Rod monochromacy and the coevolution of cetacen retinal opsins

Christopher A. Emerling
Robert W. Meredith
John Gatesy
Vincent M. York
Mark S. Springer

Follow this and additional works at: https://poetcommons.whittier.edu/bio

Part of the Biology Commons
Rod Monochromacy and the Coevolution of Cetacean Retinal Opsins

Robert W. Meredith¹², John Gatesy¹, Christopher A. Emerling¹, Vincent M. York¹³, Mark S. Springer¹

1 Department of Biology, University of California Riverside, Riverside, California, United States of America, 2 Department of Biology and Molecular Biology, Montclair State University, Montclair, New Jersey, United States of America, 3 Keck School of Medicine, University of Southern California, Los Angeles, California, United States of America

Abstract

Cetaceans have a long history of commitment to a fully aquatic lifestyle that extends back to the Eocene. Extant species have evolved a spectacular array of adaptations in conjunction with their deployment into a diverse array of aquatic habitats. Sensory systems are among those that have experienced radical transformations in the evolutionary history of this clade. In the case of vision, previous studies have demonstrated important changes in the genes encoding rod opsin (RH1), short-wavelength sensitive opsin 1 (SWS1), and long-wavelength sensitive opsin (LWS) in selected cetaceans, but have not examined the full complement of opsin genes across the complete range of cetacean families. Here, we report protein-coding sequences for RH1 and both color opsin genes (SWS1, LWS) from representatives of all extant cetacean families. We examine competing hypotheses pertaining to the timing of blue shifts in RH1 relative to SWS1 inactivation in the early history of Cetacea, and we test the hypothesis that some cetaceans are rod monochromats. Molecular evolutionary analyses contradict the “coastal” hypothesis, wherein SWS1 was pseudogenized in the common ancestor of Cetacea, and instead suggest that RH1 was blue-shifted in the common ancestor of Cetacea before SWS1 was independently knocked out in baleen whales (Mysticeti) and in toothed whales (Odontoceti). Further, molecular evidence implies that LWS was inactivated convergently on at least five occasions in Cetacea: (1) Balaenidae (bowhead and right whales), (2) Balaenopteroidea (rorquals plus gray whale), (3) Mesoplodon bidens (Sowerby’s beaked whale), (4) Physeter macrocephalus (giant sperm whale), and (5) Kogia breviceps (pygmy sperm whale). All of these cetaceans are known to dive to depths of at least 100 m where the underwater light field is dim and dominated by blue light. The knockout of both SWS1 and LWS in multiple cetacean lineages renders these taxa rod monochromats, a condition previously unknown among mammalian species.

Introduction

Cetacea [dolphins, porpoises, and whales] represents a remarkable example of aquatic specialization within Mammalia [1]. With their return to river and marine environments, the ancestors of modern toothed cetaceans (Odontoceti) and baleen whales (Mysticeti) underwent extensive modifications that included the evolution of novel structures [e.g., baleen plates, tail flukes], major anatomical rearrangements [e.g., telescoping of the skull, development of fore-flippers], the loss or reduction of typical mammalian traits [e.g., olfactory structures, hair, hindlimbs], and associated behavioral changes [echolocation, filter-feeding, deep-diving] [2–4]. At the genetic level this restructuring includes evidence of positive selection in loci related to high-frequency audition [5–7], brain size [8,9], and flipper development [10], as well as degradation of genes related to olfaction [11–13], taste [14], tooth enamel formation [15–17], and vomeronasal chemoreception [18].

In the case of vision, aquatic environments impose challenging constraints, and the cetacean eye exhibits both morphological and molecular specializations that enhance underwater sight [19]. Possible morphological adaptations include an extensive reflective tapetum lucidum, a spherical lens with high refractive power, a relatively large cornea, and a rod-dominated retina, all of which enhance visual capabilities under dim light conditions [20,21]. At the molecular level, most mammals have dichromatic color vision based on presence of three visual pigments, each of which is a G protein-coupled receptor that consists of an opsin protein moiety linked via a Schiff base to a retinal chromophore [22]. The three opsins that characterize most mammals include a rod opsin (RH1) and two cone opsins, short wavelength-sensitive opsin (SWS1) and long wavelength-sensitive opsin (LWS). Rods mainly function in dim light conditions (scotopic/night vision) whereas cones require more light (photopic vision) and are necessary with color vision. By contrast with most other mammals, all cetaceans that have been investigated are thought to be L-cone monochromats that possess an inactivated copy of SWS1 and two functional opsins, RH1 and LWS, which are expressed in rod and L-cone cells of the retina, respectively [19,23,24].

Griebel and Peichl [19] and Peichl [24] suggested that retinal S-cones, which express SWS1 and are sensitive to blue wavelengths, were lost during an early, coastal period of cetacean evolution. Near-shore waters commonly have an underwater light spectrum that is red shifted owing to the absorption of blue light by organic and inorganic debris, and the loss of ‘jobless’ S-cones may have constituted an economical advantage in this environment by...
Opsin genes

McFarland [20] suggested that some cetaceans are probably rod opposed to L-cone monochromacy, in at least some cetaceans. In dim light conditions [22]. Vision is based entirely on rods and the rod opsin gene LWSDelphinidae and Phocoenidae [26]. More recently, Fasick et al. [21] sequenced the cone opsins of representative odontocete species belonging to the families mysticete, and suggested the presence of L-cones in immunocytochemical studies have failed to support this hypothesis and instead demonstrated the presence of L-cones in the common ancestor of Cetacea. Specifically, Bischoff et al. [25] speculated that the ancestral cetacean RH1 possessed 83Asn, 292Ser, and 299Ala at the Schiff’s base counterion site, but stopped short of using explicit methods to reconstruct the ancestral RH1 sequence of Cetacea. If RH1 was blue-shifted in the common ancestor of Cetacea prior to the independent inactivation of SWS1 on the stem mysticete and odontocete branches. LWS, in turn, was pseudogenized convergently in five different cetacean lineages [right whale plus bowhead, rorquals plus gray whale, Sowerby’s beaked whale, giant sperm whale, pygmy sperm whale], all of which are deep divers that feed on bioluminescent organisms. The tandem inactivation of SWS1 and LWS in these taxa presumably renders them rod monochromats, a condition that was previously unknown within Mammalia.

Author Summary

The emergence of Cetacea (whales, dolphins, porpoises) represents a profound transition in the history of life. Living cetaceans have evolved a spectacular array of adaptations in association with their return to aquatic habitats. Aquatic environments impose challenging constraints on sensory systems, including vision, and the cetacean eye exhibits both anatomical and molecular specializations that enhance underwater sight. Most mammals have one photopigment (RH1) for dim-light vision and two photopigments (long wavelength-sensitive opsin [LWS], short wavelength-sensitive opsin [SWS1]) for daytime, color vision. By contrast, cetaceans have an inactivated copy of the gene that encodes SWS1. Here, we show that LWS is also inactivated in several cetacean lineages including the giant sperm whale, Sowerby’s beaked whale, and balaenopteroids (rorquals plus gray whale). These cetaceans dive to depths of at least 100 meters where the underwater light field is dominated by dim, blue light. The knockout of both cone pigments renders these taxa rod monochromats, a condition that is previously unknown among mammalian species. Rod opsin remains functional in these taxa and is blue-shifted to increase its sensitivity to the available blue light that occurs in deep water conditions. These results further elucidate the molecular blueprint of modern cetacean species.

Results

Phylogenetic Analyses

Maximum likelihood trees based on SWS1 exons plus introns, SWS1 exons, RH1 exons, and LWS exons are shown in Figures S1, S2, S3, S4. With a few exceptions, clades with high bootstrap support percentages (≥90) on individual gene trees are in agreement with the species tree in Figure 1. All of the gene trees recovered Cetaceodonta [Cetacea + Hipposomatidae], Cetacea, Mysticeti, Balaenidae, Balaenopteroida, Physeteroidea, Ziphiidae, Iniidae + Pontoporiidae, Phyocoenidae, Delphinidae, Delphinocolidae, Iniidae + Pontoporiidae + Delphinocidae [Delphinidae]. Odontoceti was only recovered in the SWS1 analyses, but conflicting nodes in the RH1 and LWS trees had low bootstrap support values (≤53%).

SWS1 Evolution

Inactivating mutations (frameshift indels, premature stop codons, disrupted intron splice sites, amino acid replacement at the Schiff’s base counterion site) were apparent for all cetacean species in the SWS1 alignment (Table 1), but were lacking in SWS1 from the semiaquatic outgroup species, Hipposomatidae amphibiatus (Figure 1). Although the SWS1 genes of all cetacean species show evidence of mutational decay, no inactivating mutations map to the last common ancestral branch of Cetacea (Figure 1, node 26 to node 27). Instead, different molecular lesions define various sublineages of Cetacea (Table 1). An amino acid replacement (E113G; bovine RH1 numbering) at the Schiff’s base counterion site that is thought to disrupt opsin-chromophore binding [23] optimizes to the stem branch of Odontoceti (Figure 1; node 27 to node 35), and a four base-pair frameshift deletion was derived on the stem branch of Mysticeti (Figure 1; node 27 to node 28). These independent inactivating mutations imply that SWS1 was pseudogenized convergently in the two major subclades of Cetacea (Figure 1, Figure S5).
Estimates of $v$ (dN/dS) on different branches of the cetacean tree are consistent with parallel knockouts of $SWS1$ in Odontoceti and in Mysticeti. The $v$ estimate for $SWS1$ on the stem Cetacea branch (Figure 1, node 26 to node 27), just prior to the two inferred inactivating mutations, suggests a pattern of purifying selection based on analyses with two different codon frequency models ($v = 0.31, 0.35$). Likewise, a signature of strong purifying selection ($v = 0.16, 0.17$) was inferred on the stem Odontoceti branch (node 27 to node 35) (Table 2). Neutrality is predicted on the stem odontocete branch if $SWS1$ had previously been inactivated on the stem cetacean branch [19], but statistical tests rejected this hypothesis (Table 2). The $v$ estimate (0.75, 0.83) for the stem mysticete branch (node 27 to node 28), in turn, is only slightly lower than expected for complete neutrality ($v = 1.0$) and suggests that pseudogenization occurred very early on this branch. Estimates of $v$ for crown odontocete + crown mysticete branches are in agreement with expectations for neutrality (Table 2), and in conjunction with numerous frameshift indels within these clades (Table 1) imply a release from selective constraints after the occurrence of inactivating mutations on the stem odontocete and stem mysticete branches (Figure 1).

**RH1 Evolution**

No inactivating mutations (frameshift indels, splice site mutations) were apparent in the RH1 alignment, implying that RH1 is functional in all of the species that were surveyed (Figure 1).

Ancestral amino acid sequences at key tuning sites (83, 292, 299) in Cetacea [21,25] are shown in Table 3 for internal nodes with inferred blue or red shifts in $\lambda_{\text{max}}$. Amino acid changes from DAS to NSS on the stem cetacean branch [node 26 to node 27] resulted in an inferred blue shift from 501 to 484 nm. Additional blue shifts (484 to 479 nm) are inferred in $Caperea$, in stem Physeteroidea (node 35 to node 36), and in stem Ziphiidae (node 38 to node 39) based on an amino acid changes at site 299 (NSS to NSA) that occurred independently in these three lineages. Seven red shifts were reconstructed in Cetacea, including three in Mysticeti (stem Balaenidae [node 28 to node 29], $Megaptera$, $Eschrichtius$) and four in Odontoceti (stem Iniidae + Pontoporiidae [node 42 to node 43]).
Table 1. Summary of inactivating mutations in cetacean SWS1 and LWS genes.

<table>
<thead>
<tr>
<th>SWS1 Mutations</th>
<th>Number of Exon or Intron</th>
<th>Inactivating Mutation Including Alignment Number and Nucleotide Position(s)</th>
<th>Location on Tree (Ancestral Node Number: Descendant Node Number)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tursiops truncatus</td>
<td>Exon 1</td>
<td>1-bp frameshift deletion (Alignment1:34)</td>
<td>49:24</td>
</tr>
<tr>
<td>Mysticeti, Mesoplodon bidens, Platanista minor, Phocoenoides dalli (polymorphic), Delphinapterus leucas (polymorphic)</td>
<td>Exon 1</td>
<td>4-bp frameshift deletion (Alignment1:270–273)</td>
<td>27:28, 41:13, 37:12, 47:20, 45:22</td>
</tr>
<tr>
<td>Physeter macrocephalus</td>
<td>Exon 1</td>
<td>1-bp frameshift insertion (Alignment1:310)</td>
<td>36:10</td>
</tr>
<tr>
<td>Globicephala melas</td>
<td>Exon 1</td>
<td>50-bp frameshift deletion (Alignment1:261–311)</td>
<td>48:25</td>
</tr>
<tr>
<td>Pontoporia blainvillei</td>
<td>Exon 1</td>
<td>2-bp frameshift deletion (Alignment1:308–309)</td>
<td>43:17</td>
</tr>
<tr>
<td>Odontoceti</td>
<td>Exon 1</td>
<td>E113G amino acid replacement at the Schiff's base counterion site (bovine rhodopsin numbering) (Alignment1:338–340)</td>
<td>27:35</td>
</tr>
<tr>
<td>Berardius bairdii</td>
<td>Intron 1</td>
<td>GT to GC splice site mutation (Alignment1:363–364)</td>
<td>39:16</td>
</tr>
<tr>
<td>Eschrichtius robustus</td>
<td>Exon 2</td>
<td>8-bp frameshift deletion (Alignment1:808–815)</td>
<td>33:6</td>
</tr>
<tr>
<td>Globicephala melas</td>
<td>Exon 3</td>
<td>4-bp frameshift insertion (Alignment1:1438–1441)</td>
<td>48:25</td>
</tr>
<tr>
<td>Caperea marginata</td>
<td>Intron 3</td>
<td>GT to GC splice site mutation (Alignment1:1489–1490)</td>
<td>30:4</td>
</tr>
<tr>
<td>Platanista minor</td>
<td>Exon 4</td>
<td>1-bp frameshift insertion (Alignment1:2201)</td>
<td>37:12</td>
</tr>
<tr>
<td>Kogia breviceps</td>
<td>Exon 4</td>
<td>1-bp frameshift deletion (Alignment1:2336)</td>
<td>36:11</td>
</tr>
<tr>
<td>Pontoporia blainvillei</td>
<td>Exon 4</td>
<td>2-bp frameshift insertion (Alignment1:2337–2338)</td>
<td>43:17</td>
</tr>
<tr>
<td>Physeteroidea</td>
<td>Intron 4</td>
<td>GT to GA splice site mutation (Alignment1:2382–2383)</td>
<td>35:36</td>
</tr>
<tr>
<td>Tursiops truncatus</td>
<td>Exon 5</td>
<td>1-bp deletion (Alignment1:4079)</td>
<td>49:24</td>
</tr>
<tr>
<td>Balaenoptera acutorostrata</td>
<td>Exon 5</td>
<td>2-bp frameshift deletion (Alignment1:4103–4104)</td>
<td>31:9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LWS Mutations</th>
<th>Number of Exon or Intron</th>
<th>Inactivating Mutation Including Alignment Number and Nucleotide Position(s)</th>
<th>Location on Tree (Ancestral Node Number: Descendant Node Number)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balaenopteroidea</td>
<td>Exon 1</td>
<td>22 bp deletion that includes first two bp of initiation codon in exon 1 (Alignment1:99–120)</td>
<td>30:31</td>
</tr>
<tr>
<td>Kogia breviceps</td>
<td>Exon 1</td>
<td>44-bp frameshift deletion (Alignment4:120–163)</td>
<td>36:11</td>
</tr>
<tr>
<td>Balaenoptera musculus</td>
<td>Exon 1</td>
<td>1-bp frameshift deletion (Alignment4:169)</td>
<td>32:5</td>
</tr>
<tr>
<td>Balaenopteroidea</td>
<td>Exon 2</td>
<td>1-bp frameshift deletion (Alignment5:238)</td>
<td>30:31</td>
</tr>
<tr>
<td>Mesoplodon bidens</td>
<td>Exon 2</td>
<td>4-bp frameshift insertion (Alignment5:409–412)</td>
<td>41:13</td>
</tr>
<tr>
<td>Kogia breviceps</td>
<td>Intron 2</td>
<td>GT to CT splice site mutation (Alignment5:504–505)</td>
<td>36:11</td>
</tr>
<tr>
<td>Eubalaena australis</td>
<td>Intron 2</td>
<td>AG to CG splice site mutation (Alignment6:5–6)</td>
<td>29:2</td>
</tr>
<tr>
<td>Physeter macrocephalus</td>
<td>Exon 5</td>
<td>28-bp frameshift deletion (Alignment8:186–213)</td>
<td>36:10</td>
</tr>
<tr>
<td>Taxon</td>
<td>Number of Exon or Intron</td>
<td>Inactivating Mutation Including Alignment Number and Nucleotide Position(s)</td>
<td>Location on Tree (Ancestral Node Number: Descendant Node Number)</td>
</tr>
<tr>
<td>--------------------------</td>
<td>--------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------</td>
</tr>
<tr>
<td>Physeter macrocephalus</td>
<td>Exon 5</td>
<td>1-bp frameshift deletion (Alignment1:251)</td>
<td>36:10</td>
</tr>
<tr>
<td>Balaenidae</td>
<td>Intron 5</td>
<td>AG to GG splice site mutation (Alignment9:32–33)</td>
<td>28:29</td>
</tr>
<tr>
<td>Eubalaena australis</td>
<td>Exon 6</td>
<td>7-bp frameshift deletion (Alignment9:47–53)</td>
<td>29:2</td>
</tr>
</tbody>
</table>

Inactivating mutations include frameshifts, Schiff’s base counterion site mutations, and splice site mutations. All alignments are provided in Text S3. Inactivating mutations are also cross-referenced to branches in Figure 1. doi:10.1371/journal.pgen.1003432.t001

Among ten other amino acid sites that have been linked to spectral tuning in vertebrates [28], eight (sites 96, 102, 122, 183, 253, 261, 289, 317) are invariant among the cetaceans and the hippopotamid that were included in our taxon sampling, site 194 exhibits four amino acid replacements within Cetacea, and site 195 shows an amino acid replacement [L to P] on the stem Cetacea branch and four replacements within Cetacea.

Analyses with Codeml rejected site models 2 and 8, which add an extra category for positively selected sites, in favor of models 1 and 8a, respectively. By contrast, branch-site analyses with two different codon frequency models (CF) provided statistically significant support for a bin of five positively selected sites (7, 83, 123, 266, 292) on branches with $\lambda_{max}$ changes (CF2: P = 0.00036, $\omega$ = 5.53; CF3: P = 0.00018, $\omega$ = 6.43). Three of the five positively selected sites (83, 266, 292) have probabilities >0.95 of membership in this bin.

**LWS Evolution**

Inactivating mutations are apparent in *LWS* sequences from ten cetacean species (Figure 1, Figure S6, Table 1). Reconstructions of ancestral sequences imply eight frameshift indels and three splice site disruptions within Cetacea, with convergent inactivation of *LWS* on the following five branches (Figure 1): *Physeter macrocephalus, Kogia breviceps, Mesoplodon bidens*, stem Balaenidae (node 28 to node 29), and stem Balaenopteroidea (node 30 to node 31). All of these separate knockouts of *LWS* postdate prior inactivations of *SWSI* and therefore result in rod monochromacy (Figure 1).

Estimates of $\omega$ throughout the species tree generally are consistent with multiple, independent knockouts of *LWS* within Cetacea. Branches reconstructed as functional for *LWS* exhibit a strong signature of purifying selection ($\omega$ = 0.09, 0.10). By contrast, $\omega$ estimates on “transitional” branches [16], where inactivating mutations in *LWS* were reconstructed, generally show elevated rates of nonsynonymous substitution (Physeter: $\omega$ = 0.38, 0.41; Kogia: $\omega$ = 0.73, 0.78, stem balaenopteroid branch: $\omega$ = 0.34, 0.37). Exceptions are the short transitional branches for stem Balaenidae (1.5 to 1.7 inferred substitutions, $\omega$ = 0.0001) and *Mesoplodon* (2.4 to 3.1 inferred substitutions, $\omega$ = 0.13, 0.19). Branches within crown Balaenopteroida (node 31 and descendant branches) plus crown Balaenidae (node 29 and descendant branches), which are interpreted as pseudogenic based on the prior occurrence of inactivating mutations, have an $\omega$ value based on two codon models (0.69, 0.70) that does not deviate significantly from neutral expectations ($\omega$ = 1.00) based on $\chi^2$-tests.

Reconstructions of ancestral amino acid sequences at five key tuning sites (amino acids 180, 197, 277, 285, 308) [21,25] are shown in Table 3 for branches with inferred shifts in $\lambda_{max}$. For the five tuning sites, AHYTA ($\lambda_{max}$ = 552 nm) is the inferred ancestral condition for Cetancodontia (node 26) and for the last common ancestor of extant cetaceans (node 27). Three changes at *LWS* tuning sites were reconstructed within Cetacea. Parallel changes from AHYTA to AHYTS on the stem Mysticeti branch (node 27 to node 28) and on the stem Delphinoidea branch (node 42 to node 44) imply blue shifts from 552 nm to 522–531 nm. A change from AHYTA to AHYTP was reconstructed on the terminal *Inia* branch, but the functional effect of A308P is unknown (Table 3).

Analyses with Codeml rejected site models 2 and 8, which add an extra category for positively selected sites, in favor of models 1 and 8a, respectively. Similarly, positively selected sites were not identified in branch-site analyses.

**Discussion**

**Opsin Evolution in Cetacea**

Here, we assembled complete or nearly complete protein-coding sequences for *RH1, SWS1*, and *LWS* for representatives of all extant cetacean families. These sequences, in combination with molecular evolutionary analyses, permit a detailed, synthetic reconstruction of opsin evolution in Cetacea (Figure 1).

Recent phylogenetic hypotheses imply that the aquatic ancestry of Cetacea extends back to its last common ancestor with the semi-aquatic Hippopotamidae in the early Eocene, >50 Ma [4,29,30]. Whales and hippos share a variety of “aquatic” specializations including sparse hair, loss of sebaceous glands, and the ability to birth and nurse underwater [4,31,32], but these features traditionally have been interpreted as parallel evolutionary derivations in these two lineages. Given the hypothesis that the common ancestor of cetaceans and hippos was aquatic/semi-aquatic (Figure 1, node 50 to node 26), shared mutations in opsin genes that enhance vision in aquatic environments might be expected in whales and hippos. ML reconstructions of ancestral opsin sequences imply only two replacement substitutions (*LWS: E41D; RH1: L216M*) on the stem lineage. These replacements are not at key tuning sites and fail to provide compelling evidence for an aquatic shift in opsin properties in the common ancestor of hippos and whales.

Following divergence from Hippopotamidae, the unique evolutionary history of Cetacea began on the stem cetacean branch (Figure 1, node 26 to node 27). The fossil record indicates that the stem cetacean lineage was marked by a profound...
Table 2. Results of \( \chi^2 \)-tests for nonsynonymous and synonymous substitutions in SWS1.

<table>
<thead>
<tr>
<th>CF = 3 (774.4 non-synonymous sites: 239.6 synonymous sites)</th>
<th>Observed</th>
<th>Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Branches</td>
<td>N</td>
<td>S</td>
</tr>
<tr>
<td>1. Crown odontocete + crown mysticete branches</td>
<td>193.3</td>
<td>47.8</td>
</tr>
<tr>
<td>2. Odontocete stem branch</td>
<td>4.66</td>
<td>1.44</td>
</tr>
<tr>
<td>3. Mysticete stem branch</td>
<td>6.2</td>
<td>2.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CF = 2 (765.4 non-synonymous sites: 248.6 synonymous sites)</th>
<th>Observed</th>
<th>Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Branches</td>
<td>N</td>
<td>S</td>
</tr>
<tr>
<td>1. Crown odontocete + crown mysticete branches</td>
<td>195.2</td>
<td>48.9</td>
</tr>
<tr>
<td>2. Odontocete stem branch</td>
<td>4.60</td>
<td>1.50</td>
</tr>
<tr>
<td>3. Mysticete stem branch</td>
<td>6.79</td>
<td>2.21</td>
</tr>
</tbody>
</table>

\( \chi^2 \)-tests compared the observed versus expected numbers of nonsynonymous and synonymous substitutions for SWS1 based on a neutral model of evolution (\( \omega = 1 \)). All three different branch categories are expected to exhibit neutral \( \omega \) values according to the "coastal" hypothesis of SWS1 evolution. CF = codon frequency model; N = nonsynonymous; S = synonymous; \( \omega = dN/dS \); * = significant \( p \) values.

doi:10.1371/journal.pgen.1003432.t002
Table 3. Amino acids at key tuning sites for RH1 [21,25] and LWS [36].

<table>
<thead>
<tr>
<th>Branch Node Numbers</th>
<th>Branch Name</th>
<th>RH1</th>
<th>LWS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Amino Acid Changes at Key Tuning Sites</td>
<td>Amino Acid Changes at Key Tuning Sites</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(83 292 299)</td>
<td>(180 197 277 285 308)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\lambda_{max}$ (nm) Change on Branch</td>
<td>$\lambda_{max}$ (nm) Change on Branch</td>
</tr>
<tr>
<td>26 to 27</td>
<td>Stem Cetacea</td>
<td>DAS to NSS</td>
<td>AHYTA to AHYTS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>501 to 484</td>
<td>552 to 522–531</td>
</tr>
<tr>
<td>27 to 28</td>
<td>Stem Mysticeti</td>
<td>No change</td>
<td>No change</td>
</tr>
<tr>
<td>28 to 29</td>
<td>Stem Balaenidae</td>
<td>NSS to NAS</td>
<td>No change</td>
</tr>
<tr>
<td>30 to 4</td>
<td>Caperea</td>
<td>484 to 493</td>
<td>No change</td>
</tr>
<tr>
<td>33 to 6</td>
<td>Eschrichtius</td>
<td>484 to 493</td>
<td>No change</td>
</tr>
<tr>
<td>34 to 7</td>
<td>Megaptera</td>
<td>484 to 492</td>
<td>No change</td>
</tr>
<tr>
<td>35 to 36</td>
<td>Stem Physeteroidea</td>
<td>484 to 492</td>
<td>No change</td>
</tr>
<tr>
<td>38 to 39</td>
<td>Stem Ziphiidae</td>
<td>484 to 479</td>
<td>No change</td>
</tr>
<tr>
<td>42 to 43</td>
<td>Stem Iniidae + Pontoporiida</td>
<td>484 to 479</td>
<td>No change</td>
</tr>
<tr>
<td>43 to 17</td>
<td>Pontoporia</td>
<td>493 to 501</td>
<td>No change</td>
</tr>
<tr>
<td>43 to 18</td>
<td>Inia</td>
<td>NAS to NAT$^*$</td>
<td>AHYTA to AHYTP$^*$</td>
</tr>
<tr>
<td>44 to 45</td>
<td>Stem Monodontidae + Phocoenida</td>
<td>484 to 492</td>
<td>No change</td>
</tr>
<tr>
<td>45 to 22</td>
<td>Delphinapterus</td>
<td>492 to 501</td>
<td>No change</td>
</tr>
<tr>
<td>42 to 44</td>
<td>Stem Delphinidea</td>
<td>No change</td>
<td>AHYTA to AHYTS</td>
</tr>
</tbody>
</table>

Node numbers correspond to Figure 1. RH $\lambda_{max}$ values are based on Fasick et al. [21] for DAS and Bischoff et al. [25] for all other amino acid combinations. LWS $\lambda_{max}$ values are based on Fasick et al. [21]. $^*$=$\lambda_{max}$ value unknown. doi:10.1371/journal.pgen.1003432.t003

transition in anatomy from primitive semi-aquatic forms to obligately aquatic taxa with vestigial hindlimbs [3,33–35]. Ancestral reconstructions imply that stem cetaceans retained dichromatic color vision with functional SWS1, LWS, and RH1 as in *Hippopotamus* and more distantly related artiodactyls; a blue shift in RH1 also occurred on the stem cetacean branch (Figure 1). Specifically, the amino acid array at three key tuning sites (83, 292, 299) [21,25] changed from DAS to NSS, with an inferred $\lambda_{max}$ shift from 501 to 484 nm. Our ML reconstruction supports Bischoff et al.’s [25] hypothesis that RH1 was blue shifted on the stem cetacean branch. This change from a polar amino acid to a positively charged residue has been retained in the deep-diving physetoideas.

In addition to replacements at sites 83 and 292, a change at tuning site 195 (K to T) of RH1 occurred on the stem cetacean branch. This change from a polar amino acid to a positively charged residue has been retained in the deep-diving physetoideas (giant sperm whale, pygmy sperm whale). The inferred shift in $\lambda_{max}$ that results from a K to T replacement at this site, if any, remains to be investigated with mutagenesis studies. Unlike tuning sites 83, 292, and 299, that are situated in transmembrane regions of RH1 and are in close proximity to the chromophore, site 195 is positioned in the luminal face of RH1 [36]. The nature of long distance interactions between this amino acid site and the chromophore are unknown [36].

The basal split in Cetacea defines the separation of Odontoceti from Mysticeti, and also marks the evolution of profound changes in anatomy/feeding strategy in both clades [4,37]. Echolocation capabilities and degradation of olfactory structures were derived on the stem odontocete branch (Figure 1, node 27 to node 35), whereas the transition to bulk filter feeding with a keratinous baleen sieve evolved on the stem mysticete branch (Figure 1, node 27 to node 28). These divergent specializations represent changes in feeding style that would be expected to impact demands on visual systems.

Following the blue shift in RH1 on the stem cetacean branch, SWS1 was inactivated independently in stem odontocetes and in stem mysticetes, coincident with the evolution of divergent specializations in these two clades (Figure 1). Two lines of evidence support this reconstruction and argue against an earlier knockout of SWS1 in the common ancestor of Cetacea. First, comprehensive sequencing of SWS1 exons and introns revealed no shared inactivating mutations common to all extant cetaceans. Odontocetes have a common missense mutation at the Schiff’s base counterion site (E113G) that disrupts opsin-chromophore binding [23]. Mysticetes, in turn, share a 4-bp frameshift mutation in exon 1 of SWS1 that results in a premature stop codon. Frameshift indels in the same position occur in several odontocetes, but these deletions are most parsimoniously reconstructed as convergent between Mysticeti and multiple odontocete subclades (Text S1). Several mutations that disrupt intron boundaries were identified, but in all cases these substitutions map to branches within Odontoceti or within Mysticeti. Second, dN/dS values on the stem odontocete and stem mysticete branches should indicate an absence of selective constraints if SWS1 was inactivated earlier in the common ancestor of Cetacea. Estimates of dN/dS (0.75, 0.83) for the stem mysticete branch are consistent with neutral evolution (dN/dS = 1.00), but neutrality was rejected given the low dN/dS estimates (0.16, 0.17) for the stem odontocete branch, indicative of purifying selection and thus functionality after the split between Odontoceti and Mysticeti (Figure 1, node 27; Table 2).

Together, our reconstructions for the evolution of RH1 and SWS1 contradict the coastal knockout hypothesis [19,24]. This
scenario postulates that SWS1 was inactivated during an early amphibious phase of cetacean history when semi-aquatic whales occupied coastal waters that absorbed blue light, and that RH1 was subsequently blue shifted in crown cetaceans that moved to open ocean environments dominated by blue light. The coastal knockout hypothesis requires an as yet undiscovered inactivating mutation in SWS1, perhaps in the promoter region of this gene or at splice sites [23]. Instead, our results fit the hypothesis that RH1 was blue shifted in the common ancestor of Cetacea, and that SWS1 was convergently knocked out in Odontoceti and in Mysticeti after cetaceans had invaded open ocean habitats. This is perhaps surprising given that SWS1 is well suited to detect the blue light that dominates the open ocean. However, the relative scarcity of S cones in the mammalian retina, which diminishes the efficiency of photon capture under dim light conditions, may have predisposed S-cones to eventual loss through relaxed selection [22]. By contrast, rods are much more efficient at photon capture under dim light conditions because of their higher density in the mammalian retina and their integration with large, sparsely distributed ganglion cells that sum photon detection over huge receptive fields [30]. The preferential retention of a functional copy of LWS rather than SWS1 in some cetaceans may reflect the higher density of L-cones and their greater impact on visual acuity [22].

In addition to the blue shift of RH1 in the ancestral cetacean branch, several amino acid replacements in RH1 imply further adjustments in \( \lambda_{max} \) (Figure 1). These shifts are generally consistent with the photic environments that are occupied by different cetacean species [20,25]. Among these changes are blue shifts in deep-diving physeteroids (sperm whales) and in zibhiids (beaked whales), a red shift in the common ancestor of Inia and Pontoporia, both of which are found in shallow water environments, and red shifts in several mysticetes [e.g., Eschrichtius, Megaptera]. Bischoff et al. [25] suggested that the red-shifted pigment that occur in some mysticetes are better adapted to relatively shallower foraging environments than the ancestral mysticete pigment. The blue shifts in Physeteroidea and Ziphiidae occurred in parallel and in both cases involve amino acid replacements [serine to alanine] at alanine site 299 [21,25]. Sperm whales and beaked whales rank among the deepest diving mammals and specialize on a cephalopod-rich diet [39]. Several phylogenetic studies of anatomical evidence grouped these suction feeding species, presumably based on convergent character states related to their deep diving habits [40,41], but most recent work indicates that zibhiids are more closely related to dolphins and porpoises than to physeteroids [37,42,43].

Nozawa et al. [44] suggested that Yang's [45] codeml program is not useful for identifying adaptive sites in visual pigments. Our results support Nozawa et al.'s [44] finding that site analyses fail to identify adaptive changes in visual pigments. However, branch-site tests identified five codons in RH1 that have evolved under positive selection on branches with inferred changes in \( \lambda_{max} \). The 95% confidence intervals for these sites are well above the lowest 2.5% (5.53–6.43), and supports the hypothesis that changes affecting \( \lambda_{max} \) in cetacean RH1 proteins are adaptive. The failure of site analyses to detect positively selected sites in RH1 may be a consequence of mixing positive selection on foreground branches with purifying selection on background branches. Nozawa et al. [44] criticized branch-site tests [45–47] for their proclivity to generate false positive results based on simulations, but Yang et al. [48] correctly noted that false positives only occurred in 32/14,000 cases, which is much lower than the nominal significance level (5%) and demonstrates that the branch-site test is conservative. Among the positively selected sites, two (83, 292) are known tuning sites that in part are the basis for inferring changes in \( \lambda_{max} \) (Figure 1). Changes at site 83 may also be important in dim-light conditions because the amino acid at this position affects the rate at which photoreceptor cells generate electrical signals [49]. The other three sites (7, 123, 266) have not been predicted to affect \( \lambda_{max} \). Site 7 occurs in the extracellular domain, site 123 occurs in transmembrane helix III, and site 266 occurs in transmembrane helix 6 [50]. The functional consequences of mutations at these amino acid positions in cetacean RH1 sequences remain unknown, although conformational changes associated with transmembrane domains III and VI of G protein-coupled receptors may be important in receptor activation [51].

Changes in LWS spectral sensitivity coincide with deployment of cetaceans to diverse aquatic habitats (Figure 1). A blue shift in LWS in stem mysticetes co-occurs with an SWS1 frameshift mutation on the same branch, although the sequence of these events is unclear. An additional LWS blue shift in \( \lambda_{max} \) maps to the common ancestor of Delphinoidea [dolphins, porpoises, beluga], but the most striking feature of LWS evolution in Cetacea is the convergent knockout of this gene in five different lineages: Balaenopteroidea (rorquals and grey whale), Balaenidae (bowhead and right whale), Mesoplodontidae (Sowerby’s beaked whale), Physeter macrorynchus (giant sperm whale), and Kogia breviceps (pygmy sperm whale) (Figure 1). Given that SWS1 is also debilitated in each of these species (Figure 1), the genetic data imply that these taxa are rod monochromats. This iterated degeneration of cetacean LWS was not apparent in earlier studies because complete protein-coding LWS sequences had been generated for only a few cetacean species [21].

**Rod Monochromacy in Cetacea**

Historically, the pure rod retina has been proposed as the “extreme” adaptation to low light levels [52]. Walls [52] and McFarland [20] suggested the possibility of rod monochromacy in at least some cetaceans. More generally, early studies on retinal anatomy hinted at this condition in a variety of nocturnal and aquatic mammalian species with rod dominated retinas, including night monkeys, lemurs, tarsiers, chinchillas, seals, and bats [52–57]. Recent work has shown that representative cetaceans are instead L-cone monochromats and retain a functional copy of LWS [21,26]. Similarly, primates, rodents, pinnipeds, and bats that were previously hypothesized to be rod monochromats are now known to be L-cone monochromats with functional LWS or even cone dichromats with functional LWS and SWS1 [22,58–61]. The present survey of cetacean opsins, which documents pseudogenization of both SWS1 and LWS in multiple cetacean lineages, vindicates McFarland’s [20] hypothesis that some cetaceans are rod monochromats (Figure 1). To our knowledge these are the only known examples of rod monochromacy in Mammalia or even Amniota. The observation that five independent derivations of mammalian rod monochromacy are all clustered within Cetacea is striking, and suggests that one or more features of cetacean biology have been pivotal in driving the degenerative pattern of opsin evolution in this aquatic clade.

The naked mole rat (*Heterocephalus glaber*) is the only other mammal, aside from the cetacean species characterized here, that is known to lack a functional copy of LWS, but *H. glaber* retains an intact SWS1 and is interpreted as a cone monochromat [62]. This condition contrasts with other mammalian cone monochromats [pinnipeds, dolphins, porpoises, some procyonids, some rodents, some bats] that combine a pseudogenic SWS1 with a functional copy of LWS [22,23,28,60,63–71]. It has been suggested that cetacean cone monochromats [e.g. bottlenose dolphin, *Tursiops truncatus*] can distinguish colors, possibly via interactions between LWS and RH1 [30,72], but any vestiges of color vision
presumably have been lost in the various rod monochromatic cetacean species documented here (Figure 1).

Among other vertebrates, rod monochromatic taxa are rare and to our knowledge have only been documented in bony and cartilaginous fishes [73–81], caecilians [22,82], and the cave salamander Proteus anguinus [83], although presumed rod monochromacy based entirely on immunocytochemistry, microscopy or spectral analysis does not preclude the possibility that other minor visual pigment classes exist [77,81,83]. Most of the rod monochromatic fish species inhabit the deep sea or are nocturnal; caecilians are generally fossorial and/or nocturnal with poorly developed eyes; and the cave salamander Proteus lives in a virtually light-free environment.

The phylogenetic evidence for multiple, independent knock-outs of both SWS1 and LWS within Cetacea raises the question of why convergent pseudogenization and rod monochromacy evolved in this clade but not in other mammalian groups. All rod monochromatic cetacean species that were genetically characterized in our survey are capable of diving to depths that exceed 100 m, with sperm and beaked whales ranking among the deepest diving mammals [39,84–88]. The selective pressures on mammalian retinal opsins in deep-water habitats are drastically different from those on land. In the open marine environment, the electromagnetic radiation of visible light is weakened with depth due to absorption and scattering [89]. In the mesopelagic zone (150–1000 m), down-welling sunlight becomes more monochromatic and the spectrum shifts towards shorter, bluer wavelengths with depth [81,82]. Below 1000 m becomes more monochromatic and the spectrum shifts towards the mesopelagic zone (150–1000 m), down-welling sunlight environment, the electromagnetic radiation of visible light.

In the background. Along these lines, the parallel pseudogenization of RH1 and RH1 for additional cetartiodactyl families [Bovidae, Cervidae, Suidae, Camelidae] were assembled from Ensembl, Pre-Ensembl, and NCBI based on availability with minor augmentation by new sequences (Table S1, Text S2).

### Alignments and Phylogenetic Analyses

Sequences were aligned manually using Se-Al [95]. The virtual mRNA alignment lengths were 1014 base pairs (bp) for SWS1, 1092 bp for LWS, and 1044 bp for RH1. The complete alignment for SWS1, including exons and introns, was 4163 bp. All alignments for phylogenetic and PAML analyses, along with alignments for non-overlapping PCR amplicons (exons plus partial introns for LWS and RH1), are provided in Text S3 in nexus format. Phylogenetic analyses were performed with RAxML 7.2.7 [96] and the GTR + \( \Gamma \) model of sequence evolution. Additional details are provided in Text S2.

### Inactivating Mutations

Opsin alignments were manually inspected for putative inactivating mutations, including substitutions that result in stop codons, changes at intron splice donor/acceptor sites, and frameshift indels. We also examined SWS1 sequences for a missense mutation at Schiff’s counterion site (E113G; bovine RH1 numbering) that disrupts opsin-chromophore binding [23].

### Ancestral Sequence Reconstructions and Character State Mapping

Ancestral DNA sequences for SWS1, LWS, and RH1 were reconstructed with the BaseML program implemented in PAML 4.4b [45]. We used the REV model and a composite species tree based on McGowen et al. [42] for cetaceans and Gatesy [97] for all other cetartiodactyls. Frameshift mutations and other indels

### Conclusions

The emergence of Cetacea represents a profound macroevolutionary transition that entails comprehensive remodeling at both the genetic and morphological levels [4]. Our results elucidate key events in the evolutionary history of cetacean opsins, including an initial blue shift of RH1 in stem Cetacea, parallel knockouts of SWS1 in Odontoceti and Mysticeti, and five independent inactivations of LWS in deep-diving cetacean lineages. As correctly surmised by McFarland [20], some cetaceans are rod monochromats and have evolved eyes that are highly specialized for dim-light vision.

### Materials and Methods

#### Taxon and Gene Sampling

Previously published RH1, SWS1, and LWS sequences for Cetacea were combined with new sequences that were generated through PCR and dyeoxy sequencing. We targeted complete coding regions of all three opsin genes for representatives of Hipposotomatidae and all extant cetacean families (Text S2). RH1, LWS, and SWS1 sequences for additional cetartiodactyl families [Bovidae, Cervidae, Suidae, Camelidae] were assembled from Ensembl, Pre-Ensembl, and NCBI based on availability with minor augmentation by new sequences (Table S1, Text S2).

#### PCR and Sequencing

Aligned sequences for Bos taurus, Sus scrofa, Tursiops truncatus, Vicugna pacos, and available GenBank sequences (Table S1) were used to design PCR primers for SWS1, RH1, and LWS. SWS1 (exons 1–4; partial exon 5; introns 1–4) was amplified in five overlapping segments. PCR primers for RH1 (exons 1–5) and LWS (exons 1–6) were positioned in the flanking intronic regions of each exon (see Text S2 for additional details on PCR reactions). Accession numbers for new cetartiodactyl sequences are KC676796–KC677023 (Table S1). Primer sequences are provided in Table S2.
were optimized with Fitch parsimony, as implemented in Mesquite [90].

\( \lambda_{\text{max}} \) Estimation

Spectral tuning in RH1 is influenced by at least 13 amino acid sites [28], although replacements at only three of these sites (83, 292, 299) fully explain the absorbance difference between cow RH1 (Bos taurus, \( \lambda_{\text{max}} = 500 \text{ nm} \)) and bottlenose dolphin RH1 (Tursiops truncatus, \( \lambda_{\text{max}} = 488 \text{ nm} \)) [99]. These replacements are D83N, A292S, and A299S. Different combinations of ancestral and derived amino acids at these three sites have been tested in mutagenic studies of Bos RH1 to explain the various \( \lambda_{\text{max}} \) values that occur in other cetaceans [21,25]. For LWS, Yokoyama [36] suggested a “five-sites” rule whereby \( \lambda_{\text{max}} \) values between 510 and 560 in vertebrates can be fully explained by amino acid changes S180A, H197Y, Y277F, T285A, A308S and their interactions. Here, we follow Fasick et al. [21] and Bischoff et al. [25] and provide \( \lambda_{\text{max}} \) estimates for newly determined RH1 and LWS sequences based on directly determined \( \lambda_{\text{max}} \) values from expressed RH1 and LWS pigments that possess identical amino acids at the same key sites for each of these opsins. It will be important in future studies to perform direct measurements of \( \lambda_{\text{max}} \) on reconstructions of ancestral RH1 sequences. Even without these experiments, empirical measurements on a diverse array of opsins from cetacean species and Bos taurus (wild type and mutagenesis variants) provide a strong foundation for inferring \( \lambda_{\text{max}} \) values in ancestral cetacean sequences [21,25].

d\(n/d\)s Analyses

The Codeml program in PAML 4.4b [45] was used to estimate the ratio (\( \theta \)) of the non-synonymous substitution rate (dN) to the synonymous substitution rate (dS) at individual sites (RH1, LWS) and on branches (SWS1, LWS). We also performed branch-site analyses [45–47,100] on RH1 and LWS sequences. In both cases, branches with predicted changes in \( \lambda_{\text{max}} \) (Figure 1) of the relevant opsin were assigned to the foreground, and all other branches were assigned to the background. We used a composite species tree for all cetartiodactyl taxa as detailed above. Statistical tests of neutrality [complete absence of functional constraints] for branches and sets of branches were executed as in Meredith et al. [16]. See Text S2 for details.

Supporting Information

Figure S1 Maximum likelihood phylogram based on SWS1 exons and introns. (PDF)

Figure S2 Maximum likelihood phylogram based on SWS1 exons. (PDF)

Figure S3 Maximum likelihood phylogram based on RH1 exons. (PDF)

Figure S4 Maximum likelihood phylogram based on LWS exons. (PDF)

Figure S5 Parsimony reconstruction of the 4-bp frameshift deletion in SWS1. Branch colors are as follows: gray, odontocetes; black, mysticetes, blue, stem Cetacea; green, non-cetacean. Plio = Pliocene; P = Pleistocene. Paintings are by Carl Buell. Also see Text S1. (PDF)

Figure S6 Chromatograms that illustrate inactivating mutations found in cetacean LWS sequences. Taxa exhibiting the deleterious mutations for the indicated exon are in red font. Deletions are highlighted in red and insertions are highlighted in green. (PDF)


Table S2 Primers used in this study (all 5’ to 3’). SWS1 primer pairs were designed to amplify four complete exons (1–4), four complete introns (1–4), and part of exon 5 in five overlapping segments. RH1 and LWS primers were designed to amplify each exon and part of the flanking 5’ and 3’ introns. Ex = exon. (PDF)

Text S1 Additional details on the 4-bp frameshift deletion in SWS1. (PDF)

Text S2 Additional Materials and Methods. (PDF)

Text S3 Fourteen alignments in nexus format. Alignment 1 = SWS1 Introns + Exons (RAxML); Alignment 2 = SWS1 Exons (PAML, RAxML); Alignment 3 = LWS Exons (PAML, RAxML); Alignment 4 = LWS Exon 1 ampiclon; Alignment 5 = LWS Exon 2 ampiclon; Alignment 6 = LWS Exon 3 ampiclon; Alignment 7 = LWS Exon 4 ampiclon; Alignment 8 = LWS Exon 5 ampiclon; Alignment 9 = LWS Exon 6 ampiclon; Alignment 10 = RH1 Exons (PAML/RAxML); Alignment 11 = RH1 Exon 1 ampiclon; Alignment 12 = RH1 Exon 2 ampiclon; Alignment 13 = RH1 Exons 3+4 ampiclon; Alignment 14 = RH1 Exon 5 ampiclon. (TXT)

Acknowledgments

We thank Southwest Fisheries Science Center, South Australian Museum, North Slope Borough (Barrow, Alaska), The Marine Mammal Center (Sausalito), Smithsonian Institution, New York Zoological Society, World Wildlife Fund, Greenland Institute of Natural Resources, Alaska Department of Fish and Game, P. Morin, K. Robertson, S. Clive, A. Diaz, M. Miškić, G. Amato, G. Schaller, M. Cronin, W. Murphy, M. P. Heide-Jörgensen, U. Arnason, H. Rosenbaum, and G. Braulik for providing DNA samples. Two anonymous reviewers provided helpful comments on an earlier version of this manuscript.

Author Contributions

Conceived and designed the experiments: MSS RWM JG. Performed the experiments: RWM CAE VMY MSS. Analyzed the data: RWM JG MSS. Contributed reagents/materials/analysis tools: JG MSS RWM. Wrote the paper: MSS JG RWM. Provided comments on manuscript: CAE VMY.
References


