

Evolution of the Ribosomal RNA Internal Transcribed Spacer One (ITS-1) in Cichlid Fishes of the Lake Victoria Region

Gregory C. Booton,* Les Kaufman,† Mark Chandler,‡ Richard Oguto-Ohwayo,§ Wenrui Duan,¶ and Paul A. Fuerst*,¹

*Department of Molecular Genetics and ¶Department of Zoology, The Ohio State University, Columbus, Ohio 43210; †Boston University Marine Program, Department of Biology, Boston, Massachusetts 02215; ‡New England Aquarium, Boston, Massachusetts; and §Ugandan Freshwater Fisheries Research Organization, Jinja, Uganda

Received January 12, 1998; revised May 30, 1998

The nucleotide sequences of the first internal transcribed spacer (ITS-1) of the ribosomal RNA gene cluster have been determined for 11 species of closely related endemic cichlid fishes of the Lake Victoria region (LVR) and 6 related East African cichlids. The ITS-1 sequences confirmed independently derived basal phylogenies, but provide limited insight within this species flock. The line leading to *Pseudocrenilabrus multicolor* arose early, close to the divergence event that separated the tilapiine and haplochromine tribes of the "African Group" of the family Cichlidae. In this phylogeny, *Astatoreochromis alluaudi* and the riverine *Astatotilapia burtoni* are sister taxa, which together are a sister group to a monophyletic assemblage including both Lake Victoria and Lake Edward taxa. The ITS-1 data support the monophyly of haplochromine genera across lakes. Since Lake Victoria is believed to have been dry between 14,500 and 12,400 BPE, the modern assemblage must have been derived from reinvasion by the products of earlier cladogenesis events. Thus, although the regional superflock is monophyletic, the haplochromines of Lake Victoria itself did not evolve *in situ* from a single ancestor. © 1999 Academic Press

Key Words: cichlid; evolution; ITS-1; Lake Victoria; rRNA.

INTRODUCTION

The cichlid species flocks of the Great Lakes of Africa represent some of the most remarkable events of explosive speciation ever documented in vertebrates (Fryer and Iles, 1972; Mayr, 1984; Stiassny, 1991). The endemic haplochromine species flock of Lake Victoria

shows extraordinary morphological and ecological diversity (Greenwood, 1981, 1984a,b, 1991) despite a monophyletic origin postulated as being not older than 225,000 years BPE (Meyer *et al.*, 1990) by mtDNA clock and possibly less than 14,500 years old, if the lake dried out completely as is currently suspected (Livingston, 1980; Stager *et al.*, 1986; Johnson *et al.*, 1996). The detailed nature of the evolutionary relationships among the hundreds of recently divergent species, as well as the evolutionary position of the group with respect to other species of Cichlidae, is an area of active investigation (Kornfield, 1991). The short evolutionary time period for the group makes molecular analysis difficult, since only the most rapidly evolving sequences will provide phylogenetic information. The determination of a robust phylogeny for the endemic taxa of Lake Victoria is the last major obstacle to its use as a powerful instrument to study the mechanics of vertebrate evolution.

Morphological, behavioral, and paleontological data have all been used to estimate phylogenetic relationships among cichlids (Van Couvering, 1982; Dominey, 1984; Lippitsch, 1993). More recently, protein and nucleic acid sequences have been combined with morphological data to investigate cichlid relationships in the African Great Lakes (Sage *et al.*, 1984; Meyer *et al.*, 1990; Sturmbauer and Meyer, 1992, 1993; Ono *et al.*, 1993a,b; Booton, 1995; Sülmann *et al.*, 1995). These studies have demonstrated the difficulty of identifying sequences of use for discrimination among the recently divergent taxa of Lake Victoria.

The choice of a genomic region for molecular analysis depends strongly on the suspected time since divergence of the groups being studied. In the case of the Lake Victoria fish fauna, critical time windows are likely to include 5000 years or less for recent speciation events (peripheral satellite lake differentiation), around 12,000 years for events following reflooding of the lake basin following the most recent desiccation event, and up to several hundred thousand years for initiation of

¹ To whom correspondence should be addressed at Department of Molecular Genetics, The Ohio State University, 484 West 12th Avenue, Columbus, OH 43210. Fax: (614) 292-4466. E-mail: fuerst.1@osu.edu.

the regional fauna and origination of the major component clades (Meyer *et al.*, 1990; Verheyen *et al.*, 1996). To be useful, a genetic region must accumulate changes to yield measurable genetic differentiation with sensitivity geared toward such critical time windows. Mitochondrial DNA (mtDNA) sequences have been used extensively for studies of fish at both intraspecific and interspecific levels because of the rapid evolution of the mitochondrial genome (Brown, 1985; Kocher *et al.*, 1989). However, within Lake Victoria, the analysis of mtDNA sequences, though informative for the very early divergence of this fauna from the ancestral stock, failed to provide insight into the phylogenetic patterns among taxa that subsequently evolved within the species flock (Meyer *et al.*, 1990).

Comparison of the rates of nucleotide substitution for various gene segments (Li *et al.*, 1985; Gillespie, 1986) suggests that sequences such as introns, with reduced coding constraints, may accumulate changes rapidly enough to warrant study. These sequences are among the most rapidly evolving regions of nuclear genomic DNA, thus having the potential to accumulate changes in sequence very soon after populations diverge. Included by analogy in the class of noncoding sequences are the internal transcribed spacer (ITS) sequences located between the ribosomal RNA genes. The target of this study is the primary DNA sequence of the first internal transcribed spacer (ITS-1) of the ribosomal gene cluster, which is located between the 18S and 5.8S ribosomal RNA genes.

METHODS AND MATERIALS

Taxa studied and DNA preparation. Species examined and their geographic range are summarized in Table 1. DNA was extracted from fish muscle tissue using standard phenol/chloroform techniques (Ausebel *et al.*, 1987). DNA was quantified by spectroscopy and quantifications were confirmed by electrophoresis on a 1% agarose gel.

PCR amplification and DNA sequencing. ITS-1 amplification and sequencing primers, as well as PCR conditions, are summarized in the legend of Fig. 1. Pooled PCR products were used directly for cycle sequencing with no further preparation. Sequencing products were electrophoresed on either a 6 or 8% polyacrylamide gel, and scoring was done manually. GenBank accession numbers are presented in Table 1.

Phylogenetic data analysis. The primary sequences from the taxa in this study were aligned using the sequence alignment program EyeBall Sequence Editor (ESEE) for the PC (Cabot and Beckenbach, 1989). This alignment was used to produce data sets which were suitable for analysis in PAUP and MEGA (Swofford, 1990; Kumar *et al.*, 1993). Phylogenetic analysis using distance methods (implementing the Kimura two-parameter nucleotide substitution algorithm; Kimura, 1980) were done using the computer program MEGA. *Oreochromis* species were defined as outgroup taxa. Distances derived by this method were then used to

TABLE 1
ITS 1 Study Taxa

	GenBank accession no.	MCZ Ref. no. ^a	Distribution ^e
<i>Astatoreochromis alluaudi</i> (Pellegrin, 1904)	U67332	14551	LVR
<i>Astatotilapia burtoni</i> (Günther, 1893)	U67333	LK ^c	Riverine hap. from Lake Tanganyika basin
<i>Astatotilapia nubila</i> (Boulenger, 1906)	U67334	14789	Widespread LVR
<i>Gaurochromis angustifrons</i> (Greenwood, 1980)	AF005093	program ^b	LVR
<i>Gaurochromis</i> sp. (Greenwood, 1980) ^d	AF005094	97409	L. Kachira
<i>(Harpagochromis) "kachira deep"</i> (Greenwood, 1980) ^d	U67335	97096	L. Edward
<i>Lipochromis taurinus</i> (Trewavas, 1933)	U67336	97829	L. Victoria
<i>(Neochromis) "kruising"</i> (Regan, 1920) ^d	U67337	program ^b	L. Victoria
<i>Neochromis nigricans</i> (Boulenger, 1906)	U67338	16879, 2 program ^b	L. Victoria
<i>Oreochromis esculentus</i> (Graham, 1928)	U67339	16496	LV Basin
<i>Oreochromis niloticus</i> (Linné, 1757)	U67340	16132	Widespread
<i>(Paralabidochromis) "rock kribensis"</i> (Greenwood, 1956) ^d	U67341	16579	L. Victoria
<i>Pseudocrenalibrus multicolor</i> (Scholler, 1903)	U67342	14716	Widespread, L. Vic. Subspecies
<i>Ptyochromis xenognathus</i> (Greenwood, 1957)	U67343	16260	L. Victoria
<i>Xystichromis phytophagous</i> (Greenwood, 1966)	U67344	14847	L. Victoria
<i>Yssichromis fusiformis</i> (Greenwood & Gee, 1969)	U67345	19698	L. Victoria
<i>Yssichromis laparogramma</i> (Greenwood & Gee, 1969)	U67346	14704	L. Victoria

^a Wild caught specimens are archived at The Museum of Comparative Zoology, Harvard University under these accession numbers.

^b These specimens were obtained from the Lake Victoria Species Survival Plan breeding program and samples are not archived at MCZ.

^c This was a whole specimen identified and provided by L.K. It is not archived at MCZ.

^d Authority given is for genus in those species which have not yet been formally described.

^e LVR, Lake Victoria region taxa.

produce a phylogenetic gene tree using the neighbor-joining algorithm (Saitou and Nei, 1987). Bootstrapping (1000 replicates) was performed in MEGA to examine the consistency of the data. Maximum parsimony analysis was performed using the cladistic analysis program PAUP 3.0, also with *Oreochromis* taxa chosen as outgroups. Identical sequences were excluded from the analyses. In PAUP, gene tree reconstruction was performed using the branch and bound search option and using the MULPARS option, which maintains all minimal length trees. Initial upper bounds were calculated using the stepwise option. Zero branch lengths were collapsed. The addition of taxa order was set to the furthest option. Consistency and homoplasy statistics were calculated and are presented with the maximum parsimony gene tree.

RESULTS

The sequences of ITS-1 were obtained for 2 tilapiine and 15 haplochromine cichlid species (Fig. 1). The aligned ITS-1 region spans 553 bases. Variation in overall size is due to multiple insertion/deletion (indels) events. A total of 103 positions were variable, including 56 nucleotide positions which involved indels and 50 sites at which nucleotide substitutions had occurred (the total is greater than 103 because some sites contained substitutions as well as indel events in different taxa). Three of the sites had multiple substitution events harboring 3 different nucleotides, for a total of 53 nucleotide substitutions. Among the nucleotide substitutions occurring in ITS-1, 28 changes were transitions and 25 were transversions. The transition/transversion ratio for the total data set was 1.12, unexpectedly low for such closely related sequences. The limited degree of variation did have the ancillary benefit of permitting alignment with little ambiguity along the entire ITS-1 among the species studied. The final alignment was used to calculate the corrected proportion of nucleotide substitutions.

As representatives of the subfamily Tilapiinae within the African Group of Stiassny's cichlid phylogeny (Stiassny, 1991; Lippitsch, 1995), the two *Oreochromis* species were chosen as outgroup taxa for these analyses. The calculated distances in most comparisons are small. *P. multicolor* showed the most divergence from the main group of Lake Victoria haplochromine species due to the accumulation of nine unique nucleotide substitutions compared to all other taxa. No differences in the ITS-1 sequences were observed between the two tilapiine species of *Oreochromis* or between three different haplochromine pairs of taxa: *Xystichromis phytophagous* from Lake Kanyaboli (near Victoria), *Ptyochromis xenognathus* (Victoria), *Harpagochromis* "Kachira deep" (Lake Kachira, Kooki Lakes, western flank of Victoria Basin), (*Neochromis*) "kruising" (Victoria), *Gaurochromis angustifrons* (Lakes Edward and George), and

an undescribed *Gaurochromis* species from Lake Edward. The largest distances were found between *Pseudocrenilabrus* and other taxa, ranging from 4.97 to 6.96%. The percentage nucleotide divergence between *Astatoreochromis alluaudi* and the other haplochromine genera range from 4.90% for *Pseudocrenilabrus* to 0.99% for *Astatotilapia burtoni*. Distances between *A. alluaudi* and the rest of the Lake Victoria/Edward/Kivu assemblage of haplochromines is approximately 1.45% on average.

One potential problem that must be considered when examining recently derived taxa is the possibility of incomplete lineage sorting, in which observed variation actually predates the speciation events (Pamilo and Nei, 1988; Bowers *et al.*, 1994; Moran and Kornfield, 1995; Moore, 1995). In this study, however, overall variation at the ITS-1 locus was low, and sequences observed from multiple individuals of selected taxa were identical to each other. Three different individuals of *N. nigricans* were sequenced. All had identical ITS-1 sequences. In addition, two individuals of *Y. fusiformis* were also sequenced and were found to be identical. Further, the two taxa of *Gaurochromis* sequenced were identical and maintained a change unique to this genus. While ancestral polymorphism is still a possibility in these recently derived taxa, data from these experiments did not reveal it. Nonetheless, sample sizes examined here are insufficient to rule out the persistence of ancestral polymorphisms and a larger sampling of this locus would be required to fully address this issue.

Maximum parsimony methods were implemented using PAUP to produce a cladogram for ITS-1 (Fig. 2). The branch and bound search produced one maximum parsimony tree consisting of 89 steps (as did a heuristic search). The distance matrix was used to produce a gene tree using the neighbor-joining (NJ) algorithm in MEGA, with 1000 bootstraps. Bootstrap values greater than 50% obtained from the neighbor-joining gene tree reconstruction are presented on the corresponding nodes of the maximum parsimony tree in Fig. 2.

The primary result of both methods of gene tree reconstruction is that the haplochromine taxa of the Lake Victoria region represent a monophyletic group based on the ITS-1 data. The riverine *Astatotilapia burtoni*, together with *A. alluaudi*, represents a sister group to the Lake Victoria endemics, which include the marginal lacustrine *A. nubilia*. Second, as shown by the multifurcations and low bootstrap support values, these data fail to resolve relationships within the LVR taxa.

DISCUSSION

Sequence divergence in ITS-1 was low among the 17 taxa involved in this study, but provided a few key insights nonetheless. In addition to nucleotide substitutions, a large number of insertion and deletion events

```

1 1111111112 222222223 333333334 444444445 555555556 666666667 777777778
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
Astatoreochromis alluaudi CTGGCTACAC CGAGCGGCC CGCCTGC--- ----TG-TC TCCCT-TTTT GCCGCCGAGG GTCTCCCGCC ACCGTCGCCG
Astatotilapia burtoni .....
(Harpagochromis) "kachira deep" .....G .....- .....
(Neochromis) "kruising" .....G .....- .....
Xystichromis phytophagous .....G .....- .....
Pseudocrenilabrus multicolor .....G .....C .....--
(Paralabidochromis) "rock kribensis" .....G .....- .....
Ptyochromis xenognathus .....G .....- .....
Neochromis nigricans .....G .....- .....
Astatotilapia nubila .....G .....- .....
Lipochromis taurinus .....G .....- .....
Yssichromis lamparogramma .....G .....- .....
Yssichromis fusiformis .....G .....- .....
Gaurochromis angustfrons .....G .....- .....
Gaurochromis sp. ....G .....- .....
Oreochromis niloticus .....GCA CCCGG..T.. ....C..- .....A...
Oreochromis esculentus .....GCA CCCGG..T.. ....C..- .....A...

1 1111111111 1111111111 1111111111 1111111111 1111111111 1111111111
8888888889 9999999990 0000000001 1111111112 222222223 333333334 444444445 555555556
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
Astatoreochromis alluaudi GTGGGGGTAT CCCGAGGTCT TCGGCTCGCG CGTCCCCAC CGGAAGCTCG AGCCTTAGTC TGGGCTGGT CGCCGCCCGG
Astatotilapia burtoni ..... A..... .T..... .....
(Harpagochromis) "kachira deep" .....
(Neochromis) "kruising" .....
Xystichromis phytophagous .....
Pseudocrenilabrus multicolor ---.....CT. ....T.....
(Paralabidochromis) "rock kribensis" .....C.....
Ptyochromis xenognathus .....
Neochromis nigricans .....
Astatotilapia nubila .....
Lipochromis taurinus .....C.....
Yssichromis lamparogramma .....C.....
Yssichromis fusiformis .....C.....
Gaurochromis angustfrons .....C.....
Gaurochromis sp. ....C.....
Oreochromis niloticus .....C.....TG .....- .....A.....
Oreochromis esculentus .....C.....TG .....- .....A.....

```

FIG. 1. ITS-1 primary sequence alignment. PCR amplification of ITS-1 was performed as follows: 100 ng of genomic DNA template, 10 pM of each amplification primer, 2.5 mM MgCl₂, 1.5 mM dNTP's, and 1–2.5 U Taq DNA Polymerase (Gibco BRL Inc., Gaithersburg, MD) were used for each reaction. PCR conditions were as follows: denaturation at 94°C for 1 min, annealing at 50–52°C for 1.5 min, and extension at 72°C for 3 min, 35 cycles. PCR forward amplification primers were: 1262C (gtggtgcatggccgttcta), 1262C-cloning primer (gcggatccgtggtgcatg-cgcgttcta), 1712C (agcgcgagaagacgatcaaa). PCR reverse primers were: ITS-2 (gctgcgttctcatcgacgc) and ITS2-cloning primer (cgaagcttgtagc-

	1111111111	1111111111	1111111111	1111111112	2222222222	2222222222	2222222222	2222222222	2222222222
	6666666667	7777777778	8888888889	9999999990	0000000001	1111111112	2222222223	3333333334	
	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	
<i>Astatoreochromis alluaudi</i>	ACGAACCGAC	GGCCCCGCC	-GCCTCTGCG	CCAAAGCGAG	CCCCTGCC	CGAC-GGCTT	CCTCCGAGGG	CCGACGGAGG	
<i>Astatotilapia burtoni</i>	
(Harpagochromis) "kachira deep"A.....G----	
(Neochromis) "kruising"A.....G----	
<i>Xystichromis phytophagous</i>G----	
<i>Pseudocrenilabrus multicolor</i>A.....T.....	
(Paralabidochromis) "rock kribensis"G----	
<i>Ptyochromis xenognathus</i>G----	
<i>Neochromis nigricans</i>A.....G----	
<i>Astatotilapia nubila</i>T.....G----	
<i>Lipochromis taurinus</i>C.....G----	
<i>Yssichromis lamparogramma</i>C.....G----	
<i>Yssichromis fusiformis</i>C.....G----	
<i>Gaurochromis angustifrons</i>C.....G----	
<i>Gaurochromis sp.</i>C.....G----	
<i>Oreochromis niloticus</i>C.....A.....C.....	
<i>Oreochromis esculentus</i>C.....A.....C.....	
	2222222222	2222222222	2222222222	2222222222	2222222222	2222222223	3333333333	3333333333	
	4444444445	5555555556	6666666667	7777777778	8888888889	9999999990	0000000001	1111111112	
	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	
<i>Astatoreochromis alluaudi</i>	AGAGTAGGGA	CCGTGGCTCG	TTGGAGGCGG	GCGCGGTCC	CCGTCCGCA	ACTACCGTA	CCGCCCCGCC	ACGAGAACCT	
<i>Astatotilapia burtoni</i>CA.....	
(Harpagochromis) "kachira deep"CA.....	
(Neochromis) "kruising"CA.....	
<i>Xystichromis phytophagous</i>CA.....	
<i>Pseudocrenilabrus multicolor</i>	C.....T.....	
(Paralabidochromis) "rock kribensis"CA.....	
<i>Ptyochromis xenognathus</i>CA.....	
<i>Neochromis nigricans</i>CA.....	
<i>Astatotilapia nubila</i>CA.....	
<i>Lipochromis taurinus</i>CA.....	
<i>Yssichromis lamparogramma</i>CA.....T.....	
<i>Yssichromis fusiformis</i>CA.....	
<i>Gaurochromis angustifrons</i>CA.....G.....	
<i>Gaurochromis sp.</i>CA.....G.....	
<i>Oreochromis niloticus</i>CA.....T.....A.....	
<i>Oreochromis esculentus</i>CA.....T.....A.....	

FIG. 1—Continued gctgcttctcatcgacgc). Additional forward ITS-1 sequencing primers were: 1/F (cacaccgccgtcg), SSU2 (gtgaacctgcg-gaaggatca), and 236C-ITS (ggaccgtggctcgttg). An additional reverse ITS-1 sequencing primer was: 538RE-ITS (ttgccacattcgtagacggg). Primers 1262C, 1262C-cloning primer, 1712C, 1/F, and SSU2 are located in the 18S rDNA gene. Primers ITS2 and ITS2-Cl are located in the 5.8s rDNA gene. Primers 236C-ITS and 538RE-ITS are located within the ITS-1 locus.

	3333333333	3333333333	3333333333	3333333333	3333333333	3333333333	3333333333	3333333333	3333333333	3333333333
	2222222223	3333333334	4444444445	5555555556	6666666667	7777777778	8888888889	9999999990		
	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
<i>Astatoreochromis alluaudi</i>	CGACCGAAAG	CGCGGGCTGG	CGGTCTCGCC	-TGGCCGCTG	CCCGCGCGCC	TCCGGGTACC	CAACTCTCCT	CCCTCCTGCG		
<i>Astatotilapia burtoni</i>		
(Harpagochromis) "kachira deep"A.....C.....		
(Neochromis) "kruising"A.....C.....		
<i>Xystichromis phytophagous</i>A.....		
<i>Pseudocrenilabrus multicolor</i>	.A.....A.....	...CG.....		
(Paralabidochromis) "rock kribensis"A.....		
<i>Ptyochromis xenognathus</i>A.....		
<i>Neochromis nigricans</i>A.....		
<i>Astatotilapia nubila</i>A.....		
<i>Lipochromis taurinus</i>A.....	C.....		
<i>Yssichromis lamparogramma</i>A.....	C.....		
<i>Yssichromis fusiformis</i>A.....	C.....		
<i>Gaurochromis angustfrons</i>A.....	C.....		
<i>Gaurochromis sp.</i>A.....	C.....		
<i>Oreochromis niloticus</i>A.....T.....A.....T.....		
<i>Oreochromis esculentus</i>A.....T.....A.....T.....		
	4444444444	4444444444	4444444444	4444444444	4444444444	4444444444	4444444444	4444444444	4444444444	4444444444
	0000000001	1111111112	2222222223	3333333334	4444444445	5555555556	6666666667	7777777778		
	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
<i>Astatoreochromis alluaudi</i>	GAGG-AGCAC	GGGGGGTCA	ATGTCTCCTC	TCCCCC--G	CCTCGG---	----?AGCG	CCCGGGGTTT	TTTTTCCCTT		
<i>Astatotilapia burtoni</i>CC.....GG.....		
(Harpagochromis) "kachira deep"CC.....		
(Neochromis) "kruising"CC.....		
<i>Xystichromis phytophagous</i>CC.....		
<i>Pseudocrenilabrus multicolor</i>A.....-		
(Paralabidochromis) "rock kribensis"CC.....		
<i>Ptyochromis xenognathus</i>CC.....		
<i>Neochromis nigricans</i>CC.....		
<i>Astatotilapia nubila</i>CC.....		
<i>Lipochromis taurinus</i>	...G.....CC.....CGGG	GGAAGG.AA.		
<i>Yssichromis lamparogramma</i>	...G.....CC.....CGGG	GGAAGG.....		
<i>Yssichromis fusiformis</i>	...G.....CC.....CGGG	GGAAGG.....		
<i>Gaurochromis angustfrons</i>	...G.....CC.....CGGG	GGAAGG.....		
<i>Gaurochromis sp.</i>	...G.....CC.....CGGG	GGAAGG.....		
<i>Oreochromis niloticus</i>TGCC.....A...	GGAAGG.....G..-		
<i>Oreochromis esculentus</i>TGCC.....A...	GGAAGG.....G..-		

FIG. 1—Continued

	4444444444	4444444445	5555555555	5555555555	5555555555	5555555555	5555555555	5555555555	555
	8888888889	9999999990	0000000001	1111111112	2222222223	3333333334	4444444445	555	
	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	123	
<i>Astatoreochromis alluaudi</i>	C---AAACC	CTTTTA-CCC	GTCTACGAAT	GTGGCAACCC	ACAGTGAAAC	GAAAA-----	--CAAAAACA	ATA	
<i>Astatotilapia burtoni</i>C.....	
(Harpagochromis) "kachira deep"C.....	
(Neochromis) "kruising"C.....	
<i>Xystichromis phytophagous</i>C.....	
<i>Pseudocrenilabrus multicolor</i>AAC-..AA.G..G	.G.....	..A....C..	
(Paralabidochromis) "rock kribensis"C.....	
<i>Ptyochromis xenognathus</i>C.....	
<i>Neochromis nigricans</i>C.....	
<i>Astatotilapia nubila</i>C.....	
<i>Lipochromis taurinus</i>C.....AG.....	
<i>Yssichromis lamparogramma</i>C.....A.....	
<i>Yssichromis fusiformis</i>C.....A.....	
<i>Gaurochromis angustfrons</i>C.....A.....	
<i>Gaurochromis sp.</i>C.....A.....	
<i>Oreochromis niloticus</i>	.CTTT.....	..C...T...G..A.....	-----	..AA.....	
<i>Oreochromis esculentus</i>	.CTTT.....	..C...T...G..A.....	-----	..AA.....	

FIG. 1—Continued

was observed in ITS-1. These were most prevalent in comparisons between haplochromines and tilapiines and to a lesser extent in the comparisons involving *Pseudocrenilabrus* relative to other taxa. This is similar to other studies which have observed numerous indels in ITS-1 in organisms ranging from salmon to beetles (Pleyte *et al.*, 1992; Vogler and DeSalle, 1994).

In retrospect, the low levels of observed variation at ITS-1 may not have been unexpected, given recent confirmation of the relatively recent origin of the Lake Victoria Basin (estimated at <750,000 years BP from mtDNA sequences; Beadle, 1981), Lake Victoria (<13,000 years BP; Johnson *et al.*, 1996), and the haplochromine radiations in the Lake Victoria region (<225,000 years BP; Meyer *et al.*, 1990; but see also Kaufman *et al.*, 1997). While the low levels of variation may not be surprising now, at the time the ITS-1 study was initiated, older dates for the most recent desiccation event in Lake Victoria were the prevalent hypotheses. This led to the choice of an intron like region (ITS-1) as a potentially informative nuclear locus to examine closely related LVR cichlid phylogenetic relationships. Finally, the hypothesis that the species flock of Lake Victoria should be 10 to 20 times older than the lake itself (Meyer *et al.*, 1990) is a puzzling inconsistency addressed below.

It is particularly noteworthy that taxa drawn from

Lake Edward/George fell within the Victorian taxa. This supports Greenwood's contention that the Lake Victoria radiation is one part of a larger Lake Victoria-Edward/George-Kyoga-Kivu regional fauna (see also Lippitsch, 1995; Kaufman *et al.*, 1997; Kaufman, 1997). It also suggests that Meyer *et al.*'s evidence for monophyly of the Lake Victoria haplochromines applies to the regional haplochromine fauna as a whole, but not to Lake Victoria specifically. This would help explain why the fauna is older than the lake (Victoria), but leaves open the issue of what was going on between the formation of the Victoria basin and the putative onset of radiation based on the mitochondrial clock (Meyer *et al.*, 1990; Kaufman *et al.*, 1997).

One aspect of our results that is most difficult to interpret is a seemingly incongruous recency for certain major divergence events. As stated above, the average distance between *A. alluaudi* and the other LVR haplochromines is 1.45%. A recent study examined this same ITS-1 region in six species of *Salvelinus* (Pleyte *et al.*, 1992). The two most closely related species in that study, *Salvelinus alpinus* and *Salvelinus malma*, had a sequence divergence between them of 0.53%. The authors estimated by other methods that these two taxa diverged approximately 10,000 years ago (Savvaitova, 1980). As long as precursor rRNA processing sites are maintained within the spacer

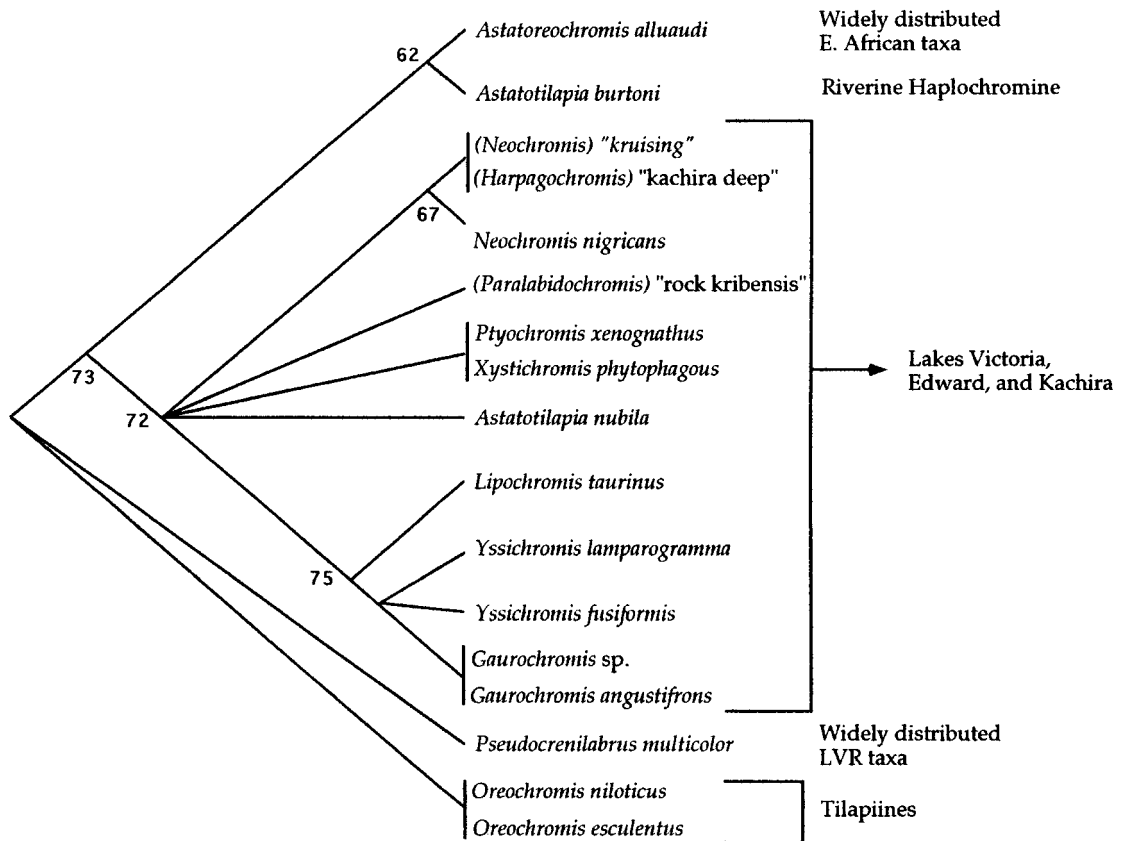


FIG. 2. Strict consensus maximum parsimony ITS-1 gene tree. Tree shown was produced in PAUP (see text). Species range is shown at right of the tree. Search was done using a heuristic search algorithm, random sequence addition, TBR swapping algorithm. This tree is a strict consensus of the 75 maximum parsimony trees which were produced. In the case of identical sequences, one of the replicates was removed prior to the search being done. The tree consists of 58 steps; CI = 0.914, HI = 0.086. Salient results of the neighbor-joining reconstruction are presented here by the addition of the confidence values shown at the nodes on this tree. These numbers represent the bootstrap confidence levels (1000 bootstraps) observed that were greater than 50%.

region, it is not unreasonable to postulate that this region might evolve at similar rates in salmonids and cichlids. The inability to align sequences from divergent fish groups, for instance between cichlids and salmonids, also suggests that there are not many selective constraints on ITS-1 sequences, excluding processing sites. In addition, the G + C content is similar in both cases, 63% in the salmonids and 70% in the cichlids.

While molecular clocks are certainly not universal (Korey, 1981; Li, 1981), the *Salvelinus* clock does provide one method with which to interpret the cichlid ITS-1 data. Based on this presumption, we attempted to apply the *Salvelinus* clock to the LVR cichlid data. The ITS-1 "clock" provided independently by the salmon data suggests a 0.053% sequence divergence/1000 years for the ITS-1. Given the average sequence divergence between *A. alluaudi* and the remainder of the LVR taxa (1.45%), we calculate a time of divergence between these groups of approximately 30,000 years b.p. The general lack of resolution within the LVR taxa precludes any application of this "clock" to comparisons

between individual taxa. However, if we examine the taxa together, we calculate an average divergence within the LVR taxa of 0.53%. Applying the *Salvelinus* ITS-1 "clock," this would correspond to a divergence data for the group of approximately 10,000 years ago. This data is of the same order of magnitude as the recent seismic reflection studies that have concluded that Lake Victoria was completely dry as little as 12,400 years ago (Johnson *et al.*, 1996). But, acceptance of this date from the ITS-1 data creates a quandary.

Since *A. alluaudi* lies outside the riverine haplochromines phylogenetically, the entire LVR fauna (including among them, the *Neochromis*, *Yssichromis*, *Lipochromis*, *Prognathochromis*, and *Xystichromis* examined here) shared among lakes Victoria, Edward/George, Kivu, and Kyoga must, in accordance with this clock estimate, be less than 14,000 years old. However, *Astatoreochromis* also lies outside the Lake Malawi haplochromine species flock, thus requiring that the species flocks of both lakes, in their entireties, must have evolved within the last 30,000 years. This date is clearly in conflict with mitochondrial results that show

Astatoreochromis to be basal to both Victoria and Malawi haplochromine taxa, but with a divergence time of approximately three million years ago based on a mitochondrial clock (Meyer *et al.*, 1990, 1996). Acceptance of such a recent origin would require unreasonable levels of faunal exchange among highly disparate lakes in relatively recent times. Thus, while it appears possible that taxa within the LVR may be of recent origin, the divergence estimate from ITS-1 between *Astatoreochromis* and the remaining taxa using the *Salvelinus*-derived ITS-1 rate is problematic.

It is also possible to conceive an alternative interpretation of the rate of divergence by considering the average rate of nucleotide substitutions in introns. Li *et al.* (1985) estimated the average substitution rate of small introns as 3.7×10^{-9} , similar to the average rate for animal mitochondrial DNA: 5×10^{-9} substitutions per site per year in a lineage (Brown, 1985). The latter corresponds to a 1% divergence per million years. The average rate of substitutions for introns is 74% of the average for the mitochondrial genes. We can use this correspondence to estimate the divergence of the haplochromine taxa. The average corrected sequence divergence between *Astatoreochromis alluaudi* and the remaining LVR haplochromines is 1.45%. This leads to an estimated divergence time between *A. alluaudi* and the remaining haplochromines of ~ 1.95 MYa. This is much more recent than the ~ 3.5 Mya estimated by Meyer *et al.* (1990) using mtDNA, but far greater than that calculated using the salmonid ITS-1 "clock." The discrepancy between the divergence time of *Astatoreochromis* and the haplochromine taxa discussed above remains, although it is not as striking a problem in this case.

Again, we can consider the LVR taxa as a largely unresolved group and use the average substitution rate within the group of 0.53%. Using the rate correspondence as above we would estimate a burst of speciation at approximately 716,000 years ago. This calculated date is within the ballpark of the date estimated for formation of the Lake Victoria basin ($<750,000$ years ago; Beadle, 1981) and older than that estimated by Meyer *et al.* of 225,000 to 250,000 years ago (1990) for the divergence of the Lake Victoria haplochromine flock. It certainly predates the most recent desiccation event in Lake Victoria.

There are several reasons why the interpretation of the earliest divergence dates of ca. 225,000–250,000 years ago for the radiation of Lake Victoria haplochromines should be considered suspect. Lake Victoria was much smaller or totally dry between about 14,000 and 12,400 years ago; yet congeners occur in other lakes. Thus, the modern assemblage is probably derived, at least in part, from reinvasion by the products of earlier cladogenesis events. Thus, although the regional super-flock is monophyletic, the haplochromines of Lake Victoria itself did not evolve *in situ* from a single ancestor. Until more data, and the right data, are

available, the complex and relatively rapid-fire sequence of events that produced the LVR haplochromines shall remain obscured in an evolutionary "fog of war." Several points have emerged from this fog, however. First, a star phylogeny bursting at $\sim 225,000$ BPE does not accurately or insightfully represent the actual sequence of events. More likely, what Meyer *et al.* detected was the trace of one of several successive bursts of cladogenesis tied to landscape changes in the region. Its importance lies in it being one of the earlier such events. The bursts that gave rise, for example, to the endemic genera of lakes Victoria and Edward–George (Kivu and Kyoga now appear also to house endemic genera) must have left a phylogenetic signal that we have yet to detect on a molecular level.

In sum, the ITS-1 sequence data contests the hypothesis that the haplochromines of Lake Victoria evolved *in situ* 225,000 years ago. Furthermore, the molecular clock for the ITS-1 locus as estimated in salmonids does not appear applicable to cichlids. Finally, phylogenetic resolution of the haplochromine species within the various lakes of the region will require additional data and new or combined methodologies.

ACKNOWLEDGMENTS

Thanks are given to Doug Warmolts and the curators of the Johnson Aquatic Complex of the Columbus Zoo. Work was partially supported by DEB-9300065 from the National Science Foundation (to P.A.F. and L.K.), grants from the Columbus Zoo (to P.A.F.), an Ohio State University Graduate School Alumni Research Award (to G.B.), NSF Grant INT-93-08276 (to L.K.), and a Pew Scholarship for Conservation and the Environment (to L.K.) (Pew Charitable Trusts). All material was collected through the efforts of the Lake Victoria Research Team (LVRT), involving authors and researchers of several institutions. We thank the staff of the Ugandan Freshwater Fisheries Research Organization at Jinja (FIRI) and the Kenya Marine and Fisheries Research Institute (KMFRI) for their assistance with the collection of LVR samples.

REFERENCES

- Ausubel, F., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1987). "Current Protocols in Molecular Biology," Wiley, New York.
- Avisé, J. (1994). "Molecular Markers, Natural History, and Evolution," Chapman & Hall, New York.
- Beadle, (1981). "The Inland Waters of Tropical Africa" 2nd ed., Longman, London.
- Booton, G. (1995). "Molecular Genetic Analysis of the Phylogenetic Relationships of Lake Victoria Cichlid Fish," Ph.D. thesis, The Ohio State University, Columbus, Ohio.
- Bowers, N., Stauffer, J. R., and Kocher, T. D. (1994). Intra- and interspecific mitochondrial DNA sequence variation within two species of rock-dwelling cichlids (Teleostei: Cichlidae) from Lake Malawi, Africa. *Mol. Phylogenet. Evol.* 3: 75–82.
- Brown, W. M. (1985). The mitochondrial genome of animals. In "Molecular Evolutionary Genetics" (R. MacIntyre, Ed.), pp. 95–130. Plenum, New York.
- Cabot, E. L., and Beckenbach, A. T. (1989). Simultaneous editing of multiple nucleic and protein sequences with ESEE. *Comput. Appl. Biosci.* 5: 233–234.

- Dominey, W. J. (1984). Effects of sexual selection and life history on speciation: Species flocks in African Cichlids and Hawaiian *Drosophila*. In "Evolution of Fish Species Flocks" (A. A. Echelle and I. Kornfield, Eds.), pp. 231–250. Univ. of Maine Press, Orono, ME.
- Fryer, G., and Iles, T. D. (1972). "The Cichlid Fishes of the Great Lakes of Africa: Their Biology and Evolution," Oliver and Boyd, London.
- Gillespie, J. H. (1986). Rates of molecular evolution. *Annu. Rev. Ecol. Syst.* **17**: 637–665.
- Greenwood, P. H. (1981). "The Haplochromine Fishes of the East African Lakes," Kraus Int. Publ., Munchen.
- Greenwood, P. H. (1984a). African cichlids and evolutionary theories. In "Evolution of Fish Species Flocks" (A. A. Echelle and I. Kornfield, Eds.), pp. 141–154. Univ. of Maine Press, Orono, ME.
- Greenwood, P. H. (1984b). What *is* a species flock? In "Evolution of Fish Species Flocks" (A. A. Echelle and I. Kornfield, Eds.), pp. 13–19. Univ. of Maine Press, Orono, ME.
- Greenwood, P. H. (1991). Speciation. In "Cichlid Fishes: Behaviour, Ecology, and Evolution" (M. H. Keenlyside, Ed.), pp. 86–102. Chapman & Hall, London.
- Johnson, T. C., Scholz, C. A., Talbot, M. R., Kelts, K., Ricketts, R. D., Ngobi, G., Beuning, K., Ssemmanda, I., and McGill, J. W. (1996). Late Pleistocene desiccation of Lake Victoria and rapid evolution of cichlid fishes. *Science* **273**: 1091–1093.
- Kaufman, L. S. (1997). Asynchronous taxon cycles in haplochromine fishes of the greater Lake Victoria region. *S. Afr. J. Sci.*, **93**: (11–12) 601–606.
- Kaufman, L. S., Chapman, C. A., and Chapman, L. J. (1997). Evolution in fast forward: Haplochromine fishes of the Lake Victoria region. *Endeavour* **21**: 23–30.
- Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitution through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**: 111–120.
- Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Paabo, S., Villablanca, F. X., and Wilson, A. C. (1989). Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. *Proc. Natl. Acad. Sci. USA* **86**: 6196–6200.
- Korey, K. A. (1981). Species number, generation length, and the molecular clock. *Evolution* **35**: 139–147.
- Kornfield, I. (1991). Genetics. In "Cichlid Fishes: Behaviour, Ecology, and Evolution" (M. H. Keenlyside, Ed.), pp. 103–128. Chapman & Hall, London.
- Kumar, S., Tamura, K., Nei, M., and The Penn State University (1993). Molecular Evolutionary Genetics Analysis Version 1.02. computer program.
- Li, W-H. (1981). What about the molecular clock hypothesis? *Curr. Opin. Genet. Dev.* **3**: 896–901.
- Li, W-H., Luo, C. C., and Wu, C. I. (1985). Evolution of DNA sequences. In "Molecular Evolutionary Genetics" (R. J. MacIntyre, Ed.), pp. 1–94. Plenum, New York.
- Lippitsch, E. (1990). Scale morphology and squamation patterns in cichlids (Teleostei; Perciformes): A comparative study. *J. Fish Biol.* **37**: 265–291.
- Lippitsch, E. (1993). A phyletic study on lacustrine haplochromine fishes (Perciformes, Cichlidae) of East Africa, based on scale and squamation characters. *J. Fish Biol.* **42**: 903–946.
- Lippitsch, E. (1995). Scale and squamation character polarity and phyletic assessment in the family Cichlidae. *J. Fish Biol.* **47**: 91–106.
- Livingston, D. A. (1980). Environmental changes in the Nile Headwaters. In "The Sahara and the Nile" (M. A. J. Williams and H. Faune, Eds.), pp. 339–359. Balkema, Rotterdam.
- Mayr, E. (1984). Evolution of fish species flocks: A commentary. In "Evolution of Fish Species Flocks" (A. A. Echelle and I. Kornfield, Eds.), pp. 3–11. Univ. of Maine Press, Orono, ME.
- 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.
- Meyer, A., Montero, C. M., and Spreinat, A. (1996). Molecular phylogenetic inferences about the evolutionary history of East African cichlid fish radiations. In "The Limnology, Climatology, and Paleoclimatology of the East African Lakes" (T. C. Johnson and E. Odada, Eds.), pp. 303–324. Gordon & Breach, Newark, NJ.
- Moore, W. (1995). Inferring phylogenies from mtDNA variation: Mitochondrial-gene trees versus nuclear-gene trees. *Evolution* **49**: 718–726.
- Moran, P., and Kornfield, I. (1995). Were population bottlenecks associated with the radiation of the Mbuna species flock (Teleostei: Cichlidae) of Lake Malawi? *Mol. Biol. Evol.* **12**: 1085–1093.
- Ono, H., O'hUigin, C., Vincek, V., and Klein, J. (1993a). Exon-intron organization of fish major histocompatibility complex II B genes. *Immunogenetics* **38**: 223–234.
- Ono, H., O'hUigin, C., Tichy, H., and Klein, J. (1993b). Major-histocompatibility-complex variation in two species of cichlid fishes from Lake Malawi. *Mol. Biol. Evol.* **10**: 1060–1072.
- Pamilo, P., and Nei, M. (1988). Relationships between gene trees and species trees. *Mol. Biol. Evol.* **5**: 568–583.
- Pleyte, K. A., Duncan, S., and Phillips, R. (1992). Evolutionary relationships of the salmonid fish genus *Salvelinus* inferred from DNA sequences of the first internal transcribed spacer (ITS-1) of ribosomal DNA. *Mol. Phylogenet. Evol.* **1**: 223–230.
- Sage, R. D., Loisel, P. V., Basasibwaki, P., and Wilson, A. C. (1984). Molecular versus morphological change among cichlid fishes of Lake Victoria. In "Evolution of Fish Species Flocks" (A. A. Echelle and I. Kornfield, Eds.), pp. 185–202. Univ. of Maine Press, Orono, ME.
- Saitou, N., and Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**: 406–425.
- Savvaitova, K. A. (1980). Taxonomy and biogeography of charrs in the Paleartic. In "Charrs: Salmonid Fishes of the Genus *Salvelinus*" (E. K. Balon, Ed.), pp. 281–294. Junk, The Hague.
- Stager, J. C., Reinthal, P. N., and Livingston, D. A. (1986). A 25,000-year history for Lake Victoria, East Africa, and some comments on its significance for the evolution of cichlid fishes. *Fresh. Biol.* **16**: 15–19.
- Stiassny, M. L. J. (1991). Phylogenetic intrarelationships of the family Cichlidae: An overview. In "Cichlid Fishes: Behaviour, Ecology, and Evolution" (M. H. Keenlyside, Ed.), pp. 1–35. Chapman & Hall, London.
- Sturmbauer, C., and Meyer, A. (1992). Genetic divergence, speciation and morphological stasis in a lineage of African cichlid fish. *Nature* **358**: 578–581.
- Sturmbauer, C., and Meyer, A. (1993). Mitochondrial phylogeny of the endemic mouthbrooding lineages of cichlid fishes from Lake Tanganyika in eastern Africa. *Mol. Biol. Evol.* **10**: 751–768.
- Sültmann, H., Mayer, W. E., Figueroa, F., Tichy, H., and Klein, J. (1995). Phylogenetic analysis of cichlid fishes using nuclear DNA markers. *Mol. Biol. Evol.* **12**: 1033–1047.
- Swofford, D. L. (1990). PAUP—Phylogenetic Analysis Using Parsimony (Version 3.0). Computer program distributed by the Illinois Natural History Survey, Champaign, Illinois.
- Van Couvering, J. A. H. (1982). Fossil cichlid fishes of Africa. *Spec. Papers Paleontol. (Paleontological Ass.)* **29**: 1–103.
- Verhennen, E., Rüber, L., Snoeks, J., and Meyer, A. (1996). Mitochondrial phylogeography of rock-dwelling cichlid fishes reveals evolutionary influence of historical lake level fluctuations of Lake Tanganyika, Africa. *Phil. Trans. R. Soc. Lond. B* **351**: 797–805.
- Vogler, A. P., and DeSalle, R. (1994). Evolution and phylogenetic information content of the ITS-1 region in the Tiger Beetle *Cicindela dorsalis*. *Mol. Biol. Evol.* **11**: 393–405.