WILEY ethology

Evolutionary transitions to cooperative societies in fishes revisited

Hirokazu Tanaka[1,2] | Joachim G. Frommen[1] (iD) | Stephan Koblmüller[3] (iD) |
Kristina M. Sefc[3] (iD) | Matthew McGee[4] | Masanori Kohda[2] | Satoshi Awata[2] |
Michio Hori[5] | Michael Taborsky[1] (iD)

[1]Institute of Ecology and
Evolution, University of Bern,
Hinterkappelen, Switzerland

[2]Department of Biology and
Geosciences, Osaka City University, Osaka,
Japan

[3]Institute of Biology, University of Graz,
Graz, Austria

[4]Monash University, Melbourne, Victoria,
Australia

[5]Department of Zoology, Kyoto University,
Kyoto, Japan

Correspondence
Joachim G. Frommen, Institute of Ecology
and Evolution, University of Bern,
Hinterkappelen, Switzerland.
Email: joachim.frommen@iee.unibe.ch

Abstract

Unravelling the evolution of complex social organization in animals is an important aim, not least because it helps to understand the evolutionary roots of human sociality. Recent advances in comparative methods allow to approach this question in a phylogenetic context. The validity of such comparative approaches depends strongly on the quality of information regarding the behaviour, sociality, and reproduction of animals in natural systems, and on the quality of the phylogenetic reconstruction. Applying a novel comparative approach, a recent study of Dey et al. (2017, Nature Ecology & Evolution, 1, 137) concluded that evolutionary transitions to cooperative breeding in cichlid fishes were not associated with the social mating pattern. Here we argue that this result was adversely affected by equivocal classifications of mating patterns, and inadequate phylogenetic data. In order to illustrate the impact of the mating system misclassifications, we scored mating patterns as reported in the original literature and re-analysed the dataset based on Dey et al.'s tree topology. The result suggests that the mating system does in fact significantly explain the evolutionary transition to cooperative breeding in lamprologine cichlids, but we submit that a reliable conclusion cannot be reached before improving the behavioural information and the underlying phylogenetic reconstruction. The problems identified in this case study are not unique and we urge caution in the interpretation of results from comparative phylogenetic studies in general. We do agree with Dey et al. (2017) though that the lamprologine cichlids of Lake Tanganyika may constitute a fundamental test case for the theory of social evolution, but better information on their behaviour and phylogenetic relationships is needed to allow meaningful analyses.

KEYWORDS
comparative analyses, cooperative breeding, direct fitness benefits, lamprologine cichlids, phylogeny, social organization

1 | INTRODUCTION

Unveiling the evolution of the astounding diversity of social organization in animals is a worthwhile endeavor. Among other merits, it helps to understand the evolutionary roots of human sociality, arguably a major source of our ecological success. Recently, elaborate comparative analyses of data from several major taxa have revealed intriguing insight into the probable pathways of social evolution and

the importance of ecological factors and biological attributes. The success of such comparative approaches depends strongly on the quality of information regarding the behavior, sociality, and reproduction of animals in natural systems, and on the quality of the phylogenetic reconstruction. It is important to be aware of the potential problems and pitfalls involved in the comparative approach when interpreting results.

In a recent paper on the evolution of cooperative breeding in fishes, Dey et al. (2017) concluded that direct fitness benefits are the primary driver for evolutionary transition to complex sociality in lamprologine cichlid fishes, exceeding the outstanding importance of social monogamy and indirect fitness benefits as revealed by comparative analyses in insects, birds, and mammals (Cornwallis, West, Davis, & Griffin, 2010; Hughes, Oldroyd, Beekman, & Ratnieks, 2008; Lukas & Clutton-Brock, 2012). Revealing an alternative evolutionary route to complex social organization constitutes a major advance in our understanding of the evolution of sociality, and cooperation in general. When scrutinizing the comparative analysis of Dey et al. (2017), however, several issues emerge that may challenge the conclusions drawn from the analyses. This includes (a) the classification of traits, (b) the proxy used for within-group relatedness, and (c) the quality of the contained phylogenetic information and the applied comparative analysis procedure. Therefore, we urge caution for the interpretation of the reported results. In the following, we will outline the major issues.

1.1 | Classification of traits

The categorization of several parameters used by Dey et al. (2017) is precarious, partly due to erroneous confounding of data collected from the literature, and partly due to problematic binary classification. For instance, a third of the mating system classifications by Dey et al. (2017) are opposite to categories given by the original literature they cite (Table 1). As an example, data used from a published table listing the mating system of 28 lamprologine species (Heg & Bachar, 2006) were used opposite to their original meaning. A column with the heading "M" was incorrectly treated as signifying "monogamy," when in fact this category represented the reverse, namely "multi-male or multi-female groups: lives in polygynous, polyandrous and/or polygynynandrous groups" Heg and Bachar (2006), p. 278, table legend). We reanalyzed the Dey et al. (2017) dataset based on the original topology of Dey et al, 2017, utilizing the mating pattern classification as graded in the original literature sources (Table 1). This new analysis reveals that the mating system does in fact significantly explain the evolutionary transition to cooperative breeding (Figure 1), which is in sharp contrast to the inference of Dey et al. (2017).

Another problem with the classification of mating systems by Dey et al. (2017) is their categorization into the strictly binary categories "monogamous" and "polygamous" (the latter was referred to as both "promiscuous" and "non-monogamous" in the main manuscript and the supporting material). In reality, the degree of polygamy ranges from rare to common between and even within different populations of the same species (e.g., Matsumoto

& Kohda, 1998), which casts doubts on assigning a species to one of these discrete mating categories. Naturally, simplifications are often required for large-scale comparative analyses. However, disregarding available quantitative information for the included species makes the applied binary classification inevitably arbitrary.

1.2 | Proxies of within-group relatedness

The main hypothesis tested by Dey et al. (2017) was that high within-group relatedness favors the evolution of cooperative breeding. Here, a fundamental problem is that social monogamy was taken as a measure of high within-group relatedness. Like in many birds (e.g., Cornwallis et al., 2010), social mating system in cichlids is a poor proxy for within-group relatedness due to several reasons. First, in most polygamous lamprologine cichlids, breeders of one sex mate with only one opposite-sex partner (e.g., Yanagisawa, 1987), rendering within-group kinship patterns similar to monogamous species. For example, a polygynous male will mate with several females in his harem. However, each of these females' only mates with the respective harem owner, leading to high levels of relatedness within each female group. Considering that the cooperative helping behaviors take place within the breeder female's group, the unit that needs to be examined for testing the monogamy hypothesis should be the female group. Here, the within-group relatedness among offspring, as well as between offspring and breeders, will be similar between polygynous and monogamous species.

Second, as in many other animals, within-group relatedness in lamprologine cichlids is significantly affected by reproductive parasitism (Sefc, 2011; Taborsky, 1994), which can be considerable in socially monogamous species (e.g., Sefc, Mattersdorfer, Sturmbauer, & Koblmüller, 2008) and has been well demonstrated in cooperatively breeding lamprologine cichlids (Awata, Munehara, & Kohda, 2005; Hellmann et al., 2015; Taborsky, 2016). Furthermore, adoption of foreign offspring occurs regularly in a number of lamprologine species, several of which are monogamous (e.g., Nagoshi, 1987; Schaedelin, Dongen, & Wagner, 2012). All this illustrates that estimating within-group relatedness from social mating system classifications is dubious. Instead, this would require the use of genetic/parentage data (cf. Petrie & Kempenaers, 1998).

1.3 | Quality of phylogenetic information

Lake Tanganyika cichlids have been subject to numerous phylogenetic studies. Most of these studies were based on mitochondrial phylogenies, which in many lineages do not approximate the true species tree due to ancient incomplete lineage sorting (Takahashi, Terai, Nishida, & Okada, 2001) and hybridization/introgression, particularly in lamprologine cichlids (Huang, Tran, & Knowles, 2014; Koblmüller et al., 2007; McGee et al., 2016; Sturmbauer, Salzburger, Duftner, Schelly, & Koblmüller, 2010; Takahashi & Koblmüller, 2011). Hence, for trustable reconstruction of species trees in these cichlids, nuclear multilocus or genome-wide phylogenies are required. Dey et al. (2017) incorporated five nuclear genes into a large dataset of three commonly used mitochondrial

TABLE 1 Re-classified social mating system of lamprologine cichlids. A total of 23 out of 69 species were classified differently by Dey et al. (2017) as compared to the original literature (of these, 13 species were cooperative breeders)

Species	Mating system from Dey et al. (2017)	Re-classified mating system	Source(s) for reclassified mating system data
Altolamprologus calvus	Monogamous	n.a.	No description of mating system in Sefc (2011)
Chalinochromis brichardi	Monogamous	Non-monogamous	Heg and Bachar (2006) reported "polygynous, polyandrous and/or polygynandrous groups"
Chalinochromis popelini	Monogamous	Non-monogamous	Heg and Bachar (2006) reported "polygynous, polyandrous and/or polygynandrous groups"
Julidochromis dickfeldi	Monogamous	n.a.	No description of mating system in Sefc (2011)
Julidochromis marlieri	Monogamous	Non-monogamous	Heg and Bachar (2006) reported "polygynous, polyandrous and/or polygynandrous groups"
Julidochromis regani	Monogamous	Non-monogamous[a]	Heg and Bachar (2006) reported "polygynous, polyandrous and/or polygynandrous groups"
Lamprologus teugelsi	Non-monogamous	n.a.	No reliable description about mating system found
Lepidiolamprologus attenuatus	Monogamous	Non-monogamous[a]	Nagoshi & Gashagaza (1988) reported frequent polygamy
Lepidiolamprologus kendalli	Monogamous	Monogamous[a]	Koblmüller et al (2007) reported "monogamy?"
Lepidiolamprologus nkambae	Monogamous	Monogamous[a]	Koblmüller et al (2007) reported "monogamy?"
Neolamprologus boulengeri	Non-monogamous	Monogamous[a]	Koblmüller et al (2007) reported "monogamy?"
Neolamprologus cylindricus	Monogamous	Monogamous[a]	Kuwamura (1997) reported "monogamy?", no description of mating system in Goodwin et al (1998)
Neolamprologus falcicula	Monogamous	Non-monogamous	Heg and Bachar (2006) reported "polygynous, polyandrous and/or polygynandrous groups"
Neolamprologus gracilis	Monogamous	Non-monogamous	Heg and Bachar (2006) reported "polygynous, polyandrous and/or polygynandrous groups"
Neolamprologus hecqui	Non-monogamous	n.a.	No description of mating system in Sefc (2011); Koblmüller et al (2007) reported "?"
Neolamprologus helianthus	Monogamous	Non-monogamous	Heg and Bachar (2006) reported "polygynous, polyandrous and/or polygynandrous groups"
Neolamprologus marunguensis	Monogamous	Non-monogamous	Heg and Bachar (2006) reported "polygynous, polyandrous and/or polygynandrous groups"
Neolamprologus moorii	Non-monogamous	Monogamous	Karino (1998) reported monogamous
Neolamprologus niger	Monogamous	Non-monogamous	Heg and Bachar (2006) reported "polygynous, polyandrous and/or polygynandrous groups"
Neolamprologus olivaceous	Monogamous	Non-monogamous	Heg and Bachar (2006) reported "polygynous, polyandrous and/or polygynandrous groups"
Neolamprologus prochilus	Monogamous	n.a.	No reliable description about mating system found
Neolamprologus splendens	Monogamous	Non-monogamous	Heg and Bachar (2006) reported "polygynous, polyandrous and/or polygynandrous groups"
Neolamprologus walteri	Monogamous	n.a.	No description of mating system in Verburg and Bills (2007)

[a]We conducted analyses where these species were both considered monogamous and non-monogamous, and the classification of these species did not influence the result.

genes to account for this problem. Unfortunately, these five nuclear genes are available for only a small number of species, and, in contrast to the mitochondrial data, contain only little phylogenetic information (e.g., Koblmüller et al., 2017). Therefore, Dey et al.'s (2017) tree topology mainly reflects mitochondrial relationships, contradicting recent phylogenomic evidence (McGee et al., 2016; Takahashi & Sota, 2016; Irisarri et al., 2018) and concordant morphological classifications. This causes an erroneous increase in the number of transitions to cooperative breeding by up to

50%. Enforcing some topological constraints, as done by Dey et al. (2017), adds to the problem, as this produces a tree that is not backed up by the input data. It is important to realize that these enforced constraints do not just determine the phylogenetic placement of the constrained taxa, but also affect the branching order of many other nodes, especially due to the large amounts of missing data for the majority of loci, and a high prevalence of introgression/hybridization and ancient incomplete lineage sorting. To scrutinize the effect of potential flaws in the phylogenetic
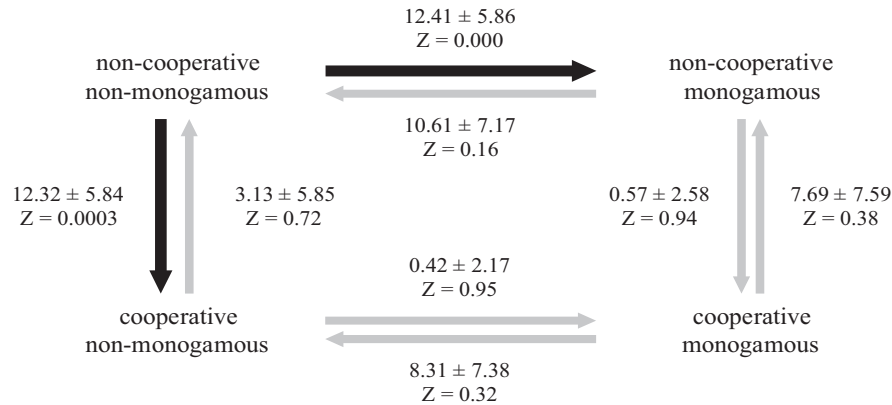
FIGURE 1 The direction of probable evolutionary transitions between cooperative breeding and mating system based on the reclassified datasets. Mean ± standard deviation of the transition parameter (q_{xy}) and the proportion of transitions classified as zero (Z) were reported in each transition parameter. Black arrows depict Z < 0.05, while Z > 0.05 is shown in gray

backbone of Dey et al. (2017), we have reanalyzed their sequence data. This reanalysis revealed major shortcomings in the phylogenetic analysis of Dey et al. (2017), which are critical for the results and their interpretation. Inevitably, this rendered erroneous conclusions regarding the evolutionary transitions to cooperative breeding in this lineage of cichlid fishes. We provide a detailed account on the problems with the phylogenetic analyses of Dey et al. (2017) in the Appendix below. Notably, the path analysis based on phylogenetic regressions as used by Dey et al. (2017) is a largely unverified approach. The only test of the error rate and power of this method with a simulated dataset has revealed that type I error rates exceeded 5% (von Hardenberg & Gonzalez-Voyer, 2013).

We should like to stress that we do not challenge Dey et al.'s (2017) inference that direct fitness benefits are most likely of major importance for the evolution of cooperative breeding in cichlids, as has been suggested by previous work (Quiñones, Doorn, Pen, Weissing, & Taborsky, 2016; Zöttl, Heg, Chervet, & Taborsky, 2013). However, we think that this hypothesis needs further specific and comparative study to unveil the factors responsible for the evolutionary pathways of cichlid sociality. Importantly, more data need to be considered for categorizing the reproductive, social, and ecological parameters than used by Dey et al. (2017), and to achieve a more trustworthy phylogenetic backbone. Nonetheless, we agree with Dey et al.'s (2017) conjecture that the lamprologine cichlids of Lake Tanganyika may constitute a fundamental test case for the theory of social evolution.

ACKNOWLEDGEMENTS

ORCID

Joachim G. Frommen (iD) http://orcid.org/0000-0002-1752-6944

Stephan Koblmüller (iD) https://orcid.org/0000-0002-1024-3220

Kristina M. Sefc (iD) https://orcid.org/0000-0001-8108-8339

Michael Taborsky (iD) https://orcid.org/0000-0002-1357-4316

REFERENCES

Awata, S., Munehara, H., & Kohda, M. (2005). Social system and reproduction of helpers in a cooperatively breeding cichlid fish (*Julidochromis ornatus*) in Lake Tanganyika: Field observations and parentage analyses. *Behavioral Ecology and Sociobiology*, *58*, 506–516.

Cornwallis, C. K., West, S. A., Davis, K. E., & Griffin, A. S. (2010). Promiscuity and the evolutionary transition to complex societies. *Nature*, *466*, 969–972.

Day, J. J., Santini, S., & Garcia-Moreno, J. (2007). Phylogenetic relationships of the Lake Tanganyika cichlid tribe Lamprologini: The story from mitochondrial DNA. *Molecular Phylogenetics and Evolution*, *45*, 629–642.

Dey, C. J., O'Connor, C. M., Wilkinson, H., Shultz, S., Balshine, S., & Fitzpatrick, J. L. (2017). Direct benefits and evolutionary transitions to complex societies. *Nature Ecology & Evolution*, *1*, 137.

Goodwin, N. B., Balshine-Earn, S., Reynolds, J. D. (1998) Evolutionary transitions in parental care in cichlid fish. Proceedings of the Royal Society B - Biological Sciences, *265*, 2265-2272

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology*, *59*, 307–321.

Heg, D., & Bachar, Z. (2006). Cooperative breeding in the lake tanganyika cichlid *Julidochromis ornatus*. *Environmental Biology of Fishes*, *76*, 265–281.

Hellmann, J. K., Ligocki, I. Y., O'Connor, C. M., Reddon, A. R., Garvy, K. A., Marsh-Rollo, S. E., … Hamilton, I. M. (2015). Reproductive sharing in relation to group and colony-level attributes in a cooperative breeding fish. *Proceedings of the Royal Society B: Biological Sciences*, *282*, 20150954.

Huang, H., Tran, L. A., & Knowles, L. L. (2014). Do estimated and actual species phylogenies match? Evaluation of East African cichlid radiations. *Molecular Phylogenetics and Evolution*, *78*, 56–65.

Hughes, W. O. H., Oldroyd, B. P., Beekman, M., & Ratnieks, F. L. W. (2008). Ancestral monogamy shows kin selection is key to the evolution of eusociality. *Science*, *320*, 1213–1216.

Irisarri, I., Singh, P., Koblmüller, S., Torres-Dowdall, J., Henning, F., Franchini, P., ... Meyer, A. (2018). Phylogenomics uncovers early hybridization and adaptive loci shaping the radiation of Lake Tanganyika cichlid fishes. *Nature Communications*, *9*, 3159.

Karino, K. (1998) Depth-related differences in territory size and defense in the herbivorous cichlid, neolamprologus moorii, in Lake Tanganyika. *Ichthyological Research*, *45*, 89 –94.

Koblmüller, S., Duftner, N., Sefc, K. M., Aibara, M., Stipacek, M., Blanc, M., ... Sturmbauer, C. (2007). Reticulate phylogeny of gastropod-shell-breeding cichlids from Lake Tanganyika–the result of repeated introgressive hybridization. *BMC Evolutionary Biology*, *7*, 7.

Koblmüller, S., Nevado, B., Makasa, L., Van Steenberge, M., Vanhove, M. P., Verheyen, E., ... Sefc, K. M. (2017). Phylogeny and phylogeography of *Altolamprologus*: Ancient introgression and recent divergence in a rock-dwelling Lake Tanganyika cichlid genus. *Hydrobiologia*, *791*, 35–50. https://doi.org/10.1007/s10750-016-2896-2

Kuhner, M. K. (2009). Coalescent genealogy samplers: Windows into population history. *Trends in Ecology & Evolution*, *24*, 86–93. https://doi.org/10.1016/j.tree.2008.09.007

Kuwamura, T. (1997) The evolution of parental care and mating systems among Tanganyikan cichlids. In H. Kawanabe, M. Hori & M. Nagoshi (Eds.) *Fish communities in Lake Tanganyika*. (pp. 59-86.) Kyoto University Press.

Lukas, D., & Clutton-Brock, T. (2012). Cooperative breeding and monogamy in mammalian societies. *Proceedings of the Royal Society B: Biological Sciences*, *279*, 2151–2156. https://doi.org/10.1098/rspb.2011.2468

Matsumoto, K., & Kohda, M. (1998). Inter-population variation in the mating system of a substrate-breeding cichlid in Lake Tanganyika. *Journal of Ethology*, *16*, 123–127. https://doi.org/10.1007/BF02769291

McGee, M. D., Faircloth, B. C., Borstein, S. R., Zheng, J., Hulsey, C. D., Wainwright, P. C., & Alfaro, M. E. (2016). Replicated divergence in cichlid radiations mirrors a major vertebrate innovation. *Proceedings of the Royal Society B: Biological Sciences*, *283*, 20151413. https://doi.org/10.1098/rspb.2015.1413

Nagoshi, M. (1987). Survival of broods under parental care and parental roles of the cichlid fish, *Lamprologus toae*, in Lake Tanganyika. *Japanese Journal of Ichthyology*, *34*, 71–75.

Nagoshi, M., Gashagaza, M. M. (1988) Growth of the larvae of a Tanganyikan cichlid, *Lamprologus attenuatus*, under parental care. *Japanese Journal of Ichthyology*. *35*, 392 -395.

Nevado, B., Koblmüller, S., Sturmbauer, C., Snoeks, J., Usano-Alemany, J., & Verheyen, E. (2009). Complete mitochondrial DNA replacement in a Lake Tanganyika cichlid fish. *Molecular Ecology*, *18*, 4240–4255. https://doi.org/10.1111/j.1365-294X.2009.04348.x

Petrie, M., & Kempenaers, B. (1998). Extra-pair paternity in birds: Explaining variation between species and populations. *Trends in Ecology & Evolution*, *13*, 52–58. https://doi.org/10.1016/S0169-5347(97)01232-9

Quiñones, A. E., van Doorn, G. S., Pen, I., Weissing, F. J., & Taborsky, M. (2016). Negotiation and appeasement can be more effective drivers of sociality than kin selection. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*. https://doi.org/10.1098/rstb.2015.0089

Ronquist, F., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, *19*, 1572–1574.

Salzburger, W., Baric, S., & Sturmbauer, C. (2002). Speciation via introgressive hybridization in East African cichlids? *Molecular Ecology*, *11*, 619–625.

Schaedelin, F. C., van Dongen, W. F., & Wagner, R. H. (2012). Nonrandom brood mixing suggests adoption in a colonial cichlid. *Behavioral Ecology*, *24*, 540–546.

Schelly, R., Salzburger, W., Koblmüller, S., Duftner, N., & Sturmbauer, C. (2006). Phylogenetic relationships of the lamprologine cichlid genus Lepidiolamprologus (Teleostei: Perciformes) based on mitochondrial and nuclear sequences, suggesting introgressive hybridization. *Molecular Phylogenetics and Evolution*, *38*, 426–438.

Sefc, K. M. (2011). Mating and parental care in Lake Tanganyika's cichlids. *International Journal of Evolutionary Biology*, *2011*, 1–20. https://doi.org/10.4061/2011/470875

Sefc, K. M., Mattersdorfer, K., Sturmbauer, C., & Koblmüller, S. (2008). High frequency of multiple paternity in broods of a socially monogamous cichlid fish with biparental nest defence. *Molecular Ecology*, *17*, 2531–2543.

Stiassny, M. L. (1997). A phylogenetic overview of the lamprologine cichlids of Africa (Teleostei, Cichlidae): A morphological perspective. *South African Journal of Science*, *93*, 513–522.

Sturmbauer, C., Salzburger, W., Duftner, N., Schelly, R., & Koblmüller, S. (2010). Evolutionary history of the Lake Tanganyika cichlid tribe Lamprologini (Teleostei: Perciformes) derived from mitochondrial and nuclear DNA data. *Molecular Phylogenetics and Evolution*, *57*, 266–284.

Taborsky, M. (1994). Sneakers, satellites, and helpers - parasitic and cooperative behavior in fish reproduction. *Advances in the Study of Behavior*, *23*, 1–100.

Taborsky, M. (2016). Cichlid fishes: A model for the integrative study of social behavior. In W. D. Koenig, & J. L. Dickinson (Eds.), *Cooperative breeding in vertebrates: Studies of ecology, evolution, and behavior* (pp. 272–293). Cambridge, UK: Cambridge University Press.

Takahashi, T., & Koblmüller, S. (2011). The adaptive radiation of cichlid fish in Lake Tanganyika: A morphological perspective. *International Journal of Evolutionary Biology*, *2011*, 620754. https://doi.org/10.4061/2011/620754

Takahashi, T., & Sota, T. (2016). A robust phylogeny among major lineages of the East African cichlids. *Molecular Phylogenetics and Evolution*, *100*, 234–242.

Takahashi, K., Terai, Y., Nishida, M., & Okada, N. (2001). Phylogenetic relationships and ancient incomplete lineage sorting among cichlid fishes in Lake Tanganyika as revealed by analysis of the insertion of retroposons. *Molecular Biology and Evolution*, *18*, 2057–2066.

Verburg, P., & Bills, R. (2007). Two new sympatric species of *Neolamprologus* (Teleostei: Cichlidae) from Lake Tanganyika, East Africa. *Zootaxa*, *1612*, 25–44.

von Hardenberg, A., & Gonzalez-Voyer, A. (2013). Disentangling evolutionary cause-effect relationships with phylogenetic confirmatory path analysis. *Evolution*, *67*, 378–387.

Yanagisawa, Y. (1987). Social organization of a polygynous cichlid *Lamprologus furcifer* in Lake Tanganyika. *Japanese Journal of Ichthyology*, *34*, 82–90.

Zöttl, M., Heg, D., Chervet, N., & Taborsky, M. (2013). Kinship reduces alloparental care in cooperative cichlids where helpers pay-to-stay. *Nature Communications*, *4*, 1341.

APPENDIX

Ancestral state reconstruction (ASR) requires information on species relationships and therefore depends on the phylogenetic approximation of the true species tree. The phylogeny used by Dey et al. (2017) suffers from several problems, including low phylogenetic

information content and a large proportion of missing data in the nuclear dataset. Consequently, the tree is heavily influenced by the mitochondrial data, which handicaps the approximation of the lamprologini species tree due to incomplete lineage sorting and prevalent hybridization/introgression (Huang et al., 2014; Koblmüller et al., 2007; McGee et al., 2016; Sturmbauer et al., 2010; Takahashi & Koblmüller, 2011; Takahashi & Sota, 2016). Measures taken to resolve these issues (topological constraints and omission of particular sequence data) unfortunately created further problems. As a result, the topology of the ASR tree and, importantly, the positions of several cooperative breeders in the tree contradict current knowledge based on genomic data. Here, we highlight our major concerns regarding the phylogenetic tree used in Dey et al. (2017).

THE PHYLOGENY MAINLY RELIES ON MITOCHONDRIAL DATA AND THUS DOES NOT APPROXIMATE THE TRUE SPECIES TREE

The phylogenetic reconstruction of Dey et al. (2017) stands in contrast to species relationships reconstructed from more comprehensive nuclear data. For instance, several nuclear multilocus datasets (AFLPs in Koblmüller et al. (2007), Sturmbauer et al. (2010); RADseq in Takahashi and Sota (2016); ultraconserved elements (UCE, >1,000 loci; >560 kb) in McGee et al. (2016) anchored phylogenomics (>500 loci; >950 kb) in Irisarri et al. (2018)) consistently placed the ossified group, a morphologically well-defined monophyletic group within the lamprologines (Stiassny, 1997), as the sister group of all other lamprologines, which is not reflected in Dey et al. Furthermore, well-recognized monophyletic groups (e.g., the *Lamprologus ocellatus*, *L. meleagris*, and *L. speciosus* species group; the group consisting of *Lepidiolamprologus cunningtoni*, *Neolamprologus christyi*, *N. modestus*, *N. mondabu*, *N. petricola*, and *N. tetracanthus*; the genus *Telmatochromis*; the *Neolamprologus pulcher* complex; Koblmüller et al., 2007; Sturmbauer et al., 2010; McGee et al., 2016; Irisarri et al., 2018) were not recovered. Dey et al. applied topological constraints to amend some of the phylogenetic peculiarities shown by their tree, which we find problematic for reasons explained further below.

Despite the inclusion of nuclear sequence data in the phylogenetic analysis, the phylogenetic reconstruction of Dey et al. (2017) is still overwhelmingly dominated by mtDNA polymorphisms, such that the resulting tree largely mirrors the mitochondrial tree topology (Day, Santini, & Garcia-Moreno, 2007; Sturmbauer et al., 2010). The deviations between single-locus trees (such as the mitochondrial tree) and multilocus trees can be attributed to (ancient) incomplete lineage sorting and (ancient) hybridization/introgression on individual loci, both particularly evident in studies of lamprologine cichlids (Day et al., 2007; Koblmüller et al., 2007, 2017 ; Nevado et al., 2009; Salzburger, Baric, & Sturmbauer, 2002; Schelly, Salzburger, Koblmüller, Duftner, & Sturmbauer, 2006; Sturmbauer et al., 2010). While Dey et al. attempted to improve their species tree approximation by the inclusion of nuclear sequence data, this commendable approach was compromised by two caveats. First, given the limited availability of data from the selected nuclear loci, the nuclear dataset

used in their study is very incomplete. Although a certain level of missing data does not generally compromise phylogenetic analyses, the amount of missing data in the present dataset is exorbitant (72% missing data). The nuclear dataset, if complete, would comprise 345 sequences (69 species times 5 loci), but 251 of these are missing (see Suppl. table 3 in Dey et al. (2017)). Moreover, the 28% available data do not cover the major lamprologine clades homogeneously, but are concentrated in the ossified group, and data (i.e., loci) overlap poorly across species. With so little nuclear data actually included in the dataset, the resulting phylogenetic tree necessarily reflects mainly the mitochondrial relationships. Second, given the low substitution rate of nuclear sequences, the five nuclear loci provide limited phylogenetic information and, for instance, fail to resolve the branching order of many lamprologine species (low node support in Figure 2a; also shown by Nevado et al. (2009); Koblmüller et al. (2017)). Consequently, even if an effort would have been made to complete the nuclear dataset, the phylogenetic tree would inevitably have been dominated by the information contained in the mtDNA data.

In order to account for hybridization/introgression demonstrated in three lamprologines (*Neolamprologus fasciatus*, *N. multifasciatus*, *N. similis*; Koblmüller et al., 2007; Nevado et al., 2009), Dey et al. discarded the mtDNA sequences of these species from the dataset. This approach is useful in principle, but ineffective with regard to several other introgressed species in the dataset and problematic given the incompleteness of the nuclear data.

To demonstrate the problems in the phylogeny applied by Dey et al. (2017), we reanalyzed the sequence data. To reduce computation time and because divergence time estimates were not required, we used MrBayes 3.2 (Ronquist & Huelsenbeck, 2003) instead of BEAST (which was used by Dey et al. (2017)) to infer phylogenetic relationships based on mtDNA data only, the nucDNA used in Dey et al. (2017), and combined mtDNA and nucDNA data. Note that MrBayes and BEAST treat missing data in the same way, so there should not be any program-specific effects on supported nodes. We employed the same partitioning scheme and substitution models as reported in Dey et al. (2017). For each dataset, two simultaneous Markov Chain Monte Carlo (MCMC) searches were conducted (8 chains each, sample frequency of 1,000) until split deviation frequencies were <0.01 and effective samples sizes (ESS) for all parameters exceeded 200, indicating that the parameter log file accurately reflected the posterior distribution (Kuhner, 2009). Since no outgroup was used (as in Dey et al. (2017)), resulting trees were midpoint-rooted.

Figures A1, A2, and A3 show the phylogenetic relationships inferred from mtDNA, nucDNA, and mtDNA and nucDNA combined, respectively. The tree topology of the nucDNA is clearly different from the mtDNA tree, but the tree based on the combined data largely mirrors the mitochondrial tree (even though statistical support for nodes, for which nuclear data were available, was lower in the combined tree). This demonstrates that the information in the nuclear data is swamped by the mtDNA data and therefore contributes very little to the tree topology of the combined data. These results are in accordance with previous studies showing that these particular nuclear loci are not able to resolve the lamprologine tree (Koblmüller et al., 2017; Nevado et al.,

2009). In the tree based on these nuclear markers, some of the phylogenetic relationships among non-ossified lamprologines are not recovered according to current knowledge. For instance, *Neolamprologus brichardi*, *N. pulcher*, and *N. savoryi*, which appear closely related based on nuclear multilocus and genomic data (e.g., Sturmbauer et al., 2010; McGee et al., 2016; Irisarri et al., 2018), as well as by morphology and behavior, do not cluster together in Dey et al. (2017). This discrepancy is most probably an artifact resulting from sampling different loci in the different non-ossified group lamprologines (see Suppl. table 3 in Dey et al. (2017)).

## TOPOLOGICAL CONSTRAINTS ARE NOT BACKED UP BY THE USED DATA AND IMPACT MANY OTHER NODES IN THE TREE

Phylogenetic tree inference typically aims at an objective reconstruction of phylogenetic relationships based on available data, and topological constraints should be employed only for the testing of alternative tree topologies. In Dey et al., however, topological constraints were applied to two clades in order to enforce consistency with published phylogenies (Irisarri et al., 2018; McGee et al., 2016; Sturmbauer et al., 2010). This not only resulted in a tree that is not backed up by the input data, but also affected branching order in other parts of the tree (e.g., compare the constrained tree of Dey et al. (2017) to our unconstrained tree in Figure A3). The latter is particularly problematic in the current context, as it affects the reconstructed transitions between cooperative and non-cooperative breeding. Specifically, Dey et al. enforced monophyly of *Neolamprologus* species with brood care helpers (constraining monophyly of *N. brichardi*, *N. gracilis*, *N. helianthus*, *N. marunguensis*, *N. olivaceous*, *N. pulcher*, *N. savoryi*, *N. splendens*), which cluster together in published multilocus phylogenies (Irisarri et al., 2018; McGee et al., 2016; Sturmbauer et al., 2010). However, in the previous studies, this monophyletic group also includes the cooperative breeders *N. falcicula* and *N. walteri*. Leaving the positions of these two species unconstrained, the tree used by Dey et al. places these two species in a different clade and suggests a loss of cooperative breeding (node 3 in Dey et al., 2017).

SOME SPECIES ARE REPRESENTED BY MITOCHONDRIAL GENE SEQUENCES ORIGINATING FROM HIGHLY DIVERGENT HAPLOGROUPS

Previous studies demonstrated the presence of divergent mitochondrial haplotype lineages within several Lamprologini species (e.g., Salzburger et al., 2002; Nevado et al., 2009; Sturmbauer et al., 2010). This phenomenon was attributed to the frequent occurrence of hybridization and introgression among lamprologines. In consequence, the mitochondrial genomes of two conspecific individuals may have experienced profoundly different evolutionary histories.

In the dataset of Dey et al., the mtDNA data are comprised by three loci, and were, for some of the species, compiled from different individuals. This bears the risk of combining sequences from different haplogroups into a "species haplotype" (note that mitochondrial genomes lack recombination and genes are therefore linked), which can severely compromise the phylogenetic analysis. Indeed, differences between trees calculated from each of the mitochondrial genes separately (Figures A4–A6; inconsistent placement of, e.g., *Neolamprologus marunguensis*, *Julidochromis ornatus*, *Lamprologus lemairii*) suggest that some of the compiled species haplotypes are mosaics of divergent haplogroups.
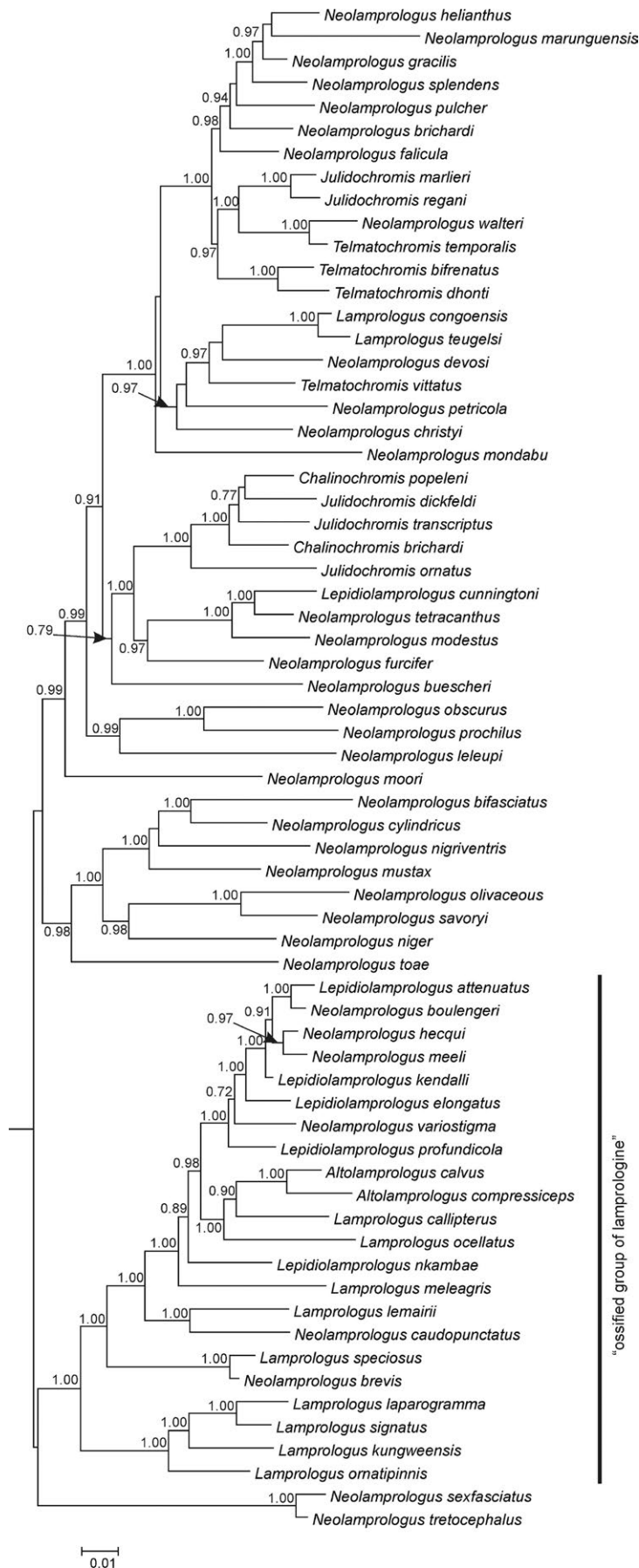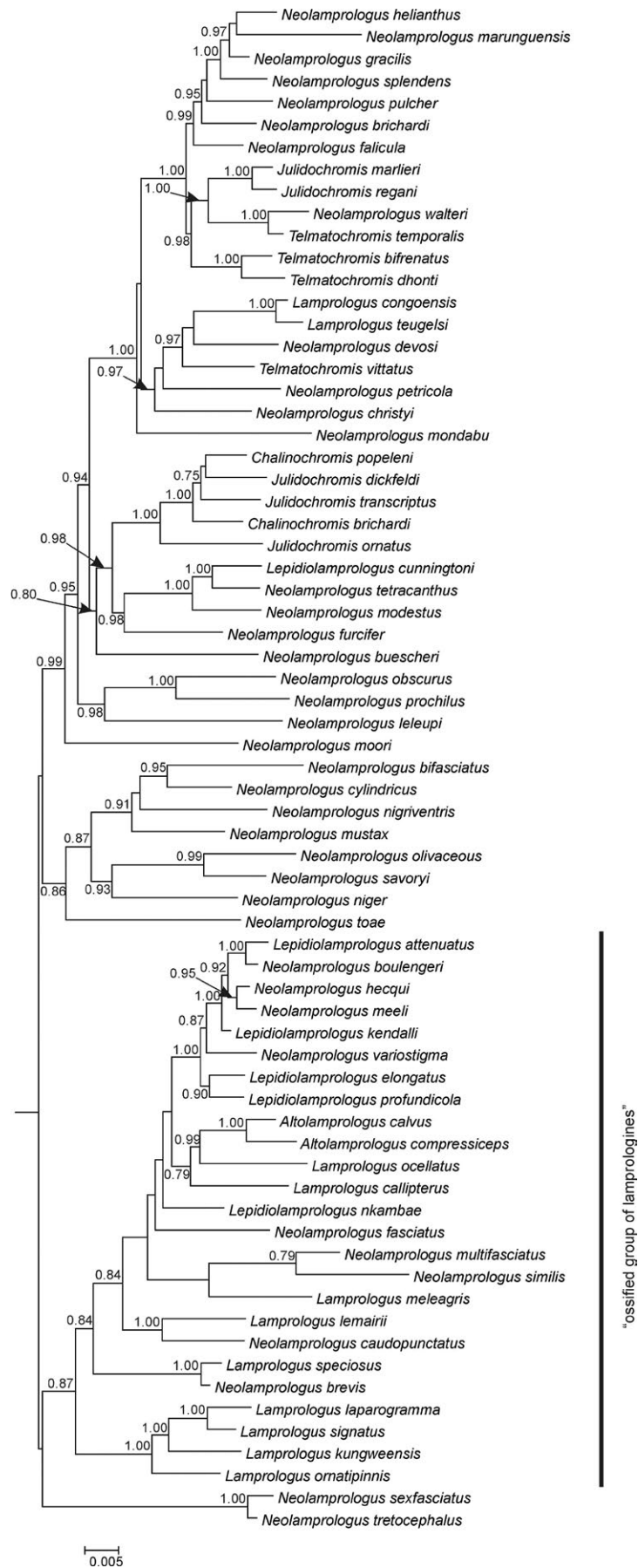
FIGURE A1 Bayesian inference tree based on the mitochondrial dataset. The tree topology of the mtDNA tree is clearly different from the nucDNA shown in Figure A2. Only posterior probabilities >0.70 are shown. Note that node support is high overall
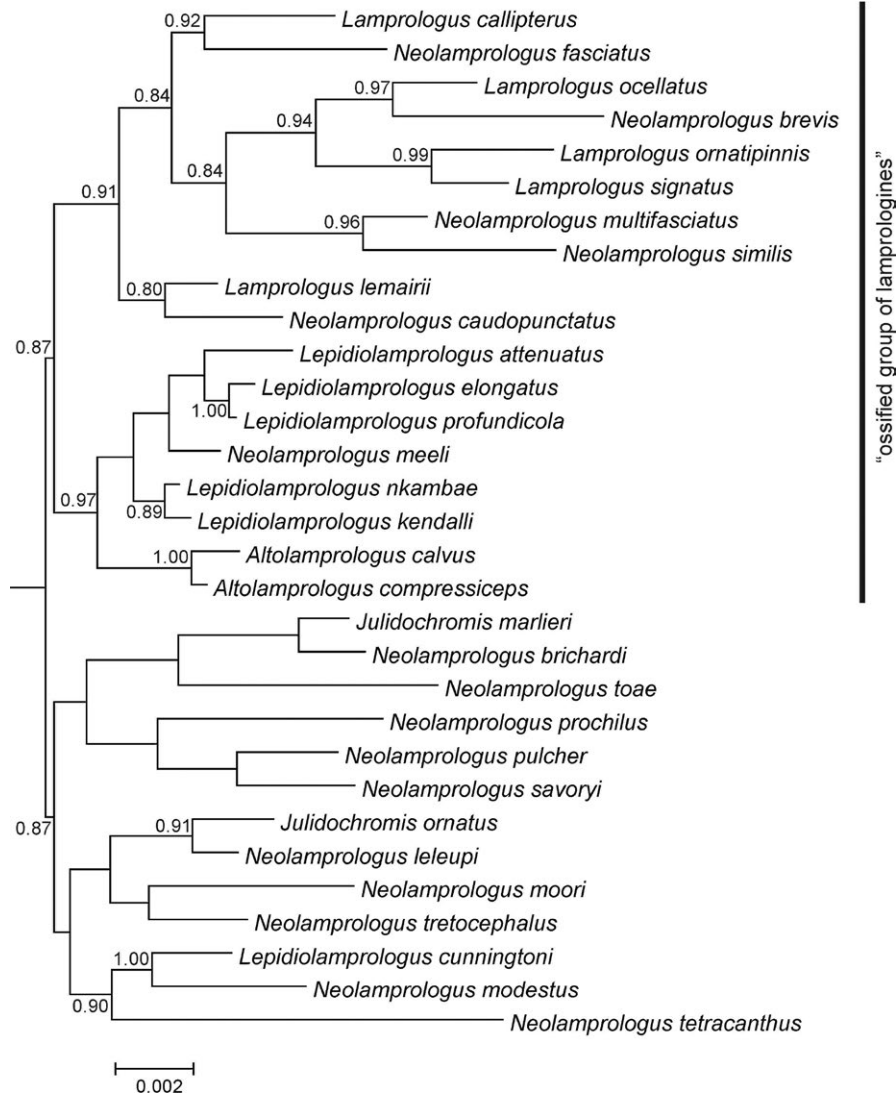
FIGURE A2 Bayesian inference tree of the nuclear dataset. Only posterior probabilities >0.70 are shown. Note that support for many nodes is low

FIGURE A3 Bayesian inference tree of the combined (nuclear plus mitochondrial) dataset. The tree largely mirrors the mitochondrial tree shown in Figure A1. Note that the tree was inferred without any topological constraint, making the topology different compared to Dey et al.'s (2017) tree. Only posterior probabilities >0.70 are shown
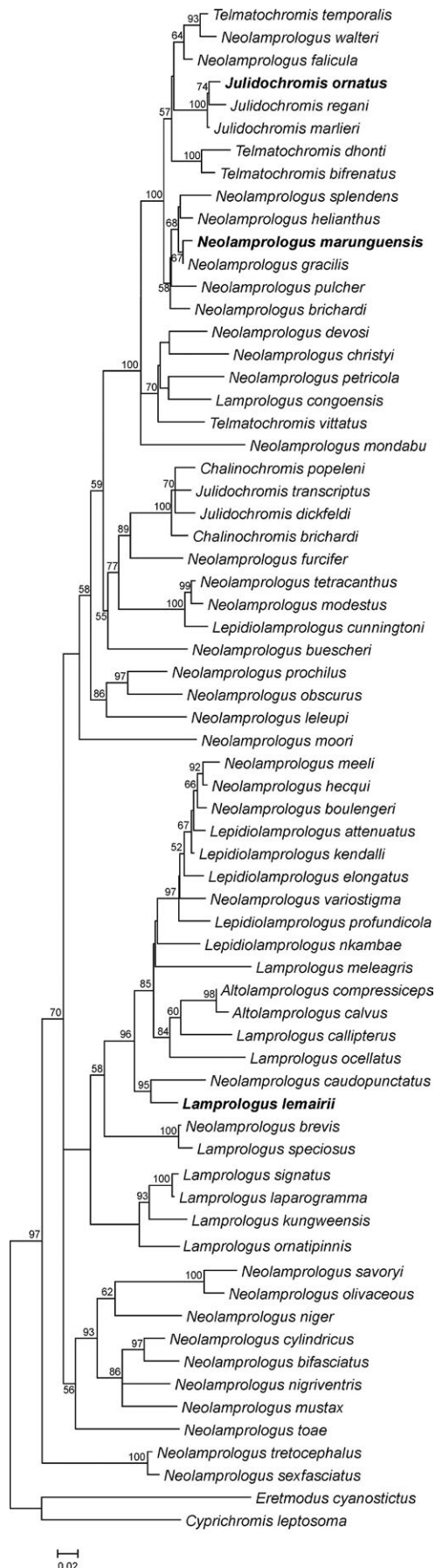
FIGURE A4 Maximum likelihood tree based on the mitochondrial ND2 (NADH dehydrogenase subunit 2) data. The tree was inferred in PhyML 3.0 (Guindon et al., 2010), rooted with the non-lamprologine cichlids *Eretmodus cyanostictus* and *Cyprichromis leptosoma*. While the tree topology should roughly be the same regardless of the particular mitochondrial gene analyzed, the topology differed when different mitochondrial genes were used (Figures A4–A6). Species that appear at clearly different positions in the three mitochondrial trees are highlighted in bold. Only bootstrap support values >50 are shown
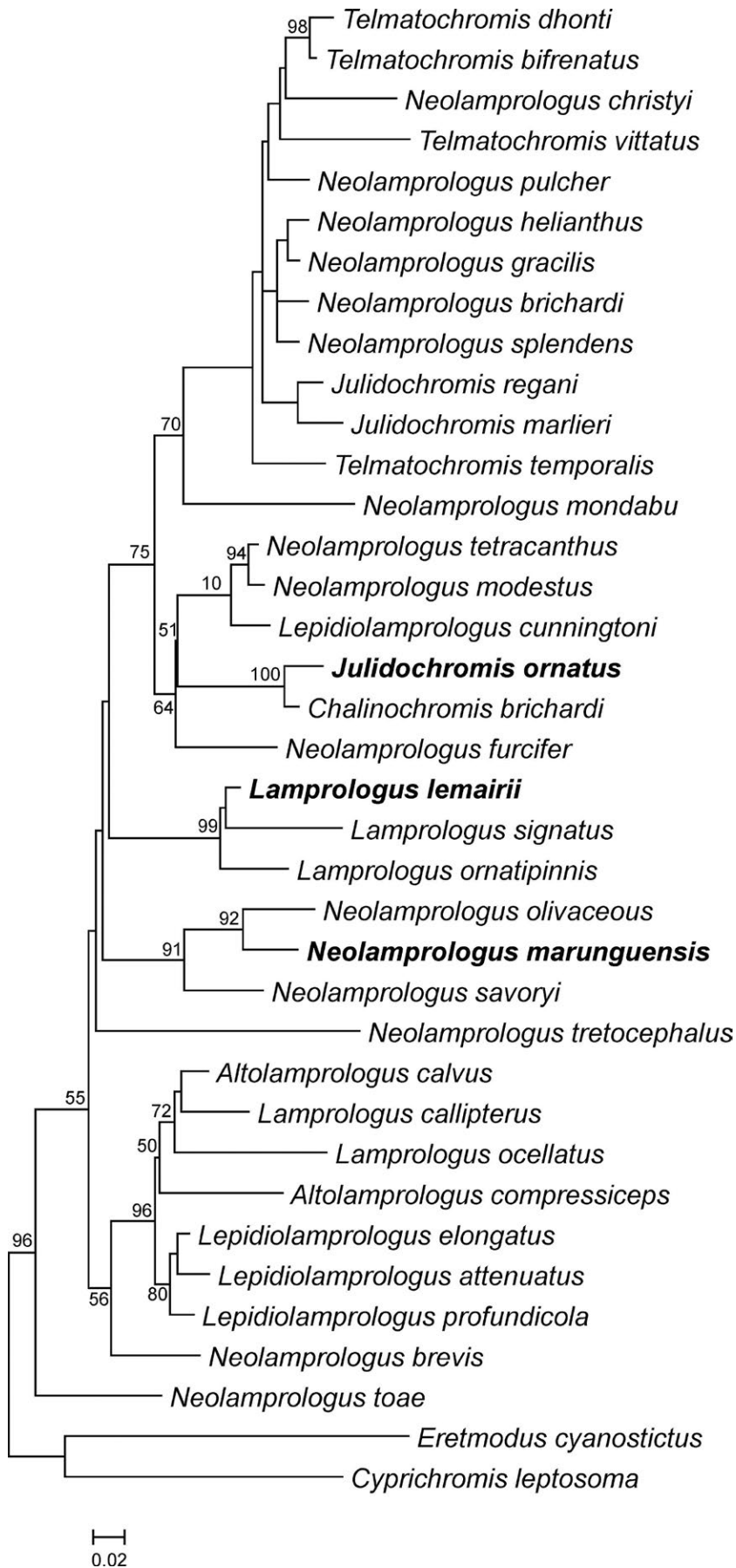
FIGURE A5 Maximum likelihood tree based on the mitochondrial cytb (cytochrome b) data. The tree was inferred in PhyML 3.0 (Guindon et al., 2010), rooted with the non-lamprologine cichlids *Eretmodus cyanostictus* and *Cyprichromis leptosoma*. Only bootstrap support values >50 are shown. Species that appear at clearly different positions in the three mitochondrial trees are highlighted in bold
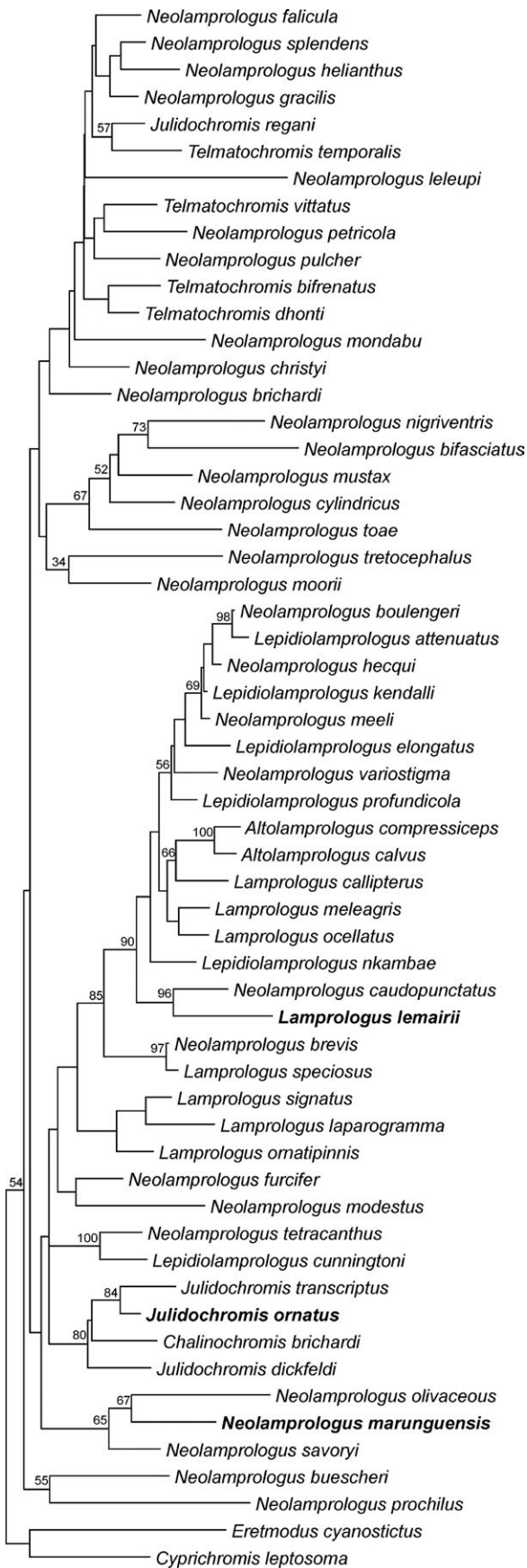
FIGURE A6 Maximum likelihood tree based on the mitochondrial control region data. The tree was inferred in PhyML 3.0 (Guindon et al., 2010), rooted with the non-lamprologine cichlids *Eretmodus cyanostictus* and *Cyprichromis leptosoma*. Only bootstrap support values >50 are shown. Species that appear at clearly different positions in the three mitochondrial trees are highlighted in bold