ARTICLE

(Discoveries)

Positive selection and inactivation in the vision and hearing genes of cetaceans

McGowen, Michael R.[1,2]; Tsagkogeorga, Georgia[1]; Williamson, Joseph[1]; Morin, Phillip A.[3]; Rossiter, Stephen J.[1]

[1] School of Biological and Chemical Sciences, Queen Mary, University of London, London, UK

[2] Department of Vertebrate Zoology, Smithsonian National Museum of Natural History, Washington, DC, USA

[3] Southwest Fisheries Science Center, National Marine Fisheries Service, NOAA, La Jolla, CA, USA

Corresponding authors:

Michael R. McGowen

(ORCID: 0000-0001-9192-3166)

mcgowenm@si.edu

Stephen J. Rossiter

s.j.rossiter@qmul.ac.uk

Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution 2020
This work is written by US Government employees and is in the public domain in the US.

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

ABSTRACT
The transition to an aquatic lifestyle in cetaceans (whales and dolphins) resulted in a radical transformation in their sensory systems. Toothed whales acquired specialized high-frequency hearing tied to the evolution of echolocation, while baleen whales evolved low-frequency hearing. More generally, all cetaceans show adaptations for hearing and seeing underwater. To determine the extent to which these phenotypic changes have been driven by molecular adaptation, we performed large-scale targeted sequence capture of 179 sensory genes across the Cetacea, incorporating up to 54 cetacean species from all major clades as well as their closest relatives, the hippopotamuses. We screened for positive selection in 167 loci related to vision and hearing, and found that the diversification of cetaceans has been accompanied by pervasive molecular adaptations in both sets of genes, including several loci implicated in non-syndromic hearing loss (NSHL). Despite these findings, however, we found no direct evidence of positive selection at the base of odontocetes coinciding with the origin of echolocation, as found in studies examining fewer taxa. By using contingency tables incorporating taxon- and gene-based controls, we show that, while numbers of positively selected hearing and NSHL genes are disproportionately high in cetaceans, counts of vision genes do not differ significantly from expected values. Alongside these adaptive changes, we find increased evidence of pseudogenization of genes involved in cone-mediated vision in mysticetes and deep diving odontocetes.

INTRODUCTION

Over the course of more than 50 million years, cetaceans (whales, dolphins, and porpoises) have transformed from a clade of terrestrial even-toed ungulates to mammals uniquely adapted for aquatic living (Gatesy and O'Leary, 2001). This transition included a radical reorganization of sensory systems that began upon entering an aquatic medium and continued through the diversification of lineages in modern cetaceans. Cetaceans acquired the improved ability to see in dim light, developed distinctive adaptations for the propagation and reception of sound underwater, and underwent reductions and losses in their chemosensory ability (McGowen et al., 2008; Hayden et al., 2010; Gatesy et al., 2013; Meredith et al., 2013; Feng et al., 2014; Kishida et al., 2015).

The differences in conduction of sound between air and water led to the development of improved underwater hearing early in cetacean history, as evidenced by changes in the ear of some of the first cetaceans such as *Pakicetus* and other archaeocetes (Nummela et al., 2004). Further evolution of both high-frequency hearing in toothed whales (Odontoceti) and low-frequency hearing in baleen whales (Mysticeti) occurred in the Late Eocene or Early Oligocene with the appearance of the two extant clades (Mourlam and Orliac, 2017). More attention has been paid to the evolution of high-frequency hearing, as this is tied to the appearance of echolocation, in which ultrasonic foraging clicks are emitted from the nasal passages and received through a heavily modified inner ear (Au, 1993; Churchill et al., 2016). Within odontocetes, high-frequency hearing has diversified as species adapted to distinct aquatic niches such as deep diving or riverine environments (Jensen et al., 2018; Galatius et al., 2019). For example, deep-diving beaked whales echolocate using frequency modulated clicks that are very distinct from those of other whales (Johnson et al., 2006; Jensen et al., 2018). Further morphological changes such as elaboration of the air sinuses, isolation of the bony ear from the skull, and increased skull asymmetry may have

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

helped to enhance echolocation in more derived odontocetes lineages such as oceanic dolphins (Fraser and Purves, 1960; Leduc, 2002). Studies of the molecular signatures of high-frequency hearing in toothed whales have revealed positive selection in genes expressed in the outer hair cells (such as *SLC26A5*) indicating a structural effect on high-frequency hearing (Li et al., 2010; Liu et al., 2010a, 2014, 2018). In *SLC26A5* and other hearing genes, convergent evolution has been documented between odontocetes and echolocating bats, emphasizing their importance in the development of ultrasonic hearing (Davies et al., 2012; Liu et al., 2010b, 2018; Li et al., 2010; Parker et al., 2013).

The aquatic environment also presented challenges for perception of light. Upon entering the water, cetaceans had to contend with lower light levels, and further evolution of deep diving forms led to less dependency on color vision in some groups (Mass and Supin, 2007; Meredith et al., 2013). For example, rod to cone densities in the retinas of many cetaceans resemble those of nocturnal terrestrial mammals (Peichl et la., 2001), and there is evidence of a shift toward perception of blue light in rhodopsin at the base of Cetacea (Meredith et al., 2013). All cetaceans have lost the shortwave-sensitive opsin (*OPN1SW*), and deep diving lineages (beaked whales, sperm whales) as well as some baleen whales, have also lost their longwave-sensitive opsin (*OPN1LW*), resulting in rod monochromatic vision (Peichl et al., 2001; Levenson and Dizon, 2003; Meredith et al., 2013). In addition, the spectral tuning of rhodopsin (*RHO*) has shifted even further toward blue light in deeper diving lineages (Fasick and Robinson, 2000; Meredith et al., 2013). Other cone-specific genes are inactivated in some baleen whales and sperm whales, providing further evidence for the loss of color vision in these lineages (Springer et al., 2016). However, most vision genes remain to be evaluated in the majority of cetacean species.

With the exception of the opsins and a handful of candidate hearing genes (Liu et al., 2010 a,b; Meredith et al., 2013; Springer et al., 2016; Dungan and Chang, 2017), large-scale analyses of selection targeting genes related to hearing and vision and incorporating a diverse group of cetaceans as well as their closest living relatives, the hippopotamuses, has not been attempted. Here we assembled a large dataset, using target sequence capture (McGowen et al., 2019), that targeted 179 genes involved in either hearing or vision from 54 cetacean species. We specifically investigated the presence of positive selection in hearing and vision genes across nodes within Cetacea and their closest relatives, the hippopotamuses. Many of these nodes have yet to be evaluated in the context of the molecular evolution of many hearing and vision genes. We predict that we would observe positive selection in hearing genes at the transition to the aquatic environment at the base of cetaceans, as well as the origin and elaboration of echolocation within odontocetes. We also predict that positive selection and/or pseudogenization in vision genes would occur at the base of cetaceans and in deep-diving lineages.

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

RESULTS

Our screens for loss-of-function of 179 protein-coding vision (Gene Ontology [GO] category GO:0007601 "visual perception") and hearing (GO:0007605 "sensory perception of sound") genes revealed eight genes in our dataset, all involved in vision, showing evidence of pseudogenization in at least two cetacean species: *OPN1SW*, *OPN1LW*, *CNGB3*, *CNGA3*, *GNAT2*, *PDE6C*, *GRK7*, and *GUCY2F* (Figure 1; Table 1). Of these loci, seven are expressed exclusively in cones (Emerling and Springer, 2014), while there is evidence of *GUCY2F* being expressed in both rods and cones (Karan et al., 2010). For one of these genes (*GUCY2F*), this is the first time that pseudogenes have been identified in cetaceans. Putative *GRK7* and *GUCY2F* pseudogenes are found within species of Ziphiidae, Kogiidae, Mysticeti, and Pontoporiidae (Table 1). The pseudogenization of *OPN1SW* independently at the base of mysticetes and odontocetes has been well characterized (Levenson and Dizon, 2003; Meredith et al., 2013; Figure 1), and will not be discussed in detail here. For the remaining genes, we identified additional species with pseudogenes based on frameshift mutations or stop codons: *OPN1LW* (*Balaenoptera borealis*, *B. edeni*, *Mesoplodon carlhubbsi*, *M. europaeus*), *CNBG3* (*M. mirus*, *Kogia sima*), *CNGA3* (*B. borealis, B. edeni, B. musculus, Eschrichtius robustus, Eubalaena glacialis, Hyperoodon ampullatus*, *K. breviceps, K. sima, M. carlhubbsi, M. europaeus, M. mirus*), *PDE6C* (*B. borealis, B. edeni, B. musculus, C. marginata, Eubalaena* spp., *M. ginkgodens*), and *GNAT2* (*C. marginata*, *E. robustus*, *M. stejnegeri*). Two frameshift mutations were identified at the end of the reading frame of *CNGA3* sequences in three *Phocoena* species (deletion of 2 bp), as well as *Feresa attenuata* (insertion of 1 bp); however, this would only serve to interrupt the reading frame of five amino acids and lengthen the resulting protein. There was some evidence of shared inactivating mutations in a few genes, for example *B. borealis* + *B. edeni* and *M. densirostris, M. europaeus,* and *M. gingkodens* in *GRK7*, and mutations in three sets of mysticetes that each do not form a monophyletic lineage in *PDE6C* and *CNGA3*. With the addition of *B. borealis* and *B. edeni* here, we confirm that a mutation in Exon 2 of *OPN1LW* identified by Meredith et al. (2013) is found in all Balaenopteroidea (Table 1).

We tested for positive selection on 167 hearing, nonsyndromic hearing loss (NSHL), and/or vision genes on five 'ingroup' branches (Cetacea, Mysticeti, Odontoceti, Ziphiidae, Delphinidae) based on predictions of where molecular changes in these genes may have occurred using the branch-sites Model A implemented in PAML 4 (Yang, 2007) and compared these to three 'outgroup' branches (Whippomorpha, Ruminantia, Cetruminantia) in which we do not expect to find changes in vision and/or hearing. We found 17 genes (nine vision, five hearing, three vision/hearing) under positive selection on at least one of five 'ingroup' branches and 14 genes (seven vision, six hearing, one vision/hearing) on the three 'outgroup' branches (Table 2; Table S1). Among genes involved in nonsyndromic hearing loss found to be under positive selection, three were on 'outgroup' branches (*MYO1A*, *OTOG*, *WFS1*) and four were on ingroup branches (*CDH23*, *LOXHD1*, *PCDH15*, *TECTA*). There were as many as seven genes found to be under positive selection on the branch leading to the deep-diving Ziphiidae as well as the terrestrial lineage Ruminantia, while surprisingly no genes were found to have undergone positive selection on the odontocete branch, where the transition to high-frequency hearing is estimated to have occurred. After correcting for false positives, four out of 30 genes were found to be statistically significant: *RBP3* in Whippomorpha (vision), *RP1* in Mysticeti (vision), *LOXHD1* in Ziphiidae (hearing; NSHL), and *PCDH15* in Delphinidae (hearing with a possible minor role in vision; NSHL).

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

For comparison, we also tested for positive selection in a random set of 167 genes that have no known major roles in hearing or vision. We implemented branch-sites Model A on the same five 'ingroup' and three 'outgroup' branches using PAML. We found episodes of positive selection on 17 genes, four of which were found to be under selection on two separate branches (Table 3; Table S2). As many as six genes were found to be under selection on the ruminant branch, while no selection was found to occur on two branches examined, Odontoceti and Delphinidae (Table 3). After correcting for false positives, only two of these were found to be significant, the kidney-related gene *CLCN5* in Ruminantia and the hormone receptor *FSHR* in Cetacea.

We also used aBSREL to simultaneously test for positive selection on hearing and vision genes across all internal branches of the tree. Within the clade of interest (Cetacea) and along the branch leading to crown Cetacea, we identified 79 episodes of positive selection on internal branches in 53 vision or hearing genes (Table S3). After correcting for multiple testing, this count was reduced to nine episodes of positive selection in seven genes, of which two are related to vision (*RP1, LUM*) and five to hearing (*CACNA1D, CDH1, LOXHD1, OTOG, PCDH15*), three of which are specifically related to nonsyndromic hearing loss (*LOXHD1*, *OTOG*, *PCDH15*). Two of these latter genes were found to be evolving under positive selection on two separate lineages, *PCDH15* (Delphinidae, Delphinoidea) and *LOXHD1* (Ziphiidae, Plicogulae). Three of these episodes of positive selection were also identified when using PAML after FDR correction (*PCDH15,* Delphinidae; *LOXHD1*, Ziphiidae; *RP1,* Mysticeti). This contrasted with 39 episodes of positive selection in 29 vision or hearing genes on internal branches outside of Cetacea and its stem lineage (Table S3). After correcting for multiple testing, only two genes, one vision (*LUM*) and one predominantly hearing, NSHL gene (*CDH23*) were found to be under positive selection on the lineage leading to the last common ancestor of Cetruminantia and Hippopotamidae respectively (Table S3).

We mapped all identified episodes of positive selection using the aBSREL method on internal branches of Cetruminatia for hearing (Figre 2A) and vision genes (Figure 2B). Colors of branches represent the cumulative number of genes that experienced episodes of positive selection. For both vision and hearing genes, the cumulative number of genes increased with number of branches, with the greatest number of genes under selection occurring on the lineage leading to Delphinoidea and its constituent families (Delphinidae, Phocoenidae, Monodontidae); this is especially evident in hearing genes, particularly those related to nonsyndromic hearing loss which are highlighted in red (Figure 2A). Although genes were observed as under selection on all branches outside of Cetacea (Whippomoprha, Hippopotamidae, Ruminantia, Cetruminatia), there were more vision genes than hearing genes under selection on each branch. Four genes involved in hearing (*HEXB*, *CACNB2*, *ESPN*, *TECTA*; Figure 2A) and four genes involved in vision (*CACNB2*, *POU6F2*, *RHO*, *RPGRIP1*; Figure 2B) are identified as under selection at the base of cetaceans before FDR correction. However, not one gene was found to have undergone positive selection at the base of Odontoceti, where echolocation and high-frequency hearing were believed to have originated. Four hearing genes each were identified on four additional branches (Mysticeti, Physeteroidea [Physeteridae + Kogiidae], Ziphiidae, Delphinidae), and three each on branches leading to Hippopotamidae, Delphinida and Delphinoidea (Figure 2A). Four vision genes were identified on the Hippopotamidae branch and three genes each on the Whippomorpha, Delphinida and Delphinoidea branches (Figure 2B).

For our control set of 167 non-sensory genes, aBSREL identified 30 episodes of positive selection in 21 genes within Cetacea or along the branch leading to crown Cetacea (Table S3).

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

After correcting for multiple testing, this count was reduced to five genes (*CEP89*, *FSHR*, *GPR50*, *MMP4*, *KCNC2*) in five separate lineages (Table S3). Only one these events, *FSHR* in Cetacea, was found to be significant after FDR correction using PAML (Table 3). This contrasted with 27 episodes of positive selection in 24 genes in branches outside Cetacea and its stem lineage. As with hearing and vision genes, we then mapped all identified episodes of positive selection using the aBSREL method on internal branches of Cetruminatia (Figure 3). For non-sensory genes, the ruminant branch had by far the greatest number of genes under positive selection, with as many as ten. In contrast, for branches within cetaceans, not more than two genes were found to be under selection (Figure 3).

We compared numbers of sensory and non-sensory genes under selection across both ingroup and outgroup lineages using two approaches based on three-dimensional contingency tables. We found that in cetaceans, numbers of loci under selection were disproportionately high for both hearing ($p = 0.0151$, Breslow-Day test; $p = 0.0144$, Log-linear model) and NSHL genes ($p = 0.0006$, Breslow-Day test; $p = 0.0004$, Log-linear model) in relation to expected values based on our sets of controls. Post hoc investigations found that an additional 29 (or 28 using Breslow-Day tests) cetacean control genes or 6 ungulate NSHL genes would have to be positively selected for NSHL tests to lose significance (Figures S1, S2). Similarly, an additional 6 cetacean control genes or 3 outgroup-lineage hearing genes would have to be positively selected for hearing tests to lose significance (Figures S1, S2). Note that additional genes come from the pool of non-PSGs, meaning that the number of genes listed here is doubled when considering net difference between two counts. Numbers of vision genes under selection did not differ from expectations ($p = 0.1251$, Breslow-Day test; $p = 0.1248$, Log-linear model).

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

DISCUSSION

Our results from 179 loci across an average of 52 cetaceans and their relatives provides the first comprehensive picture of the molecular evolution of hearing and vision genes during the origin and diversification of cetaceans. In doing so, we have expanded on earlier studies that, due to the limited availability of genomic resources, have tended to be taxonomically-limited in scope and often examining just one cetacean species (Davies et al., 2012; McGowen et al., 2012; Sun et al., 2012; Nery et al., 2013; Parker et al., 2013; Yim et al., 2014), or otherwise have focused on a specific gene or set of genes (Levenson and Dizon, 2003; Li et al., 2010; Liu et al., 2010a, 2010b; Meredith et al., 2013; Springer et al., 2016; Dungan and Chang, 2017; Liu et al., 2018). Here we found that positive selection in both hearing and vision genes occurs across the cetartiodactyl tree and is not specific to cetaceans. However, the greater proportion of positive selection events in cetaceans appears to be significant for hearing genes when compared to a set of control genes, especially those involved in nonsyndromic hearing loss. Indeed, genes involved in nonsyndromic hearing loss tend to have a more direct effect on hearing when compared to many other genes classified in GO category GO:0007605 "sensory perception of sound", such as *CDH1* or *HEXB,* which have extensive pleiotropic effects (Sango et al., 1995; Gall and Frampton, 2013). Positive selection on vision genes does not seem to correspond with the reduction in color vision seen in cetaceans, although we identified more incidences of pseudogene formation in cone-specific genes in deep diving odontocetes and mysticete whales.

Vision Pseudogenes in Cetaceans

We provide further evidence of the degradation of genes expressed within cones in some cetacean lineages, including the deep diving beaked whales (Ziphiidae) and sperm whales (Physeteridae, Kogiidae), as well as baleen whales (Mysticeti). We add the gene *GRK7* to the list of genes known to be pseudogenized in these three cetacean lineages; a previous analysis also identified the sperm whale (*Physeter microcephalus*) and common minke whale (*Balaenoptera acutorostrata*) as lacking a functional *GRK7* (Emerling, 2018). The protein product of *GRK7* phosphorylates cone opsins in many mammalian species, and there is evidence of species-specific expression differences of *GRK7* in cones (Liu et al., 2005). For example, *GRK7* is co-expressed with *GRK1* in the cones of primates, while only *GRK7* and not *GRK1* is expressed in cones from the domestic pig and dog, both laurasiatherians along with the cetaceans (Weiss et al., 2001; Osawa and Weiss, 2012). Other mammals, many of which are active in dim-light environments, such as mice and subterranean rodents, golden moles, xenarthrans, some bats, and tenrecs, also have lost *GRK7* (Weiss et al., 2001; Emerling and Springer, 2014, 2015; Fang et al., 2014; Hudson et al., 2014; Emerling, 2018). The loss of function in *GRK7* and other cone-specific genes identified here (*OPN1LW*, *CNGB3*, *CNGA3*, *GNAT2*, *PDE6C*) indicates that many more cetacean species have lost cone-mediated vision than previously known, including *Balaenoptera borealis, B. edeni, Caperea marginata*, *Kogia sima*, *Berardius bairdii*, *Hyperoodon ampullatus*, and many if not all species of *Mesoplodon*. The addition of *Caperea* to the list of rod monochromats raises the possibility that cone-mediated vision disappeared before the diversification of Mysticeti, with cone-expressed genes gradually accumulating inactivating mutations in parallel lineages over time. Based on this and previous studies, a total of ten cone-specific phototransduction genes (*ARR3*, *CNGA3*, *CNGB3*, *GNAT2*, *GNGT2, GRK7, PDE6C*, *PDE6H, OPN1SW*, *OPN1LW)* are pseudogenes in at least one cetacean (Levenson et al., 2002; Meredith et al., 2013; Springer et al., 2016; Emerling, 2018). We did not find any pseudogenization events in another cone-specific

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

gene, *SLC24A2*, which is also retained in all subterranean mammals evaluated (Emerling, 2018) and could indicate another function besides cone-mediated phototransduction. The correlation between loss of functional cone-specific genes and rod monochromatism in cetaceans highlights their value as a model for visual diseases in humans involving cone receptors (Emerling et al., 2017).

We also discovered the presence of frameshift mutations in the *GUCY2F* gene of some cetaceans, including those that have not been identified before as having mutations in cone-specific genes, such as the franciscana *Pontoporia blainvillei. GUCY2D* and *GUCY2F* both code for membrane-bound guanylate cyclases, known to be involved in the synthesis of cyclic GMP and expressed in both the rods and cones of the mammalian retina (Baehr et al., 2007; Karan et al., 2010). *GUCY2F* but not *GUCY2D* is inactivated in some subterranean mammals such as the naked mole rat and the Cape golden mole, (Emerling and Springer, 2014) as well as fourteen out of thirty mammals examined including ground squirrels and manatees (Emerling, 2018). In double knockouts of both *GUCY2D* and *GUCY2F* mice, the phototransduction cascade ceases, but in *GUCY2F* knockouts there is no noticeable effect on rod or cone physiology (Karan et al., 2010). Indeed, in cows, presence of GUCY2D in rods was found at levels 25-fold higher than GUCY2F. Perhaps GUCY2D is able to compensate for a nonfunctional GUCY2F, resulting in minimal effects on normal rod or cone photoreception.

Evolution of Vision Genes in Cetartiodactyla

Alongside evidence of degradation in some vision genes, we detected extensive positive selection of vision genes across internal branches of Cetartiodactyla. In general, the proportion of positive selection events did not differ significantly from a set of randomly selected 'control' genes, and there is no evidence that positive selection is associated with dim light levels in an aquatic environment or increased dependence on rod-mediated vision. We see substantial evidence of positive selection events on branches outside of cetaceans, particularly on the ruminant and hippopotamid lineages. The genes that passed the FDR correction using either model (aBSREL, branch-sites Model A in PAML) include lineages outside Cetacea (*LUM*, *RHO*, *RBP3*, *TULP1*) as well as within (*RP1*, *LUM*) (Table 1). Three of these genes are associated with retinitis pigmentosa (*RBP3, RP1*, *TULP1*), a retinal degenerative disease that usually results in reduced sight in dim light and reduction of the visual field (Hamel, 2006; den Hollander et al., 2009; Arno et al., 2015). However, only one of these genes (*RP1*) can be associated with a transition to a dimmer light environment (in this case mysticetes; Figure 2B). *RP1* shows signs of positive selection in other mammals which rely on vision in dim-light environments, including mole rats (Davies et al., 2015) and bats (Parker et al., 2013; Liu et al., 2015). In contrast, *RBP3* has been identified as a pseudogene in subterranean and nocturnal mammals that have reduced visual systems, such as the naked mole rat, marsupial mole, and some species of echolocating bats (Shen et al., 2013). Another visual gene, *LUM,* was found to be under selection following FDR correction at the base of *Mesoplodon* within Ziphiidae using the aBSREL method and at the base of Cetruminantia using both aBSREL and PAML methods. *LUM* is highly expressed in the cornea and sclera and is associated with myopia, as well as corneal thickness and opacity (Iglesias et al., 2018). We found significant positive selection in rhodopsin (*RHO*) on the ruminant branch (Figure 2). Although not significant after FDR correction in cetaceans, *RHO* was found to be under positive selection at the base of cetaceans as well as within Mysticeti in the aBSREL analyses (Figure 2). Previous analyses

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

of *RHO* have found evidence of both site-specific and clade-specific positive selection within Cetacea (Dungan et al., 2016).

The Evolution of Hearing Genes in Cetartiodactyla

We found significantly more hearing genes under positive selection in cetaceans in comparison to a group of control genes that are not involved in the perception of sound. Many of the genes found to be under selection are involved in nonsyndromic hearing loss (NSHL), and have documented mutations that are centered on auditory pathologies and do not have significant effects elsewhere in the body (Shearer et al., 2017). Many of these NSHL genes are expressed in the outer hair cells, which are the primary sites of sound amplification and enable high-frequency hearing (Oghalai, 2004). Below we discuss the potential effects of these genes; however, the actual effects of these modifications remain to be tested at the functional level.

Consistent with undergoing evolutionary changes in the auditory system on transitioning to an aquatic niche, we found evidence of positive selection on the ancestral cetacean branch in four hearing genes using aBSREL, two of which were also identified using branch-sites tests in PAML, although none passed FDR correction (Figure 2B; Table 2). One of these genes, *TECTA*, codes for alpha-tectorin, a major glycoprotein component of the tectorial membrane of the inner ear. Stereocilia of the sensory hair cells are imbedded in the tectorial membrane, which has a major function in cochlear amplification of sound (Legan et al., 2005; Dewey et al., 2018). As degree of stiffness is important for the function of wave propagation across the tectorial membrane (Richardson et al., 2008), structural changes in *TECTA* could have consequences on inner ear function. Recently, *TECTA* was found to be under selection at the base of pinnipeds, another branch on which there was a transition from terrestrial to aquatic living, although neither group shared amino acid changes (Park et al., 2018). Another of the loci showing molecular adaptation at the root of cetaceans was *ESPN* (Figure 2B), which encodes the protein espin. This protein has roles in the elongation of actin bundles in inner ear stereocilia, and mutations in the coding gene are implicated in non-syndromic hearing loss and vestibular dysfunction (Donaudy et al., 2006).

In contrast to these results, and contrary to our expectations, we found no evidence of positive selection in any hearing genes coinciding with the evolution of high-frequency hearing at the base of odontocetes. This absence of evidence for positive selection based on our large taxonomic dataset disagrees with some earlier studies that discovered signals of selection on this branch for the gene *SLC26A5*, which encodes the outer hair cell protein prestin (Liu et al., 2010 a,b; Parker et al., 2013). We do, however, see evidence of positive selection in *SLC26A5* on the deep-diving physeteroid lineage (Figure 2A), whose members contain species that emit distinct multi-pulse (*Physeter*) and narrow band (*Kogia*) biosonar signals (Jensen et al., 2018). Even though we did not discover evidence of positive selection on the odontocete lineage in prestin here, structural studies have confirmed that specific convergent amino acids in echolocating bats and toothed whales have created profound changes in the structure of prestin that are ultimately linked to the development of ultrasonic hearing (Liu et al., 2018). Through reconstruction and experimentation with ancestral prestin proteins, Liu et al. (2018) also inferred that the last common ancestor of cetaceans was unlikely to hear ultrasonic sounds as in modern odontocetes, adding support that high-frequency hearing exclusively evolved in odontocetes.

Given that all toothed whales use echolocation, it is perhaps surprising that none of the hearing genes examined in this study showed evidence of positive selection on the branch leading

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

to this suborder (Figure 2A). On the other hand, we found many genes under selection on major lineages within odontocetes, perhaps suggesting a time lag between the development of ultrasonic hearing and diversification of cetacean lineages. Alternatively, it may be the case that many positively selected genes are ultimately associated with major transitions in odontocetes, such as the move to deep water environments (Ziphiidae, Physeteridae, Kogiidae) or the further elaboration of the echolocation system in delphinidans (Delphinidae, Phocoenidae, Monodontidae, Iniidae, Pontoporiidae, Lipotodae). Indeed, delphinidan taxa show some of the highest call frequencies in cetaceans, requiring modifications to the sound detection apparatus (Jensen et al., 2018). At least nine of these genes (*ADGRV1*, *CLIC5*, *LOXHD1*, *PCDH15*, *SLC26A5*, *STRC*, *TMC1*, *USH1C, WHRN*) are expressed in the stereocilia of outer hair cells (Mburu et al., 2003; Dallos, 2008; Gagnon et al., 2006; Kazmierczak et al., 2007; Verpy et al., 2008; Grillet et al., 2009; Kurima et al., 2015; Yan et al., 2010; Yan et al., 2018), which are involved in the amplification of sound and enhancement of frequency sensitivity (Dewey et al., 2018). A number of these loci have previously been reported to show selection and/or convergence in echolocating taxa (e.g. *TMC1*, Davies et al. 2012; Marcovitz et al., 2017).

One of these stereocilia genes, *LOXHD1,* has been linked to progressive non-syndromic hearing loss in humans, although the distinct function of its protein product is not yet known (Grillet et al., 2009). *LOXHD1* was found to be upregulated in an echolocating bat (*Myotis ricketti*) as compared to a non-echolocating bat (*Cynopterus sphinx*) (Dong et al., 2013), inferring a possible function in high-frequency hearing. We found signals of positive selection in *LOXHD1* throughout the cetacean tree, both at nodes within echolocating odontocetes (Ziphiidae, Delphinoidea, Phocoenidae) and nodes within the low-frequency hearing mysticetes (Plicogulae, Balaenopteroidea), although only selection along the lineages leading to Plicogulae and Ziphiidae were significant after FDR correction. We found evidence of as many as seven amino acid changes exclusive to Ziphiidae. In comparison, although the lineage leading to Plicogulae was found to be under selection (Figure 2A), only one amino acid change may have been exclusive to this clade (Table 1). It is tempting to suggest that these significant changes in ziphiids coincide with the origins of deep diving and frequency modulated echolocation in this clade, but without more information concerning the function of *LOXHD1*, this is only speculation.

Two of the genes in which we detected selection (*CDH23* and *PCDH15*) encode proteins that interact to form tip-link filaments, which connect stereocilia and are crucial for mechanoelectrical transduction (Kazmierczak et al., 2007). The tip link filament acts as a tether when stereocilia are deflected due to sound, transmitting mechanical force to open ion gated channels at the tip of stereocilia (Dionne et al., 2018). The tip link consists of two homodimers each of CDH23 and PCDH15 proteins, bound together at their terminal ends (Kazmierczak et al., 2007; Dionne et al., 2018). Mutations in each of these genes cause inherited forms of deafness in humans (Ahmed et al., 2001; Bolz et al., 2001). Two individual branches show evidence of positive selection in both *PCDH15* and *CDH23*, Delphinida and Delphinidae (Figure 2A), although *CDH23* is not significant after FDR correction. At least nine exclusive amino acid changes occurred in both *PCDH15* and *CDH23* in Delphinida and at least eight occurred in each gene in Delphinidae; PAML shows evidence of 11 specific amino acid sites under selection on the delphinid branch in *PCDH15*, although three of these do not have changes exclusive to Delphinidae (Table 2). Although *CDH23* is also under selection on three other branches (Hippopotamidae, Whippomorpha, Physeteroidea), the parallel selection at Delphinidae and Delphinida in two genes with interacting products indicates a possible change in tip-link function that benefits reception of high-frequency sound. Both *PCDH15* and/or *CDH23* have been

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

identified as under selection and convergent in cetaceans and echolocating bats, but in these analyses only the bottlenose dolphin sequence was tested (Shen et al., 2012; Parker et al., 2013). Here we find evidence that selection on *PCDH15* and *CDH23* did not occur at the origin of cetaceans or odontocetes, but at branches higher up in the tree, at the origins of Delphinida and Delphinidae, clades where the bony ear is increasingly isolated from the skull and accessory air sacs are further elaborated (Fraser and Purves, 1960; LeDuc, 2002).

We also found evidence for positive selection in several hearing genes in mysticetes and hippopotamids, neither of which echolocate but are known to produce and receive infrasonic sound, an unusual ability among cetartiodactyls. Evidence suggests that both mysticetes and hippopotamuses separately developed their ability to hear low-frequencies (Mourlam and Orilac, 2017), although the evolution of infrasonic sound perception in these groups has received far less attention than the acquisition of echolocation in odontocetes. Four and three genes each were found to be under selection on the mysticete and hippo lineages, respectively, with only *CDH23* in hippopotamuses found to be significant after FDR correction. Of these seven genes, only four (*OTOF*, *USH1G*, *GRXCR1*, *CDH23*) are directly involved in mechanosensory transduction in the inner ear (Kazmierczak et al., 2007; Johnson and Chapman, 2010; Odeh et al., 2010; Yan et al., 2010). It is thus especially noteworthy that both *OTOF* and *USH1G* have both been implicated in low-frequency hearing loss (Tekin et al., 2005; Varga et al., 2006; Strenzke et al., 2016; Gallego-Martinez et al., 2019). More work is needed to assess these loci across other taxonomic groups that rely on infra-sound.

Our results suggest the molecular evolution of vision in cetaceans has involved pervasive pseudogenization of cone-adapted genes in deep diving lineages and mysticetes. We found evidence of pseudogenes in genes expressed in cones in the pygmy right whale, which infers that the evolution of rod monochromatic vision may have occurred before the diversification of mysticetes. Here we found that positive selection in both hearing and vision genes occurs across the cetartiodactyl tree and is not specific to cetaceans. However, the greater proportion of positive selection events in cetaceans appears to be significant for hearing genes when compared to a set of control genes, especially those involved in nonsyndromic hearing loss. We find the evolution of hearing genes is more complex than a direct association with echolocation, as we find multiple genes under positive selection across the tree that may be associated with the evolution of low-frequency hearing in hippos and mysticetes, deep diving in beaked whales, and the elaboration of the echolocation system in delphinidans and delphinid dolphins. Surprisingly, with the expansion of cetacean species included in our analysis, we found no direct evidence of positive selection at the base of odontocetes coinciding with the origin of echolocation. Future studies are needed to test these hypotheses by using functional analyses of resurrected proteins.

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

METHODS

We obtained alignments of 179 protein-coding hearing and vision genes from a recent phylogenomic study of cetaceans, in which target sequence capture of Illumina genomic libraries was used to generate a dataset comprising 38,167 exons from 3,191 genes and combined with existing genomic data for a dataset consisting of 100 individuals from 77 cetaceans and 12 outgroup species, including both hippopotamus species, nine other terrestrial cetartiodactyls, and the perissodactyl *Equus caballus* (McGowen et al., 2019). Details of how these sequences were obtained and how alignments were constructed can be found in McGowen et al. (2019). For the molecular analyses performed here, we pared down the number of individuals and species from McGowen et al. (2019). For all alignments, we used only one individual per species, retaining the individual with the most complete sequence. From each alignment, we also removed any species with >50% missing data (with the exception of both hippopotamids) for a total of 20 to 67 taxa per alignment ($\bar{x}$=52). A list of all species used in this study (as well as locality data) is shown in Table S4. Raw reads for this dataset were deposited in the Sequence Read Archive (SRA) of NCBI, BioProject PRJNA575269. All alignments used in this study are deposited in Dryad (doi:10.5061/dryad.63xsj3v05).

We identified the 179 hearing and vision genes from two GO categories, GO:0007605 "sensory perception of sound" or GO:0007601 "visual perception", as identified using AmiGO 2 (Carbon et al., 2009); we included genes in each category from identifications based on either *Mus musculus* or *Homo sapiens*. This included 84 genes involved in sound perception and 108 involved in visual perception with 14 involved in both. Of these genes, 33 are classified as contributing in nonsyndromic hearing loss (NSHL), meaning these genes have mutations that contribute to hearing loss with no discernible effect on other systems (Shearer et al., 2017). A description of all genes involved in this study are listed in Table S5, including notation of genes related to nonsyndromic hearing loss.

For all taxa, we screened each locus for the presence of stop codons, as well as for insertions or deletions that led to an interrupted reading frame, indicative of pseudogenization and thus loss of function. For each exon and species for which we discovered a potential pseudogenization event, we mapped raw reads using default settings in Geneious Prime 2019.1.3 (https://www.geneious.com) to the associated *Orcinus orca* exon derived from assembly Oorc_1.1 (Foote et al., 2015) to verify the event and examine its coverage. For further analyses of molecular evolution, genes identified as being pseudogenes in at least one cetacean were excluded, as well as genes with protein coding regions of less than 450 bp, leaving a total of 167 genes for further analysis.

To test for evidence of positive selection (as indicated by estimates of $dN/dS > 1$) on key branches in the cetartiodactyl tree, we implemented branch-sites model A using the codeml package of PAML version 4.8 (Yang, 2007) for both sets of genes. These were conducted separately for five 'ingroup' branches: Cetacea, Mysticeti, Odontoceti, Ziphiidae, and Delphinidae. These nodes were selected to test whether potentially adaptive changes in genes associated with sensory evolution occurred, respectively, upon invasion of the marine realm (Cetacea), upon acquisition of echolocation abilities (Odontoceti), upon evolution of filter-feeding and low-frequency hearing (Mystceti), upon the evolution of extreme deep-diving and frequency-modulated echolocation (Ziphiidae), and upon the evolution of a complex sinus system and extreme cranial asymmetry (LeDuc, 2002; Jensen et al., 2018; Delphinidae). In addition, we also

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

conducted branch-sites PAML analyses on branches outside our ingroup which have not been associated with transitions to an aquatic environment or potential shifts in sensory evolution including Whippomorpha, Ruminantia, and Cetruminatia.

We used a likelihood ratio test (LRT) with one degree of freedom to test whether each branch-sites model was significant as compared to a null model. For significant models, we identified specific sites under positive selection within each gene using Bayes Empirical Bayes posterior probabilities of >0.5, following Tsagkogeorga et al. (2015). We used the topology of the concatenated tree from Figure 2 of McGowen et al. (2019) with all species pruned that were not in each gene alignment. We conducted analyses using branch-sites and null models at least five times for each branch and each gene, taking the maximum value of both the alternative and null models for input into LRTs; all LRTs with negative values were set to 0 following Daub et al., (2017). Associated *p*-values from LRTs were corrected for multiple testing using the false discovery rate (FDR) with a significance value of $q < 0.10$ (Benjamini and Hochberg, 1995).

To gain additional insights into potential changes in selection pressure in candidate loci over the course of evolution of modern cetaceans, we also conducted tests of positive selection for each of the 167 genes using the adaptive branch-site random effects likelihood (aBSREL) model, implemented in HyPhy via Datamonkey 2.0 (Kosakovsky Pond et al., 2005; Weaver et al., 2018). This model allows selective pressures to vary across sites and branches simultaneously across a phylogenetic tree (Smith et al., 2015) and has the additional benefit that it takes into account each branch length and adapts the complexity of the model to the length of each branch. Therefore, we were able to test for selection on specific sites across every branch simultaneously without the need to identify branches of interest *a priori*. We used the same topology as in the PAML analyses described above. Episodes of inferred positive selection for both vision and hearing genes were mapped onto a phylogeny of cetaceans with topology derived from McGowen et al. (2019).

To compare our dataset with genes that have no predicted function in the auditory or visual systems, we also randomly selected 167 loci ("non-sensory genes") from the remaining 3,012 protein-coding genes within the set of loci sequenced by McGowen et al. (2019). Alignments were modified as above, and we performed the same sets of analyses using branch-site tests in PAML on our five ingroup lineages (Cetacea, Mysticeti, Odontoceti, Ziphiidae, Delphinidae) and three outgroup lineages (Cetruminantia, Ruminantia, Whippomorpha), as described above. We also used the aBSREL model in HyPhy to simultaneously examine selective pressure across the tree.

To assess whether numbers of sensory genes were disproportionately high in cetaceans, we constructed contingency tables containing observed counts of PSGs and non-PSGs from the HyPhy study. As a control in each table, we included data from non-sensory genes, and therefore accounted for any differences in the background rates of detection of positive selection. Such differences can occur due to sampling effort, tree length and topology. By including the outgroup lineage within the analysis, we also account for the fact that hearing and vision genes may be prone to higher levels of positive selection, regardless of the specific sensory adaptation of cetaceans to aquatic lifestyles. We specifically constructed contingency tables of selection (PSG vs. non-PSG) against function (hearing, vision or NSHL vs. non-sensory) for each group and combined these into 2 x 2 x 2 tables. Breslow-Day tests for homogeneity (*R* package *DescTools* (Signorell et al., 2020)) were carried out on each of the three tables to test for differing odds-ratios between the table strata. Differing odds ratios are demonstrative of a 3-way interaction between clade, function and selection. Log-linear models were also carried out on the same tables, with the saturated model compared to a model lacking a 3-way interaction term between clade, selection and function (*R*

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

package *MASS* (Venables & Ripley, 2002)). *Post hoc* data simulations were carried out on significant results to examine how changing the control and ungulate PSG ratios influenced test significance.

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

ACKNOWLEDGEMENTS

This research was funded by a Royal Society Newton International Fellowship to M.R.M. and a European Research Council Starting grant (310482 [EVOGENO]) grant to S.J.R. We would like to thank Ryan Schott, Jason de Koning, and two anonymous reviewers for providing comments on earlier versions of this study. For samples, we are grateful to R. Deaville and P. Jepson (Zoological Society of London), A. Polanowski (Australian Antarctic Division), and M. Bertelsen (Copenhagen Zoo), as well as K. Robertson (Southwest Fisheries Science Center [SWFSC]) for extractions and her help in obtaining approval from some of the original sample donors. We would like to acknowledge L. Balance, J. Barlow, K. Forney (SWFSC), T. Jefferson (Clymene Enterprises), the late P. Best (South African Museum), C. S. Baker (Oregon State University), R. Small and L. Quackenbush (Alaska Department of Fish and Game), E. Oleson (NMFS, Pacific Islands Fishery Science Center), K. West (University of Hawai'i), M. P. Heide Jørgensen (Greenland Institute of Natural Resources), and D. Kreb (Yayasan Konservasi RASI) for the use of samples obtained through SWFSC. We thank the Taranaka Iwi and the New Zealand Cetacean Tissue Archive (NZCeTA), initiated by C.S. Baker and M.L. Dalebout, and maintained by the University of Auckland (R. Constantine, curator) for directly providing the sample of *Mesoplodon gingkodens* and the help of D. Steel (Oregon State Univ.) for obtaining this sample. For all samples derived from Tasmania, we acknowledge Kris Carlyon and Rachael Alderman (Princess Melikoff Trust, Marine Conservation Program, Tasmanian Department of Primary Industries, Parks, Water and Environment [DPIPWE]). We thank C. Buell for the illustrations and J. Gatesy for their use.

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

LITERATURE CITED

Ahmed ZM, Riazuddin S, Bernstein SL, Ahmed Z, Khan Z, Griffith AJ, Morell RJ, Friedman TB, Riazuddin S, Wilcox ER. 2001. Mutations of the protocadherin gene *PCDH15* cause Usher Syndrome Type 1F. *Am J Hum Genet* 69(1): 25-34.

Arno G, Hull S, Robson AG, Holder GE, Cheetham ME, Webster AR, Plagnol V, Moore AT. 2015. Lack of interphotoreceptor retinoid binding protein caused by homozygous mutation of RBP3 is sssociated with high myopia and retinal dystrophy. *Invest Ophthalmol Vis Sci.* 56(4):2358-2365.

Au WWL. 1993. The sonar of dolphins. New York: Springer-Verlag.

Baehr W, Karan S, Maeda T, Luo DG, Li S, Bronson JD, Watt CB, Yau KW, Frederick JM, Palczewski K. 2007. The function of guanylate cyclase 1 and guanylate cyclase 2 in rod and cone photoreceptors. *J Biol Chem.* 282(12): 8837-8847.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Statist Soc. B* 57(1): 289-300.

Bolz H, von Brederlow B, Ramirez A, Bryda EC, Kutsche K, Gerd Nothwang H, Seeliger M, del C. Salcedó Cabrera M, Caballeró Vila, Pelaez Molina O, Gal A, Kubisch C. 2001. Mutation of *CDH23*, encoding a new member of the cadherin gene family, causes Usher syndrome type 1D. *Nature Genetics* 27: 108-112.

Carbon S, Ireland A, Mungall CJ, Shu SQ, Marshall B, Lewis S, the AmiGO Hub, the Web Presence Working Group. 2009. AmiGO: online access to ontology and annotation data. *Bioinformatics* 25(2): 288-289.

Churchill M, Martinez-Caceres M, de Muizon C, Mnieckowski J, Geisler JH. 2016. The origin of high-frequency hearing in whales. *Curr Biol* 26(16): P2144-P2149.

Dallos P. 2008. Cochlear amplification, outer hair cells and prestin. *Curr Opin Neurobiol* 18(4): 370-376.

Davies KTJ, Cotton JA, Kirwan JD, Teeling EC, Rossiter SJ. 2012. Parallel signatures of sequence evolution among hearing genes in echolocating mammals: an emerging model of genetic convergence. *Heredity* 108(5): 480-489.

Davies KTJ, Bennett NC, Tsagkogeorga G, Rossiter SJ, Faulkes CG. 2015. Family wide molecular adaptations to underground life in African mole-rats revealed by phylogenomic analysis. *Mol Biol Evol.* 32 (12): 3089-3107.

den Hollander AI, McGee TL, Zivello C, Banfi S, Dryja TP, Gonzalez-Fernandez F, Ghosh D, Berson EL. 2009. A homozygous missense mutation in the *IRBP* gene (*RBP3*) associated with autosomal recessive retinitis pigmentosa. *Invest Ophthalmol Vis Sci.* 50(4): 1864-1872.

Dewey JB, Xia A, Müller U, Belyantseva IA, Applegate BE, Oghalai JS. 2018. Mammalian auditory hair cell bundle stiffness affects frequency tuning by increasing coupling along the length of the cochlea. *Cell Rep* 23, 2915-2927.

Dionne G, Qiu X, Rapp M, Liang X, Zhao B, Peng G, Katsamba PS, Ahlsen G, Rubinstein R, Potter CS, et al. 2018. Mechanotransduction by Pcdh15 relies on a novel *cis*-dimeric architecture. *Neuron* 99(3): 480-492.

Donaudy F, Zheng L, Ficarella R, Ballana E, Carella M, Melchionda S, Estivill X, Bartles JR, Gasparini P. 2006. Espin gene (*ESPN*) mutations associated with autosomal dominant hearing loss cause defects in microvillar elongation or organisation. *J Med Genet* 43(2): 157-161.

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

Dong D, Lei M, Liu Y, Zhang S. 2013. Comparative inner ear transcriptome analysis between the Rickett's big-footed bats (*Myotis ricketti*) and the greater short-nosed fruit bats (*Cynopterus sphinx*). *BMC Genomics* 14:916.
Dungan SZ, Kosyakov A, Chang BSW. 2016. Spectral tuning of killer whale (*Orcinus orca*) rhodopsin: evidence for positive selection and functional adaptation in a cetacean visual pigment. *Mol. Biol. Evol.* 33:323-336.
Dungan SZ, Chang BSW. 2017. Epistatic interactions influence terrestrial–marine functional shifts in cetacean rhodopsin. *Proc R Soc B* 284:20162743.
Emerling CA. 2018. Regressed but not gone: patterns of vision gene loss and retention in subterranean mammals. *Integr Comp Biol*. 58(3):441-451.
Emerling CA, Springer MS. 2014. Eyes underground: regression of visual protein networks in subterranean mammals. *Mol Phylogenet Evol.* 78:260-270.
Emerling CA, Springer MS. 2015. Genomic evidence for rod monochromacy in sloths and armadillos suggests early subterranean history of Xenarthra. *Proc R Soc B* 282:20142192.
Emerling CA, Widjaja AD, Nguyen NN, Springer MS. 2017. Their loss is our gain: regressive evolution in vertebrates provides genomic models for uncovering human disease loci. *J. Med. Genet*. 54:787-794.
Fasick JI, Robinson PR. 2000. Spectral-tuning mechanisms of marine mammal rhodopsins and correlations with foraging depth. *Vis Neurosci.* 17:781-788.
Fang X, Seim I, Huang Z, Gerashchenko MV, Xiong Z, Turanov AA, Zhu Y, Lobanov AV, Fan D, Yim SH, Yao X, Ma S, Yang L, Lee SG, Kim EB, Bronson RT, Sumbera R, Buffenstein R, Zhou X, Krogh A, Park TJ, Zhang H, Wang J, Gladyshev VN. 2014. Adaptations to a subterranean environment and longevity revealed by the analysis of mole rat genomes. *Cell Rep* 8(5) :1354-1364.
Feng P, Zheng J, Rossiter SJ, Wang D, Zhao H. 2014. Massive losses of taste receptor genes in toothed and baleen whales. *Genome Biol Evol.* 6:1254-1265.
Foote AD, Liu Y, Thomas GWC, Vinar T, Alfoldi J, Deng J, Dugan S, van Elk CE, Hunter ME, Joshi V, et al. 2015. Convergent evolution of the genomes of marine mammals. *Nature Genet* 47(3):272-275.
Fraser FC, Purves PE. 1960. Hearing in cetaceans: evolution of the accessory air sacs and the structure and function of the outer and middle ear in recent cetaceans. *Bull Br Mus Nat Hist* 7(1): 1-140.
Gagnon LH, Longo-Guess CM, Berryman M, Shin J-B, Saylor KW, Yu H, Gillespie PG, Johnson KR. 2006. The chloride intracellular channel protein CLIC5 is expressed at high levels in hair cell stereocilia and is essential for normal inner ear function. *J Neurosci* 26(40):10188-10198.
Galatius A, Olsen MT, Steeman ME, Racicot RA, Bradshaw CD, Kyhn LA, Miller LA. 2019. Raising your voice: evolution of narrow-band high-frequency signals in toothed whales (Odontoceti). *Biol J Linn Soc* 126(2):213-224.
Gall TMH, Frampton AE. 2013. Gene of the month: E-cadherin (*CDH1*). *J Clin Pathol* 66:928-932.
Gallego-Martinez A, Requena T, Roman-Naranjo P, Lopez-Escamez JA on behalf of the Meniere Disease Consortium (MeDiC). 2019. Excess of rare missense variants in hearing loss genes in sporadic Meniere Disease. *Front Genet* 10:76.
Gao J, Cheon K, Nusinowitz S, Liu Q, Bei D, Atkins K, Azimi A, Daiger SP, farber DB, Heckenlively JR, Pierce EA, Sullivan LS, Zuo J. 2002. Progressive photoreceptor

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

degeneration, outer segment dysplasia, and rhodopsin mislocalization in mice with targeted disruption of the retinitis pigmentosa-1 (Rp1) gene. *Proc Nat Acad Sci USA* 99(8): 5698-5703.

Gatesy J, O'Leary MA. 2001. Deciphering whale origins with molecules and fossils. *Trends Ecol Evol.* 16(10): 562-570.

Gatesy J, Geisler JH, Chang J, Buell C, Berta A, Meredith RW, Springer MS, McGowen MR. 2013. A phylogenetic blueprint for a modern whale. *Mol Phylogenet Evol.* 66(2): 479-506.

Grillet N, Schwander M, Hildebrand MS, Sczaniecka A, Kolatkar A, Velasco J, Webster JA, Kahrizi K, Najmabadi H, Kimberling WJ, et al. 2009. Mutations in LOXHD1, an evolutionarily conserved stereociliary protein, disrupt hair cell function in mice and cause progressive hearing loss in humans. *Am J Hum Genet.* 85(3): 328-337.

Hayden S, Bekaert M, Crider TA, Mariani S, Murphy WJ, Teeling EC. 2010. Ecological adaptation determines functional mammalian olfactory subgenomes. *Genome Res.* 20(1):1-9.

Hudson NJ, Baker ML, Hart NS, Wynne JW, Gu Q, Huang Z, Zhang G, Ingham AB, Wang L, Reverter A. 2014. Sensory rewiring in an echolocator: genome-wide modification of retinogenic and auditory genes in the bat *Myotis davidii*. *G3(Bethesda)* 4(10):1825-1835.

Huelsmann M, Hecker N, Springer MS, Gatesy J, Sharma V, Hiller M. 2019. Genes lost during the transition from land to water in cetaceans highlight genomic changes involved in aquatic adaptations. bioRxiv, https://doi.org/10.1101/521617.

Iglesias AI, Mishra A, Vitart V, Bykhovskaya Y, Höhn R, Springelkamp H, Cuellar-Partida G, Gharahkhani P, Cooke Bailey JN, Willoughby CE, et al. 2018. Cross-ancestry genome-wide association analysis of corneal thickness strengthens link between complex and Mendelian eye diseases. *Nat Commun* 9:1864.

Jensen FH, Johnson M, Ladegaard M, Wisniewska DM, Madsen PT. 2018. Narrow acoustic field of view drives frequency scaling in toothed whale biosonar. *Curr Biol* 28, 3878-3885.

Jin M, Li S, Nusinowitz S, Lloyd M, Hu J, Radu RA, Bok D, Travis GH. The role of interphotoreceptor retinoid-binding protein on the translocation of visual retinoids and function of cone photoreceptors. *J Neurosci* 29(5): 1486-1495.

Johnson CP, Chapman ER. 2010. Otoferlin is a calcium sensor that directly regulates SNARE-mediated membrane fusion. *J Cell Biol* 191(1): 187.

Johnson M, Madsen PT, Zimmer WMX, Aguilar de Soto N, Tyack PL. 2006. Foraging Blainville's beaked whales (*Mesoplodon densirostris*) produce distinct click types matched to different phases of echolocation. *J Exp Biol.* 209:5038-5050.

Karan S, Frederick JM, Baehr W. 2010. Novel functions of photoreceptor guanylate cyclases revealed by targeted deletion. *Mol Cell Biochem.* 334(1-2): 141-155.

Kazmierczak P, Sakaguchi H, Tokita J, Wilson-Kubalek EM, Milligan RA, Müller U, Kachar B. 2007. Cadherin 23 and protocadherin 15 interact to form tip-link filaments in sensory hair cells. *Nature* 449: 87–91.

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

Kosakovsky Pond SL, Frost SDW, Muse SV. 2005. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21(5): 676-679.

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

Kurima K, Ebrahim S, Pan B, Sedlacek M, Sengupta P, Millis BA, Cui R, Nakanishi H, Fujikawa T, Kawashima Y, Choi BY, Monohan K, Holt JR, Griffith AJ, Kachar B. TMC1 and TMC2 localize at the site of mechanotransduction in mammalian inner ear hair cell stereocilia. *Cell Rep* 12(10): 1606-1617.

LeDuc R. 2002. Delphinids, overview. In: Perrin WF, Würsig B, Thewissen JGM, editors. Encyclopedia of marine mammals. San Diego: Academic Press. p. 310-314.

Levenson DH, Dizon A. 2003. Genetic evidence for the ancestral loss of short-wavelength-sensitive cone pigments in mysticete and odontocete cetaceans. *Proc R Soc B* 270:673-679.

Li Y, Liu Z, Shi P, Zhang J. 2010. The hearing gene *Prestin* unites echolocating bats and whales. *Curr Biol* 20:R55-R56.

Liu P, Osawa S, Weiss ER. 2005. M opsin phosphorylation in intact mammalian retinas. *J Neurochem* 93(1): 135-144.

Liu Y, Cotton JA, Shen B, Han X, Rossiter SJ, Zhang S. 2010a. Convergent sequence evolution between echolocating bats and dolphins. *Curr Biol* 20:R53-R54.

Liu Y, Rossiter SJ, Han X, Cotton JA, Zhang S. 2010b. Cetaceans on a molecular fast track to ultrasonic hearing. *Curr Biol* 20:1834-1839.

Liu Z, Qi FY, Zhou X, Ren HQ, Shi P. 2014. Parallel sites implicate functional convergence of the hearing gene prestin among echolocating mammals. *Mol Biol Evol.* 31:2415-2424.

Liu Z, Qi FY, Xu DM, Zhou X, Shi P. 2018. Genomic and functional evidence reveals molecular insights into the origin of echolocation in whales. *Sci Adv.* 4, eaat8821.

Marcovitz A, Turakhia Y, Gloudemans M, Braun BA, Chen HI, Bejerano G. 2017. A novel unbiased test for molecular convergent evolution and discoveries in echolocating, aquatic, and high-altitude mammals. bioRxiv, doi.org/10.1101/170985.

Mass AM, Supin AY. 2007. Adaptive features of aquatic mammals' eye. *Anat Rec* 290:701-715.

Mburu P, Mustapha M, Varela A, Weil D, El-Amraoui A, Holme RH, Rump A, Hardisty RE, Blanchard S, Coimbra RS, Perfettini I, Parkinson N, et al. 2003. Defects in whirlin, a PDZ domain molecule involved in stereocilia elongation, cause deafness in the whirler mouse and families with DFNB31. *Nat Genet* 34(4): 421-428.

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: 574-590.

McGowen MR, Grossman LI, Wildman DE. 2012. Dolphin genome provides evidence for adaptive evolution of nervous system genes and a molecular rate slowdown. *Proc R Soc B*. 279:3643-3651.

McGowen, MR, Tsagkogeorga G, Álvarez-Carretero S, dos Reis M, Struebig M, Deaville R, Jepson PD, Jarman S, Polanowski A, Morin PA, Rossiter SJ. 2019. Phylogenomic resolution of the cetacean tree of life using target sequence capture. *Syst. Biol*. Advance Access, https://doi.org/10.1093/sysbio/syz068.

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

Mourlam MJ, Orliac MJ. 2017. Infrasonic and ultrasonic hearing evolved after the emergence of modern whales. *Curr Biol*. 27: 1776-1781.

Nery MF, Gonzalez DJ, Opazo JC. 2012. How to make a dolphin: molecular signature of positive selection in cetacean genome. *PLoS ONE* 8:e65491.

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

Nummela S, Thewissen JG, Bajpai S, Hussain ST, Kumar K. 2004. Eocene evolution of whale hearing. *Nature* 430: 776-778.
Odeh H, Hunker K, Belyantseva IA, Azaiez H, Avenarius MR, Zheng L, Peters LM, Gagnon LH, Hagiwara N, Skynner MJ, Brilliant MH, et al. 2010. Mutations in *Grxcr1* are the basis for inner ear dysfunction in the pirouette mouse. *Am J Hum Genet* 86(2):148-160.
Oghalai JS. 2004. The cochlear amplifier: augmentation of the traveling wave within the inner ear. *Curr Opin Otolaryngol Head Neck Surg* 12:431-438.
Osawa S, Weiss ER. 2012. A tale of two kinases in rods and cods. In: LaVail M, Ash J, Anderson R, Hollyfield J, Grimm C, editors. Retinal degenerative diseases: Advances in experimental medicine and biology, vol. 723. Boston: Springer. p. 821-827.
Park JY, Kim K, Sohn H, Kim HW, An Y-R, Kang J-H, Kim E-M, Kwak W, Lee C, Yoo D, Jung J, Sung S, Yoon J, Kim H. 2018. Deciphering the evolutionary signatures of pinnipeds using novel genome sequences: the first genomes of *Phoca largha*, *Callorhinus ursinus*, and *Eumetopias jubatus*. *Scientific Reports* 8:16877.
Parker J, Tsagkogeorga G, Cotton JA, Liu Y, Provero P, Stupka E, Rossiter SJ. 2013. Genome-wide signatures of convergent evolution in echolocating mammals. *Nature* 502:228-231.
Peichl L, Behrmann G, Kröger RH. 2001. For whales and seals the ocean is not blue: a visual pigment loss in marine mammals. *Eur J Neurosci.* 13:1520-1528.
Richardson GP, Lukashkin AN, Russell IJ. 2008. The tectorial membrane: one slice of a complex cochlear sandwich. *Curr Opin Otolaryngol Head Neck Surg* 16:458-464.
Sango K, Yamanaka S, Hoffmann A, Okuda Y, Grinberg A, Westphal H, McDonald MP, Crawley JN, Sandhoff K, Suzuki K, Proia R. 1995. Mouse models of Tay-Sachs and Sandhoff diseases differ in neurologic phenotype and ganglioside metabolism. *Nature Genet* 11:170-176.
Shearer AE, Hildebrand MS, Smith RJH. 2017. Hereditary Hearing Loss and Deafness Overview, Updated 2017 Jul 27. In: Adam MP, Ardinger HH, Pagon RA, et al., editors. GeneReviews, Seattle.
Shen YY, Liang L, Li GS, Murphy RW, Zhang YP. 2012. Parallel evolution of auditory genes for echolocation in bats and toothed whales. *PLoS Genet* 8:e1002788.
Shen B, Fang T, Dai M, Jones G, Zhang S. 2013. Independent losses of visual perception genes *Gja10* and *Rbp3* in echolocating bats (Order: Chiroptera). *PLoS ONE* 8: e68867
Smith MD, Wertheim JO, Weaver S, Murrell B, Scheffler K, Kosakovsky Pond SL. 2015. Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol Biol Evol* 32(5): 1342-1353.
Springer MS, Emerling CA, Fugate N, Patel R, Starrett J, Morin PA, Hayashi C, Gatesy J. 2016. Inactivation of cone-specific phototransduction genes in rod monochromatic cetaceans. *Front Ecol Evol* 4:61.
Strenzke N, Chakrabarti R, Al-Moyed H, Müller A, Hoch G, Pangrsic T, Yamanbaeva G, Lenz C, Pan K-T, Auge E, Geiss-Friedlander R, Urlaub H, Brose N, Wichmann C, Reisinger E. 2016. Hair cell synaptic dysfunction, auditory fatigue and thermal sensitivity in otoferlin Ile515Thr mutants. *EMBO Journal* 35:2519-2535
Sun YB, Zhou WP, Liu HQ, Irwin DM, Shen YY, Zhang YP. 2012. Genome-wide scans for candidate genes involved in the aquatic adaptation of dolphins. *Genome Biol Evol.* 5:130-139.
Tekin M, Akcayoz D, Incesulu A. 2005. A novel missense mutation in a C2 domain of OTOF results in autosomal recessive auditory neuropathy. *Am J Med Genet A* 138:6–10.

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

Tsagkogeorga G, McGowen MR, Davies KTJ, Jarman S, Polanowski A, Bertelsen MF, Rossiter SJ. 2015. A phylogenomic analysis of the role and timing of molecular adaptation in the aquatic transition of cetartiodactyl mammals. *R Soc Open Sci*. 2:150156.

Varga R, Avenarius MR, Kelley PM, Keats BJ, Berlin CI, Hood LJ, Morlet TG, Brashears SM, Starr A, Cohn ES, Smith RJH, Kimberling WJ. 2006. OTOF mutations revealed by genetic analysis of hearing loss families including a potential temperature sensitive auditory neuropathy allele. *J Med Genet.* 43(7):576-581.

Verpy E, Weil D, Leibovici M, Goodyear RJ, Hamard G, Houdon C, Lefevre GM, Haredlin JP, Richardson GP, Avan P, Petit C. 2008. Stereocilin-deficient mice reveal the origin of cochlear waveform distortions. *Nature* 456: 255-258.

Weaver S, Shank SD, Spielman SJ, Li M, Muse SV, Kosakovsky Pond SL. 2018. Datamonkey 2.0: a modern web application for characterizing selective and other evolutionary processes. *Mol Biol Evol* 35(3): 773-777.

Weiss ER, Ducceschi MH, Horner TJ, Li A, Craft CM, Osawa S. 2001. Species-specific differences in expression of G-protein-coupled receptor kinase (GRK) 7 and GRK1 in mammalian cone photoreceptor cells: implications for cone cell phototransduction. *J Neurosci*. 21:9175-9184.

Yamashita T, Liu J, LeNoue S, Wang C, Kaminoh J, Bowne SJ, Sullivan LS, Daiger SP, Zhang K, Fitzgerald ME, Kefalov VJ, Zuo J. 2009. Essential and synergistic roles of RP1 and RP1L1 in rod photoreceptor axoneme and retinitis pigmentosa. *J Neurosci.* 29:9748-9760.

Yan J, Pan L, Chen X, Wu L, Zhang M. 2010. The structure of the harmonin/sans complex reveals an unexpected interaction mode of the two Usher syndrome proteins. *Proc Nat Acad Sci USA* 107(9):4040-4045.

Yan W, Long P, Chen T, Liu W, Yao L, Ren Z, Li X, Wang J, Xue J, Tao Y, Zhang L, Zhang Z. 2018. A natural occuring mouse model with *Adgrv1* mutation of Usher Syndrome 2C and characterization of its recombinant inbred strains. *Cell Physiol Biochem* 47:1883-1897.

Yang Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Mol Biol Evol.* 24:1586-1591.

Yim HS, Cho YS, Guang X, Kang SG, Jeong JY, Cha SS, Oh HM, Lee JH, Yang EC, Kwon KK, et al. 2014. Minke whale genome and aquatic adaptation in cetaceans. *Nat Genet.* 46:88-92.

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

Figure Legends.

Figure 1. Phylogenetic tree of the major lineages of cetaceans based on McGowen et al. (2019) with putative pseudogenization events mapped using parsimony. Lines over each node symbolize a pseudogenization event on that branch for *OPN1SW* (red), *OPN1LW* (blue), or other cone-specific genes discussed here (*CNGA3, CNGB3, GNAT2, PDE6C, GRK7*; brown) and elsewhere (*ARR3, GNGT2, PDE6H;* brown) (Meredith et al., 2013; Springer et al., 2016; Emerling, 2018). Solid lines represent a definitive silencing event on a particular branch; dashed lines symbolize that at least one species (but not all) in a particular clade is predicted to contain a pseudogene.

Figure 2. Phylogenetic trees of the major lineages of cetaceans based on McGowen et al. (2019) with genes showing evidence of positive selection using branch-site models shown above each node. All genes were identified as being under positive selection using the aBSREL model with asterisks noting genes that also passed an FDR correction. Bolded genes also show evidence of positive selection on a particular branch using branch-sites Model A conducted in PAML (Yang, 2007). Boxes to the right of clade names indicate genes with evidence of positive selection on particular branches within each clade. Each tree shows A) all genes in our analysis included within the GO category GO:0007605 "sensory perception of sound" and B) all genes in our analysis included within the GO category GO:0007601 "visual perception". The cumulative number of loci with evidence of positive selection is mapped on each tree with colors coding between under three and over 17 genes. Genes highlighted in red represent those that are also classified as nonsyndromic hearing loss (NHSL) genes.

Figure 3. Phylogenetic trees of the major lineages of cetaceans based on McGowen et al. (2019) with genes showing evidence of positive selection using branch-site models shown above each node. All genes were identified as being under positive selection using the aBSREL model with asterisks noting genes that also passed an FDR correction. Bolded genes also show evidence of positive selection on a particular branch using branch-sites Model A conducted in PAML (Yang, 2007). Boxes to the right of clade names indicate genes with evidence of positive selection on particular branches within each clade. All genes listed are non-sensory related genes. The cumulative number of loci with evidence of positive selection is mapped on each tree with colors coding between under three and over 17 genes.

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

Table 1. Inactivating mutations in cetaceans identified in this study. We have included additional species for one mutation in *OPN1LW* previously identified in Meredith et al. (2013).

Gene	Species	Length	Mutation	Region	Coverage
GRK7	*Caperea marginata*	2 bp	Deletion	Exon 1	100x
	Beradius bairdii	8 bp	Deletion	Exon 1	17x
	Balaenoptera borealis + B. edeni	13 bp	Deletion	Exon 1	17-29x
	Kogia sima	5 bp	Deletion	Exon 5	29x
	Mesoplodon europaeus + M. densirostris + M. gingkodens	13 bp	Deletion	Exon 5	36-51x
	Mesoplodon stejnegeri	7 bp	Deletion	Exon 5	129x
	Ziphius cavirostris	5 bp	Insertion	Exon 7	11x
OPN1LW	*Mesoplodon carlhubbsi*	4 bp	Insertion	Exon 2	3x
	Balenoptera edeni, B. borealis	1 bp	Deletion	Exon 2; Meredith et al., 2013	3x
	Mesoplodon europaeus	1 bp	Insertion	Exon 6	41x
CNGB3	*Mesoplodon mirus*	1 bp	Deletion	Exon 6	39x
	Balaenoptera bonaerensis	1 bp	Deletion	Exon 11	3x
	Kogia sima	4 bp	Deletion	Exon 14	16x
GNAT2	*Caperea marginata*	2 bp	Deletion	Exon 5	8x
	Eschrichtius robustus	2 bp	Deletion	Exon 5	16x
	Balaenoptera bonaerensis	1 bp	Deletion	Exon 6	3x
	Mesoplodon stejnegeri	2 bp	Deletion	Exon 6	30x
PDE6C	*Caperea marginata*	2 bp	Deletion	Exon 1	31x
	Balaenoptera edeni	10 bp	Deletion	Exon 2	11x
	Mesoplodon ginkgodens	1 bp	Deletion	Exon 2	41x
	Balaenoptera musculus, B. borealis,, Eubalaena australis	4 bp	Deletion	Exon 16	24-76x

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

	Mesoplodon ginkgodens	4 bp	Deletion	Exon 19	41x
CNGA3	*Mesoplodon europaeus*	8 bp	Deletion	Exon 5	3x
	Hyperoodon ampullatus	2 bp	Deletion	Exon 7	13x
	Mesoplodon carlhubbsi		Stop codon	Exon 7	3x
	Kogia breviceps	1 bp	Deletion	Exon 8	5x
	Kogia sima	1 bp	Deletion	Exon 8	8x
	Balaenoptera musculus	1 bp	Deletion	Exon 8	21x
	Eubalaena glacialis	1 bp	Deletion (Polymorphic)	Exon 8	80x
	Mesoplodon mirus	1 bp	Deletion	Exon 8	250x
	Balaenoptera borealis, B. edeni, Eschirchtius robustus	1 bp	Deletion	Exon 8	5-16x
	Feresa attenuata,	1 bp	Insertion	End of Exon 8 (elongation?)	11x
	Phocoena spp.	2 bp	Deletion	End of Exon 8 (elongation?)	3-16x
GUCY2F	*Mesoplodon perrini*	1 bp	Insertion	Exon 1	17x
	Caperea marginata	2 bp	Deletion	Exon 2	27x
	Mesoplodon bowdoini	1 bp	Deletion	Exon 4	20x
	Pontoporia blainvillei	4 bp	Insertion	Exon 4	19x

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

Table 2. Hearing and vision genes under positive selection using the branch-sites test A for eight branches (Cetruminantia, Ruminantia, Whippomorpha, Cetacea, Mysticeti, Odontoceti, Ziphiidae, Delphinidae) instituted in PAML 4 (Yang, 2007). Bold indicates significance after correcting for FDR. (Abbreviatons: -ln*L* = log likelihood; LRT = likelihood ratio test statistic; FDR = false discovery rate; BEB = Bayes Empirical Bayes)

Clade	Gene	-ln*L* (max)	LRT	*p* value	FDR	Positively selected sites (BEB >0.5)
Ruminantia	*TULP1*	Null: -5563.0624	10.5786	0.001	0.305	42 T 0.505; 124 E 0.713; 212 T 0.696; 226 G 0.987
		Alt: -5557.7731				
	ATP6V0A4	Null: -9757.1972	9.2362	0.002	0.413	274 I 0.516; 354 M 0.502; 434 S 0.506; 437 T 0.508
		Alt: -9752.5791				
	WFS1	Null: -10582.507	5.9014	0.015	1	43 R 0.550; 576 A 0.966; 668 Q 0.533; 728 I 0.690; 855 S 0.523
		Alt: -10579.556				
	RHO	Null: -3643.0529	5.0696	0.0243	1	26 Y 0.975
		Alt: -3640.5181				
	BBS2	Null: -7284.3812	4.7584	0.0292	1	302 H 0.747; 424 L 0.643
		Alt: - 7282.002				
	FGFR1	Null: -7275.9517	4.4384	0.0351	1	38 - 0.712
		Alt: -7273.7325				
	ATP6V1B1	Null: -4621.9055	4.0688	0.044	1	15 S 0.612
		Alt: -4619.8711				
Cetruminantia	*OTOG*	Null: -37315.488	10.0552	0.002	0.338	253 T 0.637; 1512 A 0.559; 1522 Q 0.589; 1705 A 0.635; 1949 A 0.672
		Alt: -37310.46				
	LUM	Null: -3820.653	5.366	0.021	1	7 P 0.714; 19 S 0.924; 21 - 0.796; 189 S 0.716
		Alt: -3817.97				
	MYO1A	Null: -11371.957	4.3256	0.038	1	25 I 0.634; 242 A 0.557; 357 L 0.548; 428 V 0.918; 1012 L 0.626
		Alt: -11369.794				
	GUCY2D	Null: -11370.261	4.0442	0.044	1	315 K 0.777; 477 V 0.773; 561 Y 0.981; 974 T 0.794; 1064 E 0.759
		Alt: -11368.238				
Whippomorpha	***RBP3***	**Null: -15978.1607**	**12.575**	**3.09E-04**	**0.098**	220 N 0.615; 444 R 0.653; 455 A 0.509; 1093 I 0.895; 1254 K 0.851
		Alt: -15971.8732				
	CDH1	Null: -11325.0999	5.1742	0.023	1	188 L 0.532; 189 H 0.689; 638 F 0.768
		Alt: -11322.5128				
	CDH3	Null: -10172.1386	4.737	0.03	1	112 D 0.677; 133 H 0.697; 184 E 0.712; 367 Q 0.902; 732 L 0.723
		Alt: -10169.7701				
Cetacea	*TECTA*	Null: -22571.4104	8.72	0.003	0.4505	972 Q 0.985
		Alt: -22567.0503				
	POU6F2	Null: -10172.1386	5.444	0.02	1	113 V 0.637; 160 T 0.614; 206 Q 0.735; 248 S 0.680; 250 S 0.634; 427 P 0.635
		Alt: -6315.2983				
	CACNB2	Null: -7520.7992	4.613	0.032	1	54 T 0.521; 60 R 0.917; 453 P 0.517; 453 P 0.517; 558 R 0.517
		Alt: -7518.4925				

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

	CDH1	Null: -11323.9158 Alt: -11321.8604	4.111	0.043	1	30 L 0.525; 188 L 0.585; 189 H 0.856; 269 E 0.688; 271 S 0.674; 487 L 0.674
	NIPBL	Null: -18111.7866 Alt: -18109.7941	3.985	0.046	1	99 I 0.753; 678 V 0.750; 683 L 0.925; 722 H 0.767; 890 A 0.768
Mysticeti	***RP1***	**Null: -29996.4659** **Alt: -29984.7055**	**23.521**	**1.24E-06**	**0.001**	145 R 0.595; 1472 T 0.509
	USH1G	Null: -3737.8646 Alt: -3734.926	5.877	0.015	1	98 L 0.994; 111 L 0.751; 175 A 0.618
Odontoceti	None					
Ziphiidae	***LOXHD1***	**Null: -24363.3251** **Alt: -24354.7238**	**17.203**	**3.36E-05**	**0.013**	785 A 0.723; 979 E 0.669; 1021 V 0.718; 1216 D 0.718; 1438 I 0.704; 1618 D 0.545; 1703 K 0.732; 2028 S 0.712; 2070 C 0.536
	KCNE1	Null: -1875.4199 Alt: -1870.8423	9.155	0.002	0.45	124 E 0.995
	RCVRN	Null: -2359.9133 Alt: -2359.8678	7.657	0.006	0.708	33 Q 0.730; 162 K 0.807
	SAG	Null: -4754.5279 Alt: -4753.0575	4.341	0.037	1	4 N 0.688; 12 P 0.842; 41 Q 0.555; 99 E 0.885; 100 T 0.941; 112 M 0.939; 186 R 0.836; 325 L 0.913; 344 T 0.810
	SPATA7	Null: -7544.1408 Alt: -7541.9712	4.339	0.037	1	101 Q 0.629; 489 F 0.644; 536 V 0.596
	RORB	Null: -3852.3451 Alt: -3852.3451	4.111	0.043	1	130 L 0.745; 193 S 0.769
	KRT12	Null: -6534.071 Alt: -6532.026	4.09	0.043	1	25 R 0.977; 70 S 0.988; 167 R 0.962; 305 M 0.974; 320 A 0.954
Delphinidae	***PCDH15***	**Null: -13248.7969** **Alt: -13240.3497**	**16.894**	**3.95E-05**	**0.013**	391 T 0.996; 423 V 0.993; 486 Y 0.995; 503 V 0.953; 505 A 0.996; 733 R 0.993; 1127 R 0.992; 1251 A 0.995; 1293 N 0.996; 1325 T 0.982; 1381 I 0.967
	CDH23	Null: -36231.938 Alt: -36227.5722	8.732	0.003	0.45	778 H 0.668; 977 T 0.671; 979 S 0.671; 1317 S 0.542; 1346 L 0.588; 1961 L 0.522; 2418 V 0.606; 2712 Q 0.624; 3126 Y 0.567; 3340 L 0.629
	EML2	Null: -7912.0851 Alt: -7908.4782	7.214	0.007	0.805	543 V 0.959; 631 S 0.672

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

Table 3. Non-sensory genes under positive selection using the branch-sites test A for eight branches (Cetruminantia, Ruminantia, Whippomorpha, Cetacea, Mysticeti, Odontoceti, Ziphiidae, Delphinidae) instituted in PAML 4 (Yang, 2007). Bold indicates significance after correcting for FDR. (Abbreviatons: -ln*L* = log likelihood; LRT = likelihood ratio test statistic; FDR = false discovery rate; BEB = Bayes Empirical Bayes)

Clade	Gene	-ln*L* (max)	LRT	*p* value	FDR	Positively selected sites (BEB >0.5)
Ruminantia	***CLCN5***	**Null: -6091.3432 Alt: -6083.5829**	**15.5206**	**8.16E-05**	**0.055**	1 D 0.912; 127 D 0.530; 367 M 0.518
	TNFSF4	Null: -1803.335 Alt: -1798.7251	9.2198	0.002	0.800	65 K 0.770; 117 S 0.651; 127 A 0.946
	SLC44A1	Null: -4731.5700 Alt: -4728.7774	5.5852	0.018	1	33 K 0.529; 83 L 0.543; 109 - 0.518; 114 - 0.778; 235 - 0.506; 237 - 0.524; 242 - 0.572; 264 I 0.567; 320 T 0.516; 335 V 0.801
	MC4R	Null: -4062.0351 Alt: -4059.7023	4.6656	0.030	1	172 C 0.959
	IL18R1	Null: -7677.8329 Alt: -7675.7244	4.2170	0.040	1	94 R 0.609; 203 H 0.704; 231 W 0.746; 232 K 0.515; 244 E 0.743; 294 K 0.696
	NFXL1	Null: -7577.6503 Alt: -7575.5739	4.1528	0.042	1	627 D 0.752
Cetruminantia	*ACR*	Null: -3096.0943 Alt: -3092.1387	7.9110	0.005	1	46 V 0.758; 151 R 0.545; 191 S 0.552
Whippomorpha	*HCAR1*	Null: -4162.3949 Alt: -4158.6552	7.4794	0.006	1	25 A 0.631; 79 R 0.917; 86 I 0.982; 87 P 0.604; 303 G 0.538
	PAQR7	Null: -4028.0850 Alt: -4025.5515	5.0670	0.024	1	111 T 0.800
	CMKLR1	Null: -4857.9811 Alt: -4855.6233	4.7156	0.030	1	30 S 0.940
	ATP12A	Null: -8905.0931 Alt: -8902.8400	4.5062	0.033	1	387 V 0.866; 735 E 0.924
	PDE3B	Null: -10032.0280 Alt: -10029.9020	4.2534	0.039	1	449 L 0.960
Cetacea	***FSHR***	**Null: -7575.4051 Alt: -7566.7205**	**17.3692**	**3.08E-05**	**0.041**	10 T 0.557; 216 G 0.554; 237 A 0.988; 254 I 0.973
	IL18R1	Null: -7677.8229 Alt: -7672.3984	10.8490	9.88E-04	0.44	70 H 0.945; 96 P 0.662; 298 P 0.618; 479 Q 0.521
	SLC7A10	Null: -3220.3309 Alt: -3215.9665	8.7288	0.003	0.84	19 C 0.998
	ATP12A	Null: -8905.0931 Alt: -8901.7863	6.6136	0.010	1	81 T 0.522; 319 A 0.519; 336 G 0.636; 373 V 0.558; 438 S 0.742; 506 A 0.754; 683 S 0.716

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

	PAQR7	Null: -4026.7411 Alt: -4023.7289	6.0244	0.014	1	22 M 0.855; 104 L 0.727; 175 F 0.762; 176 Y 0.757
Mysticeti	*TFAP2B*	Null: -3391.2361 Alt: -3387.6984	7.0754	0.008	1	8 - 0.871
	TRAPPC9	Null: -11975.972 Alt: -11973.691	4.5614	0.033	1	124 I 0.59; 1092 Q 0.570
Odontoceti	None					
Ziphiidae	*NFXL1*	Null: -7577.5886 Alt: -7575.0978	4.9816	0.026	1	30 G 0.680; 127 I 0.861
	DEPDC1B	Null: -4178.2103 Alt: -4176.038	4.3446	0.037	1	None
Delphinidae	None					

Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

Table 4. Contingency tables summarizing numbers of genes found to be under positive selection versus not under selection for both the ingroup and outgroup, as identified in aBSREL tests. *P*-values are generated from tests conducted on (i) hearing genes, (ii) non-syndromic hearing loss genes, and (iii) vision genes, in each case combining the data with the non-sensory control genes to form a 2x2x2 table.

Hearing genes

	Ingroup	**Outgroup**	**Totals**
PSG	54	69	123
Non-PSG	25	10	35
Totals	79	79	158

$p = 0.0151$, Breslow-Day test; $p = 0.0144$, Log-linear model

NSHL genes

	Ingroup	**Outgroup**	**Totals**
PSG	16	30	46
Non-PSG	17	3	20
Totals	33	33	66

$p = 0.0006$, Breslow-Day test; $p = 0.0004$, Log-linear model

Vision genes

	Ingroup	**Outgroup**	**Totals**
PSG	66	77	143
Non-PSG	30	19	49
Totals	96	96	192

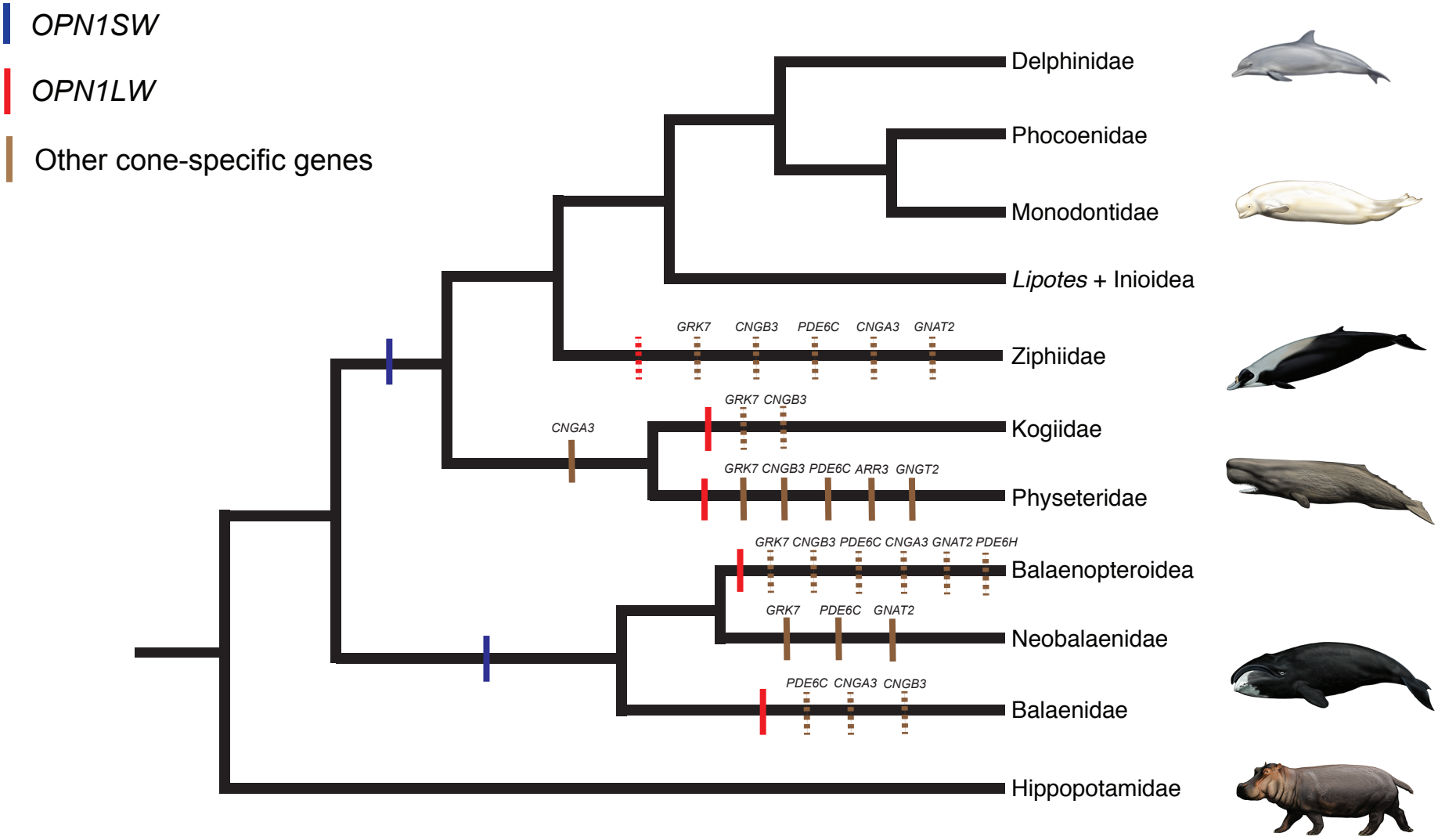
$p = 0.1251$, Breslow-Day test; $p = 0.1248$, Log-linear model

Non-sensory (control) genes

	Ingroup	**Outgroup**	**Totals**
PSG	146	144	290
Non-PSG	21	23	44
Totals	167	167	334

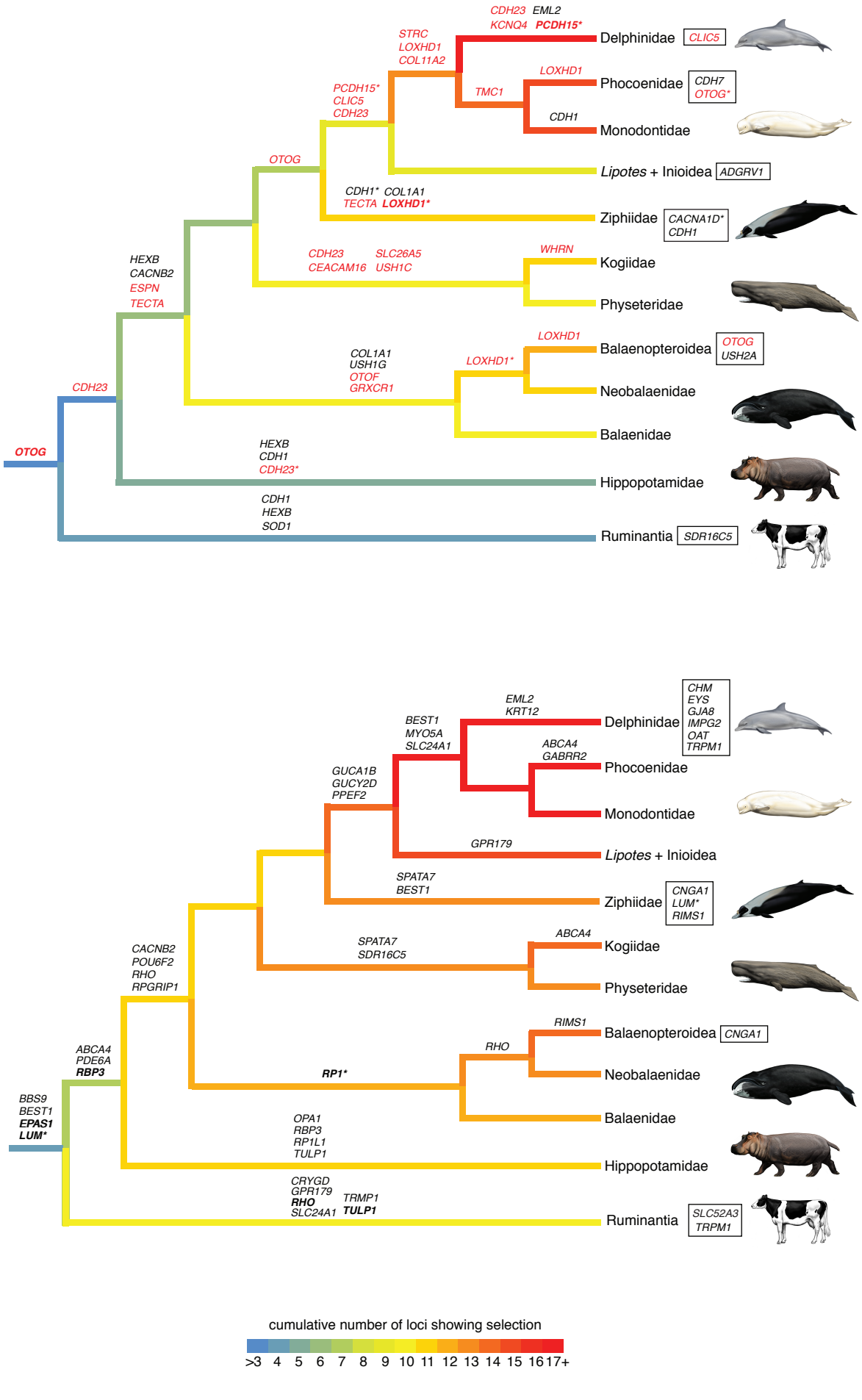
Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

Figure 1.



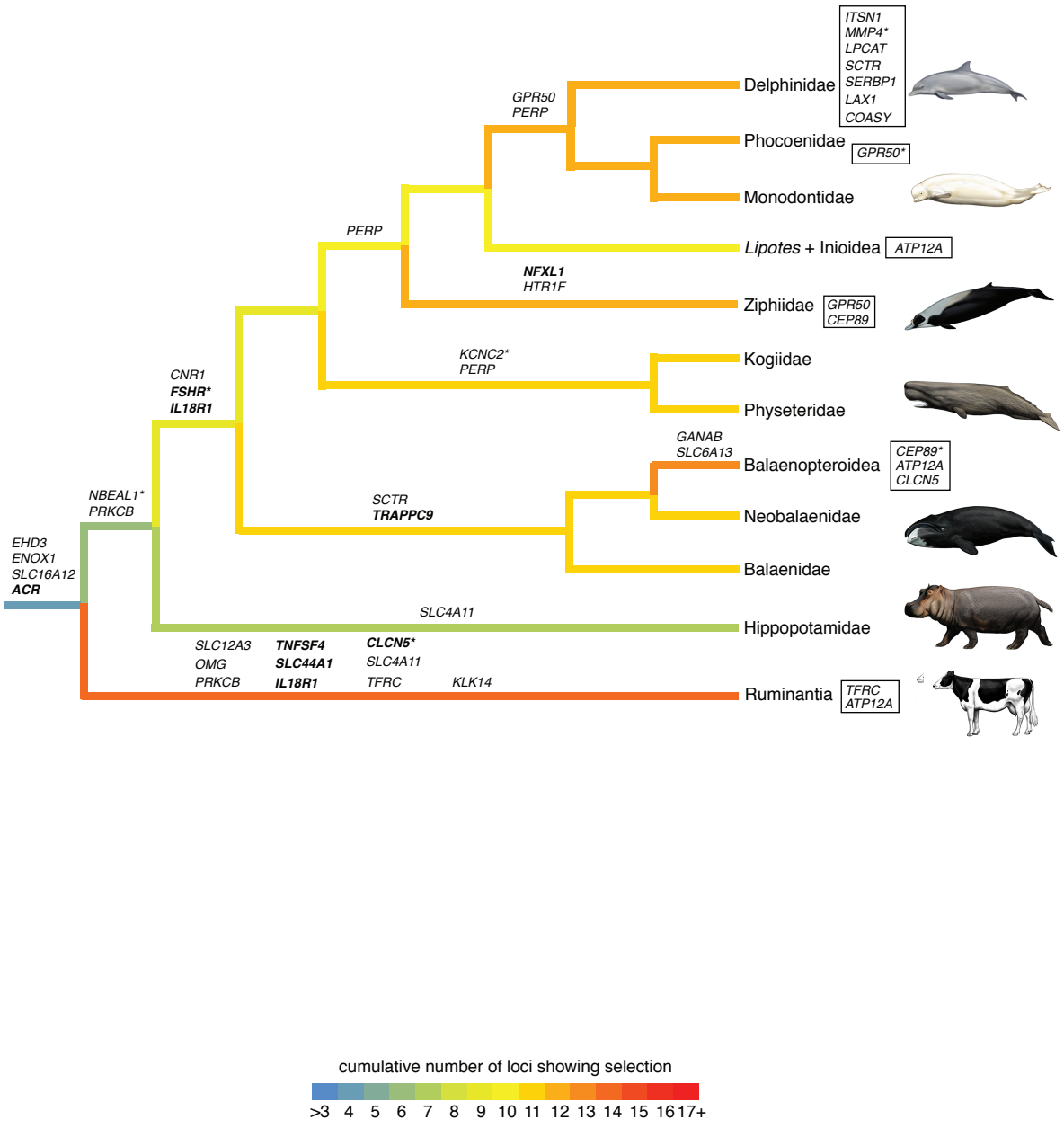
Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

Figure 2.



Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020

Figure 2.



Downloaded from https://academic.oup.com/mbe/article-abstract/doi/10.1093/molbev/msaa070/5805393 by Queen Mary University of London user on 21 April 2020