ANNUAL REVIEWS

Annual Review of Ecology, Evolution, and Systematics

A Genetic Perspective on Cetacean Evolution

Andrea A. Cabrera,[1,2] Martine Bérubé,[1,3]
Xênia M. Lopes,[1] Marie Louis,[2] Tom Oosting,[1,4]
Alba Rey-Iglesia,[2] Vania E. Rivera-León,[1]
Dóra Székely,[2] Eline D. Lorenzen,[2,*]
and Per J. Palsbøll[1,3,*]

[1]Marine Evolution and Conservation, Groningen Institute for Evolutionary Life Sciences, University of Groningen, 9747 AG Groningen, The Netherlands

[2]GLOBE Institute, University of Copenhagen, 1350 Copenhagen K, Denmark; email: andrea_ca_gt@yahoo.com

[3]Center for Coastal Studies, Provincetown, Massachusetts 02657, USA

[4]School of Biological Sciences, Victoria University of Wellington, Wellington 6140, New Zealand

Annu. Rev. Ecol. Evol. Syst. 2021. 52:131–51

First published as a Review in Advance on August 20, 2021

The *Annual Review of Ecology, Evolution, and Systematics* is online at ecolsys.annualreviews.org

https://doi.org/10.1146/annurev-ecolsys-012021-105003

*These authors contributed equally to this article

Keywords

aquatic environment, climate change, dolphins, genetics, molecular methods, whales

Abstract

Studies of cetacean evolution using genetics and other biomolecules have come a long way—from the use of allozymes and short sequences of mitochondrial or nuclear DNA to the assembly of full nuclear genomes and characterization of proteins and lipids. Cetacean research has also advanced from using only contemporary samples to analyzing samples dating back thousands of years, and to retrieving data from indirect environmental sources, including water or sediments. Combined, these studies have profoundly deepened our understanding of the origin of cetaceans; their adaptation and speciation processes; and of the past population change, migration, and admixture events that gave rise to the diversity of cetaceans found today.

1. INTRODUCTION

Secondary marine organisms: organisms of terrestrial ancestry that have recolonized marine habitats

The evolution of Cetacea (whales, dolphins, and porpoises) from an ancestral terrestrial mammal to an obligate aquatic lineage, followed by radiation into a diverse array of habitats, provides an excellent opportunity to study evolution and adaptation. Cetacea comprises one of five extant terrestrial mammal lineages (cetaceans, pinnipeds, sirenians, sea otters, and polar bears) that returned to an aquatic environment (O'Leary & Uhen 1999).

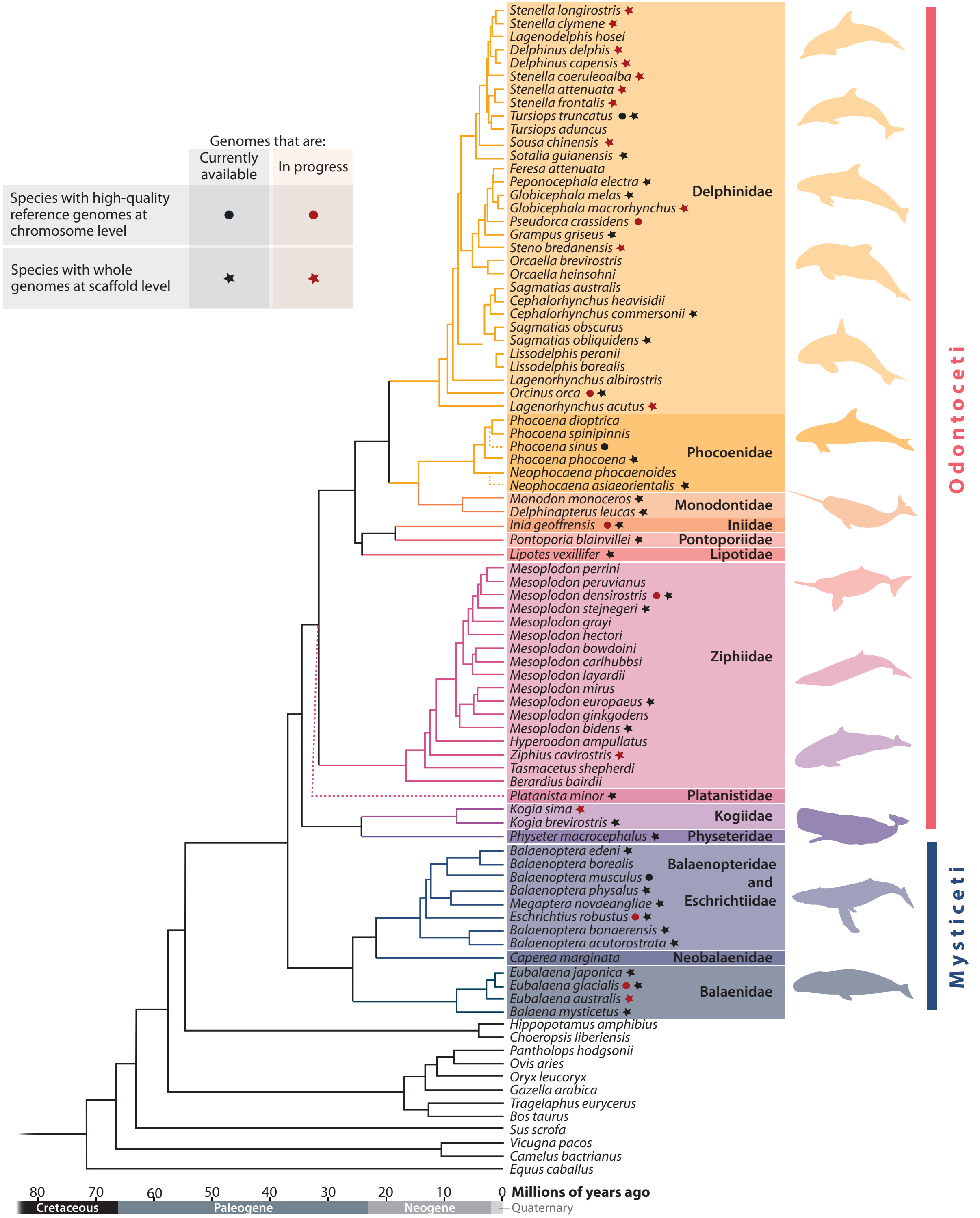
Cetacean evolution is among the best-characterized morphological and ecological transitions from a terrestrial to an aquatic existence. Cetaceans likely originated ~52.5 million years ago (Mya) during the early Eocene (Uhen 2010). The Archaeoceti, which inhabited coastal waters, represent the earliest group of mammals that evolved into the modern cetaceans (Thewissen & Williams 2002). The specific driver(s) that facilitated the transition from a terrestrial to an aquatic existence remains unresolved; current hypotheses include increased availability of marine prey, competition, and physical stressors, particularly during glacial cycles (e.g., Lipps & Mitchell 1976, Proches 2001).

The first fossil record of a pelagic cetacean is dated to ~40 Mya and belongs to the Neoceti, or crown cetaceans (Buono et al. 2016), which subsequently diversified into a speciose group including 245 fossil and extant genera (Uhen 2007). The extant cetaceans are represented by two remaining Neoceti lineages (**Figure 1**), the Mysticeti (baleen whales) and the Odontoceti (toothed whales), comprising at least 89 species across 40 genera (Pompa et al. 2011). During their early radiation, cetaceans came to occupy a wide diversity of habitats, ranging from deep oceanic to shallow coastal or freshwater habitats and from the tropics to the polar regions (Pompa et al. 2011). Multiple processes have been proposed to underlie the diversification and distribution of species, including the appearance of geographic barriers associated with the opening or closure of seaways, restructuring of ocean currents, prey distribution, climatic changes, and historical processes of speciation and extinction (e.g., Fontaine et al. 2010, Fordyce 1980, Pastene et al. 2007, Steeman et al. 2009).

As secondary marine organisms, cetaceans provide an opportunity to study evolution and adaptation associated with aquatic transitions. The presence of many taxa enables the use of a multitaxon approach to assess the response among closely related lineages to similar environmental pressures (e.g., Meredith et al. 2013). Furthermore, the complex social organization and cooperative behavior exhibited by many cetaceans facilitate the study of social and mating systems and mechanisms associated with inclusive fitness or kin selection (e.g., Amos et al. 1993, Kopps et al. 2014). Although cetaceans are less abundant compared with other marine organisms, their large body mass makes them essential to the function and structure of marine ecosystems (Bowen 1997). Whales may enhance primary productivity by recycling iron and nitrogen in the upper pelagic phase via fecal plumes (Roman & McCarthy 2010). Their high mobility and their placement at the intermediate and top of the marine food web render them indicator species of the overall ecosystem state and, consequently, habitat changes (Braithwaite et al. 2015, Cabrera et al. 2018).

The study of cetaceans in their natural environment can be challenging and costly. Many cetacean species are highly mobile, with large and remote distributions (Kaschner et al. 2011), and long generation times (Pacifici et al. 2013), making demographic and evolutionary changes difficult to detect using field studies. Molecular genetic techniques provide opportunities to investigate such changes, employing DNA from small tissue samples collected directly or indirectly from living or deceased organisms.

During the last three decades, the study of cetacean evolution has undergone significant changes. Advances in molecular technologies and bioinformatic approaches have allowed us to generate and analyze a diverse range of data, from a single small fragment of the mitochondrial



(Caption appears on following page)

Figure 1 (*Figure appears on preceding page*)

Phylogenetic tree of cetaceans showing genomes available by late 2020. Relationships and branch lengths are based on molecular dating estimates from McGowen et al. (2020). Branches with dotted lines were not included in the original tree and represent approximate positions. Genomes that are currently available (*black*) or in progress (*red*) are shown; circles indicate species with high-quality reference genomes at chromosome level, and stars indicate whole genomes at scaffold level. Odontoceti silhouette illustrations by Andrea A. Cabrera. Mysticeti silhouettes based on illustrations by Ligia E. Arreola and Frédérique Lucas.

genome to nuclear genome-wide data. Here, we review the application of genetic/omics approaches to study cetacean evolution at various organizational, spatial, and temporal scales, ranging from macroevolutionary aspects, such as early radiation, to microevolutionary processes, such as temporal changes in intraspecific genetic diversity. We provide examples of molecular studies regarding adaptation, radiation, and speciation; population structure; and demographic history. In addition, we highlight new possibilities and future perspectives in the study of cetacean evolution.

2. GENETIC/OMICS APPROACHES AS A MEANS TO STUDY CETACEAN EVOLUTION

The advent of molecular genetic approaches during the early 1990s and omics methods post-2000 have spurred numerous molecular studies of cetaceans aiming at fundamental and applied questions in evolution, ecology, and behavior (see the sidebar titled Studying Cetacean Evolution: From Casual Observation to Molecular Methods). The increasing application of molecular methods in cetacean research (**Figure 2*a***) is evident in the proportion of publications including any of the keywords molecular*, genetic*, DNA*, and genom* in the Web of Science, which has increased from 1% to 14% since 1990. Although the use of traditional markers has remained relatively constant during the past 15 years, the availability of genomic data and bioinformatics tools is growing exponentially (**Figure 2*b***). This growth is reflected in the increasing

STUDYING CETACEAN EVOLUTION: FROM CASUAL OBSERVATION TO MOLECULAR METHODS

Cetacean studies began in prehistoric times with casual observations on beaches and offshore (Allen 2014). The first formal studies, reported in the twentieth century, focused mainly on anatomical descriptions (Howell 1930). After the 1960s, research expanded to different fields, including photo-identification, telemetry, acoustics, and molecular methods.

The first genetic studies on cetaceans were published in the 1970s following the development of allozyme electrophoretic methods (Árnason 1972). This development, along with the introduction of the polymerase chain reaction (Mullis & Faloona 1987), which facilitated DNA sequencing, enabled population genetic diversity, structure, and phylogeny to be assessed (e.g., Hoelzel 1991).

The discovery of highly variable tandem repeat loci (Tautz & Renz 1984) led to studies of demography, individual identification, and kinship (e.g., Amos et al. 1993). The combined application of nuclear and mitochondrial DNA markers enhanced our understanding of the processes driving population structure and migration (e.g., Bérubé et al. 1998).

The emergence of next-generation sequencing spurred the development of single-nucleotide polymorphism genotype assays (e.g., Polanowski et al. 2011), which were followed by reduced-representation genome sequencing and whole-genome (re)sequencing (e.g., Baird et al. 2008), as a means of genotyping a large number of loci across the genome.

a

Year	Publications		
	Cetaceans[1]	Cetaceans + molecular methods[2]	% Cetaceans + molecular methods
1970s	1,174	13	1.1
1980s	1,897	18	0.9
1990s	2,777	261	9.4
2000s	4,348	519	11.9
2010–2020[3]	9,498	1,280	13.5

[1] Search criteria: TI=(whale* or dolphin* or narwhal* or vaquita* or porpoise* or cetacea*) NOT TI=("whale shark")

[2] Search criteria: Idem + AND TS=(molecular* or genetic* or DNA* or genom*)

[3] 2020 partial year

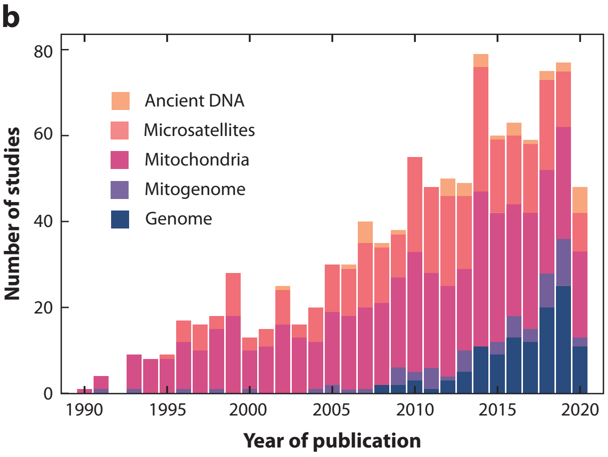


Figure 2

Publications on cetaceans and molecular methods. (*a*) Table with number of publications per time bin, including cetaceans and including both cetaceans and molecular methods, as well as the percentage of studies on cetaceans including molecular methods. (*b*) Number of publications on cetaceans from 1990 to November 22, 2020, categorized by primary genetic method: ancient genome data or paleogenomics (ancient DNA); tandem repeat loci (microsatellites); short fragments of mitochondrial genomes (mitochondria); entire mitochondrial genomes (mitogenome); or whole genomes, reduced-representation genomes, epigenomes, and transcriptomes (genome). Web of Science search criteria for each category: same as in panel *a* + AND TS = ("ancient DNA" or palaeogenetics or palaeogenomics or paleogenetics or paleogenomics or historic*); + AND TS = (microsatellite*); + AND TS = (mtDNA or mitochondrial); + AND TS = (mitogenome* or "mitochondrial genome*"); + AND TS = (genom* or radseq or gbs or epigen* or transcriptom*), respectively.

number of available cetacean genomes (**Figure 1**) and the number of publications employing genomics in cetaceans (**Figure 2*b***).

The application of molecular approaches to the study of cetaceans has been facilitated by the vast genetic and genomic resources acquired from other mammalian model species (e.g., humans, cows, and mice) whose genome structure and function are well characterized, thereby aiding inferences drawn from the observed data on gene function and effects of genetic variants. In this section, we review the current state of knowledge on cetaceans from molecular studies, with regard to selection and adaptation, speciation, intraspecific structure, and demographic history.

2.1. Selection and Adaptation to an Aquatic Existence

The transition from a terrestrial to an aquatic environment represents the defining transformation in cetacean evolution. The modern cetacean phenotype was likely a product of natural selection as an adaptation to an aquatic existence (Sun et al. 2013). This unique cetacean phenotype includes morphological, physiological, and behavioral traits for swimming, diving, thermoregulation, osmoregulation, and sensorial perception (e.g., Howell 1930).

Understanding the genetics of adaptation requires the ability to identify the genes and genetic pathways that underlie specific traits and to identify the variants subjected to natural selection (Barrett & Hoekstra 2011). Although this might prove difficult in cetaceans, as in other nonmodel organisms, available information from closely related model species, such as humans, has been used as an indirect means of obtaining insight into the function of specific genes, which is crucial for understanding adaptation to an aquatic existence. In the following subsections, we present some examples of what we have learned so far from employing genetics to study adaptation.

2.1.1. Changes that affect phenotype: gene function loss. Loss of gene function appears to be a common phenomenon in cetacean evolution (e.g., Chen et al. 2013, Meredith et al. 2011). Gene function losses have been associated with key phenotypic changes in cetaceans, such as loss of hair, functional teeth in baleen whales, taste, smell, and color vision (e.g., Chen et al. 2013; McGowen et al. 2008; Meredith et al. 2011, 2013). Kishida et al. (2007) estimated a significant increase in the proportion of nonfunctional olfactory receptor genes among cetaceans (68%) compared with the semiaquatic sea lion (37%) and terrestrial cow (17%), suggesting that loss of function of olfactory genes in cetaceans occurred in large numbers during the transition to a fully aquatic existence. Similarly, all modern cetaceans lost function in genes associated with sweet, umami, and bitter taste receptors prior to the divergence of Mysticeti and Odontoceti ∼35 Mya (Kishida et al. 2015).

Even though loss of function has been identified in the same genes across cetaceans, it could be due to either the same or different mutations. Examples of taxon-specific mutations in cetaceans include loss of function of the *SWS1* and *LWS* opsin genes, associated with color and light vision (Meredith et al. 2013), and of the *Hr* gene, associated with hair loss (Chen et al. 2013). For instance, Meredith et al. (2013) identified parallel mutations (i.e., the same mutations in independent lineages) of *SWS1* in Odontoceti and Mysticeti, as well as five independent inactivating mutations of *LWS* in deep-diving cetacean lineages.

The above examples illustrate the advantage of employing a multitaxon approach to assess the response among closely related lineages to changes in natural selection, such as convergent evolution. This multispecies approach can also be applied among marine mammal lineages. Genome-wide studies on convergent evolution among marine mammals (Foote et al. 2015, Zhou et al. 2015) found very few parallel substitutions in adaptive candidate genes. Instead, these genes exhibit distinct sequence changes in each taxonomic group (Zhou et al. 2015), suggesting that independent lineage-specific mutations led to convergent changes in coding genes associated with an aquatic existence.

2.1.2. Transition to an aquatic existence and subsequent radiations: two major events in cetacean evolution. The transition to the aquatic environment and the subsequent radiation of cetaceans have been associated with signatures of positive selection. Shen et al. (2012) studied the mechanisms and processes driving cetacean immune response to the aquatic transition. These authors found strong signatures of positive selection in the *TLR4* immune response–associated gene during two time periods. The first period, represented by the transition from a terrestrial to a semiaquatic existence and then to a fully aquatic existence, was detected along the joint lineage of hippopotamus and cetaceans and in the cetacean lineage, respectively. The signature of selection during this period likely reflected a response to different pathogens in land and water. The second period, detected along the lineage of oceanic dolphins, was associated with a rapid diversification and dispersal in waters around the world and with response to microbial pathogens from different environments (Shen et al. 2012). Similarly, evidence of positive selection acting on the branches that represent the origin of Odontoceti and the diversification of Delphinidae have been found in the *ASPM* gene, associated with brain size enlargement in cetaceans (Xu et al. 2012). Evidence of positive selection was not detected during the transition, suggesting that brain size enlargement occurred during the radiation and not immediately after the transition (Xu et al. 2012).

The transition to an aquatic existence also appears to have been characterized by changes in evolutionary rates. Chikina et al. (2016) analyzed hundreds of genes in 59 mammal genomes, including those of cetaceans, and identified hundreds of genes with an elevated number of mutations, which they inferred as evidence of accelerated evolutionary rates during the transition to an aquatic existence. These genes were associated with the control of functional adaptation to aquatic life in both loss of gene function and adaptive evolution, as observed in genes associated with skin and lung.

These examples show how the combination of single-gene and genome-wide molecular approaches with multispecies analyses can improve our understanding of the mechanisms and processes underlying the transition of cetaceans to a fully aquatic existence.

2.2. Radiation and Speciation

The transition to an aquatic existence was followed by radiation and subsequent diversification into a wide range of habitats. However, it has proven challenging to identify the factors and processes driving the initial radiation due to an incomplete spatial fossil record and the difficulty of identifying barriers in marine environments (Norris & Hull 2012, Steeman et al. 2009).

Several studies suggested that cetacean radiations coincided with large-scale oceanographic restructuring and tectonic events (e.g., Fordyce 1980, Steeman et al. 2009). Steeman et al. (2009) estimated a fossil-calibrated phylogeny of 87 extant cetacean species from mitochondrial and nuclear DNA sequences. They found that periods of pronounced physical restructuring of the oceans might have led to elevated rates of diversification of extant cetaceans during two main periods: (*a*) the opening of the Drake Passage and the establishment of the Antarctic Circumpolar Current ~35 Mya, which coincided with the initial radiation of Odontoceti, as suggested by Fordyce (1980), and (*b*) increased speciation rates ~13–4 Mya, coinciding with the diversification of Delphinidae, Phocoenidae, and Ziphiidae, which was characterized by the closure (Central American and Tethys) or restriction (Indo-Pacific) of seaways, an increase in ocean productivity, and intensification of ocean circulation (Steeman et al. 2009). These examples illustrate some key forces that have driven speciation, but what are the mechanisms involved in this process? In the following subsections, we present some examples of studies that have advanced our understanding of the mechanisms driving speciation.

Mitochondrial DNA: double-stranded DNA with uniparental (maternal) inheritance, located in the mitochondria

Nuclear DNA: double-stranded DNA with biparental inheritance, located in the cell nucleus of eukaryotic organisms

Gene flow: transfer of gene copies among different populations

Pods: stable, long-term associations with matrilineally related individuals

2.2.1. Speciation: not a simple process.

The formation of new species involves the divergence of lineages due to reproductive isolation in allopatry, sympatry, or parapatry. In allopatric speciation, geographic isolation is thought to have restricted gene flow within Mysticeti (e.g., Pastene et al. 2007) and within Odontoceti (e.g., Hare et al. 2002). Dolphin species of the genus *Lagenorhynchus*, the dusky dolphin (*Lagenorhynchus obscurus*) in the Southern Hemisphere and the Pacific white-sided dolphin (*Lagenorhynchus obliquidens*) in the Northern Hemisphere, are a clear example of allopatric speciation. These temperate species with antitropical distributions seemingly evolved from a common ancestor with a wide distribution over both hemispheres. The warm equatorial waters during the interglacial periods may have acted as a physical barrier, generating geographic and reproductive isolation facilitating the divergence of these species (Hare et al. 2002).

Killer whales (*Orcinus orca*) appeared to have evolved into reproductively isolated ecotypes in sympatry (Ford et al. 1998). Three killer whale ecotypes (transient, resident, and offshore) showing different diet specializations have overlapping distributions in the eastern North Pacific. It has been proposed that the maintenance of the three ecotypes was due to differences in prey preferences, which are culturally transmitted in matrilineal pods, in turn reducing gene flow among ecotypes (e.g., Ford et al. 1998, Hoelzel et al. 2007). However, alternative hypotheses such as divergence in allopatry and secondary contact have also been proposed (Foote et al. 2011). On the basis of genome-wide data from 50 individuals, Foote et al. (2016) suggested that differentiation into different ecotypes among killer whales worldwide was facilitated by a combination of factors. Behavioral plasticity likely facilitated the colonization of new ecological niches by a small number of individuals. These founder effects, followed by population expansion, likely promoted a rapid shift in allele frequencies between pairs of ecotypes due to genetic drift. Stable social transmission of behavior within the novel niches may have altered natural selection pressures and led to adaptive divergence in genes associated with diet, climate, and reproductive isolation (Foote et al. 2016).

Although most cetacean species arose several million years ago, studies have identified signatures of introgression between species (Árnason et al. 2018, Westbury et al. 2021), and contemporary hybrids between several cetacean species have been reported (Bérubé & Palsbøll 2018, Skovrind et al. 2019). Such hybrids occur between Mysticeti species, such as the fin (*Balaenoptera physalus*) and the blue whale (*Balaenoptera musculus*) (e.g., Árnason et al. 1991), as well as between the common (*Balaenoptera acutorostrata*) and the Antarctic minke whale (*Balaenoptera bonaerensis*) (Glover et al. 2013). Notably, pregnant female hybrids have been documented for both crosses (Árnason et al. 1991, Glover et al. 2013) and a putative second-generation fin and blue whale hybrid (Pampoulie et al. 2021). In Odontoceti, hybrids have been observed primarily between the common bottlenose dolphin (*Tursiops truncatus*) and other Delphinidae species (Bérubé & Palsbøll 2018). Hybrids have also been reported within Monodontidae and Iniidae, including a hybrid between narwhals (*Monodon monoceros*) and beluga whales (*Delphinapterus leucas*) (Skovrind et al. 2019) and between the two Amazonian freshwater dolphin species (*Inia boliviensis* and *Inia geoffrensis*) (Gravena et al. 2015).

Employing whole-genome sequences of nine species from three families (Delphinidae, Phocoenidae, and Monodontidae) of cetaceans, Westbury et al. (2021) analyzed the presence of post-divergence gene flow across their genomes and addressed the speciation processes that led to the diversity observed today. The authors found multiple ancestral gene flow events both within and among families, which continued for millions of years after initial divergence. These postdivergence gene flow events may also explain the presence of contemporaneous hybrids between several species (Westbury et al. 2021). The documented cases of hybridization, together with the identified signals of introgression among species, provide unique examples of the complex evolutionary history of cetaceans.

2.3. Demographic History

Evolution can be explained in terms of genetic changes that occur within populations, which can lead to the formation of new species, if the changes are large enough. In this section, we review how population genetic analyses of cetaceans have advanced our understanding of intraspecific population structure, size, and dispersal.

2.3.1. Genetic structuring and dispersal: result of multiple processes. Most species are spatially structured into populations that are genealogically linked. Although strong genealogical structure characterizes species with low dispersal capacities, cetaceans, which generally have high dispersal capacities and wide ranges, can be genetically structured even at relatively fine spatial scales (e.g., Bayas-Rea et al. 2018, Gravena et al. 2015). Dispersal in cetaceans can reflect either recurrent migrations, such as seasonal migrations between feeding and breeding grounds in large Mysticeti (Lockyer & Brown 1981), or dispersal from one population to another, resulting in gene flow. Here we present some of the driving forces of genetic structure and dispersal in cetaceans, and summarize some general patterns observed in both Mysticeti and Odontoceti.

Population genetic structure and dispersal are the result of contemporary and historical processes. For instance, glacial oscillations during the Pleistocene [2.5 Mya to 11.7 kya (thousand years ago)] had a large influence on the distribution, genetic structure, and connectivity of cetacean species and populations from both hemispheres (e.g., Jackson et al. 2014, Pérez-Alvarez et al. 2016). During glaciations, some cetacean populations became contracted and isolated, reducing gene flow and promoting genetic differentiation. Wang et al. (2008) proposed that the ancestral population of finless porpoises (genus *Neophocaena*) was divided by the emergence of a land bridge between Taiwan and China during the Last Glacial Maximum (LGM; 26–19 kya). However, further genomic analysis estimating an older divergence (50–40 kya) suggested that sea level changes before and after the LGM contributed to the divergence and isolation of the freshwater finless porpoise population (Zhou et al. 2018).

Physical processes that influence prey availability may have also played an important role in shaping genetic structure and dispersal. Changes in prey abundance caused by oceanographic transitions during the Pliocene and Pleistocene affected the distribution and, hence, the phylogeography of dusky dolphins (Harlin-Cognato et al. 2007). Current environmental variations such as cryptic or complex habitat breaks may drive the genetic structure of species with limited capacity to store energy, such as harbor porpoises (*Phocoena phocoena*). The genetic structure of harbor porpoises in the eastern North Atlantic has been correlated with limited dispersal abilities across less productive areas, where food resources are likely scarce (Fontaine et al. 2007).

A combination of environmental variation and behavior may drive and maintain contemporary genetic structure and dispersal in cetaceans. Many odontocetes are highly social species, and calves stay with their mothers for several years. They learn foraging techniques from their mother and conspecifics (e.g., Kopps et al. 2014). These social bonds and habitat specialization may promote philopatry, as well as genetic differentiation among individuals with different ecologies (e.g., Foote et al. 2016, Kopps et al. 2014). For instance, maternal transmission of tool use and habitat-specific feeding techniques may explain fine-scale geographic population structure of mitochondrial DNA in Indo-Pacific bottlenose dolphins (*Tursiops aduncus*) in Western Shark Bay, Australia (Kopps et al. 2014). Fidelity to feeding and mating grounds, which are transmitted by mothers to calves, likely contribute to genetic structure, particularly of mitochondrial DNA, as has been observed in right whales (*Eubalaena australis*) (Carroll et al. 2015) and humpback whales (Baker et al. 2013, Palsbøll et al. 1995).

A diverse array of mating strategies have been observed in cetaceans, which may influence patterns of genetic diversity and levels of gene flow among social groups or pods (Chesser 1991). For instance, in matrilineal long-finned pilot whales (*Globicephala melas*) and killer whale pods, both males and females stay in their natal pods. However, males do not generally mate with females inside their pods, which leads to gene flow among pods (Amos et al. 1993, Pilot et al. 2010). The opposite occurs in Atlantic spotted dolphins (*Stenella frontalis*), where mating occurs mainly within the social group or with females from the geographically closest group (Green et al. 2011), potentially reducing gene flow and hence population genetic differentiation (see the sidebar titled Sex, Individual Identification, and Kinships).

SEX, INDIVIDUAL IDENTIFICATION, AND KINSHIPS

Knowledge of sex, individual identity, and relationships among individuals is crucial for understanding cetacean evolution, including social systems and reproductive behavior. Sex determination of cetaceans using morphological traits is difficult in the wild, particularly in species without sexual dimorphism (Miller 2007). Small tissue samples from biopsies, sloughed skin, and feces can be used to determine the sex of individuals according to the presence of X and Y chromosomes (Palmer et al. 2019, Palsbøll et al. 1992). The same tissue samples can be employed to identify individuals through the use of genetic markers at multiple loci, known as genetic tagging (Palsbøll et al. 1997). Genetic tagging can also be used to identify related individuals by employing parentage and kinship analysis, as well as to estimate abundance or monitor cetacean populations (e.g., Palsbøll et al. 1997, Pilot et al. 2010).

Historical population size changes can also influence the rate of population genetic differentiation (Rogers & Harpending 1992). Huijser et al. (2018) analyzed the population genetic structure of North Atlantic and North Pacific sei whales (*Balaenoptera borealis*), and found that a recent historical population expansion might have contributed to the low levels of genetic differentiation, as genetic drift occurs more slowly in larger populations.

Effective population size: measures the stochastic rate of change in allele frequencies within an ideal population

In more solitary baleen whales, genetic structure is observed at relatively large spatial scales (Baker et al. 2013, Bérubé et al. 1998), and correlates with migratory behaviors, as detailed above. In contrast, fine-scale genetic structure is observed in many coastal delphinids with sympatric or adjacent ranges (e.g., Ansmann et al. 2012, Gariboldi et al. 2016), while offshore odontocetes living in more homogeneous habitats usually show lower levels of genetic differentiation, at ocean-basin scales (e.g., Thompson et al. 2012). This difference in the degree of fine- and large-scale genetic structure in species of coastal and offshore habitats, respectively, is also observed within the same species using both habitat types (e.g., bottlenose dolphins or Atlantic spotted dolphins; Louis et al. 2014, Viricel & Rosel 2014). Such observations reinforce the hypothesis that habitat or foraging specializations may play a role in driving dispersal and population structure.

2.3.2. Changes in population size: role of glacial–interglacial periods. The level of genetic variation within a population is dynamic and reflects the ever-changing balance between random and deterministic processes adding or removing genetic variation. Effective population size (Wright 1931) is a key determinant of genetic diversity. Several studies have employed estimates of genetic diversity to infer past, long-term changes in effective population size in order to assess the impact of the glacial cycles (e.g., Cabrera et al. 2018, Louis et al. 2020, Skovrind et al. 2021).

During glaciations, the polar regions expanded and the temperate and tropical regions contracted toward the equator. After the LGM, the population size of multiple temperate and polar cetacean species increased, most likely in response to increased prey availability and an expansion of suitable habitat as the sea ice coverage retreated and temperatures increased (Cabrera et al. 2018, Foote et al. 2013, Louis et al. 2020, Skovrind et al. 2021). Signatures of post-LGM population expansion in previously inhabitable high-latitude areas are evident in both Mysticeti and Odontoceti (e.g., Bérubé et al. 1998, Nykänen et al. 2019). Cabrera et al. (2018) detected a global increase in baleen whale abundance and their prey in the North Atlantic and Southern Oceans during the Holocene warming (12–7 kya). The timing of the increases in abundance coincided with the increasing temperatures and higher primary productivity. The observed increases in baleen whale abundance were oceanwide and continued for several millennia while global temperatures stabilized (Cabrera et al. 2018). The effects of Holocene warming also changed the environmental conditions in the Mediterranean Sea (~600 years ago), where the waters became warmer and oligotrophic (i.e., with poor nutrients). These conditions likely became unsuitable for harbor porpoises as populations began to fragment and collapse (Fontaine et al. 2010).

Industrial-scale whaling in the last few centuries severely depleted many cetacean populations, particularly the larger Mysticeti. Estimates of genetic diversity have been employed to infer prewhaling abundance in gray whales (*Eschrichtius robustus*) (Alter et al. 2007, Jackson et al. 2008) as well as in several North Atlantic baleen whales and Antarctic minke whales (Roman & Palumbi 2003). In all instances, the genetic diversity–based estimates of prewhaling abundance exceeded prewhaling abundances inferred from other types of data (Alter et al. 2007, Jackson et al. 2008). These discrepancies are probably due in part to the complexity of inferring recent abundance from genetic diversity; for example, most genetic diversity–based estimates represent long-term averages, rather than abundance estimates at a specific point in time (Palsbøll et al. 2013).

3. FUTURE PERSPECTIVES

The study of cetacean evolution with molecular methods has permitted us to address long-standing key questions and has provided new insights. Methodological developments have expanded the source material to include ancient and environmental samples, such as water or sediments (**Figure 3**). Similarly, constantly evolving molecular methods enable analyses of non-DNA biomolecules, such as transcriptomes, proteins, lipids, and metabolites (**Figure 4**). As a result, the number of available data and analytical tools are growing exponentially, a trend that will continue in the near future.

As genomic data and bioinformatics tools continue to grow, new challenges and paradigms arise (see the sidebar titled A Cautionary Note). A major challenge is balancing the capacity to generate vast quantities of data against the need to make rigorous biological interpretations (Tautz et al. 2010). Deciding which experimental and analytical approaches are best suited to address a specific

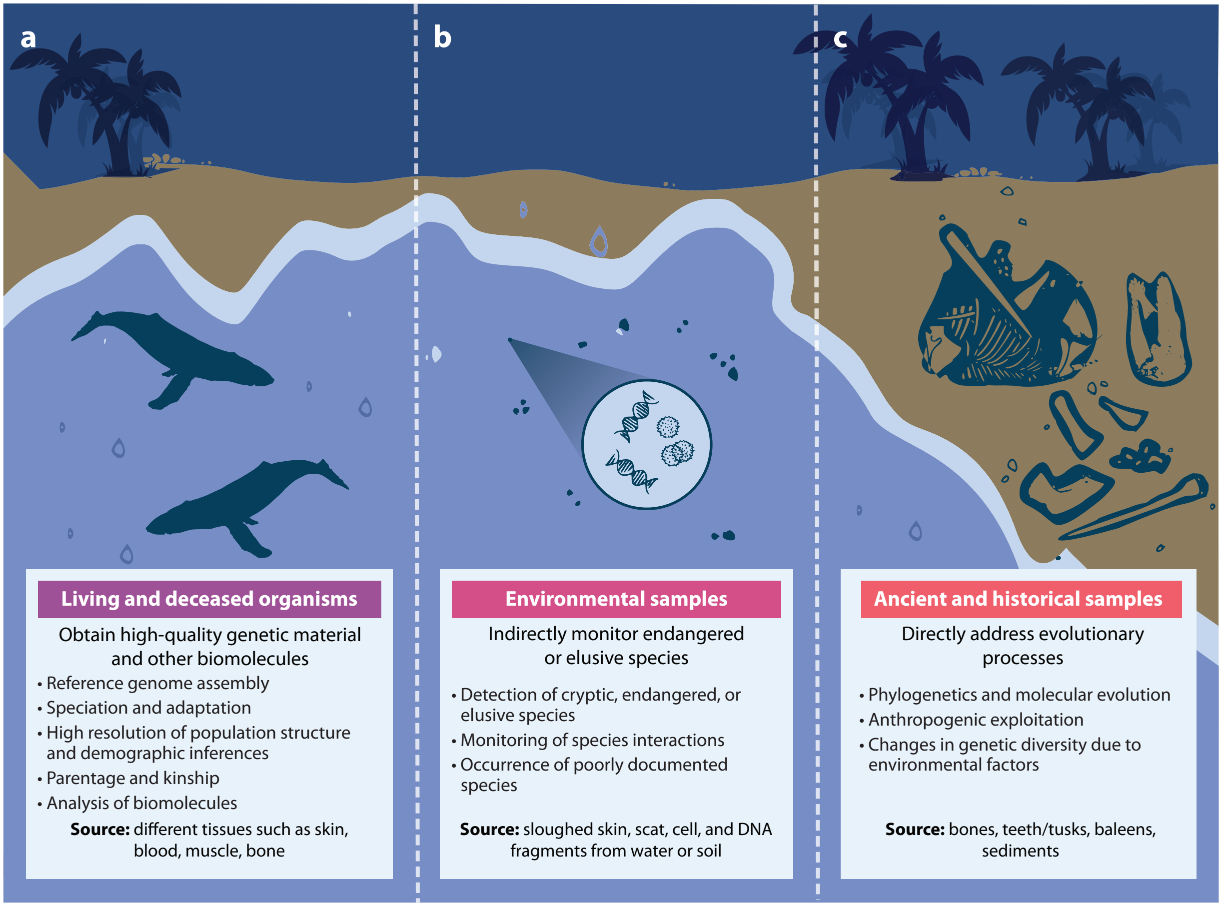


Figure 3

Advances in molecular methods enable new and deeper insights into cetacean evolution. Genetic material can be obtained from living and deceased organisms, environmental samples, or ancient and historical remains. Shown are examples of applications and biological questions that can be addressed using the three different sources of material including (*a*) living and deceased organisms; (*b*) environmental samples, and (*c*) ancient and historical samples. Examples of biological questions are shown as bullet points, and sources of the DNA material are listed at the bottom of each panel. Humpback whale silhouette based on an illustration by Ligia E. Arreola. The figure includes some elements inspired and modified from Swift et al. (2019).

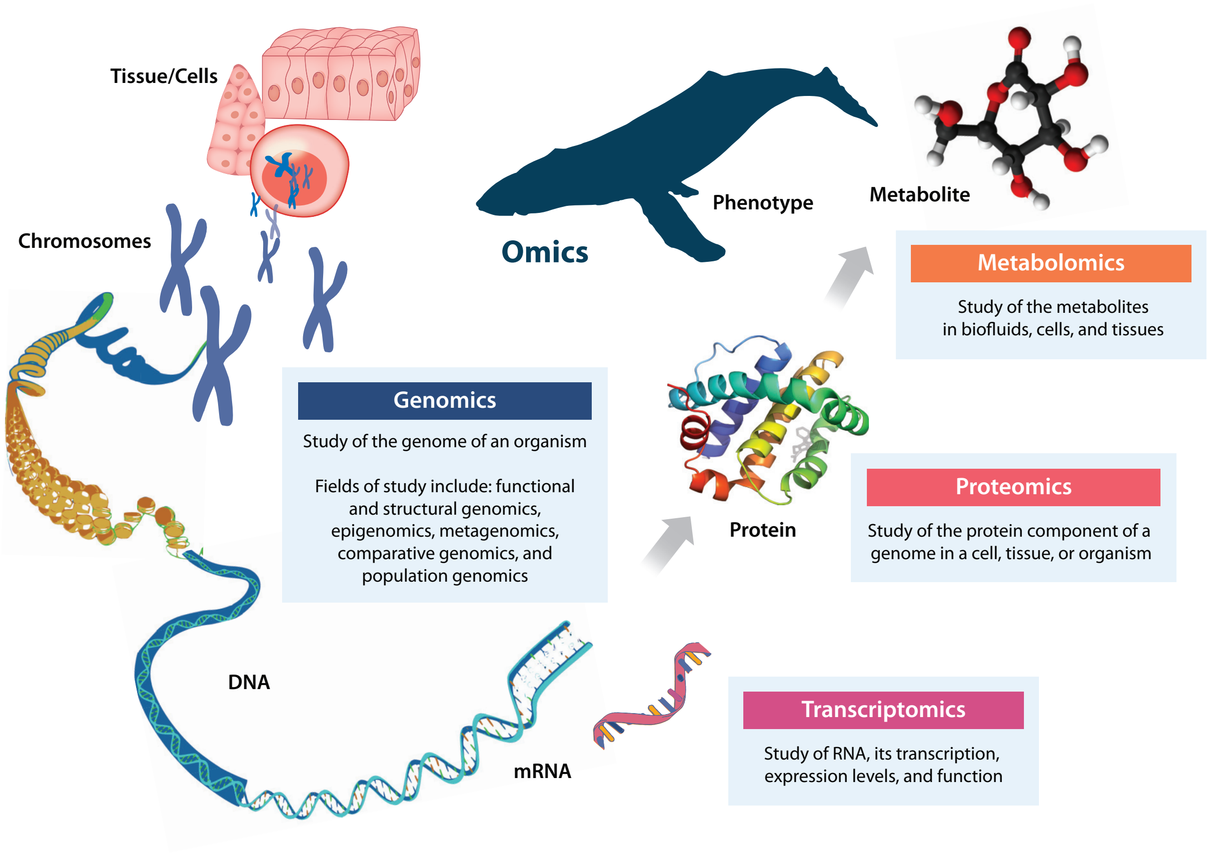


Figure 4

The omics approach is a holistic approach, which aims to understand the relationship between the genome of an organism and its phenotype. The omics approach involves acquiring a comprehensive, integrated understanding of biological processes through the identification of all involved molecules, rather than each individually. The figure shows the omics cascade and interaction across genomics, transcriptomics, proteomics, and metabolomics that can be translated into the structure, function, and dynamics of an organism. Humpback whale silhouette based on an illustration by Ligia E. Arreola.

research question is often the result of balancing resources (e.g., samples, funding), capabilities (e.g., skills and facilities), and novelty and scientific interest. For instance, Cabrera et al. (2019) analyzed the genealogy of North Atlantic, North Pacific, and Southern Hemisphere fin whales; the authors compared a previously published data set with entire mitochondrial genomes but small sample sizes ($n = 154$) against a data set with a larger sample size ($n = 1{,}676$) but only the control region of the mitochondrial genome. In the latter data set, the authors identified a North Atlantic lineage that was not detected in the former data set of longer sequences but fewer individuals and concluded that the tentative monophyly of the North Atlantic fin whales based on mitochondrial genomes was a result of small sample size.

Other issues that should be addressed relate to the quality and reproducibility of the data. These include the development of robust laboratory protocols that ensure minimal error rates and maximal reproducibility; standardized reporting of quality measures of the data, such as sequencing quality and accuracy of raw base calls; scripts to analyze the data; and detection and prevention of systematic error patterns.

A CAUTIONARY NOTE

Most genetic-based inferences in cetaceans are of an exploratory nature; consequently, the reported results should be viewed as hypotheses, rather than definitive proof. Scientific inquiries compare hypotheses to falsify those that are incorrect or assign relative probabilities to competing hypotheses. Genetic analyses of cetaceans (and most nonmodel species) compare observed data with analytical or in silico–generated null expectations derived from a parameterized statistical model. Such assessments necessitate specifications of multiple, often poorly known, nuisance (necessary but not the targeted) parameters to conduct the analysis (e.g., Haldane 1964, Wakeley 2004). Consequently, the observed scenario will differ from the null expectations and alternate hypotheses in many aspects in addition to the targeted effect, adding substantial uncertainty to the final outcome. For example, the prewhaling humpback whale abundance estimate went from 245,000 (Roman & Palumbi 2003) to 112,000 (Ruegg et al. 2012) due to an adjustment of a single nuisance parameter, the mutation rate.

Foote et al. (2016) detected genetic variants in Antarctic killer whales, which were inferred as selection on methionine metabolism. However, detecting an apparent correlation between a genotype and a specific phenotype or ecotype is not necessarily evidence of cause and effect or an increase in phenotype or ecotype fitness (Barrett & Hoekstra 2011). In rare cases, these common caveats can be partly mitigated by careful experimental consideration, such as pseudoreplication (i.e., assessing multiple populations or species, and regarding each as a replicate), additional substantiation with nongenetic data, and thorough assessment of the possible effects of realistic deviations from the assumptions made during analysis (e.g., Hoban et al. 2012).

3.1. Emerging Fields in Cetacean Research

The study of cetacean evolution is likely to break new ground in the near future. In this section, we outline a selected few research directions that are likely to make significant advances with respect to evolutionary inference.

3.1.1. Paleogenomics and paleoproteomics: a window into the past. Ancient DNA (aDNA) is the study of DNA extracted from historic or prehistoric biological materials, including bones, teeth, baleens, and sediments (e.g., Solazzo et al. 2017) (**Figure 3**). Analyses of aDNA may enable assessment of the genetic makeup of species at specific time points in the past, such as human–fauna or climate–fauna interactions. Although the field of aDNA emerged three decades ago (Higuchi et al. 1984), aDNA analyses of cetacean materials are still relatively uncommon. Analyses of aDNA enable species (Speller et al. 2016) and sex identification (Sinding et al. 2016) in materials in which molecular techniques developed for modern materials are infeasible and for which a morphological approach is nonviable. Such approaches have been applied to identify species in single cetacean bones (e.g., Rey-Iglesia et al. 2018) as well as in samples with multiple species, such as sediment cores or kitchen middens (e.g., Seersholm et al. 2016).

aDNA can be applied to compare extinct species and populations with their extant counterparts (e.g., Nichols et al. 2007), to gain further insight into the impact of past climate change (e.g., Alter et al. 2015, Foote et al. 2013), and to assess anthropogenic effects (e.g., Béland et al. 2020, Borge et al. 2007). Alter et al. (2015) combined modern and aDNA data from the Late Pleistocene to the Late Holocene with habitat modeling to analyze the demographic history of the extinct Atlantic and extant Pacific gray whale populations. These authors suggested that dispersal took place between the two populations during the Pleistocene and Early Holocene, when warming periods and sea level rise allowed the opening of the Bering Strait. Estimates of genetic diversity indicated that the Atlantic population was smaller than the Pacific one for an extended period of time and that its decline began prior to historical whaling (Alter et al. 2015).

Metabarcoding: method of species identification using a short section of DNA from a sample that contains more than one organism

Like most molecular approaches, aDNA analyses are also subject to limitations, such as temporal and spatial biases in the available fossil materials (Bernal-Casasola et al. 2016, Speller et al. 2016), difficulties in radiocarbon dating marine remains due to the marine reservoir effect (Ascough et al. 2005), and poor preservation (Carpenter et al. 2013). The pelagic habitat of many cetaceans greatly limits the fossil record compared with that of terrestrial species (Speller et al. 2016). There are only a few documented archeological sites and species with several fossil records (Alter et al. 2015, Foote et al. 2012). In addition, in cases of limited endogenous DNA, approaches that enrich the content of endogenous DNA may be required (e.g., Carpenter et al. 2013).

The study of ancient proteins or paleoproteomics is an emerging field that may open new venues for cetacean research. Proteins have a much greater postmortem longevity compared with that of DNA (Demarchi et al. 2016). Consequently, paleoproteomic analyses may yield data from very old material, providing a great opportunity to investigate systematics and macroevolutionary processes.

3.1.2. Environmental DNA: an indirect approach to detect rare and cryptic species. Environmental DNA (eDNA) is genetic material obtained from environmental samples (**Figure 4**) (Willerslev 2003). With the aid of genetic methods such as shotgun sequencing, metabarcoding, or design of species-specific primers, eDNA surveys can reveal the presence and habitat range of species (e.g., Székely et al. 2021, Valsecchi et al. 2020).

eDNA analyses can be employed to detect rare, elusive, cryptic, and endangered cetacean species (e.g., Ma et al. 2016). Studies based on eDNA can also be employed to monitor species communities when traditional, observation-based survey methods are challenging (e.g., Closek et al. 2019). For instance, Closek et al. (2019) analyzed the marine vertebrate community from water samples off the coast of California and detected an increase in marine mammal diversity from 2016 to 2017 that coincided with a change in oceanographic conditions.

Furthermore, eDNA metabarcoding of scat can provide insight into trophic interactions and prey preferences (e.g., Carroll et al. 2019, Jarman et al. 2002). Carroll et al. (2019) compared data from systematic zooplankton surveys with eDNA metabarcoding analyses of Bryde's whale (*Balaenoptera edeni*) scat and found no evidence of seasonal changes in Bryde's whale diet. These results indicated specific prey preferences in Bryde's whales (Carroll et al. 2019). Future studies on eDNA will likely increase our understanding not only of trophic interactions but also of the biodiversity interactions within an ecosystem, for instance, by connecting cetacean assemblages with microbial, invertebrate, and vertebrate assemblages.

Despite its promise, eDNA poses several challenges. eDNA is degraded and fragmented as a result of specific characteristics of the marine environment (e.g., salinity, pH, temperature, UV). Furthermore, environmental samples are dominated by the most abundant marine organisms, such as prokaryotes, and cetacean DNA fragments are lost needles in the haystack. However, an increased amount of target DNA has been retrieved from the "footprints" of diving cetaceans (Amos et al. 1992, Székely et al. 2021) that contain traces of sloughed skin. From the footprints of bowhead whales, Székely et al. (2021) were able to capture the most common mitochondrial control region haplotypes of the species. A possible though challenging use of eDNA will be to collect population genetic data (Adams et al. 2019, Sigsgaard et al. 2020), which would provide insight into the genetic composition of elusive cetaceans.

3.1.3. Omics: a holistic approach to understanding the structure, function, and dynamics of organisms. Omics is the study of biological molecules; the suffix -omics (**Figure 4**) includes analyses of the genome (genomics), transcriptome (RNA; transcriptomics), proteins (proteomics), and metabolites (metabolomics). The omics approach aims to understand the relationship between the genome of an organism and its phenotype (Fritsche-Neto & Borém 2014) by acquiring a comprehensive, integrated understanding of the biological processes through the identification of all involved molecules rather than each individually.

Metagenomics: study of genetic material recovered from environmental samples

Recent studies have shown how the use of different biomolecules can provide a more coherent approach to address certain evolutionary issues in cetaceans. These issues include analysis of adaptation using transcriptomics (Toren et al. 2020), organ or system function employing proteomics (Kershaw et al. 2018), and speciation and hybridization processes using comparative genomics (Westbury et al. 2021). They also include the study of coevolution between cetaceans and microbial communities using metagenomics (Li et al. 2019) and age estimation using epigenetics (Polanowski et al. 2014), among others. Although these single-omics studies alone will not be sufficient to address the complexity of organisms (i.e., structure, function, and dynamics), they provide insights into new possibilities.

The next step will be to describe and elucidate the evolution of cetaceans by employing a holistic approach that can be translated into the structure, function, and dynamics of an organism. Such holistic approaches include integrating analyses of different biomolecules with environmental and behavioral data in order to provide unprecedented insights into the evolution of cetaceans. As technologies and algorithms continue to improve, omics analyses are and will continue to be performed across organisms, individuals, cell types, and conditions, and eventually at the level of individual cells (Berger et al. 2013). Although this approach will encounter big challenges and it will take some time before it can easily be applied in cetaceans, it is likely to be especially fruitful in revealing the molecular landscapes underlying observed phenotypic variations.

DISCLOSURE STATEMENT

The authors are unaware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We apologize to all researchers whose work could not be appropriately cited owing to space limitations. Special thanks go to Yvonne Verkuil, Angeliki Paspati, and Katharina Fietz for their support and contributions to the introductory essay of A.A.C.'s PhD program, which was the inspiration for this review. The writing of this review was supported by the University of Groningen (to A.A.C., M.B., X.M.L., T.O., V.E.R.-L., and P.J.P.) and the University of Copenhagen (to A.A.C., M.L., A.R.-I., D.S., and E.D.L.). Funding was provided by a Rubicon-NWO 465 grant (project 019.183EN.005) to A.A.C. X.M.L. was funded by a Brazilian scholarship from CNPq (201709/2014-7), and V.E.R.-L. received funding from Consejo Nacional de Ciencia y Tecnología (CONACYT). D.S.Z. received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Actions (grant 801199). This review was also supported by a Carlsberg Foundation Distinguished Associate Professor Fellowship (grant CF16-0202) and the Villum Fonden Young Investigator program (grant 13151) to E.D.L.

LITERATURE CITED

Adams CIM, Knapp M, Gemmell NJ, Jeunen G-J, Bunce M, et al. 2019. Beyond biodiversity: Can environmental DNA (eDNA) cut it as a population genetics tool? *Genes* 10(3):192

Allen S. 2014. From exploitation to adoration: the historical and contemporary contexts of human–cetacean interactions. In *Whale-Watching: Sustainable Tourism and Ecological Management*, ed. J Higham, L Bejder, R Williams, pp. 31–47. Cambridge, UK: Cambridge Univ. Press

Alter SE, Meyer M, Post K, Czechowski P, Gravlund P, et al. 2015. Climate impacts on transocean dispersal and habitat in gray whales from the Pleistocene to 2100. *Mol. Ecol.* 24(7):1510–22

Alter SE, Rynes E, Palumbi SR. 2007. DNA evidence for historic population size and past ecosystem impacts of gray whales. *PNAS* 104(38):15162–67

Amos B, Schlötterer C, Tautz D. 1993. Social-structure of pilot whales revealed by analytical DNA profiling. *Science* 260(5108):670–72

Amos W, Whitehead H, Ferrari MJ, Glockner-Ferrari DA, Payne R, Gordon J. 1992. Restrictable DNA from sloughed cetacean skin; its potential for use in population analysis. *Mar. Mamm. Sci.* 8(3):275–83

Ansmann IC, Parra GJ, Lanyon JM, Seddon JM. 2012. Fine-scale genetic population structure in a mobile marine mammal: inshore bottlenose dolphins in Moreton Bay, Australia. *Mol. Ecol.* 21(18):4472–85

Árnason Ú. 1972. The role of chromosomal rearrangement in mammalian speciation with special reference to Cetacea and Pinnipedia. *Hered. Genet. Ark.* 70(1):113–18

Árnason Ú, Lammers F, Kumar V, Nilsson MA, Janke A. 2018. Whole-genome sequencing of the blue whale and other rorquals finds signatures for introgressive gene flow. *Sci. Adv.* 4(4):eaap9873

Árnason Ú, Spilliaert R, Pálsdóttir A, Árnason A. 1991. Molecular-identification of hybrids between the 2 largest whale species, the blue whale (*Balaenoptera musculus*) and the fin whale (*B. physalus*). *Hereditas* 115(2):183–89

Ascough P, Cook G, Dugmore A. 2005. Methodological approaches to determining the marine radiocarbon reservoir effect. *Prog. Phys. Geogr. Earth Environ.* 29(4):532–47

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, et al. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLOS ONE* 3(10):e3376

Baker CS, Steel D, Calambokidis J, Falcone E, González-Peral U, et al. 2013. Strong maternal fidelity and natal philopatry shape genetic structure in North Pacific humpback whales. *Mar. Ecol. Prog. Ser.* 494:291–306

Barrett RDH, Hoekstra HE. 2011. Molecular spandrels: tests of adaptation at the genetic level. *Nat. Rev. Genet.* 12(11):767–80

Bayas-Rea RLÁ, Félix F, Montufar R. 2018. Genetic divergence and fine scale population structure of the common bottlenose dolphin (*Tursiops truncatus*, Montagu) found in the Gulf of Guayaquil, Ecuador. *PeerJ* 6:e4589

Béland SL, Frasier BA, Darling JD, Frasier TR. 2020. Using pre- and postexploitation samples to assess the impact of commercial whaling on the genetic characteristics of eastern North Pacific gray and humpback whales and to compare methods used to infer historic demography. *Mar. Mamm. Sci.* 36(2):398–420

Berger B, Peng J, Singh M. 2013. Computational solutions for omics data. *Nat. Rev. Genet.* 14(5):333–46

Bernal-Casasola D, Gardeisen A, Morgenstern P, Horwitz LK, Piqués G, et al. 2016. Ancient whale exploitation in the Mediterranean: the archaeological record. *Antiquity* 90(352):914–27

Bérubé M, Aguilar A, Dendanto D, Larsen F, Di Sciara GN, et al. 1998. Population genetic structure of North Atlantic, Mediterranean Sea and Sea of Cortez fin whales, *Balaenoptera physalus* (Linnaeus 1758): analysis of mitochondrial and nuclear loci. *Mol. Ecol.* 7(5):585–99

Bérubé M, Palsbøll PJ. 2018. Hybridism. In *Encyclopedia of Marine Mammals*, ed. B Würsig, JGM Thewissen, KM Kovacs, pp. 496–501. Amsterdam: Elsevier

Borge T, Bachmann L, Bjørnstad G, Wiig Ø. 2007. Genetic variation in Holocene bowhead whales from Svalbard. *Mol. Ecol.* 16(11):2223–35

Bowen WD. 1997. Role of marine mammals in aquatic ecosystems. *Mar. Ecol. Prog. Ser.* 158:267–74

Braithwaite JE, Meeuwig JJ, Letessier TB, Jenner KCS, Brierley AS. 2015. From sea ice to blubber: linking whale condition to krill abundance using historical whaling records. *Polar Biol.* 38(8):1195–202

Buono MR, Fernández MS, Reguero MA, Marenssi SA, Santillana SN, Mörs T. 2016. Eocene basilosaurid whales from the La Meseta Formation, Marambio (Seymour) Island, Antarctica. *Ameghiniana* 53(3):296–315

Cabrera AA, Hoekendijk JPA, Aguilar A, Barco SG, Berrow S, et al. 2019. Fin whale (*Balaenoptera physalus*) mitogenomics: a cautionary tale of defining sub-species from mitochondrial sequence monophyly. *Mol. Phylogenet. Evol.* 135:86–97

Cabrera AA, Schall E, Bérubé M, Bachmann L, Berrow S, et al. 2018. Strong and lasting impacts of past global warming on baleen whale and prey abundance. bioRxiv 497388. **https://doi.org/10.1101/497388**

Carpenter ML, Buenrostro JD, Valdiosera C, Schroeder H, Allentoft ME, et al. 2013. Pulling out the 1%: whole-genome capture for the targeted enrichment of ancient DNA sequencing libraries. *Am. J. Hum. Genet.* 93(5):852–64

Carroll EL, Baker CS, Watson M, Alderman R, Bannister J, et al. 2015. Cultural traditions across a migratory network shape the genetic structure of southern right whales around Australia and New Zealand. *Sci. Rep.* 5:16182

Carroll EL, Gallego R, Sewell MA, Zeldis J, Ranjard L, et al. 2019. Multi-locus DNA metabarcoding of zooplankton communities and scat reveal trophic interactions of a generalist predator. *Sci. Rep.* 9:281

Chen Z, Wang Z, Xu S, Zhou K, Yang G. 2013. Characterization of hairless (*Hr*) and *FGF5* genes provides insights into the molecular basis of hair loss in cetaceans. *BMC Evol. Biol.* 13:34

Chesser RK. 1991. Influence of gene flow and breeding tactics on gene diversity within populations. *Genetics* 129(2):573–83

Chikina M, Robinson J, Clark NL. 2016. Hundreds of genes experienced convergent shifts in selective pressure in marine mammals. *Mol. Biol. Evol.* 33(9):2182–92

Closek CJ, Santora JA, Starks HA, Schroeder ID, Andruszkiewicz EA, et al. 2019. Marine vertebrate biodiversity and distribution within the Central California Current using environmental DNA (eDNA) metabarcoding and ecosystem surveys. *Front. Mar. Sci.* 6:732

Demarchi B, Hall S, Roncal-Herrero T, Freeman CL, Woolley J, et al. 2016. Protein sequences bound to mineral surfaces persist into deep time. *eLife* 5:e17092

Fontaine MC, Baird SJE, Piry S, Ray N, Tolley KA, et al. 2007. Rise of oceanographic barriers in continuous populations of a cetacean: the genetic structure of harbour porpoises in Old World waters. *BMC Biol.* 5:30

Fontaine MC, Tolley KA, Michaux JR, Birkun A, Ferreira M, et al. 2010. Genetic and historic evidence for climate-driven population fragmentation in a top cetacean predator: the harbour porpoises in European water. *PNAS* 277(1695):2829–37

Foote AD, Hofreiter M, Morin PA. 2012. Ancient DNA from marine mammals: studying long-lived species over ecological and evolutionary timescales. *Ann. Anat.* 194(1):112–20

Foote AD, Kaschner K, Schultze SE, Garilao C, Ho SYW, et al. 2013. Ancient DNA reveals that bowhead whale lineages survived Late Pleistocene climate change and habitat shifts. *Nat. Commun.* 4:1677

Foote AD, Liu Y, Thomas GWC, Vinař T, Alföldi J, et al. 2015. Convergent evolution of the genomes of marine mammals. *Nat. Genet.* 47(3):272–75

Foote AD, Vijay N, Ávila-Arcos MC, Baird RW, Durban JW, et al. 2016. Genome-culture coevolution promotes rapid divergence of killer whale ecotypes. *Nat. Commun.* 7:11693

Foote AD, Vilstrup JT, De Stephanis R, Verborgh P, Nielsen SCA, et al. 2011. Genetic differentiation among North Atlantic killer whale populations. *Mol. Ecol.* 20(3):629–41

Ford JKB, Ellis GM, Barrett-Lennard LG, Morton AB, Palm RS, Balcomb KC. 1998. Dietary specialization in two sympatric populations of killer whales (*Orcinus orca*) in coastal British Columbia and adjacent waters. *Can. J. Zool.* 76(8):1456–71

Fordyce RE. 1980. Whale evolution and Oligocene southern ocean environments. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 31(2–4):319–36

Fritsche-Neto R, Borém A. 2014. Omics: opening up the "black box" of the phenotype. In *Omics in Plant Breeding*, ed. A Borém, R Fritsche-Neto, pp. 1–11. Chichester, UK: Wiley

Gariboldi MC, Túnez JI, Failla M, Hevia M, Panebianco MV, et al. 2016. Patterns of population structure at microsatellite and mitochondrial DNA markers in the Franciscana dolphin (*Pontoporia blainvillei*). *Ecol. Evol.* 6(24):8764–76

Glover KA, Kanda N, Haug T, Pastene LA, Øien N, et al. 2013. Hybrids between common and Antarctic minke whales are fertile and can back-cross. *BMC Genet.* 14:25

Gravena W, da Silva VMF, da Silva MNF, Farias IP, Hrbek T. 2015. Living between rapids: genetic structure and hybridization in botos (Cetacea: Iniidae: *Inia* spp.) of the Madeira River, Brazil. *Biol. J. Linn. Soc.* 114(4):764–77

Green ML, Herzing DL, Baldwin JD. 2011. Reproductive success of male Atlantic spotted dolphins (*Stenella frontalis*) revealed by noninvasive genetic analysis of paternity. *Can. J. Zool.* 89(3):239–53

Haldane JBS. 1964. Defense of beanbag genetics. *Perspect. Biol. Med.* 7(3):343–59

Hare MP, Cipriano F, Palumbi SP. 2002. Genetic evidence on the demography of speciation in allopatric dolphin species. *Evolution* 56(4):804–6

Harlin-Cognato AD, Markowitz T, Wuersig B, Honeycutt RL. 2007. Multi-locus phylogeography of the dusky dolphin (*Lagenorhynchus obscurus*): passive dispersal via the west-wind drift or response to prey species and climate change? *BMC Evol. Biol.* 7:131

Higuchi R, Bowman B, Freiberger M, Ryder OA, Wilson AC. 1984. DNA sequences from the quagga, an extinct member of the horse family. *Nature* 312(5991):282–84

Hoban S, Bertorelle G, Gaggiotti OE. 2012. Computer simulations: tools for population and evolutionary genetics. *Nat. Rev. Genet.* 13(2):110–22

Hoelzel AR. 1991. *Genetic Ecology of Whales and Dolphins: Incorporating the Proceedings of the Workshop on the Genetic Analysis of Cetacean Populations*. Rep. Int. Whal.Comm., Spec. Issue 13. Cambridge, UK: Int. Whal. Comm.

Hoelzel AR, Hey J, Dahlheim ME, Nicholson C, Burkanov V, Black N. 2007. Evolution of population structure in a highly social top predator, the killer whale. *Mol. Biol. Evol.* 24(6):1407–15

Howell AB. 1930. *Aquatic Mammals: Their Adaptation to Life in the Water*. Springfield, IL: Thomas

Huijser LAE, Bérubé M, Cabrera AA, Prieto R, Silva MA, et al. 2018. Population structure of North Atlantic and North Pacific sei whales (*Balaenoptera borealis*) inferred from mitochondrial control region DNA sequences and microsatellite genotypes. *Conserv. Genet.* 19(4):1007–24

Jackson JA, Patenaude NJ, Carroll EL, Baker CS. 2008. How few whales were there after whaling? Inference from contemporary mtDNA diversity. *Mol. Ecol.* 17(1):236–51

Jackson JA, Steel DJ, Beerli P, Congdon BC, Olavarría C, et al. 2014. Global diversity and oceanic divergence of humpback whales (*Megaptera novaeangliae*). *Proc. R. Soc. B* 281:20133222

Jarman SN, Gales NJ, Tierney M, Gill PC, Elliott NG. 2002. A DNA-based method for identification of krill species and its application to analysing the diet of marine vertebrate predators. *Mol. Ecol.* 11(12):2679–90

Kaschner K, Tittensor DP, Ready J, Gerrodette T, Worm B. 2011. Current and future patterns of global marine mammal biodiversity. *PLOS ONE* 6(5):e19653

Kershaw JL, Botting CH, Brownlow A, Hall AJ. 2018. Not just fat: investigating the proteome of cetacean blubber tissue. *Conserv. Physiol.* 6(1):coy003

Kishida T, Kubota S, Shirayama Y, Fukami H. 2007. The olfactory receptor gene repertoires in secondary-adapted marine vertebrates: evidence for reduction of the functional proportions in cetaceans. *Biol. Lett.* 3(4):428–30

Kishida T, Thewissen JGM, Hayakawa T, Imai H, Agata K. 2015. Aquatic adaptation and the evolution of smell and taste in whales. *Zool. Lett.* 1:9

Kopps AM, Ackermann CY, Sherwin WB, Allen SJ, Bejder L, Krützen M. 2014. Cultural transmission of tool use combined with habitat specializations leads to fine-scale genetic structure in bottlenose dolphins. *Proc. Biol. Sci.* 281(1782):20133245

Li C, Tan X, Bai J, Xu Q, Liu S, et al. 2019. A survey of the sperm whale (*Physeter catodon*) commensal microbiome. *PeerJ* 7:e7257

Lipps JH, Mitchell E. 1976. Trophic model for the adaptive radiations and extinctions of pelagic marine mammals. *Paleobiology* 2(2):147–55

Lockyer CH, Brown JG. 1981. The migration of whales. In *Animal Migration*, ed. DJ Aidley, pp. 105–38. New York: Press Synd. Univ. Cambridge

Louis M, Skovrind M, Castruita JAS, Garilao C, Kaschner K, et al. 2020. Influence of past climate change on phylogeography and demographic history of narwhals, *Monodon monoceros*. *Proc. R. Soc. B* 2871925:20192964

Louis M, Viricel A, Lucas T, Peltier H, Alfonsi E, et al. 2014. Habitat-driven population structure of bottlenose dolphins, *Tursiops truncatus*, in the North-East Atlantic. *Mol. Ecol.* 23(4):857–74

Ma H, Stewart K, Lougheed S, Zheng J, Wang Y, Zhao J. 2016. Characterization, optimization, and validation of environmental DNA (eDNA) markers to detect an endangered aquatic mammal. *Conserv. Genet. Resour.* 8(4):561–68

McGowen MR, Clark C, Gatesy J. 2008. The vestigial olfactory receptor subgenome of odontocete whales: phylogenetic congruence between gene-tree reconciliation and supermatrix methods. *Syst. Biol.* 57(4):574–90

McGowen MR, Tsagkogeorga G, Álvarez-Carretero S, dos Reis M, Struebig M, et al. 2020. Phylogenomic resolution of the cetacean tree of life using target sequence capture. *Syst. Biol.* 69:479–501

Meredith RW, Gatesy J, Cheng J, Springer MS. 2011. Pseudogenization of the tooth gene enamelysin (*MMP20*) in the common ancestor of extant baleen whales. *Proc. R. Soc. B* 278(1708):993–1002

Meredith RW, Gatesy J, Emerling CA, York VM, Springer MS. 2013. Rod monochromacy and the coevolution of cetacean retinal opsins. *PLOS Genet.* 9(4):e1003432

Miller DL, ed. 2007. *Reproductive Biology and Phylogeny of Cetacea: Whales, Dolphins, and Porpoises.* London: Routledge

Mullis KB, Faloona FA. 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol.* 155:335–50

Nichols C, Herman J, Gaggiotti OE, Dobney KM, Parsons K, Hoelzel AR. 2007. Genetic isolation of a now extinct population of bottlenose dolphins (*Tursiops truncatus*). *Proc. Biol. Sci.* 274(1618):1611–16

Norris RD, Hull PM. 2012. The temporal dimension of marine speciation. *Evol. Ecol.* 26(2):393–415

Nykänen M, Kaschner K, Dabin W, Brownlow A, Davison NJ, et al. 2019. Postglacial colonization of northern coastal habitat by bottlenose dolphins: a marine leading-edge expansion? *J. Hered.* 110(6):662–74

O'Leary MA, Uhen MD. 1999. The time of origin of whales and the role of behavioral changes in the terrestrial-aquatic transition. *Paleobiology* 25(4):534–56

Pacifici M, Santini L, Di Marco M, Baisero D, Francucci L, et al. 2013. Generation length for mammals. *Nat. Conserv.* 5:89–94

Palmer DH, Rogers TF, Dean R, Wright AE. 2019. How to identify sex chromosomes and their turnover. *Mol. Ecol.* 28(21):4709–24

Palsbøll PJ, Allen J, Bérubé M, Clapham PJ, Feddersen TP, et al. 1997. Genetic tagging of humpback whales. *Nature* 388(6644):767–69

Palsbøll PJ, Clapham PJ, Mattila DK, Larsen F, Sears R, et al. 1995. Distribution of mtDNA haplotypes in North Atlantic humpback whales: the influence of behaviour on population structure. *Mar. Ecol. Prog. Ser.* 116:1–10

Palsbøll PJ, Peery MZ, Olsen MT, Beissinger SR, Bérubé M. 2013. Inferring recent historic abundance from current genetic diversity. *Mol. Ecol.* 22:22–40

Palsbøll PJ, Vader A, Bakke I, El-Gewely MR. 1992. Determination of gender in cetaceans by the polymerase chain reaction. *Can. J. Zool.* 70(11):2166–70

Pampoulie C, Gíslason D, Ólafsdóttir G, Chosson V, Halldórsson SD, et al. 2021. Evidence of unidirectional hybridization and second-generation adult hybrid between the two largest animals on Earth, the fin and blue whales. *Evol. Appl.* 14(2):314–21

Pastene LA, Goto M, Kanda N, Zerbini AN, Kerem D, et al. 2007. Radiation and speciation of pelagic organisms during periods of global warming: the case of the common minke whale, *Balaenoptera acutorostrata*. *Mol. Ecol.* 16(7):1481–95

Pérez-Alvarez MJ, Olavarría C, Moraga R, Baker CS, Hamner RM, Poulin E. 2016. Historical dimensions of population structure in a continuously distributed marine species: the case of the endemic Chilean dolphin. *Sci. Rep.* 6:35507

Pilot M, Dahlheim ME, Hoelzel AR. 2010. Social cohesion among kin, gene flow without dispersal and the evolution of population genetic structure in the killer whale (*Orcinus orca*). *J. Evol. Biol.* 23(1):20–31

Polanowski AM, Robbins J, Chandler D, Jarman SN. 2014. Epigenetic estimation of age in humpback whales. *Mol. Ecol. Resour.* 14(5):976–87

Polanowski AM, Schmitt NT, Double MC, Gales NJ, Jarman SN. 2011. TaqMan assays for genotyping 45 single nucleotide polymorphisms in the humpback whale nuclear genome. *Conserv. Genet. Resour.* 3(4):645–49

Pompa S, Ehrlich PR, Ceballos G. 2011. Global distribution and conservation of marine mammals. *PNAS* 108(33):13600–5

Proches S. 2001. Back to the sea: secondary marine organisms from a biogeographical perspective. *Biol. J. Linn. Soc. Lond.* 74(2):197–203

Rey-Iglesia A, Martínez-Cedeira J, López A, Fernández R, Campos PF. 2018. The genetic history of whaling in the Cantabrian Sea during the 13th–18th centuries: Were North Atlantic right whales (*Eubalaena glacialis*) the main target species? *J. Archaeol. Sci. Rep.* 18:393–98

Rogers AR, Harpending H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* 9(3):552–69

Roman J, McCarthy JJ. 2010. The whale pump: marine mammals enhance primary productivity in a coastal basin. *PLOS ONE* 5(10):e13255

Roman J, Palumbi SR. 2003. Whales before whaling in the North Atlantic. *Science* 301(5632):508–10

Ruegg K, Rosenbaum HC, Anderson EC, Engel M, Rothschild A, et al. 2012. Long-term population size of the North Atlantic humpback whale within the context of worldwide population structure. *Conserv. Genet.* 14:103–14

Seersholm FV, Pedersen MW, Søe MJ, Shokry H, Mak SST, et al. 2016. DNA evidence of bowhead whale exploitation by Greenlandic Paleo-Inuit 4,000 years ago. *Nat. Commun.* 7:13389

Shen T, Xu SX, Wang XH, Yu WH, Zhou KY, Yang G. 2012. Adaptive evolution and functional constraint at TLR4 during the secondary aquatic adaptation and diversification of cetaceans. *BMC Evol. Biol.* 12:39

Sigsgaard EE, Jensen MR, Winkelmann IE, Møller PR, Hansen MM, Thomsen PF. 2020. Population-level inferences from environmental DNA—current status and future perspectives. *Evol. Appl.* 13(2):245–62

Sinding M-HS, Tervo OM, Grønnow B, Gulløv HC, Toft PA, et al. 2016. Sex determination of baleen whale artefacts: implications for ancient DNA use in zooarchaeology. *J. Archaeol. Sci. Rep.* 10:345–49

Skovrind M, Castruita JAS, Haile J, Treadaway EC, Gopalakrishnan S, et al. 2019. Hybridization between two high Arctic cetaceans confirmed by genomic analysis. *Sci. Rep.* 9:7729

Skovrind M, Louis M, Westbury MV, Garilao C, Kaschner K, et al. 2021. Circumpolar phylogeography and demographic history of beluga whales reflect past climatic fluctuations. *Mol. Ecol.* 30:2543–59

Solazzo C, Fitzhugh W, Kaplan S, Potter C, Dyer JM. 2017. Molecular markers in keratins from Mysticeti whales for species identification of baleen in museum and archaeological collections. *PLOS ONE* 12(8):e0183053

Speller C, van den Hurk Y, Charpentier A, Rodrigues A, Gardeisen A, et al. 2016. Barcoding the largest animals on Earth: ongoing challenges and molecular solutions in the taxonomic identification of ancient cetaceans. *Philos. Trans. R. Soc. B* 371(1702):20150332

Steeman ME, Hebsgaard MB, Fordyce RE, Ho SYW, Rabosky DL, et al. 2009. Radiation of extant cetaceans driven by restructuring of the oceans. *Syst. Biol.* 58(6):573–85

Sun YB, Zhou WP, Liu HQ, Irwin DM, Shen YY, Zhang YP. 2013. Genome-wide scans for candidate genes involved in the aquatic adaptation of dolphins. *Genome Biol. Evol.* 5(1):130–39

Swift JA, Bunce M, Dortch J, Douglass K, Faith JT, et al. 2019. Micro methods for megafauna: novel approaches to late Quaternary extinctions and their contributions to faunal conservation in the Anthropocene. *BioScience* 69:877–87

Székely D, Corfixen NL, Mørch LL, Knudsen SW, McCarthy ML, et al. 2021. Environmental DNA captures the genetic diversity of bowhead whales (*Balaena mysticetus*) in West Greenland. *Environ. DNA* 3(1):248–60

Tautz D, Ellegren H, Weigel D. 2010. Next generation molecular ecology. *Mol. Ecol.* 19:1–3

Tautz D, Renz M. 1984. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Res.* 12(10):4127–38

Thewissen JGM, Williams EM. 2002. The early radiations of Cetacea (Mammalia): evolutionary pattern and developmental correlations. *Annu. Rev. Ecol. Syst.* 33:73–90

Thompson K, Baker CS, van Helden A, Patel S, Millar C, Constantine R. 2012. The world's rarest whale. *Curr. Biol.* 22(21):R905–6

Toren D, Kulaga A, Jethva M, Rubin E, Snezhkina AV, et al. 2020. Gray whale transcriptome reveals longevity adaptations associated with DNA repair and ubiquitination. *Aging Cell* 19(7):e13158

Uhen MD. 2007. Evolution of marine mammals: back to the sea after 300 million years. *Anat. Rec. Adv. Integr. Anat. Evol. Biol.* 290(6):514–22

Uhen MD. 2010. The origin(s) of whales. *Annu. Rev. Earth Planet. Sci.* 38:189–219

Valsecchi E, Bylemans J, Goodman SJ, Lombardi R, Carr I, et al. 2020. Novel universal primers for metabarcoding environmental DNA surveys of marine mammals and other marine vertebrates. *Environ. DNA* 2(4):460–76

Viricel A, Rosel PE. 2014. Hierarchical population structure and habitat differences in a highly mobile marine species: the Atlantic spotted dolphin. *Mol. Ecol.* 23(20):5018–35

Wakeley J. 2004. Recent trends in population genetics: More data! More math! Simple models? *J. Hered.* 95(5):397–405

Wang JY, Frasier TR, Yang SC, White BN. 2008. Detecting recent speciation events: the case of the finless porpoise (genus *Neophocaena*). *Heredity* 101(2):145–55

Westbury MV, Cabrera AA, Rey-Iglesia A, De Cahsan B, Duchêne DA, et al. 2021. A genomic assessment of the marine-speciation paradox within the toothed whale superfamily Delphinoidea. bioRxiv 2020.10.23.352286. **https://doi.org/10.1101/2020.10.23.352286**

Willerslev E. 2003. Diverse plant and animal genetic records from Holocene and Pleistocene sediments. *Science* 300(5620):791–95

Wright S. 1931. Evolution in Mendelian populations. *Genetics* 16:97–159

Xu S, Chen Y, Cheng Y, Yang D, Zhou X, et al. 2012. Positive selection at the *ASPM* gene coincides with brain size enlargements in cetaceans. *Proc. R. Soc. B* 279(1746):4433–40

Zhou X, Guang X, Sun D, Xu S, Li M, et al. 2018. Population genomics of finless porpoises reveal an incipient cetacean species adapted to freshwater. *Nat. Commun.* 9:1276

Zhou X, Seim I, Gladyshev VN. 2015. Convergent evolution of marine mammals is associated with distinct substitutions in common genes. *Sci. Rep.* 5:16550