

SINGLE COPY DNA HOMOLOGIES IN *ATRIPLEX*. I. CROSS REACTIVITY ESTIMATES AND THE ROLE OF DELETIONS IN GENOME EVOLUTION*

HEATHER S. BELFORD† and WILLIAM F. THOMPSON

Department of Plant Biology, Carnegie Institution of Washington, Stanford, CA 94305, U.S.A.

Received 21.vii.80

SUMMARY

Genome sizes and single copy complexity values have been estimated for eight *Atriplex* species and spinach by analysis of DNA reassociation kinetics. These values, together with measurements of interspecific hybridization carried out with purified single copy tracers, have been used to estimate the absolute amount of single copy DNA which is composed of homologous sequences in various species. The data show a large variation in cross reactivity for different species pairs which is best explained by postulating that these genomes were subject to extensive deletion during evolution of different lineages. At most, only about 5×10^7 nucleotide pairs of single copy DNA (about 10 times the amount in an *E. coli* genome) appear to be necessary to specify phenotypic features common to *Atriplex* species.

1. INTRODUCTION

VERY little is known about the relationships of speciation and genome evolution, especially among plants. Theoretically, when populations of organisms become separated into genetically isolated subgroups, differences begin to accumulate in the DNA of the subgroups. These differences include changes resulting from deletions or amplifications of large DNA segments as well as point mutations altering individual base pairs. Deletion and duplication are unpredictable events, while the accumulation of single base changes appears to be a clocklike process (Wilson *et al.*, 1977) such that the fraction of bases differing between two species is related to the time elapsed since the two species last shared a common ancestor.

DNA hybridization techniques can be used to detect both of these mutation processes. The kinetic analysis of DNA reassociation can be used to characterize individual genomes as to size and single copy sequence complexity, and to measure the extent to which different genomes share common sequences. Thermal stability profiles of interspecific DNA hybrids can be used to calculate the extent to which single base changes have accumulated in the DNA forming hybrid duplexes. To date, hybridization has been employed to examine DNA sequence evolution in several animal groups such as primates (Kohne, 1970), artiodactyls (Laird *et al.*, 1969),

* This research was supported in part by NSF grant DEB76-83405 and funds from the Carnegie Institution of Washington. This paper is Publication No. 717 from the Department of Plant Biology, and is based on a portion of a dissertation submitted by H.S.B. in partial fulfilment of the Ph.D requirements in the Department of Botany, University of Massachusetts, Amherst.

† Present address: Department of Biochemistry, Boston University, 80 E. Concord St., Boston MA 02118, U.S.A.

mice (Rice, 1972), birds (Shields and Straus, 1975; Sibley and Ahlquist, 1980), flies (Sohn *et al.*, 1975), and salamanders (Mizuno and MacGregor, 1974). In general, it is observed that within a genus nearly all single copy DNA sequences (80–100 per cent) are represented in every species, though with somewhat (0–3 per cent) modified sequence. The few instances in which cross reactivity has been greatly reduced among members of a single genus (salamander: Mizuno and MacGregor, 1974; sea urchin: Angerer *et al.*, 1976, and Harpold and Craig, 1978; *Xenopus*: Galau *et al.*, 1976) have involved genera with prolonged evolutionary histories and conserved phenotypes. Very little work has yet been reported on higher plant genera. In view of the fact that higher plants as a group exhibit a greater range of DNA contents than animals, it is of interest to know whether this tolerance reflects a different rate of DNA sequence amplification or deletion in plant evolution. In the present paper we discuss hybridization kinetic analyses designed to assess the extent to which deletions have contributed to genome evolution within the higher plant genus *Atriplex*. Thermal stability studies will be the topic of a second paper (Belford and Thompson, 1980).

Atriplex is a cosmopolitan genus often associated with halophytic, marginal environments. It has recently been shown to contain species with both C_3 and C_4 photosynthetic pathways (Björkman *et al.*, 1973). This discovery brought the previously accepted, morphologically-based phylogenetic scheme (fig. 1; Hall and Clements, 1923) under review since it implies a polyphyletic origin for the C_4 carbon-fixation pathway even within *Atriplex* itself. Considering the complexity of the C_4 pathway, a polyphyletic origin seems improbable, though not impossible, within a group so closely related as a single genus. Fossil data are not available for *Atriplex*

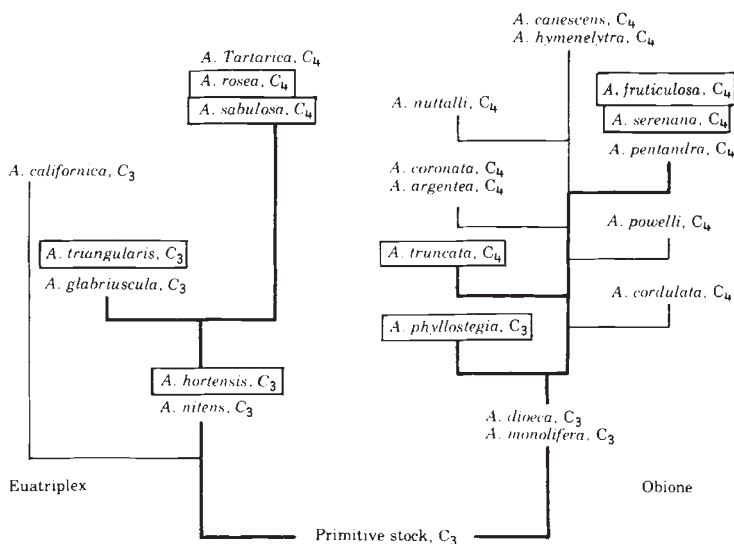


FIG. 1.—Phylogenetic scheme for the genus *Atriplex* based on morphological criteria. This fig. is adapted from fig. 29 of Hall and Clements (1923) monograph. C₃ and C₄ designations indicate the photosynthetic carbon fixation pathway in each species as determined by Björkman *et al.* (1973). The species in boxes are those for which single copy DNA sequence comparisons were made. The greater the distance of a species from the "Primitive stock", the fewer are the primitive traits displayed.

and mating barriers make verification of the morphologically-based phylogeny by traditional genetic methods extremely difficult (Nobs, 1976); phenotype is notoriously variable even with a single species. Thus a molecular approach to measuring evolutionary relationships in *Atriplex* seemed highly desirable.

In this paper, eight species of the genus *Atriplex* are characterized with respect to genome sizes, single copy sequence complexities and the extent to which single copy sequences are shared among the species. The DNA of spinach, *Spinacia oleracea*, is included as an extrageneric comparison. (Spinach, like *Atriplex*, is a member of the family *Chenopodiaceae*). We chose to focus on single copy, or unique, sequences rather than repeat sequences because of our intent to use the results of this study as a basis for a subsequent quantitative analysis of base sequence divergence and phylogenetic relationships.

2. MATERIALS AND METHODS

(i) *Plant material*

Diploid *Atriplex* seed stocks ($n = 9$) were provided by Dr Malcolm Nobs. Plants were grown in the green house or the field at the Department of Plant Biology, Carnegie Institution, Stanford, CA., under natural light and temperature conditions. Spinach was obtained fresh from a local market. Leaves were harvested in lots of 100–200 grams, frozen in liquid nitrogen, and lyophilized (Stein *et al.*, 1979).

(ii) *DNA extraction and purification*

Lyophilized tissue was homogenized with a Waring blender in two volumes extraction buffer (2 ml/g fresh weight) containing 8 M urea, 1.0 M sodium perchlorate, 100 mM tris-HCl (pH 8.4), 1 per cent 2-mercapto-ethanol, and 1 per cent octanol. The homogenate was centrifuged for 5 min at 5000 g. After re-extraction of the pellet with 1.0 volume of buffer, the supernatants were pooled and incubated at 23° for 30 min with 30 mg pronase (previously self-digested for 2 hr at 37° at a concentration of 10 mg/ml in 50 mM tris-HCl (pH 7.5), 100 mM Na-acetate) per 100 grams initial fresh weight. The supernatant was extracted twice with an equal volume of chloroform-octanol (24:1 v/v), filtered with suction through a bed of cellulose (25 g/100 g fresh weight tissue, moistened with extraction buffer) and precipitated with 2 volumes 95 per cent ethanol. The ethanol precipitate was dissolved in 5 mM EDTA, 50 mM tris-HCl (pH 7.5) and briefly dialyzed before addition of ethidium bromide (200 g/ml) and CsCl (final density 1.56–1.57). The CsCl-EtBr mixture was centrifuged in a Sorval TV850 vertical rotor for 15 hours or in a Beckman Ti50 rotor for 40 hours at 40,000 and 21°. The DNA band was removed from the gradient with a recurved 17-gauge needle, extracted with isopropanol (saturated with buffered CsCl), dialyzed, and ethanol precipitated. The precipitate was taken up in 120 mM PB, treated with AG50×8 cation exchange resin (Bio-Rad) to remove residual traces of ethidium, and then incubated with RNase A (25 µg/ml, 1 hour, 37°) and self digested pronase (100 µg/ml, 2 hours, 37°). After addition of an equal volume of 10 M urea, the DNA was

bound to 5–10 g HAP,* purified according to Britten *et al.* (1974), dialyzed and concentrated by ethanol precipitation.

Purified DNA was sheared to a length of about 350–400 NTP³ in the Virtis “60” homogenizer at 50K rpm for 1 hr following the procedure of Stein *et al.*, (1979).

(iii) Isolation of single copy fractions

Sheared DNA was dissolved at *ca.* 5 mg/ml in 600 mM PB³, 0.1 mM EDTA (pH 7.0), denatured at about 105° for 3 minutes and reassociated at 66° to an equivalent C_0t of 325 (all C_0t values are corrected to standard conditions of 120 mM PB, 60°). The reaction was stopped by chilling in ice water followed by dilution into chilled 120 mM PB. The dilution was warmed several minutes at 60° and loaded onto HAP (4 ml/mg DNA) equilibrated in 120 mM PB at the same temperature. The unbound DNA containing unreassociated fragments equal to 20–35 per cent of the original mass was recovered as the initial single copy fraction. To reduce repeated sequence contamination, a second round of reassociation and fractionation was usually added (see fig. 2). After each fractionation the samples were treated with AG50×8 cation exchange resin and Chelex 100 chelating resin. About 16 per cent of the original DNA mass normally remained unbound following the second fractionation.

(iv) Preparation of radioactive tracers

Labelled complementary DNA fragments were prepared using *E. coli* DNA polymerase I essentially as described by Murray *et al.* (1977). Each reaction mixture contained 25 µg/ml denatured DNA template, 250 µg/ml denatured oligonucleotide primer, 250 µM each of the three unlabelled dNTPs, 50 µM labelled dNTP, 5 mM MgCl₂, 70 mM tris-HCl (pH 7.5), 3 mM (NH₄)₂SO₄, 1 mM 2-mercaptoethanol, 30 mg/ml BSA, and 125–250 units/ml *E. coli* DNA polymerase I (Boehringer-Mannheim or Miles Laboratories), in 130 µl total volume. After 7–8 hours incubation at 15°, reactions were terminated by adding an equal volume of 10 mM EDTA and assayed for incorporation of label into acid-insoluble material. The labelled copy was purified by HAP chromatography in 8 M urea (Britten *et al.*, 1974). The samples were then dialyzed into 120 mM PB, denatured for 5 min at 100°C and loaded directly onto a 0.25 ml HAP column at 60°C to remove self-complementary (foldback) sequences. Purified, foldback-free tracers were extensively dialyzed against 10 mM NH₄ acetate, treated with Chelex 100 chelating resin, and stored frozen at –20°C. Preparations used for hybridization had specific activities of *ca.* 2.3×10^5 cpm ¹⁴C/µg DNA or 3.7×10^6 cpm ³H/µg DNA (measured under double label counting conditions). Average tracer fragment length was 215 NT.

(v) Alkaline agarose gel sizing

DNA fragments were sized by electrophoresis in alkaline 1.5 per cent agarose horizontal slab gels following the procedure of McDonnell *et al.*

* Abbreviations: HAP, hydroxyapatite (Bio Rad, Lot # 16020); PB, equimolar Na₂HPO₄ and NaH₂PO₄, pH 6.8; NTP, nucleotide pairs; dNTP, deoxynucleoside triphosphate.

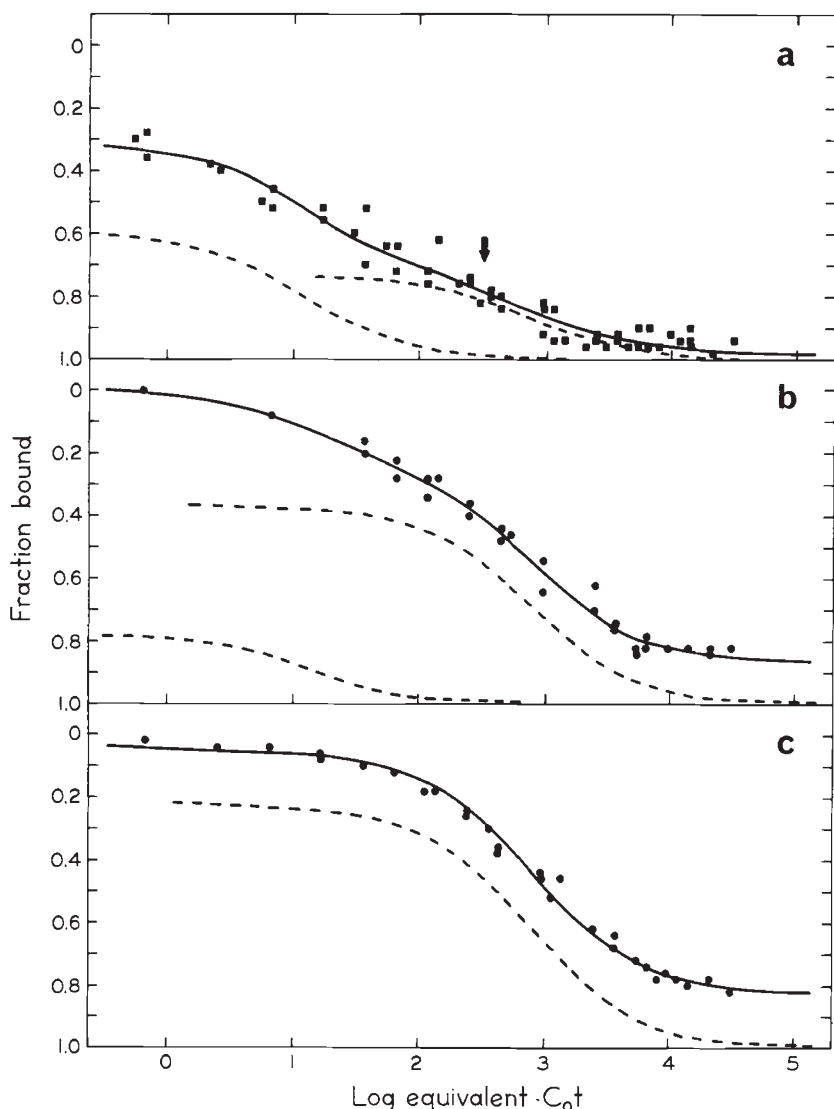


FIG. 2.—Reassociation kinetics of *A. serenana* total genome DNA and once- or twice-fractionated single copy tracers. All DNAs were reassociated and fractionated on HAP as described in the text. C_0t values have been corrected to their equivalent values in 0.12 M PB at 60°. Solid lines through the data points are computer-derived least squares fits including one or two ideal second order components (dashed lines). (a) Total *A. serenana* DNA reassociation kinetics. Note that reassociation was not assayed at low C_0t values and therefore the theoretical repeated sequence component is a rough estimate only. The arrow indicates the C_0t at which single copy sequences were isolated. (b) Kinetics for an *A. serenana* single copy tracer isolated by one cycle of reassociation and HAP fractionation, and reassociated with total DNA from *A. serenana*. Note that the data points are best fitted by two components, approximately 25 per cent repeated sequences and 75 per cent single copy sequences. (c) Kinetics of *A. serenana* single copy sequences isolated by two cycles of reassociation and fractionation, reassociated with *A. serenana* total DNA. The data are best fit as one component.

(1977) as adapted by Murray and Thompson (1977a). All tracers were sized before and/or after use in hybridization experiments. Each gel was calibrated with a pair of restriction digests or one digest plus a previously sized plant DNA standard. The most precise gel standard was found to be a Hae III digest of ϕ X-174 RFI DNA (Lee and Sinsheimer, 1974).

(vi) *Reassociation procedures*

Driver: tracer mass ratios were about 2000:1 and 9000:1 for ^{14}C - and ^3H -tracers, respectively. Appropriate mixtures in 10 mM NH_4 acetate were filtered through Metrical filters (Gelman GA-8), concentrated by a series of lyophilizations, and dissolved at 4.5–5.0 mg/ml in 600 mM PB reassociation buffer (600 mM PB, 0.1 mM EDTA) to a final concentration of 4.5–5.0 mg/ml. *E. coli* DNA (fragment length 250 NTP, specific activity 2.5×10^3 cpm $^{14}\text{C}/\mu\text{g}$ DNA) was included as an internal kinetic standard in some reassociation mixtures. Aliquots of 5–10 μl containing approximately 20–25 μg DNA, 2000 cpm ^{14}C and 10,000 cpm ^3H were sealed in disposable glass micropipettes for incubations covering equivalent C_0t values between 0.5 and 3×10^5 . For lower C_0t values, other aliquots were diluted 1:10 in 600 mM PB or 1:100 in 120 mM PB reassociation buffer and sealed into 100 μl micropipettes or 1 ml ampoules. Denaturation was for 2–4 minutes at 104–110°C for samples in 600 mM PB and 100° for 120 mM PB. Reassociations were carried out at 66° or 60°, respectively. Reactions were stopped by chilling in ice water followed by dilution into chilled 120 mM PB. Under these conditions, DNA fragments must have at least 75–80 per cent base sequence homology in order to hybridize.

Diluted samples were warmed briefly at 60°C and loaded onto 0.25 ml HAP columns pre-equilibrated in 120 mM PB at 60°C. Bound DNA was eluted with 400 mM PB, 60°C. The presence of tenaciously bound DNA was assayed by increasing the HAP column temperature to 98–100°C and washing with 120 mM PB. Column fractions were precipitated with cold 10 per cent TCA in the presence of carrier yeast RNA. Each C_0t curve was constructed from 18–21 time points plus an estimate of foldback sequence DNA. Foldback was assayed by fractionating the most dilute aliquots immediately after denaturation. Driver and homologous tracer reassociation kinetics were fitted with ideal second order components by a least squares procedure, using a computer program supplied by Dr E. Davidson (Pearson *et al.*, 1977) and adapted to the Hewlett Packard 2116C computer by Mr G. Ford. Heterologous tracer reactions were fitted by a similar program modified to take into account the reduced rate at which mismatched tracer heteroduplexes form, using the equation derived by Galau *et al.* (1976). Retardation coefficients for this equation were estimated on the basis of interspecific hybrid thermal stability measurements (Belford and Thompson, 1980) and the data of Bonner *et al.* (1973).

3. RESULTS

Hybridizations were performed as described under Methods with all eight Atriplex species and spinach using both *A. serenana* and *A. rosea* DNAs as drivers. Homologous tracers (single copy tracer from the same species as the

driver DNA) were included in all reactions. With few exceptions, homologous tracer kinetics were best fit with a single ideal second order component,* indicating that most tracers were essentially free from repeated sequence contamination. Table 1 summarizes these results. Tracer reactivity was usually about 80 per cent, and reactivities of individual tracer preparations were shown to be predictable and constant over at least 2-4 months. Similar tracer reactivities are commonly observed, especially for short fragments labelled *in vitro* (e.g., Kiper and Herzfeld, 1978; Zimmerman and Goldberg, 1977).

TABLE 1.

Single copy rate constants and genome sizes, together with the fractional and absolute amounts of single copy DNA in each genome. Calculations were made as described in the text from reassociation data for single copy tracers and unlabelled total DNA. Standard deviations are given for the measured values. Genome sizes and the absolute amounts of single copy DNA were calculated from average values of the measured parameters

	Single copy rate constant ($\times 10^{-3}$)	(N)	Haploid genome NTP $\times 10^{-8}$	Single copy fraction	(N)	Single copy per genome NTP $\times 10^{-8}$
<i>A. serenana</i>	1.71 ± 0.48	(5)	4.1	0.30 ± 0.03	(6)	1.2
<i>A. fruticulosa</i>	1.97 ± 0.17	(2)	3.5	0.40 ± 0.05	(2)	1.4
<i>A. rosea</i>	1.72 ± 0.22	(11)	4.0	0.37 ± 0.03	(7)	1.5
<i>A. sabulosa</i>	1.69 ± 0.62	(2)	4.1	0.39 ± 0.04	(2)	1.6
<i>A. phyllostegia</i>	1.49 ± 0.46	(3)	4.7	0.41 ± 0.08	(3)	1.9
<i>A. hortensis</i>	0.95 ± 0.08	(2)	7.3	0.28 ± 0.18	(2)	2.0
<i>A. truncata</i>	1.23 ± 0.26	(4)	5.6	0.38 ± 0.17	(4)	2.1
<i>A. triangularis</i>	1.00 ± 0.23	(2)	6.9	0.31 ± 0.05	(2)	2.1
Spinach	0.88 ± 0.14	(3)	7.9	0.26 ± 0.05	(3)	2.1

E. coli DNA included as an internal standard in each of these reactions served as a control for possible acceleration or retardation of renaturation rates sometimes observed with high concentrations of plant DNA (Murray and Thompson 1977b; Murray *et al.*, 1978). Based on the assumption that *E. coli* DNA fragments 400 NTP long, will reassociate under standard conditions with a rate of $0.25 \text{ M}^{-1} \text{ s}^{-1}$ (Britten, 1968), the rate expected for our *E. coli* tracer of 250 NT length was calculated to be $0.16 \text{ M}^{-1} \text{ s}^{-1}$ (Wetmur and Davidson, 1968). The experimentally observed average rate for the *E. coli* standard was 0.154 with a standard error of 0.002 in seven trials, each with DNA from a different Atriplex species. This value is in good agreement with the expected rate.

(i) Genome size and single copy complexity

The reaction rate for single copy sequences in total DNA is inversely proportional to the genome size (Britten and Kohne, 1968). We estimated Atriplex and spinach genome sizes by comparing their single copy rates with that for the *E. coli* DNA internal standard (fig. 3). Assuming the *E. coli*

* We attempted to fit each data set with two components. In all but two cases this procedure resulted in less than 3 per cent being included in the second (repeat) component, and in many cases a nonsense solution was generated for the second component.

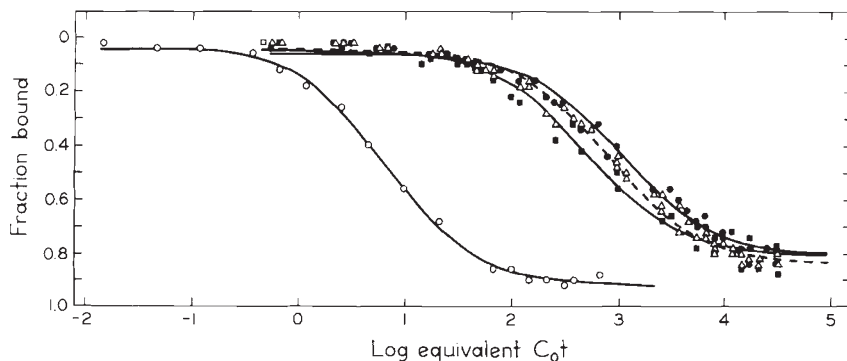


FIG. 3.—Kinetics of *E. coli* DNA reassociation (O) and several homologous Atriplex single copy tracer reactions. The reassociation of single copy tracers with homologous unlabelled DNA was carried out as described under Methods for *A. truncata* (Δ), *A. hortensis* (\bullet), and *A. fruticulosa* (\blacksquare). The curves are least-squares fits to pooled data from two separate reactions. The *E. coli* curve shown was obtained for tracer reassociated simultaneously with *A. serenana* DNA.

genome contains 4.5×10^6 NTP (Cairns, 1963) and using the rate constants in table 1, plant genome sizes were calculated from the relationship

$$G_{\text{plant}} = G_{E. coli} \times K_{E. coli} / K_{\text{plant}}$$

where G is the genome size in NTP and K is the reassociation rate constant. These estimates are listed in table 1.

The reassociation rates measured for purified single copy tracers were also used in conjunction with the data on total DNA reassociation to calculate the fraction of each genome composed of single copy sequences. For this purpose, the total DNA reassociation data from each species were fitted with two or three components by the same least squares procedure used in fitting the single copy tracer reaction, except that the rate of the slowest component was fixed at the rate observed for the corresponding purified single copy tracer (fig. 4). The result is an estimate of the fraction of the total DNA fragments reacting with single copy kinetics. It is necessary to fix the single copy rate used in the analysis when dealing with species like Atriplex in which the reaction of single copy sequences is not kinetically distinct from that of repeated sequences. If this rate is not fixed, the curve fitting program cannot accurately distinguish the reassociation of repeated DNA from that of single copy sequences and displays a tendency to overestimate both the reassociation rate and fraction of a single copy DNA. Estimates of the absolute amount of single copy DNA (single copy complexity) were obtained as the product of the values for single copy fraction and total genome size (table 1).

(ii) *Heterologous tracer reassociations*

When a single copy tracer is reacted with driver DNA from a divergent species, the hybrid duplexes are mismatched. The hybridization rate is reduced as a function of increasing base pair mismatch (Bonner *et al.*, 1973), a fact which can sometimes complicate the analysis of heterologous tracer reactivity. Not only will a heterologous tracer take longer to reach its

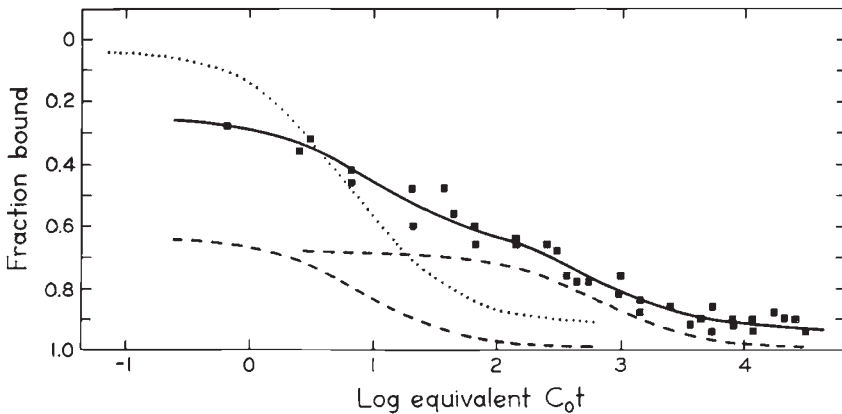


FIG. 4.—Reassociation kinetics of total DNA from *A. rosea*. The data have been fitted (solid line) by the least squares procedure described under Methods using two theoretical second order components (dashed line). The rate of the slow component has been fixed at the value measured for purified *A. rosea* single copy tracer reassociated simultaneously with the total *A. rosea* DNA. This component accounts for about 36 per cent of the total DNA reaction when foldback and unreactable sequence fractions are normalized out. The dotted curve is the fit for the internal standard *E. coli* DNA reassociated simultaneously with the Atriplex DNA.

maximum extent of reaction with driver sequences, but the more rapid driver self-reassociation will be removing sequences from the unreacted pool, leaving some potentially reactable tracer sequences unhybridized for lack of driver single strands with which to react. In principle, this problem can be minimized by isolating the unhybridized tracer at the end of the reaction and testing it for reaction with fresh driver DNA. However, such an approach is both technically difficult and time-consuming. Thus the method adopted in the present study was to fit the heterologous tracer data using the retarded rate equation developed by Galau *et al.*, (1976).^{*} The equation yields an estimate of the potential interspecific reactivity, *i.e.*, the fraction of heterologous tracer which would have hybridized with driver had the reaction proceeded at the same rate as the homologous reaction. The potential effect of the retardation factor on the calculation of Atriplex tracer cross reactivity is shown in fig. 5. Retardation coefficients in the range used in the present study increase apparent tracer reactivity by only 2-10 per cent over the values calculated assuming no retardation. Examples of heterologous tracer hybridization kinetics fitted with the retarded rate equation are shown in fig. 6. Table 3 presents the results of such fits for all the heterologous tracer reactions.

All but two heterologous reactions yielded good fits with a single component indicating that the heterologous tracers reacted primarily, if not exclusively, with single copy sequences in the driver DNA. Only *A. hortensis* and *A. triangularis* tracers reacting with *A. serenana* driver yielded kinetics best fit with a repeated sequence component. The *A. hortensis* reaction was fitted with a repeat component of 7.6 per cent while that of

^{*} The formula given by Galau *et al* (1976) contains a typographical error which can be corrected by substituting $(1 - V^{0.56})$ for V in the expression as given. See Belford and Thompson (1977).

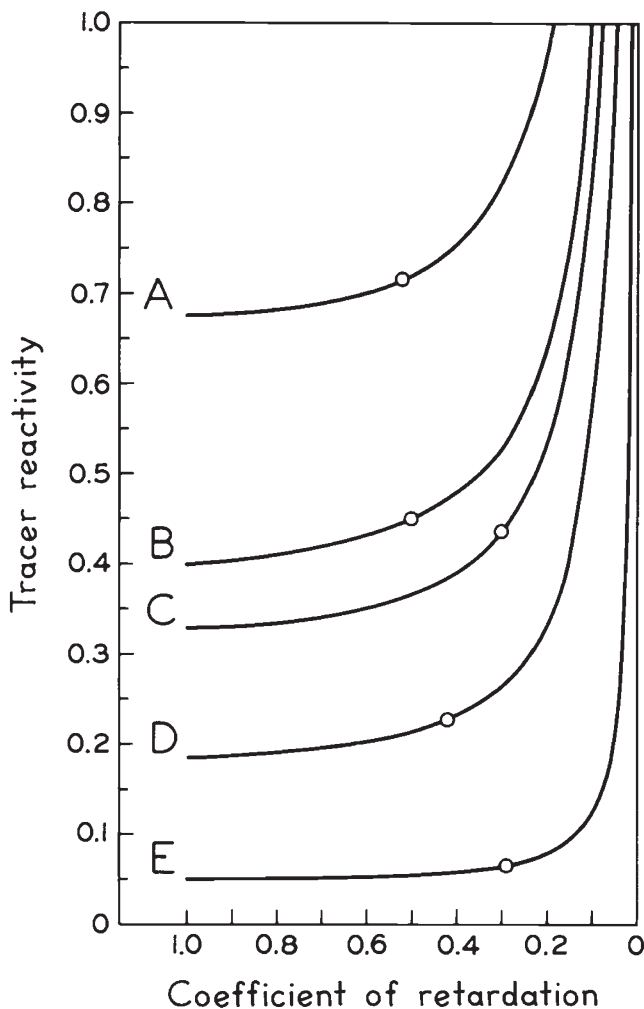


FIG. 5.—The effect of the coefficient of retardation on the estimation of theoretical heterologous tracer cross reactivity. Each curve represents the theoretical heterologous tracer cross reactivity of a heterologous single copy tracer as a function of (1) the experimentally observed cross reactivity and (2) the coefficient of retardation. Circles indicate the retardation coefficient chosen for each tracer on the basis of estimates for percent mismatch in the interspecific hybrids (Belford and Thompson, 1980) and the data of Bonner *et al.* (1973). The curves illustrated have been chosen to show the range of variation encountered in both cross reactivity and ΔT_E values. From top to bottom they represent the following reactions (source of unlabelled DNA in parenthesis): A. *fruticulosa* (A. *serenana*); A. *triangularis* (A. *rosea*); A. *fruticulosa* (A. *rosea*); A. *phyliostegia* (A. *rosea*); Spinach (A. *rosea*). Data have not been normalized for tracer reactivity with homologous unlabelled DNA.

A. triangularis fitted well with a 16 per cent repeat sequence component. These tracers were prepared from single copy sequences isolated after only one cycle of HAP fractionation; similarly prepared tracers were shown to contain repeat sequence contamination upon homologous reassociation (e.g., fig. 2). Thus the *A. triangularis* and *A. hortensis* tracers probably also

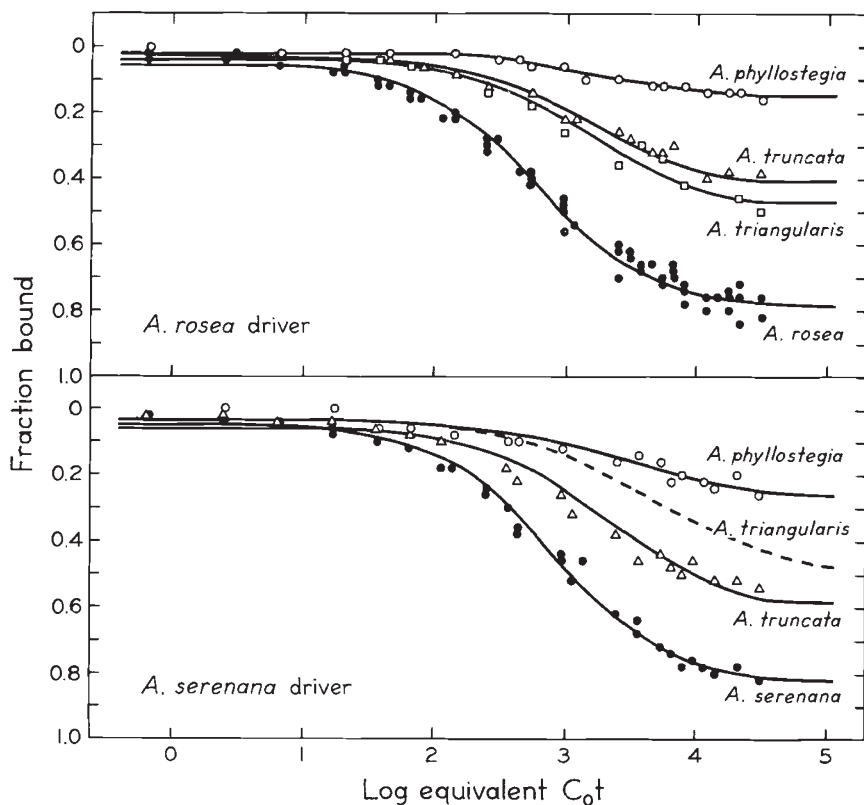


FIG. 6.—Reassociation kinetics of homologous and heterologous single copy tracers with excess total DNA from *A. rosea* (top panel) or *A. serenana* (bottom panel). Data were obtained and fitted as described in Methods, using the retarded rate equation to fit the data for heterologous reactions. The *A. rosea* tracer curve is the composite fit for pooled data from three separate reactions, while the *A. serenana* tracer curve is fit to data pooled from two reactions. Heterologous tracer data are derived from single experiments. The *A. triangularis* heterologous tracer curve in the lower panel is calculated from the theoretical single copy component of a two-component reaction.

contained residual repeated sequences able to hybridize with repeats in the *A. serenana* driver.

Percentage cross reactivity values will not always be equal in reciprocal experiments, since different species may have different amounts of single copy DNA. However, the *absolute* amount of cross-reactable heterologous tracer should be the same in reciprocal reactions. In table 2, cross reactivity percentages have been converted to absolute DNA amounts using the single copy complexity values listed in table 1. Most experimentally determined cross reactivity measurements are consistent with this prediction of reciprocity.

4. DISCUSSION

Our estimates of Atriplex haploid genome sizes range between 3.5 and 7.3×10^8 NTP (0.36 – 0.75 pg). The spinach haploid genome size was

TABLE 2

Summary of interspecific cross reactivities. Cross reactivities are given both as the fraction and as the amount (NTP) of DNA from each heterologous tracer which are able to hybridize with the indicated driver DNA sequences. The amount of DNA is calculated from the single copy complexities in table 1 using the average fractional cross reactivity where replicate values are available. In the first two columns of data, single copy tracers from species A and B have been reacted with total DNA of species A. The two columns on the right refer to reciprocal reactions, in which the same tracers have been driven by total DNA from species B. Cross reactivities have been normalized to 100 per cent tracer reactivity in the homologous reaction

Species A	Species B	Cross Reactivity			
		Driver A		Driver B	
		Fraction	10^8 NTP	Fraction	10^8 NTP
<i>A. rosea</i>	<i>A. sabulosa</i>	0.72	1.1	0.81	1.2
	<i>A. triangularis</i>	0.55	1.2	0.75	1.1
		0.59			
	<i>A. hortensis</i>	0.58	1.2	0.64	1.0
	<i>A. fruticulosa</i>	0.55	0.8	0.60	0.9
		0.56			
	<i>A. truncata</i>	0.53	1.1	0.62	0.9
	<i>A. serenana</i>	0.64	0.8	0.47	0.7
		0.56		0.45	
		0.63			
	<i>A. phyllostegia</i>	0.26	0.6	0.26	0.4
		0.33			
	Spinach	0.12	0.2	0.21	0.3
<i>A. serenana</i>	<i>A. fruticulosa</i>	0.96	1.3	—	—
		0.89			
	<i>A. sabulosa</i>	0.81	1.3	—	—
	<i>A. truncata</i>	0.72	1.5	0.93	1.2
	<i>A. triangularis</i>	0.56	1.4	—	—
	<i>A. hortensis</i>	0.46	0.9	—	—
	<i>A. rosea</i>	0.47	0.7	0.64	0.8
		0.45		0.56	
				0.63	
	<i>A. phyllostegia</i>	0.28	0.5	0.41	0.5
	Spinach	0.18	0.4	0.22	0.3

determined to be 7.9×10^8 NTP (0.81 pg), just larger than the *A. hortensis* genome. Such a two-fold variation within a genus is not unusual, especially among higher plants. For example, *Vicia* species cover a seven-fold range and species of *Lathyrus* show a three-fold variation (Bennett and Smith, 1976; Narayan and Rees, 1977). On the other hand, all the *Atriplex* genomes and spinach are small by comparison with many plant and animal genomes for which sizes have been measured previously (Bachmann *et al.*, 1972; Bennett and Smith, 1976). Taking the average *Atriplex* haploid genome size as roughly 5.4×10^8 NTP, it corresponds to about 12 per cent of the pea genome (Bennett and Smith, 1976; Murray *et al.*, 1978), or 20 per cent of an average mammalian genome (Bachmann *et al.*, 1972).

The single copy fraction of each genome, determined from the fraction of total DNA reassociating with the rate measured for the corresponding purified single copy tracer, included 26 per cent of the spinach genome and 28-40 per cent of the various *Atriplex* genomes. It is likely that our hydroxyapatite measurements somewhat underestimate the actual amount of single copy DNA since, in this assay, fragments containing both repeated

and single copy elements will appear to reassociate entirely with repetitive kinetics. The total leaf DNA preparations we used also contained chloroplast DNA, which will reassociate as though it were a repetitive sequence and thus reduce the apparent fraction of single copy sequences. In mung beans, which have a nuclear genome similar in size to those of the *Atriplex* species, chloroplast sequences were found to comprise about 10 per cent of a similar preparation of total DNA (Murray *et al.*, 1979). However the average *Atriplex* single copy fraction of 35 per cent is in agreement with the observation of Flavell *et al.* (1974) that angiosperm species with less than 2 pg of DNA per haploid genome average about 38 per cent single copy DNA when assayed by hydroxyapatite techniques similar to ours.

All such estimates of genome size and single copy DNA content are subject to errors in both experimental and analytical procedures. However we believe that inclusion of an internal standard of known complexity in each reaction, together with the calculation of single copy fractions from fits using fixed single copy rate constants determined with purified tracers, help to improve the precision and reproducibility of these measurements. The standard deviations for the rate constants listed in table 1 provide an indication of the reproducibility of our genome size estimates. In this connection, we also note that using our estimates of single copy DNA content to correct the cross reactivity data leads to a significant improvement in the degree of reciprocity (table 2). It seems unlikely that such an improvement would have been achieved if our estimates of genome size or single copy fraction were seriously in error.*

Both total and single copy DNA contents for the different *Atriplex* species show a tendency to cluster into two groups. *A. rosea*, *A. sabulosa*, *A. serenana*, and *A. fruticulosa* have the smallest genomes and lower contents of single copy DNA in spite of having a somewhat higher average fraction of single copy sequences. It is interesting to note that these four species are considered to be more derived or specialized on the basis of morphological and physiological criteria (cf. fig. 1), which is in agreement with several previous observations (e.g., Sparrow *et al.* 1972; Hinegardner, 1976) that increasing phenotypic specialization within related groups of organisms is often accompanied by a decline in DNA content.

(i) *Interspecific cross reactivity*

Perhaps the most striking feature of the cross reactivity data is the wide range of values obtained within a group of closely related species. Reactions between *Atriplex* species vary between about 30 per cent and 100 per cent. Although both low and high cross reactivity values have been previously observed for intrageneric comparisons, only in the case of *Plethodontid*

* This argument is strongest for *relative* rather than absolute values. Dr M. D. Bennett (personal communication) has estimated the total nuclear DNA content of three of our species by cytophotometry and obtained values about twice (1.6 to 2.4-fold) as large as those we report. We do not know the reason for this apparent discrepancy. Since all species used have 9 pairs of chromosomes, we do not believe tetraploidy to be a likely explanation. Although cytophotometric estimates for such small genomes would be subject to large relative errors from small absolute errors in the measurements, it is also possible that our kinetic data systematically underestimate absolute genome sizes for reasons not yet understood.

salamanders (Mizuno and Macgregor, 1974) and the plant genus *Lathyrus* (Narayan and Rees, 1977) do we find a range of different values comparable to that observed here. To our knowledge, the work of Narayan and Rees is the only previous report which systematically treats the question of intrageneric cross reactivity in higher plant single copy DNA. It will be of interest to discover if similar variations are found in other plant groups.

It appears that *A. triangularis* and *A. sabulosa* tracers contain sequences complementary to about 75-80 per cent of the *A. rosea* single copy fraction, while the corresponding value for *A. phyllostegia* is only about 25 per cent. *A. serenana*, *A. fruticulosa*, *A. truncata*, and *A. hortensis* show complementarity with 45-65 per cent of the *A. serenana* reference; *A. phyllostegia* contains sequences complementary to about 40 per cent of the *A. serenana* single copy tracer, while the *A. rosea* and *A. hortensis* genomes complement about 60-70 per cent.

Surprisingly, there are four species—*A. fruticulosa*, *A. sabulosa*, *A. truncata* and *A. triangularis*—which contain sequences complementary to ca. 100 per cent of the single copy DNA in the *A. serenana* genome (compare tables 1 and 2). High cross reactivity might be expected between *A. fruticulosa* and *A. serenana* since these species are interfertile (Nobs, 1976). Also a relatively high cross reactivity might be expected between *A. rosea* and *A. sabulosa* or *A. triangularis* because sexual crosses between these species produce a vigorous F_1 , albeit of reduced fertility (Nobs, 1976). However, the large amount of cross reaction of *A. sabulosa*, *A. truncata* or *A. triangularis* with *A. serenana* sequences is not forecast by sexual hybridization data and is not in accord with expectations based on phenotypic characters. Thus, while sexually hybridizable species do show high DNA cross reactivity, high cross reactivity alone does not guarantee that species are sexually compatible.

(ii) *The deletion hypothesis*

Both deletion events and random base substitutions can lead to decreases in cross reactivity between single copy DNAs in different species. Given sufficient time after genetic isolation, random base substitutions will accumulate to the point at which some sequences no longer have enough homology to form interspecific hybrids under normal *in vitro* reassociation conditions. This process is expected to be gradual, since base substitutions are thought to accumulate at relatively uniform average rates over very long periods (Wilson *et al.*, 1977). If the extent of substitution is randomly distributed over the entire single copy sequence population, there should be a correlation between the amount of cross reactable DNA and the degree of base substitution in the resulting hybrids. Thus if the variations in Atriplex cross reactivity were the result of random base substitution we would expect to find large differences in hybrid thermal stability paralleling the differences we observe in cross reactivity.

Thermal stability measurements for Atriplex interspecific duplexes (Belford and Thompson, 1980) do not exhibit the wide variation we see in the cross reactivity data, and fig. 7 shows that what variation there is, is not correlated with cross reactivity. Thus we do not believe that variations in the extent of sequence divergence by random base substitution have contri-

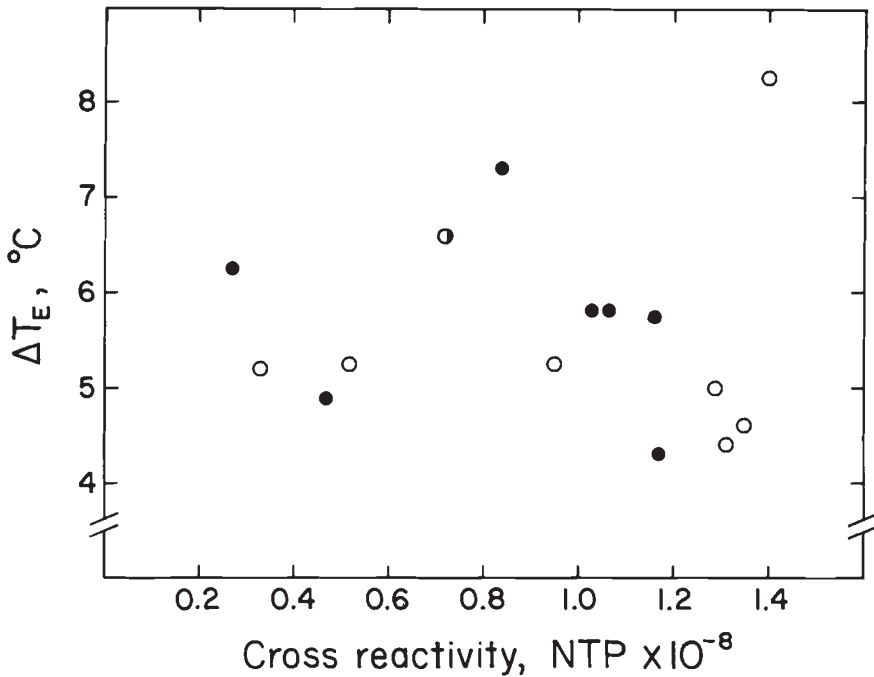


FIG. 7.—Scatter diagram to illustrate the absence of any clear correlation between cross reactivity and thermal stability of intraspecific duplexes. Open circles are for data obtained with *A. serenana* as the reference species, while closed circles represent data obtained with *A. rosea* as the reference. Values from replicate and reciprocal experiments have been averaged.

buted measurably to the variations in cross reactivity.* We therefore postulate that the variation in cross reactivity reflects varying amounts of deletion in the different lineages after their separation from a common ancestor. In contrast to base substitution, deletion events affect large blocks of DNA and may occur sporadically rather than gradually during evolution. Although there is no evidence that deletions occur preferentially in single copy regions, deletion of individual repetitive sequences would not greatly

* A more complicated version of the sequence divergence hypothesis would postulate that the non-crossreactive sequences are actually members of ancient or "fossil" repeat families (Thompson and Murray, 1980) which were already extensively diverged in the ancestral genome and have accumulated additional substitutions during the time since the *Atriplex* speciation events occurred. Such sequences would no longer reassociate as repeats, nor would they be expected to form interspecific hybrids. However, this hypothesis also predicts that cross reactivity values should parallel thermal stability values since the effect is still based on gradual accumulation of base substitutions.

It is also worth noting that amplification events may remove sequences from the single copy fraction. Since the amplified DNA would be removed from the tracers during purification, amplification of a given sequence in one species but not another would lead to the same results as deletion of the same sequence. However, most repeats are present on the order of 100 or more times per haploid genome. Since *Atriplex* genomes do not differ in size by more than 2-fold, and 1 per cent of the total DNA amplified 100 times would be sufficient to double the genome size, it is difficult to account for any significant differences in single copy cross reactivity on this basis.

affect the total interspecific cross reactivity of repetitive DNA, while deletion of a single copy sequence from one genome would completely prevent the reaction of its complement in another genome. Thus our experiments with single copy tracers should provide a particularly sensitive assay for deletion events. We visualize a large ancestral Atriplex genome giving rise to lineages which have since experienced extensive deletion and decreases in net DNA content. Random deletions would be expected to remove different sequences from the different genomes, so that the single copy DNA remaining in one genome would include sequences which were deleted from another genome and *vice versa*.

This model would account for the fact that percentage cross reactivity values are usually well below 100 per cent in both directions of reciprocal experiments. The one species pair showing close to 100 per cent reaction (*A. fruticulosa* and *A. serenana*) may be explained by assuming that most deletions occurred prior to separation of these two lineages. It is reasonable on other grounds to suppose that this separation took place somewhat later than most of the other major speciation events in this group (Belford and Thompson, 1980). If our view is correct, it follows that only relatively small amounts of single copy DNA are required to specify phenotypic characters common to all Atriplex species. The lowest cross reactivity we observe (about 5×10^7 NTP for reactions involving *A. phyllostegia*; table 3) sets an upper limit for the amount of DNA with such functions. This amount of DNA is less than half the amount in a *Drosophila* genome and only about 10 times more than in *E. coli*.

This view is consistent with an accumulating body of evidence that much of the typical eukaryotic genome is "secondary DNA" not required for coding and gene regulation (*e.g.*, Hinegardner, 1976; Thompson and Murray, 1980; Murray *et al.*, 1980). Such DNA may influence the phenotype in other ways, however, many of which do not depend on any particular sequence composition (*e.g.*, Bennett, 1972; Cavalier-Smith, 1978). Since one possible function of secondary DNA might be to increase genetic flexibility (*e.g.*, by providing a reservoir of variant sequences, facilitating chromosomal rearrangements, *etc.*) we may conjecture that its presence might be favoured by selective pressures promoting variation and rapid speciation (*e.g.*, Hinegardner, 1976). Adaptive evolution following periods of speciation may involve selection for decreases in secondary DNA since, as previously noted, increasing phenotypic specialization is often associated with decreasing DNA content. Under these circumstances, deletion events would be expected to play a prominent role in genome evolution, as we believe they have in Atriplex.

Acknowledgments.—We are greatly indebted to Glenn Ford for his skillful and cooperative assistance with the computer analysis. Sue Reed, Ann Kimber, and Steve Graff provided technical assistance at various times in the course of this work. Michael Murray, Richard Jorgensen, and Diana Stein contributed many critical comments and considerable encouragement.

5. REFERENCES

- ANGERER, R. C., DAVIDSON, E. H., AND BRITTEN, R. J. 1976. Single copy DNA and structural gene sequence relationships among four sea urchin species. *Chromosoma*, 56, 213-226.

- BACHMANN, K., GOIN, O. B., AND GOIN, C. J. 1972. Nuclear DNA amounts in vertebrates. *Brookhaven Symp. in Biol.*, 23, 419-450.
- BELFORD, H. S., AND THOMPSON, W. F. 1977. Single copy DNA sequence comparisons in *Atriplex*. *Carnegie Inst. Wash. Year Book*, 76, 246-252.
- BELFORD, H. S., AND THOMPSON, W. F. 1980. Single copy DNA homologies in *Atriplex*. II. Hybrid thermal stabilities and molecular phylogeny. *Heredity*, 46,
- BENNETT, M. D. 1972. Nuclear DNA and minimum generation time in herbaceous plants. *Proc. Royal Society London, B*, 181, 109-135.
- BENNETT, M. D., AND SMITH, J. B. 1976. Nuclear DNA amounts in angiosperms. *Philos. Trans. R. Soc. London, Ser. B*, 274, 227-273.
- BJÖRKMAN, O., TROUGHTON, J., AND NOBS, M. A. 1973. Photosynthesis in relation to leaf structure. *Brookhaven Symp. in Biology*, 25, 206-226.
- BONNER, T. I., BRENNER, D. J., NEUFELD, B. R., AND BRITTEN, R. J. 1973. Reduction in the rate of DNA reassociation by sequence divergence. *J. Mol. Biol.*, 81, 122-135.
- BRITTEN, R. J. 1968. The arithmetic of nucleic acid reassociation. *Carnegie Inst. Wash. Yearbook*, 67, 332-335.
- BRITTEN, R. J., GRAHAM, D. E., AND NEUFELD, B. R. 1974. Analysis of repeating DNA sequences by reassociation. *Methods in Enzymology*, 29, 363-418.
- BRITTEN, R. J., AND KOHNE, D. E. 1968. Repeated sequences in DNA. *Science*, 161, 529-540.
- CAIRNS, J. 1963. The chromosome of *E. coli*. *Cold Spr. Harb. Quant. Biol.*, 28, 43-46.
- CAVALIER-SMITH, T. 1978. Nuclear volume control by nucleoskeletal DNA, selection for cell volume and cell growth rate, and the solution of the C-value paradox. *J. Cell Sci.*, 34, 247-278.
- FLAVELL, R. B., BENNETT, M. D., SMITH, J. B., AND SMITH, D. B. 1974. Genome size and the proportion of repeated nucleotide sequence DNA in plants. *Biochem. Genetics*, 12, 257-269.
- GALAU, G. A., CHAMBERLIN, M. E., HOUGH, B. R., BRITTEN, R. J., AND DAVIDSON, E. H. 1976. Evolution of repetitive and nonrepetitive DNA. In: *Molecular Evolution*, ed. Ayala, F. J., pp. 200-224. Sinauer Assoc., Inc. Sunderland, Mass.
- HALL, H. M., AND CLEMENTS, F. E. 1923. The genus *Atriplex*. In *The North American Species of Artemisia, Crysothamnus, and Atriplex*. *Carnegie Inst. of Wash. Publ.* 326, pp. 235-347. Judd and Detweiler, Inc., Washington, D.C.
- HARPOLD, M. M., AND CRAIG, S. P. 1978. The evolution of nonrepetitive DNA in sea urchins. *Differentiation*, 10, 7-11.
- HINEGARDNER, R. 1976. Evolution of genome size. In *Molecular Evolution*. Ed. Ayala, F. J., pp. 179-199. Sinauer Assoc., Inc., Sunderland, Mass.
- KIPER, M., AND HERZFELD, F. 1978. DNA sequence organization in the genome of *Petroselinum sativum* (Umbelliferae). *Chromosoma*, 65, 335-351.
- KOHNE, D. E. 1970. Evolution of higher organism DNA. *Quart. Rev. Biophysics*, 33, 327-375.
- LAIRD, C., MCCONAUGHY, B. L., AND MCCARTHY, B. J. 1969. Rate of fixation of nucleotide substitution in evolution. *Nature*, 224, 149-154.
- LEE, A. S., AND SINSHEIMER, R. L. 1974. A. Cleavage map of bacteriophage ϕ X174 genome. *Proc. Nat. Acad. Sci. U.S.A.*, 71, 2882-2886.
- MCDONNELL, M. W., SIMON, M. N. AND STUDIER, E. W. 1977. Analysis of restriction fragments of T7 DNA and determination of molecular weight by electrophoresis in neutral and alkaline gels. *J. Mol. Biol.*, 110, 119-146.
- MIZUNO, S., AND MacGREGOR, H. C. 1974. Chromosomes, DNA sequences, and evolution in salamanders of the genus *Plethodon*. *Chromosoma*, 48, 239-296.
- MURRAY, J. G., BELFORD, H. S., AND THOMPSON, W. F. 1977. *In vitro* labelling of single-stranded DNA. *Carnegie Inst. Wash. Yearbook*, 76, 262-267.
- MURRAY, M. G., CUELLAR, R. E., AND THOMPSON, W. F. 1978. DNA sequence organization in the pea genome. *Biochemistry*, 17, 5781-5790.
- MURRAY, M. G., PALMER, J. D., CUELLER, R. E. AND THOMPSON, W. F. 1979. Deoxyribonucleic acid sequence organization in the mung bean genome. *Biochemistry*, 18, 5259-5266.
- MURRAY, M. G., PETERS, D. L., AND THOMPSON, W. F. 1981. Ancient repeated sequences in the pea and mung bean genomes and a model for genome evolution. *J. Mol. Evol.* (in press).
- MURRAY, M. G., AND THOMPSON, W. F. 1977a. Contaminants affecting plant DNA reassociation. *Carnegie Inst. Wash. Yearbook*, 76, 255-259.

- MURRAY, M. G., AND THOMPSON, W. F. 1977b. Single-stranded DNA fragment length determination by alkaline agarose electrophoresis: Applications for DNA renaturation studies. *Carnegie Inst. Wash. Yearbook*, 76, 259-267.
- NARAYAN, R. K. J., AND REES, H. 1977. Nuclear DNA divergence among *Lathyrus* species. *Chromosoma*, 63, 101-107.
- NOBS, M. A. 1976. Hybridization in *Atriplex*. *Carnegie Inst. Wash. Yearbook*, 75, 421-423.
- PEARSON, W. R., DAVIDSON, E. H., AND BRITTEN, R. J. 1977. A program for least squares analysis of reassociation and hybridization data. *Nucleic Acids Research*, 4, 1727-1735.
- RICE, N. 1972. Changes in repeated DNA in evolution. *Brookhaven Symp. Biol.*, 23, 44-78.
- SHIELDS, G. F., AND STRAUS, N. A. 1975. DNA-DNA hybridization studies of birds. *Evolution*, 29, 159-166.
- SIBLEY, C. G., AND AHLQUIST, J. E. 1980. The relationships of the "primitive insect eaters" (Aves: Passeriformes) as indicated by DNA \times DNA hybridization. *Proc. 17th International Ornithological Congress* (in press).
- SOHN, U.-I., ROTHFELS, K. H., AND STRAUS, N. A. 1975. DNA:DNA hybridization studies in black flies. *J. Mol. Evol.*, 5, 75-85.
- SPARROW, A. H., PRICE, H. J., AND UNDERBRINK, A. G. 1972. A survey of DNA content per cell and per chromosome of prokaryotic and eukaryotic organisms: Some evolutionary considerations. *Brookhaven Symp. Biol.*, 23, 451-494.
- STEIN, D. B., THOMPSON, W. F., AND BELFORD, H. S. 1979. Studies on DNA sequences in the Osmundaceae. *J. Mol. Evol.*, 13, 215-232.
- THOMPSON, W. F., AND MURRAY, M. G. 1980. Sequence organization in pea and mung bean DNA and a model for genome evolution. In *The Plant Genome* (Fourth John Innes Symposium), eds. Davies, D. R. and Hopwood, D. A. John Innes Charity, Norwich, England.
- WETMUR, J. G., AND DAVIDSON, N. 1968. Kinetics of renaturation of DNA. *J. Mol. Biol.*, 31, 349-370.
- WILSON, A. C., CARLSON, S. S., AND WHITE, T. J. 1977. Biochemical evolution. *Ann. Rev. Biochem.*, 46, 573-639.
- ZIMMERMAN, J. L., AND GOLDBERG, R. B. 1977. DNA sequence organization in the genome of *Nicotiana tabacum*. *Chromosoma*, 59, 227-252.