# Secrets of the Deep:
# The Molecular Genetics of Cryptic Beaked Whales

Submitted by Kirsten Freja Thompson, to the University of Exeter as a thesis for the degree of PhD by Publication in Biological Sciences,
January 2017.

# Abstract

Beaked whales are comparatively unknown social mammals due to their deep-ocean distribution and elusive habits. The deep-ocean is the largest biome on Earth and the final frontier for human expansion. Since their first discovery, beaked whales have remained largely hidden from science. In this era of rapid technological advancement, genetic and genomic methods are key tools for population biologists and are particularly useful in describing rarely seen species.

Using DNA-barcoding and nuclear markers, the publications in this thesis provide data on the distribution and external appearance of two species of beaked whale: the spade-toothed (*Mesoplodon traversii*) and Derinayagala's whale (*Mesoplodon hotaula*). These whales were previously known from only a handful of tissue and bone specimens.

Long-term efforts have facilitated the collection of samples of Gray's beaked whale (*Mesoplodon grayi*) and we have used shot-gun sequencing to characterise the mitochondrial genome and isolate species-specific nuclear microsatellite loci. Using genetic species and sex identification, together with museum specimens and multivariate analyses, we provide clear evidence of sexual dimorphism in cranial dimensions and geographic variation in external morphology.

No genetic differentiation was evident in Gray's beaked whales across a large study area (~ 6,000 km). With a large female effective population size ($N_e$) and genetic homogeneity, we hypothesise that gene flow is facilitated by large-scale oceanographic features, such as the sub-tropical convergence. Genetic kinship analyses within Gray's beaked whale groups suggest that the whales that strand together are not related. Both sexes disperse from their parents and these groups are not formed through the retention of kin. These results are consistent with a 'fission-fusion' social system that has been

observed in some oceanic dolphin species. Taken together, these data provide the first insights into the population dynamics, dispersal and social organisation in Gray's beaked whales. These publications highlight the value of using genetics alongside other techniques to describe inter- and intraspecific diversity. For beaked whales, the dead can tell us much about the living.

# Acknowledgements

All the publications included in this thesis are part of collaborations with a number of talented scientists in various parts of the world. I would like to thank all the co-authors that have enriched my understanding of science. In particularly I thank Craig Millar, Rochelle Constantine, Selina Patel, Scott Baker, Anna Santure, Kathy Ruggiero, Debbie Steel, Merel Dalebout, Anton van Helden and Judith Robbins.

At the University of Exeter, I would especially like to thank my primary supervisor Jamie Stevens and all the members of the Molecular Ecology and Evolution Group. Thanks to Josie Paris for reading my thesis, Emma Carroll and Stefan Bräger for comments and the many anonymous reviewers and editors who greatly improved the initial manuscripts. Much of the writing would not have been possible without additional support of David Santillo and Paul Johnston. This research was funded by several institutions and organisations that are credited in the respective acknowledgements within the publications in the Appendices.

Lastly, but by no mean least, thanks to my family – Paul, Rumi, Felix, Merlin, Moss, Sula and Tui – for all the support over the years that I have been toiling and writing. I have appreciated every moment of your laughs and love.

# Contents

# List of Tables

# List of Figures

# Appendices

The following publications are listed in chronological order, and follow a progressive sequence of research between 2012 and 2017:

I. **Thompson K**, Baker S, van Helden A, Patel S, Millar C, Constantine R. (2012). The world's rarest whale. Current Biology 22: 905–906.

II. Baker CS, Hutt A, **Thompson K**, Dalebout ML, Robins J, Brownell RL, Stone GS. (2013). Species identity and human consumption of beaked whales in the Gilbert Islands, Republic of Kiribati. Animal Conservation 16: 641- 647.

III. Dalebout ML, Baker CS, Steel D, **Thompson KF**, Robertson K, Chivers SJ, Perrin WF, Goonatilake M, Anderson RC, Mead JG, Potter CW, Yamada TK, Thompson TK, Jupiter D. (2013). Resurrection of *Mesoplodon hotaula* Deraniyagala 1963: A new species of beaked whale in the tropical Indo-Pacific. Marine Mammal Science 30: 1081-1108.

IV. **Thompson KF**, Patel S, Williams L, Tsai P, Constantine R, Millar CD. (2016). High coverage of the complete mitochondrial genome of the rare Gray's beaked whale (*Mesoplodon grayi*) using Illumina next generation sequencing. Mitochondrial DNA Part A 27. Early online January 2014.

V. Patel S, **Thompson KF**, Williams L, Tsai P, Constantine R, Millar CD. (2014). Mining microsatellites for Gray's beaked whale from next generation sequencing data. Conservation Genetics Resources 6: 657-659.

VI. **Thompson KF**, Ruggiero, K, Millar CD, Constantine R, van Helden A. (2014). Large-scale multivariate analysis reveals sexual dimorphism and geographic differences in the Gray's beaked whale. Journal of Zoology 294: 13-21.

VII. **Thompson KF***, Patel S*, Baker CS, Constantine R, Millar CD. (2016). Bucking the trend: genetic analysis reveals high diversity and little differentiation in a deep ocean cetacean. *Heredity* 116, 277–285.

VIII. Patel S*, **Thompson KF***, Santure A, Constantine R, Millar CD. Genetic kinship analyses reveal that Gray's beaked whale strands in unrelated groups. Final acceptance *Journal of Heredity* Early online 9th March 2017. doi: 10.1093/jhered/esx021

* Denotes joint first authorship.

These publications are referred to as I – VIII throughout the thesis.

# Author's Declaration

| | Publication | KFT contribution |
|---|---|---|
| I. | **Thompson, K**, Baker, CS, van Helden, A, Patel, S, Millar, C, Constantine, R. (2012). The world's rarest whale. *Current Biology* 22: R905-R906. | Planning, laboratory analysis, editing and writing manuscript. |
| II. | Baker CS, Hutt A, **Thompson K**, Dalebout ML, Robins J, Brownell RL, Stone GS. (2013). Species identity and human consumption of beaked whales in the Gilbert Islands, Republic of Kiribati. *Animal Conservation* 16: 641- 647. | Laboratory analysis, contributing and editing manuscript. |
| III. | Dalebout ML, Baker CS, Steel D, **Thompson KF**, Robertson K, Chivers SJ, Perrin WF, Goonatilake M, Anderson RC, Mead JG, Potter CW, Yamada TK, Thompson TK, Jupiter D. (2013). Resurrection of *Mesoplodon hotaula* Deraniyagala 1963: A new species of beaked whale in the tropical Indo-Pacific. *Marine Mammal Science* 30: 1081-1108. | Minor laboratory analysis and editing manuscript. |

| Publication | KFT contribution |
|---|---|
| IV. **Thompson KF**, Patel S, Williams L, Tsai P, Constantine R, Millar CD. (2016). High coverage of the complete mitochondrial genome of the rare Gray's beaked whale (*Mesoplodon grayi*) using Illumina next generation sequencing. *Mitochondrial DNA* Part A 27. Early online January 2014. | Planning, study design, annotated mitochondrial genome, wrote draft and edited manuscript. |
| V. Patel S, **Thompson KF**, Williams L, Tsai P, Constantine R, Millar CD. (2014). Mining microsatellites for Gray's beaked whale from next generation sequencing data. Conservation Genetics Resources 6: 657-659. | Partial laboratory and other analyses, commented on manuscript. |
| VI. **Thompson KF**, Ruggiero, K, Millar CD, Constantine R, van Helden A. (2014). Large-scale multivariate analysis reveals sexual dimorphism and geographic differences in the Gray's beaked whale. Journal of Zoology 294: 13-21. | Planning, study design, all analysis (with assistance from KR), wrote first draft and edited manuscript. |

| Publication | KFT contribution |
|---|---|
| VII. **Thompson* KF**, Patel S*****, Baker CS, Constantine R, Millar CD. (2016). Bucking the trend: Genetic analysis reveals high diversity, large population size and low differentiation in a deep ocean cetacean. *Heredity* 116: 287–115. *****Myself and Selina Patel are joint first authors. | Planning and study design, laboratory work for all mtDNA data, partial for msat data. All analysis in combination with Selina (who called genotypes), estimates of $N_e$ and $F_{st}$ statistics, wrote first draft and edited manuscript. |
| VIII. Patel S******, **Thompson KF****, Santure A, Constantine R, Millar CD. (2017). Genetic kinship analyses reveal that Gray's beaked whale strands in unrelated groups. *Journal of Heredity* Early online 9th March 2017. ****** Myself and Selina Patel are joint first authors. | Planning, study design, analysis and interpretation of results in combination with Anna Santure. Selina Patel undertook genotyping. I wrote, edited and revised manuscript. |

# Abbreviations

| | |
|---|---|
| ABC | Approximate Bayesian Computation |
| AMOVA | Analysis of molecular variance |
| DAPC | Discriminant analysis of principal components |
| EEZ | Exclusive economic zone |
| GCSC | Geneological Concordance Species Concept |
| ILS | Incomplete lineage sorting |
| mtDNA | Mitochondrial DNA |
| NZCeTA | New Zealand Cetacean Tissue Archive |
| PSMC | Pairwise sequential Markovian coalescent |
| SFS | Site-frequency spectrum |
| SMM | Stepwise mutation model |
| SNP | Single nucleotide polymorphism |

# Chapter 1

## Introduction

The deep-sea is the largest biome on Earth. Deep-sea regions (>200 m depth) cover around 65% of the Earth's surface and 95% of the global biosphere (Smith et al., 2009; Danovaro et al., 2014). Knowledge of the diversity of species and habitats in this vast area is exceptionally poor and discovery rates for new species are high (Ramirez-Llodra et al. 2010). In the last 40 years, more than 1000 animal species have been described from hydrothermal vents and hydrocarbon seeps, even though the distribution of sampling is sparse and highly patchy (Vrijenhoek 2009). Most of the species described in such targeted and enigmatic deep-sea ecosystems are sessile (e.g. tubeworms) or only moderately motile (e.g. clams).

For highly mobile marine vertebrates, such as marine mammals and elasmobranches, obtaining meaningful sampling distributions with which to describe populations can be challenging. Collecting sufficient baseline life history and even taxonomic information on these elusive animals is also problematic. For example, cryptic species have been found within several elasmobranch groups: hound sharks (*Mustela* spp., Gardener & Ward 2002); scalloped hammerhead sharks (*Sphyrna* spp., Quattro et al. 2006); wobbegong (*Orectolobus* spp., Corrigan et al. 2008), common skates (*Dipterus* spp., Griffiths et al. 2010) and dogfish (Veríssimo et al. 2016). These recent taxonomic discoveries coincide with steep population declines in elasmobranches globally and

many species are now considered critically endangered primarily due to overfishing (Randhawa et al. 2015). Unfortunately, this may also be the case with many contemporary taxonomic discoveries (Lees et al. 2015; McCauley et al. 2015; McClenachan et al. 2015). The loss of such apex predators can induce broad community level changes that may be difficult to predict. For example, the removal of large predators may result in perturbation of top-down predatory control, and an overall reduction in omnivory (Stevens et al. 2000). Estes et al. (2011) described the global loss of apex predators and the impact this has on broad-scale ecological interactions in what is now routinely termed 'trophic downgrading'.

Many wide-ranging marine vertebrates have been described as rare, for example Omura's whale (*Balaenoptera omurai*) (Cypriano-Souza et al. 2017) or certain elasmobranch species (Smart et al. 2013; Kousteni et al. 2016). Whether these species are 'rare' or 'rarely seen' is often unknown. Rarity is one of the metrics used to assess extinction risk in species and attempts to quantify how unique a species or population is, though it does not necessarily include information on population trends. Historically, rarity has been broadly defined based on either low population size or small geographic range (Gaston 1994; Hartley & Kunin 2003). Some studies, for example Cerqueira et al. (2013), assess relative rarity of Amazonian bird species by integrating these two measures – population size and range – using occupancy modelling, where rarity is defined as low occupancy, i.e. a species is defined as rare if it has a low probability of occupying a site in comparison to another species. For the purpose of conservation, rarity is often defined using a similar synthesis of species information that includes population size, range, fragmentation or presence in a limited number of locations. These descriptors, in combination with information on whether a species is at risk of steep population decline, inform extinction risk assessments (Collen et al. 2016). For marine species that are wide ranging and rarely observed, rarity is often assumed, in

some cases incorrectly, based on a lack of records rather than any accurate assessment of abundance or understanding of population distribution or fragmentation. Generating data on abundance and distribution of species that are difficult to observe is costly. These data will also take considerable time to collect – decades – if we are reliant on opportunistic bycatch, sightings or beach-cast animals. Even with such a collection of data points, there will be inherent bias as a result of fishing, survey and reporting effort that can result in skewed distributions. In the case of air-breathing species, such as cetaceans, systematic line transect sighting surveys have difficulty in capturing data for deep divers due to 'unavailability bias' (Marsh & Sinclair 1989) and this can lead to and underestimates of abundance. Considerable research effort has been applied to statistically adjusting survey data to deal with such bias (for example see Conn et al. 2012; Okamura et al., 2012; Borchers et al. 2013). One way of overcoming these barriers is to integrate multiple lines of evidence – systematic sightings and acoustic surveys, mark-recapture methods, molecular analyses – and identify the strengths and weaknesses of each.

Toothed whales (odontocetes) are long-lived, apex predators with extremely low fecundity, complex social systems and extended periods of parental care. Many species that have undergone catastrophic population declines as a result of whaling are relatively well studied, e.g. sperm whales (*Physeter macrocephalus*) (Whitehead 2009). In some areas, these whales appear to be recovering from heavy exploitation but most populations remain far from pre-exploitation abundance levels (Whitehead 2002). However, contrary to what might be expected there are also several odontocete species for which only a handful of records exist. Most species have never been the target of whaling and there are few historical records of their populations. Many of these species are in the Ziphiidae family, which comprises all members of what are commonly known as beaked whales.

The beaked whales are a little-known and mysterious mammalian group; most species have been described during the last 150 years and some within the previous 15 years (e.g. Longman's beaked whale (*Indopacetus pacificus*), Dalebout et al. 2003). Only two species of beaked whale have been systematically hunted by whalers, Baird's beaked whale (*Berardius bairdii*) and northern bottlenose whales (*Hyperoodon ampullatus*). There are also sparse catch records for southern bottlenose whales (*Hyperoodon planifrons*). For species that have not been the target of hunting, there are few historical records to draw data from and no known locations suggesting where the whales predictably aggregate. Furthermore, data collected from direct observation of beaked whales are scant. These whales have highly oceanic distributions and elusive surface behaviours with long dive times and short rest periods at the surface. Field identification of beaked whales is also problematic, particularly for the *Mesoplodon* species that are relatively small (4 – 6 m) and have a similar overall appearance. In addition, visual detection rates of beaked whales during boat-based sighting surveys decline swiftly with increasing Beaufort Sea State and therefore, obtaining robust distribution data is difficult. However, even with these numerous barriers to field research, systematic aerial surveys for marine species have described numerous beaked whale sightings in offshore areas (V. Ridoux, Université de La Rochelle, pers. comm.). These offshore beaked whale sightings, though not identified to species level, can be more numerous even than those of dolphins (or other odontocetes) suggesting that beaked whales may be more common than previously thought and potentially an important part of deep ocean ecosystems.

For most beaked whale species descriptions, data on external appearance and distributions are derived from stranded specimens. Since at least the time of Aristotle, 2000 years ago, scientists have studied the species involved in cetacean stranding

events, as well as the patterns of these strandings (Thompson 1910). However, strandings of beaked whales are sporadic and relatively rare so it has, therefore, taken decades to build up sufficient knowledge of most species (Fig. 1A). Strandings can be in an advanced state of decomposition and, due to their unusual morphology, some have historically peaked the curiosity of the public and cryptozoologists (Figure 1B, 1C).

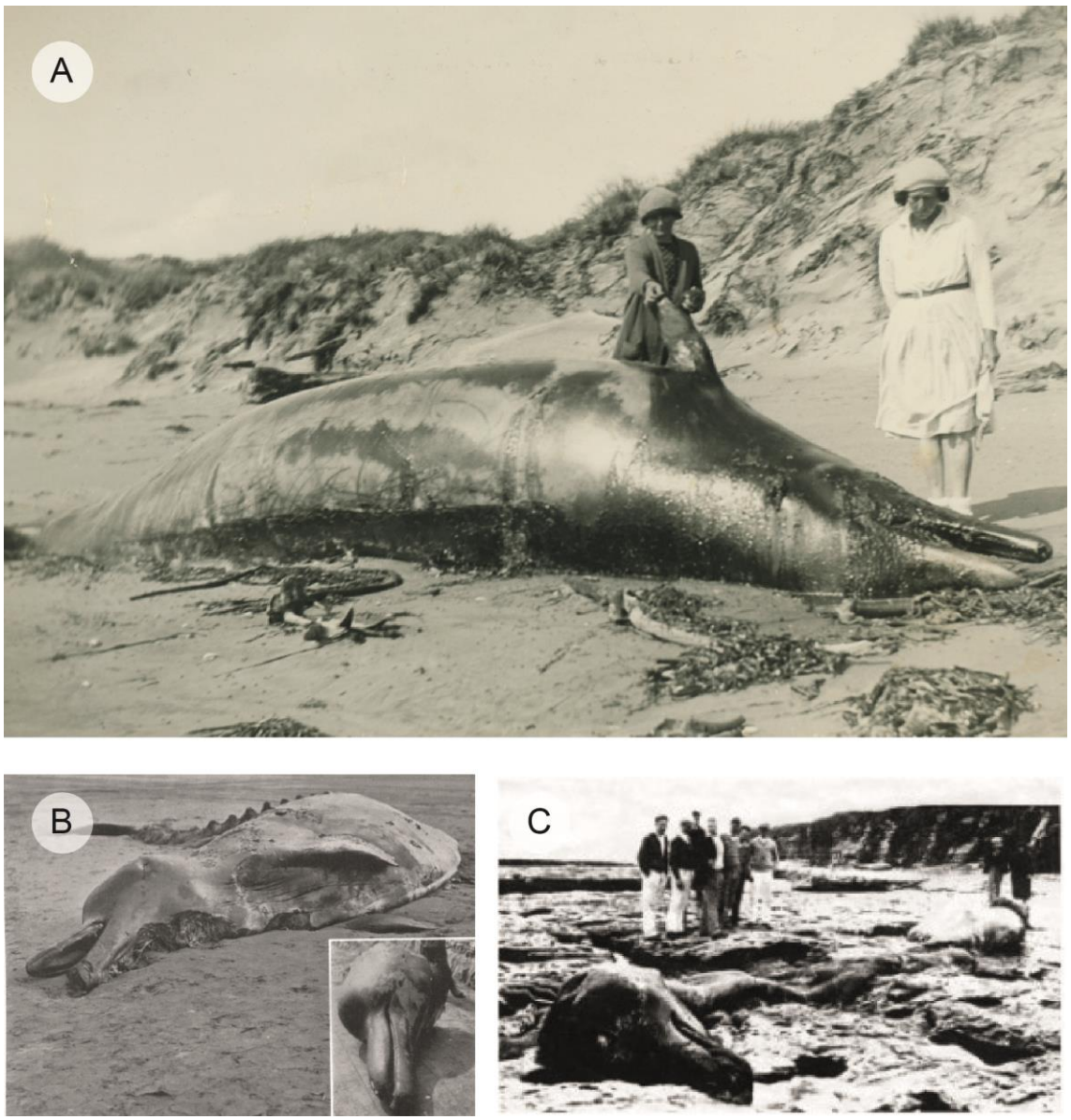



**Figure 1.** Initial beaked whale species descriptions were from stranded specimens that were found sporadically and often in a highly decomposed state, sometimes creating confusion as to their provenance. A) The first described specimen of the Shepherd's beaked whale (*Tasmacetus shepherdii*) found stranded in 1937 off the coast of New Zealand (Photo credit: Whanganui Regional Museum). B) Decomposing Baird's beaked whale (*Berardius bairdii*) from Ocean City, Washington in 1950. The insert is the head of one from a whaling station. (Photo credits: Slipp & Wilke 1953). C) Santa Cruz's 'duck billed elephant monster'. A particularly decomposed Baird's beaked whale found stranded on the Californian coast in 1925 which aroused much local curiosity (Photo credit: Special Collections, University of California at Santa Cruz).

The majority of beaked whale species have a similar external morphology, and even for those animals that are beach-cast or stranded, accurate species identification can be confusing, especially for females. Traditional species identification relied on the shape and position of mandibular "tusk' teeth, which are generally absent in females (Moore 1963). Ziphiids are thought to lack functional mandibular teeth, except for Shepherd's (*Tasmacetus shepherdii*) and Gray's beaked whale (*Mesoplodon grayi*). Some research suggests that, particularly within the *Mesoplodon* genus, the position of these tusk-teeth in males can provide information on the evolutionary relationships within the group and additionally implicates sexual selection as a driving force in their radiation (Fig. 2, Dalebout et al. 2008).

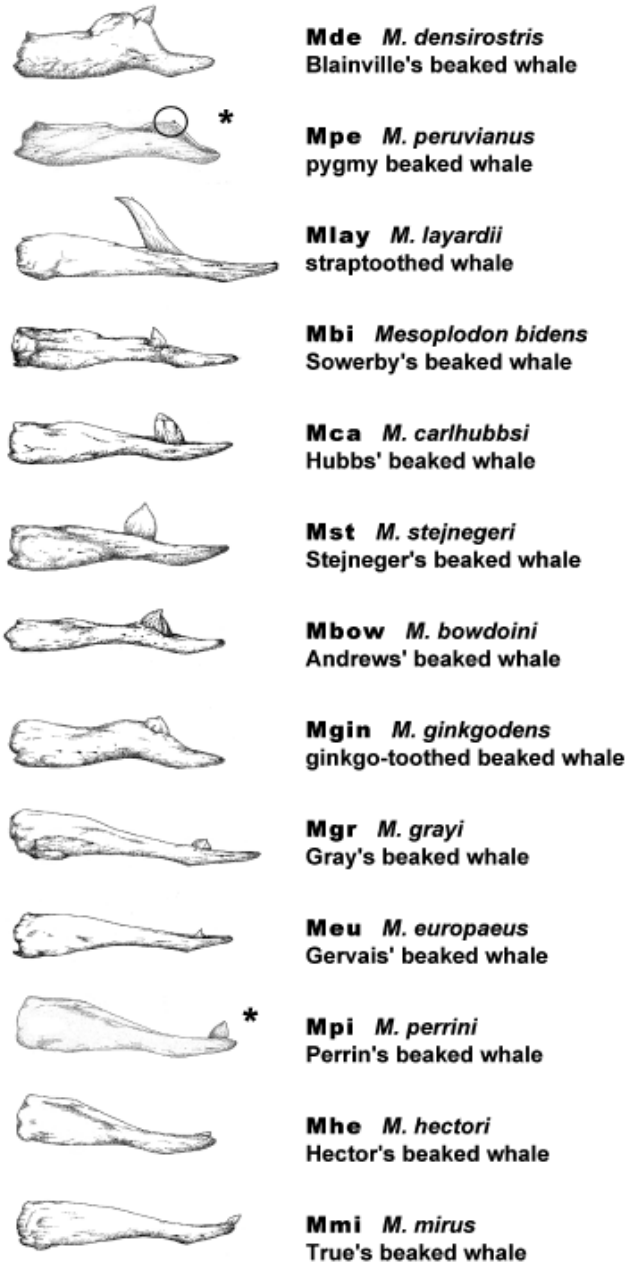



**Figure 2.** Adult male *Mesoplodon* beaked whale tusk morphology (lateral views of lower jaws), arranged in a linear progression from ancestral to derived, as perceived by Moore (1968) based on a combination of tusk position, size, and angle of inclination. (From: Dalebout et al. (2008)).

Since the year 1998, and the development of new DNA-sequencing technologies, there has been a particular motivation to develop molecular tools to identify and study beaked whales in terms of the evolutionary relationships within the group. With significant

knowledge gaps on their basic biology, these studies are also needed to inform conservation efforts. Several broad research recommendations have been suggested with the aim of addressing these questions (Dalebout et al. 2004; Barlow et al. 2006). Studies assessing population structure using genetic, morphological and tagging data were recommended in order to shed light on basic beaked whale biology and habitat requirements. In addition, estimates of abundance and global distributions are also lacking for almost all species. Tackling these knowledge gaps has been increasingly urgent in recent years given that beaked whales are vulnerable to the impacts of anthropogenic noise which can induce mass strandings (Frantzis 1998; Cox et al. 2006; Moore and Barlow 2013).

### 1.1 Beaked whales as cryptic deep ocean megafauna

The definitions of cryptic species in the literature are broad and vary widely. Some authors describe cryptic species as 'sibling species' that appear morphologically to be the same and imply a sister-species relationship with a shared common ancestry (Knowlton 1989). When distinctive features are found between these two species, this definition loses its utility (Knowlton 1993). Some authors also refer to 'cryptic species' if they are camouflaged or particularly secretive and difficult to observe (Claridge et al. 2004). In taxonomic terms, cryptic species are two or more species that are, or have been, classified as a single nominal species because they are at least superficially morphologically indistinguishable (Bickford et al. 2007).

The beaked whales are a speciose group that live in the vast areas of the deep ocean. In the context of this thesis, I have drawn from several of these definitions in describing beaked whales in that species, particularly within the genus *Mesoplodon*, are often morphologically indistinguishable if they are females or immature, and several species

have been classified as a single species due to insufficient data on morphology and distributions. Adding to this confusion, most species are difficult to observe and are cryptic in terms of their behaviour and by the inaccessibility of the habitat in which they live.

Beaked whale taxonomy is continually under review, and prior to the research conducted in this thesis, 21 species of beaked whale were recognised, present in two sub-families (Ziphiinae and Hyperoodontinae) and six genera (Table 1). Most species are rarely sighted, and in some cases these animals are known from only a handful of records, e.g. the spadetoothed whale (*Mesoplodon traversii*) described in publication I. Publications II and III give new information on the distribution of a species, Deraniyagala's whale (*Mesoplodon hotaula*), which was taxonomically resurrected as separate from its sympatric sister species, the ginkgotoothed whale (*Mesoplodon ginkgodens*). In addition, Morin et al. (2017) describe the presence of a new unnamed species of *Berardius* spp. in the North Pacific that is more divergent from Baird's beaked whale than the congeneric Arnoux's beaked whale (*Berardius arnuxii*) found in the Southern Ocean.

**Table 1.** Nomenclature of all (currently 23) recognised species within the Family Ziphiidae and their mean total body length and geographic distribution.

| Sub family | Species | Beaked whale | Mean length (m) | Distribution | Reference |
|---|---|---|---|---|---|
| Ziphiinae | *Berardius arnuxii* | Arnoux's | 9.2 | Temperate and polar waters of the Southern Hemisphere. | Kasuya 2009 |
| | *Berardius bairdii* | Baird's | 10.1 | Cold temperate waters of the North Pacific. | Kasuya 2009 |
| | *Berardius spp.?* | 'black' Baird's | < 7 | Cold temperate waters of the North Pacific | Morin et al. 2017<br>Kitamura et al. 2013 |
| | *Tasmacetus shepherdi* | Shepherd's | 6.5 | Limited records suggest temperate waters of the Southern Hemisphere. | Mead 2009b |
| Hyperoodontinae | *Hyperoodon ampullatus* | Northern bottlenose | 7.5 | Subpolar and temperate waters of the North Atlantic. | Gowans 2009 |
| | *Hyperoodon planifrons* | Southern bottlenose | 7.5 | Temperate waters of the Southern Hemisphere. | Gowans 2009 |

**Table 1 continued.** Nomenclature of all (currently 23) recognised species within the Family Ziphiidae and their mean total body length and geographic distribution.

| Sub family | Species | Beaked whale | Mean length (m) | Distribution | Reference |
|---|---|---|---|---|---|
| Hyperoodontinae | *Ziphius cavirostris* | Cuvier's | 6.1 | Cosmopolitan distribution in all oceans except polar waters. | Heyning and Mead 2009 |
| | *Indopacetus pacificus* | Longman's | ~7 | Tropical Indian and Pacific oceans. | Dalebout et al. 2003 |
| | *Mesoplodon bidens* | Sowerby's | 5.5 | Temperate North Atlantic. | Pitman 2009 |
| | *Mesoplodon bowdoini* | Andrew's | 4.4 | Temperate waters of Southern Hemisphere. | Pitman 2009 |
| | *Mesoplodon carlhubbsi* | Hubb's | 5.3 | Temperate North Pacific. | Pitman 2009 |
| | *Mesoplodon europeaus* | Gervais' | 5.2 | Warm temperate and tropical waters of North Atlantic. | Pitman 2009 |

**Table 1 continued.** Nomenclature of all (currently 23) recognised species within the Family Ziphiidae and their mean total body length and geographic distribution.

| Sub family | Species | Beaked whale | Mean length (m) | Distribution | Reference |
|---|---|---|---|---|---|
| Hyperoodontinae | *Mesoplodon hectori* | Hector's | 4.4 | Temperate waters of the Southern Hemisphere. | Mead 2009a |
| | *Mesoplodon ginkgodens* | Ginkgotoothed | 5.3 | Tropical and warm temperate waters of Indian and (mainly western) Pacific oceans | Pitman 2009 |
| | *Mesoplodon grayi* | Gray's | 4.7 | Temperate and sub-Antarctic waters of Southern Hemisphere. | Publication VII |
| | *Mesoplodon layardii* | Straptoothed whale | 6.2 | Temperate and sub-Antarctic Southern Hemisphere. | Mead 2009a |
| | *Mesoplodon mirus* | True's | 5.3 | Temperate waters of North Atlantic, Indian Ocean (also several strandings in Australia and New Zealand). | Constantine et al. 2014<br>Aguilar de Soto et al. 2016 |

**Table 1 continued.** Nomenclature of all (currently 23) recognised species within the Family Ziphiidae and their mean total body length and geographic distribution.

| Sub family | Species | Beaked whale | Mean length (m) | Distribution | Reference |
|---|---|---|---|---|---|
| | *Mesoplodon perrini* | Perrin's | 4.4 | Known only from strandings in California. | Pitman 2009 |
| | *Mesoplodon peruvianus* | Peruvian | 3.9 | Mostly tropical waters of eastern Pacific. | Pitman 2009 |
| | *Mesoplodon stejnengeri* | Stejneger's | 5.7 | Subarctic and temperate North Pacific. | Pitman 2009 |
| | *Mesoplodon traversii* | Spadetoothed whale | 5.3 | Known from five strandings in New Zealand and Chile. | Publication I |
| | | | | | Publication II and III |
| | *Mesoplodon hotaula* | Deraniyagala's | 4.6 | Tropical Indo-Pacific | Lacsmana et al. 2015 |

The 15 species within the genus *Mesoplodon* are particularly cryptic in external appearance and behaviour. All species have a spindle-shaped body, a small falcate dorsal fin, small flippers and a 'beak' or rostrum of various lengths. Detecting the surface behaviours of these whales is challenging. As these whales surface, the beak breaks the water first, and they generally show a low profile with small blows that are difficult to observe except on particularly calm days.

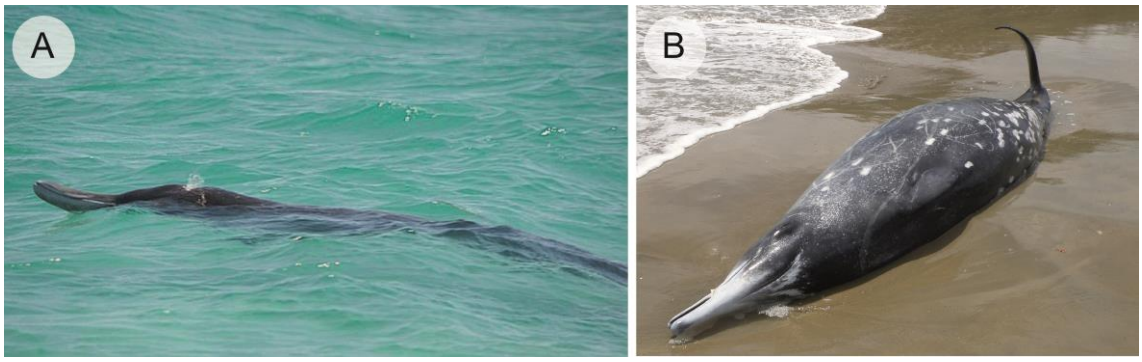



**Figure 3.** Gray's beaked whale (*Mesoplodon grayi*) showing surfacing behaviour and overall external morphology. A) During surfacing the beak breaks the water first and the whale has a small blow and low surface profile (Photo credit: Jon Gall). B) This male stranded Gray's beaked whale shows long beak, or rostrum, spindle-shaped body and small pectoral flippers (Photo credit: New Zealand Department of Conservation).

The crypticity in both external morphology and in behaviour, as well as the inaccessibility of their habitat, makes the beaked whales a difficult group of animals to study. Whilst some species are potentially killed as a result of bycatch in oceanic fisheries, there are no systematic programmes to monitor this mortality and presently genetic samples are rare unlike for other oceanic cetaceans (Madsen et al. 2014).

## 1.2 Molecular methods for studying beaked whales: Sample collections that underpin research

The ziphiids are a family that likely includes several cryptic species, yet because obtaining specimen information is extremely difficult, definitively describing the phylogeography and taxonomic relationships within the family has been a slow process. The lack of specimen information means that it is necessary to draw on many different types of sample repositories.

National museums hold important collections, such as voucher specimens that give the first morphological descriptions. In the case of beaked whales these collections are sometimes the only records for a species, e.g. the spadetoothed whale (*M. traversii*). Analyses of archaeological collections have also shown that whale remains are common in prehistoric and historical middens (Smith 2011; Foote et al. 2012a), though whether these hold ziphiid specimens is currently unknown. These middens give a historical snapshot of species distributions, as well as the relationship that communities had with these animals throughout time. Given the rapid advancement in cost effective next-generation sequencing, especially applied to the field of ancient DNA, there is considerable scope to further explore diversity throughout long time spans in the form of retrospective genetic monitoring (Hofreiter et al. 2015; da Fonseca et al. 2016). Local museums are also central focal points for contemporary collections and, with sufficient facilities, can be a valuable resource for contemporary genetic monitoring.

Until now, these various types of collections represent an under-utilised resource. These museum collections depend heavily on whether there is an initial investment in sampling and subsequent storage. For the study of beaked whales, these natural history collections are highly valuable.

In New Zealand, a programme of long-term systematic collection of tissue samples from stranded cetaceans has facilitated researchers in answering many important conservation questions. This collection was initiated in 1991 and is known as the New Zealand Cetacean Tissue Archive (NZCeTA). This provided the basis for an online genetic DNA-barcoding tool for molecular identification of specimens, known as *DNA-surveillance* (Ross et al 2003). During the period between 1995 to 2016 researchers have used this resource for the first assessments of beaked whale phylogeny, multiple beaked whale species descriptions, and an initial assessment of Cuvier's beaked whale genetic population structure using mitochondrial DNA (mtDNA) markers (e.g., Dalebout et al. 1998; 2003; 2004; 2005; 2007; van Helden et al. 2002).

More than twenty years on from its conception, a more recent study of the species diversity within NZCeTA highlighted the collection as an important tool for researchers globally, particularly given the high diversity of cetacean strandings (Thompson et al. 2013). Containing tissue samples from 13 of the 23 species of beaked whale, and eight of the 15 species of mesoplodonts, this collection has now provided an opportunity to examine more detailed aspects of beaked whale population biology (Thompson et al. 2013; Constantine et al. 2014).

### 1.3 Population genetics in wide ranging marine organisms: markers and sampling

Assessing genetic population structure in species that are oceanic, wide-ranging and difficult to sample is particularly challenging. For some species, ocean current patterns, sea floor topology and other environmental features provide barriers to dispersal, though some areas of the deep ocean lack these obvious features that might drive differentiation. Many marine fish species are characterised by low levels of genetic divergence between populations and high migration rates, presumably as a result of high

levels of genetic and demographic connectivity (Ward et al. 1994; Waples 1998; Reiss et al. 2009). Such high levels of gene flow between populations can mask the signal to noise ratio for divergence estimates (Waples 1998). As a consequence there can be a tendency to underestimate subtle differentiation between populations and this will have negative consequences for species that are commercially exploited or in decline, or give a misleading picture of population biology for those that are not.

To estimate population genetic structure, ideally individuals would be sampled from the same generation because allele frequencies vary over time as well as space (Balloux & Lugon-Moulin 2002) .For marine species, such as the beaked whales, not only are distributions, foraging or migration not well understood, but sampling is so sparse that it is highly likely to incorporate overlapping generations. Such factos may lead to a lack of power in any analyses and the challenge is estimating whether *a priori* populations are collections of individuals that are sampled across seasons and mixed as a result of migration.

Samples collected from stranded or opportunistically sampled marine species with high gene flow introduce a further complexity that leads to a combination of bias that makes population genetic estimates especially difficult. Stranded individuals have been sampled from a population that is geographically unknown and in the absence of any complementary data we must make assumptions in assigning *a priori* regions – individuals found together are assumed to be spatially from the same population in adjacent offshore waters. This assumption can introduce noise into any analyses and potentially reduce our ability to detect the true number of populations. Waples and Gaggioti (2006) evaluated methods to detect population structure and found that the power of analyses that cluster individuals without using *a priori* sampling information dropped significantly when identifying the true number of populations, particularly where gene flow was high and sampling of individuals or loci were low. To disentangle

the complexity and bias introduced by stranding data, some authors have attempted to combine such data with drift and spatially explicit modelling to gain a clearer picture of sampling representativity (Peltier et al. 2013; 2014). Assigning individuals to theoretical *a priori* populations based on complementary data, such as stable isotopic ratios or fatty acid analyses, could help to reduce such confounding factors introduced by the use of stranded individuals.

For marine species that are particularly understudied, such as cetaceans and elasmobranches, basic vital rates – for example longevity, generation time, dispersal and demographic connectivity – are often unknown (Lowe & Allendorf 2010). Many of these are required for population genetic software input parameters and making 'best guess' estimates for these vital rates can introduce further bias into analyses of differentiation. These vital rates are also key in making estimates of female effective population size, where one underlying assumption is that generations are demographically stratified, i.e. not overlapping, which is rarely true in natural populations (Hare et al. 2011).

In some cases, the number of genetic markers is too low to provide enough resolution in estimating weak population structure (O'Reilly et al. 2004). Species-specific microsatellite markers are selected for high levels of locus polymorphism to enable increased statistical power for assessing divergence. However, for such multi-allelic markers, $F_{st}$ based estimates are constrained by within-population variation making inferring subtle structure difficult (Meirmans & Hedrick 2011).

Where microsatellite loci are under selection, i.e. non-neutral, and exhibit high mutation rates, this can have a significant influence on the ability to detect population differentiation (O'Reilly et al. 2004). Microsatellites are commonly assumed to evolve according to a simple stepwise mutation model (SMM) where mutations result in the

deletion or addition of one to several of the repeats (Kimura & Ohta 1978). In truth, microsatellite mutation may be a great deal more complex and this can have knock on effects on population genetic estimates. Willems et al. (2014) analysed several features of nearly 700,000 microsatellite loci across more than 1000 humans as part of the 1000 Genomes Project and found that shorter motifs and those that reside outside coding regions showed higher variability. Motifs showed a unimodal distribution with one common allele and a number of other alleles with rapidly declining frequencies suggesting that the SMM could hold true. However, complex patterns of variation, such as insertions or deletions, within the motifs or flanking regions, or allele size constraints may mean that the SMM is oversimplified (Garza et al. 1995; Van Oppen et al. 2000). Without extensive genomic data and the analyses of the exact mutation rate per generation for each locus, it is difficult to assess the bias introduced by such an oversimplification. Theoretically, it would be possible to address this question with a large number of samples and an analysis of the squared differences in the repeat size between two alleles of the same locus (Slatkin 1995). However, these types of analyses are often not possible in non-model species.

In addition to gaps in our understanding of microsatellites across the genome and the number of loci required to answer a question on a given populations, size homoplasy (i.e. fragments that are identical in size but not descent) can also present errors and bias. Size homoplasy will reduce the number of observed alleles per population, gene diversity and the proportion of heterozygous individuals (Estoup et al. 2004; O'Reilly et al. 2004). These effects can produce overestimates of population divergence, though for most population genetic analyses the number of loci and their variability will compensate for size homoplasy. For species characterised as having high mutation rates and large population sizes with allele size constraints, the confounding effects of size homoplasy are more important (Estoup et al. 2004). In some cases for marine species,

genetic data alone is insufficient to resolve population structure and much greater resolution is obtained when multiple streams of data link to illuminate a consistent picture of a species' biology (Waples 1998).

### 1.4 Kinship analysis and reliability

Examining genetic kinship in animal aggregations can give crucial information on mating systems, dispersal, migration, natal philopatry, and social organisation (Ortega-Ortiz et al. 2012; von der Heyden et al. 2014; Heupel et al. 2015; Quintanilla et al. 2015). With these data we can make also inferences on how species use geographic areas for feeding or mating which can be critical in making decisions on the selection of sites, or seasons, for protection.

Methods to identify kinship broadly rely on one of the following four principles: allelic exclusion; categorical assignment; pedigree reconstruction or fractional assignment (Blouin 2002; Jones et al. 2010). The most common methods used are strict exclusion and categorical assignment where the genotype of each offspring is compared to that of all candidate parents. For strict exclusion, any parent that does not share at least one allele at a given locus is excluded and if more than one parent cannot be excluded, categorical assignments will assess the likelihood of each parent-offspring pair being true given their genotype. These tests also sit within the context of the observed allele frequencies in the population (Kalinowski et al. 2007). Categorical assignment approaches can more easily accommodate scoring errors, missing data or null alleles that are often a reality in microsatellite data sets (Wang 2010). In addition to evaluating pairwise relationships, individuals can also be clustered into family groups and then the likelihood of different clusters is estimated to identify the most parsimonious pattern (for example, Wang & Santure 2009; Wang 2012).

Sampling in that includes both sets of candidate parents will strengthen results, but this is not always practical in difficult to study species. In open systems, where all the parents are not sampled, pure exclusion methods can fail to identify parents or may assign false parents if the genetic information is limited. This is particularly true if these data contain a high proportion of genotyping errors (Blouin 2002). In such cases, a range of statistical likelihood methods using simulations, Monte Carlo permutations or Bayesian approaches can be applied in combination with other biological information to overcome these difficulties (Jones et al. 2010).

A number of different molecular markers can be used for parentage and kinship analyses. Microsatellites are generally the marker of choice as they are highly polymorphic and co-dominant, though their development is fairly labour intensive. The variability in such microsatellite loci has been described as delivering approximately six times the power of single-nucleotide polymorphisms (SNPs) when comparing the number of loci required for any given analyses (Städele & Vigilant 2016). The use of SNPs, however, is becoming more common with the proliferation of genomic data (Hauser et al. 2011). SNPs are less variable and many more loci are required, though generating SNP data is comparatively inexpensive and easy to replicate across multiple platforms.

The accuracy of parentage analyses and family reconstructions is influenced by the number and variability of the loci used, and this is true for both microsatellite and SNP markers (Neilsen et al. 2001; Harrison et al. 2013). Bernatchez & Duschesne (2000) suggest that these two factors – number of loci and allelic diversity – contribute interactively in achieving assignment success. Harrison et al. (2013) calculated the performance of parentage analyses using microsatellite loci in simulations with open populations where not all candidate parents were sampled. As previously reported, accuracy declined with the number and diversity of loci and these authors suggest that

20 or more loci (with 100 % sampled parents) yielded the most robust assignments. The proportion of the population sampled and the effect of genotyping errors were also evaluated and the full likelihood approach, as implemented by COLONY (Wang 2004; Jones and Wang 2010), was found to outperform the other methods tested (Harrison et al. 2013). The performance of COLONY was reduced in situations where large family groups were infrequent, as is common in wild animals systems. In natural populations where only a small proportion of the population is sampled, this can affect accuracy in kinship assignments in several ways. Sampling higher proportions of the population can increase the likelihood of falsely assigning a parent (type Ia error) or falsely excluding a parent (type II error) when the parent is in the sample. Sampling more and more adults exponentially increases the number of possible pairwise comparisons and can lead to erroneous assignments. Harrison et al. (2013) conclude that the effect of these errors will have differing consequences on the conclusions of the study depending on the central question and that, in general, increasing the number and diversity of loci can overcome the effects of incomplete sampling.

For species that are difficult to observe, kinship analyses can be combined with other techniques, such as tracking, or another molecular analysis to illuminate a variety of aspects of biology. For example, using fatty acid signatures that reflect key prey choices alongside kinship data and can reveal the transient nature of social groupings. Watt et al. (2015) describe the use of both molecular genetic and dietary markers to identify the social structure of narwhal (*Monodon monoceros*). Their data indicated that these whales form fission-fusion societies similar that have been suggested in a number of other cetaceans (for example, bottlenose dolphins (*Tursiops truncatus*) Tsai & Mann (2013), and Baird's beaked whale, Fedutin et al. (2015)). Such dietary data can complement those derived from genetic analyses and provide insights into a species' vital rates and strengthen the validity of outputs from models estimating population

size, demographic trends and population viability.

## 1.5 Genomic information from rare species: Resolving phylogeny, incomplete lineage sorting and hybridisation

Genomic approaches can provide powerful insights into population history, overcoming the hurdles introduced through low sample numbers due to the sheer number of loci available. Full and even partial genomic sequencing allows us access to information from layer upon layer of generations providing a much greater resolution in the study of evolution.

Genomic studies generally follow a typical pipeline for generating data and analyses: design of sequence strategy, generation of data, mapping sequence reads to the assembly, variant calling and downstream analyses (Ellegren 2014). The accuracy of the final genome assembly will depend on all of these factors including the depth of coverage, assembly software and the type of platform used for sequencing. Until now, many assemblies were highly fragmented with gaps, ambiguities and errors (Bickhart et al. 2017). However, new sequencing technologies have now improved our ability to sequence long reads. Currently, the Pacific Biosciences PacBio RSII platform that uses single-molecule sequencing can routinely achieve read lengths of 14 kb with maximum lengths of >60 kb (Eid et al. 2009; Gordon et al. 2016). Such high fidelity single-molecule sequencing data, analysed in tandem with those generated by more affordable sequencing platforms, for example Illumina HiSeq, and scaffolding by optical and chromatin interaction mapping, means that many of the previous issues confounding genomic analyses, such as gaps or errors, can be largely overcome (Bickhard et al. 2017).

Improving the accuracy of full genomic data has been pursued in the drive to improve

agricultural variants, such as cattle and goats (for example, Bickhard et al. 2017) but will also facilitate our understanding of non-model organisms. There are now numerous full genomes for such non-model organisms, most likely of differing completeness and quality. Some of these genomic data have been generated to answer conservation questions. For example, full genomes now exist for numerous bird species (for example Puerto Rican parrot (*Amazona frittata*), Oleksyk et al. (2012), multiple species of conservation concern, Zhang et al. (2014)) and mammals (for example, giant panda (*Ailuropoda melanoleuca*), Zhao et al. (2013), western lowland gorilla (*Gorilla gorilla gorilla*), Gordon et al. (2016)). One drawback to this work can be the $n = 1$ phenomena, where inferences are made from a single sample, relating to a single outcome of evolution.

Full genomes give a better understanding of the biology of the genome itself and we can see how genes are organised, packaged and other aspects, such as the abundance of transposable elements and chromatin marks (Ellegren 2014). It is possible to estimate recombination rates (the amount of recombination per unit of DNA) and allele frequency spectra with much greater accuracy, in comparison with previous linkage mapping. Recombination rates influence the efficacy of selection through Hill-Roberston interference that predicts that selection at linked sites interferes with selection at a given position on the genome (Hill & Robertson 1966; McVean & Charlesworth 2000). When recombination is high, linkage with extend over shorter physical distances and adaptive evolution should be more common at these sites. Such studies of recombination across the genomic landscape have shown that rates across the genome are heterogeneous, with hotspots for recombination that can tell us much about the selective forces that may have acted upon a species or population in the past.

Reverse genomic studies use likelihood methods to estimate the strength of selection from a mixture of signals. Selective sweep mapping can show that a region has reduced

nucleotide diversity and extended linkage disequilibrium (LD) (where the association of alleles from different loci is non-random), runs of homozygosity and high recombination rates can be highly informative (Ellegren 2014). These analyses must also incorporate information on how much diversity would be expected through neutral processes, i.e. background mutation.

Many studies focus on quantifying lineage-specific adaptation to a local environment or to a particular lifestyle such as diet or high altitude. The forward genomics approach involves identifying the genes for a specific phenotype and, in mammals, this is facilitated by the existence of databases (e.g. OrthoMaM) that categorise orthologous exons and coding sequences to improve the chance that the user may identify the function of a given shared gene (Douzery et al. 2014; Caspermeyer 2016; Prudent et al. 2016). Genome-wide association studies aim to find the underlying loci for particular phenotypic traits by sequencing cohorts of individuals with extreme phenotypes. The whole genome can be scanned for enrichment of certain alleles that are thought to be responsible for these phenotypes. Such studies have become increasingly important in predicting the genetic basis for an organisms' ability to adapt to a changing environment (Rönnegård et al. 2016).

Patterns of genetic diversity between and within populations are driven by demography, divergence and the degree of reproductive incompatibility. Whole-genome data offer much greater insights into changes in population size and to what extent gene flow and introgression (transfer of genetic information from one species to another as a result of hybridization) influences genetic diversity (Sousa & Hey 2013). With data from across the genome, it is now possible to assess how different regions of the genome are more or less prone to gene flow.

Allele frequency spectra, sometimes referred to as site-frequency spectra (SFS), give the

distribution of allele frequencies at a given set of loci (often SNPs) in a population or sample. The shape the SFS is displayed as a histogram of that reports the total number of loci with a derived allele frequency and distributions that are strongly skewed can indicate selection (for an example in honeybees (*Apis mellifera*) see Wallberg et al. 2015). Patterns in SFS for a species or population will also be sensitive to demography, such as changes in population size, migration and structure.

There are a number of ways in which population history can be assessed using SNP frequency spectra. Likelihood model-based diffusion approximation across SFS is a popular method (Gutenkunst et al. 2009), as are approximate Bayesian computation (ABC) approaches. ABC uses summary statistics to characterise patterns of variation observed in the data (Beaumont 2010). Reconstructing historic changes in effective population size ($N_e$) using SNP frequency spectra is now possible using stairway plots in a method that is model-flexible and, by simulation, will allow hundreds of genomes to be incorporated into the analyses (Liu & Fu 2015). Such approaches have been used in combination to illuminate population history in several mammalian species, for example, to make inferences of the population history of Bornean (*Pongo pygmaeus*) and Sumatran (*Pongo abellii*) orang-utans (Locke et al. 2011). This work used SNP-based SFS estimates from one draft genome and short-read sequence data from five genomes of each species to investigate diversity, differentiation and $N_e$ for the two species and secondary gene flow after divergence. Such high-resolution data revealed surprisingly counter-intuitive results that suggest that Sumatran orang-utans have greater genetic diversity despite a much smaller (approximately sevenfold) population census size.

Foote et al. (2016) use SNP-based SFS to investigate population structure and functionally inherited alleles in killer whales (*Orcinus orca*). This species exists in ecologically and genetically divergent ecotypes that some suggest should be regarded as

separate species (Ford et al. 2009; Morin et al. 2010). Foote et al. (2016) use a pairwise sequential Markovian coalescent (PSMC) approach to two high coverage (≥ 20 x) genome assemblies and a further 48 individuals where low-coverage data was obtained through re-sequencing. The PSMC method allows a pairwise comparison of chromosomes to facilitate population history hypothesis testing with the ability to simulate scaled mutation and recombination rates (Li & Durban 2011). The resolution provided by the genomic data in Foote et al. (2016) has shed light on the complex evolutionary history of killer whale populations. Behavioural plasticity has facilitated colonization of novel ecological niches that are followed by founder effects (for example, population bottlenecks) and rapid reproductive isolation that then resulted in population expansions. Each population contraction and expansion event has influenced the genomic landscape of killer whales and a functional association of genes that reflects adaptation to a specific diet and climate for each ecotype.

Where a dramatic change in environmental conditions has facilitated accelerated adaptive radiations in species groups, sequential speciation can occur more rapidly than the phylogenomic markers can be fixed. In such cases, analyses of these markers will be incongruent with species trees as a result of incomplete lineage sorting (ILS). Genomic data can provide greater resolution in phylogeny, particularly where there is a prevalence of ILS. An example of ILS in a species group includes the signature of rapid diversification that is imprinted on the genomic analyses of neoavian birds. All avian species, with the exception of the Paleaognathea (ratites and tinamous) and Galloanserae (ducks and pheasants), underwent one of the most rapid of adaptive radiations of all extant species at a time coinciding with the Cretaceous-Paleogene boundary resulting in 36 bird lineages (Alfaro et al. 2009; Suh et al. 2015). Simulations suggest that such a rapid diversification has given rise to high levels of gene tree-species tree incongruences (known as hemiplasy) due to ILS that make inferences on

phylogeny at the level of individual loci difficult. Suh et al. (2015) analysed ~130,000 retrotransposons (transposable elements in the genome that show homology with retroviruses) in 48 bird genomes to identify presence/absence patterns of insertions that are found genome-wide. This study tracked the fate of retrotransposons as they are virtually homoplasy free and the probability of independent insertions of these elements is very rare. Suh et al. (2015) characterized the extent of ILS using these retrotransposon elements and found that a third were affected giving persistent polymorphisms across multiple speciation events throughout the bird lineages.

The genomic consequences of hybridization are also complex. These data allow us to investigate the history of gene flow between nascent species and what genetic barriers might maintain their divergence. Diversity patterns at high numbers of loci can be used as replicate samples of hybridization history. Unusual patterns can indicate areas of the genome that contribute to reproductive isolation or adaptive introgression.

Populations can be sampled to identify whether they belong to one diverging lineage (species) or another. In some cases, sampled populations can reflect recent hybridization (for example, hybrid zones) or other measure of admixture such as geographic clines. After investigating measures of differentiation, the observed genomic distributions can be compared to those of expected outcomes from models to infer hybridization or speciation. There are various modeling and statistical methods that include using underlying coalescent, Bayesian or likelihood approaches that can be applied to whole genome, RADseq or transcriptome data (see Payseur & Rieseberg 2016). These approaches can also be applied to ancient DNA to shed light on historical demographic processes (Schaefer et al. 2016). All methods need to encompass the dynamic nature of gene flow and speciation, particularly in terms of its timing and the rate of recombination across lineages, which are very often unknown.

A study conducted by vanHoldt et al. (2016) provides an example of admixture mapping using SNP data that has been applied to identify introgression between North American wolves (*Canus lupus*) and coyotes (*Canus latrans*). This study focused on 3102 ancestry informative SNP markers in samples from individuals in a hybrid zone that has formed over the last 100 years. Animals in this zone have intermediate phenotypes with some populations appearing more wolf-like and others more like coyotes. Admixture mapping is one way of understanding the evolutionary dynamics within hybrid zones and identifying which genes are likely to be under selection. It assumes that hybrid individuals have a genome with a mosaic pattern of ancestry where the frequency and size of ancestry blocks are dictated by the intensity of selection as well as the direction, duration and rate of gene flow between the two species (Tang et al. 2006; Buerkle & Lexer 2008; Winkler et al. 2010). Admixture mapping also assumes that some of these ancestry blocks will be linked to a phenotype that is under selection and that foreign ancestry will develop for introgressed blocks under selection (Buerkle & Lexer 2008; Winkler et al. 2010). By characterising differentially introgressed blocks in individuals throughout the wolf/coyote hybrid zone, vanHoldt et al. (2016) identified 10 regions of the genome where genes contribute to phenotypic variation with certain regions that are likely to be functionally linked to skeletal variation and dentition.

Even lower coverage genomic approaches such as RADseq sequencing can give greater resolution of a species population history, given that these analyses will testing divergence in thousands, rather than tens, of markers (Andrews et al. 2016). SNP discovery though advanced software pipelines filter out poor quality reads, classify pools of individuals based on barcodes, identify loci and alleles *de novo*, align reads to an index to discover polymorphisms, and score genotypes (Catchen et al. 2013). Genotyping-by-sequencing using restriction digests can produce unique data that can

bias analyses and there are basic recommendations for data filtering to mitigate and correct such bias (Narum et al. 2013).

Novel ways of analysing genomic data are emerging that may change the way we approach population genetics. Past evolutionary processes can also be identified using linkage disequilibrium network analyses (LDna), though this method is in the early stages of development (Kemppainen et al. 2015). These analyses use LD in combination with mathematical network analysis to identify clusters that have a greater connection between their members (loci) than the remainder of the network. Evolutionary processes can result in elevated LD amongst loci leading to distinct clusters in these LD networks. Trials of the method on sticklebacks (*Gasterosteus aculeatus*) that have a known population demographic history, and characterised local adaptation, using SNP data suggest that this method is able to identify loci that are associated with these processes. In future, LDna may be useful in identifying current geographic partitioning of populations in non-model organisms that have not undergone full genome sequencing, particularly in combination with other methods. With the sheer volume of data that can be generated it is also possible to move away from a previously linear approach and incorporate such network analyses more routinely (Morrison 2014).

Genomic sequencing provides vast opportunities to those studying rare or difficult of sample species. The barriers that are intrinsic to low sampling may to some extent be overcome by the ability to sample across the whole genome and effectively gain information from multiple generations in doing so. Most approaches require stringent assembly methods and well-designed hypothesis and model testing. In beaked whales, one of the greatest difficulties in studying most species is the lack of basic information on how and where these animals live in their deep ocean world. Comparative genomics across deep-diving species (and their shallow-diving relatives) could provide significant insights into aspects of this extreme lifestyle. For example, estimates of selection on

certain functional genes related to dietary proteins or diving may mean that we can better predict how certain species live or identify convergence of characters across the group (Foote et al. 2015; Foote et al. 2016). Genomic analyses are also likely to shed light on beaked whale speciation and phylogeography, which are currently are not well defined, potentially as a result of ILS and the fact that the beaked whales may have undergone a rapid diversification during their adaptive radiation throughout the oceans.

## 1.6 Outline and aims of this study

The research within this thesis contributes to the overall knowledge of several species of beaked whale using multiple genetic methods and several existing tissue collections. Through collaboration, the overall objective was to apply these molecular methods to:

1) Investigate the species identity of previously unknown species of beaked whales and, in doing so, provide both external descriptions and an update of their taxonomy and global distributions,
2) Examine differences in morphological variation in Gray's beaked whales – between sexes and geographical areas – using the molecular identification of species and sex of specimens,
3) Evaluate genetic diversity in Gray's beaked whales and investigate population structure and demographic trends in New Zealand and western Australian waters,
4) Examine group composition and patterns of genetic kinship within and between groups of stranded Gray's beaked whales.

To achieve these aims, several different techniques were used to provide the basis for eight publications (I – VIII, see details below). These publications have updated our taxonomic knowledge of two beaked whale species, the spade-toothed beaked whale (I)

and Deriniyagala's whale (II and III), and provide a more detailed investigation into the life history of Gray's beaked whale. This work includes a molecular characterisation of Gray's beaked whale mitochondrial genome and the development of novel species-specific molecular tools with which to investigate demographic history, female effective population size, population structure, gene flow and social organisation (IV, V, V, VII and VIII).

## 1.7 Summary of publications

### I. The world's rarest whale?

The spadetoothed beaked whale (*Mesoplodon traversii*) is a species that was initially taxonomically resurrected in 2002 following phylogenetic analysis of mtDNA sequences that included the only three bone specimens recorded globally (van Helden et al. 2002). Though the species had been genetically characterised there were no descriptions of its external morphology and the whale had never knowingly been seen before. In publication I, I used sequences from ~ 750 bp of the mitochondrial control region, and 400 bp of the cytochrome *b* gene, with the online tool *DNA-surveillance* to confirm this specimens' species identity. We confirmed the spadetoothed whales' external morphology from specimen photographs taken at the time of stranding.

### II. Human consumption of cetaceans and the identity of the 'unknown beaked whale' in South Pacific equatorial waters?

The Republic of Kiribati consists of 32 coral atolls and one raised island that form three island clusters: the Gilbert Islands, the Phoenix Islands and the Line Islands. Many of these atolls are uninhabited, or sparsely populated. In total, the Republic of Kiribati covers an area of the equatorial South Pacific Ocean that covers approximately 3.4 million $km^2$ of territorial EEZ waters. The diversity of cetacean species in this vast

oceanic area is particularly understudied, as is the degree of hunting, although there are numerous examples of traditional subsistence hunting throughout the South Pacific (Robards and Reeves 2011). In 2003, dried meat served at a local feast on the island of Tabitua (one of the Gilbert Islands) was reported by locals to be from several 'long whales' that were captured in the local lagoon using the practice of 'drive hunting'. Analyses of mtDNA sequences from the control region and the cytochrome *b* gene suggested that these whales were genetically distinct from other whales that had been previously genetically characterised. It was thought that these whales possibly represented a new species, or sub-species of beaked whale (Dalebout et al. 2007). Interestingly, a separate study using acoustic surveys conducted around Palmyra Island (Line Islands) over two field seasons (2007-2008) also identified vocalisations within the frequency range used by beaked whales that could not be attributed to any known species (Baumann-Pickering et al. 2010). These vocalisations were also thought to be from an unknown *Mesoplodon* sp, potentially that which was described in Dalebout et al. (2007).

During 2009, two co-authors (Publication II) visited the Gilbert Islands and interviewed elders on their cetacean hunting practices. They showed villagers images of species for identification, and collected bones and artefacts from recent strandings. Using the online tool *DNA-surveillance*, comparisons of mtDNA sequences from the control region and cytochrome *b* gene identified four different cetacean species. A further species, sperm whale (*Physeter macrocephalus*), was also identified though morphological examination of a tooth. Three species of beaked whale were identified: Cuvier's and Blainville's beaked whale and the unidentified *Mesoplodon* sp., with sequences matching the previous dried meat samples collected in 2003. This publication provided the first evidence of human consumption and the diversity of cetaceans in the Republic of Kiribati. It also provided further evidence of the existence of this potentially new

species of beaked whale, though at the time of publication there was no evidence giving the external morphology of this whale. However, in 2009 a beaked whale stranded in the islands of the Seychelles in the Indian Ocean appeared to be morphologically different from those that had previously stranded (M. Dalebout, pers. comm.). Sequences from the mitochondrial control regions, cytochrome *b* and cytochrome oxidase *c1*, and seven nuclear autosomal introns and one Y-chromosome region were amplified for all suspected specimens of the unknown *Mesoplodon* sp. stranding in an equatorial band from the Indian Ocean to the South Pacific. These samples included the Seychelles specimen for which there were good images. The specimens (seven whales) also included the holotype specimen of *Mesoplodon hotaula*, found in Sri Lanka, a species previously described in 1963 (Derinayagala 1963a; 1963b) that was later synonymised with the ginkgotoothed whale by Moore and Gilmore (1965).

**III. Resurrection of Derinayagala's whale**

Phylogenetic analyses sequences from the unknown *Mesoplodon* sp. and all other *Mesoplodon* spp. showed that this whale and the ginkgotoothed whale specimens clustered together in two highly supported clades that were reciprocally monophyletic to each other. Using the Genealogical Concordance Species Concept (GCSC) of Avise and Ball (1990), publication III presents multiple lines of evidence from both the molecular and morphological cranial data, to support the hypothesis that the unknown *Mesoplodon* sp. is, in fact, *Mesoplodon hotaula*, as described by Derinayagala in Sri Lanka. We suggest that it be taxonomically resurrected as a species that is distinct from its closely related sympatric sister species, the ginkgotoothed whale, and named Derinayagala's whale. We also presented the first photographic evidence of the external morphology of this whale and made descriptions of the cranial and distinct tooth morphology of the species, which has since also been recorded in the Philippines (Lacsmana et al. 2014).

According to one researcher at the Davao del Norte State College, Philippines, Derinayagala's whale may strand fairly regularly around the coasts of the islands (N. A. Abreo, pers. comm.).

**IV. Characterisation of Gray's beaked whale mitochondrial genome**

To further investigate the molecular ecology of Gray's beaked whale we characterised the mitochondrial genome. Using an Illumina Miseq platform we shot-gun sequenced a single Gray's beaked whale (*M. grayi*) to an average depth of coverage of 152X (Publication IV). We performed a *de novo* assembly and were able to determine that the mitogenome of this species is 16,347 bp and has an organization similar to other cetaceans. At the time of publication, this mitochondrial genome was the first to be published for a *Mesoplodon* species.

**V. Species-specific microsatellite characterisation**

Using multiple second-generation sequencing platforms (GS Junior 454, Ion Torrent and Illumina Miseq) we designed a suite of microsatellite primers that are specific to Gray's beaked whales. No previous specific molecular work had been undertaken for Gray's beaked whales, apart from genetic identification of stranded samples. Therefore, we were required to develop these molecular methods to ensure that any further work would be based on robust protocols.

We also compared different sequencing platforms and their respective data to give technical support to other researchers in microsatellite discovery and primer design (Publication V). For example, we suggest that eliminating low-quality reads and trimming read lengths, so that 80% of bases in each sequence have >20 Phred score, was essential in the design of reliable primers. In addition, we found that for consistent

peak calling, and to reduce stuttering, tri- and tetra-repeats produced more consistent genotyping.

### VI. A multidisciplinery approach to investigate geographic variation and sexual dimorphism

In publication VI I used genetic sex and species identification on stranded Gray's beaked whales, together with specimen records, to examine sexual dimorphism and geographic variation in this species. Integrating measurements obtained from museum specimens, we used multivariate analyses to test for these differences and found a high degree of sexual dimorphism in cranial anatomy but only subtle morphological difference between animals stranded on different coasts of New Zealand. These results were then used to build hypotheses to test for levels of genetic connectivity in Gray's beaked whales.

### VII. An assessment of past demographic processes and contemporary gene flow in Gray's beaked whales

To test the hypothesis that there is restricted gene flow between the east and west coasts of New Zealand, and further to West Australia, we analysed data from 530 bp of the mitochondrial control region and 12 highly polymorphic microsatellite loci in 94 Gray's beaked whales (Publication VII). Genetic diversity estimates were found to be moderately high in comparison to that reported in other beaked whales. Using analysis of molecular variance (AMOVA), Bayesian Inference and discriminant analysis of principal components (DAPC) of genotypes we found a complete lack of genetic structure and a pattern of panmixia across the study area of approximately 6000 km. Contrary to our hypothesis, it appears that gene flow is unrestricted across this spatial scale and morphological differences exist in the presence of gene flow. Tests for departure from neutral equilibrium (Fu's $F_s$) revealed no evidence of a population

bottleneck and estimates of effective population size ($N_e$) using a Bayesion Skyline Plot gave a likely estimate of a large and stable population of whales. In this context, our results suggest that Gray's beaked whales most likely exist as a panmictic population, with movements over large spatial scales, possibly facilitated by broad-scale oceanographic features, such as the sub-tropical convergence that runs through the centre of this species distribution. However, in the absence of any available data from live sightings, these results precipitate emerging questions into possible mating systems, and other life history parameters such as dispersal, that would maintain such high levels of gene flow and genetic diversity.

**VIII. Inferences from molecular data on beaked whale societies**

Cetaceans are unique among social mammals in that certain species regularly die together in mass-strandings. Sperm whales (*Physeter macrocephalus*), short- and long-finned pilot whales (*Globicephala macrorhynchus* and *Globicephala melas* respectively) and false killer whales (*Pseudorca crassidens*) are known to mass-strand frequently (Evans et al. 2005; Ferreira et al. 2014; Téllez et al. 2014). Humans have pondered the significance of cetacean mass-strandings and the commonly held view is that the powerful bonds of kinship between whales overcome their individual need to survive. When one group member becomes unwell, or disoriented, potentially a whole family can die on the beach. Mass-stranding species have been described as having relatively stable, and in some cases, complex kin-based social structures, e.g. sperm whales form several hierarchically organized tiers of female social structure.

In an analysis of both mitochondrial control region haplotypes and data from 16 microsatellite markers, we used the software COLONY, CERVUS and Coancestry to estimate levels of kinship between Gray's beaked whales stranding together in groups (Jones and Wang 2010; Kalinowski et al 2007; Wang et al. 2011) (Publication VIII). In

the 19 stranded groups (ranging from two to 10 individuals), we found no incidences of kinship between adults, indicating that both sexes disperse from their mothers and there were only six statistically well-supported mother-calf pairs.

Most groups consisted of unrelated females, some of which share mitochondrial haplotypes. No cases of paternity were found within the groups and mature males most often stranded as solitary animals, particularly in the autumn. The temporal stability of these female groups remains unclear, and given new tracking technologies could be investigated with re-floating of animals post live stranding.

Only one group contained a female dependent (2.6 m) and a mature male (4.45 m) likely to be half-sibs, a shared haplotype between the two suggesting these individuals could share a mother. Why these half-sibs, of such different age-classes, might have stranded together remains a mystery, though it is not inconceivable that they were part of a larger group where the other members did not strand, or were not discovered.

Finally, our results suggest that there was only one case of a relationship between strandings – in this case a group stranding and an individual. A male and female sampled 17 years apart, and over 1800 km away by sea, appear to be most likely a half-sib relationship, with the whales sharing a father. This suggests that males remain reproductively active for at least 17 years, most likely longer, and range widely throughout their reproductively active period.

This is the first study to assess genetic kinship in beaked whales. Other research using photo-identification suggests that beaked whale species exhibit a range of social systems. Northern bottlenose show both sex- and age-class segregation, whereas in contrast, Blainville's beaked whales form 'harem' groups of multiple females with at least one male (Gowans et al. 2001; Claridge 2013; Dunn 2014). Baird's beaked whales appear to form fission-fusion societies with some stable associations (Fedutin et al.

2014). Unfortunately, in the absence of DNA analyses, we do not know the kinship relationships within these groups. Though other beaked whales do strand, Gray's are the only species to regularly strand in groups.

In summary, we found no evidence of kin associations in these stranded Gray's beaked whale groups. These groups may be formed opportunistically for foraging or reproduction and we cannot exclude a fission-fusion system as observed in Baird's beaked whale. Our results are also consistent with the fission-fusion systems reported in several oceanic delphinids, for example spinner dolphins (*Stenella longirostris*) (Karczmarski et al. 2005), common dolphins (*Delphinus delphis*) (Viricel et al. 2008) and Atlantic white-sided dolphins (*Lagenorhynchus acutus*) (Mirimin et al. 2010). Mirimin et al. (2010) reported stranded groups to consist of unrelated adults with calves that were also not closely related (e.g., half-sibs). This suggested that females had mated with males that were not present within the group and that these dolphins most likely have a promiscuous mating system. In the Gray's beaked whale groups that we have sampled sample sizes are, unfortunately, too low to provide further information on potential mating systems.

**Summary**

In the publications presented in this thesis I, and co-authors, have used multiple molecular genetic approaches to maximize the hidden data within collections of tissue samples that have been accumulated over many years. Through employing several technical and analytical skills, this integrative approach has allowed us to make inferences into where, and how, several beaked whales live in their deep ocean world.

I present a body of work that provides new data on the physical appearance, identification and key population parameters of several species of beaked whale. Our work on Gray's beaked whales shows that there is physical variation within the New

Zealand population that exists in the presence of significant gene flow. Preliminary results suggest that this species exists as a large population and that gene flow extends across to West Australia. These whales may range over large (~6000 km) spatial scales, perhaps facilitated by broad scale oceanographic features. Although, we still know little of how these whales exist in terms of their social structure, our genetic kinship analyses suggest that the whales that strand, do so in groups where individuals are linked only in space and time, rather than by any familial bonds.

The molecular genetic methods used in this work have allowed a unique window into the lives of these rarely observed whales. There is, however, much work to be done on these cryptic, and elusive animals, particularly in the modern context of a changing marine environment.

## 1.8 A reflection on the genetic markers used in this study

The publications in this thesis present the results of analyses that are based on short fragments of DNA – both mitochondrial and nuclear microsatellite markers. In the case of DNA-barcoding, the hyper-variable fragment of the mitochondrial control region, and cytochrome-*b* coding region, is highly diagnostic in beaked whale species identification (Dalebout et al. 2004). These whales generally show low intraspecific nucleotide diversity yet relatively high levels of interspecific divergence. However, for taxonomic descriptions, nuclear loci and species-specific morphological characters are also required to validate any potential new discoveries. Regardless of how stringent the analyses or quality control of sequence data, these loci are still only small fragments of the genomic landscape of a species. In an ideal study, genomic approaches would give better resolution in describing divergence of sister taxa and a groups' phylogenetic history. It is possible that, due to a rapid radiation event within the beaked whale group,

incomplete lineage sorting may persist. Predicting and characterising such ILS is difficult with so few markers across the genome.

A genomic approach is also preferable for population genetic studies for the reasons outlined in Sections 1.3. Full genomes, in combination with analyses of gene expression (transcriptome or even proteomic data), are highly informative for identifying population histories. Acknowledging the various limitations in assembly quality that can confound analyses, high-resolution genomic data can reveal that different regions show distinctive aspects of a species population history. These analyses give exciting opportunities to track a species' evolution, particularly when only small numbers of samples are available, as is the case in beaked whale. In the case of this study, our ability to take a genomic approach was limited. In addition, had we the ability to add an further data stream, such as information from stable isotopes, it may have been possible to better assign beaked whale samples to *a priori* populations or at least test our assumptions on the provenance of the strandings.

Genotypes derived from SNPs are defined by sequence differences rather than estimated allele sizes. Though SNPs only have two alleles, rather than the multiple alleles found in microsatellites, their sheer number throughout the genome make them statistically able to outperform microsatellites. As previously mentioned, mutation rates and models can be difficult to predict in microsatellites and the presence of size homoplasy can mask signals for population genetic structure. Homoplasy is also known to exist in SNPs, but at lower frequencies due to their lower mutation rate and more simple mutation model. This makes divergence-based estimates of $F_{st}$ simpler to interpret from SNP data (Meirmens and Hedrick 2011).

Our analyses of genetic kinship in Gray's beaked whales yielded a surprising lack of familial relationships within groups. Although these results are consistent with that

found in our analyses of population structure, it is possible that our study has suffered from a lack of power due to the number and variability of microsatellites used. Although there is currently no known possibility to improve the sampling of Gray's beaked whales, standardised power analyses could provide a method of assessing the strength of our results (e.g. MacCallum et al. 1996; Cornuet & Luikart 1996; Kraemer & Bladey 2015).

In summary, though we were not able to use full genomic approaches for our analyses of population structure and genetic kinship in Gray's beaked whales, we used shot-gun sequencing to identify highly variable species-specific microsatellite loci. We know that these markers have inherent limitations, though probably more limiting is our inability to gain access to larger sample sizes and distributions. Given either low-coverage genomic RADseq data, or a full genome approach, in combination with an extended sampling range and other streams of data, it may be possible to gain a higher resolution to our study of population structure and genetic kinship.

## 1.9 Future research on beaked whales

The global study of beaked whales currently relies on observations from only a few species and a sparsely distributed collection of samples. Therefore, more focus needs to be made on establishing international sample collections. These collaborative collections need to extend current sampling distributions for known species and increase effort into cataloguing unknown and understudied species.

Throughout the past decade there have been several attempts to draw together current knowledge on beaked whale species, particularly in outreach to scientists and the public on accurate species identification. The Smithsonian National Museum of Natural History Marine Mammal Program, with funding from the Marine Mammal

Commission, has collated species descriptions that can be used to identify stranded specimens. Genetic identification is possible using sequence repositories such as GenBank. These services are important in both informing the public, and researchers, on general baseline biology of species and their distributions.

A particular regional focus should be made on islands of South Pacific and Indian Ocean, where there is known to be the least research to date on marine mammals (Jarić et al. 2015). There are potentially many more beaked whale records, and species, that could be detected from bones, strandings and sightings. Drawing on the experience and availability of on-the-ground experts in local communities and non-government organisations could facilitate the collection of these data. Better sample availability, and by applying newer genomic techniques, may give more resolution on the phylogeny and evolution of beaked whales.

The beaked whales are the extreme divers of the Cetacea and studies using comparative genomics are likely to reveal the genetic underpinning of their remarkable adaptation to foraging in a deep-sea environment. These data could give some insightful information on the differences between species in terms of dietary specialisation, the oxygen carry capacity of haemoglobin or other indications of ecological niche separation that have not yet been quantified.

Where sufficient samples already exist, e.g. Gray's beaked whales, extending the spatial range of this sampling effort combined with genomic analyses may highlight underlying cryptic population structure and connectivity in this species' southern hemisphere circumpolar distribution. The sub-tropical convergence zone is a dynamic oceanographic feature that is known to contain large populations of squid (Butler et al. 1992) and may be a factor in why New Zealand is a global hotspot for beaked whales. Using several streams of molecular data (genetic, fatty acid and isotopic signatures) as

well as acoustic surveys may also elucidate how this oceanographic feature influences Gray's and other beaked whale foraging patterns and population dynamics (Marques et al. 2009; Newsome et al. 2010; Ramos et al. 2012; Zimmer et al. 2008).

Acoustic data may also prove useful. For example, passive acoustic monitoring devices have been used to record the vocalisations of Blainville's beaked whales off the coast of Hawai'i. These recordings have showed clear foraging patterns and habitat use (Henderson et al. 2016). Though New Zealand is a global hotspot for beaked whale strandings, to date no effort has been made to systematically investigate whether these whales can be detected and studied acoustically in these waters.

Because of the high diversity of species found in New Zealand waters, it is also a potential location to trial emerging environmental DNA (eDNA) monitoring techniques (Barnes and Turner 2016; Bohman et al. 2014; Foote et al., 2012b). Gray's beaked whales live-strand around the coast of New Zealand regularly and attaching tags on animals that are re-floated could provide also insights into aspects of their behaviour, albeit directly after a traumatic event.

Research focusing on species that maintain predictable aggregations, e.g. northern bottlenose whales off Nova Scotia, Blainville's and Cuvier's beaked whales off the coast of the Canary Islands and Bahamas, has enabled investigations into social associations, diving and sensitivity to anthropogenic noise (Martín López et al. 2015; Miller et al. 2015). Biopsy sampling of these populations would allow analyses of genetic kinship, and interchange of individuals between locations. Whilst it is still unknown how social systems and movements compare between species, given the difficulties in making assessments of the more difficult to observe beaked whales, it is important to continue these efforts with the aim of providing more insight into the biology of group.

This thesis provides information on the distributions, population history and social organisation of three mesoplodont species. Nevertheless, there are many more questions yet to be addressed.

# Bibliography

Aguilar de Soto N, et al. (2016). True´s beaked whale (*Mesoplodon mirus*) in Macaronesia. Peer J 4: p.e1785v1.

Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. Nature Reviews Genetics 17: 81-92.

Avise JC, Ball RM. (1990). Principles of genealogical concordance in species concepts and biological taxonomy. Oxford Surveys in Evolutionary Biology 7: 45-67.

Balloux F, Lugon- Moulin N. (2002). The estimation of population differentiation with microsatellite markers. Molecular Ecology 11: 155-165.

Barlow et al. (2006). Abundance and densities of beaked and bottlenose whales (family Ziphiidae). Journal of Cetacean Research Management 7: 263-270.

Barnes MA, Turner CR. (2016). The ecology of environmental DNA and implications for conservation genetics. Conservation Genetics, 17: 1-17.

Baumann-Pickering et al. (2010). Echolocation signals of a beaked whale at Palmyra Atoll. The Journal of the Acoustical Society of America 127: 3790-3799.

Beaumont MA. (2010). Approximate Bayesian Computation in evolution and ecology. Annual Reviews Ecology and Evolutionary Systematics 41: 379-406.

Bernatchez L, Duchesne P. (2000). Individual-based genotype analysis in studies of parentage and population assignment: how many loci, how many alleles? Canadian Journal of Fisheries and Aquatic Sciences 57: 1–12.

Bickford D et al. (2007). Cryptic species as a window on diversity and conservation. Trends in Ecology and Evolution 22: 148-155.

Bickhart DM, et al. (2017). Single-molecule sequencing and chromatin conformation capture enable de novo reference assembly of the domestic goat genome. Nature Genetics 49: 643-650.

Blouin MS. (2003). DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. Trends in Ecology and Evolution 18: 503-511.

Bohmann K, et al. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. Trends in Ecology and Evolution 29: 358-367.

Borchers DL, Zucchini W, Heide- Jørgensen MP, Cañadas A, Langrock R. (2013). Using hidden Markov models to deal with availability bias on line transect surveys. Biometrics 69: 703-713.

Butler C, Butt J, Lindstrom E, Tildesley P. (1992). Oceanography of the subtropical convergence zone around southern New Zealand. New Zealand Journal of Marine and Freshwater Research 26: 131-154.

Buerkle CA, Lexer C (2008) Admixture as the basis for genetic mapping. Cell 23: 686-694.

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. (2013) Stacks: an analysis tool set for population genomics. Molecular Ecology 22: 3124-3140.

Caspermeyer J. (2016). New “Forward Genomics” approach to identify keys to loss of vision in blind mammals. Molecular Biology and Evolution 33: 2175.

Cerqueira MC et al. (2013). Rare or elusive? A test of expert knowledge about rarity of Amazon forest birds. Diversity and Distributions 19: 710-721.

Claridge AW et al. (2004). Use of infrared digital cameras to investigate the behaviour of cryptic species. Wildlife Research 31: 645‑650.

Claridge DE. 2013. Population ecology of Blainville's beaked whales (*Mesoplodon densirostris*) (Doctoral dissertation, University of St Andrews).

Collen B et al. (2016). Clarifying misconceptions of extinction risk assessment with the IUCN Red List. Biology Letters 12: 20150843.

Conn PB, Johnson DS, London JM, Boveng P. (2012). Accounting for missing data when assessing availability in animal population surveys. Methods in Ecology and Evolution 3: 1039–1046.

Constantine R, Carroll E, Stewart R, Neale D. (2014) First record of True's beaked whale *Mesoplodon mirus* in New Zealand. Marine Biodiversity Records 7: e1.

Cornuet JM, Luikart G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144: 2001-2014.

Corrigan S, Huveneers C, Schwartz TS, Harcourt RG, Beheregaray LB. (2008). Genetic and reproductive evidence for two species of ornate wobbegong shark *Orectolobus* spp. on the Australian east coast. Journal Fish Biology 73: 1662-1675.

Cox TM et al. (2006). Understanding the impacts of anthropogenic sound on beaked whales. Journal Cetacean Research Management 7: 177-187.

Cypriano- Souza AL, Meirelles ACO, Carvalho VL, Bonatto SL (2017). Rare or cryptic? The first report of an Omura's whale (*Balaenoptera omurai*) in the South Atlantic Ocean. Marine Mammal Science 33: 80-95.

da Fonseca et al. (2016). Next-generation biology: Sequencing and data analysis approaches for non-model organisms. Marine Genomics 30: 3-13.

Dalebout ML, van Helden A, Van Waerebeek K, Baker CS. (1998). Molecular genetic identification of southern hemisphere beaked whales (Cetacea: Ziphiidae). Molecular Ecology 7: 687-694.

Dalebout et al. (2003). Appearance, distribution, and genetic distinctiveness of Longman's beaked whale, *Indopacetus pacificus*. Marine Mammal Science 19: 421-461.

Dalebout ML, Baker CS, Mead JG, Cockroft VG, Yamada TK. (2004). A comprehensive and validated molecular taxonomy of beaked whales, Family Ziphiidae. Journal of Heredity 65: 459-473.

Dalebout et al. (2005). Worldwide structure of mtDNA diversity among Cuvier's beaked whales (*Ziphius cavirostris*): implications for threatened populations. Molecular Ecology 14: 3353-3371.

Dalebout et al. (2007). A divergent mtDNA lineage among Mesoplodon beaked whales: molecular evidence for a new species in the tropical Pacific? Marine Mammal Science 23: 954-966.

Dalebout ML, Steel D, Baker CS. (2008). Phylogeny of the beaked whale genus Mesoplodon (Ziphiidae: Cetacea) revealed by nuclear introns: Implications for the evolution of male tusks. Systematic Biology 57: 857-875.

Danovaro R, Snelgrove P, Tyler P. (2014). Challenging the paradigms of deep-sea ecology. Trends in Ecology and Evolution 29: 465-475.

Deraniyagala PEP. (1963a). A new beaked whale from Ceylon. Ceylon Today 12: 13-14.

Deraniyagala PEP. (1963b). Mass mortality of the new subspecies of little piked whale *Balaenoptera acutorostrata thalmaha* and a new beaked whale *Mesoplodon hotaula* from Ceylon. Spolia Zelanica 30: 70-84.

Douzery EJ, et al. (2014). OrthoMaM v8: a database of orthologous exons and coding sequences for comparative genomics in mammals. Molecular Biology and Evolution 31: 1923-1928.

Dunn C. (2014). Insights into Blainville's beaked whale (*Mesoplodon densirostris*) communication. (Doctoral dissertation, University of St Andrews).

Eid J. et al. (2009). Real-time DNA sequencing from single polymerase molecules. Science 323: 133-138.

Ellegren H. (2014). Genome sequencing and population genomics in non-model organisms. Trends in Ecology & Evolution 29: 51-63.

Estes JA, et al. (2011). Trophic downgrading of planet Earth. Science 333: 301-306.

Estoup A, Jarne P, Cornuet JM. (2002). Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. Molecular Ecology 11: 1591-1604.

Evans K, Thresher R, Warneke RM, Bradshaw CJA, Pook M, Thiele D, Hindell MA. (2005). Periodic variability in cetacean strandings: links to large-scale climate events. Biology Letters 1: 147–150.

Fedutin I, Filatova O, Mamaev E, Burdin A, Hoyt E. (2014). Occurrence and social structure of Baird's beaked whales, *Berardius bairdii*, in the Commander Islands, Russia. Marine Mammal Science 31: 853-865.

Ferreira IM, Kasuya T, Marsh H, Best PB. (2014). False killer whales (*Pseudorca crassidens*) from Japan and South Africa: Differences in growth and reproduction. Marine Mammal Science 30: 64–84.

Foote AD, Hofreiter M, Morin PA. (2012a). Ancient DNA from marine mammals: Studying long-lived species over ecological and evolutionary timescales. Annals of Anatomy-Anatomischer Anzeiger 194:, 112-120

Foote et al. (2012b). Investigating the potential use of environmental DNA (eDNA) for genetic monitoring of marine mammals. PLoS ONE 7: e41781-6.

Foote AD, et al. (2015). Convergent evolution of marine mammal genomes. Nature Genetics 47: 272-275.

Foote AD, et al. (2016). Genome-culture coevolution promotes rapid divergence of killer whale ecotypes. Nature Communications 7. doi:10.1038/ncomms11693

Ford JKB, (2009). *Killer whales*. In: Encyclopedia of marine mammals. pp 650-657. Perrin, W.F., Würsig, B. & Thewissen, J.G.M. (Eds). San Diego, CA: Academic Press.

Frantzis A. (1998). Does acoustic testing strand whales? Nature 392: 29.

Gardner M, Ward R. (2002). Taxonomic affinities within Australian and New Zealand *Mustelus* sharks (Chondrichthyes: Triakidae) inferred from allozymes, mitochondrial DNA and precaudal vertebrae counts. Copeia 2: 356-363.

Garza JC, Slatkin M, Freimer NB. (1995). Microsatellite allele frequencies in humans and chimpanzees, with implications for constraints on allele size. Molecular Biology and Evolution 12: 594–603.

Gaston KJ. (1994). *Rarity*, 1st edn. Chapman & Hall, London.

Gordon D, et al. (2016). Long-read sequence assembly of the gorilla genome. Science 352: aae0344.

Gowans S, Whitehead H, Hooker SK. (2001). Social organization in northern bottlenose whales, *Hyperoodon ampullatus*: not driven by deep-water foraging? Animal Behaviour 62: 369–377.

Gowans S. (2009). Bottlenose whales: *Hyperoodon ampullatus* and *Hyperoodon planifrons*, In: Encyclopedia of marine mammals. pp 129–131. Perrin, WF, Würsig B & Thewissen JGM. (Eds). San Diego, CA: Academic Press.

Griffiths et al. (2010). Molecular markers reveal spatially segregated cryptic species in a critically endangered fish, the common skate (*Dipterus batis*). Proceedings of the Royal Society B 277: 1497-1503.

Gutenkunst RN, et al. (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. PLoS Genetics 5: e1000695.

Hare MP, et al. (2011). Understanding and estimating effective population size for practical application in marine species management. Conservation Biology 25: 438-449.

Harrison H, Saenz-Agudelo P, Planes S, Jones GP, Berumen ML. (2013). Relative accuracy of three common methods of parentage analysis in natural populations. Molecular Ecology 22: 1158–1170.

Hartley S, Kunin WE. (2003). Scale dependence of rarity, extinction risk, and conservation priority. Conservation Biology 17: 1559–1570.

Hauser L, Baird M, Hilborn R, Seeb LW, Seeb JE. (2011). An empirical comparison of SNPs and microsatellites for parentage and kinship assignment in a wild sockeye

salmon (*Oncorhynchus nerka*) population. Molecular Ecology Resources 11: 150-161.

Henderson EE, Martin SW, Manzano-Roth R, Matsuyama BM. (2016). Occurrence and Habitat Use of Foraging Blainville's Beaked Whales (*Mesoplodon densirostris*) on a US Navy Range in Hawaii. Aquatic Mammals 42: 549.

Heupel MR, et al. (2015). Conservation challenges of sharks with continental scale migrations. Frontiers in Marine Science 2: 12.

Heyning JE, Mead JG. Cuvier's beaked whale: *Ziphius cavirostris*. In: Encyclopedia of marine mammals. pp 294–296. Perrin, W.F., Würsig, B. & Thewissen, J.G.M. (Eds). San Diego, CA: Academic Press.

Hill WG, Robertson A. (1966). The effect of linkage on limits to artificial selection. Genetic Research 8: 269-294.

Hofreiter M, et al. (2015). The future of ancient DNA: Technical advances and conceptual shifts. Bioessays 37: 284-293.

Jarić I, Knežević-Jarić J, Gessner J. (2015). Global effort allocation in marine mammal research indicates geographical, taxonomic and extinction risk-related biases. Mammal Review 45: 54-62.

Jones AG, Small CM, Paczolt KA, Ratterman RL. (2010). A practical guide to methods of parentage analysis. Molecular Ecology Resources 10: 6-30.

Jones OR, Wang J. (2010). COLONY: A program for parentage and sibship inference from multilocus genotype data. Molecular Ecology Resources 10: 551-555.

Kalinowski ST, Taper ML, Marshall TC. (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology 16: 1099-1106.

Kasuya T. (2009). Giant beaked whales: *Berardius bairdii* and *B. arnuxii*. In: Encyclopedia of marine mammals. pp 498–500. Perrin, W.F., Würsig, B. & Thewissen, J.G.M. (Eds). San Diego, CA: Academic Press.

Karczmarski L, Wursig B, Gailey G, Larson KW, Vanderli C. (2005). Spinner dolphins in a remote Hawaiian atoll: social grouping and population structure. Behavioral Ecology 16: 675–685.

Kemppainen P, et al. (2015). Linkage disequilibrium network analysis (LDna) gives a global view of chromosomal inversions, local adaptation and geographic structure. Molecular Ecology Resources 15: 1031-1045.

Kimura M, Ohta, T. (1978). Stepwise mutation model and distribution of allelic frequencies in a finite population. Proceedings of the National Academy of Sciences 75: 2868-2872.

Kraemer HC, Blasey C. (2015). *How many subjects?: Statistical power analysis in research*. Sage Publications, USA. ISBN 978-1-4833-1954-4.

Kitamura S, et al. (2013). Two genetically distinct stocks in Baird's beaked whale (Cetacea: Ziphiidae). Marine Mammal Science 29: 755-766.

Knowlton N. (1986). Cryptic and sibling species among the decapod Crustacea. Journal of Crustacean Biology 6: 356–363

Knowlton N. (1993). Sibling species in the sea. Annual Review of Ecological Systematics 24: 189‑216.

Kousteni V, Megalofonou P. (2016). Observations on the biological traits of the rare shark *Oxynotus centrina* (Chondrichthyes: Oxynotidae) in the Hellenic Seas. Journal of Fish Biology 89: 1880-1888.

Lacsmana JK, Ventolero MF, Blatchley D, Santos MD. (2015). First record of a rare beaked whale *Mesoplodon hotaula* in the Philippines. Marine Biodiversity Records 8: e77. doi:10.1017/S1755267215000457

Li H, Durbin R. (2011). Inference of human population history from individual whole-genome sequences. Nature 475: 493–496.

Liu X, Fu Y-X. (2015). Exploring population size changes using SNP frequency spectra. Nature Genetics 47: 555-559.

Lees AC, Pimm SL. (2015). Species, extinct before we know them? Current Biology 25: R177-R180.

Locke DP, et al. (2011). Comparative and demographic analysis of orang-utan genomes. Nature 469: 529-533.

MacCallum RC, Browne MW, Sugawara HM. (1996). Power analysis and determination of sample size for covariance structure modeling. Psychological Methods 1: 130.

Madsen PT, de Soto NA, Tyack PL, Johnson M. (2014). Beaked whales. Current Biology 24: R728-R730.

Marques T, Thomas L, Ward J, DiMarzio N, Tyack PL. (2009). Estimating cetacean population density using fixed passive acoustic sensors: An example with Blainville's beaked whales. The Journal of the Acoustical Society of America 125: 1982-1994.

Marsh H, Sinclair DF. (1989). Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. Journal of Wildlife Management 53: 1017–1024.

Martín López LM, Miller P, Aguilar de Soto N, Johnson M. (2015). Gait switches in deep-diving beaked whales: biomechanical strategies for long-duration dives. Journal of Experimental Biology 218: 1325-1338.

McCauley et al. (2015). Marine defaunation: Animal loss in the global ocean. Science 349: 247-255.

McClenachan L, Cooper AB, Carpenter KE, Dulvy NK. (2011). Extinction risk and bottleneck in the conservation of charismatic marine species. Conservation Letters 5: 73-80.

McVean GA, Charlesworth B. (2000). The effects of Hill-Robertson interference between weakly selected mutations on patterns of molecular evolution and variation. Genetics 155: 929-944.

Mead JG. (2009a). Beaked whales, Overview: Ziphiidae. In: Encyclopedia of marine mammals. pp 94–97. Perrin, W.F., Würsig, B. & Thewissen, J.G.M. (Eds). San Diego, CA: Academic Press.

Mead JG. (2009b). Shepherd's beaked whale: *Tasmacetus shepherdi*. In: Encyclopedia of marine mammals. pp 1011–1014. Perrin, W.F., Würsig, B. & Thewissen, J.G.M. (Eds). San Diego, CA: Academic Press.

Meirmans PG, Hedrick PW. (2011). Assessing population structure: FST and related measures. Molecular Ecology Resources 11: 5-18.

Miller et al. (2015). First indications that northern bottlenose whales are sensitive to behavioural disturbance from anthropogenic noise. Royal Society Open Science, 2: 140484.

Mirimin L, Banguera-Hinestroza E, Dillane E, Hoelzel AR, Cross TF, Rogan E. (2010). Insights into genetic diversity, parentage, and group composition of Atlantic white-sided dolphins (*Lagenorhynchus acutus*) off the west of Ireland based on nuclear and mitochondrial genetic markers. Journal of Heredity 102: 79–87.

Moore JC. (1963). Recognizing certain species of beaked whales of the Pacific Ocean. The American Midland Naturalist 70: 369-428.

Moore JC, Gilmore RM. (1965). A beaked whale new to the western hemisphere. Nature 205: 1239-1240.

Moore JC. (1968). Relationships among the living genera of beaked whales. Field Zoologist 53: 209–298.

Moore JE, Barlow JP. (2013). Declining abundance of beaked whales (Family Ziphiidae) in the California Current large marine ecosystem. PLoS ONE 8: e52770. doi:10.1371/journal.pone.0052770

Morin PA, et al. (2010). Complete mitochondrial genome phylogeographic analysis of killer whales (*Orcinus orca*) indicates multiple species. Genome Research 20: 908-916.

Morin P, et al. (2017). Genetic structure of the beaked whale genus *Berardius* in the North Pacific, with genetic evidence for a new species. Marine Mammal Science. doi/10.1111/mms.12345/epdf

Morrison DA. (2014). Next generation sequencing and phylogenetic networks. EMBnet. Journal 20: e760.

Narum SR, Buerkle CA, Davey JW, Miller MR, Hohenlohe PA. (2013). Genotyping-by-sequencing in ecological and conservation genomics. Molecular Ecology 22: 2841-2847.

Nielsen R, Mattila DK, Clapham PJ, Palsbøll PJ. (2001). Statistical approaches to paternity analysis in natural populations and applications to the North Atlantic humpback whale. Genetics 157: 1673-1682.

Newsome S, Clementz MT, Koch P. (2010). Using stable isotope biogeochemistry to study marine mammal ecology. Marine Mammal Science 26: 509-572.

Okamura H, Minamikawa S, Skaug HJ, Kishiro T. (2012). Abundance estimation of long-diving animals using line transect methods. Biometrics 68: 504–513.

Oleksyk TK, et al. (2012). A locally funded Puerto Rican parrot (*Amazona vittata*) genome sequencing project increases avian data and advances young researcher education. GigaScience 1: 14.

Ortega-Ortiz JG, Engelhaupt D, Winsor M, Mate BR, Hoelzel AR. (2012). Kinship of long-term associates in the highly social sperm whale. Molecular Ecology 21: 732-744.

O'Reilly PT, Canino MF, Bailey KM, Bentzen P. (2004). Inverse relationship between $F_{st}$ and microsatellite polymorphism in the marine fish, walley pollock (*Theragra chalcogramma*): implications for resolving weak population structure. Molecular Ecology 13: 1799-1914.

Payseur B, Rieseberg LH. (2016). A genomic perspective on hybridization and speciation. Molecular Ecology 25: 2337-2360.

Peltier H, et al. (2013). The stranding anomaly as population indicator: the case of harbour porpoise *Phocoena phocoena* in North-Western Europe. PLoS One *8*: e62180.

Peltier H, et al. (2014). The contribution of stranding data to monitoring and conservation strategies for cetaceans: Developing spatially explicit mortality

indicators for common dolphins (*Delphinus delphis*) in the eastern North-Atlantic. Ecological Indicators 39: 203-214.

Pitman R. (2009). Mesoplodont whales (*Mesoplodon* spp.), In: Encyclopedia of marine mammals. p721–726. Perrin, W.F., Würsig, B. & Thewissen, J.G.M. (Eds). San Diego, CA: Academic Press.

Prudent X, Parra G, Schwede P, Roscito JG, Hiller M. (2016). Controlling for phylogenetic relatedness and evolutionary rates improves the discovery of associations between species' phenotypic and genomic differences. Molecular Biology and Evolution 33: 2135-2150.

Quattro et al. (2006). Genetic evidence of cryptic speciation within hammerhead sharks (Genus *Sphyrna*). Marine Biology 148: 1143-1155.

Quintanilla S, Gómez A, Mariño-Ramírez C, Sorzano C, Bessudo S, Soler G, Bernal J, Caballero S. (2015). Conservation genetics of the scalloped hammerhead shark in the Pacific coast of Colombia. Journal of Heredity 106: 448-458.

Ramirez-Llodra E, et al. (2010). Deep, diverse and definitely different: unique attributes of the world's largest ecosystem. Biogeosciences 7: 2851-2899.

Randhawa HS, Poulin R, Krkošek M. (2015). Increasing rate of species discovery in sharks coincides with sharp population declines: implications for biodiversity. Ecography 38: 96-107.

Ramos R, González-Solís J. (2012). Trace me if you can: the use of intrinsic biogeochemical markers in marine top predators. Frontiers in Ecology and the Environment 10: 258-266.

Reiss H, Hoarau G, Dickey- Collas M, Wolff WJ. (2009). Genetic population structure of marine fish: mismatch between biological and fisheries management units. Fish and Fisheries 10: 361-395.

Robards MD, Reeves RR. (2011). The global extent and character of marine mammal consumption by humans: 1970-2009. Biological Conservation 144: 2770-2786.

Rönnegård L, McFarlane SE, Husby A, Kawakami T, Ellegren H, Qvarnström A. (2016). Increasing the power of genome wide association studies in natural populations using repeated measures–evaluation and implementation. Methods in Ecology and Evolution 7: 792-799.

Ross et al. (2003). DNA Surveillance: Web-based molecular identification of whales, dolphins, and porpoises. Journal of Heredity 94: 111–114.

Schaefer NK, Shapiro B, Green RE. (2016). Detecting hybridization using ancient DNA. Molecular Ecology 25: 2398-2412.

Slatkin M. (1995). A measure of population subdivision based on microsatellite allele frequencies. Genetics 139: 457–462.

Smart JJ, Harry AV, Tobin AJ, Simpfendorfer CA (2013). Overcoming the constraints of low sample sizes to produce age and growth data for rare or threatened sharks. Aquatic Conservation: Marine and Freshwater Ecosystems 23: 124-134.

Smith K, Ruhl H, Bett B, Billet D, Lampitt R, Kaufmann R. (2009). Climate, carbon cycling, and deep-ocean ecosystems. Proceedings of the National Academy of Sciences 106: 19211–19218.

Smith I. (2011). Pre-European Maori exploitation of marine resources in two New Zealand case study areas: species range and temporal change. Journal of the Royal Society of New Zealand 43: 1-37.

Sousa V, Hey J. (2013). Understanding the origin of species with genome-scale data: modelling gene flow. Nature Reviews Genetics 14: 404-414.

Städele V, Vigilant L. (2016). Strategies for determining kinship in wild populations using genetic data. Ecology and Evolution 6: 6107-6120.

Stevens JD et al. (2000). The effects of fishing on sharks, rays, and chimaeras (chondrichthyans), and the implications for marine ecosystems. ICES Journal Marine Science 57: 476 – 494.

Tang H, Coram M, Wang P, Zhu X, Risch N. (2006). Reconstructing genetic ancestry blocks in admixed individuals. The American Journal of Human Genetics 79: 1-12.

Téllez R, Mignucci-Giannoni AA, Caballero S. (2014). Initial description of short-finned pilot whale (*Globicephala macrorhynchus*) genetic diversity from the Caribbean. Biochemical Systematics and Ecology 56: 196–201.

Thompson D. (1910). *Aristotle (c. 350 BC) The History of Animals.* Clarendon Press, London. 231pp.

Thompson et al. (2013). A novel conservation approach provides insights into the management of rare cetaceans. Biological Conservation 157: 331–340.

Tsai YJJ, Mann J. (2013). Dispersal, philopatry, and the role of fission- fusion dynamics in bottlenose dolphins. Marine Mammal Science 29: 261-279.

van Helden et al. (2002). Resurrection of *Mesoplodon traversii* (Gray, 1874), Senior Synonym of *M. bahamondi* (Cetacea: Ziphiidae). Marine Mammal Science 18: 609–621.

Van Oppen MJH, Rico C, Turner GF, Hewitt GM. (2000). Extensive homoplasy, nonstepwise mutations, and shared ancestral polymorphism at a complex microsatellite locus in Lake Malawi cichlids. Molecular Biology and Evolution 17: 489–498.

Veríssimo A, et al. (2016). Molecular diversity and distribution of eastern Atlantic and Mediterranean dogfishes *Squalus* highlight taxonomic issues in the genus. Zoologica Scripta. doi:10.1111/zsc.12224

Viricel A, Strand AE, Rosel PE, Ridoux V, Garcia P. (2008). Insights on common dolphin (*Delphinus delphis*) social organization from genetic analysis of a mass-stranded pod. Behavioural Ecology and Sociobiology 63: 173.

von der Heyden S, et al. (2014). The application of genetics to marine management and conservation: examples from the Indo-Pacific. Bulletin of Marine Science 90: 123-158.

vonHoldt B, Kays R, Pollinger JP, Wayne RK. (2016). Admixture mapping identifies introgressed genomic regions in North American canids. Molecular Ecology 25: 2443-2453.

Vrijenhoek, Robert C. (2009). Cryptic species, phenotypic plasticity, and complex life histories: assessing deep-sea faunal diversity with molecular markers. Deep Sea Research Part II: Topical Studies in Oceanography 56: 1713-1723.

Wallberg A, Glémin A, Webster M. (2015). Extreme recombination frequencies shape genome variation and evolution in the honeybee, *Apis mellifera*. PLoS Genetics 11: e1005189.

Wang J. (2004). Sibship reconstruction from genetic data with typing errors. Genetics 166: 1963–1979.

Wang J. (2010). COANCESTRY: A program for simulating, estimating and analysing relatedness and inbreeding coefficients. Molecular Ecology Resources 2010: 141-145.

Wang J. (2012). Computationally efficient sibship and parentage assignment from multilocus marker data. Genetics 191: 183–194.

Wang J, Santure AW. (2009). Parentage and sibship inference from multilocus genotype data under polygamy. Genetics 181: 1579–1594.

Waples, RS. (1998). Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. Journal of Heredity 89: 438-450.

Waples RS, Gaggiotti O. (2006). What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Molecular Ecology 15: 1419-1439.

Ward RD, Woodwark M, Skibinski DOF. (1994). A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes. Journal of Fish Biology 44: 213–232.

Watts C, Petersen SD, Ferguson SH. (2015). Genetics and fatty acids assist in deciphering narwhal (Monodon monoceros) social groupings. Polar Biology 38: 1971-1981.

Whitehead H. (2002). Estimates of the current global population size and historical trajectory for sperm whales. Marine Ecology Progress Series 242: 295-304.

Whitehead H. (2009). Sperm whale: *Physeter macrocephalus*. In: Encyclopedia of marine mammals. p1091–1097. Perrin, W.F., Würsig, B. & Thewissen, J.G.M. (Eds). San Diego, CA: Academic Press.

Willems T, et al. (2014). The landscape of human STR variation. Genome Research 24: 1894-1904.

Winkler CA, Nelson GW, Smith MW. (2010). Admixture mapping comes of age. Annual Review of Genomics and Human Genetics 11: 65-89.

Wit P, et al. (2012). The simple fool's guide to population genomics via RNA‐Seq: an introduction to high‐throughput sequencing data analysis. Molecular Ecology Resources 12: 1058-1067.

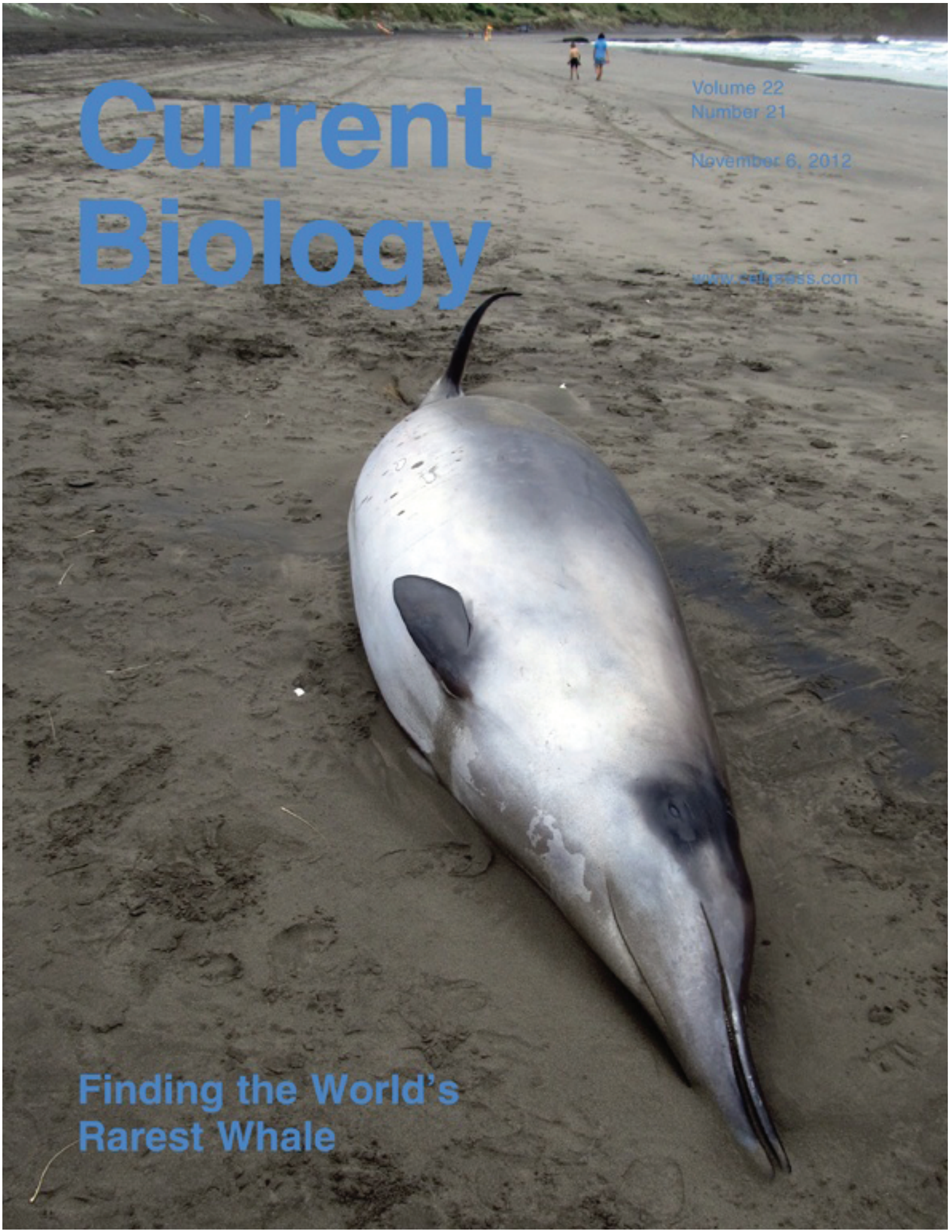
Zhang G, et al. (2014). Comparative genomics reveals insights into avian genome evolution and adaptation. Science 346: 1311-1320.

Zhao S, et al. (2013). Whole-genome sequencing of giant pandas provides insights into demographic history and local adaptation. Nature Genetics 45: 67–71

Zimmer WM, Harwood J, Tyack PL, Johnson MP, Madsen PT. (2008). Passive acoustic detection of deep-diving beaked whales. The Journal of the Acoustical Society of America 124:2823-2832.

# Appendices – Publications I – VIII

Publications are bound in sequential order with respective supplementary information.



# Current Biology

Volume 22
Number 21

November 6, 2012

www.cellpress.com

## Finding the World's Rarest Whale

# The world's rarest whale

**Kirsten Thompson[1,2], C. Scott Baker[1,3], Anton van Helden[4], Selina Patel[1,2], Craig Millar[1,2], and Rochelle Constantine[1,*]**

The vast expanses of the South Pacific Ocean have, until recently, concealed the identity of the world's rarest whale, the spade-toothed beaked whale (*Mesoplodon traversii*). Based on the scarcity of records and the total absence of previous sightings, this species is the least known species of whale and one of the world's rarest living mammals. Two individuals of this species, previously known from only two skull fragments and a mandible, were recently discovered beach-cast in New Zealand. Although initially misidentified, we have used DNA analysis to reveal their true identity. We provide the first morphological description and images of this enigmatic species. This study highlights the importance of DNA typing and reference collections for the identification of rare species.

The South Pacific Ocean represents approximately 85 million km$^2$, covering around 14% of the Earth's surface [1]. This massive and poorly surveyed habitat has some of the deepest ocean trenches. Within this area are many rare deep-water species, including the enigmatic beaked whales (Family Ziphiidae). Very little is known of the life history of these cetaceans and whilst there are 21 recognized species, many are described from only a small number of records [2]. Beaked whales are thought to be exceptionally deep divers, foraging for squid and small fish and spending little time at the surface [3]. Due to similarities in their external morphology, species are very difficult to distinguish and, given their elusive habits, are rarely seen at sea.

New Zealand has an extensive coastline and is a known hotspot for whales stranding. It has one of the highest rates and greatest diversities of stranded cetacean species in the world [4], including 13 species of beaked whale, one of which is the spade-toothed whale [5]. The only previously known specimens of this whale were a single mandible with teeth from an adult male (the holotype), collected from the Chatham Islands, New Zealand in 1872 (Figure 1A), and two skulls without mandibles, one from White Island, collected in the 1950s (Figure 1A, S1) and one from Robinson Crusoe Island, Chile, collected in 1986 (Figure 1A).

In December 2010, a female (5.3 m) and a male (3.5 m) beaked whale stranded and subsequently

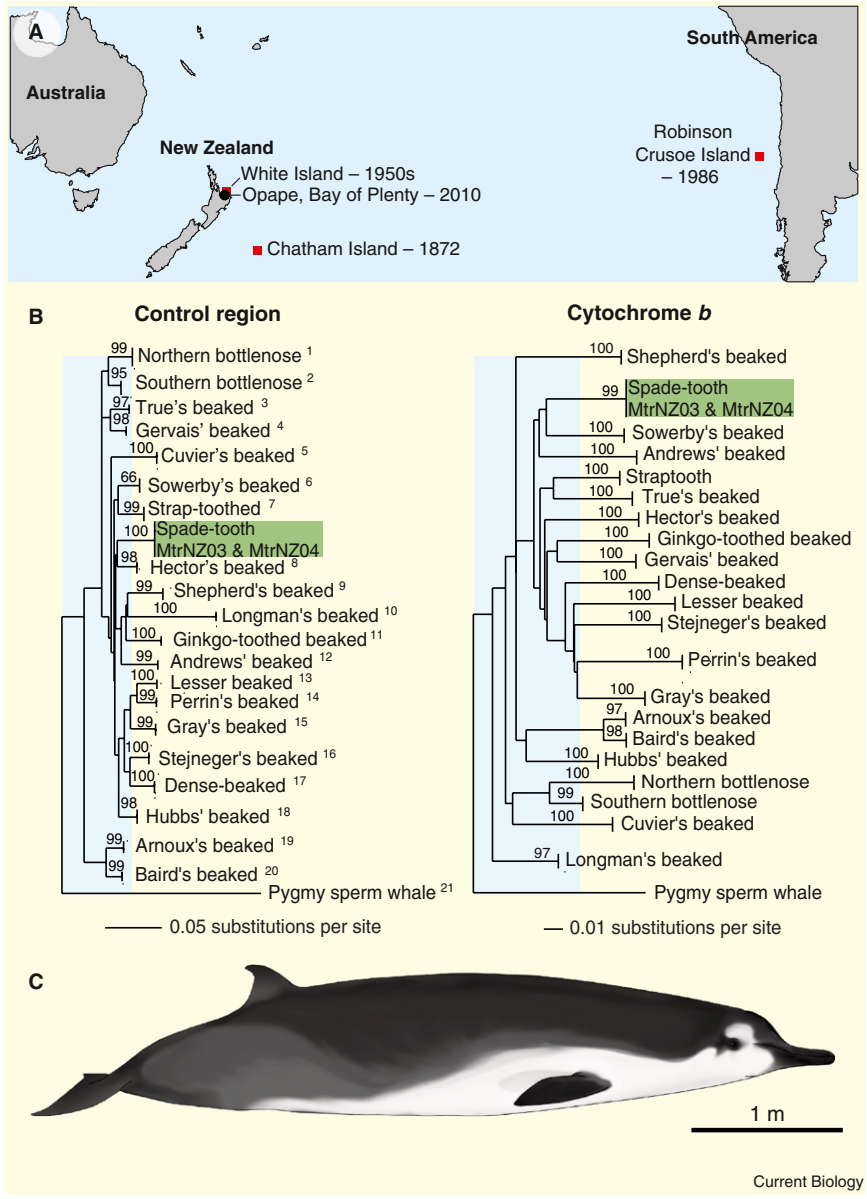



Figure 1. Location, species identification and morphology of the spade-toothed beaked whale (*Mesoplodon traversii*).
(A) Location of partial skulls found on the Chatham and White Islands, New Zealand, on Robinson Crusoe Island, Chile (red squares) and the two recently stranded specimens discovered on Opape Beach, New Zealand (black circle). (B) Neighbour-joining tree of control region and cytochrome *b* sequences from ziphiid species and the two Opape specimens (MtrNZ03 and MtrNZ04) group in a monophyletic clade (green box). Higher-level relationships within the blue shading are not well resolved. Bootstrap values are shown as percentages for a total of 1000 bootstrap replicates. The vertical bars denote several specimens included in the analysis. Scientific names are as follows: 1: *Hyperoodon ampullatus*; 2: *H. planifrons*; 3: *M. mirus*; 4: *M. europaeus*; 5: *Ziphius cavirostris*; 6: *M. bidens*; 7: *M. layardii*; 8: *M. hectori*; 9: *Tasmacetus shepherdi*; 10: *Indopacetus pacificus*; 11: *M. ginkgodens*; 12: *M. bowdoini*; 13: *M. peruvianus*; 14: *M. perrini*; 15: *M. grayi*; 16: *M. stejnegeri*; 17: *M. densirostris*; 18: *M. carlhubbsi*; 19: *Berardius arnuxii*; 20: *B. bairdii*; 21: *Kogia breviceps*. (C) Illustration depicting a generalized external morphology derived from photographs of the adult female spade-toothed beaked whale (see supplemental information for photographs).

died on Opape Beach, New Zealand (38°5′S, 177°17′E). Staff from the New Zealand Department of Conservation photographed the animals after death and collected morphometric measurements and tissue samples.

From their initial description the whales were identified as Gray's beaked whales (*M. grayi*), the most commonly stranding ziphiid around New Zealand. However, to confirm their identity, we amplified and sequenced two mitochondrial DNA regions (control region and cytochrome *b*) (GenBank JX901028, JX901029). The sex of the specimens was confirmed by amplification of a short region of the male-specific SRY gene with a ZFX/Y gene control found in both males and females.

Surprisingly, a comparison with sequences from the holotype and the other two reference specimens revealed that both whales were not Gray's beaked whales, but rather the previously unseen spade-toothed whale [5]. Control region sequences for both beach-cast specimens matched the holotype sequences (99% pairwise identity, GenBank AF439992), and a neighbour-joining tree provided robust support (100% bootstrap) for a species-specific clade (Figure 1B). This result was verified with a 400-bp region of cytochrome *b*, which showed 100% pairwise identity to the holotype sequence (GenBank AY579555) with 99% bootstrap support (Figure 1B).

It is particularly difficult to distinguish between different species of beaked whales using external morphological characters alone, especially in New Zealand where diversity is high [6]. Traditional descriptions of beaked whale species focus on the position and characteristic shape of the erupted teeth of mature males. However, this is not a useful diagnostic character for females and juveniles as the teeth are not erupted. Species diagnosis using color patterns is also problematic due to their rapid deterioration *post mortem*. Consequently, over the last two decades an increased emphasis has been placed on genetic information to complement morphometric data and photographic records. The results of these collections continue to reveal surprises about this rare and cryptic family of cetaceans.

Photographs show that the adult female spade-toothed whale has a similar mouth-line, dorsal fin and pectoral flipper shape to the Gray's beaked whale. However, it can be distinguished by the following features. The melon is more prominent and more similar to the straptoothed beaked whale (*M. layardii*), and the coloration of the rostrum is dark gray or black, rather than white as in adult Gray's beaked whales. Also notable is a dark eye-patch, the white belly and dark flippers (Figure 1C; Supplemental information). However, the long rostrum of this animal with its dark coloration does not distinguish this species from juvenile Gray's beaked whales (Supplemental information). The second animal has a color pattern characteristic of most juvenile mesoplodont beaked whales, making species identification in the field more difficult.

Based on its scarcity, only two intact animals having been seen in the last 140 years, the spade-toothed whale is the world's rarest whale. Once the stranded animals were genetically identified, the skeletal remains were exhumed and taken to the Museum of New Zealand Te Papa Tongarewa for further morphological analysis. The discovery of these specimens highlights the importance of DNA identification in conjunction with the collection of specimens and photographs from beach-cast animals. New Zealand is unique in that it has developed a co-ordinated national response to cetacean strandings. The public notifies the Department of Conservation, and after rescue attempts, stranding information is recorded by the Department of Conservation and the Museum of New Zealand Te Papa Tongarewa, and tissue samples are retained in the New Zealand Cetacean Tissue Archive [7]. This long-term strategy operates in collaboration with indigenous peoples and university scientists and has resulted in the accumulation of 20 years of specimens and records on rare species.

Rapid advances in DNA technology are having a profound effect on our understanding of the natural world and have added value to museum and other reference collections. This has been particularly important in the field of conservation biology, where losses in biodiversity are increasingly evident [8]. We can now confirm that the spade-toothed whale is extant (Supplemental information), and for the first time we have a description of the world's rarest and perhaps most enigmatic marine mammal.

## Supplemental Information

Supplemental Information including experimental procedures and two figures can be found with this article online at http://dx.doi.org/10.1016/j.cub.2012.08.055.

## Acknowledgments

We thank Whakatohea Iwi Māori Trust, Ngai Tama Haua hapu, H. Barsdell, P. Livingston, K. Chamberlain and other staff of the Department of Conservation, Opotiki, D. Steel and M. Dalebout for NZCeTA curation and management. The research is supported by the University of Auckland, G. Mason Charitable Trust Scholarship and an OMV New Zealand Ltd Scholarship.

## References

1. Eakins, B.W., and Sharman, G.F. (2010). Volumes of the World's Oceans from ETOPO1, NOAA National Geophysical Data Center, Boulder, CO, USA.
2. Dalebout, M.L., Baker, C.S., Cockroft, V.G., Mead, J.G., and Yamada, T.K. (2004). A comprehensive and validated molecular taxonomy of beaked whales, Family Ziphiidae. J. Hered. *95*, 459–473.
3. MacLeod, C.D., Santos, M., and Pierce, G. (2003). Review of data on diets of beaked whales: evidence of niche separation and geographic segregation. J. Mar. Biol. Assoc. UK *83*, 651–665.
4. Baker, C.S., Chilvers, B.L., Constantine, R., Du Fresne, S., Mattlin, R.H., van Helden, A., and Hitchmough, R. (2010). Conservation status of New Zealand marine mammals (suborders Cetacea and Pinnipedia). N.Z. J. Mar. and Freshwater Res. *44*, 101–115.
5. van Helden, A.L., Baker, A.N., Dalebout, M.L., Reyes J.C., Van Waerebeek, K., and Baker, C.S. (2002). Resurrection of *Mesoplodon traversii* (Gray, 1874), senior synonym of *M. bahamondi* Reyes, Van Waerebeek, Cárdenas and Yañez, 1995 (Cetacea: Ziphiidae). Mar. Mamm. Sci. *18*, 609–621.
6. MacLeod, C.D., and Mitchell, G. (2006). Key areas for beaked whales worldwide. J. Cetacean Res. Manage. *7*, 309–322.
7. Thompson, K.F., Millar, C.D., Baker, C.S., Dalebout, M., Steel, D., van Helden, A., and Constantine, R. (2012). A novel conservation approach provides insights into the management of rare cetaceans. Biol. Conserv. 10.1016/j.biocon.2012.07.017.
8. Sala, O.E. *et al*. (2000). Global biodiversity scenarios for the year 2100. Science *287*, 1770–1774.

[1]School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand. [2]Allan Wilson Centre for Molecular Ecology and Evolution, School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand. [3]Department of Fisheries and Wildlife and Marine Mammal Institute, Oregon State University, OR 97365, USA. [4]Museum of New Zealand Te Papa Tongarewa, Te Aro, Wellington 6011, New Zealand.
*E-mail: r.constantine@auckland.ac.nz

## Supplemental Information

# The World's Rarest Whale

Kirsten Thompson, C. Scott Baker, Anton van Helden, Selina Patel, Craig Millar, Rochelle Constantine

### Supplemental Experimental Procedures

Total genomic DNA was isolated from tissue using proteinase K digestion followed by a standard phenol: chloroform: isoamyl extraction as described by Sambrook et al. [S1] and modified in Baker et al. [S2]. Fragments of the mitochondrial control region (~750 bp) and the cytochrome *b* gene (~400 bp) were amplified via the polymerase chain reaction (PCR) using primer pairs M13Dlp1.5 [S3] and Dlp8G [S4], and Gludg-L and CB2-L [S5]. Amplification conditions were as follows: 10–40 ng template DNA, 2.5 mM MgCl2, 0.4 μM forward and reverse primers and 1 mg/ml BSA (New England Biosystems). PCR conditions were as described by Palumbi [S6]. PCR products were purified using the Agencourt AMPure XP PCR purification system. Cycle sequencing was carried out using standard BigDye protocol (Applied Biosystems Inc.). The products were purified using CleanSEQ protocol (Agencourt) and run on an ABI 3130 Genetic Analyzer (Applied Biosystems Inc.). Samples were sequenced in both directions and consensus sequences edited manually using the program Geneious [S7]. Species identifications were determined using the web-based alignment program www.DNA-Surveillance [S8], which includes the sequences from the holotype used in the taxonomic revision of *M. traversii* as well as the other two known specimens [S3].

In order to confirm the sex of the specimens, a 97-bp fragment of the male SRY gene was amplified using primers EsryF and EsryR [S9]. In addition, a 72-bp control fragment of the ZFX/Y gene known to be present in both males and females was amplified using the primer pair: P23-3EZ [S10] and CetZFXF4 5'-CTAAGCATAGTAAAGAGATGCCATT- 3'. This new primer was specifically designed for this study to allow the identification of sex from modern or ancient samples where the DNA is considerably degraded. The test was shown to correctly identify the sex of two samples from three other species: Gray's beaked whale (*M.grayi*); straptoothed whale (*M.layardii*) and Cuvier's beaked whale (*Ziphius cavirostris*).

The SRY amplification conditions were as follows: 10 - 40 ng template DNA, 2.5 mM $MgCl_2$, 0.8 μM forward and reverse primers and 1 mg/ml BSA. PCR profile followed 5 mins at 94 °C with 40 cycles of 45 sec at 94 °C, 1 min at 60 °C and 1.5 mins at 72 °C with 7 mins at 72 °C extension phase. The ZFX/Y amplification conditions were as follows: 2.5 mM $MgCl_2$, 0.6 μM forward and reverse primers and 1mg/ml BSA with the PCR profile following a 5 min denaturing step at 94 °C, with 10 cycles of 45 sec at 94 °C, 1 min at 56 °C and 1.5 mins at 72 °C, followed by 30 cycles of 45 sec at 94 °C, 1 min at 50 °C and 7 min extension at 72 °C. Amplified products were detected by gel electrophoresis in 1 X tris-borate EDTA buffer using a 2% agarose gel stained with ethidium bromide and visualized under UV light. Each sample was amplified six times alongside controls of known sex.

The putative male (MtrNZ04) in all six cases amplified both the SRY and ZFX/Y fragments. In contrast the SRY fragment was not amplified in any of the six tests in the putative female while the ZFX fragment appeared in 5 of the 6 tests. These tests conclusively demonstrate that sample MtrNZ03 is a female and MtrNZ04 is a male.

**Supplemental References**


S1. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, Volume 1 (New York: Cold Spring Harbor Laboratory Press).

S2. Baker, C.S., Slade, R.W., Bannister, R.B., Abernethy, R.B., Weinrich, M.T., Lien, J., Urban, R.J., Calambokidis, J., Vasquez, O., and Palumbi, S.R. (1994). Hierarchical structure of mitochondrial DNA gene flow among humpback whales, *Megaptera novaeangliae*, world-wide. Mol. Ecol. *3*, 313–327.

S3. Baker, C.S., Medrano Gonzalez, L., Calambokidis, J., Perry, A., Pichler, F., Rosenbaum, H., Straley, J., Urban Ramirez, J., Yamaguchi, M., and Von Ziegesar, O. (1998). Population structure of nuclear and mitochondrial DNA variation among humpback whales in the North Pacific. Mol. Ecol. *7*, 695–707.

S4. Lento, G.M., Patenaude, N.J., and Baker, C.S. (1997). Molecular genetic identification of whale and dolphin products for sale in Japan and Korea, 1995-97. Report to the Scientific Committee of the International Whaling Commission. (SC/49/O21).

S5. Dalebout, M.L., van Helden, A., Van Waerebeek, K., and Baker, C.S. (1998). Molecular identification of southern hemisphere beaked whales (Cetacea: Ziphiidae). Mol. Ecol. *7*, 687–694.

S6. Palumbi, S.R. (1996). Nucleic Acids II: The Polymerase Chain Reaction. In Molecular Systematics, D.M. Hillis, C. Moritz, and Mable, B., eds. (Sunderland, MA: Sinauer Associates), pp. 205–247.

S7. Drummond, A.J., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones-Havas, S., and Wilson, A. (2009). Geneious v4.8.Available from http://www.geneious.com

S8. Ross, H.A., Lento, G.M., Dalebout, M.L., Goode, M., Ewing, G., McLaren, P., Rodrigo, A.G., Lavery, S., and Baker, C.S. (2003). DNA Surveillance: Web-based molecular identification of whales, dolphins, and porpoises. J. Hered. *94*, 111–114.

S9. Gupta, S.K., Thangaraj, K., and Singh, L. (2006). A simple and inexpensive molecular method for sexing and identification of the forensic samples of elephant origin. J. Forensic Sci. *5,*: 805-807.

S10. Aasen, E., and Medrano, J.F. (1990). Amplification of the ZFX and ZFX genes for sex identification in humans, cattle, sheep and goats. Biotech. *8,* 1279-1281.

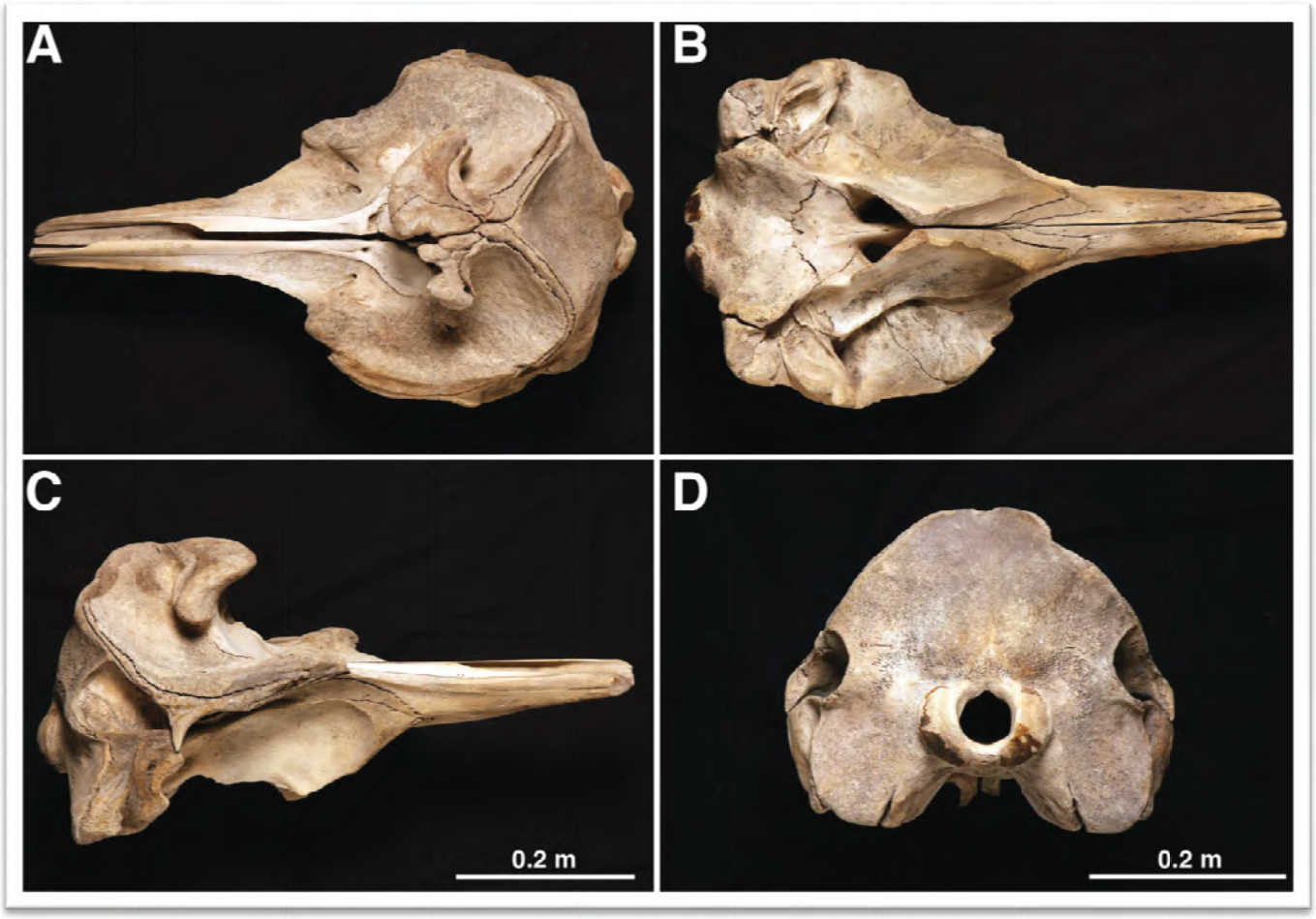



Figure S1. Spade-toothed beaked whale skull without mandible found on White Island, New Zealand in the 1950s, held in The McGregor Museum, University of Auckland. (A) Dorsal, (B) Ventral, (C) Lateral and (D) Posterior views.

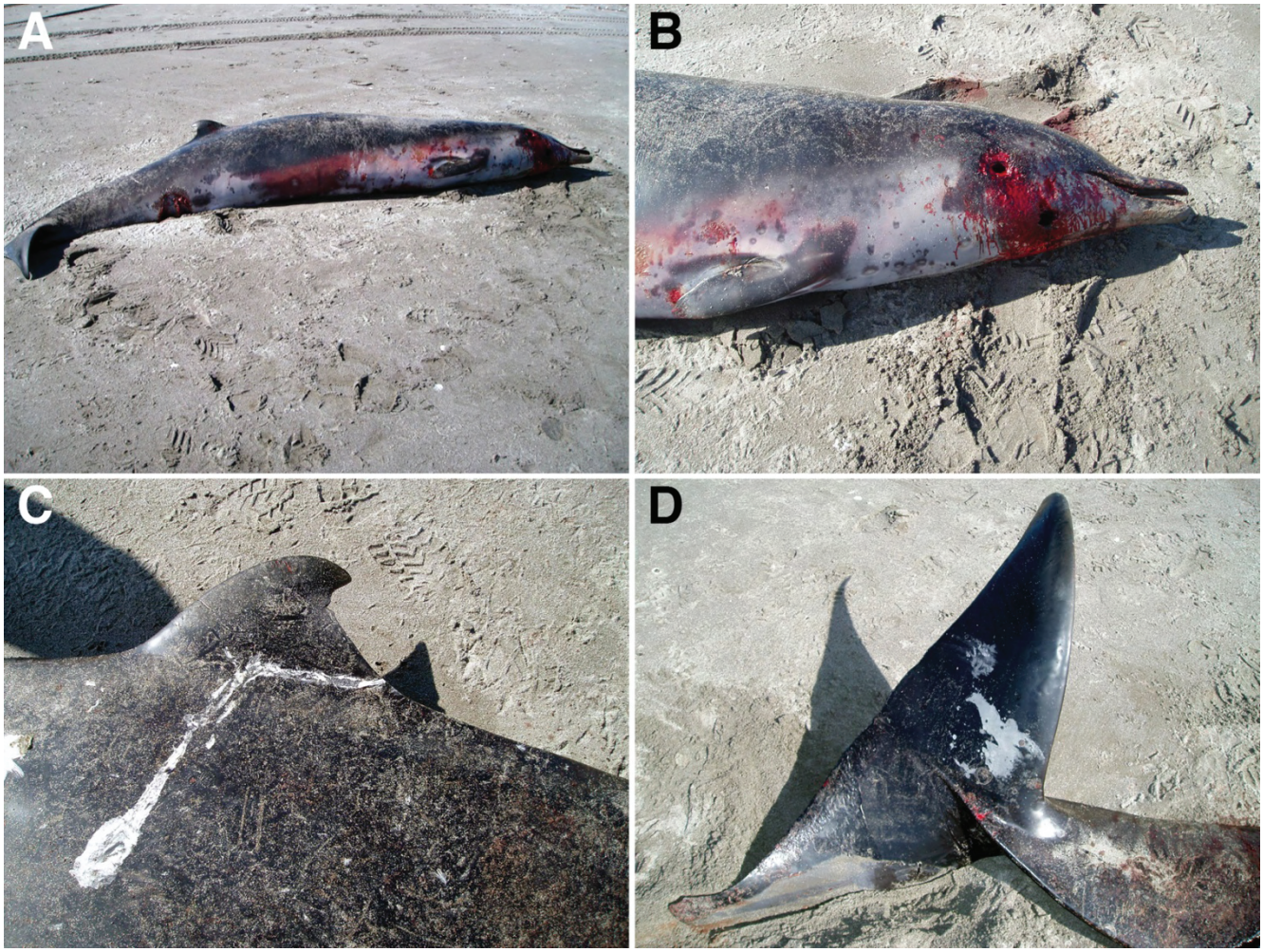



Figure S2. Photographs of the Opape specimens found in 2010. (A) Adult female, (B) jaw-line and right pectoral flipper of adult female, (C) dorsal fin of male, (D) tail fluke of male (Photo credits: New Zealand Department of Conservation).

Animal Conservation

Animal Conservation. Print ISSN 1367-9430

# Species identity and human consumption of beaked whales in the Gilbert Islands, Republic of Kiribati

C. S. Baker[1,2], A. Hutt[3], K. Thompson[2], M. L. Dalebout[4], J. Robins[5], R. L. Brownell Jr[6] & G. S. Stone[7]

1 Marine Mammal Institute, Oregon State University, Newport, OR, USA
2 School of Biological Sciences, University of Auckland, Auckland, New Zealand
3 Department of Conservation, Akaroa, New Zealand
4 School of Biological, Earth, and Environmental Sciences, University of New South Wales, Sydney, NSW, Australia
5 Department of Anthropology, University of Auckland, Auckland, New Zealand
6 NOAA Fisheries, Southwest Fisheries Science Center, Pacific Grove, CA, USA
7 Conservation International, Arlington, VA, USA

**Keywords**
molecular taxonomy; DNA surveillance; biodiversity; subsistence; marine bushmeat.

**Correspondence**
C. Scott Baker, Marine Mammal Institute, Oregon State University, Newport, Oregon 97365, USA. Tel: +1 541-867-0255; Fax: +1 541-867-0345
Email: scott.baker@oregonstate.edu

Received 24 October 2012; accepted 18 February 2013

doi:10.1111/acv.12039

## Abstract

We investigated the species identity and local use of cetaceans on the Gilbert Islands, Republic of Kiribati. Working with the Kiribati Ministry of Environment, Lands and Agricultural Development and Fisheries Division, we visited the islands of Tarawa, Tabiteuea (North), Butaritari and Onotoa from June to July 2009, and collected 24 bones, bone fragments or teeth attributed to recent strandings. The mitochondrial DNA control region or cytochrome *b* was successfully amplified from 12 bones or bone fragments and used to identify four species: *Mesoplodon* sp. representing a new species or subspecies of beaked whale, the dense-beaked whale *Mesoplodon densirostris*, Cuvier's beaked whale *Ziphius cavirostris* and the pygmy sperm whale *Kogia breviceps*. This is the first confirmed identification of the dense-beaked, Cuvier's and pygmy sperm whales from the Gilbert Islands. All specimens were reportedly used for human consumption.

## Introduction

The Gilbert Islands are a north to south chain of 16 atolls and coral islands distributed across the equator in the central Pacific Ocean. Along with the largely uninhabited Line Islands and Phoenix Islands, the Gilbert Islands and Banaba (i.e. Ocean Island) form the Republic of Kiribati, with its capital on Tarawa (in the Gilberts). Although the land area of Kiribati is only 726 km$^2$, its exclusive economic zone (EEZ) is over 3.4 million km$^2$, an area slightly larger than the total land area of India (Lanteigne, 2012). Despite this vast EEZ, little is known about the diversity or exploitation of cetaceans in the waters of Kiribati. The region around the Gilbert Islands (formerly known as the Kingsmill Group) was recognized by British and American whalers for its abundance of sperm whales *Physeter macrocephalus* (Beale, 1839; Smith *et al*., 2012) and 'whales' were hunted by locals during the first half of the 19th century (Wilkes, 1845). The hunting of 'porpoise' (actually dolphins or other small cetaceans) for human consumption was reported during the 20th century colonial history of the Gilbert Islands (Turbott, 1949) and the traditional shamanistic practice of 'porpoise calling' (again, dolphins or other small cetaceans), suggests a long history of such exploitation (Grimble, 1952). Drive hunting of small cetaceans for human consumption in Kiribati was reported until at least the early 1990s, including melon-head whales *Peponocephala electra* (Brownell, Nowacek & Ralls, 2008) and other unidentified delphinids (Robards & Reeves, 2011), but there are no historical or contemporary records of the exploitation of beaked whales.

Consequently, it came as a surprise when molecular identification of dried whale meat served at a local feast in July 2003 on Tabiteuea Island (southern Gilbert Islands, 1°12′ 14" S, 174°44′ 51" E), provided evidence of a new species (or subspecies) of beaked whale, *Mesoplodon* sp. (Dalebout *et al*., 2007). At the time, members of the local community reported that the meat came from seven 'long whales' driven ashore in the shallow water of the atoll in October 2002 (R. Grace, pers. comm.). Here we report on efforts to investigate exploitation of this unknown species of beaked whales and to better document cetacean diversity in the Gilbert Islands by collecting cetacean artifacts, including dried meat, teeth and bones, during visits to the islands in 2009. Our findings, based on molecular identification of discarded bones and bone fragments, contribute to what is otherwise a dearth of information on cetacean diversity and local use around the Gilbert Islands.

## Methods

### Collection of bones from the Gilbert Islands

Two coauthors (A. H. and G. S.) visited Bairiki, on the Island of Tarawa, in April 2006 and again in June 2008, to arrange permits from the Kiribati Ministry of Environment, Lands and Agricultural Development (MELAD) and logistical support for visits to the outer islands. From 18 June to 14 July 2009, one of us (A. H.), accompanied by an interpreter and a local Biodiversity Officer, visited three of the outer Gilbert Islands: Onotoa Island, Tabiteuea North and Butaritari Island (Fig. 1). On each island, village elders were shown a copy of *Whales and Dolphins of New Zealand and Australia, an Identification Guide* (Baker, 1999) and asked about local use of whales and dolphins and the location of any meat or skeletal remains.

### DNA extraction and species identification

Bones and teeth collected in the Gilbert Islands were exported to New Zealand (with appropriate permits) and stored at the School of Biological Sciences, University of Auckland. The bones were isolated from potential contamination with DNA from other samples of whales and dolphins. The primary laboratory analysis was conducted in a laboratory isolated from cetacean DNA in the School of Biological Sciences, University of Auckland. Each bone was drilled in several areas with a 1.5-mm drill bit as described in Pichler, Dalebout & Baker (2001) and approximately 50 mg of material was collected on sterile foil. Total genomic DNA was isolated from the powdered bone using a Qiagen Blood and Tissue Extraction Kit (Qiagen, Inc., Valencia, CA, USA), following the manufacturer's protocols. Amplification via the polymerase chain reaction (PCR) and sequencing of the mitochondrial (mt) DNA control region (D-loop) and cytochrome *b* gene generally followed protocols described in detail by Dalebout *et al*. (2004). For the control region, a 500-bp fragment was first amplified using the primers M13Dlp1.5 and Dlp5R (Dalebout *et al*., 2004). For most samples (except KI024), this was followed by a nested amplification of a 300-bp fragment using primers M13Dlp1.5 and Dlp4R (Dalebout *et al*., 2004). For amplification of the cytochrome *b* gene, DNA was first concentrated using an Amicon Ultra-0.5 mL centrifugal filter (Millipore, Billerica, MA, USA). Non-nested amplification was then attempted for a 260-bp fragment using the primer pairs, Cyb140-160F, 5′-GATACCTRCACGCAAAYGGG GC-3′ and Cyb305-328R (5′-CACCTCAGAATGATAT TTGTCCTC-3′ (M. L. Dalebout, unpubl. data), or a 190-bp fragment using M13CybMLDF (Dalebout *et al*., 2002) and Cyb196-217R, 5′- AGCCGTAATATAGTCCA CGTCC-3′ (M. L. Dalebout, unpubl. data). PCR products

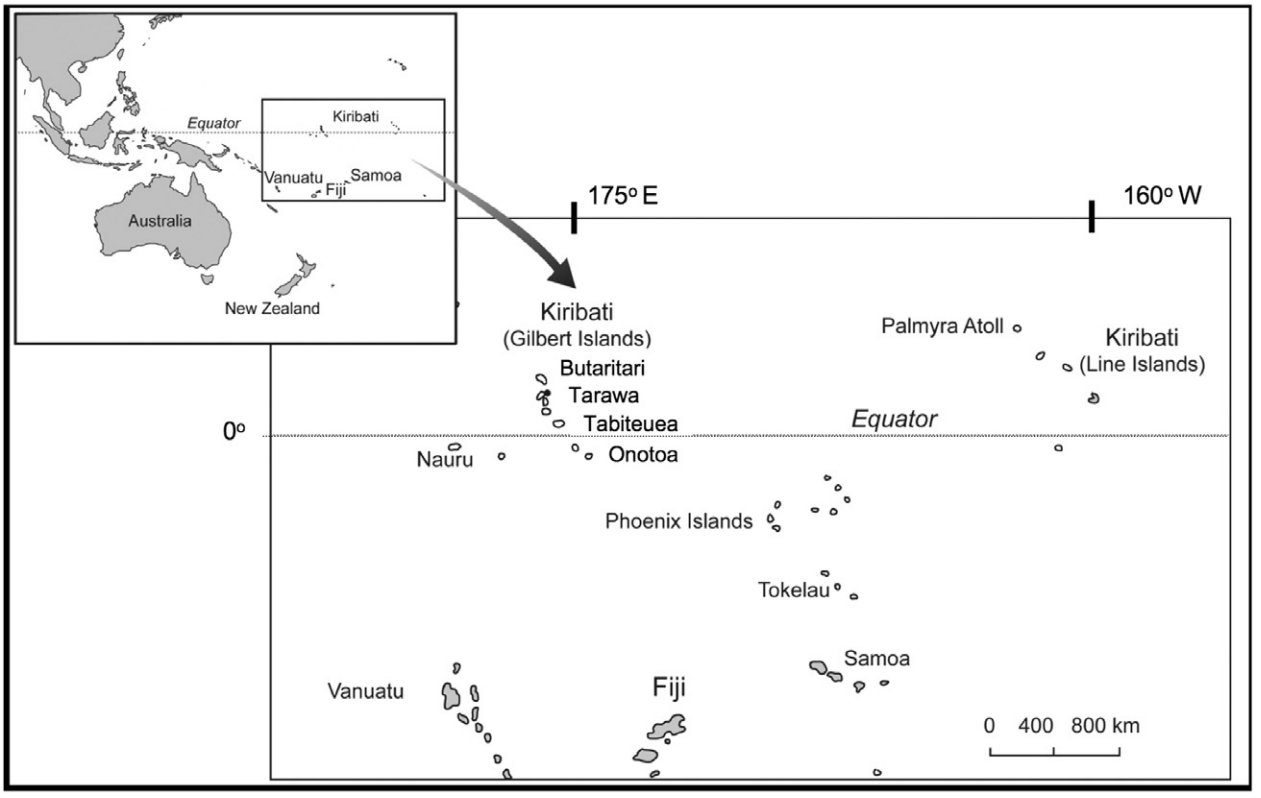



**Figure 1** The location of the Gilbert Islands, Republic of Kiribati, relative to other features of the central Pacific Ocean. Shown are the Island of Tarawa (with the capital city of Bairiki) and the three outer islands (Butaritari, Tabiteuea and Onotoa) surveyed in this study.

were purified using Agencourt AMPure XP and sequenced by standard BigDye 3.1 protocol (Applied Biosystems Inc., Foster City, CA, USA) on an ABI 3130 Genetic Analyzer (Applied Biosystems Inc.). Sequences were edited manually using the program GENEIOUS (Drummond *et al*., 2009).

Based on initial success with sequencing and identification, a subset of six bones was transferred to an ancient DNA facility in the Department of Anthropology, University of Auckland for independent replication of DNA extraction and amplification (by coauthor J. R.). The methods for DNA extraction used for these independent bone extractions were modified from that of Matisoo-Smith *et al*. (1997). The methods for amplification and sequencing follow those described earlier.

The species were identified by submitting control region or cytochrome *b* sequences to the web-based program *DNA-surveillance* (Ross *et al*., 2003), as well as by Basic Local Alignment Search Tool (BLAST) searches of GenBank, and by comparison with the sequences from other specimens of *Mesoplodon* sp., as presented in Dalebout *et al*. (2007) and now available as GenBank JX470545. Species identification was based on the grouping of sequences from the bones with the reference sequences, using a neighbor-joining reconstruction as implemented in *DNA-surveillance* (Ross *et al*., 2003), and on matching with GenBank sequences by BLAST search.

# Results

## Local use and skeletal remains

### Tarawa Island

Attempts to collect information on cetaceans from around Tarawa proved unrewarding – there seemed to be little local or institutional knowledge of whales or dolphins. Of some interest was the whaling try pot in use as a water cistern, presumably an artifact of 19th century whaling contact (see Supporting Information Fig. S1). No cetacean artifacts or skeletal material were collected from Tarawa Island.

### Tabiteuea North Island

Tabiteuea Island has an extensive coastline and discussions with the locals suggested that strandings were common. Unfortunately, the teeth and bones are considered by the islanders to have no value and are often used as fuel for fires. There was no evidence of remains from the seven whales reportedly killed in 2002 and which provided the dried meat served at the feast in July 2003 (Dalebout *et al*., 2007). The islanders described a recent stranding of two beaked whales, about 3–4 m in length, with two teeth about the size of an index finger. When shown a copy of the identification guide (Baker, 1999), the elders could not make a positive species identification, but described the whale as looking like a cross between Andrew's beaked whale *Mesoplodon bowdoini* and Longman's beaked whale *Indopacetus pacificus*. The islanders reported that the teeth were sold to fisherman from a foreign vessel. No bones could be found and no samples were collected from Tabiteuea North Island.

### Butaritari Island

Villagers reported incidental strandings of what were described as 'minke whales' and sperm whales, based on illustrations in the identification guide (Baker, 1999). The teeth and ribs from a sperm whale were displayed by the islanders (see Supporting Information Fig. S1). Islanders also reported the stranding of an unusual whale of approximately 4 m in length, a pointed jaw and a full set of teeth in its lower jaw. Although the stranding had occurred in 2008, the jaw and teeth were still available and were contributed to the collection, along with a sperm whale tooth.

### Onotoa Island

The islanders reported that a whale stranded sometime in the previous two months (*c.* April, 2009) at the northern tip of island. The whale was flensed and the meat was dried for eating. The bones were left on site as they are not considered of practical use by the islanders. The whale was described by an elder who participated in the flensing, as about 4–5 m in length, dark in color, lacking baleen or teeth and with a head like a swordfish. When shown a copy of the identification guide (Baker, 1999) the elder described the whale as looking like a cross between Andrew's beaked whale and Longman's beaked whale, similar to the description by elders on Tabiteuea North Island. A number of vertebrae were collected from the site of the stranding and beachcombing revealed older material, including one partial skull and other vertebrae. Some strips of dried meat and 21 skeletal samples were collected from Onotoa Island: 11 vertebrae, 9 bone fragments, and one definite skull fragment.

## Species identification of bones

In total, the June to July 2009 expedition to the outer islands was able to collect 24 samples, including one partial skull and a partial lower jaw with teeth and some dried skin attached, and one sperm whale tooth (Table 1). The strips of dried meat prepared for human consumption were also collected on Onotoa, but were destroyed by rats during storage in Tarawa and are not included in Table 1. The majority of samples ($n = 21$) were vertebrae and likely skull fragments collected from Onotoa Island.

Of the 24 samples of bone or teeth, 12 yielded DNA of sufficient quality for PCR amplification and sequencing of mtDNA fragments from the control region and/or cytochrome *b* (Table 1). DNA from six of these 12 samples were independently extracted, amplified and sequenced by the collaborating laboratory. Sequences from seven of the 12 samples were of sufficient length to confirm identity of four species with 100% bootstrap support using *DNA-surveillance* (see Supporting Information Fig. S2). These sequences (six control region and two cytochrome *b*) have

**Table 1** Cetacean material collected from the outer islands of the Gilbert group, in the Republic of Kiribati, with species identification based on mtDNA control region (DLP) or CYB sequences

| Sample code | Description | Species identification | Island | DLP | CYB |
|---|---|---|---|---|---|
| KI001 | Fragment | Failed | Onotoa | | |
| KI002* | Fragment | *Mesoplodon densirostris* | Onotoa | **310 bp*** | **219 bp*** |
| KI003 | Fragment | Failed | Onotoa | | |
| KI004 | Vertebra | *Mesoplodon* sp. | Onotoa | | 255 bp |
| KI005 | Vertebra | Failed | Onotoa | | |
| KI006 | Fragment | Failed | Onotoa | | |
| KI007 | Fragment | Failed | Onotoa | | |
| KI008 | Vertebra | *Mesoplodon* sp. | Onotoa | 293 bp | |
| KI009* | Vertebra | *Ziphius cavirostris* | Onotoa | | **236 bp*** |
| KI010 | Vertebra | Failed | Onotoa | | |
| KI011* | Vertebra | *Mesoplodon* sp. | Onotoa | **293 bp*** | 141 bp |
| KI012* | Vertebra | *Mesoplodon* sp. | Onotoa | **293 bp*** | 178 bp |
| KI013 | Vertebra | *Mesoplodon* sp. | Onotoa | 239 bp | |
| KI014 | Fragment – rib? | Failed | Onotoa | | |
| KI015 | Vertebra | *Mesoplodon* sp. | Onotoa | 239 bp | |
| KI016 | Vertebra | Failed | Onotoa | | |
| KI017 | Vertebra | *Mesoplodon* sp. | Onotoa | | 171 bp |
| KI018* | Fragment – skull? | *Mesoplodon densirostris* | Onotoa | **309 bp*** | 109 bp |
| KI019* | Skull fragment | *Mesoplodon densirostris* | Onotoa | **311 bp*** | 136 bp |
| KI020 | Fragment – skull? | Failed | Onotoa | | |
| KI021 | Fragment – scapula? | Failed | Onotoa | | |
| KI022[a] | Tooth – sperm whale | Failed | Butaritari | | |
| KI023 | jaw – teeth | *Kogia breviceps* | Butaritari | | |
| KI024 | jaw – dried skin | *Kogia breviceps* | Butaritari | **487 bp*** | |
| UKIRI[b] | dried meat | *Mesoplodon* sp. | Tabiteuea | **300 bp** | |

[a]The sperm whale tooth was identified from its appearance (see Supporting Information).

[b]The sample UKIRI is represented by a partial fragment of the sequence described in Dalebout *et al.* (2007) and now available as GenBank JX470545.

*Mesoplodon* sp. refers to an unrecognized species or subspecies of beaked whale first identified by Dalebout *et al.* (2007). For samples where DNA extraction or subsequent polymerase chain reaction amplification of mtDNA fragments were not successful, species identification is listed as 'failed'. Sequence lengths in bp are shown for successful DNA identification. An asterisks (*) indicates the six samples for which DNA extraction and amplification were replicated by an independent laboratory (see text). Sequences in bold are included as Supplementary Material and submitted to Dryad.

bp, base pair; DLP, D-loop or control region; CYB, cytochrome *b*; mtDNA, mitochondrial DNA.

been submitted to GenBank as KC540691-KC540698. Sequences from the other five samples were sufficient in length for likely species identification, based on internal matching to the specimens with confident identification. Given the low levels of intra-specific diversity in many beaked whales (Dalebout *et al.*, 2004) and the quality of the sequence fragments, there was not sufficient evidence to determine if the bones represented more than one individual of each species.

Sequences from seven of the vertebrae showed an exact (or near exact) match for the control region and/or cytochrome *b* sequences, with the dried whale meat collected on Tabiteuea Island in 2003 (sample code: UKIRI, Dalebout *et al.*, 2007). Of these seven vertebrae, two yielded control region sequences of sufficient quality for confident species identification using *DNA-surveillance* (see Supporting Information Fig. S2a). Although recent investigation has matched the mtDNA sequences of the samples from Tabiteuea Island with the holotype of *Mesplodon hotaula* (see Discussion), we refer to this taxon here as *Mesoplodon* sp., pending a comprehensive review and proposal for revised taxonomic ranking.

Three of the bone fragments (including a very dense fragment thought to be part of a skull) were identified as a dense-beaked whale *Mesoplodon densirostris*. All three of these provided confident identification based on control region or cytochrome *b* sequences. A single vertebra was identified as a Cuvier's beaked whale *Ziphius cavirostris*, based on a cytochrome *b* sequence (see Supporting Information Fig. S2b). The partial lower jaw from Butaritari was confirmed to be from a pygmy sperm whale *Kogia breviceps* based on control region sequences from both teeth and dried skin. We were unsuccessful in extracting DNA from the large tooth from Butaritari, but, based on size and shape, it is clear that it represents a sperm whale.

# Discussion

## Molecular taxonomy of beaked whales

The beaked whales (Family: Ziphiidae) are among the least known of all vertebrate groups. Found in deep oceanic waters, typically far from shore, several species were initially described only from partial skeletal remains found stranded on remote islands and shorelines. Previously thought to include 20 extant species, a recent survey of the molecular systematics of this family resulted in the discovery of a new species, *Mesoplodon perrini*, first identified by phylogenetic analysis of mtDNA sequences (Dalebout *et al*., 2002). Development of a comprehensive molecular taxonomy of the family Ziphiidae resulted in two additional molecular discoveries (Dalebout *et al*., 2004). First, previously misidentified specimens from South Africa were linked to the holotype of Longman's beaked whale from tropical north Queensland, Australia, providing the first evidence of the physical appearance of this species (Dalebout *et al*., 2003). Second, a single tooth and partial skull held in New Zealand institutions were linked to the holotype of Bahamonde's beaked whale *Mesoplodon bahamondi* Reyes, van Waerebeek, Cárdenas and Yáñez, 1995 held in the Museo Nacional de Historia Natural in Santiago, Chile. This finding confirmed the genetic distinctiveness of the species and uncovered a prior description of this species as *Mesoplodon traversii*, dating back to the recovery of a tooth from the Chatham Islands of New Zealand in the 1870s (van Helden *et al*., 2002). Recently, a female and juvenile beaked whale found stranded in New Zealand were identified from mtDNA as *M. traversii*, providing the first evidence of the physical appearance of this species (Thompson *et al*., 2012a). Together, these discoveries and taxonomic revisions demonstrate the power of integrating information from traditional museum specimens and morphological descriptions of stranded specimens with molecular taxonomy for identifying rare and cryptic species of cetaceans, particularly beaked whales (Baker *et al*., 2003; Dalebout *et al*., 2004; Thompson *et al*., 2012b).

Molecular taxonomy also provides a powerful tool for monitoring species exploitation (Baker, 2008), and through market surveys, of documenting species diversity (Baker *et al*., 2006). This can be particularly important where traditional use of cetaceans, including consumption, is increasing (Robards & Reeves, 2011), or where cetacean species, once taken only as 'bycatch', have become targets of artisanal fisheries (Clapham & Van Waerebeek, 2007). However, because of their pelagic habitat and relatively low abundance, beaked whales have seldom been the target of commercial or traditional exploitation (Mead, 2009).

Here we used molecular taxonomy, as implemented in the web-based program *DNA-surveillance* (Ross *et al*., 2003), to confirm the genetic distinctiveness and document local use of a new species or subspecies of beaked whales, first described from dried strips of meat served at a village feast on Tabiteuea Island in 2003 (Dalebout *et al*., 2007). At the time, analyses of mtDNA control region and cytochrome *b* sequences showed the meat came from an unknown member of the beaked whale genus *Mesoplodon*. Subsequent collaborative comparisons showed an exact match of the Tabiteuea sequences with two specimens collected in 2005 on Palmyra Atoll Wildlife Refuge, 2600 km to the northeast (Dalebout *et al*., 2007). Although grouping most closely with *Mesoplodon ginkodens* in phylogenetic reconstructions, the sequences, now including those reported here from Onotoa Island (see Supporting Information Fig. S2a), show a degree of genetic divergence consistent with species-level classifications among mesoplodons (Dalebout *et al*., 2007). Further investigation (Dalebout *et al*., 2012) has matched the mtDNA of these specimens with sequences from the holotype specimen of *M. hotaula*, initially described from a single specimen held at the National Museum, Colombo, Sri Lanka (Deraniyagala, 1963). Shortly after this description, however, *M. hotaula* was synonymized with *M. ginkgodens* (Moore & Gilmore, 1965), which had been described just a few years earlier (Nishiwaki & Kamiya, 1958). With the addition of these new specimens, the ranking of *M. hotaula* as a species or subspecies requires reconsideration, but remains undecided, pending further morphological and genetic comparisons with the small number of existing *M. ginkgodens* specimens (Dalebout *et al*., 2012).

## Local use of cetaceans by Gilbertese

We encountered no obvious evidence of the shamanistic practice of 'porpoise calling' documented by Sir Author Grimble during his experiences in the Gilbert Islands prior to World War II (Grimble, 1952). However, on each of the three islands of the Gilbert group, discussions with the elders (through translators with MELAD) confirmed the ongoing local use of cetaceans, including as food for human consumption. Descriptions offered by islanders suggested that cetaceans taken for human consumption are found stranded or are driven ashore if they enter the shallows near the atolls. At the time the meat of the *Mesoplodon* sp. was offered to the visitors of Tabiteuea Island in July 2003, villagers reported that they regularly (several times a year) chased and killed 'long whales' when they come into the shallow waters of the nearby lagoon (R. Grace, pers. comm.). To our knowledge, the drive hunting of beaked whales has never before been reported anywhere in the world. Further investigation is required to document the circumstances under which the beaked whales enter the shallow water of the atolls, and the methods used by the Gilbertese for herding the whales.

Historically, the hunting of small cetaceans was important to island and atoll societies throughout the Pacific Islands, representing the cultural divisions of Micronesia, Polynesia and Melanesia. Some islands with a documented history of hunting cetaceans include: the Mariana Islands (Costenoble, 1905), the Gilbert Islands (Grimble, 1952), Woleai Atoll in Federated States of Micronesia (Alkire, 1968), the Hawaiian Islands (Wilkes, 1845), the Marquesas Islands in French Polynesia (Reeves *et al*., 1999) and the Solomon Islands (Dawbin, 1966). Aside from food, the

teeth and bone from sperm whales and small cetaceans were an important component of traditional jewelry, often in the form of necklaces. These adornments also function to show social status such as rank and leadership (Neich & Pereira, 2004). Unfortunately, there appeared to be little or no cultural use of either the bones or teeth of cetaceans in the islands we visited and so few artifacts were available for confirming the full extent of species involved in subsistence use. In some villages, the bones are burned for fuel and, if teeth are recovered, these are sometimes sold to fishermen from passing foreign vessels.

### Species diversity and threats

Our surveys and species identifications provide new, validated records of cetaceans around the remote Gilbert Islands in the Republic of Kiribati. Other than the recent report of *Mesoplodon* sp. (Dalebout *et al*., 2007) and the historical accounts of sperm whales (Townsend, 1935), there are few confirmed records of cetacean species in the Gilbert Islands (Reeves *et al*., 1999). None of the species that we identified were listed in a recent summary of cetaceans used for human consumption in Kiribati (Robards & Reeves, 2011). Our finding of further material from *Mesoplodon* sp. on Onotoa suggests that, like Palmyra Atoll in the Line Islands (Dalebout *et al*., 2007, 2012), the Gilbert Islands are likely to be a hotspot for this species of beaked whale, and raises concerns about its frequency of stranding or potential exploitation for local consumption. The reported large size of the group killed on Tabiteuea and the apparently regular occurrence near small islands (including Palmyra Atoll, Baumann-Pickering *et al*., 2010) suggest an unusual social organization and habitat preference for this previously undescribed species or subspecies.

## Acknowledgements

We thank the many members of the Kiribati Ministry of Environment, Lands and Agricultural Development (MELAD) and the Fisheries Division for their assistance with the surveys of the Gilbert Islands. We also thank Sue Taei, Michael Donoghue, and Rochelle Constantine for their interest and assistance with the project; Craig Millar, Shane Lavery, Matt Goddard, and Jo Putteril for access to laboratory facilities; and Patty Rosel for review of an earlier draft of the paper. Roger Grace provided details of his visit to Tabiteuea Island and the whale meat collected in July 2003. This research was supported by Grant #7858-05 to CSB from the National Geographic Society. The final analysis and preparation of the paper was supported in part by a Pew Marine Conservation Fellowship to CSB for study of '*A Pattern of Dolphins*' in Oceania.

## References

Alkire, W.H. (1968). Porpoises and taro. *Ethnology* **7**, 280–289.

Baker, A.N. (1999). *Whales and dolphins of New Zealand & Australia*. 3rd edn. Wellington: Victoria University Press.

Baker, C.S. (2008). Invited review: a truer measure of the market: the molecular ecology of fisheries and wildlife trade. *Mol. Ecol.* **17**, 3985–3998.

Baker, C.S., Dalebout, M.L., Lavery, S. & Ross, H.A. (2003). Applied molecular taxonomy for species conservation and discovery. *Trends Ecol. Evol.* **18**, 271–272. www.dna-surveillance.auckland.ac.nz.

Baker, C.S., Lukoschek, V., Lavery, S., Dalebout, M.L., Yong-Un, M., Endo, T. & Funahashi, N. (2006). Incomplete reporting of whale, dolphin and porpoise 'bycatch' revealed by molecular monitoring of Korean markets. *Anim. Conserv.* **9**, 474–482.

Baumann-Pickering, S., Wiggins, S.M., Roth, E.H., Roch, M.A., Schnitzler, H.-U. & Hildebrand, J.A. (2010). Echolocation signals of a beaked whale at Palmyra Atoll. *J. Acoust. Soc. Am.* **127**, 3790–3799.

Beale, T. (1839). *The natural history of the sperm whale*. London: John van Voorst.

Brownell Jr, R.L., Nowacek, D.P. & Ralls, K. (2008). Hunting cetaceans with sound: a worldwide review. *J. Cetacean Res. Manage.* **10**, 81–88.

Clapham, P.J. & Van Waerebeek, K. (2007). Bushmeat and bycatch: the sum of the parts. *Mol. Ecol.* **16**, 2607–2609.

Costenoble, H.H.L.W. (1905). Die Marianen [The Marianas]. *Globus: illustrierte Zeitschrift für Länder und Völkerkunde* **88**, 4–9, 72–81, 92–94.

Dalebout, M.L., Mead, J.G., Baker, C.S., Baker, A.N. & Van-Helden, A.L. (2002). A new species of beaked whale *Mesoplodon perrini* sp. n. (Cetacea: Ziphiidae) discovered through phylogenetic analyses of mitochondrial DNA sequences. *Mar. Mamm. Sci.* **18**, 577–608.

Dalebout, M.L., Baker, C.S., Anderson, R.C., Best, P.B., Cockcroft, V.G., Hinsz, H.L., Peddemors, V. & Pitman, R.L. (2003). Appearance, distribution, and genetic distinctiveness of Longman's beaked whale, *Indopacetus pacificus*. *Mar. Mamm. Sci.* **19**, 421–461.

Dalebout, M.L., Baker, C.S., Cockroft, V.G., Mead, J.G. & Yamada, T.K. (2004). A comprehensive molecular taxonomy of beaked whales (Cetacea: Ziphiidae) using a validated mitochondrial and nuclear DNA database. *J. Hered.* **95**, 459–473.

Dalebout, M.L., Baker, C.S., Steel, D., Robertson, K.M., Chivers, S.J., Perrin, W.F., Mead, J.G., Grace, R.V. & Schofield, D.T. (2007). A divergent mtDNA lineage among *Mesoplodon* beaked whales: molecular evidence for a new species in the tropical Pacific? *Mar. Mamm. Sci.* **23**, 954–966.

Dalebout, M.L., Baker, C.S., Steel, D., Thompson, K., Robertson, K.M., Chivers, S.J., Perrin, W.F., Goonatilake, M., Anderson, R.C., Mead, J.G., Potter, C.W., Yamada, T.K., Thompson, L. & Jupiter, D. (2012). A newly recognised beaked whale (Ziphiidae) in the tropical Indo-Pacific: *Mesoplodon hotaula* or *M. ginkgodens*

*hotaula*. Report (SC/64/SM3) to the Scientific Committee of the International Whaling Commission, available from the Secretariat, International Whaling Commission, The Red House, 135 Station Road, Impington, Cambridge, Cambridgeshire CB24 9NP, UK.
Dawbin, W.H. (1966). Porpoises and porpoise hunting in Malaita. *Aust. Nat. Hist.* **15**, 207–211.
Deraniyagala, P.E.P. (1963). A new beaked whale from Ceylon. *Ceylon Today* **12**, 13–14.
Drummond, A.J., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones-Havas, S. & Wilson, A. (2009). Geneious, v4.8. Available from http://www.geneious.com.
Grimble, A. (1952). *A pattern of islands*. London: John Murray.
Lanteigne, M. (2012). Water dragon? China, power shifts and soft balancing in the South Pacific. *Polit. Sci.* **64**, 21–38.
Matisoo-Smith, E., Allen, J.S., Ladefoged, T.H., Roberts, R.M. & Lambert, D.M. (1997). Ancient DNA from Polynesian rats: extraction, amplification and sequence from single small bones. *Electrophoresis* **18**, 1534–1537.
Mead, J.G. (2009). Beaked whales, overview. In *Encyclopedia of marine mammals*. 2nd edn. 94–97. Perrin, W.F., Würsig, B. & Thewissen, J.G.M. (Eds). New York: Academic Press.
Moore, J.C. & Gilmore, R.M. (1965). A beaked whale new to the Western Hemisphere. *Nature* **205**, 1239–1240.
Neich, R. & Pereira, F. (2004). *Pacific jewelry and adornment*. Honolulu, HI: University of Hawaii Press.
Nishiwaki, M. & Kamiya, T. (1958). A beaked whale *Mesoplodon* stranded at Oiso Beach, Japan. *Sci. Rep. Whales Res. Inst., Tokyo* **13**, 53–83.
Pichler, F.B., Dalebout, M.L. & Baker, C.S. (2001). Nondestructive DNA extraction from sperm whale teeth and scrimshaw. *Mol. Ecol. Notes* **1**, 106–109.
Reeves, R.R., Leatherwood, S., Stone, G.S. & Eldredge, G. (1999). Marine mammals in the area served by the South Pacific Regional Environment Programme (SPREP). Apia, Samoa: South Pacific Regional Environment Programme. 48 pp.
Robards, M.D. & Reeves, R.R. (2011). The global extent and character of marine mammal consumption by humans: 1970–2009. *Biol. Conserv.* **144**, 2770–2786.
Ross, H.A., Lento, G.M., Dalebout, M.L., Goode, M., McLaren, P., Rodrigo, A.G., Lavery, S. & Baker, C.S. (2003). DNA surveillance: web-based molecular identification of whales, dolphins and porpoises. *J. Hered.* **94**, 111–114.
Smith, T.D., Reeves, R.R., Josephson, E.A. & Lund, J.N. (2012). Spatial and seasonal distribution of American whaling and whales in the age of sail. *PLoS ONE* **7**, e34905.
Thompson, K., Baker, C.S., Van Helden, A., Patel, S., Millar, C. & Constantine, R. (2012a). The world's rarest whale. *Curr. Biol.* **22**, R905–R906.
Thompson, K.F., Millar, C.D., Baker, C.S., Dalebout, M., Steel, D., Van-Helden, A.L. & Constantine, R. (2012b). A novel conservation approach provides insights into the management of rare cetaceans. *Biol. Conserv.* **157**, 331–340.
Townsend, C.H. (1935). The distribution of certain whales as shown by logbook records of American whaleships. *Zoologica* **19**, 1–50.
Turbott, I.G. (1949). Diets, Gilbert and Ellice Islands Colony. *J. Polyn. Soc.* **58**, 36–46.
Van Helden, A.L., Baker, A.N., Dalebout, M.L., Reyes, J.C., Waerebeek, K.V. & Baker, C.S. (2002). Resurrection of *Mesoplodon traversii* (Gray, 1874), senior synonym of *M. bahamondi* Reyes, Van Waerebeek, Cárdenas and Yañez, 1995 (Cetacea: Ziphiidae). *Mar. Mamm. Sci.* **18**, 609–621.
Wilkes, C. (1845). *Narrative of the United States Exploring Expedition during the years 1838, 1839, 1840, 1841, 1842*. Philadelphia, PA: Lea and Blanchard.

## Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

**Figure S1** Whaling artifacts and cetacean bones collected during surveys of the Gilbert Islands from 18 June to 14 July 2009: (a) whaling 'try-pot', a remnant of 19th-century whaling contact; (b) vertebra (KI004, identified as *Mesoplodon* sp., an unrecognized species or subspecies) and rib bone (KI014, no DNA identification) collected on Onotoa Island; and (c) mandible, rib and teeth of a sperm whale on display in Butaritari Island. Photographs courtesy of Al Hutt.
**Figure S2** Representative identification of mtDNA sequences from cetacean material collected from the outer islands of the Gilbert group, in the Republic of Kiribati, using reference sequences available through the web-based application, *DNA-surveillance* (Ziphiidae Vs4.3): (a) control region (DLP) of KI0o2, KI011, KI012, KI018 and KI019; (b) cytochrome *b* (CYB) of KI002 and KI009. The neighbor-joining tree reflects species-specific grouping with bootstrap values based on 500 replications. The program is available at http://www.cebl.auckland.ac.nz:9000/page/whales/title.
**Table S1** Sequences of the mtDNA control region (DLP) used in species identification of bones from Gilbert Islands. Shown in FASTA format with sample code followed by primer information
**Table S2** Sequences of the mtDNA cytochrome *b* used in species identification of bones from Gilbert Islands. Shown in FASTA format with sample code followed by primer information

## Supplementary Material

Title: "Species identity and human consumption of beaked whales in the Gilbert Islands, Republic of Kiribati"

Authors: C. Scott Baker, Al Hutt, Kirsten Thompson, Merel L. Dalebout, Judith Robins, Robert L. Brownell Jr. and Greg S. Stone

**SupTable 1**: Sequences of the mtDNA control region (DLP) used in species identification of bones from Gilbert Islands. Shown in FASTA format with sample code followed by primer information.

>KI002 Dlp1.5 to Dlp4R
CATAAACTATTCCCTGAAAAGTCTTGTTATAGAATCACTATAACCCCACAGTACTACGTCAGTATTGAAAAAAA
AAATCCTACAGTACATTTACTGTATTAATTATAGAGGCACACCTACCTACACGCTAATGTATAGCGCCTCTCT
AGGACTGTATGTATATATATACTATGTATAACTGTGCATTCATTTATTTTCACTACGGAGAGTTAAAGCTCGTA
ATTAATTTTTTTAATTTTACATAAGTACATAATTTGCATTATTCGTACATGTGCCCGTTCCATTAAATCACAAGC
TTAATCACCATGCCG

>KI011 Dlp1.5 to Dlp4R
CATAAACTATTCCCTGAAAAAGTCTTATTGTAGAATCACCACAACCCTACAGTGCTATGTCAGTATTGAAAAAA
AATATCCTACAGTACATTTACTGTATTAATCATACAAGCATACTTACCTACACGCTAATATATAGCGTCTCTCC
AAGACTGTATGTATATATATACTATGTATAACTGTGCATTCATTTATTTTCACTACGGAGAGTTAAAGCTCGTA
ATTAATTTTTTTTATTTTACATAAGTACATAATTTGCATTACTTGTACATGTGCCCGTTCCATTAGATCACGAGC
TTAATCACCATGCCG

>KI012 Dlp1.5 to Dlp4R
AACTATTCCCTGAAAAAGTCTTATTGTAGAATCACCACAACCCTACAGTGCTATGTCAGTATTGAAAAAAAATA
TCCTACAGTACATTTACTGTATTAATCATACAAGCATACTTACCTACACGCTAATATATAGCGTCTCTCCAAGA
CTGTATGTATATATATACTATGTATAACTGTGCATTCATTTATTTTCACTACGGAGAGTTAAAGCTCGTAATTAA
TTTTTTTTATTTTACATAAGTACATAATTTGCATTACTTGTACATGTGCCCGTTCCATTAGATCACGAGCTTAAT
CACCATGCCG

>KI018 Dlp1.5 to Dlp4R
ATTCCCTGAAAAGTCTTGTTATAGAATCACTATAACCCCACAGTACTACGTCAGTATTGAAAAAAAAAATCCTA
CAGTACATTTACTGTATTAATTATAGAGGCACACCTACCTACACGCTAATGTATAGCGCCTCTCTAGGACTGT
ATGTATATATATACTATGTATAACTGTGCATTCATTTATTTTCACTACGGAGAGTTAAAGCTCGTAATTAATTTT
TTTAATTTTACATAAGTACATAATTTGCATTATTCGTACATGTGCCCGTTCCATTAAATCACGAGCTTAATCAC
CATGCCG

>KI019 Dlp1.5 to Dlp4R
CATAAACTATTCCCTGAAAAGTCTTGTTATAGAATCACTATAACCCCACAGTACTACGTCAGTATTGAAAAAAA
AAATCCTACAGTACATTTACTGTATTAATTATAGAGGCACACCTACCTACACGCTAATGTATAGCGCCTCTCT
AGGACTGTATGTATATATATACTATGTATAACTGTGCATTCATTTATTTTCACTACGGAGAGTTAAAGCTCGTA
ATTAATTTTTTTAATTTTACATAAGTACATAATTTGCATTATTCGTACATGTGCCCGTTCCATTAAATCACGAGC
TTAATCACCATGCCG

>UKiribati_Dlp300bp
GAAAAAGTCTTATTGTAGAATCACCACAACCCTACAGTGCTATGTCAGTATTGAAAAAAAATATCCTACAGTA
CATTTACTGTATTAATCATACAAGCATACTTACCTACACGCTAATATATAGCGTCTCTCCAAGACTGTATGTAT
ATATATACTATGTATAACTGTGCATTCATTTATTTTCACTACGGAGAGTTAAAGCTCGTAATTAATTTTTTTTAT
TTTACATAAGTACATAATTTGCATTACTTGTACATGTGCCCGTTCCATTAGATCACGAGCTTAATCACCATGCC
GCGT

>KI024_dlp1.5 to Dlp5R
CCTTATTGTAAATAACCACAAACCTCTAGGGCTATGTCAGTATTAAAATTAATCTAGTCCAATTACATTCTCAT
CGTTAAAAAACCATACAAACATACACCCCATCCAATAAAATAGCGTTCTCCTCGTAGATGTATGTATATACATA
GCTATGTATAATAGTGCATTCAATTATTTTCCACACGAGAAGTTAAAGCCCGTATTAGATTTTATTAATTTTACA
TATTACATAATATTATTGATCGTACATAAGACATACTTCTAAATCAGTTCCAGTCCCTTAACAGTATGGCCGCT
CCATTAGATCACGAGCTTAACCACCATGCCGCGTGAAAANNNNAACCCGCNANGCTGGGATCCCCCTCCTC
GCACCGGGCCCATTAACTGTGGGGGTAGCTATTTTATGCCTTTTACAATACATCTGGTTCTTACTTCAGGGC
CATATTCACTTAAAATCGCCCACTCGTTCCTCTTAAATAAGACAT

**SupTable 2**: Sequences of the mtDNA cytochrome *b* used in species identification of bones from Gilbert Islands. Shown in FASTA format with sample code followed by primer information.

>KI002_CybMLDF to Cyb305-328R
AAATATTTCCTCATGATGAAATTTCGGCTCCTTACTCGGCCTCTGCCTAATTATACAAATTCTCACAGGACTAT
TCCTGGCAATACACTACACACCAGACACCACAACAGCCTTTTCATCTGTTACACATATCTGCCGAGACGTTA
ACTATGGCTGAATTATCCGATATCTACATGCAAATGGGGCTTCTATATTTTTNNNNNNNNNNNATGCACATAT
TGGACGTGGTCTGTACTACGGCTCTTATATCTTCCAAGAAACATGAAATATCGGAGTAATCTTACTCTTTACA
GTTATAGCCACTGCATTTGTAGGCTATGTCCTACCAT

> KI009_CybMLDF to Cyb305-328R
ATCTCCTCATGATGAAACTTCGGCTCCTTACTCGGCCTCTGCCTCATCATACAAATTCTCACAGGCCTGTTCT
TAGCAATACACTATACACCAGACACAACAACAGCCTTCTCATCCGTTGCACACATTTGCCGAGACGTCAACT
ATGGCTGAATCATTCGATACCTACACGCAAACGGGGCCTCCATATTCTTCATCTGCCTTTACGCCCATATCG
GACGTGGACTATATTACGGCTCCTATATCTTTCAAGAAACATGAAACATCGGAGTAATCCTACTCCTTGCAGT
TATAGCTACCGCATTTGTGGGCTATGTCCTACCTT

**SupFigure 1:** Whaling artifacts and cetacean bones collected during surveys of the Gilbert Islands from 18 June to 14 July 2009: A) whaling 'try-pot', a remnant of 19th century whaling contact; B) vertebra (KI004, identified as *Mesoplodon* sp., an unrecognized species or subspecies) and rib bone (KI014, no DNA identification) collected on Onotoa Island; and C) mandible, rib and teeth of a sperm whale on display in Butaritari Island. Photographs courtesy of Al Hutt.

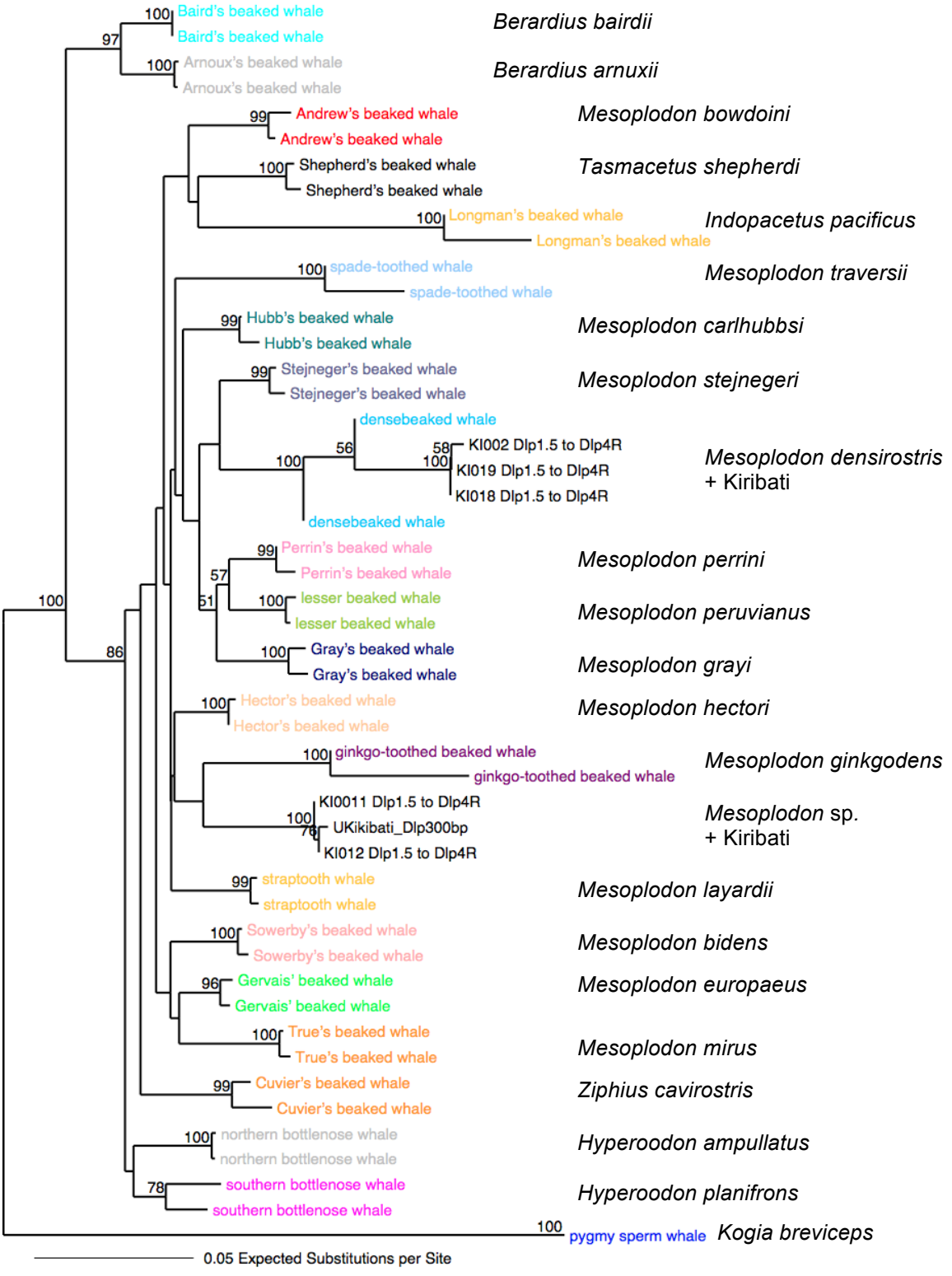
A

B

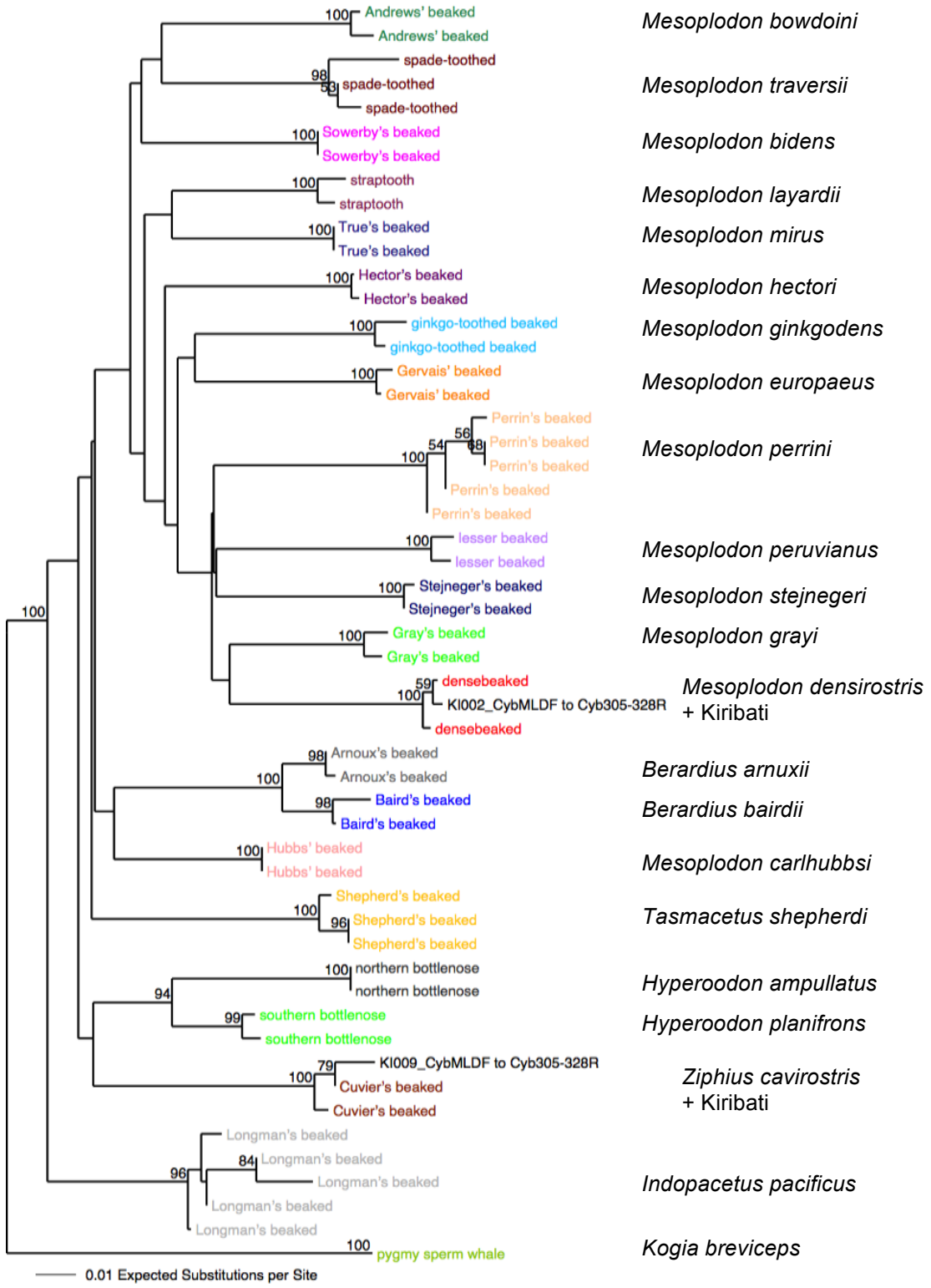
C

**SupFigure 2:** Representative identification of mtDNA sequences from cetacean material collected from the outer islands of the Gilbert group, in the Republic of Kiribati, using reference sequences available through the web-based application, *DNA-surveillance* (Ziphiidae Vs4.3): A) control region (DLP) of KI0o2, KI011, KI012, KI018 and KI019; B) cytochrome *b* (CYB) of KI002 and KI009. The neighbor-joining tree reflects species-specific grouping with bootstrap values based on 500 replications. The program is available at http://www.cebl.auckland.ac.nz:9000/page/whales/title

A) Control region



B) Cytochrome *b*



Marine Mammal Science

MARINE MAMMAL SCIENCE, **(*): ***–*** (*** 2014)

DOI: 10.1111/mms.12113

# Resurrection of *Mesoplodon hotaula* Deraniyagala 1963: A new species of beaked whale in the tropical Indo-Pacific

**Merel L. Dalebout**,[1] School of Biological, Earth, and Environmental Sciences, University of New South Wales, Sydney, New South Wales 2052, Australia; **C. Scott Baker** and **Debbie Steel**, School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland 1000, New Zealand and Marine Mammal Institute, Hatfield Marine Science Center, Oregon State University, 2030 Marine Science Drive, Newport, Oregon 97365, U.S.A.; **Kirsten Thompson**, School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland 1000, New Zealand; **Kelly M. Robertson**, **Susan J. Chivers** and **William F. Perrin**, Southwest Fisheries Science Center, NMFS, NOAA, 8901 La Jolla Shores Drive, La Jolla, California 92037, U.S.A.; **Manori Goonatilake**, Department of National Museums, PO Box 854, Colombo 7, Sri Lanka; **R. Charles Anderson**, Manta Marine Pvt. Ltd., PO Box 2074, Malé, Republic of Maldives; **James G. Mead** and **Charles W. Potter**, National Museum of Natural History, Mailstop NHB 108, Smithsonian Institution, Washington, DC 20560, U.S.A.; **Lisa Thompson** and **Danielle Jupiter**, Island Conservation Society Fondation pour la Conservation des Iles Pointe Larue, PO Box 775, Victoria, Seychelles; **Tadasu K. Yamada**, National Science Museum, Tokyo, 3-23-1 Hyakunin-cho, Shinjuku-ku, Tokyo 169-0073, Japan.

## Abstract

We present genetic and morphological evidence supporting the recognition of a previously synonymized species of *Mesoplodon* beaked whale in the tropical Indo-Pacific, *Mesoplodon hotaula*. Although the new species is closely-related to the rare ginkgo-toothed beaked whale *M. ginkgodens,* we show that these two lineages can be differentiated by maternally (mitochondrial DNA), biparentally (autosomal), and paternally (Y chromosome) inherited DNA sequences, as well as by morphological features. The reciprocal monophyly of the mtDNA genealogies and the largely parapatric distribution of these lineages is consistent with reproductive isolation. The new lineage is currently known from at least seven specimens: Sri Lanka (1), Gilbert Islands, Republic of Kiribati (1+), Palmyra Atoll, Northern Line Islands, U.S.A. (3), Maldives (1), and Seychelles (1). The type specimen (Sri Lanka) was described as a new species, *M. hotaula*, in 1963, but later synonymized with *M. ginkgodens.* This discovery brings the total number of *Mesoplodon* species to 15, making it, by far, the most speciose yet least known genus of cetaceans.

Key words: speciation, taxonomy, species delimitation, mtDNA, nuclear introns, Y-chromosome, morphology, Mesoplodon, beaked whale.

[1]Corresponding author: (e-mail: mereldalebout@gmail.com).

On 26 January 1963, a female beaked whale washed ashore at Ratmalana, Sri Lanka. The stranding was reported by the director of the National Museums of Ceylon, P. E. P. Deraniyagala (1963*a*, *b*), who described the whale as a new species, *Mesoplodon hotaula*, deriving the species name from the local Sinhala words for "pointed beak." Deraniyagala provided no diagnosis by which *M. hotaula* could be differentiated from the other *Mesoplodon* beaked whales known at that time, except to note that the position of the teeth differed from that of *M. bidens* and *M. hectori.* Two years after it was described, *M. hotaula*, still known only from the holotype, was synonymized with the ginkgo-toothed beaked whale, *M. ginkgodens* Nishiwaki and Kamiya 1958, by Moore and Gilmore (1965). Although Deraniyagala was apparently unaware of the existence of *M. ginkgodens* when he described *M. hotaula,* now it seems he was correct regarding its uniqueness.

Here we present genetic and morphological evidence for the distinctiveness of *M. hotaula*, now known from at least seven specimens. We consider the taxonomic ranking of the new taxon using the Genealogical Concordance Species Concept (GCC; Avise and Ball 1990), and discuss its sister-species relationship with *M. ginkgodens*. In choosing the GCC, we are aware that there are many definitions of what constitutes a species (De Queiroz 2007). The Biological Species Concept (BSC) defines a species as a group of actually or potentially interbreeding natural populations that are reproductively isolated from other such groups (Mayr 1963); a criterion that is generally difficult if not impossible to assess in many wild populations. The Phylogenetic Species Concept (PSC), a character-based approach originating from cladistic principles, defines a species as the smallest cluster of organisms that can be diagnosed as distinct from other clusters based on fixed character differences, showing a parental pattern of ancestry and descent (Cracraft 1989, Davis and Nixon 1992). The GCC attempts to reconcile these two approaches by requiring multiple lines of evidence, including phylogenetic analysis of DNA sequences and morphology, to establish the distinctiveness of the evolutionary lineages in question. This addresses the problem that application of a strict PSC, especially using molecular data, could lead paradoxically to a vast increase in the number of recognized species at a time when most biologists agree that global biodiversity is decreasing (Avise 2000, Zachosa *et al.* 2013).

The GCC also accounts for the fact that gene phylogenies can differ greatly from locus to locus due to recombination in meiosis, mating patterns, and varying reproductive success of the individuals through which alleles are transmitted. The GCC therefore stresses that phylogenetic diagnoses should be based on broad agreement at multiple loci. A group of organisms is considered to constitute a distinct species under the GCC if the following criteria are met (Avise and Ball 1990, Avise 2000): (1) concordance across sequence characters within a genetic locus leading to conclusive exclusion; (2) concordance in these genealogical patterns across multiple loci, both mitochondrial and nuclear; (3) concordance with biogeographical patterns; and (4) concordance with morphological characters. These criteria were reviewed and supported by a specialist workshop on shortcomings in cetacean taxonomy (Reeves *et al.* 2004) and the GCC has subsequently been used by several authors to describe new species of cetaceans (*e.g.*, Dalebout *et al.* 2004, Caballero *et al.* 2007).

The GCC is especially useful for rare or poorly described taxa. Beaked whales (family Ziphiidae) are deep-diving odontocetes that live in the offshore waters of all the world's oceans except the highest latitudes of the Arctic. They are rarely seen at sea due to their elusive habits, long dive capacity, and, for some species, probable low abundance (Reeves *et al.* 2002). Most information has come from stranded animals, and several species are known from only a handful of specimens. To assist with beaked

whale identification and discovery, a comprehensive, validated DNA taxonomy for all known species in this group was established using sequences from mitochondrial DNA (mtDNA) control region (CR) and cytochrome *b* (CYB) genes (Dalebout *et al.* 2004). This database was subsequently expanded for the most speciose genus, *Mesoplodon*, using up to six specimens per species from throughout the range where possible, to assess the robustness of genetic patterns observed at these loci (Dalebout *et al.* 2007). The patterns originally observed were again evident in this expanded sample: intraspecific variation within species was generally low (mean; CR $0.6\% \pm 0.06\%$, CYB $0.8\% \pm 0.09\%$), while interspecific divergence was generally high (mean; CR $7.4\% \pm 0.04\%$, CYB $11.8\% \pm 0.04\%$), with little overlap. In phylogenetic analyses, these mtDNA sequences formed strongly supported, species-specific clades that were reciprocally monophyletic with respect to all other such clades. In short, sequences from each of the known *Mesoplodon* species clustered together to the exclusion of sequences from other known *Mesoplodon* species. Each *Mesoplodon* species also possessed multiple diagnostic nucleotide substitutions at these genes distinguishing each species from all other species in the group, *sensu* Davis and Nixon (1990). While intraspecific variation may be underestimated because of the small sample sizes available, Dalebout *et al.* (2007) sampled as many specimens as possible, which in some cases included all currently known specimens for that species. Overall these mitochondrial markers were found to be well suited for DNA taxonomy in this genus, with the results from phylogenetic analyses concordant with morphological diagnoses and other requirements for species distinctiveness under the GCC. The application of this DNA taxonomy to beaked whales has already led to some significant discoveries: the description of a new species from the North Pacific (Perrin's beaked whale *M. perrini*; Dalebout *et al*. 2002); the resurrection of a long-forgotten species in the Southern Hemisphere (the spade-toothed whale *M. traversii*; van Helden *et al.* 2002, Thompson *et al.* 2012); and, confirmation of the identity of the enigmatic "tropical bottlenose whale" (*Indopacetus pacificus*; Dalebout *et al.* 2003).

Within the robust framework offered by this comprehensive DNA taxonomy and the guidelines provided by the GCC, the discovery of a divergent lineage could therefore indicate the existence of an unrecognized species or subspecies. Just such a lineage was reported by Dalebout *et al.* (2007), based on analyses of several specimens which appeared to be related to *M. ginkgodens* yet were genetically distinct from this species. Further specimens representing this divergent mtDNA lineage have since been discovered. One of these was the Sri Lankan specimen described as a new species, *M. hotaula*, by Deraniyagala (1963*a*, *b*) and subsequently synonymized with *M. ginkgodens* (Moore and Gilmore 1965). *M. ginkgodens* is one of the least-known of beaked whale species. It is known from less than 30 strandings and there has yet to be a confirmed sighting of a living whale at sea.

To assess the taxonomic status of "*M. hotaula*," we analyzed three mtDNA genes, seven nuclear autosomal introns, and one Y-chromosome intron, as well as morphological characters. Detection of genetic differences among recently diverged taxa can be difficult with slowly evolving, single-copy nuclear autosomal loci (Hare 2001). Our inclusion of data from a Y-chromosome intron therefore has several advantages. Firstly, under random mating, the effective population size of this nonrecombining chromosome is ¼ that of single-copy autosomal markers. Therefore the accumulation of mutations through genetic drift occurs far more rapidly. Secondly, the Y-chromosome is subject to mutations that have arisen only in the male germline, giving us a male-specific marker to compare to the female-specific mtDNA. Based on diagnostic genetic characters and morphological features consistent with the GCC and the

criterion of "irreversible divergence," as recommended by a workshop on cetacean taxonomy (Reeves *et al.* 2004), we present a formal proposal for the recognition of *M. hotaula* as a valid species.

## Materials and Methods

### *Material Examined*

Seven specimens of *M. hotaula* were examined (Table 1, Fig. 1)[2] and compared to all other known *Mesoplodon* species *via* phylogenetic analyses of mtDNA and nuclear gene sequences. Museums and institutions holding specimens of *M. hotaula* are as follows: the National Museum, Colombo, Sri Lanka (1), Smithsonian National Museum of Natural History, Washington, DC, U.S.A. (USNM, 3), a private collection in the Republic of Maldives (1), and the Island Conservation Society, Seychelles (1). Specimen 2 from the Gilbert Islands, Republic of Kiribati, is known only from a soft-tissue sample held in the University of Auckland DNA and Tissue Archive, Auckland, New Zealand. Information on additional specimens identified by DNA analysis from fragmentary osteological material from Kiribati can be found in Baker *et al.* (2013).

Genetic and morphological comparisons were made to six specimens of *M. ginkgodens* (Table 1, Fig. 1), including the holotype (Nishiwaki and Kamiya 1958). For genetic comparisons to other *Mesoplodon* species, up to six specimens per species were sampled (see Dalebout *et al.* 2007 for details).

### *DNA Extraction, PCR, and Sequencing*

Six of the seven specimens of *M. hotaula* (Nos. 1, 3–7) were represented only by osteological material requiring the use of "ancient DNA" methods. A hand-held electric drill with a 2 mm diameter drill bit was used to obtain 0.01–0.02 g of bone or tooth powder from each specimen as described by Pichler *et al.* (2001*a*). DNA was extracted using the silica-guanidinium thiocyanate method (Boom *et al.* 1990, Höss and Pääbo 1993, Matisoo-Smith *et al.* 1997) as modified by Rohland and Hofreiter (2007). These methods were also used to extract DNA from the holotype and California specimens of *M. ginkgodens* (Dalebout *et al.* 2004). Only the Kiribati specimen of *M. hotaula* (No. 2) was represented by soft tissue that was several months old and dried for preservation.[3] DNA was extracted from this sample using standard phenol: chloroform methods (Sambrook *et al.* 1989), as modified for small samples by Baker *et al.* (1994).

Specimens 1 and 6 were analyzed at the University of New South Wales, Sydney, Australia (by MLD). Specimens 2 and 7 were analyzed at the University of Auckland, New Zealand (by DS, KT, and MLD). Specimens 3 through 5 were analyzed at the NOAA/NMFS Southwest Fisheries Science Center, La Jolla, California (by KMR).

The polymerase chain reaction (PCR) was used to amplify fragments from three mitochondrial genes (control region–CR, cytochrome *b*–CYB, cytochrome *c* oxidase I–COXI), seven nuclear autosomal introns (biglycan–BGN, catalase–CAT, rhodopsin–RHO, cytotoxic T-lymphocyte-associated serine esterase 3–CTLA3, cholinergic

[2]The initial conclusions of Dalebout *et al.* (2007) were based on specimen Nos. 2–4.

[3]Personal communication from R. Grace, 56 Bertram Street, Warkworth, New Zealand, 19 August 2003.

*Table 1.* Specimens of *Mesoplodon hotaula* and *M. ginkgodens* examined for this study. See text for details of museum holdings. H, holotype.

| | Specimen code | Other codes | Date found | Location | Coordinates | Total length (cm) | Sex |
|---|---|---|---|---|---|---|---|
| *Mesoplodon hotaula* | | | | | | | |
| 1 | 3WZS (H) | | 26 January 1963 | Ratmalana, Sri Lanka | 6°49′N, 79°52′E | 445 | F |
| 2 | UKIRI | | 11 July 2003 | Tabiteuea Atoll, Republic of Kiribati | 1°07′S, 174°40′E[a] | ? | M[b] |
| 3 | USNM593418 | SW53473, PANWR 12533-06001[c] | 9 November 2005 | Palmyra Atoll, northern Line Islands, USA | 5°52′N, 162°06′W | 480 | F[b] |
| 4 | USNM593414 | SW53474, PANWR 12533-06002[c] | 9 November 2005 | Palmyra Atoll, northern Line Islands, USA | 5°52′N, 162°06′W | 470 | F? |
| 5 | USNM593426 | SW70984, PANWR 12533- 06003[c] | 8 July 2006 | Palmyra Atoll, northern Line Islands, USA | 5°52′N, 162°06′W | 386 | M |
| 6 | MDV-X | | January 2007 | Hulhudhuffaru, Raa Atoll, Republic of Maldives | 5°45′N, 73°00′E | ? | M |
| 7 | MM-0001 | | 20 June 2009 | Desroches Island, Seychelles | 5°67′S, 53°65′E | 432 | M |
| *Mesoplodon ginkgodens* | | | | | | | |
| 1 | MginUSNM298237 | | 10 June 1954 | Del Mar, California | 32°57′N, 117°15′W | ? | F |
| 2 | MginTSM8744 (H) | NSMT M8744 | 22 September 1957 | Oiso, Tokyo, Japan | 35°18′N, 139°18′E | 472 | M |
| 3 | MginMV29623 | | 26 June 1983 | Cape Reamur, Victoria, Australia | 38°23′S, 142°08′E | ? | ? |
| 4 | MginTW01 | NMNS-SU-94-29 | 1994 | northeast coast Taiwan | 23°46′N, 121°00′E | ? | F |
| 5 | MginNZ03 | NMNZ 2901 | 11 April 2003 | Taranaki, New Zealand | 39°18′S, 174°08′E | 480 | M |
| 6 | MginNZ04 | NMNZ2618 | 1 November 2004 | Pakawau, Nelson, New Zealand | 40°48′S, 172°48′E | 496 | M |

[a]Note that coordinates given in Dalebout *et al*. (2007) were incorrect.
[b]Determined or confirmed by molecular sexing.
[c]US Fish and Wildlife Service "Palmyra Atoll National Wildlife Refuge" Accession number.

Note: other specimens which may represent *M. hotaula*: adult male (BMNH1957.4.5.1) from Malaysia, November 1954, material held at the Natural History Museum, London, U.K.; identification by MLD based on skull morphology from photographs courtesy of R. Sabin; neonate female (End002) from Phuket, Thailand, 7 June 1988 or 1989 (Andersen and Kinze 1999, Chantrapornsyl *et al*. 1999); based on stranding location only.

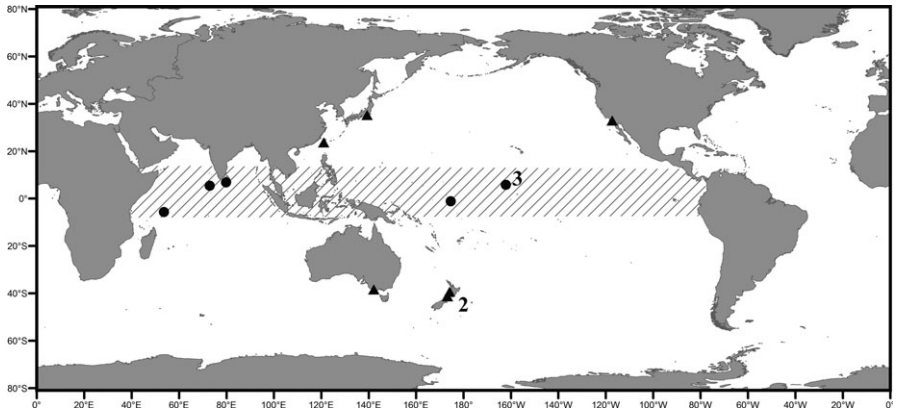



*Figure 1.* Map showing stranding locations of specimens sampled for this study: *M. hotaula* (black circles), *M. ginkgodens* (black triangles). Area with diagonal lines indicates suggested distribution of *M. hotaula.* Each symbol represents one specimen per location, except where noted (see Table 1).

receptor-nicotinic alpha polypeptide 1–CHRNA1, muscle actin–ACT, major histocompatibility complex class II–DQA) and one nuclear Y-chromosome intron (DBY7). For COXI amplification, we used the primers, BatL5310 (5′-CCT ACTCRGCCATTTTACCTATG-3′) and Bat6871tSer (5′-GTTCGATTC CTTCCTTTCTT-3′), courtesy of the Alan Wilson Centre, Massey University, New Zealand (T. McLenachan). For DBY7 amplification, we used the primers, DBY7-F (5′-GGTCCAGGAGARGCTTTGAA-3′) and DBY7-R (5′-CAGCCAATTCTC TTGTTGGG-3′), from Hellborg and Ellegren (2003). PCR information for COXI and DBY7 can be found in the online supplementary material. Information for other loci is in Dalebout *et al.* (2004, 2008*b*). PCR products were prepared for sequencing by enzymatic purification, using shrimp alkaline phosphatase and exonuclease I (Werle *et al.* 1994). Products were sequenced on an ABI 377, modified ABI 373, or ABI 3700 automated sequencer (Applied Biosystems, Inc.) using BigDye Dye Terminator Chemistry Vs. 3.1. Fragments were sequenced at least twice in both directions for confirmation in the majority of cases. Sequences were edited manually and aligned using the program SEQUENCHER Vs. 4.0 (Gene Codes Corporation, Inc.).

### *Genetic Analyses*

*Mitochondrial DNA*—Phylogenetic reconstructions of individual CR and CYB data sets were presented by Dalebout *et al.* (2007). Here, we concatenated these two data sets (819 base pairs, bp) in an attempt to obtain a stronger phylogenetic signal. Cuvier's beaked whale *Ziphius cavirostris* was used as an outgroup. COXI data were not included in these analyses as only a single representative of *M. hotaula*, and only a subset of other ziphiid species have been sequenced for this locus to date. Maximum likelihood (ML) analyses were performed using PAUP* 4.0b10 (Swofford 2003), with parameters estimated by Modeltest (Posada and Crandall 1998) and starting trees for heuristic searches obtained *via* neighbor joining. Full model details for ML analyses can be found in the online supplementary material. The robustness of the nodes was assessed using 1,000 full heuristic, nonparametric ML bootstrap replicates.

Bayesian analyses were performed using MrBayes 3.1.2 (Huelsenbeck and Ronquist 2001) using an ML model with six substitution types and empirical base frequencies. Rate variation across sites was modelled using a gamma distribution, with a proportion of sites estimated as being invariant. The Markov chain Monte Carlo search was run with four chains for 1 million generations, with trees being sampled every 100 generations (first 1,000 trees were discarded as burn-in). Details on methods used to assess convergence can be found in the online supplementary material.

Building on the results of Dalebout *et al.* (2007), we made pairwise comparisons of the CR, CYB, and COXI gene fragments to determine the number of diagnostic nucleotide substitutions (putative fixed differences) distinguishing *M. hotaula* from *M. ginkgodens.* The program MEGA3 (Kumar *et al*. 2004) was used to calculate net divergence (dA ± SE; Nei 1987) between taxa using the Kimura 2-parameter model, and determine the proportion of synonymous and nonsynonymous substitutions for the two protein-coding genes.

*Nuclear introns*—Phylogenetic reconstruction of the relationships among *Mesoplodon* beaked whales based on autosomal nuclear introns was presented by Dalebout *et al.* (2008*b*) but did not include *M. hotaula.* Initial screening of these slowly evolving loci for the present study revealed only limited differentiation between *M. hotaula* and *M. ginkgodens,* such that phylogenetic assessment would not be able to differentiate between them. Therefore, we focused on a character-based diagnostic approach (Davis and Nixon 1992) for these autosomal loci, as well as for the Y-chromosome intron, DBY7. Under a character-based approach, a clade is characterized by one or more synapomorphies, defined as shared derived character states inferred to have been present in the first member of the taxon (most recent common ancestor), inherited by its descendants (unless secondarily lost), and not inherited by any other taxa.

### *Morphological Data*

Cranial and mandibular measurements were obtained for five specimens, following the methods of Moore (1963), as adapted by JGM for the Smithsonian US National Museum of Natural History collections. Specimens USNM593418, USNM593414, and USNM593426 were measured by JGM and CWP. Measurements for the holotype (3SWZ) were taken from Deraniyagala (1963*a*), and checked by RCA and MG in consultation with JGM. Specimen MDV-X was also measured by RCA. Comparative measurements for the *M. ginkgodens* holotype were obtained from Nishiwaki and Kamiya (1958). Measurements from additional *M. ginkgodens* were provided by TKY. Measurements were obtained using calipers and rounded to the nearest whole mm.

## RESULTS

### *Genetics—Mitochondrial DNA*

CR fragments (658 bp) were successfully sequenced from all seven specimens of *M. hotaula.* CYB fragments (384–706 bp) were successfully sequenced from only four specimens due to the degraded nature of the material available. A COXI fragment (987 bp) was successfully sequenced only from the dried meat from Kiribati (Table 2). These CR, CYB, and COXI fragments were also successfully sequenced from up to six specimens of *M. ginkgodens*.

In phylogenetic analyses of the combined CR and CYB (819 bp) including all known *Mesoplodon* species, the *M. hotaula* and *M. ginkgodens* specimens clustered together in two strongly supported clades (bootstrap scores, BS 100%, posterior probabilities, BPP 1.00) that were reciprocally monophyletic to one another, consistent with the proposed species differences (Fig. 2). The sister-species relationship of these taxa was also strongly supported (bootstrap 91%, posterior probability 1.00). All other recognized *Mesoplodon* species formed similar strongly supported, species-specific clades, with branch lengths reflecting the relatively low genetic diversity observed *within* species and the comparatively large genetic divergence observed *between* species (see also Dalebout *et al*. 2002, 2004, 2007). Individual analyses of the CR and CYB data sets revealed the same pattern (Dalebout *et al.* 2007).

It is worth noting the deep mtDNA divergence observed among True's beaked whales, *M. mirus* (Fig. 2). This is the only *Mesoplodon* species with a disjunct, allopatric distribution, with populations found in the North Atlantic and in the Southern Hemisphere (South Africa and Australia). These northern and southern populations also have different color patterns (Ross 1969) and likely represent unique subspecies or species in their own right. This divergence, however, is far less than that observed between the *M. ginkgodens* and *M. hotaula* lineages.

For the CR, pairwise comparisons between *M. hotaula* and *M. ginkgodens* over 658 bp revealed 35 variable sites, of which 18 appear to represent diagnostic characters distinguishing these taxa from one another (Table S1A). In pairwise comparisons between all *Mesoplodon* species (435 bp), net divergence (dA) ranged from 3.1 % to 8.3% (mean, 5.4% ± 0.99%). Net divergence between *M. hotaula* and *M. ginkgodens* was 3.6% ± 0.91% over this fragment. Similar levels of genetic divergence were observed between other sister-species pairs (as identified by Dalebout *et al.* (2008*b*) based on nuclear introns, the results of which were concordant with cranial morphology): *M. perrini* and *M. peruvianus* (3.2% ± 0.83%); *M. densirostris* and *M. stejnegeri* (3.6% ± 0.90%); *M. mirus* and *M. europeaus* (3.7% ± 0.87%); and *M. bowdoini* and *M. carlhubbsi* (5.2% ± 1.10%; Table S2A).

For CYB (384 bp), comparisons between *M. hotaula* and *M. ginkgodens* revealed 31 variable sites, of which 26 appear to represent diagnostic characters distinguishing these taxa from one another, including four nonsynonymous substitutions (Table S1B). In pairwise comparisons between all *Mesoplodon* species (384 bp), net divergence ranged from 5.5% to 16.6% (mean, 11.4% ± 1.97%). Net divergence between *M. hotaula* and *M. ginkgodens* was 8.2% ± 1.79%. This is slightly lower than that observed between other recognized sister-species pairs: *M. mirus* and *M. europeaus* (8.7% ± 1.85%), *M. densirostris* and *M. stejnegeri* (10.4% ± 2.01%), *M. perrini* and *M. peruvianus* (11.7% ± 2.24%), and *M. bowdoini* and *M. carlhubbsi* (11.8% ± 2.24%; Table S2B).

For COXI (987 bp), comparisons between *M. hotaula* and *M. ginkgodens* revealed 64 variable sites, of which 49 appear to represent diagnostic characters distinguishing these taxa from one another (all synonymous substitutions, Table S1C). Although COXI is not considered a good "DNA barcode" for cetaceans due to significant overlap between intra- and interspecific variation in some groups, Viricel and Rosel (2012) similarly observed species-specific sequences for *Mesoplodon* beaked whales. In pairwise comparisons between a subset of species (*M. hotaula, M. ginkgodens, M. europaeus, M. mirus,* and *M. densirostris*), net divergence ranged from 5.5% to 10.0% (mean, 8.5% ± 1.24%) over 958 bp. Net divergence between *M. hotaula* and *M. ginkgodens* was 5.5% ± 0.76%. The only other sister-species pair sampled in our

*Table 2.* Summary of genetic data used in the comparison of *Mesoplodon hotaula* and *M. ginkgodens*. Auto, autosomal; CR, control region; COXI, cytochrome oxidase I; CYB, cytochrome *b*; DBY7,[a] sex intron; MtDNA, mitochondrial DNA.

| | | MtDNA | | | Nuclear introns | | |
|---|---|---|---|---|---|---|---|
| Specimen code | Source of DNA | CR | CYB | COXI | Auto | DBY7 | GenBank accession Numbers |
| *Mesoplodon hotaula* | | | | | | | |
| 1 3WZS | osteological material | Y | | | | | KF027298 |
| 2 UKIRI | dried soft tissue | Y | Y | Y | Y | Y | JX470545, JX470546, KF027311, KF027315-KF027321, KF027328 |
| 3 USNM593418 | osteological material | Y | Y | | | | JX470543, JX470547 |
| 4 USNM593414 | osteological material | Y | | | | | JX470544 |
| 5 USNM593426 | osteological material | Y | Y | | | | KF027299, KF027306 |
| 6 MDV-X | osteological material | Y | | | | | KF027300 |
| 7 MM-0001 | osteological material | Y | Y | | | | KF027301, KF027307 |
| *Mesoplodon ginkgodens* | | | | | | | |
| 1 MginUSNM298237 | osteological material | Y | | | | | KF027302 |
| 2 MginTSM8744 | osteological material | Y | | | | | AY579518 |
| 3 MginMV29623 | osteological material | Y | Y | | | | KF027303, KF027310 |
| 4 MginTW01 | soft tissue | Y | Y | Y | Y | | AY579517, AY579544, KF027312, EU447764, EU476111, EU476141 |
| 5 MginNZ03 | soft tissue | Y | Y | Y | Y | Y | KF027304, KF027308, KF027313, EU447749, EU476126, EU476156, EU476171, KF027329 |
| 6 MginNZ04 | soft tissue | Y | Y | Y | | | KF027305, KF027309, KF027314 |

[a]DBY7 GenBank Accession Numbers for other *Mesoplodon* species: KF027328-KF027337.

study, *M. mirus* and *M. europaeus*, differed by 7.8% ± 1.05% at this locus (Table S2C).

Note that for the CR and CYB data sets, the lowest divergences were observed between taxa that are not recognized sister species (CR, *M. hectori* and *M. hotaula,* 3.11% ± 0.83%; CYB, *M. layardii* and *M. mirus,* 5.53% ± 1.42%). This is the due to the rapid rate of accumulation of mutations at these highly variable loci, resulting in multiple substitutions/site (saturation). Nucleotide substitutions in ancestral lineages are undetectable as subsequent substitutions erase the evidence. This issue is particularly apparent when genetic differences are reduced to pairwise distances, as phylogenetic divergence between taxa will be underestimated even further (Avise 1994). For this reason, while well suited to addressing questions of species identity, the CR and CYB are generally not well suited for resolving higher-level relationships in the genus *Mesoplodon*, including some of the deeper, older divergences between sister species (Fig. 2, gray-shaded area, bootstrap scores <50%; see also Dalebout *et al.* 2004, 2007). A robust, higher-level phylogeny for this group has been provided by phylogenetic analyses of more slowly evolving nuclear markers (Dalebout *et al.* 2008*b*).

Intraspecific diversity for *M. hotaula* at the CR and CYB was low, in line with trends observed for these loci in other *Mesoplodon* species (Dalebout *et al.* 2007). For CR, the Kiribati and Palmyra specimens shared the same haplotype, the Maldives specimen differed from this by 4 bp (0.61%), and the holotype of *M. hotaula* differed from this by 7 bp (1.06%). For CYB, two Palmyra specimens shared the same haplotype, and the Kiribati specimen differed from this by 1 bp.

*Genetics-Nuclear Introns*

Due to the degraded nature of much of the material available, nuclear intron sequences were obtained only from the Kiribati specimen of *M. hotaula.* For this same reason, other *Mesoplodon* species were also generally represented by only a single specimen for these analyses. Partial introns were successfully amplified from seven nuclear genes: BGN, 706 bp; CAT, 559 bp; RHO, 166 bp; CTLA3, 305 bp; CHRNA1, 366 bp; ACT, 925 bp; and, DQA, 456 bp (Dalebout *et al.* 2008*b*). Over all these introns combined (3,348 bp), each previously recognized *Mesoplodon* species possessed between one (*M. grayi, M. ginkgodens* proper) and 13 (*M. bidens*) diagnostic nucleotide substitutions, *sensu* Davis and Nixon (1990), distinguishing them from all other species in the group (Table 4). One nucleotide substitution distinguished *M. ginkgodens* proper from all other species including *M. hotaula*, and four nucleotide substitutions distinguished the *M. ginkgodens–M. hotaula* complex from all other species. In sister-species comparisons, one nucleotide substitution (position 2199, ACT) distinguished *M. ginkgodens* from *M. hotaula*, while 10 substitutions distinguished *M. perrini* from *M. peruvianus,* 12 substitutions distinguished *M. bowdoini* from *M. carlhubbsi,* 14 substitutions distinguished *M. densirostris* from *M. stejnegeri,* and 17 substitutions distinguished *M. europeaus* from *M. mirus.* The divergence date estimates for these latter species pairs range from 5.3 to 10.4 Mya (Dalebout *et al*. 2008*b*), while the split of the *M. ginkgodens* and *M. hotaula* lineages appears to be a more recent occurrence.

DBY7 fragments (241 bp) were successfully amplified from 12 *Mesoplodon* species. Due to the poor quality of much of the material available and the male-only nature of this marker, each species was represented by only a single specimen for these analyses. For this marker, most though not all, *Mesoplodon* species sampled possessed at least

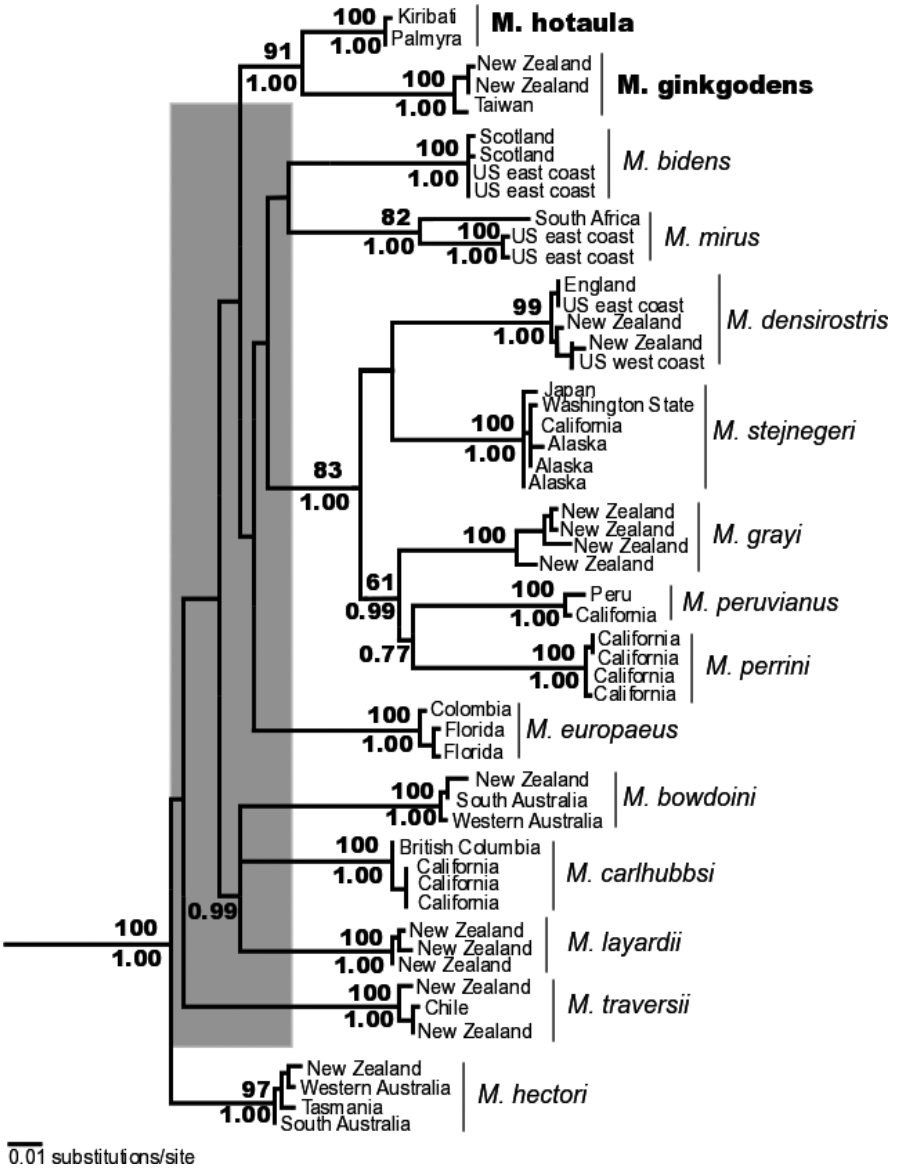



*Figure 2.* Maximum likelihood (ML) reconstruction of phylogenetic relationships among *Mesoplodon* beaked whales based on combined CR and CYB mtDNA sequences. Clade robustness is shown by bootstrap scores, BS (≥60%, above branches) and Bayesian posterior probabilities, BPP (≥0.90, below branches). Note strong support of all species-specific groupings (majority of bootstrap scores >80%, posterior probabilities >0.95) and consistent patterns of low intraspecific genetic variation and high interspecific genetic divergence in this group. Higher-level relationships between species were generally not well resolved by these markers (gray-shaded regions, most bootstrap scores <50%). A single tree was retained from the search (score –ln 4363.84942). Cuvier's beaked whale *Ziphius cavirostris* was used as an outgroup.

one diagnostic nucleotide substitution that distinguished them from the other species in the group (Table S3). Although three taxa did not yield species-specific sequences for this locus, both *M. ginkgodens* and *M. hotaula* represented unique lineages. *M. ginkgodens* possessed one nucleotide substitution (position 156) that distinguished it from all other *Mesoplodon* species, including *M. hotaula,* and *M. hotaula* possessed one nucleotide substitution (position 191) that reciprocally distinguished it from all other *Mesoplodon* species, including *M. ginkgodens.* For the other sister-species pairs sampled, two substitutions distinguished *M. europeaus* from *M. mirus*, and no substitutions distinguished *M. bowdoini* from *M. carlhubbsi.*

## Description

Order Cetartiodactyla Montgelard, Catzefils and Douzery 1997
Cetacea (Brisson 1762)
Family Ziphiidae Gray 1865
Genus *Mesoplodon* Gervais 1850
*Mesoplodon hotaula* Deraniyagala 1963*a*, *b*

### *Holotype*

Adult female (3WZ5): skull, mandible, tooth, and some postcranial elements held at the National Museum, Colombo, Sri Lanka. Specimen washed ashore "in a dying condition" on 26 January 1963. Collected, described, and named by P. E. P. Deraniyagala (1963*a*, *b*).[4] See also Deraniyagala (1965).

### *Type Locality*

Ratmalana (6°49′N, 79°52′E), approxiately 8 km south of Colombo, on the west coast of Sri Lanka.

### *Paratypes*

Adult female (USNM593418): skull, jaw, teeth, and postcranial elements held at the Smithsonian National Museum of Natural History, Washington, DC. Tissue sample (bone powder, SW53473) held at US National Marine Fisheries Southwest Fisheries Science Center (SWFSC). Collected from Cooper Island, Palmyra Atoll, one of the Northern Line Islands, *ca.* 1,770 km SW of Honolulu, Hawaii, by staff of The Nature Conservancy.[5]

Subadult, possible female (USNM593414): skull only held at the Smithsonian National Museum of Natural History, Washington, DC. Tissue sample (bone powder, SW53474) held at SWFSC. Collected from Eastern Island, Palmyra Atoll, by staff of The Nature Conservancy.

[4]Note typographical error in Deraniyagala (1963*a*); figure is labeled incorrectly as *Mesoplodon ulhota.*

[5]Palmyra Atoll was purchased by The Nature Conservancy in 2000. In 2001, the US Fish and Wildlife Service designated the coral reef habitat surrounding this atoll as a National Wildlife Refuge. Baumann-Pickering *et al.* (2010) recorded echolocation signals from an unknown species of beaked whale at Palmyra Atoll, which they suggest could be attributed to *M. hotaula.*

Adult male (USNM593426): skull and mandible held at the Smithsonian National Museum of Natural History, Washington, DC. Tissue sample (bone powder, SW70984) held at SWFSC. Collected from eastern lagoon, Palmyra Atoll by staff of The Nature Conservancy. Teeth present at time of collection but were lost in transit to the Smithsonian.

Adult male (MDV-X): skull, mandible, teeth, and postcranial elements held in a private collection in the Republic of Maldives.

Adult male (MM-0001): skull, mandible, teeth, and postcranial elements held in the collection of the Island Conservation Society (ICS) Fondation pour la Conservation des Iles in Victoria, Seychelles. Collected from Desroches Beach, Seychelles on 20 June 2009 by L. and W. Thompson with assistance from ICS members.

Male (UKIRI): tissue sample held in the University of Auckland DNA and Tissue Archive, New Zealand, collected by Roger V. Grace from Tabiteuea Atoll in the Gilbert Islands, Republic of Kiribati in 2003. The dried meat was a gift from the islanders, a leftover from a recent festival feast. It was reportedly obtained from one of seven whales driven onto the beach and killed in October 2002 when the whales came into the shallow water of the lagoon (Baker *et al*. 2013). The description of this hunt is extremely unusual. Beaked whales generally only come into shallow water prior to stranding, and such behavior usually involves only single individuals or cow-calf pairs. With the exception of *M. grayi* (von Haast 1876, Reeves *et al.* 2002), beaked whales do not generally mass strand. The islanders reported that such events occurred several times a year and provided a common source of food, not just for ceremonial occasions. The whales were described as "long ones," *ca.* 15–20 ft (457–609 cm) in length. An expedition to Kiribati in June–July 2009 did not find any further remains on Tabiteuea Atoll, but recovered osteological material from one or more whales identified as *M. hotaula* through DNA analysis from nearby Onotoa Island (Baker *et al*. 2013).

### *Etymology*

The specific name, *hotaula*, is derived from the Sinhala words, *hota* = beak, and *ula* = pointed (Deraniyagala 1963*a*). Recommended pronunciation is as follows: *ho* as in hot, *ta* as in tuppence, *ul* as in school, and *a* as in uh. We propose that this species be known by the common name, "Deraniyagala's beaked whale."

## Diagnosis

### *Molecular Characters*

*M. hotaula* can be differentiated from *M. ginkgodens* and all other species of *Mesoplodon* beaked whales based on molecular genetic characters (Fig. 2, Tables S1–S3).

*Mitochondrial DNA*—In phylogenetic analyses of combined CR and CYB sequences (Fig. 2), *M. hotaula* specimens cluster together in a strongly supported clade (BS 100%, BPP 1.00) that is reciprocally monophyletic to the clade formed by the *M. ginkgodens* specimens. The number of apparently fixed nucleotide substitutions (diagnostic characters) that distinguish *M. hotaula* from *M. ginkgodens* (Table S1), together with the overall degree of genetic differentiation (% net divergence, Table S2), is similar to what is observed between other recognized *Mesoplodon* species: CR, $n = 18$ (dA, 3.6% ± 0.91%), CYB, $n = 26$ (dA, 8.2% ± 1.79%), COXI, $n = 49$

(dA, 5.5% ± 0.76%). Branch lengths in phylogenetic reconstructions using these sequences reflect this trend.

*Nuclear DNA*—Autosomal introns: a nucleotide substitution at position 2199 (ACT) distinguishes *M. ginkgodens* from *M. hotaula* and all other *Mesoplodon* species sampled. Y-chromosome DBY7: nucleotide substitutions at positions 156 and 191 distinguish *M. ginkgodens* and *M. hotaula* from one another, and from all other *Mesoplodon* species sampled (Table S3).

### *Morphological Characters*

The following characters of the teeth and skull are, when combined, diagnostic for *M. hotaula* (Fig. 3–6).

(1) Single pair of very large, triangular, laterally compressed mandibular teeth.
(2) Alveoli of the teeth fully posterior to the mandibular symphysis.
(3) Teeth with vertical growth form, taller than they are wide, asymmetric; posterior margin convex, anterior margin almost planar.
(4) Short mandibular symphysis (distal portions of mandibles appear "stubby").
(5) Greatest transverse span of combined premaxillary bones in adults ≥60 mm (Table S4, measurement 32) and "flattened" in cross section.

Character 1 is shared with *M. ginkgodens, M. bowdoini*, *M. carlhubbsi, M. densirostris,* and *M. stejnegeri.* Character 2 is shared with *M. ginkgodens, M. densirostris, M. peruvianus,* and *M. stejnegeri.* Characters 3, 4, and 5 distinguish *M. hotaula* from *M. ginkgodens*. Tooth form in adult males is particularly distinct. In contrast, the teeth of *M. ginkgodens* are generally wider than they are tall, both the posterior and anterior margins are convex, and they are nearly symmetrical (Fig. 5). In *M. ginkgodens*, the distal portion of the mandibles appears long and gracile (Fig. 4), and the greatest transverse span of the combined premaxillary bones at the midpoint of the length of the beak is greater than 40 mm but less than 60 mm (diagnostic feature 5, Moore and Gilmore 1965; also TKY, unpublished data). Further, the premaxillary bones in *M. ginkgodens* are angled upwards (*ca.* 30°–45°) rather than flattened (*ca.* 10°–15°) as in *M. hotaula* (Fig. 6). Adult male *M. ginkgodens* also appear to be larger in size (total length, 472–496 cm) than adult male *M. hotaula* (total length, 386–432 cm; Table 1). Additional images of *M. hotaula* and *M. ginkgodens,* together with details of skull and mandibular measurements, can be found in the online supplemental material (Fig. S1–S6).

### *External Appearance*

To date, we only have information on external appearance for two specimens; the holotype (3WZS), an adult female from Sri Lanka, and an adult male (MM-0001) from the Seychelles. The holotype, which was freshly dead when examined, was described as having a relatively compressed body, a strong lateral ridge, a slender head with an elongate beak, and eyes located about half a beak length behind the angle of the gape. It was blue gray ventrally and the tail had a median lobe with a small caudal notch. There was a single, unerupted pair of teeth in the mandible, located slightly behind the symphysis (Deraniyagala 1963*a*, *b*, 1965).

The Seychelles specimen (Fig. 7), also freshly dead, was similar in overall appearance, though the tail lacked a median notch. It was examined in the early morning, shortly after its discovery, by W. and L. Thompson. The specimen was blue-black dor-

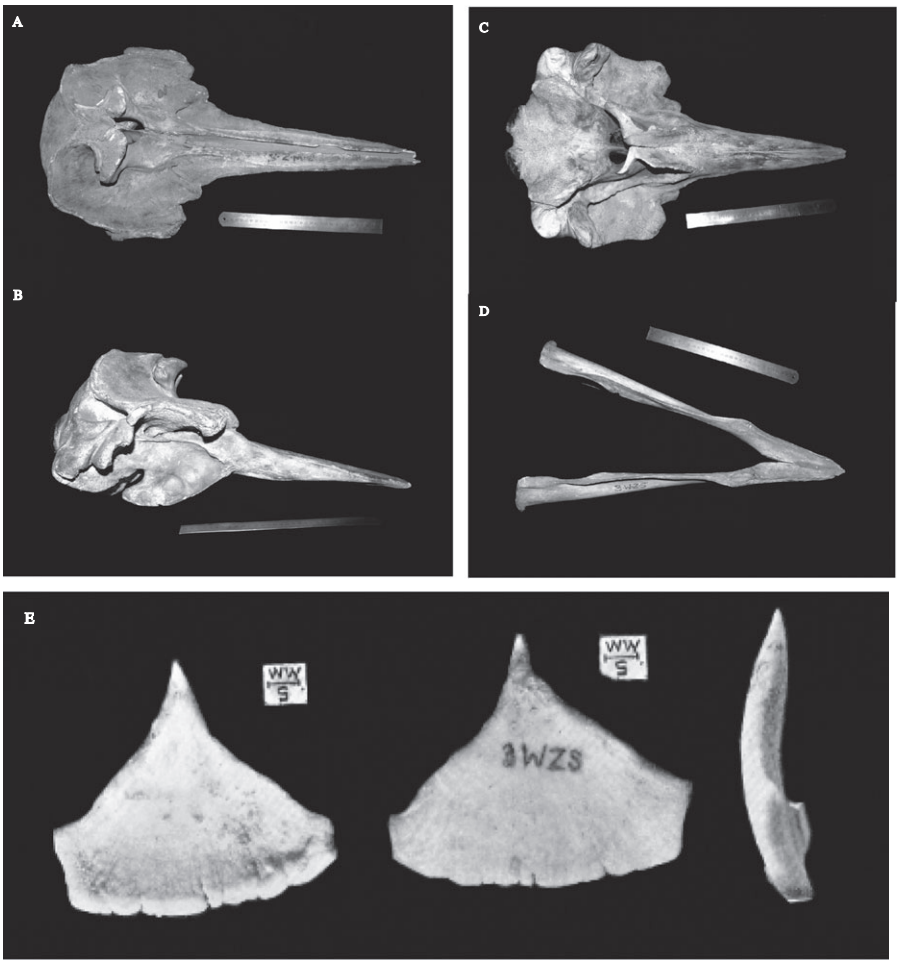



*Figure 3.* Holotype of *M. hotaula,* 3WZS, adult female, Sri Lanka: (A) Dorsal view, (B) lateral view, (C) ventral view, (D) mandibles with left tooth *in situ*, right tooth missing, (E) right tooth, labial, lingual, and posterior views. Scale for (A)–(D), 300 mm metal ruler. Scale for (E), 5 mm card tag. Photo credits: (A–D), R. C. Anderson, courtesy of National Museum of Colombo; (E) reproduced from Deraniyagala 1965, Plate III.

sally, grading to a slightly lighter shade ventrally. There were a small number of white cookie-cutter shark (*Isistius* spp.) scars on its ventral surface, predominantly at the posterior end. There were also two large, fresh shark bites; one out of the ventral peduncle and one out of the head-neck area, behind the blowhole. The blue-black color of the body continued on the head, forming a dark cap that extended along the anterior surface of the rostrum and to the posterior end of the mouth line. Coloration around the eye was a lighter mottled gray, becoming lighter ventrally. The tip of the lower jaw was gray but the lower jaw itself was predominantly white. This white color pattern extended on the lower jaw to behind the tooth and continued above the mouthline to the rostrum. The upper lips were whitish, grading to gray and blue-black on the rostrum. The gray mottling of the cheek and eye area formed a distinct wedge of color

against the white of the ventral chin and throat region, cutting across the posterior ends of the throat grooves (see also Fig. S5). Note that the tips of the teeth of both this and the Palmyra adult male (USNM593426) were broken (see also Fig. S3). This suggests that male:male combat using the teeth as weapons does occur in this species, although the Seychelles specimen did not have any of the white linear tooth rake scars that appear to result from such behavior (Mead *et al*. 1982, Heyning 1984).

In contrast, the two New Zealand adult male *M. ginkgodens* were brownish-gray dorsally (though blue-black coloration has been described from Japanese animals; Nishiwaki and Kamiya 1958, Nishiwaki *et al.* 1972), grading to lighter tones ventrally (Fig. 7, see also Fig. S6). There was a darker patch around the eye that extended further in front and a bit below the eye. The beak was white-tipped, both upper and lower jaws. This white coloration reappeared on the upper lip behind the tooth. Where *M. hotaula* appears to have a gray tip to the lower jaw and a white chin and throat region (Fig. 7, arrows), *M. ginkgodens* appears to have a white tip to the lower jaw and a gray-brown chin and throat region (Fig. 8, arrows). However it is difficult to tell from such a small number of animals whether these are fixed color pattern differences or individual variation. These suggested differences need to be confirmed from additional fresh strandings or sightings of living whales at sea.

### *Distribution*

*M. hotaula* has an equatorial distribution in the Indo-Pacific (Fig. 1), which broadly overlaps, or is mostly parapatric, with the more temperate distribution of *M. ginkgodens*. The majority of confirmed records of *M. ginkgodens* are from temperate and cold-temperate waters such as those around Japan, Taiwan, and New Zealand. Based on tooth form, the record from the Galapagos (Palacios 1996) also appears to represent *M. ginkgodens* proper, which is not surprising given the cold Humboldt Current that flows around these islands. There are three further tropical records for *M. ginkgodens*; one from the Federated States of Micronesia and one from the Marshall Islands (both based on DNA analysis; Dalebout *et al.* 2008*a*; KMR, unpublished data), and one from the Republic of Maldives, based on tooth form (Anderson *et al.* 1999). The latter specimen consists of a single tooth held in the Maldives National Museum, with no information on its provenance (Anderson *et al.* 1999). The species identity of this specimen should be re-examined, but could suggest a zone of overlap in distribution of the two species around the Maldives.

## Discussion

The genetic and morphological evidence presented here supports recognition of *M. hotaula* and *M. ginkgodens* as full species based on the four criteria of the GCC (Avise and Ball 1990, Avise and Wollenberg 1997, Avise 2000):

(1) Concordance across sequence characters within genetic locus leading to conclusive exclusion—Over the three mtDNA loci, 93 nucleotide substitutions (CR 18, CYB 26, COX1 49) were observed between *M. hotaula* and *M. ginkgodens*, which appear to represent fixed differences based on a limited number of specimens. For the CYB, this included four nonsynonymous substitutions, which would translate to amino-acid level differences in this key metabolic protein. Two nucleotide substitutions further distinguished these species at the Y-chromosome and

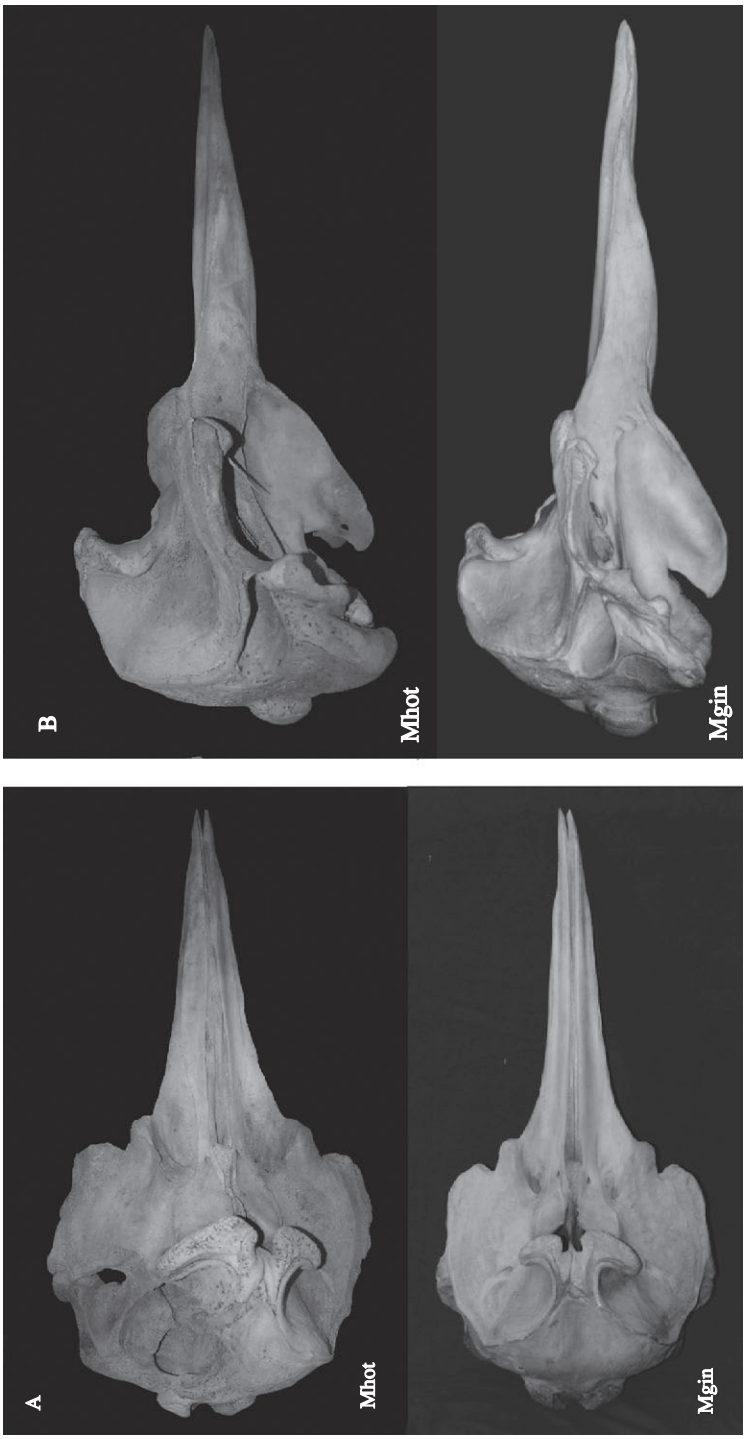



*Figure 4.* Comparison of adult males; crania and mandibles: *M. botaula* (MDV-X, Maldives) and *M. ginkgodens* (MginTSM8744, holotype). (A) Dorsal view, (B) lateral view, (C) ventral view, (D) mandibles. Photo credits: *M. botaula*, R. C. Anderson; *M. ginkgodens*, T. K. Yamada.

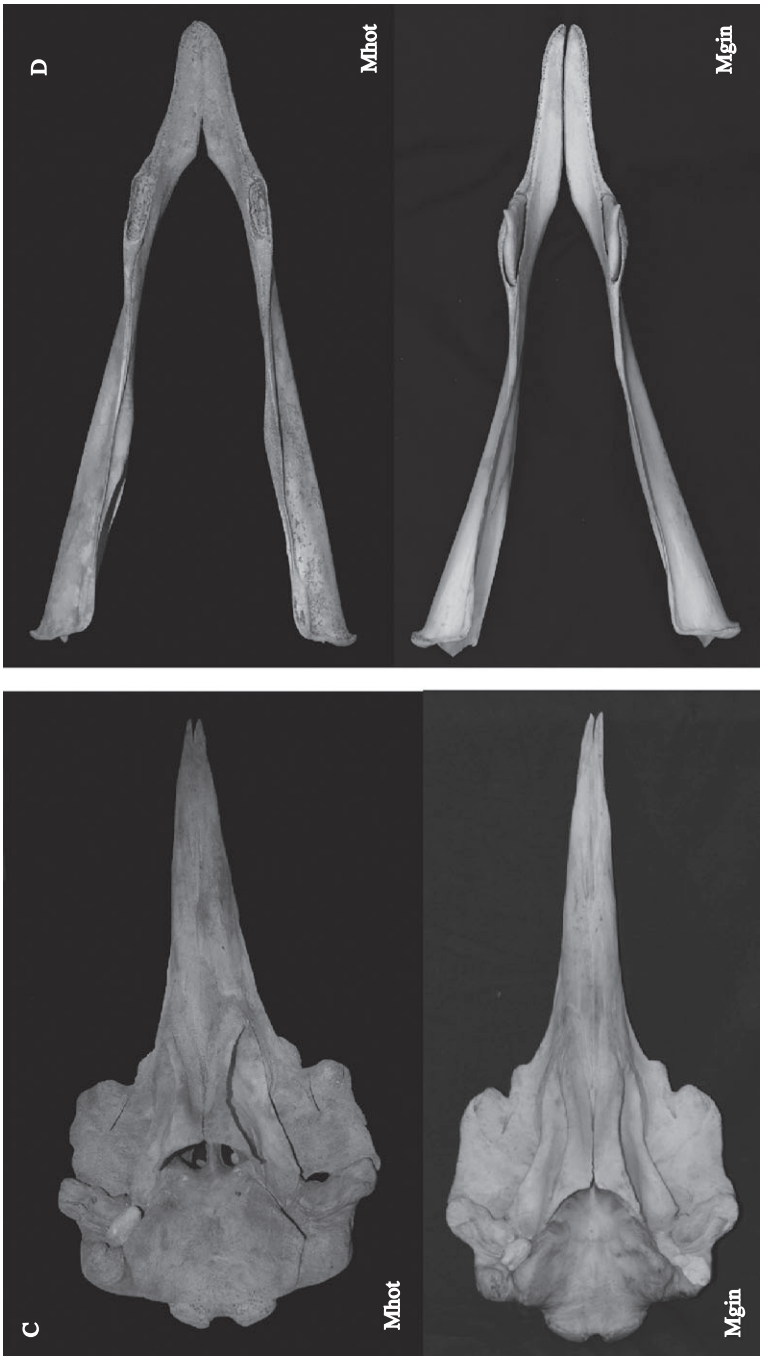



*Figure 4.* (Continued).

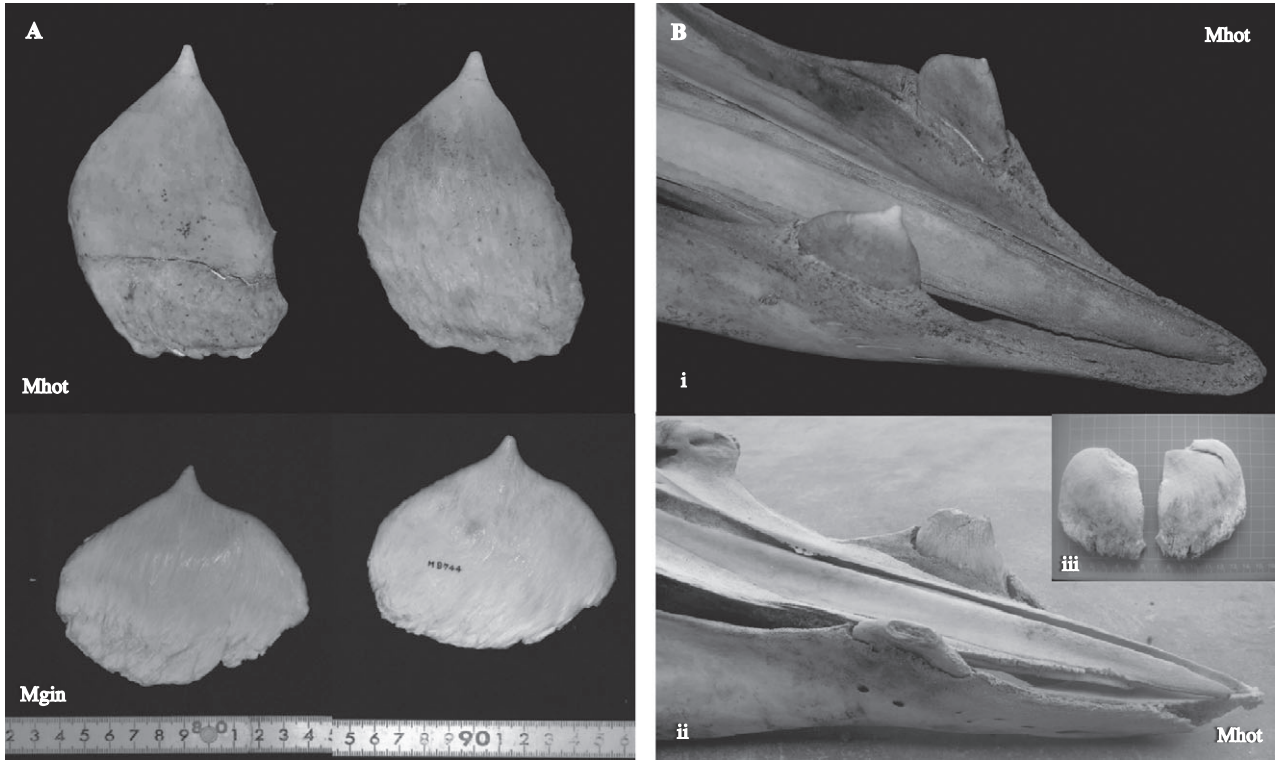



*Figure 5.* Comparison of adult male teeth: *M. hotaula* (MDV-X, Maldives) and *M. ginkgodens* (MginTSM8744, holotype). (A) L tooth, medial face; R tooth, lateral face. Leading edges facing right. (B) *M. hotaula* adult males, distal end of skull and mandibles with teeth *in situ*. (i) MDV-X, Maldives; (ii) MM-0001, Seychelles; (iii) teeth of MM-0001. Photo credits: MDV-X, R. C. Anderson; MM-0001, L. Thompson; MginTSM8744, T. K. Yamada.

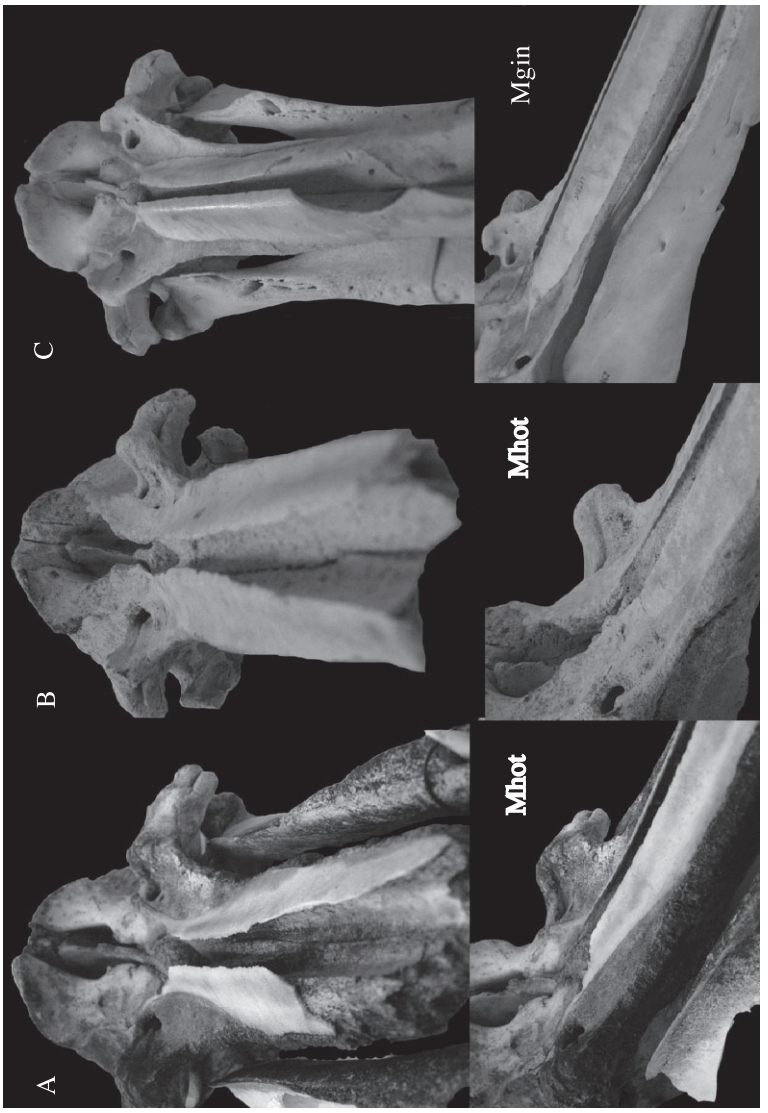



*Figure 6.* Comparison of premaxilla morphology; distal and oblique lateral views of rostrum: (A) *M. botaula* USNM593418, (B) *M. botaula* USNM593414, (C) *M. ginkgodens* USNM298237. Photo credit: M. L. Dalebout.

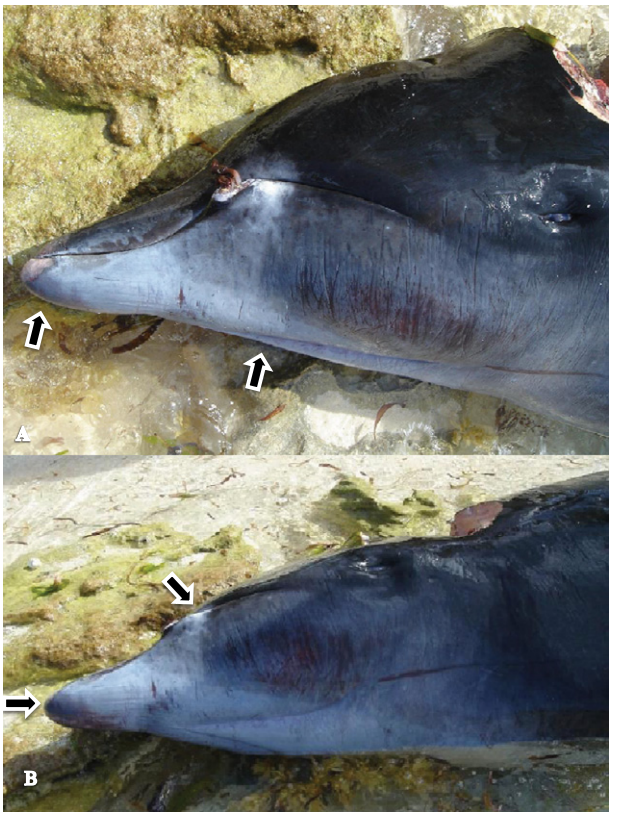



*Figure 7.* External appearance of *M. hotaula*: MM-0001, adult male, Seychelles. (A) Lateral view of head, (B) antero-ventral view of head. Photo credit: L. Thompson.

autosomal introns surveyed. Lower levels of divergence are expected at single copy nuclear loci due to the slower rate of evolution of these markers (Hare 2001). Overall, the levels of divergence observed were similar to but somewhat lower than those between other known sister-species pairs in this genus. Divergence date estimates for these other species range from 5.3 (*M. perrini* and *M. peruvianus*) to 10.4 Mya (*M. bowdoini* and *M. carlhubbsi*; Dalebout *et al.* 2008*b*), while the split of the *M. ginkgodens* and *M. hotaula* lineages appears to be a more recent occurrence.

(2) Concordance in genealogical patterns across multiple loci, both mitochondrial and nuclear—Phylogenies constructed from mtDNA lineages (haplotypes) showed a strongly supported pattern of reciprocal monophyly. Divergence at the nuclear loci was too low to be detected by such analyses but the character-based analyses were concordant with patterns observed in the mtDNA. Based on these results, there was no evidence of mitochondrial (maternal), Y-chromosome (paternal), or autosomal (bi-parental) gene flow between these taxa.

(3) Concordance with biogeographical patterns—The distributions of *M. hotaula* and *M. ginkgodens* appear to be largely parapatric, with *M. hotaula* found in more

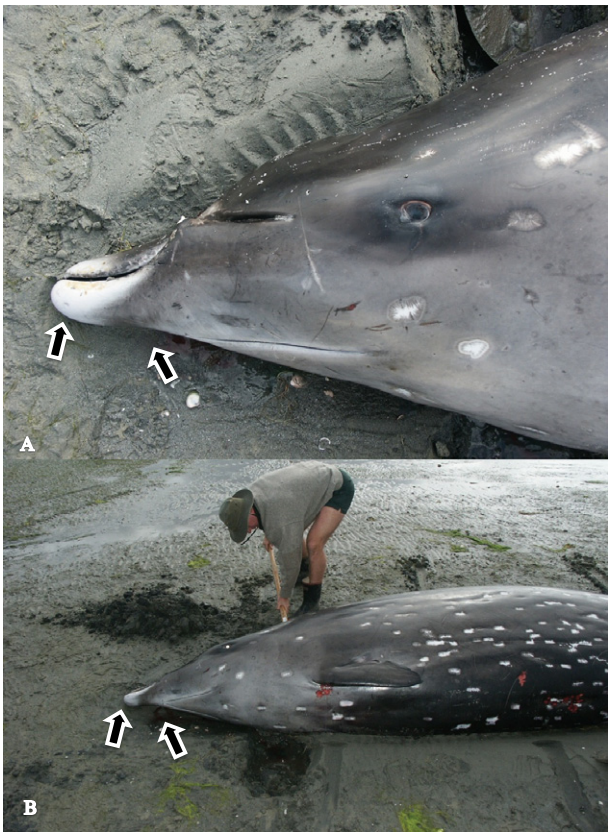



*Figure 8.* External appearance of *M. ginkgodens*: MginNZ04, adult male, New Zealand. (A) Lateral view of head, (B) antero-ventral view of body. Photo credit: H. Stoffregen, courtesy of New Zealand Department of Conservation.

tropical waters and *M. ginkgodens* found predominantly in more temperate waters, but with some zones of potential overlap (*e.g.,* in the Republic of Maldives).

(4) Concordance with morphological characters—*M. hotaula* and *M. ginkgodens* differ in features of the teeth, mandibles, and cranium, as well as potentially in color pattern. Together, these differences, while perhaps comparatively subtle, indicate that there is indeed morphological divergence between the species.

It is recognized that small sample sizes and limited geographic sampling can lead to underestimates of intra-specific genetic variability (Meyer and Paulay 2005). When evaluating the utility of the COXI gene for cetacean species identification, Viricel and Rosel (2012) also faced this issue. Although unable to consistently differentiate between closely related taxa in the *Stenella-Delphinus-Tursiops* complex, that study also found that diagnostic characters (species-specific sequences) distinguished all *Mesoplodon* species sampled. Using regression analyses, they also found that there was no significant relationship between the number of individuals analyzed per species and mean intraspecific diversity in their data sets (Viricel and Rosel 2012).

Further, they noted that cetaceans generally show relatively low levels of genetic variability (Shimura and Numachi 1987; Schlotterer *et al.* 1991; Dalebout *et al.* 2004, 2007; Kingston and Rosel 2004). In mysticetes, low rates of molecular evolution have been attributed to a combination of low metabolic rate, large body size, and long generation times (Jackson *et al.* 2009). Slower rates of molecular evolution will limit the amount of genetic diversity that can accumulate within a species and the amount of divergence that can accumulate between species over a given time. Previous studies of mtDNA CR and CYB in beaked whales have shown that intraspecific diversity is generally low, while inter-specific divergence is generally considerably higher, with little overlap (Dalebout *et al.* 2002, 2004, 2006, 2007). So while examination of additional specimens of *M. hotaula* and *M. ginkgodens* would be very useful, the low intraspecific diversity of beaked whales, and the low level of genetic variability in cetaceans overall, will together have reduced the potential biases of limited sampling in this study.

At a comprehensive, specialist workshop on cetacean taxonomy, evidence of "irreversible divergence" was considered of primary importance in the recognition and delimitation of species (Reeves *et al.* 2004). Irreversible divergence was considered to require at least two independent lines of evidence. Genetic characters from unlinked loci were considered to represent multiple lines of evidence. Confirming the independence of multiple morphological characters is difficult, and morphology was therefore considered to represent only a single line of evidence. Based on these guidelines, the proposed species-level ranking of both *M. hotaula* and *M. ginkgodens* is supported by multiple lines of evidence: mtDNA, nuclear autosomal DNA, Y-chromosome DNA, and morphology.

No guidelines were offered by the workshop on the degree of genetic divergence required to warrant species status. Here we have used a comparative approach based on the patterns of divergence observed between other recognized *Mesoplodon* species in phylogenetic reconstructions and pairwise distances to evaluate the proposed species status of *M. hotaula*. For the mtDNA, the divergence of *M. hotaula* and *M. ginkgodens* (CR 3.6%, CYB 8.2%, COXI 5.5%) was within the range observed for other *Mesoplodon* species, though generally on the lower end of the scale. MtDNA divergence between species in this group is, however, considerably higher on average than that observed between many other recognized cetacean species. For example, the net CR divergence between Chilean and Commerson's dolphins (*Cephalorynchus commersoni* and *C. eutropia*) was 2.5% over 442 bp, with three fixed differences (Pichler *et al.* 2001*b*), while the net CYB divergence between dusky and Pacific white-sided dolphins (*Lagenorhynchus obscurus* and *L. obliquidens*) was 1.2% over 496 bp, with five fixed differences (Hare *et al.* 2002). The lower levels of mtDNA divergence among these delphinids are consistent with speciation events initiated in the Pleistocene, approximately 1–1.5 Mya (Avise *et al.* 1998, Caballero *et al.* 2007). While not as deep as some of the other sister species in the genus *Mesoplodon*, it is clear that the divergence of the lineages representing *M. hotaula* and *M. ginkgodens* began well before the Pleistocene, and constitutes an ongoing and "irreversible" trend.

In conclusion, we present genetic and morphological evidence demonstrating that *M. hotaula* is a valid species, closely related to, but distinct from, *M. ginkgodens*. It is known to occur from at least the Seychelles to Palmyra Atoll, and likely ranges right across the tropical Indo-Pacific. This discovery brings the total number of *Mesoplodon* species to 15, making this by far the most speciose cetacean genus, although it remains among the least known.

## Acknowledgments

For collection and access to samples and specimens, we thank Roger V. Grace (Kiribati), NOAA Pacific Islands Regional Office, U.S. Fish and Wildlife Service, and the Nature Conservancy (Palmyra Atoll), New Zealand Department of Conservation field center staff (NZ DoC), John Wang, FormosaCetus, Taiwan, Janette Norman and Wayne Longmore, Museum Victoria, Melbourne, Australia, and Abdullah Asif Waheed, Maldives. We thank the Director General of the Department of Wildlife Conservation of Sri Lanka, and the Director of Departments of the National Museums of Sri Lanka for permission to sample the *M. hotaula* holotype and tissue-export permits. We thank Bob Pitman, US NMFS Southwest Fisheries Science Center, for discussion regarding color pattern differences, Anton van Helden, National Museum of New Zealand Te Papa Tongarewa for photographs and discussion regarding *M. ginkgodens*. This manuscript benefitted from comments by Randall Reeves, IUCN Cetacean Specialist Group, and three anonymous reviewers. For additional photographs, we thank Hans Stoffregen and Bryan Williams (NZ DoC). Partial funding for laboratory analyses was provided by grants to CSB from the US Marine Mammal Commission and the National Geographic Society. MLD is a Visiting Fellow at the University of New South Wales, Sydney, Australia.

## Literature Cited

Andersen, M., and C. C. Kinze. 1999. Annotated checklist and identification key to the whales, dolphins, and porpoises (Order Cetacea) of Thailand and adjacent waters. Natural History Bulletin of the Siam Society 47:27–62.

Anderson, R. C., A. Shaan and Z. Waheed. 1999. Records of cetacean 'strandings' from the Maldives. Journal of South Asian Natural History 4:187–202.

Avise, J. C. 1994. Molecular markers, natural history and evolution. Chapman & Hall, New York, NY.

Avise, J. C. 2000. Phylogeography: The history and formation of species. Harvard University Press, Cambridge, MA.

Avise, J. C., and R. M. Ball. 1990. Principles of genealogical concordance in species concepts in a biological taxonomy. Oxford Surveys in Evolutionary Biology 7:45–67.

Avise, J. C., and K. Wollenberg. 1997. Phylogenetics and the origin of species. Proceedings of the National Academy of Sciences of the United States of America 94:7748–7755.

Avise, J. C., D. Walker and G. C. Johns. 1998. Speciation durations and Pleistocene effects on vertebrate phylogeography. Proceedings of the Royal Society of London, Series B 265:1707–1712.

Baker, C. S., R. W. Slade, J. L. Bannister, *et al*. 1994. Hierarchical structure of mitochondrial DNA gene flow among humpback whales *Megaptera novaeangliae*, world-wide. Molecular Ecology 3:313–327.

Baker, C. S., A. Hutt, K. Thompson, M. L. Dalebout, J. Robins and G. S. Stone. 2013. Species identity and human consumption of beaked whales in the Gilbert Islands, Republic of Kiribati. Animal Conservation 16:641–647.

Baumann-Pickering, S., S. M. Wiggins, E. H. Roth, M. A. Roch, H.-U. Schnitzler and J. A. Hildebrand. 2010. Echolocation signals of a beaked whale at Palmyra Atoll. Journal of the Acoustic Society of America 127:3790–3799.

Boom, R., C. J. A. Sol, M. M. M. Salimans, C. L. Jansen, P. M. E. Wertheim-van Dillen and J. van der Noorda. 1990. Rapid and simple method for purification of nucleic acids. Journal of Clinical Microbiology 28:495–503.

Caballero, S., F. Trujillo, J. A. Vianna, *et al*. 2007. Taxonomic status of the genus *Sotalia*: Species level ranking for "tucuxi" (*Sotalia fluviatilis*) and "costero" (*Sotalia guianensis*) dolphins. Marine Mammal Science 23:358–386.

Chantrapornsyl, S., K. Adulyanukosol and K. Kittiwa'itanawong. 1999. Stranded cetaceans from Thailand. IBI Reports 9:55–72.

Cracraft, J. 1989. Speciation and its ontology: The empirical consequences of alternative species concepts for understanding patterns and processes of differentiation. Pages 28–59 *in* D. Otte and J. A. Endler, eds. Speciation and its consequences. Sinauer Associations, Sunderland, MA.

Dalebout, M. L., J. G. Mead, C. S. Baker, A. N. Baker and A. L. van Helden. 2002. A new species of beaked whale *Mesoplodon perrini* sp. n. (Cetacea: Ziphiidae) discovered through phylogenetic analyses of mitochondrial DNA sequences. Marine Mammal Science 18:577–608.

Dalebout, M. L., G. J. B. Ross, C. S. Baker, *et al*. 2003. Appearance, distribution and genetic distinctiveness of Longman's beaked whale, *Indopacetus pacificus*. Marine Mammal Science 19:421–461.

Dalebout, M. L., C. S. Baker, V. G. Cockcroft, J. G. Mead and T. K. Yamada. 2004. A comprehensive and validated molecular taxonomy of beaked whales, Family Ziphiidae. Journal of Heredity 95:459–473.

Dalebout, M. L., D. E. Ruzzante, H. Whitehead and N. I. Øien. 2006. Nuclear and mitochondrial markers reveal the distinctiveness of a small population of bottlenose whales (*Hyperoodon ampullatus*) in the western North Atlantic. Molecular Ecology 15:3115–3129.

Dalebout, M. L., C. S. Baker, D. Steel, *et al*. 2007. A divergent mtDNA lineage among *Mesoplodon* beaked whales: Molecular evidence for a new whale in the Tropical Pacific? Marine Mammal Science 23:954–966.

Dalebout, M. L., K. M. Robertson, S. J. Chivers and A. Samuels. 2008*a*. DNA identification and the impact of illegal, unregulated, and unreported (IUU) fishing on rare whales in Micronesian waters. Micronesica 40:139–147.

Dalebout, M. L., D. Steel and C. S. Baker. 2008*b*. Phylogeny of the beaked whale genus *Mesoplodon* (Ziphiidae: Cetacea) revealed by nuclear introns: Implications for the evolution of male tusks. Systematic Biology 57:857–875.

Davis, J. I., and K. C. Nixon. 1992. Populations, genetic variation and the delimitation of phylogenetic species. Systematic Biology 41:421–435.

De Queiroz, K. 2007. Species concepts and species delimitation. Systematic Biology 56:879–886.

Deraniyagala, P. E. P. 1963*a*. A new beaked whale from Ceylon. Ceylon Today 12: 13–14.

Deraniyagala, P. E. P. 1963*b*. Mass mortality of the new subspecies of little piked whale *Balaenoptera acutorostrata thalmaha* and a new beaked whale *Mesoplodon hotaula* from Ceylon. Spolia Zeylanica 30:79–84.

Deraniyagala, P. E. P. 1965. Comparison of *Mesoplodon hotaula* Deraniyagala with *Ziphius cavirostris indicus* (van Beneden). Spolia Zeylanica 30:3–12.

Hare, M. 2001. Prospects for nuclear gene phylogeography. Trends in Ecology and Evolution 16:700–706.

Hare, M. P., F. Cipriano and S. R. Palumbi. 2002. Genetic evidence on the demography of speciation in allopatric dolphin species. Evolution 56:804–816.

Heyning, J. E. 1984. Functional morphology involved in intraspecific fighting of the beaked whale, *Mesoplodon carlhubbsi*. Canadian Journal of Zoology 62:1645–1654.

Hellborg, L., and H. Ellegren. 2003. Y chromosome conserved anchor tagged sequences (YCATS) for the analysis of mammalian male-specific DNA. Molecular Ecology 12:283–291.

Höss, M., and S. Pääbo. 1993. DNA purification from Pleistocene bones by a silica-based purification method. Nucleic Acids Research 21:3913–3914.

Huelsenbeck, J. P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.

Jackson, J. A., C. S. Baker, M. Vant, D. J. Steel, L. Medrano-Gonzales and S. R. Palumbi. 2009. Big and slow: Phylogenetic estimates of molecular evolution in baleen whales (Suborder Mysticeti). Molecular Biology and Evolution 26:2427–2440.

Kingston, S. E., and P. E. Rosel. 2004. Genetic differentiation among recently diverged delphinid taxa determined using AFLP markers. Journal of Heredity 95:1–10.

Kumar, S., K. Tamura and M. Nei. 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Briefings in Bioinformatics 5:150–163.

Mayr, E. 1963. Animal species and evolution. Harvard University Press, Cambridge, MA.

Matisoo-Smith, E., J. S. Allen, T. N. Ladefoged, R. M. Roberts and D. M. Lambert. 1997. Ancient DNA from Polynesian rats: Extraction, amplification and sequence from single small bones. Electrophoresis 18:1534–1537.

Mead, J. G., W. A. Walker and W. J. Houck. 1982. Biological observations on *Mesoplodon carlhubbsi* (Cetacea: Ziphiidae). Smithsonian Contributions to Zoology 344. 25 pp.

Meyer, C. P., and G. Paulay. 2005. DNA barcoding: Error rates based on comprehensive sampling. PLoS Biology 3:e422.

Moore, J. C. 1963. Recognizing certain species of beaked whales of the Pacific Ocean. American Midland Naturalist 70:396–428.

Moore, J. C., and R. M. Gilmore. 1965. A beaked whale new to the Western Hemisphere. Nature 205:1239–1240.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York, NY.

Nishiwaki, M., and T. Kamiya. 1958. A beaked whale *Mesoplodon* stranded at Oiso Beach, Japan. Scientific Reports of the Whales Research Institute, Tokyo 13:53–83.

Nishiwaki, M., T. Kasuya, K. Kureha and N. Oguro. 1972. Further comments on *Mesoplodon ginkgodens*. Scientific Reports of the Whales Research Institute, Tokyo 24:45–56.

Palacios, D. M. 1996. On the specimen of the ginkgo-toothed beaked whale, *Mesoplodon ginkgodens*, from the Galapagos Islands. Marine Mammal Science 12:444–446.

Pichler, F. B., M. L. Dalebout and C. S. Baker. 2001*a*. Non-destructive DNA extraction from sperm whale teeth and scrimshaw. Molecular Ecology Notes 1:106–109.

Pichler, F. B., D. Robineau, R. N. P. Goodall, M. A. Meyer, C. OLavarria B. and C. S. Baker. 2001*b*. Origin and radiation of Southern Hemisphere coastal dolphins (genus *Cephalorhynchus*). Molecular Ecology 10:2215–2223.

Posada, D., and K. A. Crandall. 1998. MODELTEST: Testing the model of DNA substitution. Bioinformatics Applications Note 14:817–818.

Reeves, R. R., S. S. Brent, P. J. Clapham, J. A. Powell and P. A. Folkens. 2002. National Audubon Society guide to marine mammals of the world. Chanticleer Press, New York, NY.

Reeves, R. R., W. F. Perrin, B. L. Taylor, C. S. Baker and S. L. Mesnick. 2004. Report of the workshop on shortcomings of cetacean taxonomy in relation to needs of conservation and management, 30 April–2 May 2004, La Jolla, CA. NOAA Technical Memorandum, NOAA-TM-NMFS-SWFSC-363. 94 pp.

Rohland, N., and M. Hofreiter. 2007. Comparison and optimization of ancient DNA extraction. Biotechniques 42:343–352.

Ross, G. J. B. 1969. Evidence for a southern breeding population of True's beaked whale. Nature 222:585.

Sambrook, E., F. Fritsch and T. Maniatis. 1989. Molecular cloning. Cold Spring Harbor Press, Cold Spring Harbor, NY.

Shimura, E., and K.-I. Numachi. 1987. Genetic variability and differentiation in toothed whales. Scientific Reports of the Whales Research Institute, Tokyo 38:141–163.

Schlotterer, C., C. B. Amos and D. Tautz. 1991. Conservation of polymorphic simple sequence loci in cetacean species. Nature 354:63–65.

Swofford, D. L. 2003. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, MA.

Thompson, K., C. S. Baker, A. van Helden, S. Patel, C. Millar and R. Constantine. 2012. The world's rarest whale. Current Biology 22:R905–R906.

van Helden, A. L., A. N. Baker, M. L. Dalebout, J. C. Reyes, K. Van Waerebeek and C. S. Baker. 2002. Resurrection of *Mesoplodon traversii* (Gray, 1874), senior synonym of *M. bahamondi* Reyes, Van Waerebeek, Cárdenas and Yañez, 1995 (Cetacea: Ziphiidae). Marine Mammal Science 18:609–621.

Viricel, A., and P. E. Rosel. 2012. Evaluating the utility of cox1 for cetacean species identification. Marine Mammal Science 28:37–62.

von Haast, J. 1876. On a new ziphioid whale. Proceedings of the Zoological Society of London 4 January:7–13.

Werle, E., C. Schneider, M. Renner, M. Volker and W. Fiehn. 1994. Convenient single-step, one tube purification of PCR products for direct sequencing. Nucleic Acids Research 22:4354–4355.

Zachosa, F. E., M. Apollonio, E. V. Bärmann, *et al*. 2013. Species inflation and taxonomic artefacts—a critical comment on recent trends in mammalian classification. Mammalian Biology 78:1–6.

Received: 1 March 2013
Accepted: 16 November 2013

## Supporting Information

The following supporting information is available for this article online at http://onlinelibrary.wiley.com/doi/10.1111/mms.12113/suppinfo.

*Analytical Methods*

*Figure S1*. External appearance of *M. hotaula* holotype adult female (3WZS). Clockwise from top; right ventro-lateral view of anterior half (length of section, 205.7 cm); dorsal view of flukes with caudal notch; dorsal view of head; left side of head. Reproduced from Deraniyagala (1965), Plate I.

*Figure S2.* External appearance, artist's impression of *M. hotaula* holotype (3WZS). Reproduced from Deraniyagala (1965), fig. 2.

*Figure S3.* Skulls of *M. hotaula* paratypes. (A) USNM593418, adult female, Palmyra Atoll: (i) dorsal view, (ii) lateral view, (iii) ventral view, (iv) mandibles with teeth *in situ.* Scale, pen 138 mm (i–iii), US nickel, 21 mm (iv). Photo credit: M. L. Dalebout. (B) USNM593414, subadult female?, Palmyra Atoll: (i) dorsal view (scale, US nickel, 21 mm), (ii) lateral view (scale, 138 mm pen). Scale, pen 138 mm (i), US nickel, 21 mm (ii). Photo credit: M. L. Dalebout. (C) USNM593426, adult male, Palmyra Atoll: (i) dorsal view of skull and mandibles, teeth removed, (ii) lateral view of skull and mandibles, teeth removed, (iii) mandibles with teeth *in situ*. Photo credit: A. Hoke. (D) MM-0001, adult male, Seychelles: (i) dorsal view of skull and mandibles with teeth *in situ,* (ii) lateral view of skull and mandibles with teeth *in situ,* (iii) ventral view of skull, (iv) lateral view of mandibles with teeth *in situ.* Photo credit: L. Thompson.

*Figure S4*. Other *M. ginkgodens* skulls. (A) MginNZ04, adult male, New Zealand: (i) dorsal view, (ii) lateral view, (iii) mandibles with teeth *in situ* lateral view, (iv) mandibles with teeth *in situ* ventral view. Courtesy of A. van Helden and the Museum of New Zealand Te Papa Tongarewa. (B) MginUSNM298237, adult female, California: (i) dorsal view of skull (scale, 300 mm ruler), (ii) dorsal view of mandibles with right tooth *in situ* (scale, US nickel, 21 mm). Photocredit: M. L. Dalebout.

*Figure S5.* Additional images, external appearance of *M. hotaula,* MM-0001, adult male, Seychelles. (A) Ventral view of stranded animal, (B) postero-lateral view, (C) antero-dorsal view of head. Photo credit: L. Thompson.

*Figure S6.* Additional images, external appearance of *M. ginkgodens,* MginNZ03, adult male, New Zealand. (A) Lateral view of head, (B) anterior view of head. Photo-credit: B. Williams, courtesy of New Zealand Department of Conservation.

*Table S1*. MtDNA data. (A) Control region (CR), variable sites over 658 bp. (B) Cytochrome *b* (CYB), variable sites over 384 bp. (C) Cytochrome oxidase I (COXI), variable sites over 987 bp. Gray shading highlights the nucleotide substitutions differentiating *M. hotaula* from *M. ginkgodens*.

*Table S2*. Pairwise net divergence between species, Kimura 2-parameter distances, as percentages, below diagonal. SE, above diagonal. Values for sister-species pairs highlighted in gray with bold type. (A) Control region, CR (435 bp); (B) cytochrome *b*, CYB (384 bp); (C) cytochrome oxidase I, COXI (958 bp). See footnotes of Table S2A for translation of species codes.

*Table S3*. Y-chromosome intron, DBY7. Variable sites over 241 bp. The diagnostic nucleotide substitutions that distinguish *M. ginkgodens* and *M. hotaula* from each another, and from the other *Mesoplodon* species sampled are highlighted in gray.

*Table S4A*. Cranial measurements for *Mesoplodon hotaula*. Measurements (in mm) are taken on the right hand side (R) where possible, following Moore (1963). Where two measurements are given, R, then L. E, estimated length. See Table S3B for definitions of measurements.

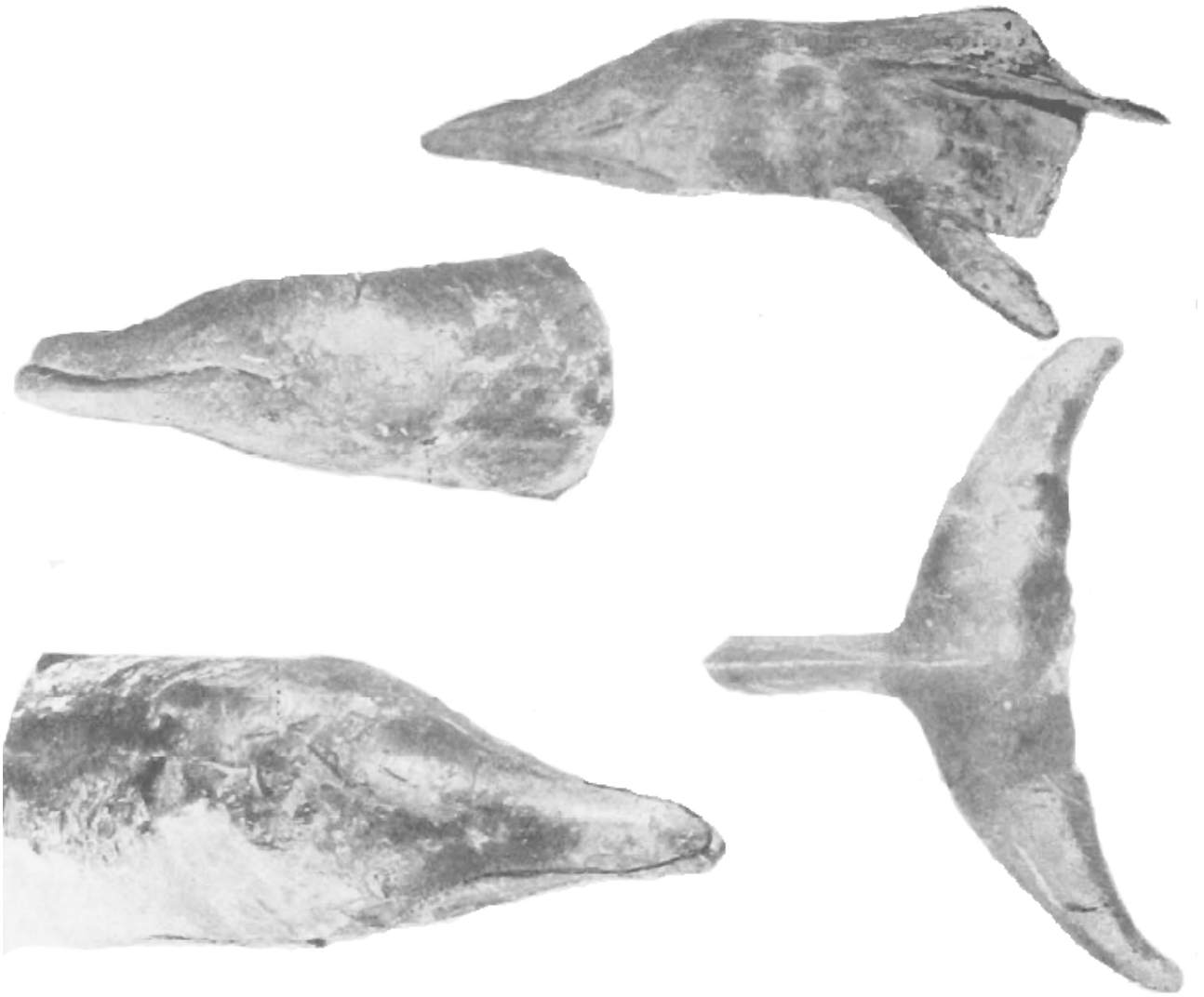
*Table S4B*. Definitions of cranial measurements. Numbers in parentheses refer to Moore (1963).

*Table S5A*. Mandibular measurements for *Mesoplodon hotaula*. Measurements (in mm) are taken on the right hand side (R) where possible, following Moore (1963). Where two measurements are given, R, then L. E, estimated length. See Table S4B for definitions of measurements.

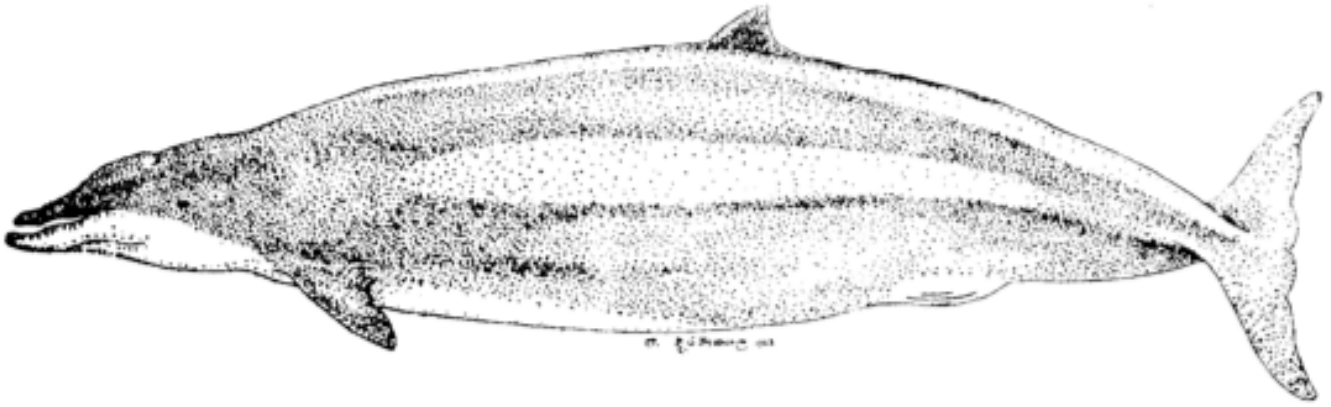
*Table S5B*. Definitions of mandibular measurements. Numbers in parentheses refer to Moore (1963).

**Supplementary Material – Figures**

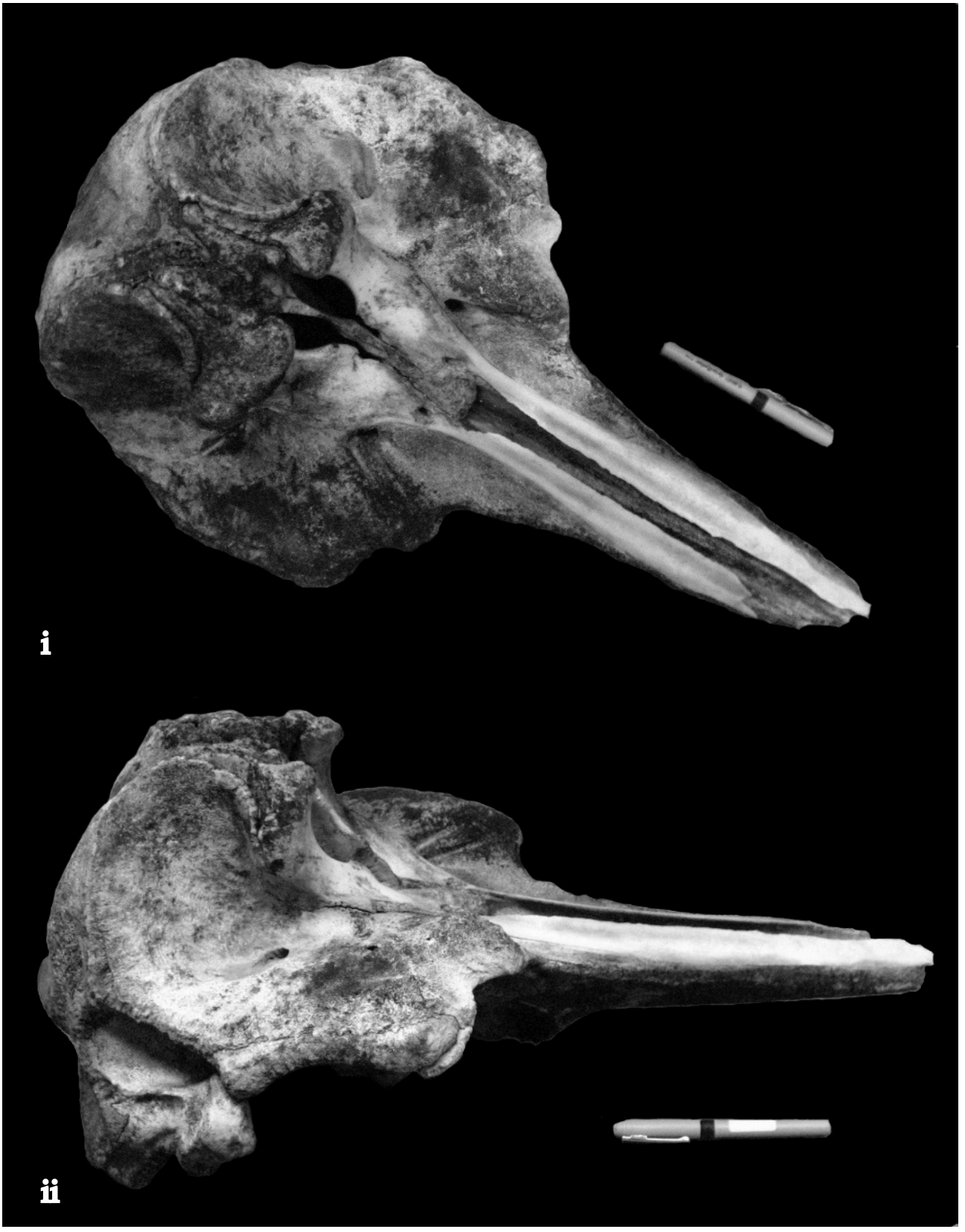
*Supplementary Figure 1*

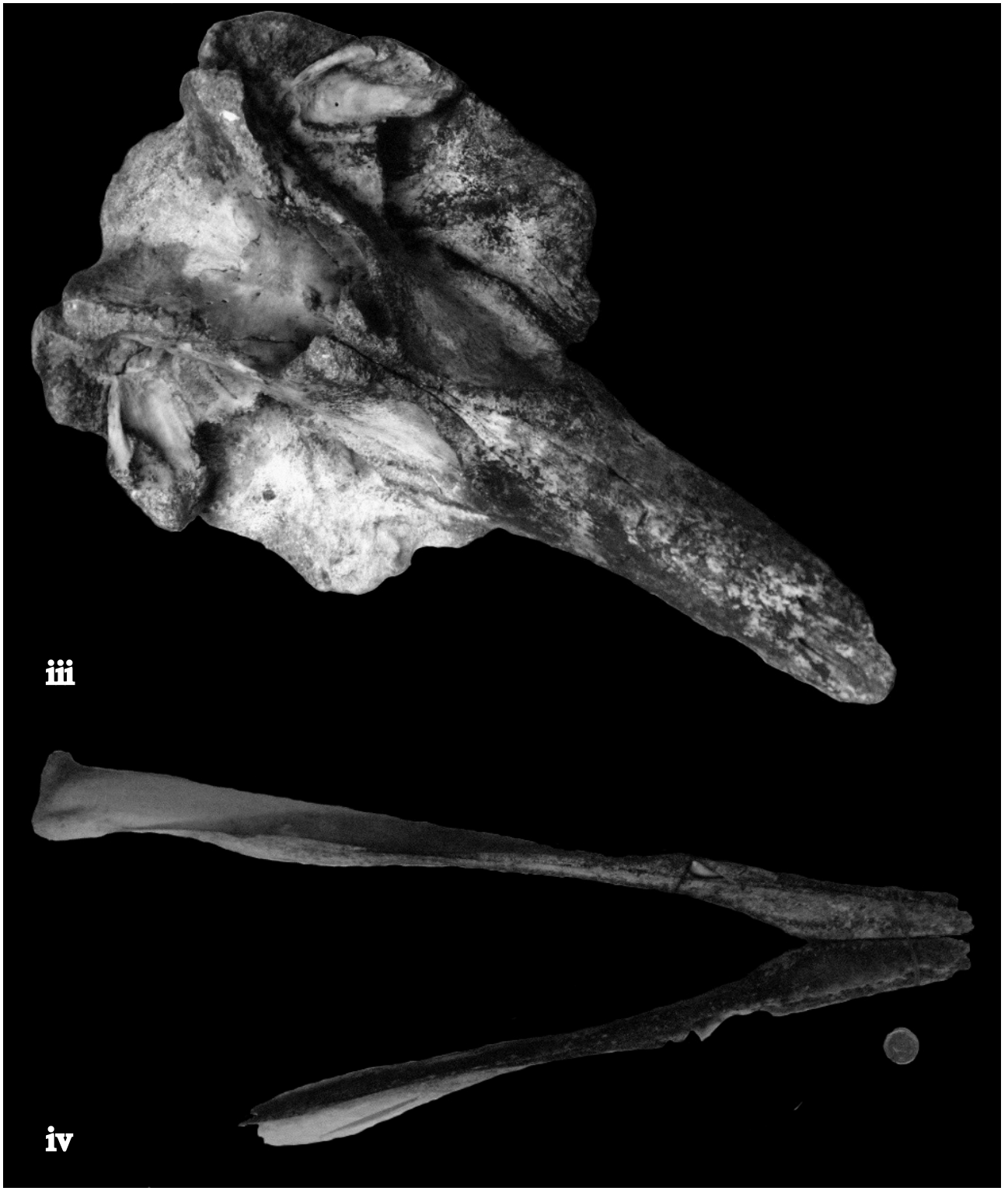


*Supplementary Figure 2*

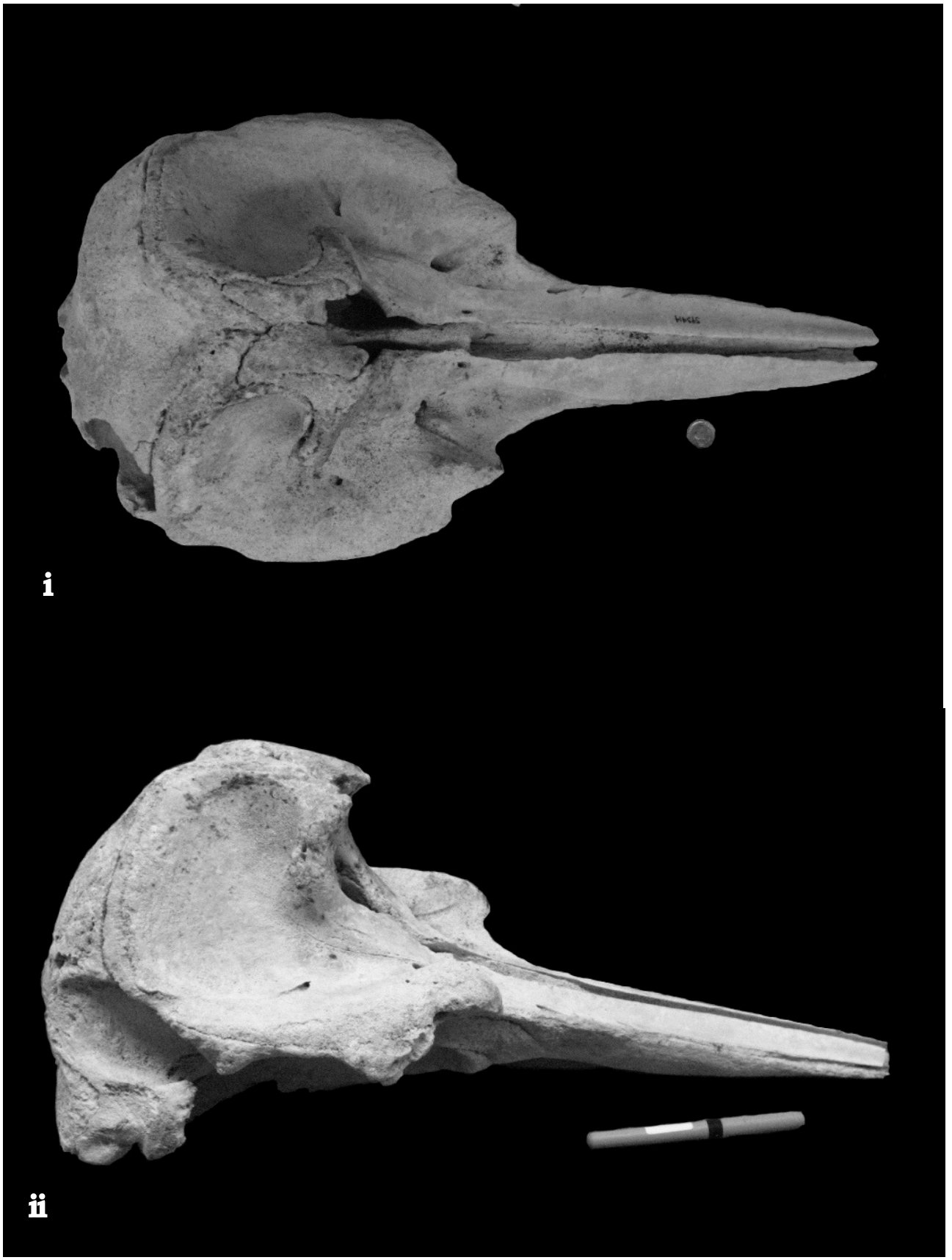


*Supplementary Figure 3A*




iii
iv

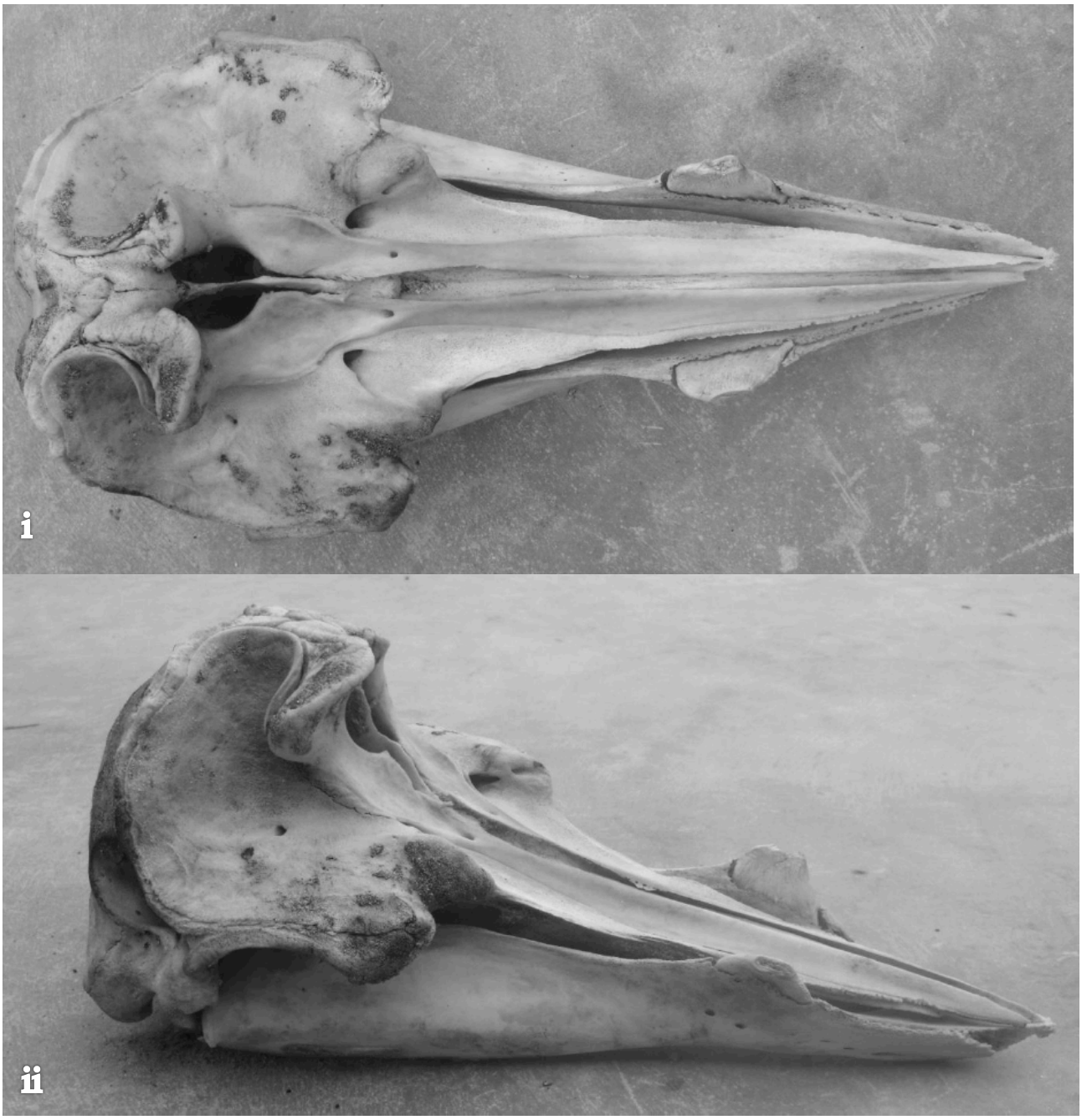
*Supplementary Figure 3B*



*Supplementary Figure 3C*

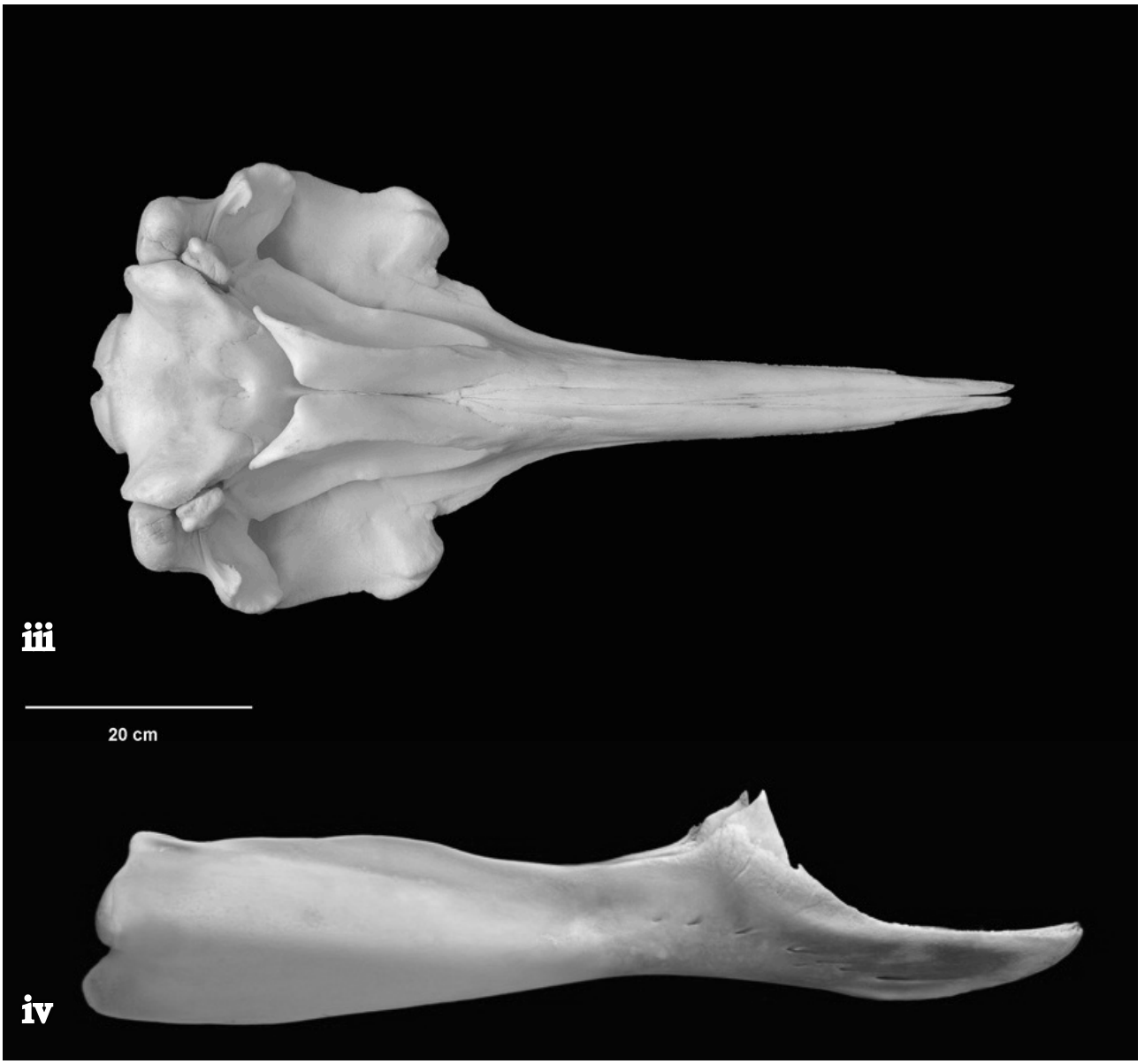

iii

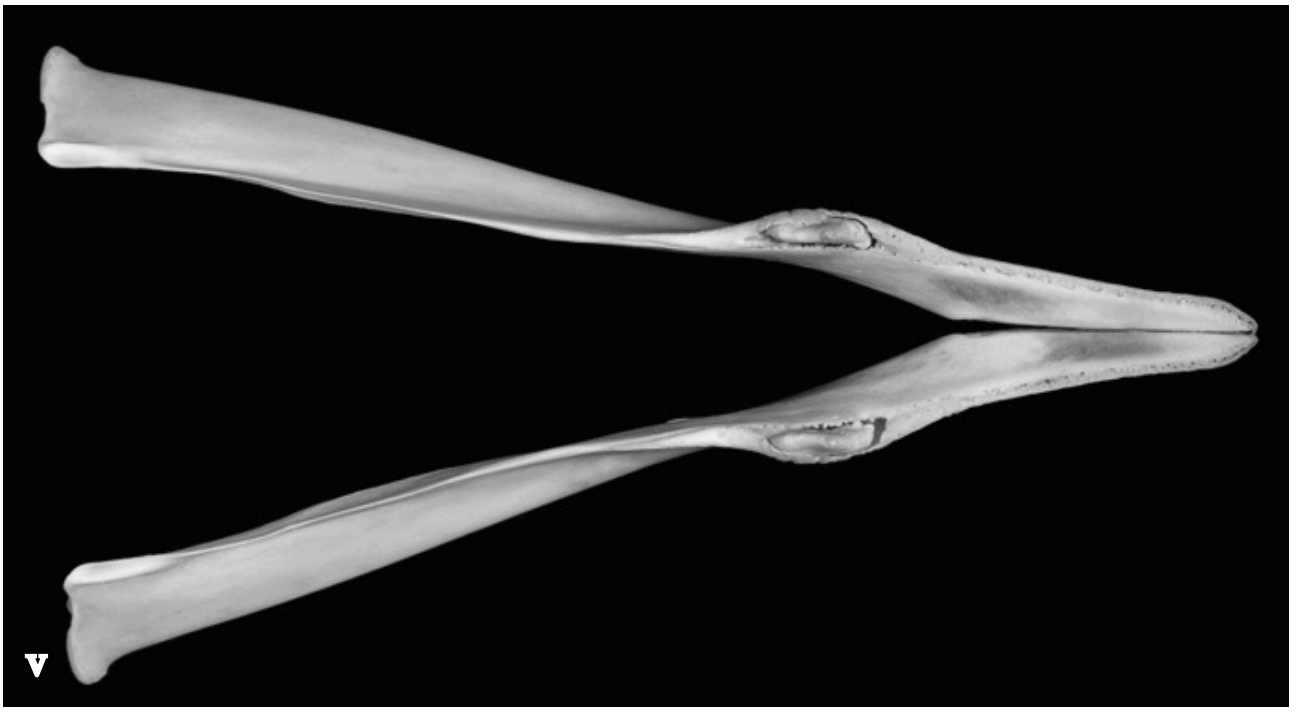
*Supplementary Figure 3D*



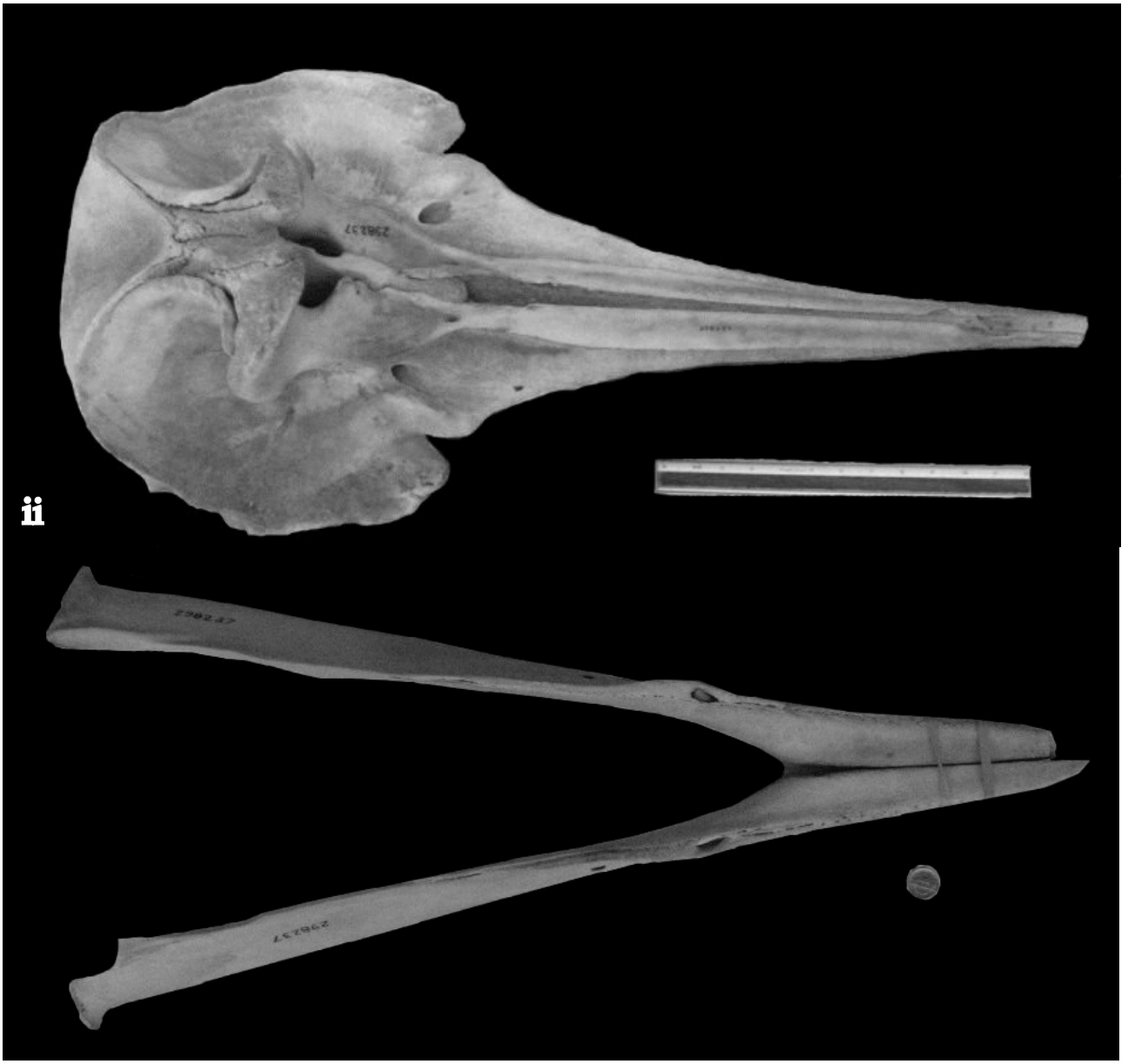

iii
iv

*Supplementary Figure 4A*


iii
20 cm
iv


V

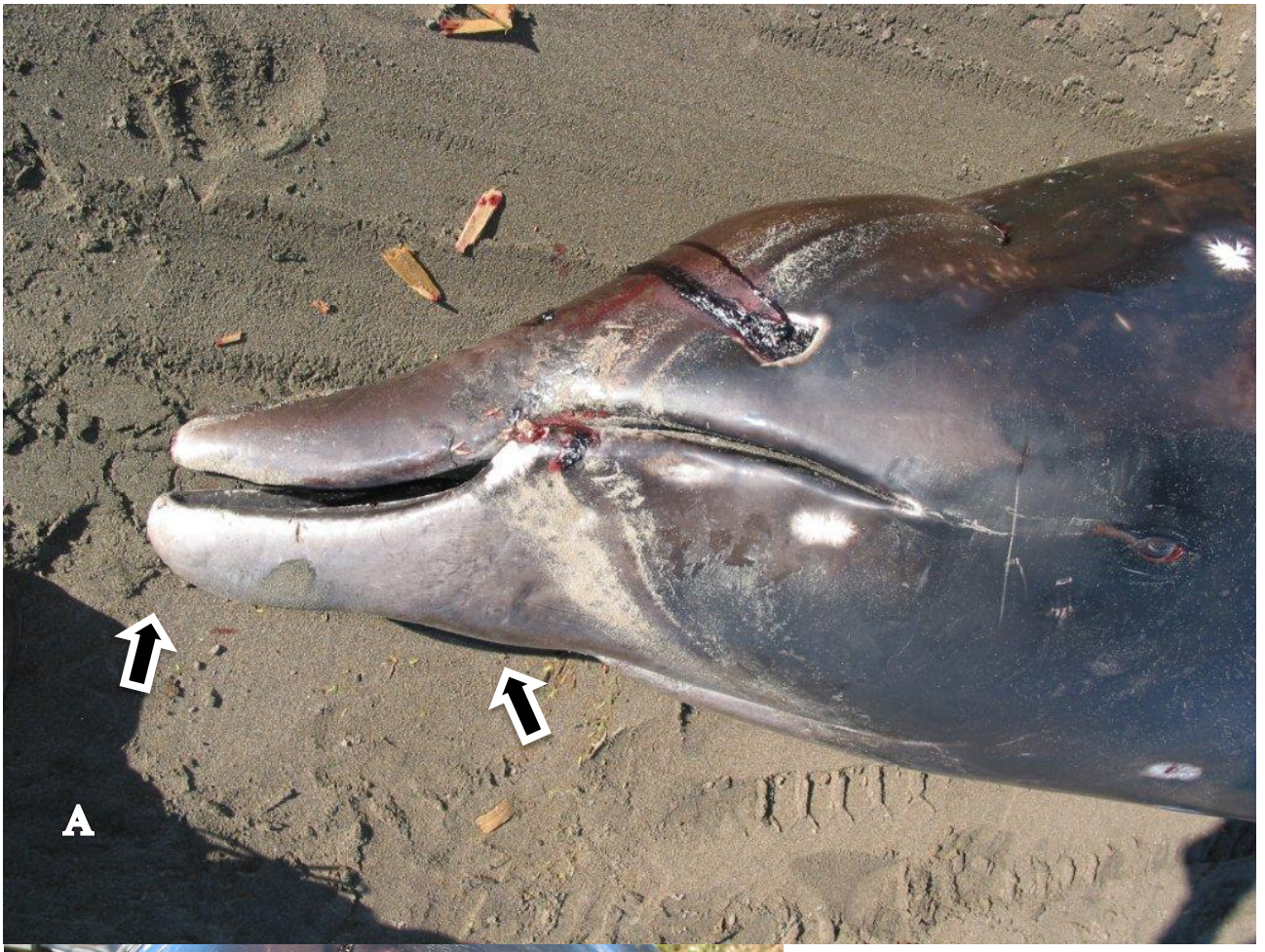
*Supplementary Figure 4B*

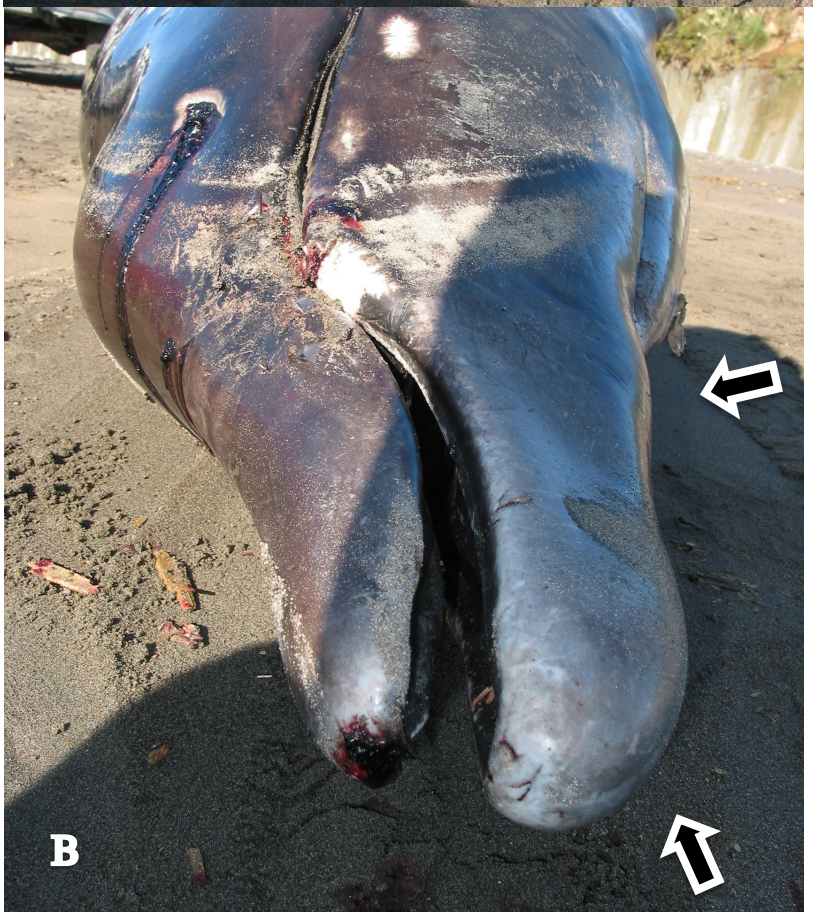


*Supplementary Figure 5*


C

*Supplementary Figure 6*



**Supplementary Tables**

*Supplementary Table 1*

MtDNA data. A) Control region (CR), variable sites over 658 bp.
B) Cytochrome *b* (CYB), variable sites over 384 bp. C) Cytochrome oxidase I (COXI), variable sites over 987 bp. Gray shading highlights the nucleotide substitutions differentiating *M. hotaula* from *M. ginkgodens.*

A) CR

| | 1 | 3 | 5 | 5 | 9 | 9 | 9 | 9 | 9 | 9 | 1 0 | 1 0 | 1 0 | 1 0 | 1 0 | 1 0 | 1 0 | 1 1 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Specimens | 5 | 9 | 6 | 9 | 0 | 2 | 4 | 7 | 8 | 9 | 0 | 1 | 4 | 5 | 6 | 7 | 9 | 0 |
| *M. hotaula* | | | | | | | | | | | | | | | | | | |
| 3WZS | G | G | A | A | C | T | C | G | C | A | T | A | T | A | C | C | A | C |
| UKIRI | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| USNM593418 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| USNM593414 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| USNM593426 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| MDV-X | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| *M. ginkgodens* | | | | | | | | | | | | | | | | | | |
| MginTSM8744 | . | A | G | . | A | C | T | A | T | G | C | G | C | G | T | _ | G | T |
| MginTW01 | . | A | G | . | A | C | T | . | T | G | C | G | C | G | T | _ | G | T |
| MginNZ03 | . | A | G | . | A | C | T | A | T | G | C | G | C | G | T | _ | G | T |
| MginNZ04 | . | A | G | . | A | C | T | A | T | G | C | G | C | G | T | _ | G | T |
| MginUSNM298237 | . | A | G | . | A | C | T | A | T | G | C | G | C | G | T | _ | G | T |
| MginMV29623 | A | A | G | C | A | C | T | A | T | G | C | G | C | G | T | _ | G | T |

(A) continued.

| Specimens | 1 1 1 1 1 2 2 2 3 3 3 4 4 4 5 5 5<br>1 1 3 4 4 4 5 7 1 5 6 0 0 3 0 2 9<br>1 3 3 7 8 9 1 9 8 2 3 5 7 9 0 2 9 |
|---|---|
| *M. hotaula* | |
| 3WZS | A G C T A T C A T A A C A T A C T |
| UKIRI | . . . . . C T G . G T G . C . . . |
| USNM593418 | . . . . . C T G . G T G . C . . . |
| USNM593414 | . . . . . C T G . G T G . C . . . |
| USNM593426 | . . . . . C T G . G T G . C . . . |
| MDV-X | . . . . . C T G . . T G . . G T . |
| *M. ginkgodens* | |
| MginTSM8744 | G . T C G . . G . G T G . ? ? ? ? |
| MginTW01 | G A . C G . . G . G T G . ? ? ? ? |
| MginNZ03 | G . T C G . . G C G T G . ? ? ? ? |
| MginNZ04 | G . T C G . . G C G T G . ? ? ? ? |
| MginUSNM298237 | G . T C G . . G C G T G . ? ? ? ? |
| MginMV29623 | G . T C G . . G C G T G G C . . C |

B) CYB

| Specimens | 306 | 309 | 345 | 366 | 381 | 390 | 412 | 435 | 436 | 438 | 465 | 490 | 491 | 498 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *M. hotaula* | | | | | | | | | | | | | | |
| UKIRI | C | A | T | T | T | T | A | T | C | C | C | C | T | C | T |
| USNM593418 | . | . | . | . | . | . | . | C | . | . | . | . | . | . | . |
| USNM593426 | ? | ? | ? | ? | ? | ? | . | C | . | . | . | . | . | . | . |
| MM-0001 | ? | ? | ? | ? | ? | ? | . | C | . | . | . | . | . | . | . |
| *M. ginkgodens* | | | | | | | | | | | | | | |
| MginTW01 | T | G | C | C | C | C | G | C | T | A | T | T | C | T | . |
| MginNZ03 | T | . | C | C | C | C | . | C | T | A | T | T | C | T | C |
| MginNZ04 | T | . | C | C | C | C | . | C | T | A | T | T | C | T | C |
| MginMV29623 | ? | ? | ? | ? | ? | ? | ? | C | T | A | T | T | C | T | C |

(B) continued.

| Specimens | 2 2 5 | 2 4 1 | 2 4 3 | 2 7 9 | 2 8 2 | 2 8 5 | 2 9 4 | 3 0 4 | 3 2 1 | 3 2 4 | 3 2 7 | 3 3 3 | 3 5 4 | 3 5 5 | 3 6 6 | 3 7 5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *M. hotaula* | | | | | | | | | | | | | | | | |
| UKIRI | T | A | C | T | T | T | C | T | C | T | C | T | C | C | G | C |
| USNM593418 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| USNM593426 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| MM-0001 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | ? |
| *M. ginkgodens* | | | | | | | | | | | | | | | | |
| MginTW01 | C | T | . | C | C | C | T | C | T | C | T | A | T | T | A | T |
| MginNZ03 | C | T | T | C | C | C | T | C | T | C | T | A | T | T | A | T |
| MginNZ04 | C | T | . | C | C | C | T | C | T | C | T | A | T | T | A | T |
| MginMV29623 | C | T | . | C | C | C | T | C | T | C | T | A | T | T | A | T |

C) COX I

| | | | | | | | | | | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 2 | 2 | 3 | 7 | 8 | 9 | 9 | 9 | 1 | 1 | 4 | 7 | 0 | 1 | 3 | 4 | 6 | 8 | 9 | 9 | 5 |
| Specimens | 5 | 0 | 7 | 3 | 1 | 6 | 5 | 6 | 8 | 4 | 9 | 3 | 3 | 6 | 9 | 6 | 2 | 9 | 5 | 3 | 9 | 9 |
| *M. hotaula* | | | | | | | | | | | | | | | | | | | | | | |
| UKIRI | C | T | T | C | C | C | A | T | A | C | T | A | C | C | T | T | A | C | C | C | T | A |
| *M. ginkgodens* | | | | | | | | | | | | | | | | | | | | | | |
| MginTW01 | T | . | C | T | . | . | G | C | G | T | C | . | T | A | C | . | C | T | T | . | A | G |
| MginNZ03 | T | C | C | T | . | T | G | C | G | T | C | G | T | A | C | C | . | T | T | T | A | . |
| MginNZ04 | T | C | C | T | T | T | G | C | G | T | C | . | T | A | C | C | . | T | T | T | A | . |

| | 3 | 3 | 3 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 6 | 6 | 6 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 6 | 7 | 8 | 0 | 1 | 1 | 2 | 2 | 2 | 5 | 7 | 8 | 9 | 1 | 2 | 2 | 3 | 4 | 9 | 0 | 2 | 2 |
| Specimens | 8 | 7 | 0 | 7 | 3 | 6 | 3 | 5 | 9 | 0 | 3 | 5 | 4 | 5 | 1 | 7 | 0 | 5 | 9 | 3 | 6 | 9 |
| *M. hotaula* | | | | | | | | | | | | | | | | | | | | | | |
| UKIRI | T | A | T | T | C | C | T | G | T | T | G | T | T | G | C | T | A | T | A | C | A | T |
| *M. ginkgodens* | | | | | | | | | | | | | | | | | | | | | | |
| MginTW01 | C | . | C | C | T | T | C | A | . | C | A | C | . | A | T | C | G | C | . | T | G | C |
| MginNZ03 | C | G | C | C | T | T | C | A | C | C | A | C | C | A | T | C | G | C | G | T | . | C |
| MginNZ04 | C | G | C | C | T | T | C | A | . | C | A | C | C | A | T | C | G | C | G | T | . | C |

(C) continued.

| Specimens | 66 | 68 | 71 | 71 | 73 | 75 | 77 | 78 | 80 | 81 | 82 | 82 | 83 | 86 | 87 | 87 | 88 | 91 | 92 | 95 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 29 | 93 | 71 | 12 | 98 | 65 | 12 | 30 | 25 | 77 | 03 | | | | | | | | | |
| *M. hotaula* | | | | | | | | | | | | | | | | | | | | |
| UKIRI | T | T | A | T | A | T | T | A | A | T | C | T | T | C | T | T | T | T | T | G |
| *M. ginkgodens* | | | | | | | | | | | | | | | | | | | | |
| MginTW01 | C | C | G | C | G | C | C | G | G | C | T | C | C | T | . | C | C | C | C | . |
| MginNZ03 | C | C | G | C | G | C | C | G | G | C | T | C | C | T | C | C | C | C | C | A |
| MginNZ04 | C | C | G | C | G | C | C | G | G | C | T | C | C | T | C | C | C | C | C | A |

*Supplementary Table 2.* Pairwise net divergence between species, Kimura 2-parameter distances, as percentages, below diagonal. SE, above diagonal. Values for sister-species pairs highlighted in gray with bold type. (A) Control region, CR (435 bp); (B) cytochrome *b*, CYB (384 bp); (C) cytochrome oxidase I, COXI (958 bp). See footnotes of *Supplementary Table 2A* for translation of species codes.

(A) CR

| | Mhot | Mgin | Mbi | Mbow | Mca | Mde | Meu | Mhe | Mlay | Mgr | Mmi | Mperu | Mpi | Mst | Mtr |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mhot | | 0.910 | 1.030 | 1.110 | 0.970 | 1.080 | 1.050 | 0.830 | 1.060 | 1.300 | 1.020 | 1.070 | 1.050 | 1.130 | 1.180 |
| Mgin | **3.57** | | 1.230 | 1.220 | 1.140 | 1.280 | 1.220 | 1.060 | 1.170 | 1.380 | 1.160 | 1.250 | 1.250 | 1.360 | 1.300 |
| Mbi | 4.40 | 5.89 | | 1.130 | 1.060 | 1.260 | 0.860 | 1.010 | 1.080 | 1.180 | 1.000 | 1.100 | 1.180 | 1.140 | 1.140 |
| Mbow | 5.48 | 6.18 | 5.75 | | 1.050 | 1.300 | 1.080 | 1.080 | 1.210 | 1.300 | 1.210 | 1.240 | 1.250 | 1.480 | 1.230 |
| Mca | 4.85 | 5.92 | 5.25 | **5.21** | | 1.200 | 1.050 | 1.010 | 1.070 | 1.210 | 1.120 | 0.990 | 1.120 | 1.200 | 1.020 |
| Mde | 4.87 | 6.26 | 6.08 | 6.67 | 5.92 | | 1.240 | 1.220 | 1.240 | 1.060 | 1.250 | 0.950 | 0.990 | 0.900 | 1.190 |
| Meu | 4.95 | 6.23 | 3.36 | 5.54 | 5.19 | 6.49 | | 0.870 | 1.080 | 1.090 | 0.870 | 1.100 | 1.200 | 1.170 | 1.190 |
| Mhe | 3.11 | 4.50 | 4.25 | 5.22 | 4.85 | 5.91 | 3.32 | | 0.950 | 1.120 | 0.960 | 1.110 | 1.170 | 1.060 | 1.130 |
| Mlay | 4.88 | 5.78 | 4.89 | 5.76 | 4.65 | 6.37 | 4.78 | 3.80 | | 1.240 | 1.170 | 1.210 | 1.240 | 1.180 | 1.170 |
| Mgr | 6.22 | 7.20 | 5.84 | 6.80 | 6.13 | 4.86 | 5.59 | 5.05 | 6.50 | | 1.160 | 0.950 | 1.070 | 1.170 | 1.400 |
| Mmi | 4.77 | 5.94 | 4.80 | 6.73 | 5.97 | 6.80 | **3.67** | 4.39 | 6.05 | 5.90 | | 1.160 | 1.120 | 1.020 | 1.180 |
| Mperu | 4.90 | 6.25 | 5.31 | 6.43 | 4.33 | 3.75 | 5.59 | 5.43 | 6.21 | 3.92 | 6.15 | | 0.830 | 0.960 | 1.210 |
| Mpi | 4.66 | 5.95 | 5.53 | 6.68 | 5.08 | 4.32 | 6.00 | 5.60 | 6.37 | 4.53 | 5.99 | **3.20** | | 1.020 | 1.180 |
| Mst | 5.22 | 6.90 | 5.01 | 8.33 | 5.70 | **3.58** | 5.57 | 4.29 | 5.39 | 5.62 | 5.02 | 4.32 | 4.75 | | 1.120 |
| Mtr | 5.72 | 6.67 | 5.50 | 6.41 | 4.47 | 6.17 | 5.90 | 5.08 | 5.67 | 8.06 | 6.27 | 6.39 | 6.10 | 5.48 | |

Mhot, *M. hotaula*; Mgin, *M. ginkgodens*; Mbi, *M. bidens*; Mbow, *M. bowdoini*; Mca, *M. carlhubbsi*; Mde, *M. densirostris*; Meu, *M. europeaus*; Mhe, *M. hectori*; Mlay, *M. layardii*; Mgr, *M. grayi*; Mmi, *M. mirus*; Mperu, *M. peruvianus*; Mpi, *M. perrini*; Mst, *M. stejnegeri*; Mtr, *M. traversii*.

(B) CYB

| | Mhot | Mgin | Mbi | Mbow | Mca | Mde | Meu | Mgr | Mhe | Mlay | Mmi | Mpi | Mperu | Mst | Mtr |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mhot | | 1.790 | 2.290 | 2.300 | 2.010 | 2.110 | 1.900 | 2.060 | 2.300 | 1.890 | 1.330 | 2.310 | 2.330 | 2.220 | 2.310 |
| Mgin | **8.24** | | 2.420 | 2.450 | 2.370 | 2.180 | 2.040 | 2.120 | 2.060 | 2.220 | 1.940 | 2.040 | 2.120 | 2.020 | 2.590 |
| Mbi | 12.19 | 13.86 | | 2.250 | 2.160 | 2.050 | 2.480 | 2.400 | 2.110 | 2.380 | 1.650 | 2.610 | 2.210 | 1.940 | 2.040 |
| Mbow | 11.95 | 13.43 | 12.13 | | 2.240 | 2.110 | 2.150 | 2.370 | 2.550 | 2.240 | 1.830 | 2.610 | 2.580 | 2.350 | 2.480 |
| Mca | 10.66 | 13.99 | 11.22 | **11.77** | | 1.970 | 2.120 | 2.040 | 2.130 | 1.860 | 1.990 | 2.290 | 2.190 | 2.070 | 2.020 |
| Mde | 11.58 | 12.38 | 9.94 | 11.32 | 10.58 | | 2.130 | 1.680 | 2.040 | 2.030 | 1.740 | 2.080 | 2.050 | 2.010 | 2.400 |
| Meu | 8.87 | 9.97 | 13.92 | 10.30 | 11.49 | 11.58 | | 2.000 | 2.350 | 2.080 | 1.850 | 2.180 | 1.980 | 2.210 | 2.270 |
| Mgr | 10.59 | 11.72 | 12.37 | 12.48 | 10.48 | 7.39 | 10.15 | | 2.150 | 2.130 | 1.830 | 1.950 | 1.810 | 1.920 | 2.430 |
| Mhe | 12.38 | 10.40 | 10.67 | 13.78 | 11.11 | 10.01 | 12.35 | 11.14 | | 2.000 | 2.000 | 2.280 | 2.520 | 2.270 | 2.130 |
| Mlay | 9.49 | 12.47 | 12.62 | 11.85 | 8.82 | 10.41 | 10.72 | 10.78 | 9.87 | | 1.420 | 2.280 | 2.220 | 2.470 | 2.280 |
| Mmi | 5.54 | 9.76 | 7.21 | 8.65 | 10.32 | 8.31 | **8.68** | 9.04 | 9.83 | 5.53 | | 2.120 | 2.000 | 1.920 | 1.990 |
| Mpi | 13.75 | 10.88 | 15.54 | 16.13 | 13.55 | 11.15 | 12.58 | 10.17 | 12.47 | 13.28 | 11.61 | | 2.240 | 2.080 | 2.720 |
| Mperu | 13.53 | 11.40 | 11.69 | 14.58 | 11.91 | 10.43 | 10.04 | 8.87 | 14.13 | 11.40 | 10.01 | **11.68** | | 2.100 | 2.630 |
| Mst | 12.80 | 10.49 | 9.76 | 12.77 | 10.76 | **10.39** | 11.24 | 9.60 | 12.45 | 13.64 | 10.10 | 11.18 | 10.89 | | 2.130 |
| Mtr | 12.77 | 13.85 | 10.95 | 13.25 | 10.32 | 13.65 | 12.43 | 13.42 | 11.03 | 12.52 | 10.26 | 16.60 | 15.71 | 11.41 | |

(C) COXI

| | Mhot | Mgin | Meu | Mmi | Mde |
|---|---|---|---|---|---|
| Mhot | | 0.760 | 1.050 | 0.830 | 1.040 |
| Mgin | **5.52** | | 0.930 | 0.950 | 1.040 |
| Meu | 9.36 | 8.03 | | 0.900 | 1.100 |
| Mmi | 7.59 | 9.06 | **7.76** | | 0.920 |
| Mde | 9.58 | 9.34 | 9.95 | 8.45 | |

*Supplementary Table 3*. Y-chromosome intron, DBY7. Variable sites over 241 bp. The diagnostic nucleotide substitutions that distinguish *M. ginkgodens* and *M. hotaula* from each another, and from the other *Mesoplodon* species sampled are highlighted in gray.

| | | | | | | | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | 5 | 6 | 6 | 6 | 8 | 0 | 2 | 2 | 2 | 4 | 5 | 8 | 9 | 9 | 0 |
| Species[a, b] | 4 | 3 | 0 | 2 | 7 | 1 | 5 | 3 | 4 | 5 | 6 | 6 | 9 | 1 | 2 | 3 |
| *M. bowdoini* | G | C | G | T | T | A | C | _ | _ | _ | C | G | C | C | C | T |
| *M. carlhubbsi* | . | . | . | . | . | . | . | _ | _ | _ | . | . | . | . | . | . |
| *M. layardii* | . | . | . | . | . | . | A | _ | _ | _ | . | . | . | . | . | . |
| *M. stejnegeri* | . | . | . | C | . | . | . | T | T | T | . | . | . | . | . | . |
| *M. grayi* | . | . | . | . | . | . | T | _ | _ | _ | . | . | . | . | . | . |
| *M. perrini* | . | . | . | . | . | . | . | _ | _ | _ | T | . | . | . | . | . |
| *M. hectori* | . | . | T | . | . | . | . | _ | _ | _ | . | . | . | . | . | . |
| *M. europaeus* | . | T | . | . | . | G | . | _ | _ | _ | . | . | . | . | . | . |
| *M. mirus* | . | . | . | . | . | . | . | _ | _ | _ | . | . | . | . | . | . |
| *M. hotaula* | . | . | . | . | C | . | . | _ | _ | _ | . | C | . | . | . | C |
| *M. ginkgodens* | . | . | . | . | C | . | . | _ | _ | _ | . | . | . | T | . | C |
| *M. bidens* | A | . | . | . | . | . | . | _ | _ | _ | . | . | . | . | . | . |

[a]Missing *M. densirostris, M. peruvianus,* and *M. traversii*.

[b]GenBank Accession No's: KF027328-KF027337.

*Supplementary Table 4A.* Cranial measurements for *Mesoplodon hotaula.* Measurements (in mm) are taken on the right hand side (R) where possible, following Moore (1963). Where two measurements are given, R, then L. E, estimated length. See *Supplementary Table 3B* for definitions of measurements.

| | 3WZS | USNM593418 | USNM593414 | USNM593426 | MDV-X |
|---|---|---|---|---|---|
| | Deraniyagala (1963a), Anderson, this paper | Mead, this paper | Mead, this paper | Mead, this paper | Anderson, this paper |
| Measurement number | adult F | adult F | subadult F? | adult M | adult M |
| 1 | 735 | 680E | 615E | 735 | 721 |
| 2 | | 604E | 528E | 678, 677 | |
| 3 | | 463E | 406E | 521 | |
| 4 | | 450E | 404E | 553, 550 | |
| 5 | | 516E | 462E | 606 | |
| 6 | | 620E, 612E | 581E, _ | 695, 693 | |
| 7 | | 339E | 323E | 403 | |
| 8 | | 323 | 282 | 348 | |
| 9 | | 346 | 284 | 372 | |

| | | | | |
|---|---|---|---|---|
| 10 | 355 | 361E | 284 | 360 |
| 11 | | 233 | _ | 237 |
| 12 | | 288 | 262 | _ |
| 13 | | 114 | 91 | 111 |
| 14 | | 44 | 31 | 41, 47 |
| 15 | | 71, 73 | 56, 56 | 70, 74 |
| 16 | | 42 | 36 | 42 |
| 17 | | 65 | 69 | 89 |
| 18 | | 42 | 45 | 48 |
| 19 | | 3 | -10 | 0 |
| 20 | | 47 | 53 | 58 |
| 21 | | 28 | _ | 21 |
| 22 | | 151 | 137 | 161 |
| 23 | | 47 | 58 | 57 |
| 24 | | 116 | 104 | 112 |
| 25 | | 115 | 104 | 117 |
| 26 | | 126E | 13 | 158 |

| | | | | |
|---|---|---|---|---|
| 27 | | 211 | 197 | 288 |
| 28 | | 89 | _ | 94 |
| 29 | | 34 | 33 | 38 |
| 30 | | 75 | 62 | 71 |
| 31 | | _ | _ | 8 |
| 32 | 60.2 | 67 | 63 | 61 |
| 33 | | 58 | 39 | 45 |
| 34 | | 183 | 194 | 20 |
| 35 | 280 | 284 | 249 | 279 |
| 36 | | 102, 100 | 104E | 103, 97 |
| 37 | | 53, 54 | _ | 55, 67 |
| 38 | | 99, 94 | 82 | 87, 100 |
| 39 | | 372E | 352E | 364 |
| 40 | | 327E, 336E | 290, 290 | 411, 410 |
| 41 | | 298E | 275 | 364 |
| 42 | | 459E | _ | 573 |
| 43 | | _ | _ | 607 |

| | | | |
|---|---|---|---|
| 44 | _ | _ | 117 |
| 45 | 107 | _ | 91 |
| 46 | 79 | 36 | 0 |
| Notes | skull extremely damaged, burned | incomplete skull | excellent skull, pathology on right lateral exo occipital (abcess) |

*Supplementary Table 4B*. Definitions of cranial measurements. Numbers in parentheses refer to Moore (1963).

1 – condylobasal length (1)

2 – tip rostrum to posterior extension maxillary plate (7)

3 – tip rostrum to anterior margin superior nares (8)

4 – tip rostrum to anterior point maxillary crest (9)

5 – tip rostrum to posterior extension premaxilla on lateral tip of right premaxillary crest (11)

6 – tip rostrum to posterior extension temporal fossa (10)

7 – tip rostrum to apices of antorbital notches (2)

8 – breadth skull across orbital centres (19)

9 – breadth skull across postorbital process frontals (17)

10 – breadth skull across zygomatic processes squamosals (18)

11 – least breadth skull across posterior margins temporal fossae (20)

12 – greatest breadth skull across ex-occipitals (25)

13 – greatest span occipital condyles (21)

14 – greatest width of an occipital condyle (22)

15 – greatest length of an occipital condyle (23)

16 – greatest breadth foramen magnum (24)

17 – greatest length of right nasal on vertex (15)

18 – length nasal suture (16)

19 – extension right premaxilla posterior to right nasal on vertex (28)

20 – greatest breadth nasals on vertex (26)

21 – least distance between anterior prominences of the synvertex (27)

22 – greatest span premaxillary crests (29)

23 – greatest transverse width of superior nares (37)

24 – least width premaxillae where narrow opposite superior nares (30)

25 – greatest width premaxillae anterior to position of previous (31)

26 – width rostrum in apices of antorbital notches (33)

27 – width rostrum in apices of prominential notches (34)

28 – least distance be- tween main maxillary foramina (41)

29 – least distance between premaxillary foramina (42)

30 – distance posterior margin of left maxillary foramina to anterior margin maxillary prominence (43)

31 – width rostrum at mid-length of rostrum (35)

32 – width premaxillae at mid-length of rostrum (32)

33 – depth rostrum at mid-length rostrum (36)

34 – height of skull (39)

35 – external cranial height

36 – greatest length of temporal fossa (13)

37 – width of temporal fossa (40)

38 – length of orbit taken from mid-point of frontals (14)

39 – tip rostrum to posterior extension of maxilla between pterygoids (6)

40 – tip rostrum to anterior extension of pterygoid sinus (12)

41 – tip rostrum to most anterior extension of pterygoids (5)

42 – tip rostrum to posterior margin of pterygoid mid-line (3)

43 – tip rostrum to posterior extension of wing of pterygoid (4)

44 – length of vomer visible at surface of palate (44)

45 – width between pterygoid notches (38)

46 – amount added to rostrum because of breakage (45)

*Supplementary Table 5A.* Mandibular measurements for *Mesoplodon hotaula*. Measurements (in mm) are taken on the right hand side (R) where possible, following Moore (1963). Where two measurements are given, R, then L. E, estimated length. See *Supplementary Table 4B* for definitions of measurements.

| Measurement number | 3WZS<br>Deraniyagal (1963a)<br>adult F | USNM593418<br>Mead, this paper<br>adult F | USNM593426<br>Mead, this paper<br>adult M | MDV-X<br>Anderson, this paper<br>adult M |
|---|---|---|---|---|
| 47 | 631 | 618E | 620, 610 | 610, 615 |
| 48 | | 491 | 493, 496 | |
| 49 | | 427 | 383, 377 | |
| 50 | | 137 | 139, 139 | |
| 51 | | 108 | 110, 114 | |
| 52 | | 49 | 88, 98 | |
| 53 | | 50, 49 | _, 74 | |
| 54 | | 23, _ | 89, 93 | |
| 55 | | 10, _ | _, 18 | |

| | | | | |
|---|---|---|---|---|
| 56 | | 168, _ | 168, 167 | |
| 57 | 45 | 54 L | _ | 88, 90 |
| 58 | 54 | 60 L | _ | 57, 60 |
| 59 | | 17 L | _ | |
| 60 | | 20 L | _ | |
| 61 | | 143 g L | _ | |
| 62 | | 170 | 200 | |
| Notes | | mandible broken, left missing 170 mm | 200 mm added to mandibles due to breakage | |

*Supplementary Table 5B*. Definitions of mandibular measurements. Numbers in parentheses refer to Moore (1963).

47 –mandibular length (1)

48 – length from posterior extension of symphysis to condyles (6)

49 – length posterior margin of alveolus to condyles (7)

50 – greatest length of symphysis (2)

51 – greatest height of mandible at coronoid processes (3)

52 – outside height of mandible at midlength of alveolus (4)

53 – inside height of mandible at midlength of alveolus (5)

54 – length of alveolus (8)

55 – width of alveolus (9)

56 – tip of mandible to alveolus (10)

57 – greatest tooth length (11)

58 – greatest tooth width (12)

59 – greatest tooth breadth (13)

60 – height of crown of tooth

61 – tooth weight

62 – amount added to mandibles due to breakage

Mitochondrial DNA

http://informahealthcare.com/mdn
ISSN: 1940-1736 (print), 1940-1744 (electronic)

Mitochondrial DNA, Early Online: 1–2
 DOI: 10.3109/19401736.2013.878908

informa
healthcare

MITOGENOME ANNOUNCEMENT

# High coverage of the complete mitochondrial genome of the rare Gray's beaked whale (*Mesoplodon grayi*) using Illumina next generation sequencing

Kirsten F. Thompson[1,2], Selina Patel[1,2], Liam Williams[2], Peter Tsai[2], Rochelle Constantine[2], C. Scott Baker[2,3], and Craig D. Millar[1,2]

[1]*The Allan Wilson Centre, School of Biological Sciences, University of Auckland, Auckland, New Zealand,* [2]*School of Biological Sciences, University of Auckland, Auckland, New Zealand, and* [3]*Department of Fisheries and Wildlife and Marine Mammal Institute, Oregon State University, Newport, OR, USA*



**Abstract**

Using an Illumina platform, we shot-gun sequenced the complete mitochondrial genome of Gray's beaked whale (*Mesoplodon grayi*) to an average coverage of 152X. We performed a *de novo* assembly using SOAPdenovo2 and determined the total mitogenome length to be 16,347 bp. The nucleotide composition was asymmetric (33.3% A, 24.6% C, 12.6% G, 29.5% T) with an overall GC content of 37.2%. The gene organization was similar to that of other cetaceans with 13 protein-coding genes, 2 rRNAs (12S and 16S), 22 predicted tRNAs and 1 control region or D-loop. We found no evidence of heteroplasmy or nuclear copies of mitochondrial DNA in this individual. Beaked whales within the genus *Mesoplodon* are rarely seen at sea and their basic biology is poorly understood. These data will contribute to resolving the phylogeography and population ecology of this speciose group.

**Keywords**

Annotation, assembly, cetacea, mitochondrial genome

**History**

Received 10 December 2013
Accepted 18 December 2013
Published online 22 January 2014

Within the cetaceans there are approximately 67 species of toothed whales (odontocetes) whose systematic relationships remain unclear (Perrin et al., 2013). One of the largest families of odontocetes, the Ziphiidae or beaked whales, contain 22 comparatively unknown oceanic species. Genetic analyses suggest that there may be a number of other cryptic species that exist within this group (Baker et al., 2013; Thompson et al., 2012).

We sequenced a sample from a Gray's beaked whale (*Mesoplodon grayi*) found in New Zealand in 2012 (code: Mgr159). Though this species is rarely seen alive, it is one of the more common beach cast whales in New Zealand (Thompson et al., 2013). Genomic DNA was extracted according to Thompson et al. (2013) and fragmented to ~400 bp using the Covaris® M220 (Woburn, MA). A library was constructed using the Illumina TruSeq DNA PCR-Free Sample Preparation Kit and sequenced using the MiSeq Reagent Kit 500 cycle PE. A single run on the Illumina MiSeq® (San Diego, CA) platform generated 8.6 Gb of sequence. Raw reads were trimmed using SolexaQA removing regions with a Phred score of <20 (Cox et al., 2010). High-quality reads were assembled using SOAPdenovo2 into 514,905 contigs (Luo et al., 2012). The largest of these contigs was identified as mitochondrial by a BLAST search. A total of 12,055 individual mitochondrial reads gave an average coverage of 152X. Annotation was performed in GENEIOUS v.6.0.5 created by Biomatters available from http://www.geneious.com/ by comparing with *Mesoplodon densirostris* (GenBank: NC021974). Transfer RNA genes were predicted using both ARWEN (Laslett & Canbäck, 2008) and tRNAscan-SE v1.21 (Lowe & Eddy, 1997).

We estimated the mitochondrial genome to be 16,347 bp in length with a nucleotide composition (33.3% A, 24.6% C, 12.6% G, 29.5% T) and an overall GC content of the H-strand of 37.2%. The gene organization was as follows: 13 protein-coding genes; 2 rRNAs (12S and 16S); 22 tRNAs and 1 D-loop (Table 1). The most common start codon was ATG followed by ATA and a single instance of GTG. Seven genes were terminated with a complete TAA stop codon, and a further three were incomplete. All genes were encoded on the heavy strand with the exception of ND6. There were eight instances of intergenic nucleotides ranging from 1 to 7 bp. Eleven out of the 13 genes overlapped between 1 and 40 bp with the largest overlap found between ATP8 and ATP6 genes. The ND5 and ND6 genes overlapped by 17 bp and were encoded by opposing strands. The 22 tRNA genes were between 66 and 74 bp in length and the origin of replication was 35 bp located on the light strand between $tRNA^{Asn}$ and $tRNA^{Cys}$. The two ribosomal RNAs were 971 bp in length and 1583 bp, respectively. The D-loop was 880 bp long and was located between $tRNA^{Pro}$ and $tRNA^{Phe}$.

Heteroplasmy is common in cetaceans and can have a tendency to increase genetic diversity (Vollmer et al., 2011). Excluding positions with a single sequence difference, the remaining had 2–3 discordant bases. With an average of 152X coverage, this suggests no heteroplasmy or nuclear copies of mitochondrial DNA were present in our data.

Given the rarity of Gray's beaked whales, the complete mitogenome sequence of this species will be valuable for larger

Correspondence: Kirsten F. Thompson, The Allan Wilson Centre, School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland Mail Centre, Auckland 1142, New Zealand. Tel: +64 (0)9 373 7599. E-mail: ktho076@aucklanduni.ac.nz

Table 1. Characterization of the mitochondrial genome of Gray's beaked whale (*M. grayi*).

| | | Position | | | | Codons | | |
|---|---|---|---|---|---|---|---|---|
| Gene | Length (bp) | Start | Stop | Intergenic nucleotides[a] | Anti-codon | Start | Stop[b] | Strand[c] |
| $tRNA^{Phe}$ | 73 | 1 | 73 | 0 | GAA | | | H |
| 12SrRNA | 971 | 74 | 1043 | 0 | | | | H |
| $tRNA^{Val}$ | 67 | 1045 | 1111 | +1 | TAC | | | H |
| 16SrRNA | 1583 | 1112 | 2693 | 0 | | | | H |
| $tRNA^{Leu}$ (UUR) | 75 | 2694 | 2768 | 0 | TAA | | | H |
| ND1 | 957 | 2771 | 3727 | +2 | | ATG | TAG | H |
| $tRNA^{Ile}$ | 69 | 3732 | 3800 | +4 | GAT | | | H |
| $tRNA^{Gln}$ | 73 | 3870 | 3798 | −3 | TTG | | | L |
| $tRNA^{Met}$ | 71 | 3871 | 3941 | 0 | CAT | | | H |
| ND2 | 1042 | 3941 | 4982 | −1 | | ATA | T(aa) | H |
| $tRNA^{Trp}$ | 70 | 4983 | 5052 | 0 | TCA | | | H |
| $tRNA^{Ala}$ | 73 | 5124 | 5054 | +1 | TGC | | | L |
| $tRNA^{Asn}$ | 74 | 5198 | 5125 | 0 | GTT | | | L |
| $O_L$ | 35 | 5233 | 5199 | 0 | | | | L |
| $tRNA^{Cys}$ | 68 | 5298 | 5231 | −3 | GCA | | | L |
| $tRNA^{Tyr}$ | 66 | 5364 | 5299 | 0 | GTA | | | L |
| COI | 1551 | 5365 | 6915 | 0 | | ATG | AGA | H |
| $tRNA^{Ser}$ | 69 | 6979 | 6911 | −5 | TGA | | | L |
| $tRNA^{Asp}$ | 68 | 6987 | 7054 | +7 | GTC | | | H |
| COII | 684 | 7055 | 7738 | 0 | | ATG | TAA | H |
| $tRNA^{Lys}$ | 67 | 7742 | 7808 | +3 | TTT | | | H |
| ATP8 | 198 | 7810 | 8007 | +1 | | ATG | TAA | H |
| ATP6 | 681 | 7968 | 8648 | −40 | | ATG | TAA | H |
| COIII | 785 | 8648 | 9432 | −1 | | ATG | TA(a) | H |
| $tRNA^{Gly}$ | 69 | 9433 | 9501 | 0 | TCC | | | H |
| ND3 | 347 | 9502 | 9848 | 0 | | ATA | TA(a) | H |
| $tRNA^{Arg}$ | 70 | 9849 | 9918 | 0 | TCG | | | H |
| ND4L | 297 | 9919 | 10,215 | 0 | | GTG | TAA | H |
| ND4 | 1378 | 10,209 | 11,586 | −7 | | ATG | T(aa) | H |
| $tRNA^{His}$ | 69 | 11,587 | 11,655 | 0 | GTG | | | H |
| $tRNA^{Ser}$ (AGY) | 60 | 11,656 | 11,715 | 0 | GCT | | | H |
| $tRNA^{Leu}$ (CUN) | 70 | 11,717 | 11,786 | +1 | TAG | | | H |
| ND5 | 1821 | 11,787 | 13,607 | 0 | | ATA | TAA | H |
| ND6 | 528 | 14,118 | 13,591 | −17 | | ATG | TAA | L |
| $tRNA^{Glu}$ | 70 | 14,187 | 14,118 | −1 | TTC | | | L |
| Cytb | 1140 | 14,191 | 15,330 | +3 | | ATG | AGA | H |
| $tRNA^{Thr}$ | 74 | 15,330 | 15,403 | −1 | TGT | | | H |
| $tRNA^{Pro}$ | 69 | 15,469 | 15,401 | −3 | TGG | | | L |
| D-loop | 880 | 15,468 | 16,347 | −2 | | | | – |

[a]Numbers correspond to the nucleotides separating different genes. Negative numbers indicate overlapping nucleotides between adjacent genes.
[b]The TAA stop codon is presumed to be completed through polyadenylation of the RNA message after cleavage (Nardi et al., 2001).
[c]H and L denote heavy and light strands, respectively.

phylogeographical studies. Annotated consensus sequence is available on GenBank KF981442 and all reads are available in the Sequence Read Archive SRX420724.

## Acknowledgements

The authors would like to thank the following: Lester Bridson of the New Zealand Department of Conservation; New Zealand Genomics Ltd; the Centre for Genomics, Proteomics and Metabolomics; the New Zealand Cetacean Tissue Archive at the University of Auckland.

## Declaration of interest

This study was supported by funding from University of Auckland. The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the article.

## References

Baker CS, Hutt A, Thompson K, Dalebout ML, Robins J, Brownell RL, Stone GS. (2013). Species identity and human consumption of beaked whales in the Gilbert Islands, Republic of Kiribati. Anim Con 16:641–7.

Cox MP, Peterson DA, Biggs PJ. (2010). SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. BMC Bioinformatics 11:485.

Laslett D, Canbäck B. (2008). ARWEN: A program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. Bioinformatics 24: 172–5.

Luo R, Liu B, Xie Y, Li Z, Huang W, Yuang J, He G, et al. (2012). SOAPdenovo2: An empirically improved memory-efficient short-read *de novo* assembler. GigaScience 1:18.

Lowe TM, Eddy SR. (1997). tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. Nucl Acids Res 25:955–64.

Nardi F, Carapelli A, Fanciulli PP, Dallai R, Frati F. (2001). The complete mitochondrial DNA sequence of the basal hexapod *Tetrodontophora bielanensis*: Evidence for heteroplasmy and tRNA translocations. Mol Biol Evol 18:1293–304.

Perrin WF, Rosel PE, Cipriano F. (2013). How to contend with paraphyly in the taxonomy of the delphinine cetaceans? Mar Mamm Sci 29: 567–88.

Thompson K, Baker CS, van Helden A, Patel S, Millar C, Constantine R. (2012). The world's rarest whale. Curr Biol 22:R905–6.

Thompson KF, Millar CD, Baker CS, Dalebout ML, Steel D, van Helden AL, Constantine R. (2013). A novel conservation approach provides insights into the management of rare cetaceans. Biol Cons 157:331–40.

Vollmer NL, Viricel A, Wilcox L, Moore MK, Rosel PE. (2011). The occurrence of mtDNA heteroplasmy in multiple cetacean species. Curr Gen 57:115–31.

Conservation Genet Resour
DOI 10.1007/s12686-014-0173-0

MICROSATELLITE LETTERS

# Mining microsatellites for Gray's beaked whale from second-generation sequencing data

**Selina Patel · Kirsten Thompson · Liam Williams · Peter Tsai · Rochelle Constantine · Craig Millar**

Received: 11 February 2014 / Accepted: 24 March 2014


**Abstract** We used three second-generation sequencing platforms to develop 12 microsatellite markers for Gray's beaked whale (*Mesoplodon grayi*). There was a strong correlation between the amount of sequence obtained from each platform and the number of microsatellites recovered. In order to design reliable primers we found that it was important in all cases to eliminate low quality reads and trim sequences so that 80 % of each sequence had a PHRED score of >20. We found that tri- and tetranucleotide repeats produced less stutter enabling robust peak-calling and consistent genotyping. The microsatellites developed here will be useful for the population genetics and conservation studies of Gray's beaked whales.

**Keywords** Microsatellite loci · Roche 454 GS Junior · Illumina MiSeq · Life Technologies Ion Torrent · *Mesoplodon grayi* · Cetacea

The Gray's beaked whale is a medium sized ziphiid that is rarely seen alive (Thompson et al. 2013). This species is unique among beaked whales in that it frequently strands in groups yet little is known of its biology. We developed 12 microsatellite markers for the analysis of population structure and social relationships within this species so as to inform conservation management.

Three second-generation technologies; Roche 454 GS Junior, Illumina MiSeq and Life Technologies Ion Torrent were used to shot gun sequence high molecular weight DNA from a Gray's beaked whale sample. Library preparation was performed specific to each platform. For the Roche 454 GS Junior a library was constructed using the GS Rapid Library Prep Kit. Genomic DNA was fragmented by nebulization and size selected to 300–800 bp fragments. For the Illumina Miseq, genomic DNA was fragmented to 400 bp using a Covaris M220, and a library made using the Illumina TruSeq DNA PCR-Free Sample Preparation Kit. The library was sequenced using a MiSeq Reagent 500 cycle PE Kit, with loading at 15 pM and 1 % PhiX. For Ion Torrent library construction, DNA was fragmented using both the 300 bp and the 400 bp protocols on the Covaris M220 and libraries made using the Ion Torrent Kapa Standard Library Preparation Kit. The libraries were sequenced using 314 chips and 300 bp sequencing reagents.

All three platforms contain software with specific internal trimming protocols where primer, adaptor sequences and low quality reads are eliminated. Sequences from the Roche 454 GS Junior were imported into GENEIOUS (http://www.geneious.com/) to visualise quality and assess read length distribution. Further trimming and sequence elimination was conducted in GENEIOUS so that 80 % of each sequence had a PHRED score >20. Post-trimmed reads were analysed in QDD (Meglécz et al. 2010). This program allows for a simple and convenient method for detecting microsatellites and primer design from a large number of sequences reads. The minimum number of repeats was set at seven and product size limited to 90–250 bp. For all other parameters default

S. Patel (✉) · K. Thompson · C. Millar
The Allan Wilson Centre, School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland Mail Centre, Auckland 1142, New Zealand
e-mail: spat107@aucklanduni.ac.nz

S. Patel · K. Thompson · L. Williams · P. Tsai · R. Constantine · C. Millar
School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland Mail Centre, Auckland 1142, New Zealand

Published online: 01 April 2014

**Table 1** Details of second-generation sequencing data and microsatellite mining

| Platform | Total no. of reads (bp) | Average read length (bp) | No. of microsatellites detected (>7 repeats) | | | |
|---|---|---|---|---|---|---|
| | | | Di | Tri | Tetra | >Penta |
| Roche 454 | 139,995 | 500 | 50 | 13 | 12 | 1 |
| Ilumina Miseq | 17,268,568[a] × 2 | 250[a] | 294 | 27 | 41 | 3 |
| Ion Torrent 300 | 489,821 | 129 | 55 | 2 | 5 | 0 |
| Ion Torrent 400 | 313,065 | 123 | 58 | 4 | 5 | 0 |

[a] Paired-end reads

settings were used. Sequences generated from the IIlumina Miseq and Ion Torrent runs were assembled using SOAPdenovo2 and imported into GENEIOUS. The same 80 % sequence criterion and QDD pipeline was applied to the data. Table 1 contains a summary of sequence information and microsatellite mining data generated from all three technologies.

Primers were selected based on their location being at least 10 bp from the beginning or end of a repeat. Hexa-, penta- and tetranucleotide repeats were preferentially selected, followed by trinucleotide repeats and finally dinucleotide repeats.

Levels of polymorphisms were assessed using seven individuals. Initial screening involved PCR and visualisation of products on a 3 % agarose gel. A total of 111 primer pairs were tested and 17 were subsequently selected to be fluorescently labelled with NED, VIC or 6-FAM. Of these 12 were found to provide good quality genotyping data. PCR was performed in a 10 µl reaction mixture containing 10 ng of genomic DNA, 1 mg/ml BSA, 2.5 mM $MgCl_2$, 0.4 µM of each primer, 0.25 U of HotStarTaq *Plus* DNA polymerase (Qiagen), 1× Q solution (Qiagen). Cycling conditions were as follows; 5 cycles of 95 °C for 30 s, 60 °C for 45 s, and 72 °C for 60 s. The annealing temperature was reduced by 1 °C every cycle. This was followed by 23 cycles of 95 °C for 30 s, 55 °C for 45 s, and 72 °C for 60 s, then 30 min at 60 °C for the final extension. The exceptions to the above conditions were; locus Mg-27 with 0.2 µM of each primer, 56 °C annealing temperature for 27 cycles in total and locus Mg-32 with an annealing temperature of 65 °C dropping 1 °C per cycle for 10 cycles followed by 18 cycles at 55 °C. Amplified products were electrophoresed using an ABI 3130

**Table 2** Summary data of 12 microsatellite loci for Gray's beaked whale (*Mesoplodon grayi*)

| Locus | *A* | $H_o$ | $H_s$ | Repeat | Primer sequence 5′–3′ |
|---|---|---|---|---|---|
| Mg-1 | 16 | 0.854 | 0.819 | TCTT | F-TTTGTTTTGCAGCCATTTGA |
| | | | | | R-TCAATCCTTGGTCAGGGAAA |
| Mg-27 | 22 | 0.909 | 0.885 | CA | F-AGAGAGCCCACAAATTTCCC |
| | | | | | R-CTGAGTTTGGGATTCAAGGG |
| Mg-29 | 7 | 0.575 | 0.727 | CATC | F-TTCATCCATTCACCTACCCG |
| | | | | | R-TCTGAAGGTGGGAGGAGGTA |
| Mg-32 | 9 | 0.432 | 0.435 | ATT | F-TGGCAAATGGAACTAAGCCT |
| | | | | | R-CCACAGCCCTCATCTACCTG |
| Mg-37 | 9 | 0.666 | 0.607 | ATCC | F-GGTGAATCAGTCAGCCCTTG |
| | | | | | R-AGATGCTCAATGGGTGTTGG |
| Mg-58 | 10 | 0.737 | 0.795 | TATT | F-CATCCTCACTGGGGACACTT |
| | | | | | R-GCGTTCAAAAGAGGCAGAAC |
| Mg-75 | 2 | 0.306 | 0.367 | TTA | F-CCTCCTAGTGTTACTTGTCTGATCTC |
| | | | | | R-ACAGAGACAAGGCAATTCCG |
| Mg-78 | 7 | 0.626 | 0.741 | ATCC | F-TCACTTTCACTGTCTCTGCTCTG |
| | | | | | R-GCTTCTGCCAATGGGTAACA |
| Mg-88 | 7 | 0.850 | 0.783 | TTTA | F-GTATGGGTGGTGCCATGATT |
| | | | | | R-CACATGAGATCCCAGTGGAAA |
| Mg-95 | 8 | 0.596 | 0.643 | AAAT | F-AACAGTGAGAGGCCCGTGT |
| | | | | | R-TGTCATGAAATGAGTAAATCAACTG |
| Mg-98 | 8 | 0.658 | 0.567 | TAT | F-GCTGCTCTTTGATAGACGGG |
| | | | | | R-CCACGTTTAAGGAGCCAGAA |
| Mg-99 | 9 | 0.569 | 0.531 | TAAA | F-CATTCATTGTAAAGCAATTATACTCCA |
| | | | | | R-TTGACTGCAAAGAGACAGAGAA |

*A*, number of alleles; $H_o$, observed heterozygosity; $H_s$, expected heterozygosity after correction for sampling bias

sequencer (Applied Biosystems) and fragment sizes were analyzed using the microsatellite plugin in GENEIOUS. Linkage disequilibrium and deviation from Hardy–Weinberg equilibrium (HWE) were assessed in ARLEQUIN ver 3.5 (Excoffier and Lischer 2010) and GENODIVE (Meirmans and Van Tienderen 2004). A Bonferroni correction was applied to all data. Linkage disequilibrium was found to be non-significant between loci. Table 2 provides primer and allele information as well as the observed heterosygosity ($H_o$) and expected heterozygosity after correction for sampling bias ($H_s$). No loci were found to significantly deviate from HWE. No evidence of scoring error due to stuttering, large allele drop out and null alleles could be found using MICRO-CHECKER (Van Oosterhout et al. 2004).

**Acknowledgments** The authors would like to thank the following: the New Zealand Department of Conservation; New Zealand Genomics Ltd; the Centre for Genomics, Proteomics and Metabolomics; the University of Auckland for financial support.

## References

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10:564–567

Meglécz E, Costedoat C, Dubut V, Gilles A, Malausa T, Pech N, Martin JF (2010) QDD: a user-friendly program to select microsatellite markers and design primers from large sequencing projects. Bioinformatics 26:403–404

Meirmans PG, Van Tienderen PH (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. Mol Ecol Notes 4:792–794

Thompson K, Millar C, Baker CS, Dalebout M, Steel D, van Helden AL, Constantine R (2013) A novel conservation approach provides insights into the management of rare cetaceans. Biol Conserv 157:331–340

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538

Journal of Zoology

# Journal of Zoology

Volume 294 | September 2014 | Number 1

Volume 294 | Number 1 | September 2014 | Pages 1–76

WILEY Blackwell

ZSL
LIVING CONSERVATION

Journal of Zoology

Journal of Zoology. Print ISSN 0952-8369

# Large-scale multivariate analysis reveals sexual dimorphism and geographic differences in the Gray's beaked whale

K. F. Thompson[1,2], K. Ruggiero[3], C. D. Millar[1,2], R. Constantine[1,4] & A. L. van Helden[5]

1 School of Biological Sciences, University of Auckland, Auckland, New Zealand
2 The Allan Wilson Centre for Molecular Ecology and Evolution, School of Biological Sciences, University of Auckland, Auckland, New Zealand
3 Department of Statistics, University of Auckland, Auckland, New Zealand
4 Institute of Marine Science, University of Auckland, Auckland, New Zealand
5 20 Clunie Ave, Raumati South 5032, Wellington, New Zealand

**Keywords**
*Mesoplodon grayi*; morphology; cranial; linear discriminant analysis; sparse PCA; geographic variation.

**Correspondence**
Kirsten F. Thompson, School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand.
Email: ktho076@aucklanduni.ac.nz

Editor: Andrew Kitchener

Received 18 December 2013; revised 04 April 2014; accepted 24 April 2014

doi:10.1111/jzo.12151

## Abstract

Gray's beaked whales (*Mesoplodon grayi*) are medium-sized odontocete (toothed) cetaceans that are members of the family Ziphiidae. These animals inhabit deep oceanic waters and are rarely seen at sea. Most information is derived from stranded animals and there has been no systematic study of their morphology. We present a multivariate analysis of the morphology of Gray's beaked whales using 80 cranial measurements from 22 individuals and 13 external measurements from 50 individuals. Sparse principal component and linear discriminant function analyses were used to classify samples into sexes. Males and females have markedly different cranial morphology. In particular, females have longer skulls with longer more slender rostra (beaks) in comparison to males. Two variables, depth of the rostrum at mid-length and tip of rostrum to the right temporal fossa, can classify sex with 100% accuracy. The external body measurements used in this study are more prone to error as they were recorded by a number of observers on carcasses in differing states of decomposition and this is reflected in the level of variance in most measurements. However, analyses of these measurements showed a significant difference between sexes in the distance between (1) the tip of the rostrum to the genital slit, (2) the tip of the rostrum to the blowhole, as found in the cranial analyses and (3) tail fluke width where males have absolutely wider tail flukes than females. Differences in these same measurements were also found between animals stranded on the east and west coasts suggesting a degree of population separation across New Zealand. Finally, we present two linear models that enable the assignment of sex from either skull or external measurements. These models will be useful for future studies as well as the management of these whales and can be applied to archived data where genetic sex assignment is not possible.

## Introduction

Gray's beaked whales (*Mesoplodon grayi*) are one of at least 15 species within the genus *Mesoplodon*. These animals are rarely encountered at sea and most of the information on the biology of these whales is derived from stranded animals (Pitman, 2009). Most species are described from fewer than 40 specimens and one particularly rare species, the spade-toothed beaked whale (*M. traversii*), is known from only four strandings worldwide (Thompson *et al.*, 2012).

Gray's beaked whales are medium-sized odontocetes (toothed whales; Fig. 1). Most records of these whales are from south of 30° latitude, particularly along the coast of New Zealand, Australia, South Africa and South America, with one extralimital record from the Netherlands (Boschema, 1950; Dalebout *et al.*, 2004; IUCN, 2013; Thompson *et al.*, 2013). There are also two documented sightings from Antarctic and sub-Antarctic waters (Nishiwaki *et al.*, 1999; Scheidat *et al.*, 2011). Like other beaked whales, Gray's are thought to primarily occur in deep water along the continental shelf edge although some sightings have been made in shallow waters (Gales, Dalebout & Bannister, 2002; Dalebout *et al.*, 2004). Seasonality in these strandings suggests that these inshore movements are associated with calving or nursing (Thompson *et al.*, 2013).

Morphological descriptions of Gray's beaked whales are virtually non-existent. As in other beaked whales, species descriptions rely on the distinct position of mandibular tusk teeth in adult males (Fig. 1b). However, unlike other mesoplodonts, Gray's beaked whales of both sexes are commonly seen with small maxillary teeth in addition to these mandibular teeth. Most species of beaked whale are thought

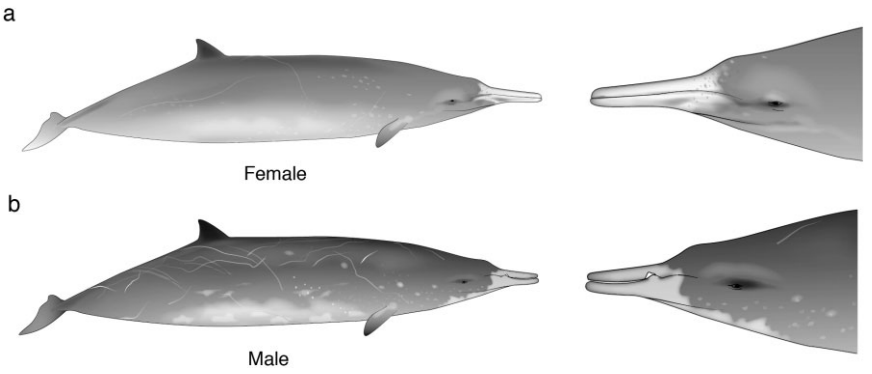



**Figure 1** External morphology of Gray's beaked whales (*Mesoplodon grayi*) showing a spindle-shaped body with a small, falcate dorsal fin and light grey coloration of the rostrum. (a) Female, (b) male.

to feed on squid and small fish. Although, MacLeod, Santos & Pierce (2003) suggested that Gray's beaked whale feed exclusively on fish there is other evidence from New Zealand to suggest they may prefer small squid species in this region (A. van H., unpubl. data). The adults of both sexes are known to possess a long rostrum (beak) that is generally white to light grey in colour in relation to other sympatric species of similar size such as the straptoothed whale (*M. layardii*; Fig. 1).

Sexual dimorphism has been recorded in several beaked whale species including Cuvier's (*Ziphius cavirostris*) and Blainville's (*M. densirostris*) beaked whales and both species of bottlenose whales (*Hyperoodon ampullatus* and *H. planifrons*; Beharse, 1971; Connor *et al*., 1998; Ralls & Mesnick, 2009). In Baird's beaked whales (*Berardius bairdii*), females are slightly larger than males (Ralls & Mesnick, 2009). Conversely, female northern bottlenose whales are smaller than males, who have much larger foreheads, or maxillary crests, that they use to head-butt during agonistic interactions (Gowans & Rendell, 1999). In the mesoplodonts, known sexual differences fall into three general categories: (1) eruption of tusk or mandibular teeth from the gums occurs only in adult males, (2) the mesorostral canal in the skull (a groove between the elongated premaxillae found in odontocetes) ossifies to a much greater extent in adult males and (3) the degree of body scarring in adult males, thought to indicate higher levels of aggressive interactions (Heyning, 1984; MacLeod, 1998). Sexual differences in cranial morphology have also been noted in Sowerby's (*M. bidens*) and Hubb's beaked whale (*M. carlhubbsi*), particularly in a bony abutment found in the mandibles of both species (MacLeod & Herman, 2004). In Blainville's beaked whale, adult females not only show a lesser degree of ossification of the mesorostral groove, but the rostrum is notably longer and less developed in terms of width and depth (Beharse, 1971). Male ginkgo-toothed beaked whales (*M. ginkgodens*) have less ossification of the mesorostral groove and less body scarring than in other beaked whale species. This is thought to suggest a lesser degree of intra-specific aggression (Heyning, 1984; MacLeod, 1998).

Although odontocetes show few sexual differences in external features, or overall size, there are exceptions (Ralls & Mesnick, 2009). The most extreme size differences occur in sperm whales, where females reach around 11 m in length (15 tonnes) but males regularly grow to 16 m in length and 45 tonnes (Ralls & Mesnick, 2009). Differences in the shape and size of appendages, head and teeth are seen in many species such as narwhals (*Monodon monoceros*), killer whales (*Orcinus orca*) and belugas (*Delphinapterus leucas*; Clark & Odell, 1999).

In addition to sexual dimorphism, morphological variation in mammals has been used to delineate population subdivisions (Elton, Dunn & Cardini, 2010; Mazák, 2010; Chen, Watson & Chou, 2011). Some species show distinct morphological differences over relatively small spatial scales, reflecting long-term barriers to gene flow. These barriers can occur as a result of a population being associated with a certain habitat or niche. For example, in the marine environment, certain species are associated with particular oceanographic features where their primary prey species aggregate (Bilgmann *et al*., 2007). In other cases, differences in foraging strategies or social behaviour can result in population divergence and hence distinct morphological types (Pitman *et al*., 2010). In rare or cryptic species, such baseline morphological data are often sparse and difficult to collect over sufficient spatial and temporal scales, particularly where there is ambiguity over accurate specimen identification.

Significant morphological variation has been shown in several cetaceans. These studies have largely focused on external features, such as total body length and dorsal fin height as well as cranial dimensions. Common dolphins (*Delphinus delphis*) in the eastern North Atlantic show cranial variation throughout their geographic range as do finless porpoises (*Neophocaena phocaenoides*) in South-East Asia (Jefferson, 2002; Murphy & Rogan, 2006; Murphy *et al*., 2006).

In this study, we report the results from multivariate analyses of cranial and external morphological measurements of Gray's beaked whales found stranded around the coast of New Zealand. First, we investigate if Gray's beaked whales exhibit sexual dimorphism in their cranial and external morphology. Second, we use the external measurements to identify any geographical variation in this species. Third, we provide predictive models with which to classify sex from both cranial and external morphology.

## Material and methods

### Study area and data collection

Gray's beaked whales commonly strand around the coast of New Zealand. A systematic programme of cetacean tissue and specimen data collection has operated throughout New Zealand since 1991 (Thompson *et al.*, 2013). In addition to collecting a tissue sample from each stranded beaked whale, the head of the animal was sometimes sent to the Museum of New Zealand Te Papa Tongarewa for curation. For 22 adult whales (12 females, 10 males), 80 cranial measurements were collected using standard vernier callipers (Vernier, Beaverton, OR, USA) with an accuracy of 0.1 mm. For eight individuals (five females, three males), species identification was confirmed from tissue samples according to Thompson *et al.* (2013) and sex was genetically identified by amplification of the SRY gene multiplexed with a ZFX/ZFY positive control (Aasen & Medrano, 1990; Gilson *et al.*, 1998; Thompson *et al.*, 2012). In all eight instances, both sex and species were concordant with external morphology and the degree of ossification of the mesorostral canal of the skull (Heyning, 1984). For the other 14 animals, sex was based solely on the degree of ossification in the mesorostral canal.

For external morphological data, the GPS coordinates for 50 individuals (30 females, 20 males) found between January 1991 and August 2011 were generated from stranding locations and plotted using ArcMAP 10.0 (www.esri.com) (Fig. 2). Both the species and sex of all these individuals were confirmed genetically using the methods outlined above. The New Zealand Department of Conservation (DOC) staff record 15 standard external morphometric measurements during examination of stranded whales and tissue sample collection. However, in some cases, some measurements were not collected due to the position of the carcass or its state of decomposition (Thompson *et al.*, 2013). Thus, 13 measurements have been selected for analyses and where there were missing values for variables (16% of data points) the mean for that sex and region (east and west coasts) were imputed where necessary (Supporting Information Fig. S1).

Individuals of both sexes exceeding 4.5 m in length were defined as adult following conservative estimates of maturity (Thompson *et al.*, 2013). The presence of erupted teeth are rarely reported, therefore this character was not used to determine maturity in males.

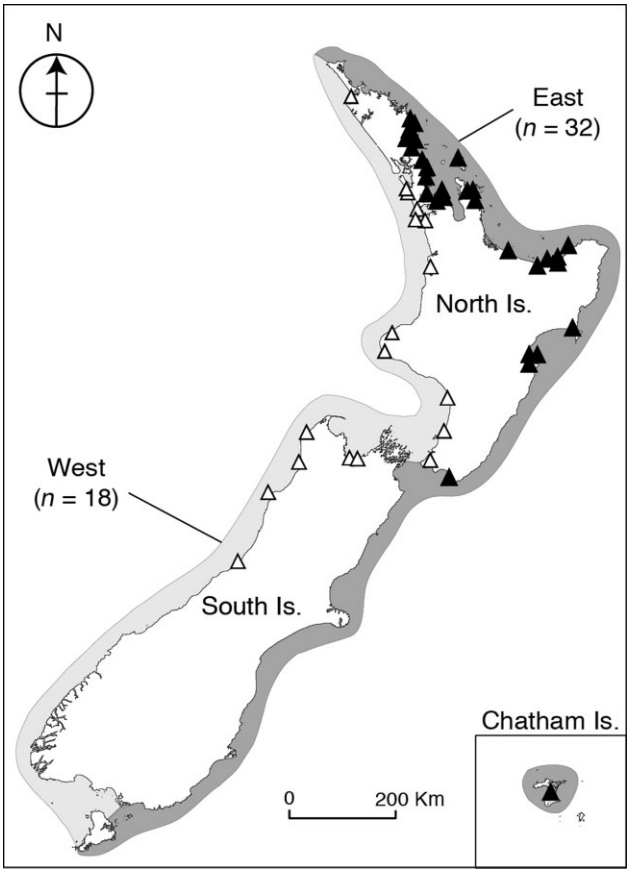



**Figure 2** Locations of all stranded Gray's beaked whales (*Mesoplodon grayi*) where external morphometric measurements were available: 32 from the east (18 females, 14 males) and 18 from the west (12 females, 6 males). East and west geographical divisions are depicted by shading and black and white symbols with the Chatham Islands included in the east.

### Data analyses

Multiple two-tailed two-sample *t*-tests using Welch's method (i.e. equal group variances is not assumed) were conducted to assess statistical significance of sexual differences in cranial measurements (Welch, 1947). To control the proportion of variables incorrectly declared as sexually dimorphic, *P*-values from these *t*-tests were adjusted using a false discovery rate of 0.05 (Benjamini & Hochberg, 1995). Two-way analyses of variance (ANOVAs) were performed on each of the 13 external measurements to assess their dependence on sex and region. Since the data were unbalanced (i.e. unequal numbers of whales in each sex-region combination), adjustments were made for order of fitting of sex and region in the linear model. Thus, the reported ANOVA results are sex adjusted for region (i.e. sex is fitted after region) and region adjusted for sex (i.e. region is fitted after sex).

A biplot associated with the principal component analysis (PCA) on the data was generated to provide a simple two-dimensional visualization of the salient features of the data, presenting information on both individual whales (points) and variables (arrows). The angle between arrows indicates the strength of correlation between pairs of variables. The arrangement of the points in two-dimension space is informative of similarities and/or differences between whales with respect to the set of physical attributes measured.

Cranial measurements were analysed for differences between males and females using a combination of sparse PCA (sPCA; Zou, Hastie & Tibshirani, 2006) and linear discriminant function analysis (LDA). The sPCA algorithm was developed for dimensionality reduction of wide datasets, that

is, the number of variables (80 cranial measurements) is large relative to the number of samples (22 whales) upon which measurements were made. It is based on the reformulation of traditional PCA as a regression-optimization problem, thereby enabling penalization of the absolute magnitude of the regression coefficients via least absolute shrinkage and selection operator (Tibshirani, 1996). It is convenient for datasets comprising of highly correlated predictor variables, as is the case here, where standard regression would likely result in coefficients that are 'too large'. The application of a penalty, or sparsity constraint, yields principal components (PCs) having most loadings set to zero. Consequently, sPCA leads to the identification of a very small subset of the original variables, which explain a substantial proportion of the variation in the data. The sPCA was performed using the R package mixOmics (Lê Cao, González & Déjean, 2009), although it can also be performed online using a web interface (see http://mixomics.qfab.org/). The relationships between the variables selected by the sPCA are visualized by projecting them onto concentric circles centred at zero and having radii 0.5 and 1, known as a correlation circle (Gonzales *et al.*, 2013). Variables appearing close to one another and forming acute angles (where lines drawn from the origin to the projected points) are positively correlated, while variables far from one another and forming obtuse angles are negatively correlated. LDA was subsequently used to derive a sex classifying function (McLachlan, 2004), or classifier, based on the sPCA-identified variables. The performance of the classifier was assessed using leave-one-out cross-validation predictions of each class.

For external measurements a different method was required as the sPCA method showed little dimensional reduction and no variables were identified as contributing the highest variance. Therefore, LDA based on ANOVA-identified variables was used to assign samples into sex. All statistical analyses were performed in R 3.0.1 and a $P$-value of $<0.05$ was considered significant (R Development Core Team, 2011).

## Results

### Do male and female Gray's beaked whales have different cranial morphology?

Many of the 80 cranial measurements were highly correlated. It was not surprising, therefore, to find over one-third of these measurements (32 variables) show significant differences between males and females even after applying a false discovery rate of 5% (Benjamini & Hochberg, 1995) (adj. $P \leq 0.001$, Supporting Information Table S1). For example, the following variables fall into this category (mean ± SE are given for each sex), the overall condylobasal, or skull length, is greater for adult females than males (females = 859.4 ± 39.2 mm, males = 792.4 ± 13.1 mm). All measures of rostrum length (nested within condylobasal length) and mandible length were significantly larger in females than males. Depth of the rostrum at mid-point was greater in males than in females (females = 35.5 ± 3.3 mm, males = 41.3 ± 1.3 mm). Males also have deeper mandibles at the mid-point along the alveolus (females = 51.0 ± 4.9 mm, males = 61.3 ± 1.9 mm) and larger alveoli both in terms of length (left alveolus, females = 36.0 ± 13.8 mm, males = 72.5 ± 4.7 mm) and width (left alveolus, females = 9.8 ± 1.7 mm, males = 13.3 ± 0.9 mm).

Overall, females have longer, more slender skulls with significantly longer and more slender rostra, while males have shorter and wider skulls. Not only are the males' rostra more robust, in that they are shorter and deeper at the mid-point, but the mandibles are also deeper at the position of the alveolus.

PCA of 80 cranial measurements from 22 individuals (12 females, 10 males) showed that the first two PCs explained 64% of the variation in the data. A biplot of the first two PCs showed that many of the variables, for example, all measures of rostrum length and length of the skull (condylobasal length) were strongly positively correlated (Fig. 3a). Arrows describing overall length measurements of the skull and rostrum measurements were parallel with each other and pointing towards positive values of PC1, indicating that these variables were highly positively correlated with each other and PC1. Other arrows describing the variables measuring the height, width and breadth of the skull, as well as aspects of the mandibular teeth, form an arc from positive values of PC1 to negative values of PC1. Although many variables show positive correlation to one another several are independent and many contribute a low degree of variance within the data (i.e. arrows are comparatively short). The distance between the nasals on the vertex was negatively correlated with skull height and width variables and independent of skull length and teeth measurements. However, this variable did not account for a large proportion of variance in the PCA. Points indicate that males and females are separated in terms of their multivariate skull dimensions.

Sparse PCA indicated that two variables, (1) depth of the rostrum at the mid-point, and (2) tip of the rostrum to the right temporal fossa accounted for the highest variance (Fig. 3b). The model derived from discriminant function analysis, using these two variables predicted sex accurately in 100% of cases (Fig. 3c). Therefore, depth of the rostrum at the mid-point (as measured at half distance from the antorbital notches) and the distance from the tip of the rostrum to the right temporal fossa, can be used as predictors of sex in Gray's beaked whale skulls (Supporting Information Fig. S2) using the formula below.

$$\text{Predicted Sex} = \begin{cases} \text{Female} & \text{if } f(x) < -12.7 \\ \text{Male} & \text{otherwise} \end{cases}$$

where

$$f(x, y) = -0.031x + 0.314y$$

and where $x$ and $y$ denote the depth of rostrum at the mid-point and distance from the tip of the rostrum, respectively.

### Do New Zealand Gray's beaked whales differ in their external morphology by sex or region?

Fifty adults (30 females, 20 males) were tested for differences in 13 external body measurements. The body measurements

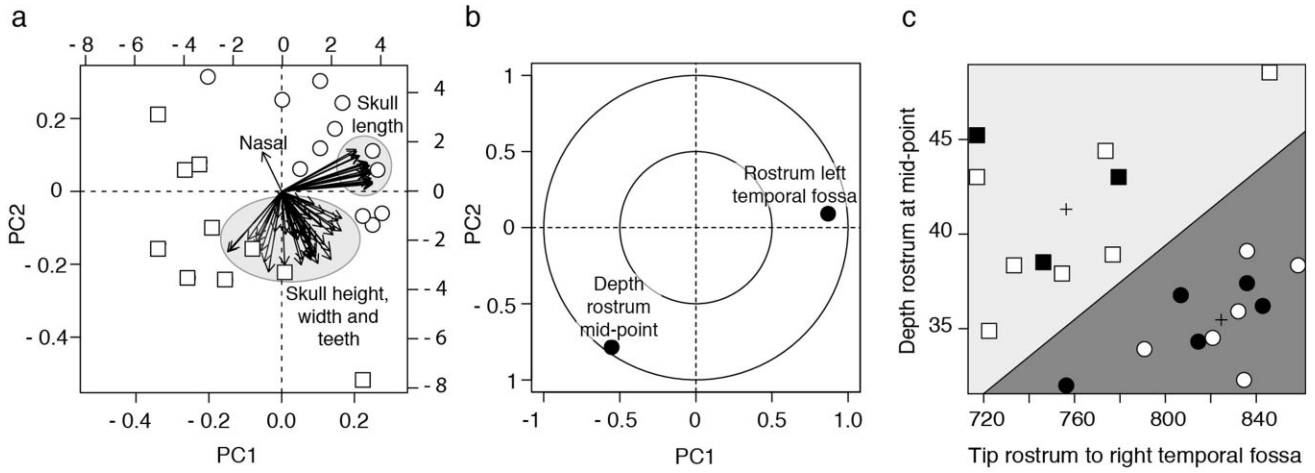



**Figure 3** Results of sparse principal component analysis (sPCA) used to test for differences in the cranial morphology of Gray's beaked whales (*Mesoplodon grayi*) found stranded around New Zealand (12 females, 10 males). (a) Biplot of principal components with sample cases where squares denote males and circles denote females. The arrow describing the depth of the nasals on the vertex bone is shown by nasal, while various measurements of skull and rostrum length are shown by shading as are all other variables measured. (b) Correlation circles for sPCA. Inner and outer circles denote correlations of 0.5 and 1, respectively. (c) Partition plot illustrating results of the linear discriminant analysis where filled shapes denote individuals with a verified genetic sex and the crosses denote the centroid.

**Table 1** Summary of results from two-way ANOVAs relating each external measurement variable to sex and region, showing common numerator degrees of freedom (DF) but different *F*-statistic and *P*-value for sex adjusted for region and region adjusted for sex

| | | Sex | | | | Region | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | Female | Male | | | East | West |
| External measurement | DF | *F* | *P* | Mean ± SE | Mean ± SE | *F* | *P* | Mean ± SE | Mean ± SE |
| Total body length | 46 | 1.14 | NS | 4784.6 ± 38.0 | 4706.0 ± 49.8 | 1.14 | NS | 4762.7 ± 36.4 | 4727.9 ± 51.0 |
| Tip of lower jaw to tip dorsal fin | 43 | 3.11 | NS | 3239.7 ± 46.0 | 3090.1 ± 58.7 | 3.11 | NS | 3162.7 ± 43.8 | 3167.2 ± 60.3 |
| Tip of lower jaw to anus | 40 | 1.37 | NS | 3520.5 ± 43.3 | 3440.0 ± 63.4 | 1.37 | NS | 3441.0 ± 40.9 | 3519.5 ± 65.0 |
| Tip of lower jaw to genital slit | 37 | 7.44 | 0.010 | 3313.0 ± 53.1 | 3078.0 ± 77.0 | 7.44 | 0.010 | 3188.5 ± 51.6 | 3202.5 ± 78.0 |
| Tip of lower jaw to insertion of flipper | 41 | 3.39 | NS | 1125.9 ± 35.4 | 1248.4 ± 47.5 | 3.39 | NS | 1118.5 ± 33.7 | 1255.7 ± 48.7 |
| Tip of lower jaw to blowhole | 32 | 9.25 | 0.005 | 675.8 ± 11.9 | 615.4 ± 17.6 | 9.25 | 0.005 | 655.6 ± 11.4 | 635.7 ± 18.0 |
| External length of flipper | 42 | 0.20 | NS | 501.4 ± 10.3 | 509.8 ± 15.4 | 0.20 | NS | 501.4 ± 9.8 | 509.8 ± 15.7 |
| Internal length of flipper | 16 | 0.20 | NS | 348.3 ± 16.3 | 362.5 ± 28.7 | 0.20 | NS | 334.7 ± 24.8 | 376.1 ± 21.7 |
| Flipper width | 39 | 1.19 | NS | 162.7 ± 6.0 | 169.1 ± 8.4 | 1.19 | NS | 160.2 ± 5.6 | 171.6 ± 8.7 |
| Tail fluke width | 38 | 9.79 | 0.003 | 1147.5 ± 18.1 | 1225.8 ± 22.7 | 9.79 | 0.003 | 1134.2 ± 17.5 | 1239.2 ± 23.2 |
| Length of rostrum | 37 | 0.19 | NS | 379.1 ± 14.4 | 372.8 ± 17.7 | 0.19 | NS | 368.8 ± 12.9 | 383.1 ± 18.7 |
| Length of gape | 33 | 2.48 | NS | 508.0 ± 15.2 | 469.2 ± 20.5 | 2.48 | NS | 468.5 ± 14.4 | 508.8 ± 21.1 |
| Height of dorsal fin | 39 | 2.54 | NS | 223.6 ± 9.2 | 247.8 ± 12.3 | 2.54 | NS | 234.1 ± 8.9 | 237.3 ± 12.4 |

Mean ± standard error of the mean (SE) are shown for each sex and region. All measurements are in mm and non-significant values ($P > 0.05$) are denoted by NS.

from animals of different sexes from the two regions were compared using an ANOVA (Table 1). This showed no interaction between sex and region in that observed differences between sexes or regions were of the same magnitude in both east and west regions or sexes, respectively.

Sex and regional differences were found in the distance between the tip of the lower jaw to both the genital slit and the blowhole and in the width of the tail fluke. The genital slit opening is known to be more anterior in male cetaceans (Clark & Odell, 1999). The larger rostrum to blowhole measurement found in females is concordant with the findings of the analyses of cranial measurements in that females clearly have longer rostra than males. The width of the tail fluke was also found to be significantly larger in males than in females. However, these sexual differences do not result in significant dimorphism in the overall total body length of males and females. Mean total body length (± SE) of females was 4.78 (± 0.04) m (maximum = 5.36 m) and males 4.72

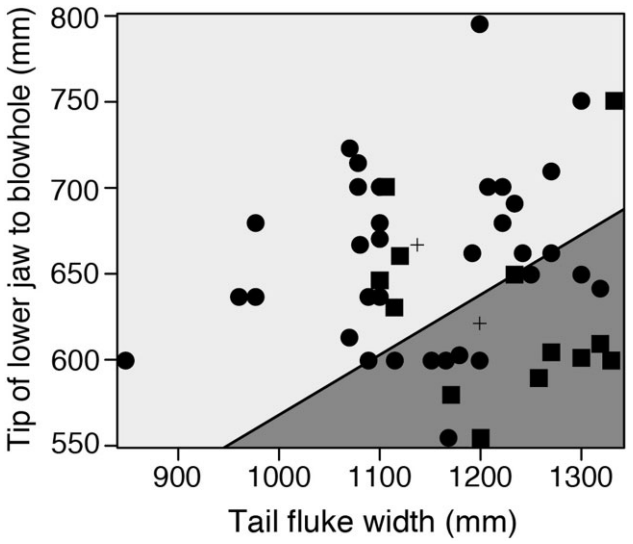



**Figure 4** Partition plot illustrating the results of linear discriminant analysis. All individuals ($n = 50$) were genetically sexed. Squares denote males and circles denote females, crosses denote the centroid.

(± 0.04) m (maximum = 5.2 m; Supporting Information Fig. S3). In general, whales on the west were slightly longer in the mid-body region and had wider tail flukes. Differences in these variables were found between regions but the additive effect of these differences is constant for both sexes.

PCA of external measurements showed that the first three principal components explained only 51% of the variation within the data. No meaningful dimensional reduction was achieved through this method. All externally measured variables had a high degree of variance reflecting both the difficulty in measuring these large, decomposing animals and the error introduced by multiple, sometimes non-specialist, observers.

We used DFA to construct a classifier of sex using the measurements shown to be significantly different between sexes using ANOVA. As the distance between the tip of the rostrum and the blowhole and that between the rostrum and genital slit were correlated, only (1) the tip of the rostrum to blowhole and (2) tail fluke width were required in the model to predict sex with almost 80% accuracy using the formula below (Fig. 4).

$$\text{Sex} = \begin{cases} \text{Female} & \text{if } f(x) < 346.3 \\ \text{Male} & \text{otherwise} \end{cases}$$

where

$$f(x, y) = -0.019x + 0.007y$$

and where $x$ and $y$ denote the distance from the tip of the lower jaw to the blowhole and the tail fluke width, respectively.

# Discussion

## Do male and female Gray's beaked whales have different cranial morphology?

There are size and shape differences between male and female Gray's beaked whale cranial morphology. Females have longer skulls with longer rostra that are shallower at the mid-point, giving an overall appearance of a long, slim skull in comparison to males. This is concordant with the reported sexual differences in the cranial morphology of Blainville's beaked whales (Beharse, 1971). Male Gray's beaked whales have shorter, more robust skulls, with mandibles that are deeper at the position of the alveolus, or tooth socket. Moreover, the alveolus is larger in males than in females. This heavier mandible is similar to that found in Sowerby's beaked whales (MacLeod & Herman, 2004). Several ontogenic changes associated with male maturity were noted in Sowerby's beaked whale including an overall thickening of the rostrum and mesorostral canal in addition to a deepening of the mandibles posterior to the alveolus that become longer and wider with maturity.

Male beaked whales are presumed to use their teeth in intra-specific aggressive interactions (Heyning, 1984; MacLeod, 1998). The thickening and increased ossification seen in the rostrum, as well as the heavy mandibles, are thought to give strength and support during these interactions between male beaked whales (Beharse, 1971; MacLeod, 1998; MacLeod & Herman, 2004). The sexual differences in terms of cranial morphology shown here for Gray's beaked whales imply that these social interactions are also likely in this species. This is supported by anecdotal observations of males with more heavy body scarring in comparison to females (A. van H. pers. obs.). Other potential causes of sexual differences in cranial morphology may be those associated with differences in the acoustic properties of the skull and variation in vocalizations between sexes. Many odontocetes show these types of differences and this may also be the case in Gray's beaked whales (Antunes *et al.*, 2011).

## Do New Zealand Gray's beaked whales differ in their external morphology by sex or region?

Our results show females and males are very similar in overall size with regards to total body length (although the largest animals found were females) and also in terms of the size of most appendages, that is, dorsal and pectoral fin length. However, male Gray's beaked whales have shorter rostra and larger fluke widths. In addition, the variance recorded within measurements may highlight the difficulty in collecting standardized data using non-specialist observers on specimens that are often degraded. The model, developed from a DFA, will facilitate the sexing of individuals in the field and can be particularly useful when mandibular teeth have not erupted in males or where genetic testing is not possible and only morphological data are available.

Interestingly, the results of this study indicate there are east-west geographic differences in the morphology of appendages, but not overall body size of Gray's beaked whales found stranded around the coast of New Zealand. In cetacean species where morphological differences have been recorded, differentiation was attributed to reproductive isolation. For example, common dolphins in the north-eastern Atlantic show significant differences across a large geographical area relative to the range of the species (Murphy *et al*., 2006). In other cases such as in the bottlenose dolphin of the Black Sea, a known physical barrier prevents dispersal of individuals to adjacent areas (Perrin, 1984; Viaud-Martinez *et al*., 2008). Similarly, bottlenose dolphin (*Tursiops truncatus*) populations in Australia show significant geographical variation across their range (Bilgmann *et al*., 2007; Ansmann *et al*., 2012). In these dolphins, and in pilot whales (*Globicephala melas*) that strand around the coast of Tasmania, it is thought that specific habitat and/or social specializations between populations have led to a barrier to gene flow over relatively fine geographic scales (Oremus *et al*., 2013).

## Conclusions

This research provides some of the first analyses of the morphology of Gray's beaked whales and clear evidence of sexual dimorphism in this species. Moreover, we are the first to evaluate geographical variation in morphology in Gray's beaked whales anywhere in the world. Despite DNA sexing becoming more readily accessible in many countries, there are not always, samples available nor the expertise to conduct these analyses. Therefore, we provide a simple method that can be used to sex this species both in the field and from retrospective analyses of skulls and external measurements.

## Acknowledgements

The authors would like to thank C. Scott Baker for initiating the New Zealand Cetacean Tissue Archive and all Department of Conservation staff who collected data from stranded whales, and Laura Boren and Laura Wakelin for database assistance. Thanks to Vivian Ward, Kai Xiong and Selina Patel and Sinéad Murphy for comments. Financial assistance was provided by the University of Auckland Faculty Research Development Fund and the Department of Conservation. K.F.T. was funded by a University of Auckland Masters Scholarship, George Mason Charitable Trust Scholarship 2011 and OMV NZ Ltd Scholarship 2011. All samples are held under DOC permit Rnw/HO/2009/03 at the University of Auckland.

## References

Aasen, E. & Medrano, J.F. (1990). Amplification of the ZFX and ZFY genes for sex identification in humans, cattle, sheep and goats. *Nat. Biotechnol.* **8**, 1279–1281.

Ansmann, I.C., Guido, J.P., Lanyon, J.M. & Seddon, J.M. (2012). Fine-scale genetic population structure in a mobile marine mammal: inshore bottlenose dolphins in Moreton Bay, Australia. *Mol. Ecol.* **21**, 4472–4485.

Antunes, R., Schulz, T., Gero, S., Whitehead, H., Gordon, J. & Rendell, L. (2011). Individually distinctive acoustic features in sperm whale codas. *Anim. Behav.* **81**, 723–730.

Benjamini, Y. & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B* **57**, 289–300.

Beharse, J.C. (1971). Maturity and sexual dimorphism in the skull, mandible, and teeth of the beaked whale, *Mesoplodon densirostris*. *J. Mamm.* **52**, 297–315.

Bilgmann, K., Möller, L.M., Harcourt, R., Gibbs, S.E. & Beheregaray, L.B. (2007). Genetic differentiation in bottlenose dolphins from South Australia: association with local oceanography and coastal geography. *Mar. Ecol. Prog. Ser.* **341**, 265–276.

Boschema, H. (1950). Maxillary teeth in specimens of *Hyperoodon rostratus* (Muller) and *Mesoplodon grayi* Von Haast stranded on the Dutch coasts. *Koninklijke Nederlandse Akademie Wetenschappen* **53**, 775–786.

Chen, I., Watson, A. & Chou, L.-S. (2011). Insights from life history traits of Risso's dolphins (*Grampus griseus*) in Taiwanese waters: shorter body length characterizes northwest Pacific population. *Mar. Mamm. Sci.* **27**, E43–E64.

Clark, S.T. & Odell, D.K. (1999). Allometric relationships and sexual dimorphism in captive killer whales (*Orcinus orca*). *J. Mamm.* **80**, 777–785.

Connor, R.C., Mann, J.J., Tyack, P.L. & Whitehead, H. (1998). Social evolution in toothed whales. *Trends Ecol. Evol.* **13**, 228–232.

Dalebout, M.L., Russell, K.G., Little, M.J. & Ensor, P. (2004). Observations of live Gray's beaked whales (*Mesoplodon grayi*) in Mahurangi Harbour, North Island, New Zealand, with a summary of at-sea sightings. *J. R. Soc. N. Z.* **34**, 347–356.

Elton, S., Dunn, J. & Cardini, A. (2010). Size variation facilitates population divergence but does not explain it all: an example study from a widespread African monkey. *Biol. J. Linnean Soc.* **101**, 823–843.

Gales, N., Dalebout, M.L. & Bannister, R.L. (2002). Genetic identification and biological observation of two free-swimming beaked whales: Hector's beaked whale (*Mesoplodon hectori*, Gray, 1871), and Gray's beaked whale (*Mesoplodon grayi*, von Haast, 1876). *Mar. Mamm. Sci.* **18**, 551–554.

Gilson, A., Syvanen, M., Levine, K. & Banks, J. (1998). Deer gender determination by polymerase chain reaction: validation study and application to tissues, bloodstains, and hair forensic samples from California. *Calif. Fish Game* **84**, 159–169.

Gonzales, I., Lê Cao, K.-A., Davis, M.J. & Dejean, S. (2013). Visualising associations between paired 'omics' data sets. *BioData Min.* **5**, 1–23.

Gowans, S. & Rendell, L. (1999). Head-butting in northern bottlenose whales (*Hyperoodon ampullatus*): a possible function for big heads? *Mar. Mamm. Sci.* **15**, 1342–1350.
Heyning, J.E. (1984). Functional morphology involved in intraspecific fighting of the beaked whale, *Mesoplodon carlhubbsi*. *Can. J. Zool.* **62**, 1645–1654.
IUCN. (2013). IUCN Red List of Threatened Species. Retrieved February 15, 2013, from: www.iucnredlist.org
Jefferson, T.A. (2002). Preliminary analysis of geographic variation in cranial morphometrics of the finless porpoise (*Neophocaena phocaenoides*). *Raffles Bull. Zool.* **10**, 3–14.
Lê Cao, K.-A., González, I. & Déjean, S. (2009). integrOmics: an R package to unravel relationships between two omics data sets. *Bioinformatics* **25**, 2855–2856.
MacLeod, C., Santos, M. & Pierce, G. (2003). Review of data on diets of beaked whales: evidence of niche separation and geographic segregation. *J. Mar. Biol. Assoc. U. K.* **83**, 651–665.
MacLeod, C.D. (1998). Intraspecific scarring in odontocete cetaceans: an indicator of male 'quality' in aggressive social interactions? *J. Zool. (Lond.)* **244**, 71–77.
MacLeod, C.D. & Herman, J.S. (2004). Development of tusks and associated structures in *Mesoplodon bidens* (Cetaceae, Mammalia). *Mammalia* **68**, 175–184.
Mazák, J.H. (2010). Craniometric variation in the tiger (*Panthera tigris*): implications for patterns of diversity, taxonomy and conservation. *Mamm. Biol.* **75**, 45–68.
McLachlan, G. (2004). *Discriminant analysis and statistical pattern recognition*. London: John Wiley & Son.
Murphy, S. & Rogan, E. (2006). External morphology of the short-beaked common dolphin, *Delphinus delphis:* growth, allometric relationships and sexual dimorphism. *Acta Zool.* **87**, 315–329.
Murphy, S., Herman, J.S., Pierce, G.J., Rogan, E. & Kitchener, A.C. (2006). Taxonomic status and geographical cranial variation of common dolphins (*Delphinus*) in the eastern North Atlantic. *Mar. Mamm. Sci.* **22**, 573–599.
Nishiwaki, S., Tohyama, D., Yuzu, S., Bando, T., Watanabe, M., Kitajima, A., Takeda, S., Murase, H., Otose, S., OKubo, J., Tsutsui, S., Takatsuki, M. & Kinoshita T. (1999). Cruise Report of the Japanese Whale Research Program under special Permit in the Antarctic (JARPA) Area V and Western Part of Area VI in 1998/1999. Paper SC/51/020 presented to the IWC Scientific Committee.
Oremus, M., Gales, R., Kettles, H. & Baker, C.S. (2013). Genetic evidence of mulitple matrilines and spatial disruption of kinship bonds in mass strandings of long-finned pilot whales, *Globicephala melas*. *J. Hered.* **104**, 301–311.
Perrin, W.F. (1984). Patterns of geographical variation in small cetaceans. *Acta Zool. Fennica* **172**, 137–140.
Pitman, R. (2009). Mesoplodont whales (*Mesoplodon spp.*), In *Encyclopedia of marine mammals*. 721–726. Perrin, W.F., Würsig, B. & Thewissen, J.G.M. (Eds). San Diego, CA: Academic Press.
Pitman, R.L., Durban, J.W., Greenfelder, M., Guinet, C., Jorgensen, M., Olson, P.A., Plana, J., Tixier, P. & Towers, J.R. (2010). Observations of a distinctive morphotype of killer whale (*Orcinus orca*), type D, from sub Antarctic waters. *Polar Biol.* **34**, 303–306.
R Development Core Team (2011). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. ISBN 3-900051-07-0, (URL: http://www.R-project.org/).
Ralls, K. & Mesnick, S.L. (2009). Sexual Dimorphism, In *Encyclopedia of marine mammals*. 1005–1011. Perrin, W.F., Würsig, B. & Thewissen, J.G.M. (Eds). San Diego: Academic Press.
Scheidat, M., Friedlaender, A., Kock, K.-H., Lehnert, L., Boebel, O., Roberts, J. & Williams, R. (2011). Cetacean surveys in the Southern Ocean using icebreaker-supported helicopters. *Polar Biol.* **34**, 1513–1522.
Thompson, K., Baker, C.S., van Helden, A., Patel, S., Millar, C. & Constantine, R. (2012). The world's rarest whale. *Curr. Biol.* **22**, R905–R906.
Thompson, K., Millar, C.D., Baker, C.S., Dalebout, M., Steel, D., van Helden, A.L. & Constantine, R. (2013). A novel conservation approach provides insights into the management of rare cetaceans. *Biol. Conserv.* **157**, 331–340.
Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO. *J. R. Stat. Soc. Series B* **58**, 267–288.
Viaud-Martinez, K.A., Brownell, R.L. Jr., Komnenou, A. & Bohonak, A.J. (2008). Genetic isolation and morphological divergence of Black Sea bottlenose dolphins. *Biol. Conserv.* **141**, 1600–1611.
Welch, B.L. (1947). The generalization of 'Student's' problem when several different population variances are involved. *Biometrika* **34**, 28–35.
Zou, H., Hastie, T. & Tibshirani, R. (2006). Sparse principal component analysis. *J. Comput. Graph. Stat.* **15**, 265–286.

## Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

**Figure S1.** External morphometric measurements taken by Department of Conservation (DOC) staff after cetacean strandings. Numbers refer to measurements: (1) total body length; (2) tip of jaw to tip of dorsal fin; (3) tip of the jaw to genital slit; (4) tip of jaw to anus; (5) tip of the jaw to the forward insertion of the flipper; (6) tip of jaw to centre of blowhole; (7) external length of flipper; (8) internal length of flipper (9) greatest width of flipper; (10) length of rostrum; (11) length of gape; (12) tail fluke width; (13) dorsal fin height. Genital slits and anus also shown in male and female.
**Figure S2.** Cranial measurements used to predict sex of Gray's beaked whale (*Mesoplodon grayi*) based on linear models.

(a) Depth of the rostrum at the mid-point. (b) Distance from the tip of the rostrum to the right temporal fossa.

**Figure S3.** Total body length for adult Gray's beaked whales (*Mesoplodon grayi*) found stranded around the coast of New Zealand (30 females, 20 males). Note: lower limit is based on conservative estimate of maturity (4.5 m for both sexes) in the absence of necropsy data.

**Table S1.** Mean size and ranges for 80 cranial measurements from adult Gray's beaked whales (*Mesoplodon grayi*) (10 males, 12 females). Sexual differences are shown by pairwise Student *t*-tests with *P*-values adjusted according to the Benjamini & Hochberg (1995) method. All measurements are in mm and non-significant values ($P > 0.05$) are denoted by NS.

## Supplementary Information

# Large-scale multivariate analysis reveals sexual dimorphism and geographic differences in the Gray's beaked whale.

K. F. Thompson[1,2], K. Ruggiero[3], C. D. Millar[1,2], R. Constantine[1,4] & A. L. van Helden[5]

1 School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

2 The Allan Wilson Centre for Molecular Ecology and Evolution, School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

3 Department of Statistics, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

4 Institute of Marine Science, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

5 20 Clunie Avenue, Raumati South 5032, Wellington, New Zealand

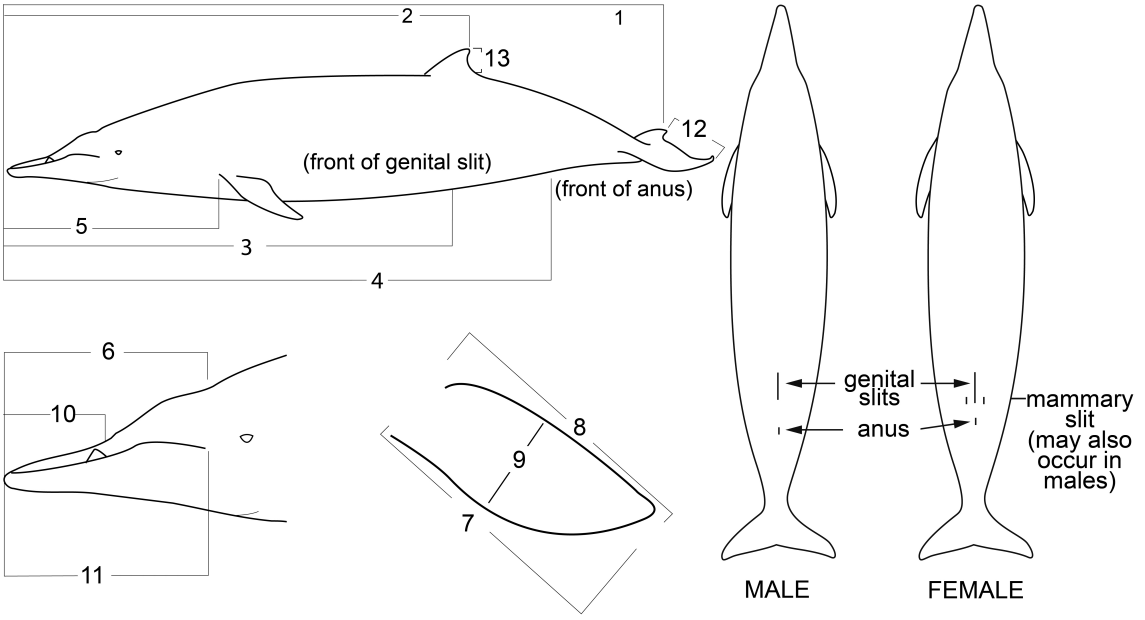



Figure S1. External morphometric measurements taken by Department of Conservation (DOC) staff after cetacean strandings. Numbers refer to measurements: (1) total body length; (2) tip of jaw to tip of dorsal fin; (3) tip of the jaw to genital slit; (4) tip of jaw to anus; (5) tip of the jaw to the forward insertion of the flipper; (6) tip of jaw to centre of blowhole; (7) external length of flipper; (8) internal length of flipper (9) greatest width of flipper; (10) length of rostrum; (11) length of gape; (12) tail fluke width; (13) dorsal fin height. Genital slits and anus also shown in male and female.

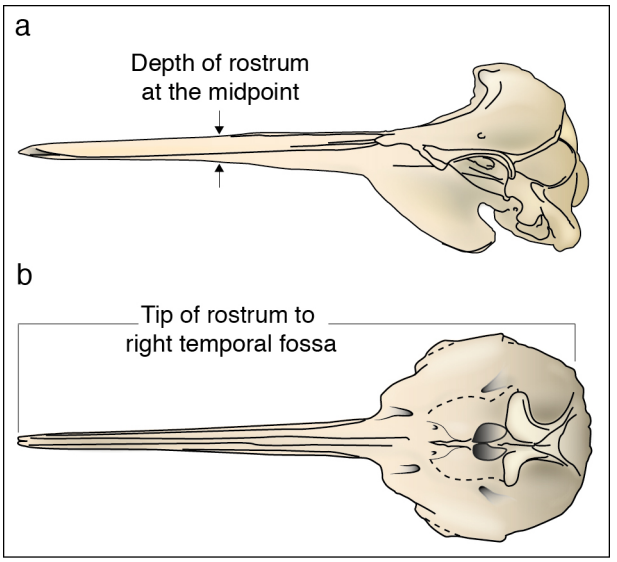



Figure S2 Cranial measurements used to predict sex of Gray's beaked whales (*Mesoplodon grayi*) based on linear models. (a) Depth of the rostrum at the mid-point. (b) Distance from the tip of the rostrum to the right temporal fossa.

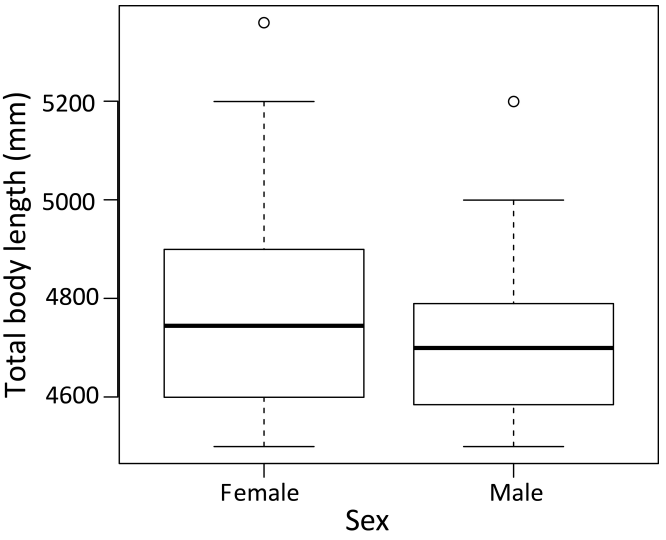



Figure S3. Total body length for adult Gray's beaked whales (*Mesoplodon grayi*) found stranded around the coast of New Zealand (30 females, 20 males). Note: lower limit is based on conservative estimate of maturity (4.5 m for both sexes) in the absence of necropsy data.

Table S1. Mean size and ranges for 80 cranial measurements from adult Gray's beaked whales (*Mesoplodon grayi*) (10 males, 12 females). Sexual differences are shown by pairwise Student t-tests with *p* values adjusted according to the Benjamini & Hochberg (1995) method. All measurements are in mm and non-significant values ($p > 0.05$) are denoted by NS.

| | Cranial measurement | Female mean ± SE (range) | Male mean ± SE (range) | T-test p | Adjusted p |
|---|---|---|---|---|---|
| 1 | Condylobasal length | 859.4 ± 39.2 (782.5 – 904.0) | 792.4 ± 13.1 (739 – 879) | <0.001 | 0.001 |
| 2 | Rostrum length to posterior extension of maxillary plate | 805.0 ± 38.5 (735.0 – 838.5) | 736.0 ± 13.9 (637 - 819) | <0.001 | 0.002 |
| 3 | Rostrum length to anterior margin of superior nares | 659.4 ± 36.4 (598 – 696) | 589.8 ± 10.9 (543 – 654.5) | <0.001 | <0.001 |
| 4 | Rostrum length to anterior point of premaxillary crest | 697.5 ± 36.4 (639 – 732) | 623.6 ± 10.9 (576 – 691) | <0.001 | <0.001 |
| 5 | Rostrum length to posterior premaxilla on tip of right premaxillary crest | 724 ± 35 (657 – 759) | 659.7 ± 10.9 (613 – 724) | <0.001 | <0.001 |
| 6 | Rostrum length to posterior extension of right temporal fossa | 823.6 ± 35.9 (791 – 895) | 755.5 ± 12.6 (713 – 845) | <0.001 | <0.001 |
| 7 | Rostrum length to posterior extension of left temporal fossa | 822.3 ± 35 (755 – 858) | 754.9 ± 11.7 (714 – 842.5) | <0.001 | <0.001 |
| 8 | Rostrum length to line across anterior of maxillary crest | 693.6 ± 35.8 (629.5 – 727) | 622.9 ± 11.4 (571 – 693.5) | <0.001 | <0.001 |
| 9 | Rostrum length across maxillary prominence | 558 ± 34.8 (489.5 – 592) | 488.1 ± 9.3 (447 – 532.5) | <0.001 | <0.001 |
| 10 | Rostrum length to line across antorbital notches | 600.8 ± 33.4 (539.5 – 629) | 529 ± 9.1 (488 – 584.5) | <0.001 | <0.001 |
| 11 | Breadth of skull across orbital centres | 289.3 ± 11.8 (266 – 317.5) | 294 ± 5.1 (273 – 318) | NS | NS |
| 12 | Breadth of skull across preorbital process frontals | 281.8 ± 11.1 (258 – 315) | 285.2 ± 4.5 (267 – 311) | NS | NS |
| 13 | Breadth of skull across postorbital process frontals | 302.8 ± 12 (238 – 322) | 312 ± 4.8 (291.5 – 338.5) | NS | NS |
| 14 | Breadth of skull across zygomatic processes squamosals | 300.8 ± 9.3 (284.5 - 318) | 307.5 ± 3.8 (289.5 – 325.5) | NS | NS |
| 15 | Least breadth of skull across posterior margins of temporal fossae | 194.4 ± 14.1 (164.5 – 215) | 199.6 ± 5.5 (182 – 240) | NS | NS |
| 16 | Greatest breadth of skull across exoccipitals | 244 ± 7.1 (266 – 254) | 248.3 ± 2.2 (239 – 257) | NS | NS |
| 17 | Greatest span occipital condyles | 105.8 ± 4.6 (97.9 – 113) | 105.7 ± 2.2 (94 – 114.8) | NS | NS |
| 18 | Greatest width of an occipital condyle | 39 ± 3.2 (34.3 – 45.3) | 39 ± 1.3 (33 – 45.1) | NS | NS |
| 19 | Greatest length of an occipital condyle | 27.2 ± 3.9 (18.8 – 33.6) | 25.5 ± 1.4 (19.7 – 33.4) | NS | NS |
| 20 | Greatest long dimension of right condyle | 69.1 ± 4.6 (60.6 – 75.9) | 70.3 ± 2 (58.6 – 79.1) | NS | NS |

| | Cranial measurement | Female mean ± SE (range) | Male mean ± SE (range) | T-test *p* | Adjusted *p* |
|---|---|---|---|---|---|
| 21 | Greatest breadth foramen magnum | 42.8 ± 3.5 (37 – 50.5) | 43.1 ± 0.8 (39.4 – 47.3) | NS | NS |
| 22 | Greatest length of right nasal on vertex | 47.5 ± 4.8 (39.9 – 56.1) | 49.7 ± 2.2 (39.4 – 61.3) | NS | NS |
| 23 | Length of nasal suture | 31.8 ± 4.5 (23.6 – 43) | 30.6 ± 1.8 (22.6 – 40.4) | NS | NS |
| 24 | Extension right premaxilla posterior to right nasal on vertex | 6.6 ± 4.5 (0 – 43) | 5.5 ± 1.3 (0 – 12.9) | NS | NS |
| 25 | Length premaxilla from right premaxilla to supraoccipital | 62.5 ± 5.9 (28 – 72.3) | 60.8 ± 2.1 (52.4 – 72.5) | NS | NS |
| 26 | Length premaxilla from left premaxilla to supraoccipital | 53 ± 8.6 (36.6 – 67.4) | 49.5 ± 2.6 (40.7 – 66.2) | NS | NS |
| 27 | Greatest breadth of nasals on vertex | 40.1 ± 20.8 (17.2 – 123.4) | 28.6 ± 2.2 (21 – 45.5) | NS | NS |
| 28 | Least distance between anterior prominences on synvertex | 11.7 ± 4.1 (3.9 – 23.2) | 10.6 ± 1 (4.4 – 13.9) | NS | NS |
| 29 | Greatest span premaxillary crests on synvertex | 131.7 ± 7.1 (121.2 – 146.6) | 134 ± 2.6 (125.3 – 154.2) | NS | NS |
| 30 | Greatest span right premaxillary at crest | 76.6 ± 6.2 (67.7 - 93.9) | 78 ± 2.6 (64.2 – 95.2) | NS | NS |
| 31 | Greatest span left premaxillary at crest | 51 ± 5.9 (39 – 62.3) | 54.3 ± 2.2 (46.4 – 70.9) | NS | NS |
| 32 | Transverse width of superior nares | 52.6 ± 4.0 (45.5 - 57.8) | 52.4 ± 1.9 (46.9 - 68.3) | NS | NS |
| 33 | Least width premaxillae where narrows opposite superior nares | 108.3 ± 6.5 (96.1 – 121.9) | 109.1 ± 2.9 (96.9 – 127.6) | NS | NS |
| 34 | Greatest width premaxillae anterior to position of previous | 115.9 ± 6.8 (106.0 - 132.4) | 118.2 ± 3.2 (102.7 - 135.3) | NS | NS |
| 35 | Width right premaxilla on premaxillary plate | 59.7 ± 3.8 (53.9 – 69.3) | 62.4 ± 1.5 (56.3 – 70.3) | NS | NS |
| 36 | Width left premaxilla on premaxillary plate | 47.1 ± 3.6 (40.6 – 53.2) | 48.5 ± 1.6 (43.7 – 58.2) | NS | NS |
| 37 | Rostrum width in apices of antorbital notches | 191.1 ± 13.1 (159 - 213.6) | 193.5 ± 5.0 (160.9 - 216.1) | NS | NS |
| 38 | Rostrum width in apices of prominential notches | 95.1 ± 12.2 (68.2 – 118) | 91.4 ± 3.7 (72.6 – 108.2) | NS | NS |
| 39 | Least distance between main maxillary foramena | 64.1 ± 5.5 (56.8 - 73.0) | 70.4 ± 2.3 (57.7 – 79.3) | 0.029 | NS |
| 40 | Least distance between premaxillary foramena | 32.9 ± 3.1 (28.1 – 37.5) | 32.3 ± 1.5 (25.7 – 42) | NS | NS |
| 41 | Margin left maxillary foramena to anterior maxillary prominence | 44.7 ± 6 (31.1 – 55.5) | 45.6 ± 1.3 (39.2 – 52) | NS | NS |
| 42 | Rostrum width at mid-point (½ distance 10 from posterior) | 37.9 ± 3.3 (31.4 - 43.8) | 38.6 ± 1.4 (31.3 - 44.9) | NS | NS |
| 43 | Width of premaxillae at mid-point of rostrum | 27.5 ± 2.1 (23.1 – 30.8) | 28.4 ± 0.8 (24.3 – 32.5) | NS | NS |
| 44 | Depth of rostrum at mid-point | 35.5 ± 3.3 (31.7 - 39.1) | 41.3 ± 1.3 (34.9 – 48.8) | 0.002 | 0.005 |
| 45 | Height of skull | 277.9 ± 13.3 (255 – 303) | 278.6 ± 5.3 (248.5 – 308) | NS | NS |

| | Cranial measurement | Female mean ± SE (range) | Male mean ± SE (range) | T-test *p* | Adjusted *p* |
|---|---|---|---|---|---|
| 46 | External cranial height | 206.5 ± 13 (186 -241) | 207.6 ± 6.2 (182 – 236) | NS | NS |
| 47 | Greatest length of right temporal fossa | 97.7 ± 7.9 (81.5 – 108.3) | 100.8 ± 3.8 (90.9 – 125) | NS | NS |
| 48 | Greatest length of left temporal fossa | 98.3 ± 8.4 (81.5 107.4) | 100.8 ± 4.1 (85.6 – 128.1) | NS | NS |
| 49 | Width of right temporal fossa | 63.6 ± 7 (54.1 – 75.3) | 67 ± 3.7 (49.8 – 90.7) | NS | NS |
| 50 | Width of left temporal fossa | 60.4 ± 7.5 (50.3 - 71.4) | 64.7 ± 3.6 (46.5 – 84) | NS | NS |
| 51 | Length of right orbit taken from mid-point of frontals | 99.6 ± 5 (90.3 – 108.3) | 98.7 ± 1.7 (92.6 – 109.6) | NS | NS |
| 52 | Length of left orbit taken from mid-point of frontals | 99. 1± 5.6 (88.6 – 109.2) | 99.4 ± 1.5 (92.6 – 108.9) | NS | NS |
| 53 | Rostrum to posterior extension of maxilla between pterygoids | 578.4 ± 32.8 (511 – 607) | 515.8 ± 8.2 (472 – 551) | <0.001 | <0.001 |
| 54 | Rostrum to anterior extension of right pterygoid sinus | 565.2 ± 36.2 (489.5 - 608) | 439.9 ± 8.2 (457 – 536) | <0.001 | <0.001 |
| 55 | Rostrum to anterior extension of left pterygoid sinus | 564.5 ± 36.2 (491 – 557) | 491.8 ± 7.7 (457 – 526) | <0.001 | <0.001 |
| 56 | Rostrum to most anterior extension of pterygoids | 525.3 ± 33.9 (459 – 557) | 455.4 ± 8.2 (414 – 483) | <0.001 | <0.001 |
| 57 | Rostrum to posterior margin of pterygoid midline | 707 ± 38.1 (624 – 745) | 633.8 ± 11.5 (580 – 697.5) | <0.001 | <0.001 |
| 58 | Rostrum to posterior extension of wing of right pterygoid | 729.4 ± 34.4 (653.5 – 759) | 665.2 ± 12.6 (619 – 741) | <0.001 | <0.001 |
| 59 | Rostrum to posterior extension of wing of left pterygoid | 737.1 ± 32.2 (706 – 768) | 659.1 ± 11.2 (604 – 705) | <0.001 | <0.001 |
| 60 | Length of vomer | 66.4 ± 12.2 (37.8 – 87.1) | 59.2 ± 3.4 (42.8 – 72.5) | NS | NS |
| 61 | Width of pterygoid notches | 76.2 ± 3.2 (68.7 – 83.5) | 76.4 ± 1.4 (70 – 83.6) | NS | NS |
| 62 | Mandibular length | 756.2 ± 32.7 (694 – 786) | 692.6 ± 10.5 (638 – 759) | <0.001 | <0.001 |
| 63 | Posterior extension of symphasis to right condyle | 469.2 ± 17.1 (443 – 495) | 460.3 ± 8.1 (423 – 518) | NS | NS |
| 64 | Posterior extension of symphasis to left condyle | 470.4 ± 17.1 (443 – 497) | 459.4 ± 7.2 (424.5 – 508) | NS | NS |
| 65 | Posterior margin or alveolus to right condyle | 475.5 ± 22.7 (445 – 505) | 426.1 ± 6.9 (396 – 474.5) | <0.001 | <0.001 |
| 66 | Posterior margin or alveolus to left condyle | 476 ± 23 (444 – 505) | 425.7 ± 6.7 (396 – 474) | NS | NS |
| 67 | Greatest length of symphasis on right | 301 ± 26.3 (254 – 328) | 244.8 ± 6 (215 – 279) | NS | NS |
| 68 | Greatest length of symphasis on left | 300.6 ± 25.3 (266 – 328) | 244.2 ± 6.2 (213 – 279) | NS | NS |
| 69 | Height of mandible at coronoid processes on right side | 113 ± 3.9 (105.8 – 119) | 111.3 ± 1.8 (103.4 – 121.6) | NS | NS |
| 70 | Height of mandible at coronoid processes on left side | 110.4 ± 3.7 (104.8 -115) | 11.2 ± 1.9 (100.1 – 120.3) | NS | NS |

| | Cranial measurement | Female mean ± SE (range) | Male mean ± SE (range) | T-test *p* | Adjusted *p* |
|---|---|---|---|---|---|
| 71 | Outside height of mandible at mid-point of alveolus on right side | 50.8 ± 4.2 (46.2 – 55.9) | 58 ± 1.6 (48.6 – 64.6) | 0.001 | 0.004 |
| 72 | Outside height of mandible at mid-point of alveolus on left side | 51.0 ± 4.9 (46.0 - 56.8) | 61.3 ± 1.9 (55.0 - 69.4) | <0.001 | <0.001 |
| 73 | Inside height of mandible at mid-point of alveolus on right side | 47.9 ± 4.4 (35.5 – 52.3) | 54.3 ± 1.1 (49 – 61) | 0.002 | 0.006 |
| 74 | Inside height of mandible at mid-point of alveolus on left side | 49.7 ± 3.1 (45 – 53.2) | 54.9 ± 1.2 (49 – 61) | 0.002 | 0.005 |
| 75 | Length of alveolus on right side | 35.9 ± 13.3 (29.7 – 47.7) | 69.5 ± 4.6 (46.2 – 96.1) | <0.001 | <0.001 |
| 76 | Length of alveolus on left side | 36.0 ± 13.8 (29.7 - 46.1) | 72.5 ± 4.7 (51.5 - 93) | <0.001 | <0.001 |
| 77 | Width of alveolus on right side | 10 ± 1.8 (8.4 – 11.8) | 12.7 ± 1 (9 – 18.3) | 0.024 | NS |
| 78 | Width of alveolus on left side | 9.8 ± 1.7 (8.8 - 11) | 13.3 ± 0.9 (9.5 - 17.9) | <0.001 | <0.001 |
| 79 | Tip of mandible to right alveolus | 257.2 ± 23.5 (224.9 – 287.7) | 207.9 ± 5.2 (179.2 – 231.4) | <0.001 | <0.001 |
| 80 | Tip of mandible to left alveolus | 257.6 ± 22.9 (228.7 – 289.8) | 207.8 ± 5.1 (179.2 -231.4) | <0.001 | <0.001 |

Heredity (2015), 1–9

www.nature.com/hdy

ORIGINAL ARTICLE

# Bucking the trend: genetic analysis reveals high diversity, large population size and low differentiation in a deep ocean cetacean

KF Thompson[1,2,4], S Patel[1,2,4], CS Baker[1,3], R Constantine[1] and CD Millar[1,2]

Understanding the genetic structure of a population is essential to its conservation and management. We report the level of genetic diversity and determine the population structure of a cryptic deep ocean cetacean, the Gray's beaked whale (*Mesoplodon grayi*). We analysed 530 bp of mitochondrial control region and 12 microsatellite loci from 94 individuals stranded around New Zealand and Australia. The samples cover a large area of the species distribution (~6000 km) and were collected over a 22-year period. We show high genetic diversity ($h = 0.933–0.987$, $\pi = 0.763–0.996\%$ and $R_s = 4.22–4.37$, $H_e = 0.624–0.675$), and, in contrast to other cetaceans, we found a complete lack of genetic structure in both maternally and biparentally inherited markers. The oceanic habitats around New Zealand are diverse with extremely deep waters, seamounts and submarine canyons that are suitable for Gray's beaked whales and their prey. We propose that the abundance of this rich habitat has promoted genetic homogeneity in this species. Furthermore, it has been suggested that the lack of beaked whale sightings is the result of their low abundance, but this is in contrast to our estimates of female effective population size based on mitochondrial data. In conclusion, the high diversity and lack of genetic structure can be explained by a historically large population size, in combination with no known exploitation, few apparent behavioural barriers and abundant habitat.
*Heredity* advance online publication, 2 December 2015; doi:10.1038/hdy.2015.99

## INTRODUCTION

Population history, demography and behaviour all interact to shape the genetic diversity of a species. Species with fragmented or reduced populations often have low levels of genetic diversity. In contrast, high genetic diversity is consistent with long-term stability in population size, whereas low levels of population differentiation suggest connectivity or recent population expansion. As a rule, cetaceans are known to have low genetic diversity in comparison with terrestrial mammals and this is hypothesized to be due to slow mutation rates (Jackson *et al.*, 2009) and demographic factors such as recent population expansion following bottlenecks or behaviour (Oremus *et al.*, 2009).

Patterns of population structure are evident in most cetacean species, even those with seemingly continuous distributions and high mobility. For example, many baleen whales are highly philopatric, returning to calving or feeding grounds each year, leading to patterns of population structure between these grounds (Alter *et al.*, 2009). The long lifespan of these animals, and extended period of maternal care, allows the cultural transmission of this philopatry over long time periods despite significant population depletion. In killer whales (*Orcinus orca*), population differentiation is thought to result from a highly matrifocal social system (Hoelzel *et al.*, 1998). It has been suggested that the lack of gene flow between such matrifocal groups has, in the longer term, led to the development of sympatric subspecies that are specific to a particular habitat or prey (Morin *et al.*, 2010). Furthermore, some wide-ranging cetacean species show local specialization. For example, common dolphins (*Delphinus delphis*) are a highly mobile pelagic species, yet genetic differentiation has been detected between animals from South Australia and those from the eastern coast of Tasmania, ~1500 km apart (Bilgmann *et al.*, 2008). This differentiation is likely due to a dependence on specific regional oceanographic features such as upwellings that influence prey distributions, for example, the Bonney Upwelling (Butler *et al.*, 2002). Whether these trends are evident in beaked whales has previously been unknown, and difficult to quantify given the problem with obtaining genetic samples.

Although patterns of population structure are known in other cetaceans, beaked whales remain enigmatic, with few published studies describing their populations. The ziphiids, or beaked whales, are one of the most speciose families of cetaceans, second only to the delphinids. Of the 22 species of beaked whale, 15 can be found within the genus *Mesoplodon*. Members of this genus are cryptic in their appearance and behaviour (Pitman, 2009). It has been assumed that their general biology is similar and most are thought to be deep-diving squid eaters that live in small groups along continental shelf edges. However, much of the information on the biology of these whales is derived from stranded animals combined with extrapolation from data collected on the few species that can be observed at sea (e.g., Wimmer and Whitehead, 2004; McSweeney *et al.*, 2007).

[1]School of Biological Sciences, University of Auckland, Auckland, New Zealand; [2]The Allan Wilson Centre, School of Biological Sciences, University of Auckland, Auckland, New Zealand and [3]Department of Fisheries and Wildlife and Marine Mammal Institute, Oregon State University, Newport, OR, USA
[4]These authors contributed equally to this work.
Correspondence: Associate Professor C Millar, School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand.
E-mail: cd.millar@auckland.ac.nz
Received 19 March 2015; revised 2 October 2015; accepted 5 October 2015

Long-term cetacean sighting surveys off the east coast of the United States suggest that beaked whales cluster into ecological niches that are different from all other odontocetes (Schick *et al.*, 2011). For example, Cuvier's (*Ziphius cavirostris*) and Sowerby's (*Mesoplodon bidens*) beaked whales occupy a different area of the continental shelf than do other squid-eating species such as sperm whales (*Physeter macrocephalus*). Moreover, there is evidence that these two beaked whale species may, in turn, occupy slightly different habitats within this area. Tagging data from Blainville's beaked whales (*Mesoplodon densirostris*) have revealed a relationship between foraging behaviour and oceanographic features, for example, water depth, movements of the deep scattering layer of mesopelagic prey and seabed topography (Johnson *et al.*, 2008). Indeed, tagging data suggest that Cuvier's beaked whales perform the deepest dives of all mammals, almost 3000 m (Schorr *et al.*, 2014). Furthermore, modelling implies that beaked whales require larger, higher quality habitats than other cetaceans to meet the energetic needs of such deep diving (Wright *et al.*, 2011; New *et al.*, 2013).

Where surveys have facilitated population size estimates, for example, west coast United States, it has been found that many beaked whale species have undergone significant population declines (Moore and Barlow, 2013). As beaked whales are particularly vulnerable to anthropogenic noise, it is speculated that these population declines are because of an increase in anthropogenic disturbance, further highlighting the need for a better understanding of the basic spatial requirements of all beaked whale species globally (Weilgart, 2007). However, in most areas of the world, and for most beaked whale species, these data are not available. The status of species inhabiting the remote areas of the Southern Ocean and seas around New Zealand is, as yet, undetermined.

New Zealand has the highest recorded number of species of stranded beaked whale in the world; 13 of the 22 recognized species, and some of the most rarely sighted (Thompson *et al.*, 2012; Thompson *et al.*, 2013; Constantine *et al.*, 2014). One of the most frequent species to strand around the coast is the Gray's beaked whale (*Mesoplodon grayi*) (Figure 1). This whale is a medium-sized (4.0–5.5 m) mesoplodont with a circumpolar southern hemisphere distribution (Figure 2a). Distributions have been primarily inferred from analyses of stranding data and live sightings are extremely rare. Most records are generally from south of 33° latitude, particularly on New Zealand, Australian, South African and South American coasts, including the sub-Antarctic and Antarctic waters, with one record from the Netherlands (Boschema, 1950; Dalebout *et al.*, 2004; Taylor *et al.*, 2008; Van Waerebeek *et al.*, 2010; Scheidat *et al.*, 2011). Similar to other beaked whales, Gray's are assumed to live along the continental shelf edge, although there are occasional sightings of animals in shallow waters (e.g., Dalebout *et al.*, 2004). An analysis of stranding patterns around New Zealand suggests that summer peaks are associated with inshore movements related to calving or nursing, particularly around the North Island (Thompson *et al.*, 2013). Moreover, Gray's beaked whales have subtle morphological differences between the east and west coasts of New Zealand (Thompson *et al.*, 2014) and this might indicate restricted gene flow between the two coasts.

**Figure 1** A stranded male Gray's beaked whale (*M. grayi*) at Pataua beach, in the North East region of New Zealand in December 2009.

Behavioural studies of other beaked whales, for example, the northern bottlenose whale (*Hyperoodon ampullatus*), suggest specific dependencies on oceanographic features, such as submarine canyons, have resulted in genetically isolated populations (Dalebout *et al.*, 2006). To date, there is no information on the foraging habitat or prey of Gray's beaked whales, although MacLeod *et al.* (2003) have speculated that this species relies more on small benthic fish than other beaked whales. The seabed topography around New Zealand is diverse, supporting a variety of mesopelagic squid and fish (De Leo *et al.*, 2010). Around the continental shelf edge, there are several areas of periodic high primary productivity resulting from upwelling of slope-associated deep water (MacDiarmid *et al.*, 2013). Many marine mammals are known to take advantage of these upwellings for foraging (Torres, 2013; Sagnol *et al.*, 2014). In the case of Gray's beaked whales, it is unclear whether the species take advantage of these upwellings, although stranding patterns appear to indicate use of the highly productive areas of the continental shelf of the north east of the North Island of New Zealand, particularly in summer (Thompson *et al.*, 2013). We hypothesize that, given the spatial scale from the east to west coast of New Zealand, we would expect genetic structure in Gray's beaked whales in line with morphological differences, with habitat dependency, as a result of specialization to local prey or breeding areas, as the driver of differentiation. Furthermore, we suggest that such genetic divergence is likely to be greater over a larger spatial scale (6000 km) between the Chatham Islands and Western Australia.

To test this hypothesis, we analysed sequence data from 530 bp of mitochondrial control region and 12 variable microsatellite loci to investigate diversity, population structure and effective population size from 94 Gray's beaked whale samples. These samples were collected from strandings around the coast of New Zealand, with an additional six samples from Western Australia representing the largest global collection of this species. We provide novel insights into the population dynamics of this enigmatic, and rarely sighted, species.

## MATERIALS AND METHODS

### Study area and sample collection

We collected samples from a region spanning more than 6000 km extending from the west coast of Australia to the Chatham Islands (New Zealand) in the east (Figure 2). Samples from New Zealand were obtained from the New Zealand Cetacean Tissue Archive and cover the period from 1991 to 2013. Samples from Australia were collected over a 3-year period and were obtained from the Western Australian Museum (for specimen details see Supplementary Tables S1 and S2 and Supplementary Information). Further samples from South Australia and Tasmania were not available in sufficient numbers to allow any meaningful analyses. Sex of all samples was determined by amplification of the SRY gene multiplexed with a ZFX/ZFY-positive control (Aasen and Medrano, 1990; Gilson *et al.*, 1998; Thompson *et al.*, 2012). Samples from New Zealand were divided into four *a priori* regional areas according to where

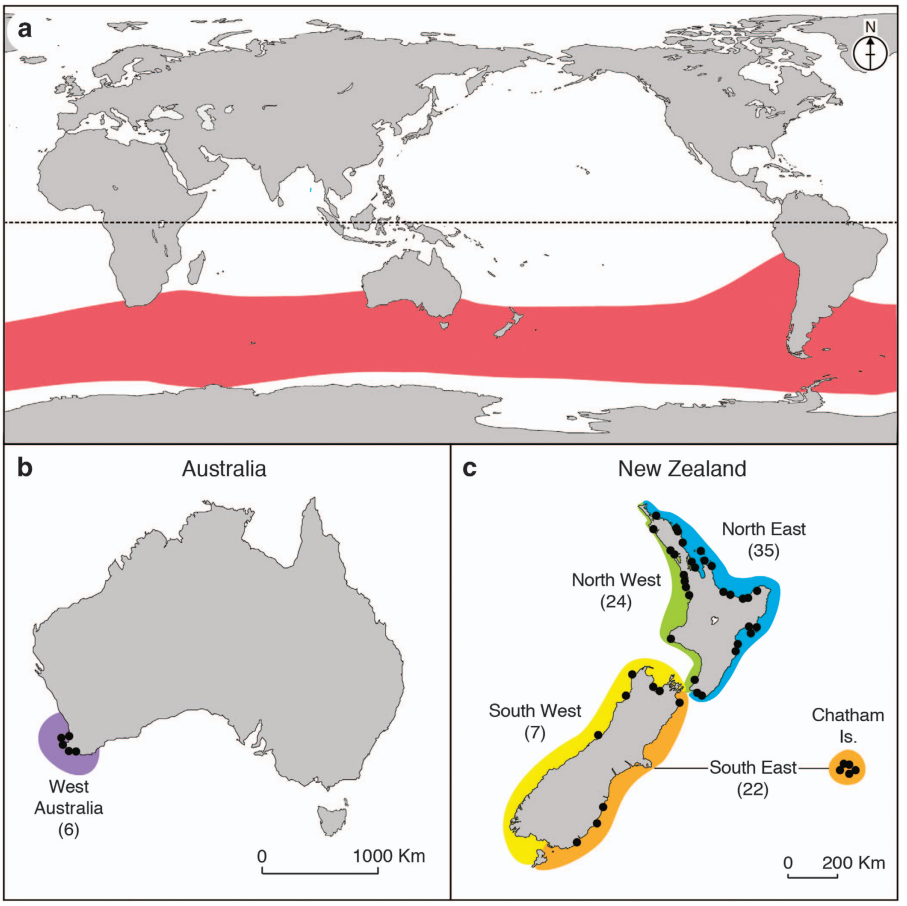



**Figure 2** (**a**) The likely global distribution of Gray's beaked whales (*M. grayi*) based on both sightings and stranding records (follows International Union for Conservation of Nature listing, www.redlist.org). (**b** and **c**) Location of Gray's beaked whale samples and *a priori* geographic regions. Stranding locations for samples are shown by black circles and sample numbers are given within parentheses. *A priori* regions are shown by colour (West Australia (purple), in New Zealand, North East (blue), North West (green), South East (orange) and South West (yellow)). For details of actual stranding locations, see Supplementary Information.

on the coast the animal was found and the Australian samples provided a fifth regional grouping (Figures 2b and c). These areas were based on the location of known marine biogeographic barriers resulting from seabed topography and oceanographic currents (Ayers and Waters, 2005).

## DNA extraction, sequencing and genotyping

Genomic DNA was isolated from tissue using proteinase K digestion followed by a standard 25:24:1 Phenol:Chloroform:Isoamyl protocol as described by Sambrook *et al.* (1989) and modified by Baker *et al.* (1994), followed by ethanol precipitation. A 530 bp fragment of DNA from the mitochondrial control region was amplified and sequenced in both directions according to methods described in Thompson *et al.* (2013) using the primer pair Dlp1.5 and Dlp8G. Sequences were trimmed by eye in the programme GENEIOUS v7.1 (www.geneious.com), and only those sequences that reached a PHRED score of 40 or above for at least 70% of individual bases were deemed acceptable for analyses (Kearse *et al.*, 2012). The first base of the control region was designated to be position 15 468 in reference to the Gray's beaked whale whole mitogenome (GenBank accession no. KF981442). Sequences were aligned using MAFFT (Multiple Alignment using Fast Fourier Transform) multiple sequence alignment tool (Katoh *et al.*, 2002).

Genotype data from 12 microsatellite loci (one di-, three tri- and eight tetra-repeats) were obtained using primers and methods developed by Patel *et al.* (2014). MICROCHECKER (Van Oosterhout *et al.*, 2004) was used to assess evidence of scoring error due to stuttering, large allele dropout and null alleles. Deviation from Hardy–Weinberg equilibrium and linkage disequilibrium between loci were tested in ARLEQUIN 3.5 (Excoffier and Lischer, 2010).

## Genetic diversity

For mitochondrial DNA (mtDNA) data, haplotypic ($h$) and nucleotide ($\pi$) diversities were calculated using ARLEQUIN. The nucleotide substitution model used to calculate genetic distance was the Tamura and Nei model with a gamma correction of $\alpha = 0.219$ as determined in jModelTest using the corrected Akaike information criterion (Tamura and Nei, 1993; Posada, 2008). For genotype data, average allelic richness ($R_s$), observed and expected heterozygosities were calculated per microsatellite locus and per putative population using GENODIVE 2.3b23 (Meirmans and Van Tienderen, 2004). Measures of genetic diversity can be highly dependent on sample size, in that larger populations are likely to have more alleles than smaller populations; therefore, allelic richness values were also calculated per population using the rarefaction method implemented in HP-RARE 1.0 (Kalinowski, 2005). This method statistically adjusts for sample size by calculating the number of alleles as a function of the sample size per population.

To identify any genetic signature of demographic expansion or population bottleneck in the mtDNA, Fu's $F_s$ statistic was calculated, as implemented in ARLEQUIN. Fu's $F_s$ is one of the more sensitive indicators of deviation from neutral population equilibrium (Ramos-Onsins and Rozas, 2002). Departure from neutral expectation was inferred by randomization using a coalescent algorithm run for 10 000 steps (Hudson, 1990). Negative values of Fu's $F_s$

statistic are indicative of historical population expansion or genetic hitchhiking, and a positive value is evidence of a recent population bottleneck and a deficiency of alleles at this locus (Fu, 1997).

### Population structure

To visualize the geographic distribution of mtDNA haplotypes and their relationships, the program POPART was used to construct a median joining network (University of Otago, Dunedin, New Zealand; http://popart.otago.ac.nz). We estimated a phylogenetic tree of samples using a variant of Bayesian inference (Mr Bayes) with two Markov chain Monte Carlo sampler runs of $1.1\times10^6$ generations and the nucleotide substitution model as determined by jModelTest (Huelsenbeck and Ronquist, 2001). Blainville's beaked whale (*M. densirostris*) and Gervais' beaked whale (*Mesoplodon europeaus*) were selected as outgroups. Trees were sampled every 200 generations and it was determined by visual inspection of posterior traces of both runs that stationarity was attained by $1.1\times10^5$ generations. The first $1.1\times10^5$ generations were discarded as burn-in leaving the remaining samples to estimate a consensus tree and posterior probabilities.

An analysis of molecular variance and pairwise F-statistics were calculated for mtDNA and microsatellites in ARLEQUIN and GENODIVE, respectively. Three F-statistics were used: standard $F_{st}$ based on mtDNA haplotype and microsatellite allele frequencies; $\Phi_{st}$ that incorporates molecular sequence divergence in mtDNA and, $F'_{st}$ that is a standardised $F_{st}$ statistic that takes into account within-population genetic variation for microsatellites (Meirmans and Hedrick, 2011). Pairwise exact tests were also carried out and significance of all F-statistics was tested using 10 000 permutations. Analysis of molecular variance and F-statistics were analysed by each sex separately (data not shown) and for the combined data set.

Population structure was also investigated using a Bayesian clustering analysis to estimate the most probable number of populations using STRUCTURE 2.3.4 (Pritchard *et al.*, 2000). Analysis of microsatellite data was conducted with and without sampling location priors using the admixture model. The number of clusters ($K$) with the highest posterior probability was identified using replicate runs assuming $K$ from 1 to 5. The burn-in length was set at 100 000 steps, followed by 1 000 000 steps with a total of 10 replicates for each value of $K$. The most likely number of homogeneous clusters was assessed using the second-order rate of change or $\Delta K$ method following Evanno *et al.* (2005) and implemented in STRUCTURE HARVESTER (Earl and vonHoldt, 2012). Results were then combined in the program CLUMPP to average individual clustering outputs between runs (Jakobsson and Rosenberg, 2007) and visualized in DISTRUCT (Rosenberg, 2004). A principal component analysis was used to further visualize differences in genotypic variation between populations and individuals as implemented in R using the package ADE-GENET (Jombart, 2008).

### Estimating effective population size

Mitochondrial control region sequences were used to estimate effective female population size ($N_{ef}$) using a Bayesian skyline plot approach implemented in BEAST v.1.8.0 (Drummond *et al.*, 2012). The substitution model (TN93) with discrete gamma distribution with four rate categories was selected as the model of evolution having been previously determined in jModelTest. A strict molecular clock approach was used that assumed a control region mutation rate of $0.9\times10^{-8}$ bp per year (Cuvier's beaked whales; Dalebout *et al.*, 2005), and $2\times10^{-7}$ bp per year derived from ancient DNA sampling (bowhead whales (*Balaena mysticetus*); Ho *et al.*, 2007, 2011). The Markov chain Monte Carlo chains were run with $3\times10^7$ iterations and samples were drawn every 30 000 iterations with the first 10% being discarded as burn-in. Population history was inferred using the Bayesian skyline plot with 10 groups of coalescent intervals. Two independent BEAST analyses were combined and in all cases convergence to stationary distribution and sufficient sampling were visually checked in TRACER v.1.6 (Rambaut *et al.*, 2013).

## RESULTS

### Genetic diversity

A total of 94 individuals were sequenced resulting in 38 mitochondrial haplotypes defined by 26 variable sites (Supplementary Table S3 and Supplementary Information; GenBank accession numbers: KJ767593–KJ767630). Diversity statistics suggest that Gray's beaked whales have moderately high levels of variation within the study areas and both haplotype ($h$) and nucleotide ($\pi$) diversity were found to be similar between regions (Table 1). The same 94 individuals were genotyped at 12 microsatellite loci. The average number of alleles ($k$), allelic richness and private allelic richness were similar among regions. The only exception being $k$, which was lower in both the south west of New Zealand and Western Australia where there were fewer samples than in the other areas (Table 1). No microsatellite loci deviated from Hardy–Weinberg equilibrium and there was no significant linkage disequilibrium between loci after Bonferroni correction (Supplementary Table S4 and Supplementary Information). Loci showed no evidence for null alleles, large allelic dropout or scoring errors due to stutter peaks. The average amount of missing allelic data per locus was 0.35%. Diversity statistics per loci are given in Supplementary Table S5.

Fu's $F_s$ value was negative and highly significant ($-23.01$, $P<0.001$) and indicative of historical demographic expansion or selective sweep, and an excess of rare substitutions and haplotypes at this locus. These results suggest that it is unlikely that Gray's beaked whales have suffered any historical genetic bottleneck.

### Population structure

The median joining network of haplotypes showed no phylogeographic structure and common haplotypes were shared across the study area (Supplementary Figure S1). Moreover, the network is highly reticulated and most haplotypes are sister lineages in that they

**Table 1 mtDNA control region and microsatellite diversity statistics[a]**

| *Region* | N | *Female* | *Male* | *mtDNA* | | | *Microsatellites* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | *Number of haplotypes* | h | *π (%)* | k | *Allelic richness* | *Private allelic richness* | $H_o$ | $H_e$ |
| North East New Zealand | 35 | 21 | 14 | 21 | 0.949±0.019 | 0.996±0.547 | 7.750 | 4.31 | 0.44 | 0.624 | 0.664 |
| North West New Zealand | 24 | 20 | 4 | 14 | 0.949±0.023 | 0.827±0.471 | 6.667 | 4.35 | 0.38 | 0.659 | 0.675 |
| South East New Zealand | 22 | 6 | 16 | 18 | 0.987±0.017 | 0.962±0.541 | 6.667 | 4.37 | 0.31 | 0.659 | 0.659 |
| South West New Zealand | 7 | 1 | 6 | 6 | 0.952±0.095 | 0.763±0.496 | 4.417 | 4.22 | 0.3 | 0.702 | 0.668 |
| West Australia | 6 | 5 | 1 | 5 | 0.933±0.122 | 0.931±0.609 | 4.333 | 4.33 | 0.35 | 0.597 | 0.624 |
| All regions | 94 | 41 | 53 | 38 | 0.963±0.007 | 0.871±0.478 | 9.500 | 4.32 | — | 0.648 | 0.659 |

Abbreviations: $h$, haplotype diversity; $H_o$, observed heterozygosity; $H_e$, expected heterozygosity; $k$, average number of alleles; mtDNA, mitochondrial DNA; $\pi$, nucleotide diversity; $R_s$, allelic richness. $k$ is not corrected for sample size differences; therefore, $R_s$ and private allelic richness were also calculated using the rarefaction method implemented in HP-RARE.
[a]For Gray's beaked whales (*Mesoplodon grayi*) sampled within each *a priori* region and overall.

**Table 2 Summary statistics of pairwise comparisons assessing population structure within five *a priori* regions[a]**

| *Region* | *North East New Zealand* | *North West New Zealand* | *South East New Zealand* | *South West New Zealand* | *West Australia* |
|---|---|---|---|---|---|
| North East New Zealand | | 0.027 | 0.01 | 0.013 | 0.012 |
| | | −0.009 | −0.016 | 0.011 | −0.035 |
| North West New Zealand | −0.006 | | −0.022 | −0.011 | −0.005 |
| | −0.021 | | −0.011 | 0.006 | −0.013 |
| South East New Zealand | −0.008 | −0.005 | | −0.018 | −0.042 |
| | −0.025 | −0.017 | | −0.046 | −0.014 |
| South West New Zealand | 0.004 | 0.006 | 0.001 | | 0.014 |
| | 0.012 | 0.012 | 0.001 | | −0.081 |
| West Australia | −0.006 | −0.002 | 0.000 | 0.023 | |
| | −0.017 | −0.015 | −0.004 | 0.065 | |

[a]Mitochondrial DNA above the diagonal (in individual cells $F_{st}$ top, $\Phi_{st}$ below), microsatellite data below the diagonal (in individual cells $F_{st}$ top, $F'_{st}$ below). Note that no *P*-values were significant at $P<0.05$.

differ by only a single substitution. This pattern is also reflected in the Bayesian tree (Supplementary Figure S2).

Pairwise comparisons between populations showed no significant differentiation in either mtDNA or microsatellites at the $P<0.05$ level (Table 2). This lack of significance held true whether the sexes were combined or separated (data not shown). Bayesian clustering analyses implemented in STRUCTURE showed no population structure for microsatellite data. The highest average posterior probability occurred at $K=1$ and graphical outputs from DISTRUCT showed that with increasing values of $K$, all populations became increasingly subdivided into multiple clusters approximately proportional to the sample size for *a priori* regions (Figure 3). Both STRUCTURE analyses, with and without priors, revealed the same results. Principal component analysis showed all populations overlapping in genotypes with no visible differentiation between any of the *a priori* regions (Supplementary Figure S3).

### Estimate of effective population size

Using the mutation rate derived from Cuvier's beaked whale, the product of female $N_{ef}$ and generation time was calculated to be 10.14 million with 95% credibility intervals of 0.39–51.79 million. Using the faster mutation rate from bowhead whales, the product of female $N_{ef}$ and generation time was calculated to 0.46 million with 95% credibility intervals of 0.02–2.25 million (Figure 4). Estimation of $N_{ef}$ from microsatellite data using programs such as NeEstimator have not been reported, as these methods produced unreliable estimates and are known to be inappropriate for estimating the size of large populations, particularly with high levels of gene flow.

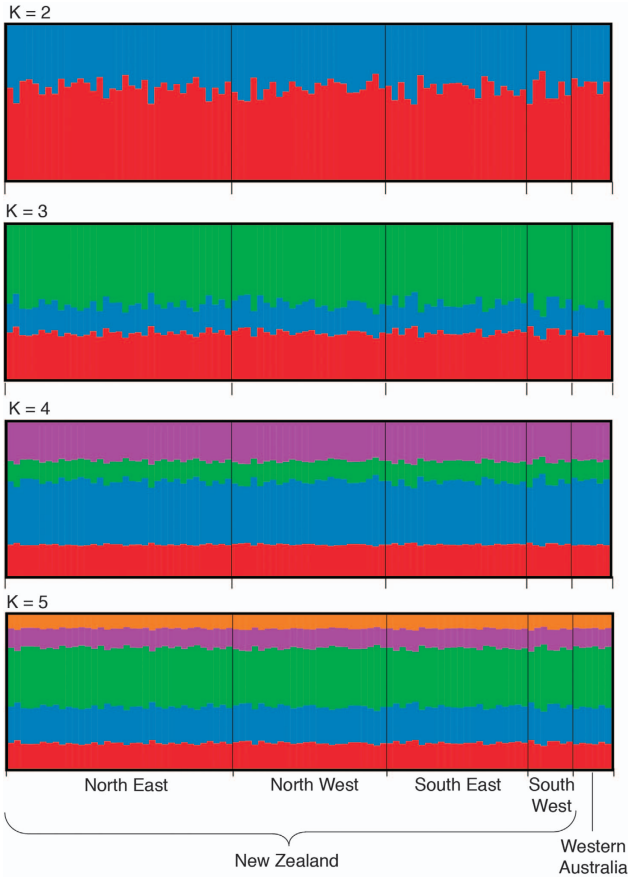



**Figure 3** Bayesian STRUCTURE analysis of 12 Gray's beaked whale (*M. grayi*) microsatellite loci from five *a priori* regions. Each bar represents the likelihood of an individual's assignment to a particular population cluster as indicated by the colours for $K=2$–5.

## DISCUSSION

We analysed samples collected from around the coast of New Zealand and Western Australia in the largest study on beaked whale population genetics to date. Our findings show that Gray's beaked whales have high mitochondrial haplotype and nucleotide diversity relative to most beaked whales (Gray's (530 bp): $h=0.93$–$0.94$, $\pi=0.76$–$0.99\%$; Blainville's (362 bp): $h=0.87\pm0.07$; $\pi=0.49\pm0.35\%$; northern bottlenose (434 bp): $h=0.57$; $\pi=0.15\%$), with the exception of the southern bottlenose whale (*Hyperoodon planifrons*) (238 bp): $h=0.97$; $\pi=3.73\%$ (Dalebout, 2002; Dalebout *et al*., 2001, 2005) (Table 3). Southern bottlenose whales also have a distribution extending throughout the Southern Ocean and have never been a target of whaling. The level of diversity observed in both these beaked whales contrasts with that found in pilot whales (*Globicephala melas*) and false killer whales (*Pseudorca crassidens*), where social factors are thought to contribute to low diversity (Whitehead, 1998). Spinner dolphins (*Stenella longirostris longirostris*) in the waters of French Polynesia are also pelagic, with a distribution concentrated around particular island groups and significant gene flow between these areas (Oremus *et al*., 2007). Our diversity statistics are comparable to this species and

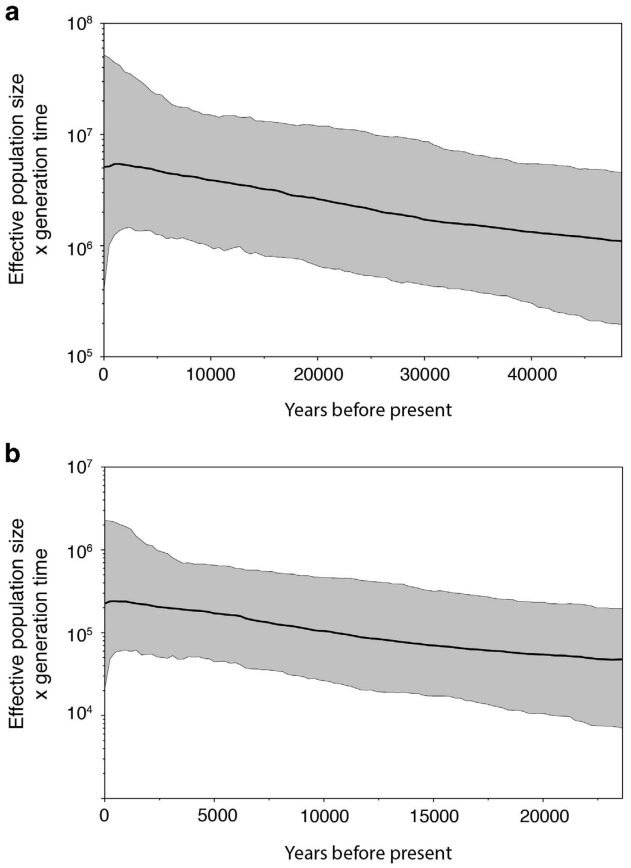



**Figure 4** Bayesian skyline plots showing temporal changes in genetic diversity in Gray's beaked whales (*M. grayi*) estimated from mitochondrial control region sequences. (**a**) Using the mutation rate from Dalebout *et al.* (2005). (**b**) Using the mutation rate from Ho *et al.* (2007, 2011). The x axis is in calendar years; the y axis is the product of effective population size and generation time ($N_{ef}\tau$). Grey shading indicates the 95% credibility intervals.

it is likely that Gray's beaked whales show similar levels of movement and gene flow.

Our estimation of Fu's $F_s$ is negative and highly significant, indicating a population expansion or a selective sweep. A rapid radiation of Gray's beaked whales during their divergence from the most recent common ancestor could potentially explain these signatures. However, the phylogeny of the ziphiids is currently in question as new and more informative genomic markers enable its revision (Morin *et al.*, 2013). Gray's beaked whales are unlikely to have undergone any recent genetic population bottleneck, although our data reflect long-term historical demographic patterns, and cannot determine more recent population changes. This species has no documented history of human consumption in this region and, therefore, these results are perhaps unsurprising (Robards and Reeves, 2011). However, there is the potential that a pelagic species such as Gray's beaked whales is impacted upon by fisheries by-catch and, given the difficulties in carcass recovery and species identification, assessment data is currently unavailable (Madsen *et al.*, 2014). To detect more recent population changes both census data and alternative genetic markers would be required, and current by-catch rates would be helpful in assessing potential human-induced mortality.

Our study is limited by small sample size from Australia, and therefore our results comparing Gray's beaked whale population structure across to New Zealand are preliminary. However, in general, our analyses of data from both mtDNA and microsatellite markers indicate a lack of genetic structure across the ~6000 km-wide study area. None of the pairwise comparisons of genetic differentiation based on $F_{st}$ were significant at the $P<0.05$ level, and therefore our results are consistent with a single Gray's beaked whale population. However, further samples from Australia are needed to confirm these findings.

Overall, this result contrasts with our original hypothesis predicting restricted gene flow between east and west coasts of New Zealand. Studies of population structure in beaked whales are inherently difficult because of the paucity of material available for genetic analysis; however, in northern bottlenose whales significant genetic structure was detected across a distance of ~2000 km between The Gully, off Nova Scotia, and the Labrador Sea (Dalebout *et al.*, 2006). This structure is thought to result from a combination of habitat specificity, that is, the need to associate with submarine canyons, and a genetic bottleneck because of hunting (Dalebout *et al.*, 2001). Cuvier's and Blainville's beaked whales are both cosmopolitan species that are broadly distributed throughout the world's oceans. These species show clear differentiation between ocean basins, with little contemporary interoceanic gene flow (Dalebout *et al.*, 2005; Morin *et al.*, 2013). This differentiation is thought to reflect patterns of long-term divergence as a result of the species' radiation, habitat preferences and/or social organization. Such genetic structure is not unusual for marine organisms with either site fidelity to breeding grounds, for example, white sharks (*Carcharodon carcharias*) (Bonfil *et al.*, 2005), or feeding grounds, for example, herring (*Clupea harengus*) (Gaggiotti *et al.*, 2009).

In contrast, given the results of our study, interoceanic gene flow is highly likely in the case of Gray's beaked whales, particularly as there are no large continents that restrict movement throughout their distribution. This pattern is the first described in the genus *Mesoplodon,* and while our samples cover approximately one-third of the species' range, further samples are needed from South Africa, South America and the Southern Ocean to confirm this finding. There are both fish and squid species that exhibit similar levels of connectivity across comparable spatial scales (e.g., orange roughy (*Hoplostethus atlanticus*), Varela *et al.*, 2012; giant squid (*Architeuthis* spp.), Winkelmann *et al.*, 2013) and it is likely that there are aspects of these species' population biology that are common.

Thompson *et al.* (2013) suggests that, given stranding patterns, seasonal shifts in distribution associated with the calving season are likely in Gray's beaked whales, perhaps in relation to a dependency on inshore waters. However, should these preferences exist they are clearly not driving long-term genetic differentiation or there is enough habitat of sufficient quality within the study area to support multiple calving grounds. Interestingly, the morphological differences seen in Gray's beaked whales between the east and west coasts of New Zealand are not reflected in the genetic data (Thompson *et al.*, 2014). This suggests that such morphological differences occur in the presence of gene flow and could perhaps result from phenotypic plasticity and/or dietary preferences. There are several examples of such phenotypic plasticity in cetaceans, for example, bottlenose dolphins (*Tursiops truncatus*) (Viaud-Martinez *et al.*, 2008) and killer whales (Foote *et al.*, 2009). These examples are thought to indicate ecological differences and are accompanied by an associated genetic divergence, but this is not the case in our study.

Given the degree of genetic homogeneity found between all regions, these results suggest that it is unlikely that the whales found off the coast of Western Australia are distinct from those found around the

**Table 3 Levels of mitochondrial DNA control region haplotype diversity and nucleotide diversity[a]**

| *Species* | *Sequence length (bp)* | *Sample size* | *Sampling location* | *Haplotype diversity (h)* | *Nucleotide diversity π (%)* | *Source* |
|---|---|---|---|---|---|---|
| *Beaked whale species* | | | | | | |
| Gray's (*Mesoplodon grayi*) | 530 | 94 | NZ/AUS | 0.933–0.987 | 0.763–0.996 | This study |
| Straptoothed (*Mesoplodon layardii*) | 361 | 22 | AUS/SA | 0.87±0.07 | 0.58±0.37 | Dalebout (2002) |
| Cuvier's (*Ziphius cavirostris*) | 290 | 87 | Global | 0.926±0.0154 | 1.27±0.723 | Dalebout *et al*. (2005) |
| Blainville's (*Mesoplodon densirostris*) | 362 | 11 | Global | 0.87±0.07 | 0.49±0.35 | Dalebout (2002) |
| Arnoux's (*Berardius arnuxii*) | 434 | 45 | North Atlantic | 0.73±0.15 | 0.20±0.19 | Dalebout (2002) |
| Baird's (*Berardius bairdii*) | 370 | 43 | North Pacific | 0.52±0.09 | 0.29±0.22 | Dalebout (2002) |
| Northern bottlenose (*Hyperoodon ampullatus*) | 434 | 45 | North Atlantic | 0.57±0.07 | 0.15±0.13 | Dalebout (2002) |
| Southern bottlenose (*Hyperoodon planifrons*) | 238 | 9 | Southern Ocean | 0.97±0.06 | 3.73±2.16 | Dalebout (2002) |
| *Other odontocete species* | | | | | | |
| Commerson's dolphin (*Cephalorhynchus commersonii*) | 466 | 196 | South America | 0.807 | 0.40 | Pimper *et al*. (2010) |
| Long-finned pilot whales (*Globicephala melas*) | 358 | 620 | NZ | 0.22±0.03 | 0.09±0.11 | Oremus *et al*. (2009) |
| False killer whale (*Pseudorca crassidens*) | 945 | 62 | Hawaii | 0.34±0.07 | 0.09±0.07 | Chivers *et al*. (2007) |
| Spinner dolphin (*Stenella longirostris longirostris*) | 555 | 70 | Moorea, FP | 0.93±0.01 | 1.62±0.84 | Oremus *et al*. (2007) |
| *Balaenoptera species:* | | | | | | |
| Antarctic blue whale (*Balaenoptera musculus*) | 410 | 184 | Southern Ocean | 0.97±0.01 | 1.4±0.70 | Sremba *et al*. (2012) |
| *Other mammals* | | | | | | |
| Gray seal (*Halichoerus grypus*) | 489 | 34 | Sweden | 0.97±0.01 | 1.62±0.9 | Graves *et al*. (2008) |

Abbreviations: AUS, Australia; FP, French Polynesia; NZ, New Zealand; SA, South Africa.
[a]Reported in beaked whales and other mammalian species.

coast of New Zealand. We speculate that Gray's beaked whales may move freely between these areas perhaps following the subtropical convergence, the boundary between cold sub-Antarctic and warmer subtropical waters, that dominates the centre of this species distribution (Garner, 1959; Heath, 1981). A number of marine mammal species are known to take advantage of this convergence, which is associated with areas of high primary productivity. Sightings surveys off the coast of south Australia have detected several beaked whales (Gill *et al*., 2015) with one particular sighting involving a single group of 20 unidentified mesoplodonts largely fitting the description of Gray's beaked whales (P. Gill, pers. comm.). It is highly possible that such an oceanographic feature, which can be as productive as the Benguela Upwelling (van Ruth *et al*., 2010), may facilitate movements of Gray's beaked whales and act as a gene flow 'conveyer belt' between New Zealand and Australia.

The panmictic pattern in Gray's beaked whales may also be a result of social factors that promote gene flow, as has been suggested in common dolphins in the North Atlantic. Gray's are unique among the beaked whales in that they commonly strand in groups. The holotype specimen was one of 28 animals stranded in the Chatham Islands in 1875 (von Haast, 1876), and other large strandings (4–6 animals) occur frequently around New Zealand (New Zealand Department of Conservation, unpublished data). It has been proposed that these larger strandings are breeding aggregations as in some cases they include multiple adult males, although further behavioural evidence would be required to confirm this (Dalebout, 2002). Whether these larger groups are formed for the purpose of mating is unknown but it is possible that Gray's beaked whales have a mating system that is distinct from other ziphiids and more akin to what has been described in the delphinids. The high levels of genetic diversity and a lack of differentiation across the geographical range of our study may imply a promiscuous and/or polygynous mating system that promotes gene flow.

There are many limitations to estimates of female effective population size; several assumptions of the coalescent model are violated because of the lack of basic knowledge of this species biology. However, based on mitochondrial data, our analyses imply that Gray's beaked whales have existed as an increasing population with no historical population bottleneck.

Given a plausible generation time for Gray's beaked whales of 15 years, as is estimated for Cuvier's beaked whales (Dalebout *et al*., 2005) and spinner dolphins (Oremus *et al*., 2007), our estimate of mean female effective population size ranges from 676 000 (26 000–3.45 million, 95% CI) to a lower estimate of 30 600 (1333–150 000, 95% CI). In all estimates, our credibility intervals highlight the high degree of uncertainty, and upper limits of female effective population size of whales do not generally reach into the millions, for example, Cuvier's beaked whales in the Southern Ocean have an upper limit of 189 000 (Dalebout *et al*., 2005) and the minke whale (*Balaenoptera acutorostrata*) upper limit is 800 000 (Alter and Palumbi, 2009). Although we have applied a coalescent approach, which tends to be more accurate than deterministic methods, there are still considerable limitations. Effective population size estimates are strongly influenced by the mutation rate, with underestimation of rates resulting in large overestimates in population size (Luikart *et al*., 2010). In general, $N_{ef}$ is most difficult to estimate in large populations with moderate gene flow and this difficulty can lead to extremely large confidence intervals as are seen in our estimates (Hare *et al*., 2011). In this context, we suggest that our estimates of $N_{ef}$ should be considered as indicative of a large population with no bottleneck. This contradicts the basic assumption that, in general, beaked whales exist at naturally low abundances and, hence, are rarely observed at sea (Pitman, 2009). In the case of Gray's beaked whales, the rarity in sightings is more likely due to their offshore distribution and cryptic behaviour, together with a paucity of dedicated oceanic surveys.

In conclusion, our results suggest that Gray's beaked whales form a large panmictic population. It is most likely that significant genetic connectivity exists between the waters of New Zealand and Western Australia. Although there are limitations in our sampling, and consequently our analyses, our inability to detect genetic heterogeneity throughout the study area suggests that there is an absence of any particular habitat dependencies, social factors or historical population depletion that have restricted gene flow. We suggest, given the strength of our findings, that Gray's beaked whales in New Zealand and Australian waters be managed as a single management unit. Our study highlights the value of long-term stranding collections in studying populations of elusive, long-lived, slow-breeding species. With more extensive sampling, and higher resolution genetic markers (e.g., single-nucleotide polymorphisms), we suggest that future research that helps to elucidate any cryptic population structure in Gray's beaked whales would be a valuable contribution to the study of this species.

## DATA ARCHIVING

Reference DNA sequences are available under GenBank accession nos. KJ767593–KJ767630. The genotype–haplotype assignments and information about sample location are available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.f47f6.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGEMENTS

We thank the staff of the New Zealand Department of Conservation for tissue collection and database management, particularly L Boren and L Wakelin; local iwi and hapu, and Massey University necropsy teams. Samples from west Australia were supplied by: R O'Shea and J Bannister, Western Australian Museum; and N Gales, Australian Antarctic Division. We thank M Dalebout and D Steel for sample collection and archiving; A Stuckey, A Veale and R Bouckhaert for assistance with analyses; the Centre for Genomics, Proteomics and Metabolomics at the University of Auckland, and New Zealand Genomics Ltd. We also thank three anonymous reviewers for their useful comments. This project was funded by a University of Auckland Faculty Research Development Fund Grant 3702180 (CDM and RC). Samples are curated in the New Zealand Cetacean Tissue Archive at the University of Auckland under Department of Conservation Permit Rnw/HO/2009/03 and CITES institutional permit NZ010.

Aasen E, Medrano JF (1990). Amplification of the ZFX and ZFY genes for sex identification in humans, cattle, sheep and goats. *Nat Biotech* **8**: 1279–1281.

Alter SE, Palumbi SR (2009). Comparing evolutionary patterns and variability in the mitochondrial control region and cytochrome *b* in three species of baleen whales. *J Mol Evol* **68**: 97–111.

Alter SE, Ramirez SF, Nigenda S, Ramirez JU, Bracho LR, Palumbi SR (2009). Mitochondrial and nuclear genetic variation across calving lagoons in eastern North Pacific gray whales (*Eschrichtius robustus*). *J Hered* **100**: 34–46.

Ayers KL, Waters JM (2005). Marine biogeographic disjunction in central New Zealand. *Mar Biol* **147**: 1045–1052.

Baker CS, Slade RW, Bannister RB, Abernethy RB, Weinrich MT, Lien J *et al.* (1994). Hierarchical structure of mitochondrial DNA gene flow among humpback whales, *Megaptera novaeangliae,* world-wide. *Mol Ecol* **3**: 313–327.

Bilgmann K, Möller LM, Harcourt RG, Gales R, Beheregaray LB (2008). Common dolphins subject to fisheries impacts in Southern Australia are genetically differentiated: implications for conservation. *Anim Conserv* **11**: 518–528.

Bonfil R, Meÿer M, Scholl MC, Johnson R, O'Brien S, Oosthuizen H *et al.* (2005). Transoceanic migration, spatial dynamics, and population linkages of white sharks. *Science* **310**: 100–103.

Boschema H (1950). Maxillary teeth in specimens of *Hyperoodon rostratus* (Muller) and *Mesoplodon grayi* Von Haast stranded on the Dutch coasts. *Koninklijke Nederlandse Akademie van Wetenschappen* **53**: 775–786.

Butler A, Althaus F, Furlani D, Ridgway K (2002). *Assessment of the conservation values of the Bonney Upwelling. A component of the Commonwealth Marine Conservation Assessment Program 2002–2004. CSIRO report to Environment Australia.* Hobart, Tasmania, Australia: CSIRO Marine and Atmospheric Research.

Chivers SJ, Baird RW, McSweeney DJ, Webster DL, Hedrick NM, Salinas JC (2007). Genetic variation and evidence for population structure in eastern North Pacific false killer whales (*Pseudorca crassidens*). *Can J Zool* **85**: 783–794.

Constantine R, Carroll E, Stewart R, Neale D, van Helden A (2014). First record of True's beaked whale *Mesoplodon mirus* in New Zealand. *Mar Biodivers Rec* **7**: e1.

Dalebout ML. (2002). Species identity, genetic diversity and molecular systematic relationships among the Ziphiidae (beaked whales). PhD thesis, University of Auckland, New Zealand.

Dalebout ML, Hooker SK, Christensen I (2001). Genetic diversity and population structure among northern bottlenose whales, *Hyperoodon ampullatus*, in the western North Atlantic Ocean. *Can J Zool* **79**: 478–484.

Dalebout ML, Robertson KM, Frantzis A, Engelhaupt D, Mignucci-Giannoni AA, Rosario-Delestre RJ *et al.* (2005). Worldwide structure of mtDNA diversity among Cuvier's beaked whales (*Ziphius cavirostris*): implications for threatened populations. *Mol Ecol* **14**: 3353–3371.

Dalebout ML, Russell KG, Little MJ, Ensor P (2004). Observations of live Gray's beaked whales (*Mesoplodon grayi*) in Mahurangi Harbour, North Island, New Zealand, with a summary of at-sea sightings. *J R Soc NZ* **34**: 347–356.

Dalebout ML, Ruzzante DE, Whitehead H, Øien NI (2006). Nuclear and mitochondrial markers reveal distinctiveness of a small population of bottlenose whales (*Hyperoodon ampullatus*) in the western North Atlantic. *Mol Ecol* **15**: 3115–3129.

De Leo FC, Smith CR, Rowden AA, Bowden DA, Clark MR (2010). Submarine canyons: hotspots of benthic biomass and productivity in the deep sea. *Proc R Soc Ser B* **277**: 2783–2792.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* **29**: 1969–1973.

Earl DA, vonHoldt BM (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Res* **4**: 359–361.

Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* **14**: 2611–2620.

Excoffier L, Lischer HEL (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Res* **10**: 564–567.

Foote AD, Newton J, Piertney SB, Willerslev E, Gilbert MTP (2009). Ecological, morphological and genetic divergence of sympatric North Atlantic whale populations. *Mol Ecol* **18**: 5207–5217.

Fu XY (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**: 915–925.

Gaggiotti OE, Bekkevold D, Jørgensen HBH, Foll M, Carvalho GR, Andre C *et al.* (2009). Disentangling the effects of evolutionary, demographic, and environmental factors influencing genetic structure of natural populations: Atlantic herring as a case study. *Evolution* **63**: 2939–2951.

Garner DM (1959). The sub-tropical convergence in New Zealand surface waters. *New Zeal J Geol Geop* **2**: 315–337.

Gill PC, Pirzl R, Morrice MG, Lawton K (2015). Cetacean diversity of the continental shelf and slope off Southern Australia. *J Wildl Manage* **79**: 672–681.

Gilson A, Syvanen M, Levine K, Banks J (1998). Deer gender determination by polymerase chain reaction: validation study and application to tissues, bloodstains, and hair forensic samples from California. *Calif Fish Game* **84**: 159–169.

Graves JA, Helyar A, Biuw M, Jüssi M, Jüssi I, Karlsson O (2008). Microsatellite and mtDNA analysis of the population structure of grey seals (*Halichoerus grypus*) from three breeding areas in the Baltic Sea. *Conserv Genet* **10**: 59–68.

Hare MP, Nunney L, Schwartz MK, Ruzzante DE, Burford M, Waples RS *et al.* (2011). Understanding and estimating effective population size for practical application in marine species management. *Conserv Biol* **25**: 438–449.

Heath RA (1981). Oceanic fronts around Southern New Zealand. *Deep-Sea Res* **28**: 547–560.

Ho SYW, Kolokotronis S-Y, Allaby RG (2007). Elevated substitution rates estimated from ancient DNA sequences. *Biol Lett* **3**: 702–705.

Ho SYW, Lanfear R, Phillips MJ, Barnes I, Thomas JA, Kolokotronis S-O *et al.* (2011). Bayesian estimation of substitution rates from ancient DNA sequences with low information content. *Syst Biol* **60**: 1–10.

Hoelzel AR, Dahlheim M, Stern SJ (1998). Low genetic variation among killer whales (*Orcinus orca*) in the eastern North Pacific and genetic differentiation between foraging specialists. *J Hered* **89**: 121–128.

Hudson RR (1990). Gene genealogies and the coalescent process. In: Futuyama F, Antonovics JD (eds). *Oxford Surveys in Evolutionary Biology*, Vol 7. Oxford University Press: New York, NY, USA.

Huelsenbeck JP, Ronquist F (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**: 754–755.

Jackson JA, Baker CS, Vant M, Steel DJ, Medrano-González L, Palumbi SR (2009). Big and slow: Phylogenetic estimates of molecular evolution in baleen whales (suborder *Mysticeti*). *Mol Biol Evol* **26**: 2427–2440.

Jakobsson M, Rosenberg NA (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**: 1801–1806.

Johnson M, Hickmott LS, Aguilar Soto N, Madsen PT (2008). Echolocation behaviour adapted to prey in foraging Blainville's beaked whale (*Mesoplodon densirostris*). *Proc R Soc Ser B* **275**: 133–139.

Jombart T (2008). Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**: 1403–1405.
Katoh K, Misawa K, Kuma K, Miyata T (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**: 3059–3066.
Kalinowski ST (2005). HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Mol Ecol Notes* **5**: 187–189.
Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S *et al*. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**: 1647–1649.
Luikart G, Ryman N, Tallmon DA, Schwartz MK, Allendorf FW (2010). Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. *Conserv Genet* **11**: 355–373.
MacLeod CD, Santos MB, Pierce GJ (2003). Review of data on diets of beaked whales: evidence of niche separation and geographic seqregation. *J Mar Biol Assoc UK* **83**: 651–665.
MacDiarmid AB, Law CS, Pinkerton M, Zeldis J (2013). New Zealand marine ecosystem services. In: Dymond JR (ed). *Ecosystem Services in New Zealand—Conditions and Trends*. Manaaki Whenua Press: Lincoln, New Zealand.
Madsen PT, Aguilar de Soto N, Tyack PL, Johnson M (2014). Beaked whales. *Curr Biol* **24**: R728–R730.
McSweeney D, Baird R, Mahaffy S (2007). Site fidelity, associations, and movements of Cuvier's (*Ziphius cavirostris*) and Blainville's (*Mesoplodon densirostris*) beaked whales off the island of Hawai'i. *Mar Mammal Sci* **23**: 666–687.
Meirmans PG, Hedrick PW (2011). Assessing population structure: $F_{st}$ and related measures. *Mol Ecol Resour* **11**: 5–8.
Meirmans PG, Van Tienderen PH (2004). GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Mol Ecol Notes* **4**: 792–794.
Moore JE, Barlow JP (2013). Declining abundance of beaked whales (Family Ziphiidae) in the California current large marine ecosystem. *PLoS One* **8**: e52770.
Morin PA, Archer FI, Foote AD *et al*. (2010). Complete mitochondrial genome phylogeographic analysis of killer whales (*Orcinus orca*) indicates multiple species. *Genome Res* **20**: 908–916.
Morin PA, Duchene S, Lee N, Durban J, Claridge D (2013). Preliminary analysis of mitochondrial genome phylogeography of Blainvilles', Cuvier's and Gervais' beaked whales. *Report to the Scientific Committee of the International Whaling Commission (No. SC/64/SM14)*, p 17.
New LF, Moretti DJ, Hooker SK, Costa DP, Simmons SE (2013). Using energetic models to investigate survivial and reproduction of beaked whales (family Ziphiidae). *PLoS ONE* **8**: e68725.
Oremus M, Gales R, Dalebout ML, Funahashi N, Endo T, Kage T *et al*. (2009). Worldwide mitochondrial DNA diversity and phylogeography of pilot whales (*Globicephala* spp.). *Biol J Linn Soc* **98**: 729–744.
Oremus M, Poole M, Steel D, Baker CS (2007). Isolation and interchange among insular spinner dolphin communities in the South Pacific revealed by individual identification and genetic diversity. *Mar Ecol Prog Ser* **336**: 257–289.
Patel S, Thompson K, Williams L, Tsai P, Constantine R, Millar C (2014). Mining microsatellites for Gray's beaked whale from second-generation sequencing data. *Conserv Genet Resour* **6**: 657–659.
Pimper LE, Baker CS, Goodall RNP, Olavarría C, Remis MI (2010). Mitochondrial DNA variation and population structure of Commerson's dolphins (*Cephalorhynchus commersonii*) in their southernmost distribution. *Conserv Genet* **11**: 2157–2168.
Pitman R (2009). Mesoplodont whales (*Mesoplodon* spp.). In: Perrin WF, Würsig B, Thewissen JGM (eds). *Encyclopedia of Marine Mammals*, 2nd edn. Academic Press: San Diego, CA, USA, pp 721–726
Posada D (2008). jModelTest: phylogenetic model averaging. *Mol Biol Evol* **25**: 1253–1256.
Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.
Rambaut A, Suchard MA, Xie D, Drummond AJ (2013). Tracer v1.5. Available from http://beast.bio.ed.ac.uk/Tracer (accessed 12th May 2015).
Ramos-Onsins SE, Rozas J (2002). Statistical properties of new neutrality tests against population growth. *Mol Biol Evol* **19**: 2092–2100.
Robards MD, Reeves RR (2011). The global extent and character of marine mammal consumption by humans: 1970–2009. *Biol Conserv* **144**: 2770–2786.
Rosenberg NA (2004). DISTRUCT: a program for the graphical display of population structure. *Mol Ecol Notes* **4**: 137–138.
Sagnol O, Richter C, Field LH, Reitsma F (2014). Spatio-temporal distribution of sperm whales (*Physeter macrocephalus*) off Kaikoura, New Zealand, in relation to bathymetric features. *NZ J Zool* **41**: 234–247.
Sambrook J, Fritsch EF, Maniatis T (1989). *Molecular Cloning: A Laboratory Manual*, Vol 3. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA.
Scheidat M, Friedlaender A, Kock K-H, Lehnert L, Boebel O, Roberts J *et al*. (2011). Cetacean surveys in the Southern Ocean using icebreaker-supported helicopters. *Polar Biol* **34**: 1513–1522.
Schick RS, Halpin PN, Read AJ, Urban DL, Best BD, Good CP *et al*. (2011). Community structure in pelagic marine mammals at large spatial scales. *Mar Ecol Prog Ser* **434**: 165–181.
Schorr GS, Falcone EA, Moretti DJ, Andrews RD (2014). First long-term behavioral records from Cuvier's beaked whales (*Ziphius cavirostris*) reveal record-breaking dives. *PLoS One* **9**: e92633.
Sremba AL, Hancock-Hanser B, Branch TA, LeDuc RL, Baker CS (2012). Circumpolar diversity and geographic differentiation of mtDNA in the critically endangered Antarctic blue whale (*Balaenoptera musculus intermedia*). *PLoS One* **7**: e32579.
Tamura K, Nei M (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* **10**: 512–526.
Taylor BL, Baird R, Barlow J *et al*. (2008). *Mesoplodon grayi*. The IUCN Red List of Threatened Species 2008: e.T13247A3428839. Available at: www.iucnredlist.org (last accessed 27 February 2014).
Thompson K, Baker S, van Helden A, Patel S, Millar C, Constantine R (2012). The world's rarest whale. *Curr Biol* **22**: 905–906.
Thompson K, Millar C, Baker C, Dalebout M, Steel D, van Helden A *et al*. (2013). A novel conservation approach provides insights into the management of rare cetaceans. *Biol Conserv* **157**: 331–340.
Thompson K, Ruggiero K, Millar C, Constantine R, van Helden A (2014). Large-scale multivariate analysis reveals sexual dimorphism and geographic differences in the Gray's beaked whale. *J Zool* **294**: 13–21.
Torres LG (2013). Evidence for an unrecognised blue whale foraging ground in New Zealand. *NZ J Mar Fresh* **47**: 235–248.
Van Oosterhout C, Hutchinson WF, Wills DP, Shipley P (2004). MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* **4**: 535–538.
Van Waerebeek K, Leaper R, Baker A, Papastavrou V, Thiele D, Findlay K *et al*. (2010). Odontocetes of the southern ocean sanctuary. *J Cet Res Man* **11**: 315–346.
Varela AI, Ritchie PA, Smith PJ (2012). Low levels of global genetic differentiation and population expansion in the deep-sea teleost *Hoplostethus atlanticus* revealed by mitochondrial DNA sequences. *Mar Biol* **159**: 1049–1060.
van Ruth PD, Ganf GG, Ward TM (2010). Hot-spots of primary productivity: an alternative interpretation to conventional upwelling models. *Estuar Coast Shelf S* **90**: 142–158.
Viaud-Martinez KA, Brownell RL, Komnenou A, Bohonak AJ (2008). Genetic isolation and morphological divergence of Black Sea bottlenose dolphins. *Biol Cons* **141**: 1600–1611.
von Haast J (1876). On a new ziphioid whale. *Proc Zool Soc Lond* **1876**: 7–13.
Weilgart L (2007). The impacts of anthropogenic ocean noise on cetaceans and implications for management. *Can J Zool* **85**: 1091–1116.
Whitehead H (1998). Cultural selection and genetic diversity in matrilineal whales. *Science* **282**: 1708–1711.
Wimmer T, Whitehead H (2004). Movements and distribution of northern bottlenose whales, *Hyperoodon ampullatus*, on the Scotian Slope and in adjacent waters. *Can J Zool* **82**: 1782–1794.
Winkelmann I, Campos PF, Strugnell J, Cherel Y, Smith PJ, Kubodera T *et al*. (2013). Mitochondrial genome diversity and population structure of the giant squid *Architeuthis*: genetics sheds new light on one of the most enigmatic marine species. *Proc R Soc Ser B* **280**: 20130273.
Wright AJ, Deak T, Parsons ECM (2011). Size matters: management of stress responses and chronic stress in beaked whales and other marine mammals may require larger exclusion zones. *Mar Pollut Bull* **63**: 5–11.

Supplementary Information accompanies this paper on Heredity website (http://www.nature.com/hdy)

# Bucking the trend: Genetic analysis reveals high diversity and low differentiation in a deep ocean cetacean

KIRSTEN F. THOMPSON[*1,2], SELINA PATEL[*1,2], C. SCOTT BAKER[1,3], ROCHELLE CONSTANTINE[2] and CRAIG D. MILLAR[1,2]

[1] School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

[2] The Allan Wilson Centre, School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

[3] Department of Fisheries and Wildlife and Marine Mammal Institute, Oregon State University, OR 97365, USA

* Authors contributed equally

**Table S1**. Gray's beaked whale (*Mesoplodon grayi*) specimen data key for age-class and condition of animal at time of stranding.

| Key | Age-class | Condition of animal |
|---|---|---|
| 0 | Unknown | Unknown |
| 1 | Foetus | Live stranded |
| 2 | Calf | Wounded or sick |
| 3 | Mature female | Freshly dead |
| 4 | Mature female in group with dependents | |
| 5 | Mature male | |
| 6 | Sub-adult female | |
| 7 | Sub-adult male | |
| 8 | Sub-adult unknown sex | |
| 9 | Mature unknown sex | |

*NA = data not available, NDR = no Department of Conservation record

**Table S2**. Gray's beaked whale (*Mesoplodon grayi*) specimen data.

| NZ CeTA code | Date | Location | Latitude | Longitude | Genetic sex | Group code | Group size | Age class | Haplotype | Total body length (m) | Condition of animal |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Mlay03 | 23 Jan 1993 | Oamaru | -45.1081 | 170.9719 | F | NA | 1 | 0 | T | NDR | 0 |
| MgrH02 | 1 Dec 1993 | Waiheke | -36.8152 | 175.0862 | F | Mgrm1 | 4 | 0 | P | NA | 4 |
| MgrH04 | 1 Dec 1993 | Waiheke | -36.8152 | 175.0862 | F | Mgrm1 | 4 | 0 | A | NA | 4 |
| Mgr01 | 6 Dec 1994 | Mahia | -39.0821 | 177.9107 | F | Mgrcc1 | 2 | 4 | D | NDR | 0 |
| Mgr05 | 22 Aug 1995 | Hawkes Bay | -39.0945 | 177.2783 | M | NA | 1 | 0 | W | NDR | 0 |
| Mgr06 | 12 Jan 1996 | Muriwai | -36.7657 | 174.3869 | M | Mgrm16 | 3 | 7 | L1 | 3.45 | 1 |
| Mgr07 | 12 Jan 1996 | Muriwai | -36.7657 | 174.3869 | F | Mgrm16 | 3 | 4 | J | 4.76 | 1 |
| Mgr08 | 12 Jan 1996 | Muriwai | -36.7657 | 174.3869 | F | Mgrm16 | 3 | 4 | L1 | 4.7 | 1 |
| Mgr10 | 13 May 1995 | Wellington | -41.3439 | 174.7724 | F | NA | 1 | 3 | F | 4.9 | 3 |
| Mgr12 | 6 Feb 1996 | McGregor's Bay | -35.7291 | 174.5524 | F | Mgrm2 | 6 | 4 | W1 | 4.6 | 1 |
| Mgr13 | 6 Feb 1996 | McGregor's Bay | -35.7291 | 174.5524 | F | Mgrm2 | 6 | 4 | L | 4.6 | 1 |
| Mgr14 | 6 Feb 1996 | McGregor's Bay | -35.7291 | 174.5524 | M | Mgrm2 | 6 | 7 | AH | 3.45 | 1 |
| Mgr17 | 22 Sep 1996 | Gt. Barrier Is. | -36.2422 | 175.4787 | M | Mgrm3 | 2 | 5 | Y | 4.96 | 1 |
| Mgr18 | 22 Sep 1996 | Gt. Barrier Is. | -36.2422 | 175.4787 | M | Mgrm3 | 2 | 7 | NA | 3.2 | 1 |
| Mgr21 | 11 Oct 1996 | Te Werahi | -34.656 | 173.0274 | M | NA | 1 | 5 | K | 4.5 | 4 |
| Mgr22 | 1 Mar 1996 | Waitangi | -43.9543 | -176.554 | M | NA | 1 | 0 | A1 | NDR | 0 |
| Mgr24 | 18 Jan 1997 | Whangarei | -35.828 | 174.3949 | F | Mgrcc3 | 1 | 4 | W1 | NDR | 0 |
| Mgr31 | 7 May 1997 | Chatham Is. | -43.8599 | -176.562 | M | Mgrm4 | 4 | 0 | A | NDR | 0 |
| Mgr32 | 7 May 1997 | Chatham Is. | -43.8599 | -176.562 | M | Mgrm4 | 4 | 0 | C | NDR | 0 |
| Mgr33 | 1 Sep 1997 | Chatham Is. | -43.8599 | -176.562 | M | NA | 1 | 0 | L | NA | 3 |
| Mgr38 | 18 Mar 1998 | Chatham Is. | -43.8599 | -176.562 | M | Mgrm5 | 10 | 0 | I | NA | 3 |
| Mgr39 | 18 Mar 1998 | Chatham Is. | -43.8599 | -176.562 | M | Mgrm5 | 10 | 0 | U | NA | 3 |
| Mgr40 | 18 Mar 1998 | Chatham Is. | -43.8599 | -176.562 | F | Mgrm5 | 10 | 0 | O1 | NA | 3 |
| Mgr41 | 18 Mar 1998 | Chatham Is. | -43.8599 | -176.562 | F | Mgrm5 | 10 | 0 | H | NA | 3 |
| Mgr42 | 18 Mar 1998 | Chatham Is. | -43.8599 | -176.562 | M | Mgrm5 | 10 | 0 | W | NA | 3 |
| Mgr43 | 18 Mar 1998 | Chatham Is. | -43.8599 | -176.562 | M | Mgrm5 | 10 | 0 | O | NA | 3 |
| Mgr44 | 18 Mar 1998 | Chatham Is. | -43.8599 | -176.562 | M | Mgrm5 | 10 | 0 | O3 | NA | 3 |

**Table S3**. Gray's beaked whale specimen data (continued).

| NZ CeTA code | Date | Location | Latitude | Longitude | Genetic sex | Group code | Group size | Age class | Haplotype | Total body length (m) | Condition of animal |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Mgr45 | 18 Mar 1998 | Chatham Is. | -43.8599 | -176.562 | M | Mgrm5 | 10 | 0 | O | NA | 3 |
| Mgr46 | 18 Mar 1998 | Chatham Is. | -43.8599 | -176.562 | M | Mgrm5 | 10 | 0 | A1 | NA | 3 |
| Mgr47 | 18 Mar 1998 | Chatham Is. | -43.8599 | -176.562 | M | Mgrm5 | 10 | 0 | M | NA | 3 |
| Mgr48 | 21 Apr 1998 | Clifton | -39.6408 | 177.0059 | F | NA | 1 | 0 | O1 | NDR | 0 |
| Mgr49 | 1 Sep 1997 | Blenheim | -41.482 | 174.0372 | M | NA | 1 | 0 | V | NDR | 0 |
| Mgr58 | 10 Aug 1999 | Wellington | -41.2905 | 174.7894 | M | NA | 1 | 7 | L1 | 4.34 | 3 |
| Mgr59 | 11 Oct 1999 | Lake Ferry | -41.4006 | 175.1465 | M | NA | 1 | 5 | L | 5.2 | 1 |
| Mgr60 | 12 Jan 2000 | Opotiki | -37.6797 | 177.7271 | M | Mgrm6 | 3 | 7 | O1 | 4.1 | 1 |
| Mgr61 | 12 Jan 2000 | Opotiki | -37.6797 | 177.7271 | F | Mgrm6 | 3 | 6 | D | 3.72 | 1 |
| Mgr62 | 13 Jan 2000 | Opotiki | -37.6797 | 177.7271 | F | Mgrm6 | 3 | 0 | O | NA | 1 |
| Mgr63 | 15 Mar 2000 | Torere Beach | -37.9495 | 177.4885 | F | NA | 1 | 0 | NA | NA | 3 |
| Mgr64 | 10 Feb 2000 | Opunake | -39.4595 | 173.8577 | F | NA | 1 | 4 | X | 4.8 | 4 |
| Mgr66 | 11 Apr 2000 | Rangirewa | -37.6189 | 177.9124 | F | NA | 1 | 6 | NA | 4.9 | 4 |
| Mgr67 | 18 May 2000 | Marsden Point | -35.8523 | 174.4886 | F | NA | 1 | 3 | L | 4.7 | 1 |
| Mgr68 | NA | Opotiki | -37.9911 | 177.289 | M | NA | 1 | 0 | O1 | NA | 0 |
| Mgr76 | 21 Mar 2002 | Napier | -39.635 | 176.9935 | F | NA | 1 | 6 | Z | 3.86 | 4 |
| Mgr78 | 9 Apr 2002 | Karamea | -41.2413 | 172.0971 | M | NA | 1 | 5 | A1 | 4.6 | 3 |
| Mgr81 | 12 Feb 2003 | Kauri Mountain Beach | -35.7662 | 174.5538 | F | Mgrm7 | 3 | 3 | L | 4.53 | 3 |
| Mgr82 | 12 Feb 2003 | Kauri Mountain Beach | -35.7662 | 174.5538 | F | Mgrm7 | 3 | 4 | K | 5.2 | 3 |
| Mgr83 | 12 Feb 2003 | Kauri Mountain Beach | -35.7662 | 174.5538 | M | Mgrm7 | 3 | 7 | L | 3.3 | 3 |
| Mgr86 | 6 Mar 2003 | Whangarei | -35.7793 | 174.407 | M | NA | 1 | 7 | D | 4.25 | 1 |
| Mgr87 | 12 Feb 2003 | Nelson | -41.2671 | 173.2576 | M | NA | 1 | 7 | O2 | 3.9 | 0 |
| Mgr88 | 7 Apr 2003 | Dunedin | -45.8828 | 170.5174 | M | NA | 1 | 0 | AB | NA | 3 |
| Mgr89 | 22 Apr 2003 | Nelson | -41.2795 | 173.238 | M | NA | 1 | 5 | L | 4.8 | 1 |
| Mgr90 | 5 Feb 2004 | Glink's Gully | -36.2544 | 173.9912 | F | Mgrm8 | 2 | 6 | O3 | 3.47 | 3 |
| Mgr91 | 5 Feb 2004 | Glink's Gully | -36.2544 | 173.9912 | F | Mgrm8 | 2 | 6 | A1 | 4.04 | 3 |
| Mgr92 | 14 Feb 2004 | Whangarei Heads | -35.8597 | 174.5625 | F | Mgrm9 | 4 | 6 | S | 4.3 | 3 |

**Table S1**. Gray's beaked whale specimen data (continued).

| NZ CeTA code | Date | Location | Latitude | Longitude | Genetic sex | Group code | Group size | Age class | Haplotype | Total body length (m) | Condition of animal |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Mgr93 | 14 Feb 2004 | Whangarei Heads | -35.8597 | 174.5625 | F | Mgrm9 | 4 | 3 | W1 | 4.5 | 3 |
| Mgr94 | 14 Feb 2004 | Whangarei Heads | -35.8597 | 174.5625 | F | Mgrm9 | 4 | 3 | D | 4.7 | 3 |
| Mgr95 | 14 Feb 2004 | Whangarei Heads | -35.8597 | 174.5625 | F | Mgrm9 | 4 | 3 | D | 4.5 | 3 |
| Mgr96 | 31 Jan 2004 | Kawau Is. | -36.4102 | 174.8339 | F | NA | 1 | 6 | O2 | 4.05 | 3 |
| Mgr99 | 8 Jun 2004 | Napier | -39.4713 | 176.8804 | F | NA | 1 | 6 | W1 | 3.45 | 1 |
| Mgr103 | 26 Aug 2004 | Ruapuke | -37.8931 | 174.7648 | F | NA | 1 | 3 | A1 | 4.5 | 3 |
| Mgr105 | 20 Nov 2004 | Surat Bay | -46.4768 | 169.7392 | M | NA | 1 | 2 | A1 | 2.05 | 0 |
| Mgr112 | NA | Akaroa | -43.8279 | 172.6907 | M | NA | 1 | 0 | E1 | NA | 0 |
| Mgr116 | 21 Dec 2005 | Mahia | -39.093 | 177.867 | F | NA | 1 | 0 | R1 | NDR | 0 |
| Mgr117 | 4 Feb 2006 | Waipu Beach | -35.9749 | 174.472 | F | Mgrm10 | 5 | 3 | L | 4.65 | 3 |
| Mgr118 | 4 Feb 2006 | Waipu Beach | -35.9749 | 174.472 | F | Mgrm10 | 5 | 3 | W3 | 4.73 | 3 |
| Mgr123 | 5 May 2006 | Muriwai | -36.8063 | 174.4127 | F | Mgrm12 | 2 | 6 | AL | 3.9 | 4 |
| Mgr124 | 5 May 2006 | Muriwai | -36.8063 | 174.4127 | F | Mgrm12 | 2 | 6 | D | 3.72 | 4 |
| Mgr125 | 25 May 2006 | Motueka | -41.201 | 173.0851 | M | NA | 1 | 5 | AH | 4.7 | 3 |
| Mgr127 | 27 Mar 2006 | Te Arai Beach | -36.1478 | 174.6382 | M | Mgrm11 | 2 | 5 | O2 | 4.55 | 4 |
| Mgr129 | 26 Jul 2006 | Kahurangi Lighthouse | -40.7743 | 172.2195 | M | NA | 1 | 5 | O1 | 4.7 | 4 |
| Mgr130 | 16 Feb 2007 | Chatham Is. | -43.8167 | -176.706 | F | NA | 1 | 6 | C | 3.62 | 3 |
| Mgr132 | 25 Feb 2007 | Chatham Is. | -43.8108 | -176.71 | F | NA | 2 | 6 | O2 | 3.87 | 3 |
| Mgr134 | 16 Apr 2007 | Ruatapu | -42.8039 | 170.8824 | F | NA | 1 | 3 | R1 | 5.05 | 3 |
| Mgr135 | 22 Sep 2007 | Te Awanga | -39.6323 | 176.9885 | M | NA | 1 | 5 | AF | 4.78 | 1 |
| Mgr138 | 12 Apr 2009 | Opotiki | -37.6618 | 177.8424 | M | Mgrm13 | 2 | 0 | W1 | NA | 0 |
| Mgr139 | 27 May 2009 | Mahia | -39.0699 | 177.8053 | M | NA | 1 | 0 | AG | NA | 0 |
| Mgr141 | 12 Apr 2009 | Opotiki | -37.6618 | 177.8424 | M | Mgrm13 | 2 | 0 | S | NA | 0 |
| Mgr144 | 29 Jul 2009 | Kariotahi | -37.2841 | 174.6539 | F | Mgrm14 | 2 | 3 | O1 | 5.06 | 1 |
| Mgr145 | 29 Jul 2009 | Kariotahi | -37.2841 | 174.6539 | M | Mgrm14 | 2 | 5 | A | 4.69 | 1 |
| Mgr148 | 13 Jan 2011 | Sunset Beach | -37.3931 | 174.7099 | F | Mgrm15 | 5 | 4 | K | 4.8 | 3 |
| Mgr149 | 13 Jan 2011 | Sunset Beach | -37.3931 | 174.7099 | F | Mgrm15 | 5 | 4 | L1 | 4.7 | 3 |
| Mgr150 | 13 Jan 2011 | Sunset Beach | -37.3931 | 174.7099 | F | Mgrm15 | 5 | 6 | O1 | 3.3 | 3 |

**Table S3**. Gray's beaked whale specimen data (continued).

| NZ CeTA code | Date | Location | Latitude | Longitude | Genetic sex | Group code | Group size | Age class | Haplotype | Total body length (m) | Condition of animal |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Mgr151 | 13 Jan 2011 | Sunset Beach | -37.3931 | 174.7099 | F | Mgrm15 | 5 | 4 | O1 | 4.8 | 3 |
| Mgr152 | 18 Jan 2011 | Sunset Beach | -37.3931 | 174.7099 | F | Mgrm15 | 5 | 6 | L | 4.15 | 3 |
| Mgr153 | 27 Jan 2009 | Raumati | -40.913 | 174.9763 | F | NA | 1 | 3 | L2 | 4.82 | 1 |
| Mgr155 | 18 Oct 2010 | Mahia | -39.0699 | 177.8053 | M | NA | 1 | 5 | L2 | NA | 4 |
| Mgr156 | 20 Feb 2009 | Carter's Beach | -41.7499 | 171.549 | F | NA | 1 | 3 | S | 4.7 | 3 |
| Mgr157 | 2 May 2011 | Chatham Is. | -43.8552 | -176.566 | M | NA | 1 | 7 | L1 | 4.5 | 0 |
| Mgr159 | 5 Feb 2012 | Ninety Mile beach | -34.9917 | 173.1465 | F | NA | 1 | 8 | J | 4.35 | 0 |
| Mgr161 | 21 Mar 2012 | Bethell's Beach | -36.8966 | 174.4457 | F | NA | 1 | 9 | A1 | 5.6 | 0 |
| Mgr162 | 30 Apr 2012 | Hamilton's Gap | -37.1307 | 174.5736 | M | NA | 1 | 9 | K | 4.78 | 0 |
| Mgr164 | 6 Jan 2013 | Ninety Mile Beach | -34.8863 | 173.0736 | F | Mgrm17 | 2 | 8 | C | 3.3 | 0 |
| Mgr165 | 6 Jan 2013 | Ninety Mile Beach | -34.8863 | 173.0736 | F | Mgrm17 | 2 | 8 | L2 | 4.3 | 0 |
| M49874 | 19 Jan 2003 | Bussleton, West Australia | -33.2662 | 115.642 | F | NA | 1 | 3 | L | NA | 0 |
| M49871 | 16 Jan 2003 | Yallingup Beach | -33.6759 | 114.9927 | F | NA | 1 | 3 | R1 | NA | 0 |
| M49873 | 15 Jan 2003 | Eagle Bay | -33.5631 | 115.0662 | F | NA | 1 | 3 | O1 | NA | 0 |
| WA04 | 27 Dec 2000 | Mandurah, West Australia | -33.2662 | 115.642 | M | NA | 2 | 0 | L2 | NA | 0 |
| TM15 | 24 Jan 2002 | Leighton Beach, West Australia | -33.2662 | 115.642 | F | NA | 1 | 6 | O1 | NA | 0 |
| TM16 | 2 Mar 2003 | Wonnerup Beach, West Australia | -33.2662 | 115.642 | F | NA | 1 | 3 | W2 | 4.9 | 0 |
| TM16 | 2 Mar 2003 | Wonnerup Beach, West Australia | -33.2662 | 115.642 | F | NA | 1 | 3 | W2 | 4.9 | 0 |

**Table S3.** Gray's beaked whale *(Mesoplodon grayi*) mitochondrial DNA haplotypes and the variable positions in control region (530 bp).

| Haplotype | 2 | 20 | 22 | 39 | 56 | 59 | 91 | 92 | 98 | 101 | 104 | 109 | 123 | 126 | 133 | 134 | 149 | 193 | 252 | 289 | 319 | 353 | 370 | 408 | 413 | 418 | 519 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Hap A | A | G | C | A | A | G | T | G | A | C | C | T | T | C | C | T | A | A | C | C | T | A | A | G | T | C | C |
| Hap A1 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | G | . | A | . | . | . |
| Hap C | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | G | . | . | . | G | . | A | . | . | . |
| Hap D | . | . | . | . | . | . | . | . | . | . | . | . | G | . | . | . | . | G | . | . | . | G | G | A | . | . | . |
| Hap E1 | . | A | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | G | . | . | . | G | . | A | . | . | . |
| Hap F | . | . | . | . | . | . | . | . | . | . | . | . | . | T | . | . | . | . | T | . | . | G | . | A | . | T | . |
| Hap H | . | A | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | G | . | A | . | . | . |
| Hap I | – | . | . | . | . | . | . | A | . | . | . | . | . | . | . | . | . | . | T | . | . | . | G | A | . | T | . |
| Hap J | . | . | . | . | G | . | . | . | . | T | . | . | . | . | . | . | . | . | T | . | . | . | . | A | . | . | . |
| Hap K | . | . | . | . | G | . | . | . | G | . | . | . | . | . | . | . | . | . | T | . | . | . | . | A | . | . | . |
| Hap L | . | . | . | . | . | . | . | . | . | . | . | C | . | . | . | . | . | . | T | . | . | . | . | A | . | . | . |
| Hap L1 | . | . | . | . | . | . | . | . | . | . | . | C | . | . | . | . | . | . | T | . | . | G | . | . | . | . | . |
| Hap L2 | . | . | . | . | . | . | . | . | . | . | . | C | . | . | . | . | . | . | T | . | . | G | . | A | . | . | . |
| Hap M | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | T | . | . | | | | | | |
| Hap O | . | . | . | . | . | . | . | . | . | . | . | . | . | T | . | . | . | G | T | . | . | . | . | A | . | . | . |
| Hap O1 | . | . | . | . | . | . | . | . | . | . | . | . | . | T | . | . | . | G | T | . | . | G | . | A | . | T | . |
| Hap O2 | . | . | . | . | . | . | . | . | . | . | . | . | . | T | . | . | . | G | T | . | . | G | . | . | . | T | . |
| Hap O3 | . | . | . | . | . | . | . | . | . | . | . | . | . | T | . | . | . | G | T | . | . | . | . | A | . | T | . |
| Hap P | . | . | . | . | . | . | . | . | . | . | . | . | . | T | . | . | . | G | T | . | C | G | . | A | . | . | T |
| Hap R | . | A | . | . | . | A | . | . | . | . | . | . | . | T | . | . | . | G | T | . | . | G | . | . | . | . | . |
| Hap R1 | . | A | . | . | . | A | . | . | . | . | . | . | . | T | . | . | . | G | T | . | . | G | . | A | . | . | . |
| Hap S | . | . | . | . | . | A | . | . | . | . | . | . | . | T | . | . | . | G | T | . | . | G | . | A | . | . | . |
| Hap S1 | . | . | . | . | . | A | . | . | . | . | . | . | . | T | . | . | . | G | T | . | . | G | . | A | . | T | . |
| Hap T | . | . | . | . | G | . | . | . | . | . | . | . | . | . | T | C | . | G | T | . | . | . | . | A | . | . | . |
| Hap U | . | A | . | . | G | . | . | . | . | . | G | . | . | . | T | C | . | . | T | . | . | . | . | A | . | . | . |
| Hap V | . | A | . | . | G | . | . | . | . | . | . | . | . | . | T | C | . | . | T | . | C | . | G | A | . | . | . |
| Hap W | . | . | . | . | G | . | . | . | . | . | . | . | . | . | T | C | . | . | T | . | . | G | . | A | . | . | . |
| Hap W1 | . | . | . | . | G | . | . | . | . | . | . | . | . | . | T | C | . | . | T | . | . | . | . | A | . | . | . |
| Hap W2 | . | . | . | . | G | . | . | . | . | . | . | . | . | . | T | C | . | . | T | . | . | G | . | A | C | . | . |
| Hap W3 | . | . | . | . | G | . | . | . | . | . | . | . | . | . | T | C | . | . | T | . | . | G | . | . | . | . | . |
| Hap X | . | . | . | . | G | . | . | . | . | . | . | . | . | . | T | C | . | G | . | . | . | . | . | A | . | . | . |
| Hap Y | . | . | . | . | G | . | . | . | . | . | . | . | . | . | T | C | G | . | . | . | . | G | . | A | . | . | . |
| Hap Z | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | T | . | . | G | . | A | . | T | . |
| Hap AB | . | . | . | G | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | G | . | . | . | . | . |
| Hap AF | . | A | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | T | . | . | . | . | A | . | . | . |
| Hap AG | . | . | T | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | T | . | . | . | . | . | . | . |
| Hap AH | – | . | . | . | . | . | . | . | . | . | . | . | . | T | . | . | . | G | T | . | . | G | . | . | . | T | . |
| Hap AL | . | A | . | . | . | . | . | . | . | . | . | . | . | T | . | . | . | G | T | . | . | G | . | A | . | . | |

**Table S4**. Tests for Hardy-Weinberg Equilibrium per locus per population. No significance was found after Bonferroni correction.

| | North East New Zealand | | | North West New Zealand | | | South East New Zealand | | | South West New Zealand | | | West Australia | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Locus | N | $H_o$ | $H_e$ | N | $H_o$ | $H_e$ | N | $H_o$ | $H_e$ | N | $H_o$ | $H_e$ | N | $H_o$ | $H_e$ |
| 1 | 35 | 0.8 | 0.824 | 24 | 0.833 | 0.846 | 22 | 1 | 0.852 | 7 | 0.571 | 0.659 | 6 | 1 | 0.894 |
| 27 | 35 | 0.912 | 0.895 | 24 | 0.875 | 0.882 | 22 | 0.955 | 0.905 | 7 | 0.857 | 0.901 | 6 | 0.833 | 0.894 |
| 29 | 35 | 0.714 | 0.751 | 24 | 0.750 | 0.742 | 22 | 0.773 | 0.763 | 7 | 0.571 | 0.769 | 6 | 0.333 | 0.561 |
| 32 | 35 | 0.457 | 0.463 | 24 | 0.542 | 0.545 | 22 | 0.409 | 0.429 | 7 | 0.571 | 0.495 | 6 | 0.333 | 0.303 |
| 37 | 35 | 0.657 | 0.657 | 24 | 0.583 | 0.616 | 22 | 0.409 | 0.587 | 7 | 0.857 | 0.736 | 6 | 0.667 | 0.546 |
| 58 | 35 | 0.743 | 0.808 | 24 | 0.792 | 0.819 | 22 | 0.591 | 0.776 | 7 | 0.714 | 0.835 | 6 | 0.833 | 0.758 |
| 75 | 35 | 0.286 | 0.388 | 24 | 0.333 | 0.337 | 22 | 0.273 | 0.304 | 7 | 0.571 | 0.440 | 6 | 0.167 | 0.410 |
| 78 | 35 | 0.743 | 0.77 | 24 | 0.625 | 0.736 | 22 | 0.714 | 0.791 | 7 | 0.714 | 0.791 | 6 | 0.333 | 0.758 |
| 88 | 35 | 0.857 | 0.788 | 24 | 0.708 | 0.670 | 22 | 0.864 | 0.746 | 7 | 0.857 | 0.714 | 6 | 0.833 | 0.849 |
| 95 | 35 | 0.571 | 0.635 | 24 | 0.542 | 0.681 | 22 | 0.546 | 0.623 | 7 | 0.714 | 0.626 | 6 | 0.667 | 0.561 |
| 98 | 35 | 0.543 | 0.529 | 24 | 0.542 | 0.575 | 22 | 0.571 | 0.568 | 7 | 0.571 | 0.440 | 6 | 0.667 | 0.561 |
| 99 | 35 | 0.6 | 0.581 | 24 | 0.542 | 0.507 | 22 | 0.571 | 0.565 | 7 | 0.429 | 0.385 | 6 | 0.500 | 0.455 |

**Table S5**. Diversity statistics for microsatellite loci. $H_o$ denotes the observed heterozygosity, $H_s$ denotes the within population heterozygosity and $H_t$ denotes the total heterozygosity.

| Locus | Number of alleles | Effective number of alleles | $H_o$ | $H_s$ | $H_t$ |
|---|---|---|---|---|---|
| 1 | 16 | 4.66 | 0.854 | 0.819 | 0.837 |
| 27 | 22 | 6.58 | 0.909 | 0.855 | 0.885 |
| 29 | 7 | 3.22 | 0.575 | 0.727 | 0.751 |
| 32 | 9 | 1.71 | 0.432 | 0.435 | 0.439 |
| 37 | 9 | 2.40 | 0.666 | 0.607 | 0.597 |
| 58 | 10 | 4.14 | 0.737 | 0.759 | 0.801 |
| 75 | 2 | 1.54 | 0.306 | 0.367 | 0.356 |
| 78 | 7 | 3.38 | 0.629 | 0.741 | 0.739 |
| 88 | 7 | 4.04 | 0.850 | 0.783 | 0.776 |
| 95 | 8 | 2.59 | 0.596 | 0.643 | 0.634 |
| 98 | 8 | 2.21 | 0.658 | 0.567 | 0.560 |
| 99 | 9 | 2.04 | 0.569 | 0.531 | 0.530 |
| Overall | 9.5 | 3.21 | 0.648 | 0.658 | 0.659 |

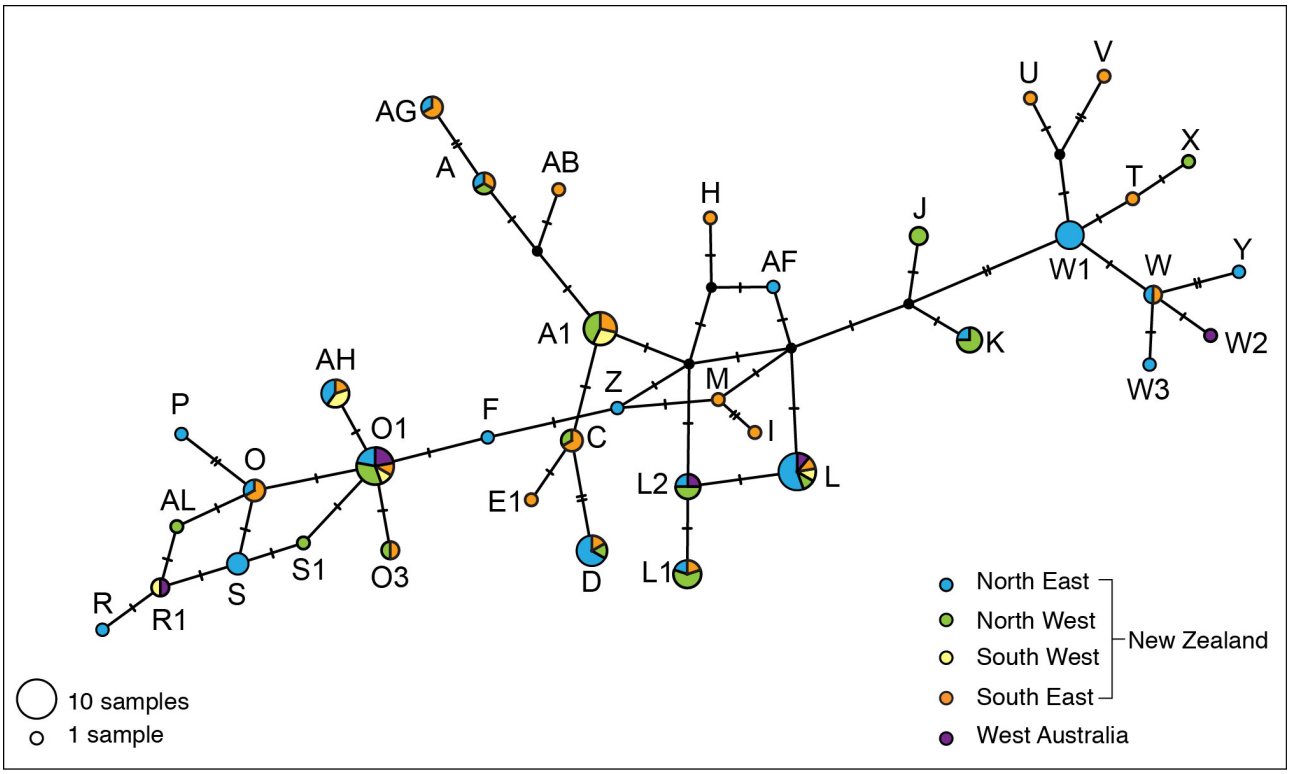



**Figure S1.** A median-joining network of the genealogical relationships among Gray's beaked whale (*Mesoplodon grayi*) mtDNA control region haplotypes from New Zealand and Australia. Branch lengths are not proportional but the number of inferred mutational steps is indicated by a hatch on the branches. The diameter of each circle is proportional to the number of individuals of each haplotype. Haplotype names are shown beside circles. Black circles represent median vectors corresponding to theoretical unsampled haplotypes.

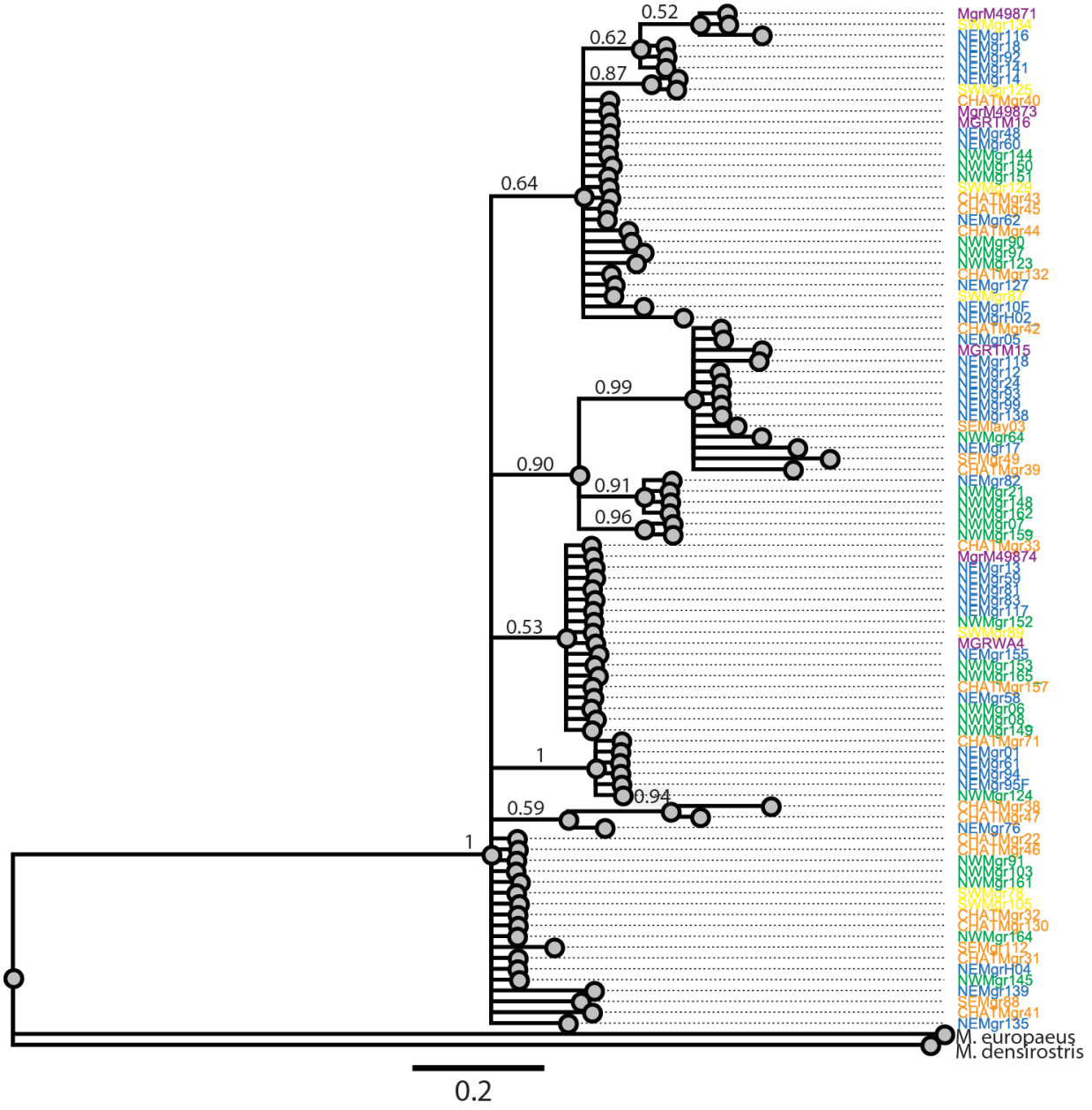



**Figure S2.** A Bayesian tree presenting the phylogenetic reconstruction of the mitochondrial DNA control region haplotypes of Gray's beaked whales (*Mesoplodon grayi*). Samples found in the North East are shown in blue, north west samples are shown in green, South West samples are shown in yellow, South East samples are shown in orange and West Australian samples are shown in purple.

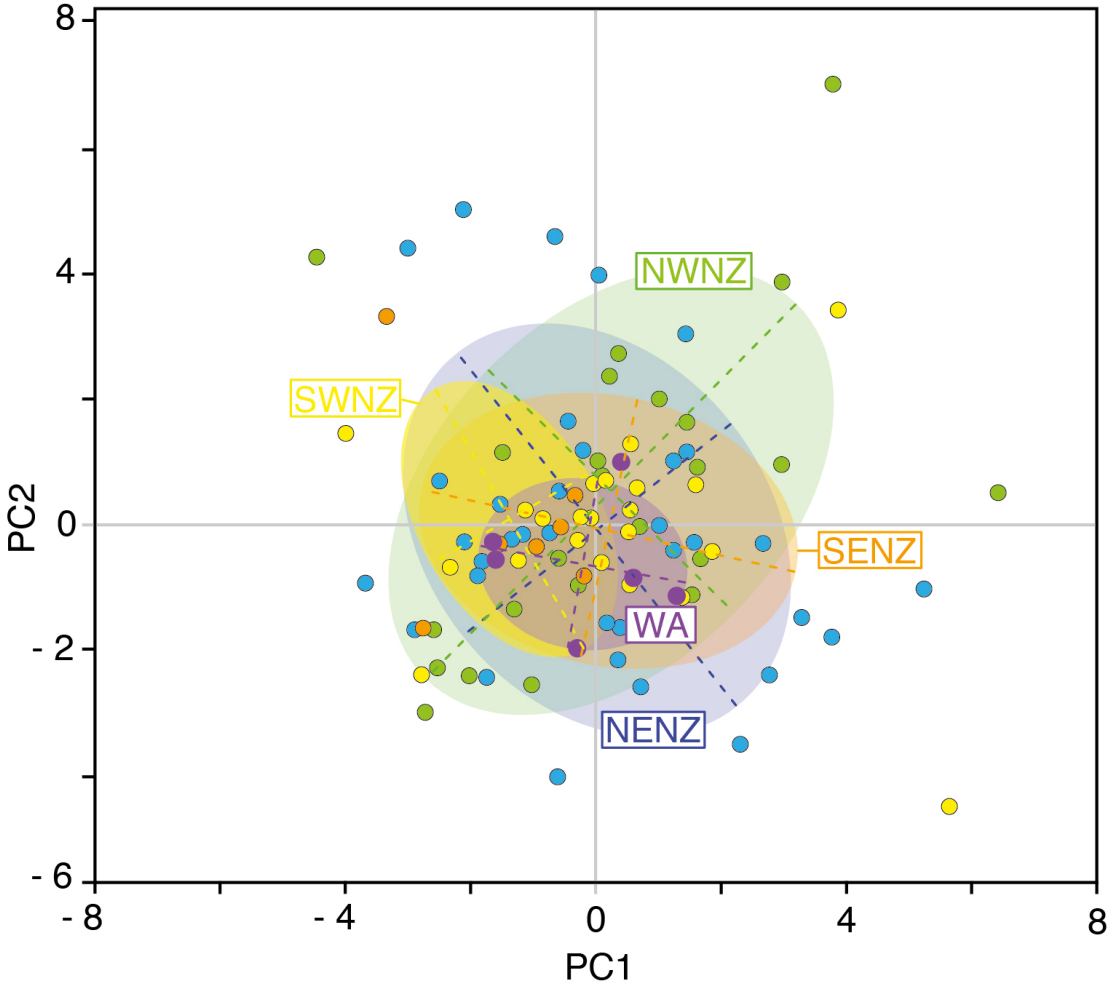



**Figure S3**. Principal components 1 and 2 of individual genotypes (dots) and assumed populations from a priori regions (circles). Colours indicate a priori regions, West Australia (purple), North East (blue), North West (green), South East (orange) and South West (yellow). The centre of the distribution of each region is marked by the intersection of dashed lines.

American Genetic Association

*Journal of Heredity*, 2016, 1–6
doi:10.1093/jhered/esx021
Brief Communication

OXFORD

Brief Communication

# Genetic Kinship Analyses Reveal That Gray's Beaked Whales Strand in Unrelated Groups

**Selina Patel*, Kirsten F. Thompson*, Anna W. Santure, Rochelle Constantine, and Craig D. Millar**

From the School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand (Patel, Thompson, Santure, Constantine, and Millar); and Biosciences, College of Life and Environmental Sciences, University of Exeter, Geoffrey Pope, Stocker Road, Exeter EX4 4QD, UK (Thompson).

*These authors contributed equally to the work.

Address correspondence to Craig D. Millar at the address above, or e-mail: cd.millar@auckland.ac.nz.

Received June 2, 2016; First decision August 25, 2016; Accepted February 16, 2017

Corresponding editor: William Sherwin

## Abstract

Some marine mammals are so rarely seen that their life history and social structure remain a mystery. Around New Zealand, Gray's beaked whales (*Mesoplodon grayi*) are almost never seen alive, yet they are a commonly stranded species. Gray's are unique among the beaked whales in that they frequently strand in groups, providing an opportunity to investigate their social organization. We examined group composition and genetic kinship in 113 Gray's beaked whales with samples collected over a 20-year period. Fifty-six individuals stranded in 19 groups (2 or more individuals), and 57 whales stranded individually. Mitochondrial control region haplotypes and microsatellite genotypes (16 loci) were obtained for 103 whales. We estimated pairwise relatedness between all pairs of individuals and average relatedness within, and between, groups. We identified 6 mother–calf pairs and 2 half-siblings, including 2 whales in different strandings 17 years and 1500 km apart. Surprisingly, none of the adults stranding together were related suggesting that groups are not formed through the retention of kin. These data suggest that both sexes may disperse from their mothers, and groups consisting of unrelated subadults are common. We also found no instances of paternity within the groups. Our results provide the first insights into dispersal, social organization, and the mating system in this rarely sighted species. Why whales strand is still unknown but, in Gray's beaked whales, the dead can tell us much about the living.

**Subject area:** Reproductive strategies and kinship analysis
**Key words:** beaked whales, DNA, genetic kinship, relatedness, social systems, strandings

Cetaceans are unique among social mammals in that certain species regularly die together in mass-strandings. Sperm whales (*Physeter macrocephalus*), short- and long-finned pilot whales (*Globicephala macrorhynchus* and *Globicephala melas*, respectively), and false killer whales (*Pseudorca crassidens*) are known to mass-strand frequently (Evans et al. 2005; Ferreira et al. 2014; Téllez et al. 2014). Humans have pondered the significance of cetacean mass-strandings since the time of Aristotle and the commonly held view is that the powerful bonds of kinship between whales overcome their individual need to survive. When group members become unwell, or disoriented, potentially a whole family can die on the beach. Mass-stranding species have been described as having relatively stable, and in some cases, complex kin-based social structures, for example, sperm whales form several hierarchically organized tiers of female social structure.

At the base of this hierarchical organization is the "unit" which is made up of several adult females from one or more matrilines, their dependent calves, and juveniles (Mesnick 2001; Gero et al. 2008). These units can assemble into groups over periods of hours or days. Over several thousand kilometers, units can also be classified into "clans" based on the similarity of their vocalizations (Rendell and Whitehead 2003). Similarly, long-finned pilot whales also exhibit a matrilineal social structure with social units forming the base of a hierarchical social organization (Amos et al. 1993; Ottensmeyer and Whitehead 2003). However, according to Whitehead et al. (2012) and Oremus et al. (2013) both these deep-diving species show differences in such associations between the Pacific and the Atlantic. Other delphinids also mass-strand, for example, common dolphins (*Delphinus delphis*), and melon-headed whales (*Peponocephala electra*) (Viricel et al. 2008; Amano et al. 2014). Samples from stranding events present us with an opportunity to characterize rare species' diet, reproduction, and other life-history parameters as well as group composition, trophic level, and aspects of the environment such as background contaminant levels (Newsome et al. 2010; Ramos and González-Solís 2012; Thompson et al. 2012; Amano et al. 2014; Jepson et al. 2016).

New Zealand is a recognized global hotspot for strandings (Pyenson 2011) and some of the most common species to strand are among the most rarely observed—the beaked whales. These whales are known to be the deepest divers of all mammals, foraging for small fish and squid, spending little time at the surface or in coastal waters (Schorr et al. 2014). Most species are rarely observed and nothing is known of their social organization. Gray's beaked whales (*Mesoplodon grayi*) are no exception. However, Gray's are unique among the beaked whales in that they regularly strand in groups. The holotype was 1 of 28 animals found stranded on the Chatham Islands, New Zealand in 1874 (von Haast 1876). Though their global distribution stretches throughout Southern Hemisphere circumpolar waters, most strandings of Gray's beaked whales are around New Zealand. Curiously, even though this species is the second most common cetacean to strand, these whales are exceptionally elusive and are almost never seen alive. In New Zealand, a national collaborative sampling regime has facilitated the collection of both morphological data and tissue samples from stranded Gray's beaked whales (Thompson et al. 2013). Nothing is known of this species' social structure although previous research suggests that genetic diversity is high, with considerable gene flow implying a homogenous population between New Zealand and Australia (Thompson et al. 2016).

We investigated the social organization of Gray's beaked whales by examining group composition and genetic kinship within, and between, stranded groups. We used specimen data to determine maturity and used genetic sex identification, mitochondrial haplotype data, and microsatellite genotyping to estimate familial relationships between whales. To our knowledge, this is the first assessment of genetic kinship in beaked whales. These mass-strandings provide a unique opportunity to examine the social structure of these enigmatic whales.

## Material and Methods

### Sample Collection and Genotyping

We examined a total of 113 stranded Gray's beaked whales where samples were collected over a 20-year period (1993–2013) (103 from New Zealand, 10 from Australia). Of these 113 whales, 56 individuals stranded in 19 groups (2 or more individuals), and 57 whales stranded individually. Five of these groups were live strandings and the remaining whales were either thought to be freshly dead (7 groups) or no data were recorded as to their condition (7 groups). Details of sampling regime, age-class estimates, and morphological data can be found in Thompson et al. (2013) and Thompson et al. (2014). Samples from Australia were from whales that stranded as individuals.

Mitochondrial control region DNA sequence data (530 bp) and genetic sex identification were determined and samples were genotyped at 16 microsatellite loci. A full description of methods used for sex identification, microsatellite genotyping, and control region sequencing are detailed in Thompson et al. (2016) with additional information in Supplementary Material accompanying this publication. Twelve microsatellite loci are detailed in Patel et al. (2014), and a further 4 are in Supplementary Material. Summary diversity statistics for these loci are provided in Thompson et al. (2016) and in Supplementary Material. All individuals were genotyped and genetically sexed apart from 6 from different strandings that were refloated and not sampled, 3 that were not genotyped, 1 that could not be sexed, and 1 that could not be genotyped or sexed. These additional 11 whales have been included in the broader analyses as they provide information on group sizes, sex, and composition of individuals within groups. Fifty-six whales were genotyped for 16 loci (100% genotyping success for 51 individuals; 93.7% for 4; 81.2% for 1) and 47 for 12 loci (100% genotyping success). Full details of the number and locations of whales that mass-stranded are provided in the Results section (see Figure 1).

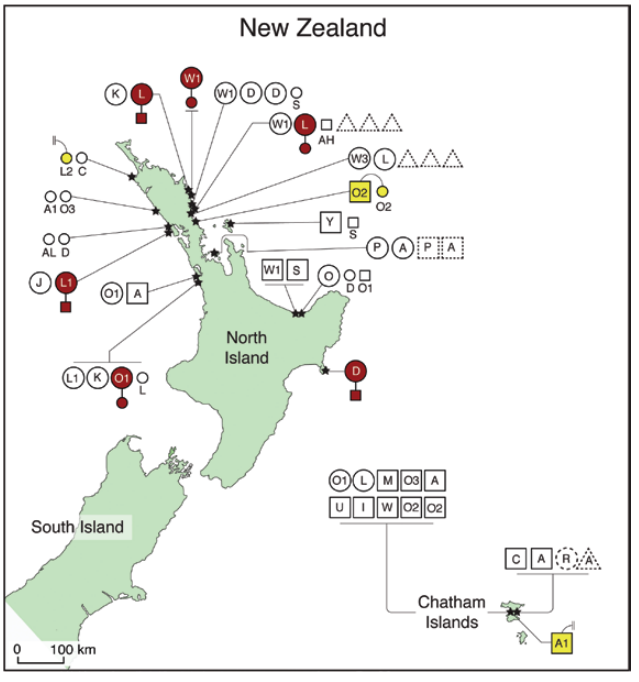



**Figure 1.** Location of Gray's beaked whale (*Mesoplodon grayi*) group strandings and a single individual stranding, where squares denote males, circles denote females, and triangles represent individuals that could not be sexed. Dashed symbols represent individuals not genotyped and smaller symbols indicate dependents or subadults. Mother–calf relationships are shaded black (red) with lines connecting symbols, whereas half-sibling relationships are shown in gray (yellow). Text overlaid or below each symbol denotes mitochondrial haplotypes.

### Kinship Analyses

Genetic kinship was estimated between all 103 genotyped whales. Given no evidence for population structure (Thompson et al. 2016), it was assumed that samples were derived from a single, homogeneous population. In addition, according to Thompson et al. (2016) loci showed no evidence for null alleles, allelic dropout, or scoring errors due to stutter peaks.

Pairwise relatedness between every pair of individuals, as estimated by the triadic likelihood estimator (Wang 2007), was calculated using COANCESTRY v1 (Wang 2011). All analyses included 16 loci, and for individuals only genotyped for 12 loci, the additional 4 were set as unknown. To test whether relatedness was higher within groups than between groups, we restricted the data set to only individuals in group strandings and calculated average relatedness within and between groups. Group allocation was then randomized 5000 times to determine whether the observed difference in relatedness within groups compared with between groups was significant. A histogram was used as a density estimator with binwidth derived using Scott's method (Scott 1979) and a cubic spline interpolation using the spline function in R (R Core Team 2015).

A 2-step approach was taken to recover genetic kinship between pairs of individuals. First, CERVUS v3.0 (Kalinowski et al. 2007) was run to identify putative first-degree relationships between all 103 genotyped individuals. It is possible that strandings included both parents and offspring, so we assigned all 103 individuals as potential offspring, restricting candidate mothers and fathers to females and males not identified as "fetus" (1 individual), "dependent" (<3 m total body length, 4 individuals), or "subadult" (3–4.4 m total body length, 25 individuals). In total, we identified 39 candidate mothers and 35 candidate fathers.

Second, any putative relationships identified by CERVUS were tested in COLONY v2.0.5.5 (Wang and Santure 2009; Wang 2013) using a full likelihood approach to determine whether they were more likely to be parent–offspring, full-sibling, half-sibling, or unrelated. The putative relationships tested comprised of all pairs identified at >80% confidence in CERVUS, and pairs identified at 50% confidence that also had high pairwise relatedness values as estimated in COANCESTRY. Allele frequencies were recalculated after adjusting for biologically feasible parent–offspring pairs that were consistently identified over a range of parameter assumptions in CERVUS. The relationships were then fixed in the input parameters for COLONY, and the adjusted allele frequencies were used to calculate the likelihood of each of the relationships (for further details see Supplementary Material). We combined information from the 2-step CERVUS–COLONY approach with the pairwise relatedness coefficients, sharing of haplotypes and, finally, biological feasibility of relationships to determine the most likely relationships between individuals.

## Results

Across the entire data set (103 stranded whales) we identified 38 mitochondrial control region haplotypes. Within the 19 groups we identified 6 statistically well-supported mother–calf pairs, 2 of which stranded without other whales (Table 1, Cases 1–5 and 8). Excluding these 2 mother–calf pairs, the remaining 17 groups (20 males, 32 females), excluding all calves, had a mean group size of 3.4 (± 0.5 standard error [SE]).

Two possible half-sibling relationships were also identified (Table 1, Cases 6 and 7). In one stranding, a mature male and a female neonate were most likely to be half-siblings, with a shared haplotype suggesting these individuals may have shared a mother (Case 6). The other half-sibling relationship was between a male (stranded 1 March 1996) and a female subadult (stranded 6 January 2013) (Case 7). These whales were sampled 17 years apart, and at least 1500 km away from each other (Figure 1). A parent–offspring relationship was strongly supported by COLONY in this case, though this is biologically unfeasible given the stranding dates. However, a half-sibling relationship had the second highest likelihood; the fact that these whales do not share a haplotype supports the hypothesis that they share a father.

None of the adults within any of the groups were related. When parent–offspring and sibling relationships were excluded, there was no evidence that the average relatedness within groups (0.067) was significantly higher than between groups (0.065); $P < 0.414$. The mean estimated relatedness between all pairs of individuals was low (0.078) (Figure 2) and only 397 pairs (7.6%) had relatedness values of 0.25 or more.

## Discussion

The lack of evidence of parent–offspring relationships or relatedness between adults implies that both sexes disperse from their parents. Curiously, the adult whales that die together in these mass-strandings are unrelated and, therefore, these groups have not formed by the retention of close kin.

**Table 1.** The 8 relationships with strong support from both CERVUS and COLONY analyses.

| Case | Individual 1/code | Individual 2/code | Relatedness | CERVUS confidence | COLONY most likely relationship | Shared haplotype | Same stranding |
|---|---|---|---|---|---|---|---|
| 1 | Male dependent/ Mgr02 | Female adult/Mgr01 | 0.5132 | 95 | Offspring–mother | Yes | Yes |
| 2 | Male subadult/Mgr06 | Female adult/Mgr08 | 0.6157 | 95 | Offspring–mother | Yes | Yes |
| 3 | Female fetus/Mgr11 | Female adult/Mgr13 | 0.5330 | 95 | Offspring–mother | Yes | Yes |
| 4 | Female dependent/ Mgr23 | Female adult/Mgr24 | 0.5749 | 95 | Offspring–mother | Yes | Yes |
| 5 | Female subadult/ Mgr150 | Female adult/Mgr151 | 0.5718 | 95 | Offspring–mother | Yes | Yes |
| 6 | Female dependent/ Mgr122 | Male adult/Mgr127 | 0.5129 | 95 | Half-sibling | Yes | Yes |
| 7 | Female subadult/ Mgr165 | Male unknown age/ Mgr22 | 0.5197 | 95 | Offspring–father; half-sibling second most likely | No | No |
| 8 | Male subadult/Mgr83 | Female adult/Mgr81 | 0.5595 | 85 | Offspring–mother | Yes | Yes |

Half-sibling relationships are highlighted in gray; the remaining are mother–offspring relationships.

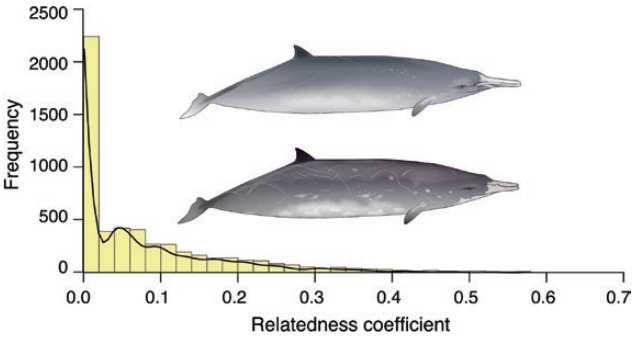



**Figure 2.** Distribution of relatedness coefficients among the 103 individual Gray's beaked whales. The bars give the number of observations within binwidths that were calculated using Scott's method and a cubic spline interpolation is shown as the overlayed black line. Inserted images show a female (above) and male (below).

We provide the first indications of group size (3.4 [± 0.5 SE]) in Gray's beaked whales. MacLeod and D'Amico (2006) suggest that patterns of group size for beaked whales may fall into 2 categories. Some species have smaller groups (2.5–3.5 individuals, maximum 20), for example, northern bottlenose (*Hyperoodon ampullatus*) and southern bottlenose (*Hyperoodon planifrons*) whales, Cuvier's (*Ziphius cavirostris*) and Blainville's (*Mesoplodon densirostris*) beaked whales. Other species have much larger groups, for example, 7.9 individuals for *Berardius* spp. and 19.4 for Longman's beaked whales (*Indopacetus pacificus*). The larger group size category for beaked whales, while based on limited observations, can include groups of up to 100 animals (Friedlaender et al. 2010; Fedutin et al. 2014). Our estimate for Gray's beaked whale mean group size falls within the smaller range but must be interpreted with some caution given uncertainty in whether the whole group, or just part of a group, stranded and that genetic information is just one source of information.

In particular, inferring group composition and social structure solely from genetic data can be problematic. As with other studies of wide-ranging oceanic species (Viricel et al. 2008, common dolphins), obtaining observational data is difficult and costly. In the case of Gray's beaked whales obtaining behavioral data is currently unfeasible without extensive survey work to identify key areas of their distribution. In studies where making direct observations of group structure is possible, for example, bottlenose dolphins (Möller et al. 2001; Krützen et al. 2003; Frère et al. 2010), detailed information on associations between animals, genetic kinship, and paternity can be built over many years. These data can show detailed patterns in social alliances, reproductive strategies, and, more critically, variation among individuals.

Nonetheless, our data represent the most comprehensive collection for this species. These strandings are a "snapshot" of this rarely sighted species at the time of death. However, we have little knowledge of whether Gray's beaked whale groups may have consisted of additional animals that did not strand, or whether those animals may have stranded along the coast and were not discovered. There may also have been other strandings around the broader region that were not sampled throughout our study period. Because of these uncertainties mean group sizes may be underestimated, and group composition may not entirely reflect that of live groups.

Our data suggest that subadults were common within these stranded groups, and it is conceivable that the presence of these immature individuals is a contributing factor to the group stranding. As we have no data on the composition of live Gray's beaked whale groups, it may be that groups with immature individuals have a higher likelihood of stranding. In Case 6 (Table 1), why these 2 half-siblings, a mature male and neonate female, of very different age- and sex-classes, might have stranded together remains unknown. It is highly possible that they were part of a larger group in which the other individuals did not strand or were not discovered.

The mating system of Gray's beaked whales is entirely unknown and the fact that we found no fathers within the groups, or few mixed-sex groups, is intriguing. Gray's beaked whale strandings around New Zealand peak in the austral summer, coinciding with the calving season (September–December), suggesting that these inshore movements may be associated with reproduction (Thompson et al. 2013). We know that solitary mature males tend to strand between March and May (unpublished data). It is possible that these "roving males" strand during an autumnal mating season as they move to coastal waters in search of receptive females. The only genotyped group consisting of multiple males and females was found during autumn on the Chatham Islands (Figure 1). Unfortunately, no specimen data were obtained during this stranding so no estimates of age-class could be implied. However, it is possible that this larger group of 10 whales was a breeding aggregation and that males may form social alliances to consort females, a behaviour seen in bottlenose dolphins (*Tursiops* spp.), but not previously described in beaked whales (Connor et al. 1992).

This is the first study to assess genetic kinship in beaked whales. Other research using photo-identification suggests that beaked whale species exhibit a range of social systems. Northern bottlenose show both sex- and age-class segregation, whereas in contrast, Blainville's beaked whales form "harem" groups of multiple females with at least 1 male (Gowans et al. 2001; Claridge 2013; Dunn 2014). Baird's beaked whales (*Berardius bairdii*) appear to form fission–fusion societies with some stable associations (Fedutin et al. 2014). Unfortunately, in the absence of DNA analyses, we do not know the kinship relationships within these groups. Though other beaked whales do strand, Gray's are the only species to regularly strand in groups.

We found no evidence of kin associations in these stranded Gray's beaked whale groups. These groups may be formed opportunistically for foraging or reproduction and we cannot exclude a fission–fusion system as observed in Baird's beaked whale. Our results are also consistent with the fission–fusion systems reported in several oceanic delphinids, for example, spinner dolphins (*Stenella longirostris*) (Karczmarski et al. 2005), common dolphins (Viricel et al. 2008), and Atlantic white-sided dolphins (*Lagenorhynchus acutus*) (Mirimin et al. 2011). Stranded groups of white-sided dolphins consist of unrelated adults with calves that were also not closely related (e.g., half-sibs) (Mirimin et al. 2011). This suggested that females had mated with males that were not present within the group and that these dolphins most likely have a promiscuous mating system. Unfortunately, in the Gray's beaked whale groups that we have sampled, the sample sizes are too small to provide further information on potential mating systems.

We found that both sexes may disperse, and subadults can be frequently found with unrelated individuals. We also observed groups of mixed age-classes and potential sex segregation. More data would confirm whether the sex segregation we observed in these strandings persists across Gray's beaked whale society. Affiliations of unrelated females are also observed in sperm whales and long-finned pilot whales—species with complex and hierarchical social structures (Ortega-Ortiz et al. 2012; Oremus et al. 2013). Both these species

strand regularly in groups and are known to be deep-divers that feed on similar prey to beaked whales. We hypothesize that there could be similarities in terms of social organization between these oceanic species that induce them to strand so regularly together, but clear differences in terms of their social complexity. A potentially temporally flexible social structure is most likely driven by an extreme deep oceanic lifestyle where prey resources are dispersed and unpredictable. We currently have no data on the temporal nature of Gray's beaked whale associations but, given the suite of new tracking technologies, tracking refloated whales would prove useful. We would also suggest continued effort to collect data from all stranded Gray's beaked whales, particularly full necropsy analyses and further genetic sampling.

## Supplementary Material

Supplementary data are available at *Journal of Heredity* online.

## Funding

The study was funded by University of Auckland FDRF Grant 3702180 (C.D.M., R.C.) and Department of Conservation funding (R.C.).

## Acknowledgments

We thank the New Zealand Department of Conservation; L. Boren, L. Wakelin; iwi; hapu; C.S. Baker; M. Dalebout; D. Steel; A. van Helden; Massey University necropsy teams; Centre for Genomics, Proteomics and Metabolomics (University of Auckland); New Zealand Genomics Ltd. We also thank the following for Australian samples: R. O'Shea and J. Bannister, Western Australian Museum; N. Gales, Australian Antarctic Division; R. Gales and D. Pemberton, Tasmanian Museum and Art Gallery. Samples are held at University of Auckland (NZ Department of Conservation Permit Rnw/HO/2009/03, CITES institutional permit NZ010). We also thank 2 anonymous reviewers and the editor for their comments that improved the manuscript.

## Data Availability

Data deposited at Dryad: http://dx.doi.org/10.5061/dryad.m3q59.

## References

Amano M, Yamada TK, Kuramochi T, Hayano A, Kazumi A, Sakai T. 2014. Life history and group composition of melon-headed whales based on mass strandings in Japan. *Mar Mamm Sci*. 30:480–493.

Amos B, Schlötterer C, Tautz D. 1993. Social structure of pilot whales revealed by analytical DNA profiling. *Science* 260:670–672.

Claridge DE. 2013. *Population ecology of Blainville's beaked whales* (*Mesoplodon densirostris*) [doctoral dissertation]. St Andrews (Scotland): University of St Andrews.

Connor RC, Smolker RA, Richards AF. 1992. Two levels of alliance formation among male bottlenose dolphins (*Tursiops* sp.). *Proc Natl Acad Sci USA*. 89:987–990.

Dunn C. 2014. *Insights into Blainville's beaked whale* (*Mesoplodon densirostris*) *communication* [doctoral dissertation]. St Andrews (Scotland): University of St Andrews.

Evans K, Thresher R, Warneke RM, Bradshaw CJ, Pook M, Thiele D, Hindell MA. 2005. Periodic variability in cetacean strandings: links to large-scale climate events. *Biol Lett*. 1:147–150.

Fedutin I, Filatova O, Mamaev E, Burdin A, Hoyt E. 2014. Occurrence and social structure of Baird's beaked whales, *Berardius bairdii*, in the Commander Islands, Russia. *Mar Mamm Sci*. 31:853–865.

Ferreira IM, Kasuya T, Marsh H, Best PB. 2014. False killer whales (*Pseudorca crassidens*) from Japan and South Africa: differences in growth and reproduction. *Mar Mamm Sci*. 30:64–84.

Friedlaender A, Nowacek DP, Johnston DW, Read AJ, Tyson RB, Peavey L, Revelli EMS. 2010. Multiple sightings of large groups of Arnoux's beaked whales (*Berardius arnuxii*) in the Gerlach Strait, Antarctica. *Mar Mamm Sci*. 26:246–250.

Frère CH, Krützen M, Mann J, Watson-Capps JJ, Tsai YJ, Patterson EM, Connor R, Bejder L, Sherwin WB. 2010. Home range overlap, matrilineal and biparental kinship drive female associations in bottlenose dolphins. *Anim Behav*. 80:481–486.

Gero S, Engelhaupt D, Whitehead H. 2008. Heterogeneous social associations within a sperm whale, *Physeter macrocephalus*, unit reflect pairwise relatedness. *Behav Ecol Sociobiol*. 63:143–151.

Gowans S, Whitehead H, Hooker SK. 2001. Social organization in northern bottlenose whales, *Hyperoodon ampullatus*: not driven by deep-water foraging? *Anim Behav*. 62:369–377.

Jepson PD, Deaville R, Barber JL, Aguilar À, Borrell A, Murphy S, Barry J, Brownlow A, Barnett J, Berrow S, *et al*. 2016. PCB pollution continues to impact populations of orcas and other dolphins in European waters. *Sci Rep*. 6:18573.

Kalinowski ST, Taper ML, Marshall TC. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol Ecol*. 16:1099–1106.

Karczmarski L, Wursig B, Gailey G, Larson KW, Vanderli C. 2005. Spinner dolphins in a remote Hawaiian atoll: social grouping and population structure. *Behav Ecol*. 16:675–685.

Krützen M, Sherwin WB, Connor RC, Barré LM, Van de Casteele T, Mann J, Brooks R. 2003. Contrasting relatedness patterns in bottlenose dolphins (*Tursiops* sp.) with different alliance strategies. *Proc Biol Sci*. 270:497–502.

MacLeod CD, D'Amico A. 2006. A review of beaked whale behaviour and ecology in relation to assessing and mitigating impacts of anthropogenic noise. *J Cetacean Res Manage*. 7:211–211.

Mesnick SL. 2001. Genetic relatedness in sperm whales: evidence and culture implications. *Behav Brain Sci*. 24:346–347.

Mirimin L, Banguera-Hinestroza E, Dillane E, Hoelzel AR, Cross TF, Rogan E. 2011. Insights into genetic diversity, parentage, and group composition of Atlantic white-sided dolphins (*Lagenorhynchus acutus*) off the west of Ireland based on nuclear and mitochondrial genetic markers. *J Hered*. 102:79–87.

Möller LM, Beheregaray LB, Harcourt RG, Krützen M. 2001. Alliance membership and kinship in wild male bottlenose dolphins (*Tursiops aduncus*) of southeastern Australia. *Proc Biol Sci*. 268:1941–1947.

Newsome SD, Clementz MT, Koch PL. 2010. Using stable isotope biogeochemistry to study marine mammal ecology. *Mar Mamm Sci*. 26:509–572.

Oremus M, Gales R, Kettles H, Baker CS. 2013. Genetic evidence of multiple matrilines and spatial disruption of kinship bonds in mass strandings of long-finned pilot whales, *Globicephala melas*. *J Hered*. 104:301–311.

Ortega-Ortiz JG, Engelhaupt D, Winsor M, Mate BR, Hoelzel AR. 2012. Kinship of long-term associates in the highly social sperm whale. *Mol Ecol*. 21:732–744.

Ottensmeyer CA, Whitehead H. 2003. Behavioural evidence for social units in long-finned pilot whales. *Can J Zool*. 81:1327–1338.

Patel S, Thompson K, Williams L, Tsai P, Constantine R, Millar C. 2014. Mining microsatellites for Gray's beaked whale from second-generation sequencing data. *Conserv Genet Resour*. 6:657–659.

Pyenson ND. 2011. The high fidelity of the cetacean stranding record: insights into measuring diversity by integrating taphonomy and macroecology. *Proc Biol Sci*. 278:3608–3616.

R Core Team. 2015. *R: a language and environment for statistical computing*. Vienna (Austria): R Foundation for Statistical Computing. Available from: http://www.R-project.org/.

Ramos R, González-Solís J. 2012. Trace me if you can: the use of intrinsic biogeochemical markers in marine top predators. *Front Ecol Environ*. 10:258–266.

Rendell LE, Whitehead H. 2003. Vocal clans in sperm whales (*Physeter macrocephalus*). *Proc Biol Sci*. 270:225–231.

Schorr GS, Falcone EA, Moretti DJ, Andrews RD. 2014. First long-term behavioral records from Cuvier's beaked whales (*Ziphius cavirostris*) reveal record-breaking dives. *PLoS ONE*. 9:e92633.

Scott DW. 1979. On optimal and data-based histograms. *Biometrika*. 66:605–610.

Téllez R, Mignucci-Giannoni AA, Caballero S. 2014. Initial description of short-finned pilot whale (*Globicephala macrorhynchus*) genetic diversity from the Caribbean. *Biochem Syst Ecol*. 56:196–201.

Thompson K, Baker CS, van Helden A, Patel S, Millar C, Constantine R. 2012. The world's rarest whale. *Curr Biol*. 22:R905–R906.

Thompson KF, Millar CD, Baker CS, Dalebout M, Steel D, van Helden AL, Constantine R. 2013. A novel conservation approach provides insights into the management of rare cetaceans. *Biol Cons*. 157:331–340.

Thompson KF, Patel S, Baker CS, Constantine R, Millar CD. 2016. Bucking the trend: genetic analysis reveals high diversity, large population size and low differentiation in a deep ocean cetacean. *Heredity (Edinb)*. 116:277–285.

Thompson KF, Ruggiero K, Millar CD, Constantine R, van Helden AL. 2014. Large-scale multivariate analysis reveals sexual dimorphism and geographic differences in the Gray's beaked whale. *J Zool*. 294:13–21.

Viricel A, Strand AE, Rosel PE, Ridoux V, Garcia P. 2008. Insights on common dolphin (*Delphinus delphis*) social organization from genetic analysis of a mass-stranded pod. *Behav Ecol Sociobiol*. 63:173.

von Haast J. 1876. On a new ziphioid whale. *Royal Proc Zool Soc Lond*. 7–13.

Wang J. 2007. Triadic IBD coefficients and applications to estimating pairwise relatedness. *Genet Res*. 89:135–153.

Wang J. 2011. COANCESTRY: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Mol Ecol Resour*. 11:141–145.

Wang J. 2013. An improvement on the maximum likelihood reconstruction of pedigrees from marker data. *Heredity (Edinb)*. 111:165–174.

Wang J, Santure AW. 2009. Parentage and sibship inference from multilocus genotype data under polygamy. *Genetics*. 181:1579–1594.

Whitehead H, Antunes R, Gero S, Wong SNP, Engelhaupt D, Rendell L. 2012. Multilevel societies of female sperm whales (*Physeter macrocephalus*) in the Atlantic and Pacific: why are they so different? *Int J Primatol*. 33:1142–1164.

## Supplementary Material

# Genetic kinship analyses reveal that Gray's beaked whales strand in unrelated groups

Selina Patel*[1], Kirsten Thompson*[1, 2], Anna Santure[1], Rochelle Constantine[1], Craig Millar**[1]

[1] School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand.

[2] Molecular Ecology and Evolution Group, Biosciences, College of Life and Environmental Sciences, Geoffrey Pope, Stocker Rd, University of Exeter, EX4 4QD, UK

* Joint first author with equal contribution

** Author for correspondence: Craig D. Millar, email: cd.millar@auckland.ac.nz

### Supplemental experimental procedures

We used three second-generation sequencing technologies (Roche 454 GS Junior, Illumina MiSeq and Life Technologies Ion Torrent) to shot-gun sequence high molecular weight DNA from a Gray's beaked whale sample (Patel et al. 2014). The sequences generated were then assembled using SOAPdenovo2 (Luo et al. 2012) and imported into GENEIOUS (Kearse et al. 2012). Strict quality control criteria were applied to sequences so that 80 % of each sequence has a PHRED score of > 20, and post-trimmed reads were analysed through a pipeline in QDD (Meglécz et al. 2010) to identify microsatellite repeats, as described in Patel et al. (2014). Twelve microsatellite markers were genotyped on 47 individuals from group and individual strandings (Thompson et al. 2016).

In this manuscript, four additional loci were genotyped on 56 individuals (Table S1). These additional loci were obtained using genomic resources detailed in Patel et al. (2014). Polymerase chain reactions for the four additional loci were performed in 10µL reaction mixtures containing 10 ng of genomic DNA, 1mg/ml BSA, 2.5mM $MgCl_2$, 0.4 µM of each primer, 0.25U of HotStarTaq Plus DNA polymerase (Qiagen), 1x Q solution (Qiagen). Cycling conditions for Loci Mg-11 and Mg-61 were as follows; 10 cycles of 95°C for 30 s, 65°C for 45 s, and 72°C for 60 s. The annealing temperature was reduced by 1°C every cycle. This was followed by 18 cycles of 95°C for 30 s, 56°C for 45 s, and 72°C for 60 s, then 30 min at 60°C for the final extension. For locus Mg-23, the conditions were the same except the annealing temperature of 54°C was kept constant for 30 cycles in total and for locus Mg-93 an annealing temperature of 62°C was kept constant for a total of 28 cycles. Amplified products were electrophoresed using an ABI 3130 sequencer (Applied Biosystems) and fragment sizes were analysed using the microsatellite plugin in GENEIOUS (Kearse et al. 2012). Linkage disequilibrium and deviation from Hardy-Weinberg equilibrium (HWE) were assessed in ARLEQUIN ver 3.5 (Excoffier and Lischer 2010) and GENODIVE (Meirmans and Van Tienderen 2004). A Bonferroni correction was applied to all data. Linkage disequilibrium was found to be non-significant between loci, including the 12 other loci detailed in Patel et al. (2014).

**Table S1.** Summary data of four microsatellite loci for Gray's beaked whale (*Mesoplodon grayi*) additional to those reported in Patel et al. (2014), *n* denotes the number of individuals tested per locus, *A* denotes the number of alleles, $H_o$ is the observed heterozygosity and $H_s$ denotes expected heterozygosity after correction for sampling bias.

| Locus | *n* | A | $H_o$ | $H_s$ | Repeat | Primer sequence 5' – 3' |
|---|---|---|---|---|---|---|
| Mg-11 | 56 | 10 | 0.732 | 0.81 | CA | F-TAGGCCTGTGTTGTGCATGT<br>R-TGTGCATTCAGACTGGCAAT |
| Mg-23 | 53 | 19 | 0.868 | 0.913 | TG | F-ACTCATTGCAAGGATGGTCC<br>R-GCCTGGAGGGACACAGTTAC |
| Mg-61 | 56 | 5 | 0.321 | 0.36 | ATTT | F-TGAAAATGCAAGGTCAGACAA<br>R-ATAAATGTTTGCGCAGGAGG |
| Mg-93 | 56 | 8 | 0.768 | 0.782 | AAAT | F-GACAGGCAATTCTGCACTCA<br>R-GAGGCCACAACAGTGAGAGG |

**Kinship Analyses**

For COANCESTRY v1 analyses, the number of reference individuals for the triadic likelihood estimator was set to 100, allele frequencies were estimated from all genotyped individuals with an error rate for loci assumed to be 0.05 (Wang 2011).

CERVUS v3.0 uses a pairwise likelihood approach to parentage assignment, with simulation to determine critical thresholds for likelihood ratios, and assign confidence to putative parentage assignments (Kalinowski et al. 2007). Strict and relaxed critical delta thresholds were determined by simulating 100,000 offspring given 39 mothers, 35 fathers, a minimum of 12 loci typed (of a total of 16 loci), with mistyping rates per locus of 0.05. To first determine the impact of simulation input parameters on parentage

assignment, we ran CERVUS with a strict confidence of 95% and varied the proportion of sampled candidate parents and the relatedness and inbreeding rates. Relationships identified at 95% confidence were consistently recovered when small changes were made to the simulation input parameters. Changes to the overall inbreeding level and relatedness of true parents (0.00, 0.04, 0.08, 0.09, 0.10) and proportion of parents sampled (0.04 both parents, 0.1 mothers and 0.05 fathers) did not affect the output. Given the consistency of relationships identified, allele frequencies were recalculated adjusting for parent-offspring pairs identified at strict confidence that had strong biological justification. CERVUS was then re-run twice, first with strict confidence at 95% and relaxed confidence at 85%, and second at 80% (strict) and 50% (relaxed) confidence, using the adjusted allele frequencies, with proportion of sampled candidate parents of 0.04. Inbreeding rates, as estimated from COANCESTRY, were set to 0.07 for relatedness of true parents and overall rate of inbreeding in the population.

**Supplementary References**

Excoffier L, Lischer HEL. 2010. ARLEQUIN suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Res.* 10: 564 – 567.

Kalinowski ST, Taper ML, Marshall TC. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.* 16: 1099 – 1106.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28: 1647 – 1649.

Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Lui Y, et al. 2012. SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler. *GigaScience* 1:18.

Meglécz E, Costedoat C, Dubut V, Gilles A, Malausa T, Pech N, Martin JF. 2010. QDD: A user friendly program to select microsatellite markers and design primers from large sequencing projects. *Bioinformatics* 26: 403 – 404.

Meirmans PG, Van Tienderen PH. 2004. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Mol. Ecol. Notes* 4: 792 – 794.

Patel S, Thompson K, Williams L, Tsai P, Constantine R, Millar C. 2014. Mining microsatellites for Gray's beaked whale from second-generation sequencing data. *Conserv. Genet. Res.* 6: 657 – 659.

Thompson KF, Patel S, Baker CS, Constantine R, Millar CD. 2016. Bucking the trend: Genetic analysis reveals high diversity and low differentiation in a deep ocean cetacean. Heredity 116: 287 – 115.

Wang J. 2011. COANCESTRY: A program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Mol. Ecol. Res.* 11: 141 – 145.