Reprinted from Nature, Vol. 347, No. 6293, pp. 550–553, 11th October, 1990

Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences

Axel Meyer*, Thomas D. Kocher*, Pereti Basasibwaki* & Allan C. Wilson

Division of Biochemistry and Molecular Biology, University of California, Berkeley, California 94720, USA

LAKE Victoria, together with its satellite lakes, harbours roughly 200 endemic forms of cichlid fishes that are classified as 'haplochromines'[1,2] and yet the lake system is less than a million years old. This 'flock' has attracted attention because of the possibility that it evolved within the lake from one ancestral species[3] and that biologists are thus presented with a case of explosive evolution. Within the past decade, however, morphology has increasingly emphasized the view that the flock may be polyphyletic[4,5]. We sequenced up to 803 base pairs of mitochondrial DNA from 14 representative Victorian species and 23 additional African species. The flock seems to be monophyletic, and is more akin to that from Lake Malawi than to species from Lake Tanganyika; in addition, it contains less genetic variation than does the human species, and there is virtually no sharing of mitochondrial DNA types among species. These results confirm that the founding event was recent.

Mitochondrial DNA (mtDNA) sequences were obtained from 14 species from 9 endemic genera of the Lake Victoria flock. These included two of the four monotypic genera described by Greenwood[6] as well as other representatives of the five main ecological and morphological groups, namely eaters of insects, fish, fish larvae, molluscs and algae (Table 1). A 363-base-pair (bp) part of the cytochrome *b* gene and a 440-bp segment of mtDNA that bears part of the threonine transfer RNA gene, all of the proline tRNA gene and the most variable part of the control region were sequenced after enzymatic amplification[7] (Fig. 1). There is little variation in these mtDNA segments among the 32 specimens examined; only 15 sites out of the 803 surveyed show base-substitutional differences (Table 1). The average number of differences between pairs of the 16 mtDNA types found was less than that occurring within human populations in this part of the control region[8]. A comparable result was evident from electrophoretic comparisons of proteins encoded by nuclear genes[3,9]. Despite the low genetic variation, there was virtually no sharing of mtDNA types among species (Table 1), a result that is exceedingly unlikely to arise by chance. This finding seems to rule out the possibility that most of the morphological species in Lake Victoria are merely alternative morphs of a few biological species[10,11]. No length mutations were detected in this survey and no variation was found in either the cytochrome *b* or the tRNA genes. In addition, our findings contrast with studies of other fishes, which generally show more intraspecific variation in conservative regions of mtDNA than is observed in our intergeneric comparisons within Lake Victoria in the control region[12,13].

To place these Victorian cichlids in relation to other cichlids, we sequenced both mtDNA segments in 26 additional species (Fig. 1, Table 2). As shown in Table 2, the endemic Victoria haplochromine species, which differ from each other by an average of only 3 substitutions (Table 1), differ from those of Lake Malawi by 54 to 55 substitutions and by at least 77 substitutions from other species. Likewise, the two groups of haplochromines from Lake Malawi, which differ from each other by an average of 24 substitutions, differ by at least 54 substitutions from cichlids elsewhere. Genealogical analyses using *Hemichromis*, from West Africa, as an outgroup confirmed this strong indication of monophyly for the Victoria flock (Fig. 2). Some of the morphologically convergent forms (for example, *Macropleurodus* from Lake Victoria, and *Chilotilapia* from Lake Malawi) that served before to support the notion of polyphyly[5] are represented in this study; they confirm our monophyly conclusion. In addition, two species endemic to Lake Tanganyika (*Julidochromis* and *Lamprologus*) are more distantly related to the taxa from Lakes Malawi and Victoria than is *Astatoreochromis*. This nonendemic Victorian haplochromine is the sister group to the endemic cichlid faunas of both Lakes Victoria and Malawi.

TABLE 1 Variation in the control region among 32 endemic fishes of the Lake Victoria flock

Species	No. of individuals	Bases at 15 variable positions
1. *Astatotilapia piceatus*	3	TGGTGCCATCACCGG
2. *Astatotilapia elegans*	1	....A....T..T..
3. *Astatotilapia nubilis*	1	...C.....T.T..
4. *Ptyochromis sauvagei*	3	....C.....C.T..
5. *Ptyochromis xenognathus**	3	..A.....C.T.A A A
6. *Labrochromis ishmaeli**	3	.....T.. C T
7. *Platytaeniodus degeni*	3	.....T..T..
8. *Macropleurodus bicolor*	2	.A.....C.T..
9. *Lipochromis obesus*	5	.....TG....T..
10. *Neochromis nigricans*	4	.....T..
11. *Prognathochromis longirostris*	1	.....T..
12. *Prognathochromis dentex*	1	..A.....T..
13. *Prognathochromis paraguiarti*	1	.....CT.TTA.
14. *Harpagochromis guiarti*	1	C.....T..

Variable nucleotide positions are indicated with numbers 1–15 in Fig. 1*a*; dots indicate identity of a nucleotide with the first species and letters below lines indicate that more than one nucleotide has been observed at this position in the polymorphic species (*). Specimens of species 1, 2 and 4–8 were collected in the Mwanza Gulf, Tanzania (southern Lake Victoria) and the others at Jinja, Uganda (northern Lake Victoria). For species 3, 13 and 14 only positions 84–440 were sequenced. Species 2 and 3 are not strictly endemic to Lake Victoria but also occur in its surrounding lakes, therefore supporting Greenwood's notion of a 'species superflock' of Victoria haplochromines. The species include eaters of algae (7, 10), insects (1, 2, 7), fish larvae (9), fish (11–14) and molluscs (4–6). All endemic haplochromine cichlids of Lake Victoria and the sand-dwelling cichlids of Lake Malawi (group A in Figs 1, 2) were once placed into four monotypic genera plus the genus *Haplochromis*. Greenwood[19,20] revised the taxonomy of the Lake Victoria haplochromines and assigned ~170 species into 20 endemic genera, reducing the number of species in *Haplochromis* to six. The use of the term 'haplochromines' here is not intended to imply any taxonomic rank assignments to this group of cichlids[6,14,19,20].

The 24 members of the Lake Malawi flock surveyed fall into two monophyletic groups, sand-dwellers and rock-dwellers, here called haplochromine groups A and B, respectively (Fig. 2). As each of these two ecological groups seems to consist of about 200 species[14] and our sampling of the faunal diversity of Lake Malawi may not be adequate, it would be premature to propose the existence of an exact association between phylogenetic position and ecological group for all Malawian species.

The patterns of base changes in the two regions sequenced behave as expected from surveys of other groups of vertebrates. Transitions outnumber transversions in both segments, and silent changes outnumber replacement changes in the cytochrome *b* gene. Moreover, the rate of change in the control region exceeds that in the cytochrome *b* gene. The results give no reason to suppose that the rates of mtDNA evolution are uneven or accelerated among East African cichlid fishes.

The likelihood of extreme recency for the Victorian flock is underscored by the finding that species within the Neotropical genus *Cichlasoma* differ by as much as 11% in their cytochrome *b* sequences[7] (Fig. 1*b*), whereas the Victoria flock differs by only 5% from the Malawi flock. Indeed, no diversity in the cytochrome *b* segment (legend to Table 2) was evident within the Victoria flock. An estimate of the divergence times among the flocks is based on the assumption that the mean rate of divergence in the cytochrome *b* gene is 2.5% per million years. Our molecular results date the origin of the extant Victorian flock at less than 200,000 years ago (Fig. 2). Lake Victoria's age has been estimated to be between 250,000 and 750,000 years[15],

* Present addresses: Department of Ecology and Evolution, State University of New York, Stony Brook, New York 11794, USA (A.M.); Department of Zoology, University of New Hampshire, Durham, New Hampshire 03824, USA (T.D.K.), and Freshwater Fisheries Research Organisation, Jinja, Uganda (P.B.).

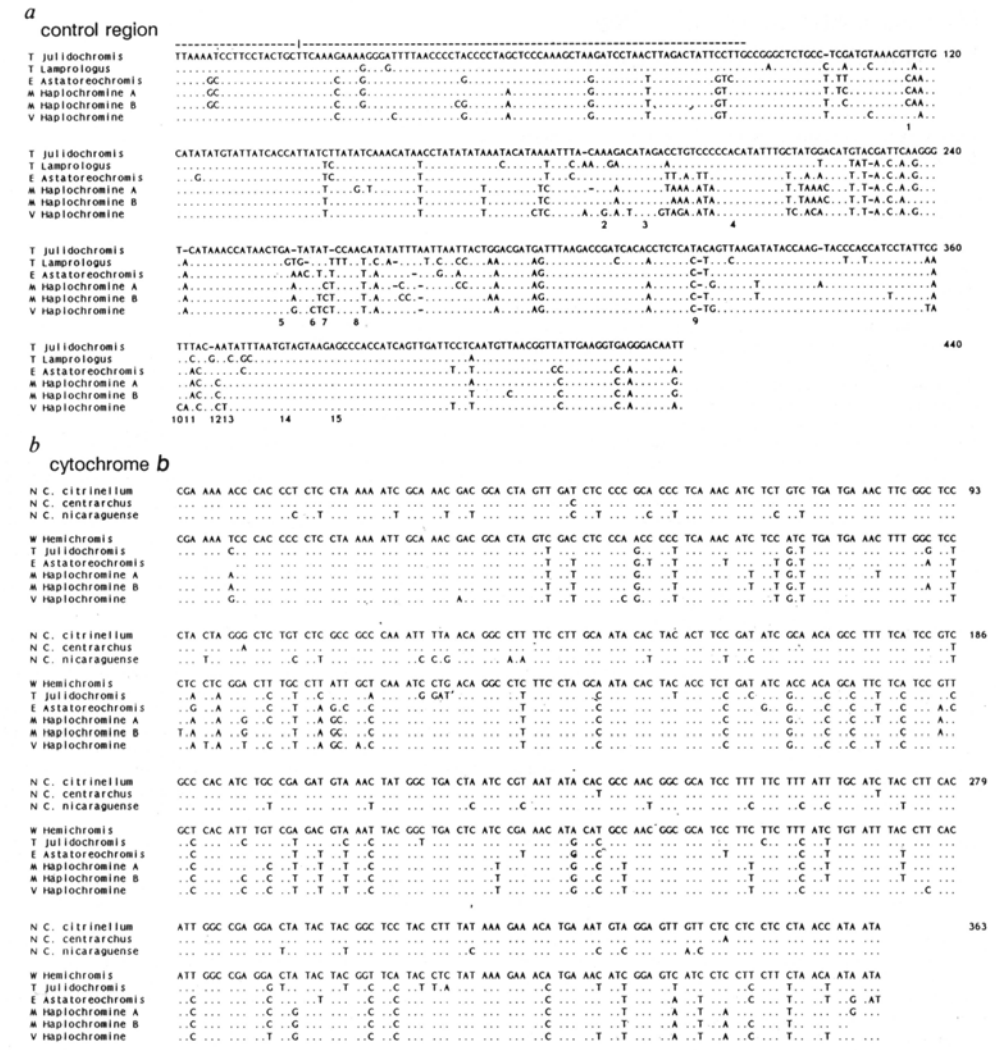

FIG. 1 Mitochrondrial DNA sequences from *a*, part of the threonine tRNA gene (18 bp), the proline tRNA gene (71 bp) and a 350-bp segment of the control region of African cichlids and *b*, part of the cytochrome *b* gene for African and Central American cichlids. In *a*: dots, sequence identity with *Julidochromis*; dashes, deletions; 1–15, nucleotide positions that are variable among the endemic cichlids of Lake Victoria (see Table 1). In *a* and *b*, consensus sequences for all specimens sequenced are shown for the endemic Lake Victoria haplochromines (see Table 1), the haplochromine groups A and B from Lake Malawi, and *Astatoreochromis* (see Table 2). For position 342 in *a*, one species of the Malawi group A (*Buccochromis*) had an A inserted (as shown). The sequences in *a* correspond approximately to positions 15,937–16,497, and in *b*, to positions 14,759–15,121 in the human mtDNA sequence[22]. The gene arrangement in the cichlids is similar to that in sturgeon[23] but seems to differ from that in cod, in which a 74-bp insertion exists between the threonine and proline tRNAs[24]. In *b*: dots, sequence identity either with *Cichlasoma citrinellum* (one of three species of *Cichlasoma* sequenced from the Central American Neotropics) for the three *Cichlasoma* congeners or with *Hemichromis*, from West Africa, for the African taxa. Blanks, missing sequence; N, Neotropics; T, Lake Tanganyika; E, East Africa; M, Lake Malawi; V, Lake Victoria.

METHODS. DNA was extracted from tissues of frozen or ethanol-preserved specimens[7,8]. Amplifications were performed in 25-μl volumes of Tris buffer (67 mM, pH 8.8) containing 2 mM $MgCl_2$, 1 mM of each dNTP, 1 μM of each primer, 10–1,000 ng template DNA and *Taq* polymerase (1 unit; Cetus). The primers used for the cytochrome *b* gene were L14724 (5′-CGAAGCTTGATATGAAAAACCATCGTTG-3′, designed by S. Pääbo) and H15149 (ref. 7); for the control region they were L15926 (ref. 7) and H16498 (5′-CCTGAAGTAGGAACCAGATG-3′). L and H refer to the light and heavy strands, respectively, and the numbers refer to the position of the 3′ base of the primers in human mtDNA[22]. The temperature profile for 30 cycles of double-stranded and 35 cycles of single-stranded amplification was 45 s at 94 °C (denaturation), 1 min at 55 °C (annealing of primers), 2 min at 72 °C (polymerization). The rest of the protocol is reported in detail in ref. 7.

TABLE 2 Base differences in two segments of mtDNA from African cichlid fishes

Species	Geographical origin	No. of base-substitutional differences B	C	D	E	F
A. *Hemichromis*	W. Africa	55	57	55	55	58
B. *Julidochromis*	L. Tanganyika	—	40	44	43	41
C. *Astatoreochromis*	E. Africa	58	—	24	23	32
D. *Buccochromis*	L. Malawi	63	38	—	7	18
E. *Pseudotropheus*	L. Malawi	63	38	17	—	18
F. 'Haplochromines'	L. Victoria	63	45	37	36	—

Mean number of differences in a 363-bp segment of the cytochrome *b* gene (from Fig. 1*b*) appear above the diagonal and in a 440-bp segment bearing tRNA genes and the control region below the diagonal (from Fig. 1*a*). The cytochrome *b* region was sequenced for a total of 18 specimens, which included 3 from 3 species of the Neotropical genus *Cichlasoma*, 1 *Hemichromis bimaculatus* from West Africa, 1 *Julidochromis regani*, 1 *Astatoreochromis alluaudi* (from Lake Victoria; this species also occurs in rivers of East Africa), 1 *Buccochromis atritaeniatus* (Malawi haplochromine group A), 1 each of 3 species of the *Pseudotropheus tropheops* complex from Lake Malawi (often called 'mbuna', and here called haplochromine group B), and 8 specimens of 7 haplochromine species from Lake Victoria (1, 2 and 4–8 in Table 1). No variation was detected among the specimens from Lake Victoria or among haplochromines of group B from Lake Malawi. The control region was sequenced from 66 specimens, which included 1 *Julidochromis regani* and 1 *Lamprologus birchardi* from Lake Tanganyika (which differ by 53 base substitutions, Fig. 1*a*), 8 specimens of *Astatoreochromis alluaudi* from Lake Victoria (3 from southern localities in Tanzania and 5 from northern Uganda; no fixed differences were found between northern and southern populations), 9 specimens of 9 species of group A from Lake Malawi (variation not shown), 15 specimens from Malawi of 10 group B species (variation not shown), and the 32 Lake Victoria haplochromines from 14 species whose results appear in Table 1. The Malawi group A species were *Chilotilapia rhoadesi, Buccochromis atritaeniatus, Nimbochromis polystigma, Dimidiochromis compressiceps, Maravichromis labidodon, Champsochromis spilorhynchus, Protomelas annectens, 'Haplochromis' diaboli* and *Sciaenochromis gracilis*[14]; the Malawi group B representatives consisted of eight specimens from seven undescribed species of the *Pseudotropheus tropheops* species complex[21] (P. Reinthal and A. M., unpublished data), 5 of *Pseudotropheus zebra* and 1 each of *Melanochromis auratus* and *Labeotropheus fuelleborni*. Consensus sequences appear in Fig. 1 for *Astatoreochromis* and Malawi groups A and B as well as for the Victoria haplochromines.

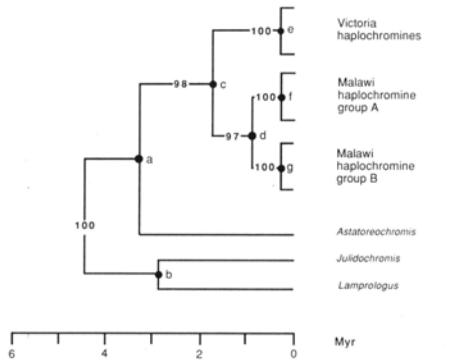

FIG. 2 Evolutionary tree for 36 African cichlid species based on a parsimony analysis[25] of mtDNA sequences. The tree is a composite based on two approaches. Only amino-acid replacement changes in cytochrome *b* were used to confirm the two deepest nodes, a and b, as well as to place the root on the ab branch, using *Hemichromis* as an outgroup. The four equally short trees obtained (each with 14 steps, consistency index CI = 0.75) did not sort out the branching order inside the group stemming from node a. A second approach used control region data with two cichlids of Lake Tanganyika (*Julidochromis* and *Lamprologus*) as the outgroup and established the branching pattern shown, with high boot-strap values (97–100)[26] confirming each of the internal branches (ab, ac, cd, ce, df and dg). The data did not allow the resolution of the relationships within any of the three assemblages stemming from nodes e, f and g (accordingly, parsimony analysis yielded nine equivalent solutions, each requiring 178 mutations; CI = 0.826). Other studies suggest that the cytochrome *b* gene diverges at a rate of at least 2.5% per million years in mammals[29]. On the basis of this rate and the per cent differences obtainable from Table 2, *Astatoreochromis* shared a common ancestor with the endemic Lake Victoria cichlids ~3.5 Myr ago, in exact agreement with the date inferred from studies of proteins encoded by nuclear genes[3]. The two main groups of cichlid fishes from Lake Malawi seem from this approach to have had a common ancestor ~700,000 years ago, the geological age of this lake being 1–2 Myr[16]. The low genetic distances observed among the proteins of the Malawi haplochromines B are consistent with this inference[27,28].

but younger ages are considered likely for the present-day fauna[3]. This date and the other estimated divergence times (Fig. 2) fit with the estimated geological ages of the three lakes (1–2 million years (Myr) for Lake Malawi and 2–4 Myr for Lake Tanganyika[16]). So it is likely that the Victoria flock arose within the lake.

Furthermore, this result weakens the hypothesis that similarly specialized species from different lakes are more closely related to each other than to morphological generalists from the same lake. These specializations probably evolved repeatedly and independently. The establishment of the monophyly of this species flock draws attention to the remarkable speed of the morphological diversification in Lake Victoria without an acceleration of molecular evolution.

Unfortunately, we are losing the opportunity to study the cichlid fauna of Lake Victoria, because much of it is going or has gone extinct as a result of the introduction of a non-endemic predatory fish.[17,18]. □

Received 30 July; accepted 20 August 1990.

1. Fryer, G. & Iles, T. D. *The Cichlid Fishes of the Great Lakes of Africa. Their Biology and Evolution* (Oliver & Boyd, Edinburgh, 1972).
2. Regan, C. T. *Proc. zool. Soc. Lond.* **1922,** 157–191 (1922).
3. Sage, R. D., Loiselle, P. V., Basasibwaki, P. & Wilson, A. C. in *Evolution of Fish Species Flocks* (eds Echelle, A. A. & Kornfield, I.) 185–201 (University of Maine Press, Orono, 1984).
4. Greenwood, P. H. *Bull. Br. Mus. nat. Hist. (Zool.)* **44,** 249–290 (1983).
5. Greenwood, P. H. *Bull. Br. Mus. nat. Hist. (Zool.)* **45,** 209–231 (1983).
6. Greenwood, P. H. *Bull. Br. Mus. nat. Hist. (Zool.)* **3,** 295–333 (1956).
7. Kocher, T. D. *et al. Proc. natn. Acad. Sci. U.S.A.* **86,** 6196–6200 (1989).
8. Vigilant, L., Pennington, R., Harpending, H., Kocher, T. D. & Wilson, A. C. *Proc. natn. Acad. Sci. U.S.A.* **86,** 9350–9354 (1989).
9. Nei, M. & Roychoudhury, A. K. *Evol. Biol.* **14,** 1–59 (1982).
10. Sage, R. D. & Selander, R. K. *Proc. natn. Acad. Sci. U.S.A.* **72,** 4669–4673 (1975).
11. Meyer, A. *Evolution* **41,** 1357–1369 (1987).
12. Avise, J. C., Bermingham, E., Kessler, L. G. & Saunders, N. C. *Evolution* **38,** 931–941 (1984).
13. Thomas, W. K. & Beckenbach, A. T. *J. molec. Evol.* **29,** 233–245 (1989).
14. Eccles, D. H. & Trewavas, E. *Malawian Cichlid Fishes. The Classification of Some Haplochromine Genera* (A. W. Diekhoff Lake Fish Movies, Herten, 1989).
15. Temple, P. H. *Biol. J. Linn. Soc.* **1,** 363–371 (1969).
16. Banister, K. E. & Clarke, M. A. *J. nat. Hist.* **14,** 483–542 (1980).
17. Barel, C. D. N. *et al Nature* **315,** 19–20 (1985).
18. Miller, D. J. *Trends Ecol. Evol.* **2,** 56–59 (1989).
19. Greenwood, P. H. *Bull. Br. Mus. nat. Hist. (Zool.)* **35,** 265–322 (1979).
20. Greenwood, P. H. *Bull. Br. Mus. nat. Hist. (Zool.)* **39,** 1–101 (1980).
21. Ribbink, A. J., Marsh, B. A., Marsh, A. C., Ribbink, A. C. & Sharp, B. J. *S. Afr. J. Zool.* **18,** 149–310 (1983).
22. Anderson, S. *et al. Nature* **290,** 457–465 (1981).
23. Gilbert, T. L. *et al. Nucleic Acids Res.* **16,** 11825 (1988).
24. Johansen, S., Guddal, P. H. & Johansen, T. *Nucleic Acids Res.* **18,** 411–419 (1990).
25. Swofford, D. L. *PAUP: Phylogenetic Analysis Using Parsimony, Version 3* (Illinois Natural History Survey, Champaign, Illinois, 1989).
26. Felsenstein, J. *Evolution* **39,** 783–791 (1985).
27. Kornfield, I. L. *Experientia* **34,** 335–336 (1978).
28. Kornfield, I. L., McKaye, K. & Kocher, T. *Isozyme Bull.* **18,** 76 1985.
29. Irwin, D. M., Kocher, T. D. & Wilson, A. C. *J. molec. Evol.* (in the press).

ACKNOWLEDGEMENTS. We thank K. Barel, R. Hoogerhoud, I. Kornfield, P. Reinthal, M. Stiassny and F. Witte for the collection and identification of specimens, I. Kornfield and P. Moran for mtDNAs of six Malawian cichlids and D. Irwin, I. Kornfield, M. Nishida, P. Reinthal and M. Stiassny, and especially E. Prager, for discussion. This work was supported by the NSF and NIH (A.C.W.), a Sloan fellowship to A.M., and an NIH fellowship to T.D.K.