Available online at www.sciencedirect.com
SCIENCE DIRECT
Molecular Phylogenetics and Evolution 38 (2006) 65–78
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Phylogenetic relationships and divergence time estimate of African anguilliform catfish (Siluriformes: Clariidae) inferred from ribosomal gene and spacer sequences

G. Jansen [a,*,1], S. Devaere [a], P.H.H. Weekers [b,c], D. Adriaens [a]

[a] Ghent University, Evolutionary Morphology of Vertebrates, K.L. Ledeganckstraat 35, B-9000 Gent, Belgium
[b] Ghent University, Research Group Limnology, K.L. Ledeganckstraat 35, B-9000 Gent, Belgium
[c] Centre for Molecular Phylogeny and Evolution (CeMoFE), Ghent University, K.L. Ledeganckstraat 35, B-9000 Gent, Belgium

Received 11 March 2005; revised 12 September 2005; accepted 12 September 2005
Available online 14 November 2005

Abstract

The catfish family Clariidae comprises species in which the body shape ranges from fusiform to anguilliform. Recent studies have shown that this body elongation is the result of convergent evolution. This paper aims to study the evolution towards anguilliformity in a phylogenetic framework. Sequences of 29 taxa were analyzed using the neighbor-joining, maximum-likelihood, maximum-parsimony, and Bayesian inference algorithms and the parsimony algorithm in POY. The study yields phylogenetic hypotheses showing well-supported clades. Anguilliformity appears to have arisen at least four times, each time having a sister group relation with a fusiform *Clarias*-like ancestor. Divergence time estimation indicates that the African Clariidae started radiating between 123 and 56 My ago.

Keywords: Molecular phylogeny; 18S; 5.8S; ITS1+2; Ribosomal DNA; Dating; Divergence times

1. Introduction

Clariidae are one of the about 31 families belonging to the teleostean order Siluriformes (catfish) (Teugels, 2003; but see Sabaj et al., 2004; for a more up to date number). Its representatives can be found all over Africa, the Middle East and parts of Asia (Greenwood, 1961; Teugels, 2003). They have a suprabranchial organ, permitting them to extract oxygen directly from the atmosphere (Greenwood, 1961; Graham, 1997). This adaptation and their thick skin with mucous pores could explain their distribution in swamps, flood plains and periodically dry pools (Burgess, 1989; Matthes, 1964; Teugels, 1986). The family consists of fifteen genera, 12 endemic to Africa and two endemic to Asia, comprising a total of 93 species (Sabaj et al., 2004). *Clarias* is the only genus occurring on both African and Asian continents and can be divided into seven subgenera (Teugels, 2003). Anguilliform clariids belong to the genera *Channallabes*, *Gymnallabes*, *Dolichallabes*, *Platyallabes*, and *Platyclarias*. They have a long, eel-like body, and often the pectoral and pelvic fins are reduced. They show hypertrophied jaw-muscles and a narrow skull roof (Cabuy et al., 1999; Devaere et al., 2001). These eel-like forms can typically be found in the swamps of the Western and Central African rain forests.

This family presents a unique example among Teleostei since it is the only one where an evolutionary transformation from fusiform to anguilliform morphs has been observed at the species level (Pellegrin, 1927). Initially, anguilliformity was considered to have evolved gradually in this family, in a series starting with *Heterobranchus* Geoffrey St. Hilaire, 1809 and ending with the extreme anguilliform *Dolichallabes* Poll, 1942 (Boulenger, 1902; David, 1935; Pellegrin, 1927). This intuitional gradualism was first doubted by Poll (1942, 1977). More recent studies

* Corresponding author. Fax: +358 9 191 57694.
E-mail address: gunther.jansen@helsinki.fi (G. Jansen).
[1] Present address: University of Helsinki, Department of Biological and Environmental Sciences, P.O. Box 65 (Viikinkaari 1), FIN-00014 Helsinki, Finland.

1055-7903/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2005.09.011

on clariid phylogeny, based on the structure of the suprabranchial organ (Graham, 1997), molecular data (Agnèse and Teugels, 2001a,b), a review of the literature on morphological data (Teugels and Adriaens, 2003), and mitochondrial DNA (Agnèse and Teugels, 2005) all confirm the polyphyletic nature of anguilliformity, and the monophyletic origin of the family, although the interrelations of different species and genera of Clariidae are not yet resolved. *Bathyclarias* species appear to have evolved recently from *Clarias gariepinus* ancestors in Lake Malawi (Agnèse and Teugels, 2001b). A close relationship between *Heterobranchus* and *Clarias gariepinus* is suggested by Legendre et al. (1992), Teugels et al. (1992), Agnèse and Teugels (2001a), and Teugels and Adriaens (2003).

DNA sequence data play an essential role in the reconstruction of evolutionary relationships among organisms, resulting in insights in genetic affinities that may confirm or conflict with traditional taxonomy. Because of its attractive properties, ribosomal DNA (rDNA) is popular for examining phylogenetic relationships and for studying genetic variability and divergence within and between species. Such properties are secondary structure features, differential rates of evolution between different regions, and tandemly repeated genes (Arnheim, 1983; Gebri, 1985). As in other eukaryotes, the rDNA of catfish contains tandemly repeated transcriptional units, which are separated by intergenic spacers. Within each transcriptional unit, the internal transcribed spacer 1 (ITS1) separates the 18S small subunit (18S rDNA) from the 5.8S, while the internal transcribed spacer 2 (ITS2) separates the 5.8S from the 28S large subunit (28S rDNA). The coding regions show little sequence divergence among closely related species, whereas the spacer regions may exhibit higher variability as mutations occur at a relatively rapid rate in internal transcribed spacers. Therefore, these regions may resolve the relationships between closely related species that otherwise show little genetic divergence (Fritz et al., 1994b; Porter and Collins, 1991; Tang et al., 1996).

This study concentrates on the nature of the phylogenetic relationships between the anguilliform African clarids only, using molecular data. Parsimony based morphological evolution would imply that anguilliform taxa should form a monophyletic group. However, a homoplasic origin of body elongation has been suggested based on mtDNA (e.g., Agnèse and Teugels, 2005). The latter study, focusing on overall Clariidae phylogeny, included only two anguilliform taxa: *Channallabes apus* and *Gymnallabes typus*. A recent taxonomic revision of the former species, however, demonstrated the existence of six species, of which three new to science and two rehabilitated nominal taxa (*Channallabes apus*, *Ch. longicaudatus*, *Ch. ogooensis*, *Ch. sanghaensis*, *Ch. alvarezi*, and *Ch. teugelsi*, Devaere et al., unpublished data). With the main objective to reveal the phylogenetic affinities of the anguilliform taxa (and not the Clariidae as a whole), these new taxa are included in this study (with the exception of *Ch. ogooensis*). Additionally, as morphological data is not included in this phylogenetic analyses, and as this is part of a separate study (Devaere et al., unpublished data), no discussion on the morphological evolution towards anguilliformity is given in this paper. An estimate of divergences between the anguilliform taxa and their sister groups is performed, to reveal the time frame of this vast body transformation.

2. Materials and methods

2.1. DNA extraction, PCR amplification, and sequencing reactions

The origins of the samples used in this study (28 new samples, 3 GenBank sequences) are listed in Table 1. The Gabon specimens were collected during expeditions in 1999 and 2000 to assess the complete species richness in the Gabon area. Several samples of different regions of what was then still known as *Channallabes apus* were included in this analysis. Muscular tissue was isolated and total DNA was prepared according to the protocol of the Puregene DNA isolation kit type D-5000A (Gentra Systems, BIOzym, Landgraaf, The Netherlands). The complete region of the ribosomal spacers (ITS1 and ITS2) and the ribosomal 18S, 5.8S and part of the 28S genes was amplified using the polymerase chain reaction (PCR) with Qiagen DNA polymerase (Westburg, Leusden, The Netherlands). Eukaryote-specific external primers complementary to the 5′-terminus of the 18S rDNA gene (5′-TYCCTGGTTGATYYTGCC AG-3′) and the 5′-terminus of the 28S rDNA gene (5′-T GA TCC ATC TGC AGG TTC ACC T-3′) were used to amplify the entire 18S–ITS1–5.8S–ITS2 and part of the 28S region. As a new reverse primer, we used (5′-AAT CCT GGT TAG TTT CTT TTC C-3′). Internal primers were used as described previously (Samraoui et al., 2003; Weekers et al., 1994). PCR amplifications, purification of the PCR products, and DNA sequencing was done according to standard procedures (Samraoui et al., 2003). External and internal primers in conserved regions of the 18S and 5.8S rDNA were used for sequencing (Samraoui et al., 2003; Weekers et al., 1994).

2.2. Sequence alignment and the construction of datasets

The DNA sequences covering the complete 18S–ITS1–5.8S–ITS2–28S (partial) region were aligned with CLUSTALW 1.8 (Thompson et al., 1997) using default settings, resulting in an initial dataset. A second dataset was created by fine-tuning the alignment of the initial dataset based on secondary structural information, using GeneDoc 2.6.002 (Nicholas et al., 1997). The alignment of the 18S gene region was manually optimized with published 18S rDNA sequences based on the conservation of both primary sequence data and inferred secondary structural features (Nelles et al., 1984) (the rRNA WWW Server: http://www.psb.ugent.be/rRNA/ssu/index.html) (The Ribosomal Database Project: http://geta.life.uiuc.edu/index2.html.) The small and highly conserved 5.8S gene region and the small

Table 1
Taxon names, geographical origin, collectors name, and EMBL accession number of the ribosomal DNA sequences (18S, ITS1, 5.8S, ITS2, partial 28S) that are used in this study

Full species name	Geographic location	Coordinates	Collector	Collection date	Accession Number
Outgroup					
Clupea harengus					X98845
Cyprinus carpio					U87963
Ictalurus punctatus					AF021880
Kryptopterus bicirrhis	Commercially obtained	—	—	—	AJ876375 (18S, ITS1, 5.8S, ITS2)
Pangasianodon hypophthalmus	Commercially obtained	—	—	—	AJ876376 (18S, ITS1, 5.8S, ITS2)
Heteropneustes fossilis	Cultured in Hungary	—	Donated by Lazlo Orban to Jean-François Agnèse	—	AJ876377 (18S, ITS1, 5.8S, ITS2)
Clarias					
Subgenus Anguilloclarias					
Clarias pachynema 1	Mopia, Gabon	1°48,92′S 13°36,68′E	Adriaens, Devaere & Herrel	2000	AJ876379 (18S, ITS1, 5.8S, ITS2)
Clarias pachynema 2	Ebeigne, Gabon	1°48,92′S 3°36,68′E	Adriaens, Devaere & Herrel	2000	AJ876392 (18S (partial), ITS1, 5.8S, ITS2)
Clarias submarginatus	Abenelang, Gabon	1°28,92′N 11°35,41′E	Adriaens, Devaere, & Herrel	2000	AJ876395 (18S (partial), ITS1, 5.8S, ITS2)
Clarias theodorae 1	Zambia	—	Donated by Jean-François Agnèse		AJ876396 (18S (partial), ITS1, 5.8S, ITS2)
Clarias theodorae 2	Swampy channel near Drotsky's Camp, Botswana	18°24,68′S 21°52,8′E	SAIAB	02-02-2003	AJ876397 (18S (partial), ITS1, 5.8S, ITS2)
Subgenus Clarioides					
Clarias buthupogon	Zambia	—	Donated by Jean-François Agnèse		AJ876378 (18S, ITS1, 5.8S, ITS2)
Subgenus Clarias					
Clarias gariepinus	Raised in lab	—	Adriaens, Devaere & Herrel		AJ876383 (18S (partial), ITS1, 5.8S, ITS2)
Subgenus Dinotopteroides					
Clarias ngamensis 1	Okavango, Botswana	—	Roger Bills, SAIAB	03-05-1999	AJ876384 (18S, ITS1, 5.8S, ITS2)
Clarias ngamensis 2	Mpika-stream (Tanganyika)	11°49,87′S 1°14,82′E	SAIAB	18-03-2004	AJ876398 (18S (partial), ITS1, 5.8S, ITS2)
Clarias stappersi	Mofwe, Zambia	—	Donated by Jean-François Agnèse		AJ876381 (18S, ITS1, 5.8S, ITS2)
Subgenus Platycephaloides					
Clarias jaensis	Oyem, Gabon	1°28,37′N 11°35,34′E	Adriaens, Devaere & Herrel	2000	AJ876386 (18S, ITS1, 5.8S, ITS2)
Clarias platycephalus	Ebeigne, Gabon	1°28,37′N 11°35,34′E	Adriaens, Devaere & Herrel	2000	AJ876400 (18S (partial), ITS1, 5.8S, ITS2)
Subgenus Brevicephaloides					
Clarias camerunensis	Ebeigne, Gabon	1°28,37′N 11°35,34′E	Adriaens, Devaere & Herrel	2000	AJ876402 (18S (partial), ITS1, 5.8S, ITS2)
Heterobranchus					
Heterobranchus isopterus	Côte d'Ivoire	—	Donated by Jean-François Agnèse		AJ876382 (18S, ITS1, 5.8S, ITS2)
Dinotopterus					
Dinotopterus cunningtoni	Nsumbu Island	08°30,56′S,30°28,47′E	SAIAB	07-03-2004	AJ876385 (18S, ITS1, 5.8S, ITS2)
Tanganikallabes					
Tanganikallabes mortiauxi	Cape Kachese harbour	08°29,9′S,30°28,47′E	SAIAB	08-03-2004	AJ876380 (18S, ITS1, 5.8S, ITS2)
Clariallabes					
Clariallabes longicauda 1	Makokou, Gabon	0°33′N 12°51′E	Adriaens, Devaere & Herrel	2000	AJ876389 (18S, ITS1, 5.8S, ITS2)
Clariallabes longicauda 2	Mefange, Gabon	1°30,65′N 11°35,57′E	Adriaens, Devaere & Herrel	2000	AJ876390 (18S, ITS1, 5.8S, ITS2)
Channallabes					
Channallabes alvarezi 1	Abenelang, Gabon	1°28,92′N,11°35,41′E	Adriaens, Devaere & Herrel	2000	AJ876401 (18S (partial), ITS1, 5.8S, ITS2)
Channallabes alvarezi 2	Abenelang, Gabon	1°28,92′N,11°35,41′E	Adriaens, Devaere & Herrel	2000	AJ876388 (18S, ITS1, 5.8S, ITS2)

(*continued on next page*)

Table 1(*continued*)

Full species name	Geographic location	Coordinates	Collector	Collection date	Accession Number
Channallabes apus	Kinshasa, D.R. Congo	4°18′S,15°18′E	Biofish Bassleer	1999	AJ876387 (18S, ITS1, 5.8S, ITS2)
Channallabes longicaudatus	Etakaniabe, Gabon	0°32,16′N,12°57,49′E	Adriaens, Devaere & Herrel	2000	AJ876391 (18S, ITS1, 5.8S, ITS2)
Channallabes sanghaensis 1	Ntchouo, River Mbessy, D.R. Congo		Adriaens, Devaere & Herrel	2000	AJ876394 (18S (partial), ITS1, 5.8S, ITS2)
Channallabes sanghaensis 2	Ntchouo, River Mbessy, D.R. Congo	0°46′S–14°32′E	Adriaens, Devaere & Herrel	2000	AJ876393 (18S (partial), ITS1, 5.8S, ITS2)
Gymnallabes					
Gymnallabes typus	Okaka, in delta of the Niger, Nigeria	4°56′N,6°18′E	C.B. Powell, obtained via KMMA/MRAC	01-06-1993	AJ876399 (18S (partial), ITS1, 5.8S, ITS2)

Collector's institutional affiliations are provided in the acknowledgments.

portion of the 28S gene were easy to align, and were used to position the highly variable ITS1 and ITS2 regions. The boundaries of the ITS1 and ITS2 were determined by comparison of the aligned dataset with ITS sequences of fish taxa available in the EMBL databank (e.g., *Cyprinus carpio*–Cyprinidae). The ITS regions were manually optimized based on conservation of both primary sequence data and inferred secondary structural features. The secondary structures of the ITS1 and ITS2 regions were predicted using the Mfold webserver for nucleic acid folding and hybridization prediction (Zuker, 2003) (http://www.bioinfo.rpi.edu/applications/mfold), and compared with published data (Fritz et al., 1994a; May and Coleman, 1997; Morgan and Blair, 1998).

2.3. Sequence and phylogenetic analyses

The Akaike information criterion (AIC) in MODELTEST 3.6 (Posada and Crandall, 1998; Posada and Buckley, 2004) was used to select an appropriate substitution model of DNA evolution. The corresponding nucleotide frequencies, substitution rates and types, and Ti/Tv ratios were used in the neighbour-joining (NJ), maximum-parsimony (MP) and the maximum-likelihood (ML) algorithms in PAUP* 4.0b10 (Swofford, 2003) and Mrbayes 3.0b4 (Huelsenbeck and Ronquist, 2001) to resolve the phylogenetic relationships.

Pairwise sequence divergence data between taxa were computed for the complete 18S–ITS1–5.8S–ITS2–28S (partial) region. Absolute distance values and distances based on a maximum-likelihood distance matrix (PAUP*), with appropriate parameters for the DNA evolution model (MODELTEST) were calculated (data not shown).

After selection of the appropriate substitution model of DNA evolution with corresponding parameters that best fitted the data, the ML analysis was conducted in PAUP*. Heuristic search settings were stepwise taxon addition, TBR branch swapping, MulTrees option in effect, no steepest descent, and rearrangements limited to 10,000. The nonparametric-bootstrap analysis with 100 replicates was used to assess the reliability of individual branches in the phylogenetic trees obtained by heuristic search with stepwise sequence addition (Felsenstein, 1985).

Minimum-evolution analysis was performed with PAUP* by application of the selected ML substitution model to the NJ algorithm. The nonparametric bootstrap analysis used 10,000 replicates to assess the reliability of individual branches in the phylogenetic tree.

Settings of the ML parameters in MrBayes were determined by MODELTEST for each individual data partition. The Bayesian analysis was performed with MrBayes version 3.0b4 (Huelsenbeck and Ronquist, 2001). Parameters for the appropriate substitution model were specified for each data partition to enable site-specific rate variation, allowing a mixed model approach for the heterogeneous dataset. In a first analysis the length of the Bayesian run was tested in order to be certain of convergence. The

Mnsarkov Chain Monte Carlo process was set so that four chains ran simultaneously for 5,000,000 generations, with trees being sampled every 100 generations for a total of 50,000 trees in the initial sample. For the final analysis five independent Bayesian runs where performed in order to confirm that there was adequate convergence and mixing. Each MCMC process started from random starting points and was set so that four chains ran simultaneously for 1,500,000 generations, with trees sampled every 100 generations for a total of 15,000 trees in the initial sample. Variation in the ML scores in the samples was examined by inspecting the MrBayes-logfile, and the position where the ML scores stopped improving was determined. The portion of the trees before the position (tree number) where the ML score stopped improving dramatically and only fluctuated around a plateau, was discarded. The posterior probability of the phylogeny and its branches was determined for all those trees in the plateau phase with nearly the best ML scores. The Bayesian estimates of posterior probability and bootstrap analyses were included to assess support.

Two different types of parsimony analysis were conducted. First, equally weighted MP analyses were performed with PAUP*. Heuristic search settings were: stepwise taxon addition, tree bisection–reconnection branch swapping, multiple trees retained, no steepest descent, rearrangements limited to 100,000, and accelerated transformation. Treating gaps as characters as in Swofford et al. (1996), or Lutzoni et al. (2001) would have provided more information from these sites, but we treated gaps as missing data so that the MP analysis could be directly compared to the ML analyses. The nonparametric-bootstrap analysis used 1000 replicates to assess the reliability of individual branches in the phylogenetic trees obtained by heuristic search with stepwise sequence addition (Felsenstein, 1985). Trees were displayed with TREEVIEW 1.6.6 (Page, 1996).

A second parsimony analysis was done under the optimality criterion of parsimony with equal weights (i.e., gaps, transitions, and transversions all given a weight of 1). The parsimony analysis was conducted using direct optimization and iterative pass as implemented in the program POY (Wheeler et al., 2002) and run on the American Museum of Natural History Parallel Computing Cluster. The analysis began by generating 30 random addition sequences (RAS) per random replicate for five replicates. These 150 RAS were improved with TBR branch swapping during the searches, an additional round of TBR branch swapping and tree fusing (Goloboff, 1999) at the end. These random replicates resulted in one most parsimonious tree. The resulting tree was submitted to POY for further tree searching using the commands "iterative pass" and "exact." This second step of the analysis began by tree fusing (Goloboff, 1999) the submitted topology, and it was followed by an additional round of tree fusing and TBR branch swapping to reduce heuristics in the first-step analysis.

The lengths of the resulting implied alignments were verified in NONA (Goloboff, 1998) and WinClada (Nixon, 2002). To estimate the "robustness" of the clades recovered in the phylogenetic hypothesis, bootstrap percentages (200 replications, 10 random addition sequences per replicate) were calculated in NONA based on the resulting implied alignment. Character evolution on the recovered topology was examined using NONA and WinClada.

2.4. Divergence time estimation

Many methods have been proposed for phylogenetic dating (e.g. Britton et al., 2002; Thorne and Kishino, 2002; Yang and Yoder, 2003). Here the program r8s (Sanderson, 2002, 2003) was used to do the dating. r8s performs a semiparametric rate smoothing, using a penalized likelihood approach applied to the distances inferred from the ML tree with branch lengths. It combines a model-based likelihood approach with a roughness penalty that prevents too much rate variation across the tree. A cross-validation procedure is done to obtain a smoothing parameter that specifies the size of the roughness parameter. Cross validation was performed for the trees with branch lengths, obtained by multiplying the per site values as reported by PAUP* with the number of sites. This procedure provides an objective method for model selection and choice of optimal smoothing value (Sanderson, 2002, 2003). Comparison of seven independent dating analyses (ML trees with smoothing factors 0.63, 1.00, 1.58, 2.51, 3.98, 6.31, and 10.00), using optimal (= lowest) and sub-optimal smoothing factors, was used as a measure of confidence; average node date and confidence intervals (SD values) were calculated (Dumont et al., 2005).

We used two reference fossils, the first appearance of the African Clariidae in the Lower Eocene (34–56 MYA, 45 MYA average) and the first appearance of *Heterobranchus* in the Miocene (16–23 MYA, 19.5 MYA average) to estimate divergence times (data obtained from Gayet and Meunier, 2003). All fossil calibration points were used simultaneously, using the "fixage" command for the Miocene *Heterobranchus* group, and the "constrain" command with "minage" and "maxage" limits for the Eocene Clariida. Alternatively, we tested dating each with one fossil as calibration point using the "fixage" command, and the other one as a constraint specifying minimum ages using the "constrain" command.

3. Results

3.1. Sequence analysis and alignments

Length of ribosomal genes (18S, 5.8S) and spacers (ITS1, ITS2) are given in Table 2. The length of the 18S and 5.8 genes showed little variation (1869–1870 bp in the ingroup, 1836–1882 bp in the outgroup and 156–157 bp in the ingroup, 157–158 bp in the outgroup, respectively). Length variation in the spacer regions was much higher (ITS1: ingroup, 454–567; outgroup, 368–409; ITS2: ingroup, 384–490; outgroup, 310–386). The GC percentage shows the

Table 2
Length and G+C content of the ribosomal 18S, ITS1, 5.8S and ITS2 DNA regions for the taxa used in this study

Taxon	18S		ITS1		5.8S		ITS2		28S	
	GC%	Length	GC%	Length	GC%	Length	GC%	Length	GC%	Length
Clupea harengus	0.5418	1836	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Cyprinus carpio	0.5368	1882	0.6603	368	0.5759	158	0.7746	386	0.6538	26
Ictalurus punctatus	0.5551	1880	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Kryptopterus bicirrhis	0.5600	1870	0.5369	406	0.5860	157	0.7316	339	0.6154	26
Pangasianodon hypophthalmus	0.5545	1873	0.5278	396	0.5860	157	0.6581	310	0.6154	26
Heteropneustes fossilis	0.5531	1869	0.6479	409	0.5860	157	0.6862	341	0.6154	26
Clarias pachynema 1	0.5538	1869	0.6854	553	0.5860	157	0.6835	436	0.6154	26
Clarias submarginatus	n.a.	n.a.	0.6866	552	0.5860	157	0.6835	436	0.6154	26
Clarias theodorae 1	n.a.	n.a.	0.6838	544	0.5860	157	0.6902	468	0.6154	26
Clarias theodorae 2	n.a.	n.a.	0.6915	564	0.5860	157	0.6880	468	0.6154	26
Clarias buthupogon	0.5548	1869	0.6750	560	0.5860	157	0.6798	456	0.6154	26
Clarias gariepinus	0.5545	1869	0.6872	567	0.5860	157	0.6793	449	0.6154	26
Clarias ngamensis 1	0.5543	1869	0.6710	567	0.5860	157	0.6839	446	0.6154	26
Clarias ngamensis 2	n.a.	n.a.	0.6721	567	0.5860	157	0.6839	446	0.6154	26
Clarias stappersi	0.5554	1869	0.6970	472	0.5860	157	0.6848	441	0.6154	26
Clarias jaensis	0.5554	1869	0.6826	523	0.5732	157	0.6339	439	0.6154	26
Clarias platycephalus	n.a.	n.a.	0.6493	499	0.5669	157	0.6563	384	0.6154	26
Clarias camerunensis	n.a.	n.a.	0.6852	521	0.5860	157	0.7114	440	0.6154	26
Heterobranchus isopterus	0.5545	1870	0.6888	556	0.5860	157	0.7157	490	0.6154	26
Dinotopterus cunningtoni	0.5548	1869	0.6795	567	0.5860	157	0.6906	459	0.6154	26
Tanganikallabes mortiauxi	0.5543	1869	0.6827	498	0.5860	157	0.6879	439	0.6154	26
Clariallabes longicauda 1	0.5548	1869	0.6761	528	0.5860	157	0.7047	430	0.6154	26
Clariallabes longicauda 2	0.5554	1869	0.6471	529	0.5860	157	0.6912	434	0.6154	26
Channallabes alvarezi 1	0.5556	1869	0.6858	522	0.5860	157	0.7193	431	0.6154	26
Channallabes alvarezi 2	n.a.	n.a.	0.6597	523	0.5860	157	0.7186	430	0.6154	26
Channallabes apus	n.a.	n.a.	0.7117	555	0.5860	157	0.6761	457	0.6154	26
Channallabes longicaudatus	0.5554	1869	0.6772	474	0.5860	157	0.7172	435	0.6154	26
Channallabes sanghaensis 1	n.a.	n.a.	0.6524	492	0.5796	157	0.6715	417	0.6154	26
Channallabes sanghaensis 2	0.5559	1869	0.6454	485	0.5732	157	0.6424	439	0.6154	26
Gymnallabes typus	n.a.	n.a.	0.6366	454	0.5705	156	0.7015	402	0.6154	26

same pattern of variation, being higher in spacer regions than in the genes (Table 2). Both spacer regions have almost equal GC percentages. The sequences of representatives of *Clarias ngamensis* originating from different geographical locations were almost identical, but with a slight difference in ITS1 GC percentage. In *C. theodorae* the length and GC percentage varied considerably between the two specimens, which was even larger in the two *C. pachynema* specimens.

3.2. Phylogenetic analysis

A first dataset, including the 18S sequences only, comprised two new (*Kryptopterus bicirrhis*—Siluridae and *Pangasianodon hypophthalmus*—Pangasiidae) and three Genbank sequences (*Clupea harengus*—Clupeidae, *Cyprinus carpio*—Cyprinidae, and *Ictalurus punctatus*—Ictaluridae) as an outgroup, and 20 taxa as an ingroup. Phylogenetic analyses using four methods, NJ, MP, ML, and Bayesian, showed a rather consistent topology (Fig. 1). The studied ingroup catfish (Clariidae) are a monophyletic group, well separated from the outgroup specimens. *Heteropneustes* is sister group to these Clariidae. The 18S gene tree (Fig. 1, Table 2) and combined 18S + 5.8S gene tree (not shown) did not show enough variation to resolve relationships within the ingroup, as could be expected since 18S is known to be suited for relationships at species level or higher (Weekers et al., 1994, 2002). The choice of the outgroup taxa, i.e., *Cyprinus carpio*, *Kryptopterus bicirrhis*, and *Pangasianodon hypophthalmus*, is valid; these taxa could be used with confidence as outgroup in the analysis of the ITS1–5.8S–ITS2 dataset. The ingroup of African Clariidae is monophyletic, which is confirmed by all analyses of both the 18S-dataset and the dataset containing the complete sequences (18S-ITS1–5.8S–ITS2–28S) (latter not shown).

The second dataset contained the ITS1, 5.8S and ITS2 sequences of all ingroup and four outgroup taxa, one being a GenBank sequence (*Cyprinus carpio*), others being newly sequenced (*Kryptopterus bicirrhis*, *Pangasianodon hypophthalmus*, and *Heteropneustes fossilis*). The phylogenetic analysis was performed using the five different methods (Figs. 2A–D and 3). A consensus is presented in Fig. 4. The neighbor-joining analysis with distance measurement set to maximum likelihood used a gamma correction for among site rate variation (GTR+G), with the following values: R = (1.4238, 2.2369, 1.3615, 1.0339, and 2.2369), and gamma shape parameter = 0.7744. The tree is shown in Fig. 2A. The maximum likelihood analysis with GTR+G, gamma correction, R = (1.4238, 2.2369, 1.3615, 1.0339, and 2.2369), and gamma shape parameter = 0.7744 generated a tree with a log likelihood of −9874.58847. Branch lengths were corrected according to the settings mentioned above. A 50%

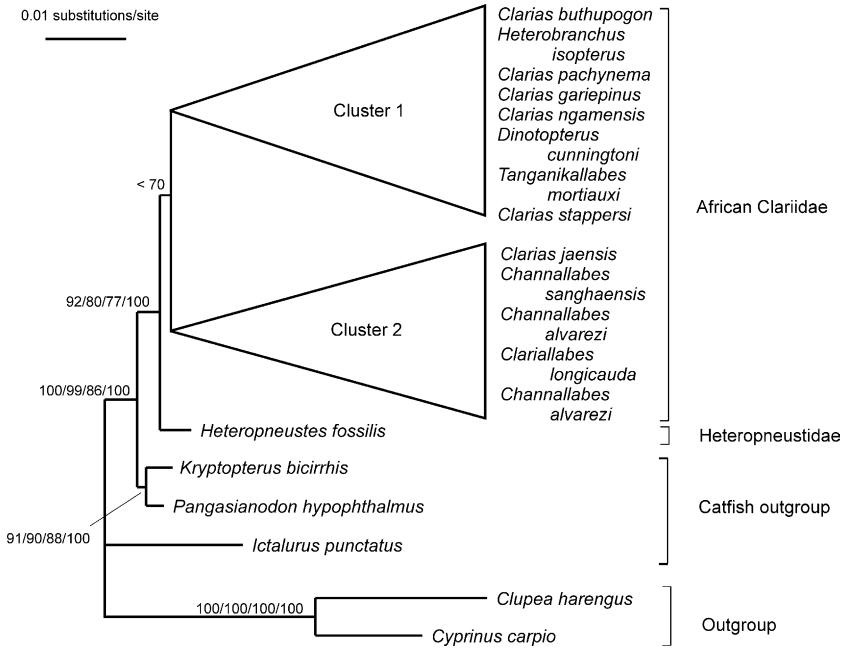

Fig. 1. Consensus tree of the phylogenies based upon 18S data. Methods are described individually in Fig. 2.

majority rule consensus tree from 127 trees is shown in Fig. 2B. Five independent Bayesian runs where performed, and in each run one out of every 100 trees was sampled for 15,000,000 generations. The point of stationary ML scores ("burnin") was after 350 trees; these first 350 trees were discarded and the posterior probability of the phylogeny was determined from 14,650 trees. A 50% majority-rule consensus tree with Bayesians probability values (Fig. 2C) was calculated in PAUP* using the 14,650 trees with lowest ML scores. Branch lengths were corrected for ML distances using the GTR+G model. The maximum parsimony analysis generated two most parsimonious trees (MPTs) of 1719 Steps (CI = 0.7533, RI = 0.6570, RC = 0.4949). Bootstrap method with heuristic search, stepwise taxon addition, TBR branch swapping, MULTREES option, no steepest descent, rearrangements limited to 100,000, and accelerated transformation. Gaps were treated as missing data. Bootstrap support, calculated from 1000 replicates, is expressed as percentage (Fig. 2D).

Heteropneustes fossilis is a close relative to the studied Clariidae (BS = 100, 100, 100, 89, and 100; NJ, MP, POY, ML, and MrB, respectively), as could be expected, but is generally not considered part of it (Diogo, 2005; Teugels and Adriaens, 2003).

The first clade (cluster 1a) is ambiguously supported (NJ, MP, ML, respectively, 44, 47, 52, and MrB, POY, respectively, 97, 100), but still suggest a sister group relationship between *Clarias buthupogon–Channallabes apus* and *Clarias pachynema–Clarias submarginatus*. Cluster 1b unites the two specimens of *Clarias ngamensis* with *Clarias gariepinus*, *Dinotopterus cunningtoni*, and *Heterobranchus isopterus*. This clade is well supported by all methods (Bootstrap support for NJ, MP, POY, ML, and MrB, respectively, 73, 84, 100, 79, and 100).

In clade 2a *Clarias camerunensis* is the sister group of two specimens of *Clariallabes longicauda*, and two specimens of *Channallabes alvarezi* always cluster together. The most basal taxon is *C. stapersii* in MP, ML, and MrB. In NJ, however, it is the sister group to the *C. theodorae* specimens. *Channallabes longicaudatus* appears in the clade, twice as a sister group of the *Channallabes alvarezi* specimens (MP, MrB), once as an outgroup of the rest of the 2a clade (NJ, where *C. stappersi* is not included in the cluster), once as the sister group of the subclade of *C. camerunensis* and *C. longicauda* (ML) and once as the closest relative of *Clarias stappersi* forming the sister group of *Channallabes alvarezi* (POY). We can find a monophyletic grouping (cluster 2b) of the two specimens of *Channallabes sanghaensis* and *Clarias platycephalus* in the POY analysis. However, one of the two *Channallabes sanghaensis* specimens is most basal in the cluster in the NJ, MP, ML, and MrB trees. The NJ, MP, ML, and MrB analyses all show, contrary to the POY tree, *Clarias jaensis* as an outgroup of *Clarias platycephalus* and one of the *Channallabes sanghaensis* specimens (BS = 98, 93, 90, and 100). The two representatives of *C. theodorae*, originating from very distant locations, always cluster together but differ in position. The positions of *C. stappersi*, *C. jaensis* and *Tanganikallabes mortiauxi* remain uncertain, and the two *C. pachynema* taxa do not cluster together. The representative from Ebeigne, Northern Gabon (*C. pachynema 2*) is always a member of cluster 1a, whereas the other one, from Mopia, Southern Gabon, does not have an unambiguous position in the phylogenies; it occurs as a sister group of all taxa of the ingroup except 1a, of 2a + 2b in MP, of 1a + 1b + *C. theodorae* in ML or of 1a + 1b + *C. theodorae* + *T. mortiauxi* in MrB.

In POY, the most primitive taxon of the studied ingroup appears to be *Gymnallabes typus*, a situation that is not

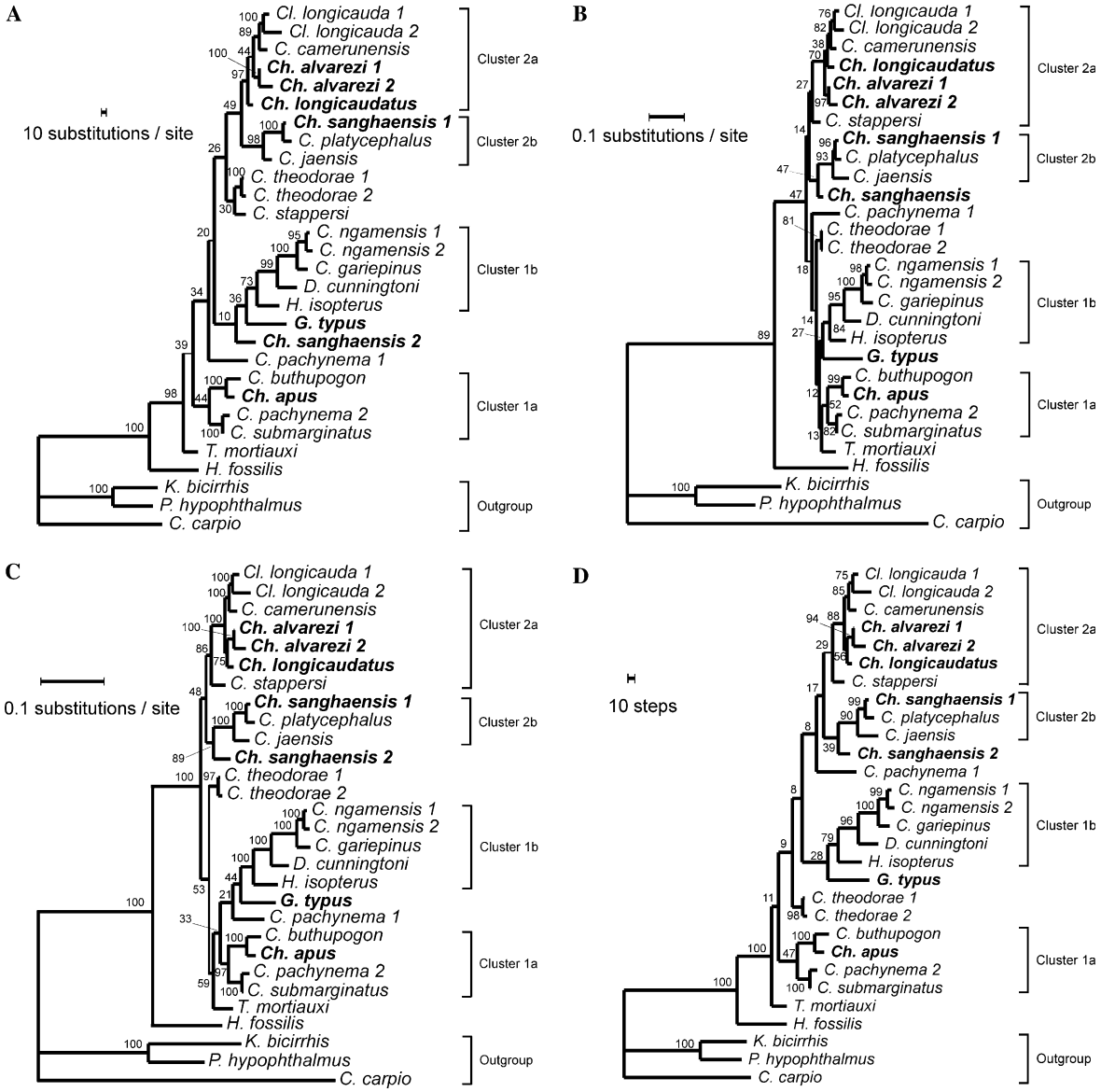

Fig. 2. Phylogenetic trees based on sequences of the ribosomal ITS1–5.8S–ITS2 region. (A) Neighbor-joining estimate of the phylogeny. The scale bar represents 10 substitutions/site, and bootstrap support, calculated from 10,000 replica's, is expressed as a percentage. Anguilliforms are indicated in bold. (B) Maximum likelihood estimate of the phylogeny showing a 50% majority rule consensus tree from 127 trees. The scale bar represents 0.1 substitutions/site, and bootstrap support, calculated from 100 replicates, is expressed as percentage. Anguilliform taxa are indicated in bold. (C) Bayesian probability estimate of the phylogeny showing a 50% majority-rule consensus tree with Bayesians probability values. The tree is rooted with *Cyprinus carpio*, *K. bicirrhis*, *P. hypophthalmus*, and *Heteropneustes fossilis* as outgroup. Branch lengths are corrected for ML distances using the GTR+G model. The scale bar represents 0.1 substitutions/site. The numbers along the branches are support indicated by Bayesian probability analysis, which are expressed as percentage. Anguilliform taxa are indicated in bold. (D) Maximum parsimony estimate of the phylogeny; the analysis generated two most parsimonious trees (MPTs) of 1719 Steps. Bootstrap support, calculated from 1000 replicates, is expressed as percentage. The scale bar represents 10 steps. Anguilliform taxa are indicated in bold.

sunspported by the other four analyses, where it clusters with 1b (NJ, MP, ML, and MrB).

3.3. Dating analysis

Application of the two available fossil calibration points in the penalized likelihood procedure applied to the ML tree, provided us with a range of data. Initial results were obtained with the default settings for dating analysis in the r8s program, with cross validation function enforced. The rate smoothing parameters with optimal (=lowest) and sub-optimal cross validation scores were selected, and the dating procedure was then repeated. The result of the time divergence estimation is shown in Fig. 5, and age estimates with rather low SD values for all internal nodes are shown in Table 3. The analyses using the ML tree topology and different smoothing factors, resulting from optimal and sub-optimal cross validation scores, yielded small deviations in age estimates. The use of several reference fossils is expected to reduce variation due to error. The age

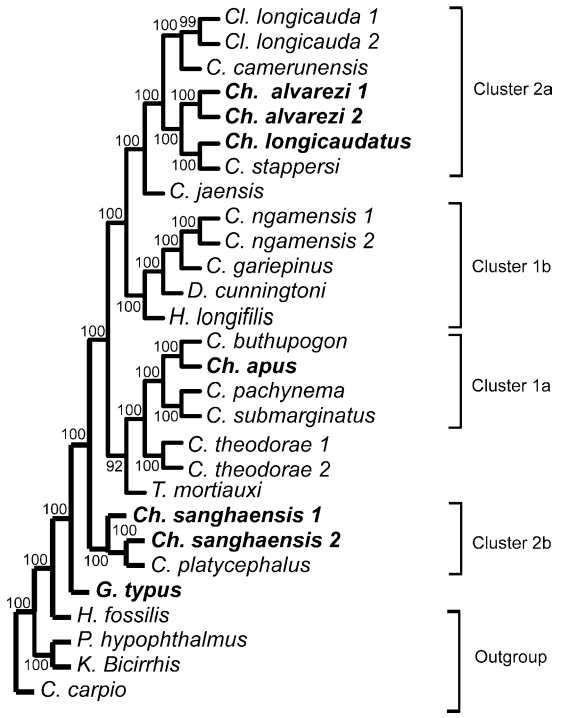

Fig. 3. Estimate of the ITS1–5.8S–ITS2 phylogeny tree under the optimality criterion of parsimony with equal weights, i.e., gaps, transitions, and transversions all given a weight of (1) using POY. Random replicates resulted in one most parsimonious tree. The resulting tree was submitted to POY for further tree searching using the commands "iterative pass" and "exact." 993 of 1900 characters were phylogenetically informative, CI = 0,80; RI = 0,85; TL = 2823 steps. Anguilliform taxa are indicated in bold.

constraints imposed by the two fossils in different parts of the tree are likely to restrict variation caused by a variety of other factors. Divergence time estimation based on the ribosomal ITS1–5.8S–ITS1 region and fossil constraints indicate that the African Clariidae started radiating between 123 and 56 My ago. *Heteropneustes fossilis* (a close relative to Clariidae) is older (about 123 My) and may be a survivor of old clades that went extinct at the K/T boundary. The extant Clariidae originated in the Miocene, and much of the species richness probably extended much wider in preglacial times.

4. Discussion

4.1. Sequence variation

The length and GC% of the 18S and 5.8S genes of the studied Clariidae fall within the range of known sequences available in GenBank/EMBL. The same goes for the length and GC content of the spacers. A comparable GC% between ITS1 and ITS2 is considered to be a general property, as suggested by Torres et al. (1990). The GC% is quite high, and could be due to the fact that these animals occur in the tropics: higher GC content can be an adaptation to warmer climates, as has been shown in tetrapods (Bernardi et al., 1985).

4.2. Phylogenetic analysis

The 18S dataset and complete sequence analysis (data not shown) show that the African Clariidae considered here form a monophyletic group, well separated from the outgroup. *Heteropneustes* is the sister group of these African Clariidae. The 18S region contained insufficient information to resolve the terminal clades.

The analyses of the ITS1–5.8S–ITS2 dataset still supports *Heteropneustes* being the sister group to the studied Clariidae, which confirms the results of Teugels and Adriaens (2003) but could contradict a recent phylogenetic analysis based on morphology (Diogo, 2005). In the latter, *Heteropneustes* is considered as the sister group of all Clariidae except *Uegitglanis* (the latter genus is not included in our study).

According to the POY analysis, the most basal taxon within the Clariidae is *Gymnallabes typus*, an eel-like species occurring in the Niger delta. In all other analyses, however, it is the sister group to cluster 1b. It seems very unlikely that anguilliformity would be the plesiomorphic condition for the clariid family as the outgroup taxa are non-anguilliform and as this would imply a secondary reduction of body length, with a subsequent elongation. Further sampling, both of specimens and of gene sequences, might resolve this.

Clade 1a, containing two different clusters (*C. pachynema 2*, *C. submarginatus*, and *C. buthupogon*, *Channallabes apus*) was consistent in every analysis. *Clarias pachynema*, *C. submarginatus*, but also *C. theodorae* all belong to the *Clarias* (*Anguilloclarias*) subgenus sensu Teugels (1986), thus rendering the subgenus paraphyletic. In Teugels and Adriaens (2003) and Agnèse and Teugels (2005), *Clarias* (*Anguilloclarias*) was considered the sister clade of *Gymnallabes*, but this is not found in any of the trees presented here.

The study of Agnèse and Teugels (2005), using the cytochrome b gene (Cyb), did not yield a close relationship between *C. theodorae* (see below) and *C. pachynema*. Agnèse and Teugels (2005) used four representatives of the subgenus *Clarioides*, which formed a monophyletic group. Teugels and Adriaens (2003) and Agnèse and Teugels (2005) both show a close relationship between the *Clarioides* species and *Channallabes apus*, which is confirmed here. A close relationship between *Clarioides* and *Clariallabes s.l.*, as suggested by Poll (1942), could not be corroborated.

Cluster 1b is a very interesting one, uniting two specimens of *C. ngamensis* with *C. gariepinus*, *Dinotopterus cunningtoni* and *Heterobranchus isopterus*. The close relationship between *Clarias* (*Clarias*), comprising *C. gariepinus* and *C. anguillaris*, *Clarias* (*Dinotopteroides*), comprising *C. ngamensis*, and *Heterobranchus*, was also suggested by Teugels and Adriaens (2003) based on morphological evidence. It has also been shown that *Clarias* (*Clarias*) is the ancestral group of *Bathyclarias*, as a result of a recent speciation event in Lake Malawi (Agnèse and Teugels, 2001b). The close relationship between *Clarias*

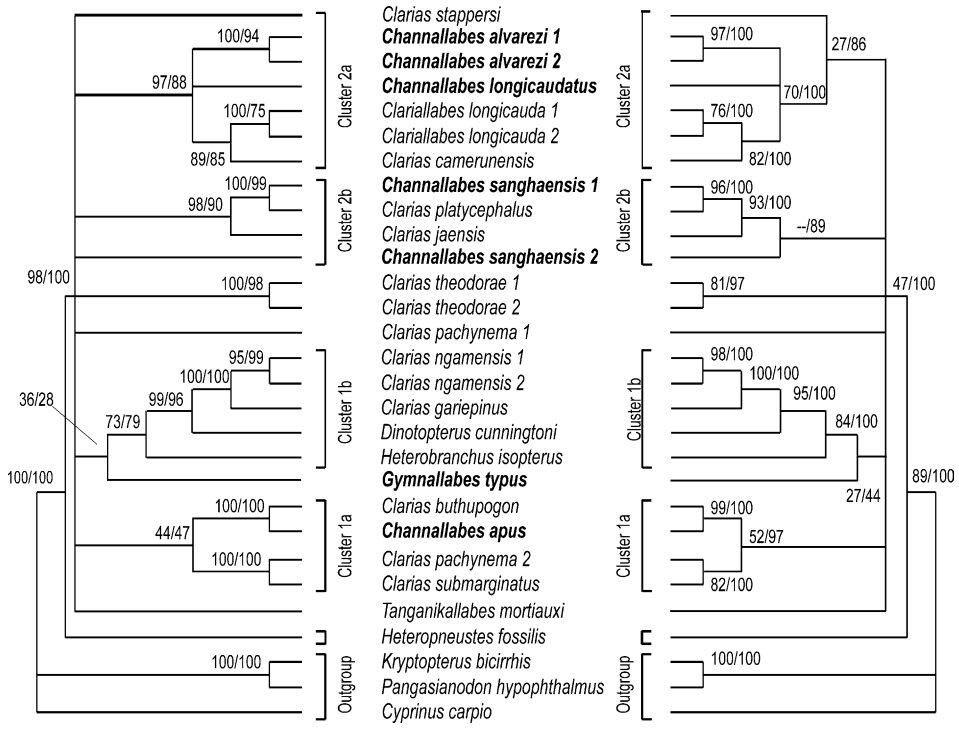

Fig. 4. Consensus tree of four analyses (NJ, MP, ML, MrB) based on sequences of the ribosomal ITS1–5.8S–ITS2 region. Methods are described individually in Fig. 2. Only nodes supported by all methods are shown. Left-hand side: consensus of NJ and MP; right-hand side: consensus of ML and MrB analyses. Anguilliform taxa are indicated in bold.

(*Clarias*) and *Heterobranchus* was also confirmed by Legendre et al. (1992), Teugels et al. (1992), and Agnèse and Teugels (2005). Finally, Agnèse and Teugels (2005), Teugels and Adriaens (2003) and Graham (1997) also found a sister group relationship between *Heterobranchus* and *Dinotopterus*. This clade is thus supported both by general morphology, morphology of the suprabranchial organ, cytochrome *b* and nuclear ribosomal DNA analyses.

The two *C. theodorae* specimens (one from Zambia and one from Botswana) cluster together; their relationship to the other taxa remains ambiguous (Figs. 2A–D and 3). *Tanganikallabes* was considered part of a trichotomy with *Gymnallabes* and *Clariallabes* by Graham (1997). This relationship can't be confirmed here. The apparent polyphyletic nature of *C. pachynema* remains unexplained at this point, and deserves further study with additional specimens.

Cluster 2a shows two monophyletic groups. In the first subclade, *Clariallabes longicauda* is the closest relative to *C. camerunensis* (subgenus *Clarias* (*Brevicephaloides*)). This is in contrast to what has previously been suggested: Poll (1942) placed *Clariallabes* with *Clarias* (*Clarioides*); Graham (1997) considered it within a polytomy with *Channallabes* and *Dolichallabes*. However, Agnèse and Teugels (2005) saw a close relationship between the *Brevicephaloides* and the *Platycephaloides* species, where the sister clade of *C. jaensis* and *Clariallabes longicauda* is nested within representatives of both the *Brevicephaloides* and *Platycephaloides* subgenera.

The second subclade of cluster 2a unites two monophyletic groups according to the POY-analysis, the first one comprising two *Channallabes alvarezi* specimens (supporting their designation to a new species), and the second one comprising *Clarias stappersi* and *Channallabes longicaudatus*. In the other analyses, the second subgroup consists of the two *Channallabes alvarezi* specimens in relation to *Channallabes longicaudatus* at variable positions. Again, it seems that another lineage of anguilliform species (*Channallabes longicaudatus* and *Channallabes alvarezi*) has a close relationship with a non-anguilliform *Clarias* species, this time a representative of the polyphyletic subgenus *Platycephaloides* (see above).

The last cluster (2b) unites taxa from the Congo region, i.e., *Clarias platycephalus* with two specimens of *Channallabes sanghaensis*. Remarkably, even though both *Channallabes sanghaensis* specimens come from the same region, one of the specimens seems to be more closely related to *C. platycephalus* than to the other *Ch. sanghaensis* specimen. Even more surprising is that in the NJ, MP, ML and MrB trees, the second *Ch. sanghaensis* forms the sister taxon of the latter two with *Clarias jaensis*. Both *Ch. sanghaensis* specimens were collected by local fishermen, however, no information on the exact localities could be obtained. As a consequence, they could originate from different localities, hence they show genetic variability (but any assumption on whether they represent two different species is too speculative at this point). Both *C. jaensis* and *C. platycephalus* are classified in the subgenus *Platycephaloides*, together with C. *stappersi* (cluster 2a), which implies a polyphyletic nature of this subgenus sensu Teugels (1986). In Agnèse and Teugels (2005), *C. jaensis* forms a monophyletic group

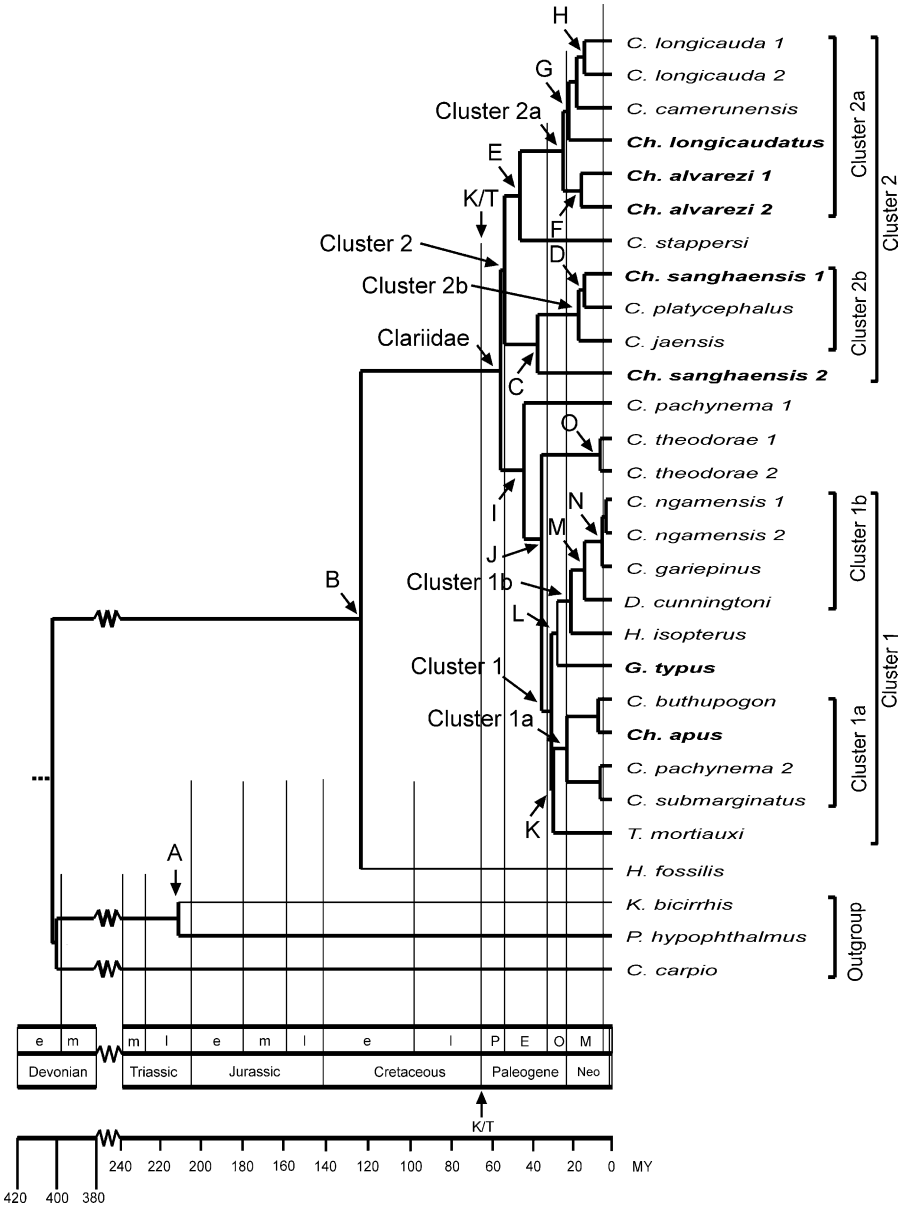

Fig. 5. Dated phylogenetic tree of African Clariidae obtained from semiparametric rate smoothing (penalized likelihood; Sanderson, 2002, 2003) of a 29-taxon phylogenetic tree, with maximum likelihood branch lengths from sequences of the ribosomal ITS1–5.8S–ITS2 region, and calibrated with two reference fossils. Reference fossils are listed in Table 3, and further information on age estimates is found in Table 3. Anguilliform taxa are indicated in bold. The time scale shows ages in million years (My) before present. The bar represents a stratigraphy showing geological era, periods and epoches. E, eocene; O, oligocene; M, miocene; e, early; m, middle; l, late; K/T, cretaceous/tertiary boundary.

together with *Clariallabes longicauda* and *Clarias* (*Brevicephaloides*) (with a.o. *C. camerunensis*), which is not confirmed in this study.

The splitting of the original *Channallabes apus* into six species, i.e., *Channallabes apus, Ch. longicaudatus, Ch. alvarezi, Ch. sanghaensis, Ch. ogoouensis* and *Ch. teugelsi*, was originally based on morphological data (Devaere et al, submitted) and is confirmed by these molecular analyses (with exception for *C. ogooensis*, for which tissue was not available for this study).

Even though the clades mentioned above are well defined and largely supported by all topologies, the mutual relationships of the clades is less straight forward. It is even so that the topologies of the ML and Bayesian analysis seem to present an opposite result to that of the POY, MP and NJ analysis. The bootstrap supports for the more basal nodes are very low in all analyses, except for POY. During direct optimization, the POY alignment is influenced by the relationships on the phylogeny. Therefore indices of support calculated from that combination of phylogeny and alignment can be expected to be high. The lower support of trees (except POY) may be the result of rapid speciation events that occurred in a very short interval of time. When we look at the consensus data of the NJ, MP, ML, and MrB analyses, the basal nodes all collapse (Fig. 4). In this case both topologies show a consistent pattern, with, most importantly, the same terminal clades.

Table 3
Stem node ages (sna) and crown node ages (cna) in My for mayor clusters and families of African catfish obtained from semiparametric rate smoothing (penalized likelihoods; Sanderson, 2002, 2003) of a 29-taxon catfish tree, with maximum likelihood branch lengths from the ribosomal ITS1-5.8S-ITS2 region, and calibrated with two reference fossils

Cluster (node)	sna		cna	
	Age	SD	Age	SD
Outgroup	—	—	403.41	20.46
A	403.41	20.46	213.75	10.59
B	403.41	20.46	122.87	4.88
Clariidae	122.87	4.88	56.00**	0
Cluster 2	56.00**	0	53.80	0.09
C	53.80	0.09	37.62	0.32
Cluster 2b	37.62	0.32	16.36	0.28
D	16.36	0.28	13.77	0.61
E	53.80	0.09	46.17	0.47
Cluster 2a	46.17	0.47	23.96	1.07
F	23.96	1.07	14.44	1.22
G	23.96	1.07	20.96	0.93
H	17.30	0.69	13.11	0.44
I	56.00**	0	43.81	0.18
J	43.81	0.18	35.06	0.30
Cluster 1	35.06	0.30	30.69	0.27
K	30.69	0.27	29.48	0.30
Cluster 1a	29.48	0.30	22.60	0.44
L	30.69	0.27	27.25	0.26
Cluster 1b	27.25	0.26	20.00*	0
M	20.00*	0	12.96	0.40
N	12.96	0.40	4.99	0.35
O	35.06	0.30	5.65	0.79

The age with an asterisk is the average fossil date (see Table 3), set to fixed age, and the age with two asterisks is the minimum–maximum constrained fossil date in the divergence time estimation.

4.3. Divergence time estimation

The divergence time estimation, based on ML distances, is shown in Fig. 5. Stem and crown ages are given in Table 3, including the standard deviation intervals. The table indicates that these are not too broad. According to our analyses the African Clariidae used in this study have emerged between 56 MYA (crown node age) and 123 MYA (stem node age). Two hypothetical scenarios for the timing of speciation events are envisaged here.

A late radiation would have occurred around the K/T crisis, starting about 65 millions years ago. Rapid speciation of the Clariidae during this event could explain the ambiguity and low bootstrap support of the nodes connecting the terminal clusters (1a,b and 2a,b).

However, since the currently available data from the fossil record does not allow a more accurate estimation of the origin of oldest clariids (i.e. between 34 and 56 MYA), the split between Clariidae and the monotypic Heteropneustidae, and thus the origin of the Clariidae, may have occurred somewhere between 123 and 56 MYA. A possible origin of the Clariidae as early as 123 MYA would be possible, taking into consideration that in our analysis two basal clariid taxa are not included. Phylogenetic analyses, based on morphology, suggested a most basal position of the African, monotypic *Uegitglanis* within the Clariidae (De Pinna, 1993; Diogo, 2005). A recent molecular analysis has shown that Asian *Clarias* species are the sister group of all African clariids (Agnèse and Teugels, 2005). Both taxa were not used in this study because of unavailability of tissue (*Uegitglanis*), whereas Asian clariids were not used in this study because of the focus on the phylogeny of African, anguilliform taxa. Subsequently, the period between 123 and 56 MYA may have included the speciation of those basal groups (the Asian clariids also show a substantial number of species) (Teugels, 2003).

If the latter hypothesis would be correct, and the origin of Clariidae and Heteropneustidae is to be situated during the early Cretacious, the extant *Heteropneustes fossilis* may thus be the sole survivor of an old clade that could have gone extinct, e.g., at the K/T boundary. The fact that *Heteropneustes fossilis* has a morphology resembling that of the generalized clariids quite strikingly (according to some authors, it should even be considered a clariid, e.g., De Pinna, 1993; Diogo, 2005; Shrinivashagar, 1958), the origin of the first Clariidae could very well have been closer to the stem node than the crown node, and thus quite earlier than the K/T-boundary. To test the validity of these two hypothesis, a phylogenetic analysis including all the Asian representatives and *Uegitglanis* would have to be performed.

The position of *Gymnallabes typus* according to the POY analysis seems to indicate an early split towards anguilliformity. However, the other analyses place it as a sister group of cluster 1b. A possible explanation for the ambiguous position could be that the lineage towards *Gymnallabes typus*, as represented here, only reflects one branch of evolution towards anguilliformity and that sister taxa within the lineage are missing (only one specimen from the Niger basin was used) or have become extinct. Other studies place *Gymnallabes* close to *Channallabes* (Poll, 1942), *Clariallabes* and *Tanganikallabes* (Graham, 1997); or close to *Platyallabes*, *Tanganikallabes* and *Clarias* (*Anguilloclarias*) (Teugels and Adriaens, 2003). According to Agnèse and Teugels (2005), *G. typus* is nested within a clade representing the *Clarias* (*Anguilloclarias*) subgenus.

Within the clariid family, body elongation occurred at least four times according to this analysis (Fig. 5): (1) the branch to *Gymnallabes typus*, (2) the branch of *Channallabes apus* in cluster 1a, (3) the branch of *Channallabes sanghaensis*, and (4) the branch to *Channallabes alvarezi* and *Ch. longicaudatus*. The DTE does not provide an exact dating for these morphological events; it could lie anywhere between the stem and crown node ages. Clades with multiple anguilliform taxa (nodes H and F) give a minimum age as they probably share a common anguilliform ancestor. Clades uniting a fusiform with an anguilliform (nodes D and L) provide maximum dates, assuming that fusiform morphology is not the result of a secondary shortening of the body. The branch of *Clariallabes longicauda* (node H) is a special case, since the latter is morphologically intermediate between fusiforms and anguilliforms: infraorbitalia and suprapraeopercularia are slightly reduced, the jaw muscles

are slightly hypertrophied and, as could be anticipated, their length is intermediate (Boulenger, 1902; Cabuy et al., 1999). However, it does not seem to be an evolutionary intermediate, as fully elongated anguilliforms in this analysis never emanated from this lineage (more *Clariallabes* species would have to be included to support this). The taxonomic position of *Clariallabes longicauda* has thus always been a problematic puzzle. Originally, it was described by Boulenger (1902) as *Allabenchelys longicauda*, distinct from *Clariallabes*, and intermediate between *Clarias* and *Clariallabes s.s.* David (1935) placed it in the genus *Clarias*, i.e., in *Clarias* (*Allabenchelys*), but since Teugels (1986) it is transferred to *Clariallabes*. Here it seems to have a close relationship with *Clarias camerunensis*, but the exact relationship with the other *Clariallabes* species remains unresolved.

The rather unexpected polyphyly of *Channallabes sanghaensis* could indicate that the inclusion of more samples could be useful to come to a complete understanding of the species complex. This could clear out at least some of the problematic relationships concerning, e.g., *C. pachynema* and *G. typus* as well.

Acknowledgments

We wish to thank Jean-François Agnèse of the University of Montpellier 2 (France), Roger Bills and Paul Skelton of the South African Institute for Aquatic Biology, and the Royal Museum for Central Africa (Belgium) for providing tissue samples. Leo Smith (American Museum for Natural History) helped us with the POY analysis. We are grateful to Mireille Gayet, Jean Gaudant and René Zaraguëta i Bagils (MNHN, France), who helped us with fossil data. We would also like to thank Melanie Stiassny (New York), Carl Hopkins (Cornell) and IRAF (Gabon) for help during the collections of the samples. We thank three anonymous reviewers for their useful comments. Research was funded by FWO (G.0388.00).

References

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

Agnèse, J.F., Teugels, G.G., 2001b. The *Bathyclarias–Clarias* catfish species flock. A new model to understand rapid speciation in African Great Lakes. Compt. Rend. Acad. Sci. Paris Sciences de la vie 325, 683–688.

Agnèse, J.F., Teugels, G.G., 2005. Insight into the phylogeny of African Clariidae (Teleostei, Siluriformes): Implications for their body shape evolution, biogeography, and taxonomy. Mol. Phylogenet. Evol. 36, 546–553.

Arnheim, N., 1983. Concerted evolution of multigene families. In: Nei, M., Koehn, R.K. (Eds.), Evolution of genes and proteins. Sinauer, Sunderland, MA, pp. 38–61.

Bernardi, G., Oloffson, B., Filipsli, J., Zerial, M., Salinas, J., Cuny, G., Meunier-Rotival, M., Rodier, M., 1985. The mosaic genome of warm-blooded vertebrates. Science 228, 935–958.

Boulenger, G.A., 1902. Matériaux pour la faune du Congo, additions à la faune ichthyologique du basin du Congo. Ann. Mus. Congo 2 (1), 19–57. pl. VII–XV.

Britton, T., Oxelman, B., Vinnersten, A., Bremer, K., 2002. Phylogenetic dating with confidence intervals using mean path-lengths. Mol. Phylogenet. Evol. 24, 58–65.

Burgess, W.E., 1989. An atlas of Freshwater and Marine Catfishes. A preliminary survey of the Siluriformes. Berkshire TFH Publications, Neptune City, NJ.

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*. J. Morph. 240, 169–194.

David, L., 1935. Die Entwicklung der Clariiden und ihre Verbreitung. Rev. Zool. Bot. Afr. 28, 77–147.

De Pinna, M.C.C., 1993. Higher-level phylogeny of Siluriformes, with a new classification of the order (Teleostei, Ostariophysi), New York City, University of New York, Unpublished PhD thesis.

Devaere, S., Adriaens, D., Verraes, W., Teugels, G.G., 2001. Cranial morphology of the anguilliform clariid *Channallabes apus* (Günther, 1873) (Teleostei: Siluriformes): adaptations related to a powerful biting? Zool. J. Linn. Soc. 255, 235–250.

Diogo, R., 2005. Morphologicl evolution, aptations, homoplasies, constraints and evolutionary trends. Catfishes as a case study on general phylogeny and macroevolution. Science Publishers Inc., Enfield, USA.

Dumont, H.J., Vanfleteren, J.R., De Jonckheere, J.F., Weekers, P.H.H., 2005. Phylogenetic relationships, divergence time estimation and global biogeographic patterns of Calopterygoid Damselflies (Odonata, Zygoptera) inferred from ribosomal DNA sequences. Syst. Biol. 54 (3).

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

Fritz, H.J., Brundl, K., Frisch, C., Kolmar, H., 1994a. Immunoglobulin folding stability genetically screened in *Escherichia coli* and construction of a disulfide-free variable domain. J. Cell. Biochem. Suppl. 18, 213.

Fritz, G.N., Conn, J., Cockburn, A., Seawright, J., 1994b. Sequence analysis of the ribosomal DNA internal transcribed spacer 2 from populations of *Anopheles nuneztovari* (Diptera: Culicidae). Mol. Biol. Evol. 11, 406–416.

Gayet, M., Meunier, F.J., 2003. Paleontology and Palaeobiogeography of Catfishes. In: Arratia, G., Kapoor, B.G., Chardon, M., Diogo, R. (Eds.), Catfishes, vol. 2. Science Publishers, Inc., Enfield, USA, pp. 291–522, Chapter 17.

Gebri, S.A., 1985. Evolution of ribosomal DNA. In: MacIntyre, R.J. (Ed.), Molecular evolutionary genetics. Plenum, New York, pp. 419–517.

Goloboff, P., 1998. NONA (NO NAME) ver. 2. Published by the author, Tucumán, Argentina.

Goloboff, P., 1999. Analyzing large data sets in reasonable times: solutions for composite optima. Cladisitcs 15 (4), 415–429.

Graham, J.B., 1997. Air-breathing Fishes. Evolution, Diversity and Adaptation. Academic Press, San Diego.

Greenwood, P.H., 1961. A revision of the genus Dinotopterus Blgr. (Pisces, Clariidae) with notes on the comparative anatomy of the suprabranchial organs in the Clariidae. Bull. Brit. Mus. Nat. His. (Zool.) 7 (4), 217–241.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17 (8), 754–755.

Legendre, M., Teugels, G.G., Cauty, C., Jalabert, B., 1992. A comparative study on morphology, growth rate and reproduction of *Clarias gariepinus* (Burchell 1822), *Heterobranchus longifilis* Valenciennes, 1840, and their reciprocal hybrids. J. Fish Biol. 40, 59–79.

Lutzoni, F., Pagel, M., Reeb, V., 2001. Major fungal lineages are derived from lichen symbiotic ancestors. Nature 411 (6840), 937–940.

Matthes, H., 1964. Les poissons du lac Tumba et la région de d'Ikela. Ann. Mus. Afr. Centr. Zool., 1–199.

May, J.C., Coleman, A.W, 1997. The internal transcribed spacer 2 exhibits a common secondary structure in green algae and flowering plants. J. Mol. Evol. 44, 258–271.

Morgan, J.A.T., Blair, D., 1998. Trematode and monogenean rRNA ITS2 secondary structures support a four-domain model. J. Mol. Evol. 47 (4), 406–419.

Nelles, L., Fang, B.L., Volckaert, G., Vandenberghe, A., De Wachter, R., 1984. Nucleotide sequence of a crustacean 18S ribosomal RNA gene and secondary structure of eukaryotic small subunit ribosomal RNAs. Nucleic Acids Res. 12, 8749–8768.

Nicholas, K.B., Nicholas Jr., H.B., Deerfield, D.W., 1997. GeneDoc: analysis and visualization of genetic variation. EMBNEW NEWS 4, 14.

Nixon, K.C., 1999–2002. WinClada ver. 1.0000. Published by the author, Ithaca, NY, USA.

Page, R.D.M., 1996. TREEVIEW: an application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. 12, 357–358.

Pellegrin, J., 1927. La disparation des nageoires paires chez les poissons africains du groupe des clariinés. Ann. Sci. Nat. (Zool.) Paris 10, 209–222.

Poll, M., 1942. Note sur l'ostéologie de *Dolichallabes microphthalmus* Poll et remarques sur l'èvolution des Clariidae. Ann. Soc. R. (Zool.) Belge 73, 222–235.

Poll, M., 1977. Les genres nouveaux *Platyallabes* et *Platyclarias* comparés au genre *Gymnallabes* GTHR. Synopsis nouveau des genres de Clariidae. Bull. Acad. Roy. Belg. Classe Sci. 63, 122–149.

Porter, C.H., Collins, F.H., 1991. Species-diagnostic differences in a ribosomal DNA internal transcribed spacer from sibling species *Anopheles freeborni* and *Anopheles hermsi* (Diptera: Culicidae). Am. J. Trop. Med. Hyg. 45, 271–279.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Posada, D., Buckley, T.R., 2004. Model selection and model averaging in phylogenetics: advantages of the AIC and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53, 793–808.

Sabaj, M.H., Page, L.M., Lundberg, J.G., Ferraris, C.J., Jr., Armbruster, J.W., Friel, J.P., Morris, P.J., 2004. All Catfish Species Inventory Website. Internet address: http://clade.acnatsci.org/allcatfish.

Samraoui, B., Weekers, P.H.H., Dumont, H.J, 2003. Two taxa within the North African *Lestes virens* complex (Zygoptera: Lestidae). Odonatologica 32 (2), 131–142.

Sanderson, M.J., 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol. Biol. Evol. 19, 101–109.

Sanderson, M.J., 2003. r8s: inferring absolute rates of molecular evolution and divergence Times in absence of a molecular clock. Bioinformatics 19, 301–302.

Shrinivashagar, H.R., 1958. Development in the skull of catfishes: Part V: development of skull in *Heteropneustes fossilis* (Bloch) (Heteropneustidae). Proc. Nat. Inst. Sci. India 24B, 165–190.

Swofford, D.L., 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Swofford, D.L., Thorne, J.L., Felsenstein, J., Wiegmann, B.M., 1996. The topology-dependent permutation test for monophyly does not test for monophyly. Syst. Biol. 45, 575–579.

Tang, J., Toè, L., Back, C., Unnasch, T.R., 1996. Intra-specific heterogeneity of their DNA internal transcribed spacer in the *Simulium damnosum* (Diptera: Simuliidae) complex. Mol. Biol. Evol. 13, 244–252.

Teugels, G.G, 1986. A systematic revision of the African species of the genus Clarias (Pisces: Clariidae). Zool. Wet. Ann. Sci. Zool. KMMA 247.

Teugels, G.G, 2003. State of the art of recent Siluriform systematics. In: Arratia, G., Kapoor, B.G., Chardon, M., Diogo, R. (Eds.), Catfishes, vol. 1. Science Publishers, Inc., Enfield, USA, pp. 317–352, Chapter 11.

Teugels, G.G., Adriaens, D., 2003. Taxonomy and phylogeny of Clariidae-An overview. In: Arratia, G., Kapoor, B.G., Chardon, M. and Diogo, R. (Eds.), Catfishes, vol. 1. Science Publishers, Inc., Enfield, USA, pp. 465–487, Chapter 16.

Teugels, G.G., Ouzof-Costaz, C., Legendre, M., Parrent, M., 1992. A karyological analysis of the artificial hybridization between *Clarias gariepinus* (Burchell, 1822) and *Heterobranchus longifilis* Valenciennes, 1840 (Pisces; Clariidae). J. Fish Biol. 40, 81–86.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24, 4876–4882.

Thorne, J.L., Kishino, H., 2002. Divergence time and evolutionary rate estimation with multilocus data. Syst. Biol. 51, 689–702.

Torres, R.A., Ganal, M., Hemleben, V., 1990. GC balance in the internal transcribed spacers ITS1 and ITS2 of nuclear ribosomal DNA. J. Mol. Evol. 30, 170–181.

Weekers, P.H.H., Gast, R.J., Fuerst, P.A., Byers, T.J., 1994. Sequence variations in small-subunit ribosomal RNAs of *Hartmannella vermiformis* and their phylogenetic implications. Mol. Biol. Evol. 11, 684–690.

Weekers, P.H.H., Murugan, G., Vanfleteren, J.R., Belk, D., Dumont, H.J., 2002. Phylogenetic analysis of Anostraca (Branchiopoda, Anostraca) inferred from 18S rDNA sequences. Mol. Phylogenet. Evol. 25, 535–544.

Wheeler, W.C., Gladstein, D., DeLaet, J., 2002. POY. American Museum of Natural History, New York.

Yang, Z., Yoder, A.D, 2003. Comparison of likelihood and Bayesian methods for estimating divergence times using multiple gene loci and calibration points with application to a radiation of cute-looking mouse lemur species. Syst. Biol. 52, 705–716.

Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 31 (13), 3406–3415.