Journal of Fish Biology (2008) **73,** 1139–1148
doi:10.1111/j.1095-8649.2008.01935.x, available online at http://www.blackwell-synergy.com

Molecular phylogeny of the clariid fishes of Lake Victoria, Tanzania, inferred from cytochrome *b* DNA sequences

C. J. Mwita*† and G. Nkwengulila‡

**Department of Fisheries Science and Aquaculture, University of Dar es Salaam, P. O. Box 35064, Dar es Salaam, Tanzania and ‡Department of Zoology and Wildlife Conservation, University of Dar es Salaam, P. O. Box 35064, Dar es Salaam, Tanzania*

(Received 20 November 2007, Accepted 14 April 2008)

Mitochondrial DNA cytochrome *b* (cyt *b*) sequence variation among the clariid fishes of Lake Victoria and the Malagarasi wetland Tanzania were studied between August 2003 and February 2005. Seven species were sampled and together with 26 cyt *b* sequences from GenBank were used to reconstruct phylogenetic relationships in the family Clariidae. The study revealed two clades: one consisting of the big-head species, *Clarias gariepinus* and *Heterobranchus longifilis*, and the other of small-sized species, *Clarias werneri*, *Clarias alluaudi*, *Clarias liocephalus* and *Clariallabes petricola*. The study further revealed that the genus *Clarias* is paraphyletic and that *H. longifilis* clusters deeply inside the *Clarias* group. Although *H. longifilis* is thought to be the oldest in evolutionary age in Tanzania, it is not close to the ancestor of the Clariidae family. The results also showed that *C. werneri* and *C. alluaudi* are genetically distinct from each other.
© 2008 The Authors
Journal compilation © 2008 The Fisheries Society of the British Isles

Key words: catfish; Clariidae; freshwater fishes; mtDNA.

INTRODUCTION

The Lake Victoria ecosystem has undergone dramatic shifts caused by the introduction of the Nile perch *Lates niloticus* (L.) in the late 1950s and early 1960s (Witte & Van Densen, 1995) and the water hyacinth in the past few decades (Odada & Olago, 2002). The ecogeological history suggests that at one time the lake dried up completely and thus the hundreds of species of fish in modern Lake Victoria may have evolved recently (Johnson *et al.*, 1996; Odada & Olago, 2002), and these include members of the family Clariidae. In Tanzania, the family Clariidae is represented by six genera, *Dinotopterus*, *Bathyclarias*, *Xenoclarias*, *Clariallabes*, *Clarias* and *Heterobranchus*, with 21 representative species occurring in various habitats in streams, rivers and lakes (Eccles, 1992). Lake Victoria alone has three clariid genera (*Xenoclarias*, *Clariallabes*

†Author to whom correspondence should be addressed. Tel.: +255 22 2410462/+255 22 2650011; fax: +255 22 2410480/2650244; email: mwitachacha@udsm.ac.tz

© 2008 The Authors
Journal compilation © 2008 The Fisheries Society of the British Isles

and *Clarias*) with six species, and *Dinotopterus* and *Bathyclarias* are endemic to Lakes Tanganyika and Malawi, respectively (Skelton, 1993).

Clariid fishes are adapted to different habitats. Species of *Bathyclarias* (Lake Malawi), *Dinotopterus* (Lake Tanganyika) and *Xenoclarias* (Lake Victoria) occupy deeper waters. *Clariallabes* inhabits stony habitats in rivers and wave-washed shores, whereas *Clarias* and *Heterobranchus* species thrive in relatively shallow, muddy and vegetated margins of lakes and rivers (Eccles, 1992; Skelton, 1993).

The clariids are characterized by long dorsal and anal fins, four pairs of barbels, absence of a dorsal spine and pectorals with a strong spine that is serrated in some species (Skelton, 1993), a taxonomic feature of the genus (Eccles, 1992). However, delimitation of species by this latter feature is sometimes ambiguous. For instance, Eccles (1992) and Witte & Van Densen (1995) defined *Clarias liocephalus* Boulenger differently. On the other hand, *Clarias werneri* Boulenger is considered to be synonymous with *Clarias alluaudi* Boulenger (Eccles, 1992). In a phylogenetic study of clariids, Agnese & Teugels (2001*a*, 2005) concluded that the current systematics of the clariid catfishes can no longer be maintained; however, more clariid species will need to be studied before introducing new nomenclature. One objective of this study was, therefore, to estimate phylogenetic relationships among clariid catfishes, particularly those from the Lake Victoria Basin and other freshwater bodies in Tanzania, and to relate phylogeny to the wider context of the genus worldwide. The second objective was to determine the genetic distances (hence the age) between clariid fish species in Lake Victoria and hence their evolutionary relationships.

MATERIALS AND METHODS

COLLECTION AND IDENTIFICATION OF FISH SPECIES

Sampling was carried out in diverse environments inhabited by clariid species in Lake Victoria on the Tanzanian side and the Malagarasi wetland (Table I). *Clarias gariepinus* (Burchell) ($n = 290$), *C. werneri* ($n = 126$) and *C. alluaudi* ($n = 81$) were collected from the muddy vegetated lake margins of the Mwanza Gulf, Speke Gulf, eastern part of Ukerewe Island and the delta of the Mara River. *Clarias liocephalus* ($n = 139$) were collected from the delta of Kagera River in the western part of the lake. *Clariallabes petricola* Greenwood ($n = 2$) were collected from the wetlands of Nyashishi River, *Bagrus docmac* (Forsskål) ($n = 19$) from Bukerebe Island, and *Heterobranchus longifilis* Valenciennes ($n = 1$) was obtained from Malagarasi wetland (Fig. 1).

Fish were caught by handlines and longlines with baited hooks, scoop and gill nets and transported to the laboratory in an ice-cooled box. Fish identification was based on phenotypic characteristics, such as serration of the pectoral fin spine, shape and position of the anterior fontanelle relative to the position of the eyes and the position of the dorsal and anal fins relative to the caudal fin (Teugels, 1986; Eccles, 1992; Witte & Van Densen, 1995).

About 1–2 g of white muscle tissue from each fish was taken for DNA analysis. The tissue was initially soaked in 1·0 ml 95% ethanol, transferred into another vial with fresh 1·0 ml 95% ethanol after 1 h and stored until required for DNA extraction.

DNA EXTRACTION, AMPLIFICATION AND SEQUENCING

Total genomic DNA was extracted from *c.* 5 g of tissue using the Qiagen DNeasy Extraction kit™ (Qiagen, Valencia, CA, U.S.A.) following the manufacturer's protocol

© 2008 The Authors

TABLE I. Clariid–bagrid species, specimen voucher numbers, GenBank accession numbers and sample localities of specimens used in this study

Species name	Locality	Specimen voucher numbers	GenBank accession numbers
Bagrus docmac	Mwanza Gulf	MGUDZM3481104	DQ646345
	Kagera River	KRUDZM3631104	DQ646346
	Mwanza Gulf	MGUDZM3501104	DQ646347
Clarias gariepinus	Ukerewe Island	UKUDZM3431104	DQ646360
	Ukerewe Island	UKUDZM2971004	DQ646361
	Magu Bay	MGBUDZM4091204	DQ646362
	Magu Bay	MGBUDZM3941104	DQ646365
	Mwanza Gulf	MGUDZM4350105	DQ646366
	Mwanza Gulf	MGUDZM4230105	DQ646367
	Mara Bay	MBUDZM3581104	DQ646368
	Mara Bay	MBUDZM3541104	DQ646369
	Kagera River	KRUDZM3701104	DQ646370
	Kagera River	KRUDZM3671104	DQ646371
	Kagera River	KRUDZM3661104	DQ646372
	Magu Bay	MGBUDZM4021204	DQ646363
	Magu Bay	MGBUDZM3961104	DQ646364
Heterobranchus longifilis	Malagarasi River	MLGUDZM4520105	DQ646348
Clarias alluaudi	Mwanza Gulf	MGUDZM4400105	DQ646373
	Mwanza Gulf	MGUDZM2881004	DQ646374
Clarias werneri	Mwanza Gulf	MGUDZM2931004	DQ646350
	Mwanza Gulf	MGUDZM1840704	DQ646351
	Mwanza Gulf	MGUDZM1680604	DQ646352
Clarias liocephalus	Mara Bay	MBUDZM3191104	DQ646353
	Mara Bay	MBUDZM3171104	DQ646354
	Kagera River	KRUDZM3831104	DQ646356
	Kagera River	KRUDZM3771104	DQ646357
	Kagera River	KRUDZM3861104	DQ646355
	Kagera River	KRUDZM3741104	DQ646356
Clariallabes petricola	Mwanza Gulf	NYUDZM0001104	DQ646359

and stored at 4° C until needed. A portion of mtDNA containing the cytochrome *b* (cyt *b*) gene was amplified using primers L15267, 5′-AAT GAC TTG AAG AAC CAC CGT-3′ and H15891, 5′-GTT TGA TCC CGT TTC GTG TA-3′, previously described by Briolay *et al.* (1998) and used by Agnese & Teugels (2001*a*, 2005) for similar work on clariid fishes. One microlitre of template DNA was used per 25 μl amplification reaction, which contained 0·5 μl of 10 mM dNTPs, 1 μl of 10 μM of each primer, 0·2 μl (5 U μl^{-1}) of HotStar *Taq* DNA polymerase (Qiagen) and its amplification buffer. Polymerase chain reactions (PCR) were run in an Eppendorf Mastercycler Thermocycler (Eppendorf, Hamburg, Germany) with an initial 'hot start' at 95° C for 15 min, followed by 40 cycles of denaturation at 93° C for 30 s, annealing at 50° C for 30 s and extension at 72° C for 1 min. Amplified products were cleaned with a QIAquick column (Qiagen), following manufacturer's instructions. DNA was recovered in a total volume of 30 μl double distilled H_2O. DNA was sequenced with the dideoxy-termination method with fluorescent-labelled primers using an ABI 377 Prism Automated DNA Sequencer (PE Applied Biosystem, Foster City, CA, U.S.A.). Both strands were sequenced with each of the two initial PCR primers, L15267 and H15891 (Agnese & Teugels, 2001*a*). Sequences were deposited in GenBank under accession numbers DQ646345–DQ646374 (Table I).

© 2008 The Authors

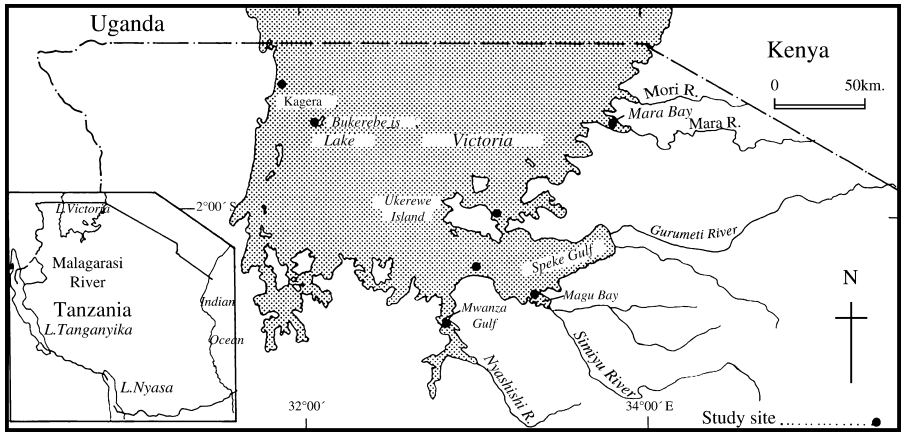


FIG. 1. Map of Lake Victoria showing the sampling sites.

DNA SEQUENCES AND PHYLOGENETIC ANALYSIS

Sequences were analysed using SEQUENCHER 3.0 (Gene Codes Corporation, Ann Arbor, MI, U.S.A.). The forward and reverse sequences were aligned to form a composite fragment of cyt *b* for each fish. For comparative analysis, 26 cyt *b* sequences from GenBank (Agnese & Teugels, 2005) were combined with the sequences of the present study and used to infer the phylogenetic relationship of clariid catfishes in the major freshwater bodies in Tanzania and elsewhere in Africa. For phylogenetic inference, maximum parsimony (MP), neighbour joining (NJ) with Tamura–Nei distances and maximum likelihood (ML) (likelihood parameters $A = 0{\cdot}2741$, $C = 0{\cdot}3047$, $G = 0{\cdot}1379$, $T = 0{\cdot}2831$, $\Gamma = 0{\cdot}6320$ and $I = 0{\cdot}333$) were applied using PAUP* 4b10 for Macintosh (Swofford, 2000). The robustness of the nodes in each phylogeny was assessed by bootstrap analysis (Felsenstein, 1985) with 1000 replicates.

RESULTS

The nucleotide compositions of cyt *b* sequences in the species examined in the present study are similar to those of other fishes, including those in the Clariidae (Briolay *et al.*, 1998; Agnese & Teugels, 2001*a*, *b*, 2005); there is low G content (14·7%) and almost equal amounts of A, T and C (28, 27·8 and 29·5%, respectively). An 1122 bp composite segment of cyt *b* included 814 invariant and 300 phylogenetically informative positions. Most of the variable nucleotide sites occurred in the third codon position (228, 75·3%), and the rest were distributed in the second codon (12, 4·0%) and the first codon positions (60, 20·7%). Uncorrected sequence divergences ranged from 0·3 to 11·5% between the clariid species, but ranged between 18·3 and 20·6% in *B. docmac*.

The cyt *b* nucleotide sequences were used to reconstruct the phylogeny of the six clariid species from Lake Victoria and *H. longifilis* from the Malagarasi River. Four trees were derived separately with MP and NJ, based on Tamura–Nei distances, with *B. docmac* as an outgroup (Fig. 2). The topology of a likelihood tree was similar to the topologies of the MP and the NJ trees (Fig. 3).

Based on these trees, two major clades were consistently observed. The first clade included *C. gariepinus* (the big-head clade) and the second included

© 2008 The Authors

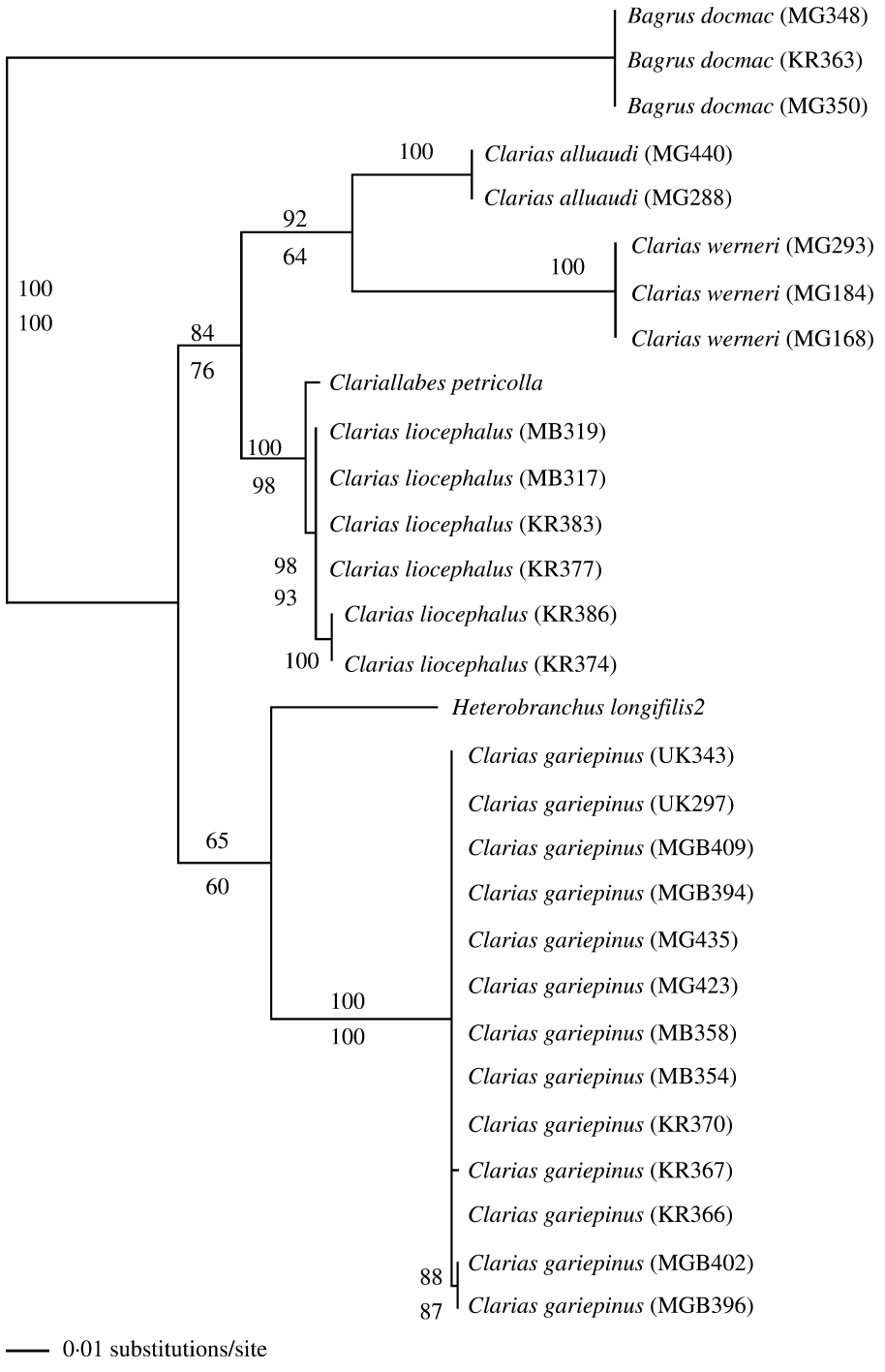


FIG. 2. Phylogenetic tree based on maximum parsimony (MP) and neighbour joining (NJ) methods, numbers on branches are percentage bootstrap values based on 1000 replicates, top (MP) and lower (NJ). Abbreviations in brackets refer to sampling sites: KR, Kagera River; MB, Mara Bay; MG, Mwanza Gulf; MGB, Magu Bay; UK, Ukerewe Island.

© 2008 The Authors
Journal compilation © 2008 The Fisheries Society of the British Isles,

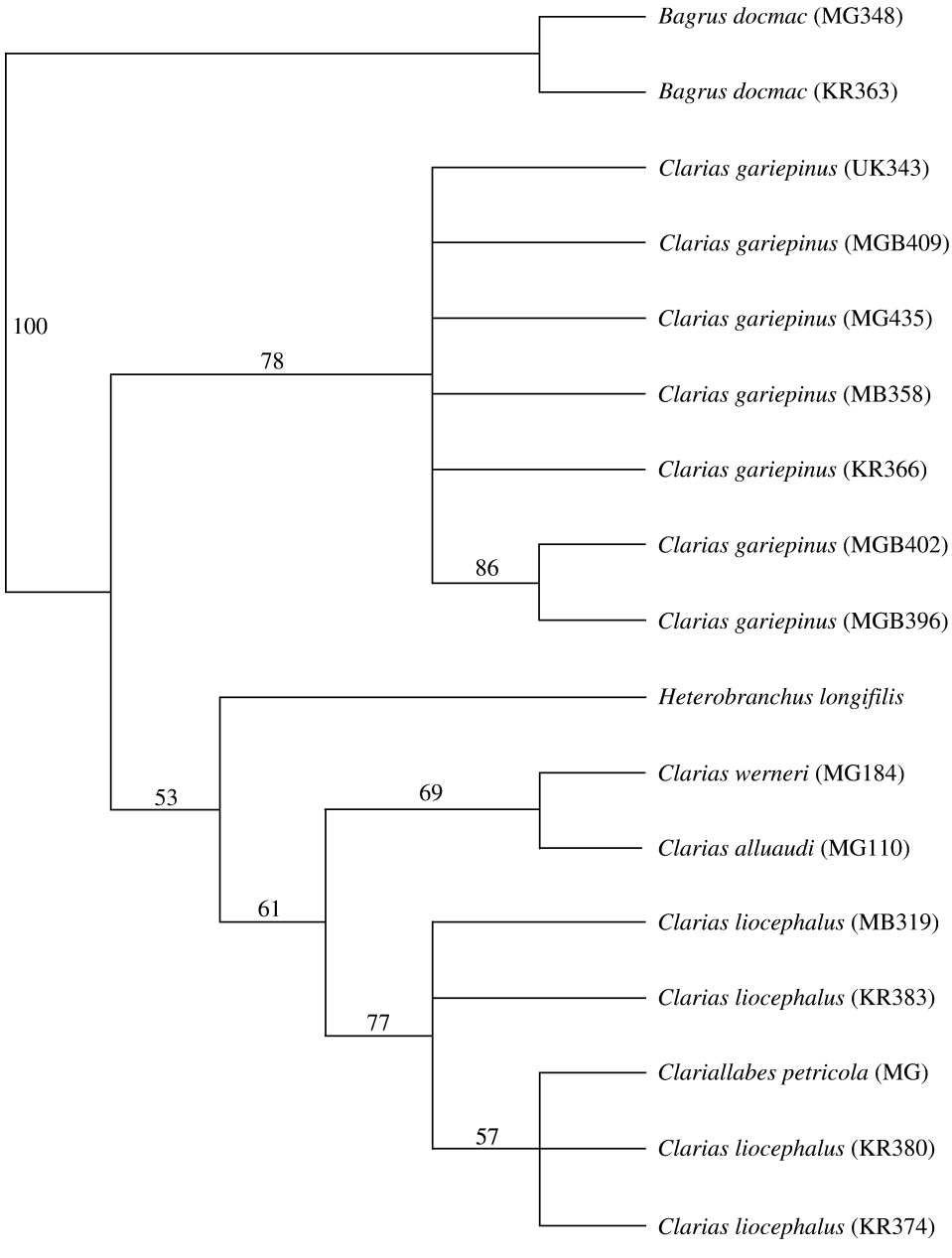


FIG. 3. Phylogenetic tree based on maximum likelihood method, numbers on branches are percentage bootstrap values based on 1000 replicates. Abbreviations in brackets refer to sampling sites: KR, Kagera River; MB, Mara Bay; MG, Mwanza Gulf; MGB, Magu Bay; UK, Ukerewe Island.

C. werneri, *C. alluaudi*, *C. liocephalus* and *C. petricola* (small-sized clariids) (Figs 2 and 3). Each species clustered independently to form a separate monophyletic clade. In addition, *C. gariepinus* and *C. liocephalus* from different localities of the lake each formed two subclades in the MP, NJ and ML trees

© 2008 The Authors

(Figs 2 and 3). *Clarias liocephalus* is closely related to *C. petricola*; likewise, *C. werneri* is closely related to *C. alluaudi*. However, the NJ tree showed that *C. werneri* and *C. alluaudi* are distantly related to each other, as they are each more closely related to other species represented by sequences from GenBank (Fig. 4).

DISCUSSION

Analysis of mtDNA cyt *b* sequences yielded a well-supported clariid phylogenetic tree revealing two clades: a big-head clade including *H. longifilis*, *C. gariepinus* (Figs 2 and 3), *Dinotopterus cunningtoni* Boulenger and *Bathyclarias* species (Fig. 4) and a monophyletic clade including smaller clariid species (Figs 2 and 4). In trees rooted with the outgroup species *B. docmac*, *H. longifilis* was placed deep inside the *Clarias* group, which is paraphyletic, as previously reported by Agnese & Teugels (2001*a*). Based on the body morphology, *C. gariepinus*, *Clarias anguillaris* (L.) and *H. longifilis* were considered as close relatives (Cabuy *et al.*, 1999) and the anguilliform clariid species as the most recent in the line of evolutionary divergence (Agnese & Teugels, 2001*a*). Recently, however, Agnese & Teugels (2005) demonstrated that morphological and osteological evolution in the Clariidae did not follow the orthogenic series, suggesting that the body forms observed in clariid species represent adaptations to habitats. Some forms, *e.g.* eel-like bodies, have evolved independently several times in the Clariidae. Molecular evidence from the present analysis clearly supports the above argument. The results also show that *C. gariepinus*, *C. werneri*, *C. alluaudi*, *C. liocephalus* and *C. petricola*, as a group, are distinct from *H. longifilis*, which appears deep inside the *Clarias* group.

Eccles (1992) and Witte & Van Densen (1995) considered the species *C. werneri* and *C. alluaudi* to be synonymous. Similarly, morphometric analysis in the present study (unpubl. data) could not separate these two species (Mwita, 2006). Cytochrome *b* analysis, however, showed that although *C. werneri* and *C. alluaudi* are closely related (Figs 2 and 3), they are genetically distinct from each other (Fig. 4) and hence should be retained as two separate species. These results, however, should be interpreted with caution, as efforts are underway to further review the taxonomy of the Clariidae by the barcoding molecular method. *Clarias liocephalus* and *C. gariepinus*, on the other hand, each forms two separate subclades. These results support previous studies that suggest the existence of genetically distinct populations in *C. gariepinus* (Agnese & Teugels, 2001*b*, 2005; Giddelo *et al.*, 2002). The present results, however, also demonstrate genetic population differences within *C. liocephalus* for the first time.

Several hypotheses may explain the above findings. One of the best explanations to support the possibility of different genetic subpopulations in either *C. liocephalus* or *C. gariepinus* is that of extensive adaptive radiation, as reported for species in the Lake Victoria Basin (Greenwood, 1991). Populations using different habitats in the same water bodies are likely to diverge genetically from one another, as is the case in the Lake Victoria Basin in which different clariid species are known to use a variety of habitats (Eccles, 1992; Witte & Van Densen, 1995).

A second way in which genetic variation might have occurred in Lake Victoria could be through introductions of different gene pools of clariid

© 2008 The Authors

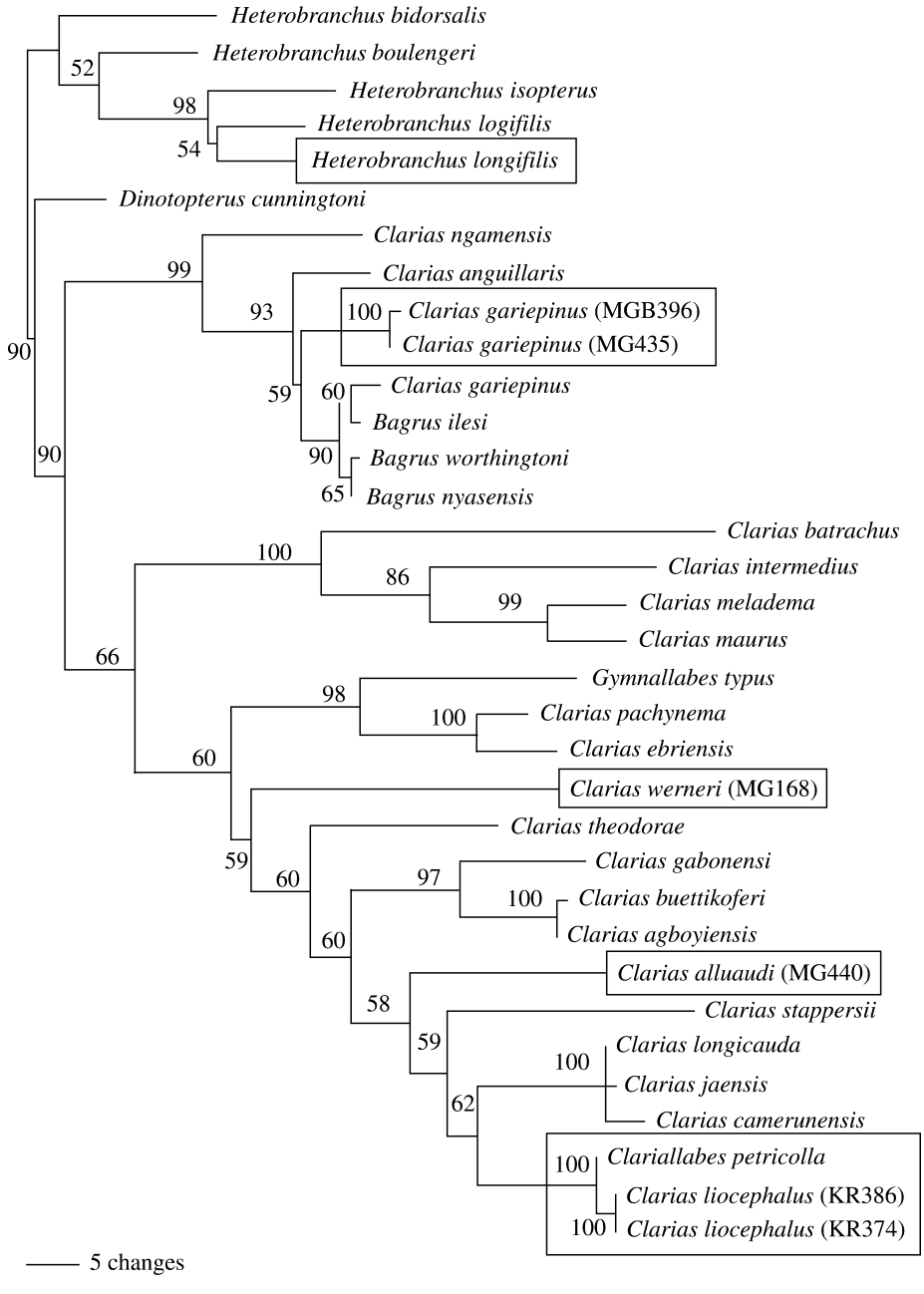


FIG. 4. Phylogenetic tree (unrooted) based on neighbour joining methods with sequences from the GenBank, numbers on branches are percentage bootstrap values based on 1000 replicates. Sequences for the species in boxes are from the present material, the rest are from GenBank (Agnese & Teugels, 2005). Abbreviations in brackets refer to sampling sites: KR, Kagera River; MG, Mwanza Gulf; MGB, Magu Bay.

© 2008 The Authors
Journal compilation © 2008 The Fisheries Society of the British Isles,

population(s) from the major rivers, including the Kagera River in the west (specimens KR386 and KR374), the Mara in the East (specimens MB319 and MB317) and the Simiyu in the south-east of the lake (specimens MGB396 and MGB402) (Fig. 1). Similarly, as *C. gariepinus* transverses great distances, the observed genetic variability may be because of dispersal, followed by restriction to some habitats or watersheds because of hydrogeographic cycles. This may well be the case in the Lake Victoria ecosystem, which has had a long and eventful ecogeological history (Meyer *et al.*, 1990).

Lake Victoria is *c.* 250 000–750 000 years old (Meyer *et al.*, 1990). However, clariid species in the lake are estimated to be 0·5–12 million years old (present study), based on a 1% sequence divergence per million years (Agnese & Teugels, 2005). The lake is believed to have dried out *c.* 17 500 years BP and refilled again *c.* 12 500 years BP (Johnson *et al.*, 1996). The timings of these hydrographic events indicate that the present-day clariid species flock in Lake Victoria is older than the lake itself. This supports the hypothesis that the different gene pools of clariid populations observed in the present study were introduced into the lake from the major rivers, which serve as refugia, as suggested by Katunzi (1996). The timeframe also nearly coincides with the time estimated for the invasion of Clariidae into the African continent from Asia (Agnese & Teugels, 2005). It is therefore possible that *H. longifilis* was the first to invade the freshwater bodies in Tanzania. However, the species has disappeared from the present Lake Victoria Basin because of ecohydrological modification of the ichthyofauna of the lake (Giddelo *et al.*, 2002).

The authors thank all those who made this work possible. Most important is the Lake Victoria Environment Management Project, University of Dar es Salaam subcomponent who financed this study, Tanzania Fisheries Research Institute (TAFIRI), Mwanza centre for provision of laboratory space and S. M. Carr of Memorial University, Newfoundland in whose laboratory the molecular work was conducted.

References

Agnese, J. F. & Teugels, G. G. (2001*a*). Genetic evidence for monophyly of the genus *Heterobranchus* and paraphyly of the genus *Clarias* (Siluriformes, Clariidae). *Copeia* **2001,** 548–552.

Agnese, J. F. & Teugels, G. G. (2001*b*). The *Bathyclarias–Clarias* species flock. A new model to understand rapid speciation in African great lakes. *Comptes Rendus de l'Académie des Sciences de Paris, Sciences de la vie* **34,** 683–688.

Agnese, J. F. & Teugels, G. G. (2005). Insight into the phylogeny of African Clariidae (Teleostei: Siluriformes): implications for their body shape evolution, biogeography and taxonomy. *Molecular Phylogenetics and Evolution* **36,** 546–553.

Briolay, J., Galtier, N., Brito, R. M. & Bouvet, Y. (1998). Molecular phylogeny of Cyprinidae inferred from cytochrome b DNA sequences. *Molecular Phylogenetics and Evolution* **9,** 100–108.

Cabuy, E., Adriaens, D., Verraes, W. & Teugels, G. G. (1999). Comparative study on the cranial morphology of *Gymnallabes typus* (Siluriformes: Clariidae) and their less anguilliform relatives, *Clariallabes melas* and *Clarias gariepinus*. *Journal of Morphology* **240,** 169–194.

Eccles, D. H. (1992). *FAO Species Identification Sheets for Fishery Purposes. Field Guide to the Freshwater Fishes of Tanzania*. Rome: UNDP Project URT/87/016.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39,** 783–791.

© 2008 The Authors
Journal compilation © 2008 The Fisheries Society of the British Isles,

Giddelo, C. S., Arndt, A. D. & Volckaert, F. A. M. (2002). Impact of rifting and hydrogeography on the genetic structure of *Clarias gariepinus* in eastern Africa. *Journal of Fish Biology* **60,** 1252–1266.

Greenwood, P. H. (1991). Speciation in cichlid fishes. In *Cichlid Fishes Behaviour, Ecology and Evolution* (Keenleyside, M. H. A., ed.), pp. 86–102. London: Chapman & Hall.

Johnson, T. C., Scholz, C. A., Talbot, M. R., Kelts, K., Ricketts, R. D., N'gobi, G., Beuning, K. R. M., Ssemmanda, I. & McGill, J. W. (1996). Late Pleistocene desiccation of Lake Victoria and rapid evolution of cichlid fishes. *Science* **273,** 1091–1093.

Katunzi, E. F. B. (1996). Conservation of satellite lakes as refugia for the endangered species of Lake Victoria – a case study of Lake Malimbe. *TAFIRI Research Bulletin* **4,** 72–80.

Mwita, C. J. (2006). Studies on the determinants of the parasites' community structure of the clariid fishes in Lake Victoria. PhD Thesis, University of Dar es Salaam.

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

Odada, E. O. & Olago, D. O. (2002). *The East African Great Lakes: Limnology, Palaeolimnology and Biodiversity*. London: Kluwer.

Skelton, P. (1993). *A Complete Guide to the Freshwater Fishes of Southern Africa*. Johannesburg: Southern Book Publishers.

Swofford, D. L. (2000). *PAUP*. Phylogenetic Analysis Using Parsimony and Other Methods. Ver. 4.b10*. Sunderland, MA: Sinauer Associates.

Teugels, G. G. (1986). A systematic revision of the African species of the genus *Clarias* (Pisces; Clariidae). *Annales Musée Royal de l'Afrique Centrale* **247,** 1–199.

Witte, F. & Van Densen, W. L. T. (1995). *Fish Stock and Fisheries of Lake Victoria. A Handbook for Field Observation*. Cardigan: Samara Publishing House.

© 2008 The Authors
Journal compilation © 2008 The Fisheries Society of the British Isles,