J Mol Evol (2005) 61:75–89
DOI: 10.1007/s00239-004-0289-z

JOURNAL OF MOLECULAR EVOLUTION



# Rod and Cone Opsin Families Differ in Spectral Tuning Domains but Not Signal Transducing Domains as Judged by Saturated Evolutionary Trace Analysis

**Karen L. Carleton,[1,2] Tyrone C. Spady,[1,2] Rick H. Cote[3]**

[1] Hubbard Center for Genome Studies, University of New Hampshire, Durham, NH 03824, USA
[2] Department of Zoology, University of New Hampshire, Durham, NH 03824, USA
[3] Department of Biochemistry and Molecular Biology, University of New Hampshire, Durham, NH 03824, USA

Received: 24 September 2004 / Accepted: 15 December 2004 [*Reviewing Editor:* Dr. Rafael Zardoya]

**Abstract.** The visual receptor of rods and cones is a covalent complex of the apoprotein, opsin, and the light-sensitive chromophore, 11-*cis*-retinal. This pigment must fulfill many functions including photoactivation, spectral tuning, signal transmission, inactivation, and chromophore regeneration. Rod and cone photoreceptors employ distinct families of opsins. Although it is well known that these opsin families provide unique ranges in spectral sensitivity, it is unclear whether the families have additional functional differences. In this study, we use evolutionary trace (ET) analysis of 188 vertebrate opsin sequences to identify functionally important sites in each opsin family. We demonstrate the following results. (1) The available vertebrate opsin sequences produce a definitive description of all five vertebrate opsin families. This is the first demonstration of sequence saturation prior to ET analysis, which we term saturated ET (SET). (2) The cone opsin classes have class-specific sites compared to the rod opsin class. These sites reside in the transmembrane region and tune the spectral sensitivity of each opsin class to its characteristic wavelength range. (3) The cytoplasmic loops, primarily responsible for signal transmission and inactivation, are essentially invariant in rod versus cone opsins. This indicates that the electrophysiological differences between rod and cone photoreceptors cannot be ascribed to differences in the protein interaction regions of the opsins. SET shows that chromophore binding and regeneration are the only aspects of opsin structure likely to have functionally significant differences between rods and cones, whereas excitatory and adaptational properties of the opsin families appear to be functionally invariant.

**Key words:** Opsin — Phototransduction — Evolutionary trace — Saturated ET — Spectral tuning — Activation

*Correspondence to:* Karen L. Carleton, Hubbard Center for Genome Studies, 438 Gregg Hall, 35 Colovos Road, University of New Hampshire, Durham, NH 03824, USA; *email:* karen.carleton@unh.edu

## Introduction

The majority of vertebrates have a duplex retina, containing both rod and cone photoreceptors. Photoreceptors convert incident light into a neural signal through a complex phototransduction pathway. Rods provide high sensitivity and operate down to the single-photon detection limit. Cones operate over a wide range of illumination without saturation and have faster rates of response and adaptation (Miller et al. 1994; Yau 1994; Pugh and Lamb 2000).

The classification of photoreceptors as either rods or cones is generally based on several criteria including cell morphology, neural wiring, and electrophysiology (Ebrey and Koutalos 2001; Burns and Lamb 2003). In addition, rod and cone photorecep-

tors each contain separate but parallel pathways made up of unique proteins specific to rod and cone phototransduction (Ebrey and Koutalos 2001; Hisatomi and Tokunaga 2002). Differences in the enzymatic rates or binding efficiencies of the proteins in each pathway are one likely explanation for differences between rod and cone electrophysiology.

The G protein coupled receptor, opsin, is the first protein in the phototransduction pathway (Sakmar et al. 2002; Filipek et al. 2003). The opsin apoprotein is bound to its ligand, 11-*cis*-retinal, to make a functional visual pigment. The seven-transmembrane (TM) alpha helices of the opsin protein surround retinal as shown in the crystal structure of bovine rhodopsin (Palczewski et al. 2000; Teller et al. 2001; Okada et al. 2002). 11-*cis*-Retinal is the light-sensitive component of the visual pigment. Interactions between neighboring amino acids of the opsin and the chromophore shift the absorption to different wavelengths, tuning the peak sensitivity of visual pigments (Kochendoerfer et al. 1999; Yokoyama 2000; Sakmar et al. 2002).

Early in the evolution of vertebrates, distinct classes of opsins arose through various gene duplication events (Okano et al. 1992; Yokoyama 2000). These include a rod opsin (RH1) and four cone opsin classes: very short wavelength sensitive (SWS1), short wavelength sensitive (SWS2), rhodopsin-like (RH2), and medium/long wavelength sensitive (M/LWS). The amino acid sequence of each opsin class has evolved to produce visual pigments with peak absorptions ($\lambda_{max}$) in different parts of the spectrum, although there is some spectral overlap between opsin classes. The spectral ranges of the opsin classes, when combined with 11-*cis*-retinal, are as follows: RH1, 470–510 nm; SWS1, 358 to 425 nm; SWS2, 420 to 474 nm; RH2 opsins, 466 to 511 nm; and M/LWS, 521 to 575 nm (Ebrey and Koutalos 2001; Hisatomi and Tokunaga 2002).

The phototransduction pathway in vertebrate rod photoreceptors is well understood and represents a model for G-protein coupled signaling pathways (Pugh and Lamb 2000; Ebrey and Koutalos 2001). The absorption of light by the visual pigment results in the isomerization of 11-*cis*-retinal to all-*trans*-retinal (Filipek et al. 2003). This causes a conformational change in the opsin protein which enables it to activate the G protein, transducin (Ebrey and Koutalos 2001). The activated opsin is subsequently inactivated when phosphorylated by G-protein receptor kinase (GRK) and then capped with arrestin (Pugh et al. 1999; Maeda et al. 2003). The cytoplasmic loops of opsin are critical for these protein interactions. Transducin is activated by interactions with loops C-2, C-3, and the eighth helix (Konig et al. 1989; Shi et al. 1995). GRK and arrestin also interact with the three cytoplasmic loops as well as the C-terminal tail (Palczewski et al. 1989; Shi et al. 1995; Thurmond et al. 1997).

The molecular mechanisms responsible for the physiological differences between rod and cone photoreceptors have not been determined. There are a number of possible explanations including differences in cell morphology/cell volume, differences in the relative concentration of phototransduction proteins (differential gene expression), and differences in the function of the rod and cone-specific proteins in the pathway. The latter hypothesis is attractive in light of the fact that there are rod and cone-specific homologs for almost all pathway proteins which could differ in photoactivation or deactivation properties. Experimental support for functionally unique rod vs. cone proteins is inconclusive at present. Reconstitution experiments showed similar activation of purified rod transducins by cone SWS1 and RH1 ospins (Starace and Knox 1997). Expression of nonnative opsins in rod or cone cells of transgenic animals resulted in photoresponses with identical signal amplification and kinetics (Kefalov et al. 2003). In contrast, purified rod and cone cell membrane preparations from fish retina displayed marked differences in their efficiency in transducin activation and opsin phosphorylation (Tachibanaki et al. 2001), suggesting that rod and cone opsins may have inherently different photoactivation and deactivation properties. In this previous work, functional differences among each of the five opsin families have been largely unexplored. Hence, there is a need to understand to what extent structural differences among the five opsin families can account for differences in each class of retinal rod and cone photoreceptors.

Although site-directed mutagenesis has been an important tool to probe structure–function relationships, nature has also performed mutant screens of protein function through millions of years of evolution. Sites which are free to vary within a protein evolve at a relatively constant rate, according to a molecular clock (Jukes and Cantor 1969; Li 1997). By comparing homologous sequences, sites which are invariant and therefore essential to protein function can be identified. To detect functionally significant sites, the compared sequences should be sufficiently divergent to allow nonessential sites the necessary time to accumulate substitutions.

A number of methods based on natural mutant screens have been developed to identify sites which are critical to protein function (Casari et al. 1995; del Sol Mesa et al. 2003; Oliveira et al. 2003; Kalinina et al. 2004). Because of the well-defined phylogenetic relationships of opsin classes, their high degree of conservation, and the availability of an x-ray crystal structure, we chose to utilize the tree-based method known as evolutionary trace (ET) analysis (Lichtarge et al. 1996; Lichtarge and Sowa 2002; Lichtarge et al. 2002). For ET analysis, sequences are first phylogenetically grouped into classes which differ by protein

function. A consensus sequence is then determined for each class consisting of all amino acids which are unanimous within the class. Comparisons of these unanimity sequences identifies invariant sites (identical in all classes) and class-specific sites (unique to a class). Finally, these sites are mapped onto a three-dimensional crystal structure. This often reveals clusters of sites which are important for protein–protein interactions and protein function.

In this study, we use the most restricted form of ET analysis where sites must be unanimous within a class to be considered. While other less restrictive methods have been developed (Landgraf et al. 1999), the high conservation observed in opsin sequences motivated our examination of the rod and cone opsins using this method. Indeed, ET analysis has proved insightful for several studies of phototransduction proteins (Sowa et al. 2000; Imanishi et al. 2002; Lichtarge et al. 2002). In studies of G-protein $\alpha$ subunits, for example, comparisons to a large-scale mutagenesis study showed that over 75% of the sites identified by ET were found to have important functional roles in the enzymatic activity of Gt$\alpha$ (Lichtarge et al. 1996, 2002; Onrust et al 1997). Dean et al. (2001) used ET to examine hundreds of G-protein coupled receptors and found evidence for sites important in receptor dimerization. Madabushi et al. (2004) examined the transmembrane regions of several hundred G-protein coupled receptors and identified a network of contacts from retinal to the cytoplasmic loops which couple to the G-protein. In both these studies, rod and cone opsins were grouped together, thereby obscuring any differences between these groups. A recent report using a smaller set of opsin sequences provided insights into residues responsible for signal transmission within the opsin protein during photoactivation of the visual receptor (Teller et al. 2003). However, these authors did not consider large scale tuning effects nor differences in rod and cone phototransduction.

In this work, we use ET analysis to compare 188 vertebrate rod and cone opsins to test for rod–cone differences in the first step of the phototransduction pathway. Invertebrate opsins were excluded because they couple to a different signaling pathway (Pepe 2001). Since focusing on vertebrate opsins limits the number of sequences, we first prove that the available sequences are phylogenetically diverse enough to have reached saturation, with a clearly defined set of conserved sites. We then perform saturated ET (SET) analysis for the rod and the four cone opsin families to identify sites which are important for distinguishing rod and cone opsin function. The results are substantiated by comparison to mutagenesis studies as well as sites identified from disease states.

We also compared SET analyses with statistical methods based on maximum likelihood (ML). The ML methods have successfully identified sites contributing to the molecular evolution of lysozyme (Yang and Nielsen 1998, 2002), fertilization proteins (Swanson et al. 2001; Civetta 2003), dopamine receptors (Ding et al. 2002), and lepidopteran LWS opsins (Briscoe 2001). However, we found that these methods did not identify positively selected sites which evolved to distinguish the opsin classes. Therefore, SET is a more powerful means for identifying these sites which arose over 500 MY ago when the opsin classes diverged.

Our SET findings identify those sites in the transmembrane regions which are important in spectral tuning of the opsin classes. However, there are no key differences in the cytoplasmic regions between the rod and cone opsins, suggesting that opsin–protein interactions are not a contributing factor to electrophysiological differences between rods and cones. SET analysis is therefore an excellent method for determining whether and in what ways proteins contribute to unique, signal transduction pathways. Future studies should examine additional phototransduction proteins to identify those that make major contributions to the unique physiological differences between rods and cones.

## Materials and Methods

### *Opsin Phylogenies*

Vertebrate opsin sequences were identified by BLAST and downloaded from GenBank. Taxa included all of the major vertebrate groups (mammals, birds, reptiles, amphibians, fish, and deeper vertebrates such as skates, coelecanths, and lampreys). There were 48 M/LWS, 27 RH2, 26 SWS2, 32 SWS1, and 55 RH1 opsins included with no partial sequences used (see Supplementary Table 1). For the RH2 and SWS2 classes, there were no mammalian sequences available, as mammals have lost these two opsin classes (Yokoyama 2000). The lamprey RH1 and RH2 sequences were not included in these analyses because of their equivocal phylogenetic relationships and evidence that they both reside in cone photoreceptors (Collin et al. 2003).

Nucleotide sequences were translated to amino acid sequence and then aligned using ClustalW in DNAstar (Burland 2000). Alignments were visually inspected and fine tuned as needed. This was most necessary at the N- and C-termini, where ClustalW does not properly align homologous sites. The C-terminus alignment is important as it has been shown to be key to opsin trafficking (Concepcion et al. 2002).

Phylogenies were generated by several methods. Trees based on amino acid sequences were constructed by both parsimony and the neighbor-joining algorithm using amino acid distances. Distances were corrected for saturation using distance = $-\ln(1 - \text{proportion of amino acid site differences})$ (Graur and Li 2000). For trees based on nucleotide sequence, the maximum likelihood method was used to estimate the transition to transversion ratio, the fraction of invariant sites and the gamma distribution shape parameter. These data were then used to constructed trees using Tamura–Nei (1993) corrected distances with the neighbor-joining method or the maximum likelihood method. Because of the computational requirments of the maximum likelihood method, different subsets of the sequences were analyzed together to test the various segments

of the tree. All methods of tree construction were implemented with PAUP* 4.0b2 (Swofford 1999). The significance of each node was assessed from 1000 bootstrap replicates. The opsin classes determined from this phylogenetic analyses were used in the ET analysis.

### *Sequence Saturation*

For each of the opsin classes, we examined whether the available sequences were sufficient to define the set of conserved sites for that class. This was determined by whether a set of sequences had reached saturation. To test for saturation, the number of unanimous sites was calculated when increasing numbers of taxa were added to each class. If saturation had been reached, then adding more taxa did not significantly change the number of unanimous sites. The fraction of sites which were unanimous was plotted as a function of total tree length, where tree length was used as a proxy for total divergence time and was calculated in PAUP for the varying number of taxa. For each class of opsin, these plots were fit to

$$f = f_o + (1 - f_o)[(1 - b) + be^{-2\lambda T/b}] \quad (1)$$

where $f_o$ is the fraction of sites which do not vary, $\lambda$ is the rate at which amino acid sites diverge with time, $1-b$ is 0.05 (the probability that two amino acids are the same by chance), and $T$ is the tree length. This equation is based on a Poisson process for sequence divergence (Jukes and Cantor 1969) with two modifications. First, to account for amino acid divergence, $1-b$ is based on the number of amino acids (1/20) instead of the number of nucleotides (1/4). Second, the equation has been modified to include a fraction of sites which remain invariant (Nei 1987, p 42). For large $T$, the fraction of sites which are unanimous at saturation, $f_s$, is

$$f_s = f_0 + (1 - f_0)(1 - b) \quad (2)$$

For each opsin class, the percentage saturation was determined by comparing the actual fraction of unanimous sites for the available sequences, $f_a$, with the predicted fraction at saturation (Eq. 2), $f_s$, by

$$\%\,\text{saturation} = (1 - f_a)/(1 - f_s) * 100 \quad (3)$$

### *SET Analysis*

SET analysis was performed in two ways. In the first method, the rod opsins were compared with all the cone opsin classes combined. In the second method, the rod opsin class was compared separately to each of the cone opsin classes. To compute the trace, amino acid positions which were identical for all sequences within an opsin class (which we denote as unanimous sites) were identified to make a unanimity sequence. A unanimity sequence was also determined for the combined cone opsin classes (denoted "all cone"). The unanimity sequences for the five opsin classes were then compared to generate a trace sequence that identified the invariant sites and the class-specific sites. These sites were mapped onto the three-dimensional crystal structure of bovine rhodopsin: 1L9H (Palczewski et al. 2000; Teller et al. 2001; Okada et al. 2002). Cytoplasmic, transmembrane, and extracellular regions were determined from the crystal structure. For the 1L9H structure, the amino acid positions of the seven TM regions are as follows: I, 34–64; II, 71–100; III, 106–139; IV, 150–172; V, 200–225; VI, 244–276; and VII, 286–309. The eighth helix is at 311–323 (Filipek et al. 2003).

### *Likelihood Ratio Tests for Detecting Unique Sites*

In an attempt to independently identify sites that were unique to each class, we utilized the statistical methods based on maximum likelihood (Yang 1994, 1998). In particular, we used codon-based maximum likelihood methods (Code ML; Yang et al. 2000) as implemented in the Phylogenetic Analysis by Maximum Likelihood package (PAML version 3.14). In these analyses, $\omega$ is the ratio of nonsynymous-to-synonymous substitution rates ($\omega$ = dN/dS). The value of $\omega$ indicates the selective regime with sites under purifying ($\omega < 1$), neutral ($\omega = 1$), and positive ($\omega > 1$), selection. Nested pairs of models were compared using likelihood rates tests (LRT) to look for evidence of positive selection and the sites under such selection. The goal was to determine if such sites could be identified during the early divergence of the opsin classes.

Three LRTs were performed. The first LRT tests for heterogeneity in $\omega$ across sites (comparing models M0 and M3). The second LRT tests for positive selection by comparing a two-rate class model (M1) with a model where a variable $\omega$ class is added (M2). The third LRT is an additional test for positive selection (comparing models M7 and M8) where the rate classes are fit to a $\beta$ distribution (M7) versus a model which adds an additional class of variable $\omega$. For both M1/M2 and M7/M8 comparisons, positive selection is implicated only if the extra variable $\omega$ class has a value of $\omega > 1$.

Because of the large computational requirements for these maximum likelihood calculations, only two opsin classes were considered at a time. Each cone opsin class was compared with the rod opsin class, so that the maximum likelihood comparisons mirrored the SET comparisons. Preliminary analyses used 5–7 sequences for each opsin class, for a total of 10–12 sequences in any one LRT. Additional tests were performed with as many as 33 sequences at a time.

## Results and Discussion

### *Phylogenetic Analysis Defines Five Opsin Classes*

The phylogenetic tree containing 188 vertebrate opsin genes consists of five major classes—RH1, SWS1, SWS2, RH2, and M/LWS—all with high bootstrap support (95–100%). Figure 1 shows the neighbor-joining, nucleotide distance tree. Distances were corrected for multiple hits using the Tamura–Nei gamma distribution with empirical base frequences and the following parameters fit using maximum likelihood: transition/transversion ratio = 0.986, fraction invariant sites = 0.15, and gamma shape parameter = 0.97. This tree is in good agreement with the different maximum likelihood trees and the parsimony and distance trees based on amino acid sequences. The only difference between the trees based on nucleotide and amino acid sequence was lack of reciprocal monophyly of the two RH classes. In the trees based on amino acid sequence, the RH1 class is embedded within the RH2 class, though the RH1, fish RH2, and tetrapod RH2 nearly form a trichotomy. This paraphyly is likely the result of our inability to correct amino acid distances as well as nucleotide distances. The Poisson corrections for protein evolution are less sophisticated than those available for nucleotide sequence. Both nucleotide distance and maximum likelihood trees provide good support for monophyly of the five opsin classes. This monophyly

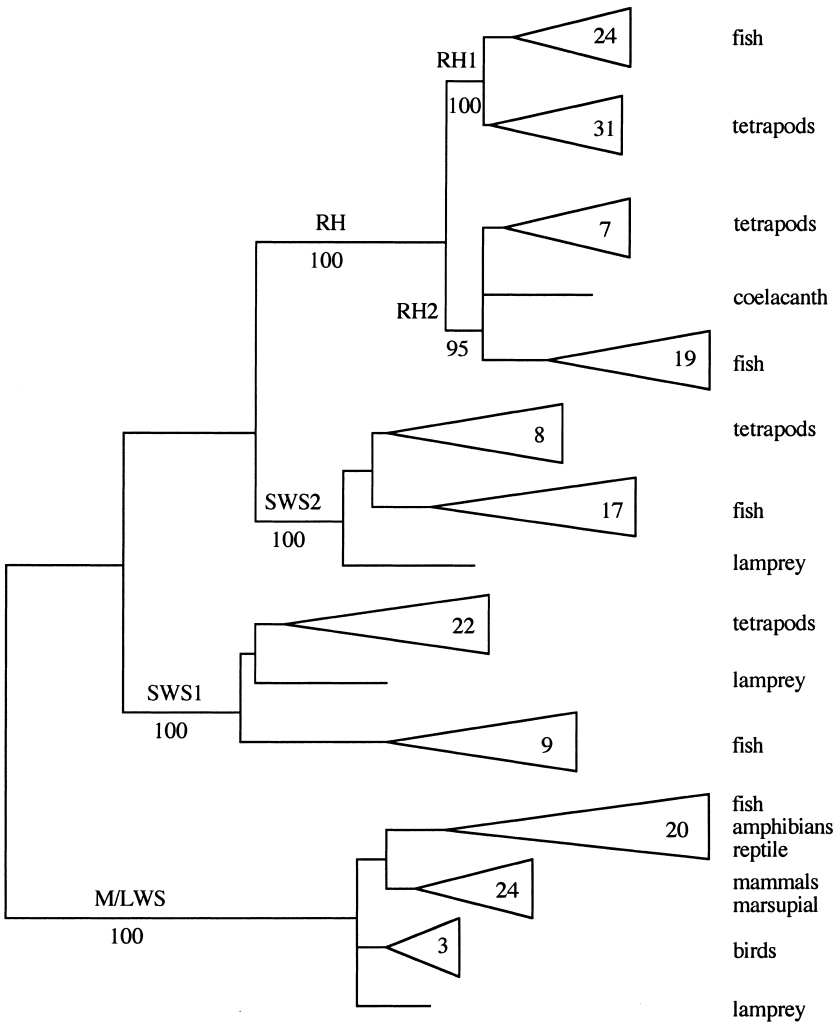



**Fig. 1.** Phylogenetic tree showing the relationships of the 188 opsin sequences used in this study. This neighbor-joining tree is based on nucleotide distances which have been corrected for multiple hits by a Tamura–Nei gamma distribution. The opsin classes are labeled on the branches leading to each clade, with the corresponding bootstrap values given below. The number of sequences within each group is given in the triangles.

is consistent with previous work suggesting that the M/LWS opsin class evolved first, followed by the SWS1, SWS2, and RH classes (Okano et al. 1992; Yokoyama 2000). This is evident when an invertebrate opsin is used as an outgroup (data not shown). The RH1 class is the youngest class, which supports the idea that rods have evolved from cones (Okano et al. 1992; Ebrey and Koutalos 2001).

Within each class, sequences are generally related as expected based on known organismal relationships, with tetrapods (amphibians, birds, mammals) separate from fish. However, there were a few exceptions. For the M/LWS class, birds group outside of the mammals and fish/amphibia group for the nucleotide tree, while mammals group separately from fish/birds/amphibia in the amino acid tree. This may indicate an ancient gene duplication event or it may be due to the difficulties of correcting for saturation over these long ($>500$ MY) timescales. The only other anomaly was in the SWS1 class; lamprey groups with the tetrapods rather than outside of the fish/tetrapod split. This may be the result of long branch attractions as lamprey is the oldest vertebrate diverging $\sim$560 MY ago from other vertebrates (Kumar and Hedges 1998).

This phylogenetic analysis identifies five major opsin classes for the SET analysis: M/LWS, SWS1, SWS2, RH2, and RH1. Because of the loss of the SWS2 and RH2 genes in the mammalian lineage, there is some difference in phylogenetic diversity within each class. If different lineages evolved at different rates, this might bias the analyses. In spite of the fact that the number of taxa included in each tree is different, the total tree lengths are similar for each class (RH1, 2.2; SWS1, 2.3; SWS2, 2.3; RH2, 1.5; and LWS, 2.1 amino acid substitutions per site). In addition, maximum likelihood analyses did not suggest that the mammalian lineages had significantly different rates of evolution from those of other taxa. Because of the similar tree lengths and the fact that each class is close to saturation (see below), we conclude that there is no significant bias as a result of differences in class phylogenetic diversity.

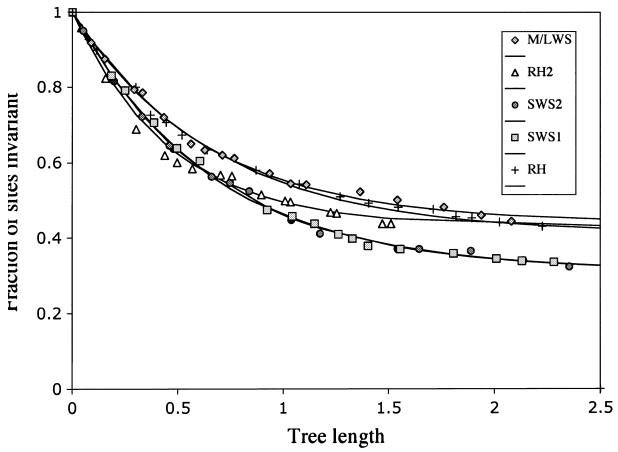



**Fig. 2.** The fraction of sites which are fixed in each opsin class as a function of increasing tree length (sequence diversity). The points are determined by subsampling different subsets of the sequences in each opsin class. The lines are fit to Eq. 1.

### *Sequence Divergence Has Saturated in All Opsin Classes*

ET first determines sites which are unanimous within a given protein class and then compares these unanimity sequences between classes (Lichtarge et al. 1996, 2002). To achieve this, a sufficient number of phylogenetically diverse sequences must be compared to ensure that amino acid identity is significant and not the result of similarity by descent. Past studies have failed to quantify the extent to which "site saturation" has been achieved. To determine whether the available sequences within each family are sufficient for ET analysis, we subsample the data to determine the relationship between sequence unanimity and sequence diversity (or tree length). In this way, we can compare the unanimity for the entire data set with that predicted by extrapolating the data to infinite tree length. For the available data, we find the sequence saturation is 96–99% for all five opsin classes (Fig. 2). The estimated fits predict that there are at most two to eight sites that are currently identified as unanimous which are potentially variable. Therefore, the current data defines the unanimous sites of each opsin class with better than 95% degree of confidence. Additional sequences are not needed to define the quintessential sites for each class. This is the first time that sequence saturation has been demonstrated, enabling us to perform saturated ET (SET) analysis.

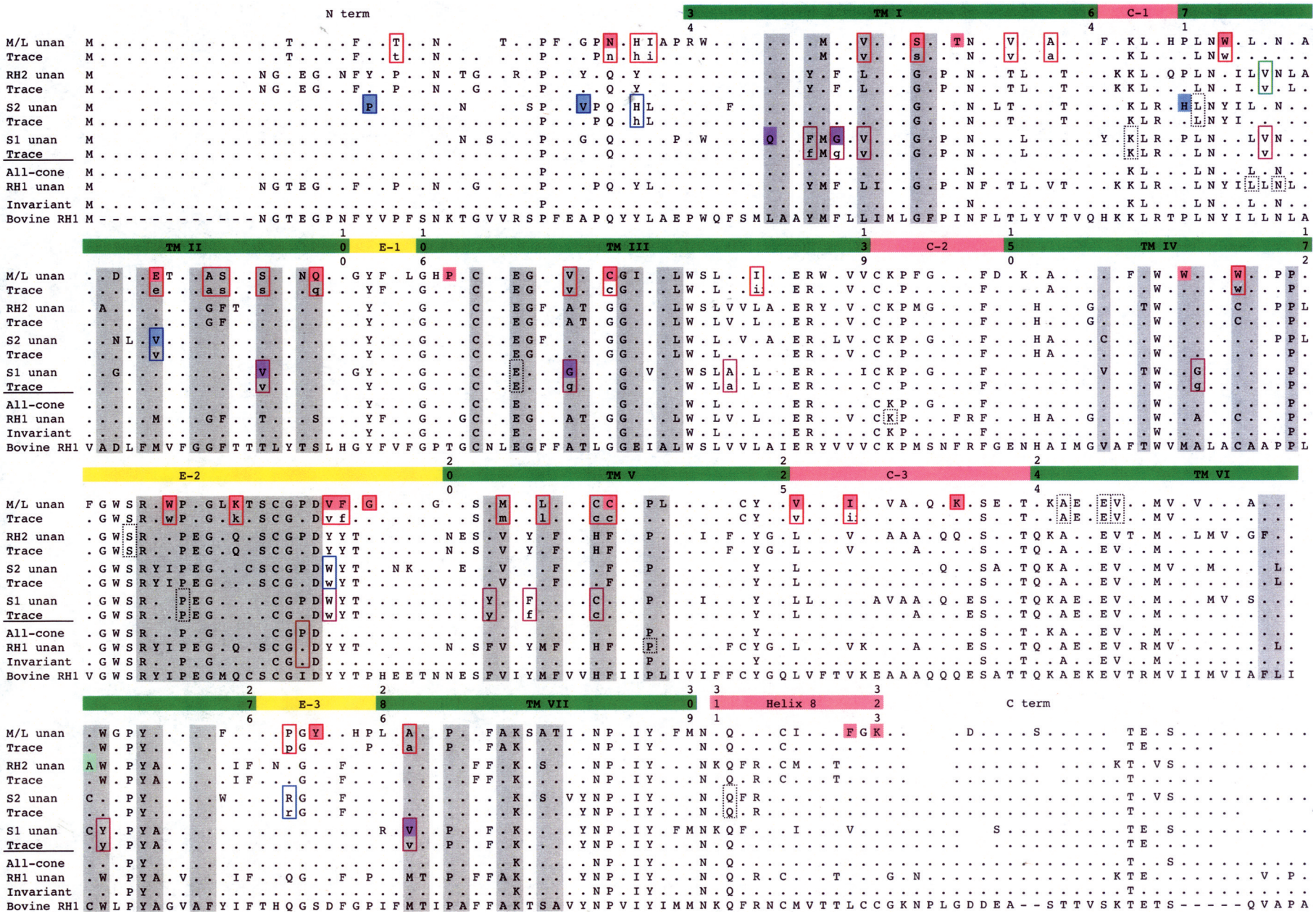
### *Each Cone Opsin Class Has a Signature of Key Sites but Rod Opsins Do Not*

Figure 3 shows the unanimity sequences for each of the five opsin classes aligned with respect to bovine rhodopsin. By comparing the amino acids present at each site with those which occur in the other opsin classes, we can identify sites where the amino acid is unique to that opsin class and does not occur in any other class (Supplementary Table 2). These are the result of fixed differences that evolved along the branch leading to the opsins in that class. These sites provide a class "signature" and are shaded in color in Fig. 3. The M/LWS opsin class has 22 signature sites, suggesting that it has unique characteristics (see below). The other cone opsin classes have far fewer signature sites (SWS1, 5; SWS2, 4;and RH2, only 1). Surprisingly, the RH1 class does not have any signature sites. Although almost half (155) of this class's sites are unanimous, the amino acid at any given position in rod opsin is present in at least one of the cone opsin classes.

### *SET Analysis Identifies Sites That Are Invariant in All Opsin Classes*

SET analysis of the unanimity sequences for the rod class and the all-cone opsin class (unanimity sequence of all four cone opsin classes combined) identifies 38 amino acids which are invariant for all retinal opsins (Fig. 3; "Invariant" row). In addition to these 38 invariant sites identified, 13 additional sites are unanimous for all but 1 of the 188 sequences (outlined with dotted lines in Fig. 3). Single base differences could be the result of DNA sequencing errors, so we consider these thirteen sites also to be invariant. This makes 51 sites (14%) which are unanimous for all opsins and therefore critical for maintaining protein structure and function. A few of these sites are unique to opsins relative to other G-protein coupled receptors (Dean et al. 2001, Madabushi et al. 2002). The invariant sites are most common in the cytoplasmic and extracellular loops (20 and 25% invariant), while 14% of the sites in the TM region and $< 5\%$ of the N- or C-terminal residues are invariant.

**Fig. 3.** Unanimity and trace sequences determined from the SET analysis. For each cone opsin class, the unanimity sequence is shown, and below it is given the trace sequence based on comparison of the corresponding cone opsin class with the rod opsin class. Invariant sites are shown in capital letters. Sites which are class-specific are shown in lowercase letters and outlined with a colored box. Signature sites are shaded in color. For simplicity, there are 10 variable sites after the start codon of the M/LWS class which are not included in this alignment. These sites do not overlap with any other opsin class. In the second part of the alignment, sites that are unanimous for the all-cone class ("All-cone") or unanimous in the RH1 class ("RH1 unan") are listed. Sites that are invariant in all opsin classes ("Invariant") are shown above the bovine RH1 sequence (used as a reference for numbering sites). The transmembrane helices defined by the x-ray crystal structure are outlined (Filipek et al. 2003) and the sites which are directed into the retinal binding pocket close to the level of retinal are shaded gray. The β sheet inserted into the retinal binding pocket in loop E-2 is also shaded. Sites identical in all but one opsin sequence are outlined with a dotted line. The primary domains are numbered and marked with colored bars: transmembrane (green), cytoplasmic (pink), and extracellular (yellow).


N term
TM I
C-1
M/L unan
Trace
RH2 unan
S2 unan
S1 unan
All-cone
RH1 unan
Invariant
Bovine RH1
TM II
E-1
TM III
C-2
TM IV
E-2
TM V
C-3
TM VI
E-3
TM VII
Helix 8
C term

*Class-Specific Sites*

Two sets of analyses were done to look for class-specific sites. In the first case, we compared the rod class with the all-cone class. This analysis examined the hypothesis that rod and cone opsins had distinctive signatures with unique sites contributing to differences in rod and cone phototransduction. In the second case, we performed pairwise comparisons between each of the opsin classes. This analysis looked for evidence of class-specific function which would likely include spectral sensitivity but might also involve other functions.

*Few Class-Specific Sites Distinguish Rod and Cone Opsins*

When the 133 cone opsin sequences are analyzed as a single class, there are 55 sites that are identical in all four cone opsin classes. (This includes 8 sites which vary in only 1 of the 133 sequences). This is far fewer than the 155 sites (42%) which are conserved within the rod opsin class itself. Comparisons of the all-cone and rod unanimity sequences reveals no class-specific sites. However, there are four sites which are unanimous in the all-cone opsin class but variable in rods. For three of these sites, the cone opsin amino acid also occurs in some rhodopsins (G144 to A, S, T, or G in RH1; K245 to K or R in RH1; and S343 to S or T in RH1). However, the fourth site has a unique amino acid in cone opsins different from those in rhodopsin (P189 in cones versus I/V in rods). This site has been shown in mutagenesis experiments to increase the visual pigment's meta II decay rate (Shichida and Imai 1999; Imai et al. 2001; Kuwayama et al. 2002). While the meta II decay rate has been ruled out as the rate limiting step in opsin deactivation (Lyubarsky et al. 2000; Kefalov et al. 2003), we hypothesize that it may be important in distinguishing rod and cone opsin regeneration kinetics. In all other respects, SET analysis reveals no significant differences between the visual receptor properties of rods and cones (see more below).

*Likelihood Ratio Tests to Identify Unique Sites*

The three maximum likelihood LRTs were performed for each pairwise combination of a cone opsin class with the rod opsin class. (For results see Supplementary Table 3.) The more complex models which contained extra rate classes did fit the data better for both the M1/M2 and the M7/M8 comparisons. Unfortunately, for these cases the values for $\omega$ of these extra rate classes were either much less than 1 or the fraction of sites in the extra class was 0.000. Therefore, there was no evidence of positive selection, or of sites subject to selection, in comparisons between these opsin classes. This was further confirmed with models where $\omega$ was allowed to vary in a branch-specific manner. No evidence for positive selection was detected.

Although the maximum likelihood method has been used to identify sites under positive selection in other molecular studies, no sites were identified in opsin comparisons. These other studies involved more recent evolutionary events, such as evolution of lysozyme in New World monkeys which occurred in the past 30 MY (Yang 1998). However, the opsin classes diverged extremely early in the evolution of vertebrates, prior to the split between jawless and jawed vertebrates (Collin et al. 2003). The tree lengths used in this study ranged from 10 to 50 nucleotide substitutions per codon across the tree. This amount of divergence is considered high and can make positive selection quite difficult to detect (Anisimova et al. 2001). Increasing the number of taxa can sometimes improve detection. However, when we tripled the number of taxa included, the results remained the same with no sites being identified as under positive selection (data not shown). We conclude that the SET method is more powerful at detecting sites which evolved to distinguish the opsin classes over 500 MY ago and which were then retained by purifying selection.

*Comparisons with Mutational Studies*

To test the predictive ability of the SET analysis to reveal functionally important sites in opsins, we first compared the invariant sites with sites which have been studied by mutational analyses (Supplementary Table 4). Of the 51 invariant sites, 30 have been shown to impact some aspect of opsin signal transduction, including chromophore regeneration, transducin activation, phosphorylation, and arrestin binding. Mutations of eight more of the invariant sites have been shown to impact other G-protein coupled receptors but have not yet been studied in opsins (Madabushi et al. 2002). Twelve of the sites have not been studied. The mutation W161L has been shown in rhodopsin to have no effect. However, other mutations at this site in 10 different G-protein coupled receptors has been shown to impact ligand binding (Madabushi et al. 2002). Therefore, all of the invariant sites identified by SET analysis that have been studied experimentally have been shown to impact protein function.

*Comparisons with Disease States*

As an additional test of this analysis, we compared both the class-specific and the invariant sites with known disease states. In this analysis, only single point mutations were considered ignoring deletions and premature terminations. Specific amino acids causing human disease were obtained from OMIM

**Table 1.** Pairwise comparisons of unanimity amino acid sequences for all five opsin classes

| | M/LWS | RH2 | SWS2 | SWS1 | RH1 |
|---|---|---|---|---|---|
| M/LWS | — | 68 | 60 | 67 | 67 |
| RH2 | 30 | — | 85 | 75 | 111 |
| SWS2 | 22 | 7 | — | 65 | 77 |
| SWS1 | 16 | 15 | 4 | — | 64 |
| RH1 | 30 | 1 | 4 | 14 | — |

The number of sites that are identical between two classes is shown above the diagonal and the number of class-specific sites for each pairwise comparison is shown below.

for rhodopsin as well as for the red, green, and blue cone opsins (Supplementary Table 5).

Of the 44 different mutations in the rhodopsin protein which cause retinal disease, 38 are the result of point mutations at 28 different amino acid sites. Comparisons with the SET analyses show that 9 of these sites are invariant in all opsin proteins and 14 are invariant in the rod opsin class. The remaining five sites are variable in the RH1 class. Two (T17M and V87I) differ for only 2 of 55 sequences in the class with quite conservative changes. The remaining three sites (G89D, E150K, and A292E) are more highly variable. However, the naturally occurring variation retains amino acid charge, while the disease-causing mutation results from a change in R-group charge. This correlation between highly conserved sites and disease-causing mutations in rhodopsin agrees with a recent comparison of vertebrate rhodopsins and retinal disease states (Briscoe et al. 2004).

For the 12 disease states in cone opsins, there are seven unique sites with point mutations leading to loss of function (ignoring site S180A, which causes spectral shifts in the red opsins but not loss of function). For these seven disease-causing sites, two are invariant in all opsins and three are invariant within the particular opsin gene class involved.

These comparisons of rod and cone disease states with SET sites suggest that mutations at invariant or class-conserved sites are excellent predictors of disease-causing mutations. Of the 35 unique sites in rod and cone opsins causing disease by point mutation, 82% are unanimous and 86% are very highly conserved as judged by SET analysis. In those cases where SET fails to identify a site, the amino acid disease causing mutation is one which alters the charge of the R group, which would likely have a major impact on opsin structure and function.

In conclusion, comparisons with mutational studies and disease states demonstrate that SET is an accurate predictor of critical residues for protein function. Having demonstrated the predictive potential of our SET analysis, we now apply this approach to identify class-specific sites important for visual pigment wavelength modulation, photoactivation properties, and downstream signal transduction.

### *Comparisons of Opsin Families to Identify Class-Specific Sites Important for Spectral Tuning*

Class-specific sites were identified for each of the pairwise comparisons between the opsin classes. The number of identical and class-specific sites for each of these comparisons is given in Table 1. The RH1 and RH2 classes are most similar to each other, having the highest number of unanimous sites. The M/LWS class is the most dissimilar to both the RH1 and the other cone opsin classes. These differences are a direct reflection of the phylogenetic relationships shown in Fig. 1.

For simplicity, we focus our discussion on comparisons of class-specific sites in each cone opsin class versus the rod opsin class. These comparisons capture all the critical differences between these classes, particularly in the TM regions. The class-specific sites determined from each cone opsin class compared with the RH1 class are outlined in Fig. 3 and highlighted on the bovine rhodopsin crystal structure in Fig. 4. In addition, class-specific sites in the retinal binding pocket (both in the transmembrane domains and in extracellular loop 2) are listed in Table 2 for comparison with site-directed mutagenesis studies.

The RH2 opsins are most similar to RH1. These two groups are distinguished by only one class-specific site, which is not in the retinal binding pocket and involves a conservative change (L77V, where the first amino acid occurs in RH1 and the second occurs in the class in question, in this case RH2). The RH2 opsins also have a signature site, A264 (C or S in RH1), which is in the retinal binding pocket. This signature site cannot have a major effect on spectral tuning, because the RH2 and RH1 opsins completely overlap in spectral sensitivity (RH1, 470–510 nm; RH2, 466 to 511 nm). These two opsins are essentially identical in their predicted structure and function and are distinguished only by the cell types where they are expressed.

The SWS2 opsins have four class-specific sites, of which M86V (TM II) is in the retinal binding pocket and Y191W (loop E-2) is close to it. Site 86 has been shown to be important in tuning the SWS1 opsins, where it can produce shifts of up to 70 nm (Lin et al.

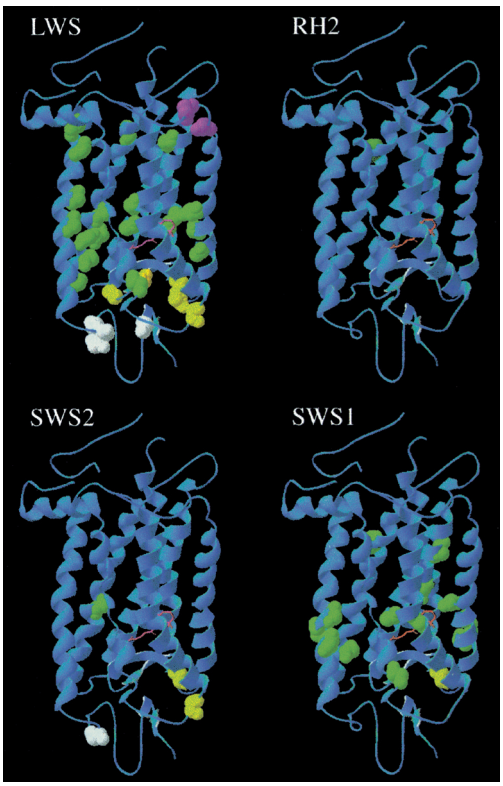



**Fig. 4.** Class-specific sites for each of four cone opsin classes determined by comparison with the rod opsin class. Sites are labeled by region as transmembrane (green), cytoplasmic (pink), extracellular (yellow), and N- or C-termini (white). 11-cis retinal is shown in red.

1998; Yokoyama et al. 2000; Cowing et al. 2002b; Fasick et al. 2002). Site Y191W is next to the β sheet inserted into the retinal binding pocket. The SWS2 class is sensitive at 420 to 474 nm. Both M86V and Y191W must contribute to the short wavelength sensitivity of this class. These sites need to be tested by site-directed mutagenesis to confirm their importance. The functional significance of the other two class-specific sites (Y30H in the N-terminus and Q279R in loop E-3) is uncertain.

The SWS1 opsins have 14 class-specific sites with respect to RH1. Thirteen are in the TM regions, of which nine are in the retinal binding pocket itself (Table 2). They are also strategically located within 4.5 Å of retinal (A117G, H211C), close to the E113 counterion for the Schiff base (Y43F and T94V) or close to retinal's indole ring (W265Y) (Sakmar et al. 2002). In addition, site Y191W (the same site as in the SWS2 opsins) is close to the retinal binding pocket and likely contributes to short wavelength shifts. Spectral tuning effects of A117G (−4 nm [Nakayama and Khorana 1991; Han et al. 1996; Lin et al. 1998]), H211C (−6 nm [Nathans 1990]), and W265Y (−13 to 15 nm [Nakayama and Khorana 1991; Lin et al. 1998]) have been demonstrated in the RH1 opsins by site-directed mutagenesis. In addition, mutations at site T94 have also been shown to cause short wavelength shifts of between 4 and 22 nm (Ramon et al. 2003). The SWS1 opsin class is the shortest wavelength sensitive class, with sensitivities at 358 to 425 nm. Three of the class-specific sites (A117G, H211C, and W265Y) are known to cause shifts to lower wavelengths. However, the effects of mutations in SWS1 opsins are known to be nonlinear and complex (Shi et al. 2001). Site-directed mutagenesis of single and multiple sites is needed to identify the relative contributions of the nine binding pocket sites plus Y191W in modulating the SWS1 wavelength sensitivities.

The M/LWS opsins have 30 class-specific sites with respect to RH1. Nineteen of these sites are in the TM, of which 14 are in the retinal binding pocket. In addition, four sites in loop E-2 are also close to the binding pocket site (I179W, Q184K, Y191V, and Y192F; Table 2). Several of these sites involve changes in amino acid polarity or charge (G51S, M86E, S98Q, Q184K, Y192F, V204M, M207L). There are also sites that are either within 4.5 Å of retinal (A117V, M207L, H211C, and F212C) or close to the E113 counter ion (L47V T94S) (Sakmar et al. 2002). Sites Q184K and Y181H in loop E-2 are known to cause long wavelength shifts as part of a chloride binding site (Wang et al. 1993). Although Y181H causes a 20-nm shift to longer wavelengths (Yokoyama and Radlwimmer 2001), it varies (H or Y) within the M/LWS class and so is not a class-specific site. Several of these sites (86, 117, and 207) are known to cause spectral tuning, though the exact mutations important for the M/LWS class have not been examined. Site-directed mutagenesis will be required to determine how these sites contribute to the sensitivity of the M/LWS class, which has $\lambda_{max}$ from 521 to 575 nm.

It is interesting that several of the class-specific sites reoccur in more than one opsin class. Classes differ in which mutations occur, suggesting that these sites can cause opposite shifts. This includes site 86 (M86V in SWS2, M86E in M/LWS), 94 (T94V in SWS1, T94S in M/LWS), 117 (A117G in SWS1, A117V in M/LWS), 191 (Y191W in both SWS1 and SWS2, Y191V in M/LWS), and 288 (M288V in SWS1, M288A in LWS). The fact that these sites are alternatively fixed in classes at the wavelength extremes offers more support that these sites are critical to opsin tuning.

Analysis of the class-specific SET sites leads to several conclusions. First, the majority of the class-specific sites identified for each of the cone opsin classes are in the retinal binding pocket, and so spectral tuning must be one of their primary functions. Second, SET analysis has produced a comprehensive survey of sites likely to modulate wavelength and so explain the large-scale tuning differences between the opsin classes. Third, the importance of these sites for spectral tuning has

**Table 2.** Comparison of SET class tuning sites in the retinal binding pocket (in TM regions or extracellular loop 2) with site-directed mutagenesis and expression of visual pigments

| Opsin class | Class tuning site SET | Related mutagenesis | Opsin class studied | Effect (nm) | Reference(s) |
|---|---|---|---|---|---|
| RH2 | None | | | | |
| SWS2 | M86V | F86Y | SWS1 | 60–66 | Cowing et al. 2002, Fasick et al. 2002 |
| | | M86L | RH1 | −2 | Lin et al. 1998 |
| | Y191W | | | | |
| SWS1 | Y43F | | | | |
| | L47V | | | | |
| | T94V | T94I | RH1 | −22 | Ramon et al. 2003 |
| | | T94D | | −4 | |
| | | T94S | | −6 | |
| | A117G | **A117G** | RH1 | −4 | Lin et al. 1998 |
| | A164G | A164S | RH1 | +2 | Chan et al. |
| | | **A164G** | SWS2 | 0 | Yokoyama and Tada 2003 |
| | Y191W | | | | |
| | F203Y | | | | |
| | H211C | **H211C** | RH1 | −6 | Nathans 1990 |
| | W265Y | **W265Y** | RH1 | −13 | Nakayama and Khorana 1991, Lin et al. 1998 |
| | M288A | | | | |
| M/LWS | L47V | | | | |
| | G51S | | | | |
| | M86E | F86Y | SWS1 | 60–66 | Cowing et al. 2002, Fasick et al. 2002 |
| | G90A | G90S | RH1 | −11 to −13 | Janz and Farrens 2001, Lin et al. 1998. |
| | F91S | S91P | SWS2 | 10 | Yokoyama and Tada 2003 |
| | T94S | A94S | SWS2 | 14 | Yokoyama and Tada 2003 |
| | S98Q | | | | |
| | A117V | A117W | RH1 | +4 | Han et al. 1996 |
| | | A117G | | −4 | Lin et al. 1998 |
| | C167W | | | | |
| | I179W | | | | |
| | Q184K | | | | Wang et al. 1993 |
| | Y191V | | | | |
| | Y192F | | | | |
| | V204M | | | | |
| | M207L | L207M | RH2 | +6 | Yokoyama et al. 1999 |
| | | L207I | SWS2 | −6 | Yokoyama and Tada 2003 |
| | H211C | **H211C** | RH1 | −6 | Nathans 1990 |
| | F212C | | | | |
| | M288A | | | | |

*Note*. Class-specific tuning sites are given as RH1 amino acid, the RH1 site number, and the cone class amino acid. Mutations exactly matching the SET class tuning sites are in boldface.

been confirmed in several instances by mutagenesis studies. (Indeed, three of the SWS1 sites have already been demonstrated to cause the exact shift predicted by this analysis [Nathans 1990; Nakayama and Khorana 1991; Lin et al. 19984]). Additional sites predicted by SET to be critical to wavelength tuning need to be examined with site-directed mutagenesis.

While SET analysis is very useful for identifying large-scale tuning effects between opsin families, it is not informative in identifying amino acids that contribute to the small-scale spectral tuning within a given opsin class, (Teller et al. 2003). This is because within-class spectral differences (40–70 nm) are determined by sites which are variable among the opsins of a given class. Site variability means they are not unanimous in a class and therefore not considered by SET. Fine-scale spectral tuning within an opsin class is likely to have evolved multiple times during vertebrate evolution as organisms adapted to different photic environments (Yokoyama and Yokoyama 1996). Other approaches such as comparisons of closely related species and site-directed mutagenesis have successfully identified many fine-scale tuning sites: M/LWS (Yokoyama and Radlwimmer 2001), RH2 (Yokoyama et al. 1999; Yokoyama and Tada 2000), SWS2 (Cowing et al. 2002a; Takahashi and Ebrey 2003; Yokoyama and Tada 2003), SWS1 (Wilkie et al. 2000; Shi et al. 2001; Cowing et al. 2002b; Fasick et al. 2002), and RH1 (Nakayama and Khorana 1991; Lin et al. 1998; Yokoyama et al. 1999).

*Uniqueness of the M/LWS Opsin Class*

The M/LWS class is the oldest vertebrate opsin class, having differentiated early during the evolution of vertebrates (Yokoyama 2000). This class has retained 22 signature sites where the amino acid in the M/LWS class is not found in any other opsin classes. As described above, 23 of the 30 class-specific sites are in or close to the retinal binding pocket. The seven class-specific sites in the non-TM region include Q279P in the extracellular loops, L226V and V230I in the cytoplasmic loops, and P12T, Q28N, Y30H, and L31I in the N-terminus, which are all of unknown function.

These class-specific sites may contribute to the unique properties of the M/LWS opsins. Studies in tiger salamander (Perry and McNaughton 1991) and striped bass (Miller and Korenbrot 1993) have shown that the M/LWS opsins respond fastest to light stimulation and with the least sensitivity. Studies have also found evidence for increased thermal activation in M and L cones in salamanders (Rieke and Baylor 2000; Sampath and Baylor 2002). This agrees with the more rapid spontaneous isomerization rates quantified in *Xenopus* M/LWS opsin relative to that of RH1 (Kefalov et al. 2003). The spontaneous isomerization contributes to the decreased sensitivity of the red cones and their increased ability to adapt. The increased facility for 11-*cis*-retinal to isomerize may be the result of a less constrained retinal binding pocket created by the M/LWS opsin helices. We predict that some of the SET class-specific residues contribute to binding pocket conformation and the ease of 11-*cis*-retinal isomerization.

*Cytoplasmic Loops of Opsin Fail to Reveal Class-Specific Differences in Ability to Interact With Other Phototransduction Proteins*

The three cytoplasmic loops (C-1, C-2, and C-3) have a significant number of invariant sites, with 33, 40, and 11% invariance, respectively. Furthermore, examination of all the variable sites reveals that the amino acids occurring in the rod opsin are not unique to RH1, as they occur in some fraction of the cone opsins for all sites except one. This one site differs between cone (144G) and rod (A, S, or T) opsins though by a conservative change. There are only two class-specific sites in the cytoplasmic loops. These two sites occur in the M/LWS class and involve very conservative changes (L226V, V230I). We therefore conclude that the rod and cone opsin families are indistinguishable within the cytoplasmic loops. In addition to similarities in the cytoplasmic loops, the cytoplasmic ends of TM II, III, V, and VI each have several invariant sites which contribute to rod and cone similarities on the cytoplasmic side of the membrane. These sequence similarities are in stark contrast to the large number of class-specific sites identified in the retinal binding pocket.

The similarities of rod and cone opsin sequence in the cytoplasmic loops support the idea that rod and cone opsins have similar functionality in this region. We therefore predict that rod and cone opsins are equivalent in their interactions with other phototransduction proteins, including transducin, rhodopsin kinase, and visual arrestin.

These results are in agreement with the work of Kefalov et al. (2003), who demonstrated that when either a rod or a cone opsin was photoactivated in the same photoreceptor cell, the amplification and recovery kinetics were virtually identical. Our work makes it unlikely that the biochemical differences detected in phototransduction assays of isolated rod and cone membrane preparations (Tachibanaki et al. 2001) can be ascribed to intrinsic differences in opsin activation or deactivation. Rather, our SET analysis supports the alternative hypothesis that differences between other rod and cone phototransduction proteins are responsible for biochemical differences in light responsiveness of the rod and cone pathway. Future SET analyses of the proteins of the entire phototransduction cascade will reveal the most likely candidates whose structure/function have diverged in rods and cones. It is these candidate proteins, not opsins, which are most likely to have different binding affinities and/or enzymatic activities that result in the observed physiological differences in light responsiveness of rods and cones.

Our SET analysis of opsin sequences does not consider that chemically similar amino acids at any given site might fulfill the same function. Thus, our conservative implementation of SET analysis may underestimate the number of *functionally* identical sites within or between opsin families. The use of weighted ET analyses that consider the chemical similarity of amino acids in determining unanimous and class-specific sites (Landgraf et al. 1999) is certainly helpful when sequence saturation has not been achieved. However, it is unlikely that such an analysis will distinguish the cytoplasmic loops of rod and cone opsin classes which SET analysis has shown to be so similar.

## Conclusions

In summary, SET analysis of the vertebrate opsin family of G-protein coupled receptors has identified numerous sites which are key to this visual receptor's function, and leads to the following conclusions.

1. SET analysis of 188 vertebrate opsin sequences displays ≥96% site saturation for all opsin fami-

lies and defines 96–99% of the class-specific sites distinguishing these families. This comprehensive treatment is definitive in that analysis of additional sequences will not significantly change the results.

2. Class-specific sites were identified for all cone opsin classes compared to the rod opsin class. The majority of these class-specific sites reside in the retinal binding pocket close to retinal, where they can contribute to the large scale differences in wavelength sensitivities between these opsin classes. Several of the class-specific sites have been shown to shift the peak spectral sensitivity in mutational studies. However, the majority of the class-specific sites have not been tested by site-directed mutagenesis, and future work must experimentally examine their functional significance both individually and in combination.
3. The M/LWS opsin class has the most class-specific sites, with 30 in total, and is the only class with a significant number of class-specific sites outside of the spectral tuning region. These sites may alter the conformation of the M/LWS opsin structure and contribute to the greater ease with which 11-*cis*-retinal can be photoisomerized. This may explain why the M/LWS visual pigments have higher thermal isomerization rates and respond more quickly and less sensitively to light.
4. Invariant sites were identified in all opsin domains, with a significant number in the cytoplasmic and extracellular loops. Comparisons with site-directed mutagenesis and disease states confirm that these invariant sites are critical to opsin structure and function. The prevalence of invariant sites—and the lack of class-specific sites—in the cytoplasmic loops leads to the conclusion that rod and cone opsins have similar structure and hence similar function as far as protein–protein interactions are concerned.

We conclude that the only significant differences between the five classes of rod and cone opsins reside in the chromophore binding pocket. Therefore, differences in rod and cone opsin structure are unlikely to contribute to the substantial electrophysiological differences between rods and cones. Phototransduction proteins downstream from opsin will require close examination to identify their possible contributions to differences in rod and cone phototransduction pathways.

*Acknowledgments.* This work was supported by National Science Foundation Grant IBN-0131285 (to K.L.C), National Eye Institute Grant EY 05798 (to R.H.C.) and grants from the New Hampshire Agricultural Experiment Station. Thanks to Tom Kocher and two anonymous reviewers for comments on the manuscript. This is Scientific Contribution Number 2215 from the New Hampshire Agricultural Experiment Station.

## References

Anisimova M, Bielawski JP, Yang Z (2001) Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. Mol Biol Evol 18:1585–1592

Briscoe AD (2001) Functional diversification of lepidopteran opsins following gene duplication. Mol Biol Evol 18:2270–2279

Briscoe AD, Gaur C, Kumar S (2004) The spectrum of human rhodopsin disease mutations through the lens of interspecific variation. Gene 332:107–118

Borland TG (2000) DNASTAR's Lasergene sequence analysis software. Methods Mol Biol 132:71–91

Burns ME, Lamb TD (2003) Visual transduction by rod and cone photoreceptors. In: Chalupa LM, Werner LS (eds) The visual neurosciences. MIT Press, Boston, MA, pp 215–233

Casari G, Sander C, Valencia A (1995) A method to predict functional residues in proteins. Nat Struct Biol 2:171–178

Civetta A (2003) Positive selection within sperm-egg adhesion domains of fertilin: an ADAM gene with a potential role in fertilization. Mol Biol Evol 20:21–29

Collin SP, Knight MA, Davies WL, Potter IC, Hunt DM, Trezise AEO (2003) Ancient colour vision: multiple opsin genes in the ancestral vertebrates. Curr Biol 13:R864–R865

Concepcion F, Mendez A, Chen J (2002) The carboxyl-terminal domain is essential for rhodopsin transport in rod photoreceptors. Vision Res 42:417–426

Cowing JA, Poopalasundaram S, Wilkie SE, Bowmaker JK, Hunt DM (2002a) Spectral tuning and evolution of short wave-sensitive cone pigments in cottoid fish from Lake Baikal. Biochemistry 41:6019–6025

Cowing JA, Poopalasundaram S, Wilkie SE, Robinson PR, Bowmaker JK, Hunt DM (2002b) The molecular mechanism for the spectral shifts between vertebrate ultraviolet- and violet-sensitive cone visual pigments. Biochem J 367:129–135

Dean MK, Higgs C, Smith RE, Bywater RP, Snell CR, Scott PD, Upton GJ, Howe TJ, Reynolds CA (2001) Dimerization of G-protein-coupled receptors. J Med Chem 44:4595–4614

del Sol Mesa A , Pazos F, Valencia A (2003) Automatic methods for predicting functionally important residues. J Mol Biol 326:1289–1302

Ding YC, Chi HC, Grady DL, Morishima A, Kidd JR, Kidd KK, Flodman P, Spence MA, Schuck S, Swanson JM, Zhang YP, Moyzis RK (2002) Evidence of positive selection acting at the human dopamine receptor D4 gene locus. Proc Natl Acad Sci USA 99:309–314

Ebrey T, Koutalos Y (2001) Vertebrate photoreceptors. Prog Retin Eye Res 20:49–94

Fasick JI, Applebury ML, Oprian DD (2002) Spectral tuning in the mammalian short-wavelength sensitive cone pigments. Biochemistry 41:6860–6865

Filipek S, Stenkamp RE, Teller DC, Palczewski K (2003) G protein-coupled receptor rhodopsin: a prospectus. Annu Rev Physiol 65:851–879

Graur D, Li WH (2000) Fundamentals of molecular evolution. Sinauer Associates, Sunderland, MA

Han M, Lin SW, Smith SO, Sakmar TP (1996) The effects of amino acid replacements of glycine 121 on transmembrane helix 3 of rhodopsin. J Biol Chem 271:32330–32336

Hisatomi O, Tokunaga F (2002) Molecular evolution of proteins involved in vertebrate phototransduction. Comp Biochem Physiol B Biochem Mol Biol 133:509–522

Imai H, Hirano T, Kandori H, Terakita A, Shichida Y (2001) Difference in molecular structure of rod and cone visual pigments studied by Fourier transform infrared spectroscopy. Biochemistry 40:2879–2886

Imanishi Y, Li N, Sokal I, Sowa ME, Lichtarge O, Wensel TG, Saperstein DA, Baehr W, Palczewski K (2002) Characterization of retinal guanylate cyclase-activating protein 3 (GCAP3) from zebrafish to man. Eur J Neurosci 15:63–78

Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN (ed) Mammalian protein metabolism. Academic Press, New York, pp 21–132

Kalinina OV, Mironov AA, Gelfand MS, Rakhmaninova AB (2004) Automated selection of positions determining functional specificity of proteins by comparative analysis of orthologous groups in protein families. Protein Sci 13:443–456

Kefalov V, Fu Y, Marsh-Armstrong N, Yau KW (2003) Role of visual pigment properties in rod and cone phototransduction. Nature 425:526–531

Kochendoerfer GG, Lin SW, Sakmar TP, Mathies RA (1999) How color visual pigments are tuned. Trends Biochem Sci 24:300–305

Konig B, Arendt A, McDowell JH, Kahlert M, Hargrave PA, Hofmann KP (1989) Three cytoplasmic loops of rhodopsin interact with transducin. Proc Natl Acad Sci USA 86:6878–6882

Kumar S, Hedges SB (1998) A molecular timescale for vertebrate evolution. Nature 392:917–920

Kuwayama S, Imai H, Hirano T, Terakita A, Shichida Y (2002) Conserved proline residue at position 189 in cone visual pigments as a determinant of molecular properties different from rhodopsins. Biochemistry 41:15245–15252

Landgraf R, Fischer D, Eisenberg D (1999) Analysis of heregulin symmetry by weighted evolutionary tracing. Protein Eng 12:943–951

Li WH (1997) Molecular evolution. Sinauer Associates, Sunderland, MA

Lichtarge O, Bourne HR, Cohen FE (1996) Evolutionarily conserved Galphabetagamma binding surfaces support a model of the G protein-receptor complex. Proc Natl Acad Sci 93:7507–7511

Lichtarge O, Sowa ME (2002) Evolutionary predictions of binding surfaces and interactions. Curr Struct Biol 12:21–27

Lichtarge O, Sowa ME, Philippi A (2002) Evolutionary traces of functional surfaces along G protein signaling pathway. Methods Enzymol 344:536–556

Lin SW, Kochendoerfer GG, Carroll KS, Wang D, Mathies RA, Sakmar TP (1998) Mechanisms of spectral tuning in blue cone visual pigments. Visible and raman spectroscopy of blue-shifted rhodopsin mutants. J Biol Chem 273:24583–24591

Lyubarsky AL, Chen C, Simon MI, Pugh EN Jr (2000) Mice lacking G-protein receptor kinase 1 have profoundly slowed recovery of cone-driven retinal responses. J Neurosci 20:2209–2217

Madabushi S, Gross AK, Philippi A, Meng EC, Wensel TG, Lichtarge O (2004) Evolutionary trace of G protein-coupled receptors reveals clusters of residues that determine global and class-specific functions. J Biol Chem 279:8126–8132

Madabushi S, Yao H, Marsh M, Kristensen DM, Philippi A, Sowa ME, Lichtarge O (2002) Structural clusters of evolutionary race residues are statistically significant and common in proteins. J Mol Biol 316:139–154

Maeda T, Imanishi Y, Palczewski K (2003) Rhodopsin phosphorylation: 30 years later. Prog Retin Eye Res 22:417–434

Miller JL, Korenbrot JI (1993) Phototransduction and adaptation in rods, single cones, and twin cones of the striped bass retina: a comparative study. Vis Neurosci 10:653–676

Miller JL, Picones A, Korenbrot JI (1994) Differences in transduction between rod and cone photoreceptors: an exploration of the role of calcium homeostasis. Curr Opin Neurobiol 4:488–495

Nakayama TA, Khorana HG (1991) Mapping of the amino acids in membrane-embedded helices that interact with the retinal chromophore in bovine rhodopsin. J Biol Chem 266:4269–4275

Nathans J (1990) Determinants of visual pigment absorbance: identification of the retinylidene Schiff's base counterion in bovine rhodopsin. Biochemistry 29:9746–9752

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 USA 99:5982–5987

Okano T, Kojima D, Fukada Y, Shichida Y, Yoshizawa T (1992) Primary structures of chicken cone visual pigments: vertebrate rhodopsins have evolved out of cone visual pigments. Proc Natl Acad Sci 89:5932–5936

Oliveira L, Paiva PB, Paiva AC, Vriend G (2003) Identification of functionally conserved residues with the use of entropy-variability plots. Proteins 52:544–552

Onrust R, Herzmark P, Chi P, Garcia PD, Lichtarge O, Kingsley C, Bourne HR (1997) Receptor and betagamma binding sites in the alpha subunit of the retinal G protein transducin. Science 275:381–384

Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Le Trong I, Teller DC, Okada T, Stenkamp RE, Yamamoto M, Miyano M (2000) Crystal structure of rhodopsin: A G protein-coupled receptor. Science 289:739–745

Palczewski K, McDowell JH, Jakes S, Ingebritsen TS, Hargrave PA (1989) Regulation of rhodopsin dephosphorylation by arrestin. J Biol Chem 264:15770–15773

Pepe IM (2001) Recent advances in our understanding of rhodopsin and phototransduction. Prog Retin Eye Res 20:733–759

Perry RJ, McNaughton PA (1991) Response properties of cones from the retina of the tiger salamander. J Physiol 433:561–587

Pugh EN, Lamb TD (2000) Phototransduction in vertebrate rods and cones: Molecular mechanisms of amplification, recovery and light adaptation. In: Stavenga DG, de Grip WJ, Pugh EN (eds) Handbook of biological physics. Elsevier Science, New York, pp 183–255

Pugh EN Jr, Nikonov S, Lamb TD (1999) Molecular mechanisms of vertebrate photoreceptor light adaptation. Curr Opin Neurobiol 9:410–418

Ramon E, del Valle LJ, Garriga P (2003) Unusual thermal and conformational properties of the rhodopsin congenital night blindness mutant Thr-94 → Ile. J Biol Chem 278:6427–6432

Rieke F, Baylor DA (2000) Origin and functional impact of dark noise in retinal cones. Neuron 26:181–186

Sakmar TP, Menon ST, Marin EP, Awad ES (2002) Rhodopsin: insights from recent structural studies. Annu Rev Biophys Biomol Struct 31:443–484

Sampath AP, Baylor DA (2002) Molecular mechanism of spontaneous pigment activation in retinal cones. Biophys J 83:184–93

Shi W, Osawa S, Dickerson CD, Weiss ER (1995) Rhodopsin mutants discriminate sites important for the activation of rhodopsin kinase and Gt. J Biol Chem 270:2112–2119

Shi Y, Radlwimmer FB, Yokoyama S (2001) Molecular genetics and the evolution of ultraviolet vision in vertebrates. Proc Natl Acad Sci USA 98:11731–11736

Shichida Y, Imai H (1999) Amino acid residues controlling the properties and functions of rod and cone visual pigments. Novartis Found Symp 224:142–153; discussion 153–157

Sowa ME, He W, Wensel TG, Lichtarge O (2000) A regulator of G protein signaling interaction surface linked to effector specificity. Proc Natl Acad Sci USA 97:1483–1488

Starace DM, Knox BE (1997) Activation of transducin by a Xenopus short wavelength visual pigment. J Biol Chem 272:1095–1100

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 USA 98:2509–2514

Swofford DL (1999) PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, MA

Tachibanaki S, Tsushima S, Kawamura S (2001) Low amplification and fast visual pigment phosphorylation as mechanisms characterizing cone photoresponses. Proc Natl Acad Sci USA 98:14044–14049

Takahashi Y, Ebrey TG (2003) Molecular basis of spectral tuning in the newt short wavelength sensitive visual pigment. Biochemistry 42:6025–6034

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10:512–526

Teller DC, Okada T, Behnke CA, Palczewski K, Stenkamp RE (2001) Advances in determination of a high-resolution three-dimensional structure of rhodopsin, a model of G-protein-coupled receptors (GPCRs). Biochemistry 40:7761–7772

Teller DC, Stenkamp RE, Palczewski K (2003) Evolutionary analysis of rhodopsin and cone pigments: connecting the three-dimensional structure with spectral tuning and signal transfer. FEBS Lett 555:151–159

Thurmond RL, Creuzenet C, Reeves PJ, Khorana HG (1997) Structure and function in rhodopsin: peptide sequences in the cytoplasmic loops of rhodopsin are intimately involved in interaction with rhodopsin kinase. Proc Natl Acad Sci USA 94:1715–1720

Wang Z, Asenjo AB, Oprian DD (1993) Identification of the Cl (−)-binding site in the human red and green color vision pigments. Biochemistry 32:2125–2130

Wilkie SE, Robinson PR, Cronin TW, Poopalasundaram S, Bowmaker JK, Hunt DM (2000) Spectral tuning of avian violet- and ultraviolet-sensitive visual pigments. Biochemistry 39:7895–7901

Yang Z (1998) Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. Mol Biol Evol 15:568–573

Yang Z, Nielsen R (1998) Synonymous and nonsynonymous rate variation in nuclear genes of mammals. J Mol Evol 46:409–418

Yang Z, Nielsen R (2002) Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. Mol Biol Evol 19:908–917

Yau KW (1994) Phototransduction mechanism in retinal rods and cones. The Friedenwald Lecture. Invest Ophthalmol Vis Sci 35:9–32

Yokoyama S (2000) Molecular evolution of vertebrate visual pigments. Prog Retin Eye Res 19:385–419

Yokoyama S, Radlwimmer FB, Blow NS (2000) Ultraviolet pigments in birds evolved from violet pigments by a single amino acid change. Proc Natl Acad Sci USA 97:7366–7371

Yokoyama S, Radlwimmer FB (2001) The molecular genetics and evolution of red and green color vision in vertebrates. Genetics 158:1697–1710

Yokoyama S, Tada T (2000) Adaptive evolution of the African and Indonesian coelacanths to deep-sea environments. Gene 261:35–42

Yokoyama S, Tada T (2003) The spectral tuning in the short wavelength-sensitive type 2 pigments. Gene 306:91–98

Yokoyama S, Yokoyama R (1996) Adaptive evolution of photoreceptors and visual pigments in vertebrates. Annu Rev Ecol Syst 27:543–567

Yokoyama S, Zhang H, Radlwimmer FB, Blow NS (1999) Adaptive evolution of color vision of the Comoran coelacanth (Latimeria chalumnae). Proc Natl Acad Sci USA 96:6279–6284