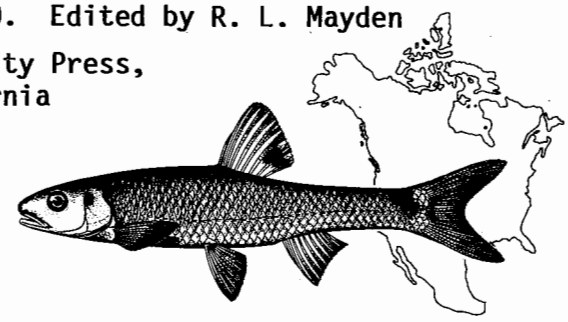
Reprinted from: Systematics, Historical Ecology, & Freshwater Fishes
1992, pp. 534-550. Edited by R. L. Mayden
Stanford University Press,
Stanford, California

19



Evolution of Genome Size in North American Fishes

John R. Gold, Chara J. Ragland, and James B. Woolley

A long standing problem in the evolutionary genetics of eukaryotes regards the quantitative variation in genome size or DNA content. Abundant data are now available which show that variation in genome size among eukaryotes spans several orders of magnitude and that sizeable differences often occur even between closely related species (Bachmann et al., 1972; Sparrow et al., 1972; Hinegardner and Rosen, 1972). Early suggestions (Kauffman, 1971) were that the variation in genome size was related directly to organismal and/or genetic complexity. It is now clear, however, that there are no correlations between genome size (the C-value) and organismal or genetic complexity (Cavalier-Smith, 1985a; Price, 1988). This has been termed the C-value paradox.

Hypotheses invoked to explain the C-value paradox have focused primarily on an "adaptive" function for the genome size variation and have been based on studies where genome size was found to vary in relation to parameters associated with organismal growth, metabolism, or differentiation (Bennett, 1971, 1972; Cavalier-Smith, 1985a, b; Szarski, 1983; Sessions and Larson, 1987). Other suggestions offered to explain interspecific genome size variation include (i) differences in specialization of body form and design (Hinegardner and Rosen, 1972), (ii) accumulations of phenotypically inconsequential or "selfish" DNAs (Orgel and Crick, 1980; Doolittle and Sapienza, 1980), and (iii) variation in structural gene heterozygosity (Pierce and Mitton, 1980). Of these, only the last has been seriously criticized (Larson, 1981; Parker and Krietman, 1982), although criticisms of the selfish DNA hypothesis have been published (Cavalier-Smith, 1980; Price, 1988). Finally, because of the large differences in genome size encountered between even closely related species, it has 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 are the assumptions that (i) the variations in genome size are non-randomly distributed, at least among species, and (ii) genome

size is constant within species. A significant consequence of the latter is that almost all of the quantitative data on genome size variation in eukaryotes are from fully differentiated species and/or higher taxa. Several recent studies, however, have shown that intraspecific variation in genome size may be substantial and often approximates the average genome size differences found between species (Sherwood and Patton, 1982; Price et al., 1986; Johnson et al., 1987). These findings suggest that the assumption of constancy of genome size within species may be invalid, and moreover, that variation in genome size within species is neither well understood nor adequately characterized. As succinctly noted by Sherwood and Patton (1982), it will be difficult to evaluate the meaning or significance of genome size variation among species without an understanding of genome size variation within species.

A last point to note is that relatively little attention has been given to the mechanisms by which DNA might be gained or lost from a genome. The findings that species within cohesive taxonomic groupings (e.g., genera) often differ substantially in genome size and that genome sizes among species are frequently distributed discontinuously have promoted the suggestion that genome size change may occur in a "quantized" manner by a succession of large-scale changes (Narayan, 1982; Cavalier-Smith, 1985b). Subsumed within this issue is the hypothesis noted previously that changes in genome size may be associated with the process of speciation.

In this chapter, we examine and summarize quantitative variation in genome size within and among North American species of the fish families Cyprinidae and Centrarchidae. Much of the data is published (Gold and Price, 1985; Ragland, 1986; Gold and Amemiya, 1987; Ragland and Gold, 1989; Gold et al., 1990), although the entire study (as initially envisaged) is by no means finished. The subjects of primary interest are the pattern and magnitude of genome size variation within populations and among species, and the question of whether genome size changes are associated with speciation events. Because the detection of changes occurring during speciation events necessitates comparing sister groups, the latter was examined using a genealogy of species relationships derived from phylogenetic analysis and represents a significant departure from our previous attempts (Ragland, 1986; Gold et al., 1990) to examine this issue. Finally, the distribution of the genome size variation in relation to life-history and habitat differences in some of the cyprinid species is also briefly considered.

MATERIALS AND METHODS

To date, genome size determinations have been carried out on 464 individuals representing 49 North American cyprinid species and 23 North American centrarchid species. Collection localities for all of the species examined may be found in Ragland (1986), Gold and Amemiya (1987), and Gold et al. (1990). All fish were collected by either seining or electroshocking from natural populations. Individuals sampled more than 200 or so miles from College Station were processed in the field or at facilities made available by colleagues (see Acknowledgments); those sampled near College Station were returned live to our laboratory and then processed. Genome sizes were determined by scanning microdensitometry of Feulgen-stained erythrocyte nuclei using chicken blood as an internal control. The latter was obtained from a highly inbred, pathogen free strain obtained from the Texas A&M College of Veterinary Medicine. Details of slide preparation, staining, and microden-

sitometry may be found in Gold and Price (1985) and Ragland and Gold (1989). 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 and 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. Statistical analyses of the data as described below were carried out using either SAS (1982) or our own programs on the Texas A&M mainframe computer.

RESULTS AND DISCUSSION

Intraspecific Genome Size Variation

Descriptive statistics (mean $\pm$ standard errors and ranges) for the 49 cyprinid species and 23 centrarchid species may be found in Gold et al. (1990) and Ragland (1986), respectively. Genome sizes among the cyprinids ranged from 2.06 picograms of DNA in *Cyprinella callistia* to 3.26 picograms of DNA in *Phenacobius castostomus*, a difference of approximately 58%. Genome sizes among the centrarchids ranged from 1.88 picograms of DNA in *Lepomis macrochirus* to 2.56 picograms of DNA in *Ambloplites rupestris*, a difference of approximately 36%.

A summary of genome size variation within populations of both cyprinids and centrarchids is shown in Table 1. Briefly, the data were organized into a number of different sampling distributions and each was tested for normality using the g_1 and g_2 indices for skewness and kurtosis, respectively (Sokal and Rohlf, 1969). The distributions tested included (i) all measurements (nuclei) within each population or sample (73 sampling distributions), and (ii) two rankit distributions (one for cyprinids and one for centrarchids) reflecting the distribution of individuals within populations (summed over all populations). The latter was generated following Equation (1) in Gold and Amemiya (1987) in order to remove scaling effects due to individuals being drawn from different species (Sokal and Rohlf, 1969). The majority of the distributions of measurements (nuclei) within populations of cyprinids were normal, although the incidence of non-normal distributions was higher than expected by chance at $\alpha = 0.05$. All of the distributions of measurements (nuclei) within populations of centrarchids were normal. Both rankit distributions, however, were significantly platykurtic although visual inspection of the frequency histograms [shown in Ragland (1986) and Gold et al. (1990)] indicated that the deviations from normality were slight.

The normality (or near normality) of genome size distributions within populations of both cyprinids and centrarchids suggests that changes in DNA quantity at the populational level are small in amount, involve both gains and losses of DNA, and are cumulative and independent in effect. This is based on the assumption that genome size variation within populations of species follows the premises of the normal probability density function (Sokal and Rohlf, 1969). Of significance is that no instance of a large-scale or "quantized" (Cavalier-Smith, 1985b) difference in genome size among individuals has been found in the 73 populations of cyprinids and centrarchids thus far studied. The platykurtosis observed in a few of the cyprinid population genome size distributions, and in both rankit distributions, however, deserves further comment. Since the deviations from normality in each case appear

Table 1. Summary of genome size variation within populations of North American cyprinids and centrarchids

Parameter	Cyprinids	Centrarchids
Number of populations surveyed	49	24
Measurements (nuclei) within populations	37 Normal 6 PK* 14 SR 1 SL 1 PK, SR	24 Normal — — —
Individuals within populations**	PK	PK
Average number of significantly different groups within populations§	4.48† 2.79††	2.79†† —
Average percent difference between individuals within populations	4.86 ± 0.31	5.15 ± 0.50

* PK= platykurtic; SR= skewed towards higher values; SL= skewed towards lower values; all deviations from normality were significant at $\alpha > 0.05$.

** Rankit distributions of DNA values of individuals (see text).

§ From Duncan's multiple range tests (see text).

† N = 10 (21 populations).

†† N = 5 (28 populations) for cyprinids; N=5 (23 populations) and N=4 (1 population) for centrarchids.

to be small (Ragland, 1986; Gold et al., 1990), we have hypothesized that the platykurtosis may be due to non-random sampling. Our normal methods of capture are to seine or electroshock individuals from only one or two "beaches" in a stream or pond, such that many of the individuals sampled from the same species could represent close relatives (e.g., full sibs) rather than individuals drawn at random from the population. Consistent with this hypothesis are the data in Ragland (1986) where DNA values of individuals within populations of centrarchids often appeared "clumped" on either or both sides of the populational DNA mean value.

Separate single classification analyses of variance were used to test for heterogeneity of genome sizes of individuals within populations using the distribution of measurements (nuclei) of that species. All F-tests were significant at $\alpha = 0.05$, and on average, approximately half of the individuals sampled from any given population differed significantly in genome size (Table 1). The magnitude of genome size difference among individuals within populations was estimated as the percent difference between the high and low individuals of each population. Among cyprinids, this value ranged from 1.5% in *Lythrurus bellus* to 13.49% in *Notemigonus crysoleucas* (Gold and Amemiya, 1987; Gold et al., 1990); among centrarchids, the range was from 1.01% in *Pomoxis nigromaculatus* to 12.18% in *Lepomis cyanellus* (Ragland, 1986). The average values of 4.86 ± 0.31% (cyprinids) and 5.15 ± 0.50% (centrarchids) represent the average maximum variation in genome size between individuals within populations. In both groups, this represents approximately 100

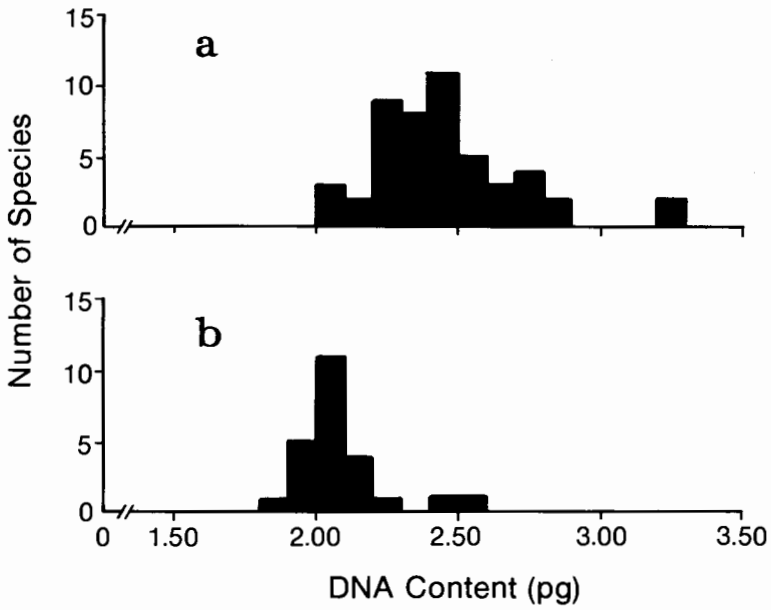


Figure 1.
Frequency distributions of species DNA values for cyprinids (A) and centrarchids (B).

million base pairs of DNA (Price, 1976). Assuming both cyprinid and centrarchid genomes contain roughly 50,000 structural genes, and 1,500 coding DNA base pairs per gene, the quantity of DNA that apparently is free to vary quantitatively among individuals within populations is approximately one-third greater than the amount of DNA theoretically needed to code for the entire structural gene component. Finally, the finding that both the range and average magnitude of difference in genome size among individuals in both families is similar suggests that the mechanism(s) and/or potential for genome size change is similar in both families.

Interspecific Genome Size Variation

Plots of the distribution of DNA values of species for both families are shown in Figure 1. With the exception of two species of *Phenacobius* in Cyprinidae, and two species of *Ambloplites* and one species of *Enneacanthus* in Centrarchidae, the interspecies distributions appear continuous and overlapping. Single classification analysis of variance using the distribution of DNA values of individuals was used to test for significant heterogeneity in genome size variation among species in both families. Significant heterogeneity of DNA values among species at $\alpha = 0.05$ was found in both families. The results of Duncan's multiple range tests may be found in Ragland (1986) and Gold et al. (1990). Again, with the exception of the species noted above, interspecies genome sizes were continuously distributed in both families.

Nested analyses of variance were used to examine the magnitude of genome size variation in both families at each experimental level from between slides within individuals to among species. Significant heterogeneity at $\alpha = 0.05$ was found at each level (i.e., between slides within individuals, among individuals within species, and among species) in both families (Ragland, 1986; Gold et al., 1990). The majority of the genome size variation in both families, however, occurred among species. In

cyprinids, more than 88% of the variation occurs among species (Gold et al., 1990); and in centrarchids, more than 78% of the variation occurs among species (Ragland, 1986).

To examine further the magnitude of genome size variation, average genome size differences (or distances) between individuals drawn from ascending taxonomic levels were computed using Equations [2] and [3] of Gold and Amemyia (1987). In brief, Equation [2] generates a genome size distance (GSD) value that represents the average of all possible pairwise differences in genome size between all individuals of any two species. Equation [3] generates a GSD value that represents the average of all possible pairwise differences in genome size between all individuals from any single population of a species. The latter, when averaged over all species within a family, yields an estimate of the average genome size difference (distance) between individuals within populations of species for that family. Both GSD values are minimum linear difference or distance estimates and hence may underestimate the true difference (Sneath and Sokal, 1973). A point to note is that the taxonomic groupings used in the analysis were based upon phylogenetic hypotheses of relationships, rather than upon a non-phylogenetic classification scheme.

The average genome size differences (distances) between individuals drawn from successive levels of taxonomic divergence in both families are shown in Table 2. The values for "Within populations" were derived from Equation [3] in Gold and Amemiya (1987) and represent the average difference in genome size between any two individuals drawn at random from any population of a species. The values for "Species in genera" and "Species in family" were derived from Equation [2] in Gold and Amemiya (1987) and represent the average difference in genome size between any two species in the same genus and any two species in the same family, respectively. As shown, individuals drawn at random from a population of the same cyprinid or centrarchid species will differ, on average, by approximately 0.05 picograms of DNA. A striking difference between two families, however, is observed in the average genome size difference between species at both the generic and familial levels. In cyprinids, congeneric species differ, on average, by about 0.22 picograms of DNA, whereas in centrarchids, congeneric species differ, on average, by about 0.10 picograms of DNA. Stated differently, the average genome size difference between any two species in the same cyprinid genus is more than twofold greater than the average genome size difference between any two species in the same centrarchid genus. A similiar difference, although of lesser magnitude, is observed at the level of species within the family (Table 2). Taken together, the results in both families (especially in cyprinids) indicate that the majority of genome size divergence occurs above the level of individuals within populations, thus corroborating the results of the nested analyses of variance (Ragland, 1986; Gold et al., 1990).

In Table 3, the average genome size difference (distance) between species in the different cyprinid and centrarchid genera sampled to date are shown. Although incomplete in terms of the number of extant species sampled, especially in some cyprinid genera, significant ($P<0.05$) correlations are found overall ($r = 0.68$, $t_{[2]} = 3.04$) and within Cyprinidae ($r = 0.77$, $t_{[2]} = 3.00$) between the average genome size difference between species in a genus and the number of extant species in that genus. The tentative implication of the correlations may be that a positive relationship exists between the number of species within a genus and average divergence in genome size.

Table 2. Average genome size difference (distance) between individuals drawn from successive levels of taxonomic divergence.

Level	Mean genome size difference ± s.e.*	Number of pairwise comparisons
Within populations		
Cyprinidae	0.05 ± 0.00	1,225
Centrarchidae	0.05 ± 0.01	236
Species in genera		
Cyprinidae**	0.22 ± 0.02	116
Centrarchidae†	0.10 ± 0.01	46
Species in family		
Cyprinidae	0.29 ± 0.01	1,176
Centrarchidae††	0.17 ± 0.01	253

* In picograms of DNA.

** Does not include: *Notropis baileyi* whose generic placement has recently been questioned (Mayden and Matson, 1988); and *Opsopoeodus emiliae* which is now thought to be related to the genera *Cyprinella* and *Pimephales* rather than the *N. volucellus* species-group of *Notropis* (Cavender and Coburn, 1986; Amemiya and Gold, 1990; Mayden, pers. comm.).

† Includes two subspecies of *Micropterus salmoides*, but not *Lepomis gulosus*.

†† Includes two subspecies of *Micropterus salmoides*, but not *Elassoma zonatum*.

We have argued elsewhere (Gold and Price, 1985; Ragland, 1986; Ragland and Gold, 1988; Gold et al., 1990) that the data on genome size variation within populations of cyprinids and centrarchids strongly imply the action of stabilizing or normalizing selection through the truncation of deleterious extremes (Stebbins, 1966; Mettler and Gregg, 1969). Both the normality of genome size distributions and the apparent constraints on the quantity of DNA that can vary within cyprinid and centrarchid populations suggest that selection may be eliminating individuals whose genome sizes are too large or too small for efficient growth and development. Very possibly, this could reflect the accidental gain or loss of structural or regulatory gene DNAs which would be expected to interfere with normal cellular processes. Although consistent with the hypotheses that genome size is strongly influenced by natural selection (Bennett, 1971, 1972; Cavalier-Smith, 1985a, b; Szarski, 1983), our data do not necessarily indicate that selection for (or against) some organismal parameter influenced by genome size is operating. This point is to be emphasized since the "adaptive" hypotheses were initially proposed to account for genome size differences among species as a function of selection acting on parameters directly or indirectly related to genome size. Gold and Amemiya (1987) examined this premise by testing associations between interspecific variation in genome size and several life-history parameters (including body size) in twenty cyprinid species. No significant correlations or associations were found, suggesting that species DNA values were randomly distributed with respect to the variation in life-history parameters examined. How-

Table 3. Average genome size difference (distance) between species in different genera.

Taxon	Number of species examined	Number of extant species in genus*	Mean genome size difference**
Cyprinidae			
Cyprinella	8	27	0.27
Luxilus	2	10 †	0.17
Lythrurus	3	8	0.11
Notropis	13	46	0.21
Pimephales	3	4	0.05
Campostoma	2	4	0.06
Nocomis	2	7	0.02
Phenacobius	2	5	0.06
Centrarchidae			
Ambloplites	2	4	0.09
Enneacanthus	2	3	0.19
Lepomis	8	10 †	0.13
Micropterus	6	7 †	0.05
Pomoxis	2	2	0.02

* Phylogenetic relationships and placement of species in cyprinid genera follow Mayden (1989, pers. comm.) with the exception that *N. baileyi* and *N.* (=*Opsopoeodus*) *emiliae* were not included in the genus *Notropis*. Taxonomy of and placement of species in centrarchid genera follow Lee et al. (1980).

** In picograms of DNA.

† *Luxilus* includes two subspecies of *L. chrysocephalus*; *Lepomis* does not include *L. gulosus*; *Micropterus* includes two subspecies of *M. salmoides*.

ever, as noted by Gold and Amemiya (1987), the findings that significant differences in genome size occur within populations of cyprinids and that species genome sizes overlap considerably confounds the problem of identifying significant correlations or associations even if they exist.

Speciation and Rates of Genome Size Evolution

Our findings that the majority of genome size variation in both cyprinids and centrarchids appears to occur above the level of individuals within populations, and that a positive relationship may exist between the number of species within a clade and divergence in genome size, suggest that genome size change could be concentrated in speciation episodes. The hypothesis that genome size change might be associated with speciation is not new (Hinegardner, 1976; Morescalchi, 1977; Cavalier-Smith, 1978), although a positive correlation between genome size change and speciation has not been demonstrated empirically due to the difficulty in discriminating changes occurring during speciation episodes from those occurring gradually over long periods of evolutionary time between speciation events.

To circumvent this problem in a general sense, Avise and Ayala (1975, 1976) and Avise (1978) developed models which contrast expected means and variance of

genetic difference or distance among living members of rapidly versus slowly speciating lineages. The models predict that if genetic divergence is gradual and a function of time, the ratio of mean distance between species-rich versus species-poor lineages should be approximately one, and the ratio of variances should be less than one. Conversely, if genetic divergence is "punctuated" (*sensu* Eldridge and Gould, 1972) and proportional to the number of speciation episodes, the ratio of mean distances should be greater than one, and the ratio of variances should be much greater than one. There are several assumptions underlying the models, the major one of which is that the lineages being evaluated are of comparable evolutionary age (Avise and Ayala, 1975; Avise, 1978). The models have been used previously by Avise (1977) and Douglas and Avise (1982) to test whether divergence in structural genes and morphology occurs as a function of time or as a function of speciation episodes. In both studies, the North American cyprinid genus *Notropis* (as formerly defined) was used as a species-rich lineage and the centrarchid genus *Lepomis* was used as a species-poor lineage. Avise (1977) also compared North American cyprinids (as a species-rich lineage) to North American centrarchids (as a species-poor lineage). The reasons why these groupings were chosen for comparison may be found in Avise and Ayala (1975) and Avise (1977). In both studies, differentiation (in structural genes and morphology) was found to occur primarily as a function of elapsed time.

We compared means and variances of average genome size differences (distances) between 32 species of *"Notropis"*-like shiners and eight species of *Lepomis* and between 45 species of Cyprinidae and 23 species of Centrarchidae. The *"Notropis"*-like shiner assemblage (*sensu* Mayden, 1989) included all but two of the species assayed for genome size that were formerly placed in the genus *Notropis*. To that extent, the *"Notropis"*-like shiner lineage was more-or-less analogous to *"Notropis"* as used by Avise (1977) and Douglas and Avise (1982). Two *Notropis* species (*N. atrocaudalis* and *N. ludibundus* [formerly *N. stramineus*]) were omitted from the calculations because they may not belong to a putatively monophyletic lineage which includes the remaining species (Mayden, 1989, pers. comm.). For similar reasons (Mabee, 1987; G. V. Lauder, pers. comm.), *Lepomis gulosus* was not included in the calculations of mean and variance values for the genus *Lepomis* and *Elassoma zonatum* was not included in the calculations for Centrarchidae. The calculations for Cyprinidae included only those species belonging to a lineage characterized by an opening in the posterior myodome (Mayden, 1989). The ratios of mean distances in both comparisons (viz., *"Notropis"*-like shiners versus *Lepomis* and Cyprinidae versus Centrarchidae) were greater than one as were the ratios of variances. According to the models, these results indicated that changes in genome size are correlated with (and hence concentrated in) speciation episodes.

Mayden (1986) criticized the models of Avise and Ayala (1975, 1976) and Avise (1978) as used by Avise (1977) and Douglas and Avise (1982). Although Mayden (1986) listed a number of concerns, his major criticisms involved the assumptions regarding the monophyly of the groups used, their evolutionary ages, and the types of data employed. In brief, Mayden (1986) noted that there was little evidence to support the monophyly of *Notropis*, *Lepomis*, North American Cyprinidae, and Centrarchidae, and that current evidence suggested that each group (especially *Notropis*) was either paraphyletic or artificial. Mayden's (1986) concern about the assumption of equal antiquity of *Notropis* and *Lepomis* and of North American Cyprinidae and Centrarchidae was predicated in part on the general problem of age

determinations based on the fossil record, in part on peculiarities and possible incorrect classifications of North American cyprinid and centrarchid fossils, and in part on the absence of evidence for monophyly of extinct and extant forms. Finally, Mayden (1986) raised several points regarding the "phenetic" (rather than "phylogenetic") nature of the approach in that (i) the methods employed in the models appear to have a bias toward averaging, and hence towards gradualism, and (ii) the methods failed to distinguish between primitive and derived characters and to account for character transformation.

In our analysis above, we compared only putatively monophyletic assemblages, viz., *"Notropis"*-like shiners versus *Lepomis* (without *L. gulosus*) and cyprinids with an opening in the posterior myodome versus centrarchids (without *E. zonatum*). We did, however, make the assumption that the lineages under comparison were of equal evolutionary age and we did not take character transformation along given branches into account. In order to overcome the criticisms of Mayden (1986) regarding the latter two problems, we employed the "relative rate test" logic of Sarich and Wilson (1967) and Wilson et al. (1977) and examined changes in genome size along the branches of two sister clades in each of two putatively monophyletic cyprinid groups. The comparison of sister clades alleviates the problems associated with age since sister clades are by definition the result of an historic speciation event rendering the clades equal in age (Wiley, 1981). The changes in genome size were partitioned among branches using the optimization (for maximum parsimony) procedure of Farris (1970) as employed by Larson (1984) and Sessions and Larson (1987). Hypothetical genome sizes at the basal node of each phylogeny were inferred by optimization with respect to appropriate outgroups.

The first group examined was a lineage comprised of the genera *Pimephales*, *Opsopoeodus*, and *Cyprinella*. According to Cavender and Coburn (1986) and Mayden (pers. comm.), *Pimephales* and *Opsopoeodus* are sister clades and comprise a monophyletic group that is sister to *Cyprinella*. The *Pimephales-Opsopoeodus* clade contains five extant species (four species of *Pimephales* plus one species of *Opsopoeodus*), whereas *Cyprinella* contains 27 extant species (Mayden, 1987, 1989). If the *Pimephales-Opsopoeodus* clade is sister to *Cyprinella*, both lineages are descended from the same common ancestor and both are of the same evolutionary age *relative* to one another. *Cyprinella*, however, appears to have experienced a greater number of speciation episodes. The second group examined was the lineage referred to as the *"Notropis"*-like shiners by Mayden (1989). In terms of the species examined for genome size, this lineage can be divided into two sister clades, one of which contains the genus *Notropis* (*sensu* Mayden, 1989), while the other contains the genera *Lythrurus*, *Luxilus*, *Cyprinella*, *Pimephales*, and *Opsopoeodus*. The *Notropis* clade contains over 70 extant species, some of which have yet to be formally described (Mayden, 1989, pers. comm.). The *Lythrurus-Luxilus-Cyprinella-Pimephales-Opsopoeodus* clade contains approximately 49–50 extant species (Mayden, 1989). Our systematic assessment of the two clades was based primarily on Mayden (1989) with the following exceptions: (i) *Opsopoeodus emiliae* was placed in the *Lythrurus* et al. clade as sister to the genus *Pimephales* (after Cavender and Coburn, 1986); and (ii) *Notropis rubellus* was placed in the subgenus *Notropis* (in a trichotomy with *Notropis atherinoides* and *Notropis stilbius*) instead of as a member of the subgenus *Hydrophlox* (Mayden and Matson, 1988). In addition, *Notropis baileyi* was not used in the analysis since its true affinities may not be with other members of the subgenus *Hydrophlox*,

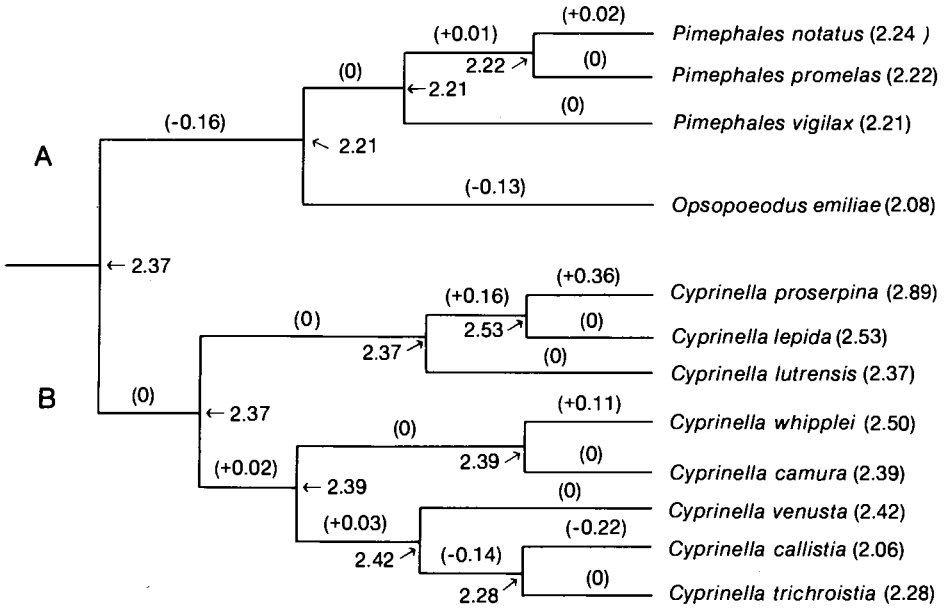


Figure 2.
Phylogenetic hypothesis used to estimate changes in genome size in the *Pimephales-Opsopoeodus* clade (A) and the *Cyprinella* clade (B). Amounts of inferred change in genome size along branches are shown in parentheses. DNA values for hypothetical ancestral nodes are indicated at each node (with arrows). Genome sizes for each species are also shown in parentheses. The ancestral DNA value for the *Pimephales-Opsopoeodus-Cyprinella* clade was taken from the analysis leading to Figure 3.

where it is currently placed (Mayden and Matson, 1988). Given that the two clades (i.e., the genus *Notropis* versus *Lythrurus* et al.) are sister to one another, both are putatively descended from the same common ancestor and both are of the same evolutionary age *relative* to one another. The *Notropis* clade, however, appears to have experienced more speciation episodes. As outgroups to the "*Notropis*"-like shiner lineage we used nine species assayed for genome size from the "chub" lineage (*sensu* Mayden, 1989) and a lineage comprised of *Ericymba buccata* and *Notropis ludibundus*. The choice of outgroups was based on Mayden (1989). The hypothetical genome size at the basal node of the *Cyprinella-Opsopoeodus-Pimephales* clade was obtained from the optimization analysis of the "*Notropis*"-like shiner assemblage.

The inferred changes in genome size in the first group are shown in Figure 2. Overall, 22 branches are represented: seven in the *Pimephales-Opsopoeodus* clade and 15 in the *Cyprinella* clade. Within both clades, gains or losses of DNA appear to occur at equal frequencies, although a considerable reduction in DNA content appears to have occurred in the *Pimephales-Opsopoeodus* lineage relative to the hypothetical ancestral value of 2.39 picograms. The inferred changes in genome size in the second group are shown in Figure 3. Overall, 54 branches are represented: 21 in the *Notropis* clade and 33 in the *Lythrurus* et al. clade. The latter is fully resolved, whereas the former is not and contains three unresolved polychotomies. As in the first comparison group, gains or losses of DNA appear to occur at approximately the same frequency in both clades.

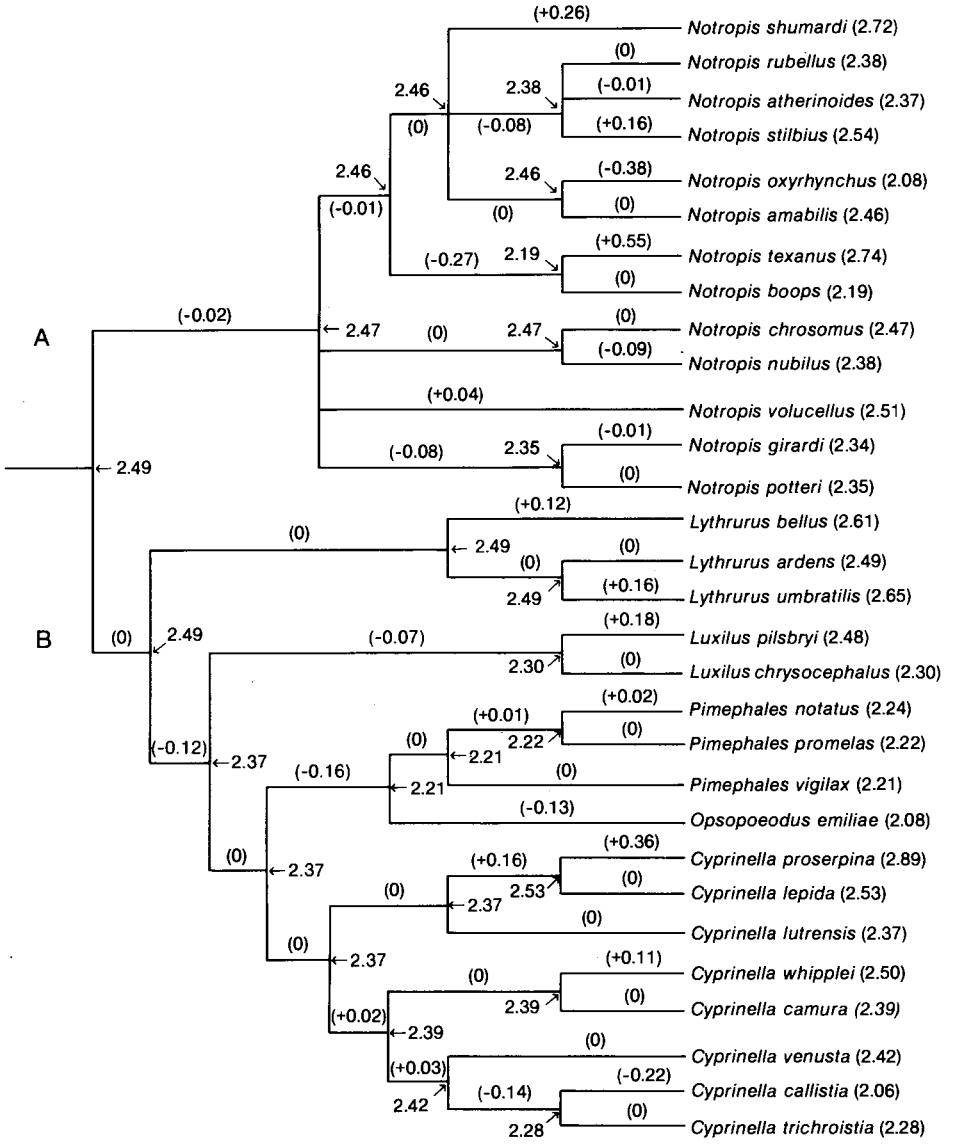


Figure 3.

Phylogenetic hypothesis used to estimate changes in genome size in the *Notropis* (A) and the *Lythrurus-Luxilus-Cyprinella-Pimephales-Opsopoeodus* clade (B). Amounts of inferred change in genome size along branches are shown in parentheses. DNA values for hypothetical ancestral nodes are indicated at each node (with arrows). Genome sizes for each species are also shown in parentheses. Topology, including outgroup taxa, is as follows: ((Clade A, Clade B), (*Ericymba bucatta, Notropis ludibundus*), (((*Phenacobius catostomus, Phenacobius mirabilis*), *Extrarius aestivalis*), ((*Nocomis asper, Nocomis leptocephalus*), (*Hybognathus placitus* (*Dionda episcopa* (*Campostoma anomalum, Campostoma oligolepis*)))))). DNA values for outgroup taxa may be found in Gold and Amemiya (1987). Tree topology follows Mayden (1989) with the exception of the placement of *Opsopoeodus emiliae*, the genus *Pimephales*, and *Notropis rubellus*. The last was placed by Swift (1970) in the subgenus *Hydrophlox* (of *Notropis*). More recent evidence (Mayden and Matson, 1988; Mayden, pers. comm.; T. E. Dowling, pers. comm.) suggests that *N. rubellus* is closely related to *N. atherinoides* and belongs in the subgenus *Notropis* (of *Notropis*).

Table 4. Comparisons of mean absolute genome size change per branch in species-rich versus species-poor clades in North American cyprinidae

Test group	Average genome size change/branch ± s.e.*	N†
Cyprinella[a]	0.069 ± 0.030	15
Pimephales-Opsopoeodus[b]	0.046 ± 0.026	7
Notropis[a]	0.093 ± 0.033	21
Lythrurus et al.[b]	0.061 ± 0.015	33

* Data are in picograms of DNA.
† Number of branches.
[a] Species-rich lineage; [b] species-poor lineage.

In order to contrast the expected amount of genome size change depending on whether divergence in genome size is proportional to elapsed time or to speciation episodes, we computed the mean absolute genome size change per branch in both species-rich and species-poor lineages. This was accomplished simply by summing the branch lengths in each lineage shown in Figures 2 and 3 and dividing by the number of branches. Our logic is that if genome size divergence is primarily a function of time, the absolute genome size change per branch should be *less* in a species-rich lineage since each branch represents less time, on average, relative to branches in the appropriate species-poor lineage. Alternatively, if genome size divergence is primarily a function of the number of speciation episodes, the absolute genome size change per branch should be the same in appropriate species-rich and species-poor lineages provided the same quantum of genome size change occurs per speciation episode in each lineage. At present, we have no way of knowing if the latter is true, although it is reasonable to expect that genome size change at speciation would not be less in a species-rich lineage if genome size change occurs primarily during speciation. We chose this approach over others (e.g., Mindell et al., 1989) for reasons to be detailed fully in a subsequent paper. In brief, the advantages to this approach are that (i) the total amount of genome size divergence per lineage remains the same regardless of how changes are partitioned (optimized) along branches, and (ii) the parameter (i.e., absolute genome size change per branch) is independent of when speciation events occur within a lineage.

Estimates of absolute genome size divergence per branch in the two species-rich versus species-poor lineage comparisons are shown in Table 4. In the *Pimephales-Opsopoeodus* clade, the estimated genome size change per branch is 0.046 $\pm$ 0.026 picograms of DNA, whereas in the *Cyprinella* clade, the estimated change per branch is 0.069 $\pm$ 0.030 picograms of DNA. This constitutes a 1.50-fold difference and suggests that genome size change does not occur primarily as a function of elapsed time, and moreover, that changes may be occurring disproportionally in the *Cyprinella* clade. In the *Notropis* clade, the estimated change per branch is 0.093 $\pm$ 0.033 picograms of DNA, whereas the estimated change per branch in the *Lythrurus*

et al. clade is 0.061 ± 0.015 picograms of DNA. This constitutes a 1.52-fold difference and suggests again that genome size change does not occur primarily as a function of elapsed time and that changes may be occurring disproportionally in the species-rich clade. It should be noted that this disproportionality does not change even if four branches are added to the *Notropis* clade in order to produce a fully resolved, bifurcating tree, i.e., the estimated change per branch in the *Notropis* clade would be 0.078 picograms of DNA constituting a 1.28-fold difference in genome size divergence between the two clades.

The above findings suggest that divergence in genome size, at least in cyprinids, may be concentrated in or associated with speciation episodes. This interpretation is contingent on at least three assumptions. The first is that the phylogenetic hypotheses used in our analysis are robust. Work on cyprinid phylogeny is currently ongoing in several laboratories and some modifications of Mayden's (1989) hypotheses may occur in the future. The second assumption is that the relative number of speciation events for the comparison lineages are representative of the actual number, i.e., that rates of extinction are randomly distributed and proportional to rates of speciation. However, as noted elsewhere (Gold, 1980; Mindell et al., 1989), estimated rates of speciation and extinction appear to be highly correlated among vertebrates, at least to the extent that both speciation and extinction are similarly affected by biogeographic and climatic factors. The last assumption is that genome size changes are essentially neutral to natural selection and occur only in relation to time or speciation. As discussed by Avise (1977), the time or speciation hypotheses (with regard to genetic or genomic change) represent extremes with a wide range of intermediate possibilities. One of these is that species DNA values may be determined, at least in part, by selection acting indirectly on genome size as proposed by several authors (Bennett, 1971, 1972; Szarski, 1983; Cavalier-Smith, 1985a, b; Sessions and Larson, 1987). Whether or not selection directly affects genome sizes, however, has yet to be empirically demonstrated.

SUMMARY

Our studies to date on genome size variation in cyprinids and centrarchids have shown that the patterns and magnitude of genome size evolution within populations of species are essentially the same in both families. Genome sizes at this hierarchical level are normally distributed, suggesting that changes in DNA quantity are small in amount, involve both gains and losses of DNA, and are cumulative and independent in effect. No large-scale or "quantized" change in genome size appears to occur, and the average maximum variation in genome size among individuals within populations is about 5% of the genome. The latter is of interest since this quantity of DNA that apparently is free to vary quantitatively among individuals within populations represents only a small part of the fraction of the genome that has no known coding function. Collectively, the data indicate that the mechanism(s) and potential for changes in genome size among individuals within populations of species in both families is essentially the same.

Alternatively, there is evidence, at least in cyprinids, that divergence in genome size may be concentrated in or associated with speciation events. This hypothesis is consistent with our findings that the majority of genome size divergence in cyprinids occurs above the level of individuals within populations and that there are no

apparent associations between species genome sizes and various life history characteristics. However, the evidence is essentially correlative, and it will be difficult to determine empirically whether the correlation is one of cause and effect. Moreover, it is difficult to conceive exactly how a genome size change might precipitate a speciation episode, particularly if differences in genome size have little to no detectable effect on organismal phenotypes. Without question, future work on genome size variation will need to address this latter issue.

ACKNOWLEDGMENTS

A number of individuals have provided assistance in several ways throughout the course of our studies on genome size and have been acknowledged in our published papers. We would like to again thank the following individuals for their continued support and help: Chris Amemiya, Tom Buchanan, Mike Howell, Bill Matthews, and Bob Stiles. We also acknowledge the use of facilities at the Oklahoma University Biological Station near Lake Texoma and at the Department of Biology, Samford University, Birmingham, Alabama. The scanning microdensitometer used in the research was made available by H. J. Price of the Soil and Crop Sciences Department at Texas A&M University. The chicken blood used as an internal standard was provided by S. A. Naqi of the Texas A&M College of Veterinary Medicine. Finally, we thank Tim Schmidt, Rick Mayden, and two anonymous reviewers for constructive criticism of the manuscript, and Dave Swofford and Jim Carpenter for advice and assistance with optimization methods. The work has been supported by projects H-6187 and H-6703 of the Texas Agricultural Experiment Station and by National Science Foundation grant BSR-8415428 and its renewal.

LITERATURE CITED

Amemiya, C. T., and J. R. Gold. 1990. Chromosomal NOR phenotypes of seven species of North American Cyprinidae, with comments on cytosystematic relationships of the *Notropis volucellus* species-group, *Opsopoeodus emiliae*, and the genus *Pteronotropis*. Copeia 1990:68–78.

Avise, J. C. 1977. Is evolution gradual or rectangular? Evidence from living fishes. Proc. Natl. Acad. Sci. USA 74:5083–5087.

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

Avise, J. C., and F. J. Ayala. 1975. Genetic change and rates of cladogenesis. Genetics 81:757–773.

Avise, J. C., and F. J. Ayala. 1976. Genetic differentiation in speciose versus depauperate phylads: Evidence from the California minnows. Evolution 30:46–58.

Bachmann, K., O. B. Goin, and C. J. Goin. 1972. Nuclear DNA amounts in vertebrates. Brookhaven Symp. Biol. 23:419–450.

Bennett, M. 1971. The duration of meiosis. Proc. R. Soc. London Ser. B 178:277–299.

Bennett, M. 1972. Nuclear DNA content and minimum mitotic time in herbaceous plants. Proc. R. Soc. London Ser. B 181:109–135.

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. How selfish is DNA? Nature 285:617–618.

Cavalier-Smith, T. 1985a. Introduction: The evolutionary significance of genome size, p. 1–36. *In*: The evolution of genome size. T. Cavalier- Smith (ed). John Wiley and Sons, New York, New York.

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

Cavender, T. M., and M. M. Coburn. 1986. Cladistic analysis of eastern North American Cyprinidae. Abst. Ohio J. Sci. 86:1.

Doolittle, W. F., and F. Sapienza. 1980. Selfish genes, the phenotype paradigm and genome evolution. Nature 284:617–618.

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

Eldredge, N., and S. J. Gould. 1972. Punctuated equilibria: An alternative to phyletic gradualism, p. 82–155. *In*: Models in paleobiology. T. J. M. Schopf (ed). Freeman, Cooper and Co., San Francisco, California.

Farris, J. S. 1970. Methods for computing Wagner trees. Syst. Zool. 19:83–92.

Gold, J. R. 1980. Chromosomal change and rectangular evolution in North American cyprinid fishes. Genet. Res. 35:157–164.

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

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

Gold, J. R., C. J. Ragland, and L. J. Schliesing. 1990. Genome size variation and evolution in North American cyprinid fishes. Genetique, Selection, Evolution 22:11–29.

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

Hinegardner, R., and D. Rosen. 1972. Cellular DNA content and the evolution of teleostean fishes. Amer. Nat. 106:621–644.

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

Kauffman, S. 1971. Gene regulation networks: A theory for their global structure and behaviors, p. 145–182. *In*: Current topics in developmental biology. A. A. Moscona and A. Monroy (eds.). Academic Press, New York, New York.

Larson, A. 1981. A reevaluation of the relationship between genome size and genetic variation. Amer. Nat. 118:119–125.

Larson, A. 1984. Neontological inferences of evolutionary pattern and process in the salamander family Plethodontidae. Evol. Biol. 17:119– 217.

Lee, D. S., C. R. Gilbert, C. H. Hocutt, R. E. Jenkins, D. E. McAllister, and J. R. Stauffer, Jr. 1980. Atlas of North American freshwater fishes. North Carolina State Museum of Natural History, Raleigh, North Carolina.

Mabee, P. M. 1987. Phylogenetic change and ontogenetic interpretation in the family Centrarchidae (Perciformes:Centrarchidae). Unpubl. Ph.D. dissert., Duke University, Durham, North Carolina.

Mayden, R. L. 1986. Speciose and depauperate phylads and tests of punctuated and gradual evolution: Fact or artifact? Syst. Zool. 35:591–602.

Mayden, R. L. 1987. Historical ecology and North American highland fishes: A research program in community ecology, p. 211–222. *In*: Community and evolutionary ecology of North American stream fishes. W. J. Matthews and D. C. Heins (eds.). University of Oklahoma Press, Norman, Oklahoma.

Mayden, R. L. 1989. Phylogenetic studies of North American minnows, with emphasis on the genus *Cyprinella* (Teleostei:Cypriniformes). Misc. Publ. Mus. Nat. Hist. Univ. Kansas 80:1–189.

Mayden, R. L., and R. H. Matson. 1988. Evolutionary relationships of eastern North American cyprinids: An allozyme perspective. Abst. Amer. Soc. Ichthyol. Herpetol., Ann Arbor, Michigan.

Mettler, L. E., and T. G. Gregg. 1969. Population genetics and evolution. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Mindell, D. P., J. W. Sites, Jr., and D. Graur. 1989. Speciational evolution: A phylogenetic test with allozymes in *Sceloporus* (Reptilia). Cladistics 5:49–61.

Morescalchi, A. 1977. Phylogenetic aspects of karyological evidence, p. 149–167. *In*: Major patterns in vertebrate evolution. M. K. Hecht, P. C. Goody, and B. M. Hecht (eds.). Plenum Press, New York, New York.

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

Orgel, L. E., and F. H. C. Crick. 1980. Selfish DNA: The ultimate parasite. Nature 284:645–646.

Parker, E. D., Jr., and M. Krietman. 1982. On the relationship between genome size and genetic variation. Amer. Nat. 119:749–752.

Pierce, B. A., and J. B. Mitton. 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. 1988. DNA content variation among higher plants. Ann. Missouri Bot. Gard. 75:1248–1257.

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

Ragland, C. J. 1986. Nuclear DNA content variation in the North American sunfish (Pisces, Centrarchidae). Unpubl. M. S. thesis, Texas A&M University, College Station, Texas.

Ragland, C. J., and J. R. Gold. 1989. Genome size variation in the North American sunfish genus *Lepomis* (Pisces:Centrarchidae). Genet. Res. 53:173–182.

Sarich, V. M., and A. C. Wilson. 1967. Rates of albumin evolution in primates. Proc. Natl. Acad. Sci. USA 58:142–148.

SAS Institute, Inc. 1982. SAS User's Guide: Statistics. Cary, North Carolina.

Sessions, S. K., and A. Larson. 1987. Developmental correlates of genome size in plethodontid salamanders and their implications for genome evolution. Evolution 41:1239–1251.

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

Sneath, P. H. A., and R. R. Sokal. 1973. Numerical taxonomy. W. H. Freeman & Co., San Francisco, California.

Sokal, R. R., and F. J. Rohlf. 1969. Biometry. W. H. Freeman and Co., San Francisco, California.

Sparrow, A. H., H. J. Price, and A. G. Underbrink. 1972. A survey of DNA content per cell and per chromosome of prokaryotic and eukaryotic organisms: Some evolutionary considerations. Bookhaven Symp. Biol. 23:451–494.

Stebbins. G. L. 1966. Processes of organic evolution. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Swift, C. C. 1970. A review of the eastern North American fishes of the *Notropis texanus* species group (subgenus *Alburnops*), with a definition of the subgenus *Hydrophlox* and materials for a revision of the subgenus *Alburnops*. Unpubl. Ph.D. dissert., Florida State University, Tallahassee, Florida.

Szarski, H. 1983. Cell size and the concept of wasteful and frugal evolutionary strategies. J. Theor. Biol. 105:201–209.

Wiley, E. O. 1981. Phylogenetics. The theory and practice of phylogenetic systematics. John Wiley and Sons, New York, New York.

Wilson, A. C., S. S. Carlson, and T. J. White. 1977. Biochemical evolution. An. Rev. Biochem. 46:573–639.