Genet. Res., Camb. (1989), **53**, *pp.* 173–182 *With* 1 *text-figure Printed in Great Britain*

Genome size variation in the North American sunfish genus *Lepomis* (Pisces: Centrarchidae)

CHARA J. RAGLAND* AND JOHN R. GOLD†
Department of Wildlife and Fisheries Sciences, Texas A & M University, College Station, Texas 77843 U.S.A.

(Received 17 August 1988 and in revised form 22 November 1988)

Summary

Genome sizes (nuclear DNA contents) were documented spectrophotometrically from individuals of each of nine species of the North American centrarchid (sunfish) genus *Lepomis*. The distributions of DNA values within and among the nine species were essentially normal and continuous, suggesting that changes in DNA quantity in *Lepomis* are small in amount, involve both gains and losses of DNA, and are cumulative and independent in effect. Significant differences in mean genome size were found between individuals within populations in all nine species and between species. Nested analysis of variance and comparisons of average genome size difference or distance between individuals drawn from different levels of taxonomic organization revealed that the majority of genome size divergence in *Lepomis* occurs above the hierarchical level of individuals within populations. The *Lepomis* data when compared to similar data from North American cyprinid fishes appear to suggest that: (i) genome size evolution in these fishes at least follows a continuous rather than a discontinuous mode; (ii) the general predictions of hypothetical models relating genome size variation as a function of organismal position along adaptive continua may be oversimplified, or not applicable to complex, higher eukaryotes; and (iii) changes in genome size in these fishes may be concentrated in speciation episodes.

1. Introduction

Quantitative variation in nuclear DNA content or genome size has remained an enigma in eukaryotic evolution since the seminal paper by Mirsky & Ris (1951). Interspecific variation in genome size among eukaryotes ranges over several orders of magnitude, and comparatively large DNA content differences are often found even between closely related species (Hinegardner & Rosen, 1972; Bachmann *et al.* 1972; Sparrow *et al.* 1972). Early suggestions to explain the observed genome size differences were that the variation was related directly to organismal complexity and/or the number of genes within a given species (Kauffman, 1971). It is now apparent, however, that among eukaryotes genome size is independent of organismic and/or genetic complexity (Cavalier-Smith, 1985*a*; Price, 1988). Other suggestions offered to explain interspecific genome size variation among eukaryotes include (i) natural selection acting on nucleotypic parameters presumably affected by genome size, (ii) natural selection acting on ecological parameters correlated with genome size, (iii) the proliferation of 'selfish' DNAs which presumably are constrained by selection, and (iv) an inverse relationship between genome size and the amount of genetic variation in structural genes (Bennett, 1971, 1972; Cavalier-Smith, 1980, 1985*a*; Doolittle & Sapienza, 1980; Orgel & Crick, 1980; Pierce & Mitton, 1980; Szarski, 1983). Because of the large differences in DNA content between species, it has also been suggested that changes in genome size may be associated with speciation episodes (Hinegardner, 1976; Morescalchi, 1977; Cavalier-Smith, 1978). Inherent to most of these hypotheses is the assumption that genome size is constant within a species. As a result, almost all of the quantitative data on eukaryotic genome size variation are from fully differentiated species and/or higher taxa. Studies examining variation at lower taxonomic levels, i.e. between individuals within populations or between populations within species, are few.

Recently, several studies in both plants (Price *et al.* 1981) and animals (Sherwood & Patton, 1982; Gold & Price, 1985; Gold & Amemiya, 1987; Johnson *et al.* 1987) have focused on genome size variation within

* Present address: Department of Biochemistry and Biophysics, Texas A & M University, College Station, Texas 77843 U.S.A.
† Corresponding author.

and between populations or closely related species. Gold & Price (1985) and Gold & Amemiya (1987) studied genome size variation in North American cyprinid fishes and found differences between individuals within populations of species to average 6–7 % of the total genome. Johnson *et al.* (1987) studied genome size variation in salmonid fishes and reported mean levels of variation of 4·6 % between individuals within populations. In mammals of the genus *Thomomys*, Sherwood & Patton (1982) observed intraspecific genome size variation as high as 35 %. Finally, Price *et al.* (1980, 1981 *a*, *b*, 1986) have found significant levels of intraspecific variation in genome size in at least two species of the plant genus *Microseris*, some of which is in excess of 20 %. Taken together, these studies would appear to indicate that (i) substantial variation in genome size within species occurs and is neither well understood nor adequately characterized, (ii) different patterns of genome size variation may occur in different groups, and (iii) the assumption of the constancy of genome size within species may be invalid. As noted by Sherwood & Patton (1982), it will be difficult to evaluate the meaning of genome size differences between species without a thorough understanding of genome size variation within species.

In this paper, we document genome size variation within and between species of the North American centrarchid fish genus *Lepomis*. The initial purpose of the study was to test the generality of the model of genome size evolution developed by Gold and Amemiya (1987) for cyprinid fishes. Briefly, Gold & Amemiya (1987) hypothesized that changes in DNA quantity within populations of cyprinids were small in amount, involved both gains and losses of DNA, and were cumulative and independent in effect. This hypothesis was based on the observation that genome size distributions within cyprinid populations were essentially normal and continuous, and on the assumption that genome size variation within populations followed the premises of the normal probability density function. Gold & Amemiya (1987) also hypothesized that genome size changes in cyprinids were concentrated in speciation episodes. The latter was based on the finding that the majority of genome size variation among the species examined was distributed above the hierarchical level of individuals within populations of species. Centrarchids are an appropriate system in which to test Gold & Amemiya's (1987) hypotheses in that centrarchids differ trenchantly from cyprinids in evolutionary history and several ecological and life history parameters (Lee *et al.* 1980). Because of the latter, it was anticipated that a comparison of genome sizes and their patterns between the two groups might also allow considerations or tests of some of the adaptive hypotheses relating genome size variation and natural selection. Finally, centrarchids are considerably less speciose than cyprinids, and for reasons discussed elsewhere in this paper provide a means to test whether changes in genome size are concentrated in speciation episodes.

2. Materials and methods

All fish were collected either by seining or by electroshocking. The species (collection localities) were as follows: *Lepomis auritus* (Lake Eustis, Lake Co., FL); *Lepomis cyanellus*, *Lepomis gulosus*, *Lepomis macrochirus*, *Lepomis microlophus*, and *Lepomis megalotis* (Camp Creek Lake, Robertson Co., TX); *Lepomis humilis* (Texas A & M University Aquaculture Research Center, Brazos Co., TX); and *Lepomis marginatus* and *Lepomis punctatus* (Lake Conroe, Montgomery Co., TX). The *L. auritus* specimens (collected in Florida) were processed in the field; all other specimens were returned to College Station and processed in our laboratory. Voucher specimens will ultimately be deposited in the Texas Cooperative Wildlife Collections at Texas A & M University.

Relative genome sizes of five individuals from each of the nine species were determined by scanning microdensitometry of Feulgen stained erythrocyte nuclei. Blood was taken from live fish by direct cardiac puncture and smeared on one end of a glass microscope slide. Chicken blood, obtained from a highly inbred, pathogen-free strain was smeared on the opposite end of each slide and served as an internal control. Slide preparation, fixation, and staining followed Fand (1970) and Gold & Price (1985). Schiff's reagent was prepared after Humason (1979). An acid hydrolysis time of thirty minutes was determined empirically from optimal absorbency curves of two different centrarchid species. All stained slides were coded by number, randomized, and stored in the dark at 4 °C until analysed.

Measurements of nuclear DNA content were taken using a Zeiss Universal-II scanning microdensitometer. Microspectrophotometry was carried out at a scanning speed of 8/64, using a 100 × planachromat oil immersion objective at 560 nm light. Fifteen fish nuclei and ten chicken nuclei were measured from each of two slides per individual (= 30 nuclei per individual). Choice of scanned nuclei followed the recommendations of Gold & Price (1985).

Absorbency values of individual fish nuclei from each slide were standardized as a percentage of the mean absorbency value of chicken erythrocyte nuclei on that slide and coded by multiplying the percent of chicken standard by twenty. The latter was the mean absorbency value of chicken nuclei over all slides. Coding of data in this fashion does not distort subsequent statistical analysis (Sokal & Rohlf, 1969) and was used to ease data handling. For conversion to picograms (pg) DNA, the coded data are reconverted to standardized data (percent of chicken standard) and then multiplied by 2·5, the generally accepted DNA value of diploid chicken erythrocyte nuclei (Rasch *et al.* 1971). Statistical analyses of the data as

described below were carried out using either SAS or our own programs on the Texas A & M mainframe computer.

3. Results

The coded absorbency data were initially subjected to descriptive statistical analysis which included generation of sample means, variances, ranges, and coefficients of variation (CVs). The descriptive statistics from the distributions of DNA measurements (nuclei) and of DNA values for all individuals over all populations (species) are shown in Tables 1 and 2, respectively. Note that the standard errors are higher, variances lower, and ranges smaller in Table 2 as compared to Table 1. This is to be expected because of the differences in sample size, and also because the distribution of individuals is essentially a distribution of means.

Genome sizes of the nine *Lepomis* species ranged from 1·88–2·16 pg of DNA. Ohno & Atkin (1966) previously reported the genome size of one individual of *L. cyanellus* to be 2·16 pg of DNA, a value which is slightly higher than the range of genome sizes of *L. cyanellus* individuals found in this study (Table 2). The differences are probably insignificant, however, given that different acid hydrolysis times were used in the two studies. Overall, the genome sizes of the nine *Lepomis* species are essentially the same as the 'average' teleost fish genome size of 2·00 pg of DNA as reported by Hinegardner & Rosen (1972).

Homogeneity of variances of both distributions was tested by Bartlett's method (Sokal & Rohlf, 1969). Significant heterogeneity ($\chi^2_{[8]} = 68{\cdot}27$, $P < 0{\cdot}05$) was found among variances from the distribution of measurements (nuclei), but not from the distribution of DNA values of individuals ($\chi^2_{[8]} = 9{\cdot}95$, $P > 0{\cdot}05$). Coefficients of variation (CVs) based on the distribution of measurements (nuclei) were generated for each slide, individual, and population (species). The mean CVs (±S.E.) at each of these three levels were 3·21±0·76 ($n = 90$), 3·34±0·59 ($n = 45$), and 4·03±0·05 ($n = 9$), respectively. This means that at an α probability level of 0·05 there is a 90% chance of detecting a 2–3% difference between mean genome sizes at each of the three levels (Sokal & Rohlf, 1969; Gold *et al.* 1975).

The coded absorbency data were organized into

Table 1. *Descriptive statistics from the distribution of measurements (nuclei)*

Taxon	*n*	Mean ± S.E.	Variance[a]	Range
Lepomis				
auritus	150	2·05±0·01	0·006	1·90–2·23
cyanellus	150	1·96±0·01	0·010	1·73–2·22
gulosus	150	2·00±0·01	0·006	1·84–2·22
humilis	150	2·01±0·01	0·009	1·81–2·20
macrochirus	150	1·88±0·01	0·005	1·70–2·05
marginatus	150	2·16±0·01	0·007	1·97–2·35
megalotis	150	2·12±0·01	0·007	1·91–2·31
microlophus	150	1·97±0·01	0·008	1·76–2·20
punctatus	150	1·92±0·01	0·003	1·77–2·03

[a] Variances heteroscedastic at $\alpha = 0{\cdot}05$.
Data are in picograms of DNA.

Table 2. *Descriptive statistics from the distribution of DNA values of individuals*

Taxon	*n*	Mean ± S.E.	Variance[a]	Range
Lepomis				
auritus	5	2·05±0·02	0·002	2·01–2·12
cyanellus	5	1·96±0·04	0·007	1·86–2·08
gulosus	5	2·00±0·01	0·001	1·96–2·04
humilis	5	2·01±0·03	0·005	1·94–2·12
macrochirus	5	1·88±0·01	0·001	1·86–1·92
marginatus	5	2·16±0·03	0·004	2·07–2·23
megalotis	5	2·12±0·02	0·002	2·08–2·18
microlophus	5	1·97±0·02	0·002	1·92–2·03
punctatus	5	1·92±0·01	0·001	1·89–1·95

[a] Variances homoscedastic at $\alpha = 0{\cdot}05$.
Data are in picograms of DNA.

Table 3. *Distribution normality statistics*

Distribution	n	Skewness (g_1)	Kurtosis (g_2)
Measurements (nuclei) among species	1350	0·259[a]	−0·400[a]
Individuals among species	45	0·380	−0·639
Measurements (nuclei) within species			
Lepomis			
auritus	150	−0·015	−0·589
cyanellus	150	0·247	−0·140
gulosus	150	0·225	−0·220
humilis	150	0·066	−0·762
macrochirus	150	−0·012	−0·514
marginatus	150	0·156	−0·594
megalotis	150	0·109	−0·639
microlophus	150	0·276	0·046
punctatus	150	−0·100	−0·577
Individuals within populations[b]	45	0·181	−1·156

[a] Significance at $\alpha = 0{\cdot}05$. Positive g_1 values indicate skewness towards higher values; negative g_2 values indicate platykurtosis.
[b] Rankit distribution (*cf* text).

twelve different sampling distributions which were tested using the g_1 and g_2 indices of distribution normality (Sokal & Rohlf, 1969). The distributions tested included (i) all measurements (nuclei) over all nine samples ($n = 1350$), (ii) all DNA values of individuals over all samples ($n = 45$), (iii) all measurements (nuclei) within each sample (nine sampling distributions, $n = 150$ each), and (iv) a rankit distribution reflecting the distribution of individuals within samples summed over all samples ($n = 45$). The last was generated according to equation (1) in Gold & Amemiya (1987) in order to remove scaling effects due to individuals being drawn from different species (Sokal & Rohlf, 1969), and reflects the distribution of DNA values of individuals within populations of species. As shown in Table 3, only the distribution of all measurements (nuclei) over all samples was significantly non-normal; the deviations from normality, however, appear to be slight (Fig. 1). All other distributions, including both nuclei (measurements) and individuals within populations of species (i.e. the rankit distribution) were normal. These findings indicate that genome size variation is essentially normally distributed within and among the nine *Lepomis* species.

Single classification analysis of variance (ANOVA) was used to test for significant heterogeneity in genome size variation among species and among individuals within samples (populations). The distribution of DNA values of individuals was used for the test of heterogeneity among species in part for reasons outlined in Gold & Price (1985), and in part because the distribution was normal (Table 3) and the variances were homoscedastic. Significant heterogeneity ($F_{[8,9]} = 7{\cdot}21$, $P < 0{\cdot}05$) in mean genome size was detected among species, and the results of a Duncan's multiple range test are shown in Table 4. All nine species means appear to form a continuous, overlapping series of groupings, and given the range of DNA values of individuals within populations (species), it is clear that no single species can be unequivocally differentiated solely on the basis of genome size.

Separate single classification ANOVA's were used to test for heterogeneity of DNA values of individuals within each population (species) using the distribution of measurements (nuclei) of that population. Note that all of these distributions were normal (Table 3). All F-tests were significant at $\alpha = 0{\cdot}05$, and a synopsis of the results of Duncan's multiple range tests on each population is shown in Table 5. On the average, more than half ($\bar{X} = 2{\cdot}88$) of the five individuals sampled from each population differed significantly in genome size.

The maximum variation in genome size between individuals within populations ranged from 3·18% in *L. machrochirus* to 12·18% in *L. cyanellus* and averaged $6{\cdot}24 \pm 0{\cdot}46$%. Assuming an average *Lepomis* genome size of 2·10 pg DNA, this represents approxi-

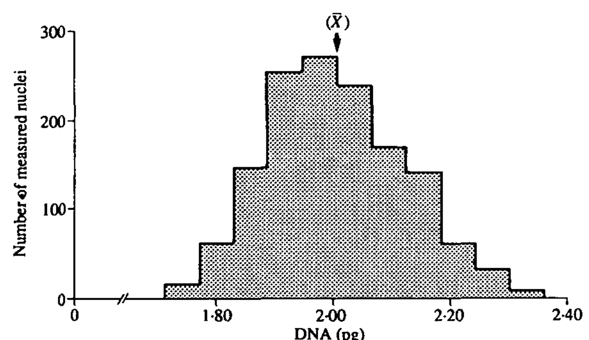

Fig. 1. Frequency distribution of DNA measurements (nuclei) over all nine *Lepomis* species. $\bar{X} = 2{\cdot}01$; S.D. = 0·07; $n = 1350$.

Table 4. *Results of Duncan's multiple range test of the distribution of DNA values of individuals*

Species	DNA values[a]
Lepomis	
marginatus	2·16a
megalotis	2·12ab
auritus	2·05bc
humilis	2·01cd
gulosus	2·00cd
microlophus	1·97de
cyanellus	1·96de
punctatus	1·92ef
macrochirus	1·88f

[a] Mean DNA values of species with the same letter (Duncan's test grouping) are not significantly different at $\alpha = 0{\cdot}05$. DNA values are in picograms of DNA.

Table 5. *Results of single classification analyses of variance and Duncan's multiple range tests for heterogeneity of DNA values of individuals within populations*

Species	*F* value[a]	No. of significantly different groups[b]
Lepomis		
auritus	10·46	2
cyanellus	23·63	4
gulosus	2·98	2
humilis	15·42	4
macrochirus	5·40	2
marginatus	12·92	4
megalotis	7·01	2
microlophus	5·20	3
punctatus	2·89	3
		$\bar{X} = 2{\cdot}88$

[a] All *F* values are significant at $\alpha = 0{\cdot}05$. Degrees of freedom are (4, 5) for all populations.
[b] Refers to groups of significantly different means using Duncan's multiple range test at $\alpha = 0{\cdot}05$.

mately 0·13 pg or about $1{\cdot}3 \times 10^8$ base pairs (bp) DNA. Interestingly, this maximum quantity of DNA which varies between individuals within *Lepomis* populations is essentially the same as the average DNA quantity difference between *Lepomis* species. The latter was estimated as 5·6% (roughly 0·12 pg DNA) and was calculated as the average percent difference between species means using all 36 possible pairwise comparisons. Of interest also is that $1{\cdot}3 \times 10^8$ bp DNA is nearly as large as the total quantity of DNA needed for the structural gene component of the *Lepomis* genome if one assumes there are 50000 structural genes and 1500 coding base pairs per structural gene.

A nested analysis of variance was carried out to determine the significance and proportion of genome size variation at each experimental level from between slides within individuals to among species. The results are shown in Table 6 and demonstrate that significant heterogeneity existed at each experimental level. The proportion of the total variation explained at each level, however, differed markedly. Over half (*ca.* 51%) of the variance appears to be due to variation among species, while only *ca.* 17% is attributed to differences among individuals within populations of species. To examine this further, genome size divergence at the two levels (i.e. among individuals within populations versus among species) was evaluated by computing genome size difference or distance values between and within each of the nine *Lepomis* samples. We used equations (2) and (3) of Gold & Amemiya (1987) which estimate genome size distance (GSD_{min}) values at each level. Briefly, Equation (2) generates a GSD_{min} value between two taxa (species) which represents the average of all pairwise differences in genome size between all individuals in each taxon or species (i.e. with $n = 5$ individuals for each of two taxa, there are 25 possible pairwise comparisons). Equation (3) generates a GSD_{min} value among individuals within populations and represents the average of all possible pairwise differences in genome size between all individuals of any one population averaged over all populations. A point to note is that both GSD_{min} values are minimum linear distance estimates and as such can underestimate the true distance if reversed or reticulated patterns of change occur. Alternatively, such linear distances are the only metrics suitable for non-rooted quantitative data such as genome sizes (Sneath & Sokal, 1973).

The average genome size differences (distances) between individuals drawn from the two levels of evolutionary divergence are shown in Table 7. As

Table 6. *Nested analysis of variance*

Variance source	D.F.	M.S.	*F*	Variance component	%
Total	1349	0·89	—	0·94	100·00
Species	8	77·10	14·53[a]	0·48	50·83
Individuals	36	5·31	8·73[a]	0·16	16·63
Slides	45	0·61	2·13[a]	0·02	2·29
Error	1260	0·28	—	0·28	30·24

[a] Significance at $\alpha = 0{\cdot}05$.

Table 7. *Average genome size difference (distance) between individuals at two levels of evolutionary divergence*

Level	Mean genome size difference ± S.E.[a]	No. of pair-wise comparisons
Individuals within populations	0·476 ± 0·192	90
Species within the genus	0·943 ± 0·079	36

[a] In coded DNA values. Picograms are estimated by dividing coded values by eight.

indicated, individuals drawn at random from two different *Lepomis* species will differ on the average by 0·943 GSD_{min} units (approximately 0·118 pg DNA), whereas any two individuals drawn at random from a population of the same *Lepomis* species will differ on the average by only 0·476 GSD_{min} units (approximately 0·059 pg). These data indicate that a sizable fraction of genome size divergence occurs above the level of individuals within populations of species and corroborates the results of the nested analysis of variance.

4. Discussion

The initial purpose of this study was to examine the applicability of the model of genome size evolution developed by Gold & Amemiya (1987) for cyprinid fishes. Briefly, Gold & Amemiya (1987) studied genome size variation within and among 20 species of North American cyprinid fish (including 12 species from the highly speciose cyprinid genus *Notropis*) using essentially the same methodology as employed here. They found that the distributions of DNA values of individuals within populations of the 20 species were essentially continuous and normal, and that the distribution of DNA values among the 20 species was continuous and overlapping. They suggested, based on the assumption that genome size variation followed the premises of the normal probability density function (Sokal & Rohlf, 1969), that changes in DNA quantity in cyprinids were small in amount, involved both gains and losses of DNA, and were cumulative and independent in effect. We chose to employ the centrarchid genus *Lepomis* as a comparison group since centrarchids differ from cyprinids in a number of parameters including evolutionary history, taxonomy, and several life history features (Lee *et al.* 1980 and references therein).

The patterns of genome size variation both within and between the nine *Lepomis* species are identical to those observed in cyprinids. Genome sizes are normally distributed within species (both in terms of measurements (nuclei) and DNA values of individuals within populations), and continuously distributed across species. As noted by Gold & Amemiya (1987), this pattern of continuous genome size change appears to typify many animal groups (Bachmann *et al.* 1972, 1985; Bianchi *et al.* 1983; Johnson *et al.* 1987) and differs markedly from the discontinuously distributed species DNA values found in many plant groups (Narayan, 1983; Cavalier-Smith, 1985*b*; Sims & Price, 1985; Raina *et al.* 1986; Labani & Elkington, 1987). In the latter, many species or species-groups are differentiated by quantum differences in genome size which is not attributable to polyploidy. As pointed out by Narayan (1982, 1983) and Cavalier-Smith (1985*b*), this difference in pattern could be due to either (i) a difference in the mechanism(s) by which DNA sequences are gained or lost, (ii) a difference in the rates of DNA quantity change, or (iii) a difference in the selection intensity on individuals with intermediate DNA values. A fourth possibility could be differences in the types of DNA sequences which are gained or lost from a genome given that plants in general possess greater quantities of repeated DNAs than do animals, and that it is often the repeated DNAs which vary quantitatively in plant genomes (Hutchinson *et al.* 1980; Flavell *et al.* 1974; Flavell, 1986; Narayan, 1988). Alternatively, there are plant groups where interspecific DNA contents are continuously distributed (Bennett *et al.* 1977; Price *et al.* 1981*a*, *b*; Kenton, 1983), and as noted by Labani & Elkington (1987) it is highly possible that some of the discontinuous DNA distributions may have resulted from a non-random or incomplete sampling of species within a given plant group.

A major constraint affecting consideration of the 'continuous versus discontinuous' modes of genome size evolution is that most reports of species DNA values are taken from only one or a few individuals of a species, and data on intraspecific genome size variation are relatively sparse. In cyprinid and centrarchid fish, genome size differences between individuals within populations and among species appear to result from small changes in DNA quantity. The only other comprehensive studies of intraspecific genome size variation of which we are aware are those of H. J. Price and colleagues on the plant genus *Microseris* (Price *et al.* 1980, 1981*a*, *b*, 1986). In *Microseris* at least, genome size variation among individuals within populations, among geographic populations within species, and among (annual) species is clearly continuous. Based on the above, we predict that as more plant and animal groups are

comprehensively surveyed for intraspecific as well as interspecific genome size variation, the discontinuous mode of genome size evolution will ultimately be falsified.

A second parallel between the cyprinid and centrarchid genome size data sets regards the proportion or quantity of DNA which apparently is free to vary within the genomes of these two fish groups. Gold & Amemiya (1987) estimated this fraction at the populational level as the maximum genome size difference between individuals within populations averaged over all populations surveyed. Their estimate for cyprinids of approximately 6% is nearly the same as the value of 6·2% for the nine *Lepomis* populations examined here. Although the cyprinids examined by Gold & Amemiya (1987) have about 19% more DNA than do the nine *Lepomis* species examined in this study, these estimates of roughly 6% represent approximately $1{\cdot}3 \times 10^8$ bp DNA in both groups. In theory, this quantity of DNA is essentially the same as that needed for the structural gene component in both genomes if one assumes liberal figures of 50000 structural nuclear genes per genome and 1500 coding DNA base pairs per gene.

Comparable data, i.e. genome sizes of several individuals sampled from each of several different populations, from other organisms are few, and are limited to Price *et al.*'s (1981*a*, 1986) studies of *Microseris* and Johnson *et al.*'s (1987) study of salmonid fishes. Using the data in those papers, we estimated the average maximum genome size difference between individuals within populations to be 4·5% (*ca.* $1{\cdot}3 \times 10^8$ bp DNA) for *Microseris** and 5·6% (*ca.* $2{\cdot}6 \times 10^8$ bp DNA) for salmonids. Although limited in terms of the number of organismal groups surveyed, these estimates imply that the quantity of DNA which is free to vary within populations is small compared to the genome as a whole. Of interest in the future will be to examine populational genome size variation in groups with considerably larger genome sizes.

Our finding that the patterns and quantity of genome size variation are similar in cyprinid and centrarchid fishes is of interest relative to the historical question of whether genome size *per se* has an adaptive function. The major hypotheses forwarded to date are (i) the 'nucleotype' theory of Bennett (1971, 1972), (ii) the *r*- versus *K*-selection theory of Cavalier-Smith (1980), and (iii) the wasteful (W) versus frugal (F) adaptive strategy theory of Szarski (1983). Briefly, Bennett (1971, 1972) proposed that genome size evolved in response to selection acting directly on the 'nucleotype' which he defined as certain biophysical parameters such as cell or nuclear size (or area) and minimum meiotic and mitotic cycle times. In general, organisms with small genome sizes should tend to have small cells and rapid reproductive and growth rates; whereas organisms with larger genome sizes should tend to have large cells and slow reproductive and growth rates. Cavalier-Smith (1980) proposed a similar theory and suggested that observed variations in genome size could be the secondary result of a varying balance between *r*-selection for rapid reproductive rates and *K*-selection for larger cell volume. In his view, *r*-selected species would have smaller genome sizes as a result of selection for smaller cells and more rapid growth and reproduction, *K*-selected species would be the reverse, and the genome size of an organism would essentially reflect its position along an *r*- and *K*-selection continuum. A point to note is that Cavalier-Smith's (1980) hypothesis was based primarily on data from protists with a life-cycle lacking cell differentiation. Finally, Szarksi (1983) proposed that an organism's genome size may also reflect its position along a different continuum, that of wasteful (W) versus frugal (F) adaptive strategies. In his view, species whose metabolic and developmental rates are reduced because of lower oxygen or other environmental conditions are F-adapted and should have larger genome sizes; W-adapted species, conversely, have higher metabolisms and faster development and should have smaller genome sizes.

Although no direct, standardized experimental data on either growth or metabolic rates are available for both cyprinids and centrarchids, the basic ecologies and life histories of species in both families would appear to suggest that relative to one another cyprinids would be considered as *r*- and W-strategists, whereas centrarchids would be considered as *K*- and F-strategists. In brief, cyprinids are typically small, short-lived, highly fecund opportunists that grow rapidly, mature quickly, and prefer more oxygen-rich habitats. Centrarchids (including *Lepomis*), alternatively, are generally larger, longer-lived, more slowly growing and later maturing fish that prefer less oxygen-rich habitats (Lee *et al.* 1980 and references therein). Although exceptions to the above generalities occur in both families, one would predict, based on extrapolations of the adaptive hypotheses to complex (higher) organisms, that cyprinids should have smaller genome sizes. The reverse, however, appears to be the case in that the range of DNA values from the 20 cyprinid species examined by Gold & Amemiya (1987) was 2·19–2·72 pg as compared to the range of 1·92–2·05 pg for the nine *Lepomis* species examined in this study. Stated differently, the nine *Lepomis* species have, on the average, about 19% less DNA than do the 20 cyprinid species. On the surface, this would appear to falsify at least the general predictions of genome sizes as related to adaptive continua based on growth and/or metabolic rates. Alternatively, neither the *r*- versus *K*-selection or the W-versus F-adaptive strategy hypotheses have been subjected to critical, experimental testing, and both are based either on data from organisms lacking cell differentiation in their life-

* The estimate for *Microseris* (from Price *et al.* 1981*a*) was based only on populations where five or more individuals were examined.

cycle, on rather 'loose' correlations between genome size and various organismic and/or ecological factors, or both. We therefore suspect that the general predictions of the models relating genome size variation to organismal position along adaptive continua may be oversimplified or not applicable to complex, higher eukaryotes.

On the other hand, the normality of the genome size distributions within both cyprinid and centrarchid species does indicate that stabilizing or normalizing selection may be operating through the truncation of deleterious extremes (Stebbins, 1966; Mettler & Gregg, 1969). Assuming that selection acts at the level of individuals within populations, one could hypothesize that individuals with genome sizes which are too large or too small for efficient growth and development have reduced fitness. This would not necessarily mean that selection for (or against) some organismal or biophysical parameter is operating, but rather that accidental gain or loss of coding structural or regulatory gene DNAs could be interfering significantly with normal cellular processes.

A comparatively recent suggestion put forward to account for at least part of the variation in genome size observed among organisms is the so-called 'selfish' or 'parasitic' DNA hypothesis (Doolittle & Sapienza, 1980; Orgel & Crick, 1980). The basis for this hypothesis is that many (if not most) eukaryotic genomes contain DNA sequences that can increase their copy number in a genome through differential replication. Presumably, such sequences can increase in quantity within a genome at least to the point where the amount of energy expended in replicating their DNA begins to infringe on the metabolic and energy needs of the organism (Doolittle & Sapienza, 1980). To some extent, the cyprinid and centrarchid genome size data are consistent with the 'selfish' DNA hypothesis in that (i) a significant fraction of both genomes appears to vary quantitatively within species and have little or no observable phenotypic consequence, (ii) species DNA values appear to be randomly distributed within the variation which occurs, and (iii) individuals at the high end of the genome size distribution may be removed by normalizing selection. However, the 'selfish' DNA hypothesis cannot explain how or why individuals at the low end of the genome size distribution also appear to be removed by normalizing selection, nor can it account for the normality of the genome size distributions within nearly every cyprinid or centrarchid species surveyed. Other inconsistencies with the 'selfish' DNA hypothesis were discussed by Price (1988), and as noted by Gold & Amemiya (1987), empirical evidence confirming a direct relationship between genome size and 'selfish' DNA proliferation will be exceedingly difficult to obtain.

The final point of interest to consider is the historical question of whether genome size changes are involved in the process of speciation. The past assumptions that genome sizes in organisms are non-randomly distributed, as well as the reports of discontinuously distributed species DNA values (see above), have led to the inference that genome size change may precipitate or at least be associated with new species formation (Hinegardner, 1976; Morescalchi, 1977; Cavalier-Smith, 1978). In both cyprinids and *Lepomis*, the majority of the genome size variation is distributed above the hierarchical level of individuals within populations of species. This is apparent from both nested analyses of variance and estimated genome size difference (or distance) values between individuals at the different levels of taxonomic organization (Gold & Amemiya, 1987; this paper). Without question, these data would appear to indicate that considerable genome size divergence occurs during the evolution of species. However, whether genome size differentiation occurs at speciation in a 'punctuated' manner (*sensu* Eldredge & Gould, 1972) cannot be ascertained from this perspective. One could hypothesize, for example, that genome size divergence occurs gradually as part of the divergence which typically accompanies speciation.

To circumvent this problem, Avise & Ayala (1975) and Avise (1978) developed a set of theoretical models which contrast expected means and variances of genetic distance among living members of rapidly versus slowly speciating taxa. Briefly, the models contrast expected means and variances of genetic distance among living members of rapidly versus slowly speciating taxa. If divergence is gradual and a function of time, the ratios of distance between species-rich versus species-poor taxa should be very nearly one, and the ratio of variances should be less than one. Alternatively, if divergence is punctuated and a function of speciation episodes, the ratio of both distances and variances should be greater than one. There are several assumptions underlying the models, the major one of which is that the taxa under comparison be of approximately the same evolutionary age. Previous tests of the models (Avise, 1977; Avise & Gold, 1977; Douglas & Avise, 1982) have employed cyprinid and centrarchid fish, specifically the cyprinid genus *Notropis* (species-rich) versus the centrarchid genus *Lepomis* (species-poor). The reasons why these fish are appropriate for comparison are discussed fully in Avise & Ayala (1975) and Avise (1977).

The mean and variance of genome size distance among the nine *Lepomis* species as determined in this study were 0·943 and 0·225, respectively. Comparable values for the twelve *Notropis* species studied by Gold & Amemiya (1987) when normalized for the difference in genome size between the two groups were 1·136 and 0·719. For the comparison of *Notropis* (species-rich) versus *Lepomis* (species-poor), the ratio of mean distances is 1·20 and the ratio of variances is 3·19. According to the models, these results indicate that changes in genome size are correlated with speciation

episodes. At present, however, the results should be viewed only as tentative, in part because of several recent objections to the models themselves (Mayden, 1986), and in part because the *Notropis* sample included only twelve of the more than 100 living *Notropis* species. In addition, since the evidence is essentially correlative, it would be difficult to determine whether the correlation is one of cause and effect or one of association. Finally, our finding in both centrarchids and cyprinids that intraspecific variation in genome size can, in some instances, be as great as interspecific variation does cast some doubt on the idea of a strong relationship between genome size change and speciation. One might hypothesize, for example, that the generally lower intraspecific genome size variation compared with interspecific variation could stem from gene flow within species which might tend to homogenize genome sizes. Obviously, direct empircal evidence confirming a relationship between genome size change and speciation *sensu strictu* will be difficult to obtain.

We thank Joe Crumpton, Bill Johnston, Wes Porak, and Allen Rudd for assistance in collecting specimens; Chuck Cichra and Terrence Walters for help in data analyses; Jennifer Riggs for help with the scanning microdensitometer; and John Bickham, Jim Price, and an anonymous reviewer for constructive comments on the manuscript. The scanning microdensitometer used in the research was made available for our use by Dr Jim Price of the Soil and Crop Sciences Department at Texas A & M University. The chicken blood used as an internal standard was provided by Dr Syed Naqi of the Texas A & M College of Veterinary Medicine. The work was supported by project H-6703 of the Texas Agricultural Experiment Station and by National Science Foundation grant BSR-8415428.

References

Avise, J. C. (1977). Is evolution gradual or rectangular? Evidence from living fishes. *Proceedings of the National Academy of Sciences U.S.A.* **74**, 5083–5087.

Avise, J. C. (1978). Variances and frequency distributions of genetic distance in evolutionary phylads. *Heredity* **40**, 225–237.

Avise, J. C. & Ayala, F. J. (1975). Genetic change and rates of cladogenesis. *Genetics* **81**, 757–773.

Avise, J. C. & Gold, J. R. (1977). Chromosomal divergence and speciation in two families of North American fishes. *Evolution* **31**, 1–13.

Bachmann, K., Chambers, K. L. & Price, H. J. (1985). Genome size and natural selection: observation and experiments in plants. In *The Evolution of Genome Size* (ed. T. Cavalier-Smith), pp. 267–276. New York: John Wiley & Sons.

Bachmann, K., Goin, O. B. & Goin, C. J. (1972). Nuclear DNA amounts in vertebrates. *Brookhaven Symposia in Biology* **23**, 419–450.

Bennett, M. D. (1971). The duration of meiosis. *Proceedings of the Royal Society of London* **B178**, 277–299.

Bennett, M. D. (1972). Nuclear DNA content and minimum mitotic time in herbaceous plants. *Proceedings of the Royal Society of London* **B181**, 109–135.

Bennett, M. D., Gustafson, J. P. & Smith, J. B. (1977). Variation in nuclear DNA in the genus *Secale*. *Chromosoma* **61**, 149–176.

Bianchi, N. O., Redi, C., Garagna, C., Capanna, E. & Manfredi-Romanini, M. G. (1983). Evolution of genome size in *Akodon* (Rodentia, Cricetidae). *Journal of Molecular Evolution* **19**, 362–370.

Cavalier-Smith, T. (1978). Nuclear volume control by nucleoskeletal DNA, selection for cell volume and growth rate, and the solution of the DNA C-value paradox. *Journal of Cell Science* **34**, 247–278.

Cavalier-Smith, T. (1980). *r*- and *K*-tactics in the evolution of protist developmental systems: cell and genome size, phenotype diversifying selection and cell cycle patterns. *Biosystems* **12**, 43–59.

Cavalier-Smith, T. (1985*a*). Introduction: the evolutionary significance of genome size. In *The Evolution of Genome Size* (ed. T. Cavalier-Smith), pp. 1–36. New York: John Wiley & Sons.

Cavalier-Smith, T. (1985*b*). Cell volume and the evolution of eukaryotic genome size. In *The Evolution of Genome Size* (ed. T. Cavalier-Smith), pp. 105–184. New York: John Wiley & Sons.

Doolittle, W. F. & Sapienza, F. (1980). Selfish genes, the phenotype paradigm and genome evolution. *Nature (London)* **284**, 617–618.

Douglas, M. E. & Avise, J. C. (1982). Speciation rates and morphological divergence in fishes: tests of gradual versus rectangular modes of evolutionary change. *Evolution* **36**, 224–232.

Eldredge, N. & Gould, S. J. (1972). Punctuated equilibria: an alternative to phyletic gradualism. In *Models in Paleobiology* (ed. T. J. M. Schopf), pp. 82–116. San Francisco: Freeman, Cooper & Co.

Fand, S. B. (1970). Environmental conditions for optimal Feulgen hydrolysis. In *Introduction to Quantitative Cytochemistry*, vol. II (ed. G. L. Wied & G. F. Bahr), pp. 209–221. New York: Academic Press.

Flavell, R. B. (1986). Repetitive DNA and chromosome evolution in plants. *Philosophical Transactions of the Royal Society of London.* **B312**, 227–242.

Flavell, R. B., Bennett, M. D., Smith, J. B. & Smith, D. B. (1974). Genome size and the proportion of repeated nucleotide sequence DNA in plants. *Biochemical Genetics* **12**, 257–269.

Gold, J. R. & Amemiya, C. T. (1987). Genome size variation in North American minnows (Cyprinidae). II. Variation among 20 species. *Genome* **29**, 481–489.

Gold, J. R. & Price, H. J. (1985). Genome size variation among North American minnows (Cyprinidae). I. Distribution of the variation in five species. *Heredity* **54**, 297–305.

Gold, J. R., Bennett, L. F. & Gall, G. A. E. (1975). A set of tables for determining minimum sample sizes necessary to show statistically significant differences among treatment groups using analysis of variance. Division of Agricultural Sciences, University of California Special Publication Number 3051, pp. 1–9, with Tables I–XII.

Hinegardner, R. (1976). Evolution of genome size. In *Molecular Evolution* (ed. F. J. Ayala), pp. 179–199. Sunderland, Mass.: Sinauer Press.

Hinegardner, R. & Rosen, D. E. (1972). Cellular DNA content and the evolution of teleostean fishes. *American Naturalist* **106**, 621–644.

Humason, G. L. (1979). *Animal Tissue Techniques*. San Francisco: W. H. Freeman & Company.

Hutchinson, J., Narayan, R. K. J. & Rees, H. (1980). Constraints on the composition of supplementary DNA. *Chromosoma* **78**, 137–145.

Johnson, O. W., Utter, F. M. & Rabinovitch, P. S. (1987). Interspecies differences in salmonid cellular DNA identified by flow cytometry *Copeia:* 1001–1009.

Kauffman, S. (1971). Gene regulation networks: a theory for their global structure and behaviours. In *Current*

Topics in Developmental Biology (ed. A. A. Moscona & A. Monroy), pp. 145–182. New York: Academic Press.

Kenton, A. (1983). Qualitative and quantitative chromosome changes in the evolution of *Gibasis*. In *Kew Chromosome Conference*, vol. II (ed. P. E. Brandham & M. D. Bennett), pp. 273–282. London: Allen & Unwin.

Labani, R. M. & Elkington, T. T. (1987). Nuclear DNA variation in the genus *Allium* L. (Liliaceae). *Heredity* **59**, 119–128.

Lee, D. S., Gilbert, C. R., Hocutt, C. H., Jenkins, R. E., McAllister, D. E. & Stauffer, J. R., Jr (1980). *Atlas of North American Freshwater Fishes*. Raleigh, NC: North Carolina Biological Survey Publication Number 1980–12.

Mayden, R. L. (1986). Speciose and depauperate phylads and tests of punctuated and gradual evolution: fact or artifact? *Systematic Zoology* **35**, 591–602.

Mettler, L. E. & Gregg, T. G. (1969). *Population Genetics and Evolution*. Englewood Cliffs, NJ: Prentice-Hall.

Mirsky, A. E. & Ris, H. (1951). The deoxyribonucleic acid content of animal cells and its evolutionary significance. *Journal of General Physiology* **34**, 451–462.

Morescalchi, A. (1977). Phylogenetic aspects of karyological evidence. In *Major Patterns in Vertebrate Evolution* (ed. M. K. Hecht, P. C. Goody & B. M. Hecht), pp. 149–167. New York: Plenum Press.

Narayan, R. K. J. (1982). Discontinuous DNA variation in the evolution of plant species. The genus *Lathyrus*. *Evolution* **36**, 877–891.

Narayan, R. K. J. (1983). Chromosome change in the evolution of *Lathyrus* species. In *Kew Chromosome Conference*, vol. II (ed. P. E. Brandham & M. D. Bennett), pp. 243–250. London: Allen & Unwin.

Narayan, R. K. J. (1988). Constraints upon the organization and evolution of chromosomes in *Allium*. *Theoretical and Applied Genetics* **75**, 319–329.

Ohno, S. & Atkin, N. B. (1966). Comparative DNA values and chromosome complements of eight species of fishes. *Chromosoma* **18**, 455–466.

Orgel, L. E. & Crick, F. H. C. (1980). Selfish DNA: the ultimate parasite. *Nature* (*London*) **284**, 645–646.

Pierce, B. A. & Mitton, J. B. (1980). The relationship between genome size and genetic variation. *American Naturalist* **116**, 850–861.

Price, H. J. (1988). Plant genome size and the DNA C-value paradox. *Plant Genetics Newsletter* **4**, 18–24.

Price, H. J., Bachmann, K., Chambers, K. L. & Riggs, J. (1980). Detection of intraspecific variation in nuclear DNA content in *Microseris douglasii*. *Botanical Gazette* **141**, 195–198.

Price, H. J., Chambers, K. L. & Bachmann, K. (1981*a*). Geographic and ecological distribution of genomic DNA content variation in *Microseris douglasii* (Asteraceae). *Botanical Gazette* **142**, 415–426.

Price, H. J., Chambers, K. L. & Bachmann, K. (1981*b*). Genome size variation in diploid *Microseris bigelovii* (Asteraceae). *Botanical Gazette* **142**, 156–159.

Price, H. J., Chambers, K. L., Bachmann, K. & Riggs, J. (1986). Patterns of mean nuclear DNA content in *Microseris douglasii* (Asteraceae) populations. *Botanical Gazette* **147**, 496–507.

Raina, S. N., Srivastava, P. K. & Ramo Rao, S. (1986). Nuclear DNA variation in *Tephrosia*. *Genetica* **69**, 27–33.

Rasch, E. M., Barr, H. J. & Rasch, R. W. (1971). The DNA content of sperm of *Drosophila melanogaster*. *Chromosoma* **33**, 1–18.

Sherwood, S. W. & Patton, J. L. (1982). Genome evolution in pocket gophers (genus *Thomomys*). II. Variation in cellular DNA content. *Chromosoma* **85**, 163–179.

Sims, L. E. & Price, H. J. (1985). Nuclear DNA content variation in *Helianthus* (Asteraceae). *American Journal of Botany* **72**, 1213–1219.

Sneath, P. H. A. & Sokal, R. R. (1973). *Numerical Taxonomy*. San Francisco: W. H. Freeman & Sons.

Sokal, R. R. & Rohlf, F. J. (1969). *Biometry*. San Francisco: W. H. Freeman & Sons.

Sparrow, A. H., Price, H. J. & Underbrink, A. G. (1972). A survey of DNA content per cell and per chromosome of prokaryotic and eukaryotic organisms: some evolutionary considerations. *Brookhaven Symposia in Biology* **23**, 451–494.

Stebbins, G. L. (1966). *Processes of Organic Evolution*. Englewood Cliffs, NJ: Prentice-Hall.

Szarski, H. (1983). Cell size and the concept of wasteful and frugal evolutionary strategies. *Journal of Theoretical Biology* **105**, 201–209.