Chromosoma (Berl.) 47, 309–326 (1974)
© by Springer-Verlag 1974

Genome DNA Content and Chromosome Organization in *Gossypium*

G. A. Edwards, J. E. Endrizzi, and R. Stein

From the Committee of Genetics and Departments of Agronomy and Plant Genetics and Mathematics, University of Arizona, Tucson

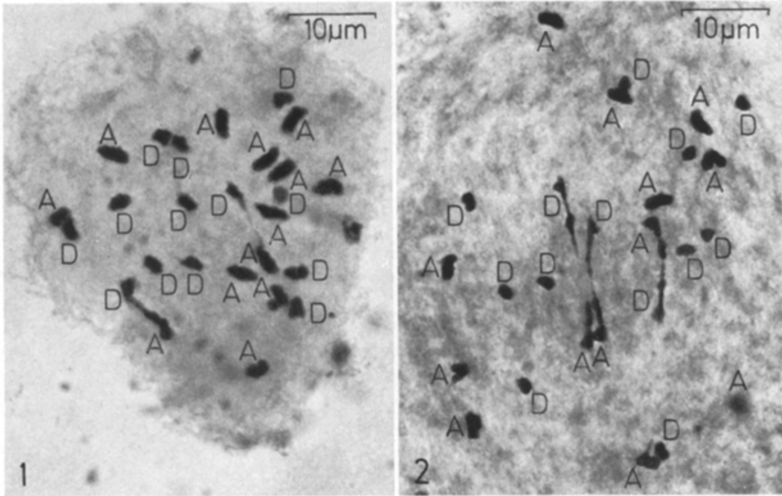
Abstract. The ascending genome size in *Gossypium* is assumed to be D, A, B, E and F, and C. Feulgen cytophotometry revealed that mean value of DNA content for each genome was D = 10.95, B = 13.88, F = 14.31, E = 18.24, A = 18.66, and C = 20.30, and that there is a close relationship of genomic chromosome size and DNA content. Evidence suggests that the five genomes with large chromosomes arose from a D genome-like progenetor by large scale, saltatory replication of repetitive DNA distributed uniformly through the ancestral genome. Corresponding adjustment in recombination units did not accompany the two-fold divergence in DNA value of the two homoeologous A and D genomes in the allotetraploid species. (Summary see p. 323.)

Introduction

Thirty-two diploid species ($2n = 2x = 26$) and five tetraploid species ($2n = 4x = 52$) are currently recognized in the genus *Gossypium*. On the basis of cytotaxonomic relationships, the diploid species have been classified into six genome groups which have been designated A through F (Beasley, 1942; Phillips and Strickland, 1966). The tetraploid species are made up of two genome groups, the A and the D.

Measurements of the total complement of somatic chromosomes have been reported in only two species, the amphidiploid *G. barbadense* (A_b and D_b genomes) and the diploid *G. arboreum* (A_2 genome) (Skovsted, 1934). He reported the average size of the chromosomes of the D_b genome to be 1.25 μ and of the A_b genome to be 2.25 μ. From his measurements of the 26 chromosomes of *G. arboreum*, we calculated an average length of 2.61 μ for this A genome species. The approximate 1:2 relationship for the D and A genome chromosomes is evident in the haploid of *G. hirsutum* (Fig. 1) and in the $F_1A_2D_5$ hybrid (Fig. 2). An examination of the somatic metaphase plates recorded by Skovsted (1935) for both parents and the hybrid of *G. sturtianum* (C genome) and *G. davidsonii* (D genome) shows that the C genome chromosomes are 2–3 times larger than the D genome chromosomes.

Cytological studies of pollen mother cells have been reported on almost all species, either individually or in hybrid combination, and they revealed that each genome group has chromosomes of a characteristic size and that all species in the same genome group have chromo-



Figs. 1 and 2. Metaphase I in PMCs of haploids of *G. hirsutum* and the F_1 A_2D_5 hybrid. Fig. 1 has 2 AD bivalents and 11A and 11D univalents. Fig. 2 has 3 AD bivalents and 10A and 10D univalents. Note that the *A* chromosomes are approximately twice the size of the *D* chromosomes

somes of similar size. Katterman and Ergle (1970) have summarized this information as relative meiotic chromosome size for the six diploid genomes, which is as follows: (1) the C genome species have very large chromosomes; (2) the E and F genome species have large chromosomes which are slightly larger than those of the A genome; (3) the B genome species have large chromosomes and some are slightly larger than the A genome; (4) the A genome species have moderately large chromosomes; and (5) the D genome species have small chromosomes. An examination of MI plates in a number of publications reporting chromosome associations in species hybrids shows that the most pronounced difference in chromosome size is between the D genome on one hand and the A, B, C, E and F genomes on the other.

Katterman and Ergle (1970) made a quantitative determination, by the diphenylamine method, of the DNA content of species in the genus and reported a reasonably good relationship between chromosome size and DNA content; however, several major discrepancies were noted. The DNA was isolated from decorticated seed and they assumed that cell and nuclear size were constant in the genus. Recent studies show that cell and nuclear size varies with genome size in *Gossypium* (Edwards, 1973; Edwards and Endrizzi, 1973). Such variation could have contributed to the discrepancies they noted. Until the present report, additional studies, more definitive in methodology, have not

been made to determine a more exact relationship between DNA content and chromosome size within *Gossypium*.

The present study reports the relative DNA content of representative *Gossypium* species and of a haploid and two hybrids as determined by the Feulgen cytophotometric technique. These data are utilized to determine the relationships between DNA content and genomic chromosome size in *Gossypium*. Some of the current concepts of chromosome structure are discussed in light of this relationship. This information is then discussed in relation to the gene linkages in the duplicate linkage groups of the amphidiploid species.

Materials and Methods

Plant Materials and Technique

Species selected for study included two representatives of each diploid genome group, except the F genome which has only one species, and two of the five tetraploid species. One further exception to selecting two representatives from each diploid genome group is noted in the C genome. Three representatives of this genome group were selected because Fryxell (1971) has raised the question whether *G. bickii* and *G. australe* belong in the C genome. A haploid of the tetraploid *G. hirsutum*, an F_1 hybrid of *G. arboreum* (A_2) $\times$ *G. raimondii* (D_5), and the colchicine doubled derivative of the F_1 hybrid, $2(A_2D_5)$, were also studied. The haploid and the hybrids were included primarily as additional reference points for assessing the reliability or accuracy of the DNA determinations. Table 1 lists the species and forms examined. Five selected plants of each species or type, cultured in six inch clay pots in a glass house, were utilized for the collection of root tips.

Standardized procedures, developed with preliminary tests, were followed throughout the study. Root tips from control plants, *G. herbaceum* var. *africanum*, and one other type were collected and processed in the same vial. Fixation was in Cornoy's 3:1 alcohol-acetic acid for one day at room temperature, hydrolized in 5N HCl at $20 \pm 0.25°$ C for 70 minutes, and stained by the Feulgen method. The root tips were softened in 5% pectinase for 3 hours before squashing. Prepared slides were assigned a random number randomly arranged to eliminate any possible bias during the readings.

Hydrolysis

Hydrolysis is one of the key steps in obtaining optimal binding of the Feulgen dye. Fox (1969) reported maximum Feulgen binding in *Vicia faba* root tips after about 40 minutes hydrolysis in 5N HCl at 20° C.

Gossypium herbaceum var. *africanum* root tips were hydrolyzed in 5N HCl at $20 \pm 0.25°$ C for various times from 10 to 90 minutes to determine the optional hydrolysis time for maximum Feulgen binding of the *Gossypium* material. Maximum dye binding first occurred at about 40 minutes and the plateau of staining was maintained through the final hydrolysis time of 90 minutes. Differences between hydrolysis times were examined by means of Duncan's New Multiple-Range Test, and significant differences of means were detected at the 0.05 level for 10, 20, 30 and 40 minutes. Results showed no statistical differences at hydrolysis times between 40 and 90 minutes. An *F* test was performed to examine differences between root tips within each of the eight hydrolysis times and no statistical differences were found, indicating that cells in root tips from different plants of the same

species have approximately the same amount of dye bound to the DNA. Since 70 minutes was a median, optimal time for hydrolysis, this was selected as the hydrolysis time for all species.

Cytophotometry

It has been recognized for sometime that within a single population, cells with chromosomes in a condensed state give lower Feulgen-DNA values than do cells with chromosomes in a less condensed state (Swift, 1950; Halkka, 1965; Garcia, 1970). Endrizzi (1962) has reported that the different genomes of *Gossypium* may exhibit differential coiling of their chromosomes. A more accurate comparison, therefore, would be obtained from interphase cells where the chromatin is in a dispersed state. However, problems exist in measuring interphase cells because one cannot distinguish between cells in the G1(2C), S(replication), and G2(4C) stages. It was therefore decided to take a population of cells from which the 2C cells could be separated by statistical methods. In order to measure a population of such cells, two major assumptions were made: (1) there are no differences in DNA content in root tips from the same plant or from different plants of a species, as suggested by the tests of *G. herbaceum* and reported above under Hydrolysis; (2) cells undergoing replication occurred in approximately equal numbers throughout the experiment and were not present in such abundance as to skew the 2C distribution from normal. This assumption appears to be quite valid as evidenced by the normal distributions of the 2C curves throughout the experiment.

The cells included in this population were all interphase-like and selected at random. An exception to random selection occurred when some 4C cells were excluded from measurement when they were found in groups. The data showed that this method of selection gives a fairly normal distribution of 2C cells, but a more irregular distribution of S and G2 cells. The 2C distribution was the one of interest, however, and this could be separated from the second distribution by the method of Bhattacharya (1967). This method uses the fact that the normal density can be transformed to be log-linear in the values x assumed by its random variable. The midpoints, x, of class intervals $(x-h/2,\ x+h/2)$ are plotted on the abscissa against $\Delta\log_e y$ where $y(x)$ is the observed probability density in the class interval centered at x and $\Delta\log_e y$ is equal to $\log_e y(x+h)$-$\log_e y(x)$. For a distribution composed of purely gaussian components, this function plots against x as a set of connected straight lines of negative slope forming angles Θ_i with the abscissa.

The mean of any particular gaussian component is given as the x intercept of the corresponding line plus $h/2$, and the variance is given by $(\mathrm{d}h \cot \Theta/b)$–$(h^2/12)$, where d and b denote the relative scales for $\Delta\log_e y$ and x respectively.

In general, any particular straight line is expected to fit most closely where the frequency or density is large. When, as was the case of this experiment, there are only two gaussian components, the classes with high and low x values contain considerable information in determining the 2C and the 4C distribution.

In order to plot the 2C distribution, $\Delta\log_e y$ was determined for each midpoint value and these points were used to fit a line by least squares to obtain the abscissa intercept and slope of the fitted line.

The method of Bhattacharya (1967) also permits the estimation of the number of observations in each class interval as well as the total number of observations associated with the 2C and the 4C distributions. This separation is necessary in order to proceed with any of the statistical separation of means techniques.

For each species, 75 cells were selected arbitrarily from each of seven root tips, giving a total of 525 cells. In each case the control consisted of 75 cells from each of three root tips, giving a total of 225 cells. Measurements were made on a Barr and Stroud Integrating Microdensitometer, type GN2, at a wavelength of 560 nm.

Results and Discussion

Feulgen-DNA Values for all Species

The populations of Feulgen-DNA values for each of the 17 species are illustrated as histograms and shown in Fig. 3 and 4. Each column of a histogram represents an interval of 2.0 and the Feulgen-DNA values are midpoints of selected intervals. All Feulgen-DNA values in these figures are unadjusted for the control.

The mean and standard deviation for each species were obtained from the left-hand section of the histogram, which represents the 2C distribution of Feulgen-DNA values. The method of separating the left-hand distribution is that outlined by Bhattacharya (1967). The unadjusted Feulgen-DNA means and standard deviations obtained from this method are shown below in Table 1.

In addition to the unadjusted Feulgen-DNA means, Table 1 shows also the adjusted Feulgen-DNA means for all species. The standard deviation is the same for the adjusted and unadjusted means of each species. The approximate mean value of all control means was 18, and therefore all controls were given the arbitrary mean value of 18.00. This selection of an arbitrary mean has no influence on the statistical separation of means. The value of each species was adjusted the same amount as its control by determining $\Delta \bar{X}$ as $\bar{X}$ of the species minus $\bar{X}$ of the control and adding this value to 18.00. The Feulgen-DNA means are adjusted to their control on the assumption that for any particular processing of root tips both the species mean and control mean are affected to the same degree and in the same direction by the same uncontrollable variables.

In order to indicate the presence of uncontrolled variation over and beyond expected statistical variation of the control species, *G. herbaceum* var. *africanum*, a separation of means at the 0.01 confidence level using the Scheffe procedure was done. The Scheffe method is appropriate in this instance because of its firm theoretical grounding and the fact that we are not interested in finding as many real differences as possible; but rather we are interested in assuring that if differences are found, they are real. Thus, there were 16 different replications of the control over different Feulgen batches, 5N HCl batches, fixative batches, etc. A measure of the variance of these control means was then assumed to be an estimator of the treatment variance of all species had they been replicated a like number of times.

The separations of the adjusted Feulgen-DNA means of all species are presented in Table 1 and are based on pair-wise comparison error rates using the least significant difference or LSD test at the 0.01 confidence level. Because of the large number of observations, a separation of means from distributions of unequal variance which accounts for the adjustment of means procedure previously described, can be accomplished without difficulty, since the ratio of the difference of any pair of means divided by the estimated standard deviation of this difference can be assumed to be normally distributed. Thus the LSD test can be made using tables of the normal distribution. The z values for the LSD tests give valid probability statements under the assumption that the variance between control means is an accurate measure of the treatment variance of the uncontrollable variable mentioned above. The z value was computed as $(\bar{X}_1 - \bar{X}_{c1})\text{-}(\bar{X}_2 - \bar{X}_{c2})/\sqrt{V_1 + 4V_2}$ where $\bar{X}_{c1}$ and $\bar{X}_{c2}$ are the control means and $\bar{X}_1$ and $\bar{X}_2$ are the species means. V_1 is the variance of the difference of species means and V_2 is a pooled variance associated with any control mean.

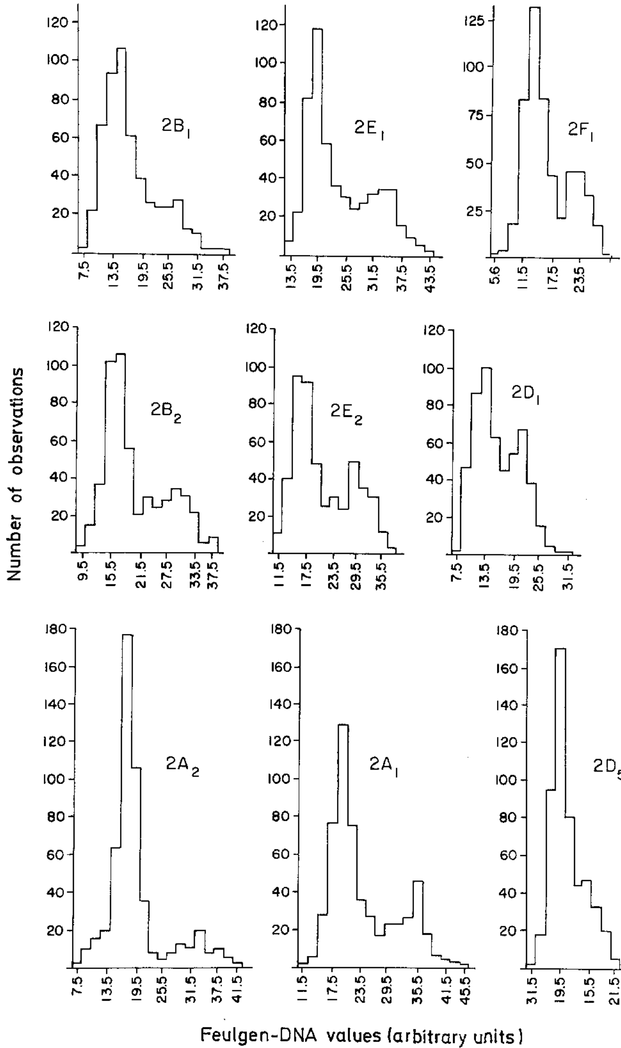


Fig. 3. Feulgen DNA values (arbitrary units) of interphase nuclei of diploid species of *Gossypium*

DNA Content in the Genus Gossypium

The relative DNA values for the species range from a high of 26.83 in the tetraploid, *G. hirsutum*, to a low of 10.38 in the diploid, *G. raimondii* (Table 1). This represents more than a two-fold difference in DNA values. Within the diploid genomes, *G. stocksii* has the highest relative

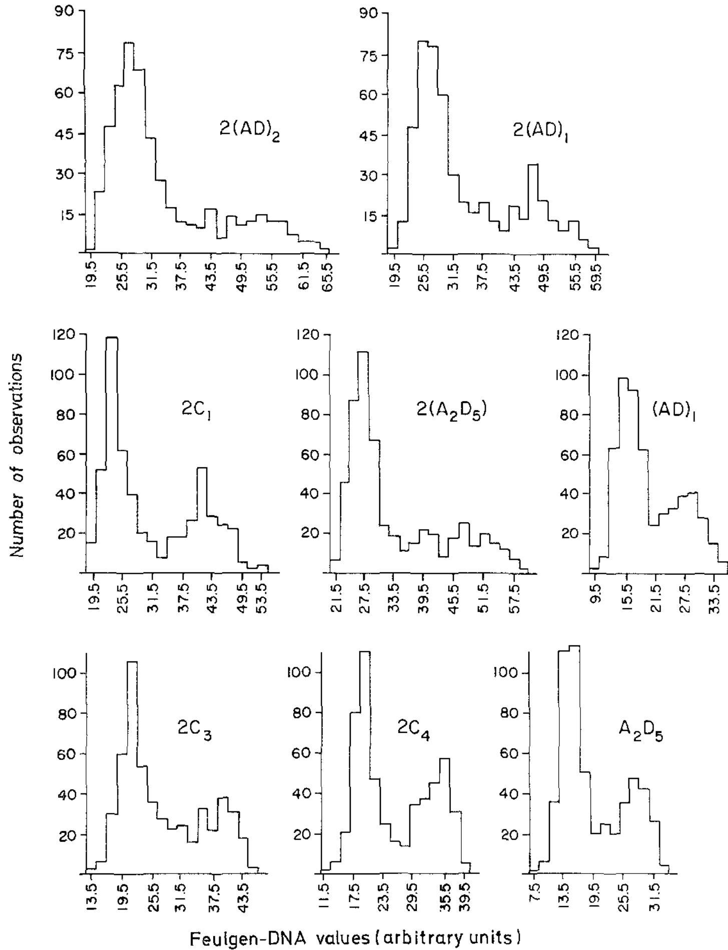


Fig. 4. Feulgen DNA values (arbitrary units) of interphase nuclei of tetraploid, haploid, and diploid species or forms *Gossypium*

DNA content with 21.46 and *G. raimondii* the lowest with 10.38, which is a two-fold difference.

A comparison of intragenomic values reveal that the DNA content of species within a genome group, with one major exception, are very similar. The A genome species have similar amounts of DNA and are not significantly different from each other. The same is true for both the B genome and the D genome species. In the C genome species *G. sturtianum* ($2C_1$) is not significantly different from *G. australe* ($2C_3$), and *G. australe*

Table 1. Feulgen-DNA values of 2C cells of 17 species and forms of *Gossypium*. Separation of adjusted Feulgen-DNA means based on pair-wise comparisons using the LSD test at the 0.01 confidence level

Species	Genome		Adjusted Feulgen-DNA means	Unadjusted Feulgen-DNA means	Standard deviations
Doubled Hybrid	$2(A_2D_5)$	a	26.97	26.93	2.51
G. hirsutum	$2(AD)_1$	a	26.83	27.29	3.31
G. barbadense	$2(AD)_2$	a	26.56	27.61	4.01
G. stocksii	$2E_1$	d	21.46	19.27	2.58
G. sturtianum	$2C_1$	d	21.45	23.86	2.28
G. australe	$2C_3$	df	20.61	21.43	2.59
G. arboreum	$2A_2$	fg	19.32	16.89	2.90
G. bickii	$2C_4$	fg	18.83	19.12	2.35
Haploid	$(AD)_1$	g	18.55	16.84	2.24
G. herbaceum	$2A_1$	g	18.00	19.78	2.78
G. somalense	$2E_2$	k	15.02	16.80	2.58
G. anomalum	$2B_1$	kl	14.46	14.83	2.53
G. longicalyx	$2F_1$	kl	14.31	13.83	2.31
Hybrid	A_2D_5	l	13.35	14.96	2.05
G. triphyllum	$2B_2$	l	13.29	16.65	2.62
G. thurberi	$2D_1$	p	11.52	13.54	1.90
G. raimondii	$2D_5$	p	10.38	9.63	2.23

is not significantly different from *G. bickii* ($2C_4$), but *G. sturtianum* is significantly different from *G. bickii.* This suggests, that as a group, these three C genome species are quite similar in DNA content and are members of the C genome species group. The two E genome species show a wider divergence of DNA content than any of the other genomes, and are significantly different from each other. *Gossypium stocksii* ($2E_1$) has about 1.4 times as much DNA as *G. somalense* ($2E_2$). Since there is only one F genome species no comparison is possible in this genome.

An intergenomic comparison is also possible for most of the genomes because of the similarity of DNA content within genomes. The exception to this similarity is in the E genome where the two species have a wide divergence of DNA content. With this point in mind, the relative DNA content of each genome can be illustrated as a mean of the species within the genome groups. The C genome would have the highest value with a mean of 20.30, followed by the A genome with a mean of 18.66, the E genome with a mean of 18.24, the F genome with a mean of 14.31, the B genome with a mean of 13.88, and the D genome with a mean of 10.95.

The comparison of *G. hirsutum* and its haploid is also possible. *G. hirsutum* has a relative DNA content of 26.83, and its haploid has a relative DNA content of 18.55, or a ratio of 1.4:1. The haploid plants

used in this experiment have been analyzed cytologically and each contains the expected number of 26 chromosomes. The 1.4:1 ratio observed is a discrepancy since a 2:1 relationship would be expected to exist. At the present time no suitable explanation can be given. This discrepancy cannot be resolved without further experimentation where the two sets of plants would be run together, one as the control for the other.

The two natural allotetraploid species, *G. hirsutum* and *G. barbadense*, and the synthetic $2(A_2D_5)$ hybrid have the same DNA content—26.83, 26.56, and 26.97, respectively. Cytogenetic data have established that the A and D genomes of the natural allotetraploids are more similar to the genomes of *G. herbaceum* $(2A_1)$ and *G. raimondii* $(2D_5)$ than to other diploid genomes in the A and D groups. Adding the DNA values of these two species gives a value of 28.38 which is almost the same as that determined for the allotetraploids, indicating that there has been little change, if any at all, of the DNA content of the two genomes of the allotetraploids since the origin of the tetraploid cottons.

When the hybrid, A_2D_5, is compared to its doubled derivative, $2(A_2D_5)$, the 1:2 relationship is very evident. The doubled hybrid has 2.04 times as much DNA as the F_1 hybrid. When the means of *G. arboreum* $(2A_2)$ and *G. raimondii* $(2D_5)$ are added, a value of 29.60 is obtained, and half this value is 14.85. These values are quite close to the mean values of 26.97 and 13.35 obtained for the $2(A_2D_5)$ hybrid and the A_2D_5 hybrid, respectively.

DNA Content and Relative Chromosome Size of the Diploid Genomes

Since the length of the chromosomes of only two species has been measured, a precise comparison between DNA content and chromosome size cannot be made. However, the relative chromosome sizes of the diploid genomes, which are based on the visual observations by cotton cytologists and reported by Katterman and Ergle (1970) and the few measurements of Skovsted (1934), can be used as a guide for a general comparison. In the relative classification of Katterman and Ergle, the C genome is considered as having "very large chromosomes," the E, F, and B genomes as having "large chromosomes," the A genomes as having "moderately large chromosomes, " and the D genome as having "small chromosomes."

The C genome is considered to have very large chromosomes, and this is reflected to a large extent in relative DNA content (see Table 1). They have the highest Feulgen-DNA average of the six diploid genomes. However, *G. stocksii*, a species from the E genome which is considered to have large chromosome as opposed to the very large chromosomes of

C genome species, has a higher mean value than any of the C genome species, though not significantly different than the two highest C genome species. The A species are considered to have moderately large chromosomes. One of the A genome species, *G. arboreum*, has a higher DNA content than *G. bickii*, the C genome species with the least amount of DNA of three species tested, but is not significantly different from it.

The E genome species have chromosomes which are considered large, slightly larger than A genome species. *Gossypium stocksii* ($2E_1$) has a higher DNA content than either of the A genome species, and is significantly different from them. However, *G. somalense* ($2E_2$) has significantly less DNA than either of the A species. Skovsted (1935) noted that root-tip chromosomes of the two A genomes species and $2E_1$ are of the same size, but his drawings indicate that the latter may be slightly larger.

The F genome species is considered to have large chromosomes, slightly larger than the A genome species. The relative chromosome size is poorly reflected in the DNA content of this species because it has less DNA than either of the A genome species, and is significantly different from them.

The B genome species are considered to have large chromosomes, some slightly larger than those of the A genome species. The two B genome species have significantly less DNA content than the two A species. These results are in agreement with Skovsted's (1935) observation that the somatic metaphase chromosomes of $2B_1$ are somewhat smaller than the chromosomes of the two A genomic species.

The D genome species have small chromosomes, and this is reflected in the DNA content of the D genome species. These two species have less DNA than any of the other species and are significantly different from them, though not different from each other. The significant difference of the DNA content of the D genome from all other diploid genomes was expected since the chromosomes of the D genome are one-half to one-third the size of the chromosomes of the other genomes.

Variation of DNA content occurs within a genome group, and some of the genome groups overlap each other in DNA content. The overlapping includes species in the A, B, C, E, and F genomes, which is not surprising since these five genomes have "large" chromosomes with the relative classification of moderately large, large, and very chromosomes.

The Feulgen-DNA values given in Table 1 show that chromosome size is related to DNA content in the genus *Gossypium*. This indicates that differences in chromosome size can be accounted for by differences in DNA content. Therefore, the proposal by Endrizzi (1962) that differences in chromosome size in the genus *Gossypium* are due to differential

condensation of chromosomes is untenable. However, exact measurements of chromosome length and volume are needed to rule out the possibility that some degree of differential condensation does occur which may account for some of the visual differences in relative chromosome size noted in squash preparations of pollen mother cells.

DNA Content and Chromosome Structure

The unineme theory of the chromatid is the widely held view of chromosome structure (Thomas, 1971; Callan, 1972; Rees and Jones, 1972; Huberman, 1973). It is assumed in the present discussion that each chromosome in the species of *Gossypium* consists of a single double helix of DNA. Among the different genomes of *Gossypium* there is no indication of DNA increases by whole strand doubling since the DNA content does not fall into an orderly series of increases. The genome groups of the diploid species cover a range of DNA values which is best explained as a reflection of longitudinal increases or decreases in DNA content.

The classic studies by Britten and Waring (1965) and Britten and Kohne (1967, 1968) indicated that the genomes of eukaryotic organisms probably evolved by a continuing process of addition and/or amplification of repeated sequences of DNA. Numerous studies by other workers with different organisms have since verified their observations, all of which indicate that large populations of repeated DNA sequences, in some cases as high as 80–90% of the DNA, exist universally in the genomes of eukaryotes above the fungi (Britten and Kohne, 1969; Britten and Davidson, 1971; Bostock, 1971; Rice, 1972). This repetitive DNA can have a wide variety of distribution in the genomes of different organisms (Eckardt, 1972) and can be grouped into three general categories: (a) localized in the centromeric heterochromatic region of the genome, (b) distributed at multiple sites in the genome, and (c) distributed in a more generalized or symmetrical pattern through the genome.

We believe that the data in *Gossypium* best fit the hypothesis that the difference in the amounts of DNA between the different genome groups is the result of amplification of the highly repetitive sequences that are distributed throughout the whole genome (Edwards and Endrizzi, 1973). The data that support this contention are discussed below.

With the use of the Wagner Divergence Index for determining phylogenetic relationships, Fryxell (1971) noted that the species exhibiting the least evolutionary advancement occur in the D genome group. He pointed out that his phenetic analysis conforms in broad outlines to views in species relationships derived from cytogenetic data in species hybrids. On the basis of these findings, we may reasonably assume that the amount of DNA in the D genome is representative of the ancestral

content of the genus; and that the other five genomes with the larger chromosomes evolved from a D genome-like progenitor by amplification of the repeated sequences.

Chromosome measurements of Skovsted (1934) and observations by Mursal (1973) of paired and unpaired chromosomes at Metaphase I of haploids of the 2(AD) allotetraploid species and that of F_1 of the *G. arboreum-raimondii* hybrid (A_2D_5) reveal that the chromosomes of the A genome are about twice the size of the D genome chromosomes (see Fig. 1 and 2). It is also noted that a range in chromosome size, though not differentially great, can be seen within each genome. Bivalents, associated predominantely by a single chiasma in one arm, are frequently seen at Metaphase I, primarily in the A_2D_5 hybrid where a complete complement of 13 bivalents may occur. The bivalents are asymmetrical and it is apparent in these bivalents that A and D chromosomes of corresponding rank size are paired (Endrizzi and Phillips, 1960; Mursal, 1973). Mursal (1973 and personal communication) has made extensive studies of diakinesis and pachytene in these two plant forms, and at diakinesis he observed the same pattern of chromosome size association. In pachytene, Brown (1961) and Mursal have recorded a surprisingly high amount of AD pairing, an amount much greater than what is reflected at the subsequent Metaphase I stage. Mursal informs us that pairing appears to be initiated at the telomeric ends and frequently may extend throughout the length of both chromosomes. In these cases, one of the chromosomes in the pair may show one or more loops, suggesting that it contains a greater length of DNA. In many of the AD bivalents pairing or synapsis may be incomplete and extend over only a short length of the chromosome. A difference in length of the associated members is frequently seen in these incompletely paired bivalents. In *Lolium* and *Allium*, very similar observations were reported by Rees and Jones (1967) and Jones and Rees (1968) in the bivalents of the F_1 hybrids between species with large differences in chromosome size and DNA content. These differences seen at pachytene are inconsistent with a model of chromosomes behavior based on the polynemic hypothesis with a two-fold increase in DNA helices (Rees and Jones, 1972).

The mean DNA value per genome in *Gossypium* was estimated to be C = 20.30, A = 18.66, E = 18.24, F = 14.31, B = 13.88, and D = 10.95. On the assumption that these mean values embody the actual pattern of evolutionary change in the quantitative amounts of DNA between genomes, it is then apparent that the range of DNA values is best explained as a reflection of longitudinal increases in DNA content.

The size variation of chromosomes between genomes is not unique to *Gossypium*, in fact, it is a common phenomenon that has accompanied divergence and speciation of many diploid species within genera of higher

plants (Rees, 1972; Rees and Jones, 1972). One example of many that parallels the situation in *Gossypium* is provided by the genus *Lathyrus* where there is about a three-fold difference in genome size between the diploid species (Rees and Hazarika, 1967). As in *Gossypium*, chromosome size in *Lathyrus* is closely correlated with nuclear DNA content. An increase in the proportion of repetitive DNA in species with higher DNA content has been observed in several conifers by Miksche and Hotta (1973). Species in several other genera of the *Malvaceae* have variations in chromosome size comparable to that in *Gossypium* and could reflect a similar pattern of chromosome evolution.

The results presented in *Gossypium* are consistent with a model which states that the increase of genomic DNA content in the species evolved by a process of amplification of existing sequences of repetitive DNA throughout the entire genome.

On the assumption that the D genome represents the ancestral amount of DNA, the other five genomes would then have arisen from a D genome-like progenitor by large scale, saltatory replication of the existing repeated sequences that were distributed uniformly throughout the ancestral genome. With this model, the repetitious DNA in the five genomes, A, B, C, E, and F, would differ from each other on a much smaller scale.

Fryxell's (1971) phenetic analysis shows that the D genome species exhibiting the most primitive characters are followed in order of divergence by a species in the B genome. This is in agreement with the cytophotometric data of DNA content and the visual estimates of chromosome size, which rank the B genome chromosome in between the D and the A, C, E, F genomes for these characters. These data suggest that the B genome arose directly from a D genome-like ancestor. From this point on, the origin of the A, C, E, F genomes would be pure conjecture since several schemes can be devised for their evolution beginning with the D and/or B genomes.

It has been proposed that the extra DNA may have a regulatory function affecting patterns of organization and activity at cellular and higher levels, leading to the final phenotype of the organism (Britten and Davidson, 1969, 1971; Britten, 1972). It is our contention that the amplification of the repetitive DNA played a major role in genomic evolution and speciation in the genus *Gossypium*.

One would predict that the increase in chromosome length due to amplification of repeated sequences distributed uniformly through the genome would increase the distance between the unique sequence, *i.e.*, the functional genes. Rees (1972) in fact has discussed this point and states that there is no direct evidence in any species for an adjustment

Table 2. Homoeologous linkage maps of *Gossypium hirsutum*

Genome								
A	Lc_1	32	Yg_2	4	Cl_2	16	R_2	
D	Dw	29	Yg_1	14	R_1	15	Cl_1	
A	Bw_1	15	Gl_2	32			Ne_1	[a]
D	Bw_2	20	Gl_3	38			Ne_2	[a]
A	L^L	38.1	●		Lp_1			
D	L°	32.6[b]	●		Lp_2 [c]			

[a] Gene order according to Kohel (private communication).
[b] The 32.6 map units may be under estimated since it was determined in a translocation heterozygote and it is the amount of recombination between L° and the breakpoint which is located in the short arm near the centromere (Wilson and Kohel, 1970).
[c] Preliminary linkage data suggest that Lp_2 is to the right of the centromere.
● = Centromere.

in linkage as a result of amplification of DNA. We believe that the amphidiploid *G. hirsutum* provides an excellent example of this kind. The amphidiploid is composed of the A and D genomes in which the A genome chromosomes have about twice as much DNA as the D genome chromosomes.

Ten linkage groups have been established in both the A and D genomes of *G. hirsutum* and six involve homoeologous chromosomes (Kohel, 1972; and Endrizzi, unpublished data). The six linkage maps are shown in Table 2. Each pair of homoeologs consists of an A and a D chromosome that have been identified in a series of cytogenetic tests involving monosomes, telosomes, translocations and species hybrids. Visual observation of these chromosomes as monosomes and/or telosomes at MI revealed that the A chromosomes are approximately twice the size of their D chromosome homoeolog.

Except for the inverted position of the R_1 locus and the associated difference in map distance, the gene order and map distance within the three sets of homoeologs are very similar. Some small differences in map distances are evident, but they do not in any way reflect the twofold difference in chromosome size and DNA content that was observed between the two homoeologous groups. The difference in some cases may be real; however, it is well known that the amount of crossing-over in a given region is not absolutely fixed but is variable from one experiment to another due to the normal statistical variations in random sampling, population size, and fluctuations in environmental conditions.

In addition to the duplicate loci in the linkages given here, other duplicate loci are known in *G. hirsutum* and some of these involve homoeologous linkages but have not been worked out as yet.

The above results suggest that the two-fold difference in DNA content, and concomitant chromosome size that occurred in the divergence of the A and D genomes which later were combined to form the allotetraploid *G. hirsutum*, was not accompanied by a corresponding adjustment in recombination units; in fact, the amount of recombination between loci has changed very little, if any at all. This may not be surprising in view of the recent evidence for the presence of recombination genes (*rec*) whose effects are strictly localized in specific regions of the chromosomes (Smith, 1966; D. G. Catcheside, 1968; Thomas and Catcheside, 1969; Smyth, 1971; Catcheside and Corcoran, 1973). This coupled with the hypothesis for the existence of specific sequences of nucleotides in DNA molecules, or recognition loci, that respond in some manner to the *rec* gene product and the recombinases (Angel, Austin, and Catcheside, 1970; D. E. A. Catcheside, 1970; Catcheside and Austin, 1971; Sobell, 1972), suggest that the DNA incisions initiating recombination are site specific and independent of any repetitive DNA sequences.

Such specificity of control of recombination would readily explain the lack of change in genetic map distances between duplicate loci in homoeologous chromosomes even though there is a two-fold difference in their DNA content. The evidence suggests that the increase in DNA content of the genomes of *Gossypium* occurred by amplification of the repetitive sequences throughout the genome, while at the same time the unique sequences, including the recognition sites of recombination remain unchanged and constant in each genome.

Summary

The relative DNA content for representative species of each of the six diploid genomes, for two tetraploid species, an A_2D_5 hybrid and its doubled derivative, and a haploid of one of the tetraploid species in the genus *Gossypium* was determined by Feulgen cytophotometry.

All species in the same diploid genome group have chromosomes of similar size, but each genome group has chromosomes of a characteristic size. The ascending genome size based on visual estimates is as follows: D, A, B, E and F, and C.

The intragenomic values revealed that the DNA content of species within a genome group, with one exception, are very similar. The mean value of DNA content for each genome was D = 10.95, B = 13.88, F = 14.31, E = 18.24, A = 18.66 and C = 20.30. Intergenomic comparisons show some deviation of the means of the genomic DNA value and

relative genomic size based on visual comparisons, but for the most part there was a close relationship of genomic chromosome size and DNA content.

The haploid and its tetraploid counterpart had a ratio of relative DNA content of 1:1.4, rather than the expected of 1:2.

The $F_1A_2D_5$ hybrid and its doubled derivative $2(A_2D_5)$ had a 1:2 ratio of relative DNA values. The added DNA values of *G. arboreum* $(2A_2)$ and *G. raimondii* $(2D_5)$ and half this value are very close to the DNA values obtained for $2(A_2D_5)$ and A_2D_5 hybrids, respectively.

The evidence suggests that the chromosomes of *Gossypium* are uninemic and that the DNA content of D genome represents the ancestral amount of DNA. It is presumed that the other five genomes with large chromosomes arose from a D genome-like progenitor by large scale, saltatory replication of repeated sequences of DNA that were distributed uniformly throughout the ancestral genome. The amplification of the repetitive DNA is assumed to have played a major role in genomic evolution and speciation in the genus *Gossypium*. A comparison of the linkage relationships of genes in homoeologous linkage groups with the two-fold difference in DNA content of the chromosomes of the A and D genomes of the allotetraploid *G. hirsutum* showed that a corresponding adjustment in recombination units did not accompany the two-fold divergence in DNA values of the two homoeologous genomes.

Acknowledgements. This study was carried out by the senior author (G. A. E.) in partial fulfillment for the requirements for the Ph. D. degree in genetics, and was supported in part by Regional Research Project S-77, Genetics and Cytology of Cotton II. The authors wish to thank Dr. Peter Bartels of the Department of Microbiology and Medical Technology for his assistance and use of the Barr and Stroud Integrating Microdensitometer, and Ibrahim El Jack Mursal for the use of Fig. 1 and 2. Journal paper no. 2309 of the Arizona Agricultural Experiment Station.

References

Angel, T., Austin, B., Catcheside, D. G.: Regulation of recombination at the *his*-3 locus in Neurospora crassa. Aust. J. biol. Sci. **23**, 1229–1240 (1970)

Beasley, J. O.: Meiotic chromosome behavior in species, species hybrids, haploids, and induced polyploids of *Gossypium*. Genetics **27**, 25–54 (1942)

Bhattacharya, C. G.: A simple method of resolution of a distribution into gaussian components. Biometrics **23**, 115–135 (1967)

Bostock, C.: Repetitious DNA. Advanc. Cell Biol. **2**, 153–223 (1971)

Britten, R. J.: DNA sequence interspersion and a speculation about evolution. Brookhaven Symp. Biol. **23**, 80–94 (1972)

Britten, R. J., Davidson, E. H.: Gene regulation for higher cells: A theory. Science **165**, 349–357 (1969)

Britten, R. J., Davidson, E. H.: Repetitive and non-repetitive DNA sequences and a speculation on the origin of evolutionary novelty. Quart. Rev. Biol. **46**, 111–138 (1971)

Britten, R. J., Kohne, D. E.: Nucleotide sequence repetition in DNA. Carnegie Inst. Wash. Yearbook **65**, 78–106 (1967)

Britten, R. J., Kohne, D. E.: Repeated sequences in DNA. Science **161**, 529–540 (1968)

Britten, R. J., Kohne, D. E.: In: Handbook of molecular cytology (A. Lima-de-Faria, ed.), p. 21–51. Amsterdam: North Holland Publ. 1969

Britten, R. J., Waring, M.: Repetition of nucleotide sequences. Carnegie Inst. Wash. Yearbook **64**, 331–333 (1965)

Brown, M. S.: Chromosome differentiation in *Gossypium*. Amer. J. Bot. **48**, 532 (1961)

Callan, H. G.: Replication of DNA in chromosomes of eukaryotes. Proc. roy. Soc. B **181**, 19–41 (1972)

Catcheside, D. E. A.: Control of recombination within the *nitrate*-2 locus of Neurospora crassa: An unlinked dominant gene which reduces prototroph yields. Aust. J. biol. Sci. **23**, 855–865 (1970)

Catcheside, D. G.: The control of genetic recombination in *N*. crassa. In: Replication and recombination of genetic material (W. J. Peacock and R. D. Brock, eds.), p. 216–226. Canberra: Australian Academy of Sciences: 1968

Catcheside, D. G., Austin, B.: Common regulation of recombination at the amination-1 and histidine-2 loci in Neurospora crassa. Aust. J. biol. Sci. **24**, 107–115 (1971)

Catcheside, D. G., Corcoran, D.: Control of non-allelic recombination in Neurospora crassa. Aust. J. biol. Sci. **26**, 1337–1353 (1973)

Eckhardt, R. A.: Chromosomal localization of repetitive DNA. Brookhaven Symp. Biol. **23**, 271–292 (1972)

Edwards, G. A.: DNA content by Feulgen cytophotometry and a determination of the relationship of DNA with nuclear, cell and chromosome sizes in the genus Gossypium L. Ph. D. dissertation submitted to the Faculty of the Committee of Genetics, The University of Arizona, Tucson (1973)

Edwards, G. A., Endrizzi, J. E.: Feulgen cytophotometry of DNA content and a determination of the relationship of DNA content to nuclear, cell and chromosome sizes in the genus Gossypium. (Abst.) Genetics **74**, Suppl., 69 (1973)

Endrizzi, J. E.: The diploid-like cytological behavior of tetraploid cotton. Evolution (Lawrence, Kans.) **16**, 325–329 (1962)

Endrizzi, J. E., Phillips, L. L.: A hybrid between Gossypium arboreum L. and G. raimondii, Ulb. Canad J. Genet. Cytol. **2**, 311–319 (1960).

Fox, D. P.: Some characteristics of the cold hydrolysis technique for staining plant tissues by the Feulgen reaction. J. Histochem. Cytochem. **17**, 266–272 (1969)

Fryxell, P. A.: Phenetic analysis and the phylogeny of the diploid species of Gossypium L. (Malvaceae). Evolution (Lawrence, Kans.) **25**, 554–562 (1971)

Garcia, A.: Stoichiometry of dye binding versus degree of chromatin coiling. In: Introduction to quantitative Cytochemistry (G. L. Wied and G. F. Bahr, eds.), vol. 2, p. 153–170. New York: Academic Press 1970

Halkka, O.: A photometric study of the Luzula problem. Hereditas (Lund) **52**, 81–88 (1965)

Huberman, J. A.: Structure of chromosome fibers and chromosomes. Ann. Rev. Biochem. **42**, 355–378 (1973)

Jones, R. N., Rees, H.: Nuclear DNA variation in Allium. Heredity **23**, 591–605 (1968)

Katterman, F. R. H., Ergle, D. R.: A study of quantitative variations of nucleic acids in Gossypium. Phytochemistry **9**, 2007–2010 (1970)

Kohel, R. J.: Linkage tests in Upland cotton, Gossypium hirsutum L. II. Crop Sci. **12**, 66–69 (1972)

Miksche, J. P., Hotta, Y.: DNA base composition and repetitious DNA in several conifers. Chromosoma (Berl.) **41**, 29–36 (1973)
Mursal, I. El Jack: A cytological study chromosome pairing in cotton. J. Ariz. Acad. Sci. 8, Proc. Suppl. p. 15 (1973)
Phillips, L. L., Strickland, M. A.: The cytology of a hybrid between Gossypium hirsutum and G. longicalyx. Canad. J. Genet. Cytol. 8, 91–95 (1966)
Rees, H.: DNA in higher plants. Brookhaven Symp. Biol. **23**, 80–94 (1972)
Rees, H., Hazarika, M. H.: Chromosome evolution in *Lathyrus*. Chromosomes today **2**, 158–165 (1967)
Rees, H., Jones, G. H.: Chromosome evolution in *Lolium*. Heredity **22**, 1–18 (1967)
Rees, H., Jones, R. N.: The origin of the wide species variation in nuclear DNA content. Int. Rev. Cytol. **32**, 53–92 (1972)
Rice, N. R.: Changes in repeated DNA in evolution. Brookhaven Symp. Biol. **23**, 44–79 (1972)
Skovsted, A.: Cytological studies in cotton II. Two interspecific hybrids between Asiatic and New World cottons. J. Genet. **28**, 407–424 (1934)
Skovsted, A.: Chromosome numbers in the Malvaceae I. J. Genet. **31**, 263–296 (1935)
Smith, R.: Genetic controls of recombination I. The recombination-2 gene of Neurospora crassa. Heredity **21**, 481–498 (1966)
Smyth, D. R.: Effect ot rec-3 on polarity of recombination in the amination-1 locus of Neurospora crassa. Aust. J. biol. Sci. **24**, 97–106 (1971)
Sobell, H. M.: Molecular mechanism for genetic recombination. Proc. nat. Acad. Sci. (Wash.) **69**, 2483–2487 (1972)
Swift, H.: The constancy of desoxyribosenucleic acid in plant nuclei. Genetics **36**, 643–653 (1950)
Thomas, C. A., Jr.: The genetic organization of chromosomes. Ann. Rev. Genet. **5**, 237–256 (1971)
Thomas, P. L., Catcheside, D. G.: Genetic control of flanking marker behavior in an allelic cross of Neurospora crassa. Canad. J. Genet. Cytol. **11**, 558–566 (1969)
Wilson, F. D., Kohel, R. J.: Linkage of green-lint and okra-leaf genes in a reciprocal translocation stock in Upland cotton. Canad. J. Genet. Cytol. **12**, 100–104 (1970)

Received May 13, 1974 / Accepted by H. Bauer
Ready for press May 16, 1974

Dr. G. A. Edwards
formerly graduate student
Committee of Genetics and
Department of Agronomy
and Plant Genetics
University of Arizona
present address:
County Agent
Monticello, Utah 84535 U.S.A.

Dr. J. E. Endrizzi
Department of Agronomy
and Plant Genetics
Dr. R. Stein
Department of Mathematics
The University of Arizona
Tucson, Arizona 85721 U.S.A.