A robust phylogeny among major lineages of the East African cichlids

Tetsumi Takahashi [a,b,*], Teiji Sota [a]

[a] Department of Zoology, Graduate School of Science, Kyoto University, Sakyo, Kyoto 606-8502, Japan
[b] National Institute of Genetics, Yata, Mishima, Shizuoka 411-8540, Japan

A B S T R A C T

The huge monophyletic group of the East African cichlid radiations (EAR) consists of thousands of species belonging to 12–14 tribes; the number of tribes differs among studies. Many studies have inferred phylogenies of EAR tribes using various genetic markers. However, these phylogenies partly contradict one another and can have weak statistic support. In this study, we conducted maximum-likelihood (ML) phylogenetic analyses using restriction site-associated DNA (RAD) sequences and propose a new robust phylogenetic hypothesis among Lake Tanganyika cichlid fishes, which cover most EAR tribes. Data matrices can vary in size and contents depending on the strategies used to process RAD sequences. Therefore, we prepared 23 data matrices with various processing strategies. The ML phylogenies inferred from 15 large matrices (2.0×10^6 to 1.1×10^7 base pairs) resolved every tribe as a monophyletic group with 100% bootstrap support and shared the same topology regarding relationships among the tribes. Most nodes among the tribes were supported by 100% bootstrap values, and the bootstrap support for the other node varied among the 15 ML trees from 70% to 100%. These robust ML trees differ partly in topology from those in earlier studies, and these phylogenetic relationships have important implications for the tribal classification of EAR.

© 2016 Elsevier Inc. All rights reserved.
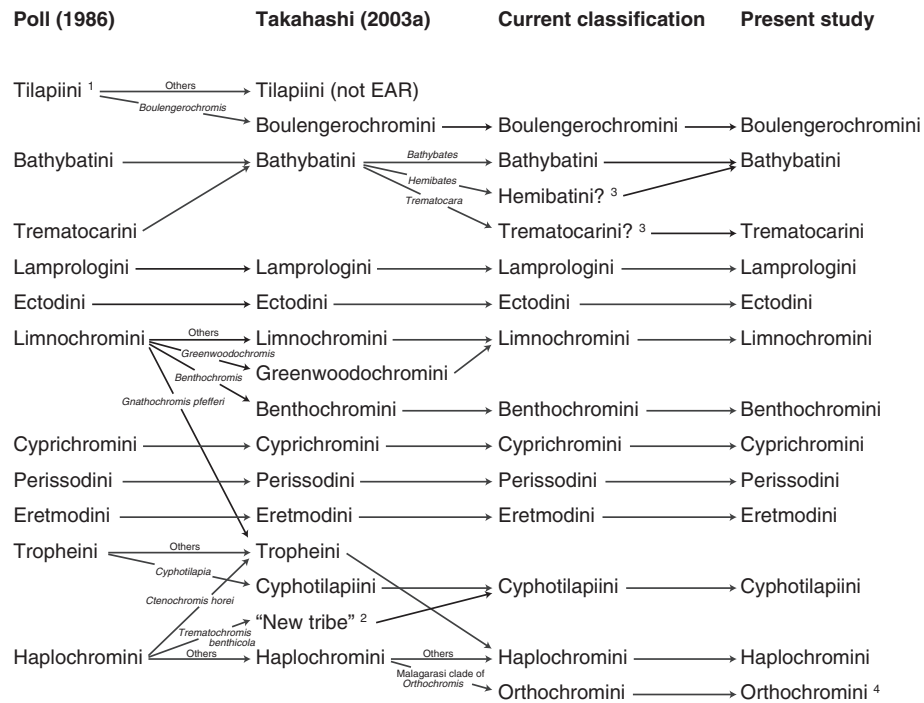
1. Introduction

The huge monophyletic group of the East African cichlid radiations (EAR, Schwarzer et al., 2009; Dunz and Schliewen, 2013) consists of thousands of species inhabiting lakes and rivers mainly in East Africa. Cichlid species of this group exhibit high morphological, ecological and genetic diversity, and large-scale adaptive radiations have taken place at least in Lake Tanganyika (Salzburger et al., 2002, but Genner et al., 2007; Schwarzer et al., 2009), Lake Malawi (Sturmbauer et al., 2001), the Lake Victoria basin (Verheyen et al., 2003) and the extinct Lake Makgadikgadi basin (Joyce et al., 2005). Lake Tanganyika, which is likely the oldest lake in Africa (9–12 million years, Cohen et al., 1993), is an evolutionary reservoir of old lineages of EAR (Nishida, 1991); indeed, many EAR lineages consist entirely of species endemic to this lake. Therefore, Lake Tanganyika cichlid fishes are important for resolving the phylogeny of the cichlid fishes of the EAR. Large young haplochromine species flocks of Lake Malawi and the Lake Victoria basin, which are well-known examples of rapid adaptive radiation, originated in the Lake Tanganyika radiation (Salzburger et al., 2005; Koblmüller et al., 2008a).

Poll (1986) first classified the Lake Tanganyika cichlid fishes into 12 tribes based on morphological features. Takahashi (2003a) revised Poll's (1986) classification and recognised 16 tribes based on cladistic analysis of anatomical data, of which 14 tribes composed the EAR. Subsequently, further revisions were made for the framework of the EAR tribes (Fig. 1). Currently, there is at least one point of debate regarding the tribal classification of EAR cichlid fishes. That is, three alternative tribal classifications exist for the genera *Bathybates* Boulenger, 1898, *Hemibates* Regan, 1920 and *Trematocara* Boulenger, 1899. Poll (1986) recognised the tribes Bathybatini Poll, 1986 (*Bathybates* and *Hemibates*) and Trematocarini Poll, 1986 (*Trematocara*; note that Poll, 1986 also included the genus *Telotrematocara* Poll, 1986 in this tribe, but this genus is currently treated as a synonym of *Trematocara*, see Takahashi, 2002); Takahashi (2003a) recognised only Bathybatini, consisting of the three genera; and Koblmüller et al. (2008b) suggested the presence of three monotypic tribes, Trematocarini, Bathybatini and Hemibatini Koblmüller et al., 2008b. The phylogenetic relationships among these three genera have not been clearly resolved (Koblmüller et al., 2005; Day et al., 2008; Weiss et al., 2015).

Phylogenetic relationships among the EAR cichlid tribes have repeatedly been inferred based on different molecular markers, although statistical support for these topologies is often weak (Table 1). These phylogenies partly agree in topology, but differ in several points. Many studies have located the tribes

* Corresponding author at: Institute of Natural and Environmental Sciences, University of Hyogo, Yayoigaoka, Sanda, Hyogo 669-1546, Japan.
   E-mail address: tetsumi@hitohaku.jp (T. Takahashi).

Fig. 1. Tribal classifications of cichlid fishes in the EAR. From left, two major tribal classifications of Poll (1986) and Takahashi (2003a), currently accepted classification referring to recent studies (Clabaut et al., 2005, 2007; Koblmüller et al., 2008a,b; Muschick et al., 2012; Takahashi, 2014; Meyer et al., 2015), and the classification used in this study. 1: Poll's (1986) Tilapiini is a paraphyletic group, and only *Boulengerochromis microlepis* (Boulenger, 1899) is included in the EAR. 2: The name of this monotypic tribe was not assigned because generic allocation of *Trematochromis benthicola* (Matthes, 1962) was not resolved at that time. 3: Early studies are in disagreement on the validity of these tribes (see text). 4: No species of the tribe Orthochromini Clabaut et al., 2005 are included in the present study.

Table 1
Robustness of phylogenies of cichlid tribes in the EAR. The minimum statistical support [bootstrap percentage for neighbour-joining (NJ), maximum-likelihood (ML) and most-parsimonious (MP) trees; posterior probability for Bayesian tree] on the nodes regarding relationships among the EAR tribes is shown for each analysis.

Study	Marker	Number of EAR tribes	Method	Minimum statistical support
Nishida (1991)	nDNA (allozyme)	10	NJ	Not estimated
			MP	Not estimated
			UPGMA	Not estimated
Sturmbauer and Meyer (1993)	mtDNA (*cytb*, control region)	7	NJ	Not estimated
			MP	<50%
Kocher et al. (1995)	mtDNA (ND2)	10	NJ	20%
Salzburger et al. (2002)	mtDNA (ND2, *cytb*)	12	ML	<50%
			MP	<50%
Clabaut et al. (2005)	mtDNA (ND2)	12	ML	<50%
			Bayes	<0.5
	nDNA (*rag1*)	12	ML	<50%
			Bayes	<0.5
	mtDNA (ND2) + nDNA (*rag1*)	12	ML	<50%
			Bayes	<0.5
Clabaut et al. (2007)	mtDNA (ND2)	11	ML	Not estimated
Day et al. (2008)	mtDNA (ND2, control region)	13	ML	<50%
			Bayes	<0.5
Muschick et al. (2012)	mtDNA (ND2) + nDNA (*ednrb1, phpt1*)	12	ML	Not estimated
			Bayes	<0.97
Meyer et al. (2015)	nDNA (44 loci)	11	ML	<50%
			Bayes	<0.75
Weiss et al. (2015)	mtDNA (ND2)	13	ML	<50%?
	nDNA (AFLP)	13	NJ	37%
			MP	<50%?
			Bayes	<0.5?
McGee et al. (2016)	nDNA (ultraconserved elements)	9	ML	52%[a]
			Bayes	1.00[a]
Present study	nDNA (RAD-seq)	12	ML	100%[b]

[a] Tree inferred from 95% complete dataset without partitioning (see McGee et al., 2016 for details).
[b] Tree produced from the matrix with W_{clust} = 95% and Min_{taxa} = 12.