Published September, 1967

A Diallel Analysis of Several Fiber Property Traits in Upland Cotton (*Gossypium hirsutum* L.)[1]

Laval M. Verhalen and Jay C. Murray[2]

ABSTRACT

Diallel analyses of upper 2.5% span length, micronaire, 1/8" gauge stelometer strength, and 0" gauge stelometer strength were conducted on 10 selected varieties of upland cotton and the 45 possible F_1 crosses among them using the procedure of Jinks and Hayman. A partial failure of the basic assumptions of the analysis was noted for length and micronaire. The assumptions were valid for 1/8" gauge stelometer and 0" gauge stelometer.

In the character 2.5% span length, long fiber was on the average partially dominant over short fiber. In micronaire, some dominant genes appeared to increase fiber coarseness while others decreased it. Additive gene action appeared to govern the 1/8" gauge stelometer. Some dominant genes appeared to increase 0" gauge stelometer while others appeared to decrease it. Based on their heritability estimates, mass selection should be an effective breeding method for improving 1/8" gauge stelometer and 0" gauge stelometer. Pedigrees, sib tests, and progeny tests may increase the efficiency of improving length and micronaire.

The experiment is being continued to study the reliability of estimates made from one year to the next at the same location.

Additional index words: diallel cross, fiber length, *Gossypium hirsutum* L., fiber strength, cotton, micronaire.

IN recent years the analysis of diallel crosses has received considerable emphasis in many plant breeding programs because it fulfills certain specific needs of plant breeders. The analysis provides a systematic approach for the detection of parents and crosses superior for the traits under investigation. At the same time it helps the plant breeder choose the most efficient method of selection by allowing estimates to be made of the magnitude and relative importance of various genetic parameters.

Methods of breeding are relatively simple for crops in which large amounts of hybrid seed can be obtained at reasonable cost. In these crops a high degree of specific combining ability is sought in the parents chosen, and crosses displaying large amounts of heterosis are desired. The outstanding crosses, of superior to those already in production, are repeated on a much larger scale, and the hybrid seed is then utilized commercially. For other crops such as cotton (*Gossypium hirsutum* L.) in which cost of hybrid seed production on a commercial scale is prohibitive, the methods of breeding are more complex. In these crops general combining ability is more easily utilized in a selection program leading towards a pure-line variety than specific combining ability, and crosses displaying large amounts of additive genetic variance are preferred over those with the more heterotic responses.

The most glaring deficiency of the upland cotton varieties grown in the Texas and Oklahoma plains area is their lack of adequate fiber properties. Stormproof varieties with their desirable boll type generally have shorter and weaker fiber than open-boll varieties. On the other hand, the open-boll varieties are unsuitable for the mechanical harvesting practices used in the plains area. Since excessive fineness of fiber in this environment is often a problem, fiber coarseness as well as fiber length and strength were included in this study. The purposes of this experiment were to investigate the genetic mechanisms controlling these traits and to suggest the most efficient procedures for the development of new stormproof varieties with desirable fiber properties.

MATERIALS AND METHODS

Ten varieties (five stormproof and five open-boll) were included in the experiment. The stormproof varieties were 'Paymaster 101,' 'Gregg,' 'Western Stormproof,' 'Lankart 57,' and '6-77.' The open-boll varieties were 'Deltapine 45,' 'Coker 100A WR,' 'Acala 44,' 'Stoneville 7,' and 'Auburn M.' Except for 6-77, all of the parental strains were standard commercial varieties of cotton. Strain 6-77 is a Bacterial Blight-resistant selection from the variety 'Stormproof No. 1.' These varieties were specifically chosen and do not represent a random sample of all upland cotton varieties. Therefore, inferences made from the data apply only to the varieties and crosses studied.

Crosses were made in Iguala, Mexico, in the winter of 1964. In the following spring the 10 varieties and the 45 possible F_1 crosses among them were planted in a 7 × 8 rectangular lattice design with three replications at Perkins, Okla. Reciprocal crosses were not included. A dummy entry, 8948, was included to complete the 7 × 8 design. Plots were single rows 7.5 m long, and plants within plots were spaced approximately 50 cm apart. Single border rows of the variety 'Kemp' were planted between adjacent plots to equalize border effects between plots. Seedling diseases reduced stands considerably. To partially compensate for the resulting differential spacing between plants, 'De Ridder Red,' a variety with the dominant marker gene, R_p, was planted in the blank hills.

Two harvests were made on the material. Six random plants from each plot were chosen for laboratory analysis. However, seven plots had only four to six plants and within those plots all plants were taken. In the laboratory fiber length was measured by upper 2.5% span length, fiber coarseness by micronaire, and fiber strength by 1/8" gauge stelometer and 0" gauge stelometer. Fiber samples from each plant were analyzed separately, and then a weighted average of each fiber measurement over the two harvests was calculated for each plant based on percentage of total lint yield per harvest. All subsequent calculations were made from these weighted averages. The analysis of the data followed the diallel procedure described by Jinks and Hayman (5, 8, 9).

[1] Contribution from the Oklahoma Agricultural Experiment Station. Published with the approval of the Director as paper No. 1461 of the Journal Series and with the approval of the Crops Research Division, ARS, USDA. Received Feb. 27, 1967.

[2] Graduate Research Assistant and Professor of Agronomy, Oklahoma State University, Stillwater.

RESULTS AND DISCUSSION

Analyses of variance were conducted on a plot mean basis for each trait. These analyses revealed highly significant differences among entries for each of the traits studied. A diallel analysis was then conducted for each trait individually.

Assumptions of the diallel analysis are diploid segregation, no reciprocal differences, homozygous parents, no multiple alleles, uncorrelated gene distributions, no genotype-environment interaction within locations and years, and no epistasis (3). Since the analysis is invalidated to some degree by failure of any of the seven assumptions, three broad, general tests were used to determine whether the assumptions were fulfilled by the characters. The tests were as follows: *(1)* analysis of variance of the quantity ($W_r - V_r$), *(2)* analysis of the (W_r, W'_r) regression, and *(3)* analysis of the (V_r, W_r) regression. V_r is the variance of the members of an individual array; W_r is the covariance of the members of an array with their nonrecurrent parents; and W'_r is the covariance of the members of an array with the array means of their nonrecurrent parents.

In the first test of the assumptions, the quantity (W_r—V_r) is expected to be constant over arrays if all assumptions of the analysis are fulfilled (7, 9). Heterogeneity of this quantity over arrays indicates that one or more of the hypotheses are not valid for that particular character. The quantity was calculated for each of the 10 arrays in each of the three replications, and then an analysis of variance was conducted upon the 30 values obtained. The results of this test are summarized in Table 1. The F values obtained indicate that the variation of (W_r—V_r) attributable to arrays was not significant at the 0.05 probability level for each fiber quality measurement, though the variation of length and micronaire did approach significance. Therefore, this test indicates that the assumptions of the analysis are valid for each of the four traits.

In the second test, the (W_r, W'_r) regression coefficient for each trait is expected to be significantly different from zero but not significantly different from 0.5 if the assumptions are valid (1). Ninety-five percent confidence limits about the regressions were calculated by the method prescribed by Steel and Torrie (13). The results presented in Table 2 indicate that the regression coefficients were significantly different from zero for all four characters. However, the regression for micronaire was also significantly different from 0.5, whereas the others were not. As a consequence, validity of the assumptions is indicated for length, 1/8" gauge stelometer, and 0" gauge stelometer, while a partial failure of the assumptions is suggested for micronaire.

In the third test, the (V_r, W_r) regression coefficient for each trait is expected to be significantly different from 0 but not significantly different from 1.0 if all of the assumptions hold true (9). The results of this test are presented in Table 3. This last test indicates that at least one of the assumptions was not valid for micronaire, since its slope was not significantly different from zero. A partial failure of the assumptions was also indicated for length, since its slope was significantly different from 1.0. Therefore, according to this last test, only 1/8" gauge stelometer and 0" gauge stelometer fulfill all assumptions of the diallel.

In view of the results from the three tests, the assumptions appear to be valid for 1/8" gauge stelometer and 0" gauge stelometer. On the other hand, length did not comply with the expectations in the last test, and micronaire did not in the last two. For these two traits, a partial failure of the assumptions of the diallel was indicated.

The pinpointing of the offending assumptions for length and micronaire cannot be accomplished with the present data. However, certain assumptions may be considered fulfilled with some degree of confidence. Endrizzi (4) and Kimber (10) have established with reasonable certainty that cotton, an amphidiploid, does undergo diploid segregation. As a rule, reciprocal crosses within *Gossypium hirsutum* L. have not been significantly different. White and Richmond (14) recently reported no significant differences between reciprocal crosses for fiber length, strength, or micronaire in a diallel cross among five widely differing *G. hirsutum* strains. The assumption of uncorrelated gene distributions may be tested (3) by the ratio, $\frac{1}{4}$ H_2/H_1. This ratio did not differ significantly from the expected for any of the characters under study (see Table 5). As a consequence, the assumptions of diploid segregation, no reciprocal differences, and uncorrelated gene distributions were considered to be fulfilled for each of the traits.

Since cotton is a predominantly self-pollinated plant and since the varieties used in this experiment were selfed for one generation before crossing and testing, the parents were probably relatively homozygous. However, since some heterozygosity may remain even after many generations of self-pollination (2), this

Table 1. Analyses of variance of (W_r—V_r) values.

Source	df	Mean squares: 2.5% span length	Micronaire	1/8" gauge stelometer	0" gauge stelometer
		($\times 10^{-6}$)	($\times 10^{-2}$)	($\times 10^{-4}$)	($\times 10^{-5}$)
Arrays	9	.132415	.212584	.0652	.05542
Replications	2	.080044	.024980	.5490**	.39453**
Error	18	.055930	.079756	.0482	.05264

** Significant at the 0.01 level of probability.

Table 2. (W_r, W'_r) regression coefficients.

Measurement	Slope	95% confidence limits
2.5% span length	.5559	.4882-.6236
Micronaire	.2874	.1000-.4748
1/8" gauge stelometer	.5047	.3870-.6224
0" gauge stelometer	.4190	.2000-.6380

Table 3. (V_r, W_r) regression coefficients.

Measurement	Slope	95% confidence limits
2.5% span length	.5785	.3903 - .7667
Micronaire	.3929	(-.3651)-1.1509
1/8" gauge stelometer	.7751	.5209 -1.0293
0" gauge stelometer	.9163	.2018 -1.6308

Table 4. Mean parameter estimates of the fiber measurements.

Parameter	2.5% span length	Micronaire	1/8" gauge stelometer	0" gauge stelometer
E_0	.000297*	.0251**	.0036**	.0070*
E_1	.000201**	.0146*	.0021**	.0046**
D	.002411*	.1256	.0316**	.0635**
F	.000047	.0391	.0101*	.0255*
H_1	.001513*	.1407	.0153	.0424
H_2	.001489*	.1128*	.0124	.0337*

*, ** Significantly different from zero at the 0.05 and 0.01 levels of probability, respectively.

Table 5. Mean estimator ratios of the four fiber characters.

Estimator	2.5% span length	95% confidence limits	Micronaire	95% confidence limits
H_1/D	.64	.40 - .88	1.16	.68 -1.64
$(H_1/D)^{1/2}$	.79	.62 - .96	1.08	.86 -1.30
$(V_{1L1}-E)/(W_{0L01}-E/n)$	.78	.56 -1.00	.98	.70 -1.26
K	1.64	(-1.20)-4.48	.37	.11 - .63
$1/4\ H_2/H_1$	.2478	.2211 - .2745	.2071	.1456- .2686
$1/4D/(1/4D + 1/4H_1-1/4F+E)$	.4942	.1728 - .8156	.3971	.0830- .7112
Estimator	1/8" gauge stelometer	95% confidence limits	0" gauge stelometer	95% confidence limits
H_1/D	.49	(-.10)-1.08	.69	(-.14)-1.52
$(H_1/D)^{1/2}$	.69	.30 -1.08	.81	.25 -1.37
$(V_{1L1}-E)/(W_{0L01}-E/n)$	.64	.25 -1.03	.77	.29 -1.25
K	.08	(-.09)- .25	.37	(-.79)-1.53
$1/4\ H_2/H_1$	.1992	.1415 - .2569	.2079	.1193 - .2965
$1/4D/(1/4D+1/4H_1-1/4F+E)$	.6711	.2890 -1.0532	.6815	.1591 -1.2039

assumption may account for at least part of the partial noncompliance of the length and micronaire data to the assumptions.

The other three assumptions cannot be tested with the present data, but additional experimentation is under way to obtain data to test the assumptions of no genotype-environment interaction within locations and years and of no epistasis. No method for testing the assumption of no multiple alleles is known to the authors at present.

When a trait exhibits a partial failure of the assumptions, estimates of the population parameters of that trait are still possible (5), although the estimates for such a trait are probably less reliable than when all assumptions are fulfilled. The parameter estimates are listed in Table 4. Here, each replication was treated as a separate experiment with its own estimate of environmental variation as suggested by Nelder (12). One estimate of each parameter for each trait was obtained from each replication. The standard errors of the mean, used in the tests of significance, were estimated by the variation of the block values around the overall mean.

The parameter, E_0, is an estimate of the parental environmental variation while E_1 is the corresponding estimate of the F_1 environmental variation. Estimates of E_0 were obtained from a between plot-within plot analysis of variance of the parental entries within each replication for each trait. Since all of the other parameter estimates in the diallel were calculated on a plot mean basis, it was necessary to convert the estimates of E_0 to an equivalent basis by dividing the within plot mean square by the average number of plants within a plot. Estimates of E_1 were obtained in the same manner using the F_1 entries rather than the parental entries.

The remaining parameters (D, F, H_1, and H_2) are as defined by Jinks and Hayman (9) using the notation of Mather (11). D is the additive genetic variance parameter which may also include a portion of the additive $\times$ additive epistatic variance as well as the additive genetic variance itself. H_1 and H_2 are dominance genetic variance parameters which may include the dominance genetic variance proper, dominance $\times$ dominance epistatic variance, and additive $\times$ dominance variance as well as the portion of the additive $\times$ additive variance not included within D. F is an indicator of the relative frequency of dominant and recessive alleles in the parents and may take sign, whereas the other parameters are expected to be positive. A negative F value results if there is an excess of recessive alleles in the parents, while a positive value indicates an excess of dominant alleles. F will equal zero if the dominant and recessive alleles of each gene are distributed equally among the parents or if no genes exhibit dominant effects (3). Estimates of these four parameters were obtained by solving the equations which follow (3, 5) where n equals the number of parents:

[1] Variance of the parents $= V_{0L0} = D + E_0$

[2] Mean covariance of arrays $= W_{0L01} = \frac{1}{2}D - \frac{1}{4}F + E_0/n$

[3] Mean variance of arrays $= V_{1L1} = \frac{1}{4}D + \frac{1}{4}H_1 - \frac{1}{4}F + [E_0 + (n-1)\ E_1]/n$

[4] Variance of array means $= V_{0L1} = \frac{1}{4}D + \frac{1}{4}H_1 - \frac{1}{4}H_2 - \frac{1}{4}F + [E_0 + (n-2)\ E_1]/n^2$

The estimates of V_{0L0}, W_{0L01}, V_{1L1}, and V_{0L1} were obtained from the diallel table, and weighted estimates of the environmental variation were used in equations [3] and [4] because parents and offspring do not make equal contributions to V_{1L1} and V_{0L1}.

All estimates of environmental variation were significantly different from zero. Furthermore, in every case the estimate of E_0 was larger than the corresponding estimate of E_1 which reinforces the statement by Hayman (6) that in cotton the variation of the parents is not equal to the variation of the F_1's. Also, the estimates of E_0 and E_1 within any replication were positively correlated to some degree for every trait studied. Nelder (12) states that this correlation is mainly a consequence of the lack of a randomization set within blocks, and he suggests a complete randomization of all plants within a replication to alleviate this difficulty. The impracticality of his suggestion with an experiment of this size prohibited its use.

All estimates of D were significantly different from zero except the one for micronaire. This lack of significance for micronaire was in part due to the lack of consistency of the estimates from replication to replication and in part to the large t value associated with two degrees of freedom. Since only one estimate of each parameter is possible for each replication, the number of estimates that can be made becomes a matter of practical concern. Because the number of replications that can usually be included is limited, degrees of freedom are therefore small, and t values used for setting confidence intervals on means are large, which in turn causes large confidence intervals.

Each of the traits showed some degree of significant dominance except 1/8" gauge stelometer. This lack of dominance for 1/8" gauge stelometer was expected, since none of the 45 crosses differed significantly from its respective mid-parent value at the 0.05 level of probability for this trait. For each of the other traits some crosses differed significantly from their respective mid-parent values. Three crosses were significantly different from their mid-parent values for 0" gauge stelometer. One was in the direction of strong fiber. For micronaire, twelve crosses were significantly different from their mid-parent values with nine of them being in the direction of coarser fiber. In 11 crosses, the lengths of the F_1's were significantly different from their mid-parent values. All 11 were in the direction of longer fiber.

The estimates of F for 1/8" gauge stelometer and 0" gauge stelometer were significantly different from zero in the positive direction. Theoretically, this

should indicate an excess of dominant alleles in the parents. However, neither of the estimates of dominance variance for 1/8" gauge stelometer was significantly different from zero, and the overall mean difference between the F_1's and their corresponding mid-parent values was exactly zero. This apparent contradiction has no ready explanation. The significance of F for 0" gauge stelometer is more easily understood since three crosses for this trait did exhibit significant differences from their mid-parent values. Since the overall mean of the difference between the F_1's and their respective mid-parent values was —0.03, an excess of dominant alleles which operate in the direction of weaker fiber is suggested. Although the F values obtained for length and Micronaire were both positive, they were not significantly different from zero. Two alternatives for these latter two traits is therefore possible: either no genes exhibit dominance or the dominant and recessive alleles are distributed equally among the parents. The former alternative must be rejected on the grounds that the variances for H_1 and/or H_2 were significantly different from zero for both traits and that the overall means of the difference between the F_1's and their corresponding mid-parent values were 0.024 and 0.1 for the length and micronaire, respectively. Therefore, the latter explanation must be the correct one.

Various ratios were calculated using the parameters estimated in Table 4 to provide further information about the genetic systems operating for each trait. An estimate of each ratio was obtained in each replication. The standard errors of the mean, used for setting confidence limits on the ratios, were estimated by the variation of the block values around the overall mean as was done for the parameter estimates in Table 4. These ratios and their 95% confidence limits are given in Table 5.

H_1/D, $(H_1/D)^{1/2}$, and $(V_{1L1}-E)/(W_{0L01}-E/n)$ are weighted overall measures of the degree of dominance (3) and are expected to range between zero and one with partial dominance, at one with complete dominance, and above one with overdominance. If there is no dominance, the estimates are not expected to be significantly different from zero. These are overall estimates of the degree of dominance, and each individual cross does not necessarily display the same degree of dominance for the same trait. For example, of the 45 crosses studied, three had significantly longer staple length than the longer of their respective parents (overdominance), eight were significantly longer than their mid-parent values but not significantly different from their longer parent (complete dominance), and the remainder of the crosses were not significantly different from their mid-parent values (no dominance). However, averaged over all crosses, the mean of the F_1 generation was 0.024 longer than the mean of the mid-parent values but 0.005 shorter than the mean of the higher parents in each cross (partial dominance). All three estimates for length, 1/8" gauge stelometer, and 0" gauge stelometer were within the partial dominance range. None of the estimates for micronaire were significantly different from one, indicating overall complete dominance for that trait.

K is an estimate of the number of effective factors operating for a certain trait as defined by Mather (11), and it measures only those factors showing some degree of dominance. The formula (8) used to obtain these estimates is as follows: $K = (\text{overall progeny mean} - \text{parental mean})^2/\frac{1}{4}H_2$. These assessments of the dominance situation are underestimated if the dominance effects of all the genes concerned are not equal in size and direction or if the distribution of the genes is correlated or both (8, 11). As expected, the estimate of K for 1/8" gauge stelometer was very low and not significantly different from zero. The estimates for micronaire and 0" gauge stelometer were small. Since these characters displayed dominance in both positive and negative directions, lack of directional dominance was at least one of the reasons for the low estimates, although differences in size effects between genes displaying dominance could also have played a role in these traits. There was no evidence that correlated gene distributions for any of these traits was a cause of the low estimates. The estimate for length was the highest of the four, though it was also somewhat smaller than expected. Dominance effects unequal in size must be the reason for the low estimates for this trait. None of the estimates were significantly different from zero except the estimate for micronaire.

The average frequency of negative vs. positive alleles in the parents was estimated by $\frac{1}{4}\ H_2/H_1$ (9); and it, like K, only provides information about those genes with some degree of dominance. When the negative versus positive alleles are distributed equally among the parents, the quantity equals $\frac{1}{4}$. When the alleles are not distributed in such a manner, they are correlated to some degree and the quantity is less than $\frac{1}{4}$. None of the estimates obtained were significantly different from $\frac{1}{4}$ at the 0.05 probability level. Therefore, there is no reason to assume that the alleles for each trait are not distributed equally among the parents used in this experiment. Consequently, the assumption of uncorrelated gene distributions is fulfilled as was stated previously.

A narrow-sense heritability estimate (3), $\frac{1}{4}D/(\frac{1}{4}D + \frac{1}{4}H_1 - \frac{1}{4}F + E)$, was calculated for each character on a plot mean basis. All of the estimates were relatively high, and each was significantly different from 0 at the 0.05 probability level. The characters studied may be ranked by their relative heritabilities as to probable ease of selection in a breeding program in the following order: micronaire $<$ upper 2.5% span length $<<$ (1/8" gauge stelometer, 0" gauge stelometer). The fact that the heritabilities for 1/8" gauge stelometer and 0" gauge stelometer were above 0.5 indicates that the majority of the variance exhibited by these traits is additive and/or additive $\times$ additive in nature. Therefore, mass selection should be an effective breeding method for improving strength within this material. Mass selection for length and micronaire is expected to be somewhat less effective. To obtain a high degree of genetic progress in these two traits, some emphasis may have to be placed on pedigrees, sib tests, and progeny tests.

LITERATURE CITED

1. ALLARD, R. W. 1956. Estimation of prepotency from lima bean diallel cross data. Agron. J. 48:537-543.
2. BROWN, H. B., and J. O. WARE. 1958. Cotton. McGraw-Hill Book Company, Inc., New York. 566 p.
3. CRUMPACKER, D. W., and R. W. ALLARD. 1962. A diallel analysis of heading date in wheat. Hilgardia 32:275-318.
4. ENDRIZZI, J. E. 1962. The diploid-like cytological behavior of tetraploid cotton. Evolution 16:325-329.

5. HAYMAN, B. I. 1954. The theory and analysis of diallel crosses. Genetics 39:789-809.
6. ———. 1958. The theory and analysis of diallel crosses II. Genetics 43:63-85.
7. ———. 1963. Notes on diallel-cross theory. p. 571-578. *In* W. D. Hanson and H. F. Robinson (ed.), Statistical genetics and plant breeding. National Academy of Sciences-National Research Council, Washington, D. C.
8. JINKS, J. L. 1954. The analysis of continuous variation in a diallel cross of *Nicotiana rustica* varieties. Genetics 39:767-788.
9. ———, and B. I. HAYMAN. 1953. The analysis of diallel crosses. Maize Genet. Co-op. Newsletter 27:48-54.
10. KIMBER, G. 1961. Basis of the diploid-like meiotic behaviour of polyploid cotton. Nature 191:98-100.
11. MATHER, K. 1949. Biometrical genetics. Dover Publications, Inc., London. 158 p.
12. NELDER, J. A. 1953. Statistical models in biometrical genetics. Heredity 7:111-119.
13. STEEL, R. G. D., and J. H. TORRIE. 1960. Principles and procedures of statistics. McGraw-Hill Book Co., Inc., New York. 481 p.
14. WHITE, T. G., and T. R. RICHMOND. 1963. Heterosis and combining ability in top and diallel crosses among primitive, foreign, and cultivated American upland cottons. Crop Sci. 3:58-63.