3-1994

The Utility of SATA Satellite DNA Sequences for Inferring Phylogenetic Relationships an1ong the Three Major Genera of Tilapiine Cichlid Fishes

Jens P. C. Franck

Irv Kornfield

Jonathan M. Wright

MOLECULAR PHYLOGENETICS AND EVOLUTION
Vol. 3, No. 1, March, pp. 10–16, 1994

The Utility of SATA Satellite DNA Sequences for Inferring Phylogenetic Relationships among the Three Major Genera of Tilapiine Cichlid Fishes

JENS P. C. FRANCK,*,[1] IRV KORNFIELD,† AND JONATHAN M. WRIGHT*,[2]

*Department of Biology, Dalhousie University, Halifax, Nova Scotia, Canada, B3H 4J1; and †Department of Zoology and Center for Marine Studies, University of Maine, 5751 Murray Hall, Orono, Maine 04469-5751

Received June 21, 1993; revised November 3, 1993

The SATA satellite DNA family of sequences, composed of three size variants of approximately 237, 230, and 209 bp, is conserved in the genomes of tilapiine and haplochromine cichlid fishes. In the present study we examined the utility of the SATA sequences for inferring phylogenetic relationships among the three major genera of tilapiine fishes, *Oreochromis, Sarotherodon,* and *Tilapia.* Hybridization of the monomer SATA repeat to genomic DNA of representative cichlid species established conservation of the sequence in the African tilapiine and haplochromine lineages and its absence from other cichlid lineages. Bootstrapped DNA parsimony and neighbor-joining analyses of derived consensus sequences revealed two distinct clades, one containing the mouthbrooding genera *Oreochromis* and *Sarotherodon,* and the other containing the substrate spawning genus *Tilapia.* These results are consistent with recent independent studies using mitochondrial DNA and establish the utility of the SATA satellite DNA family for phylogenetic reconstruction. Concerted evolution of the SATA sequences was also demonstrated within the tilapiine tribe. © 1994 Academic Press, Inc.

INTRODUCTION

Cichlids are the predominant percoid fishes of the tropical regions of the world. Members of the cichlid family are broadly distributed in the Old World (Africa) and New World (Neotropical; Central and South America), with a few additional taxa in the Indian subcontinent (Fryer and Iles, 1972). The startling explosive radiation of the cichlids, especially in the rift lakes of eastern Africa, has resulted in taxa with a wide diversity of ecology, morphology, and behavior (Stiassny, 1991). Two major tribes of the African cichlid assemblage are the tilapiine and haplochromine fishes (Regan, 1920). Taxonomic classification of tilapias has relied primarily upon morphological characters (Stiassny, 1991) and reproductive behavior (Trewavas, 1982, 1983). Trewavas identified three major tilapiine genera as maternal mouthbrooders (*Oreochromis*), biparental and paternal mouthbrooders (*Sarotherodon*), and substrate spawners (*Tilapia*).

Two models have been proposed for the evolution of brooding behaviors in tilapias. Peter and Berns (1982) argued that the mouthbrooding strains periodically diverged from the substrate spawning lineage, in which the most ancient lineage is represented by the maternal mouthbrooders while the more recently diverged species are represented by the biparental mouthbrooders. Trewavas (1982, 1983), however, proposed that the mouthbrooding lineage originated from a *Tilapia*-like substrate spawner which subsequently split into the maternal and biparental mouthbrooding lines. Allozyme analyses of the major tilapiine genera suggested a close relationship of the mouthbrooding *Oreochromis* and *Sarotherodon* genera as distinct from the substrate spawning *Tilapia* genus (Kronfield *et al.*, 1979; McAndrew and Majumdar, 1984). Mitochondrial DNA (mtDNA) restriction analysis resolved the trichotomy among these three tilapiine genera, and supported Trewavas' hypothesis (Seyoum, 1989; S. Seyoum and I. Kornfield, unpublished). Restriction endonuclease analysis of mtDNAs has been used to clarify the taxonomy of the *Oreochromis* subspecies complex (Kornfield, 1991; Seyoum and Kornfield, 1992). In addition, two recent mtDNA studies employed sequence information to elucidate the relationships of the haplochromine cichlids of Africa (Meyer *et al.*, 1990; Sturmbauer and Meyer, 1992).

Satellite DNAs have been extensively studied in many organisms including invertebrates (Miklos, 1982, 1985), amphibians (Hummel *et al.*, 1984), mammals (Singer, 1982; Arnason *et al.*, 1984), and plants (Flavell *et al.*, 1983). Satellite DNAs in tilapiine fishes were first observed as intensely staining bands of ap-

[1] Present address: Department of Organismal Biology and Anatomy, University of Chicago, 1025 East 57th Street, Chicago, Illinois 60637

[2] To whom correspondence should be addressed.

1055-7903/94 $6.00

proximately 200 bp in *Eco*Ri- and *Hae*III-digested genomic DNA of *Oreochromis mossambicus* × *Oreochromis hornorum* hybrids by agarose gel electrophoresis (Wright, 1989). The consensus sequence is 237 bp long and the satellite constitutes 7% of the haploid genome. Subsequently, the homologous sequences were cloned and characterized from representatives of *Oreochromis, Sarotherodon,* and *Tilapia* (Franck *et al.,* 1992). The SATA satellite DNA family includes three major size variants of approximately 237 bp (type I), 230 bp (type II), and 209 bp (type III).

Recent studies have demonstrated the utility of satellite DNA sequences for inferring phylogenetic relationships. The resolving power of satellite DNA sequences in systematic studies ranges from the identification of conspecific populations of the Sheepshead Minnow (Turner *et al.,* 1991) to interfamilial relationships of cetaceans (Arnason *et al.,* 1992; Gretarsdottir and Arnason, 1992). In the present study we investigated the potential utility of the SATA satellite DNA family for inferring phylogenetic relationships among the major genera of the tilapiine cichlid fishes. The results supported the close relationship of the mouthbrooding genera *Oreochromis* and *Sarotherodon* in a clade distinct from the substrate spawning *Tilapia* genus consistent with the monophyletic model of mouthbrooding evolution proposed by Trewavas (1982, 1983). Additional analyses established the concerted evolution of the SATA sequences.

MATERIALS AND METHODS

Sources of Fish Specimens and Genomic DNA

All tilapiine specimens with the exception of *O. niloticus,* which was maintained at Dalhousie University, were obtained from Dr. Brendan McAndrew at the Institute of Aquaculture, University of Stirling, Scotland. All other cichlid specimens were purchased from a local tropical fish store. Total genomic DNA was extracted either from the caudal peduncle tissue or from blood samples collected by caudal vein puncture of anesthetized fish. The extraction of DNA from blood has been described previously (Wright, 1989). For DNA extraction from tissue, approximately 1 mg of muscle from the caudal peduncle region was homogenized in 500 μl of extraction buffer (0.1 M Tris–OH, pH 8.0, 0.1 M EDTA, 0.25 M NaCl, 0.1% sodium dodecyl sulfate (SDS)). The homogenate was extracted once with one volume of TE-saturated phenol (TE = 10 mM Tris–HCl, 1 mM EDTA, pH 8.0), followed by extraction with phenol/chloroform. The DNA was precipitated by the addition of 1/10 vol of 5 M ammonium acetate and 1 vol of isopropanol. The DNA was dried and resuspended in TE buffer to a concentration of approximately 1 mg/ml.

Cloning Methodology and Sequencing

The cloning methodology for the tilapiine satellite monomer repeats has been previously described (Franck *et al.,* 1992). The *Protomelas similis* SATA sequence was cloned from a *Hinf*I digest of genomic DNA. The approximate 230-bp intensely staining band from the digest was eluted from the gel using a silica bead matrix (Geneclean II, Bio. 101). The eluted DNA was incubated with Mung Bean nuclease, and the blunt end was ligated into the *Sma*I site of M13mp18. Phage lysates were immobilized on a nylon membrane (Hybond-N, Amersham) with a slot blot apparatus and hybridized to the radiolabeled SATA monomer repeat, OplI-5, from *Oreochromis placidus.* Five of the positively hybridizing clones were sequenced and used to construct a consensus sequence. Single-stranded recombinant M13 templates were sequenced by the chain-terminating method (Sanger *et al.,* 1977) using α-35S-dATP (1000 Ci/mmol) with T7 DNA polymerase (Pharmacia).

Radiolabeling Techniques and Hybridization Conditions

The OplI-5 monomer repeat was labeled either by random priming (Feinberg and Vogelstein, 1983) or by Klenow filling of the recessed 5′ ends with [α-^{32}P]dATP (300 Ci/mmol). DNA was routinely labeled to a specific activity of 10^8 cpm/μg. Nylon membranes were incubated for at least 2 h in a prehybridization mixture (50% formamide; 5× SSPE; 1× Denhardt's solution; 100 μg/ml yeast tRNA; 0.1% SDS) (1× SSPE = 0.015 M NaCl, 10 mmol sodium dihydrogenorthophosphate, 1 mmol EDTA) at 42°C. Radiolabeled probe was added to a final concentration of 10^6 cpm/ml and hybridization was allowed to proceed for 24 to 48 h. Membranes were washed under low stringency conditions (0.2× SSC, 0.1% SDS) at room temperature (~20°C). Membranes were exposed to Kodak XAR5 X-ray film. To estimate the copy number of the repetitive DNAs, digitized images of the slot blot autoradiographs were quantified using the program ScanAnalysis Densitometry for the Macintosh (Release 2.20; Burcham, 1987).

Sequence Analyses

DNA sequences were aligned using the multiple sequence-alignment program CLUSTAL (Higgins and Sharp, 1988) for both the pairwise and multiple alignments. The phylogenetic signal present in the satellite sequences was assayed by examining tree-length distribution (Hillis, 1991). All possible unrooted trees were generated for the aligned 246-bp data set of the seven satellite DNA sequences using PAUP, Version 3.1 (Swofford, 1993). The significance of the skewness (g_1) of the distribution of tree lengths was calculated with Biom-pc (Rohlf, 1987). Percentage sequence simi-

larity was calculated on a matches/length basis with large deletions being counted as a single mutational event. Phylogenetic trees were constructed using parsimony with PAUP, Version 3.1 (Swofford, 1993) and by neighbor-joining (NEIGHBOR) Version 3.41 in PHYLIP (Felsenstein, 1991) using the Kimura two-parameter model (Kimura, 1980). Neighbor-joining analysis was also performed on a multiple alignment of the 30 individual cloned SATA sequences (data not shown). The parsimony tree generated by PAUP was analyzed using the program MacClade, Version 3.0 (Maddison and Maddison, 1992).

The spread and fixation of the SATA sequences were analyzed using the method of Strachan *et al.* (1985). The variation is classified into six classes ranging from homogeneity to complete heterogeneity after the alignment of all clones from two species.

Class 1: Completely homogeneous positions in all clones of both species in pairwise comparisons (species A: N_1, species B: N_1; where N = G, A, T, or C).

Class 2: The minority of clones have a new mutation (N_2) at a position, whereas the majority of clones remain homogeneous for the ancestor base (species A: N1 only, species B: $N_1 > N_2$).

Class 3: Positions where the ancestor bases and the mutations are in equal frequencies (species A: N_1 only, species B: $N_1 = N_2$).

Class 4: Positions where one species is homogeneous for the ancestor base but a mutation has replaced this base in the majority of clones in the other species (species A: N_1 only, species B: $N_2 > N_1$).

Class 5: Positions where the two species are homogeneous for different bases. This class represents the classical observation of concerted evolution (Strachan *et al.*, 1985) (species A: N_1 only, species B: N_2 only).

Class 6: Situations including all subsequent mutations (species A: N_1 only, species B: $N_2 > N_3$).

RESULTS

To determine the copy number of the SATA repeat and establish its distribution among species, a radiolabeled satellite DNA insert (OplI-5) was hybridized to a slot blot with graded quantities of denatured genomic DNA from various fish species, as well as OplI-5 RF DNA. Genomic DNA was used from fish of three tilapiine genera, *Oreochromis, Sarotherodon,* and *Tilapia,* as well as representatives from both Old World and New World cichlids (data not shown). Cichlid species other than the tilapiines represented on the blot include three haplochromine fishes, *Protomelas similis, Haplochromis moori,* and *Melanochromis auratus,* two West African cichlids, the hemichromine *Hemichromis bimaculatus,* and a member of the chromidotilapiine tribe, *Pelvicachromis pulcher,* an Asian cichlid, *Etroplus maculatus,* and the South American cichlid, *Cichlasoma meeki.* Genomic DNA samples were also included from rainbow trout (*Oncorhynchus mykiss*) and haddock (*Melanogrammus aegelfinus*). The SATA sequence was detected in tilapiine samples as well as the three haplochromine species, but was absent from all other cichlid species examined. The copy number of the repeat estimated from digitized autoradiographs is much lower in the haplochromines (Table 1). This result may indicate a major amplification of the SATA array after the divergence of the tilapiine fishes from the haplochromine lineage. Alternatively, the estimates for the low copy number of the SATA repeat in the haplochromine genomes may be attributed to a lower sequence identity between the tilapiine SATA repeats and haplochromine repeats. The hybridization and wash conditions for the slot blot should prevent hybridization of the tilapiine SATA probe to sequences with less than ~75–80% sequence identity; calculations for copy number are therefore conservative estimates. To establish the basis for the reduced signal intensity, the SATA sequence was determined for one of the haplochromine species, *Protomelas similis.* The monomer repeat was isolated from *Hinf*I-digested genomic DNA and cloned into the *Sma*I site of M13mp18. Five independent-cloned monomer repeats were used to construct a consensus sequence of 230 bp in length. The consensus length is shorter than the ~237-bp tilapiine SATA type I size variant due to the cloning methodology which used mung bean nuclease to blunt end the insert DNA before cloning. This resulted in the digestion of the overhanging single-stranded DNA

TABLE 1

SATA Copy Number Estimation for Selected Cichlid Species

Species	Copies/haploid genome	Genome equivalent (%)
Oreochromis hornorum	3.6×10^4	1.0
Oreochromis niloticus	5.4×10^4	1.5
Oreochromis placidus	2.6×10^4	0.7
Oreochromis mossambicus	3.9×10^4	1.1
Oreochromis andersonii	4.9×10^4	1.4
Oreochromis mortimeri	4.4×10^4	1.2
Oreochromis aureus	5.6×10^4	1.6
Sarotherodon galilaeus	5.9×10^4	1.6
Tilapia rendalli	3.9×10^4	1.1
Tilapia zillii	7.4×10^3	0.2
Tilapia mariae	1.6×10^4	0.4
Tilapia tholloni	2.8×10^4	0.8
Protomelas similis	5.1×10^3	0.1
Haplochromis moori	1.1×10^3	0.03
Melanochromis auratus	1.2×10^3	0.03

Note. Copy number estimation of SATA repeat based on densitometry of slot blot hybridization experiment (data not shown). Copy number is estimated based on a haploid genome content of 1.0 pg (Majumdar and McAndrew, 1986).

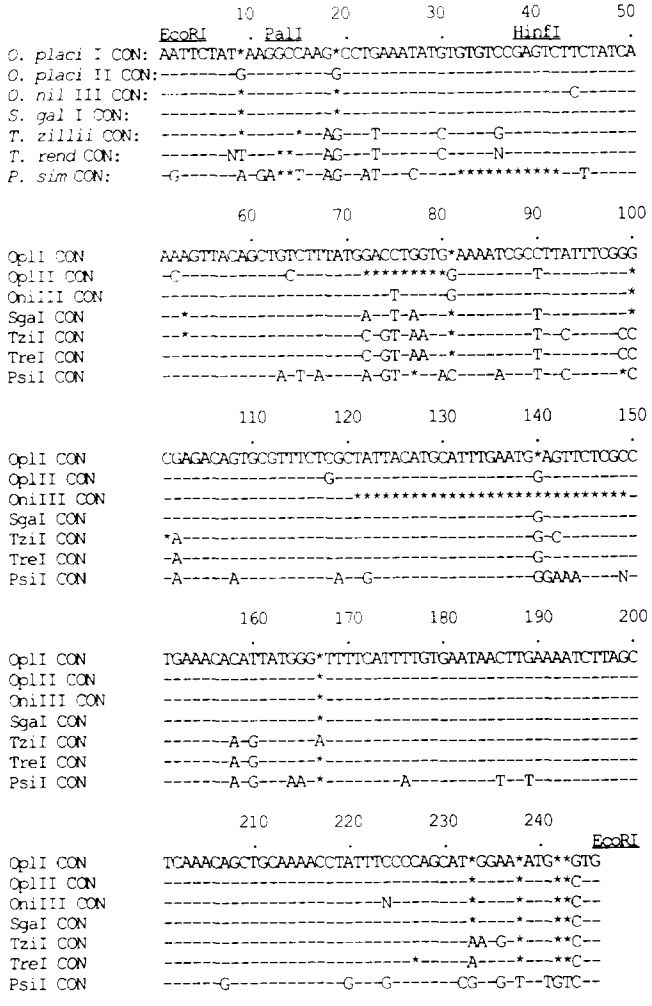

FIG. 1. Multiple alignment of the six tilapiine and *P. similis* SATA consensus sequences generated by the CLUSTAL program (Higgins and Sharp, 1988). The sequences were aligned in relation to the *Eco*RI restriction endonuclease recognition site. Dashes (-) indicate sequence identity with the above sequence and asterisks (*) indicate deletions or gaps that were introduced to maintain the alignment. Nonconsensus bases are represented by N.

and the subsequent truncation by six base pairs of the SATA repeat.

The seven consensus sequences for the tilapiine and *P. similis* SATA repeats were aligned using the CLUSTAL multiple alignment program (Fig. 1) (Higgins and Sharp, 1988); sequence similarity was calculated as the proportion of matches with the two major deletions counted as single mismatches (Table 2). The *Oreochromis* and *Sarotherodon* SATA sequences have the highest sequence similarity, with values of approximately 96%. Both *Oreochromis* and *Sarotherodon* sequences show near identical similarities to *Tilapia* SATA sequences with values of approximately 90%. The *P. similis* consensus sequence shows lower sequence identity, ranging from 73.5% for *O. placidus* type II to 80.1% for the *T. zillii* SATA sequence.

We analyzed the spread and fixation of species-diagnostic nucleotide differences in satellite DNAs for the SATA family. The comparison was made between the individual cloned monomer repeats from *O. niloticus* and the six other satellite DNA sequences. All available cloned sequences were aligned and each nucleotide position was classified according to the scheme of Strachan *et al.* (1985). In this scheme, six classes describe the transition stages of monomer repeats from homogeneity at a given nucleotide position to partial and full fixation for a nucleotide difference in one species (described under Materials and Methods). The majority of positions are homogeneous for all clones (class 1) ranging from 59.7% for the *O. niloticus* vs *P. similis* comparison, to 79.1% for the *O. niloticus* vs *O. placidus* (type I) comparison (Table 3). Class 5 nucleotide differences, which indicate complete fixation for different bases at an aligned position in each species, are highest in frequency for the *O. niloticus* vs *P. similis* comparison (14.2%) and lowest for the intrageneric *O. niloticus* vs *O. placidus* comparison (0.5%). Class 5 represents the classic case of concerted evolution for a satellite DNA sequence within a species (Dover, 1982).

The phylogenetic relationships of the six tilapiine SATA DNA consensus sequences were analyzed using the haplochromine *P. similis* sequence as the designated outgroup. The multiple-aligned sequences generated by CLUSTAL were used as the input for the boostrapped DNA parsimony and neighbor joining. Tree lengths for all possible 945 unrooted dichotomous trees ranged from 17 to 35 (mean 30.6). The distribution of tree lengths was highly skewed, $g_1 = -0.875$, $P < 0.001$. For boostrapped parsimony analysis, the data set was resampled with replacement 2000 times to generate accurate confidence intervals (Hedges, 1992) on the major nodes of the tree. The resultant tree clearly identifies two distinct clades with the two

TABLE 2

Comparison of Tilapiine and *Protomelas similis* SATA Consensus Sequences

	Percentage sequence similarity						
	O. placidus I	*O. placidus* II	*O. niloticus* III	*S. galilaeus* I	*T. rendallii* I	*T. zillii* I	*P. similis* I
O. placidus I	—	97.0	97.1	95.8	89.5	89.2	78.3
O. placidus II	—	—	96.0	96.1	90.9	89.5	73.5
O. niloticus III	—	—	—	96.2	88.1	88.2	78.6
S. galilaeus I	—	—	—	—	90.8	92.0	77.4
T. rendalli I	—	—	—	—	—	94.2	80.0
T. zillii I	—	—	—	—	—	—	80.1
%GC	40.5	40.7	41.1	39.0	36.3	38.2	34.9
Size (bp)	237	231	209	236	237	238	230

TABLE 3

Distribution of Transitional Classes (%) at Individual Nucleotide Positions in Pairwise Comparisons between the SATA Sequence from *Oreochromis niloticus* and SATA Sequences from Five Other Cichlid Species[a]

	Oreochromis niloticus vs				
Class	*Oreochromis placidus*[b]	*Sarotherodon galilaeus*	*Tilapia rendalli*	*Tilapia zillii*	*Protomelas similis*
1	79.1	79.0	61.5	57.2	59.7
2	14.0	15.4	14.7	20.5	15.6
3	—	0.9	—	2.8	—
4	—	0.5	3.9	0.9	3.8
5	0.5	1.9	4.8	3.7	14.2
6	0.9	0.5	3.9	2.8	3.8

[a] Classification follows the definition of Strachan *et al.* (1985).
[b] *Oreochromis placidus* size-variant I.

Tilapia satellite consensus sequences in one clade and the *Oreochromis* and *Sarotherodon* sequences united in the other (Fig. 2); the confidence interval on this node is 100%. The resolution between the *Sarotherodon* and *Oreochromis* sequences remains ambiguous. The unrooted parsimony analysis was repeated on the data set with the corresponding deleted regions of the *O. placidus* type II and *O. niloticus* type II size variants being removed from each sequence. Statistical confidence in the clade containing the *Tilapia* species decreased considerably. Additional analyses using neighbor-joining methodology (Saitou and Nei, 1987) produced a single tree with a topology identical to the one produced by the parsimony analysis (data not shown). Neighbor-joining analyses were also performed on the individual cloned SATA sequences that were used to derive the seven consensus sequences

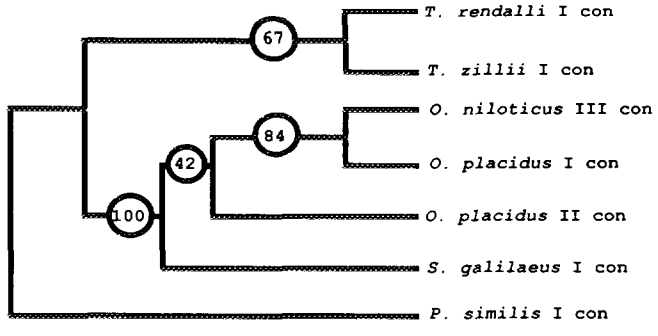

FIG. 2. Phylogenetic tree of tilapiine cichlids based on SATA consensus sequences. The *P. similis* SATA consensus sequence was used as a designated outgroup to root the tree for the six tilapiine satellite sequences. The most parsimonious trees were generated using the exhaustive search option of PAUP (Swofford, 1993). The bootstrap values as pecentage of 2000 replicates are indicated on the tree. The most parsimonious trees differed in the placement of *O. placidus* II, which is reflected in the low bootstrap value (42%) for the clade containing the sequences of *Oreochromis*. Consistency index = 0.889, retention index = 0.900.

(data not shown). The overall topology of the tree remained the same as that generated from the consensus sequences and with the exception of one of the *O. placidus* type I size variant clones (OplI 2), the cloned sequences from the same species clustered together. Minor changes in the multiple alignment did not alter the overall topology of the tree.

DISCUSSION

Previous phylogenetic analyses (Kornfield *et al.*, 1979; McAndrew and Majumdar, 1984; Sodusk and McAndrew, 1991) have suggested the close relationship of the mouthbrooding tilapiine genera *Oreochromis* and *Sarotherodon* to the exclusion of the substrate spawning *Tilapia* genus. In the present study this relationship is further supported by the phylogenetic analyses of the SATA satellite DNA family.

Examination of the distribution of tree lengths has been proposed as a method to detect the presence of phylogenetic information within comparative sequence data sets (Hillis, 1991; Huelsenbeck, 1991). In particular, significant skewness is absent from simulated random sequence, but may be present in real sequence data. The highly significant skewness produced with the SATA DNA sequences clearly indicates the presence of phylogenetic information in this data. It is noteworthy that this is an independent test of the phylogenetic utility of these satellite sequences and does not depend on tree topology.

Two different tree building methodologies were used to test the phylogenetic utility of the SATA sequences. Both the DNA parsimony and neighbor-joining methods resulted in similar trees. The *Sarotherodon* and *Oreochromis* genera were grouped together in a clade distinct from the *Tilapia* SATA sequence with complete confidence. When *Sarotherodon* was forced to become a member of the clade containing the two species of *Tilapia*, the resultant tree was five steps longer than the most parsimonious tree. The results of the SATA phylogenetic analyses thus support the monophyletic origin of the mouthbrooding tilapiines proposed by Trewavas (1982). It was not possible to resolve the position of *Sarotherodon* with respect to the *Oreochromis* SATA sequences with confidence. Though the node connecting these two genera had a bootstrap value of 42%, recent simulations of bootstrapping (Hillis and Bull, 1993) suggest that such values are extremely conservative; the true value is probably greater than 50%. The inability to resolve the *Sarotherodon* and *Oreochromis* mouthbrooding tilapiine genera suggests at least three possibilities: their divergence has occurred too recently to be detected in the evolution of the SATA sequences, the rate of evolution within the clade has been slow, or there is ambiguity in the SATA sequences of these two taxa. Comparative studies using allozymes (McAndrew and Majumdar, 1984) and

mtDNA (Seyoum, 1989) produced estimates of divergence between *Oreochromis* and *Sarotherodon* that are almost as great as those between these two genera and *Tilapia*. It is thus clear that there has probably been sufficient time for evolution in this clade. It is more difficult to evaluate the idea that there has been slow sequence evolution within the *Oreochromis–Sarotherodon* clade. Examination of the distribution of mutations among lineages (Table 3) does not suggest any rate heterogeneity between the two major clades. It is more probable that our inability to distinguish between *Oreochromis* and *Sarotherodon* lineages with high confidence is that the SATA sequence of *Sarotherodon* contains less information because of the presence of autapomorphies and convergent bases. Greater confidence might result when additional sequence becomes available for other species in this genus.

Chromosomal studies of Old World cichlids demonstrate a highly conservative level of karyotypic evolution (Kornfield *et al.*, 1979; Majumdar and McAndrew, 1984). The most distinctive difference in chromosomal banding patterns is in the distribution of constitutive heterochromatin detected by C-banding. These studies detected differences between the *T. zillii* karyotype and the karyotypes of species in the *Oreochromis* and *Sarotherodon* genera. Ten to twelve of the chromosomes in the *T. zillii* karyotype ($2n = 44$) did not stain at all for constitutive heterochromatin, and those chromosomes that did stain possessed different banding patterns from the mouthbrooding species; all of the chromosomes from the mouthbrooding species possessed constitutive heterochromatin. While these studies did not implicate karyotype evolution as a factor in tilapiine diversification, they did indicate underlying qualitative differences in the repetitive DNA component of the genomes of the mouthbrooding and substrate spawning assemblages. In the present study, the molecular characterization of the constitutive heterochromatin, of which satellite DNAs are a major component, corroborates the physical observations of the karyotype studies. The SATA sequences of the mouthbrooding species from the *Oreochromis* and *Sarotherodon* genera showed a high level of sequence identity to each other, but both have significantly diverged from the *Tilapia* SATA sequences. This result, coupled with the karyotype observations, suggests a potential molecular basis for genetic isolation of the substrate spawning and mouthbrooding lineages.

Pairwise comparison of the individual cloned sequences from different species revealed the species-specific fixation of nucleotides in the alignments, providing evidence of concerted evolutionary processes. The frequency of these species-specific fixed nucleotide positions ranged from 0.5% for the intrageneric *O. niloticus* (type III) vs *O. placidus* (type I) comparison to 14.2% for the intergeneric comparison of *O. niloticus* (type III) to the *H. similis* SATA sequence (Table 3). The frequency of the concerted nucleotide positions therefore increases with increasing phylogenetic distance. Species diagnostic nucleotide differences, which are homogeneous for each member of a repetitive DNA family, represent the classical observation of concerted evolution (Strachan *et al.*, 1985). The sequence divergence observed between the mouthbrooding and substrate spawning species may therefore serve as an effective mechanism for maintaining the genetic isolation between these two major lineages.

ACKNOWLEDGMENTS

We thank B. J. McAndrew for providing the tilapiine samples. This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada to J.M.W. Data analysis was supported, in part, by NSF Grant EHR 91-08766 to the Maine Science and Technology Commission.

REFERENCES

Arnason, U., Hoglund, M., and Widegren, B. (1984). Conservation of highly repetitive DNA in cetaceans. *Chromosoma* **89,** 238–242.

Arnason, U., Gretarsdottir, S., and Widegren, B. (1992). Mysticete (baleen whale) relationships based upon the sequence of the common cetacean DNA satellite. *Mol. Biol. Evol.* **9,** 1018–1028.

Burcham, T. (1987). "Scan Analysis: Densitometry for the Macintosh (Release 2.20)," Specom Research, December 1987, published by BIOSOFT.

Dover, G. A. (1982). Molecular drive: A cohesive mode of species evolution. *Nature* **299,** 111–117.

Feinberg, A. P., and Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132,** 6–13.

Felsenstein, J. (1991). "PHYLIP (Phylogeny Inference Package)" Version 3.4, Department of Genetics, University of Washington, Seattle.

Flavell, R., Jones, J., Lonsdale, D., and O'Dell, M. (1983). Higher plant genome structure and the dynamics of genome evolution. *In* "Advances in Gene Technology" (K. Downey, R. W. Voellmy, F. Ahmed, and J. Schultz, Eds.), pp. 47–59, Academic Press, New York.

Franck, J. P. C., Wright, J. M., and McAndrew, B. J. (1992). Genetic variability in a family of satellite DNAs from tilapia (Pisces: Cichlidae). *Genome* **35,** 719–725.

Fryer, G., and Iles, T. D. (1972). "The Cichlid Fishes of the Great Lakes of Africa: Their Biology and Evolution," Oliver and Boyd, Edinburgh.

Gretarsdottir, S., and Arnason, U. (1992). Evolution of the common cetacean highly repetitive DNA component and the systematic position of *Orcaella brevirostris*. *J. Mol. Evol.* **34,** 201–208.

Hedges, S. B. (1992). The number of replicates needed for accurate estimation of the bootstrap p value in phylogenetic studies. *Mol. Biol. Evol.* **9,** 366–369.

Higgins, D. G., and Sharp, P. M. (1988). CLUSTAL: A package for performing multiple sequence alignment on a microcomputer. *Gene* **73,** 237–244.

Hillis, D. L. (1991). Discriminating between phylogenetic signal and random noise in DNA sequences. *In* "Phylogenetic Analysis of DNA Sequences" (M. M. Miyamoto and J. Cracraft, Eds.), pp. 278–294, Oxford Univ. Press, New York.

Hillis, D. M., and Bull, J. J. (1993). An empirical test of bootstrap-

ping as a method for assessing confidence in phylogenetic analysis. *Syst. Zool.*, in press.

Huelsenbeck, J. P. (1991). Tree-length distribution skewness: An indicator of phylogenetic information. *Syst. Zool.* **40,** 257–270.

Hummel, S., Meyerhof, W., Korge, E., and Knochel, W. (1984). Characterization of highly and moderately repetitive 500 bp *Eco*RI fragments from *Xenopus laevis* DNA. *Nucleic Acids Res.* **12,** 4921–4937.

Kimura, M. (1980). A simple model for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16,** 111–120.

Kornfield, I., Ritte, U., Richler, C., and Wahrman, J. (1979). Biochemical and cytological differentiation among cichlid fishes of the Sea of Galilee. *Evolution* 33, 1–14.

Kornfield, I. (1991). Genetics. *In* "Cichlid Fishes: Behaviour, Ecology and Evolution" (M. H. A. Keenleyside, Ed.), pp. 103–128, Chapman and Hall, London.

Maddison, W. P., and Maddison, D. (1992). "MacClade: Analysis of Phylogeny and Character Evolution," Version 3.0 Sinauer, Sunderland, MA.

Majumdar, K. C., and McAndrew, B. J. (1986). Relative DNA content of somatic nuclei and chromosomal studies in three genera, *Tilapia, Sarotherodon,* and *Oreochromis* of the tribe tilapiini (Pisces: Cichlidae). *Genetica* **68,** 175–188.

McAndrew, B. J., and Majumdar, K. C. (1984). Evolutionary relationships within three tilapiine genera (Pisces: Cichlidae). *Zool. J. Linn. Soc.* 80, 421–435.

Meyer, A., Kocher, T. D., Basasibwaki, P., and Wilson, A. C. (1990). Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. *Nature* **347,** 550–553.

Milklos, G. L. (1982). Sequencing and manipulating highly repeated DNA. *In* "Genome Evolution and Phenotypic Variations" (G. A. Dover and R. B. Flavell, Eds.), pp. 41–68, Academic Press, London.

Miklos, G. L. (1985). Localized highly repetitive DNA sequences in vertebrate and invertebrate genomes. *In* "Molecular Evolutionary Genetics" (R. F. MacIntyre, Ed.), pp. 241–321, Plenum, New York.

Peters, Von H. M., and Berns, S. (1982). Die Maulbrutpflege der Cichliden: Untersuchungen zur evolution eines verhaltensmusters. *Z. Zool. Syst. Evolutionsforsch.* **20,** 18–52.

Regan, C. T. (1920). The classification of the fishes of the family Cichlidae. I. The Tanganyika genera. *Ann. Mag. Nat. Hist.* **9,** 33–53.

Rohlf, F. J. (1987). "Biompc," Exeter Publ., Setauket, New York.

Saitou, N., and Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4,** 406–425.

Sanger, F., Nicklen, S., and Coulsen, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.

Seyoum, S. (1989). Stock identification and the evolutionary relationship of the genera *Oreochromis, Sarotherodon,* and *Tilapia* (Pisces: Cichlidae) using allozyme analysis and restriction endonuclease analysis of mitochondrial DNA. PhD thesis, University of Waterloo, Waterloo, Ontario, Canada.

Seyoum, S., and Kornfield, I. (1992). Taxonomic notes on the *Oreochromis niloticus* subspecies-complex (Pisces: Cichlidae), with a description of a new subspecies. *Can. J. Zool.* **70,** 2161–2165.

Singer, M. F. (1982). Highly repeated sequences in mammalian genomes. *Int. Rev. Cytol.* **76,** 67–112.

Sodusk, P., and McAndrew, B. J. (1991). Molecular systematics of three tilapiine genera *Tilapia, Sarotherodon,* and *Oreochromis* using allozyme data. *J. Fish Biol.* **39**(SA), 301–308.

Stiassny, M. L. J. (1991). Phylogenetic intrarelationships of the family Cichlidae: An overview. *In* "Cichlid Fishes: Behaviour, Ecology and Evolution" (M. H. A. Kennleyside, Ed.), pp. 1–35, Chapman and Hall, London.

Strachan, T., Webb, D., and Dover, G. A. (1985). Transition stages of molecular drive in multiple-copy DNA families in *Drosophila. EMBO J.* **4,** 1701–1708.

Sturmbauer, C., and Meyer, A. (1992). Genetic divergence, speciation and morphological stasis in a lineage of African cichlid fishes. *Nature* **358,** 578–581.

Swofford, D. L. (1993). "PAUP: Phylogenetic Analysis Using Parsimony, Version 3.1, Illinois Natural History Survey, Champaign, Illinois.

Trewavas, E. (1982). Generic groupings of tilapiini used in aquaculture. *Aquaculture* **27,** 79–81.

Trewavas, E. (1983). "Tilapiine Fishes of the Genera *Sarotherodon, Oreochromis,* and *Danakilla,*" British Museum (Natural History), London.

Turner, B. J., Elder, J. F., and Laughlin, T. F. (1991). Repetitive DNA sequences and the divergence of fish populations: Some hopeful beginnings. *J. Fish Biol.* **39**(SA), 131–142.

Wright, J. M. (1989). Nucleotide sequence, genomic organization and evolution of a major repetitive DNA family in tilapia (*Oreochromis mossambicus/hornorum). Nucleic Acids Res.* **17,** 5071–5079.