Complete mitochondrial genome of the smallmouth hardyhead (Atherinosoma microstoma) and its phylogenetic position among the Atheriniform fishes

Davin H. E. Setiamarga\textsuperscript{a,b}

\textsuperscript{a}Department of Applied Chemistry and Biochemistry, National Institute of Technology, Wakayama College, Gobo City, Wakayama, Japan; \textsuperscript{b}The University Museum, The University of Tokyo, Tokyo, Japan

**ABSTRACT**

In this paper, I report the full mitochondrial genome sequence of the smallmouth hardyhead (Atherinosoma microstoma), an endemic marine fish from the shallow coastal waters of southeastern Australia. The mitogenome is 16,573 bp-long with the standard 37 genes all included, with a genomic structure typical of a vertebrate mitogenome. In order to confirm the phylogenetic position of this species, phylogenetic trees were inferred using a data set including publicly available 28 atherinomorph, nine percomorph, and two outgroup mitogenome sequences. The complete mitogenome data of \textit{A. microstoma} reported here will be useful for further genetics, phylogeography, and phylogenetics studies involving this species.

**ARTICLE HISTORY**

Received 10 May 2017
Accepted 21 May 2017

**KEYWORDS**

Atheriniformes; mitogenome; percomorph teleosts; small-mouth hardyhead

The smallmouth hardyhead, \textit{Atherinosoma microstoma}, is one of the only two known species of its genus. Taxonomically, it belongs to the family Atherinidae, order Atheriniformes, series Atherinomorpha. The species is native to Australia, and distributed along the shallow coastal waters of southeastern Australia. Here, we report the full mitochondrial genome (mitogenome) sequence of this species. Tissue sample was collected from an individual specimen from the Fish Collection of the Ichthyology Department of the Australian Museum (Voucher No. I.40457-001). Detailed information of the specimen can be obtained online (http://collections.australianmuseum.net.au/amweb/pages/am/Display.php?rm=8142666&QueryPage=%2Famweb%2Fpages%2Fam%2FAdvQuery.php&highlight_term=). PCR-based mitogenome sequencing using fish versatile primers was conducted in accordance to what was reported previously (Miya & Nishida 1999; Setiamarga et al. 2008). Assembled mitogenome sequence was annotated using the MitoAnnotator on the MitoFish homepage (http://mitofish.aori.u-tokyo.ac.jp/annotation/input.html).

The newly sequenced mitogenome of \textit{Atherinosoma microstoma} was 16,573 bp long (Registered to DDBJ). Its genomic structure is similar to a typical vertebrate/euteleost mitogenome with some interesting exceptions. (1) There are 13 protein-coding, two rRNA, and 22 tRNA genes, (2) Most genes are coded on the H chain, except for ND6 and eight tRNA genes (tRNAPro, tRNAGln, tRNAAla, tRNAAsn, tRNAAsp, tRNATyr, tRNASer(UCN), and tRNAIle); (3) ND4L is located on the H chain, unlike a that of a typical vertebrate. The total GC content the mitogenome was 46.4%.

In order to check the phylogenetic position of \textit{A. microstoma} molecularly, we collected publicly available mitogenomes of 12 Atheriniformes, eight Beloniformes, eight Cyprinodontiformes, nine non-atherinomorph percomorphs, and two non-percomorph outgroups from GenBank, and then built a data set for phylogenetic analyses. I conducted a maximum likelihood phylogenetic analysis, which methods are detailed in the legend of Figure 1. The resulting phylogeny was congruent with previous mitogenome (e.g. Kawahara et al. 2008; Setiamarga et al. 2008, 2009) and nuclear gene marker (e.g. Betancur et al. 2013) studies. The monophilies of the order Atheriniformes, superorder Cyprinodontea, and series Atherinomorpha are well supported. The monophilies of the two suborders of Atheriniformes, Atherinopsoidei and Atherinoidei were supported with \textit{A. microstoma} included in the latter suborder.

The result presented here will be useful for future phylogeography and population genetics studies of this Australian endemic species. Meanwhile, the addition of the full mitogenome sequence data of this species will be useful for future molecular phylogenetics studies of the Atheriniformes, Atherinomorpha, and Percomorpha in general.
The rapid bootstrap analyses were conducted with 1000 replications, with four threads running in parallel. A data matrix (11,082 positions) including all the concatenated nucleotide sequences of the mitogenomes except the third codon positions. The ND6 gene was partitioned ML analyses were performed with RAxML-GUI ver. 1-5b1 (Silvestro & Michalak 2012), with the GTR + I + ι nucleotide substitution model (Yang 1994). Aligned sequences were individually edited using the online version of GBlocks using the least stringent settings (http://molevol.es/MATLAB). Gene sequences were aligned individually using the online version of MAFFT under default settings (http://mafft.cbrc.jp/alignment/server/).

Acknowledgements
The author would like to thank Mark McGrouther (Australian Museum, Sydney) for arranging sample loan and tissue sample gift, and to Nagisa Nakaji and Satoko Nakashima (National Institute of Technology, Wakayama College) for their assistance on data-set building for the phylogenetic analyses in this study. The manuscript has been checked by a professional science and engineering English language, native-speaking editor.

Disclosure statement
The author declares that there is no conflict of interest. The author alone is responsible for the content and writing of the paper.

Funding
This study was partially supported by the National Institute of Technology, Wakayama College Competitive Internal Research Grant for Education 2015.
Setiamarga DH, Miya M, Yamanoue Y, Mabuchi K, Satoh TP, Inoue JG, Nishida M. 2008. Interrelationships of Atherinomorpha (medakas, flyingfishes, killifishes, silversides, and their relatives): The first evidence based on whole mitogenome sequences. Mol Phylogenet Evol. 49:598–605.

Setiamarga DH, Miya M, Yamanoue Y, Azuma Y, Inoue JG, Ishiguro NB, Mabuchi K, Nishida M. 2009. Divergence time of the two regional medaka populations in Japan as a new time scale for comparative genomics of vertebrates. Biol Lett. 5:812–816.

Silvestro D, Michalak I. 2012. raxmlgui: a graphical front-end for RAxML. Org Divers Evol. 12:335.

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

Yang Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J Mol Evol. 39:306–314.