

The Complete Mitogenome of Longman's Beaked Whale (*Indopacetus pacificus*) (Chordata: Ziphiidae)

Chiou-Ju Yao^{1,2}, Gilbert Audira³, Pei-Jung Wu¹, Chung-Der Hsiao^{3,4,5*}

¹ Department of Biology, National Museum of Natural Science, Taichung, Taiwan

² Department of Life Science, National Chung-Hsing University, Taichung, Taiwan

³ Department of Bioscience Technology, Chung Yuan Christian University, Chungli, Taoyuan, Taiwan

⁴ Center for Biomedical Technology, Chung Yuan Christian University, Chungli, Taoyuan, Taiwan

⁵ Center for Nanotechnology, Chung Yuan Christian University, Chungli, Taoyuan, Taiwan

(Received August 18, 2019; Accepted December 18, 2018; Published online December 28, 2019)

DOI : 10.6693/CAR201803_31(1).0007

Abstract. In this study, we used the next-generation sequencing method to deduce the complete mitogenome of Longman's beaked whale (*Indopacetus pacificus*) for the first time. The nucleotide composition was asymmetric (33.2% A, 27.0% C, 12.8% G, and 27.1% T), with an overall GC content of 39.9%. The assembled mitogenome consisted of 16,341 base pairs (bp) and followed the typical vertebrate arrangement, including 13 protein coding genes, 22 transfer RNAs, two ribosomal RNA genes, and a non-coding control region of the D-loop. The D-loop contains 873 bp and is located between tRNA-Pro and tRNA-Phe. The maximum likelihood phylogenetic tree of 10 cetacean mitogenomes showed that *I. pacificus* is located in the clade with *Hyperoodon ampullatus* and *Berardius bairdii*. The complete mitogenome of the Indo-Pacific beaked whale deduced in this study provides essential and important DNA molecular data for further phylogenetic and evolutionary analyses of cetaceans.

Key words: Ziphiidae, Longman's beaked whale, mitogenome, next generation sequencing.

INTRODUCTION

Longman's beaked whale (*Indopacetus pacificus*) is one of the world's most poorly known whales and is considered one of the rarest of all cetaceans (West *et al.* 2013). It has been known for more than 30 years as a large, unidentified species of beaked whale that inhabits tropical waters of the Pacific and Indian Oceans. It belongs to the family Ziphiidae and is among the least known cetaceans. The genus *Indopacetus* includes only one species, *I. pacificus*, with no

known subspecies.

Indopacetus pacificus, also known as the Indo-Pacific beaked whale or tropical bottlenose whale, is restricted to warmer waters of the Pacific and Indian Oceans. It can be found in deep ocean waters, and the color of this whale is poorly known but it is probably of a dark-bluish or black hue. This species has an estimated length of 5.0~5.5 m, a moderately long beak, a relatively flat head, and a distinctive dorsal fin that is low, triangular, and wide-based (Nowak 2003). It has two forward-pointing teeth present at the

*Corresponding author. E-mail: Longman's Beaked Whale Mitogenome

very front of the lower jawbone; each of which is located at the tip of its respective mandible. Only mature males have teeth that erupt (Pitman *et al.* 1999) which indicates sexual dimorphism in this species. Until recently, it was one of the least known species among living toothed whales, and only cryptic information was available on its field characteristics (Afsal *et al.* 2009). Moreover, evidence revealed that fishery interactions and loud anthropogenic sounds may threaten Longman's beaked whale. Even low levels of bycatch might cause unsustainable impacts to this naturally rare cetacean (Taylor *et al.* 2008). It is listed in Appendix II of CITES, and research is needed to determine the impacts of potential threats to this species.

Beaked whales in the Ziphiidae are generally poorly known and have some unresolved taxonomic issues (Burbidge *et al.* 2014). Not much information is known about intraspecific variations in morphology, and many species are very similar in external appearance (Dalebout 2002). Although *I. pacificus* is morphologically distinct from other ziphiid species, significant potential still exists for the misidentification of these animals, both at sea and when stranded. These problems occur due in part to its rarity, remote distribution, and similarity in external form to *Hyperoodon planifrons*, a southern bottlenose whale, as well the overall subtlety of diagnostic features in ziphiids (Dalebout *et al.* 2003). Since very little information is available concerning *I. pacificus*, it is difficult to describe its appearance (Tinker 1988). The aim of this paper was to decode the complete mitogenome of *I. pacificus*, so that a clear molecular structure of this species can be presented.

MATERIALS AND METHODS

Muscle tissue from a dead stranded *I. pacificus* (voucher no. NMNS011169) was collected in Ilan County, northeastern Taiwan on 23 July 2005. The methods for genomic DNA extraction, library construction, and next-generation sequencing (NGS) followed those described in previous publication (Shen *et al.* 2014). Using commercial software (Geneious V9, Auckland, New Zealand), about 0.004% (912 of 21,026,460) raw NGS reads generated by MiSeq (Illumina, San Diego, CA, USA.) were *de novo* assembled to produce a circular form of the complete mitogenome with

about an average of 14 coverage. To validate the phylogenetic position of *I. pacificus*, we used PHYML 3.0 (Guindon *et al.* 2010) software to construct a maximum-likelihood tree containing complete mitogenomes of seven species derived from four different genera in the Ziphiidae. The pygmy sperm whale (*Kogia breviceps*) (family Kogiidae) and sperm whale (*Physeter macrocephalus*) (family Physetrida) were used as outgroups for tree rooting. Substitution model selection for all mitogenomes was performed using PHYML as implemented in Smart Model Selection (SMS) in PhyML (Lefort *et al.* 2017). The best-fitting model according to the Bayesian information criterion (BIC) was used for the phylogenetic analyses.

DESCRIPTION AND REMARKS

The complete mitogenome of *I. pacificus* consists of 16,341 base pairs (bp) (GenBank accession no.: KY364702). The nucleotide composition was asymmetric (33.2% A, 27.0% C, 12.8% G, and 27.1% T) with an overall GC content of 39.9%. The D-loop contained 873 bp and was located between transfer (t)RNA-Pro and tRNA-Phe.

The protein coding, and ribosomal (r)RNA and tRNA genes of the mitogenome were predicted using the DOGMA (Wyman *et al.* 2004), ARWEN (Laslett and Canback 2008), and MITOS (Bernt *et al.* 2013) tools. Some ambiguous annotations were manually inspected by BLAST against closely related species. Four of 13 protein-coding genes (PCGs) terminated with incomplete stop codons of T- - (*COXI* and *ND4*). Other PCGs terminated with either TAA or TAG. The longest one was the *ND5* gene (1821 bp) among all PCGs, whereas the shortest was the *ATP8* gene (198 bp). The *12S* rRNA (1046 bp) and *16S* rRNA (2692 bp) genes were located between tRNA-Phe and tRNA-Leu (UAA) and were separated by tRNA-Val. Figure 1 illustrates the schematic map of the 16,341-bp mtDNA of *I. pacificus*.

Table 1 shows length distributions of 13 PCGs, two rRNAs, and the D-loop region among ten cetacean mitogenomes examined in this study. A length variation was not shown in the *cytochrome B*, *ND6*, *ND5*, *ND4*, *ND4L*, *ATP6*, or *COI* genes. The most variable DNA length was among the D-Loop regions and ranged

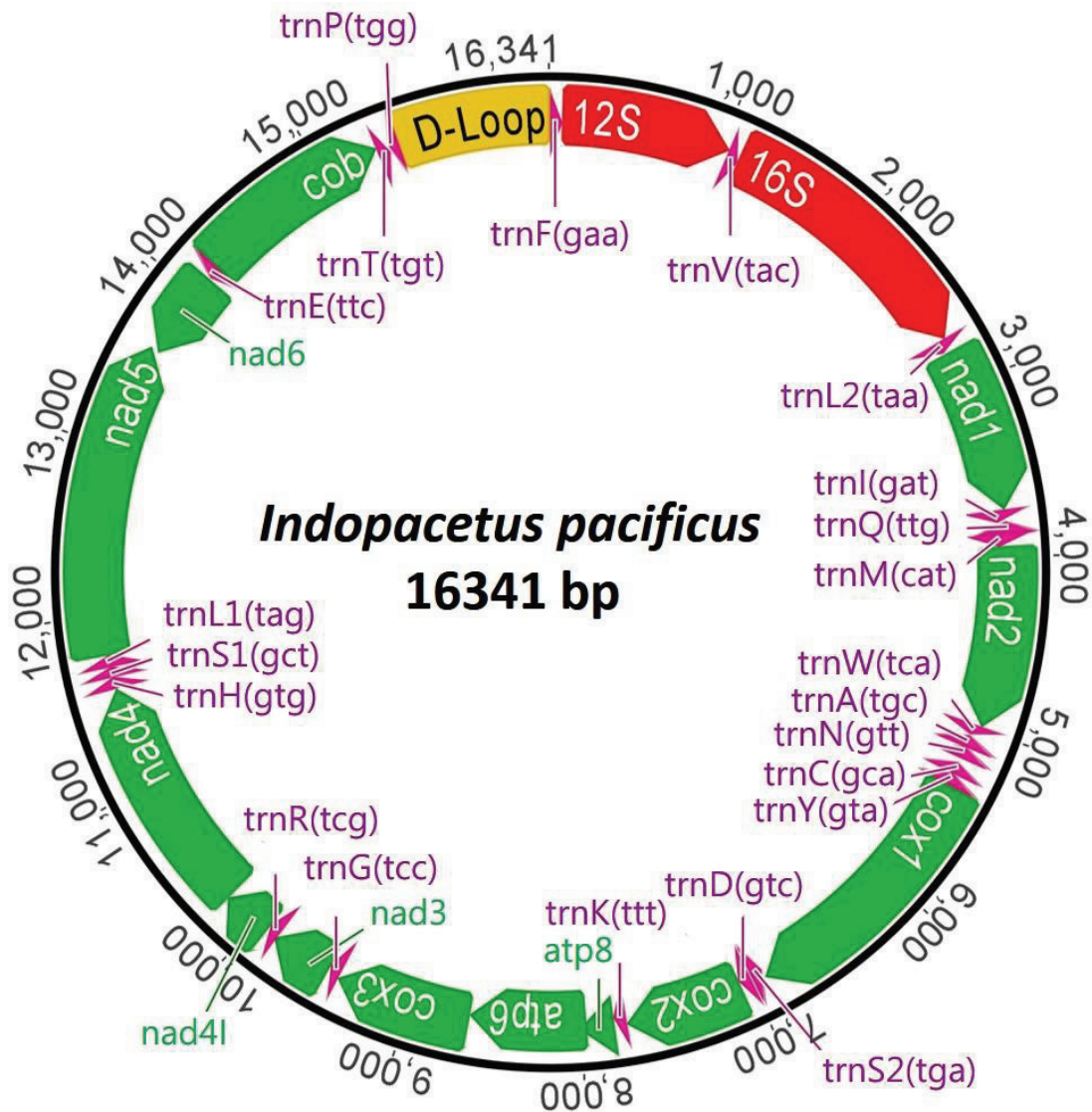


Fig. 1. Schematic map of the 16,341-bp mtDNA of *Indopacetus pacificus*. Green-colored sections represent 13 protein-coding genes: seven subunits of complex I (nd); one subunit of complex III (cob); three subunits of cytochrome c oxidase (cox); and two subunits of ATP synthase (atp6 and atp8). 12S and 16S ribosomal RNA are in red-colored sections; 22 transfer RNAs are identified at their positions; and the non-coding control region is in the brown-colored section.

from 870 bp of *M. ginkgodens* to 954 bp of *P. macrosaphalus*. In addition, DNA sequence divergences were calculated as p-distances (Nei and Kumar 2000) of the 13 PCGs, two rRNAs, and D-loops among the ten mitogenomes examined. Table 2 shows the overall mean p-distance and standard error (SE) of each gene or DNA fragment. Among them, 12S rRNA and 16S rRNA were the two least variable genes, and the D-Loop region was the most divergent fragment of the ten cetacean mitogenomes.

The maximum-likelihood phylogenetic tree (Fig. 2) shows that *I. pacificus* is located in the clade with *Hyperoodon* and *Mesoplodon*. It is not a sister species with *Mesoplodon* (Dalebout *et al.* 2004) nor *Ziphius* (Morin *et al.* 2017), a genus that was previously thought to be closely related to *I. pacificus*. Our results shed some light on relationships among the four genera of the Ziphiidae, and further research is needed to clarify ziphiid phylogeny. In conclusion, the complete mitogenome of *I. pacificus* was decoded

Table 1 DNA length distribution of 13 PCGs, 2 rRNAs, D-Loop and all tRNAs among 10 cetacean mitogenomes.

Species	GenBank Accession Number	Protein Coding Genes													rRNA		D-Loop	all tRNAs	mitgenome length
		C ^{ITB}	ND6	ND5	ND4	ND4L	ND3	COIII	ATP6	ATP8	COII	COI	ND2	ND1	12S	16S			
Bb	AJ554057	1140	528	1821	1378	297	346	786	681	201	684	1551	1044	957	970	1578	880	1504	16346
Ha	AJ554056	1140	528	1821	1378	297	347	786	681	198	684	1551	1044	957	968	1584	878	1504	16346
Ip	KY364702	1140	525	1821	1378	297	347	785	681	198	684	1546	1042	951	972	1579	873	1522	16341
Md	KF032860	1140	528	1821	1378	297	347	785	681	198	684	1551	1042	957	971	1582	877	1508	16347
Me	KC776688	1140	528	1821	1378	297	347	785	681	198	684	1551	1042	960	971	1580	876	1504	16343
Mg	KR534596	1140	528	1821	1378	297	347	785	681	198	684	1551	1042	957	972	1579	870	1509	16339
Mgr	KF981442	1140	528	1821	1378	297	347	785	681	198	684	1551	1044	957	970	1582	880	1504	16347
Zc	KC776696	1140	528	1821	1378	297	347	785	681	201	684	1551	1042	957	972	1580	878	1507	16349
Pm	AJ277029	1140	528	1821	1378	297	346	786	681	192	684	1551	1042	955	970	1588	954	1515	16428
Kb	AJ554055	1140	528	1821	1378	297	346	786	681	201	684	1551	1044	957	972	1578	939	1503	16406

The species abbreviations are listed as follows: Ip - *Indopacetus pacificus*, Me - *Mesoplodon europaeus*, Mg - *Mesoplodon ginkgodens*, Md - *Mesoplodon densirostris*, Mgr - *Mesoplodon grayi*, Ha - *Hyperoodon ampullatus*, Zc - *Ziphius cavirostris*, Bb - *Berardius bairdii*, Pm - *Physeter macrocephalus* and Kb - *Kogia breviceps*

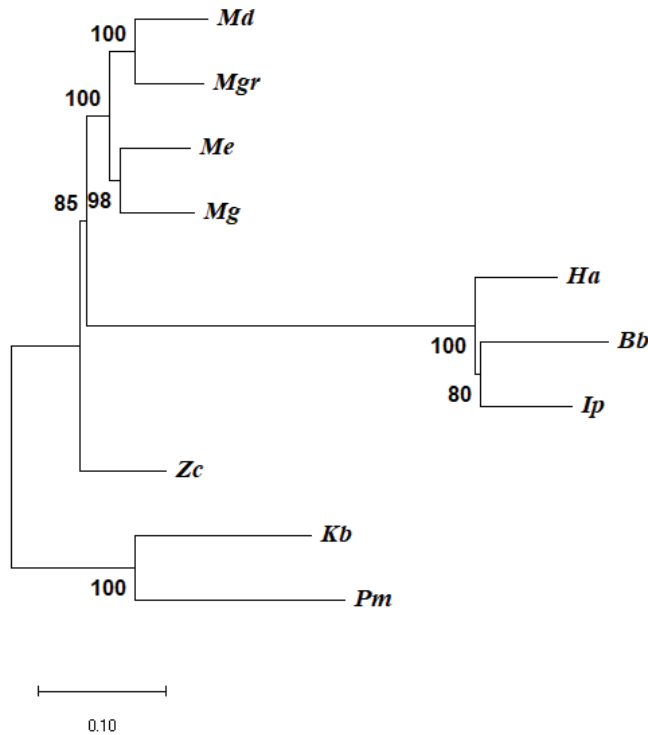


Fig. 2. Molecular phylogeny of *Indopacetus pacificus* and other related species in the Ziphiidae based on the complete mitogenome. The complete mitogenomes of additional species were downloaded from GenBank, and the phylogenetic tree was constructed by the maximum-likelihood method with the best-fitting model GTR+G and the computed approximate likelihood ratio test (SH-like branch supports). GenBank accession numbers for tree construction are listed as follows: *Ip*, *Indopacetus pacificus* (KY364702); *Me*, *Mesoplodon europaeus* (KC776688); *Mg*, *Mesoplodon ginkgodens* (KR534596); *Md*, *Mesoplodon densirostris* (KF032860); *Mgr*, *Mesoplodon grayi* (KF981442); *Bb*, *Berardius bairdii* (AJ554057); *Pm*, *Physeter macrocephalus* (AJ277029); and *Kb*, *Kogia breviceps* (AJ554055).

for the first time in this study and provides essential and important DNA molecular data for further phylogenetic and evolutionary analyses of cetaceans.

Declaration of interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

ACKNOWLEDGEMENTS

This study was supported by a grant (107-FM-1.1-C-6(5)) from the Council of

Table 2 Overall mean p-distance and standard error of 13 PCGs, 2 rRNAs and D-Loop region among 10 cetacean mitogenomes.

DNA fragment	mean p-distance	S.E.
12S	0.0879	0.0052
16S	0.0777	0.0040
ND1	0.1269	0.0049
ND2	0.1442	0.0066
CO1	0.1334	0.0052
CO2	0.1367	0.0072
ATP8	0.1794	0.0159
ATP6	0.1575	0.0081
CO3	0.1367	0.0072
ND3	0.1490	0.0108
ND4L	0.1720	0.0122
ND4	0.1456	0.0054
ND5	0.1560	0.0050
ND6	0.1499	0.0089
CTB	0.1425	0.0053
DL	0.1809	0.0065

Agriculture, Taiwan. We thank the Taiwan Cetacean Society for assistance in transporting the specimen and conducting the necropsy.

REFERENCES

- Afsal, V., P. Manojkumar, K. Yousuf, B. Anoop, and E. Vivekanandan. 2009. The first sighting of Longman's beaked whale, *Indopacetus pacificus* in the southern Bay of Bengal. *Mar. Biodivers. Rec.* 2, e133. doi:10.1017/S1755267209990510.
- Bernt, M., A. Donath, F. Juhling, F. Externbrink, C. Florentz, G. Fritzsch, [J. Putz, M. Middendorf, and P. F. Stadler/et al?]. 2013. MITOS: improved de novo metazoan mitochondrial genome annotation. *Mol. Phylogenet. Evol.* 69: 313-319.
- Burbidge, A., J. Woinarski, and P. Harrison. 2014. The action plan for Australian mammals 201. CSIRO Publishing, [city?], Australia.
- Dalebout, M.L. 2002. Species identity, genetic diversity, and molecular systematic relationships among the Ziphiidae (beaked whales). PhD thesis. University of Auckland, Auckland, New Zealand.
- Dalebout, M.L., C.S. Baker, R.C. Anderson, P.B. Best, V.G. Cockcroft, H.L. Hinsz, [V.

- Peddemors, and R.L. Pitman/et al?]. 2003. Appearance, distribution, and genetic distinctiveness of Longman's beaked whale, *Indopacetus pacificus*. Mar. Mamm. Sci. 19: 421-461.
- Dalebout, M.L., C.S. Baker, J.G. Mead, V.G. Cockcroft, and T.K. Yamada. 2004. A comprehensive and validated molecular taxonomy of beaked whales, family Ziphiidae. J. Hered. 95: 459-473.
- Guindon S., J.F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, and O. Gascuel. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Systemat. Biol. 59(3): 307-321.
- Laslett, D., and B. Canback. 2008. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. Bioinformatics 24: 172-175.
- Lefort, V., J.-E. Longueville, and O. Gascuel. 2017. SMS: Smart Model Selection in PhyML. Mol. Biol. Evolution 34: 2422-2424.
- Morin P.A., C.S. Baker, R.S. Brewer, A.M. Burdin, M.L. Dalebout, J.P. Dines, [I. Fedutin, O. Filatova, E. Hoyt, J.L. Jung, M. Lauf, C.W. Potter, G. Richard, M. Ridgway, K.M. Robertson, and P.R. Wade/et al?]. 2017. Genetic structure of the beaked whale genus *Berardius* in the North Pacific, with genetic evidence for a new species. Mar. Mamm. Sci. 33: 96-111.
- Nei X., and X. Kumar. 2000. Molecular evolution and phylogenetics. Oxford University Press, New York. 33 pp.[please give author initials]
- Nowak, R.M. 2003. Walker's marine mammals of the world. Johns Hopkins University Press, city?, ST?.
- Pitman, R.L., D.M. Palacios, P.L. Brennan, B.J. Brennan, K.C. Balcomb, and T. Miyashita. 1999. Sightings and possible identity of a bottlenose whale in the tropical Indo-Pacific: *Indopacetus pacificus*? Mar. Mamm. Sci. 15: 531-549.
- Shen, K.N., T.C. Yen, C.H. Chen, H.Y. Li, P.L. Chen, and C.D. Hsiao. 2014. Next generation sequencing yields the complete mitochondrial genome of the flathead mullet, *Mugil cephalus* cryptic species NWP2 (Teleostei: Mugilidae). Mitochondrial DNA Part A 27: 1758-1759.
- Taylor, B.L., R. Baird, J. Barlow, S.M. Dawson, J. Ford, J.G. Mead, [G. Notarbartolo di Sciara, P. Wade, and R.L. Pitman/et al?]. 2008. *Indopacetus pacificus*. The IUCN Red List of Threatened Species 2008: e.T40635A10345818.
- Tinker, S.W. 1988. Whales of the world. Brill Publishing Company, New York.
- West, K.L., S. Sanchez, D. Rotstein, K.M. Robertson, S. Dennison, G. Levine, [N. Davis, D. Schofield, C.W. Potter, and B. Jensen/et al?]. 2013. A Longman's beaked whale (*Indopacetus pacificus*) strands in Maui, Hawaii, with first case of Morbillivirus in the central Pacific. Mar. Mamm. Sci. 29: 767-776.
- Wyman, S.K., R.K. Jansen, and J.L. Boore. 2004. Automatic annotation of organellar genomes with DOGMA. Bioinformatics 20: 3252-3255.

朗氏喙鯨(*Indopacetus pacificus*) (Chordata: Ziphiidae)的完整粒線體基因體組成

姚秋如^{1,2} Gilbert Audira³ 吳佩蓉¹ 蕭崇德^{3,4,5*}

¹國立自然科學博物館生物學組 404台中市北區館前路一號

²國立中興大學生命科學系 402 台中市南區興大路145號

³中原大學生物科技系 台灣320桃園市中壢區中北路200號

⁴中原大學生醫技術研發中心 台灣320桃園市中壢區中北路200號

⁵中原大學奈米科技中心 台灣320桃園市中壢區中北路200號

本研究利用次世代定序技術首次成功獲得朗氏喙鯨(*Indopacetus pacificus*)的完整粒線體DNA序列。本種粒線體基因體的核 酸組成爲非對稱性，其組成爲33.2% A, 27.0% C, 12.8% G, 及27.1% T，整體的GC組成是39.9%。經組裝後的粒線體基因體共有 16,341鹼基對。如同典型的脊椎動物一樣，本種的粒線體DNA環狀結構的基因排列組合包含13個蛋白質編碼基因(protein coding genes), 22個轉移核糖核酸(transfer RNAs)基因, 2個核糖體核糖核酸(ribosomal RNAs)基因及一個非編碼區域的控制區(non-coding control region of the D-loop，以下簡稱D-loop)。D-loop共有873 鹼基對，座落於 tRNA-Pro與 tRNA-Phe之間。以最大似然分析分法建構10種齒鯨之粒線體基因體分子親緣樹發現，朗氏喙鯨與北瓶鼻鯨(*Hyperoodon ampullatus*)和貝氏喙鯨(*Berardius bairdii*)被歸類於同一個支系中。本研究完整組裝朗氏喙鯨的粒線體基因體，提供重要的DNA分子訊息可做爲鯨類的親緣關係及演化分析的素材。

關鍵詞：喙鯨科，朗氏喙鯨，粒線體基因體，次世代定序