

University of Groningen

Evolutionary ecology of marine mammals

Cabrera, Andrea A.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2018

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Cabrera, A. A. (2018). Evolutionary ecology of marine mammals [Groningen]: University of Groningen

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

2

Genetics and genomics in marine mammals

Palsbøll, P. J., Cabrera, A. A. & Bérubé, M.

The study of marine mammals is difficult and expensive, particularly in the oceans where most species are difficult to observe. Many species are highly mobile with large and remote distributions, which, combined with their long generation time, make evolutionary and demographic changes difficult to detect, especially using field studies. Advances in molecular genetic techniques provide an opportunity for investigating such changes and to examine interactions among populations and the roles of individuals within those populations. With a small tissue sample collected directly or indirectly from living or deceased animals, nuclear and mitochondrial genomic information can be obtained. Nowadays, we are able to assess questions about selection and adaptation, mechanisms and patterns of speciation, systematics and taxonomy of different groups, population structure demographic changes, parentage and mating systems, among others. Here, we provide a brief overview of the key applications of genetics and genomics to the study of marine mammals. We address a wide range of questions from the origin of marine mammal lineages, taxonomy and adaptation to the identification of individuals, their sex and pathogens.

Published in Encyclopedia of Marine Mammals. 3rd Edition, 2018

Introduction

Genetics has undergone several significant technological advances since the early 1960s, which have dramatically expanded the application of genetics across life sciences, most recently into the field of genomics, which is a subarea of genetics. Consequently, genetic methods are employed not only to study genetics *per se*, but also as an indirect means for inquiries into the evolution, ecology, behavior and conservation of natural populations. This chapter focuses on the current, most commonly applied methodologies.

The analysis of changes in the DNA sequence itself took a giant leap forward by the development of the polymerase chain reaction in the late 1980s (Mullis & Faloona, 1987b). PCR-based analyses require minute amounts of genetic material which, in turn, facilitated the use of nonlethal methods to obtain the required tissue for DNA extraction, such as skin biopsies (Aguilar & Nadal, 1984; Palsbøll *et al.*, 1991b), as well as noninvasive tissue samples, such as feces and even “exhalant breath” (e.g., Acevedo-Whitehouse *et al.*, 2010). The high sensitivity and specificity of PCR-based analyses meant that DNA obtained from ancient/historic samples (Sremba *et al.*, 2015) or environmental samples, such as filtered sea water (Foote *et al.*, 2012b) can be analyzed to determine species and assess genetic diversity.

The human genome project initiated the subsequent development of so-called massive “parallelized” sequencing technologies which was applied to non-model species from 2010 and onwards (Kircher & Kelso, 2010; Davey *et al.*, 2011) thereby facilitating the adoption and application of genomic approaches to the study of marine mammals (Miller *et al.*, 2012; Zhou *et al.*, 2013; Yim *et al.*, 2014; Foote *et al.*, 2015; Keane *et al.*, 2015). Current genomic approaches can be tailored to target different kinds of hereditary variation, such as the nucleotide sequence in DNA and RNA sequences, proteins as well as epigenetic modifications, such as methylation. Most studies that generate genomic data sets from marine mammal species and populations take advantage of the vast amounts of data generated to obtain more precise estimates of aspects, such as, gene flow, past demographic events as well as the detection of genes under selection (Allendorf *et al.*, 2010), providing novel, key insights into the ecology and evolution of marine mammals.

Here, we provide a brief overview of the key applications of genetics and genomics (which is a subdiscipline of genetics) to the study of marine mammals. The range of questions towards which genetics have been applied in marine mammals is very broad ranging from the origin of marine mammal lineages, taxonomy and adaptation to the identification of individuals, their sex and

pathogens. Herein, we highlight a few illustrative examples (for a broad overview see Cammen *et al.* (2016)).

Sex and age in an era of nonlethal sampling

The identification of the sex of individuals by observation alone is difficult, at best, for the vast majority of marine mammals. However, marine mammals have specific sex chromosomes in common with most other mammal species; males have one copy of the X- and one copy of the Y-chromosome, whereas females have two copies of the X-chromosome. Sex chromosome-specific DNA sequences can easily be targeted and amplified using PCR revealing the sex of the sampled individual (Baker *et al.*, 1991; Bérubé & Palsbøll, 1996b). Elucidating this very basic aspect has provided novel insights into group composition (Amos *et al.*, 1993; Pinela *et al.*, 2009), mating strategies (Clapham *et al.*, 1992; Best *et al.*, 2003; Flatz *et al.*, 2012) and sex-specific dispersal rates and patterns (Brown Gladden *et al.*, 1997; Baker *et al.*, 2013).

The age of individuals represents another elusive characteristic in most marine mammal species. Pinnipeds and toothed whales may be aged from the number of dentine layers in their teeth (Hohn *et al.*, 1989; Murphy *et al.*, 2012). However, collecting teeth for aging is highly invasive and, when possible, requires a high degree of effort. In many species, in particular cetaceans, capturing individuals for tooth extraction is not feasible. In long-term studies based upon individual photo-identification (or tagging), individuals identified as pups/calves will be of known age. Such long-term individual-based studies require a nontrivial and sustained effort to accumulate sufficient observations and are limited to relatively few uniquely studied populations (e.g., Northeast Pacific killer whales, *Orcinus orca*; Gulf of Maine humpback whales, *Megaptera novaeangliae*; and Bird Island Antarctic fur seals, *Arctocephalus gazella*). Two kinds of genetic aging assays have been applied to cetaceans. Initial attempts were based upon telomeres (Olsen *et al.*, 2012; Olsen *et al.*, 2014), which, in humans and other vertebrates, decrease in average length during an individual's lifespan. This approach performs poorly in mysticetes despite some age-related telomere changes. In addition, telomere shortening likely reflects biological age but not necessarily chronological age, and hence telomeres are probably better suited to assess cumulative stress exposure of individuals rather than age *per se* (Jarman *et al.*, 2015).

Another “short-term” change in the DNA of an individual is so-called epigenetic modifications. Many epigenetic modifications appear to correlate consistently with chronological age in humans. Polanowski and colleagues (2014) characterized several such age-related human epigenetic markers in humpback whales of known age. The degree of methylation at three of these epigenetic markers were found to correlate tightly with age in humpback whales (Polanowski *et al.*,

2014b). Unfortunately, the same epigenetic markers did not appear to correlate with age universally across a wider range of species (Polanowski *et al.*, 2014b), likely implying that the specific age-correlated epigenetic makers must be identified specifically in each marine mammal species. However, the more recent genome-wide epigenetic screening methods should make it relatively straight forward to identify species-specific age-correlated epigenetic markers (Bossdorf *et al.*, 2008). The age distribution within and among conspecific populations differs substantially between shrinking, expanding and recovering populations. Accordingly, reliable and cost efficient ageing methods that are applicable to non-invasive samples, will facilitate the assessment of the current population status and recent history, which would make epigenetic aging a key tool to inform and guide the management and conservation of marine mammal populations (Polanowski *et al.*, 2014b; Jarman *et al.*, 2015).

Identification of individuals and their close relatives

Genetic methods have also been applied to identify individuals and pairs of closely related individuals which can be utilized to gain insights into many key ecological and conservation aspects such as mating strategies and kin selection. Smaller marine mammal species can be captured and fitted with human-made tags or branded, but such an approach is effort intensive and infeasible in most marine mammal species. In particular, attaching man-made tags for long periods of time has proven challenging in large mysticetes. Alternatively, individuals may be identified from their permanent natural markings; an approach that has been applied with great success in species that differ in natural markings among individuals in species such as grey seals, *Halichoerus grypus* (Hastings *et al.*, 2012), killer whales (Würsig & Jefferson, 1990b) and humpback whales (Katona *et al.*, 1979). The genetic equivalent to unique individual markings is so-called genetic fingerprinting, which has a long tradition in human forensics. Individual identification, using genetic tagging, has been applied to many natural populations of marine mammals (e.g., Hoelzel & Amos, 1988; Palsbøll, 1999; Hoffman *et al.*, 2006), in a few cases across entire ocean basins (Baker *et al.*, 2013; Palsbøll *et al.*, 1997). Common applications of individual identification in natural population are the estimation of abundance using capture-mark-recapture methods as well as assessing connectivity across the seascape. One example of the latter was the genetic (and photographic) re-identification of a female humpback whale between off western Africa (off Gabon) in the South Atlantic, which was identified the following year off Madagascar in the Indian Ocean (Pomilla & Rosenbaum, 2005), representing a trans-oceanic migration event. Several large ocean-wide studies based upon the analysis of thousands of skin biopsies have utilized genetic tagging both to estimate abundance and to map seascape use, in particular in humpback whales (Palsbøll *et al.*, 1997a).

The identification of closely related individuals enables insights into mating strategies, i.e., by paternity analysis (e.g., Krützen *et al.*, 2004; Wiszniewski *et al.*, 2012), pod structure (Pilot *et al.*, 2010) and even abundance (Palsbøll *et al.*, 2010). In general, it is assumed that most species evolve towards behaviors that prevent excessive inbreeding, which may lead to a reduction in average population fitness due to inbreeding depression. Long-term studies of several odontocete species, e.g., killer whales and bottlenose dolphins, *Tursiops truncatus* (Amos *et al.*, 1993; Gero *et al.*, 2005; Pilot *et al.*, 2010) revealed that most individuals, including mature males, stay within their natal pod their entire life. This kind of social structure could lead to excessive inbreeding if mature (and presumably closely related) individuals from the same pod were to mate with each other. Amos and colleagues (1991; 1993) investigated this specific aspect in long-finned pilot whales, *Globicephala melas*. Employing genetic estimates of relatedness among aged (using the dentine layers in sectioned teeth) individuals from entire pods the authors affirmed that all individuals in a single pod were essentially part of the same extended matrilineal family (Amos *et al.*, 1993). In contrast, using paternity exclusion, none of the mature males in a pod sired other pod members (Amos *et al.*, 1991). In addition, it appeared that pod members that were part of the same age cohort had been sired by a few, but closely related, males. Amos and colleagues (1993) concluded that the estrous females in a pod probably mate with a few males that were “visiting” from another pod. The males subsequently returned to their natal pod after mating is completed. Since such visiting males would be related (i.e., part of the same maternal pod), the findings would explain why calves of the same age cohort appeared to be sired by a few, closely related males. Later work in other species with similar pronounced matrilineal pod structure (e.g., killer whales) have yielded similar findings, i.e., that pods appear to represent extended matrilineal families, but mating appear to take place between individuals from different pods (Hoelzel *et al.*, 1998). Several studies, particularly in pinnipeds, have suggested mate preference for individuals that are genetically diverse and/or dissimilar (Amos *et al.*, 2001; Hoffman *et al.*, 2007). All the above studies have reported findings that are consistent with the evolution of behaviors that presumably maximize outbreeding.

Estimating current and past abundance

Given the intensive human overexploitation of many marine mammal species, it is perhaps not surprising that several studies have attempted to infer historic abundance and the rate of decline from genetic data. The degree and distribution of genetic variation within and among con-specific individuals is a product of past (and current) population sizes and migration rates, although such inferences require a number of highly simplistic assumptions (Palsbøll *et al.*, 2013). In general, such genetic assessments of historic abundance in large whales (Roman & Palumbi, 2003; Alter *et al.*,

2007; Ruegg *et al.*, 2013) and pinnipeds (Hoffman *et al.*, 2011) have been consistent with the notion of a recent drastic decline due to human exploitation. In some cases, the historic abundance inferred from the amount of genetic variation in contemporary populations has been much higher than indicated by other approaches (e.g., demographic modeling). The discrepancies in inferred historic abundances may be due to a wide range of differences (and violations) of assumptions underlying each estimation approach. Perhaps the most common source of discrepancies lies with the time “point” that an estimate of historic abundance applies to, which for most genetic assessment is very wide, likely representing a mean across many hundreds to thousands of generations (depending on the approach used). Since there is a considerable temporal lag between a demographic change, and the corresponding change in genetic diversity, it is often challenging to assign a genetic-based historical abundance estimate to a specific, narrow timeframe (e.g., just prior to the onset of human exploitation). For similar reasons, it is also difficult to detect a very recent demographic change from the current degree and nature of genetic variation; unless the decline was very large (Fontaine *et al.*, 2012).

Population genetic structure and units of conservation

One key application of genetic analyses is for management and conservation, especially to delineate populations or management units (Dizon *et al.*, 1992; Moritz, 1994). In other words, many genetic assessments have been aimed toward detecting a spatial and/or temporal heterogeneous structure in the genetic variation within species. Typically, heterogeneity in genetic variation among con-specific individuals across the seascape is inferred as evidence for reduced dispersal and consequently population substructuring, where each homogenous set of samples often are equated with a management unit (Dizon *et al.*, 1992; Moritz, 1994). The identification of management units is important in order to direct monitoring efforts at the proper temporal and spatial scale, thereby hopefully facilitating the early detection of possible endangerment and facilitate local recovery. The application of genetics and genomics toward delineating populations and/or management units has a long tradition in marine mammals (Banguera-Hinestroza *et al.*, 2002; Clapham *et al.*, 2008; Bilgmann *et al.*, 2014). However, translating what are essentially abstract population genetic concepts and entities into real-life ecological processes and management units is far from trivial (Waples & Gaggiotti, 2006; Palsbøll *et al.*, 2007). These applications of genetics and genomics, also outside marine mammals, are currently undergoing substantial debate and revision in order to develop concepts and analytical approaches where the underlying assumptions and estimates obtained from genetic and genomic data are better aligned with the general time frames and effect sizes of relevance to ecology and conservation.

The recent application of genomics to the study of marine mammals

Most of the examples above were based on studies that utilized relatively few genetic markers due to experimental restrictions in the methodologies employed available at the time. The development of massive parallel DNA sequencing methods during the last 1-2 decades has enabled the sequencing of entire genomes, i.e., the generation of several orders of magnitude more data. The vast amounts of unbiased, genome-wide data should, in principle, make inferences more robust to potential biases arising from a small, nonrepresentative “sample” of a genome (Hoffman *et al.*, 2014). The complete genome has been published for several marine mammal species already (Zhou *et al.*, 2013; Liu *et al.*, 2014; Yim *et al.*, 2014; Foote *et al.*, 2015; Keane *et al.*, 2015; Humble *et al.*, 2016), and many more are underway, making marine mammals unique among mammals in terms of available, well-covered genome sequences. These genomic data have facilitated novel and detailed insights into the evolution and ecology of marine mammals (Liu *et al.*, 2014; Tsagkogeorga *et al.*, 2015).

Although genomic analyses have been applied to several different species of marine mammals, the killer whale is the species in which most genomic data have been generated and analyzed so far (Foote *et al.*, 2016). Accordingly, the work in killer whales provides a nice illustration of the kind of inferences that may be drawn from genomic-level data. Most of the works in killer whales have been aimed at the timing and adaptations resulting in divergence into different ecotypes in this species. The division of killer whales into different ecotypes is largely based upon dietary differences (Ford *et al.*, 1998). Some killer whale ecotypes prey exclusively upon fish, while others prey mainly on pinnipeds and/or large whales. In some cases, different ecotypes are morphologically distinguishable, such as in the Antarctic. Genetic studies have focused on whether the same ecotype arose independently in different ocean basins and whether the evolution into different ecotypes coincided with specific adaptations, i.e., as a consequences of the difference in diet (Ford *et al.*, 1998; Hoelzel *et al.*, 2007; Foote *et al.*, 2011b). These studies have utilized the entire gamut of genetic data ranging from genotypes collected at a dozen microsatellite loci and mitochondrial control region DNA sequences (Hoelzel *et al.*, 1998; Chivers *et al.*, 2007), complete mitochondrial genome sequences (Morin *et al.*, 2010; Foote *et al.*, 2011b) as well as complete and reduced genome data (Moura *et al.*, 2014; Foote *et al.*, 2016). The first kind of “genomic” studies in killer whales were based upon sequencing the complete mitochondrial genome (~16,500 nucleotides), as opposed to the 300 - 500 nucleotides of the mitochondrial control region, which is common to most marine mammal genetic studies. Morin and colleagues (2010) presented an analysis of 139 complete mitochondrial genome sequences from different ecotypes in multiple the

North Pacific, Antarctic as well as the North Atlantic. These authors concluded that the mammal and fish eating ecotypes had arisen multiple times and independently within each ocean basin. Subsequent analyses based upon reduced representation genome data (Moura *et al.*, 2015), as well as an analysis of 50 complete, but low coverage, killer whale genome sequences (Foote *et al.*, 2016), essentially confirmed the results inferred from previous data albeit with higher levels of precision as well insights into the dynamics of past demographic changes. Foote and colleagues (2016) concluded that the radiations of killer whales, leading to the present ecotypes, occurred recent and during the last 250,000 years. The founding populations went through severe bottlenecks, presumably due to a small number of individuals that became the ancestors of novel ecotypes in new habitats. These initial small populations diverged rapidly genetically, presumably due to a combination of random effect (caused by the small ancestral populations), cultural cohesion of ecotypes and adaptations due to diet specialization (Foote *et al.*, 2016). The analyses included an assessment of genes subjected to an above-genome-average rate of changes in key enzymes between mammal and fish eating ecotypes. In the two mammal eating ecotypes, the candidate enzyme-coding genes and processes identified as under positive selection were associated with the regulation of methionine metabolism. The observed changes in the inferred enzyme structure were interpreted by the authors as selection for coping with infrequent rich sources of dietary methionine (i.e., from mammal predation), which, in turn, would result in an additional selective pressure on methionine metabolism compared to ecotypes that feed more regularly, i.e., upon fish (Foote *et al.*, 2016). Other genes under apparent selection were involved in lipid metabolism presumably due to preying on marine mammals which have substantial lipid stores. A similar finding was previously reported in an analysis of polar bear genome sequences (Liu *et al.*, 2014), suggesting some degree of convergent adaptation driven by a diet mainly based upon marine mammals.

Convergent evolution among marine mammal lineages

The question of convergent evolution among marine mammal lineages (and even among echolocating mammals) has also been investigated by comparative genomics, in a manner similar to the killer whale analyses (Foote *et al.*, 2015). The three main lineages of marine mammals, i.e., the cetaceans, pinnipeds and sirenians, each represents an independent “return” by a terrestrial mammal lineage into the marine environment resulting a triplicated, evolutionary scale experiment of convergent evolution in mammals when subjected to a marine existence. Foote and coworkers (2015) sequenced and assembled the genome sequence from a killer whale, a bottlenose dolphin, a walrus (*Odobenus rosmarus*) and a West Indian manatee (*Trichechus manatus latirostris*). A comparative analysis of the protein-coding regions in these four genomes with terrestrial mammal

genomes revealed a number of changes in the three marine mammal genomes that displayed signs consistent with positive selection. Approximately 1% of these changes were identified in two or more marine mammal lineages, suggesting convergent evolution among the three marine mammal lineages. It is worth noting that the analysis was conservative, i.e., very stringent in terms of what constituted a so-called convergent mutation. The inferred gene function of the regions that appeared subject to convergent positive selection included genes involved in bone formation, cardiac muscle development and blood coagulation; all traits which possibly could be important in terms of adapting to diving and a denser medium (i.e., water compared to air). An interesting outcome of the study was the observation that the rate of convergent positive selected substitutions was higher among the terrestrial mammal genomes relative to the rates observed among the marine mammal genomes.

Our last example, in this overview, will be of the apparent convergent evolution of echolocation in two very diverse and taxonomically divergent groups of mammals: bats and odontocetes. A fundamental question is whether such a complex trait could have evolved in a convergent manner due to similar selective pressures, i.e., a similar rationale for convergent evolution across different lineages of marine mammals as described above. In both bats and odontocetes some species have evolved echolocation, which led Parker and colleagues (2013) to compare 2,326 coding DNA sequences across 22 mammal species, including bats and odontocetes (the common bottlenose dolphin). The authors found strong support for convergence in the same genes and direction among echolocating bats and the dolphin, and to the exclusion of nonecholocating bats. The genes inferred as subjected to (negative and positive) selection included numerous genes linked to hearing or deafness, which were presumed to play a role in the evolution of echolocation. In addition, genes linked to vision also showed signs of convergent evolution. In general, the degree of convergence in many sensory genes was found to correlate with the inferred strength of natural selection.

A final note of caution

The examples above have illustrated several uses of genetic and genomic analyses to elucidate the past and present status and evolution of marine mammals, both in terms of selection and subsequent adaptation as well as demographic changes in relation to diversification, climate change and human exploitation. However, we would like to end on a note of caution. Genetic (including genomic) analysis has had a wide and fundamental effect upon the entire field of biology and medicine. The power of modern DNA-based analyses, coupled with genomic scale data and recently *in silico* inferences about gene function, is truly impressive and has permeated biology and medicine over just a few decades. In the vast majority of studies, the reported results essentially

comprised of statistical significant correlations/signals detected using models of inference that are based upon highly simplistic assumptions about how populations, species and genes are structured and evolve. This is particularly true for difficult-to-study, nonmodel organisms such as marine mammals. The result is that most of the conclusions are more appropriately labeled as hypotheses, i.e., one (of several) possible explanations that are consistent with the findings of the particular study. Consequently, these hypotheses should ideally be subjected to subsequent rigorous and specific tests before the proposed explanation of the observed results is deemed to be correct. In many cases, such additional testing will require other experiments (e.g., controlled breeding or common garden experiments), which are difficult or impossible to conduct in marine mammals. In several cases, new experimental or analytical methods have significantly changed previous findings of earlier studies, which at the time of publication represented the state-of-the-art. There are several notable examples, which were published in leading journals at the time. One older example was the finding that sperm whales were more closely related to baleen whales than other odontocetes (implying that baleens and not teeth is the ancestral character) which has since been proven incorrect (Cerchio & Tucker, 1998; Nikaido *et al.*, 2001). Another example is the estimation of prewhaling abundance of North Atlantic humpback whales at 265,000 humpback whales (Roman & Palumbi, 2003), which was vastly above the historic abundance at 25,000–40,000 inferred using other data sources and approaches. Subsequent data analyses and adjustments of mutation rates, by the same authors, yielded estimates at less than half the original estimate (Alter & Palumbi, 2009; Ruegg *et al.*, 2013).

The convergent evolution of echolocation among echolocating bats and the common bottlenose dolphin has been questioned by Thomas and Hahn (2015). Thomas and Hahn's (2015) reanalysis revealed no excess convergence between echolocating bats and the bottlenose dolphin; they found that the degree of differentiation between these diverse taxa were well within expectations under a model of zero evolutionary convergence.

In closing, the field of genetics (including genomics) has contributed substantially to our understanding and conservation of marine mammals. The extremely rapid advances of the field of genetics in general, and of genomics in particular, have brought several new and exciting hypotheses forward in terms of the ecology and evolution of marine mammals. However, living at the “cutting edge”, scientifically and experimentally, harbors the risk of prematurity- and only through the test of time with complementary data and analyses will we eventually be able to establish which of the new and exciting hypotheses emerging from the recent genomic studies are confirmed.

