Cone Ossin Genes of African Cichlid Fishes: Tuning Spectral Sensitivity by Differential Gene Expression

Karen L. Carleton* and Thomas D. Kocher†

*Department of Zoology and †Program in Genetics, University of New Hampshire

Spectral tuning of visual pigments is typically accomplished through changes in opsin amino acid sequence. Within a given opsin class, changes at a few key sites control wavelength specificity. To investigate known differences in the visual pigment spectral sensitivity of the Lake Malawi cichlids, *Metriaclima zebra* (368, 488, and 533 nm) and *Dimidiochromis compressiceps* (447, 536, and 569 nm), we sequenced cone opsin genes from these species as well as *Labeotropheus fuelleborni* and *Oreochromis niloticus*. These cichlids have five distinct classes of cone opsin genes, including two unique SWS-2 genes. Comparisons of the inferred amino acid sequences from the five cone opsin genes of *M. zebra*, *D. compressiceps*, and *L. fuelleborni* show the sequences to be nearly identical. Therefore, evolution of key opsin sites cannot explain the differences in visual pigment sensitivities. Real-time PCR demonstrates that different cichlid species express different subsets of the available cone opsin genes. *Metriaclima zebra* and *L. fuelleborni* express a complement of genes which give them UV-shifted visual pigments, while *D. compressiceps* expresses a different set to produce a red-shifted visual system. Thus, variations in cichlid spectral sensitivity have arisen through evolution of gene regulation, rather than through changes in opsin amino acid sequence.

**Introduction**

The evolution of visual spectral sensitivity is a prime example of molecular adaptation (Yokoyama and Yokoyama 1996; Goding and Dean 1998; Yokoyama 2000). Evidence for visual pigment adaptation by amino acid substitution includes the evolution of red and green color vision in primates (Yokoyama and Radlwimmer 1998), the evolution of ultraviolet vision in birds (Wilkie et al. 2000; Yokoyama, Radlwimmer, and Blow 2000), and the blue shift of visual pigments in deep-water fishes (Hunt et al. 1996; Yokoyama et al. 1999). Visual pigments are composed of a light-absorbing chromophore, typically 11-cis retinal, which is bound to and surrounded by an opsin protein (Wald 1968). The absorption spectrum of the chromophore can be modified by amino acid substitution in the opsin protein. Key sites tend to be in one of the seven transmembrane (TM) alpha helices close to the retinal-binding pocket and usually involve changes in amino acid polarity (Yokoyama and Yokoyama 1990a; Chang et al. 1995; Yokoyama and Radlwimmer 1998; Kochendoerfer et al. 1999). Single amino acid changes can cause shifts of between 2 and 35 nm, depending on their location and interaction with the conjugated chain of the chromophore (Asenjo, Rim, and Oprian 1994; Bowmaker and Hunt 1999; Yokoyama 2000).

Five classes of retinal opsins diverged early in vertebrate evolution and include the rod opsin, RH1, and four cone opsins: SWS-1 (ultraviolet sensitive), SWS-2 (short-wave-sensitive), RH2 (mid-wavelength-sensitive and similar to rod opsin), and MWS/LWS (mid-to-long-wavelength-sensitive) (Okano et al. 1992; Hisatomi et al. 1994; Yokoyama 1994, 1995, 1997, 2000; Chang et al. 1995). Each of these opsin classes exhibits a narrow range of spectral sensitivity. Vision over a broader spectral range is accomplished through expression of several opsin genes. The majority of vertebrates have at least one gene from each of the five opsin classes, including birds (Okano et al. 1992), reptiles (Kawamura and Yokoyama 1993, 1995, 1996), and fishes (blind cavefish *Astyanax fasciatus* [Yokoyama and Yokoyama 1990a, 1990b, 1993; Register, Yokoyama, and Yokoyama 1994], goldfish *Carassius auratus* [Johnson et al. 1993], killifish *Oryzias latipes* [Hisatomi, Satoh, and Tokunaga 1997], and zebrafish *Danio rerio* [Vihtelic, Doro, and Hyde 1999]).

We are interested in the visual system of cichlids from the East African Rift lakes. These cichlids represent some of the most species-rich assemblages of fishes in the world. Lake Malawi harbors upwards of 1,000 species which have arisen in the last 1 Myr (Fryer and Iles 1972; Lewis, Reinthal, and Trendall 1986; Kornfield and Smith 2000). There are two major clades of cichlids within Lake Malawi: the brightly colored rock dwellers, which are known by their local name, the “mbuna,” and the more cryptically colored sand dwellers (Meyer et al. 1990).

Breeding for both sand and rock dwellers occurs on large leks, where hundreds of males congregate. Females visit the lek to choose a mate for fertilizing eggs and then leave to mouthbrood the young. Males utilize bright mating colorations when displaying on the lek and contribute only genes to their progeny. The sexual dimorphisms in color pattern and the differential parental investment suggest that sexual selection is an important factor in cichlid mating and perhaps in cichlid speciation (Dominey 1984; Danley and Kocher 2001). Although we do not know all the cues which females use in choosing mates, male color patterns have been shown to be important mate recognition signals, as well as key taxonomic characters (Ribbink et al. 1983; Seehausen, van Alphen, and Witte 1997; Seehausen and van Alphen 1998; Seehausen 1999).
In spite of the important role that visual cues play in cichlid mate choice, little is known about their visual system. Studies of two Malawi cichlid species suggest that their visual system is based on three visual pigments, similar to the visual systems of other African and Neotropical cichlids (Levine and MacNichol 1979; Fernald and Liebman 1980; Fernald 1981, 1984; van der Meer and Bowmaker 1995). The spectral sensitivities of double cones in *Dimidiochromis* (Haplochromis) *compressiceps* have peak absorbances of 536 and 569 nm (Levine and MacNichol 1979). Work in progress indicates that the sensitivity of single cones in this species is 447 nm (Jordan and Loew, personal communication). The double cones of *Metriaclima* (Pseudotropheus) *zebra* have peak absorbances of 488 and 533 nm (Levine and MacNichol 1979). Our recent studies show that *M. zebra* single cones contain an ultraviolet-sensitive pigment with a peak absorbance of 368 nm (Carleton, Hár nosi, and Kocher 2000). The visual systems of these two species differ significantly, with shifts of 79, 48, and 36 nm, respectively, for the three visual pigments. These large spectral shifts suggest that there might be several key differences in opsin amino acid sequence responsible for spectral tuning of each opsin gene.

This study considers the molecular basis of the visual system of these two cichlid species. Their diverse ecology suggests that their visual systems may be subject to selection. *Dimidiochromis compressiceps* is a sand dweller that inhabits the shallow marshy areas. It is a piscivore that is cryptically silver-colored and op-timized for speed and prey capture. *Metriaclima zebra* is a member of the brightly colored mbuna. Males have bright-blue body coloration with vertical black bars (blue-black). Like most of the mbuna, they eat algae, although *M. zebra* are also known to feed on plankton from the water column. To explore the diversity within the mbuna, we also examined *Labeotropheus fuelleborni*, which inhabits shallower waters and almost never feeds in the water column. Males are also bright blue, but without the black bars. To add phylogenetic context and a link to our genomic studies (Kocher et al. 1998), we also included a tilapia, *Oreochromis niloticus*, which is a riverine species which is widely distributed across Africa and is a sister group to the cichlids of the three major East African lakes (Streelman et al. 1998).

In this work, we obtained cone opsin gene sequenc-es from genomic DNA from these four species and also quantified the relative gene expression of cone opsins from retinal cDNA. Surprisingly, our studies suggest that differences in visual sensitivity in the Malawi cich-lid species are not due to spectral tuning through opsin amino acid substitution, but, rather, are the result of changes in the expression of multiple cone opsin genes.

**Materials and Methods**

**Cichlid Specimens**

Individuals of *M. zebra* and *L. fuelleborni* were ob-tained in the field and shipped to our fish facility at the University of New Hampshire. Two *M. zebra* individu-als were used for cDNA sequencing: a female orange blotch (OB) morph and a blue black (BB) male from Nkhata Bay. The female OB was used for genomic opsin sequencing. Three additional males were used for reverse transcription PCR (RT-PCR) analysis of gene ex-pression levels. All *L. fuelleborni* individuals were males from the southeast arm of the lake. One individual was used for genomic DNA sequencing, and two were used for RT-PCR. *Dimidiochromis compressiceps* and *O. niloticus* were obtained from commercial fish breed-ers. One individual of *D. compressiceps* was used for both cDNA and genomic DNA sequencing. Two addi-tional individuals were used for RT-PCR. One specimen of *O. niloticus* was used for genomic DNA sequencing, and three additional individuals were used for RT-PCR.

**Genomic DNA Sequences**

This work extends our previous efforts to sequence the SWS-1 opsin genes from African cichlids and uti-lizes the same species and DNA sequencing strategy (Carleton, Hár nosi, and Kocher 2000). Briefly, a pair of degenerate primers was used to amplify a short region of genomic DNA. This identified the cone opsin genes and provided cichlid-specific sequence for designing new primers. cDNAs were then amplified and sequenced from retinal mRNA using 3′ and 5′ rapid amplification of cDNA ends (RACE). These sequences allowed the design of primers for amplifying and sequencing the complete opsin genes from genomic DNA.

Degenerate primers were used to obtain the se-quence of a short segment of each cone opsin for both *M. zebra* and *D. compressiceps*. The primers matched two conserved regions (amino acid sequences S T Q K A E and Y N P I/V I/V Y) spanning TMs 6 and 7 (Hisatomi et al. 1994; Carleton, Hár nosi, and Kocher 2000). Their locations are shown in figure 1 as OpF and OpR. The first primer discriminates against ampli-fication of the rhodopsin gene, which in fishes has the ami-no acid sequence T T Q R A E. Genomic DNA was extracted from fin clips and amplified in a 50-μl PCR reaction for 35 cycles (94°C for 1 min, 40°C for 1 min, and 72°C for 1 min). The 210-bp product was cloned in pBluescript II SK. Sequences from three to seven clones were compared to determine the consensus sequence for each opsin class. These sequences were used to design gene-specific primers for RACE amplifications.

Sequences for the putative opsin genes were ob-tained from retinal mRNA using RACE (Frohman, Dush, and Martin 1988; Ohara, Dorit, and Gilbert 1989; Hisatomi et al. 1996). Total RNA was extracted from both retinas of a single fish using guanidine isothio-cyanate, precipitated with isopropanol, and resuspended in water. mRNA was isolated using polyT magnetic beads (Dynal) and reverse-transcribed at 42°C for 1 h. First-strand cDNA was polyG-tailed using terminal transfer-ase (Lambert and Williamson 1993) prior to second-strand cDNA synthesis using a polyC primer. This prim-er contained unique sequence for further amplification and a unique restriction site for directional cloning (Carleton, Hár nosi, and Kocher 2000).
RACE was performed in both the 3' and the 5' directions using gene-specific primers. For 3' RACE, the first-strand cDNA was used with a gene-specific forward primer, polyTamp and Tamp. For 5' RACE, the second-strand cDNA was used with a gene-specific reverse primer, polyCamp and Camp (see Carleton, Hárosi, and Kocher [2000] for Tamp and Camp primers). Ten cycles of single-sided amplification were performed using the gene-specific primer (94°C for 20 s, 54°C for 45 s, and 72°C for 2 min) to enhance PCR specificity. An additional 40 cycles of amplification were performed following the addition of Tamp or Camp primers. The products were gel-purified and digested with β-agarase. Either the resulting 3' RACE products were directly sequenced using the RACE primer or a nested primer, or the product was cloned in pBluescript. The longer 5' RACE products were typically reamplified and then cloned for sequencing. The 3' and 5' RACE products overlapped by 80–100 bp. These sequences were easily aligned to assemble the complete cDNA sequence.

Primers for amplifying the opsin genes from genomic DNA were designed from the corresponding cDNA sequences (Primer 3 of Rozen and Skaletsky 1998) with 60°C annealing temperatures. Bases were added to the 5' end of each primer to incorporate a restriction site for cloning (GCGCGGAATTC for forward primers and GCGCGCAAGCTT for reverse primers). The LWS gene was amplified using pairs of primers located to amplify across one intron at a time based on known intron locations in vertebrate opsins. The RH2 gene was amplified as one fragment (using Green F1 and Green R4). The SWS-2A gene was amplified in two overlapping pieces (Blue F3 to Blue R1a, and Blue F1a to Blue R4), while the SWS-2B gene was amplified in three pieces (Blue F3bb to Blue R2b, Blue F2 to Blue R1bb, and Blue F1b to Blue R4b).

Thirty-five cycles of PCR (94°C for 20 s, 50°C or 53°C for 45 s, and 72°C for 2 min) produced single PCR products. These products were gel-purified and the bands cut out and digested with β-agarase. Pairs of cycle-sequencing primers were placed at 400–800-bp intervals so that the entire gene sequence could be assembled from 8–12 sequencing reactions. Figure 1 shows the location of all 37 primers, with amplification primers for the RH2 and SWS-2 genes shown in bold. Following this procedure, the complete coding regions from D. compressiceps were confirmed for each opsin gene by long PCR amplification. This demonstrated that the component sequences amplified as partial gene segments were contiguous in the genome.

In order to estimate the number of unique copies of the RH2 gene, internal primers (Green F2 and R3) were used to amplify a fragment of the gene containing exons 2, 3, and 4 in D. compressiceps. The PCR product was cloned using a TA cloning kit (Promega), and seven clones were sequenced.

Sequence Analysis

The nucleotide and inferred amino acid sequences for the cichlid opsins were used to construct phylogenies including several other fish opsin sequences and those for the chicken. Amino acid sequences were aligned by eye using numerous highly conserved amino acid sites and comparisons with previous alignments (Chang et al. 1995). The amino acid alignment was used to define the nucleotide sequence alignments for the coding regions. Each opsin class had different amino and carboxy termini, creating gaps at the beginning and end of the alignment. These regions were excluded from the phylogenetic analysis. The final data set included nucleotides corresponding to the region from 17 amino acids before the beginning of transmembrane I to 3 amino acids after the end of transmembrane VII.

Phylogenetic relationships of the amino acid sequences were estimated from amino acid distances (where distance = −ln(1 − proportion of amino acid site differences) (Graur and Li 2000) by the neighbor-joining method using PAUP 4.0b4a (Swofford 1999). The significance of each node was assessed from 1,000
bootstrap replicates. Neighbor-joining trees were also built from Tamura-Nei gamma corrected nucleotide distances (Tamura and Nei 1993; Kocher and Carleton 1997). Gamma parameters, including the gamma shape parameter and the proportion of invariant sites, were estimated by maximum likelihood using empirical base frequencies and estimated transition/transversion ratios (PAUP 4.0b4a; Swofford 1999). The maximum-likelihood estimates were made in three ways: from the entire tree, from the fish opsin sequences, and from the fish SWS-2 opsin clade.

Models for the structure of rhodopsin (Baldwin 1993; Baldwin, Schertler, and Unger 1997; Palczewski et al. 2000) were used to identify the amino acid sites which make up the transmembrane regions, as well as the TM sites which point inward toward the retinal binding pocket.

Relative Gene Expression by Real-Time RT-PCR

Real-time RT-PCR was used to quantify the relative levels of mRNA expression corresponding to each of the genomic cone opsin genes (Freeman, Walker, and Vrana 1999). Total retinal RNA (3.6 μg) was reverse-transcribed using a polyT primer and Superscript II (Gibco) at 42°C to create a retinal RT cDNA mixture. Parallel RT-PCR reactions were set up for each of the five opsin genes using the same master mix such that each 30-μl RT-PCR reaction contained equal amounts of the retinal RT cDNA mixture. Gene-specific Taqman primers and probes were then added.

Primers and probes were designed for each gene (LWS, RH2, SWS-2A, SWS-2B, and SWS-1) using Primer Express, version 1.5 (Applied Biosystems; see appendix), to amplify short (60–90 bp) fragments. They were designed from the consensus sequence of all four species. All but the RH2 gene primers could be located such that there were no mismatched bases among the cichlids studied. For the RH2 gene, unique forward primers and probes were designed to match either O. niloticus or the Malawi species (appendix). Probe-binding sites were located across an exon-exon boundary to prevent unwanted signal from genomic DNA. Specificity for mRNA was confirmed by control experiments using total retinal RNA which had not been reverse-transcribed. The probes were 5′-labeled with 6-FAM and 3′-labeled with TAMRA. During PCR, the 5′–3′ exonuclease activity of Taq polymerase released the 3′ TAMRA dye molecule. The relative TAMRA fluorescence was monitored during 40 cycles of PCR on a GeneAmp 5700 Sequence Detection system (Applied Biosystems; 95°C for 15 s, 55°C for 30 s, and 65°C for 1 min). Critical cycle number was determined when the fluorescence exceeded a threshold set close to the background fluorescence. Genes having higher levels of expression have smaller critical cycle numbers than genes with lower expression. Relative gene expression was determined as a fraction of the total cone opsin genes expressed for an individual according to:

\[
\frac{T_i}{T_{all}} = \frac{(1 + E_i)^{C_{ti}}}{\sum (1 + E_i)^{C_{ti}}},
\]

where \(T/T_{all}\) is the relative gene expression ratio for a given gene normalized by the total cone opsin genes expressed, \(E_i\) is the PCR efficiency for each gene, and \(C_{ti}\) is the critical cycle number for each gene. PCR efficiencies were determined from a dilution series (over three orders of magnitude) for each of the five opsin genes for both O. niloticus and M. zebra. Data were collected for 2–10 replicates from each of 2–3 individuals for each species.

Results
Genomic DNA Sequences

The degenerate primer amplifications identified five unique cone opsin genes based on the sequencing of 34 clones from M. zebra and 17 clones from D. compressiceps. These genes were equally prevalent in the clones, with five to eight copies of each gene occurring in the 34 M. zebra clones. The short sequences were phylogenetically similar to the four known cone opsin classes, with one of each of the MWS/LWS, RH2, and SWS-1 classes and two distinct genes for the SWS-2 class. The latter genes will be referred to as SWS-2A and SWS-2B.

RACE amplification yielded cDNA sequences for the LWS, RH2, and SWS-2A genes from D. compressiceps (GenBank accession numbers AF247129–AF247131) and for SWS-2B (GenBank accession number AF317674) from M. zebra. Based on these sequences, primers were designed and used to obtain genomic DNA sequences for all four species studied (D. compressiceps, M. zebra, L. fuelleborni, and O. niloticus). The genomic sequences are deposited in GenBank (SWS-2A, AF247113–AF247116; SWS-2B, AF247117–AF247120; RH2, AF247121–AF247124; LWS, AF247125–AF247128). Including the SWS-1 gene (AF191220–AF191223; Carleton, Hárosi, and Kocher 2000), these species each have at least five complete cone opsin genes in their genomic DNA.

The sequenced genes all contained amino acid sites characteristic of an opsin protein, including lysine for forming the Schiff base linkage with retinal, glutamic acid to serve as the counterion to the protonated Schiff base, two cysteines to form a disulfide bond, an asparagine near the N-terminal end which may be glycosylated, and a serine- and threonine-rich region near the C-terminal end (Hisatomi et al. 1996). The genomic sequences and the cDNA sequences for a given species (when available) match exactly in the coding regions. Comparisons of the cDNA and genomic sequences revealed the exon/intron boundaries which were delimited by the standard GT...AG splicing junctions at locations identical to those of the introns of visual pigments in other organisms (Yokoyama and Yokoyama 1990a, 1990b, 1993; Register, Yokoyama, and Yokoyama 1994).
Sequence Analysis

Amino acid sequences were inferred from the nucleotide sequences. The numbers of amino acid sites which differed between the cichlids were summed for each opsin gene in four ways: (1) the total number of sites which differed, (2) the number of differences occurring within the transmembrane regions, (3) the number of differences at sites directed into the retinal-binding pocket, and (4) the number of sites in the retinal-binding pocket which differed in amino acid polarity (table 1). These data confirm that the amino acid sequences of the Lake Malawi cichlid opsin genes are extremely similar both within the retinal-binding pocket and over the entire gene. Although there are more differences between the Malawi species and <i>O. niloticus</i>, few differences were observed at sites directed into the retinal-binding pocket.

Phylogenetic comparison of the cichlid opsin nucleotide sequences with those of other fish is shown in figure 2. Sequences for <i>L. fuelleborni</i> are essentially identical to those for <i>M. zebra</i> and thus have been left out for clarity. The topology of this tree is similar to that obtained from amino acid sequences and is well supported by the bootstrap values. The relationships within each opsin class are compatible with currently accepted phylogenetic relationships among fishes (Nelson 1994).

The two cichlid SWS-2 genes are closely related to each other, indicating a gene duplication event. The cichlid SWS-2B gene is more closely related to the kilifish SWS-2 gene than to the cichlid SWS-2A gene. This suggests that the gene duplication occurred prior to the atherinomorph/percomorph split over 100 MYA (Kumazawa, Yamaguchi, and Nishida 1999). It is also consistent with preliminary sequences we have obtained for the Central American cichlid <i>Cichlasoma citrinellum</i>, which also has both the SWS-2A and the SWS-2B genes. Therefore, gene duplication must have occurred prior to the breakup of Gondwanaland, which separated African and Neotropical cichlids approximately 100 MYA (Kumazawa, Yamaguchi, and Nishida 1999).

The gene duplication time can also be estimated by calibrating distances to the 10 MYA divergence time for the Malawi species and <i>O. niloticus</i> (Kocher et al. 1995). Gamma-corrected nucleotide distances gave duplication times of 134 MYA (gamma parameter = 1.08, 5.01).
8.9% invariant sites estimated over entire tree), 140 MYA (gamma parameter = 0.65, estimated from fish cone opsin genes), and 133 MYA (gamma parameter = 13.5, 45% invariant sites estimated from fish blue opsin genes only). This suggests an average gene duplication time of 136 ± 8 MYA, which is consistent with a divergence predating the split of killifish and cichlids.

In addition to the SWS-2 gene duplication, we have evidence that there are at least two copies of the RH2 gene. Cloning of PCR products amplified within the coding region for *D. compressiceps* showed evidence for two unique forms of introns 2 and 3, although their amino acid sequences were identical in the coding regions. Sequences from the RH2 gene of the three other species also gave evidence of multiple gene copies.

Relative Gene Expression by Real-Time RT-PCR

Relative opsin gene expression levels are given as percentages of total opsin gene expression. The results for the four cichlid species are shown in figure 3. They are also given in table 2, where they are compared with the visual pigment sensitivities from MSP (see Discussion). Relative levels of gene expression were consistent among individuals within a species over a wide range of absolute concentrations. The PCR efficiency used to calculate relative gene expression levels was 0.9, as determined from the average of data from *M. zebra* and *O. niloticus*. There were no statistical differences in PCR efficiencies for different genes.

There were significant differences in opsin gene expression among the Malawi species. The expression patterns for the mbuna, *M. zebra* and *L. fuelleborni*, were very similar, with the RH2 and SWS-2B genes being expressed the most, followed by the SWS-1 and LWS genes. The SWS-2A gene expression was hardly measurable in these species. In contrast, the patterns observed for *D. compressiceps* were high expression of the RH2 and LWS genes and moderate expression of the SWS-2A gene. The genes preferentially expressed in *O. niloticus* were similar to those of *D. compressiceps* and included the LWS, RH2, and SWS-2A genes. However, the ratios of the LWS and RH2 genes differ considerably. *Oreochromis niloticus* preferentially expresses the LWS gene, while *D. compressiceps* expresses more RH2.

Discussion

Microspectrophotometry has shown that the visual system of Lake Malawi cichlids is based on three visual pigments. The visual pigment sensitivities of *M. zebra* and *D. compressiceps* are quite different, with large sensitivity shifts in both single and double cones (Levine and MacNichol 1979; Carleton, Hárosi, and Kocher 2000; Jordan and Loew, personal communication). Several mechanisms could cause these changes, including shifts in chromophore usage, evolution of opsin sequences, and changes in opsin gene expression. As we discuss below, differences in the visual sensitivity of these cichlids can only be explained by changes in opsin gene expression.

Chromophore Usage

Shifts in the spectral sensitivities of fish visual pigments can occur through changes in chromophore usage.
Red shifts of 10–40 nm can be achieved by changing chromophores from 11-cis retinal (vitamin A$_1$) to 3-dehydroretinal (vitamin A$_2$; Munz and McFarland 1977), with larger shifts at longer wavelengths (Hárosi 1994). Previous measurements of pigment extracts have demonstrated that the rod pigments in Lake Malawi cichlids are A$_1$-based (Muntz 1976). The A$_2$ basis of both rod and cone pigments is supported by more recent MSP studies of these fish (Carleton, Hárosi, and Kocher 2000; Jordan and Loew, personal communication). The use of A$_1$ chromophores is consistent with the stable, highly transmissive photic environment of Lake Malawi (Muntz 1976; Munz and McFarland 1997; personal observation). This suggests that the diversity of visual pigment sensitivities in Lake Malawi cichlids is not the result of shifts in chromophore usage.

Opsin Sequence Evolution

The four African cichlid species studied here have five cone opsin genes, which belong to ancient classes of vertebrate opsins. There is one opsin in each of the SWS-1, RH2, and LWS classes and two opsins in the SWS-2 class. The nucleotide and predicted amino acid sequences of homologous genes are remarkably similar among the Malawi species, differing by two or fewer amino acid substitutions. For the two mbuna species, among the Malawi species, differing by two or fewer amino acid substitutions. The nucleotide and predicted amino acid sequences of homologous genes are remarkably similar among the Malawi species, differing by two or fewer amino acid substitutions. For the two mbuna species, among the Malawi species, differing by two or fewer amino acid substitutions.

There are one or two amino acid site differences for each homologous gene in comparisons between D. compressiceps and the mbuna. Of these differences, only the SWS-2B gene has variations at sites directed into the binding pocket, and only one of these sites represents a change in amino acid polarity. The latter site is A275T, which is equivalent to the A285T difference observed between human green and red opsins (Asenjo, Rim, and Oprian 1994). This change would probably cause a 10-nm blue shift in the mbuna visual pigment relative to D. compressiceps. This is the only significant difference in opsin sequence within the retinal-binding pocket. Otherwise, the opsin sequences of the Malawi species studied are essentially identical. This is convincing evidence that the large differences in spectral sensitivity between D. compressiceps and M. zebra are not the result of opsin sequence evolution.

Opsin Gene Expression

The real-time RT-PCR experiments suggest that differences in visual sensitivities are caused by changes in opsin gene expression. These cichlids differ considerably in the relative expression levels of the five cone opsin genes. The mbuna species predominantly express the RH2 and SWS-2B genes, with weaker expression of the SWS-1 and LWS genes. In D. compressiceps and O. niloticus, the LWS and RH2 genes are expressed along with the SWS-2A gene and a small amount of SWS-2B. In each of the species, several of the gene transcripts are present in large amounts, while others are expressed minimally or not at all (fig. 3 and table 2). In this way, each species expresses a subset of the five available genes to obtain three visual pigments.

By matching the MSP spectral sensitivities with the genes which are maximally expressed (fig. 3) and utilizing the phylogenetic context of the opsin gene class, we can tentatively assign the wavelength sensitivity of each opsin gene (see table 2). In the mbuna, the wavelength sensitivity of the predominant genes are 368 nm (SWS-1), 488 nm (SWS-2B), and 533 nm (RH2). In D. compressiceps, the wavelength sensitivities are 447 nm (SWS-2A), 536 nm (RH2), and 569 nm (LWS). Although these correspondences need to be confirmed by protein expression studies (Oprian 1993; Kawamura and Yokoyama 1998), our data provide a clear explanation for the differences in visual pigment sensitivity. To obtain an ultraviolet-shifted visual system, the mbuna reduce expression of the LWS and SWS-2A genes and utilize the SWS-1 and SWS-2B genes instead. Conversely, to obtain a red-shifted visual system, D. compressiceps utilizes the LWS and SWS-2A genes while minimally expressing the SWS-1 and SWS-2B genes. RH2 gene expression is retained in all species.

This is the first time that differential opsin use has been documented when comparing closely related species. Changes in opsin use have been observed within species at different life stages, particularly for species which undergo metamorphosis (Bowmaker 1995). However, we observed differential opsin use when comparing the adult fish of different species. This is a new mechanism for “tuning” spectral sensitivities of visual pigments. When the opsin gene which is expressed in a given photoreceptor is changed, the resulting visual pigment has an entirely different spectral sensitivity. This mechanism can operate only in species which have more opsin genes than visual pigments.

These measurements are also the first to quantify opsin gene expression. The results raise several questions we would like to address: (1) what is the relationship between the measured gene expression ratios and the distribution of pigments in a retinal mosaic, (2) why have we measured more expressed gene transcripts than have been observed as visual pigments, (3) why are there unused cone opsin genes which are not highly expressed in the retina, and (4) what selective forces might be acting on the visual system to explain the differential gene expression?

Expressed Opsin Gene Ratios and Retinal Mosaics

African cichlids have retinal mosaics composed of a square array of double cones with one single cone in the center (Levine et al. 1979; Fernald 1981; van der Meer, Anker, and Barel 1995). In this pattern, there are two double-cone pairs for each single cone. If all cones express equal amounts of pigment and alternate members of the double cones express a medium- and long-wavelength-sensitive pigment, the expected proportions...
of short-, medium-, and long-wavelength visual pigments would be 20%, 40%, and 40%, respectively. Typically, single cones are smaller than double cones and would express less pigment per photoreceptor, so 20% would be an upper bound. This is consistent with the cichlid expression data, in which the shorter-wavelength transcript comprises only 2.5%–10% of the total cone opsin genes expressed.

If alternate members of the double-cone pairs express one of the longer-wavelength-sensitive pigments (RH2 and LWS for *D. compressiceps* and *O. niloticus*, SWS-2B and RH2 for the mbuna), these genes should be expressed equally. However, all four species show unequal expression of the two longer-wavelength opsin genes. This could be the result of experimental artifacts, such as differential PCR efficiencies. However, it may also result from some fraction of the retina containing double cone pairs with the same visual pigments. This pigment pairing has been observed in the New World cichlids *Cichlasoma longimanus* and *Heterotilapia multispinosa* (Levine et al. 1979). In situ hybridization studies could be performed to test for pigment pairing and to identify its spatial variation.

**Expression of Extra Opsin Genes**

There are additional transcribed messages in the retinal mRNA for which the corresponding visual pigment has not been observed by MSP. The extra transcripts, SWS-2B in *D. compressiceps* and LWS in *M. zebra* and *L. fuelleborni*, may be the result of leaky opsin gene expression. Recent long-PCR experiments indicate that the SWS-2A, SWS-2B, and LWS genes occur in a tandem array, with 4.5 kb separating SWS-2A and SWS-2B and 6 kb separating SWS-2B and the LWS gene (unpublished data). This is consistent with data from cavefish, in which the SWS-2 and LWS genes are just 6 kb apart (Yokoyama and Yokoyama 1993). This close spacing suggests that these genes may be in the same chromatin region. When the chromatin is unwound to express one of the genes, the others may be exposed and available for transcription, but at a reduced rate (Kornberg 1999; Carey and Smale 2000).

There are several possible fates for these rare transcripts. The mRNA may be generated but not translated into a visual pigment and thus not detected by MSP. Alternatively, the resulting visual pigment may be co-located with another visual pigment in a single photoreceptor, but in such a reduced amount as to not be detected by MSP. Dual pigment photoreceptors have recently been observed in mice (Applebury et al. 2000) and humans (Xiao and Hendrickson 2000). Expression of these genes might also be restricted to localized areas of the retina and therefore be difficult to find by MSP.

A recent electroretinogram study of the closely related Lake Malawi mbuna *M. thapsinogen* found evidence for a long wavelength component in the ERG b-wave component (ON-bipolar cells) which is not explained by the three visual pigments detected by MSP (Bilotta and McElroy, personal communication). This is evidence for the presence of an LWS photoreceptor somewhere in the retina of this species and is consistent with the small amount of LWS gene transcript we quantified in the mbuna. Detailed MSP surveys and in situ hybridization studies will be necessary to identify and locate photoreceptors containing weakly expressed opsin genes in Malawi cichlids.

**Unused Opsin Genes**

Each of these species contains opsin genes which are not expressed in the adult retina. These genes appear to be fully functional and indeed do function in closely related species. Unused genes would diverge through mutational processes in approximately 1 Myr (for mutation rates of 10⁻⁶ mutations per gene per year), although complete loss of function can take 30–100 Myr (Force et al. 1999). Since these cone opsin gene classes all arose over 100 MYA, their retention within the genome suggests that these genes are utilized in these fish at some other life stage or in some nonretinal tissue.

Changes in opsin gene expression have been observed at different life stages (Bowmaker 1995). Examples of ontogenetic shifts occur in migratory species such as eels (Beatty 1975) and salmon (Beatty 1966; Beaudet, Novales Flamarique, and Hawryshyn 1997), as well as fish which metamorphose from pelagic to benthic habitats, such as goatfish (Shand 1993), pollack (Shand et al. 1988), and flounder (Evans, Hårosi, and Fernald 1993). These species must all have more opsin genes than are used in the adult stage. We are currently examining the expression patterns of juvenile cichlids to look for evidence that the genes not expressed in adults are utilized in the juvenile phase.

It is also possible that the “unused” opsins are expressed in other tissues for some other function. We have identified the cichlid rhodopsin and pineal-based exo-rhodopsin genes (Mano, Kojima, and Fukada 1999; Philp et al. 2000), which are different from those we report here. The close phylogenetic relationships of the five cichlid cone opsins with other fish cone opsins and their differential expression in the retina of Malawi cichlids strongly suggest that their primary function is vision. Perhaps expression at even a very low level provides enough selection pressure to maintain the genomic integrity of the “unused” cone opsin genes. If these genes have no alternate functions, it is possible that given enough time and further decrease in expression level, the “unused” opsin genes may eventually become pseudogenes in certain species.

**Selective Forces Acting on the Visual System**

Visual pigment adaptation has occurred in these Lake Malawi species through changes in opsin gene expression. The visual capabilities of *D. compressiceps* are quite similar to those of closely related Victorian cichlids (van der Meer and Bowmaker 1995) and the more ancestral Tanganyikan cichlids (Fernald and Liebman 1980; Fernald 1981, 1984). The brightly colored mbuna have a derived pattern of cone opsin gene expression, giving them ultraviolet sensitivity. This sensitivity might provide benefits such as enhanced foraging, navigation,
...and communication (Carleton, Hárosi, and Kocher 2000). In addition, expression of both the SWS-2B (488 nm) and the RH2 (535 nm) genes could enhance color discrimination capabilities in the blue-yellow part of the spectrum. This correlates with the predominance of blues and yellows utilized in male color patterns (McElroy, Kornfield, and Everett 1991; Deutsch 1997).

These new visual capabilities may have arisen by either natural or sexual selection (Endler 1992; Endler and Basolo 1998). The evolution of different visual systems in *D. compressiceps* and the mbuna is not surprising, as they differ in almost every aspect of their ecology. *Dimidiochromis compressiceps* is a piscivorous species which inhabits the sandy reef beds and has subdued blue male mating colors typical of sand dwellers. *Metriaclima zebra* and *L. fuelleborni* are herbivorous rock dwellers, although *M. zebra* is often observed feeding on plankton in the water column. These mbuna both utilize bright blue (UV-reflective) male coloration with contrasting yellow and orange accent colors. To determine which selective mechanisms are acting on the cichlid visual system, we are in the process of examining gene expression and opsin gene sequences for a range of cichlid species which differ by one or more aspects of their ecology. These data, placed in a phylogenetic context, may help us determine how and why selection has acted to change gene expression.

Conclusions

This work establishes the molecular basis of photopic visual capabilities in several cichlid fishes. The Malawi cichlids studied have nearly identical sequences for five classes of cone opsin genes, suggesting that opsin sequence variation is not important in controlling spectral sensitivity. Rather, visual pigment sensitivity is modified by changing patterns of gene expression. It is unclear how this change in regulation is accomplished, whether through mutations in promoters or through modification of other transcription control factors. Work is in progress to examine the regulatory regions in species expressing different complements of cone opsins. In addition, we hope to unravel the selective forces acting to modify gene expression by exploring the relationships between opsin gene usage and ecological traits for a variety of cichlid genera in Lake Malawi.

The high rate of speciation in African cichlids and the large phenotypic differences among species raise the question of how these differences have evolved on such rapid timescales. This work demonstrates that adaptation can occur through changes in gene expression rather than changes in gene sequence. The generality of this observation in cichlid evolution remains to be seen.

Acknowledgments

K.L.C. gratefully acknowledges support from NSF grant IBN9720479 and the UNH Hubbard Marine Research Initiation and Infrastructure Program. Cichlid fish were collected with the support of the National Geographic Society and an NSF DIG grant (DEB-96-23881) to Jeff Markert. The University of Malawi and the Malawi government gave permission to work in Malawi, and Stuart Grant helped ship live specimens. Thanks to Craig Albertson, Pat Danley, Brooke Gowdy, Tyrone Spady, and Todd Steeleman for helpful discussions, as well as assistance with fish collection and dissection. Thanks to Britta Hult for the American cichlid opsin sequencing and long-PCR studies. Special thanks to Ellis Loew and Rebecca Jordan for sharing their unpublished microspectrophotometry results, and to Shozo Yokoyama and several anonymous reviewers for extensive comments on previous versions of the manuscript.

APPENDIX

The primers and probes used for Taqman analysis are listed for each gene in the following order: forward primer, reverse primer, and Taqman probe. The probes were 5’-labeled with 6FAM and 3’-labeled with TAMRA and purchased from Applied Biosystems. SWS-1: GGCTGT GCCGCCAC, AGAGGAGCCGAGCACTTTC, TTT CTGTTGGCAGCAGTGATAC; SWS-2A: TTT GTGTGCCATGCAG, AAGGGACACAGCTCCTAC CAT, AGATCGAAGTTCTCATGTAACACTCGTG; SWS-2B: GGCCTGCACTCCACCTG, GGCACACGG AACACGCACT, TTGGATGGACGCAGGTATACCCA GAAGG, RH2: TCTGTGCAATTGAGGATTC, CC AGGACAAAGTGGACAGAG, TGCCACACTWG GAGGTGAA GTTGC (where W = A for Malawi species and W = T for *O. niloticus*); LCS: CTGCTGACCT CCTGTGTCG, GCCCTCTGGTTGACTCTGACT, TG GCCATCCGTGCTGTGCC.

LITERATURE CITED


