

Estimates of Epistasis for Diverse Peanut Cultivars^{1, 2}

T. G. Isleib, J. C. Wynne* and J. O. Rawlings³

ABSTRACT

Progeny from a six-parent half-diallel of diverse peanut (*Arachis hypogaea* L.) cultivars were evaluated for the F₁ through F₅ generations for presence of epistatic effects. Significant variability attributable to specific combining ability (SCA) persisted over generations for yield and other seed characters. Epistasis was indicated since dominance could account for little of the variance due to SCA in the F₅ generation. Estimates of dominance and epistatic variance components were obtained using an iterative weighted least squares procedure. For all characters measured, estimates of epistatic variance were larger than those of dominance variance. It was concluded that considerable epistatic variance may exist in populations derived from crosses of diverse peanut lines.

Key Words: *Arachis hypogaea*, genetic variance, combining ability, exotic germplasm.

The effectiveness of selection for any quantitative trait is largely determined by the nature of the genetic effects governing the trait. In peanuts (*Arachis hypogaea* L.) and other self-pollinated crop species, the ultimate goal of a selection program is usually a homozygous line. Dominance effects are not fixable in a pure line, and in predicting genetic gains under selection, breeders use narrow-sense heritability estimates which do not include the portion of phenotypic variation due to dominance effects. However, other nonadditive effects are fixable, specifically additive x additive and higher order epistasis, causing deviation from additivity of single locus effects. Sampson and Tarumoto (1976) found evidence of nonallelic interaction in an eight-parent diallel cross of oats. Hanson *et al.* (1967) made estimates of additive x additive epistatic variance that comprised more than one-half the total genetic variance for yield among homozygous soybean lines. Wynne and coworkers (1970, 1975) found significant estimates of specific combining ability (SCA) for a number of characters in the F₁ and F₂ generations of a six-parent diallel cross of diverse peanut lines. This was attributed to either dominance or geometric gene action. Hammons (1973) suggested that many characters in peanuts may be based on epistatic interactions. In an evaluation of early generation testing, Wynne (1976) found that the yields of the F₂'s from a six-parent diallel cross were ineffective in predicting the performance of the bulk F₅ generations and selected F₅ lines derived by a modified pedigree program from the F₂'s. He suggested that epistatic genetic variance might be important for yield and that, if so, selection for yield would be best practiced in late generations to allow advantageous gene combinations to come together. It was the purpose of this study to examine crosses of diverse peanut lines for

evidence of epistatic genetic variance and to estimate the additive x additive genetic component for yield and other characters.

Materials and Methods

Field Procedure

Six lines of peanuts described previously (Parker *et al.*, 1970; Wynne *et al.*, 1975), representing three botanical varieties from three geographic areas of South America were crossed in diallel without reciprocals. The two lines selected from each geographic area were either Valencia (ssp. *fastigiata* var. *fastigiata*), Virginia (ssp. *hypogaea* var. *hypogaea*) or Spanish (ssp. *fastigiata* var. *vulgaris*) types when classified by branching pattern (Gregory *et al.*, 1951).

Crosses in the F₁ generation were grown at the Peanut Belt Research Station at Lewiston, NC in 1969. The 15 crosses and six parents were planted in a randomized complete block (RCB) design with five replications. Each plot consisted of a single row of 10 plants spaced 51 cm apart. Data were taken on individual plants and plot averages computed. Equal numbers of seed from 30 plants of each cross were bulked. Half of the seed was stored in a freezer at 0 C and half was used to produce an F₂ generation in the following year. In 1970 the F₂ crosses and parents were grown both at Lewiston and at the Upper Coastal Plain Research Station at Rocky Mount, NC. Each entry was replicated four times in an RCB design. Each plot consisted of two rows spaced 91 cm apart with 50 plants per row. Plants were spaced 25.4 cm apart within rows. Plots in this and subsequent years were dug and harvested mechanically. Seed from each entry was bulked and sampled to produce the F₃ generation. The F₃ and F₄ crosses were grown at Lewiston in 1971 and 1972, respectively. In both years the crosses and parents were replicated four times in an RCB design. Plots were constructed as in 1970. Seed for the F₄ crosses was obtained from the first replication of the F₃'s, and the F₅ generation was produced in the same manner from the F₄. In 1973 the stored F₂ seed was planted along with the F₅ at Lewiston and Rocky Mount in an RCB design with three replications. Each plot consisted of two rows with 35 plants per row. Spacings were the same as in the F₄.

The data measured in each generation were:

- Percentage fancy size pods (FS); pods which ride a 1.34 x 7.62-cm screen,
- Percentage extra large kernels (XL); kernels which ride a .85 x 2.54-cm screen,
- Percentage sound mature kernels (SM); kernels which ride a 60 x 2.54-cm screen,
- Number of kernels per kg (CT),
- Length of 20 randomly drawn pods in cm (FL), and
- Yield in kg per plot (YP).

Analyses

Combining ability analyses were performed for each character using the computer program DIALL (Schaffer and Usanis, 1969), which follows Griffing's (1956) Method IV, Model II. Parents were omitted from all analyses. The DIALL program pooled the SCA x environment and error sums of squares, so conventional analyses of variance were performed to obtain the error sum of squares and, by subtraction, the SCA x environment sum of squares for the F₂ and F₅ generations. For the F₂ and F₅ there were no exact tests of significance for general combining ability (GCA), so Satterthwaite approximations were constructed.

For the F₁, F₃, and F₄, DIALLC, an extension of the DIALL program, was used to calculate the variance components for general and specific combining ability. For the F₂ and F₅, and GCA and SCA variance components were calculated by equating the mean squares to their expectations in the analysis of variance. The variance components for GCA (σ_G^2) and SCA (σ_S^2) were equated to the covariance among relatives as:

$$\sigma_G^2 = \text{Cov}(1)$$

$$\sigma_S^2 = \text{Cov}(2) - 2\text{Cov}(1)$$

¹Paper no. 5611 of the Journal Series of the North Carolina Agricultural Experiment Station, Raleigh, NC 27650.

²This work was partially supported by CSRS Research Agreement 701-15-51.

³Graduate Research Assistant and Associate Professor of Crop Science and Professor of Statistics, North Carolina State University, Raleigh, NC 27650.

after Cockerham (1963) where $Cov(1)$ was the covariance of half sib relatives or their progeny formed by bulk selfing and $Cov(2)$ was the covariance of full sib relatives or their progeny formed by bulk selfing.

Genetic variance components were estimated using a modification of the method presented by Stuber (1970). The components were expressed in terms of genetic variances defined for the non-inbred generation represented by the F_1 (non-inbred) progenies of the diallel matings. Assumptions similar to those of Stuber (1970) were made in formulating covariances among inbred relatives. The variance components due to SCA and SCA x E interaction from each inbred generation were partitioned into linear combinations of dominance and additive by additive (A x A) epistatic variance (Table 1). This partitioning required no assumptions as to gene frequencies in the diallel while separation of GCA variance into genetic components would require knowledge of the gene frequencies.

Table 1. Coefficients of genetic components of the SCA variance component.

Generation	$\sigma_D^2(F=0)$	$\sigma_{AA}^2(F=0)$
F ₁	1	1/2
F ₂	1/4	1/2
F ₃	1/16	1/2
F ₄	1/64	1/2
F ₅	1/256	1/2

Since field plot size was not constant over generations, the error variances were not assumed to be homogeneous. Bartlett's test showed heterogeneity at the .01 probability level for all traits except XL. Therefore, the expectations for the mean squares for SCA and SCA x environment interaction minus their respective error terms were expressed in terms of dominance variance, additive x additive epistatic variance, dominance x environment interaction, and additive x additive epistasis x environment interaction variances. Even though the SCA x environment mean square was estimable from only the F₂ and F₅ generations, the environmental interaction components were included in all expectations. Equating the observed mean squares, minus error, for SCA and SCA x environment to their expectations led to seven equations in four unknowns — σ_D^2 , σ_{AA}^2 , σ_{DE}^2 , and σ_{AAE}^2 . Letting m_{11} and m_{12} be the mean squares minus error for SCA and SCA x environment for the F_i generation, this system of equations was written as:

$$\begin{bmatrix} m_{11} \\ m_{21} \\ m_{22} \\ m_{31} \\ m_{41} \\ m_{51} \\ m_{52} \end{bmatrix} = \begin{bmatrix} 5 & 2.5 & 5 & 2.5 \\ 3.5 & 7 & 0.875 & 1.75 \\ 0 & 0 & 0.875 & 1.75 \\ 0.25 & 2 & 0.25 & 2 \\ 0.0625 & 2 & 0.0625 & 2 \\ 0.0234375 & 3 & 0.0117188 & 1.5 \\ 0 & 0 & 0.0117188 & 1.5 \end{bmatrix} \begin{bmatrix} \sigma_D^2 \\ \sigma_{AA}^2 \\ \sigma_{DE}^2 \\ \sigma_{AAE}^2 \end{bmatrix}$$

or $M = X\Sigma$.

Iterative weighted least squares were used to obtain the estimates $\Sigma_{k+1} = (X^*W_kX)^{-1}(X^*W_kM)$, where W_k is the inverse of the estimated variance-covariance matrix using the results of the kth iteration and assuming the original mean squares are multiples of independent chi-square random variables. In all cases, negative estimates of variance components were set equal to zero before computing W_k for the next iteration.

Regression error, $s_{y,x}^2$, was computed after Draper and Smith (1966) and scalar multiplication by the final $(X^*W_kX)^{-1}$ produced the estimated variance-covariance matrix of the general component estimates.

Results and Discussion

Estimates of GCA and SCA variance components (Table 2) were mostly positive, although the GCA components for yield were negative in later generations. Except for the SCA components for extra large kernels, none of the series of estimates of a component for a trait exhibited a consistent trend of increase or decrease through the F₅ generation. The variability for GCA and SCA estimates within a series could have arisen from environmental factors unaccounted for in the generations grown at a single location or simply from the inherent difficulty in estimating variance components. Large negative estimates for components due to SCA in the F₃ for FL, GCA x E for CT in the F₅, and SCA x E for FS in the F₅ may have been due to failure of the data to fit the statistical model imposed upon them in the analysis.

Table 2. Estimates of GCA and SCA variance components with their standard deviations.

Trait	Generation	σ_{GCA}^2	σ_{SCA}^2	$\sigma_{GCA \times E}^2$	$\sigma_{SCA \times E}^2$
FS	F ₁	8.56±12.89	49.97±22.43**	--	--
	F ₂	22.36±16.46*	25.75±12.33*	2.63±2.15	0.94±3.52
	F ₃	27.87±18.96*	14.49±12.72	--	--
	F ₄	15.39±10.61*	4.95±7.76	--	--
	F ₅	24.68±17.27*	20.20±10.04**	2.76±2.41	-12.82±4.44
XL	F ₁	22.97±16.10*	25.20±11.38**	--	--
	F ₂	8.08±8.72	25.05±11.10**	3.43±1.54**	0.11±1.13
	F ₃	11.25±9.71	20.86±10.66**	--	--
	F ₄	14.43±10.59*	16.66±8.54**	--	--
	F ₅	10.24±7.54*	12.96±6.28**	-20±.49	-.51±1.60
SM	F ₁	0.46±2.26	9.81±5.18**	--	--
	F ₂	1.39±2.88	12.19±5.53**	0.36±.44	0.98±.84
	F ₃	5.71±4.69	9.62±4.76**	--	--
	F ₄	4.25±4.94	14.66±7.41**	--	--
	F ₅	3.55±3.03	5.26±3.38*	-.07±.77	-.45±2.20
CT	F ₁	43572±34691	72228±33185**	--	--
	F ₂	3541±3211	1452±2183	4303±2731	6912±3500**
	F ₃	1988±2121	3991±3050*	--	--
	F ₄	8640±5279**	2198±2074	--	--
	F ₅	8869±5923	5538±5775	-2234±1458	2950±5765
FL	F ₁	2.65±1.64**	1.47±.69**	--	--
	F ₂	7.35±4.67*	4.47±2.50**	-.62±.44	1.69±1.47
	F ₃	9.73±5.83*	-5.04±3.13	--	--
	F ₄	4.46±4.28	3.34±5.87	--	--
	F ₅	8.91±7.91	13.34±6.40**	3.41±2.27*	-.68±1.53
YP	F ₁	0.01±.02	0.10±.05**	--	--
	F ₂	0.01±.03	0.10±.06**	0.03±.03	0.06±.05
	F ₃	-.16±.12	0.86±.44**	--	--
	F ₄	-.08±.12	0.66±.37**	--	--
	F ₅	-.09±.06	0.12±.11	0.12±.10	0.04±.11

*,**Indicate that the mean square from which the component was derived was significant at the .05 or .01 probability level.

Precise genetic interpretation of GCA variance was not possible without knowledge of or assumptions about gene frequencies in the reference population. Nevertheless, the GCA component consists largely of additive-type variances with more of the additive component represented in a given generation than the A x A epistatic component. Estimates of the GCA component for sound mature kernels and yield were not significant for any generation, indicating low additive variance. Inconsistently significant estimates of the GCA component, as found for extra large kernels, count per kg, and fruit length, could again have resulted from the effect of environment in those generations grown at only one location. Also the sensitivity of the F-test of the GCA mean square was reduced by the small size of the diallel cross, which provided only 5 and 9 degrees of freedom for the test. Consistently significant estimates of GCA components, as found for fancy size pods, indicated the presence

of additive-type variance which could not, however, be resolved into additive and epistatic components.

For all traits except count per kg, estimates of SCA variance remained significant at either the .05 or .01 probability level in the F_4 and/or F_5 generation. SCA variance estimates for all these traits except yield were significant in the F_5 . Under the genetic model assumed, the SCA component contained only 1/64 or .0156 of the dominance variance of the reference population by the F_4 generation and only 1/256 or .0039 by the F_5 . If A x A epistatic variance for a trait were negligible, then the dominance variance alone would have to be of sufficient magnitude to cause significance of the SCA variance component.

The results of the least squares estimation procedure (Table 3) indicated that A x A epistatic variance was larger than dominance variance for all traits. For all traits except CT, estimates of the epistatic component were greater than their standard deviations, and for FS, XL and SM the estimates were statistically significant. This implied that epistatic effects, rather than dominance, caused the significance of late generation estimates of SCA variance. The implications for peanut breeders interested in utilization of exotic germplasm are several. First, the development of homozygous lines or blends of lines for commercial use should allow the best utilization of additive epistatic variance. Second, selection based on early generation testing would be ineffective compared to late generation testing because unique gene combinations may not come together in the early segregating generations. Large populations should be maintained until late generations to provide the best opportunity for advantageous combinations to occur. Lastly, since the number of possible homozygous genotypes in a segregating population is a geometric function of the number of segregating loci, incorporation of diverse parental lines into a breeding program should increase the amount of additive epistatic variance in the population.

Table 3. Estimates of genetic variance components with their standard deviations.

Trait	$\sigma_D^2(F=0)$	$\sigma_{AA}^2(F=0)$	$\sigma_{DE}^2(F=0)$	$\sigma_{AA-E}^2(F=0)$
FS	1.49±18.54	41.05±7.95	50.44±15.80	-25.82±6.37
XL	13.38±9.12	29.86±5.07	1.75±4.58	-.89±1.90
SM	-.50±8.93	17.67±5.02	3.15±7.42	.14±3.50
CT	-2089±26195	6140±9863	45264±29116	-2948±10000
FL	-17.58±14.85	16.77±7.66	11.44±11.12	-2.48±3.94
YP	-.25±.65	.55±.35	.03±.62	.16±.29

In conclusion, it must be recognized that the parents used in this study represented only a small fraction of the available exotic peanut cultivars and the populations in each generation were limited in size. Nevertheless, the findings of this study suggested that significant levels of additive epistatic variance for yield and fruit characters occur in populations derived from crosses of diverse peanut lines.

Additional studies using larger samples of exotic germplasm are needed to critically define the importance of epistatic genetic variance in peanut populations.

Literature Cited

1. Cockerham, C. C. 1963. Estimation of genetic variances, pp. 53-94. In W. D. Hanson and H. F. Robinson (eds.), *Statistical genetics and plant breeding*. Nat. Acad. Sci. - Nat. Res. Council. Publ. 982, Washington, D. C.
2. Draper, N. R. and H. Smith. 1966. *Applied regression analysis*. John Wiley and Sons, New York.
3. Gregory, W. C., B. W. Smith and J. A. Yarbrough. 1951. Morphology, genetics, and breeding, pp. 28-88. In *The peanut — the unpredictable legume*. National Fertilizer Assoc., Washington, D. C.
4. Griffing, B. 1956. Concept of general and specific combining ability in relation to diallel crossing systems. *Aust. J. Biol. Sci.* 9:463-493.
5. Hammons, R. O. 1973. Genetics of *Arachis hypogaea*, pp. 135-173. In *Peanuts—culture and uses*. American Peanut Research and Education Assoc., Stillwater, Oklahoma.
6. Hanson, W. D., A. H. Probst and B. E. Caldwell. 1967. Evaluation of a population of soybean genotypes with implications for improving self-pollinated crops. *Crop Sci.* 7:99-103.
7. Parker, R. C., J. C. Wynne and D. A. Emery. 1970. Combining ability estimates in *Arachis hypogaea* L. I. F_1 seedling responses in a controlled environment. *Crop Sci.* 10:429-432.
8. Sampson, D. R. and I. Tarumoto. 1976. Genetic variances in an eight parent half diallel of oats. *Can. J. Genet. Cytol.* 18: 419-427.
9. Schaffer, H. E. and R. A. Usanis. 1969. General least squares analysis of diallel experiments, a computer program—DIALL. Genetics Dep. Res. Rep. No. 1, North Carolina State University at Raleigh.
10. Stuber, C. W. 1970. Estimation of genetic variances using inbred relatives. *Crop Sci.* 10:129-135.
11. Wynne, J. C. 1976. Evaluation of early generation testing in peanuts. *Peanut Sci.* 3:62-66.
12. Wynne, J. C., D. A. Emery and P. W. Rice. 1970. Combining ability estimates in *Arachis hypogaea* L. II. Field performance of F_1 hybrids. *Crop Sci.* 10:713-715.
13. Wynne, J. C., J. O. Rawlings and D. A. Emery. 1975. Combining ability estimates in *Arachis hypogaea* L. III. F_2 generation or intra- and intersubspecific crosses. *Peanut Sci.* 2: 50-54.