

# Genome size variation among north american minnows (Cyprinidae). I. distribution of the variation in five species

John R. Gold and  
H. James Price

Department of Wildlife and Fisheries Sciences and  
Department of Soil and Crop Sciences,  
Texas A&M University, College Station, Texas 77843

Genome sizes (nuclear DNA contents) were examined spectrophotometrically from ten individuals of each of five species of North American cyprinid fishes (minnows). The distributions of DNA values both within and between the five species were essentially continuous and normal. Differences between individuals within populations were significant and contributed to approximately 16 per cent of the total variation. Variation between individuals within species ranged from 4.7–13.5 per cent and averaged *ca.* 7.4 per cent. Variation between species ranged from 0–9.5 per cent and the average difference between any species pair was *ca.* 4.6 per cent. Statistical analyses showed that the methodology used was sufficient to detect significant differences in genome size as small as 2–3 per cent. Consideration of these data lead to the following tentative conclusions: (i) changes in genome size in cyprinids appear small in amount, frequent in occurrence, to involve both gains and losses of DNA, and to be cumulative and independent in effect; (ii) differences within and between cyprinid taxa are likely the result of accumulations of small changes in DNA quantity; and (iii) the primary focus of quantitative DNA variation in cyprinids is between individuals within populations. The extent of DNA quantity variation which occurs within species would appear to preclude any direct relationship between genome size variation and many of the organismal parameters (including speciation) which differentiate the five species. In short, the data suggest that a significant fraction of the cyprinid genome, perhaps more than 10 per cent, is free to vary quantitatively without phenotypic constraint or biological consequence. This fraction is considerably larger than that theoretically needed for the structural gene component.

## INTRODUCTION

A long-standing problem in evolutionary genetics regards the quantitative variation in genome size or nuclear DNA content (the C-value) among eukaryotic organisms. Abundant data are now accumulated which show that large, often spectacular differences in genome size commonly occur between even closely related taxa, and that increases in genome size are not necessarily associated with evolutionary advancement (Bachmann *et al.*, 1972; Rees and Jones, 1972; Hinegardner, 1976; Price, 1976). Early suggestions (Kauffman, 1971) were that the variation was related to either the number of genes in an organism, its organismal complexity, or both. The general consensus now, however, is that there are no significant correlations between genome size (the C-value) and organismal or genetic complexity (Cavalier-Smith,

1978). This is called the C-value paradox. Other pertinent findings which have emerged are that: (i) increasing organismal specialization in body form and design may often be associated with decreasing genome size (Hinegardner and Rosen, 1972; Hinegardner, 1976); (ii) the quantitative differences in genome size, even between close relatives, generally exceed the most generous theoretical estimates of the amount of DNA needed to code for all the structural or regulatory RNA molecules or proteins of the organism (Mizuno and Macgregor, 1974; Murray *et al.*, 1981); and (iii) the differences appear to reflect primarily gains or losses of repeated sequence DNAs (Flavell *et al.*, 1974; Hutchinson *et al.*, 1980).

The underlying causes of genome size variation are not well understood. Much of the data has been interpreted as supporting the idea that the variation has an adaptive basis and is strongly

influenced by natural selection (Sparrow *et al.*, 1972; Cavalier-Smith, 1978; 1980; Price *et al.*, 1981a). This in turn has led to suggestions regarding genome size variation and its possible relationship to speciation and phylogeny (Hatch *et al.*, 1976; Hinegardner, 1976; Cavalier-Smith, 1978). There also are the recent suggestions (Ohno, 1972; Doolittle and Sapienza, 1980; Orgel and Crick, 1980) that much of the variation may reflect gain or loss of phenotypically inconsequential DNAs. As pointed out by Sherwood and Patton (1982), however, almost all the data on genome size variation among both animals and plants are from comparisons of distinct species or higher taxa. A few studies (Price *et al.*, 1981a, b; Sherwood and Patton, 1982) have shown that substantial differences in genome size do occur between isolated populations of the same species, and that significant differences also exist between individuals of the same population. On the whole, however, the dynamics or patterns of change in genome size at lower hierarchical levels, especially at and around the species level of differentiation, remain poorly understood, and it will be difficult to evaluate the meaning of quantitative DNA differences between higher taxa without data on the extent of the variation within those taxa (Sherwood and Patton, 1982).

This paper represents the first of a series of investigations on genome size variation in the cyprinid fishes (minnows) endemic to North America. These fishes are extremely prolific in terms of number of individuals and number of (recently diverged) species, are broadly distributed in North America, and display a wealth of diversity in habitats, morphological adaptations, and behaviors. As such, they will provide an excellent model system in which to study DNA quantity variation at lower hierarchical levels. In this report, a highly reproducible protocol for measuring genome size in cyprinids is described, and quantitative data on the magnitude and distribution(s) of DNA variation within and between five species are presented and discussed.

#### MATERIALS AND METHODS

Four of the five cyprinid species examined were collected by seine from natural populations and returned live to the laboratory; these included *Camptostoma anomalum* (Boardhouse Cr., Blanco R. drainage, Blanco Co., TX); *Notropis venustus* (Bull Cr., Colorado R. drainage, Travis Co., TX); and *Notropis lutrensis* and *Pimephales vigilax*

(Little Brazos R., Brazos R. drainage, Brazos Co., TX). The sample of one species, *Notemigonus crysoleucas*, was purchased live from a local bait shop in Bryan, TX, and originated from a minnow ranch in Arkansas. After processing, all specimens were deposited in the Texas Cooperative Wildlife Collection at Texas A&M University.

Relative genome sizes of individual fish (10 individuals per species) were determined microspectrophotometrically using Feulgen-stained erythrocyte interphase nuclei. Blood was obtained by cardiac puncture of single fish and smeared near the frosted end on each of three slides; on the far end of each slide a (freshly made) smear of chicken blood served as the internal standard. The chicken blood was obtained by vein puncture from individuals of a very highly inbred, pathogen-free strain (SPAFAS) available from the Pathology Department at the Texas A&M College of Veterinary Medicine. The slides were then fixed 20 min in 9:1 methanol-formaldehyde (37 per cent), rinsed twice (10 min each) in distilled water, dehydrated in 70 per cent ETOH (2 min) and 95 per cent ETOH (2 min), and stored overnight under desiccated conditions at 4°C. The following day, individual batches of 20 (randomized) slides were hydrolysed 35 min in 3.5 N HCl at 37°C, rinsed briefly in distilled water, and stained two hours in Schiff's reagent (Feulgen stain). The conditions of acid hydrolysis followed Fand (1970); the hydrolysis time was determined empirically as the point of maximum absorbancy in a hydrolysis curve. Following staining, the slides were rinsed twice (10 min each) in SO<sub>2</sub> water and once (10 min) in distilled water, air dried in the dark, cleared in xylene, and mounted in Permount. All slides were coded by number, randomized, and stored in the dark until analysed. Schiff's reagent and SO<sub>2</sub> water preparation followed standard recipes (Humason, 1979).

The microspectrophotometric apparatus used was a Zeiss Universal-II scanning microscope with a 03 photometer system, 0.5 µm scanning stage, modified 45-control unit, modified PMI-indicator, and a value averaging module. All measurements were performed at a scanning speed of 8/64 with a damping setting of 0.1, and using a 100× oil-immersion objective at 560 nm light. For each individual fish, 15 nuclei were measured from each of two slides (=30 nuclei per individual), and all slides (over all fish) were examined in random (blind) sequence. The third slide prepared from each specimen served as a backup in case of breakage. Only nuclei which were roughly spherical, homogeneously Feulgen-stained and from clear

areas on the slide, and which had homogeneous background absorbancies were chosen for measurement. Small, positively heteropycnotic nuclei were avoided for reasons outlined in Garcia (1970). The decision to measure 30 nuclei per individual was based on preliminary experiments (measurements of 20–50 cells per slide and over 50 slides) that showed an average coefficient of variation (per slide and per individual) of from 3–4 per cent. This means that measuring 30 nuclei per fish should differentiate between a 2–3 per cent difference in mean genome size at  $\alpha$ - and  $\beta$ -probability levels of 0.05 (Sokal and Rohlf, 1969; Gold *et al.*, 1975). The experimental design of measuring two slides and blind slide selection for staining and scanning should randomise between slide and between staining batch variation.

Absorbancy values of fish and chicken nuclei were recorded into data books, loaded on to mini-discs using a small laboratory computer, and transferred to the University main-frame computer. Absorbancy values of fish nuclei from each slide were then standardised as a percentage of the mean absorbancy value of chicken nuclei on that slide and then coded (for convenience) by multiplying the percentage chicken standard (for each fish nucleus) by 20. The latter was chosen simply because it is the average absorbancy of chicken erythrocyte nuclei under our experimental conditions. For conversion to picograms of DNA, the coded data are simply reconverted to standardised data (percentage of chicken standard) then multiplied by 2.5, the generally accepted DNA value of diploid chicken erythrocyte nuclei (Rasch *et al.*, 1971). The coded data were subjected to descriptive statistical analysis which included generation of sample means, variances, ranges, coefficients of variation, and the  $g_1$  and  $g_2$  indices of distribution normality. Homogeneity of variances was tested by Bartlett's method using the correction factor in Sokal and Rohlf (1969). Homogeneity of means was tested using one-way analysis of variance and mean separation was accomplished using Duncan's multiple range test. Hierarchical levels of the variation were analysed using a nested analysis of variance. Except for the tests of variance homogeneity which were done by hand, all statistical analyses were carried out on the University main-frame computer using SAS programs.

## RESULTS

The coded absorbancy data initially were organised into seven different sampling distribu-

tions and each was tested for normality. The distributions included all DNA measurements (nuclei) among species ( $N = 1500$ ), all DNA measurements within each species (five distributions,  $N = 300$  each), and DNA values of individuals among species ( $N = 50$ ). The latter is essentially a distribution of means (*i.e.*, the DNA value for an individual is the mean of all 30 measured nuclei from that individual). The distributions of DNA values of individuals within each species were not tested because of their small ( $N = 10$ ) sample sizes. Only the distributions of DNA measurements among species and within *Notemigonus crysoleucas* were significantly non-normal (table 1); the deviations

**Table 1** Distribution normality statistics

Distribution	$N$	Skewness ( $g_1$ )	Kurtosis ( $g_2$ )
Measurements (nuclei) among species	1500	0.012	-0.645*
Measurements (nuclei) within species			
<i>Notropis venustus</i>	300	0.086	-0.452
<i>Notropis lutrensis</i>	300	-0.028	-0.289
<i>Campostoma anomalum</i>	300	0.155	-0.214
<i>Notemigonus crysoleucas</i>	300	0.376*	-0.091
<i>Pimephales vigilax</i>	300	0.018	-0.343
Individuals among species	50	0.088	-0.703

\* Significance at  $\alpha = 0.05$ . Positive  $g_1$  values indicate skewness toward higher values; negative  $g_2$  values indicate platykurtosis.

from normality in both cases, however, were only slight (figs. 1 and 2). Distribution normality tests also were carried out on log transformed data, primarily in response to a previous report (Bachmann *et al.*, 1972) that genome size variation in teleost fishes became normal only when log transformed. No significant change in the shape of any of the distributions was found except for measurements (nuclei) within *N. crysoleucas* which became normal when log transformed. This is to be expected for a distribution slightly skewed towards higher values (table 1).

Descriptive statistics from the distributions of all DNA measurements (nuclei) and of DNA values for all individuals over all species are shown in tables 2 and 3, respectively. The Bartlett's tests revealed significant heterogeneity ( $\chi^2_{[4]} = 58.4$ ,  $P < 0.05$ ) among species variances from the distribution of measurements, but not from the distribution of individuals ( $\chi^2_{[4]} = 7.4$ ,  $P > 0.05$ ). One-way analyses of variance revealed significant heterogeneity of species means in both distributions. Duncan's multiple range tests, however,

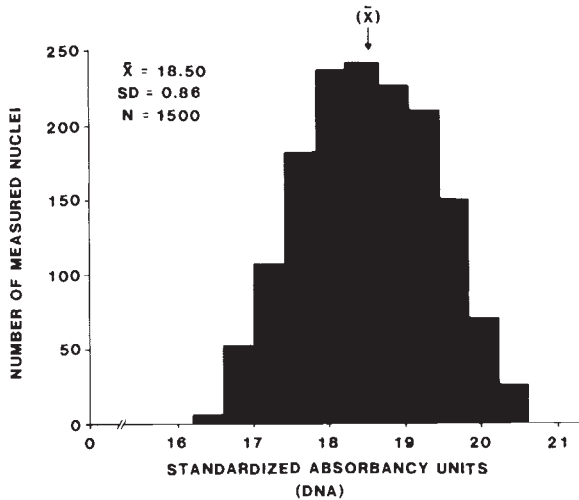


Figure 1 Frequency distribution (coded data) of DNA measurements (nuclei) over all five cyprinid species.

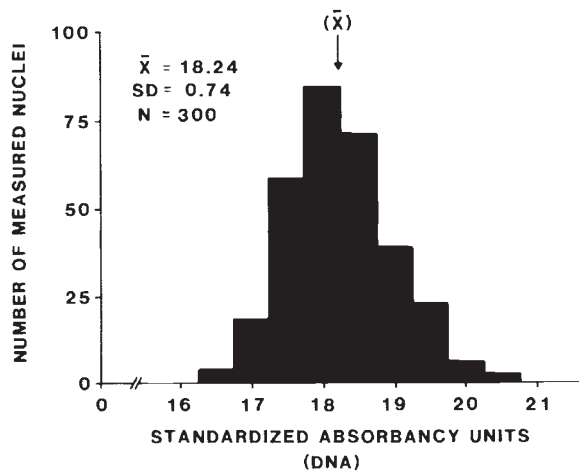


Figure 2 Frequency distribution (coded data) of DNA measurements (nuclei) from *Notemigonus crysoleucas*.

revealed that means of each species were significantly different from one another in the distribution of measurements (table 2), whereas in the distribution of individuals the means of *Notropis venustus* and *Notropis lutrensis*, and of *Campostoma anomalum* and *N. crysoleucas* were statistically identical (table 3). We interpret these results conservatively and suggest the multiple range tests from the distribution of individuals best reflect the true situation. This is based in part on the large number of degrees of freedom in the distribution of measurements which may in itself engender statistical significance, and in part on the fact that only the distribution of DNA values of individuals

Table 2 Descriptive statistics from the distribution of measurements

Taxon	Mean $\pm$ S.E.	Variance*	Range
<i>Notropis venustus</i> (N = 300)	<sup>a</sup> 19.32 $\pm$ 0.03	0.300	18.02-20.71
<i>Notropis lutrensis</i> (N = 300)	<sup>b</sup> 18.99 $\pm$ 0.03	0.316	17.53-20.49
<i>Campostoma anomalum</i> (N = 300)	<sup>c</sup> 18.30 $\pm$ 0.04	0.441	16.76-20.21
<i>Notemigonus crysoleucas</i> (N = 300)	<sup>d</sup> 18.24 $\pm$ 0.04	0.555	16.50-20.40
<i>Pimephales vigilax</i> (N = 300)	<sup>c</sup> 17.66 $\pm$ 0.03	0.263	16.23-19.25

\* Variances heteroscedastic at  $\alpha = 0.05$ .

<sup>a-c</sup> Means significantly differ at  $\alpha = 0.05$ .

Data are in coded values (cf. text).

Table 3 Descriptive statistics from the distribution of DNA values of individuals

Taxon	Mean $\pm$ S.E.	Variance*	Range
<i>Notropis venustus</i> (N = 10)	<sup>a</sup> 19.32 $\pm$ 0.09	0.087	18.87-19.74
<i>Notropis lutrensis</i> (N = 10)	<sup>a</sup> 18.99 $\pm$ 0.10	0.109	18.46-19.51
<i>Campostoma anomalum</i> (N = 10)	<sup>b</sup> 18.30 $\pm$ 0.14	0.194	17.66-19.05
<i>Notemigonus crysoleucas</i> (N = 10)	<sup>b</sup> 18.24 $\pm$ 0.20	0.400	17.20-19.49
<i>Pimephales vigilax</i> (N = 10)	<sup>c</sup> 17.66 $\pm$ 0.10	0.105	17.14-18.06

\* Variances homoscedastic at  $\alpha = 0.05$ .

<sup>a-c</sup> Means significantly differ at  $\alpha = 0.05$ .

Data are in coded values (cf. text).

was both normal and homoscedastic (i.e., had homogeneous sample variances). Since the smallest significant difference detected between species means was a little more than 3 per cent (*N. crysoleucas* vs. *P. vigilax*), whereas *N. venustus* and *N. lutrensis* (ca. 2 per cent difference) were statistically homogeneous (table 3), it follows that our methodology is sufficient to detect differences in relative genome sizes as small as 2-3 per cent. This also follows from consideration of the coefficients of variation among all 100 slides and all 50 individuals which were  $2.17 \pm 0.06$  and  $2.47 \pm 0.07$ ,

respectively, *i.e.*, the data should allow detection of 2-3 per cent differences between means at  $\alpha$ - and  $\beta$ -probability levels of 0.05 (Sokal and Rohlf, 1969; Gold *et al.*, 1975). Finally, the ranges (tables 2 and 3) of DNA measurements and of DNA values of individuals within each species show that genome size is essentially distributed continuously among the five species and that unequivocal identification of a given species based on its genome size cannot be made easily by measuring only a few nuclei from a few individuals. The only exception is the range of DNA values of individuals of *P. vigilax* which shows no overlap with the ranges of DNA values of individuals of *N. venustus* and *N. lutrensis*.

To determine whether significant intraspecific variation in genome size occurs, two analytical approaches were carried out. The first was a nested analysis of variance which showed (table 4) that significant variation exists at each hierarchical level from between slides (within individuals) to between species. While over half of the variation is explained by variation between species, an appreciable amount (*ca.* 16 per cent) is explained by variation between individuals within species.

The second approach was to test for heterogeneity of mean DNA values of individuals within each of the five species using one-way analysis of variance of the distributions of DNA measurements within each species (= five separate tests). In all five species, significant heterogeneity (at  $\alpha = 0.05$ ) of individual DNA means was detected. Duncan's multiple range tests (also at  $\alpha = 0.05$ ) revealed that on the average from four to five significantly different groupings occurred among the ten individuals within each species, and moreover, that the heterogeneity was distributed in a continuous fashion with modal DNA values within each species being approximately the same as species mean DNA values.

Summary statistics of genome size variation in diploid picograms of DNA and per cent differences within and between the five species are shown in table 5. The DNA values were taken from the distribution of DNA values of individuals. The variation over all individuals over all species ranged from 2.14 (one individual of *P. vigilax*) to 2.47 picograms of DNA (one individual of *N. venustus*) or about 15.4 per cent. The differences between species means ranged from 0.4 per cent

**Table 4** Nested analysis of variance for DNA data shown in table 2

Variance Source	D.F.	M.S.	F	Variance component	%
Total	1499	0.72	—	0.809	100
Species	4	130.50	24.3*	0.417	51.6
Individuals	45	5.37	3.6*	0.129	15.9
Slides	50	1.51	8.9*	0.089	11.0
Error	1400	0.17	—	0.174	21.5

\* Significance at  $\alpha = 0.05$ .

**Table 5** Summary statistics of genome size variation within and between five cyprinid species\*

Taxon	2N DNA Content (in picograms)		Difference from <i>Notropis venustus</i>	
	Mean $\pm$ S.E.	Range	Picograms	%
<i>Notropis venustus</i> (N = 10)	2.42 $\pm$ 0.01	2.36-2.47 (4.7%)	—	—
<i>Notropis lutrensis</i> (N = 10)	2.37 $\pm$ 0.01	2.31-2.44 (5.6%)	0.05	2.1
<i>Camptostoma anomalum</i> (N = 10)	2.29 $\pm$ 0.02	2.21-2.38 (7.7%)	0.13	5.7
<i>Notemigonus crysoleucas</i> (N = 10)	2.28 $\pm$ 0.02	2.15-2.44 (13.5%)	0.14	6.1
<i>Pimephales vigilax</i> (N = 10)	2.21 $\pm$ 0.01	2.14-2.26 (5.6%)	0.21	9.5

\* From the distribution of DNA values of individuals (*cf.* text).

Avg. % variation *within* species = 7.4% [0.16 pg =  $150 \times 10^6$  bp].

Avg. % difference *between* species = 4.6% [0.10 pg =  $94 \times 10^6$  bp].

(*C. anomalum* vs. *N. crysoleucas*) to 9.5 per cent (*N. venustus* vs. *P. vigilax*), and averaged (ten pairwise comparisons) about 4.6 per cent or roughly 0.10 picograms of DNA. Variation between individuals within species ranged from 4.7 per cent in *N. venustus* to 13.5 per cent in *N. crysoleucas* and averaged about 7.4 per cent or roughly 0.16 picograms of DNA. These comparisons of the variation within and between the five species reveal two important aspects of quantitative DNA variation in cyprinids. First, because the range and average DNA quantity variation within populations of species are as great or greater than the range and average DNA quantity difference between species, the primary focus of quantitative DNA variation would appear to be between individuals within populations. This is further supported by the low coefficients of variation (from 3–4 per cent) observed over all slides and all individuals which suggest that individuals are essentially homogeneous in genome size. The second is that the average DNA quantity variation between individuals within populations was about 7.4 per cent of the cyprinid genome or roughly 0.16 picograms of DNA which represents approximately 150 million base pairs of DNA. Assuming a liberal figure of 50,000 structural genes in the cyprinid genome and 1500 coding DNA base pairs per gene, the quantity of DNA which varies between individuals within populations is about twice that theoretically needed for the structural gene component.

## DISCUSSION

At the outset, two points should be noted. The first is that this study represents only the beginning of a long-term project on DNA variation in cyprinid fishes, and the considerations which follow should be regarded primarily as grounds for hypotheses to be tested in the future. The second is that this initial data set does reflect a random and fairly representative sample of North American Cyprinidae. Briefly, four of the five species belong to the dominant New World subfamily Leuciscinae which includes all but one endemic species; that species is *Notemigonus crysoleucas* which is thought to belong to the Old World subfamily Abramidinae (Hubbs, 1955; Miller, 1959). Of the four leuciscines, *N. venustus* and *N. lutrensis* are closely related species in the diverse subgenus *Cyprinella* of *Notropis* (Gibbs, 1957), the genus *Pimephales* is thought to be a specialised derivative of *Notropis* (Hubbs and Black, 1947), and the

genus *Campostoma* is only distantly related to both *Notropis* and *Pimephales*. The species also differ biologically. All are fairly wide-spread in North America, but *Notemigonus* prefers a markedly different habitat from the rest, *Campostoma* is a specialised herbivore (the others are essentially omnivorous), and both *Campostoma* and *Notemigonus* grow to a considerably larger size than the other three (Lee *et al.*, 1980, and references therein).

The normality (or near-normality) of all the distributions tested in the present data set along with the continuous variation of DNA values of individuals observed both within and across the five species suggest that genome size variation essentially follows the premises of the normal probability density function (see Sokal and Rohlf, 1969: pp. 104–105). From this we infer first, that changes in genome size in cyprinids are small in amount, frequent in occurrence, involve both gains and losses of DNA, and are cumulative and independent in effect; and second, that differences within and between cyprinid taxa are the result of accumulations of small changes in DNA quantity. A point to note is that all five species have  $2n = 50$  very similar chromosomes (Gold *et al.*, 1980) such that most or all the genome size changes must have occurred within chromosomes and not from chromosomal aneuploidy.

Comparison of the pattern of continuous DNA quantity variation in cyprinids with patterns in other organisms is not feasible at present because of the paucity of data on genome size variation below the species level in most organismal groups. There are of course numerous studies in both animals and plants which have shown sharp discontinuities in genome size between related species (e.g., Mizuno and Macgregor, 1974; Narayan, 1982). Whether these discontinuities arose from continuous change (as appears to be the case in cyprinids) is unknown. One reason as pointed out by Sherwood and Patton (1982) is simply that almost all past studies have tended to characterise species by single DNA values and hence failed to appreciate the extent to which genome size may vary within species. Obviously, further studies of genome size variation within a broad range of species are needed to address the issue.

The finding that both the range and average magnitude of genome size variation within populations of the five species is the same or greater than the range and magnitude of the differences between the species raises a number of interesting points. The first is that intraspecific genome size variation, in this case between individuals within popula-

tions, should not be regarded as insignificant or unimportant (*e.g.*, Bennett and Smith, 1976), but rather as the focus of genome size change itself and hence the source of the variation from which species values arise. That the focus of genetic change is at the level of individuals within populations is not contrary to general evolutionary thought, but it does seem to have been disregarded by investigators concerned primarily with assessing genome sizes only between higher taxa. One obvious corollary to the above is that species-specific DNA values in many cases may be misleading at best.

The second point concerns the underlying biological causes of DNA quantity variation. The most frequently encountered hypotheses in the literature are that the variation has an adaptive basis and is strongly influenced by natural selection (Sparrow *et al.*, 1972; Cavalier-Smith, 1978, 1980; Price *et al.*, 1981a). The data are primarily the long-standing interspecific correlations observed between genome size and certain biophysical parameters such as cell or nuclear size and minimum meiotic or mitotic cycle times (Bennett, 1971, 1972; Szarski, 1974; Cavalier-Smith, 1978; 1982). Organismal phenotypes used to demonstrate these correlations have included body size, clinal or habitat differences, and several life-history characteristics (Ebeling *et al.*, 1971; Hinegardner, 1974; Bennett, 1976; Mazin, 1980; Shuter *et al.*, 1983). Other hypotheses suggested to account for the variation have included organismal specialisation in body form and design (Hinegardner and Rosen, 1972; Hinegardner, 1974, 1976), accumulations (or losses) of phenotypically inconsequential DNAs (Ohno, 1972; Doolittle and Sapienza, 1980; Orgel and Crick, 1980), and variation in structural gene heterozygosity (Pierce and Mitton, 1980). The latter, however, has been severely criticized as being drawn from an inappropriate data base (Larson, 1981; Parker and Kreitman, 1982).

The normality of the DNA value distributions within and between these cyprinid species does suggest that stabilizing or normalizing selection may be operating through the truncation of deleterious extremes (Stebbins, 1966; Mettler and Gregg, 1969). This does not necessarily mean that selection for some organismal parameter favours a particular species DNA value, but rather that a few individuals may have genome sizes which are too large or too small for efficient growth and development. Very possibly, this may reflect accidental gain or loss of coding, structural or regulatory gene DNAs which might be expected to significantly interfere with normal cellular pro-

cesses. However, the extent of DNA quantity variation which occurs within these cyprinid species would appear to preclude any direct relationships between genome size variation and many of the organismal parameters which differentiate the five species. These parameters include systematic relationships, body size, and a few life-history characteristics. There also would seem to be no reason to suggest a positive relationship between genome size variation and higher level processes such as speciation. What the cyprinid data suggest simply is that a certain fraction of the genome, perhaps as much as 15 per cent (see below) is free to vary quantitatively without phenotypic constraint or biological consequence. Again, the salient point to be made is that it will be difficult to evaluate the significance of genome size variation between higher taxa without knowledge on the extent of the variation within them (Sherwood and Patton, 1982).

The final point which merits brief consideration is the size and the nature of the DNA fraction which apparently is free to evolve within cyprinid genomes. Based on the present data, we estimate that fraction to be minimally 7.4 per cent (the average DNA quantity variation within populations), although it could be as high as 13–15 per cent (the range of variation of DNA values of individuals in *N. crysoleucas*), or even higher. These figures are not surprising in view of Sherwood and Patton's (1982) finding that intraspecific genome size variation in the gopher *Thomomys bottae* can be as high as 35 per cent, and further demonstrate the difficulty in hypothesising a meaning for the genome size variation and the need for further studies on the DNA quantity variation within species. We have no direct data yet as to the qualitative nature of the DNA which varies quantitatively in cyprinids. However, our preliminary (unpublished) experiments on genomic DNA melting profiles and chromosomal C-bands show that cyprinids are considerably enriched in highly-repeated and heterochromatic DNAs as compared to other teleost fish. Based on data from other organisms (Flavell *et al.*, 1974; Hutchinson *et al.*, 1980), it is likely that the sequences which vary quantitatively in cyprinids are repeated DNAs.

In closing, it should be noted that the cyprinid DNA quantity data are in accord with certain predictions of the junk or selfish DNA hypothesis (Ohno, 1972; Doolittle and Sapienza, 1980; Orgel and Crick, 1980) in that (i) a significant fraction of the cyprinid genome appears to vary quantitatively within species and have little or no

phenotypic consequence, and (ii) species DNA values appear to be more or less randomly distributed within the variation which occurs. This does not, however, prove the existence of junk or selfish DNA, nor does it falsify the hypothesis that natural selection acts on the variation in genome size. What will be important in future work is to determine the qualitative nature of the DNA sequences which vary quantitatively in cyprinids as well as the hierarchical or phyletic level(s) at which natural selection could or does act on the variation in genome size.

*Acknowledgements* We thank C. T. Amemiya, J. R. Ellison, and L. J. Schliesing for help in collecting specimens. We also thank Dr S. A. Naqi for providing the chicken blood used as the internal standard. The spectrophotometric equipment used was purchased off NSF Grant DEB-7809940 to HJP; the work was funded by The Texas Agricultural Experiment Station off Projects 6187 and 6703 to J.R.G.

## REFERENCES

- BACHMANN, K., GOIN, O. B. AND GOIN, C. J. 1972. Nuclear DNA amounts in vertebrates. *Brookhaven Symp. Biol.*, **23**, 419-450.
- BENNETT, M. D. 1971. The duration of meiosis. *Proc. R. Soc. Lond. Ser. B.*, **178**, 277-299.
- BENNETT, M. D. 1972. Nuclear DNA content and minimum mitotic time in herbaceous plants. *Proc. R. Soc. Lond. Ser. B.*, **181**, 109-135.
- BENNETT, M. D. 1976. DNA amount, latitude, and crop plant distribution. *Envir. Exptl. Bot.*, **16**, 93-108.
- BENNETT, M. D. AND SMITH, J. B. 1976. Nuclear DNA amounts in angiosperms. *Phil. Trans. R. Soc. London. Ser. B.*, **274**, 227-274.
- 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. *J. Cell Sci.*, **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. 1982. Skeletal DNA and the evolution of genome size. *Ann. Rev. Biophys. Bioeng.*, **11**, 273-302.
- DOOLITTLE, W. F. AND SAPIENZA, C. 1980. Selfish genes, the phenotypic paradigm and genome evolution. *Nature*, **284**, 601-603.
- EBELING, A. W., ATKIN, N. B. AND SETZER, P. Y. 1971. Genome sizes of teleostean fishes: increases in some deep-sea species. *Amer. Nat.*, **105**, 549-561.
- FAND, S. B. 1970. Environmental conditions for optimal Feulgen hydrolysis. In Wied, G. L. and Bahr, G. F. (eds.) *Introduction to Quantitative Cytochemistry*, II, Academic Press, New York and London, pp. 209-221.
- 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. Genet.*, **12**, 257-269.
- GARCIA, A. M. 1970. Stoichiometry of dye binding versus degree of chromatin coiling. In Wied, G. L. and Bahr, G. F. (eds.) *Introduction to Quantitative Cytochemistry*, II, Academic Press, New York and London, pp. 153-170.
- GIBBS, R. H., JR. 1957. Cyprinid fishes of the subgenus *Cyprinella*, with a key to the species exclusive of the *lutrensis-ornatus* complex. *Copeia*, 185-195.
- GOLD, J. R., BENNETT, L. F. AND 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. *Div. Agric. Sci., Univ. California, Spec. Pub.* #3051, pp. 1-9, with Tables I-XII.
- GOLD, J. R., KAREL, W. J. AND STRAND, M. R. 1980. Chromosome formulae of North American Fishes. *Progr. Fish-Cult.*, **42**, 10-23.
- HATCH, F. T., BODNER, A. J. AND MAZRIMAS, J. A. 1976. Satellite DNA and cytogenetic evolution. *Chromosoma*, **58**, 155-168.
- HINEGARDNER, R. 1974. Cellular DNA content of the Mollusca. *Comp. Biochem. Physiol.*, **47A**, 447-460.
- HINEGARDNER, R. 1976. Evolution of genome size. In Ayala, F. J. (ed.) *Molecular Evolution*, Sinauer, Sunderland, MA, pp. 179-199.
- HINEGARDNER, R. AND ROSEN, D. 1972. Cellular DNA content and the evolution of teleostean fishes. *Amer. Nat.*, **106**, 621-644.
- HUBBS, C. L. 1955. Hybridization between fish species in nature. *Syst. Zool.*, **4**, 1-20.
- HUBBS, C. L. AND BLACK, J. D. 1947. Revision of *Ceratichthys*, a genus of American cyprinid fishes. *Misc. Pub. Mus. Zool., Univ. Michigan*, **66**, 1-56.
- HUMASON, G. L. 1979. *Animal Tissue Techniques*. Freeman and Sons, San Francisco, CA.
- HUTCHINSON, J., NARAYAN, R. K. J. AND REES, H. 1980. Constraints on the composition of supplementary DNA. *Chromosoma*, **78**, 137-145.
- KAUFMAN, S. 1971. Gene regulation networks: a theory for their global structures and behaviors. In Moscona, A. A. and Monroy, A. (eds.) *Current Topics in Developmental Biology*, Academic Press, New York and London, pp. 145-182.
- LARSON, A. 1981. A re-evaluation of the relationship between genome size and genetic variation. *Amer. Nat.*, **118**, 119-125.
- LEE, D. S., GILBERT, C. R., HOCUTT, C. H., JENKINS, R. E., McALLISTER, D. E. AND STAUFFER, J. R. 1980. *Atlas of North American Freshwater Fishes*. Pub. #1980-12, North Carolina Biol. Surv., North Carolina St. Mus. Nat. Hist., Raleigh, N.C.
- MAZIN, A. I. 1980. Amounts of nuclear DNA in anurans of the USSR. *Experientia*, **36**, 190-191.
- MITTLER, L. E. AND GREGG, T. G. 1969. *Population Genetics and Evolution*. Prentice-Hall, Englewood Cliffs, NJ.
- MILLER, R. R. 1959. Origin and affinities of the freshwater fish fauna of western North America. *Zoogeography, Amer. Assoc. Adv. Sci. Pub.*, **51**, 187-222.
- MIZUNO, S. AND MACGREGOR, H. C. 1974. Chromosomes, DNA sequences, and evolution in salamanders of the genus *Plethodon*. *Chromosoma*, **48**, 239-296.
- MURRAY, M. G., PETERS, D. L. AND THOMPSON, W. F. 1981. Ancient repeated sequences in the pea and mung bean genomes and implications for genome evolution. *J. Mol. Evol.*, **17**, 31-42.
- NARAYAN, R. K. J. 1982. Discontinuous DNA variation in the evolution of plant species: the genus *Lathyrus*. *Evolution*, **36**, 877-891.
- OHNO, S. 1972. So much "junk" DNA in our genome. *Brookhaven Symp. Biol.*, **23**, 366-370.



- ORGEL, L. E. AND CRICK, F. H. C. 1980. Selfish DNA: the ultimate parasite. *Nature*, 284, 604-607.
- PARKER, E. D., JR. AND KREITMAN, M. 1982. On the relationship between heterozygosity and DNA content. *Amer. Nat.*, 119, 749-752.
- PIERCE, B. A. AND MITTON, J. B. 1980. The relationship between genome size and genetic variation. *Amer. Nat.*, 116, 850-861.
- PRICE, H. J. 1976. Evolution of DNA content in higher plants. *Bot. Rev.*, 42, 27-52.
- PRICE, H. J., CHAMBERS, K. L. AND BACHMANN, K. 1981a. Geographic and ecological distribution of genomic DNA content variation in *Microseris douglasii* (Asteraceae). *Bot. Gaz.*, 142, 415-426.
- PRICE, H. J., CHAMBERS, K. L. AND BACHMANN, K. 1981b. Genome size variation in *Microseris bigelovii* (Asteraceae). *Bot. Gaz.*, 142, 156-159.
- RASCH, E. M., BARR, H. J. AND RASCH, R. W. 1971. The DNA content of sperm of *Drosophila melanogaster*. *Chromosoma*, 33, 1-18.
- REES, H. AND JONES, G. H. 1972. The origin of the wide species variation in nuclear DNA content. *Int. Rev. Cytol.*, 32, 53-92.
- SHERWOOD, S. W. AND PATTON, J. L. 1982. Genome evolution in pocket gophers (genus *Thomomys*). II. Variation in cellular DNA content. *Chromosoma*, 85, 163-179.
- SHUTER, B. J., THOMAS, J. E., TAYLOR, W. D. AND ZIMMERMAN, M. 1983. Phenotypic correlates of genomic DNA contents in unicellular eukaryotes and other cells. *Amer. Nat.*, 122, 26-44.
- STEBBINS, G. L. 1966. *Processes of Organic Evolution*. Prentice-Hall, Inc., Englewood Cliffs, NJ.
- SOKAL, R. R. AND ROHLF, F. J. 1969. *Biometry*. W. J. Freeman and Sons, San Francisco, CA.
- 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.
- SZARSKI, H. 1974. Cell size and nuclear DNA content in vertebrates. *Int. Rev. Cytol.*, 44, 93-111.