Deep genetic differentiation between two morphologically similar species of wolf herrings (Teleostei, Clupeoidei, Chirocentridae)

Sébastien Lavoué, Joris Bertrand, Kang-ning Shen, Sahat Ratmuangkhwang, Tetsuya Sado, Masaki Miya, Mohd Nor Siti Azizah

DR. SEBASTIEN LAVOUE (Orcid ID : 0000-0003-4798-6666)

Article type : Original Article

Deep genetic differentiation between two morphologically similar species of wolf herrings (Teleostei, Clupeoidei, Chirocentridae)

Running title: Mitogenomics of wolf herrings

Sébastien Lavoué[1,7,*], Joris A. M. Bertrand[2], Kang-Ning Shen[3], Sahat Ratmuangkhwang[4], Tetsuya Sado[5], Masaki Miya[5], Siti Azizah Mohd Nor[6]

[1] *School of Biological Sciences, Universiti Sains Malaysia, Penang, 11800, Malaysia*

[2] *Laboratoire Génome & Développement des Plantes (UMR 5096), Université de Perpignan Via Domitia, 58 av. P. Alduy, F-66860 Perpignan, France*

[3] *Aquatic Technology Laboratories, Agricultural Technology Research Institute, Hsinchu, Taiwan*

[4] *Andaman Coastal Research Station for Development, Faculty of Fisheries, Kasetsart University, Suksamran, Ranong 85120, Thailand*

[5] *Department of Ecology and Environmental Sciences, Natural History Museum and Institute,*

Chiba, 955-2 Aoba-cho, Chuo-ku, Chiba 260-8682, Japan

[6] *Institute of Marine Biotechnology, Universiti Malaysia Terengganu, Kuala Nerus 21030, Malaysia*

[7] *Open Researcher and Contributor ID (ORCID): 0000-0003-4798-6666*

[*] Corresponding author (email: microceb@hotmail.com)

Acknowledgements

This study was supported by research grants MOST103-2119-M-002-019-MY3 and MOST106-2119-M-002-032 from the Ministry of Science and Technology of Taiwan (to S.L.) and by JST CREST Grant Number JPMJCR13A2, Japan (to M.M.). We express our gratitude to Drs. Shu-Hui Liu, Pei-Chun Lo and Wei-Jen Chen (National Taiwan University, Taipei) for sharing samples.

Summary

Wolf herrings (Chirocentridae; Clupeoidei) are commonly found in local fish markets throughout the Indo-West Pacific region where they constitute an auxiliary source of food and income for local communities. The validity of the two species of wolf herrings, *Chirocentrus dorab* Forsskål, 1775 and *C. nudus* Swainon, 1839, is only supported by slight morphological differences. The identification of either species is challenging, especially for juveniles, and precludes accurate assessments of these natural resources at a species level. As a step towards gaining better knowledge of the genetic structure of these fishes, we examined genetic differentiation between these two species by reconstructing their entire mitogenomic sequences using high-throughput sequencing technology. We found that the mitogenome of each species shared the same gene content and order that were the same for those found in most other teleost fishes. Despite their high morphological similarity, these two species of *Chirocentrus* were genetically well differentiated (*p*-distance = 16.3% at their cytochrome oxidase I). A mitogenomic time-calibrated phylogenetic analysis showed that wolf herrings originated about 35 million years ago, and they represent a case of morphological stasis. Furthermore, comparison of published and newly determined mitochondrial COI barcode

region sequences from 22 individuals revealed species-level cryptic genetic diversity within *C. dorab*. Altogether, these mitochondrial data are effective in discriminating species within this genus and informing population genetic relationships within species of wolf herrings.

Keywords: NGS; mitogenome; fish; Clupeiformes; evolution; phylogeny

1 Introduction

Wolf herrings (Chirocentridae) are unusual fish within the suborder Clupeoidei, the latter also including sardines, anchovies, and herrings (Whitehead, 1985; Lavoué, Konstantinidis, & Chen, 2014). The wolf herrings are distinctive in their piscivorous nature, elongated and laterally compressed body, large size (up to 1 m in total length [TL]), and the absence of a full series of ventral scutes (Figure 1A) (Whitehead, 1985). These characteristics make wolf herrings easily recognizable. They are widely distributed off the continental coasts of the Indo-West Pacific (IWP) region, from southeastern Africa to the China Sea region and, further south, to the tropical Australian region (Whitehead, 1985) (Figure 1B).

According to Luther (1985a) who examined over eight thousand specimens of wolf herrings from Indian waters, they comprise only two sympatric species that are morphologically similar, *Chirocentrus dorab* Forsskål 1775 and *C. nudus* Swainon 1839. These two species can be distinguished from each other by the coloration of the dorsal fin (blackish in *C. dorab* vs. whitish in *C. nudus*) and the relative length of the pectoral fin (longer in *C. nudus*) (Luther, 1985a). In practice, however, wolf herrings are difficult to identify because 1) the interpretation of these two characters needs taxonomical practice, 2) often specimens at fish landing sites have damaged fins, and 3) these two characters are not differentiated in smaller specimens (less than 15 cm TL).

Wolf herrings are pelagic coastal fishes, and while common throughout their range, they do not form large schools. They reach sexual maturity after three years, at about 30-40 cm TL (Luther 1985b; but see Abdussamad, Pillai, Zacharia, & Jeyabalan, 2011 for different estimations), and they spawn from February to August in India (Luther 1985b). Both species have similar diets, as they mainly feed on sardines (e.g., *Sardinella* spp.) and anchovies (e.g., *Stolephorus* spp.), although Luther (1985c) reported that *C. nudus* prefers larger prey.

In the IWP region, wolf herrings represent a common by-catch for artisanal fisheries and, in some places as in India, even a targeted resource (Abdussamad et al., 2011). The annual total landing catch fluctuates, but precise data are lacking. Often, both species are

mixed together in landing catches, and only a few authors, such as Abdussamad et al. (2011), have reported separate data for each species. Both Abdussamad et al. (2011) and Pinsky et al. (2011) reported the overfishing of wolf herrings.

Genetic data of wolf herrings are limited. One nearly complete mitochondrial genome (i.e., 15,989 base-pairs [bp], with a part of the control region undetermined) of a specimen of *C. dorab*, from an unspecified geographical origin, was determined (Ishiguro, Miya, Inoue, & Nishida, 2005), and it is available in the GenBank database under accession number AP006229. In addition, partial cytochrome oxidase I (COI) reference sequences of wolf herrings have been published in several papers which attempted to provide molecular identification of the regional marine ichthyofauna of China (Zhang & Hanner, 2012), Iran (Asgharian, Sahafi, Ardalan, Shekarriz, & Elahi, 2011), India (Lakra et al., 2011), Australia (Ward & Holmes, 2007), and South Africa (Steinke, Connell, & Hebert, 2016).

The main aim of this study was to examine genetic differentiation between the two species of wolf herring by reconstructing the complete or nearly complete mitogenomic sequence of each species by high-throughput sequencing (also known as next-generation sequencing, NGS) technology. We also provide preliminary data on genetic differentiation within each species by comparing the COI reference gene for several specimens of *C. dorab* and *C. nudus* from different geographical locations.

2 Materials and Methods

2.1 Specimens sampling

Complete or nearly complete mitogenomic sequences were determined from two mid-sized (about 35 and 45 cm in total length) fresh specimens of *Chirocentrus* collected in two local fish markets in Malaysia. The specimen of *C. dorab* was collected at a fish market in Kuala Perlis, Perlis, East coast of Peninsular Malaysia. The specimen of *C. nudus* was collected at a fish market in Kota Bharu, Kelantan, West coast of Peninsular Malaysia. Species identifications were done by inspecting the dorsal fin coloration (blackish in *C. dorab* vs. whitish in *C. nudus*) and the relative length of the pectoral fin (longer in *C. nudus*).

Muscle tissue samples were taken from the two fresh specimens, immediately preserved in 95% ethanol, and subsequently stored at -20 °C. Both specimens were deposited in the zoological collections of University Sains Malaysia, Penang, Malaysia.

Genomic DNA was extracted from each muscle sample using a commercial kit (DNeasy Blood and Tissue Kit, Qiagen, Hilden, Germany).

2.2 High throughput sequencing of mitogenomes

The mitogenome of *C. dorab* was first targeted with a long polymerase chain reaction (PCR) amplification technique into two overlapping fragments. The following fish-versatile long PCR primers were used in various combinations to amplify the entire mitogenome in three or four reactions: forward L2508-16S (5´-CTC GGC AAA CAT AAG CCT CGC CTG TTT ACC AAA AAC-3´), forward L8343-lys (5´- AGC GTT GGC CTT TTA AGC TAA WGA TWG GTG-3´), forward L12321-leu (5´- GGT CTT AGG AAC CAA AAA CTC TTG GTG CAA-3´), reverse H2990-16S (5´-TGC ACC ATT RGG ATG TCC TGA TCC AAC ATC-3´), H12293-leu (5´-TTG CAC CAA GAG TTT TTG GTT CCT AAG ACC-3´), reverse H1065-12S (5´- GGC ATA GTG GGG TAT CTA ATC CCA GTT TGT-3´) and reverse HS-LA-16S (5´- TGC ACC ATT RGG ATG TCC TGA TCC AAC ATC-3´).

Reactions were carried out in 25 μl reaction volume containing 15.25 μl of sterile distilled H2O, 2.5 μl of 10× LA PCR buffer II (Takara), 4.0 μl of dNTP (2.5 mM), 1.0 μl of each primer (5 μM), 0.25 μl of 1.25-unit LA Taq (Takara), and 1.0 μl of template containing approximately 5 ng DNA. The thermal cycle profile was: denaturation at 98°C for 10 s and annealing and extension combined at the same temperature (68°C) for 16 min. Each reaction comprised 30 thermal cycles. Then, a DNA library was prepared from the long PCR products using the Nextera XT DNA Library Preparation Kit (Illumina, San Diego, CA, USA) following the manufacturer's protocol, before to be sequenced on a MiSeq Sequencing platform with MiSeq Reagent Kit v2 PE 300 cycles (Illumina) at the Natural History Museum and Institute, Chiba, Japan.

A whole genomic DNA library of *C. nudus* was constructed following the manufacturer's protocol, then massively sequenced (using 2x250bp Rapid Run chemistry; paired-end) on an Illumina Hiseq 2500 sequencer (Illumina) at the Agricultural Technology Research Institute (Hsinchu, Taiwan).

2.3 Mitogenome reconstruction and annotation

We used the baiting and iterative mapping procedure implemented in MITObim vers.1.8 (Hahn, Bachmann, & Chevreux, 2013) to reconstruct the mitochondrial genome of each of the two individuals. Raw reads were first trimmed by quality with the FASTQ Quality Trimmer script (Blankenberg et al., 2010) available at the online Galaxy portal (www.usegalaxy.org). Reads were trimmed at both the 5' and 3' ends until the aggregate

quality score was ≥ 20 (with all other settings at default values). We then performed reconstructions following the two main approaches available in the MITObim pipeline. First, we used as a starting reference the previously published mitochondrial genome of *C. dorab*. Second, we used conspecific COI sequences as a seed to initiate the process. The circularity of the mitochondrial genomes was inferred using editing features provided in Geneious vers.6.1.8 (Auckland, New Zealand). Raw reads were mapped onto the resulting consensus sequences to check for assembly success and assess coverage.

The consensus sequences were annotated using MitoAnnotator (Iwasaki et al., 2013) and exported for subsequent analyses. The two mitogenomes determined in this study were archived in the DDBJ/EMBL/GenBank databases under the accession numbers AP018761 (*C. nudus*) and AP018763 (*C. dorab*).

2.4 Mitogenomic time-calibrated phylogenetic tree reconstruction

We first built a phylogenetic matrix "taxa*characters" as previously described (Lavoué, Miya, Kawaguchi, Yoshino, & Nishida, 2008) by combining the two newly determined mitogenomes with a selection of 24 mitogenomes of clupeoid species, purposely chosen based on the current knowledge on the systematics of Clupeoidei, and that of *Denticeps clupeoides*, the sister group of Clupeoidei. The 12S and 16S ribosomal (r)RNAs were aligned using the online version of MAFFT vers.7 (Katoh & Standley 2013; Katoh, Rozewicki, & Yamada, *in press*), and default setting parameters. The matrix (15,248 positions) includes concatenated nucleotide sequences from 12 protein-coding genes (10,890 positions), 22 tRNA genes (1579 positions) and the 12S and 16S rRNA genes (2779 positions).

From this matrix, we reconstructed a time-calibrated Bayesian phylogenetic tree using the software BEAST vers.1.10.3 (Suchard et al., 2018). We selected a general time-reversible model of sequence evolution with discrete gamma-distributed rate heterogeneity [GTR + Γ], a relaxed log-normal molecular clock and two time-constraints (see below). Two independent runs of 10^8 generations each were performed using BEAST. Estimation of trees and divergence time were sampled once every 5,000 generations and the parameters of each run were checked for convergence with the software Tracer vers.1.7.1 (Rambaut, Drummond, Xie, Baele, & Suchard, 2018). After removing the burn-in part of each run (10%), the remaining tree samples from the two runs were pooled into a combined file, and the maximum clade credibility tree with mean divergence times and their 95% credibility intervals was built using TreeAnnotator vers.1.10.3.

To calibrate the molecular clock, we used the fossil *Lecceclupea ehiravaensis* (Taverne,

2007) to constrain the age of the clade containing *Gilchristella* and *Sundasalanx* to a minimum of 74.0 million years ago which corresponds to the minimum age of this fossil [exponential distribution offset = 74.0; mean = 0.5]. The age of the tree, i.e. the divergence between *Denticeps clupeoides* and Clupeoidei was constrained to 130 million years ago following the results of Lavoué et al (2013) [exponential distribution offset = 130.0; mean = 0.1].

2.5 Intraspecific genetic differentiation

We downloaded all COI sequences previously published and available in GenBank of *C. dorab* (12 sequences) and *C. nudus* (five sequences) that we combined with 10 newly determined sequences (from Taiwan, Malaysia, and Thailand; see localities in Table 1 and Figure 1B). We used the PCR technique to amplify the COI sequences using the following PCR primer pair: forward COI_F1 (5´-TCAACCAACCACAAAGACATTGGCAC-3´) and reverse COI_R2 (5´-ACTTCAGGGTGACCGAAGAATCAGAA-3´) (Ward, Zemlak, Innes, Last, & Hebert, 2005). Reactions were carried out in 25 µl reaction volume containing 9.5 µl of sterile distilled H2O, 12.5 µl of Fast-Run Advanced Taq Master Mix (Pro Tech), 1.0 µl of each primer (5 µM), and 1.0 µl of template containing approximately 5 ng DNA. The thermal cycle profile consisted of an initial 94°C denaturation step for 10 min, 35 cycles of 94 °C for 1 min, annealing for 1 min at 55 °C, extension at 72 °C for 1.5 min, followed by a final extension at 72 °C for 10 min. PCR products were sequenced by the classical Sanger method. Evolutionary relationships among these COI haplotypes were inferred with 1) a rooted maximum likelihood (ML) phylogenetic tree using RAxML vers.8 (Stamatakis 2014) and a GTR + Γ model of sequence evolution and 2) separately for each of the two species, unrooted networks constructed with the program PopArt (Leigh & Bryant, 2015) using a median-joining algorithm (Bandelt, Forster, & Rohl, 1999) and default settings.

3 Results

For *Chirocentrus nudus*, 57,624 reads were mapped to construct a single, circular, complete mitogenome with an average coverage of 840 X. For *Chirocentrus dorab*, 83,106 reads were mapped to construct a single, nearly complete mitogenome with an average coverage of 630 X. The lengths of the mitogenomes of *C. nudus* and *C. dorab* were 16,722 bp and 16,552 bp, respectively. Each contained 13 protein-coding genes, two rRNA genes, 22 transfer (t)RNA genes, and a main non-coding region, the control region. Excluding the control region, we

found that our mitogenomic sequence of *C. dorab* was overall 99% identical to the published sequence of the same species (GenBank accession number: AP006229; Ishiguro et al., 2005). The overall genetic *p*-distance between the newly determined mitogenomes of *C. dorab* and *C. nudus* was about 15%. In particular, cytochrome b diverged by 16%, and the complete COI sequence diverged by 16.3%.

The Bayesian time-calibrated phylogenetic reconstruction (Figure 2) showed that *C. dorab* is the sister group to the sequence of *C. nudus*, making the genus *Chirocentrus* monophyletic. *Chirocentrus* is a sister group to a clade comprising the two genera of small sprats, *Spratelloides* and *Jenkinsia* (subfamily Spratelloininae), and a still undescribed paedomorphic taxon (Li and Ortí, 2007; Lavoué, Miya, Musikasinthorn, Chen, & Nishida, 2013). The divergence time between the two lineages of *Chirocentrus* is estimated to about 35 million years ago [95% confidence interval: 52.7-17.3 million years ago].

Aligned sequences of 648 bp of the mitochondrial COI gene from 22 individuals of *C. dorab* (Table 1) were grouped into three main genetic groups (Figure 3A) and nine haplotypes (Figure 3B, left). Specimens of each main groups have different geographical origins (Figure 3A): Australian waters (one haplotype), South African waters (two haplotypes), and northern Indo-West Pacific waters, from west Indian Ocean to South China Sea (six haplotypes from seven localities). Each lineage differed from the others by at least 46 nucleotide differences (*p*-genetic distance > 7%) (Figure 3B, left). Aligned sequences of the 648 bp of COI from only six individuals of *C. nudus* from three localities of the northern Indo-West Pacific region grouped into five weakly differentiated haplotypes (Figure 3B, right).

4 Discussion

4.1 Interspecific genetic differentiation

Thanks to continuous progress in DNA sequencing technology combined with taxonomic sampling efforts, the mitogenomes of more than 1500 (of about 32,000) species of Teleostei have already been determined, including more than 100 (of about 400) species of sardines, anchovies, and their relatives (in the suborder Clupeoidei). The uses of mitogenomic sequences are multiple: they can be used to infer the taxonomic status, evaluate genetic differentiation, reconstruct phylogenetic relationships and evolutionary histories, and design oligonucleotide primers to target short informative genetic markers for population genetics and phylogeographic assessments. Ultimately, such data are crucial for assessing potential negative effects of anthropogenic disturbances (such as overfishing) on species and

populations and preparing management plans for the long-term sustainability of natural resources.

The gene content and order of the two newly determined mitogenomic sequences are typical of those found in most other clupeoids and teleosts (Satoh, Miya, Mabuchi, & Nishida, 2016). So far, gene rearrangements within Clupeoidei have only been discovered in one taxon, a marine paedomorphic species which is still undescribed (Lavoué et al., 2008).

Despite low morphological differentiation between the two species of wolf herring, mitogenomic information demonstrated that *C. dorab* and *C. nudus* diverged from each other about 35 million years ago. This represents a case of morphological stasis in a group, which otherwise exhibits several specialized characters (e.g., the presence of canine-like teeth).

We believe that the mitogenomes of the two species of *Chirocentrus* represent a useful genetic resource for targeting individual molecular markers (and designing specific primers to amplify them) to further study the population genetics and phylogeography of these two species.

4.2. Intraspecific genetic structure

Preliminary COI sequence data of *C. dorab* suggest both low (=intraspecific level) genetic differentiation (*p*-distance < 1%) across the northern part of the Indo-West Pacific (IWP) region and important (=species-level; Ward et al., 2005; April, Mayden, Hanner & Bernatchez, 2011) genetic divergence (*p*-distance > 7%) among this region, the Australian region, and the South African region. Several phylogeographic/phylogenetic studies on IWP marine fish have evidenced similar genetic patterns with the distinction of non-overlapping distributed lineages in similar regions (e.g., Kulbicki et al., 2013). This suggests that fauna of these regions have been isolated for long periods of time since the late Cenozoic and remain, nowadays, mostly isolated. Our results support the presence of three putative allopatric cryptic species in *C. dorab*. However, the precise limits among distributions of these three lineages remain to be identified by analyzing a geographically denser taxonomic sampling. In addition, studies of nuclear markers and morphological characters are needed to precisely determine the systematic status of these three allopatric cryptic lineages. *Chirocentrus dorab* was described by Forsskål (1775) who indicated two localities in the Red Sea: "Djiddae" (currently known as Jeddah, Saudi Arabia) and "Mochhae" (currently known as Al-Mukhā, Yemen). There is "no type known" according to Eschmeyer, Fricke, & van der Laan (2018). We did not sample the Red Sea, and therefore, we were unable to determine whether the Red Sea population is closer to the lineage of the northern part of the IWP than to the lineage of

South Africa or it represents another distinct lineage.

Although *C. nudus* also occurs in Australian and South African regions (Whitehead, 1985), we could not examine any COI sequence of this species from these two regions. We can only conclude that the genetic differentiation within the northern part of the IWP region is similar to that of *C. dorab*: at the intraspecific level. Furthermore, specimens from the Indian Ocean and those from the West Pacific Ocean are not reciprocally monophyletic, indicating either gene flow or incomplete lineage sorting. Our taxonomic and character samplings are too incomplete to provide reliable statistics into the population genetics of *C. dorab* and *C. nudus* but we expect that the results of this study will serve to develop research on these charismatic and economically important fish.

References

Abdussamad, E. M., Pillai, N. G. K., Zacharia, P. U., & Jeyabalan, K. (2011). Dorab fishery of Gulf of Mannar waters and population characteristics of the species *Chirocentrus dorab* (Forsskal, 1775) and *Chirocentrus nudus* Swainson, 1839. *Indian Journal of Fisheries, 58*, 19-23.

April J., Mayden, R. L., Hanner, R. H., & Bernatchez, L. (2011). Genetic calibration of species diversity among North America's freshwater fishes. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 10602-10607. https://doi.org/10.1073/pnas.1016437108

Asgharian, H., Sahafi, H. H., Ardalan, A. A., Shekarriz, S., & Elahi, E. (2011). Cytochrome c oxidase subunit 1 barcode data of fish of the Nayband National Park in the Persian Gulf and analysis using meta-data flag several cryptic species. *Molecular Ecology Resources, 11*, 461-472. https://doi.org/10.1111/J.1755-0998.2011.02989.X

Bandelt, H. J., Forster, P., & Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution, 16*, 37-48. https://doi.org/10.1093/oxfordjournals.molbev.a026036

Blankenberg, D., Gordon, A., Von Kuster, G., Coraor, N., Taylor, J., Nekrutenko, A., & Team, G. (2010). Manipulation of FASTQ data with Galaxy. *Bioinformatics, 26*, 1783-1785. https://doi.org/10.1093/bioinformatics/btq281

Eschmeyer, W. N., Fricke, R., & van der Laan, R. (Eds.). (2018). *Catalog of Fishes electronic*

version (May 2018). *http://research.calacademy.org/ichthyology/catalog/fishcatmain.asp*.

Forsskål, P. (1775). *Descriptiones animalium avium, amphibiorum, piscium, insectorum, vermium; quae in itinere orientali observavit* (Niebuhr Ed.): Niebuhr.

Hahn, C., Bachmann, L., & Chevreux, B. (2013). Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads-a baiting and iterative mapping approach. *Nucleic Acids Research, 41*, e129. https://doi.org/10.1093/nar/gkt371

Ishiguro, N. B., Miya, M., Inoue, J. G., & Nishida, M. (2005). *Sundasalanx* (Sundasalangidae) is a progenetic clupeiform, not a closely-related group of salangids (Osmeriformes): mitogenomic evidence. *Journal of Fish Biology, 67*, 561-569. https://doi.org/10.1111/j.0022-1112.2005.00746.x

Iwasaki, W., Fukunaga, T., Isagozawa, R., Yamada, K., Maeda, Y., Satoh, T. P., . . . Nishida, M. (2013). MitoFish and MitoAnnotator: A Mitochondrial Genome Database of Fish with an Accurate and Automatic Annotation Pipeline. *Molecular Biology and Evolution, 30*, 2531-2540. https://doi.org/10.1093/molbev/mst141

Katoh, K., & Standley, D.M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: improvements in performance and usability. *Molecular Biology and Evolution, 30*, 772-780. https://doi.org/10.1093/molbev/mst010

Katoh, K., Rozewicki, J., & Yamada, K.D. (in press). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*. https://doi.org/10.1093/bib/bbx108 (advance access publication date: 6 September 2017)

Kulbicki, M., Parravicini, V., Bellwood, D. R., Arias-Gonzalez, E., Chabanet, P., Floeter, S. R., . . . Mouillot, D. (2013). Global Biogeography of Reef Fishes: A Hierarchical Quantitative Delineation of Regions. *PLoS ONE*, 8(12). https://doi.org/10.1371/journal.pone.0081847

Li, C., & Ortí, G. (2007). Molecular phylogeny of Clupeiformes (Actinopterygii) inferred from nuclear and mitochondrial DNA sequences. *Molecular Phylogenetic and Evolution, 44*, 386-398.

Lakra, W. S., Verma, M. S., Goswami, M., Lal, K. K., Mohindra, V., Punia, P., . . . Hebert, P. (2011). DNA barcoding Indian marine fishes. *Molecular Ecology Resources, 11*, 60-71. https://doi.org/10.1111/j.1755-0998.2010.02894.x

Lavoué, S., Miya, M., Kawaguchi, A., Yoshino, T., & Nishida, M. (2008). The phylogenetic

position of an undescribed paedomorphic clupeiform taxon: mitogenomic evidence. *Ichthyological Research, 55*, 328-334. https://doi.org/10.1007/s10228-008-0044-3

Lavoué, S., Miya, M., Musikasinthorn, P., Chen, W.-J., & Nishida, M. (2013). Mitogenomic Evidence for an Indo-West Pacific Origin of the Clupeoidei (Teleostei: Clupeiformes). *PLoS ONE, 8*, e56485. https://doi.org/10.1371/journal.pone.0056485

Lavoué, S., Konstantinidis, P., & Chen, W.-J. (2014). Progress in Clupeiformes Systematics. In: Ganias, K., editor. *Biology and Ecology of Anchovies and Sardines*. Enfield, New Hampshire: Science Publishers. p. 3-42.

Leigh, J. W., & Bryant, D. (2015). PopART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution, 6*, 1110-1116. https://doi.org/10.1111/2041-210x.12410

Luther, G. (1985a). Studies on the biology and fishery of the fishes of the genus *Chirocentrus* Cuvier I. Taxonomy. *Matsya, 11*, 46-55.

Luther, G. (1985b). Studies on the biology and fishery of the fishes of the genus *Chirocentrus* Cuvier IV : Reproduction. In *Recent Advances in Marine Biology*. New Delhi, India: Today and Tomorrow Printers and Publishers. p. 439-514.

Luther, G. (1985c). Food and feeding habits of the two species of *Chirocentrus* from Mandapam. *Indian Journal of Fisheries* 32, 439-446.

Pinsky, M. L., Jensen, O. P., Ricard, D., & Palumbi, S. R. (2011). Unexpected patterns of fisheries collapse in the world's oceans. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 8317-8322. https://doi.org/doi:10.1073/pnas.1015313108

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematics Biology*, 67:901-904. https://doi.org/10.1093/sysbio/syy032

Satoh, T. P., Miya, M., Mabuchi, K., & Nishida, M. (2016). Structure and variation of the mitochondrial genome of fishes. *Bmc Genomics, 17*. https://doi.org/10.1186/S12864-016-3054-Y

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics, 30*, 1312-1313.

Steinke, D., Connell, A. D., & Hebert, P. D. N. (2016). Linking adults and immatures of South African marine fishes. *Genome, 59*, 959-967. https://doi.org/10.1139/gen-2015-0212

Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J., & Rambaut, A. (2018).

Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution, 4*, vey016. https://doi.org/10.1093/ve/vey016

Taverne L. 2011. Les poissons crétacés de Nardò. 33°. *Lecceclupea ehiravaensis* gen. et sp. nov. (Teleostei, Clupeidae). *Bollettino del Museo Civico di Storia Naturale di Verona, Geologia Paleontologia Preistoria, 35*, 3-17.

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., & Hebert, P. D. N. (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B-Biological Sciences, 360*(1462), 1847-1857. https://doi10.1098/Rstb.2005.1716

Ward, R. D., & Holmes, B. H. (2007). An analysis of nucleotide and amino acid variability in the barcode region of cytochrome c oxidase I (cox1) in fishes. *Molecular Ecology Notes, 7*, 899-907. https://doi.org/10.1111/j.1471-8286.2007.01886.x

Whitehead, P. J. P. (1985). Clupeoid fishes of the World (Suborder Clupeoidei): An annotated and illustrated catalogue of the herrings, sardines, pilchards, sprats, shads, anchovies and wolf herrings. Part 1. Chirocentridae, Clupeidae and Pristigasteridae. *FAO Fisheries Synopsis, 125*(part 1), 1-303.

Zhang, J., & Hanner, R. (2012). Molecular approach to the identification of fish in the South China Sea. *PLoS ONE, 7*, e30621. https://doi.org/10.1371/journal.pone.0030621

Figure captions

Figure 1. A) Lateral view of a specimen of *Chirocentrus dorab* (east Indian Ocean, Ranong, Thailand; horizontal scale: 5 cm) and close-up view of its head; B) Indo-West Pacific map showing distribution (yellow shading) and collection localities of *Chirocentrus dorab* (stars) and *Chirocentrus nudus* (circles). White circle and white star indicate the collection localities of the two specimens from which we reconstructed the complete mitogenomes. Localities are numbered from one to nine.

Figure 2. Mitogenome-based phylogenetic chronogram of the Clupeoidei based on a Bayesian relaxed clock analysis (using BEAST v1.10.3) showing the monophyly of wolf herrings *Chirocentrus* and their divergence time. Divergence time in this tree were calibrated with two constraints (see text for details). *Denticeps clupeoides* was used to root the tree. Horizontal timescale is in million years before present (Ma) (Paleogene Epoch abbreviations:

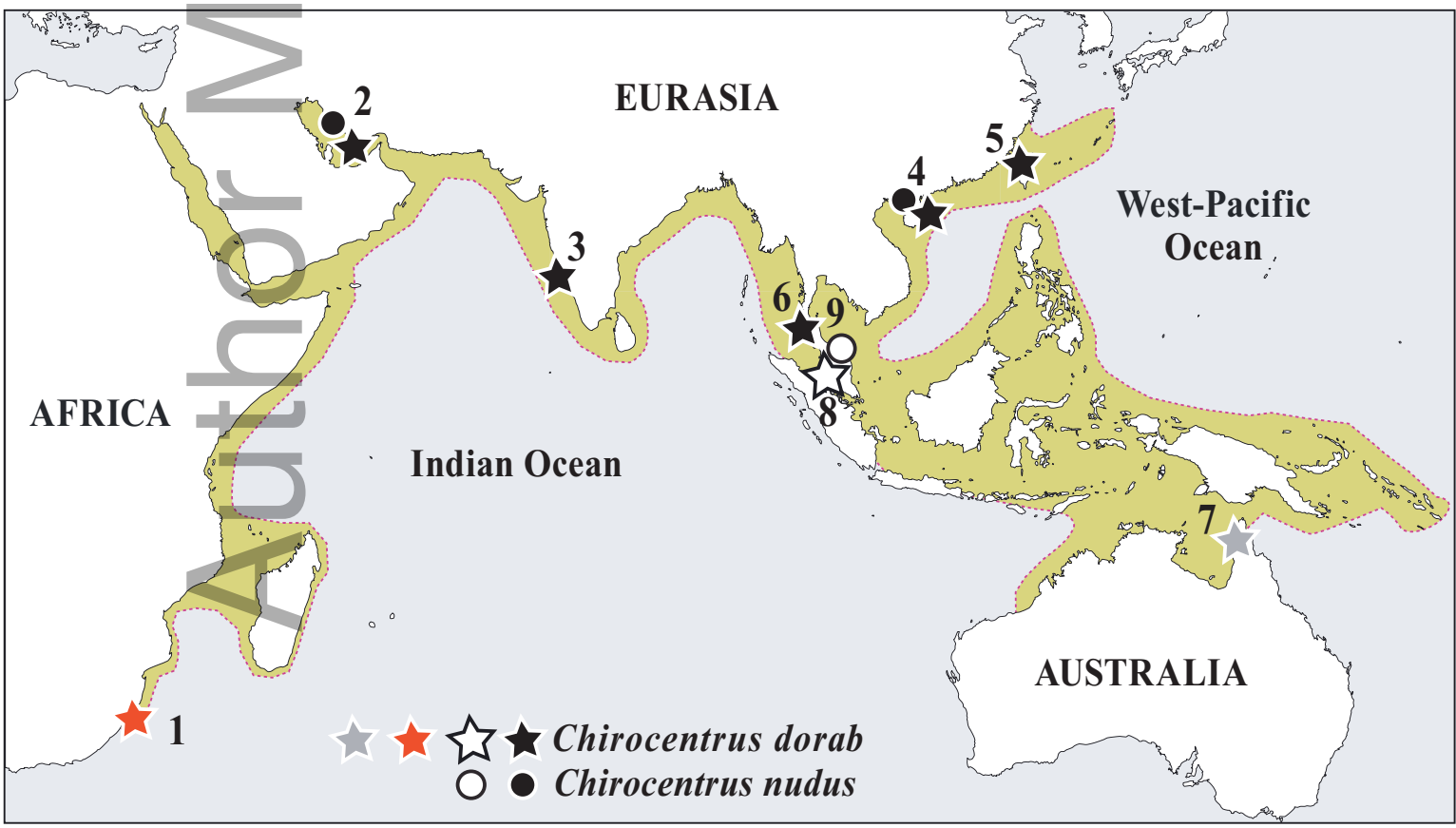
438 Paleo, Paleocene; Eo, Eocene; and Oligo, Oligocene). Black horizontal bars (indicating
439 calibration constraints on the corresponding nodes) and light grey gradient horizontal bars at
440 nodes are 95% age credibility intervals. Numbers given at nodes are the Bayesian posterior
441 probabilities when <1.
442
443 **Figure 3.** A) Maximum-likelihood phylogenetic tree of *Chirocentrus* using COI sequences.
444 Branch lengths are proportional to number of substitutions. This tree is rooted at the mid-
445 point; B) Median-joining networks showing relationships among mitochondrial DNA
446 cytochrome oxidase subunit I (COI) haplotypes for each species of *Chirocentrus* within the
447 Indo-West Pacific region (using the software PopART). Each circle represents a unique
448 haplotype and its size is proportional to its total frequency. Black cross-bars represent a single
449 nucleotide change, small black circles represent missing haplotypes, and colors denote four
450 subregions within the Indo-West Pacific region, as indicated in the embedded legend.

Table I. List of specimens of wolf herrings for which COI sequences are examined in this study along with their geographical origin (and their estimated latitude [lat.] and longitude [long.]), number of specimens per locality and species, code specimen used, and the reference where these data are published.

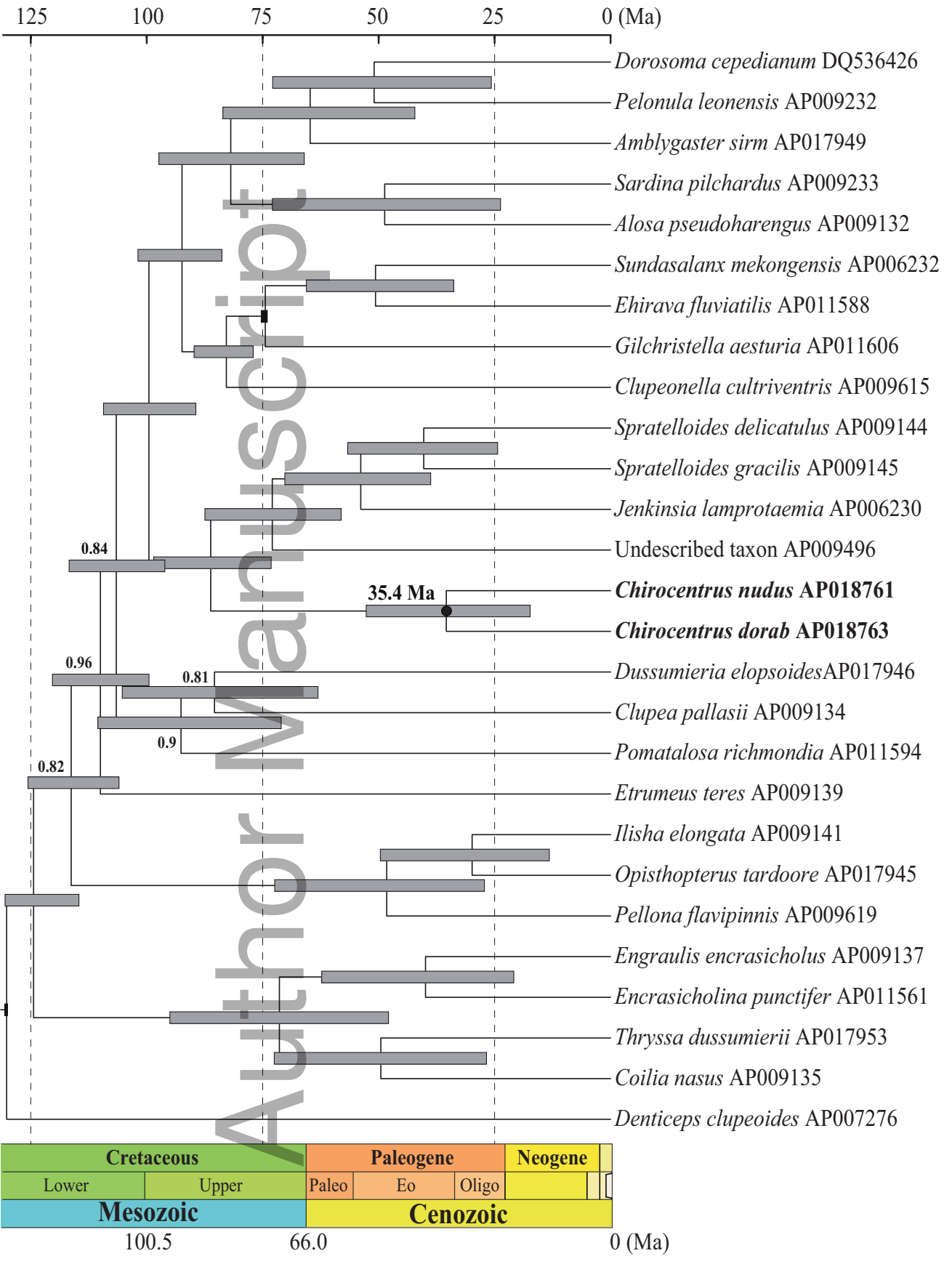
Species	Geographical origin	Estimated lat., long.	Specimen number	Code specimen	Reference
Chirocentrus dorab	South Africa, off Durban	-30.06, 31.66	3	JF493144-6	Steinke et al (2016)
	South Africa, Tugela Banks	-30.06, 31.66	2	HQ945860, GU805007	Steinke et al (2016)
	India, off Bombay	19.09, 72.26	4	FJ347874-7	Lakra et al (2011)
	Iran, Gulf Persia, off Bushehr	28.85, 50.59	2	HQ149825-6	Asgharian et al (2011)
	China, off Hainan Island	18.53, 111.22	1	EF607342	Zhang et al (2012)
	Malaysia, off Perlis	6.25, 100.02	1	MAY1 (mt)	This study
	Thailand, off Ranong	9.77, 98.32	5	Ra30, Ra32, Ra33, Ra34, Ra35	This study
	Taiwan, off Kaohsiung	22.40, 120.13	3	HH81, HH82, HH83	This study
	Australia, Queensland	-23.41, 153.33	1	EF609327	Ward et Holmes 2007
Chirocentrus nudus	Iran, Gulf Persia, off Bushehr	28.79, 50.52	1	HQ149827	Asgharian et al (2011)
	China, off Hainan Island	18.53, 111.22	4	EF607343-6	Zhang et al (2012)

A)

B)



jai_13886_f1.eps



jai_13886_f2.eps

jai_13886_f3.eps