Review

*Cell*Press

Molecular evolution tracks macroevolutionary transitions in Cetacea

Michael R. McGowen[1*], John Gatesy[2], and Derek E. Wildman[1]

[1] Center for Molecular Medicine and Genetics, Wayne State University School of Medicine, Detroit, MI 48201, USA
[2] Department of Biology, University of California-Riverside, Riverside, CA 92521, USA

Cetacea (whales, dolphins, and porpoises) is a model group for investigating the molecular signature of macroevolutionary transitions. Recent research has begun to reveal the molecular underpinnings of the remarkable anatomical and behavioral transformation in this clade. This shift from terrestrial to aquatic environments is arguably the best-understood major morphological transition in vertebrate evolution. The ancestral body plan and physiology were extensively modified and, in many cases, these crucial changes are recorded in cetacean genomes. Recent studies have highlighted cetaceans as central to understanding adaptive molecular convergence and pseudogene formation. Here, we review current research in cetacean molecular evolution and the potential of Cetacea as a model for the study of other macroevolutionary transitions from a genomic perspective.

A poster child for macroevolution

The evolution of cetaceans (see Glossary) has emerged as a poster child for macroevolution, and is one of the best-characterized morphological transitions in the fossil record, documenting the transformation from a terrestrial to obligate aquatic mammal [1]. This is apparent in the use of Cetacea as a primary example of macroevolutionary change in recent textbooks on evolution [2,3]. Major phenotypic transitions that connect disparate taxa have been a focus of evolutionary research since Darwin and are still a central concern of paleontology and evolutionary developmental biology [2,3]. The molecular basis of these radical morphological transitions remains little explored due to the paucity of genetic data in nonmodel species; however, recent advances in genomics have allowed some groups to be studied in more detail.

Given the extraordinary diversity of derived attributes, cetaceans make an excellent case study for examining the molecular correlates of phenotypic change and what molecules are able to tell us about the transition to aquatic life.

Cetacea offers notable advantages relative to other taxa that also have experienced major anatomical and ecological reorganizations. Phylogenetic hypotheses for extinct and extant taxa are well developed (Figure 1), and include molecular divergence times among extant species with integration of extensive fossil data [4–8]. Semiaquatic hippopotamids (hippopotamuses), potential extant 'intermediate forms,' share multiple aquatic traits with cetaceans (Figure 2A,B) and might provide critical genetic insights into the early evolution of Cetacea (Figure 1).

Glossary

Cetaceans: a clade of mammals that includes whales, dolphins, and porpoises; the approximately 90 extant species are characterized by a wholly aquatic lifestyle.

Cetartiodactyls: a clade of mammals that includes camels, pigs, peccaries, cattle, antelope, deer, chevrotains, giraffes, hippopotamuses, and cetaceans.

Convergent recruitment: evolutionary change whereby an orthologous gene is utilized in two or more independent lineages to effect a similar phenotypic outcome in each lineage.

Convergent substitution: independent molecular evolution of the same state at a homologous position of the same gene and/or locus in two or more evolutionary lineages, effecting similar phenotypic change in each lineage.

Echolocation: sensory system that includes the emission of high-frequency sounds and the subsequent perception of their echoes to detect nearby objects.

Enamelogenesis: the process of forming enamel on the surface of teeth.

Frameshift mutation: an insertion or deletion in a protein-coding gene that causes a change in reading frame; commonly results in nonsensical amino acid sequence and premature stop codons.

Mysticeti: monophyletic group that includes all extant baleen whales, approximately 15 species.

Nonsynonymous change: a nucleotide substitution in a protein-coding gene that results in a modification of the encoded amino acid sequence; as opposed to a synonymous change, which is a nucleotide substitution that does not result in a modification of the encoded amino acid sequence.

Odontoceti: monophyletic group that includes all extant toothed whales, approximately 75 species.

Pinnipeds: a clade of aquatic mammals within the order Carnivora that comprises approximately 35 extant species, including seals, walruses, and sea lions; living representatives return to land to rest, breed, and give birth

Positive selection: selection for the maintenance or fixation of new genetic variants. Positive selection in a gene is inferred when the ratio of the rate of nonsynonymous substitutions to the rate of synonymous substitutions (dN/dS) is greater than 1. Neutral evolution is inferred if $dN/dS = 1$, and purifying selection if $dN/dS < 1$.

Pseudogene: a DNA segment that exhibits a high degree of similarity to a functional gene but contains defects that prevent it from being expressed properly; an inactivated gene that is functionally 'dead.'

Short interspersed elements (SINEs): short transposable elements that are interspersed throughout the genome and derive from reverse-transcribed RNA molecules.

Vomeronasal organ: an auxiliary olfactory structure in some tetrapods that is used for the detection of pheromones in mammals.
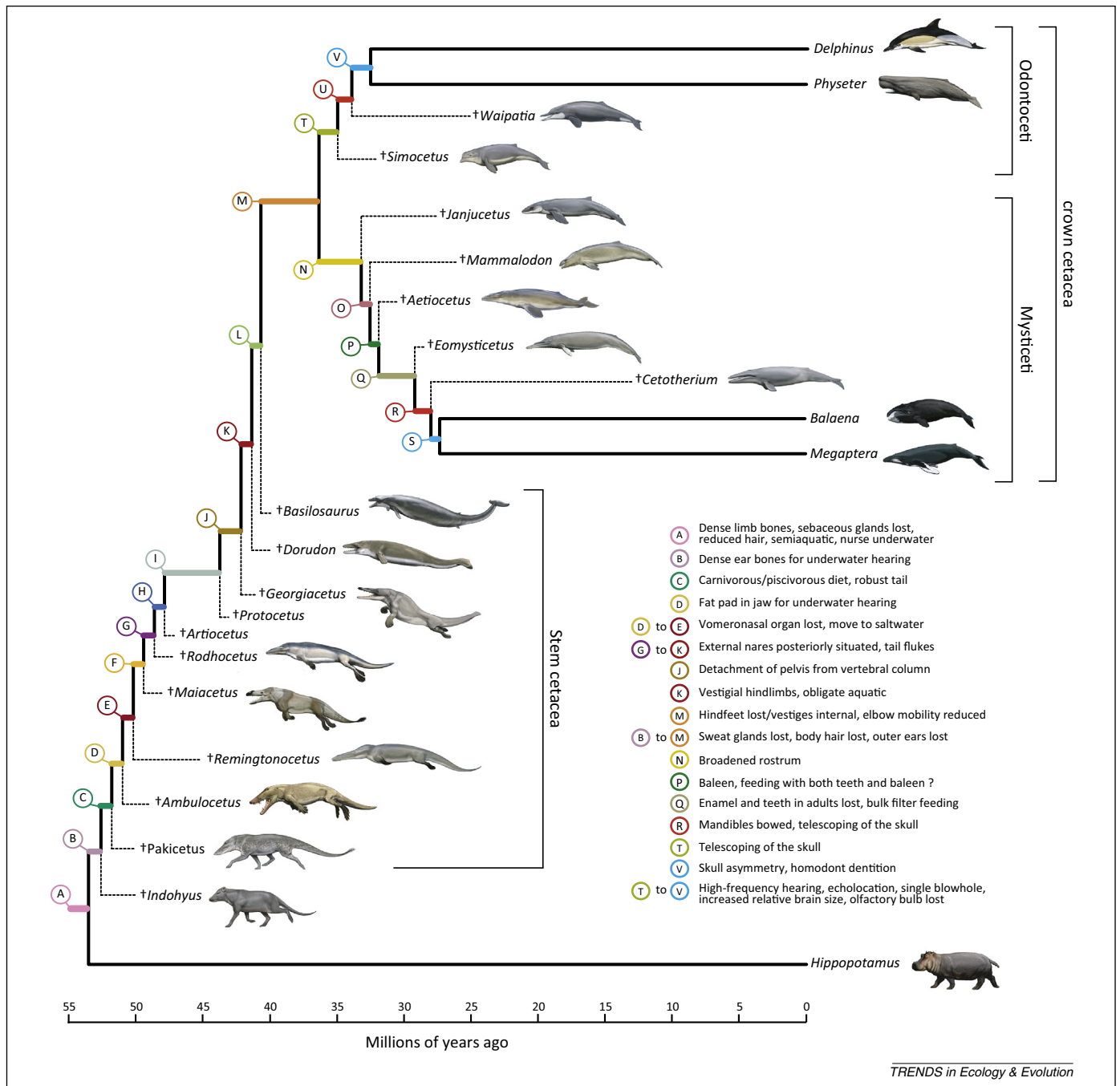
Figure 1. Phylogenetic relationships of Cetacea based on paleontological and molecular evidence showing the sequential acquisition of derived aquatic specializations in the group [5]. Internodes for outgroup branches (A,B), the stem lineage of Cetacea (C–M), the stem lineage of Mysticeti (N–S), and the stem lineage of Odontoceti (T–V) are colored and labeled with capital letters. Major character transitions and/or acquisitions are listed to the right and are ordered by internode (A–V) from the base of the tree. For character mappings that are ambiguous, a range of nodes is given. Approximate branch lengths are from [5], and brackets to the right delimit groups. Fossil lineages are dashed, and extinct genera are marked by crosses. Paintings are by Carl Buell.

In addition, developmental trajectories and vestigial or rudimentary traits help track major transformations in cetaceans (Figure 2C,D), where in several cases, ontogeny recapitulates phylogeny [1,9–12]. The fossil record of Cetacea is impressive [11,13–20] and includes primitive taxa that record sequences of change at three critical phases: initial aquatic specialization, evolution of filter feeding with baleen (Figure 2E), and derivation of echolocation (Figure 1). Whole genomes for both cetaceans [21–23] and terrestrial cetartiodactyls (e.g., Box 1, [24]) now illuminate these transitions, and candidate loci associated with aquatic features and degenerative evolution have been

characterized by several research groups (Figure 2F). Divergence times in cetaceans are conducive to detection of positive selection and inactivated loci (Figure 3); evolutionary splits in Cetacea are deep enough to provide adequate change for statistically robust analyses of molecular evolution, but are not so deep that substitution patterns are saturated and recognition of pseudogenes is hindered. Given the extensive phenotypic convergence with other taxa (e.g., echolocating bats, seals, and manatees), convincing evidence of parallel pseudogenization [25–28] and adaptive molecular convergence [29–32] have been documented.
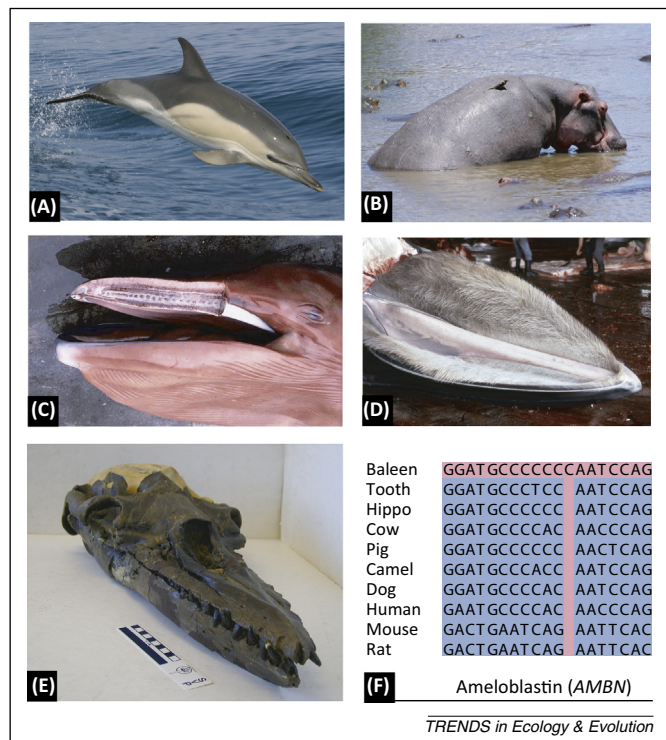
Figure 2. Representatives of Cetacea [**(A)** *Delphinus delphis*] and its extant sister group Hippopotamidae [**(B)** *Hippopotamus amphibius*;]; fetal teeth **(C)** and baleen **(D)** in *Balaenoptera physalus*; Oligocene toothed mysticete, *Aetiocetus weltoni*, which shows fossil evidence for simultaneous expression of teeth and baleen **(E)**; alignment of ameloblastin (*AMBN*) enamel gene sequences **(F)** from a baleen whale, a toothed whale, and more distantly related mammals with single base-pair frameshift insertion in the baleen whale *AMBN* pseudogene (pink). Clear-cut molecular signature of a shared aquatic ancestry for hippos and whales is currently lacking; loci related to underwater vision, deep diving, and lung surfactant do not show evidence of molecular adaptation in the common ancestor of these taxa [30,31,79]. (C–F) represent a convergence of evidence from ontogeny, fossils, and genetics for the transition from feeding with teeth (C), to feeding with teeth + baleen (E), to feeding solely with baleen racks (D) in the context of phylogeny (Figure 1, main text; internodes N–S). Reproduced, with permission, from NOAA-SWFSC Robert Pitman (A), Rochelle Campbell (B), Alex Aguilar (C,D), and Tom Deméré (E).

Box 1. Cetacean genomics

A low coverage (2.59×) bottlenose dolphin (*Tursiops truncatus*) genome was sequenced using Sanger methods and originally released in 2008. Coverage of this genome has been increased subsequently through the use of next-generation sequencing technology [21]. No less than four papers [80–83] have recently analyzed the dolphin genome to determine which genes were positively selected on the dolphin lineage. Genome-wide scans of *Tursiops* revealed that approximately 4% of 1:1 orthologous genes showed evidence for positive selection. Three scans indicated that a large percentage of these genes are expressed in the nervous system or involved in nervous system development [80,81,83]. In addition, these studies supported enrichment of positively selected loci in gene ontology categories related to immunity, lipids, heart development, hearing, and skin [80–83]. Recently, both Yangtze River dolphin (*Lipotes vexillifer*) and minke whale (*Balaenoptera acutorostrata*) genomes have been sequenced and analyzed [22,23], offering additional insights into cetacean evolution. These studies confirmed the low nucleotide substitution rate characteristic of cetaceans and revealed positive selection and/or expansion of genes related to osmoregulation, hypoxia, and DNA repair. In addition, these genomes have provided more evidence of pseudogenization in cetaceans, including genes involved in blood clotting, hair and tooth development, sight, olfaction, and taste (Figure 3, main text).

lineage leading to Cetacea, *Indohyus* (approximately 48 Mya), is interpreted as partially aquatic with dense limb bones for walking underwater, similar to hippos, and advanced capacity to hear in an aquatic medium [15]. Early cetacean fossils, such as *Pakicetus* and *Ambulocetus* (Figure 1), show progressive specializations for living in water, with *Ambulocetus* having proportionally shorter hindlimbs but broad feet likely utilized for swimming [14,16,17]. Extinct taxa, such as *Rodhocetus* (Figure 1), are characterized by posteriorly positioned external nares (nostrils), well-developed underwater hearing, and locomotion via hindlimb paddling and dorsoventral undulation [18,19]. Middle to Late Eocene Basilosauridae (e.g., *Dorudon* and *Basilosaurus*; Figure 1) were the first fully aquatic cetaceans with bony nares shifted even further posteriorly, forelimb flippers, a tail fluke, highly reduced hindlimbs, and pelvis detached from the spine [18]. Fossils from the latest Eocene to the present detail the diversification of modern cetaceans, including further aquatic specializations, such as echolocation in Odontoceti and filter feeding with baleen in Mysticeti [5]. In sum, this astonishing series of fossils illustrates an incremental transition in early cetaceans to an obligate aquatic lifestyle (Figure 1).

Cetaceans faced multiple obstacles upon entering the ocean, many of which were overcome with the derivation of anatomical and physiological novelties. For example, eyes that can see in dim light conditions, specialized underwater hearing, advanced oxygen exchange and storage required for long dives, expanded relative brain size associated with increased sociality, and massive fat reserves with metabolic and thermoregulatory functions evolved within Cetacea [33]. Recent studies have revealed the molecular basis for some of these traits and characterized associations between pseudogenes and lost or vestigial anatomical structures (Figure 3). In addition, adaptively evolving suites of genes have been identified via analyses of three complete genomes, with additional genomes available in the near future (Box 1).

The confluence of these features is not replicated in other model systems of vertebrate macroevolution. For example, birds and turtles are highly derived taxa but are ancient clades and more difficult to study from a molecular perspective, wholly extinct groups (ichthyosaurs, plesiosaurs, and pterosaurs) lack genetic data, bats and snakes have comparatively poor fossil records, and the extant diversity of obligately aquatic manatees is limited. All of these groups lack a living sister group, such as hippopotamuses, that retains numerous functional characteristics relevant to macroevolutionary patterns [4–7]. Therefore, Cetacea has emerged as a unique case study that demonstrates key concepts in molecular evolution. In this review, we outline major recent findings that have catapulted cetaceans into the evolutionary spotlight once again, this time as exemplars of broad-scale genomic change that accompanies the morphological and physiological transition to a radically different environment.

The morphological transition

Numerous fossils from the Eocene [56–34 million years ago (Mya)] record the move from land to sea [1,4,5,13,14] (Figure 1). One of the earliest branching taxa on the
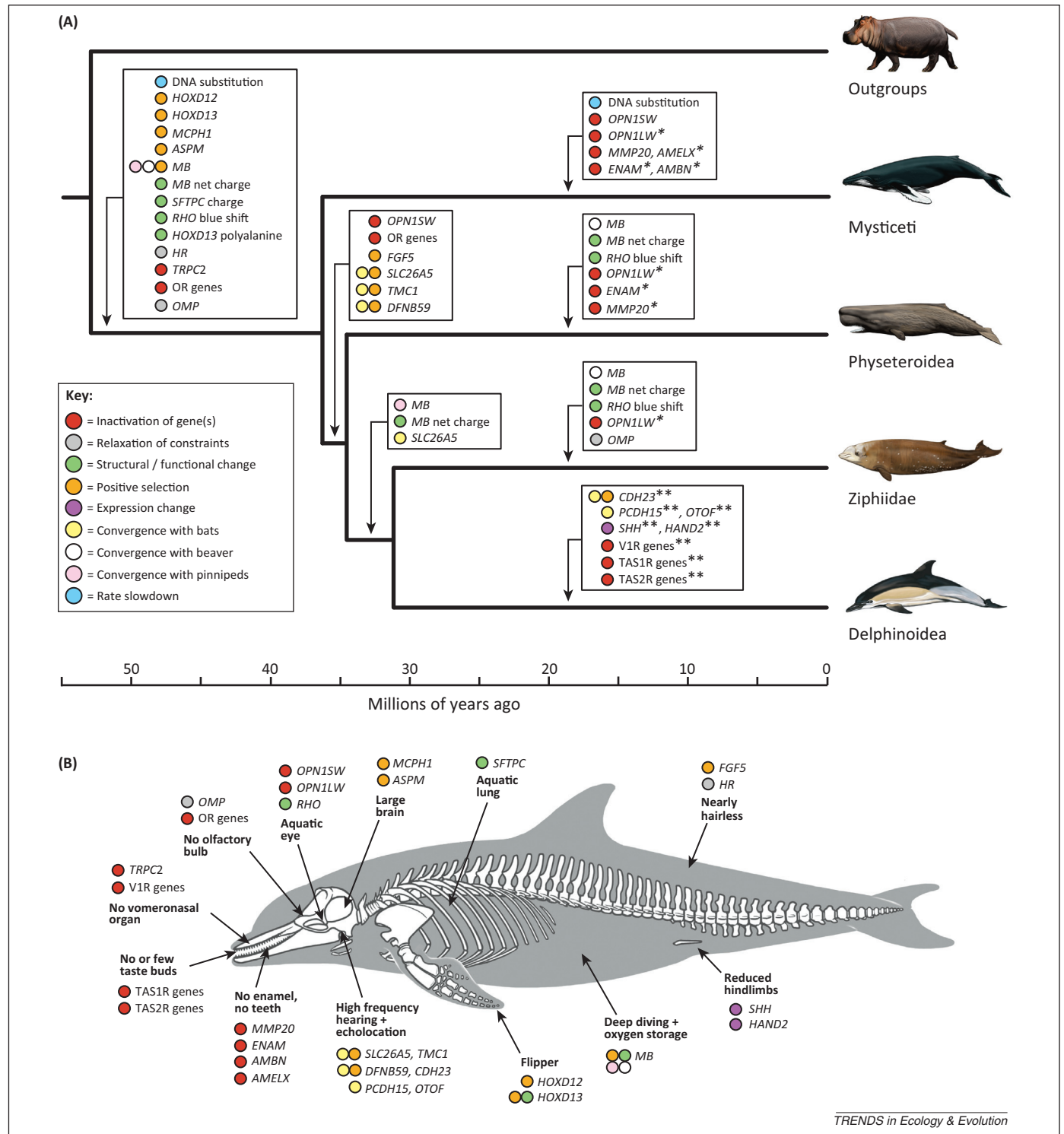
Figure 3. Major molecular changes in Cetacea mapped onto a phylogenetic hypothesis **(A)** and to corresponding anatomical features **(B)**. Colored circles identify the type of molecular change. Single asterisks show that the molecular change occurred within the given clade; for example, inactivation of *OPN1LW** on the Physeteroidea branch indicates that pseudogenization occurred within Physeteroidea (in this case, independently in *Physeter* and in *Kogia* [30]). Double asterisks mark genetic changes that have been inferred based on the bottlenose dolphin genome, and are mapped arbitrarily to the Delphinoidea branch. Inferred positive selection in *MCPH1, ASPM, MB, HOXD12,* and *HOXD13* is mapped to the stem cetacean branch and indicates general evidence for positive selection at these loci in Cetacea [31,77,84–86]. Genes not discussed in the main text are also mapped here [84–88]. Topology, branch lengths, and timescale are as in [30]. Paintings are by Carl Buell. For definitions of gene symbols, please see the main text.

Molecular convergence

Convergent evolution is defined as the independent origination of similar traits in two or more clades [34], For most of the history of evolutionary thought, discussions of convergent evolution have been confined to morphology, physiology, or behavior, but molecular convergence increasingly has been recognized as a more common phenomenon than was previously believed [32,34,35].

Extensive convergence in protein-coding genes implies that these molecular changes might cause similar phenotypic changes in divergent lineages.

Phenotypic convergence need not be due to the same molecules, but molecular convergence commonly involves change in the same (orthologous) molecules from two or more distinct lineages. Molecular convergence can be classified into several types. Deployment and modification of the same gene in different species can lead to similar phenotypes (termed 'convergent recruitment' [34]), as in sodium channel genes of electrostatic fishes or melanocortin 1 receptor (*MC1R*)-based pigmentation in tetrapods [35,36]. With convergent recruitment, there might be no specific convergent similarities at individual homologous sites of a particular gene. Perhaps the most compelling examples of molecular convergence involve parallel changes to an identical amino acid at the homologous site in two or more lineages, resulting in similar phenotypes. Christin *et al.* [34] termed this 'convergent substitution'. Examples currently are rare in the literature, and include lysozymes of foregut-fermenting mammals and birds, RNases in monkeys, phosphoenolpyruvate carboxylase (PEPC) enzymes in C_4 grasses, and mitochondrial proteins in snakes [37–41]. Similar sequence motifs in the genomes of different taxa might also be due to convergent evolution of nonorthologous genes, as in antifreeze proteins of polar fishes [42] and the silk fibroin sequences of moths and spiders [43]. Finally, parallel inactivation of genes is expected with the independent loss of a homologous anatomical structure in various clades [27].

In any given case, convergent evolution can include a complex combination of phenotypic convergence encoded by different genes in different lineages, convergent recruitment of orthologous molecules, convergent substitution at particular sites, and parallel pseudogenization of homologous genes. Below, we discuss various examples of molecular change in the aquatic specialization of cetaceans where convergent recruitment, substitution, and pseudogenization have been documented. In sum, these case studies imply that molecular convergence is a pervasive evolutionary phenomenon, and that major morphological and physiological transitions can be constrained to follow a small subset of molecular pathways [32,34,35].

Myoglobin tracks the convergent evolution of deep diving

Mammals have invaded the aquatic environment numerous times over their evolutionary history. As air-breathers, aquatic and semiaquatic mammals require modifications of the respiratory and cardiovascular systems that confer a greater tolerance to hypoxia and increased pressure when foraging at depth. In cetaceans, these diving specializations include the cessation of gas exchange, collapse of the lungs, and enhanced storage of oxygen in blood and muscle [44]. The deepest divers among cetaceans include sperm whales (Physeteroidea) and beaked whales (Ziphiidae), which primarily feed on deep-water cephalopods and descend to depths of >1000 m [45].

Myoglobin is a small 153-amino acid protein encoded by the *MB* gene and is the primary carrier and storage center of oxygen in muscle. Myoglobin concentrations in diving mammals, such as cetaceans and pinnipeds, can be 30-fold greater than in terrestrial mammals. The concentration of myoglobin is correlated with the net surface charge of the protein [31] as well as protein folding stability, and there is a signature of positive selection in Cetacea [46,47]. High net surface charge enables electrostatic repulsion that might reduce deleterious self-association of myoglobin molecules; protein-folding stability could have a similar impact by limiting the percentage of misfolded molecules at high expression levels.

Mirceta *et al.* [31] sequenced and assembled myoglobin data and then mapped changes in net surface charge across a tree of 130 mammals. Amino acid replacements that increased charge occurred overwhelmingly in lineages of extreme divers with semiaquatic or aquatic lifestyles (Figure 4). The greatest increases in charge occurred independently in cetaceans, pinnipeds, and the beaver (*Castor*), all of which share at least two convergent amino acid replacements within each lineage. Three amino acid changes increased net charge by +2.6 on the cetacean stem lineage, predicting a 3.5-fold increase in myoglobin concentration. Further amino acid changes increased charge in parallel on the sperm whale and beaked whale lineages (Figure 4). Thus, deep diving in mammals represents a case of both convergent recruitment of the same molecule and convergent substitution at particular sites.

Remarkably, by combining body mass with inferred concentration of myoglobin based on net surface charge, Mirceta *et al.* [31] were able to estimate the maximum dive times of extinct cetaceans. The earliest fossil relatives of Cetacea (e.g., *Indohyus)* were estimated to have concentrations of myoglobin akin to the hippopotamus and, thus, were not capable of long sustained dives (approximately 1.6 min). Based on ancestral reconstructions, late Eocene cetaceans such as *Basilosaurus* and *Dorudon* (Figure 1) experienced an increase in myoglobin concentration relative to the earliest whales and an inferred approximately 17.4 min dive time, within the range of modern mysticetes and dolphins, but well outside the range of elite divers. In addition to presenting a textbook example of molecular convergence, this case study illustrates how reconstructing the evolutionary history of a molecule from living species can predict behaviors of extinct species that are otherwise unknown from preserved skeletal elements.

Convergence in auditory genes between echolocating bats and odontocetes

One of the more remarkable research findings in recent years is that echolocating bats and toothed whales share extensive convergent substitution in various 'auditory genes' [29,32,48–51]. Indeed, bats and odontocetes have many similar capabilities with regards to high-frequency sound perception [52]. In the odontocete echolocation system, high-frequency sounds are produced in the nasal cavity and focused via a fatty organ called the melon located anterior to the nares [53]. These sounds are reflected back and received through the lower jaw to a modified inner ear. Basal turns of the cochlea similar to echolocating bats permit perception of high frequency sounds up to 200 kHz in some species [54]. The earliest
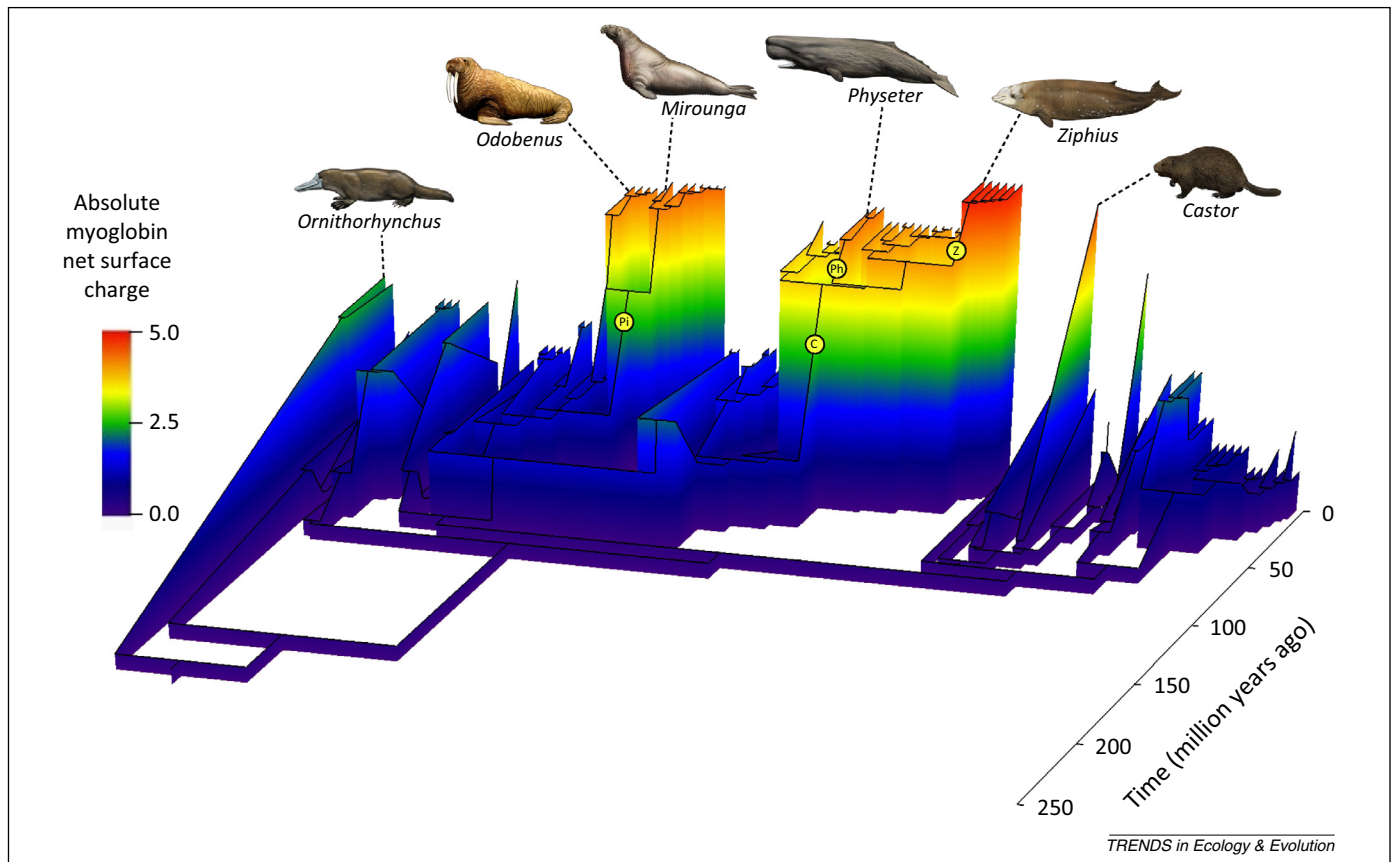
Figure 4. The evolution of absolute net surface charge in the oxygen storage molecule myoglobin (MB). Reconstruction of ancestral amino acid sequences indicates several instances of convergence in net surface charge among lineages of diving mammals [31]. The 3D heatmap time-tree shows increased surface charge in the duck-billed platypus, Pinnipedia (Pi), Cetacea (C), beaver, and muskrat (peak to the right of *Castor*). Net charge increases in MB on the stem branches of Cetacea, Physeteroidea (Ph), and Ziphiidae (Z) were coincident with parallel blue shifts in rhodopsin (RHO) (Figure 3, main text). 3D tree provided by Michael Berenbrink; paintings are by Carl Buell.

odontocete fossils already show phenotypic evidence for perception of high-frequency sound, including cochlear specializations, such as wide basal turns and an enlargement of the cochlear duct, which houses the hair cells of the inner ear [5,18,55].

In 2010, two studies simultaneously reported evolutionary analyses of the cochlear gene solute carrier family 26 (anion exchanger), member 5 (*SLC26A5*) in cetaceans and bats [48,49]. *SLC26A5* encodes prestin, which is involved in conferring electromotility to the outer hair cells of the cochlea, enabling sensitivity to sound by driving the cochlear amplifier. In each study, a phylogenetic analysis of prestin amino acid sequences supported a tree in which some odontocete cetaceans were nested within bats with high support, despite phylogenetic analysis of nucleotide sequences favoring a more distant relation (Figure 5) [48,49]. This striking pattern was driven by convergent amino acid changes that were shared between toothed whales and echolocating bats. Further phylogenetic analysis of amino acid sequences from a broader sampling of cetacean species revealed that horseshoe and Old World leaf-nosed bats (Rhinolophidae + Hipposideridae) were included in a clade with cetaceans [50]. Significant convergence was found for three pairs of branches, resulting in 26 total amino acids in prestin that converged between some or all echolocating bats and various toothed whales; many of these changes occurred on the Odontoceti branch and the

Ziphiidae + Delphinoidea branch (Figure 5). In addition, there was a significant association between the number of nonsynonymous changes in prestin and the highest frequency of hearing sensitivity in particular cetacean species, implying a tight link between amino acid change and high-frequency hearing [50]. Both the Odontoceti and Ziphiidae + Delphinoidea branches also showed evidence of positive selection, signifying that these changes were the result of adaptive evolution rather than of random drift.

Analyses of additional auditory genes and a genomewide analysis [six bats plus the bottlenose dolphin (*Tursiops truncatus*)] revealed further compelling evidence of convergent substitution between echolocating bats and toothed whales [29,32,51]. These included several genes associated with hair cells of the cochlea and, in some cases, linked to deafness in humans [transmembrane channel-like 1 (*TMC1*), deafness, autosomal recessive 59 (*DFNB59*), cadherin-related 23 (*CDH23*), protocadherin-related 15 (*PCDH15*), and otoferlin (*OTOF*)] (Figure 3). In addition, significant positive selection on the dolphin lineage was detected in *TMC1*, *DFNB59*, and *CDH23* [29,51]. Recent comparative analyses furthermore suggest that genomes of four echolocating bats and the bottlenose dolphin share convergent amino acid changes in as many as 392 homologous loci, some of which are known to be involved in hearing or linked to protein–protein interaction networks involved in cochlear development [32].
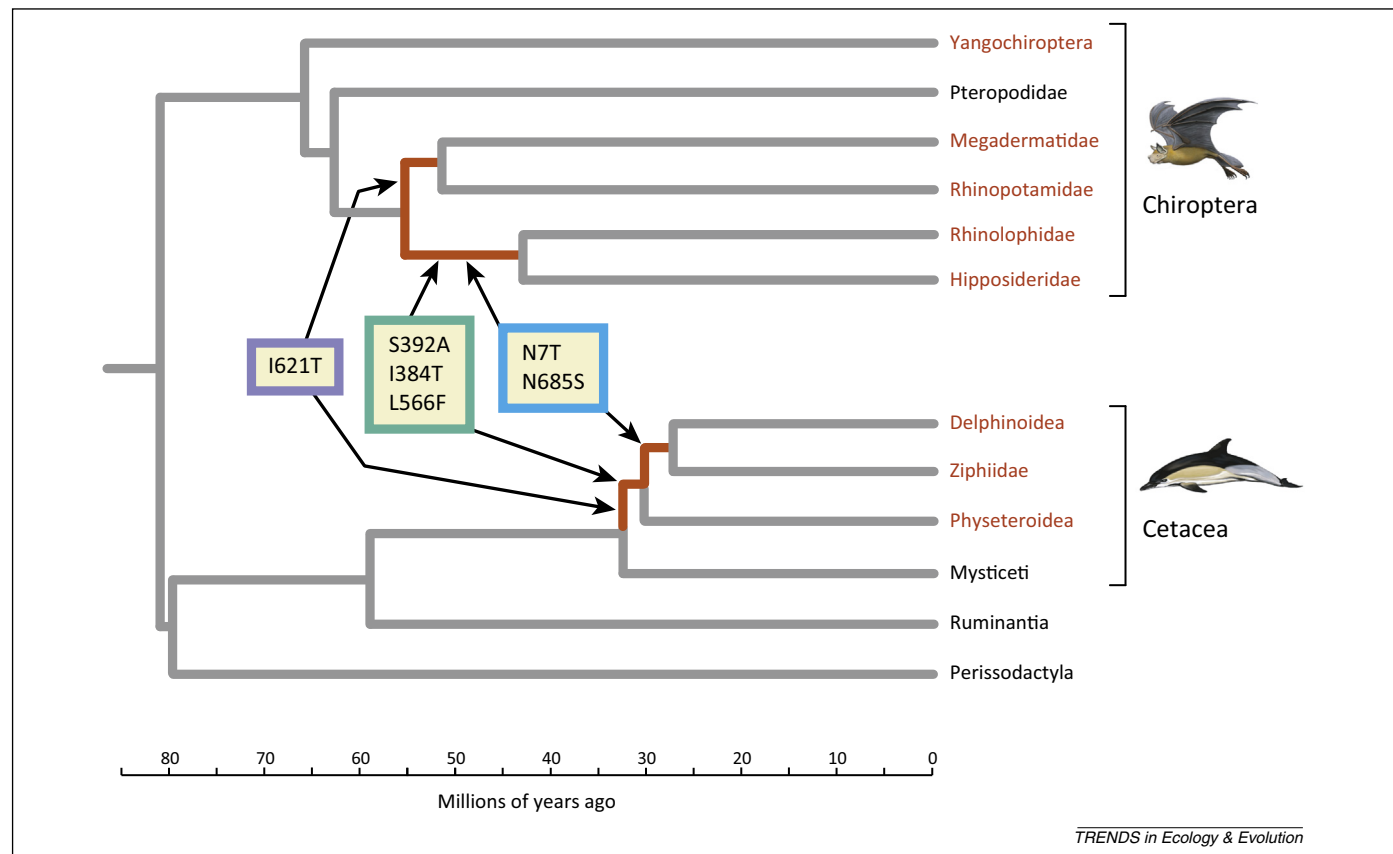
Figure 5. Species tree with selected amino acid replacements in solute carrier family 26 (anion exchanger), member 5 (*SLC26A5*; prestin) that are convergent between bats and cetaceans. Red lettering indicates taxa that echolocate and have ultrasonic hearing. The Rhinolophidae + Hipposideridae clade of bats groups within Cetacea in the tree supported by amino acid sequences of *SLC26A5*, implying a strong pattern of convergence at the molecular level between distantly related evolutionary relatives [50]. Boxes indicate unambiguous convergent amino acid changes between three pairs of branches. These three branch comparisons were identified as having at least marginally significant (*P* <0.1) signals of convergent selection by Liu *et al.* [50]. Branch lengths are based on the time-tree from [7]. Paintings are by Carl Buell.

Taken together, these results provide extraordinary evidence of adaptive convergence at the molecular level across the genome (Figure 3). The evolution of echolocation is a complex process, involving many genes, but broad genomic comparisons across large phylogenetic distances can pinpoint genes related to the function of a complex trait. As more genomes become available, this might be just one of many future examples where specific amino acid changes at multiple loci underlie complex phenotypic or physiological convergence in distantly related lineages.

Pseudogenes reflect past morphological change and ancestral function

With the move to an aquatic environment, cetaceans have become anatomically streamlined and have lost many structures. Within a phylogenetic context, the absence of anatomical structures in certain cetaceans (e.g., enamel or olfactory bulbs) and the presence of these structures in outgroups predicts inactivation of genes related to these structures. Pseudogenes are often neglected by genetic studies, but can provide critical information for reconstructing ancestral functionality and the sequence of phenotypic change. Given the radical transformation of the senses, pseudogenes abound in cetaceans, offering a glimpse at the process of large-scale pseudogenization across genes related to olfaction, taste, vomeronasal chemoreception, and sight.

The evolution and inactivation of cetacean opsins

The behavior of light in aquatic environments poses distinct challenges that are not present on land, including reduction in image quality due to scattering of light, reduced brightness, and a shift in the spectrum of light to deep blue in open-ocean environments [56]. In mammals that utilize the sense of sight, the optical pigment rhodopsin (RHO) is expressed in rod cells of the retina and is optimal in low light conditions. In addition, most mammals have dichromatic color vision due to the presence of two other opsins, short wavelength-sensitive opsin (OPN1SW) and long wavelength-sensitive opsin (OPN1LW) expressed in S-cones and L-cones, respectively [57]. Presumably due to dim light conditions, cetaceans have a rod-dominated retina, and all cetaceans analyzed lack S-cones [57]. The absence of S-cones is correlated with change at the molecular level, where cetaceans have an inactivated *OPN1SW* [30,58]. However, odontocetes and mysticetes do not share a common inactivating mutation in this gene, and the last common ancestor of modern cetaceans likely had a functional *OPN1SW* [30] (Figure 3).

It is perhaps unexpected that *OPN1SW* is knocked out in Cetacea, because in most mammals this gene is responsible for the reception of shorter wavelengths abundant in the marine environment. Initially, it was noted that some cetacean species, especially those in pelagic environments, have rhodopsin that is shifted toward blue light, implying a

compensation for the loss of *OPN1SW* [59,60]. However, Meredith *et al.* [30] analyzed complete sequences of all three opsins in a wide array of cetaceans and reconstructed a blue shift in *RHO* on the cetacean stem that was the result of amino acid replacements at two key spectral tuning sites (Figure 3). Their phylogenetic reconstruction [30] implied that this shift preceded the convergent pseudogenization of *OPN1SW* in mysticetes and in odontocetes and that *OPN1LW* was subsequently inactivated independently in five separate lineages (*Physeter*, *Kogia*, *Mesoplodon*, Balaenidae, and Balaenopteroidea), suggesting convergent recruitment by subtraction of 'unnecessary' genes that interfere with function.

Pseudogenization of both *OPN1SW* and *OPN1LW* has resulted in rod monochromacy, a unique feature among mammals (Figure 3). Rod monochromacy has been considered an extreme adaptation to low light levels, and the cetacean monochromats are all deep divers. Physeteroids and ziphiids, extreme divers that can descend to >1000 m, show evidence of further independent blue shifts in *RHO*. Convergent substitution was inferred at a key tuning site, and this is coincident with the evolution of net surface charge in myoglobin on these two lineages [30,61] (Figure 3). The overall evolutionary pattern suggests that selection has driven the independent loss of both *OPN1SW* and *OPN1LW* in deep-diving cetaceans to optimize visual acuity at depth and at night [30].

The degradation of chemosensory genes

As with sight, the transition to an obligate aquatic lifestyle impacted the ability to detect chemical signals. Chemosensory structures, including the olfactory, vomeronasal, and gustatory systems, have been substantially reduced or have disappeared altogether in cetaceans. Molecular correlates of these systems, such as olfactory receptors (ORs), vomeronasal receptors (V1Rs and V2Rs), and taste receptors (TAS1Rs and TAS2Rs), are G protein-coupled receptors (GPCRs) responsible for the detection of specific molecules and are of paramount importance for finding food, identifying volatile compounds, and regulating social cues [62].

In mammals, ORs only function in air [62], reducing their relevance for fully aquatic taxa. As a result, the olfactory systems of cetaceans have been heavily modified. Postnatally, odontocetes lack an olfactory bulb, olfactory nerve, cribriform plate, and most likely the ability to detect airborne molecules [9]. The nasal passages of odontocetes have been extensively remodeled to serve as a source of sound production for their elaborate echolocation system [10]. However, mysticetes retain all of these olfactory structures, although reduced in comparison to terrestrial mammals [10]. Anecdotal behavioral evidence [63] as well as a recent anatomical description of a small, but histologically complex, olfactory bulb in the bowhead whale [64] suggest that at least some mysticetes retain a sense of smell.

The percentage of OR pseudogenes present in a mammalian genome roughly mirrors the elaboration of olfactory anatomy and the ecology of a particular species [25,65]. Several studies have identified OR genes in diverse cetacean species, many of which were inferred to be

pseudogenes based on frameshift mutations or premature stop codons [23,25,64,66–68] (Figure 3). Odontocetes are characterized by a greater proportion of pseudogenes (74–100%) relative to Mysticeti (29–58%), consistent with loss of the olfactory apparatus early in the history of Odontoceti (Figure 1) [23,25,64,67,68]. The greater proportion of functional OR genes in mysticetes more closely resembles that of other marine mammals, (e.g., pinnipeds and sirenians), which like mysticetes retain the olfactory sense [25].

The vomeronasal organ (VNO) is completely lacking in extant cetaceans [69] and based on the absence of incisive foramina in early fossil whales, was lost >45 Mya (Figure 1). Functional V1R genes are not found in cetaceans, although the bottlenose dolphin genome contains approximately 36 V1R pseudogenes [70]. The bottlenose dolphin and fin whale do not have a functional transient receptor potential cation channel, subfamily C, member 2, pseudogene (*TRPC2*) gene, a gene exclusively expressed in VNO neurons [71], and share inactivating mutations, implying that VNO function was lost before the last common ancestor of Odontoceti and Mysticeti (Figure 3) [72]. This pattern is consistent with reconstructions of ancestral anatomy (Figure 1) and convergent degradation of *TRPC2* in multiple lineages of aquatic mammals [72].

Cetacean tongues have few or no taste buds, suggesting that taste is severely reduced or absent [69]. Sweet, umami, and bitter tastes are regulated by receptors belonging to the *TAS1R* and *TAS2R* gene families of GPCRs [73]. Three TAS1R genes generally are found in mammals, but all of these loci are pseudogenes in the bottlenose dolphin; in addition, the bottlenose dolphin genome contains at least ten TAS2R pseudogenes [26]. The loss of functional TAS1R genes also occurred independently in Otariidae (sea lions), and many carnivorous mammals convergently lack a functional sugar taste perception gene *TAS1R2* [26].

Molecular cavities in enamel genes

Some cetaceans also show evidence of pseudogenization in genes that are critical for enamelogenesis [enamelin (*ENAM*), amelogenin, X-linked (*AMELX*), ameloblastin (enamel matrix protein) (*AMBN*), matrix metallopeptidase 20 (*MMP20*)]. Mysticetes, as well as odontocetes belonging to the genus *Kogia* (pygmy and dwarf sperm whales), lack enamel [27]. Adult mysticetes are entirely toothless, although enamel-less tooth buds develop and are resorbed during fetal development [74] (Figure 2C). Mysticete fossils document a stepwise pattern of tooth loss from predators with well-developed teeth (*Janjucetus* and *Mammalodon*), through forms with both teeth and palatal foramina that imply the presence of baleen (*Aetiocetus*) (Figure 2E), to forms with no teeth postnatally and only baleen as adults (*Eomysticetus* and modern mysticetes) [11,20] (Figure 1). The phylogenetic pattern predicts that purifying selection for the maintenance of enamel should be relaxed on the stem mysticete branch [27].

Mysticete whales and *Kogia* have inactivated *ENAM* genes (Figure 3), as do other enamel-less mammals including aardvarks, pangolins, and xenarthrans, another example of convergent pseudogenization in Cetacea [11,27]. For the region sequenced, there are no frameshift mutations or

premature stop codons in *ENAM* that are shared among all mysticetes, although relaxation of constraints is evident before the diversification of crown mysticetes [27]. Frameshifts and premature stop codons also were found in the *AMBN* (Figure 2F), *AMELX*, and *MMP20* genes of mysticetes [11,28]. *MMP20* encodes a protease that processes secreted enamel proteins; insertion of a short interspersed transposable element (SINE) breaks up the reading frame of *MMP20* and maps to the ancestral mysticete branch [28]. This implies that the secretion of enamel terminated no later than this insertion event and before the branching of modern baleen whale species (Figure 3), as suggested by phylogenetic analyses that include fossils (Figure 1) and the ontogeny of extant baleen whales (Figure 2).

The genomes of cetaceans tell a tale of wholesale degradation of genes involved in chemosensory detection and enamel development (Figure 3). This is complementary to evidence from morphology, which suggests a large-scale reduction or disappearance of chemosensory structures in adult cetaceans (Figure 1). Presumably, the progressive dependence on the aquatic environment in stem cetaceans led to a relaxation of selective pressure for maintaining both the molecules and organs involved in smell and taste. The numerous pseudogenes in cetacean genomes are vestigial remnants, genetic fossils that document primitive sensory capabilities that are absent in modern taxa [11,68].

Ontogeny recapitulates phylogeny in the evolution of cetacean limb development

Extreme modifications of the limbs represent perhaps the most striking transformations in the cetacean body plan (Figure 1), and several recent studies have related changes at the molecular level to this phenotypic divergence (Figure 3). Phylogenetic analyses of fossils document a gradual reduction in the size of the hindlimbs through the Eocene [12]. In addition, the cetacean forelimb was modified into a rigid flipper before the last common ancestor of extant cetaceans (Figure 1).

To ascertain potential developmental correlates of limb loss, Thewissen *et al.* [12] documented the expression patterns of genes during hindlimb bud formation and resorption in the pantropical spotted dolphin (*Stenella attenuata*). In mouse, *FGF8* is expressed at the distal end of the limb bud. At the maximum extent of hindlimb development in dolphins, embryos express fibroblast growth factor 8 (*FGF8*) in the distal hindlimb, but *FGF8* expression is not maintained during limb bud resorption. Two other genes, Sonic Hedgehog (*SHH*) and heart and neural crest derivatives expressed 2 (*HAND2*), are not expressed in the hindlimb throughout all stages of limb development [12]. *SHH* is required for the maintenance of *FGF8* expression in mice [75]. Conditional knockout mice that lack *SHH* hindlimb expression show a similar pattern in *FGF8* expression, do not form distal limb elements, and retain more proximal elements within the body wall, a condition similar to that of modern cetaceans [12,75]. The loss of *SHH* and *HAND2* expression in the hindlimb is inferred to have occurred before the last common ancestor of modern cetaceans.

The forelimb also experienced profound change in cetaceans, including the shortening of proximal elements, the

loss of digit I in some mysticetes, and the addition of phalanges in some lineages [76]. Wang *et al.* [77] inferred an increase in polyalanine repeats from 14–15 to 17–18 in the homeobox D13 (*HOXD13*) gene of modern cetaceans; intriguingly, an increase in polyalanine repeats in *HOXD13* of mouse results in webbing between the digits [78]. Another gene involved in digit patterning, *HOXD12*, also shows evidence of adaptive evolution along the cetacean stem lineage [77]. In comparisons among mammals, sliding window analyses record extensive positive selection at specific regions of *HOXD12* and *HOXD13* in cetaceans (Figure 3). Wang *et al.* [77] suggested that the earlier loss of developmental patterning in hindlimbs [12] loosened developmental constraints on the forelimb and permitted increased diversity in forelimb shape and size within Cetacea

Concluding remarks

Cetacea represents a group of mammals that has developed remarkable adaptations to the extremes of the marine environment. This transition from a terrestrial ancestor to aquatic forms is well studied using standard paleontological methods (Figure 1), and increasingly so from a molecular and genomic perspective (Figure 3). Cetacea is a premier model of wholesale sensory evolution, in which anatomical changes related to sensory perception dovetail nicely with molecular evolutionary patterns. In addition, studies of cetacean genomes can be used as a standard to approach the analysis of other major macroevolutionary transitions in vertebrate evolution (Box 1). With the increasing availability of whole-genome sequences, future studies of secondary marine adaptation in disparate tetrapod taxa might show analogous molecular patterns of adaptive change and pseudogenization, as well as molecular convergence with Cetacea at the amino acid level. For example, detailed comparisons among a broader taxonomic sample of genomes can reveal large suites of genes that have undergone convergent recruitment and substitution among multiple clades of aquatic tetrapods. Over the past 30 years, Cetacea has emerged as an example of how phylogenetic analyses of molecules and morphology can be used to reconstruct major phenotypic transitions within a clade [5]. Recent genetic studies provide hope that comparative genomics ultimately might reveal the genetic underpinnings of these and other critical transitions in the Tree of Life.

References
1 Thewissen, J.G.M. and Bajpai, S. (2001) Whale origins as a poster child for macroevolution. *Bioscience* 51, 1037–1049
2 Futuyma, D.J. (2009) *Evolution*, Sinauer Associates
3 Zimmer, C. and Emlen, D.J. (2013) *Evolution: Making Sense of Life*, Roberts and Co
4 Gatesy, J. and O'Leary, M.A. (2001) Deciphering whale origins with molecules and fossils. *Trends Ecol. Evol.* 16, 562–571

5 Gatesy, J. *et al.* (2013) A phylogenetic blueprint for a modern whale. *Mol. Phylogenet. Evol.* 66, 479–506

6 O'Leary, M.A. *et al.* (2013) The placental mammal ancestor and the post-K-Pg radiation of placentals. *Science* 339, 662–667

7 Meredith, R.W. *et al.* (2011) Impacts of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification. *Science* 334, 521–524

8 Zhou, X. *et al.* (2011) Phylogenomic analyses and improved resolution of Cetartiodactyla. *Mol. Phylogenet. Evol.* 61, 255–264

9 Oelschläger, H.A. (1992) Development of the olfactory and terminalis systems in whales and dolphins. In *Chemical Signals in Vertebrates VI* (Doty, R.L. and Müller-Schwarze, D., eds), pp. 141–147, Plenum Press

10 Oelschläger, H.A. and Oelschläger, J.S. (2008) Brain. In *Encyclopedia of Marine Mammals* (2nd edn) (Perrin, W.F. *et al.*, eds), pp. 134–149, Academic Press

11 Deméré, T.A. *et al.* (2008) Morphological and molecular evidence for a stepwise evolutionary transition from teeth to baleen in mysticete whales. *Syst. Biol.* 57, 15–37

12 Thewissen, J.G.M. *et al.* (2006) Developmental basis for hind-limb loss in dolphins and origin of the cetacean bodyplan. *Proc. Natl. Acad. Sci. U.S.A.* 103, 8414–8418

13 Nummela, S. *et al.* (2004) Eocene evolution of whale hearing. *Nature* 430, 776–778

14 Thewissen, J.G.M. *et al.* (1994) Fossil evidence for the origin of aquatic locomotion in archaeocete whales. *Science* 263, 210–212

15 Thewissen, J.G.M. *et al.* (2007) Whales originated from aquatic artiodactyls in the Eocene epoch of India. *Nature* 450, 1190–1191

16 Gingerich, P.D. *et al.* (1983) Origin of whales in epicontinental remnant seas: new evidence from the Early Eocene of Pakistan. *Science* 220, 403–406

17 Thewissen, J.G.M. *et al.* (2001) Skeletons of terrestrial cetaceans and the relationship of whales to artiodactyls. *Nature* 413, 277–281

18 Uhen, M.D. (2010) The origin(s) of whales. *Annu. Rev. Earth Planet. Sci.* 38, 189–219

19 Gingerich, P.D. *et al.* (2001) Origin of whales from early artiodactyls: hands and feet of Eocene Protocetidae from Pakistan. *Science* 293, 2239–2242

20 Fitzgerald, E.M.G. (2006) A bizarre new toothed mysticete (Cetacea) from Australia and the early evolution of baleen whales. *Proc. R. Soc. B* 273, 2955–2963

21 Lindblad-Toh, K. *et al.* (2011) A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* 478, 476–482

22 Zhou, X. *et al.* (2013) Baiji genomes reveal low genetic variability and new insights into secondary aquatic adaptations. *Nat. Commun.* 4, 2708

23 Yim, H-S. *et al.* (2013) Minke whale genome and aquatic adaptation in cetaceans. *Nat. Genet.* 46, 88–92

24 The Bovine Genome Sequencing and Analysis Consortium (2009) The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science* 324, 522–528

25 Hayden, S. *et al.* (2010) Ecological adaptation determines functional mammalian olfactory subgenomes. *Genome Res.* 20, 1–9

26 Jiang, P. *et al.* (2012) Major taste loss in carnivorous mammals. *Proc. Natl. Acad. Sci. U.S.A.* 109, 4956–4961

27 Meredith, R.W. *et al.* (2009) Molecular decay of the tooth gene enamelin (*ENAM*) mirrors the loss of enamel in the fossil record of placental mammals. *PLoS Genet.* 5, e1000634

28 Meredith, R.W. *et al.* (2011) Pseudogenization of the tooth gene enamelysin (*MMP20*) in the common ancestor of extant baleen whales. *Proc. R. Soc. B* 278, 993–1002

29 Davies, K.T.J. *et al.* (2012) Parallel signatures of sequence evolution among hearing genes in echolocating mammals: an emerging model of genetic convergence. *Heredity* 108, 480–489

30 Meredith, R.W. *et al.* (2013) Rod monochromacy and the coevolution of cetacean retinal opsins. *PLoS Genet.* 9, e1003432

31 Mirceta, S. *et al.* (2013) Evolution of mammalian diving capacity traced by myoglobin net surface charge. *Science* 340, 1234192

32 Parker, J. *et al.* (2013) Genome-wide signatures of convergent evolution in echolocating mammals. *Nature* 502, 228–231

33 Berta, A. *et al.* (2005) *Marine Mammals: Evolutionary Biology*. (2nd edn), Academic Press

34 Christin, P.A. *et al.* (2010) Causes and evolutionary significance of genetic convergence. *Trends Genet.* 26, 400–405

35 Zakon, H.H. *et al.* (2006) Sodium channel genes and the evolution of diversity in communication signals of electric fishes: convergent molecular evolution. *Proc. Natl. Acad. Sci. U.S.A.* 103, 2675–2680

36 Hubbard, J.K. *et al.* (2010) Vertebrate pigmentation: from underlying genes to adaptive function. *Trends Genet.* 26, 231–239

37 Stewart, C.B. *et al.* (1987) Adaptive evolution in the stomach lysozymes of foregut fermenters. *Nature* 330, 401–404

38 Kornegay, J.R. *et al.* (1994) Molecular adaptation of a leaf-eating bird: stomach lysozyme of the hoatzin. *Mol. Biol. Evol.* 11, 921–928

39 Zhang, J. (2006) Parallel adaptive origins of digestive RNases in Asian and African leaf monkeys. *Nat. Genet.* 38, 819–823

40 Christin, P.A. *et al.* (2007) C4 Photosynthesis evolved in grasses via parallel adaptive genetic changes. *Curr. Biol.* 17, 1241–1247

41 Castoe, T.A. *et al.* (2009) Evidence for an ancient adaptive episode of convergent molecular evolution. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8986–8991

42 Chen, C-H.C. (1998) Evolution of diverse antifreeze proteins. *Curr. Opin. Genet. Dev.* 8, 715–720

43 Gatesy, J. *et al.* (2001) Extreme diversity, conservation, and convergence of spider silk fibroin sequences. *Science* 291, 2603–2605

44 Kooyman, G.L. and Ponganis, P.J. (1998) The physiological basis of diving to depth: birds and mammals. *Annu. Rev. Physiol.* 60, 19–32

45 Tyack, P.L. *et al.* (2006) Extreme diving of beaked whales. *J. Exp. Biol.* 209, 4238–4253

46 Nery, M.F. *et al.* (2013) Accelerated evolutionary rate of the myoglobin gene in long-diving whales. *J. Mol. Evol.* 76, 380–387

47 Dasmeh, P. *et al.* (2013) Positively selected sites in cetacean myoglobins contribute to protein stability. *PLoS Comput. Biol.* 9, e1002929

48 Li, Y. *et al.* (2010) The hearing gene prestin unites echolocating bats and whales. *Curr. Biol.* 20, R55–R56

49 Liu, Y. *et al.* (2010) Convergent sequence evolution between echolocating bats and dolphins. *Curr. Biol.* 20, R53–R54

50 Liu, Y. *et al.* (2010) Cetaceans on a molecular fast track to ultrasonic hearing. *Curr. Biol.* 20, 1834–1839

51 Shen, Y.Y. *et al.* (2012) Parallel evolution of auditory genes for echolocation in bats and toothed whales. *PLoS Genet.* 8, e1002788

52 Vater, M. and Kossl, M. (2004) The ears of whales and bats. In *Echolocation in Bats and Dolphins* (Thomas, J.T. *et al.*, eds), pp. 89–99, University of Chicago Press

53 Cranford, T.W. *et al.* (1996) Functional morphology and homology in the odontocete nasal complex: implications for sound generation. *J. Morphol.* 228, 223–285

54 Ketten, D.R. (1992) The marine mammal ear: specializations for aquatic audition and echolocation. In *The Evolutionary Biology of Hearing* (Webster, D.B. *et al.*, eds), pp. 717–750, Springer-Verlag

55 Fleischer, G. (1976) Hearing in extinct cetaceans as determined by cochlear structure. *J. Paleontol.* 50, 133–152

56 Griebel, U. and Peichl, L. (2003) Colour vision in aquatic mammals: facts and open questions. *Aquat. Mamm.* 29, 18–30

57 Peichl, L. (2005) Diversity of mammalian photoreceptor properties: adaptations to habitat and lifestyle? *Anat. Rec. A: Discov. Mol. Cell. Evol. Biol.* 287, 1001–1012

58 Levenson, D.H. and Dizon, A. (2003) Genetic evidence for the ancestral loss of short-wavelength-sensitive cone pigments in mysticete and odontocete cetaceans. *Proc. Biol. Sci.* 270, 673–679

59 Fasick, J.I. and Robinson, P.R. (2000) Spectral-tuning mechanisms of marine mammal rhodopsins and correlations with foraging depth. *Vis. Neurosci.* 17, 781–788

60 Sugawara, T. *et al.* (2010) Vertebrate rhodopsin adaptation to dim light via rapid Meta-II intermediate formation. *Mol. Biol. Evol.* 27, 506–519

61 Bischoff, N. *et al.* (2012) Deep-sea and pelagic rod visual pigments identified in the mysticete whales. *Vis. Neurosci.* 29, 95–103

62 Nei, M. *et al.* (2008) The evolution of animal chemosensory receptor gene repertoires: roles of chance and necessity. *Nat. Rev. Genet.* 9, 951–963

63 Cave, A.J.E. (1988) Note on olfactory activity in mysticetes. *J. Zool.* 214, 307–311

64 Thewissen, J.G.M. *et al.* (2011) Olfaction and brain size in the bowhead whale (*Balaena mysticetus*). *Mar. Mamm. Sci.* 27, 282–294

65 Rouquier, S. *et al.* (2000) The olfactory receptor gene repertoire in primates and mouse: evidence for reduction of the functional fraction in primates. *Proc. Natl. Acad. Sci. U.S.A.* 97, 2870–2874

66 Freitag, J. *et al.* (1998) Olfactory receptors in aquatic and terrestrial vertebrates. *J. Comp. Physiol. A* 183, 635–650

67 Kishida, T. *et al.* (2007) The olfactory receptor gene repertoires in secondary-adapted marine vertebrates: evidence for reduction of the functional proportions in cetaceans. *Biol. Lett.* 3, 428–430

68 McGowen, M.R. *et al.* (2008) The vestigial olfactory receptor subgenome of odontocete whales: phylogenetic congruence between gene-tree reconciliation and supermatrix methods. *Syst. Biol.* 57, 574–590

69 Pihlström, H. (2008) Comparative anatomy and physiology of chemical senses in aquatic mammals. In *Sensory Evolution on the Threshold: Adaptations in Secondarily Aquatic Vertebrates* (Thewissen, J.G.M. and Nummela, S., eds), pp. 95–109, University of California Press

70 Young, J.M. *et al.* (2010) Extreme variability among mammalian V1R gene families. *Genome Res.* 20, 10–18

71 Liman, E.R. *et al.* (1999) TRP2: a candidate transduction channel for mammalian pheromone sensory signaling. *Proc. Natl. Acad. Sci. U.S.A.* 96, 5791–5796

72 Yu, L. *et al.* (2010) Characterization of *TRPC2*, an essential genetic component of VNS chemoreception, provides insights into the evolution of pheromonal olfaction in secondary-adapted marine mammals. *Mol. Biol. Evol.* 27, 1467–1477

73 Chandrashekar, J. *et al.* (2006) The receptors and cells for mammalian taste. *Nature* 444, 288–294

74 Slijper, E.J. (1962) *Whales*, Basic Books

75 Kraus, P. *et al.* (2001) Some distal limb structures develop in mice lacking Sonic hedgehog signaling. *Mech. Dev.* 100, 45–58

76 Cooper, L.N. *et al.* (2007) Evolution of hyperphalangy and digit reduction in the cetacean manus. *Anat. Rec. (Hoboken)* 290, 654–672

77 Wang, Z. *et al.* (2009) Adaptive evolution of 5'HoxD genes in the origin and diversification of the cetacean flipper. *Mol. Biol. Evol.* 26, 613–622

78 Johnson, K.R. *et al.* (1998) A new spontaneous mouse mutation of Hoxd13 with a polyalanine expansion and phenotype similar to human synpolydactyly. *Hum. Mol. Genet.* 7, 1033–1038

79 Foot, N.J. *et al.* (2007) Positive selection in the N-terminal extramembrane domain of lung surfactant protein C (SP-C) in marine mammals. *J. Mol. Evol.* 65, 12–22

80 McGowen, M.R. *et al.* (2012) Dolphin genome provides evidence for adaptive evolution of nervous system genes and a molecular rate slowdown. *Proc. Biol. Sci.* 279, 3643–3651

81 Nery, M.F. *et al.* (2013) How to make a dolphin: molecular signature of positive selection in cetacean genome. *PLoS ONE* 8, e65491

82 Shen, Y.Y. *et al.* (2012) Genome-wide scan for bats and dolphin to detect their genetic basis for new locomotive styles. *PLoS ONE* 7, e46455

83 Sun, Y.B. *et al.* (2013) Genome-wide scans for candidate genes involved in the aquatic adaptation of dolphins. *Genome Biol. Evol.* 5, 130–139

84 Xu, S.X. *et al.* (2012) Positive selection at the ASPM gene coincides with brain size enlargements in cetaceans. *Proc. Biol. Sci.* 279, 4433–4440

85 Montgomery, S.H. *et al.* (2014) *ASPM* and mammalian brain evolution: a case study in the difficulty in making macroevolutionary inferences about gene–phenotype associations. *Proc. R. Soc. B* 281, 20131743

86 Chen, Z. *et al.* (2013) Characterization of hairless (*Hr*) and *FGF5* genes provides insights into the molecular basis for hair loss in cetaceans. *BMC Evol. Biol.* 13, 34

87 Kishida, T. and Thewissen, J.G.M. (2012) Evolutionary changes of the importance of olfaction in cetaceans based on the olfactory marker protein gene. *Gene* 492, 349–353

88 McGowen, M.R. *et al.* (2011) Phylogeny and adaptive evolution of the brain-development gene microcephalin (*MCPH1*) in cetaceans. *BMC Evol. Biol.* 11, 98