and mid-parent heterosis (Gardner and Eberhart, 1966; Murray et al., 2003). Heterosis is further partitioned into average, variety, and specific heterosis, but additive and dominance parameters cannot be determined in GEAN II because they are confounded with the source of variation called labelled "variety" (Murray et al., 2003; Table 12).

The following models are used to determine the sums of squares for the analysis shown in table 12:

 $(B'G)_1 = \mu + (1/2)(v_j + v_j')$

 $(B'G)_2 = \mu + (1/2)(v_j + v_j') + v\overline{h}$

 $(B'G)_3 = \mu + (1/2)(v_j + v_j') + v\overline{h} + v(h_j + h_j')$

 $(B'G)_4 = \mu + (1/2)(v_j + v_j') + v\bar{h} + v(\underline{h}_j + \underline{h}_j') + vs_{jj}'$

In each of the models, μ , v_{j} , \bar{h} , h_{j} , and s_{ij} indicate the mean, variety and heterosis effects. Because the phenomenon of heterosis is important, the analysis maximizes the information on variety performance and the expression of heterosis of their crosses. Estimates of the variety and heterosis effects can be determined for each of the constants in the models.

The GEAN III provides estimates of variety and GCA effects from an analysis that contains the following sources of variation: parents, parents versus F1 crosses, and F1 crosses. GCA effects are estimated in a manner similar to Griffing's method 4, Model I (Murray et al., 2003). Both GEAN II and III provide estimates of average heterosis and SCA (Gardner and Eberhart, 1966). Murray et al. (2003) pointed that formulas for the effects for GEAN II and GEAN III were nonintuitive both biologically and genetically and they incorporated the number of parents as multipliers. They further opined that (i) the variety effects obtained in GEAN III were "unconstrained" estimates whereas those from GEAN II were constrained estimates because of the assumption of no heterosis; and (ii) the results had implications for the use and interpretation of such effects.

Some of the commonly used diallel mating designs:

Full diallel

Here each parent is mated with every other parent in the population, including selfs and reciprocal. The number of progeny families (pf) comprises of all the n^2 progeny families, i.e. the *n* parents, n(n-1)/2 F₁ and n(n-1)/2 reciprocals (Figure 5)

Half diallel

Each parent is mated with every other parent, excluding selfs and reciprocals. The number of progeny families (pf) is n(n-1)/2 (Figure 6).

Smart diallel

The parents are sorted based on their breeding values from the best to the worst, and most crosses are made among the best (Figure 7).

Partial Diallel

The partial diallel design was developed by Kempthorne and Curnow (1961). This design is a modification of the diallel and its purpose is to increase the number of parents that can be included in a diallel mating design. The mechanical procedures for developing crosses and the principles of analyzing data are similar to those of complete diallel. The major difference between partial and complete diallel is the number of crosses made among parents. For the diallel, n(n-1)/2 combinations of crosses are made among *n* parents whereas for the partial diallel less crosses are made. Therefore, an important advantage of partial over full diallel is the estimation of genetic variances with affixed number of resources i.e. greater number of parents. The partial diallel has a total of ns/2 crosses, where n is the number of parents, s is whole number greater than or equal to 2, and k is a whole number (k = (n+1-s)/2). The partial diallel consists of selected subsets of full diallels. With a large number of inbred lines, it may be necessary to raise only a sample of all possible crosses among them. Through partial diallel, a plant breeder is not only able to estimate GCA of a large number of parental lines but can also make selection among crosses from a wide range of parents. General combining ability of each line may be estimated with relatively low precision but larger genetic gains may result from the more intense selection that can be applied to them (Kempthorne and Curnow, 1961). In a normal diallel mating, any given line is involved in (n-1) crosses. With partial diallels it is possible to draw valid inferences from s crosses, sampled from crosses where n lines are involved in a diallel mating. The value of s

is so chosen that it is less than (n-1). Both s and n should not be odd or even at the same time. With *n* lines, the total number of crosses to be analyzed in partial diallel is thus ns/2 (Gilbert, 1958) (Figure 8). The ANOVA for partial diallel is similar to that of diallel mating design (Table 13). The experimental design model, the genetic model, the analysis of variance, and the covariances of the relatives of partial diallel are the same as for diallel. However, the degrees of freedom and coefficient of expected mean squares are different because of the sampling of crosses among parents. Therefore, in addition to including a greater number of parents, the partial diallel also has the advantage of having a more even distribution of degrees of freedom for GCA and SCA, because more parents are included for a given number of crosses. As a consequence, components of variance for GCA and SCA are obtained with similar precision while in the diallel the degrees of freedom are smaller for the GCA mean square in relation to the SCA mean square. Since its introduction, the partial diallel has had limited testing (Hallauer et al., 2010). For a small number of selected parents, the partial diallel will not provide any more information than can be obtained from the complete diallel. For a larger number of parents, North Carolina design II is simper to use than partial diallel. The partial diallel provide another alternative, but the design does not seem to have potential for extensive use (Hallauer et al., 2010). Partial diallels are also used to test single crosses among selected inbred lines. Data from partial diallels are used to in best linear unbiased prediction (BLUP) and best linear unbiased estimation (BLUE) analyses to predict the untested single crosses, which are tested in future trials. The partial diallel mating design would be a more appropriate mating design, in most instances, than the diallel for estimation of genetic components of variance with similar accuracy (e.g. GCA and SCA) because a greater number of parents can be included, for the resources available, to have a better sample that is representative of the genetic variation of the population.

Disconnected half diallel

The half-diallel mating is repeated for the second diallel group. Sometimes crosses are made between parents from two diallels to have connection between two groups (Figure 9).

Advantages and drawbacks of diallel mating designs

1. Diallel designs provide good evaluation of parents and full-sib families.

2. They provide estimates of both additive and dominance genetic effects,

3. They provide estimates of genetic gains from both additive and non-additive genetic variance,

4. When the number of parents mated increases, the number of crosses increases by 2N, where N is the number of parents; the design can be costly.

5. Using the same parents as males and females make the mating design a little bit complicated to analyze.

Discussion and Conclusion

The choice of mating design should be one that detects additive genetic variation at the outset of the breeding programme with unselected material because non-additive effects become prominent as selection continues. This is because selected material has greater similarity thereby eliminating additive effects (Bernado, 2002; Hill et al., 1998). The larger the sample of parents used in a cross, the closer the sample statistic estimates the population value; the BIPs use the most parents, in fact 10 times as many as does the diallel (Kearsey, 1965). Hill et al., (1998) roughly summarized the mating designs in two ways: in terms of coverage of the population; BIPs/paired crosses > NC1 > Polycross > NCII > NC II > Diallel in order of decreasing effectiveness. However, in terms of amount of information which each mating design supplies; full diallel > NC II > NC III > NC I > BIPs. The diallel mating design is the most important for estimation of GCA and SCA.

Other considerations in the choice of a mating design are the reproductive potential of a species, for example, a diallel cross should be considered only if a single cross produces an abundant supply of F_1 seed. In species where plants do not produce sufficient flowers for a single plant to be used as a recurrent female parent, the BIPs and NCI techniques are eminently suitable because only one female per progeny family is required. (Kearsey, 1965). Other practical considerations such as cost and labour requirements also have to be borne in mind when choosing a mating design. For the breeders immediate purpose, the topcross and polycross that provide information on GCA and SCA would suffice (Hill et al., 1998).

The study of gene action has been approached in two ways: 1) through the characterization of the predominant types of genetic variance (additive vs. dominant) in populations, leading to the development and analysis of mating designs including North Carolina mating designs. 2) Generation mean analysis is the most prominent approach to determine gene action in species that are difficult to cross and thus the variance component approach cannot be used (Lamkey and Lee, 1993). Dominance has been used as a measure of heterosis if epistatic effects are negligible in the diallel analysis (Moreno-Gonzalez and Dudley, 1981). Since epistasis is computed after taking out the additive and dominance variation, epistasis effects are usually much smaller than additive or dominance effects (Moreno-Gonzalez and Dudley, 1981).

For estimation of GCA and SCA effects, none of the designs supply as much information as the diallel analyses by Griffing (1956b) or Gardner and Eberhart (1966). The diallel however should be restricted to the production of F_1 hybrid cultivars (Hill et al., 1998). Hill et al. (1998) emphasized that majority of the available designs supply statistical rather than genetical estimates of the components of variation. Only the triple test cross and Griffing's methods 1 and 2 of the diallel cross, after redefinition of the components, are capable of supplying genetical estimates.

In terms of the statistics obtained, BIP provide the fewest statistics (two) and hence allow only two parameters to be estimated. This only permits an approximate breakdown of the phenotypic variance into a genetic and environmental component. However, one can partition the genetic variance using the covariance of BIP family means on parental values (Kearsey, 1965). The North Carolina designs and the diallel all yield three statistics by allowing the between families statistics to be subdivided. This allows the genetic and environmental components to be separated and the genetic component to be broken down into additive and nonadditive components. Furthermore NCII allows one to discriminate between maternal and paternal arrays and so obtain a measure of maternal effect (Kearsey, 1965). Among the NC designs, NCI is the least restricted design for analysing all kinds of populations on plants and animals. However, it does yield the minimum genetic information and does not allow an independent test for the dominance effects. North Carolina II is more restrictive in terms of the material and cannot be applied to those species that produce either one or two progeny per female or are difficult to cross. However, it allows independent tests and estimates of the additive, dominance and maternal effects. The NCIII and TTC designs are more suited for studying the genetic control of variation in a cross of two pure lines. The designs provide tests and estimates of all major sources of phenotypic variation and allow for determination of the relative importance of the additive, dominance and epistatic effects. These designs are rather costly and are therefore restricted to a few species which are highly reproductive and easily crossable (Kearsey and Pooni, 1996).

The diallel design is more effective in measuring the additive variation than generation mean analysis of individual crosses. This is due to the fact that differences between the means of the two parents of each cross are the source of additive variation in the generation mean analysis; while some portions of the means of all generations and all crosses are responsible for additive variation in the diallel analysis (Moreno-Gonzalez and Dudley, 1981). The diallel and generation mean analysis have had agreeable results in showing the major importance of dominance in heterosis for grain yield in maize (Moreno-Gonzalez and Dudley, 1981). The basic generations also provide an inefficient design for obtaining reliable estimates of the components of variances, particularly of V_D (Kearsey and Pooni, 1996). The diallel is a design that is not very suitable for studying variation in a segregating population because the parental sample sizes are generally too small to be reliable. Its value comes in the analysis of crosses between inbred lines. The diallel and the NCII both allow the estimation of V_{A} and $V_{D}.$ Nevertheless, NCII has two advantages over diallel: First, NCII design can be used to estimate genetic variances in one population or in the cross between two populations. In contrast, a diallel can be used to estimate V_A and V_D in one population only. Secondly, fewer crosses need to be made and evaluated with NCII design than with a diallel. For example, if p parents are divided equally into male and female parents, i.e., $m = f = \frac{1}{2}p$, with p =10 parents, a diallel requires 45 crosses while NCII would only require 25 crosses. The degrees of freedom for estimating V_A are only slightly less with NCII than with a diallel if the total number of parents across sets is large. For these reasons, NCII is superior to a diallel for estimating V_A and V_D , despite the wider use of the diallel (Bernado, 2002).

From the ongoing, it must be noted that it is not the mating design *per se*, but rather the breeder who breeds a new cultivar. The implication is that the proper choice and use of a mating design will provide the most valuable information for breeding (Acquaah, 2007).

References

- Acquaah G (2007) Principles of plant genetics and breeding. Blackwell Publishing Ltd. Malden, MA.
- Bernado R (2002) Breeding for quantitative traits in plants. Stemma Press, Minnesota, USA.
- Cockerham CC (1963) Estimation of genetic variances. p. 53-94. *In* statistical genetics and plant breeding. National
- Academy of sciences, NRC.

- Comstock RE, Robinson HF (1948) The components of genetic variance in populations of biparental progenies and their use in estimating the average degree of dominance. Biometrics. 4:254-266.
- Comstock RE, Robinson HF (1952) Estimation of average dominance of genes. p. 494-516. *In* Heterosis. Iowa State College Press, Ames, Iowa.
- Dabholka AR (1992) Elements of biometrical genetics. 1st ed. Concept Publishing Company, New Delhi, India.
- Darbeshwar R (2000) Plant Breeding. Analysis and exploitation of variation. Narosa Publishing House, New Delhi, India.
- Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics. 4th ed. Pearson Prentice Hall, Harlow, England.
- Frandsen HN, Frandsen KJ (1948) Observations on the inheritance of resistance to *Drechslera dictyoides* (*Helminthosporium* dict. Drechsl.) in a population of meadow fescue (*Festuca pratensis* Huds). Acta Agriculturae. 31: 91 – 99.
- Gardner CO, Eberhart SA (1966) Analysis and interpretation of the cross diallel and related populations. Biometrics. 22:439–452.
- Geraldi IO, Miranda Filho JB (1988) Adapted models for the analysis of combining ability of varieties in partial diallel crosses. Review of Brasilia Genetics. 11: 419-430.
- Gilbert N (1958) Diallel crosses in plant breeding. Heredity. 12:477–492.
- Griffing B (1956a) Concept of general and specific combining ability in relation to diallel crossing systems. Australian Journal of Biological Science. 9: 463–493.
- Griffing B (1956b) A generalized treatment of the use of diallel crosses in quantitative inheritance. Heredity. 10:31-50.
- Hallauer AR, Carena MJ, Miranda Filho JB (2010) Quantitative genetics in maize breeding. Springer:, New York, Dordrecht, Heidelberg, London.
- Hayman BI (1954a) The analysis of variance of diallel tables. Biometrics. 10:235-244.
- Hayman BI (1954b) The theory and analysis of diallel crosses. Genetics. 39: 789-809.
- Hayman BI (1958) The theory and analysis of diallel crosses. II. Genetics. 43: 63-85.
- Hill J, Becker HC, Tigersteddt PMA (1998) Quantitative and ecological aspects of plant breeding. Chapman and Hall, London.
- Jenkins MT (1934) Methods of estimating the performance of double crosses in corn. Journal of the American Society of Agronomy. 26: 199–204.
- Jinks JL, Hayman BI (1953) The analysis of diallel crosses. Maize Genetics Cooperation News Letter. 27: 48-54.
- Johnson GR, King JN (1998) Analysis of half diallel mating designs: A practical analysis procedure for ANOVA approximation. Silvae Genetica. 47: 74-79.
- Kang H (1989) Inbreeding effective population size under some artificial selection schemes. II. Normal distribution of breeding values. Forest Science. 35:303–318.
- Kang H, Namkoong G (1988) Inbreeding effective population size under some artificial selection schemes. I. Linear distribution of breeding values. Theoretical and Applied Genetics.58:181–191.
- Kearsey MJ (1965) Biometrical analysis of a random mating population: a comparison of five experimental designs. Heredity. 20:205-235.

- Kearsey MJ (1970) Experimental sizes for detecting dominance variation. Heredity. 25: 529-542.
- Kearsey MS, Jinks JL (1968) A general method of detecting additive, dominance and epistatic variation for metrical traits. I. Theory. Heredity. 23:403-409.
- Kearsey MJ, Pooni HS (1996) The genetical analysis of quantitative traits. Chapman and Hall, London, UK.
- Kempthorne O (1957) An introduction to genetic studies. John Wiley and Sons, Inc., New York.
- Kempthorne O, Curnow RN (1961) The partial diallel crosses. Biometrics. 17: 229-250.
- Klein TW, DeFries JC, Finkbeiner CT (1973) Heritability and genetic correlation: Standard errors of estimates and sample size. Behaviour Genetics. 3:355–364.
- Koutsika-Sotiriou MS, Karagounis CA (2005) Assessment of maize hybrids. Maydica. 50:63-70.
- Lamkey KR, Lee M (1993) Quantitative genetics, molecular markers, and plant improvement, In B. C. Imrie and J. B.
 Hacker, (eds.) Focused plant improvement: Towards responsible and sustainable agriculture. Proceedings of the 10th Australian Plant Breeding Conference. Organizing committee, Australian Convention and Travel Service: Canberra, Golf Coast.
- Le Clerg EL (1966) Significance of experimental design in plant breeding. p. 243-313. *In* Plant breeding. The Iowa State University Press, Ames, Iowa.
- Lindgren D, Matheson AC (1986) Increasing the genetic quality of seed from seed orchards by using the better clones in higher proportions. Silvae Genetica. 35:173–177.
- Lindgren D, Libby WS, Bondesson FL (1989) Deployment to plantations of numbers and proportions of clones with special emphasis on maximizing gain at a constant diversity. Theoretical and Applied Genetics 77:825–831.
- Mather K, Jinks JL (1982) Biometrical Genetics, 3rd ed. Chapman & Hall, London.

- Mayo O (1987) The theory of plant breeding. Oxford University Press, New York.
- Miranda Filho JB, Geraldi IO (1984) An adapted model for the analysis of partial diallel crosses. Review of Brasilia Genetics. 7: 677-688.
- Moreno-Gonzalez J, Dudley JW (1981) Epistasis in related and unrelated maize hybrids determined by three methods. Crop Science. 21:644-651.
- Murray LW, Ray IM, Dong H, Segovia-Lerma A (2003) Clarification and re-evaluation of population-based diallel analyses: Gardner and Eberhart analyses II and III revisited. Crop Sci. 43:1930–1937.
- Namkoong G, Roberds JH (1974) Choosing mating designs to efficiently estimate genetic variance components for trees.
 I. Sampling errors of standard analysis of variance estimators. Silvae Genetica. 23:43–53.
- Nasoetion AH, Cockerham CC, Matzinger DF (1967) Simultaneous selfing and partial diallel test crossing II. An evaluation of two methods of estimation of genetic and environmental variance. Biometrics. 23:325–334.
- Pepper WD (1983) Choosing plant-mating design allocations to estimate genetic variance components in the absence of prior knowledge of the relative magnitudes. Biometrics. 39:511–521.
- Sleper DA, Poehlman JM (2006) Breeding field crops, 5th ed. Blackwell Publishing Professional. 2121 State Avenue, Ames, Iowa.
- Sprague GF, Tatum LA (1942) General versus specific combining ability in single crosses of corn. Journal of American Society of Agronomy. 34:923-932.
- Stuber CW (2004) Breeding: Mating Designs. Encyclopedia of Plant and Crop Science. 1:225–228.
- Tysdal HM, Kiesselbach TA, Westover HL, (1942) Alfalfa breeding. Nebraska Experimental Station Research Bulletin, 124:1-6. Lincoln, Nebraska.