A Diallel Analysis of Agronomic Characters in Selected Lines of Cotton, *Gossypium hirsutum* L.[1]

T. G. White and R. J. Kohel[2]

DURING the past few years several reports on the use of the diallel analysis in the study of quantitatively inherited characteristics of crop plants have appeared in the literature. Hayman (5) reviewed the development of the different approaches to the diallel analysis (of which three are now generally recognized), and Crumpacker and Allard (1) more recently discussed some of the criticisms leveled against the various approaches by other workers. While evaluation of the diallel analysis by professional biometrician is continuing, information which plant breeders and geneticists might contribute to the investigations seems valuable, wholly aside from knowledge gained specific to their materials. Toward this end, a diallel analysis of several agronomic characters of cotton, *Gossypium hirsutum* L., was undertaken.

MATERIALS AND METHODS

The data for these analyses were taken from an experiment reported by White and Richmond (7). In that experiment a complete 5-parent diallel cross (n^2 or 25 progenies), a set of top crosses, and commercial checks were employed in a study of heterosis and combining ability in cotton. For purposes of the research reported here, the 25 diallel entries were extracted without adjustment from the 36-entry, simple-lattice design (4 replications) of the original experiment. The parents of the diallel cross were MU8b, a Cambodian type of *Gossypium hirsutum* selected in India; Texas 468, a stock of *G. hirsutum* race *punctatum;* CB3150, a stock of *G. hirsutum* introduced from Russia; Texas 63, a stock of *G. hirsutum* race *latifolium;* and 2–8–7–6, the F_4 of a cross between DPL, an inbred line of *G. hirsutum* race *latifolium* and a stock of *G. hirsutum* race *palmeri.*

The characters subjected to diallel analysis were seedling height (at 11 days), first fruiting node, days to anthesis, plant height, relative fruitfulness (number of bolls per 100 g. aerial vegetative weight), fruiting index (ratio of the weight of fruiting to vegetative parts), yield (pounds of lint per acre), boll size (number per pound), lint percent, seed index, lint index, earliness (first 2 of 5 pickings expressed as a percentage of all pickings), number of bolls per plant, and weight of aerial vegetative parts of the plant.

The diallel analysis used here was that presented by Hayman (4). Hayman's model is based on several assumptions, each of which must hold true for the analysis to be valid. The assumptions, together with a discussion of the applicability of each to *G. hirsutum* and the material used in this study, follow:

Assumptions Concerning the Species

a. *Diploid segregation. G. hirsutum*, an amphidiploid, segregates in a diploidal manner (Endrizzi, 2; Kimber, 6).

b. *No differences between reciprocal crosses.* This assumption is considered valid for most characteristics of *G. hirsutum*, but certain possible exceptions were noted for some of the characters.

c. *Independent action of non-allelic genes.* This assumption can be tested with F_1 data.

Assumptions Concerning the Diallel Cross

d. *No multiple allelism.* One might not be surprised to find multiple alleles at certain loci controlling a quantitatively expressed character. However, Hayman (5) showed that the measure of dominance is not seriously disturbed by multiple allelism.

e. *Homozygous parents.* Except for 2–8–7–6, an F_4 that may have some residual heterozygosity, the parents used in this diallel cross were assumed to have been fairly homozygous on their acces-

sion, and they have been maintained by self-fertilization (for a minimum of 2 cycles) in the S-1 Collection at College Station, Texas. Heterozygosity in 2–8–7–6 would be minimized since it was pedigreed by individual plant selections.

f. *Genes independently distributed between the parents.* This assumption can be tested by the use of F_1 data.

The statistics estimated and the symbols used in this presentation are those employed by Hayman (4). They are as follows:

The covariance between the n parents and their offspring in the rth array (W_r).

The variance of the rth array (V_r).

Variance of the parents (V_{oLo}).

The covariance between the parents and the mean of their offspring. For each replication, this represents the mean of the n W_r (W_{oLo1}).

The mean of the n V_r for each replication (V_{1L1}).

Variance of the means of the arrays (V_{oL1}).

The square of the difference between the mean of the parents and the mean of the n^2 progeny ($m_{L1} - m_{Lo}$)^2.

These statistics are used to estimate the components of variation: D, F, H_1, H_2, and h^2. (See Hayman, 4, 5, for equational relationships between the statistics and their components.)

The components provide estimates as follows:

(H_1/D)½—An overall measure of the degree of dominance.

H_2/4H_1—The mean value of the product $u_i v_i$ (u_i and v_i are the frequencies of the parents with positive and negative homozygotes at the ith locus, respectively).

[(4DH_1)½ + F]/[(4DH_1)½ — F]—The ratio of the total numbers of dominant to recessive genes in all the parents.

h^2/H_2—An estimate of the number of groups of genes which control the character and exhibit dominance to some degree.

Hayman's method (4) of the diallel analysis includes an analysis of variance of the diallel cross, which he presented in an earlier publication (3). The analysis of variance was also used in this paper. The relationships between the components of variance in the *analysis of variance of the diallel* and the components of variance in the broader, more inclusive *diallel analysis* are as follows:

AOV	Diallel Analysis
a ------------------------	D — F + H_1 — H_2
b ------------------------	H_2
b_1 -------------------	h^2
b_2 -------------------	H_1 — H_2
b_3 -------------------	Residual
c ------------------------	Maternal differences
d ------------------------	Reciprocal differences not maternal

If the diallel cross conforms to the assumptions listed earlier, W_r—V_r is constant. Heterogeneity of W_r—V_r thus would indicate nonconformance to the model. Because of the replications in this study, it was possible to analyze the variance of W_r—V_r to test for failures of the hypotheses. Standard errors for the diallel analysis were calculated by use of the table of covariance matrix (4); the estimated error used in the table was the residual sum of squares from observed minus expected values of the statistics divided by 43 degrees of freedom, derived as follows: 53 statistics: V_{oLo}, V_{oL1}, ($m_{Lo} - m_{L1}$)^2, V_r (r=1–5), W_r (r=1–5) in each block, and also residual error from analysis of variance minus 10 constants: D, H_1, H_2, h^2, E, and F_r (r=1–5).

When significant dominance was found, the correlation between y_r (parents) and W_r + V_r was calculated, and if this was significant, the theoretical limits of selection among genes exhibiting dominance were then calculated.

RESULTS

Before any analyses were carried to completion, Bartlett's test was used to test homogeneity of the variances computed over replications. Significant chi-square values were obtained for the folowing characters: fruiting node, days to anthesis, relative fruitfulness, fruiting index, line percent, and number of bolls per plant. Evidence of non-normality

Table 1. Mean squares and significance of characters tested in analysis of variance.

	a(4)†	b(10)	b₁'(1)	b₂(4)	b₃(5)	c(4)	d(6)	Error(72)
Seedling height	730**	122*	605**	83	57.4	78	36	56.11
Fruiting node	45.5*	8.8	7.0	10.0	8.0	20.5	7.8	14.90
Days to anthesis	2,129**	578	428	587	601	1,214*	283	376.25
Plant height	16,678**	1,978	3,329	2,586	1,223	2,206	1,520	1,514.00
Relative fruitfulness	19,493**	2,433	9,616*	911	2,215	828	741	2,161.00
Fruiting index	20,479*	6,180	12,225*	2,042	8,280	4,331	2,072	‡
Yield	232,478**	40,469**	331,949**	1,900	13,028	820	16,348	7,358
Boll size	1,368**	42**	218**	34	12	14	51**	15
Lint %	13,742**	111	195	16	171	38	267*	89
Seed index	2,406**	91*	35	16	162**	38	39	40
Lint index	156,234**	1,963	4,369	334	2,785	2,752	4,408	2,039
Earliness	429,979**	9,352	16,731	11,297	6,320	12,086	27,505*	11,089
Bolls per plant	5,892**	1,197**	7,859**	114	731	182	508	‡
Vegetative weights	80,036**	12,623*	77,367**	1,436	8,625	9,230	3,927	5,830

* Significant at the .05 level.   ** Significant at the .01 level.   † Degrees of freedom in parentheses.   ‡ Individual interaction mean squares were used because of heterogeneity of variances.

Table 2. Statistics and components of the diallel analysis for each character studied.

	Statistics					Components					
	V₀LO	W₀LO₁	V₁L₁	V₀L₁	(mLO-mL1)²	D	F	H₁	H₂	h²	E
Seedling height	106.7	37.18	59.96	19.74	28.97	50.59*	-2.65	51.92	48.66	79.96*	56.11
Fruiting node	.1775	.0484	.0785	.0178	.0146	.0285	-.0173	-.0893	-.0552	-.0369	.149
Days to anthesis	8.29	1.83	3.18	0.75	0.48	4.52	4.74	3.89	2.17	-0.49	3.76
Plant height	2,093	762	1,324	563	213	579	-676	408	18	-118	1,514
Relative fruitfulness	4.62	1.95	1.89	1.07	0.48	2.45**	-1.18	-1.25	-1.05	0.53	2.16
Fruiting index	.126	.050	.046	.025	.006	.073**	-.012	-.030	-023	-.010	.053
Yield	23,325	11,681	13,951	6,586	13,407	15,966**	-8,905	13,274	14,743*	48,918**	7,358
Boll size	123.9	63.5	45.6	35.2	10.4	108.8**	-24.1	13.3	11.3	31.8	15.0
Lint %	14.61	6.95	3.99	3.46	0.15	13.72**	0.33	0.43	0.32	0.03	0.89
Seed index	2.32	1.16	0.84	0.62	0.036	1.92**	-0.48	-0.01	.076	-0.114	0.402
Lint index	1.67	0.797	0.497	0.401	0.023	1.47**	-0.087	-0.058	-0.025	-0.039	0.204
Earliness	476.9	220.9	174.9	119.1	14.9	366.0**	-62.9	4.7	1.7	-11.4	110.9
Bolls per plant	6.88	3.20	4.37	1.82	3.24	3.18*	-3.47	1.96	2.82	10.61**	3.70
Vegetative weights	10,126	4,259	5,826	2,340	3,773	4,296	-3,779	1,236	2,286	11,361**	5,830

* Significant at .05 level.   ** Significant at .01 level.

Table 3. Proportions, correlations of $V_r + W_r$ and y_r, and limits of selection for characters exhibiting dominance.

	$(H_1/D)^{\frac{1}{2}}$	$H_2/4H_1$	$\dfrac{[(4DH_1)^{\frac{1}{2}}+F]}{[(4DH_1)^{\frac{1}{2}}-F]}$	h^2/H_2	Corr. V_r+W_r & y_r	Limits of selection	
						Upper	Lower
Seedling height	1.01	0.23	0.95	1.64	-0.846*	134.3	56.2
Yield	0.91	0.28	0.53	3.32	-0.936**	939	432
Boll size	0.35	0.21	0.52	2.81	0.926**	48.7	95.3
Bolls per plant	0.79	0.36	0.18	3.76	-0.374	-	-
Vegetative weights	0.53	0.46	0.10	4.97	-0.204	-	-

* Significant at the .05 level.   ** Significant at the .01 level.

of the data for these characters was thus suggested, and the validity of their analysis was questionable. Furthermore, the pattern of heterogeneity did not fall into any of the three categories under which it may be handled in the analysis (4). Therefore, all except number of bolls per plant were removed from further consideration, although the results of the analyses of all characters are present in Tables 1 and 2. The data for number of bolls per plant held special interest because of the findings of White and Richmond (7) which showed a relation of this character to heterosis for yield. While any conclusions drawn from the results of the analysis of this character are tenuous, it seemed more appropriate to attempt an evaluation rather than to dismiss the data without further consideration.

In the analyses of variance of $W_r - V_r$ for the various characters, no significant variation among the n lines was found. From the standpoint of conformity to the model employed here, it therefore appears that the analyses of these characters are valid.

Table 1 shows the mean square values for the sources of variance in the analysis of variance for each of the 14 characters studied. Component a was highly significant for every character except fruiting node and fruiting index, for which it was significant at the .05 level. This component measures only additive variation when dominance is absent. Component b, the dominance component, was highly significant for yield, boll size, and bolls per plant; it was significant for seedling height, seed index, and vegetative weights. Component b₂ in the analysis of variance was not

significant for any character; thus symmetry of gene frequencies for those characters having *significant* dominance was indicated. (Component b₂ is also zero when no dominance exists.) Some maternal effects may have been associated with days to anthesis (component c under analysis of variance). In addition, some reciprocal differences were indicated for boll size, lint percent, and earliness (component d).

Among the components (Table 2) estimated by the broader *diallel analysis,* the additive component D was significant at the .05 or .01 level for all characters except fruiting node, days to anthesis plant height, and vegetative weights. Yield was the only character for which a dominance component (H₁ or H₂) was significant. However, as noted above, in the more reliable (as far as confidence limits are concerned) analysis of variance, six of the characters showed significant dominance variation. But one of these, seed index, showed an overall measure of dominance of .07 (0 — 1.00 = partial dominance) and had other anomalous values for certain statistics and proportions. As a consequence seed index was eliminated from the characters to which dominance was ascribed. Therefore it seems legitimate to say, as far as these data and analyses are concerned, that dominance was operating to control the five characters listed in Table 3. This table shows the various proportions, the correlation values of $V_r + W_r$ and y_r (parents), and the limits of selection when the correlation was significant.

The character seedling height in Table 3 displays a near-classic set of results that closely approximate the theoretical expectations for the diallel analysis. The over-all measure of dominance is 1.01 (close to complete dominance and extending slightly into the overdominance range). The proportion $H_2/4H_1$ estimates the product $u_i v_i$. The notations, u_i and v_i, refer to the frequencies of parents with positive and negative homozygotes, respectively, at the ith locus. When these frequencies are equal, $u_i v_i = \frac{1}{2} \times \frac{1}{2} = .25$. For seedling height, this figure is .23. This finding con-

firms the results already shown by component b_2 in the analysis of variance. In a like manner, the proportion $[(4DH_1)^{1/2} + F]/[(4DH_1)^{1/2} - F]$ confirms symmetry. This proportion should be unity if the total number of dominant genes in all the parents is equal to the total number of recessives (assuming equality of gene effects). Unity is approximated by the .95 value obtained.

The number of groups of genes (or, quite possibly, number of genes) which exhibit dominance to some degree and which control seedling height, is estimated by h^2/H_2, for which a value of 1.64 was obtained (again assuming equality of gene effects). The value obtained indicates that perhaps two such groups were present. It must be emphasized that this value usually underestimates the number of groups, since positive and negative effects of dominance genes cancel one another. On the other hand, dominance was sufficiently unidirectional to cause h^2 to be significant, so perhaps this estimate may approach the actual number. The correlation of $V_r + W_r$ and y_r (parents) was —.846, significant at the .05 level. Since $V_r + W_r$ is closely correlated with the recessive homozygotes, significant dominance in a positive direction, or greater seedling height, was indicated. The predicted upper and lower limits of selection *among genes exhibiting dominance* were 134.3 and 56.2 mm of height. It must be stressed that these predictions *do not* take into account genes that show additive effects.

The remaining characters in Table 3 are yield and characters related to yield. Yield, in this study, had an over-all measure of dominance of .91, which indicated partial dominance for the character. Mean direction of dominance was towards greater yield, since the F_1 mean exceeded the parental mean. Correlation of $V_r + W_r$ and y_r was —.936, highly significant; this indicated that most of the dominant genes had positive (greater yield) effects. The predicted limits of selection, 939 and 432 pounds per acre, overlapped the lower end of the range of recorded yields. The highest actual yield, 1019 pounds per acre, illustrates the fact that the predicted limits do not take into account the additive genes. The number of genes that exhibited dominance among those controlling yield was estimated at 3.32. Again, the estimate is a minimum.

Boll size showed a mean dominance level of .35, with a high correlation (.926**) between $V_r + W_r$ and y_r. Although the correlation was positive, larger boll size was the direction of dominance, since the measurement of boll size used here (number per pound) is inversely proportional to actual boll size. The number of genes that exhibit dominance was estimated at 2.81.

In the character, bolls per plant, mean level of dominance was .79, but a low level of correlation between $V_r +$ W_r and y_r indicated that some of the genes showed dominance in a positive direction for the character (greater number of bolls) while some acted in the negative direction (lesser number of bolls). The estimated number of genes exhibiting dominance was four.

Vegetative weights was similar to bolls per plant. The mean level of dominance was .53, and the correlation value was low. The estimate of the number of dominance genes was higher, about five. Because of the low correlations, the limits of selection were not predicted for vegetative weights and bolls per plant.

Although there was some evidence of asymmetry in the estimates of the product u_iv_i and the ratio of dominant to recessive genes for yield, boll size, number of bolls per plant, and vegetative weights, the fact that no such asymmetry was detected in the analysis of variance of the diallel or in the analysis of variance of $W_r - V_r$ indicated that there was insufficient bias to distort the analysis.

DISCUSSION

White and Richmond (7) noted two cases of heterosis for yield. One case was accompanied by a corresponding case in the character bolls per plant. Since fruiting efficiency (number of bolls per 100 g. of vegetative weight) had not shown a heterotic effect, vegetative weights were investigated, and again a corresponding case of heterosis was noted. The other heterotic hybrid did not behave in exactly the same fashion. There was no corresponding heterotic effect in either number of bolls per plant or vegetative weights, although there was a tendency toward heterosis in both characters. However, a non-significant heterotic effect was found in the character boll size for this latter case of heterosis and they concluded that both boll size and boll number were operating to increase yield heterotically.

In the work reported here, diallel analyses of these yeild-complex characters (including yield itself) are consistent with the results reported by White and Richmond (7). Regardless of how one views the association between vegetative wieghts and bolls per plant (i.e., whether they are co-results of a common cause or whether more bolls per plant are the direct result of greater vegetative size—hence more fruiting nodes), yield heterosis might be reasonably accounted for by the dominance genes in bolls per plant or boll size, or a combination of the two, operating with associated gene which exhibit additive effects.

In the light of the correlation values obtained for $V_r +$ W_r and y_r for bolls per plant and vegetative weights, it seems that there are dominance genes with both positive and negative effects, even though (in this material) the positive outnumber the negative (or at least the mean level

Table 4. Components of the diallel analysis and their tests of significance, when the parents are considered to be a random sample from a broader population.

	Components						Significance & variance ratio (F values)			
	D	F	H_1	H_2	h^2	E	H_1	H_2	H_1-H_2	h^2
Seedling height	50.6	-4.42	34.83	38.61	124.11	56.11	1.21	1.34	0.96	2.51
Fruiting node	.0285	-.0290	-.0925	.0042	.0192	.1490	0.79	1.01	0.61	1.16
Days to anthesis	4.53	7.91	4.85	0.13	-4.55	3.76	1.43	1.02	1.75	0.48
Plant height	579	-1,131	459	-931	-1,202	1,514	1.10	0.69	1.63	0.58
Relative fruitfulness	2.46	-1.92	-1.81	-1.89	0.96	2.16	0.72	0.56	1.03	1.39
Fruiting index	.073	-.019	-.031	-.022	-.006	.053	0.81	0.79	0.89	0.90
Yield	15,967	-14,839	-2,194	2,833	80,174	7,358	0.90	1.19	0.62	13.25*
Boll size	108.9	-40.3	3.8	-1.2	46.0	15.0	1.08	0.96	1.20	2.94
Lint %	13.72	0.59	0.56	0.35	-0.15	0.89	1.21	1.20	1.13	0.90
Seed index	1.92	-0.80	0.067	0.43	0.07	0.402	1.06	1.53	0.55	1.19
Lint index	1.470	-0.155	-0.052	0.073	0.033	0.204	0.92	1.18	0.66	1.17
Earliness	366.0	-104.8	9.8	1.5	-24.4	110.9	1.03	1.01	1.04	0.84
Bolls per plant	3.18	-5.80	-1.54	1.27	18.67	3.70	0.86	1.17	0.57	7.17*
Vegetative weight	4,296	-6,300	-2,736	262	20,124	5,830	0.84	1.02	0.69	4.70

* Significant at the .05 level.

of dominance was on the positive side). It follows that the number of genes exhibiting dominance in yield is underestimated.

The data thus analyzed provide information about the population as constituted by the parents of the diallel cross. If one considers the five parents as a proper sampling from a broader population, e.g., the cultivated types of *G. hirsutum*, then the results of the analysis are quite different. Actually, this would be improper since the diallel cross used here is so small that it cannot conceivably represent a proper sample of such a broad population. However, for illustrative purposes, the results of an analysis based on a random model (model II) presented by Hayman (5) are presented.

Table 4 shows the new values of the components, and the approximate F values along with indications of significance. No value of H_1 or $H_1 - H_2$ is significant. Two values of h^2 (yield and bolls per plant) are significant at the .05 level. The general picture emerging from this analysis is no dominance for any character, but with two cases of overall heterosis. This indicates failure of the dominance model for these two characters since H_1 and H_2 cannot be zero if h^2 remains positive, as long as the ordinary dominance model is correct.

This latter analysis is given here to emphasize that the situation as found here for this small population cannot have wide application in ascribing characteristics to Upland cotton in general.

The results obtained from these diallel analyses when combined with those obtained from more conventional analyses (7) provide a clearer picture of a biological phenomenon than would be provided by either analysis when used alone. Although the effort of making diallel analyses probably would not be worthwhile in a systematic search for good combiners, it might be employed when one seeks more information about quantitative genetic control of certain characters. This information applies to the population as constituted by the parents of the diallel cross (as in the analyses reported here), or it may apply to a larger population if appropriate sampling is made from such a population.

SUMMARY

Diallel analyses were made for several agronomic characters measured in a five-parent diallel cross. Additive variation was detected in all characters. In addition, dominance appeared to be operating in seedling height, yield, boll size, bolls per plant, and vegetative weights. Except for seedling height, for which the dominance genes seemed to be close to complete dominance, the dominance detected appeared to be partial dominance. Limits of selection among genes exhibiting dominance were calculated for certain characters.

These data were also subjected to a diallel analysis based on model II in order to illustrate the considerable difference in results that may be obtained.

Results obtained from these diallel analyses are consistent with results obtained by more conventional analyses of the same material. It seems plausible that the diallel analysis could be profitably employed as an integral part of a basic quantitative genetics study.

LITERATURE CITED

1. CRUMPACKER, D. W., and ALLARD, R. W. A diallel cross analysis of heading date in wheat. Hilgardia 32:275–318. 1962.
2. ENDRIZZI, J. E. The diploid-like cytological behavior of tetraploid cotton. Evolution 16:325–329. 1962.
3. HAYMAN, B. I. The analysis of variance of diallel tables. Biometrics 10:235–244. 1954.
4. ——————. The theory and analysis of diallel crosses. Genetics 39:789–809. 1954.
5. ——————. The theory and analysis of diallel crosses. III. Genetics 45:155–172. 1960.
6. KIMBER, G. Basis of the diploid-like meiotic behavior of polyploid cotton. Nature (London) 19:98–100. 1961.
7. WHITE, T. G., and RICHMOND, T. R. Heterosis and combining ability in top and diallel crosses among primitive, foreign and cultivated American Upland cottons. Crop Sci. 3:58–62. 1963.