MBE Advance Access published November 24, 2015

# Spectral Tuning of Killer Whale (*Orcinus orca*) Rhodopsin: Evidence for Positive Selection and Functional Adaptation in a Cetacean Visual Pigment

Sarah Z. Dungan,[1] Alexander Kosyakov,[2] and Belinda S.W. Chang*[,1,2,3]

[1]Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON, Canada

[2]Department of Cell and Systems Biology, University of Toronto, Toronto, ON, Canada

[3]Centre for the Analysis of Genome Evolution and Function, Toronto, ON, Canada

***Corresponding author:** E-mail: belinda.chang@utoronto.ca.

**Associate editor:** James McInerney

## Abstract

**Cetaceans have undergone a remarkable evolutionary transition that was accompanied by many sensory adaptations, including modification of the visual system for underwater environments. Recent sequencing of cetacean genomes has made it possible to begin exploring the molecular basis of these adaptations. In this study we use in vitro expression methods to experimentally characterize the first step of the visual transduction cascade, the light activation of rhodopsin, for the killer whale. To investigate the spectral effects of amino acid substitutions thought to correspond with absorbance shifts relative to terrestrial mammals, we used the orca gene as a background for the first site-directed mutagenesis experiments in a cetacean rhodopsin. The S292A mutation had the largest effect, and was responsible for the majority of the spectral difference between killer whale and bovine (terrestrial) rhodopsin. Using codon-based likelihood models, we also found significant evidence for positive selection in cetacean rhodopsin sequences, including on spectral tuning sites we experimentally mutated. We then investigated patterns of ecological divergence that may be correlated with rhodopsin functional variation by using a series of clade models that partitioned the data set according to phylogeny, habitat, and foraging depth zone. Only the model partitioning according to depth was significant. This suggests that foraging dives might be a selective regime influencing cetacean rhodopsin divergence, and our experimental results indicate that spectral tuning may be playing an adaptive role in this process. Our study demonstrates that combining computational and experimental methods is crucial for gaining insight into the selection pressures underlying molecular evolution.**

***Key words:*** **cetacean vision, evolution of protein structure and function, codon substitution model, $d_N/d_S$, clade model, adaptive evolution, opsins, visual ecology, absorption spectra, dim-light vision, site-directed mutagenesis.**

## Introduction

Cetaceans are the only mammals that have evolved a fully aquatic lifestyle and diversified into all major aquatic ecosystems, a remarkable transition that has been characterized by extreme morphological and physiological adaptations (Uhen 2010; Gatesy et al. 2013). The molecular mechanisms underlying these major evolutionary changes are only just beginning to be understood, largely due to advances in cetacean whole-genome sequencing projects (Lindblad-Toh et al. 2011; Sun et al. 2013; Zhou et al. 2013; Moura et al. 2014; Yim et al. 2014). Recent studies that exploit cetacean sequence data have correlated large-scale patterns of positive selection and gene loss with key aquatic adaptations including limb development, diving physiology, echolocation, and dim-light vision (Gatesy et al. 2013; McGowen et al. 2014). Because light detection is often essential for behaviours such as predator avoidance, mate selection, and foraging, its evolution has been directly linked to survivorship and reproductive fitness (Nilsson 2013). Marine environments present a particular challenge to vision because light intensity is reduced, and the spectrum is narrowed and blue-shifted relative to terrestrial environments (Warrant and Locket 2004). Compared with most terrestrial mammals (including their nearest living relatives, the hippopotamids) cetaceans have a reduced number of visual pigments (reviewed in Bowmaker 2008 and Jacobs 2013), and the ones they do possess show evidence of spectral shifts (Fasick et al. 1998; Fasick and Robinson 2000; Newman and Robinson 2005). Nevertheless, the structure–function mechanisms responsible for these differences have not been experimentally investigated in an evolutionary context in any cetacean pigment.

The killer whale (*Orcinus orca*) is one of the best-studied cetacean species, making it a model for toothed cetacean (Odontoceti) biology and behavior in general (Baird 2000). It is the largest member of Delphinidae, and though it has a cosmopolitan distribution in the world's oceans, unlike other widely distributed delphinids, it tends not to venture beyond the continental shelves into open ocean habitats (Wang et al. 2014). Most foraging takes place in epipelagic waters where it is an apex predator, with over 140 species documented as prey including fish, cephalopods, marine mammals, seabirds, and marine turtles (Wang et al. 2014). In other cetaceans, foraging has likely driven several instances of ecological convergence, such as the evolution of bathypelagic foraging







Article

(below 1,000 m) in the sperm whales and beaked whales (Tyack et al. 2006; Watwood et al. 2006), and independent freshwater invasions by the river dolphin clades (Cassens et al. 2000; Hamilton et al. 2001). Though the killer whale genome has recently become publicly available (Moura et al. 2014), its visual pigments have yet to be experimentally investigated.

Visual pigments are molecular complexes consisting of an opsin apoprotein covalently bound through a Schiff base linkage to a vitamin A-derived chromophore, 11-*cis* retinal (Wald 1968; Sakmar et al. 1989; Palczewski et al. 2000). Visual pigments are activated when incoming photons trigger isomerization of 11-*cis* retinal to all-*trans* retinal, which results in a conformational change in the associated heptahelical opsin structure, giving rise to metarhodopsin II, the biologically active form (Lamb and Pugh 2004). Visual pigment molecules are contained within the outer segments of photoreceptor cells, and are the first component of the visual transduction pathway. Similar to other mammals, cetaceans are known to possess three major classes of visual pigments: A rhodopsin (RH1) that mediates dim-light vision, and two classes of cone opsins (LWS and SWS1) that mediate daylight and color vision, though in all cetaceans that have been analyzed so far, SWS1 has been found to be pseudogenized (Bowmaker 2008; Meredith et al. 2013).

Adaptive shifts in the peak wavelength of maximal absorbance (λmax) of visual pigments have been proposed to evolve in response to light environment, particularly in aquatic and nocturnal organisms (e.g., Bowmaker 2008; Hunt et al. 2009; Jacobs 2009). Prior studies of cetacean pigments have found that they tend to be blue-shifted relative to terrestrial mammals, with the greatest shifts occurring in deep-diving species (McFarland 1971; Fasick et al. 1998; Fasick and Robinson 2000). Substitutions at three specific amino acid sites (83, 292, and 299) are thought to be largely responsible for these shifts (Fasick and Robinson 1998), but these have yet to be experimentally mutated in a cetacean pigment.

In this study, we have used in vitro expression and site-directed mutagenesis experiments to investigate spectral tuning in the rhodopsin from the killer whale. To determine the potential for adaptive molecular evolution in cetacean rhodopsin, we combined our experimental assays of killer whale rhodopsin with computational investigations of selection patterns in cetacean rhodopsin genes and their association with ecological variables such as foraging depth. We hypothesized that these ecological differences among cetaceans may be driving divergent evolution in rhodopsin. We investigated this hypothesis using codon-based random sites and clade models to test for positive selection and divergence in cetacean rhodopsin evolution. Combined with our in vitro experiments measuring the functional consequences of amino acid substitutions at spectral tuning sites, our in silico analyses support the hypothesis that cetacean rhodopsins may have been the targets of positive selection in order to adapt to changes in aquatic light environments, particularly in response to foraging depth.

## Results

### Killer Whale Rhodopsin Expression and Site-Directed Mutagenesis

Like most cetaceans, the killer whale possesses an intact gene coding for rhodopsin (RH1), although some of its other visual pigments, such as SWS1, may have been pseudogenized (Moura et al. 2014). We used heterologous expression methods to express and characterize the absorption maximum (λmax) of the wild-type killer whale rhodopsin (Morrow and Chang 2010). The expressed, purified rhodopsin was found to have a λmax that was 12 nm blue-shifted relative to the bovine control (fig. 1*A* and table 1). The killer whale pigment was also stable in hydroxylamine, similar to bovine rhodopsin (Sakmar et al. 1989; Kawamura and Yokoyama 1998) (supplementary fig. S1, Supplementary Material online). Acid bleaching of killer whale rhodopsin shifted its λmax to 440 nm (fig. 1*B*), indicative of an intact covalent linkage with the retinal chromophore (Kito et al. 1968). All RH1 pigments were capable of being bleached with light to approximately 380 nm, indicating the formation of the light-activated state, metarhodopsin II (fig. 1*A*).

We then used site-directed mutagenesis to investigate the effects of mutating spectral tuning residues 83, 292, and 299 on the λmax of cetacean rhodopsin. Single mutations of each of these sites to the bovine rhodopsin residue resulted in clear spectral shifts (fig. 1*C*). S292A resulted in the largest red-shift (10 nm), with N83D resulting in a small red-shift (2 nm), and S299A in a small blue-shift (2 nm). A mutant with both red-shifting substitutions (N83D+S292A) was found to have synergistic effects that resulted in a large red-shift (14 nm) (fig. 1*C* and table 1).

To gain insight into the mechanisms by which the substitutions at these sites may influence λmax, we constructed a homology model of the wild-type killer whale rhodopsin based on the dark state bovine rhodopsin crystal structure (Palczewski et al. 2000). Overall, the structure of killer whale rhodopsin was very similar to bovine rhodopsin (fig. 2*A*). Sites 292 and 299 on helix VII were clustered near the chromophore and counterion, within 10 Å of the Schiff base (though 292 was closer at ~5 Å) (fig. 2*B*–*E*). The serine substitutions at these sites in killer whale, unlike the alanine residues in bovine, result in polar side chains in close proximity to the protonated Schiff base end of the chromophore. Site 83 is located farther away (>10 Å) from the Schiff base on helix II, and is highly conserved as aspartic acid (D) in terrestrial vertebrate rhodopsins and other G protein-coupled receptors (GPCRs) where it participates in a hydrogen bond network with other residues that may influence chromophore activation (Okada et al. 2002). Though the N83 substitution in killer whale is not charged, the configuration of this network appears comparable to the bovine rhodopsin dark-state (fig. 2*F*).

### Molecular Evolutionary Analyses

To investigate patterns of selection in cetacean rhodopsin, we used codon-based likelihood models (reviewed in Anisimova and Koisol 2009) to estimate $d_N/d_S$ for a data

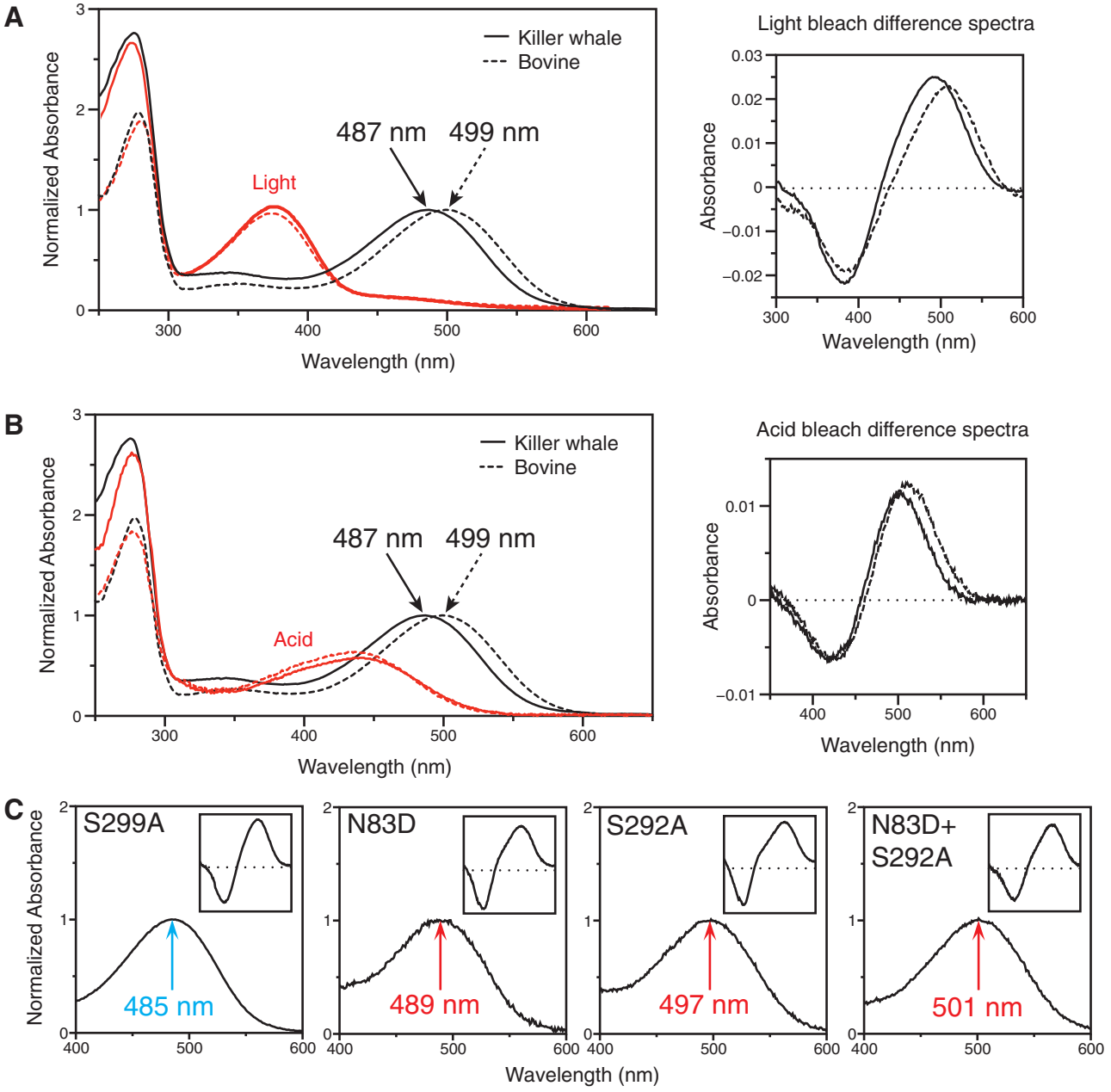


**Fig. 1.** Absorption spectra for wild-type killer whale and bovine control showing bleaching by (*A*) light and (*B*) acid, and absorption spectra for (*C*) mutant killer whale rhodopsins. The indicated λmax values were estimated according to the curve-fitting methodology of Govardovskii et al. (2000). The red lines show the shifted spectra that result after sample exposure to white light (λmax = 380 nm), or to 100 mM hydrochloric acid (λmax = 440 nm). Dark-light and dark-acid difference spectra are shown to the right in panel (*A*) and (*B*), and as insets in panel (*C*).

**Table 1.** Spectral Tuning of Killer Whale and Bovine Rhodopsins Measured In Vitro.

| Species | Mutant | λmax | 83, 292, 299 |
|---|---|---|---|
| *Bos taurus* | Wild-type | 499 | D A A |
| *Orcinus orca* | N83D+S292A | 501 | D A S |
| *Orcinus orca* | S292A | 497 | N A S |
| *Orcinus orca* | N83D | 489 | D S S |
| *Orcinus orca* | Wild-type | 487 | N S S |
| *Orcinus orca* | S299A | 485 | N S A |

Note.—λmax values presented for this study are the mean of three dark spectrum measurements. 83, 292, and 299 refer to the amino acid positions of the rhodopsin protein sequence. The amino acid identities are indicated with standard abbreviations (A, alanine; D, aspartic acid; N, asparagine; S, serine).

set of cetacean rhodopsin coding sequences. An alignment (supplementary fig. S3, Supplementary Material online) of full rhodopsin coding sequences from 23 cetacean species (supplementary table S1, Supplementary Material online) was analyzed both with random sites models implemented in PAML (Yang 2007), as well as models that independently estimate rates of $d_S$ and $d_N$ in HYPHY (Pond et al. 2005). We found significant evidence not only for variation in the nonsynonymous to synonymous substitution rate ratio, $\omega$ ($d_N/d_S$; M3 vs. M0, $P < 0.0001$; table 2 and supplementary table S2, Supplementary Material online) but also for pervasive

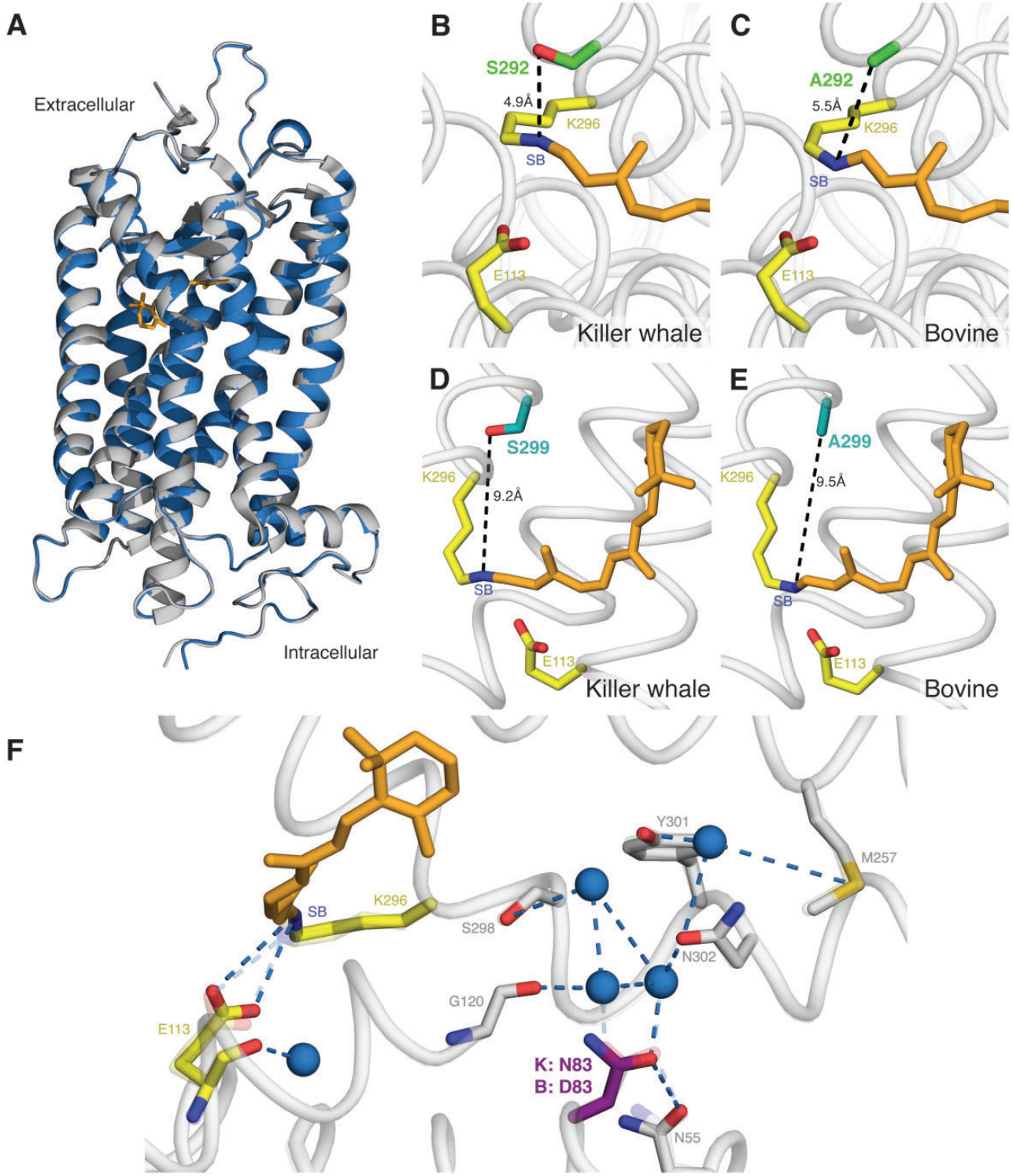


FIG. 2. Homology model of killer whale rhodopsin. (*A*) Overall, bovine (gray) and killer whale (blue) rhodopsin were closely aligned. (*B*–*E*) Killer whale residues at spectral tuning sites 292 and 299 add hydroxyl groups near the chromophore (orange) and reduce side-chain distance to the Schiff base (SB). (*F*) Superposition of killer whale (solid) and bovine (transparent) rhodopsin showing participation of site 83 in a hydrogen-bond network (waters and H-bonds shown as blue spheres and dotted lines, respectively) that may influence chromophore isomerization (Palczewski et al. 2000; Okada et al. 2002). The different residues in killer whale (N83) and bovine (D83) appear to only minimally impact the configuration of this network.

positive selection on a subset of sites (M2a vs. M1a, M8 vs. M7, M8 vs. M8a, $P < 0.01$; table 2 and supplementary table S2, Supplementary Material online). We found similar results using both a species tree (fig. 3) of commonly accepted cetacean relationships (McGowen 2011; Gatesy et al. 2013) as well as a rhodopsin gene tree (supplementary fig. S2, Supplementary Material online). HYPHY's PARRIS test (Scheffler et al. 2006), which resembles the M2a versus M1a test but with variation in synonymous rates ($d_S$) incorporated, was also significant ($P < 0.05$,

**Table 2.** LRTs for Random-Sites Models (PAML) of the Cetacean RH1 Species Tree.

| Model | np | ln *L* | $\kappa$ | Parameters[a] | | | Null | LRT | df | *P* |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | $\omega_0/p$ | $\omega_1/q$ | $\omega_2/\omega_p$ | | | | |
| M0 | 49 | −3,166.43 | 4.12 | 0.12 | | | | | | |
| M1a | 50 | −3,052.82 | 4.29 | 0.02 (88.0) | 1 (12.0) | | | | | |
| M2a | 52 | −3,047.29 | 4.35 | 0.02 (87.9) | 1 (11.6) | 7.13 (0.5) | M1a | 11.1 | 2 | 0.004* |
| M3 | 53 | −3,048.16 | 4.34 | 0.02 (87.5) | 0.92 (11.9) | 6.85 (0.6) | M0 | 236.5 | 4 | 0.000* |
| M7 | 50 | −3,054.33 | 4.27 | 0.05 | 0.33 | | | | | |
| M8a | 51 | −3,052.29 | 4.29 | 0.18 | 5.64 | 1.0 (10.4) | | | | |
| M8 | 52 | −3,048.73 | 4.35 | 0.08 | 0.76 | 2.46 (2.9) | M7 | 11.2 | 2 | 0.004* |
| | | | | | | | M8a | 7.1 | 1 | 0.008* |

NOTE.—np, number of parameters; ln *L*, ln likelihood; $\kappa$, transition/transversion ratio; df, degrees of freedom.

[a]For models M0–M3, the $\omega$ values for each site class ($\omega_0$–$\omega_2$) are shown with their proportions in parentheses. For models M7–M8, *p* and *q* describe the shape of the beta distribution, and $\omega_p$ refers to the positively selected site class (with proportion in parentheses) for models M8 and M8a (where it is constrained to one). *Indicates statistical significance ($p < 0.05$).

supplementary table S3, Supplementary Material online), indicating that this result is robust to variation in $d_S$.

## Positively Selected Sites and Spectral Tuning

Codon-based random sites models were also used to identify specific amino acid sites under positive selection. In total, we found 29 positively selected sites in cetacean rhodopsin as identified with a posterior probability of at least 80% of belonging to a positively selected site class under PAML's M8 model (8 sites), HYPHY's REL (26 sites) or FUBAR (10 sites) models (table 3 and supplementary fig. S3, Supplementary Material online). Three sites (16, 104, and 299) were identified by all three of the models, and 12 were identified by at least two of the models (fig. 4 and table 3). Six of the M8 sites and all of the FUBAR sites were identified by at least one of the other two models, but the less stringent REL model identified 15 sites that were not detected by the other two models. The two M8-only sites (194 and 196) had low posterior probabilities and $\omega < 1$ under the HYPHY models. Generally, sites identified as positively selected in both the HYPHY models but not M8 were nevertheless assigned to the positively selected site class in M8, only with posterior probabilities less than 80%.

Of the three spectral tuning sites investigated by site-directed mutagenesis in our in vitro experiments, site 299 was identified as positively selected in all three models. Site 292 was identified by M8 and REL, but not FUBAR (though $\omega$ was still slightly $> 1$). Site 83 had the weakest support for positive selection, and was only identified by REL; though slightly elevated relative to the majority of other sites, $\omega$ was estimated as less than 1 in both M8 and FUBAR. The levels of support for these three sites are likely related to their different levels of variation on the cetacean species tree. The codon-based tests for positive selection that we have used are optimized to detect diversifying rather than directional selection, and so are most powerful when sites undergo frequent substitutions, including convergences and reversals. Sites 292 and 299 both show greater variation on the cetacean tree than 83. However, despite the varying support for positive selection at these three sites, all were placed in the divergent site class under the CmC (Clade model C) depth models (see below).

## Clade Model Tests of Divergence

We used clade models to investigate the relationship between divergence and ecological variables. We hypothesized that ecological differences among cetaceans occupying different aquatic ecosystems, and that forage at different depths, may be driving divergence in rhodopsin evolution. To test these hypotheses, we implemented CmC, designing partitions to reflect differences in habitat (oceanic, coastal, and freshwater) and foraging depth (epipelagic, mesopelagic, and bathypelagic) (fig. 3). For the likelihood ratio tests (LRTs) among nested models, only the three- and four-partition depth models performed better than the simplest (null) model, M2a_rel, which suggests that foraging depth, rather than habitat or phylogeny, is the main factor driving divergence among cetacean rhodopsins. In the four-partition model $d_N/d_S$ was lowest for the mesopelagic clades (near neutral at $\omega = 0.68$), but increased above 1 for the epipelagic and bathypelagic foragers, the bathypelagic foragers having the greatest value ($\omega = 2.54$; table 4, "CmC: Depth 4-partition"). However, the three-partition model that grouped bathypelagic and epipelagic clades together was a significantly better fit than the four-partition model, as well as the two-partition model with only the bathypelagic+epipelagic clade in the foreground (table 4, "CmC: Depth 3-partition," "CmC: Depth 2-partition"). Two- and three-partition models that placed epipelagic and bathypelagic clades in the foreground separately also did not show significant divergence (data not shown). In other words, although divergent selection signals were strongest among near-surface foragers and extreme divers, mesopelagic foraging clades were also contributing to the significant divergence pattern. Evaluating the three-partition depth model against a null model where the bathypelagic+epipelagic divergent site class was constrained to equal one was also significant (table 4, "CmC: Depth 3-partitionC"), confirming the presence of positive selection within the bathypelagic and epipelagic foragers. The best fitting model overall (according to Akaike's Information Criterion, AIC) was the three-partition foraging depth model (table 4, "CmC: Depth 3-partition"), with the $\omega$ value for bathypelagic+epipelagic clade being similar to that

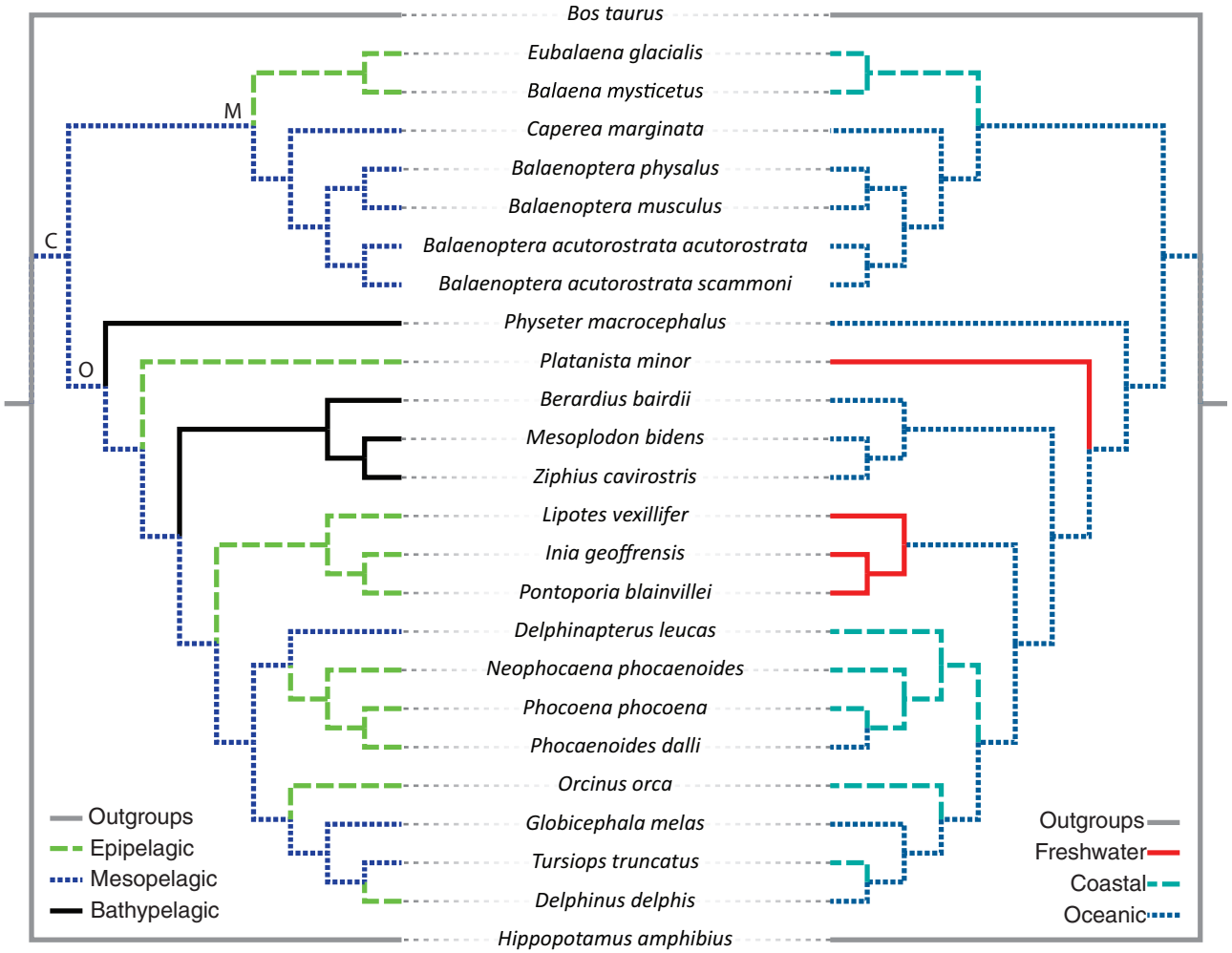


**Fig. 3.** Tree topology used in codon-based likelihood models showing lineage partitions according to foraging depth (left) and habitat (right). Zone definitions: Epipelagic (0–200 m), mesopelagic (200–1,000 m), bathypelagic (> 1,000 m), oceanic (beyond continental shelves), coastal (within continental shelves), freshwater (rivers, lakes, estuaries). Please see supplementary table S5, Supplementary Material online, for details regarding assignment of lineages to different groups. C, Cetacea; M, Mysticeti; O, Odontoceti. Note that the unrooted tree is required for model implementation.

estimated for the positively selected site class in the M8 model.

Cetacean rhodopsin evolution did not appear to be significantly driven by the evolutionary differences between toothed and baleen species, as there was no evidence for divergent selection occurring between the two cetacean suborders, Odontoceti and Mysticeti using CmC (table 4). Branch-site models were also implemented, as they are increasingly being used to investigate the evolution of both branches and clades (e.g., Spady et al. 2005; Badouin et al. 2013; Veilleux et al. 2013; Schott et al. 2014). However, the branch-site model can only accommodate two partitions, so we could not evaluate any of our higher partition configurations this way. Only the model with Odontoceti in the foreground turned out to fit significantly better than the null model ($P < 0.05$; supplementary table S4, Supplementary Material online), even though the equivalent clade model test was not significant. Of the branch-site models, the model with all Cetacea in the foreground had the best fit according to AIC, and was the only branch-site model that performed better than its equivalent CmC model under this criterion. For all but the Cetacea-foreground partition, the branch-site models were also worse fits than the random-sites M2a-rel model, which assumes no partitions in the data set (supplementary table S4, Supplementary Material online). Because these models do not allow background positive selection (even though both Odontoceti and Mysticeti appear to have $\omega$ values greater than 1, with Odontoceti being more pronounced), the poor fits relative to the null are consistent. Overall the branch-site models seemed to have a weaker performance on our data set, and our clade model results suggest that this is because divergent patterns in $\omega$ cannot be accommodated by its restrictions.

## Discussion

By combining mutation experiments and protein function assays with molecular evolutionary analyses, we have found evidence for functional adaptation in cetacean rhodopsin. The rhodopsin of the killer whale (*Orcinus orca*), a delphinid, was expressed in vitro for the first time, along with a series of mutants with substitutions at the hypothesized spectral tuning sites, 83, 292, and 299. We found a maximal

**Table 3.** $\omega$ Values for Positively Selected Sites under Random-Sites Models from PAML (M8) and HYPHY (REL, FUBAR).

| | M8 | | REL | | FUBAR | |
|---|---|---|---|---|---|---|
| Site | $\omega$ | Post. Probability | $\omega$ | Post. Probability | $\omega$ | Post. Probability |
| 7 | 1.063 | 0.437 | 2.055 | 0.962 | 1.238 | 0.533 |
| 16[a] | 1.939 | 0.968 | 1.711 | 0.950 | 2.811 | 0.877 |
| 26 | 0.547 | 0.150 | 1.703 | 0.924 | 0.921 | 0.628 |
| 39 | 0.450 | 0.100 | 1.883 | 0.910 | 1.248 | 0.625 |
| 49 | 0.449 | 0.098 | 1.933 | 0.915 | 1.339 | 0.638 |
| 83 | 0.777 | 0.246 | 1.725 | 0.951 | 0.955 | 0.682 |
| 99[b] | 1.631 | 0.760 | 2.190 | 0.996 | 4.335 | 0.877 |
| 104[a] | 1.883 | 0.931 | 1.983 | 0.978 | 8.014 | 0.966 |
| 112 | 1.641 | 0.765 | 2.079 | 0.985 | 2.649 | 0.745 |
| 137 | 0.470 | 0.110 | 1.957 | 0.921 | 1.392 | 0.646 |
| 151 | 0.346 | 0.049 | 1.548 | 0.848 | 0.659 | 0.561 |
| 159 | 0.291 | 0.027 | 1.925 | 0.937 | 1.374 | 0.651 |
| 162 | 0.250 | 0.015 | 1.719 | 0.893 | 0.952 | 0.601 |
| 165[b] | 1.463 | 0.659 | 2.174 | 0.994 | 3.344 | 0.844 |
| 194 | 1.826 | 0.891 | 0.620 | 0.595 | 0.789 | 0.375 |
| 195[b] | 1.973 | 0.998 | 0.658 | 0.630 | 5.500 | 0.887 |
| 196 | 1.936 | 0.963 | 0.294 | 0.011 | 0.320 | 0.010 |
| 198 | 0.514 | 0.085 | 2.018 | 0.980 | 1.884 | 0.751 |
| 213[b] | 1.076 | 0.423 | 2.018 | 0.981 | 3.327 | 0.869 |
| 216 | 1.153 | 0.474 | 2.028 | 0.981 | 2.098 | 0.766 |
| 259[b] | 1.748 | 0.831 | 1.227 | 0.851 | 0.524 | 0.235 |
| 266[b] | 1.303 | 0.564 | 2.132 | 0.990 | 2.779 | 0.816 |
| 282 | 1.198 | 0.498 | 1.728 | 0.952 | 1.398 | 0.772 |
| 290[b] | 0.911 | 0.314 | 1.829 | 0.963 | 2.165 | 0.814 |
| 292[b] | 1.863 | 0.910 | 1.322 | 0.883 | 1.037 | 0.498 |
| 299[a] | 1.845 | 0.903 | 2.031 | 0.982 | 7.063 | 0.959 |
| 329 | 1.644 | 0.766 | 1.955 | 0.876 | 1.733 | 0.706 |
| 333[b] | 0.885 | 0.297 | 2.031 | 0.982 | 2.611 | 0.830 |
| 335 | 0.552 | 0.108 | 2.022 | 0.980 | 1.685 | 0.728 |

NOTE.—Sites were identified as positively selected if they had a posterior probability greater than 80% of belonging to the positively selected site class.
[a]Site was identified by all three models.
[b]Site was identified by two models.

absorbance of 487 nm in the wild-type pigment, which was considerably blue-shifted relative to a terrestrial mammal (bovine, 499 nm). Site-directed mutagenesis experiments in the killer whale rhodopsin background showed that mutations at the three tuning sites were found to account for the spectral difference between the killer whale and bovine rhodopsin. These results are consistent with previous mutagenesis studies of these sites in a bovine background (Fasick and Robinson 1998), as well as with other wild-type cetacean rhodopsins with equivalent substitutions (Bischoff et al. 2012), although we did find variation in the magnitude of spectral shifts. Furthermore, we found significant evidence for positive selection in a data set of cetacean rhodopsins with up to 29 positively selected sites, including 83, 292, and 299. We also found evidence to suggest that divergent selection in the gene is driven, at least in part, by ecological differences related to diving behavior. Here, we discuss our results in the context of both rhodopsin structure-function and cetacean ecology.

## The Role of Sites 83, 292, and 299 in Cetacean Rhodopsin Spectral Tuning

We expressed a killer whale visual pigment for the first time, and found the spectral tuning of its rhodopsin (487 nm) to be substantially blue-shifted relative to bovine rhodopsin (499 nm). This is similar to other delphinids whose pigments have been expressed and assayed in vitro (488–489 nm; Fasick et al. 1998; Fasick and Robinson 2000). In the past, this difference between terrestrial and marine mammal rhodopsin has been attributed to substitutions at sites 83, 292, and 299. Not only are these sites implicated in the spectral tuning of rhodopsin in other aquatic vertebrates (e.g., Hunt et al. 1996, 2001; Hope et al. 1997; Sugawara et al. 2005), but mutations at these sites in bovine rhodopsin are also thought to mimic the range of observed cetacean rhodopsin $\lambda$max (Fasick and Robinson 1998, 2000). In addition, Meredith et al. (2013) found significant evidence for positive selection on sites 83 and 292 (but not 299) in cetacean rhodopsin using branch-sites models. Residues at these sites have been used to predict

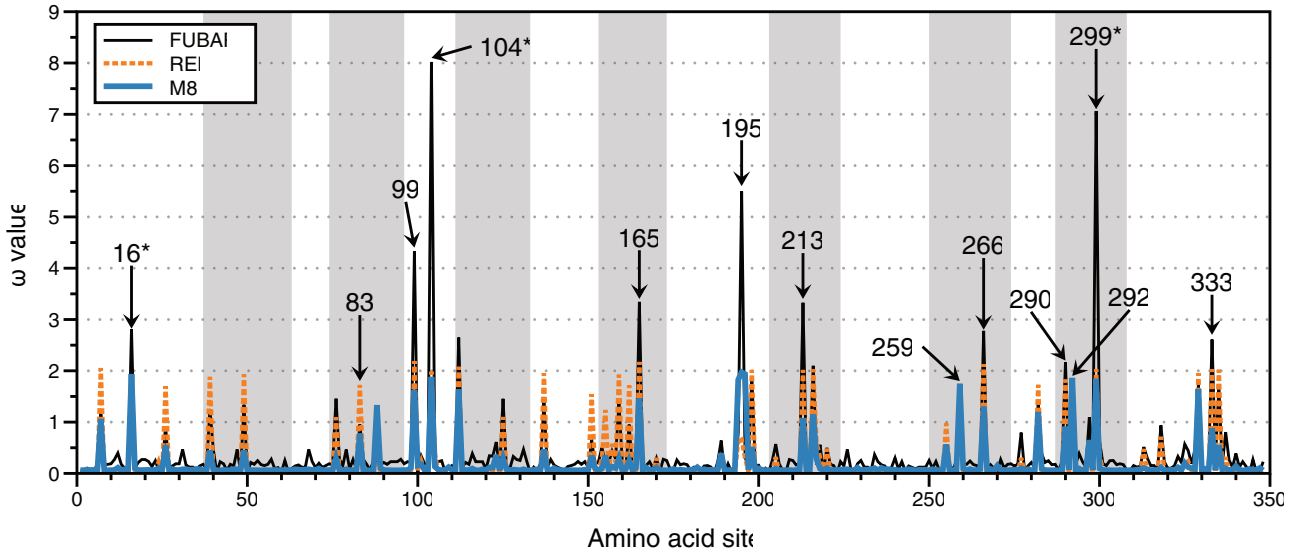


FIG. 4. $\omega$ values for each site of cetacean rhodopsin. Shaded regions indicate transmembrane helix domains. Sites identified as positively selected with $\geq$80% posterior probability in at least two of the random-sites models, M8 (PAML), REL and FUBAR (HYPHY), are labeled (exception site 83, identified by REL only, shown for spectral tuning context). Asterisks indicate that the site was identified by all three models.

the rhodopsin λmax of different cetacean species (e.g., Fasick et al. 2011; Bischoff et al. 2012), though for many cetaceans this has yet to be confirmed by in vitro expression experiments. Variation at site 292 has been associated with blue shifts in the rhodopsins of deep-dwelling fish that can range from 11 to 14 nm depending on the species (Sugawara et al. 2005). Site 299 has not been as well investigated in previous experimental studies; however, given its increased distance from the chromophore relative to site 292, its lesser effect on λmax in our study was not surprising. The killer whale residue, N83, has also long been associated with varying degrees of blue-shifts in deep-dwelling teleost fish rhodopsins (Hunt et al. 1996, 2001; Hope et al. 1997; Sugawara et al. 2005), and in the nocturnal echidna, mutating N83D red-shifts λmax by as much as 6 nm (Bickelmann et al. 2012). The mutagenesis data for site 83 in particular illustrate that predicting λmax based on sequence alone can be a risky proposition, as the spectral effects of substitutions at a site can vary substantially depending on the genetic background in which those substitutions were made (Asenjo et al. 1994; Hauser et al. 2014).

In this study, we used site-directed mutagenesis to investigate the contributions of sites 83, 292, and 299 to the blue-shifted killer whale λmax relative to terrestrial mammals. Our results support the notion that although these sites are clearly the primary determinants of spectral tuning, there may be other sites that interact with these to affect cetacean rhodopsin λmax. Though the directions of the λmax shifts in our killer whale mutants were consistent, the magnitudes of the shifts were found to differ compared with the reverse mutations in bovine rhodopsin (Fasick and Robinson 1998). The degree of shifts present in our mutant killer whale rhodopsins also accounts for some, but not all of the spectral shifts found in other expressed wild-type cetacean rhodopsins. For example, the North Atlantic right whale (*Eubalaena glacialis*) has N83, A292, and S299, which according to our mutagenesis results should result in a λmax of approximately 497 nm, but Bischoff et al. (2012) measured it as 493 nm. A wide variety of sites have been shown to influence visual pigment λmax in various vertebrate species (Hunt et al. 2001, 2009; Bowmaker and Hunt 2006; Bowmaker 2008), and some of these clearly warrant further investigation in cetaceans. In addition, spectral tuning sites (site 83 in particular) have also been implicated in nonspectral functions such as active state (Meta II) kinetics (e.g., Sugawara et al. 2010), but these functions, and their potential adaptive relevance, have yet to be investigated in cetaceans.

## Positive Selection in Cetacean Rhodopsin

Our results demonstrated significant evidence for positive selection in cetacean rhodopsins. Even though there are many mammalian groups with specialized visual abilities, particularly for dim-light environments, evidence for positive selection has not typically been found for mammalian rhodopsin data sets, and overall substitution rate ratios tend to be very low (e.g., M0 $\omega = 0.04$ in Zhao et al. [2009] vs. our data set M0 $\omega = 0.12$). One possibility is that selection signatures in such data sets are too weak to be detected, especially if taxonomic coverage is wide. Rhodopsin selection signals are generally weaker relative to other proteins due to the majority of the amino acid sequence being highly conserved. In other vertebrate groups, population-level polymorphism data have been crucial for detecting signatures of selection among closely related species for highly conserved genes, including rhodopsin (Larmuseau et al. 2010). Our random-sites models indicated that only a small portion of sites (3%) were under positive selection. This is consistent with other positively selected vertebrate rhodopsins (e.g., Spady et al. 2005; Rennison et al. 2012; Schott et al. 2014), but does not approach the high proportions and substitution rate ratios typically found in rapidly diversifying genes, such as those that encode

**Table 4.** Results from Clade Model (CmC—PAML) Tests for Divergence Partitioned by Phylogeny, Foraging depth, and Habitat.

| Model and Partition[a] | np | ln *L* | $\kappa$ | Parameters[b] | | | ΔAIC[c] | Null | LRT | df | *P* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | $\omega_0$ | $\omega_1$ | $\omega_2/\omega_d$ | | | | | |
| M2a_rel | 52 | −3,047.29 | 4.35 | 0.02 (87.9) | 1 (11.6) | 7.13 (0.5) | 4.34 | | | | |
| CmC: Phylo (4-partition) | 55 | −3,047.15 | 4.31 | 0.02 (88.2) | 1 (3.7) | 0.26 (8.1)<br>Root: 1.08<br>Myst: 1.46<br>Odon: 1.57 | 10.06 | M2a_rel<br>2-partition<br>3-partition | 0.28<br>0.26<br>0.16 | 3<br>2<br>1 | 0.964<br>0.878<br>0.689 |
| CmC: Phylo (3-partition) | 54 | −3,047.23 | 4.31 | 0.02 (88.1) | 1 (3.8) | 0.24 (8.1)<br>Myst+Root: 1.33<br>Odon: 1.56 | 8.22 | M2a_rel<br>2-partition | 0.12<br>0.10 | 2<br>1 | 0.942<br>0.752 |
| CmC: Phylo (2-partition) | 53 | −3,047.28 | 4.35 | 0.02 (87.9) | 1 (11.2) | 5.09 (0.9)<br>Odon: 7.85 | 6.32 | M2a_rel | 0.02 | 1 | 0.888 |
| CmC: Habitat (4-partition) | 55 | −3,045.78 | 4.31 | 0.02 (86.5) | 1 (6.8) | 0.0 (6.6)<br>Ocean: 0.71<br>Coast: 1.23<br>Fresh: 2.48 | 7.32 | M2a_rel<br>2-partition<br>3-partition | 3.02<br>2.88<br>2.88 | 3<br>2<br>1 | 0.389<br>0.237<br>0.090 |
| CmC: Habitat (3-partition) | 54 | −3,047.22 | 4.31 | 0.02 (86.7) | 1 (6.7) | 0.0 (6.4)<br>Salt: 0.91<br>Fresh: 2.43 | 8.20 | M2a_rel<br>2-partition | 0.14<br>0.00 | 2<br>1 | 0.932<br>1.000 |
| CmC: Habitat (2-partition) | 53 | −3,047.22 | 4.33 | 0.02 (87.3) | 1 (11.2) | 0.0 (1.5)<br>Fresh: 2.41 | 6.20 | M2a_rel | 0.14 | 1 | 0.708 |
| CmC: Depth (4-partition) | 55 | −3,042.96 | 4.33 | 0.02 (87.7) | 1 (5.0) | 0.09 (7.1)<br>Epi: 1.91<br>Meso: 0.68<br>Bath: 2.54 | 1.68 | M2a_rel<br>2-partition<br>3-partition | 8.66<br>6.84<br>0.32 | 3<br>2<br>1 | 0.034*<br>0.033*<br>0.572 |
| CmC: Depth (3-partition) | 54 | −3,043.12 | 4.33 | 0.02 (87.8) | 1 (5.1) | 0.09 (7.2)<br>Meso: 0.69<br>Epi+Bath: 2.12 | 0.00 | M2a_rel<br>2-partition<br>3-partitionC | 8.34<br>6.52<br>4.54 | 2<br>1<br>1 | 0.015*<br>0.011*<br>0.033* |
| CmC: Depth (3-partitionC) | 53 | −3,045.39 | 4.33 | 0.01 (85.4) | 1 (6.7) | 0.0 (7.9)<br>Meso: 0.47<br>Epi+Bath: 1.00 | 2.54 | M2a_rel | 3.80 | 1 | 0.051 |
| CmC: Depth (2-partition) | 53 | −3,046.38 | 4.34 | 0.02 (87.9) | 1 (5.7) | 0.40 (6.4)<br>Epi+Bath: 2.12 | 4.52 | M2a_rel | 1.82 | 1 | 0.177 |

Note.— np, number of parameters; ln *L*, ln likelihood; $\kappa$, transition/transversion ratio; df, degrees of freedom; Cet, Cetacea; Odon, Odontoceti; Myst, Mysticeti; Root, root branch of cetacean; Ocean, oceanic; Coast, coastal; Fresh, freshwater/estuarine; Meso, mesopelagic; Epi, epipelagic; Bath, bathypelagic.

[a]Partitions for Habitat and Depth are explained in figure 3, and for 2- and 3-partition categories, the model with the highest ln-likelihood is shown.

[b]$\omega$ values for each site class ($\omega_0$–$\omega_2$) are shown with the proportion of each in parentheses. $\omega_d$ refers to the divergent site class in the CmC models, which has a separate value for each partition: The first value is for the background, followed by the foreground clade(s).

[c]Minimum overall AIC (foraging depth 3-partition of CmC; 6,194.24) was used for all comparisons.

*Indicates statistical significance ($p < 0.05$).

major histocompatibility complex (MHC) proteins (Swanson et al. 2001). More recent work has also revealed that mammalian rhodopsins have elevated synonymous substitution rates and experience positive selection at synonymous sites (Du et al. 2014); nonsynonymous positive selection may be more difficult to detect as a consequence of this pattern. Important sites may not always be detected by computational analyses of selection patterns, making experimentally derived protein structure-function data critical for molecular evolutionary studies. Our recovery of a significant signal of positive selection highlights the suitability of cetaceans for studies of molecular adaptation, as they are a group that is both ecologically diverse and relatively closely related (McGowen et al. 2014).

Although a variety of other genes show evidence of positive selection in cetaceans (reviewed in McGowen et al. 2014), this has not typically been the case for the visual pigments, despite experimental evidence for functionally important variation. In particular, tests for positive selection using random-sites models have not been significant, though evidence was found for positive selection using branch-site models (Meredith et al. 2013). Our data set indicated that homoplasies (e.g., convergences and reversals) were frequent enough in the cetacean RH1 gene to result in an inaccurate tree topology, an idea that was supported by recovery of a more accurate topology when nonsynonymous nucleotide positions were removed from the alignment (supplementary fig. S2, Supplementary Material online). Interestingly, convergent evolution appears to have led to recovery of inaccurate mammalian relationships in the past, not only in rhodopsin (Zhao et al. 2009; Shen et al. 2010) but also in other positively selected genes such as prestin (SLC26A5), which is linked to high-frequency hearing (Liu, Cotton, et al. 2010; Elgoyhen and Franchini 2011). Convergence signals are so strong in mammalian prestin that the gene tree actually has a topology that places echolocating bats and cetaceans as sister lineages (Jones 2010; Li et al. 2010; Liu, Rossiter, et al. 2010). The

rhodopsin gene is also known for having sites with convergent patterns in other vertebrate data sets where it is positively selected (Larmuseau et al. 2011; Schott et al. 2014). Studies such as these suggest that, in certain cases, species trees may be more appropriate than gene trees when directional evolutionary pressure may skew phylogenetic reconstructions.

## The Role of Rhodopsin in Cetacean Visual Ecology

In addition to statistical and experimental evidence for positive selection and functional adaptation in cetacean rhodopsin, we also found evidence for higher $d_N/d_S$ in both near-surface and extremely deep foragers. Absorption of light by water increasingly narrows the light spectrum to predominantly blue wavelengths as depth increases, and in clear ocean attenuates light intensity such that deeper than approximately 200 m, the boundary of the epipelagic zone, photosynthesis is no longer supported (Wozniak and Dera 2007; Warrant and Johnsen 2013). As such, the general expectation is that organisms that live or are active in the deep-sea will have blue-shifted visual sensitivity to capitalize on the minimal amounts of available light (Lythgoe and Dartnall 1970; Hunt et al. 2001; Griebel and Peichl 2003). Codon-based clade models are increasingly being used to test hypotheses of ecological divergence (Schott et al. 2014; Van Nynatten et al. 2015), and in our case they were a useful exploratory tool for determining whether there was statistical evidence for rhodopsin divergence based on foraging depth. Significant evidence for positive selection in both the bathypelagic and epipelagic clades suggested that divergent selection signals were strongest in these lineages, likely reflecting common descent from primarily mesopelagic foraging ancestors (see supplementary table S5, Supplementary Material online). Nevertheless, the models that showed significant evidence for divergence also included a mesopelagic partition with slightly elevated $d_N/d_S$, which suggests that there is functionally important variation distinguishing this foraging profile as well. Even though the coarse partitions we used do not capture the more nuanced variation in cetacean foraging behavior, our results provide statistical evidence for an association between rhodopsin evolution and foraging depth in cetaceans that was previously only speculative. Furthermore, although prior studies have emphasized the bathypelagic foragers (e.g., Fasick and Robinson 2000), our results suggest that the same subclass of sites is also positively selected in near-surface foragers. This makes sense under a spectral tuning regime where substitutions at a few sites can result in a wide range of spectral sensitivities. In the future, with a greater sampling of cetacean species, this association could be further tested and refined.

In addition to foraging, we also investigated divergence along habitat partitions (oceanic, coastal, and freshwater). The spectrum and intensity of light can be affected by the scattering tendencies of dissolved particulate matter, which is more abundant in coastal and freshwater systems due to terrestrial runoff (e.g., the red-shifted light environment of some river systems; Costa et al. 2013). However, we did not find significant evidence for divergence with respect to this aspect of ecology, though this model was still a better fit than the model partitioned by phylogeny. If habitat variables are influencing cetacean rhodopsin evolution, it is possible we did not have enough power to detect this in our data set. In particular, our sampling of freshwater and estuarine cetaceans was low and did not include any of the independent invasions made by the Delphinidae (e.g., *Sotalia*, *Orcaella*, *Sousa*). On the other hand, this lack of significance also suggests that the significant contribution of the epipelagic clades to the foraging depth models was also not due solely to the freshwater species. Given that freshwater invasions have been implicated in the spectral tuning of rhodopsin in several cichlid fish lineages (Spady et al. 2005; Terai et al. 2006; Miyagi et al. 2012; Schott et al. 2014) and in anchovies (Van Nynatten et al. 2015), further investigations with a more complete data set of freshwater cetacean rhodopsins could be fruitful in the future.

The ways in which light can vary in underwater environments have had a profound impact on the evolution of eyes in aquatic organisms (Warrant and Locket 2004), and cetaceans are no exception. Like fish, they possess a spherical lens (Mass and Supin 2007), and like many carnivorous mammals that forage in dim-light conditions, the retina is dominated by rod photoreceptors (Peichl 2005) and is supported by a reflective *tapetum lucidum* layer (Mass and Supin 2007). Among the visual pigments, blue-shifts in λmax are present not only in rhodopsin but also in the long-wave sensitive cone opsin as well (Fasick and Robinson 1998; Fasick et al. 1998; Newman and Robinson 2005). Oddly, marine mammals (cetaceans and pinnipeds) share loss of function in the short-wave sensitive cone opsin with other dim-light inhabiting (nocturnal, fossorial) mammals (Jacobs 2013), which seems contradictory to the hypothesis of optimizing sensitivity toward downwelling light. Although other hypotheses have been put forward to explain the loss of this opsin (Peichl 2005), ultimately it remains a mystery; one that has been compounded by the recent discovery of long-wave cone opsin losses in some cetacean lineages (Meredith et al. 2013). Addressing these issues in cetacean visual pigment evolution will certainly require a more thorough understanding of pigment function even beyond spectral tuning, in the context of the varied ecologies in which cetaceans have evolved. Some have argued that the field of molecular evolution in general requires an increased emphasis on experimental studies that test potential targets of natural selection in genomic contexts by evaluating the phenotypic consequences of individual mutations (reviewed in Barrett and Hoekstra 2011). We have presented here the first results of mutation experiments on a cetacean rhodopsin in one of few studies of adaptive molecular evolution in cetaceans to combine both bioinformatic and experimental approaches. All together, our results both computationally and experimentally support adaptive evolution in cetacean rhodopsin to dim-light vision, and suggest mechanistic and ecological selection pressures that can form the basis of hypotheses for future investigations.

## Materials and Methods

### Killer Whale *RH1* Preparation and Site-Directed Mutagenesis

The wild-type killer whale rhodopsin gene (*RH1*) coding sequence, along with a C-terminal nine-amino acid epitope tag for the 1D4 antibody, was synthesized (GeneArt, Invitrogen) and inserted into the p1D4-hrGFP II expression vector (Morrow and Chang 2010). Site-directed mutagenesis (QuickChange II, Agilent) was used to generate mutants N83D, S292A, S299A, as well as a double mutant N83D/S292A in the killer whale wild-type *RH1*. All mutants were confirmed by double-stranded sequencing.

### Killer Whale Rhodopsin Expression and Spectral Tuning

Killer whale rhodopsin and mutants were transiently transfected, harvested, and purified alongside a bovine rhodopsin control, as reported previously (Morrow and Chang 2010). Briefly, vectors containing rhodopsin were transiently transfected into HEK293T cells (Lipofectamine 2000; Invitrogen), harvested, regenerated with 11-*cis*-retinal, solubilized in 1% dodecylmaltoside, and immunoaffinity purified using the 1D4 monoclonal antibody. The UV-visible absorption spectra of purified rhodopsin samples were measured using a Cary 4000 double-beam spectrophotometer (Varian) at 25 °C in the dark, and after 60 s of white light bleaching. Difference spectra were calculated by subtracting light spectra from dark spectra. λmax was estimated by fitting a template curve (Govardovskii et al. 2000) to the normalized dark spectrum. The wild-type killer whale rhodopsin was also exposed to hydrochloric acid (HCl; 100 mM) and hydroxylamine ($NH_2OH$; 50 mM) to confirm the presence of a properly functioning Schiff base link and protection of the chromophore from hydrolyzing solvent, respectively.

### Homology Modeling of Wild-Type and Mutant Killer Whale Rhodopsin

The killer whale rhodopsin 3D structure was inferred with homology modeling in the Modeller package (Sali and Blundell 1993). Wild-type bovine rhodopsin (PDB code: IU19, Okada et al. 2004) was used as a modeling template. For each structure, 100 models were generated and ranked by the Modeller objective function (optimized ten times/model), and the run with the lowest DOPE score (Shen and Sali 2006) was selected for visualization in MacPyMol (Delano Scientific). The bovine template and killer whale rhodopsin had comparable total energies (comparable *z*-scores in ProSA-web; Wiederstein and Sippl 2007), and stereochemical conformations of bonds among amino acid residues in the models had high probabilities (assessed in ProCheck; Laskowski et al. 1993), both indicators that the models were of high quality.

### Rhodopsin Sequence Alignment and Phylogenetic Analyses

Full rhodopsin coding sequences (1,044 bp) for 23 cetacean species were obtained from publicly available databases (supplementary table S1, Supplementary Material online), which covered most cetacean families. Coding sequences were also obtained from 22 other mammal species to represent major mammalian outgroup taxa during gene tree building (supplementary fig. S2, Supplementary Material online). Codons were aligned using PRANK as implemented on the webPRANK server (Löytynoja and Goldman 2010) with default settings. Our alignment of a total 45 sequences was used to generate two rhodopsin gene trees, one estimated using a maximum-likelihood method, and the other with a Bayesian method. The maximum-likelihood tree was constructed in PhyML 3.0 (Guindon et al. 2010) under the GTR+I+G model with a BioNJ starting tree, the best of either NNI or SPR tree improvement, and aBayes branch support. The GTR+I+G model was selected according to AIC values after running the alignment through MrModelTest (Nylander 2004) in PAUP 4.0 (Swofford 2002). The Bayesian analysis was run in MrBayes 3.0 (Huelsenbeck and Ronquist 2001) using a reversible jump Markov chain Monte Carlo algorithm with a gamma rate parameter (nst = mixed, rates = gamma, four chains) for two runs each of 2 million generations (25% burn-in, sampling every 100 generations). Convergence was confirmed by checking output diagnostics (standard deviation of split frequencies approached zero, scattered log-likelihood plot).

Though the resulting gene trees were highly similar, there were several major incongruences with accepted species relationships at deep nodes (supplementary fig. S2, Supplementary Material online) suggesting that convergence may have affected the gene tree, a finding that has been previously reported for mammalian rhodopsins (Zhao et al. 2009; Shen et al. 2010). As with previous studies (Shen et al. 2010), removing nonsynonymous nucleotide positions from the alignment resulted in a gene tree that was much closer to accepted species relationships (supplementary fig. S2, Supplementary Material online). However, in order to ensure that the slight topological differences did not significantly affect our results, random-sites analyses were performed on both this gene tree and the accepted species tree. Before running selection analyses in PAML, outgroups were pruned to include only the two nearest cetacean relatives, *Bos taurus* and *Hippopotamus amphibius*, so that the tree could be unrooted (as required for codeml algorithms) while maintaining the basal topology of Cetacea (fig. 3).

### Molecular Evolutionary Analyses

Patterns of selection in the data set were investigated by submitting the alignment (supplementary fig. S3, Supplementary Material online) and pruned trees (fig. 3 and supplementary fig. S2*B*, Supplementary Material online)

to the codeml program in the PAML 4.0 software package (Yang 2007). LRTs with a $\chi^2$ distribution were conducted on nested random-sites models to determine evidence for variation in $\omega$ (M3 vs. M0), and for positive selection (M2a vs. M1a, M8 vs. M7, and M8 vs. M8a). All models were run a minimum of three times with varying initial starting values for $\kappa$ (transition to transversion ratio) and $\omega$ to ensure parameter estimates did not represent local optima in the sampling space. For positive selection models, M2a and M8, PAML incorporates a Bayes' Empirical Bayes (BEB) analysis to identify specific sites (codons) under positive selection (Yang et al. 2005).

Because PAML estimates $d_N/d_S$ as a single variable, variation in this parameter is usually assumed to be due to variation in $d_N$. As such, we also used models that relax this assumption, in which $d_N$ and $d_S$ are estimated as separate parameters (REL, FUBAR, and PARRIS models in the HYPHY package; Pond and Frost 2005; Ben Murrell et al. 2013) as implemented on the Datamonkey webserver (Delport et al. 2010). The PARRIS model performs an LRT of positive selection in a manner similar to the M1a versus M2a comparison, using three $d_S$ and two $\omega$ classes ($\omega = 0$ and $\omega = 1$) in the null model, and an additional freely varying $\omega$ class in the alternate model (Scheffler et al. 2006). The REL and FUBAR models only estimate site-by-site $\omega$ values (similar to PAML's BEB analysis), and do not implement LRTs of selection patterns between nested models. FUBAR assigns sites to a much greater number of site classes than either REL or the PAML models, arguably providing more accurate site-by-site $\omega$ estimates (Murrell et al. 2013). Under the models, M8, REL, and FUBAR, sites were identified as positively selected if they had a posterior probability $\geq$80% of belonging to a positively selected site class.

We used PAML's CmC (Bielawski and Yang 2004) and branch-site models (Zhang et al. 2005) to statistically test for divergence along specific branches of the species tree used in the random-sites analyses. These models allow $\omega$ to vary not only among sites but also among foreground and background clades specified by the user. Under CmC, two classes of sites are allowed to evolve conservatively ($0 < \omega < 1$) and neutrally ($\omega = 1$) across the whole tree, but a third site class is allowed to evolve differently among two or more clade partitions. Significant divergence is indicated with an LRT comparing CmC with the null model, M2a_rel, which does not allow $\omega$ to diverge in the third site class (Weadick and Chang 2012). We designed tests for CmC based on three divergence hypotheses, each with four clade partitions: Phylogeny (Odontoceti, Mysticeti, root Cetacea, outgroups), foraging depth (bathypelagic, mesopelagic, epipelagic, outgroups), and habitat (freshwater, coastal, oceanic, outgroups) (fig. 3). Assignment of clades/nodes to habitat and depth classes was based on a priori knowledge of cetacean ecological variables for both living and extinct taxa (supplementary table S5, Supplementary Material online). For each hypothesis, the four-partition models were compared with nested three- and two-partition models using LRTs (the best fitting model in each partition category is shown in table 4). The models from each hypothesis were compared using AIC values to determine whether ecological variables had a greater influence on rhodopsin divergence than species relationships. The branch-site model, while similar to CmC, only allows one division into foreground and background lineages, so we could only test the two-partition configurations from each CmC hypothesis. In addition to the conservative and neutral site classes, there are two additional site classes where the foreground is positively selected and the background is either conserved or neutral. Of note is that the branch-site model, unlike CmC, cannot accommodate positive selection in the background lineages. Like the random-sites models, all models were run at least three times varying initial $\kappa$ and $\omega$ values to ensure local optima were avoided in the sampling space.

## Supplementary Material

Supplementary figures S1–S3 and tables S1–S5 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

## Acknowledgments

This work was supported by a Natural Sciences and Engineering Research Council (NSERC) Discovery Grant to B.S.W.C., an NSERC Postgraduate Scholarship to S.Z.D., and a Human Frontier Science Program Grant to B.S.W.C. The 11-*cis*-retinal was generously provided by Rosalie Crouch (Medical University of South Carolina).

## References

Anisimova M, Kosiol C. 2009. Investigating protein-coding sequence evolution with probabilistic codon substitution models. *Mol Biol Evol.* 26:255–271.

Asenjo AB, Rim J, Oprian DD. 1994. Molecular determinants of human red/green color discrimination. *Neuron* 12:1131–1138.

Badouin H, Belkhir K, Gregson E, Galindo J, Sundström L, Martin SJ, Butlin RK, Smadja CM. 2013. Transcriptome characterisation of the ant *Formica exsecta* with new insights into the evolution of desaturase genes in social Hymenoptera. *PLoS One* 8:e68200.

Baird RW. 2000. The killer whale: foraging specializations and group hunting. In: Mann J, Connor RC, Tyack PL, editors. Cetacean societies: field studies of dolphins and whales. Chicago (IL): University of Chicago Press. p. 127–153.

Barrett RDH, Hoekstra HE. 2011. Molecular spandrels: tests of adaptation at the genetic level. *Nat Rev Genet.* 12:767–780.

Bickelmann C, Morrow JM, Müller J, Chang BSW. 2012. Functional characterization of the rod visual pigment of the echidna (*Tachyglossus aculeatus*), a basal mammal. *Vis Neurosci.* 29:211–217.

Bielawski JP, Yang Z. 2004. A maximum likelihood method for detecting functional divergence at individual codon sites, with application to gene family evolution. *J Mol Evol.* 59:121–132.

Bischoff NN, Nickle BB, Cronin TWT, Velasquez SS, Fasick JIJ. 2012. Deep-sea and pelagic rod visual pigments identified in the mysticete whales. *Vis Neurosci.* 29:95–103.

Bowmaker JK. 2008. Evolution of vertebrate visual pigments. *Vision Res.* 48:2022–2041.

Bowmaker JK, Hunt DM. 2006. Evolution of vertebrate visual pigments. *Curr Biol.* 16:R484–R489.

Cassens I, Vicario S, Waddell VG, Balchowsky H, Van Belle D, Ding W, Fan C, Mohan RS, Simões-Lopes PC, Bastida R, et al. 2000. Independent adaptation to riverine habitats allowed survival of ancient cetacean lineages. *Proc Natl Acad Sci U S A.* 97:11343–11347.

Costa MPF, Novo EMLM, Telmer KH. 2013. Spatial and temporal variability of light attenuation in large rivers of the Amazon. *Hydrobiologia* 702:171–190.

Delport W, Poon AFY, Frost SDW, Pond SLK. 2010. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *J Gerontol.* 26:2455–2457.
Du J, Dungan SZ, Sabouhanian A, Chang BS. 2014. Selection on synonymous codons in mammalian rhodopsins: a possible role in optimizing translational processes. *BMC Evol Biol.* 14:96.
Elgoyhen AB, Franchini LF. 2011. Prestin and the cholinergic receptor of hair cells: positively-selected proteins in mammals. *Hear Res.* 273:100–108.
Fasick JI, Bischoff N, Brennan S, Velasquez S, Andrade G. 2011. Estimated absorbance spectra of the visual pigments of the North Atlantic right whale (*Eubalaena glacialis*). *Mar Mamm Sci.* 27:E321–E331.
Fasick JI, Cronin TW, Hunt DM. 1998. The visual pigments of the bottlenose dolphin (*Tursiops truncatus*). *Vis Neurosci.* 15:643–651.
Fasick JI, Robinson PR. 1998. Mechanism of spectral tuning in the dolphin visual pigments. *Biochemistry* 37:433–438.
Fasick JI, Robinson PR. 2000. Spectral-tuning mechanisms of marine mammal rhodopsins and correlations with foraging depth. *Vis Neurosci.* 17:781–788.
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.* 66:479–506.
Govardovskii VI, Fyhrquist N, Reuter T, Kuzmin DG, Donner K. 2000. In search of the visual pigment template. *Vis Neurosci.* 17:509–528.
Griebel U, Peichl L. 2003. Colour vision in aquatic mammals—facts and open questions. *Aquat Mamm.* 29.1:18–30.
Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol.* 59:307–321.
Hamilton H, Caballero S, Collins AG, Brownell RL. 2001. Evolution of river dolphins. *Proc R Soc Lond B Biol Sci.* 268:549–556.
Hauser FE, van Hazel I, Chang BSW. 2014. Spectral tuning in vertebrate short wavelength-sensitive 1 (SWS1) visual pigments: can wavelength sensitivity be inferred from sequence data? J Exp Zool B. 322:529–539.
Hope AJ, Partridge JC, Dulai KS, Hunt DM. 1997. Mechanisms of wavelength tuning in the rod opsins of deep-sea fishes. *Proc R Soc Lond B Biol Sci.* 264:155–163.
Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17(8):754–755.
Hunt DM, Carvalho LS, Cowing JA, Davies WL. 2009. Evolution and spectral tuning of visual pigments in birds and mammals. *Philos Trans R Soc Lond B Biol Sci.* 364:2941–2955.
Hunt DM, Dulai KS, Partridge JC, Cottrill P, Bowmaker JK. 2001. The molecular basis for spectral tuning of rod visual pigments in deep-sea fish. *J Exp Biol.* 204:3333–3344.
Hunt DM, Fitzgibbon J, Slobodyanyuk SJ, Bowmakers JK. 1996. Spectral tuning and molecular evolution of rod visual pigments in the species flock of cottoid fish in Lake Baikal. *Vision Res.* 36:1217–1224.
Jacobs GH. 2009. Evolution of colour vision in mammals. *Philos Trans R Soc Lond B Biol Sci.* 364:2957–2967.
Jacobs GH. 2013. Losses of functional opsin genes, short-wavelength cone photopigments, and color vision—a significant trend in the evolution of mammalian vision. *Vis Neurosci.* 30:39–53.
Jones G. 2010. Molecular evolution: gene convergence in echolocating mammals. *Curr Biol.* 20:R62–R64.
Kawamura S, Yokoyama S. 1998. Functional characterization of visual and nonvisual pigments of American chameleon (*Anolis carolinensis*). *Vision Res.* 38:37–44.
Kito Y, Suzuki T, Azuma M, Sekoguti Y. 1968. Absorption spectrum of rhodopsin denatured with acid. *Nature* 218:955–957.
Lamb TD, Pugh EN. 2004. Dark adaptation and the retinoid cycle of vision. *Prog Retin Eye Res.* 23:307–380.
Larmuseau MHD, Huyse T, Vancampenhout K, Van Houdt JKJ, Volckaert FAM. 2010. High molecular diversity in the rhodopsin gene in closely related goby fishes: A role for visual pigments in adaptive speciation? *Mol Phyl Evol.* 55(2):689–698.
Larmuseau MHD, Vanhove MPM, Huyse T, Volckaert FAM, Decorte R.. 2011. Signature of selection on the rhodopsin gene in the marine radiation of American seven-spined gobies (Gobidae, Gobiosomatini). *J Evol Biol.* 24(7):1618–1625.
Laskowski RA, Macarthur MW, Moss DS, Thorton JM. 1993. Procheck—a program to check the stereochemical quality of protein structures. *J Appl Crystallogr.* 26:283–291.
Li Y, Liu Z, Shi P, Zhang J. 2010. The hearing gene prestin unites echolocating bats and whales. *Curr Biol.* 20:R55–R56.
Lindblad-Toh K, Garber M, Zuk O, Lin MF, Parker BJ, Washietl S, Kheradpour P, Ernst J, Jordan G, Mauceli E, et al. 2011. A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* 478:476–482.
Liu Y, Cotton JA, Shen B, Han X, Rossiter SJ, Zhang S. 2010. Convergent sequence evolution between echolocating bats and dolphins. *Curr Biol.* 20:R53–R54.
Liu Y, Rossiter SJ, Han X, Cotton JA, Zhang S. 2010. Cetaceans on a molecular fast track to ultrasonic hearing. *Curr Biol.* 20:1834–1839.
Löytynoja A, Goldman N. 2010. webPRANK: a phylogeny-aware multiple sequence aligner with interactive alignment browser. *BMC Bioinformatics* 11:579.
Lythgoe JN, Dartnall HJ. 1970. A "deep sea rhodopsin" in a mammal. *Nature* 227:955–956.
Mass AM, Supin AY. 2007. Adaptive features of aquatic mammals' eye. *Anat Rec (Hoboken).* 290:701–715.
McFarland WN. 1971. Cetacean visual pigments. *Vision Res.* 11:1065–1076.
McGowen MR, Gatesy J, Wildman DE. 2014. Molecular evolution tracks macroevolutionary transitions in Cetacea. *Trends Ecol Evol.* 29:336–346.
McGowen MR. 2011. Toward the resolution of an explosive radiation—a multilocus phylogeny of oceanic dolphins (Delphinidae). *Mol Phylogenet Evol.* 60:345–357.
Meredith RW, Gatesy J, Emerling CA, York VM, Springer MS. 2013. Rod monochromacy and the coevolution of cetacean retinal opsins. *PLoS Genet.* 9:e1003432.
Miyagi R, Terai Y, Aibara M, Sugawara T, Imai H, Tachida H, Mzighani SI, Okitsu T, Wada A, Okada N. 2012. Correlation between nuptial colors and visual sensitivities tuned by opsins leads to species richness in sympatric Lake Victoria cichlid fishes. *Mol Biol Evol.* 29:3281–3296.
Morrow JM, Chang BSW. 2010. The p1D4-hrGFP II expression vector: a tool for expressing and purifying visual pigments and other G protein-coupled receptors. *Plasmid* 64:162–169.
Moura AE, van Rensburg CJ, Pilot M, Tehrani A, Best PB, Thornton M, Ploen S, de Bruyn PJN, Worley KC, Gibbs RA, et al. 2014. Killer whale nuclear genome and mtDNA reveal widespread population bottleneck during the last glacial maximum. *Mol Biol Evol.* 31:1121–1131.
Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Pond SLK, Scheffler K. 2013. FUBAR: a fast, unconstrained Bayesian approximation for inferring selection. *Mol Biol Evol.* 30:1196–1205.
Newman LA, Robinson PR. 2005. Cone visual pigments of aquatic mammals. *Vis Neurosci.* 22:873–879.
Nilsson D-E. 2013. Eye evolution and its functional basis. *Vis Neurosci.* 30:5–20.
Nylander JAA. 2004. MrModeltest v2. Uppsala (Sweden): Uppsala University.
Okada T, Fujiyoshi Y, Silow M, Navarro J, Landau EM, Shichida Y. 2002. Functional role of internal water molecules in rhodopsin revealed by x-ray crystallography. *Proc Natl Acad Sci U S A.* 99:5982–5987.
Okada T, Sugihara M, Bondar A-N, Elstner M, Entel P, Buss V. 2004. The retinal conformation and its environment in rhodopsin in light of a new 2.2Å crystal structure. *J Mol Biol.* 342:571–583.
Palczewski KK, Kumasaka TT, Hori TT, Behnke CAC, Motoshima HH, Fox BAB, Le Trong II, Teller DCD, Okada TT, Stenkamp RER, et al. 2000. Crystal structure of rhodopsin: a G protein-coupled receptor. *Science* 289:739–745.

Peichl L. 2005. Diversity of mammalian photoreceptor properties: adaptations to habitat and lifestyle? *Anat Rec* 287A:1001–1012.
Pond SLK, Frost SDW. 2005. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol Biol Evol.* 22:1208–1222.
Pond SLK, Frost SDW, Muse SV. 2005. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21:676–679.
Rennison DJ, Owens GL, Taylor JS. 2012. Opsin gene duplication and divergence in ray-finned fish. *Mol Phylogenet Evol.* 62:986–1008.
Sakmar TPT, Franke RRR, Khorana HGH. 1989. Glutamic acid-113 serves as the retinylidene Schiff base counterion in bovine rhodopsin. *Proc Natl Acad Sci U S A.* 86:8309–8313.
Sali A, Blundell TL. 1993. Comparative protein modeling by satisfaction of spatial restraints. *J Mol Biol.* 234:779–815.
Scheffler K, Martin DP, Seoighe C. 2006. Robust inference of positive selection from recombining coding sequences. *Bioinformatics* 22:2493–2499.
Schott RK, Refvik SP, Hauser FE, López-Fernández H, Chang BSW. 2014. Divergent positive selection in rhodopsin from lake and riverine cichlid fishes. *Mol Biol Evol.* 31:1149–1165.
Shen M-Y, Sali A. 2006. Statistical potential for assessment and prediction of protein structures. *Protein Sci.* 15:2507–2524.
Shen Y-Y, Liu J, Irwin DM, Zhang Y-P. 2010. Parallel and convergent evolution of the dim-light vision gene RH1 in bats (Order: Chiroptera). *PLoS One* 5:e8838.
Spady TC, Seehausen O, Loew ER, Jordan RC, Kocher TD, Carleton KL. 2005. Adaptive molecular evolution in the opsin genes of rapidly speciating cichlid species. *Mol Biol Evol.* 22:1412–1422.
Sugawara T, Imai H, Nikaido M, Imamoto Y, Okada N. 2010. Vertebrate rhodopsin adaptation to dim light via rapid meta-II intermediate formation. *Mol Biol Evol.* 27:506–519.
Sugawara T, Terai Y, Imai H, Turner GF, Koblmuller S, Sturmbauer C, Shichida Y, Okada N. 2005. Parallelism of amino acid changes at the RH1 affecting spectral sensitivity among deep-water cichlids from Lakes Tanganyika and Malawi. *Proc Natl Acad Sci U S A.* 102:5448–5453.
Sun Y-B, Zhou W-P, Liu H-Q, Irwin DM, Shen Y-Y, Zhang Y-P. 2013. Genome-wide scans for candidate genes involved in the aquatic adaptation of dolphins. *Genome Biol Evol.* 5:130–139.
Swanson WJ, Yang Z, Wolfner MF, Aquadro CF. 2001. Positive Darwinian selection drives the evolution of several female reproductive proteins in mammals. *Proc Natl Acad Sci U S A.* 98:2509–2514.
Swofford DL. 2002. Phylogenetic analysis using parsimony (and other methods) v4. Sunderland (MA): Sinauer Associates.
Terai Y, Seehausen O, Sasaki T, Takahashi K, Mizoiri S, Sugawara T, Sato T, Watanabe M, Konijnendijk N, Mrosso HDJ, et al. 2006. Divergent selection on opsins drives incipient speciation in Lake Victoria cichlids. *PLoS Biol.* 4:e433.
Tyack PL, Johnson M, Soto NA, Sturlese A, Madsen PT. 2006. Extreme diving of beaked whales. *J Exp Biol.* 209:4238–4253.
Uhen MD. 2010. The origin(s) of whales. *Annu Rev Earth Planet Sci.* 38:189–219.
Van Nynatten A, Bloom D, Chang BSW, Lovejoy NR. 2015. Out of the blue: adaptive visual pigment evolution accompanies Amazon invasion. *Biol Letters.* 11(7):20150349.
Veilleux CC, Louis EE, Bolnick DA. 2013. Nocturnal light environments influence color vision and signatures of selection on the OPN1SW opsin gene in nocturnal lemurs. *Mol Biol Evol.* 30:1420–1437.
Wald G. 1968. Molecular basis of visual excitation. *Science* 162:230–239.
Wang JY, Riehl KN, Dungan SZ. 2014. Family Delphinidae (ocean dolphins). In: Wilson DE, Mittermeier RA, editors. Handbook of the mammals of the world. Vol. 4: Sea mammals. Barcelona (Spain): Lynx Edicions. p. 410–527.
Warrant EJ, Johnsen S. 2013. Vision and the light environment. *Curr Biol.* 23:R990–R994.
Warrant EJ, Locket NA. 2004. Vision in the deep sea. *Biol Rev Camb Philos Soc.* 79:671–712.
Watwood SL, Miller PJO, Johnson M, Madsen PT, Tyack PL. 2006. Deep-diving foraging behaviour of sperm whales (*Physeter macrocephalus*). *J Anim Ecol.* 75:814–825.
Weadick CJ, Chang BSW. 2012. An improved likelihood ratio test for detecting site-specific functional divergence among clades of protein-coding genes. *Mol Biol Evol.* 29:1297–1300.
Wiederstein M, Sippl MJ. 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* 35:W407–W410.
Wozniak B, Dera J. 2007. Light absorption by suspended particulate matter (SPM) in sea water. In: Light absorption in sea water. New York: Springer. p. 167–294.
Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol.* 24:1586–1591.
Yang Z, Wong WSW, Nielsen R. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. *Mol Biol Evol.* 22:1107–1118.
Yim H-S, Cho YS, Guang X, Kang SG, Jeong J-Y, Cha S-S, Oh H-M, Lee J-H, Yang EC, Kwon KK, et al. 2014. Minke whale genome and aquatic adaptation in cetaceans. *Nat Genet.* 46:88–92.
Zhang J, Nielsen R, Yang Z. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol Biol Evol.* 22:2472–2479.
Zhao H, Ru B, Teeling EC, Faulkes CG, Zhang S, Rossiter SJ. 2009. Rhodopsin molecular evolution in mammals inhabiting low light environments. *PLoS One* 4:e8326.
Zhou X, Sun F, Xu S, Fan G, Zhu K, Liu X, Chen Y, Shi C, Yang Y, Huang Z, et al. 2013. Baiji genomes reveal low genetic variability and new insights into secondary aquatic adaptations. *Nat Commun.* 4:2708.