Genome Size Evolution in New Zealand Triplefin Fishes

A. J. R. HICKEY AND K. D. Clements

From the School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland, New Zealand.

Address correspondence to A. J. R. Hickey at the address above, or e-mail: a.hickey@auckland.ac.nz.

Abstract

The genome sizes of 18 species of New Zealand triplefin fishes (family Tripterygiidae) were determined by flow cytometry of erythrocytes. The evolutionary relationships of these species were examined with a molecular phylogeny derived from DNA sequence data based on 1771 base pairs from fragments of three mitochondrial loci (12S and 16S ribosomal RNA, and the control region) and one nuclear locus (ETS2). Haploid genome sizes ranged from .85 pg (1C) to 1.28 pg with a mean of 1.15 ± .01 pg. Genome size appeared to be highly plastic, with up to 20% variation occurring within genera and a 50% difference in size between the smallest and the largest genome. No evidence was found to indicate polyploidy as a mechanism for speciation in New Zealand triplefins. Factors suggested to influence genome sizes of other organisms, such as morphological complexity, neoteny, and longevity, do not appear to be associated with shifts in the genome sizes of New Zealand triplefins.

Introduction

Genome size varies greatly among organisms, even among closely related species (Cavalier-Smith 1985a, b), with reported variation estimates of 80,000-fold across eukaryotes (Gregory and Hebert 1999). Even among vertebrates genome size ranges from .35 pg (1C) in the smooth puffer fish Tetraodon fluviatilis (Lamatsch et al. 2000) to 142 pg in the marbled lungfish Protopterus aethiopicus (Pedersen 1971). No broad correlation is apparent between genome size and overall organismal complexity (the C-value paradox/enigma, Gregory 2001a), as clearly illustrated by the amoeba Amoebae dubia having a 200-fold larger genome than humans (Cavalier-Smith 1985a).

Positive and negative correlations have been observed between genome size and several ecophysiological traits. Rates of cellular division, differentiation, development, and nuclear division (Cavalier-Smith 1985a; Olmo 1983; Van’t Hof 1974) and the thermostability of plants (Macgillivray and Grime 1995) are examples that appear to correlate positively with genome size. Although positive correlations between genome size and longevity have been reported in birds (Monaghan and Metcalf 2000) and fishes (Griffith et al. 2003), these relationships have since been refuted (Gregory 2004).

Several studies concerning fish genome size have focused on broad correlations with nucleotypic effects (Gregory 2001b; Hardie and Hebert 2003), discussed potential associations between rapid genome size changes and speciation (Gold et al. 1990; Ragland and Gold 1989), and searched for trends in genome size evolution (Brainerd et al. 2001; Cano et al. 1982; Carvalho et al. 1998; Griffith et al. 2003; Hinegardner 1968, 1976; Hinegardner and Rosen 1972). Although insightful, these studies have either made comparisons across families or have lacked a phylogenetic framework that incorporates measures of genetic divergence. If genome size data were analyzed within a phylogenetic framework incorporating distance data, trends or potential processes of genome evolution may be better illuminated, particularly if closely related organisms are targeted.

New Zealand triplefin fishes are a prime group for comparative study. Triplefins belong to the suborder Blennioidei, which includes six families: Tripterygiidae (triplefins), Dactyloscopidae (sand stargazers), Labrisomidae (labrisomids), Chaenopsidae (pike-, tube-, and flagblennies), Clinidae (kelpfishes), and Blenniidae (combtooth blennies). New Zealand triplefins are unusually diverse in the context of both the family Tripterygiidae and the New Zealand fish fauna as a whole. The New Zealand triplefin fauna with 26 species assigned to 13 genera comprises about a fifth of the species and a third of the genera in the family (Fricke 1994). All of these species are endemic to New Zealand, although three have been accidentally introduced to Australian waters.
With the exception of two genera, New Zealand triplefins appear to make up a monophyletic radiation (Eyton 1999), with a level of endemism ~10 times that of the entire New Zealand marine fish fauna (Paulin and Roberts 1992). We examine the relationship between genome size and phylogeny with the view that a high level of variation in genome size among closely related species may suggest the involvement of genome size changes in speciation.

**Methods and Materials**

**Sampling and Sequencing Protocol**

A total of 18 species of triplefins were collected from a variety of habitats from coastal and offshore island sites in northern New Zealand (34.08' S, 172.08' E to 36.10' S, 174.45' E). Habitats sampled included estuaries, rockpools, coastal reefs and the offshore Three Kings Islands. Several capture methods were employed, including bait traps, barrier nets, trap nets, set nets, longlines, gillnets, and slurb guns. Fish were anaesthetized with clove oil (Griffiths 2000), and blood was removed by cardiac puncture using 1.7% phosphate buffered saline (NaCl PBS) containing 0.2 M KCl. Blood samples were washed in 1 ml 1.7% NaCl PBS and then pelleted by centrifugation at 3000 × g for 1 min, and the supernatant discarded. Erythrocytes were washed a further two times in 1.7% NaCl PBS before resuspension in 200 μl 1.7% NaCl PBS and fixed in 100% ice-cold methanol and stored at −20°C for at least 16 h. Fixed cells were rehydrated and equilibrated for 5 min in 1.7% NaCl PBS and allowed to equilibrate for 5 min. Rehydrated cells were further washed in 1.7% PBS as outlined. Cell numbers were estimated for each sample by hemocytometer, and an aliquot of hydrated suspended cells was removed and diluted to ~1 × 10^6 cells ml^−1 for each sample. Diluted samples were then centrifuged at 3000 × g, aspirated, resuspended in 200 μl 1.7% PBS, and incubated with 20 μl 5 mg/ml RNase A for 30 min at 20°C. Cells were again centrifuged at 3000 × g and resuspended in 1 ml 1.7% PBS with 1 M EDTA and stained with 20 μg ml^−1 propidium iodide. Chicken erythrocytes were used as internal controls and were prepared as above except with 0.8% NaCl PBS.

Flow cytometry was performed using a Becton/Dickson FSCAN with 10,000 events (cells) counted for each specimen. Internal controls were run in conjunction with a representative of each species. Values for DNA content were calibrated to values reported in the literature for the domestic chicken (1.25 pg; Vinogradov 1998).

A molecular phylogeny of 19 New Zealand triplefin species was constructed based on mitochondrial gene fragments from the control region (D-Loop) and the ribosomal subunits 12S and 16S and a nuclear locus from the putative oncogene ETS2. Typical PCR reactions for the control region, 12S and ETS2 contained ~50 ng of DNA template, 25 mM Tris–HCl, pH 8.3, 2.5 mM MgCl₂, 20 μM premixed deoxynucleotides, 0.5 U hot-start Ampli-Taq gold polymerase (Applied Biosystems), and 1 μM of each oligonucleotide primer, with a final volume of 25 μl. Standard Ampli-Taq polymerase (Applied Biosystems) was used for 16S amplifications. Amplification profiles for the control region and 12S loci involved an initial denaturation step at 94°C for 10 min, followed by 40 cycles at 92°C for 30 sec, annealing at 50°C for 1 min, and extension at 74°C, with a final extension at 74°C for 5 min. Amplification of 16S was performed using the same profile although with an initial denaturation of only 1 min. ETS2 was amplified as for the control region and 12S, although with a 60°C annealing temperature. The same primers used to amplify D-Loop (Meyer et al. 1994), 12S (Kocher et al. 1989), 16S (Palumbi et al. 1991), and ETS2 (Lyons et al. 1997) were used for PCR and sequencing.

Primers were enzymatically removed from PCR products by incubation at 37°C for 30 min with 20 U ml^−1 of exonuclease 1 (Amersham) and nucleotides phosphorylated by addition of 10 U ml^−1 shrimp alkaline phosphatase (Amersham). Exonuclease 1 and shrimp alkaline phosphatase activities were then stopped by incubation at 80°C for 15 min. Automated sequencing of PCR products was performed on Applied Biosystems 377 and 3100 DNA Sequencing Systems (Foster City, CA) using versions I and III ABI Prism Dye Terminator Cycle Sequencing Ready Reaction Kits, with Ampli-Taq DNA polymerase (Applied Biosystems).

Dye incorporation was performed as follows: 40 cycles of 96°C denaturation for 10 sec, 55°C annealing for 30 sec, with an extension at 60°C for 1 min. Incorporated products were precipitated with 100 mM sodium acetate (pH 5.2) in 80% ethanol for 1 h and pelleted by centrifugation at 3500 rpm for 45 min. Pellets were washed twice with 70% ethanol and allowed to air-dry.

Sequences were initially aligned and edited using Sequencher 4.1 (Gene Codes Corporation 2000) aligned with Bioedit and Clustal W and further aligned manually using MacClade (version 4.0). A NEXUS format file with the complete alignment of consensus sequences is available from the corresponding author.

**Tree Construction and Distance Determinations**

Maximum parsimony (MP), maximum likelihood (ML), and neighbor joining (NJ) analyses were employed to construct a phylogeny using a beta version of PAUP* 4.0b4a (Swofford 1998). The New Zealand endemic triplefin species Notolichthys compressus was used as an outgroup. Heuristic searches to find the most parsimonious tree(s) were performed using tree bisection-reconnection (TBR) branch swapping. All sites were equally weighted and gaps treated as missing characters. Branch support under MP and ML models was measured using the nonparametric bootstrap (Felsenstein 1985a). Bootstrap analyses used heuristic searches (TBR branch swapping) with 1000 and 200 replicates for MP and ML, respectively. ML analysis used a GTR substitution with an invariant sites model determined using Modeltest (Posada and Crandall 1998).
All loci amplified for all species except the mitochondrial control region (mtCR) in Notoclinops yaldwyni and ETS2 in Bellapiscis medius. As missing sequence data for N. yaldwyni (mtCR) and B. medius (ETS2) impacts genetic distance estimates between species when using all four loci, plots of pairwise divergence of 12S and 16S were plotted against pairwise divergences of the four loci, excluding N. yaldwyni and B. medius. Corrected pairwise distances were estimated for these two species following linear regression. Genetic distances were calculated under a Kimura 2-parameter model with a \( \gamma \) of .5.

Comparisons between genetic distance and genome size data sets were performed to search for evolutionary trends in genome size. The corrected distances outlined were used as measures of genetic divergence from the root species (N. compressus) and were regressed against genome size.

**Phylogenetic Signal and Evolutionary Trends**

Phylogenetic signal was inferred by testing whether the variance of genome size in an independent contrast analysis (ICA) was significantly lower than the variance of randomly permuted data (10,000 permutations). This test assumes that quantitative traits of closely related taxa will tend to resemble each other due to phylogeny (phylogenetic signal) and therefore will show a variance lower than randomly permuted data (Felsenstein 1985b; Harvey and Pagel 1991). An additional test was performed that contrasted pairwise genetic distances against the pairwise differences in genome size among species. Due to nonindependence of these pairwise measures, the relationship between these two distance measures was tested using a Mantel test (Mantel 1967). Tests for phylogenetic signal were conducted using the Phenotypic Diversity Analysis Program (PDAP; Garland et al. 1992, 1999; Garland and Ives 2000, http://cnas.ucr.edu/~bio/faculty/Garland/PDAP.html) and the subset packages PDTREE, PDRANDOM (Lapointe and Garland 2001) and PDERROR (Diaz-Uriarte and Garland 1996).

**Statistical Analyses**

Mean genome sizes were calculated where possible and compared under a general linear model (GLM) performed using the software package SPSS 7.5. Analyses were performed at the generic and species level where data were sufficient and Tukey’s honestly significant difference with Spjotvoll/Stoline tests used for between-species post hoc comparisons.

**Results**

Genome sizes of New Zealand triplefins fell within the typical size range of teleosts (Figure 1) and formed four Tukey’s groups (a, b, c, and d) with respective C-value means 1.02, 1.18, 1.20, and 1.28 pg (Figure 2). Tukey’s group a contained small genomes found in four of the eight genera examined. The smallest genome estimated was in Blennodon dorsale (.85 pg), although this species could not be assigned to any Tukey’s group due to small sample size (n = 1). The genome of B. dorsale is ~ 16.3%, 27.7%, 30%, and 33.6% smaller than the mean genome sizes of Tukey’s groups a, b, c, and d, respectively. Intraspecific variation in genome size ranged from .04% to 3%, with a mean variance of 1.01%.

To place this study in a rigorous phylogenetic framework, ML and MP analyses were conducted that resulted in similar topologies (Figure 3). ML and MP analyses both resulted in a single optimal topology (ML –In likelihood .9832; MP CI .567, RI .571). Two minor topological differences involving the placement of Matanui profundum and B. dorsale, and Grahamina capito and G. gymnota, were found between the ML/MP topology and the NJ topology (Figure 2).

A linear regression of species divergence from the root species N. compressus against genome size resulted in a highly significant relationship (Figure 4, \( P = .005 \)). However, as bivariate distributions were not normally distributed, non-parametric tests were performed. These also resulted in a significant relationship (Spearman’s rho \(-.59, P = .01\), Kendall’s tau \(-.44, P = .011,\) two-tailed). This relationship suggests DNA loss with increased distance from the root species.

Phylogenetic signal was evident among the data, as the variance of genome size under independent contrast analysis (.077, d.f. 17) was significantly smaller than the variance of permuted data (.196 ± .011, 99.9% confidence level), implying that closely related species had similar-sized genomes. In addition, a significant positive relationship was found using a Mantel test (Figure 5, log C-value divergence, \( Z = -15.53, r = .21, P <= .035, \) log-log \( Z = 176.02, r = .20,\) one-sided \( P <= .045\)), indicating that genome size diverged with genetic divergence providing further support for phylogenetic signal. No relationship between maximum length of species (standard length, SL, measured from snout to end of the vertebral column) and genome size was evident under phylogenetically independent contrast analysis (\( r^2 = .015, P < .59,\) d.f. 17).
**Discussion**

New Zealand triplefin C-values sit within the typical range observed for most teleosts (Figure 1). New Zealand triplefin genome sizes fit within the upper range of other blennioid species (.53–1.19 pg; Cano et al. 1982), which have a mean C-value of .79 pg. The only other reported genome size measurement for a triplefin is the largemouth triplefin, *U. xenogrammus*, which has a genome size of .82 pg (C-value = .81 pg; Hardie and Hebert 2003). Values for blennies and *U. xenogrammus* are ~ 30% smaller than the mean New Zealand triplefin genome size (1.15 pg), indicating that New Zealand triplefins have larger genomes relative to other blennies and triplefins sampled.

Intraspecific variation in triplefin genomes is also within the range of values reported for other species, with a 3% maximum variance in *Ruanoho whero* and 1% mean variance overall. Both intraspecific variation and measurement error may account for the observed variation. Extensive intraspecific studies of cyprinids, centarchids, and salmonids have found genome size variance of 4–13.5%, 6%, and 5.6%, respectively (C-value .35–.5 pg; Gold and Amemiya 1987; Gold et al. 1990; Hinegardner 1976; Ragland and Gold 1989), suggesting that fish genome sizes may vary substantially within species.

It has been suggested that genome size changes are associated with speciation episodes in fish (Brainerd et al. 2001). Therefore, if genome size played a role in speciation in New Zealand triplefins, one would expect variation among closely related taxa. Although some sister taxa (e.g., *Ruanoho*) show significant differences in genome size (Figure 2), few large changes in genome size are apparent between the majority of closely related species. Furthermore, the lower variance levels detected in phylogenetic contrast C-values (relative to randomized data) and a Mantel test (Figure 5) showed that closely related taxa have similar genome sizes.

Although this study does not allow a conclusive test, the quantum changes that would be consistent with polyploidy are not apparent in New Zealand triplefins. Moreover, though polyploidy is most apparent among cyprinids and salmoniformes, it does not appear to be common among...
perciformes (Gregory 2001c). Although our data suggest that gross changes in genome size have not been directly involved with speciation in New Zealand triplefin species, some substantial genome size changes among closely related taxa are evident. Across the 18 New Zealand triplefin species (Figure 2) there is an approximate variation of 50% between the smallest and largest genomes. More dramatically, an approximate 20% change in genome size has occurred between *G. gymnota* and all other taxa within the Forsterygion/Grahamina clade (Figure 2). This contrasts with a 12% difference between the more genetically divergent *Bellapiscis* species. Such large changes in genome size between sister taxa are not unusual among vertebrates, as a broad comparison of genome sizes across families found a high level of variance is not uncommon between closely related species (Gregory 2001b).

Other fish groups also show considerable variation in genome size. For example, puffer fish (family Tetraodontidae) genomes are only marginally smaller than the average teleost genome, although two puffer fish subfamilies, Tetraodontinae and Canthigasterinae, have extremely small genomes (Brainerd et al. 2001). Furthermore, within Tetraodontinae, genome size varies by 30% (Brainerd et al. 2001). Cano et al. (1982) reported genome size differences among eight *Blennius* species (Blenniidae) that ranged over 100% and found no correlation between genome size and chromosome number, but they did not report the phylogenetic relationships of taxa examined. It is therefore apparent that significant changes in genome size among closely related fish can occur and investigations of triplefin karyotypes are warranted.

Although there is evidence for phylogenetic signal, triplefin genome sizes appear to have fluctuated with no direct trend in size evolution. At least three Tukey’s groups are represented in each of the four hypothetical clades (Figures 3 and 4). This pattern differs from puffer fish genomes, which appear to show a progressive loss of DNA in more derived species.

Morphological complexity and neoteny have been suggested to influence, or have been influenced by genome size. Developmental “complexity” has been suggested to be associated with genome size, where small genomes may be typical of morphologically complex species (e.g., seahorses) and permit rapid cell division and differentiation rates (Gregory 2001b). However, New Zealand triplefins are morphologically similar despite differences in physiology (Hickey and Clements 2003), ecology (Clements 2003, table 1), and genome size (this study). We also note that similar morphological complexity does not correlate with genome sizes in puffer fishes (Brainerd et al. 2001) or blennies (Cano et al. 1982).

Neoteny in salamanders may have resulted in or driven increases in genome size following the removal of developmental constraints of metamorphosis (Jockusch 1997;
Raff 1996). The New Zealand triplefin fauna includes a paedomorphic species, \textit{Obliquichthys maryannae} (Hickey and Clements, 2003), but this species has a similar genome size to normally developing congener (Figure 2). Clearly, paedomorphy is driven by another mechanism in \textit{O. maryannae}.

Our data suggest that mechanisms such as polyploidy have not played a major role in the New Zealand triplefin speciation and radiation, because variation in genome size is greatest among more distant species, and is generally similar between closely related species. However, significant differences in genome size between some closely related taxa (e.g., \textit{Ruanoho} and \textit{Grabinina}) are apparent, which raises the question whether processes such as chromosomal rearrangements (see Rieseberg 2001) have played a role in at least some speciation events.

**Acknowledgments**

We thank Brady Doak and Murray Birch for field assistance as skippers of the \textit{R.V. Proteus}; Marie Hickey, Jacob Wedding, and Lisa Clements for the collection of animals; and Scott Baker, Howard Chao, Paul Rainey, Allen Rodrigo, Shane Lavery, and Brian Murray for helpful discussions. We also thank Tegel for allowing collection of chicken blood and the faculty of Medical and Health Sciences, University of Auckland, for use of the flow cytometer. This study met ethical requirements of the University of Auckland and was supported by the Marsden Fund of the New Zealand Royal Society.

**References**


Gregory TR, 2001b. The bigger the C-Value, the larger the cell: genome size and red blood cell size in vertebrates. Blood Cells Mol Dis 27:830–843.


Received November 17, 2004
Accepted January 31, 2005

Corresponding Editor: William Modi