ABSTRACT

Title of Document: FROM GENES TO BEHAVIOR: VARIATION IN THE VISUAL SYSTEMS OF LAKE MALAWI CICHLID FISHES

Adam Ray Smith, Doctor of Philosophy, 2011

Directed By: Dr. Karen Carleton, Department of Biology

Visual systems are ideal models for the study of sensory evolution. The cichlids of Lake Malawi possess an elaborated complex of genes (opsins) that encode chromatic visual pigments, which allows us to study the evolution and diversification of chromatic vision in great detail. In this dissertation, we investigated the molecular and behavioral properties of cichlid visual systems in order to more thoroughly understand the diversification of visual systems and the behavioral consequences of these changes. The work is organized into three research projects, with the following results:

(1) Opsin gene sequence variation, with corresponding functional sensitivity changes, were found for the SWS1 (ultraviolet-sensitive), SWS2B (violet-sensitive), RH2Aβ (green-sensitive), and LWS (red-sensitive) opsin genes. Of the two genera profiled, each had two variable genes, suggesting that diversifying selection acts on different

opsins in each genus. Furthermore, our data suggest that the variation in the SWS1 gene has arisen recently in Lake Malawi and is under rapid selection.

(2) Intraspecific cone opsin gene expression variation was found in wild populations of multiple species. Expression variation was found primarily for the LWS and SWS1 genes, while the other genes were relatively consistent within species. This finding suggests that expression can be modulated by adding genes to what may otherwise be considered a species-specific expression pattern. Quantitative models suggested that this expression variation was not the result of environmental constraint.

(3) Fish raised in different ambient developmental light environments had different cone opsin gene expression, primarily in the LWS opsin gene. These expression differences caused an increase in behavioral sensitivity in the optomotor response. Furthermore, analyses indicated that the OMR response is determined solely by the LWS cone pigment, rather than a complement of different cone types.

Taken together, these findings shed new light on how visual systems diversify over short evolutionary time-scales, and the possible linkage of early determinants of visual sensitivities (opsin genes) and processes that directly influence speciation (behavior).

# FROM GENES TO BEHAVIOR: VARIATION IN THE VISUAL SYSTEMS OF LAKE MALAWI CICHLID FISHES

By

Adam Ray Smith

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2011

Advisory Committee:
Professor Karen L. Carleton, Chair
Dr. Thomas D. Kocher
Dr. Daphne Soares
Dr. Thomas Cronin
Dr. Richard Payne
Dr. David Yager

# Preface

The Harvard Rule of Behavior (paraphrased to an unclear extent): “Under the most carefully controlled experimental conditions, an animal will do whatever it damn well pleases.”

-Originator Unknown

# Dedication

This work is dedicated to all my family and friends who have helped me to get this point today. They include, but are not limited to:

Abbi Smith

Karen Carleton

Terry, Cindi, and Christy Smith

Jeff and Jeri Fickes

My entire academic “family tree,” both past and present

# Acknowledgements

First and foremost, I would like to thank my loving wife Abbi for all of the support she has provided me. I thank my advisor Karen for her patience, understanding, and insight, along with her willingness to accept the challenge of training me. Thanks to my advisory committee for all of their efforts – Tom Kocher, Daphne Soares, Richard Payne, Tom Cronin, and Dave Yager. Finally, the various chapters in this work could not have been completed without the help of many other people, particularly those in the Carleton lab. I would like to acknowledge my numerous coauthors, along with the other individuals who helped develop each chapter.

**Chapter 2** – This work was supported by NSF grant IOS0841270 and the University of Maryland. Thanks to Tyrone Spady for providing samples of *Labidochromis*. Special thanks to Chris Hofmann and Kelly O'Quin for technical assistance and discussions of experimental design. Special thanks as well to Ad Konings for generously providing pictures of fishes. We thank all members of the Kocher Lab at UMD as well as two anonymous reviewers for helpful comments on this manuscript.

**Chapter 3 -** We would like to thank Tom Kocher, James Maluza, Richard Zatha, Kelly O'Quin, Brian Dalton, and Jennifer Ser for help with fish collection. Thanks also to the cichlid labs at University of Maryland for comments on this manuscript. This work was supported by the National Science Foundation (IOS-0841270) and the University of Maryland.

**Chapter 4 -** We thank Thomas Cronin for help with the spectrometer calibration. Also, we thank Abbi Smith for assistance with rig construction. Thanks to Thomas Kocher, David Yager, Richard Payne, Thomas Cronin, Moira van Staaden, and Martine Maan for helpful comments on the manuscript. This work was supported by NSF grant IOS-0841270 (to KLC) and a Guy Jordan Endowment grant from the American Cichlid Association (to ARS).

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1:

# The Evolution of Chromatic Visual Systems: An Introduction to the Dissertation

*Evolution and models of speciation*

Evolution is, arguably, the single most important topic in the biological sciences. As Dobzhansky (1973) famously stated, "Nothing in biology makes sense except in the light of evolution". In particular, evolutionary processes such as selection and genetic drift can contribute to species divergence and the development of new species. Different evolutionary forces may be observed as driving speciation events over ancient and recent timescales. By looking at variation across a large number of distantly-related taxa, it is possible to uncover broad, fundamental patterns of trait evolution that underlie adaptation (Dominey 1984; Yokoyama & Yokoyama 1996; Yokoyama & Radlwimmer 2001). In the cases of relatively young rapid species radiations, it may even be possible to apply these models to accurately pinpoint the specific origin of speciation traits (Seehausen *et al*. 1998; Streelman & Albertson 2006). As such, phylogenetic analyses across various timescales have been key to the development of evolutionary theory and for the study of speciation in particular.

Speciation can occur in either sympatry or allopatry depending on the specific barriers to gene flow. Allopatric models of speciation are generally very intuitive. Species, separated by large distances and geographical obstacles, diverge in response

to adaptation in different habitats or through genetic drift (Wright 1932). However, sympatric models are more contentious, and even had a noteworthy opponent in Ernst Mayr (1982). Due to the relative environmental homogeneity in a single habitat patch and high gene flow within a population, speciation in sympatry can be difficult without strong divergent selection. One such strong selective mechanism facilitating sympatric speciation is sexual selection. Divergent mate preferences, with or without ecological or habitat segregation, could result in rapid species diversification in sympatry (Fisher 1915; Lande 1981).

*Sexual selection and the visual system*

In order for an individual to choose a mate, it must first assess potential candidates. Displaying individuals either actively or passively emit sensory information to attract and solicit potential mates. Visual courtship displays are quite common, and are employed by a broad range of both vertebrates (Baerands & Baerands-van Roon 1950; Borgia 1986) and invertebrates (Rovner & Barth 1981; Hanlon *et al*. 2005). In particular, colorful male ornamentation is a common means of attracting the attention of females (Kocher 2004). Visual signals are attractive research models not only because they are easily observed, but also because of the wealth of knowledge regarding how these signals are detected from receptor sensitivities to the resulting animal behaviors.

The perception of chromatic signals is determined by both the transmission qualities of the signaling environment and the sensory capabilities of the receiver. Because of this, a broad range of hypotheses regarding the selective pressures that shape sensory systems and subsequent mate selection have been posited (Endler

1993). The sensory drive hypothesis is one such model (Boughman 2002; Seehausen *et al*. 2008). In essence, this model proposes that environmental selection imposes constraints on the sensory capabilities of individuals. Male signals then evolve to match these selected sensory characteristics, effectively linking female choice and male signal via environmental constraint. This model is particularly relevant in aquatic habitats due to the differential attenuation of various wavelengths of light traveling through water (Seehausen *et al*. 2008).

*Genes to behavior: bridging the gap*

In order to investigate the sensory drive hypothesis, it is useful to subdivide the broader hypothesis into three primary facets: (i) selective pressures mediated by habitat transmission (ii) evolutionary adaptation of sensory systems, and (iii) co-evolution of female sensory capabilities and corresponding male display traits. These facets can be tested at multiple levels, ranging from the molecular basis of receptor sensitivities, the peripheral and central neural pathways that process sensory information, and the behavioral responses derived from sensory processes. Many studies have tested for adaptation of these three aspects in various species radiations, and found particularly compelling evidence for sensory drive adaptations in teleosts (Fuller *et al*. 2004; Fuller *et al.* 2005; Seehausen *et al*. 2008; Hofmann *et al*. 2009).

For the purposes of this dissertation, we will subdivide the experimental investigation of sensory drive into three testable hypotheses: (i) changes in receptor gene sequence will be observed across species with varying ecologies and nuptial colorations, such that long- or short-wavelength shifts in spectral absorbance compared to a reference species will have long- or short-wavelength shifted nuptial

signals, (ii) cone opsin expression will vary with changes in ambient light environment (depth), with fishes from deeper waters with a more constrained ambient light spectrum expressing less LWS opsin and with lower quantitative variance in gene expression than conspecifics from shallow waters, and (iii) shifts in gene expression will result in quantitative changes in behavioral thresholds, such that fishes expressing relatively more LWS opsin gene will be more sensitive to long-wavelength stimuli while fishes expressing more RH2A and B opsin genes will be more sensitive to shorter-wavelength stimuli. Hypotheses (i) and (ii) have both been investigated in other systems with some substantial success (Fuller *et al*. 2004; Terai *et al*. 2006). While the latter hypothesis is intuitive enough, it is deceptively difficult to demonstrate, although evidence for a link between opsin gene sequence and specific behaviors is available for sister cichlid taxa from Lake Victoria (Maan *et al*. 2006).

The difficulty in demonstrating the third hypothesis is partially due to the information processing steps that occur in the nervous system. As the nervous system processes incoming sensory signals to generate appropriate behavioral responses, it segregates different components of the total signal into information streams. Many behaviors only utilize a small subset of the total information available, as has been demonstrated in the case of the visual system (Ewert 1984). For example, the optomotor response (OMR) behavioral paradigm is known to function through a luminosity-driven visual pathway (Schaerer & Neumeyer 1996). Furthermore, the OMR is mediated by medium- and long-wavelength sensitive cones, with short-wavelength sensitive cones being excluded from the pathway (Orger & Baier 2005).

Therefore, the OMR paradigm is useful for modeling the brightness response of animals under various lighting environments, but is inappropriate for studies that attempt to target contrast-based behaviors. The OMR demonstrates how the construction of even basic models relating genes and behavior requires an integrative experimental and conceptual approach accommodating the molecular, cellular, and neuronal properties of the investigated network.

*Cichlids, visual traits, and sensory drive*

The cichlids of Lake Malawi are the most speciose extant vertebrate radiation (Kornfield & Smith 2000). This monophyletic species flock is a model system for speciation via sexual selection due to the rapid diversification of male nuptial coloration within genera (Danley & Kocher 2001). Prior research has demonstrated that females can identify and select potential mates based on coloration alone, lending further support to the sexual selection hypothesis (Seehausen & van Alphen 1998; Couldridge & Alexander 2002; Kidd *et al*. 2006). Furthermore, the cichlids of Lake Malawi have one of the most elaborate visual pigment gene complexes known in vertebrates (Spady *et al*. 2006; Hofmann *et al*. 2009). Multiple gene duplication events have resulted in seven chromatically-distinct cone opsin genes: SWS1 (ultraviolet), SWS2B (violet), SWS2A (blue), RH2B (blue-green), RH2Aα and RH2Aβ (green), and LWS (red). At least six of these genes are commonly expressed, with RH2Aβ being a likely exception. Taken in concert, these data suggest that the cichlids of Lake Malawi may be an ideal model for the study of the sensory drive hypothesis.

*How and why: a brief outline of the following chapters*

This dissertation was structured to investigate both the molecular and behavioral facets of the sensory drive model. In essence, we have approached sensory processing and integration using an "outside-in" approach. The sensory processing cascade essentially consists of the following processes: (i) collection of the signal by the receptor cells, (ii) information segregation in early neuronal pathways, (iii) information processing and re-integration in higher-order neural processing regions, and (iv) the behavioral output of the organism. In chapters one and two, we used molecular techniques including gene sequencing and expression of cone opsin genes to investigate the fundamental receptor substrate in sensory step (i). In chapter three we performed a behavioral study using the optomotor response to describe cone inputs into a fundamental behavioral process such as in sensory step (iv). By taking measures at these levels, we are able to make specific predictions regarding the quantitative relationship between steps (i) and (iv), and we can use this information to generate hypotheses regarding the neuronal pathways that underlie steps (ii) and (iii). With this information, we will formulate a hypothesis regarding the cellular underpinnings of female signal identification and the implications for male trait evolution.

# Chapter 2:

# Allelic Variation in Malawi Cichlid Opsins: A Tale of Two Genera

Adam R. Smith & Karen L. Carleton

See Appendix 1 for all supplemental materials referenced in this chapter

This chapter is slightly modified from an article published as:

Smith AR, Carleton KL (2010) Allelic variation in the Malawi cichlid fish opsins: a tale of two genera. Journal of Molecular Evolution 70: 593-604

Reproduced under license number 2697210289516 from Springer

*Abstract*

The role of sequence variation in the spectral tuning of color vision is well established in many systems. This includes the cichlids of Lake Victoria where sequence variation has been linked to environmental light gradients and speciation. The cichlids of Lake Malawi are a similar model for visual evolution, but the role of gene sequence variation in visual tuning between closely related species is unknown. This work describes such variation in multiple species of two rock-dwelling genera: *Metriaclima* and *Labidochromis*. Genomic DNA for seven cone opsin genes was sequenced and the structure of the opsin proteins was inferred. Retinal binding pocket polymorphisms were identified and compared to available data regarding spectral absorbance shifts. Sequence variation with known or potential effects on absorbance spectra were found in four genes: SWS1 (UV-sensitive), SWS2B (violet-sensitive), RH2Aβ (green-sensitive), and LWS (red-sensitive). Functional variation was distributed such that each genus had both a variable short-wavelength and long-wavelength sensitive opsin. This suggests spectral tuning is important at the margins of the cichlid visual spectrum. Further, there are two SWS1 opsin alleles that differ in sensitivity by 10nm and are > 2 MY divergent. One of these occurs in a haplotype block > 1kb. Potential haplotype blocks were found around the *RH2* opsin loci. These data suggest that molecular diversification has resulted in functionally unique alleles and changes to the visual system. These data also suggest that opsin sequence variation tunes spectral sensitivities between closely related species and that the specific regions of spectral tuning are genus-specific.

*Introduction*

African cichlids are a model for understanding speciation (Kornfield & Smith 2000; Danley & Kocher, 2001; Genner & Turner 2005). Lakes Malawi, Victoria, and Tanganyika all house large flocks of endemic species. The flocks in Malawi and Victoria are each composed of large radiations of closely-related haplochromine cichlids that have evolved at least partially through hybridization (Seehausen 2004). These radiations are an excellent natural laboratory for the study of speciation because of the large numbers of species that have arisen in such a short period of time.

Numerous models have been proposed for how species are formed, and the possibility or probability of each is a matter of some contention (Kirkpatrick & Ravigne 2002). The classic model of allopatric speciation requires extrinsic (geographic) barriers to gene flow (Mayr 1982). However, there is also evidence for parapatric and sympatric speciation, in which divergence is driven primarily by ecological (Schluter 2001) or sexual selection (Zahavi 1975; Weatherhead & Robertson 1979; Lande 1981). Sympatric speciation is a particularly attractive explanation for the recent rapid radiation of haplochromine cichlids (Seehausen & van Alphen 1999).

Cichlid visual systems have been the subject of intensive study because of their direct link to mate identification and preference (Maan *et al.* 2004; Maan *et al.* 2006). Visual cues appear to be sufficient for conspecific mate recognition (Seehausen & van Alphen 1998; Kidd *et al.* 2006), and may be a metric for intraspecific mate choice (Maan *et al.* 2004). The degradation of the visual

environment in Lake Victoria has been linked to the breakdown of visual mate recognition systems and species isolation (Seehausen *et al.* 1997). Behavioral studies have demonstrated that sister species of Lake Victorian cichlids have different chromatic sensitivities, and that male nuptial coloration tends to match these sensitivities (Maan *et al.* 2006). Therefore, the inherent sensitivities of cichlid visual systems may determine mate choice and contribute to speciation (Kawata *et al.* 2007; Seehausen *et al.* 2008).

Visual sensitivities are determined by the photoreceptor visual pigments. These are comprised of a chromophore (typically 11-cis retinal) bound to an opsin protein (Palczewski *et al.* 2000). The nature of the opsin protein sets the absorption properties, and is therefore a fundamental source of functional variation in visual systems (Yokoyama & Yokoyama 1996). Amino acid changes at multiple sites can alter the characteristic spectral absorbance of the protein (Takahashi & Ebrey 2003; Sugawara *et al.* 2005). A relatively small number of sequence changes can result in a functionally novel opsin class (Yokoyama & Radlwimmer 2001). This characteristic is especially important in the context of color vision (Carleton *et al.* 2000; Spady *et al.* 2005). Gene duplication and divergence has led to the evolution of new opsin classes with unique visual sensitivities in mammals (reviewed by Ebrey & Koutalos 2001), birds (reviewed by Ebrey & Koutalos 2001), crustaceans (Porter *et al.* 2009), and teleosts (Hofmann & Carleton 2009). Evolution of opsin coding sequences has led to the diversity of opsin chromatic sensitivities that are observed in organisms today.

East African cichlid fishes are an ideal system for the study of functional variation in cone opsins (Carleton *et al.* 2006). Cichlids have seven unique cone opsin genes that have arisen through a series of gene duplications (Spady *et al.* 2006). Variation among species in the expression of these gene duplicates causes important changes in visual sensitivity among species in Lake Malawi (Parry *et al.* 2005). These fishes are diverse in male nuptial coloration, and are thought to be actively speciating (Danley & Kocher 2001). Functional diversity in cone opsin genes that cause differences in female visual sensitivity is a possible explanation for the divergence of female preferences for different male coloration.

Hofmann *et al.* (2009) proposed that, while shifts in visual sensitivity in the intermediate spectral range (~ 410-520nm) can be accomplished by altering opsin gene expression, spectral shifts at the ends of the spectral range can only be achieved by varying opsin sequence. In this study, we sequenced all seven known cone opsin genes in several species from two species rich genera: *Metriaclima* and *Labidochromis*. Our hypothesis was that sequence variation altering the amino acid residues at sites in the retinal binding pocket would be found in the SWS1 and LWS opsins as these proteins are at the extreme ends of the cichlid opsin spectral range. Conversely, we also hypothesized that the remaining five opsin genes would be conserved at these functional sites due to the potential for expression shifts amongst these genes to act as a tuning mechanism.

## *Materials and Methods*

### *DNA Samples*

Twenty-four fish from the genus *Metriaclima* were lab reared from captive-bred stocks with a known African geographical source. Caudal finclips were taken from three individuals from each of the following eight species: *M. barlowi, M. benetos, M. callainos, M. fainzilberi, M. lombardoi, M. mbenji, M. phaeos,* and *M. pyrsonotus*. Samples from six species of *Labidochromis* were taken from a mixture of lab and wild caught individuals. *L. caeruleus* and *L. chisumulae* samples were prepared from abdominal muscle tissue taken from three frozen individuals that were originally procured from Old World Exotic Fish (FL, USA). Finclips were taken from the following wild-caught fishes at Lake Malawi: one *L. flavigulus*, two *L. gigas*, two *L.ianthinus*, and three *L. vellicans* (Table S2-1). Genomic DNA was extracted using a Qiagen DNeasy kit (Valencia CA, USA).

*Gene amplification and sequencing*

Gene-specific primers were used to amplify each of the seven cone opsin genes: SWS1 (UV sensitive), SWS2b (violet sensitive), SWS2a (blue sensitive), RH2B (blue-green sensitive), RH2Aα (green sensitive), RH2Aβ (green sensitive), and LWS (red sensitive). Most genes were amplified in either two or three fragments. These fragments overlapped to produce one contiguous sequence for all but the RH2B gene, where we did not amplify and sequence across the 1.5kb of introns two. PCR products were purified using a Qiagen QIAquick kit (Valencia, CA, USA) and then cycle-sequenced using internal primers (Table S2-2) and the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA). Sequencing was performed on either a 3730xl DNA Analyzer or a 3100 Genetic Analyzer (Applied Biosystems).

*Sequence assembly and analysis*

Sequence assembly was performed using the Sequencher 4.7 software (Gene Codes Corporation, Ann Arbor, MI, USA). For a few *Labidochromis* sequences, long intronic repeats within the LWS or SWS2b genes made assembly difficult. Sequences were aligned to a genomic reference sequence from *M. zebra* and the repeat was excised to allow further analysis. Coding sequences were identified by comparison to previous sequenced cichlid opsin cDNAs and amino acid sequences were inferred. For phylogenetic analyses, gene and protein alignments of large species panels were performed using the LINSI protocol in MAFFT (European Bioinformatics Institute). Bootstrap trees were generated in PAUP (Swofford 2002), and distance trees were created using a general time reversible (GTR) model in GARLI v0.951 (Zwickl 2006; www.bio.utexas.edu/faculty/antisense/garli/Garli.html).

Haplotype blocks were analyzed using the DnaSP 5.10 software package (Librado & Rozas 2009). DnaSP 5.10 computationally separated the provided diploid sequence data to estimate the sequence structure of haploid alleles. Recombination sites were identified for both the *SWS1* locus and the tandem array of the *RH2* genes. Haplotype blocks were then defined as continuous regions of fixed sequences between these recombination sites and mapped onto the amino acid sequences.

*Allele identification*

The cDNA sequences were compared to previously sequenced opsins from cichlids in Lakes Malawi, Victoria, and Tanganyika (Table S2-3). The Nile tilapia (*Oreochromis niloticus*) was used as a standard outgroup for all gene families. All known visual pigment spectral shifts involve amino acid sites directed into the retinal

binding pocket which change amino acid polarity (Chang *et al.* 1995; Takahashi & Ebrey 2003; Yokoyama 2008). Cichlid cDNA sequences were translated to amino acids and aligned with bovine rhodopsin to identify the corresponding retinal binding pocket sites from the bovine rhodopsin crystal structure (Palczewski *et al.* 2000). (Note: we use these corresponding bovine rhodopsin site numbers for all cichlid opsin amino acid locations). Sites were annotated as to whether they were in the retinal binding pocket close to 11-cis retinal, in the transmembrane regions, or outside of these regions (see Carleton *et al.* 2005b; Figure 3 for site locations). Next amino acid changes were examined to determine whether they involved a change in polarity. Only sites that were both in the retinal binding pocket and involved a change in polarity were considered to have the potential for causing functional differences. Functional alleles were then identified for individuals and species (*e.g.* the *M. zebra* allele) and were classified by fixed residues at these retinal binding pocket sites. In addition, actual functional differences in alleles were estimated by comparisons between sequences from species with known spectral sensitivity determined by microspectrophotometry (MSP) or protein expression (Parry *et al.* 2005; Jordan *et al.* 2006).

*Results*

All of the genomic sequences were deposited in Genbank and the accession numbers are listed in Table S2-3. The RHB gene contains a long second intron (~1.5kb) that we did not sequence across. Therefore, the sequences for this gene were submitted in two segments that correspond to the regions both upstream and downstream of this intron (Table S2-3).

Sequences from the seven cone opsin genes revealed significant differences between genes in the degree of overall sequence variation. Three genes showed no likely functional variation while the other four displayed significant variation with variable sites likely to contribute to functional differences in visual pigment spectral sensitivity. The following genes could not be assembled due to an inability to obtain sufficient unambiguous sequence: one SWS2a (*M. benetos*), two RH2Aβ (*L. flavigulus* and *L. ianthinus*), two RH2Aα (*L. vellicans* and *M. benetos*), and one LWS (*L. vellicans*).

*Genes with no functional allelic variation*

Three genes lacked any amino acid sequence polymorphisms likely to cause functional shifts: SWS2a, RH2B, and RH2Aα (Figs. 2-1 and 2-2). The SWS2a gene was invariant in *Labidochromis*, and displayed no variation across the several species in the genus (Fig. 2-2). In *Metriaclima*, SWS2a had no variation at retinal binding pocket sites. Therefore, SWS2a had no sequence variation predicted to have functional significance. RH2Aα had a few variable transmembrane sites in each genus. However, none of these sites were in the retinal binding pocket and none involved changes in polarity. Therefore, RH2Aα also had no functionally significant variation. The RH2B gene had several variable transmembrane sites, but only one has a high likelihood of being functionally significant. Both genera show variation at M44I, a retinal binding pocket site that involves a change in amino acid polarity. *Labidochromis* also shows polymorphism at A124S. However, although these sites are present in or near the retinal binding pocket, neither of these polymorphic sites are associated with known absorbance shifts in other systems (reviewed by Takahashi &

Ebrey 2003). In particular, the A124S polymorphism (~8 angstroms from retinal) is two sites away from site 122 (~5 angstroms from retinal), which is known to cause shifts in the spectral sensitivity of rhodopsin pigments (Yokoyama *et al.* 2008). Because it is two-thirds of an alpha-helical turn away from a known tuning-site, it is unlikely that the A124S polymorphism faces in a direction that would result in interactions with the bound retinal. Therefore, *RH2B* is unlikely to show functional variation.

We discovered pseudogene alleles for two of the opsin genes. *M. fainzilberi* and *M. mbenji* were found to be polymorphic for an SWS2a allele that would result in a slightly truncated gene, with a stop codon being inserted eight residues prior to that in the normal sequence. Expression of the SWS2a gene has not been detected in any Malawi rock-dweller, so this truncated allele is not likely to have functional significance (Hofmann *et al.* 2009). All three individuals of *L. caeruleus* were homozygous for an RH2Aα allele in which an insertion disrupts the splice site for the third intron. A stop codon is then present in the sequence prior to the beginning of the normal fourth exon.

*SWS1*

In *Metriaclima*, fourteen nonsynonymous changes were found in SWS1 sequences, (Fig. 2-1). Eight of the polymorphisms were present in the transmembrane regions of the protein, with four in retinal binding pocket sites. All four caused a change in polarity. Based on these polymorphisms, two functional alleles were identified. One corresponds to the *M. zebra* allele which absorbs maximally (λmax) at 368nm, while the other has a λmax of 378nm and was previously described in

*Pseudotropheus acei* (Parry *et al.* 2005). These two alleles differ at four key retinal binding pocket sites at amino acid positions 83, 114, 160, and 204. The *M. zebra* allelic class (henceforth referred to as UV368) is defined by a GSTT sequence at these sites, whereas the *P. acei* allele (henceforth referred to as UV378) is defined by SAAI. Most of the *Metriaclima* have the UV368 allele. However, *M. callainos* and *M. phaeos,* as well as all of the published sequences for SWS1 from Lake Victoria cichlids, possess UV378 (Fig. 2-3). Although the *Labidochromis* are polymorphic for seven nonsynonymous changes in the gene, none of these are binding pocket polymorphisms (Fig. 2-2). *Labidochromis* are all fixed for UV368 and show none of the functional variation found in *Metriaclima*.

*SWS2b*

SWS2b showed much less variation than SWS1. In *Labidochromis*, two nonsynonymous polymorphisms were found in the violet-sensitive genes. Both cause a change in polarity in a transmembrane region (Fig. 2-2). Of these sites, only the A269T polymorphism is present in the retinal binding pocket. This polymorphism is know to cause a functional change, and is associated with an 11nm shift in the cottoid fishes of Lake Baikal (Cowing *et al.* 2002). The T269 allele was found only in *L. flavigulus*, with all of the other *Labidochromis* being fixed for A269. All of the *Metriaclima* were fixed for a variant of the A269 allele (Fig. 2-1).

*RH2Aβ*

The RH2Aβ gene showed extensive diversity. The *Metriaclima* had 17 variable sites in this gene (Fig. 2-1). Eight of the sites were in transmembrane regions. However, only two were in the retinal binding pocket, with one of those

resulting in a change in polarity. Although this gene displays a large amount of amino acid variation, the only polymorphism in a protein region likely to be functionally significant is M183L. Yokoyama *et al.* (2008) identified residue changes at position 183 as a determinant of λmax in opsins from the RH class. This residue is present in an extramembrane loop that penetrates the retinal-binding pocket of the opsin protein and is in close proximity to the bound chromophore (Palczewski *et al.* 2000). The shift from a residue containing sulfur (M) to one without it (L) at this site likely has a functional effect. Therefore, RH2Aβ was considered to have two primary functional alleles. However, current MSP data for variation at this site in cichlids is insufficient to make a definitive statement regarding the magnitude of this λmax shift. While most of the *Metriaclima* are fixed for the M183 allele, all three *M. lombardoi* and one *M. barlowi* have the leucine residue at site 183.

The *Labidochromis* had ten nonsynonymous changes in the RH2Aβ gene (Fig. 2-2). However, the large amount of polymorphism in this genus does not include any variation in the retinal binding pocket. All of the *Labidochromis* possess a variant of the M183 allele. Therefore, this gene does not show any functional variation in *Labidochromis*.

*LWS*

The *Labidochromis* had a total of seven nonsynonymous polymorphisms. Three would result in binding pocket polarity shifts (Fig. 2-2). Nine polymorphic sites were found in *Metriaclima*, with only one of those causing polarity shifts in the binding pocket (Fig. 2-1). Both genera had alleles corresponding to the H (A164) and M2 (S164) alleles described by Terai *et al.* (2006). These alleles differ by a S164A

polymorphism, which causes a 7nm shift in humans (Asenjo *et al.* 1994). However, expressed proteins of cichlid alleles differing at this site did not show spectral shifts when combined with type A1 chromophore (Terai *et al.* 2006). Since Malawi cichlids primarily use A1, this site may not be functionally significant in the cichlid LWS background.

*L. caeruleus* is polymorphic for Y261F. This substitution is associated with a 10nm shift in humans (Asenjo *et al.* 1994), but MSP data are not available for cichlids or other fishes in general. The F261 variant was found only in H type alleles, while the Y261 variant is present in both the H and M2 allelic classes.

*Haplotype blocks around opsin genes*

The observed pattern of association between amino acid polymorphisms varied both between genes and between genera as well. None of the genes showed clear haplotypes of associated polymorphisms in *Labidochromis*. The largest continuous differentiated block within the *Labidochromis* is in the *RH2Aβ* locus and extends for roughly 300 bases, which includes amino acid positions 56-162 (Fig. 2-4). Otherwise, the high proportion of heterozygosity and singleton polymorphisms in the genus disrupted any obvious physical structure.

The *Metriaclima*, however, displayed haplotypes in both the *SWS1* gene and the *RH2* tandem array. In the *SWS1* locus, the extent of variation was clearly correlated with the functional allele type. The UV378 allele seems to be highly conserved in the small number of individuals in which it was found, with no residue polymorphisms after the 21$^{st}$ residue in the protein. This continuous block extends for the remainder of the gene including the introns, which corresponds to ~1kb (Fig. 2-5).

The UV368 alleles were variable at several sites throughout the gene. Of the 14 polymorphic sites present in the SWS1 gene, the UV368 alleles only shared one common polymorphism (I165) aside from the four previously mentioned functional sites (G83, S114, T160, and T204). Perhaps most importantly, there is not a continuous block across these four sites. Sites 103 and 201 are polymorphic due to recombination, meaning the longest continuous block in the UV368 allele extends for ~380bp when the intron is included (Fig. 2-5). This haplotype structure corresponds well with the large genetic distance observed between the various UV368 alleles (Fig. 2-3). However, it is important to note that the long haplotype structure in the UV378 alleles may be the result of the relatively small number of individual (n=6) in which this allele was observed.

The RH2 genes display a potential large-scale haplotype structure that extends over a much larger portion of the genome than that observed in *SWS1*. The three *RH2* genes form a 26kb tandem array (*RH2B – RH2Aα - RH2Aβ*; Carleton *et al.* 2010). Corresponding blocks at the margins of these *RH2* genes suggest that haplotype blocks may extend through the intergenic spaces in this array, although these regions were not sequenced independently. There are two different haplotype blocks that center around the RH2Aα gene. One is based on the *RH2Aα* locus of *M. barlowi* which is correlated with a conserved block of 500bp in the *RH2B* gene. These genes are ~15kb apart, suggesting that our conserved sequences may represent the edges of a very large haplotype block. Similarly, the *RH2Aα* locus found in two *M. lombardoi* is a continuous haplotype block that is correlated with the first half of the otherwise highly-variable *RH2Aβ* gene. The genomic distance between these genes is ~10kb

(Fig. 6). This linkage breaks down approximately 300bp into the *RH2Aβ* coding sequence before the L183M polymorphic site. These possible haplotype blocks are visually apparent in the sequence, and they were confirmed computationally using DnaSP.

*Gene trees and genetic distance between alleles*

Gene trees were constructed using the sequences described in this work as well as several reference sequences from cichlids of Lakes Victoria and Tanganyika and the riverine cichlid *Oreochromis niloticus*. The topology of the SWS1 gene tree indicates that the split between the UV378 and UV368 alleles is quite old, and probably dates to very early in the Malawian radiation. The total branch length within the UV368 clade is roughly threefold greater than that in the UV378 clade, even when the Victorian variations of UV378 are included (Fig. 2-3). By assuming a model of linear molecular evolution, we can estimate the approximate genetic divergence time of these two alleles. This uses the genetic distances between these two alleles as well as those to the Nile tilapia SWS1 sequence and a range of divergence times previously estimated for Malawi and tilapia of 20 (Genner *et al.* 2007) or 45 MY (Genner *et al.* 2007; Azuma *et al.* 2008). Using these, we find that these two alleles diverged between 2.3 and 5.3 MY ago. This is close to the age of the lake, and dates back to the basal radiation of the Malawian cichlids (Genner *et al.* 2007). While possible rapid diversification of SWS1 alleles could skew our linear estimates of divergence times, the concordance of the allelic split with the early radiation of the Malawian/Victorian cichlid clade is consistent with our data.

The tree for RH2Aβ shows significant variation in divergence between alleles. Some branches are much longer than others, such as that leading to the L183 allele of only 2 species. This allele is clearly derived within the Malawian flock (Fig. 2-7). The long branch lengths within this L183 clade might suggest that this allele is rapidly evolving. Notably, the distance between the M183 alleles in Victoria and Malawi is actually less than that observed between the terminal L183 allele of *M. lombardoi* and the M183 clade (Fig. 2-7). This suggests either rapid diversification or pseudogenization of the L183 allele of the RH2Aβ gene.

The LWS gene shows a complex, intertwined relationship of sequences between Lakes Victoria and Tanganyika, but appears to be monophyletic within Malawi (Fig. 2-8). The M2 allelic clade is nested within the monophyletic H grouping. These two alleles differ by S164A, which suggests that the S164A polymorphism arose independently in the M2 alleles in Malawi and in Victoria. Also, the F261 alleles are not separated from the general H clade by any appreciable distance, indicating either a lack of selection for the allele or a very recent origin.

*Discussion*

Cichlid opsin genes have been examined at the population level for Lake Victoria cichlids in previous studies (Terai *et al.* 2006; Seehausen *et al.* 2008). A few studies have examined single individuals within species from Lake Malawi. However, this is the first study to examine multiple individuals from multiple species within two important Lake Malawi cichlid genera, *Metriaclima* and *Labidochromis*. As such, it provides new insight into the degree of both inter- and intraspecific opsin sequence

variation in these fishes. Such variation could be important for cichlid mate choice and cichlid speciation.

Opsin sequence polymorphisms with known or potential functional effects were found in four of the seven genes studied. Of these four genes, two were found to have functional variation in *Metriaclima* (SWS1 and RH2Aβ) and two in *Labidochromis* (SWS2B and LWS; examples illustrated in Fig. 2-9). The variable sites in genes SWS1, SWS2B, and LWS are all associated with absorbance shifts of 7nm or more. These data suggest that the potential for spectral tuning exists in both genera studied, and that the region of spectral tuning varies between these genera. Based on previous analyses in cichlid opsins for species from Lakes Malawi and Victoria, we hypothesized that variation would predominate in the SWS1 and LWS genes (Hofmann *et al.* 2009). While the presence of variation in the SWS1 and LWS genes was in accordance with this initial hypothesis, the distribution of this variation between genera and the variable sites in SWS2B and RH2Aβ were unexpected.

The SWS1 gene represents the shortest-wavelength sensitive opsin in the cichlid visual palette. Spectral tuning in the UV portion of the visible spectrum cannot be accomplished by changes in gene expression, it must occur via sequence variation in SWS1. The genetic distance between the two functional alleles found is fairly large. This may indicate rapid diversification within the UV368 allele, an initial divergence between UV368 and UV378 that occurred around the time of the formation of the lake, or a combination of these two scenarios (Fig. 2-3). More specifically, UV378 appears to be fairly conserved, as the branch length between the extant alleles in both Malawi and Victoria is fairly short. The UV368 allele, which

has only been described in Lake Malawi, has relatively long branch lengths leading to and within its clade, indicating rapid diversification within the lake. Our data suggest that the UV368 allele in particular is undergoing accelerated evolution. This is supported by recent SNP analyses showing that the SWS1 gene has abnormally high Fst within *Metriaclima zebra* populations (Loh *et al.* 2008). Additional testing should look for evidence of a selective sweep as has been seen in the *LWS* gene in Lake Victoria cichlids (Terai *et al.* 2006). The fact that both Malawian and Victorian taxa fall within the UV378 clade suggests that this allele was present in the riverine ancestor of both radiations. This would suggest that the UV368 allele arose and rapidly diversified within the clear waters of the newly formed Lake Malawi (Fig. 2-3). However, it is important to note that there are no obvious differences in the color patterning or ecology of the observed *Metriaclima* species that we would intuitively associate with divergent selection for these two SWS1 alleles.

Sequence variation in the LWS gene is associated with speciation in Lake Victoria (Terai *et al.* 2002; Terai *et al.* 2006; Seehausen *et al.* 2008). Variation in this gene would also be anticipated in the sand-dwelling cichlids of Lake Malawi, given that many of these species are known to actively express LWS (Carleton *et al.* 2005a; Hofmann *et al.* 2009). However, the potentially functional Y261F polymorphism is absent in all of the sand-dwellers examined to date. The *Metriaclima* possess a subset of the alleles known in Lake Victoria, but the specific polymorphism present in the genus probably causes no functional change without the A2 chromophore present in Lake Victoria cichlids (Terai *et al.* 2006). The *Labidochromis*, however, possess the additional Y261F variable site that would cause a similar or possibly larger shift in λ

max than that observed between the Victorian L and H alleles (Terai *et al.* 2006). This variable site may indicate spectral tuning in the long-wavelength pigments of this genus.

The *Metriaclima* RH2Aβ gene presents an unusual situation. It is the most polymorphic of all the genes described here, with 17 nonsynonymous changes throughout the protein. Although the L183 allele was only found in four fishes (all three *M. lombardoi* and one of three *M. barlowi*) and appears to be unique to Lake Malawi, the genetic distance within this clade of four genes is threefold greater than the distance from a Tanganyikan species (*Tropheus duboisi*) to a Malawian species (*Tramitichromis intermedius*; Fig. 2-6). These data, along with the haplotype structure observed in the *M. lombardoi* individuals, suggests one of three alternative scenarios: (i) rapid functional diversification between the M183 and L183 alleles, (ii) degradation and potential undescribed pseudogenization of the L183 allele, or (iii) recombination of the M183 allele with a separate gene of unknown origin. Future expression studies quantifying relative expression of the RH2Aβ and RH2Aα genes may provide insight into the probability of the first two scenarios. However, the third scenario would require identifying a plausible source of the recombinant sequence.

Intraspecific variation was observed for the LWS gene much as it was for RH2Aβ. For LWS, our *L. caeruleus* samples included one individual homozygous for the Y261 allele, one homozygous for the F261 allele, and one heterozygote. Given that only three individuals were sampled for each species and multiple functional alleles were still observed, it is likely that these polymorphisms represent alleles of relatively high frequency within their respective species. This has profound

implications for speciation processes, as visual diversification within a species could easily result in ecological or sexual isolation. This concept can be strongly illustrated by the presence of the variant LWS alleles in *L. caeruleus*. If the Y261F polymorphism causes a shift in spectral sensitivity similar to what it causes in humans (10nm), it would represent a dramatic change in long-wavelength sensitivity in this species. Throughout the lake, various subpopulations of *L. caeruleus* have solid yellow males, pearly white males, and males with vertical black bars (Fig. 2-10). Although speculative, it is entirely possible females in different populations possess different LWS alleles and may mate assortatively because of this. Therefore, our data suggest a scenario that would be compatible to divergence via sensory drive (Terai *et al.* 2006; Kawata *et al.* 2007; Seehausen *et al.* 2008). This is a hypothesis that both warrants and requires further investigation.

There were several instances of pseudogenes found in these sequences. These are the first pseudogenes that we have detected in Malawi cichlid opsins, though pseudogenes have been detected in Tanganyikan cichlid opsins (Spady *et al.* 2005). The issue of the pseudogenization of RH2Aα in *Labidochromis caeruleus* is an important one. It is unclear whether the duplicate genes RH2Aβ, RH2Aα, or both are actively expressed in the cichlid retina. Since both these genes have a similar λmax, there is the potential for functional redundancy in the system. Therefore, we would predict that in *L. caeruleus,* RH2Aβ would be expressed. In other species, we have predicted that RH2Aα would be the predominant RH2A gene expressed, based on comparisons of λmax of visual pigments from MSP and protein expression. Because of their sequence similarity, it is difficult to assay the expression of each gene

separately. Further studies are needed to separately quantify RH2Aα and β expression. Such studies might confirm that expression of RH2Aβ compensates for the pseudogenization of RH2Aα in *L. caeruleus*.

A final topic that deserves mention is the physical structure of polymorphisms and their presence in haplotype blocks through the genome. In the *Labidochromis*, there is no discernable haplotype pattern in *LWS*. Only two associated variable residues were found in the entire *SWS2B* locus, with the rest of the sequence being fixed across the two major alleles. This large haplotype block that is common to both alleles strongly suggests that the T269 allele was recently derived from the A269 allele. However, unlike in the *Labidochromis*, the *Metriaclima* have large, distinct haplotype blocks in or around their variable genes (Figs. 2-5 and 2-6). In *SWS1*, the UV378 allele is fixed across roughly two-thirds of the protein sequence, while the UV368 is much more variable. In *M. lombardoi*, two haplotype corresponding haplotype blocks are found at the margins of both *RH2Aβ* and *RH2Aα*. These blocks surround a 10kb intergenic region, potentially indicating the presence of a larger block. However, this block may actually represent a haplotype focused around the *RH2Aα* locus as *M. barlowi* has two corresponding blocks at the margins of *RH2Aα* and *RH2B*, with an intergenic distance of roughly 15kb. If these blocks extend through these intergenic regions, they are two to three times greater than those described for the Victorian LWS genes (Terai *et al.* 2006), suggesting very strong or very recent selection on the *RH2Aα* locus. Extensive population sampling and sequencing of the intergenic regions will be required to determine if reduced variation in this region is truly representative of a selective sweep.

In sum, the work presented here demonstrates that the sequence polymorphisms needed to tune spectral sensitivity are present not just within the Malawian flock as a whole, but within genera and even species. Both the *Metriaclima* and the *Labidochromis* possess two polymorphic genes that could tune spectral sensitivities between congeners. Therefore, the potential for spectral sensitivity shifts via opsin sequence changes was found in both genera, but functional variation within a given opsin gene was confined to a single genus in each case. This suggests that selection may act on different regions of spectral sensitivity in these genera. Although variations in gene expression are clearly key to visual sensitivity within Lake Malawi, the presence of opsin sequence polymorphisms may allow for spectral tuning in situations where gene expression is fairly constrained. As such, opsin sequence variation may contribute to sensory drive and subsequent species divergence in Lake Malawi much as it does in Lake Victoria.

*Figures*

| Opsin | Whole Protein | | Transmembrane Regions | | Binding Pocket Sites | |
|---|---|---|---|---|---|---|
| | Total | Polarity shift | Total | Polarity shift | Total | Polarity shift |
| SWS1 | 14 | 9 | 9 | 6 | 4 | 4 |
| SWS2b | 5 | 2 | 3 | 1 | 1 | 1 |
| SWS2a | 6 | 2 | 4 | 2 | 0 | 0 |
| RH2B | 7 | 3 | 5 | 2 | 1 | 1 |
| RH2Aα | 5 | 1 | 3 | 0 | 0 | 0 |
| RH2Aβ | 17 | 8 | 8 | 3 | 2 | 1 |
| LWS | 9 | 3 | 6 | 3 | 1 | 1 |

**Figure 2-1:** Summary of opsin polymorphisms in the genus *Metriaclima*. Colored rows represent genes with suspected functional variation.

| | Whole Protein | | Transmembrane Regions | | Binding Pocket Sites | |
|---|---|---|---|---|---|---|
| Opsin | Total | Polarity shift | Total | Polarity shift | Total | Polarity shift |
| SWS1 | 7 | 3 | 4 | 1 | 0 | 0 |
| SWS2b | 2 | 2 | 2 | 2 | 1 | 1 |
| SWS2a | 0 | 0 | 0 | 0 | 0 | 0 |
| RH2B | 9 | 3 | 7 | 2 | 2 | 2 |
| RH2Aα | 4 | 0 | 4 | 0 | 0 | 0 |
| RH2Aβ | 10 | 5 | 9 | 4 | 0 | 0 |
| LWS | 7 | 4 | 4 | 4 | 3 | 3 |

**Figure 2-2:** Summary of opsin polymorphisms in the genus *Labidochromis*. Colored rows represent genes with suspected functional variation.

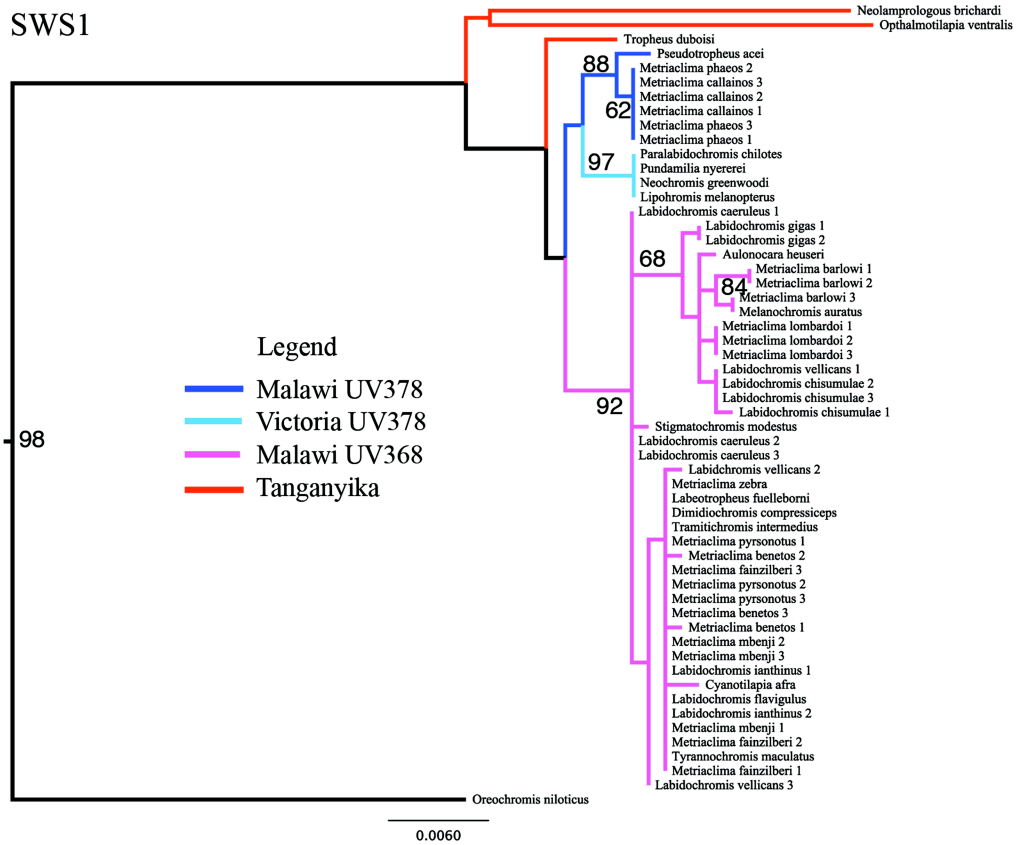


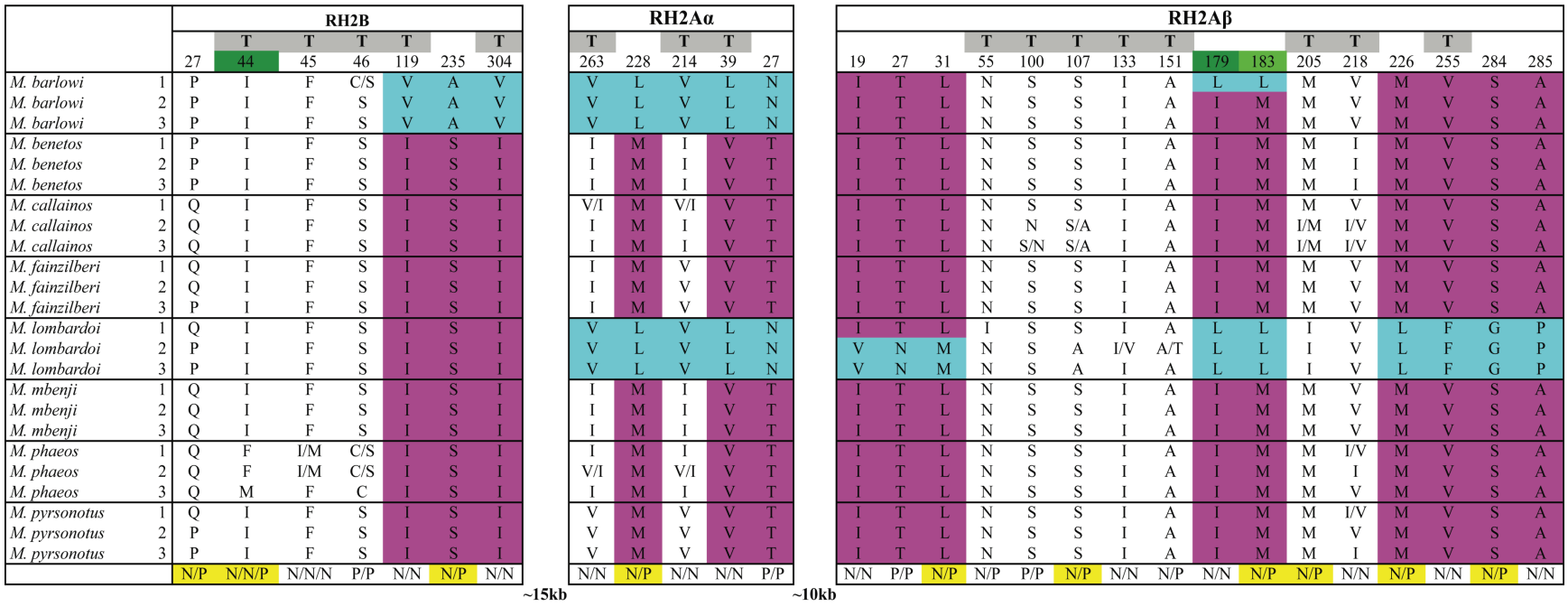

**Figure 2-3:** Gene tree for the SWS1 gene. The red branches represent reference sequences taken from Lake Tanganyika. The dark blue branches represent the UV378 allele, with the light blue clade being the Victorian subset of the UV378 allele. The pink clade represents all of the UV368 alleles described in this study, as well as some reference sequences. This clade appears to be unique to Lake Malawi.

| | | RH2Aβ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | T | T | T | T | T | T | T | T | T | |
| | | 56 | 100 | 107 | 133 | 158 | 162 | 205 | 214 | 218 | 228 |
| *L. caeruleus* | 1 | G | S | S | I | L | I | M | V | I/T | L/M |
| *L. caeruleus* | 2 | G | S | S | I | L | I | I | V | T | L |
| *L. caeruleus* | 3 | G | S | S | I | L | I | M | V | I | M |
| *L. ianthinus* | 2 | G | S | S | I | L | I | M | V | I | M |
| *L. chisumulae* | 1 | S | N | P | I | L | I | M | V | I | M |
| *L. chisumulae* | 2 | S | N | P | I | L | I | M | V | I | M |
| *L. chisumulae* | 3 | S | N | P | I | L | I | M | V | I | M |
| *L. gigas* | 1 | S | N | P | I | L/F | I/V | I | V | I | M |
| *L. gigas* | 2 | S | N | P | I | L/F | I | I | V | I | M |
| *L. vellicans* | 1 | S | N | P | I | L | I | M | V | I | M |
| *L. vellicans* | 2 | S | N | P | I/V | L | I | M | V | I | M |
| *L. vellicans* | 3 | S | N | P | I | L/F | I | M | I/V | I/T | M |
| | | N/P | P/P | N/P | N/N | N/N | N/N | N/P | N/N | N/P | N/P |

**Figure 2-4:** Table of all variable amino acid residues in the RH2Aβ gene in the genus *Labidochromis*. Pink and light blue portions represent haplotype blocks in the amino acid sequence. Shaded "T"s denote polymorphic sites in the transmembrane regions, while yellow cells represent sites with polymorphisms resulting in polarity shifts.

| | | SWS1 | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | T | T | | T | T | T | T | T | T | T |
| | | 15 | 21 | 36 | 37 | 81 | 83 | 103 | 114 | 160 | 165 | 201 | 204 | 214 | 217 |
| *M. barlowi* | 1 | Y | V | V | Y | L | G | F | S | T | I | S | T | M | F |
| *M. barlowi* | 2 | Y | V | V | Y | L | G | F | S | T | I | S | T | M | F |
| *M. barlowi* | 3 | Y | I/V | A/V | Y/F | L/F | G | F | S | T | I | T | T | I/M | F/S |
| *M.benetos* | 1 | Y | I | A | F | F | G | Y | S | T | I | T | T | I | S |
| *M.benetos* | 2 | F | I | A | F | F | G | Y | S | T | I | T | T | I | S |
| *M.benetos* | 3 | Y | I | A | F | F | G | Y | S | T | I | T | T | I | S |
| *M. fainzilberi* | 1 | Y | I | A | F | F | G | Y | S | T | I | T | T | I | S |
| *M. fainzilberi* | 2 | Y | I | A | F | F | G | Y | S | T | I | T | T | I | S |
| *M. fainzilberi* | 3 | Y | I | A | F | F | G | Y | S | T | I | T | T | I | S |
| *M. lombardoi* | 1 | Y | V | A | Y | L | G | F | S | T | I | S | T | M | S |
| *M. lombardoi* | 2 | Y | V | A | Y | L | G | F | S | T | I | S | T | M | S |
| *M. lombardoi* | 3 | Y | V | A | Y | L | G | F | S | T | I | S | T | M | S |
| *M. mbenji* | 1 | Y | I | A | F | F | G | Y | S | T | I | T | T | I | S |
| *M. mbenji* | 2 | Y | I | A | F | F | G | Y | S | T | I | T | T | I | S |
| *M. mbenji* | 3 | Y | I | A | F | F | G | Y | S | T | I | T | T | I | S |
| *M. pyrsonotus* | 1 | Y | I | A | F | F | G | Y | S | T | I | T | T | I | S |
| *M. pyrsonotus* | 2 | Y | I | A | F | F | G | Y | S | T | I | T | T | I | S |
| *M. pyrsonotus* | 3 | Y | I | A | F | F | G | Y | S | T | I | T | T | I | S |
| *M. callainos* | 1 | Y | V | A | F | F | S | Y | A | A | V | T | I | I | S |
| *M. callainos* | 2 | Y | V | A | F | F | S | Y | A | A | V | T | I | I | S |
| *M. callainos* | 3 | Y | V | A | F | F | S | Y | A | A | V | T | I | I | S |
| *M. phaeos* | 1 | Y | I/V | A | F | F | S | Y | A | A | V | T | I | I | S |
| *M. phaeos* | 2 | Y | V | A | F | F | S | Y | A | A | V | T | I | I | S |
| *M. phaeos* | 3 | Y | I/V | A | F | F | S | Y | A | A | V | T | I | I | S |
| | | N/P | N/N | N/N | N/P | N/N | N/P | N/P | N/P | N/P | N/N | P/P | N/P | N/P | N/P |

**Figure 2-5:** Table of all variable amino acid residues in the SWS1 gene in the genus *Metriaclima*. Pink and light blue portions represent invariable blocks in the amino acid sequence. Bright green residue numbers represent sites that are present in the retinal binding pocket and are suspected to be functional. Shaded "T"s denote polymorphic sites in the transmembrane regions, while yellow cells represent sites with polymorphisms resulting in polarity shifts.




| | | RH2B | | | | | | | RH2Aα | | | | | RH2Aβ | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | T | T | T | T | | T | T | | T | T | | | | | T | T | T | T | T | | | T | T | | T | | | |
| | | 27 | 44 | 45 | 46 | 119 | 235 | 304 | 263 | 228 | 214 | 39 | 27 | 19 | 27 | 31 | 55 | 100 | 107 | 133 | 151 | 179 | 183 | 205 | 218 | 226 | 255 | 284 | 285 |
| *M. barlowi* | 1 | P | I | F | C/S | V | A | V | V | L | V | L | N | I | T | L | N | S | S | I | A | L | L | M | V | M | V | S | A |
| *M. barlowi* | 2 | P | I | F | S | V | A | V | V | L | V | L | N | I | T | L | N | S | S | I | A | I | M | M | V | M | V | S | A |
| *M. barlowi* | 3 | P | I | F | S | V | A | V | V | L | V | L | N | I | T | L | N | S | S | I | A | I | M | M | V | M | V | S | A |
| *M. benetos* | 1 | P | I | F | S | I | S | I | I | M | I | V | T | I | T | L | N | S | S | I | A | I | M | M | I | M | V | S | A |
| *M. benetos* | 2 | P | I | F | S | I | S | I | I | M | I | V | T | I | T | L | N | S | S | I | A | I | M | M | I | M | V | S | A |
| *M. benetos* | 3 | P | I | F | S | I | S | I | I | M | I | V | T | I | T | L | N | S | S | I | A | I | M | M | I | M | V | S | A |
| *M. callainos* | 1 | Q | I | F | S | I | S | I | V/I | M | V/I | V | T | I | T | L | N | S | S | I | A | I | M | M | V | M | V | S | A |
| *M. callainos* | 2 | Q | I | F | S | I | S | I | I | M | I | V | T | I | T | L | N | N | S/A | I | A | I | M | I/M | I/V | M | V | S | A |
| *M. callainos* | 3 | Q | I | F | S | I | S | I | I | M | I | V | T | I | T | L | N | S/N | S/A | I | A | I | M | I/M | I/V | M | V | S | A |
| *M. fainzilberi* | 1 | Q | I | F | S | I | S | I | I | M | V | V | T | I | T | L | N | S | S | I | A | I | M | M | V | M | V | S | A |
| *M. fainzilberi* | 2 | Q | I | F | S | I | S | I | I | M | V | V | T | I | T | L | N | S | S | I | A | I | M | M | V | M | V | S | A |
| *M. fainzilberi* | 3 | P | I | F | S | I | S | I | I | M | V | V | T | I | T | L | N | S | S | I | A | I | M | M | V | M | V | S | A |
| *M. lombardoi* | 1 | Q | I | F | S | I | S | I | V | L | V | L | N | I | T | L | I | S | S | I | A | L | L | I | V | L | F | G | P |
| *M. lombardoi* | 2 | P | I | F | S | I | S | I | V | L | V | L | N | V | N | M | N | S | A | I/V | A/T | L | L | I | V | L | F | G | P |
| *M. lombardoi* | 3 | P | I | F | S | I | S | I | V | L | V | L | N | V | N | M | N | S | A | I | A | L | L | I | V | L | F | G | P |
| *M. mbenji* | 1 | Q | I | F | S | I | S | I | I | M | I | V | T | I | T | L | N | S | S | I | A | I | M | M | V | M | V | S | A |
| *M. mbenji* | 2 | Q | I | F | S | I | S | I | I | M | I | V | T | I | T | L | N | S | S | I | A | I | M | M | V | M | V | S | A |
| *M. mbenji* | 3 | Q | I | F | S | I | S | I | I | M | I | V | T | I | T | L | N | S | S | I | A | I | M | M | V | M | V | S | A |
| *M. phaeos* | 1 | Q | F | I/M | C/S | I | S | I | I | M | I | V | T | I | T | L | N | S | S | I | A | I | M | M | I/V | M | V | S | A |
| *M. phaeos* | 2 | Q | F | I/M | C/S | I | S | I | V/I | M | V/I | V | T | I | T | L | N | S | S | I | A | I | M | M | I | M | V | S | A |
| *M. phaeos* | 3 | Q | M | F | C | I | S | I | I | M | I | V | T | I | T | L | N | S | S | I | A | I | M | M | V | M | V | S | A |
| *M. pyrsonotus* | 1 | Q | I | F | S | I | S | I | V | M | V | V | T | I | T | L | N | S | S | I | A | I | M | M | I/V | M | V | S | A |
| *M. pyrsonotus* | 2 | P | I | F | S | I | S | I | V | M | V | V | T | I | T | L | N | S | S | I | A | I | M | M | V | M | V | S | A |
| *M. pyrsonotus* | 3 | P | I | F | S | I | S | I | V | M | V | V | T | I | T | L | N | S | S | I | A | I | M | M | I | M | V | S | A |
| | | N/P | N/N/P | N/N/N | P/P | N/N | N/P | N/N | N/N | N/P | N/N | N/N | P/P | N/N | P/P | N/P | N/P | P/P | N/P | N/N | N/P | N/N | N/P | N/P | N/N | N/P | N/N | N/P | N/N |

**Figure 2-6:** Table of all variable amino acid residues in the RH2 genic tandem array in the genus *Metriaclima*. Pink and light blue portions represent invariable blocks in the amino acid sequence. Bright green residue numbers represent sites that are present in the retinal binding pocket that are suspected to be functional, while dark green residues are present in the binding pocket but probably have no functional effect. Shaded "T"s denote polymorphic sites in the transmembrane regions, while yellow cells represent sites with

polymorphisms resulting in polarity shifts. The physical distance between the RH2B and RH2Aα genes is ~15kb, while the distance between the RH2Aα and RH2Aβ genes is ~10kb. The orientation of the RH2Aα gene is reversed in the genome when compared to the other two, and so its sites are listed in reverse order.

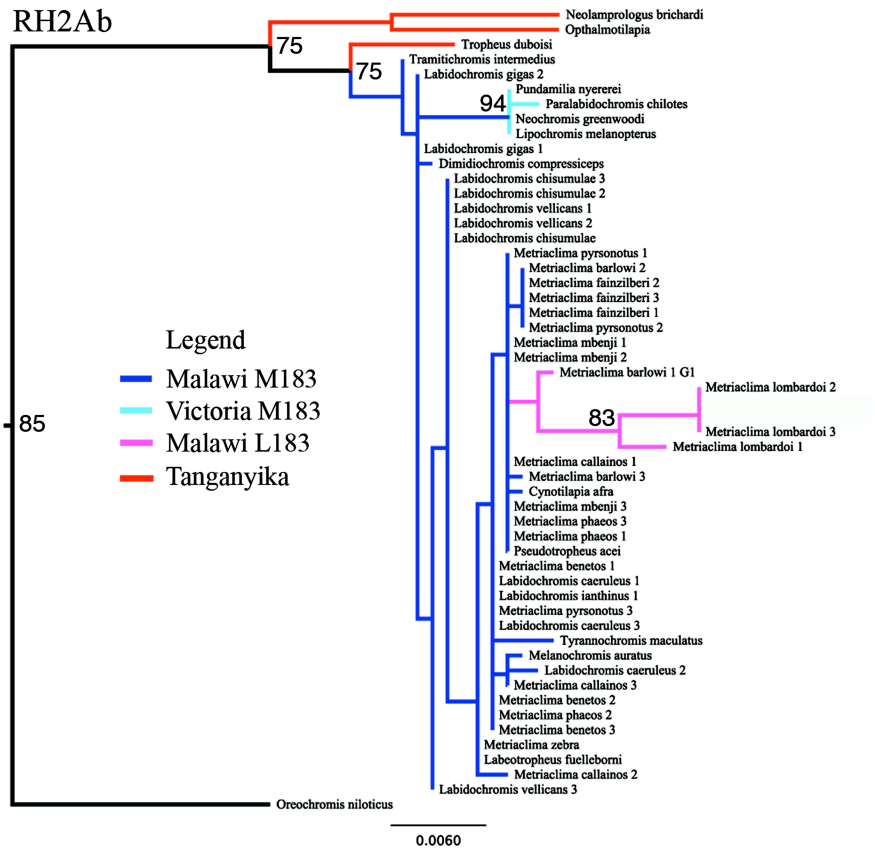



**Figure 2-7:** Gene tree for the RH2Aβ gene. The red branches represent reference sequences taken from Lake Tanganyika. The dark blue branches are Malawian versions of the M183 allele, with the light blue clade being the Victorian subset of this allele. The pink clade represents the L183 allele.

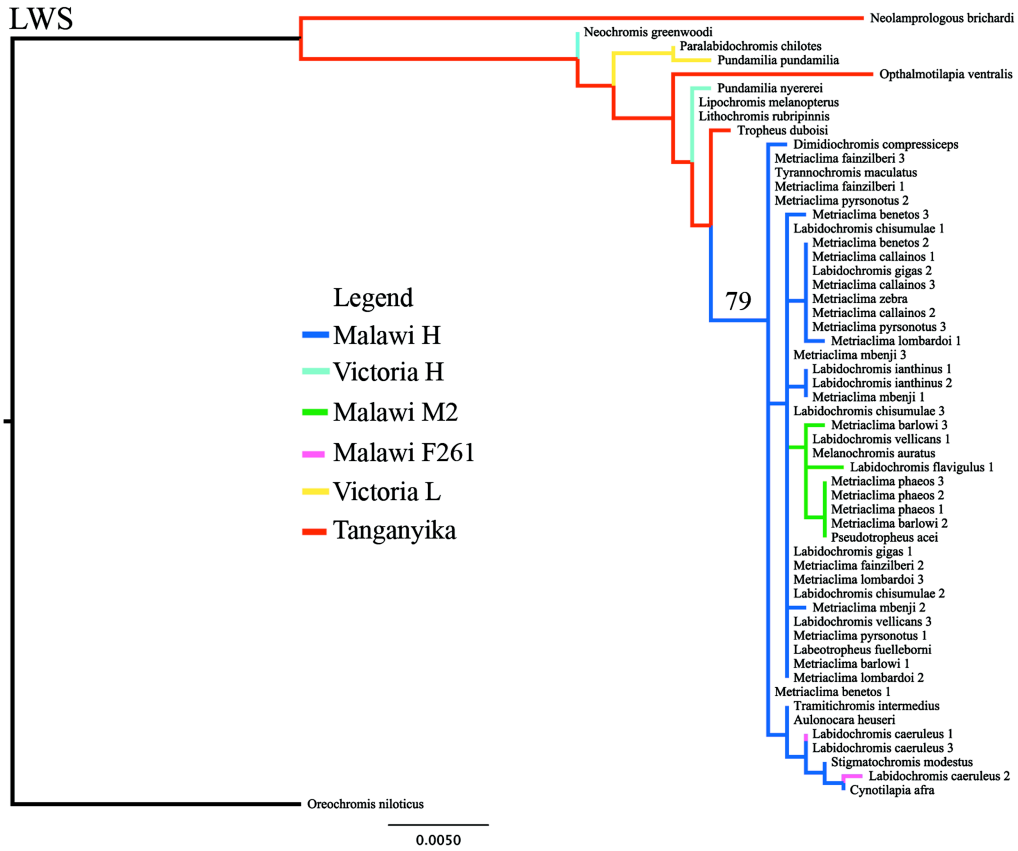



**Figure 2-8:** Gene tree for the LWS gene. The red branches represent reference sequences taken from Lake Tanganyika. The light blue branches are H-type allele in Victoria, whereas the yellow branches represent L-type alleles (Terai *et al.* 2006). The dark blue clade is the H-type alleles from Lake Malawi, with the green clade being the M2-type alleles. The two pink branches are the two individual *Labidochromis caeruleus* that possess the functional F261 polymorphism.

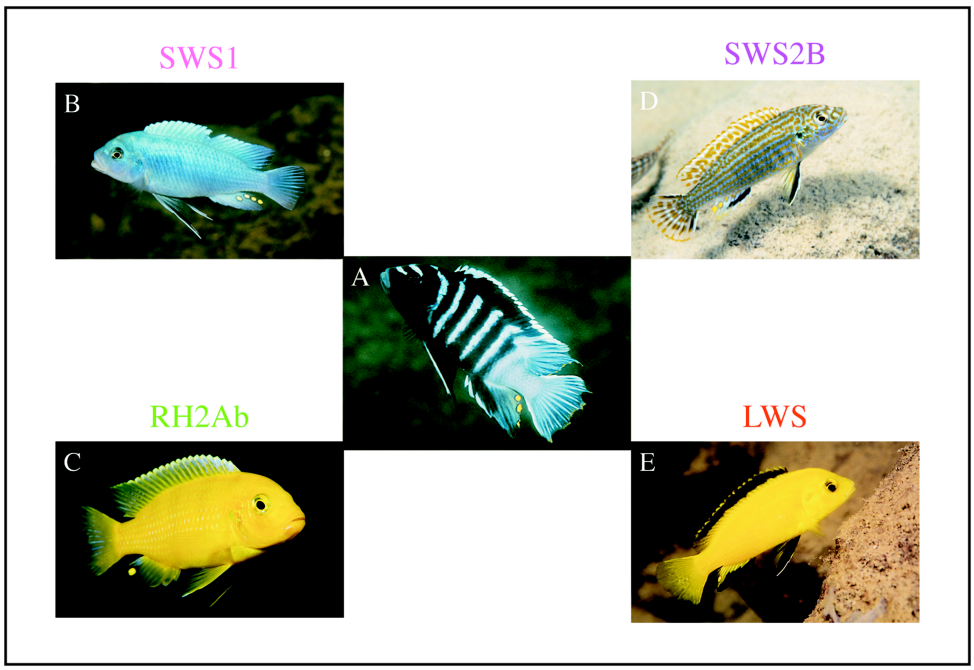



**Figure 2-9:** Panel of species containing functional variants when compared to the reference alleles from *Metriaclima zebra* (A). (B) *M callainos* - UV378 variant of SWS1. (C) *M. lombardoi* - L183 variant of the RH2Aβ gene. (D) *Labidochromis flavigulus* – T269 variant of SWS2B. (E) *L. caeruleus* – F261 variant of LWS. Photos courtesy of Ad Konings, Cichlid Press.

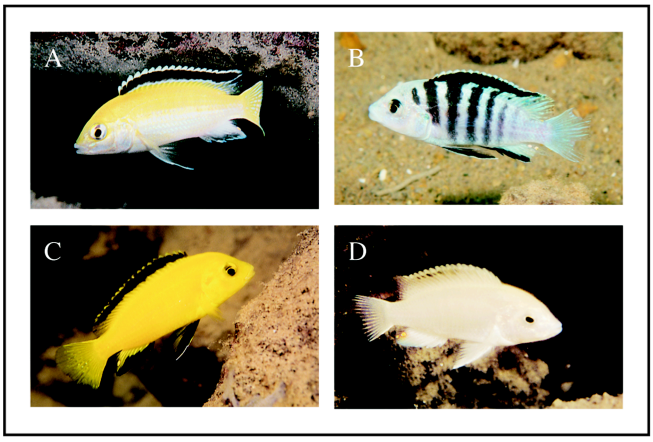



**Figure 2-10:** Variation in male nuptial coloration of *Labidochromis caeruleus* at several locations across the Lake. (A) Lion's Cove, (B) Lundo Island, (C) Mbowe, and (D) Ruarwe. Photos courtesy of Ad Konings, Cichlid Press.

# Chapter 3:

# Intraspecific Cone Opsin Expression Variation in the Cichlids of Lake Malawi

Adam R. Smith, Lindsay D'Annunzio, Abbi E. Smith, Anit Sharma, Christopher M. Hofmann, N. Justin Marshall & Karen L. Carleton

See Appendix 2 for all supplemental materials referenced in this chapter

This chapter is slightly modified from an article published as:

Smith AR, D'Annunzio L, Smith AE, Sharma A, Hofmann CM, Marshall NJ, Carleton KL (2011) Intraspecific cone opsin expression variation in the cichlids of Lake Malawi. Molecular Ecology 20: 299-310.

*Abstract*

The expression of cone opsin genes is a primary determinant of the characteristics of colour vision. Interspecific variation in opsin expression is common in African cichlids. It is correlated with foraging among cichlids from Lake Malawi, and with ambient light environment among cichlids from Lake Victoria. In this study we tested whether gene expression varied within species such that it might be important in contributing to divergence. We hypothesized that light attenuation with depth would be correlated with predictable changes in gene expression in Lake Malawi, and that this variation would tune visual sensitivities to match the ambient light environment. We observed significant differences in cone opsin expression in three different comparisons among populations of the same species. Higher LWS expression was found in shallow versus deep *Copadichromis eucinostomus*. In *Metriaclima zebra*, individuals from Zimbawe Rock expressed significantly more SWS2B than those from Thumbi West Island although these locales have similar ambient light environments. Finally, *Tropheops gracilior* from deeper water had significantly more variation in expression than their shallow counterparts. These results support that gene expression varies significantly between populations of the same species. Surprisingly, these results could not be explained by predicted visual performance as models predicted that differential expression patterns did not confer sensitivity advantages at different depths. This suggested that expression variation did not confer a local sensitivity advantage. Therefore, our findings were contrary to a primary requirement of the sensory bias hypothesis. As such, other explanations for intraspecific gene expression variation need to be tested.

*Introduction*

Several evolutionary models have been proposed to explain speciation in sympatry. One of these is Fisher's runaway model, in which random changes in female preferences lead to rapid divergence in male traits (Fisher 1930; Lande 1981). Another is sensory drive, in which the local environment sets female sensitivities that then select for particular male traits (Endler 1992; Boughman 2002; Dangles *et al.* 2009). Both of these models predict that there should be variation in sensory systems among sister taxa and perhaps even among populations. However, these models differ in whether female sensitivities will be correlated with the local environment. The Fisherian model suggests changes can occur neutrally, perhaps through drift, such that no obvious association with local environment will be observed. In contrast, the sensory drive model supposes that the local environment shapes a key sensory system and through habitat variation leads to female sensory variation. Alternatively, models of pure genetic drift require no selective processes to drive or reinforce divergence, and may occur in small, isolated populations.

African cichlid fishes are a model for rapid divergence and speciation. Replicate species flocks have arisen in each of the Great Lakes: Malawi, Tanganyika and Victoria (Kocher 2004). Lake Malawi is the most species rich of the East African rift lakes, and is estimated to have well over 500 endemic species (Kornfield & Smith 2000). Sexual selection on male nuptial colouration has been posited as one of the major forces contributing to the development of new species in this lake (Danley & Kocher 2001). Female choice for male nuptial colouration is tightly linked to the ability to perceive and distinguish different colour patterns, and therefore we expect

variation in colour vision to be an important facet of speciation. More specifically, if nuptial colouration is the substrate for selection and speciation, we should observe differences in sensory traits that influence colour vision in diverging species or populations.

Work on Lake Victoria cichlids has demonstrated that sequence changes in the LWS opsin gene shift visual sensitivities (Terai *et al.* 2002; 2006; Carleton *et al.* 2005). Variation in the LWS opsin sequences is correlated with depth. Longer wavelength-sensitive alleles are found deeper where the downwelling light spectrum shifts to longer wavelengths. LWS sequences also show signs of selection at multiple locales where differing levels of turbidity influence the quality of the ambient light environment. Male colours are also correlated with depth and visual sensitivity. These results suggest that speciation has been driven by the ecological light gradient selecting for alternate visual sensitivities, which in turn select for particular male colour patterns (Seehausen *et al.* 2008). The evolution of LWS genes in Lake Victoria cichlids is an example of the sensory drive model. Similar changes in opsin sequence have been observed in closely-related fishes from Malawi (Spady *et al.* 2006; Hofmann *et al.* 2009; Smith & Carleton 2010), but correlations with male colour have not been examined.

In addition to sequence diversity, changes in opsin gene expression act as a second genetic mechanism for visual tuning in African cichlids (Carleton 2009). Changes in opsin expression patterns are known to cause visual shifts in certain teleost fish species that are tied to variation in the surrounding light environment (Fuller *et al.* 2004; 2005; Shand *et al.* 2008). East African cichlids possess seven

distinct opsin genes that are differentially expressed to tune the visual system to different environments. For example, the relative expression rates of SWS and LWS opsin genes of Lake Victoria cichlids are tightly linked to the ambient light environment at different geographic locations (Hofmann *et al.* 2009).

Although Malawi cichlids have seven different opsin genes, previous research suggests that individual fish usually express only three opsin genes: one for the short-wavelength sensitive single cones (ultraviolet (SWS1 - $\lambda_{max}$ 368nm), violet (SWS2B - $\lambda_{max}$ 415nm), or blue (SWS2A - $\lambda_{max}$ 455nm)) and two for the medium- and long-wavelength sensitive double cones (blue-green (RH2B - $\lambda_{max}$ 488nm), two true greens (RH2A$\alpha$ and $\beta$ – $\lambda_{max}$ 518nm), and red (LWS - $\lambda_{max}$ 560nm); Parry *et al.* 2005; Spady *et al.* 2006). These are typically expressed in one of a few combinations or visual palettes. The three most common visual palettes are the short wavelength (SWS1, RH2B, RH2A), medium wavelength (SWS2B, RH2B, RH2A) and long wavelength (SWS2A, RH2A, LWS) sets. Since the total number of expressed pigments is limited, visual systems should be tuned such that the genes expressed most closely matches the local light environment.

In previous studies, we found that light environment shaped opsin expression in cichlids from Lake Victoria (Hofmann *et al.* 2009). Fish in clear water habitats expressed more SWS2B opsin than those from murkier habitats. This increase was correlated with the amount of available light for stimulating these shorter wavelength cones. In addition, more LWS double cones were observed in fish from murky habitats (Carleton *et al.* 2005). Further, hybrid-cross experiments between Malawi cichlids suggested there is a strong genetic component to opsin gene expression with

just a few genes controlling expression patterns (Carleton *et al.* 2010). However, we also found that there was some plasticity to opsin expression when fish were raised in the lab versus those caught from the wild (Hofmann *et al.* 2010). This variation was smaller, but statistically significant. Such environmental plasticity has been found in opsin expression from other species, most notably killifish (Fuller *et al.* 2004; 2005) and black bream (Shand *et al.* 2008). Clearly, both genetic and environmental mechanisms play a role in modulating expression of opsin genes in natural populations of cichlids from Lake Malawi.

In this work, we sought to determine the extent of opsin gene expression variation in natural populations and examined variation at two spatial scales. First, we profiled variation with depth at a single site in two species, *Tropheops gracilior* and *Copadichromis eucinostomus* in order to determine the effects of habitat depth on local expression patterns. The differential attenuation of wavelengths through the water column should have created an environmental gradient wherein the SWS1 and LWS opsins at the extremes of the visual spectrum should become less efficient as depth increases (Dalton *et al.* 2010; Fig. 3-1). We also profiled variation in *Metriaclima zebra* in two distinct geographic locations that occurred at similar depths with similar light environments. We had three primary hypotheses: (i) expression variation would be quantitatively more variable at shallow depths in *T. gracilior* and *C. eucinostomus* than in deeper waters due to the presence of a broader light spectrum allowing for greater pigment efficiency at the ends of the spectrum, (ii) qualitative changes in these species would involve the loss of LWS in response to the narrowing

light spectrum with depth, and (iii) the *M. zebra* populations would be relatively homogenous due to selection from similar light environments.

## Materials and Methods

### *Research Subjects*

All individuals were collected from Lake Malawi during July 2008. Fishes were sacrificed the day of capture and their eyecups were stored in RNAlater. Three species were sampled for this study: *Tropheops gracilior*, *Copadichromis eucinostomus*, and *Metriaclima zebra*. Thirty-six *T. gracilior* were collected in total at Otter Island, with twenty from 5m and sixteen from 20m depth. Twenty-eight *C. eucinostomus* were collected at Otter Point, with fifteen from 5m and thirteen from 15m. A total of twenty *M. zebra* were collected from two geographically distinct locations: ten from Thumbi West and ten from Zimbawe Rock. These sampling locations are depicted in Figure S3-1.

### *Quantification of relative expression*

RNA was isolated from the dissected retina using commercial Qiagen Qiashredder and RNeasy kits (Valencia CA, USA). The isolated RNA was then reverse-transcribed to cDNA. Real-time quantitative PCR (RT-qPCR) was performed using a set of seven cone opsin-specific primer/probe combinations. The total expression of the duplicate green genes (RH2Aα and RH2Aβ) was measured together using the aggregate RH2A primer/probe developed for prior studies (Spady *et al.* 2006). The relative expression of these two genes was then quantified using a RH2Aβ-specific primer (F:ATTCTTGGATCCACCTTCTGTGCA,

R:TACTCCAGCACCAGCATGAG) to compare RH2Aβ to aggregate RH2A expression. This seven-reaction template was run in parallel for each sample, and each sample was replicated at least twice on separate plates with separate reaction master mixes. The RT-qPCR reaction was performed on a Roche Lightcycler 480 and the critical cycle number for each gene was determined.

Reaction efficiencies for each primer were estimated using a plasmid construct containing a single copy of each opsin gene, although the RH2Aβ-specific primer could not anneal to the construct (Spady *et al.* 2006). Experiments using either RH2Aβ-specific or RH2A aggregate probes showed PCR was equally efficient on a purified RH2Aβ substrate (data not shown). Therefore, we used the same efficiency for both RH2A genes. The efficiencies of all seven genes were used in conjunction with the critical cycle data from the Lightcycler to determine relative expression values for each gene. Replicate plate measures were averaged for each individual.

*Statistical Analyses*

All analyses were performed on groups within each species (defined either by depth or by locale). Differences in gene expression were analyzed using both univariate and multivariate methods. For univariate analyses, the mean relative expression of each gene was tested between groups (depth or location) using an ANOVA. Homogeneity of variances between groups for each gene was tested using the Brown-Forsythe test of unequal variances. The Brown-Forsythe test is robust to distributions that are skewed towards the tails (Brown & Forsythe 1974), and therefore was appropriate for our data. General patterns of gene expression and variance were plotted via a triplot method based on an open-source Excel spreadsheet (Tri-plot; Graham and Midgley

2000). By triangulating individual expression patterns for both short-wavelength pigments expressed in single cones (SWS1, SWS2B, SWS2A) or medium- and long-wavelength sensitive pigments expressed in double cones (RH2B, RH2A, LWS), it is possible to observe clustering patterns within the data. However, the triplot itself is simply a graphical technique and has no associated statistical p-value.

For multivariate comparisons, groups were compared using a discriminant function analysis (DFA). The DFA analysis tests for differences in multivariate means along principal component axes, allowing for the discrimination of groups using the multivariate characteristics that describe the largest fraction of data in the model. Our DFA model was constructed using all genes except for RH2Aα, which was excluded to alleviate linear dependence in the dataset. Since RH2Aα was expressed in high levels in all individuals it was considered as generally uninformative for the multivariate models.

In order to quantify the theoretical light collection of each cichlid visual pigment in the environment, quantum catches were calculated using *in situ* water spectral data from Zimbawe Rock (Hofmann *et al.* 2009), which is similar to that at Thumbi West or Otter Point (data not shown). Quantum catch for pigments was estimated for 1m, 5m, 15m, 30m and 50m using the following equation from Hofmann *et al.* (2009):

$$Q_{abs,i} = \int_{320}^{750} I(\lambda) T_w(\lambda,d) R(\lambda) d\lambda \qquad (1a)$$

Here, $Q_{abs,i}$ is the absolute quantum catch for a particular photoreceptor visual pigment, $I(\lambda)$ is the incident solar irradiance at the surface, $T_w(\lambda,d)$ is the transmission of light of wavelength $\lambda$ through water to a depth, d, and $R(\lambda)$ is the absorbance of

the photoreceptor visual pigment based on the equations of Govardovskii *et al.* (2000). The visual pigment quantum catches were then normalized relative to each other by dividing by the sum of the total of all visual pigments to produce relative quantum catches:

$$Q_{rel,i} = \frac{Q_{abs,i}}{\sum Q_{abs,i}} \quad (1b)$$

This corrected for total light intensity differences with depth so the relative efficiencies of each visual pigment could be compared.

In order to approximate the effects of variable gene expression on visual system sensitivity, we constructed a derived model based on the $Q_i$ functions. In so doing, we made two primary simplifying assumptions regarding the function of cone opsins in the retina: (i) relative gene expression is a proxy for cone cell number and therefore proportional to sensitivity, and (ii) cone mechanisms are additive and function as a luminance channel. While (i) is a relatively fixed assumption, (ii) may only hold for certain lighting conditions and visual tasks. However, this luminance-based additive model is the simplest model that we can make without requiring knowledge of specific neural mechanisms. This enables us to estimate the total visual system sensitivity for a given combination of expressed opsins in a given ambient light environment. In performing this calculation, we weighted the quantum catch of a particular visual pigment by its relative gene expression (given as $f_i$ or the fraction of the total opsin expression). Further, we performed this weighting separately for the single cones (SC) and the double cones (DC) or combining all cones using the following equations:

$$RT-S_{sc}=\frac{f_{SWS1}Q_{SWS1}+f_{SWS2B}Q_{SWS2B}+f_{SWS2A}Q_{SWS2A}}{f_{SWS1}+f_{SWS2B}+f_{SWS2A}} \quad (2)$$

$$RT-S_{dc}=\frac{f_{RH2B}Q_{RH2B}+f_{RH2A}Q_{RH2A}+f_{LWS}Q_{LWS}}{f_{RH2B}+f_{RH2A}+f_{LWS}} \quad (3)$$

$$RT-S_{all}=f_{SWS1}Q_{SWS1}+f_{SWS2B}Q_{SWS2B}+f_{SWS2A}Q_{SWS2A}+f_{RH2B}Q_{RH2B}+f_{RH2A}Q_{RH2A}+f_{LWS}Q_{LWS} \quad (4)$$

Henceforth, we refer to the expression-scaled quantum catch as RT-S (Real-Time corrected Sensitivities). These calculations were repeated in *C. eucinostomus* and *T. gracilior* using ambient spectra from both 5m and 15m to test visual sensitivities in each microhabitat, *i.e.* RT-S values were calculated for fish from each population (depth) for both light environments and the sensitivities were then compared within that depth. Absolute quantum catch models were analyzed by ANOVA and the Brown-Forsythe test. In addition, weighted $\lambda_{max}$ calculations were performed for comparison to the results from prior work (Hofmann *et al.* 2009) where the following equations were used:

$$\lambda_{\max,SC}=\frac{f_{SWS1}\lambda_{SWS1}+f_{SWS2B}\lambda_{SWS2B}+f_{SWS2A}\lambda_{SWS2A}}{f_{SWS1}+f_{SWS2b}+f_{SWS2A}} \quad (5)$$

$$\lambda_{\max,DC}=\frac{f_{RH2B}\lambda_{RH2B}+f_{RH2A}\lambda_{RH2A}+f_{LWS}\lambda_{LWS}}{f_{RH2B}+f_{RH2A}+f_{LWS}} \quad (6)$$

The $\lambda_{max}$ calculation estimates the average visual pigment peak sensitivity weighted by gene expression, with each visual pigment represented solely by it's wavelength of maximal absorption, This is simpler than the RT-S calculations which weights quantum catch by gene expression. However, since quantum catch considers the full absorption profile of each opsin as well as the environmental light conditions, the RT-

S model is more likely to be a good mathematical predictor of *in vivo* visual sensitivities.

In addition to calculating relative quantum catch and RT-S using downwelling irradiance in Lake Malawi, we calculated these same quantities using downwelling irradiance in Lake Victoria (Hofmann *et al.* 2009). The latter used spectra taken near Makobe Island, a relatively clear water site in Lake Victoria to calculate quantum catch and compare it to that at Zimbawe Rock, Malawi. In both cases:

$$I(\lambda)T_w(\lambda,d) = I_{surface}(\lambda)\exp(-\alpha_\lambda d) \quad (7)$$

where $I_{surface}(\lambda)$ was irradiance measured at the water's surface, which was taken at Zimbawe Rock, $\alpha_\lambda$ is the attenuation coefficient measured previously for these locations in each lake (Hofmann *et al.* 2009), and d is the depth. This enabled a comparison of the depth effects in Lake Malawi with our previous results, which showed that gene expression differences change significantly with depth in Lake Victoria.

## Results

### *Variation in gene expression*

This is the first study where we have attempted to distinguished RH2Aα and β in qPCR studies from wild caught samples. The Rh2Aβ gene was either absent or weakly expressed (<3% of total opsin expression) in all samples. Therefore, the measure obtained from the aggregate RH2A primer/probe pair was considered indicative of RH2Aα expression. This makes sense in comparison to previous protein expression studies where the RH2Aα gene was shown to have longer wavelength

sensitivity (528 nm) compared to that of RH2Aβ (518 nm) (Parry *et al.* 2005; Spady et al 2006). RH2Aα gene expression is in agreement with MSP studies for Lake Malawi species where green cone peak absorbances were 525-535 nm (Jordan *et al.* 2006).

We found clear evidence that populations of the same species can differ in gene expression. Significant differences in gene expression were observed among populations in both *M. zebra and C. eucinostomus* (Fig. 3-2A,B), but not in *T gracilior* (Fig. 3-2C). Significant differences in the variability of gene expression within a population were found in both *M. zebra* and *T. gracilior* but not for *C. eucinostomus* (Fig. 3-3). More specifically, *M. zebra* displayed differential levels of variation in short-wavelength sensitive pigments (Fig. 3-3A) while *T. gracilior* had significant variation differences for both single cone (Fig. 3-3E) and double cone pigments (Fig. 3-3F).

Significant mean differences were observed between the Thumbi West and Zimbawe Rock populations of *M. zebra*. Expression differences were found for three genes (Fig. 3-2A). Both SWS1 ($F = 5.629$, $p = 0.011$) and RH2Aα ($F = 7.575$, $p = 0.013$) were expressed in greater quantities in the Thumbi West population, though there was no difference in the degree of variation between populations ($p > 0.05$). The fish from Zimbawe Rock expressed more of the violet-sensitive SWS2B gene than the fish from Thumbi West (Welch's ANOVA, $F = 13.474$, $p = 0.001$), and there was a larger variance of expression in the Zimbawe Rock population ($F = 11.500$, $p < 0.001$). Multivariate models could discriminate the fish from Zimbawe and Thumbi West based on relative opsin expression (Wilks' $\lambda = 0.286$, $p = 0.001$).

For *C. eucinostomus*, relative gene expression was qualitatively different among the populations. The fish from 5m depth expressed LWS while fish from 15m did not (Fig. 3-2B). Statistical differences were observed for both the green-sensitive RH2Aα ($F = 21.2$, $p < 0.001$; Fig. 3-2B), and the red-sensitive LWS ($F = 17.5$, $p < 0.001$; Fig. 3-2B) genes. Fish collected at 15m expressed more RH2Aα while the fish from 5m expressed more LWS. The higher level of RH2Aα in fishes from 15m was equal to the sum of RH2Aα and LWS in the shallow fish. For this reason, a multivariate model based on the expression of this pair of double-cone genes easily differentiates fish from the two depths (Wilks' $\lambda = 0.098$, $p < 0.001$). The qualitative LWS shift in *C. eucinostomus*, was similar to the SWS2B effect in *M. zebra* as the *C. eucinostumus* at depth did not express LWS while those at the surface did (Fig. 3-3C,D).

No depth-specific differences were found for cone opsin gene expression in *T. gracilior* ($p > 0.05$; Fig. 3-2C), and the multivariate model could not discriminate the two depth classes based on gene expression (Wilks' $\lambda = 0.705$, $p = 0.0517$). However, differences in the variance of gene expression were found for the UV-sensitive SWS1 gene ($F = 10.45$, $p = 0.003$), the blue-green-sensitive RH2B gene ($F = 23.47$, $p < 0.001$), and the LWS gene ($F = 6.285$, $p = 0.017$; Fig. 3-3E,F). Expression of all three genes had higher variance for fish collected at 20m than the fish collected at 5m.

*Variation in quantum catch*

Using light transmission spectra collected from downwelling light transmitted through the water column to calculate relative quantum catch for the different visual

pigments (eq. 1b), we found three distinct trends for changes in pigment quantum catch over the 1-15m range in which fish were collected for this study: (i) the SWS1 and SWS2B pigments became less efficient with depth, (ii) the RH2B and RH2Aα became more efficient, and (iii) the SWS2a and LWS pigments quantum catch had no change (Fig. 3-4). However, these changes in quantum catch were much smaller than what we would have predicted. We extended our depth series to 50m to estimate changes in efficiency at the suspected extremes of the distributions in *M. zebra* and *T. gracilior* (Ribbink *et al.* 1983; data not available for *C. eucinostomus*). Even over a depth range from 1 to 50m, the change in the relative efficiency of different pigments for four of the six cone opsins was proportionately small with the exception of SWS1 and SWS2B (Fig. 3-4). These results are further supported by calculations of the absolute quantum catch for each pigment (eq. 1a), as the attenuation rate of cone catches with depth was similar for all pigments with the exceptions of SWS1 and SWS2b (Fig. S3-2). This result is particularly surprising for the LWS pigment, as its absorbance curve overlaps a primary region of spectral attenuation with depth (Fig. 3-1).

Significant differences were found for gene expression weighted quantum catch, RT-$S_{sc}$ in *Metriaclima zebra* from Thumbi West and Zimbawe Rock (Fig. 3-5A). The additional SWS2B opsin expression increased both the RT-S and the variance of expression in the short wavelength sensitive single cones in the individuals from Zimbawe Rock in comparison to those at Thumbi West (Welch's ANOVA: $F = 19.566$, $p = 0.001$). This increase in short-wavelength sensitivity is accompanied by a decrease in both double-cone RT-$S_{dc}$ ($F = 19.768$, $p < 0.001$) and

total RT-$S_{all}$ ($F = 14.088$, $p = 0.001$) when compared with *M. zebra* from Thumbi West.

We calculated RT-S for *C. eucinostomus* and *T. gracilior* using light spectra measured for water depths of 5 and 15m, corresponding to the depths where the populations were sampled. Statistical tests found no difference between depths for single cones (RT-$S_{sc}$), double cones (RT-$S_{dc}$), or the sum of those mechanisms (RT-$S_{all}$) for either species ($p > 0.05$; Fig. 3-5B,C). While there was no variation between depths for the mean, there was a significant difference in the spread of the data between depths for *T. gracilior,* as individuals from 20m had significantly more variation in their double cone (5m: $F = 5.817$, $p = 0.024$; 15m: $F = 6.470$ $p = 0.016$) and total (5m: $F = 8.175$, $p = 0.007$; 20m: $F = 7.709$, $p = 0.001$) sensitivity. This is in accordance with the increased variance in the expression of RH2B and LWS in these fish.

We calculated relative quantum catch (eq 1b) for cichlid opsins in both Lake Malawi and Lake Victoria. The variation in relative quantum catch as a function of depth is shown in Fig 3-6A for several of the opsins. In Lake Victoria, the relative quantum catch of the LWS pigment increased rapidly with depth, while the SWS1 and SWS2B pigment quantum catch neared zero at depths as shallow as 5 m. By contrast, the clear waters of Lake Malawi allowed for relatively invariant pigment quantum catches to a depth of 10 m and beyond (Figs. 3-4, 3-6A). Therefore, there were no real changes in quantum catch as cichlids move to deeper depths in Malawi.

The changes in gene expression in *C. eucinostomus* caused very little difference in RT-S for shallow versus deep locations in Lake Malawi (Fig 3-5C).

However, there were significant differences in depth-dependent relative quantum catch between the two lakes. In light of these large differences, we wanted to see if this variation would cause significant differences in RT-S at different depths in Lake Victoria. These calculations did indeed find large differences in the double cone RT-$S_{DC}$ and total RT-$S_{all}$ for *C. eucinostomus* based on light environment from shallow (5m) vs deep (15 m) locations (Fig 3-6B).

*Discussion*

If visual sensitivities are important in mate choice and ultimately speciation, then we should be able to detect differences among populations and species. It has recently been shown that differences do occur among species (Hofmann *et al.* 2009). However, this work is the first demonstration that variation in opsin gene expression can occur within species on small spatial scales. We found pronounced qualitative gene expression differences in both *Copadichromis eucinostomus* and *Metriaclima zebra*, while quantitative variation was present in *Tropheops gracilior*. The *T. gracilior* and *C. eucinostomus* depth samples are in such close geographical proximity (~100m distance along the lake bottom) that they would typically be considered fully sympatric. As such, the degree of gene expression variation in Malawian cichlids may be of a much greater magnitude than we once thought.

In previous work, variation was examined in Lake Victoria and a link was found between gene expression and light environments (Hofmann *et al.* 2009). Fishes from locations with red-shifted ambient light environments expressed higher amounts of the medium- and long-wavelength sensitive RH2A and LWS genes. Large variation in SWS2 gene expression also occurred with changes in the SWS2B/

SWS2A ratio that correlated with quantum catch of the single cones as it varied over four orders of magnitude. Short-wavelength attenuation is so great in some locales that it renders single-cone photoreceptors essentially inactive, and this is reflected in the overall gene expression patterns (Hofmann *et al.* 2009).

Based on these previous results in Lake Victoria, we expected the differences we observe at local spatial scales in Lake Malawi to be linked to differences in local light environments. We began with three hypotheses that follow logically from the hypothesis that visual systems should be tuned to the local environment. We expected (i) greater variance in gene expression in fishes from shallow water, where the light spectrum is broader (ii) differences in opsin gene expression with depth as a result of differences in light stimulation of cone pigments, and (iii) no differences in gene expression among the two *M. zebra* populations as the populations occur at similar depths in locales with similar light environments. However, the gene expression patterns observed in this study are inconsistent with each of the three hypotheses we put forth.

Instead of observing increased variation in shallow water fishes, where the light spectrum is broader, we found evidence that gene expression is more variable in deep-dwelling individuals of *Tropheops gracilior*. Our initial hypothesis regarding the variance of gene expression was predicated on the idea that selection on visual systems at depth would be stronger due to spectral attenuation. More specifically, we expected genes outside of the primary three-cone colour palette would not be observed due to a penalty incurred by a decrease in quantum catch. Calculations of visual efficiency between the shallow *T. gracilior* and their relatively variable deep

counterparts indicate that variable gene expression does not cause a decrease in sensitivity due to the high transmission of light in Lake Malawi (Fig. 3-5E).

Perhaps more surprisingly than the nature of expression variance with depth is the observation that the presence of LWS in shallow *C. eucinostomus* would not cause a decrease in the predicted visual sensitivity of these fish if they were transplanted in deeper waters (Fig. 3-5C). The light attenuation with depth in Malawi is such that less than ~10% of the total LWS absorbance spectrum falls outside of the ambient irradiance at 15m. When combined with the 15-30% expression of LWS in the retina, this leads to a rough sensitivity decrease of less than 3% at depth. In the context of the natural range of RT-S variation between individuals, this decrease is statistically negligible. In other words, these data indicate that there is very little difference in the sensitivity of photoreceptors expressing LWS and RH2 opsins in Lake Malawi associated with depth (Fig. 3-4), and therefore the presence or absence of LWS in *Copadichromis eucinostomus* does not confer a sensitivity advantage to either group. Conversely, if these same *C. eucinostomus* were introduced into Lake Victoria where a pronounced sensitivity advantage is associated with the LWS pigment, we would expect the shallow individuals to display higher visual sensitivities. Our RT-S model supports this hypothesis, as the addition of the LWS pigment in the shallow fishes increases double-cone sensitivity relative to their deeper counterparts (Fig. 3-6B). In other words, while gene expression variation associated with total spectral sensitivity is apparent due to the steep light gradients in Lake Victoria, the relatively invariant light environment in Lake Malawi makes our initial sensitivity hypotheses less plausible (Fig. 3-5C).

We hypothesized there would be no expression difference between the *Metriaclima zebra* populations at Thumbi West Island and Zimbawe Rock given the similar ambient light environments in their respective habitats. On the contrary, the Thumbi West population expressed a typical short-wavelength sensitive palette (SWS1, RH2B, RH2A), whereas the Zimbawe Rock population expressed a modified tetrachromatic palette (SWS1, SWS2B, RH2B, RH2A). Perhaps most surprisingly, the *M. zebra* populations represented the only system we observed in which our RT-S models were able to predict a specific difference in chromatic sensitivities between the two groups. The addition of the SWS2B pigment in the Zimbawe Rock population augments the total single-cone catch in these fish. However, given the similar ecology of these two populations, we are currently unable to tie this sensitivity difference to any known environmental or behavioral correlate.

Comparisons of gene expression in both *C. eucinostomus* and *M. zebra* from two locales showed a change from expression of three visual pigments to four. This could have profound implications for cichlid visual perception. Both represent modifications of the classic short trichromatic UV, blue-green, and green template observed in many of the mbuna (Hofmann *et al.* 2009). In *C. eucinostomus*, LWS is added to the double-cone complement, creating a three pigment double-cone system. This extends the theoretical sensitivity of the *C. eucinostomus* visual system further into the red. Conversely, the addition of SWS2B to create a single-cone complement with two distinct pigments in *M. zebra* would not broaden the theoretical total spectral range of visual sensitivity. Rather, it broadens single-cone sensitivities in the direction of the double-cones, effectively augmenting sensitivity in intermediate

spectral ranges. However, more important than the broadening of spectral sensitivities is the potential for contrast (opponent) mechanisms that arise from the addition of these pigments. If these pigments are wired opponently, they may create a tetrachromatic visual system from the original trichromatic one. Tetrachromatic opponency was recently demonstrated by Lisney *et al.* (2010) through electrophysiological recordings in the tilapia *Oreochromis niloticus*. They found three potential opponent mechanisms based on four cone pigments present in this cichlid retina, including an opponent interaction between the SWS2B and SWS2A pigments.

Two primary facets of the expression data described here suggest that ongoing selection for divergent visual systems within species is plausible in Malawian cichlids: (i) the highly-variable opsin expression in the deep *T. gracilior* indicates that there is sufficient sensory variation available for selection to act upon, and (ii) the localized tetrachromatic visual systems in the other two species demonstrate that there is profound sensory diversification within the lake. These results are consistent with both sensory drive and Fisherian selection and drift. However, we are not currently able to tie these data to any known environmental or behavioural variables that would exert definitive selection. Prior work has demonstrated a correlation between feeding mode and SWS1 expression in Lake Malawi cichlids (Hofmann *et al.* 2009). The observation of SWS2B expression in *M. zebra* from Zimbawe Rock is actually contrary to this idea. Strong currents at Zimbawe Rock prevent food items from settling out of the water column as they do at Thumbi West (Kocher, personal communication). Therefore, *M. zebra* from Zimbawe should be more planktivorous

than their counterparts at Thumbi, and we would expect selection to drive a SWS1-dominated visual system at this site (Hofmann *et al.* 2009). However, *M. zebra* at this site have more SWS2B, and hence less SWS1expression than their counterparts at Thumbi. As such, evidence for the sensory drive hypothesis remains weak in the Malawi cichlids. Therefore, Fisherian mechanisms or neutral drift may be the more plausible explanations for phenotypic divergence in Lake Malawi.

The tetrachromatic visual systems in particular support the Fisherian runaway or neutral drift hypotheses. The addition of a fourth major pigment to the retinal mosaic is a considerable change to the visual system as a whole. Such a dramatic shift in expression patterns suggests that stabilizing selection on the trichromatic short palette is fairly relaxed, and relaxed selection may result in diversification through random drift. The derivation of divergent phenotypes that result from drift without any explicit environmental consequences is amenable to the development of subsequent Fisherian selection. Therefore, we posit that in the clear waters of Lake Malawi intraspecific visual diversification largely functions outside sensory drive and results from less rigid selective processes. In order to test this hypothesis, we hope to identify the loci controlling opsin gene expression (Carleton *et al.* 2010) so that we can then test these loci for evidence of selection or drift.

This study demonstrates that cone opsin gene expression in the cichlids of Lake Malawi is much more complex than previously thought. The magnitude of expression variation differs in fishes occupying different local habitats. Also, distinct qualitative differences in gene expression are present in some species. Taken together, these observations offer the possibility that Fisherian sexual selection may

be linked to colour vision in Malawian cichlids. Furthermore, the observation of two distinct tetrachromatic visual systems in different species presents a novel view of how visual systems adapt. Rather than switching from one trichromatic visual palette to another, visual tuning may occur by adding expressed genes to the current visual palette. This system has profound implications for both ecology and neuroscience, as the logistics of collecting and processing tetrachromatic information can be quite different than those for a trichromatic system (Zana *et al.* 2001). As such, visual divergence in Malawian cichlids may function through genetic and neural mechanisms that remain to be found.

*Figures*

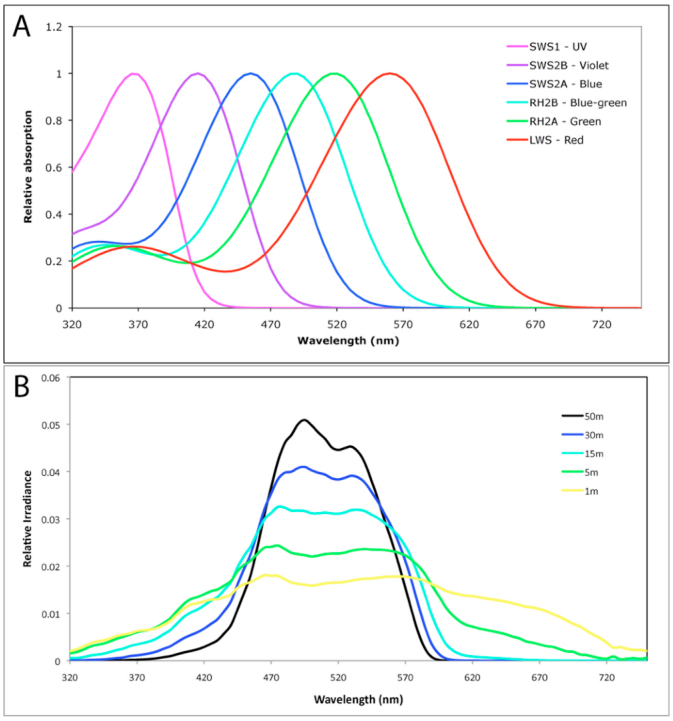



**Figure 3-1:** Spectral absorbances of the cichlid cone opsins and the corresponding ambient light environment in Lake Malawi. (A) Idealized absorbance curves for the six cone opsin genes investigated in this study. (B) Modeled spectral irradiance curves for 1m, 5m, 15m, 30m, and 50m depth at Zimbawe Rock.

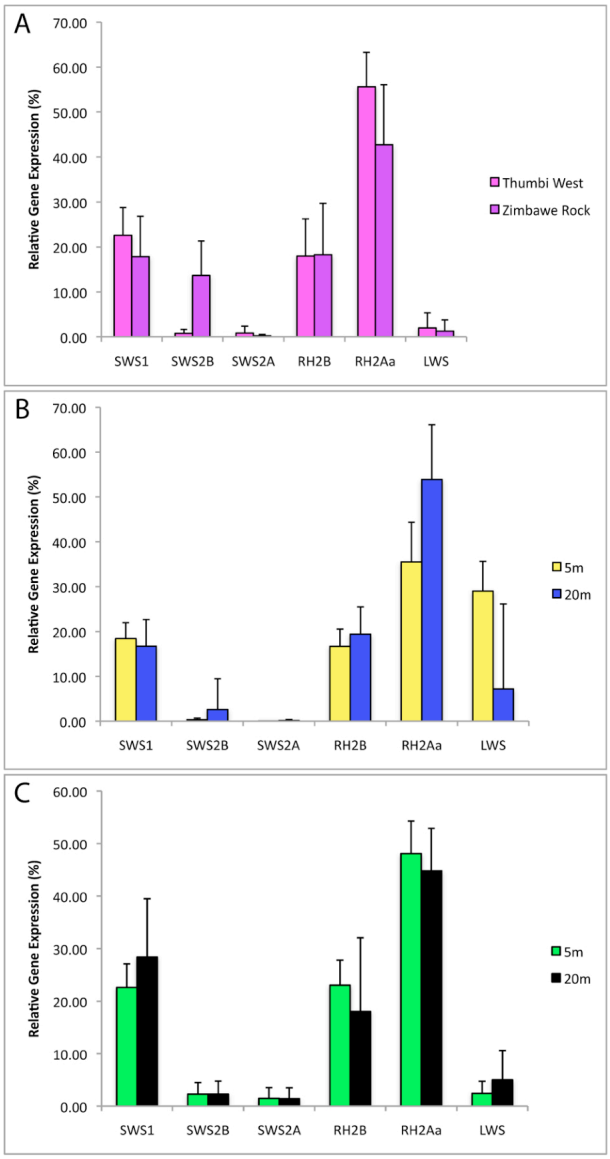



**Figure 3-2:** Mean relative opsin expression patterns for populations of (A) *Metriaclima zebra*, (B) *Copadichromis eucinostomus*, and (C) *Tropheops gracilior*.

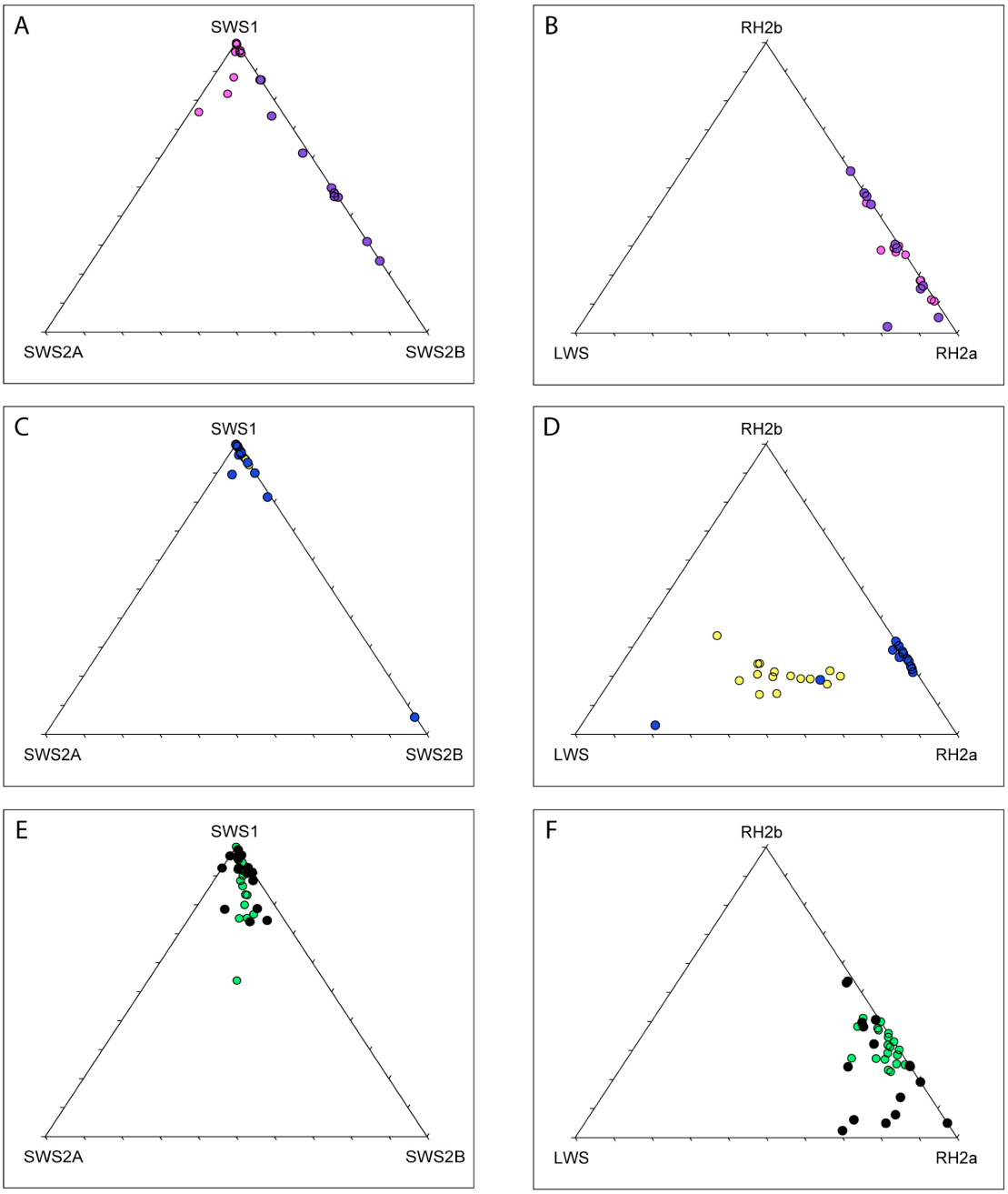



**Figure 3-3**: Gene expression triplots for all individuals included in this study. (A) Single-cone and (B) double-cone pigments for *Metriaclima zebra* (pink = Thumbi West, violet = Zimbawe Rock). (C) Single-cone and (D) double-cone pigments for *Copadichromis eucinostomus* (yellow = 5m, blue = 15m). (E) Single-cone and (F) double-cone pigments for *Tropheops gracilior* (green = 5m, black = 20m).

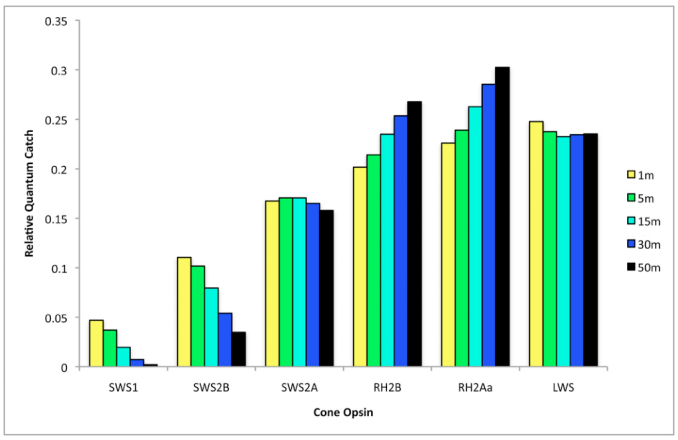



**Figure 3-4:** Relative quantum catch of the cichlid visual pigments at a series of five depths at Zimbawe Rock. The sum of gene expression for single cone genes (SWS1, SWS2B and SWS2A) and for double cone genes (RH2B, RH2A and LWS) are normalized to one.

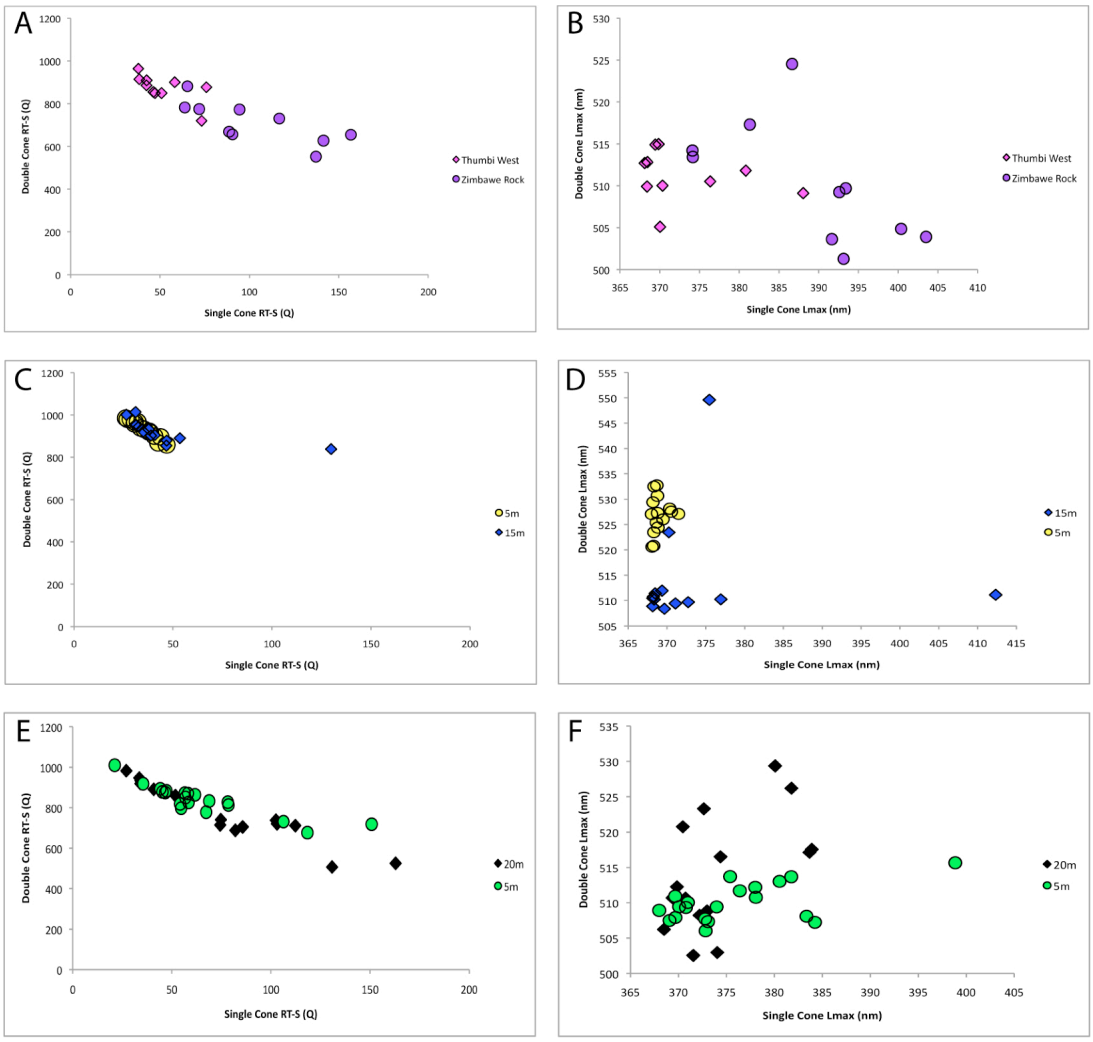



**Figure 3-5:** Comparison of the expression-weighted RT-S models at 5m depth applied in this study to the predicted single- or double cone maximal absorbances calculated in previous studies (Hofmann et al. 2009). (A) RT-S and (B) predicted λmax calculations for *Metriaclima zebra*. (C) RT-S and (D) predicted λmax calculations for *Copadichromis eucinostomus*. (E) RT-S and (F) predicted λmax calculation for *Tropheops gracilior*.

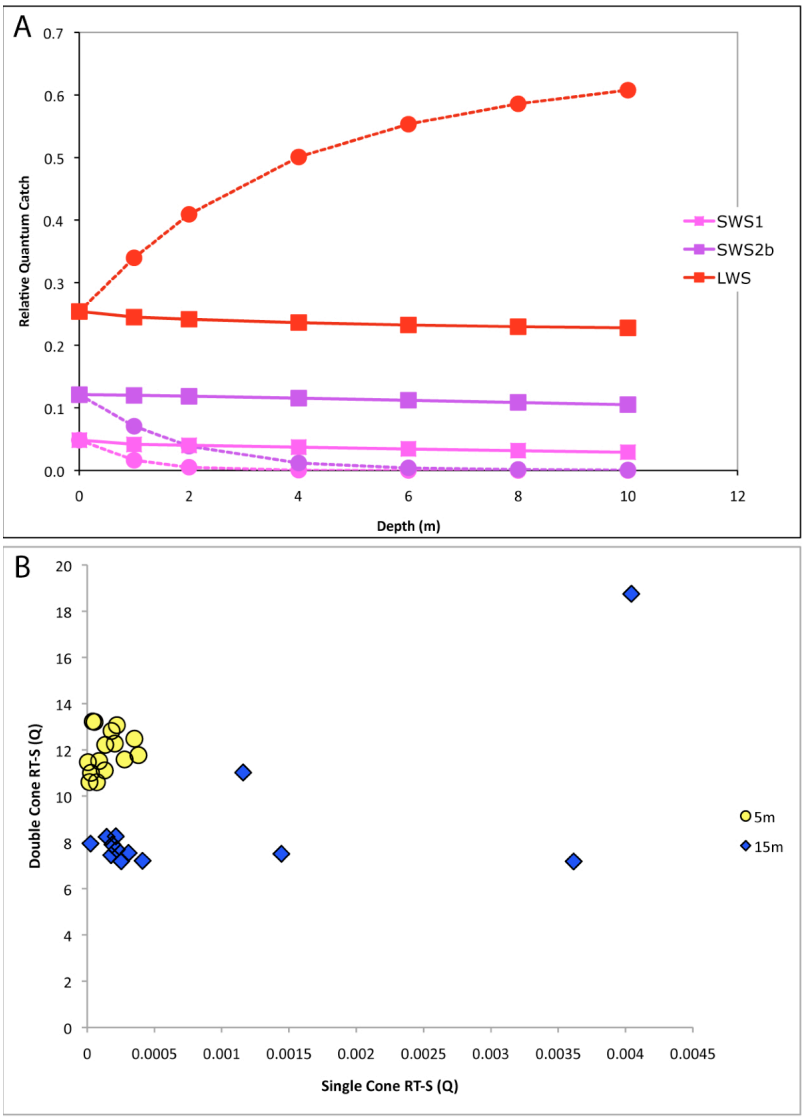



**Figure 3-6:** Comparison of the biological effects of light gradients in Lake Malawi and Lake Victoria. (A) Variation in relative quantum catch for the SWS1, SWS2B, and LWS pigments from 1-10m depth in Lake Malawi (squares with solid lines) and Lake Victoria (circles with dashed lines). Quantum catches are normalized to the full six-cone template as in Figure 2, but SWS2A, RH2B, and RH2Aα have been excluded for clarity. (B) Comparison of RT-S models for *Copadichromis eucinostomus* from 5m (yellow) and 15m (blue) under the ambient light environment

observed at 10m depth in Lake Victoria. This should be compared to panel 4C for the equivalent comparisons in the light environment of Lake Malawi.

# Chapter 4:

# Relative LWS Cone Opsin Expression Determines Optomotor Thresholds in Malawi Cichlid Fishes

Adam R. Smith, Ke Ma, Daphne Soares, Karen L. Carleton

See Appendix 3 for all supplemental materials referenced in this chapter

*Abstract*

Associating quantitative genetic traits with quantitative behaviors is a relatively unexplored region of sensory neurobiology. The visual system is an ideal place to test models associating these levels of sensory perception. In this study, we reared cichlid fishes from Lake Malawi in different ambient light environments. We then tested the visual sensitivities of these fish using the optomotor response (OMR) behavioral paradigm and measured the relative expression of cone opsin genes. We found that the light environment experienced by fishes during development can alter gene expression, particularly as it applies to the long wavelength-sensitive (LWS) opsin gene. Also, fishes from different rearing conditions exhibited different behavioral sensitivities. We combined these data with predictions of opsin pigment absorption by the different OMR stimuli to determine which cone types are most likely to influence the OMR behavior. While we hypothesized that this behavior would be controlled by a random-wiring model reflecting the expression of both medium wavelength sensitive (MWS) and LWS opsin, our models suggest that only the LWS pigment is required to predict behavior. Furthermore, analyses demonstrate that LWS expression variation accounts for ~20% of the observed behavioral variance. This work confirms that sensory gene expression influences behavior in a predictable fashion. It also suggests that the neural wiring of basal visual pathways in cichlid fishes may differ from that observed in mammals and zebrafish, but is similar to that described in goldfish. This finding has important implications for the evolution of the magnocellular neural pathway in teleosts.

*Introduction*

Sensory perception is defined by the compound effects of genetic, neural, and behavioral processes. Signals are first transduced to molecular and electrical signals in receptor cells. Following a series of neural processing stages in both the sensory organ and the brain, a final behavioral output is generated. Understanding the interaction between these signal-processing levels is a fundamental question in sensory neuroscience, and is of considerable import in evolutionary biology (Dangles *et al*. 2009).

The visual system has been the subject of some seminal studies in neuroscience, such as the neuronal tracings of Ramón y Cajal (c1900) and the electrophysiological studies of Hubel & Wiesel (1962). The rhodopsin protein was the first G-protein coupled receptor to have its physical structure determined (Palczewski *et al*. 2000). As such, an enormous amount of information is present regarding the characteristics of the visual transduction pathway. This makes the visual system amenable to studies relating multiple layers of sensory transduction processes.

A classic behavioral paradigm for studying visual sensitivities is the optomotor response (OMR), which exploits the tendency of an animal to follow a moving grating when no stronger stimuli are present. This model can be used to study visual acuity (Neumeyer 2003), photic sensitivity (Anstis *et al*. 1998), differential cone mechanisms in the visual response (Orger & Baier, 2005), and even the perception of first- and second-order signals (Orger *et al*. 2000). Cone inputs to the OMR are of particular interest in teleost systems, as studies have implicated different cone types as being necessary for the behavior. Generally, the OMR is thought to

represent the summed activity of long (LWS)- and medium (MWS)- wavelength sensitive cones in zebrafish (Orger & Baier 2005). However, Schaerer & Neumeyer suggested that the goldfish OMR is dominated by the LWS cone mechanism (1996). This suggests fundamental differences in the wiring of the magnocellular visual pathway, and deserves further investigation in alternate systems.

Lake Malawi cichlids are an ideal model system for the study of visual systems and evolution, due to their highly elaborated cone opsin gene complex (Hofmann *et al*. 2009). Researchers have proposed behavioral shifts that should accompany changes in gene sequence and expression (Hofmann *et al*. 2009), and variation in OMR sensitivities between cichlid sister taxa from Lake Victoria support these hypotheses (Maan *et al*. 2006). Smith *et al.* (2011) proposed a model predicting luminance-sensitivity of the cichlid visual system based on relative opsin gene expression and opsin pigment absorbance templates. In this study, we tested one component of that model: the luminance sensitivity of the medium- (RH2A & RH2B) and long-wavelength sensitive (LWS) double cones. The OMR is ideal for testing this model, as it relies on the luminance response of these cone types. We hypothesized that (i) behavioral sensitivities would reflect a random-wiring model equivalent to the relative expression values of all double-cone opsins combined, and (ii) fishes expressing relatively more LWS would have lower thresholds for long-wavelength dominant stimuli while fishes expressing more RH2A and B would be more sensitive to short-wavelength stimuli.

*Materials and Methods*

*Experimental subjects*

Thirty *Metriaclima lombardoi* and ten *Melanochromis auratus* were used for this study. Fishes were raised under two light rearing treatments: (i) one composed of two relatively broad-spectrum bulbs designed to mimic downwelling spectra in aquatic environments and (ii) one composed of narrow-spectrum (standard fluorescent) lighting (light spectra depicted in Fig. S4-1). Pilot developmental studies indicated that the two broad-spectrum bulbs had the same effect on gene expression through ontogeny (unpublished data). *M. lombardoi* broods were split before the age of ten days post-fertilization, with fifteen individuals being raised under each lighting treatment and then behaviorally tested. All *M. auratus* individuals were reared under the broad-spectrum lighting, as pilot experiments indicated there was no difference in gene expression between young raised in the narrow or broad lighting environments for this species (unpublished data). No more than three broods were used for each experimental grouping to minimize effects of family structure. The majority of the *M. lombardoi* in this study were full- or half- sibs, with a small number of cousins to supplement the narrow-spectrum group. All *M. auratus* individuals were full-sibs. All subject were reared and sampled under the guidelines of the University of Maryland IACUC (protocol #R-09-73).

*Experimental arena*

Fish were placed in an aquarium above a monitor used to generate the optomotor grating stimulus. The stimulus was controlled by a Macbook laptop (Apple Inc., Cupertino, CA, USA) and projected on a 61cm white-LED backlight

computer monitor with a 60Hz refresh rate (model B243HL-bmdrz, Acer America Corporation, San Jose, CA, USA). The monitor was placed face-up on a small acrylic table with an acrylic plate supported above it with ~5cm of space between the monitor screen and the tabletop. This enabled us to place colored plastic filters above the monitor during trials. A ten gallon aquarium (50.8cm x 26.5cm x 32cm) was centered above the monitor on the acrylic sheet, and an acrylic arena (46cm x 15.5cm x 14cm) with black sides was placed within the aquarium to minimize external visual cues. The tank was filled with water to 9-10 cm depth. Two white LED desk lamps (16-LEDs, total power ~60 μW at water surface, Dekcell.com) provided adapting light for the arena. These lights were attached to and powered by the control computer through a USB hub (Targus Group International Inc., Anaheim, CA, USA). A camcorder equipped with an infrared emitter (Sony HDR-CX550V, Sony Corporation of America, New York, NY, USA) was mounted 65cm above the aquarium to image the trial. Half of the camcorder lens was occluded with a long-pass filter with a cut-on of 780nm (FSQ-RG780, Newport Corporation, Irvine, CA, USA). This filter allowed for the arena with the subject to be imaged clearly under IR while the projected stimulus was visible in the other half of the visual field. The trial was viewed on an external monitor connected to the camcorder and positioned next to the control computer. The control station was isolated from the test arena by a black curtain. The apparatus is depicted in Fig. S4-2.

*Stimulus design and presentation*

Four spectral test stimuli targeting cichlid MWS and LWS cone types were defined using colored gel filters (Lee Filters USA) placed over the presentation

monitor. Henceforth, the four filters will be referred to by the primary color of the transmission spectra: red (HT109), green (HT139), blue-green (HT116), and blue (HT142). All fishes were tested under the red, green, and blue-green stimuli. Most of the individuals were also tested under the blue filter stimulus (broad-spectrum *M. lombardoi* n=5; narrow-spectrum *M. lombardoi* n=11; *M. auratus* n=9). The radiance spectra for each monitor/filter combination (Fig. 4-1) were measured using an Ocean Optics USB4000 spectrophotometer equipped with a bare collection fiber (Ocean Optics, Dunedin, FL, USA). The spectrometer was radiometrically calibrated with an LS-1 lamp traceable to the National Institute of Standards (Ocean Optics). This calibration converted radiance to photons/$cm^2$ s. Acceptance angle in steradians was constant across all measures due to the use of a single fiber.

A black and white sinusoidal grating presentation was designed using the modular program VisionEgg (Straw 2008). The program (which is available in full as supplemental material) used parameters to generate a barred pattern with an average bar width of 20mm, corresponding to a spatial frequency of 12.5 degrees. The grating moved across the screen at a temporal rate of 5.0 Hz, and switched movement direction every ten seconds. Each individual stimulus object consisted of four directional presentations (two forward, two reverse) for a total presentation time of forty seconds. Presentation objects of different intensities were selected from the keyboard, so that the intensity gradient could be manipulated by the experimenter. Three consecutive objects of decreasing intensity were sequentially presented over a total time of 120 seconds. The white adapting lights were turned on for 60s after each triplicate presentation to keep the fish light-adapted.

*Threshold determination*

At the beginning of a trial, a fish was removed from community housing and placed in the arena. The adapting lights were turned on while a stationary grating stimulus was projected on the monitor underneath the arena. The first color filter was put in place and the curtain was closed. Following a thirty minute acclimation period, the camcorder was turned on and the trial was started. Each trial consisted of three or four color stimuli. For each stimulus, the subject was exposed to a series of grating sequences of decreasing intensities. Threshold was determined using a modified stair-step experimental design. There was a ten minute acclimation period between each color stimulus. The presentation order of the stimuli was randomized for each trial.

In order for a subject to score a "positive" response for a stimulus, it had to follow at least three of the four presentations that occurred within each directional grating sequence. The initial triplicate presentation consisted of a high, intermediate, and low intensity grating sequence. The lowest intensity grating of these three to which the subject responded was then taken as the baseline for the presentation sets. This intensity was used as the intermediate intensity for the following triplicate presentation. In the following presentations, if the subject scored correctly on only the first (highest intensity) stimulus, the next triplicate presentation would be increased by one intensity step. If the subject responded correctly to the intermediate stimulus, the same triplicate presentation would be repeated. Finally, if the subject responded correctly to the lowest intensity stimulus, the next triplicate stimulus would be lowered one intensity step (procedure summarized in Fig. S4-3). Threshold was defined as the point where the subject responded correctly to the intermediate but not

the lowest intensity stimulus on at least two consecutive triplicate presentations. Once the threshold stimulus had been determined, the recording was stopped and the camcorder IR light was turned off. The stimulus was then played again and the light power (nW) was recorded perpendicular to the monitor's surface just above the water surface with a Newport Optical Power Meter (Model #1916-C, Newport Corporation, Irvine, CA, USA). The next color stimulus was then tested following a 10 minute acclimation time. Once the thresholds for all color stimuli were measured, subjects were sacrificed immediately and their retina were collected to quantify gene expression.

*Gene expression*

Gene expression analysis was performed using techniques laid out in previous work (Spady *et al*. 2006; Smith *et al*. 2011). Gene expression was quantified for the six classes of opsins including SWS1, SWS2B, SWS2A, RH2B, RH2A and LWS. All RH2A gene expression was assumed to represent the RH2Aα gene, as prior work has found no evidence that the RH2Aβ gene is expressed (Smith *et al.* 2011). RNA was extracted from retinal samples using commercial Qiagen Qiashredder and RNeasy kits (Qiagen, Valencia, CA, USA) and the isolated RNA was transcribed to cDNA. Real-time quantitative PCR (RT-qPCR) was performed using six gene-specific primer/probe pairs, and all individuals were run in parallel. All individuals were replicated at least twice on separate plates. The RT-qPCR reaction and critical-point determination was performed on a Roche Lightcycler 480 (Roche Diagnostics, Indianapolis, IN, USA). Relative primer/probe efficiencies for each gene were estimated using a plasmid construct (Spady *et al*. 2006). The efficiencies and the

critical cycle data were used to estimate relative cone opsin gene expression. Total expression values for the medium- and long-wavelength pigments (representing the double-cone cell types) were normalized to sum to one, as were the values for the SWS pigments (single-cone cells). The relative proportion of each of these different genes was then denoted as $f_{gene}$ for later analyses.

*Statistical Analyses*

Absolute sensitivity thresholds and relative gene expression data were compared across the treatment groups using ANOVAs and Tukey's HSD post-hoc comparisons. For expression tests, the RH2A gene (which is ubiquitously expressed at high levels) was excluded to alleviate linear dependence in the dataset. For the SWS cones, only the developmental treatments within *Metriaclima lombardoi* were compared, as *M. lombardoi* and *Melanochromis auratus* are already known to express different primary SWS palettes (Hofmann *et al*. 2010).

Models for behavioral predictions were derived and modified from those first proposed by Smith *et al*. (2011). The sensitivity spectra of cone opsin pigments were modeled according to Govardovskii *et al*. (2000). Quantum catch for each pigment for each light stimulus was modeled using an equation modified from Hofmann *et al*. (2009):

$$Q_{abs,i} = \int_{300}^{700} I(\lambda,g) R_i(\lambda) d\lambda \qquad (1)$$

Here $Q_{abs,i}$ is the quantum catch of a given photoreceptor visual pigment , $I(\lambda,g)$ is the radiance of the monitor (in photons/cm$^2$ s) after passing through the color gel filter, g, and $R_i(\lambda)$ is the absorbance of one of the visual pigment based on the equations of Govardovskii *et al.* (2000). The quantum catch were calculated for single-pigment

models for the RH2B, RH2A, and LWS visual pigments. The lambda max used for these three cichlid visuals pigments were 488 (RH2B), 518 (RH2A) and 560 (LWS) nm. For multiple-pigment prediction models, the sum of the visual pigment absorbances was used. For example, to include both RH2A and LWS visual pigments, the combined quantum catch is:

$$Q_{abs,i} = \frac{\int_{300}^{700} I(\lambda,g)[R_{RH2A}(\lambda) + R_{LWS}(\lambda)]d\lambda}{x} \quad (2)$$

where $x$ is the number of genes incorporated into the model. The value of $x$ prevents the model from predicting a quantum catch greater than the available light in the system.

The behavioral threshold from the green color stimulus was used as the independent predictive variable (*i.e.* the green behavioral measure was used to predict behaviors under other conditions), as fishes were generally most sensitive to the OMR stimulus under this light. Three primary single-cone models were constructed according to the absorbance spectra of the LWS, RH2A, and RH2B pigments. These models were based on the assumption that behavioral sensitivity thresholds are directly proportional to the relative photoreceptor quantum catch for that light spectrum normalized by the total light spectrum. Therefore, the single-cone models took the form:

$$S_t = \left( \frac{Q_{abs,t}}{\int_{300}^{700} I(\lambda,t)d\lambda} \right) \left( \frac{S_{green}}{\frac{Q_{abs,green}}{\int_{300}^{700} I(\lambda,green)d\lambda}} \right) \quad (3)$$

Here $S$ represents the predicted behavioral threshold under a given color stimulus, $Q_{abs}$ represents the quantum catch of the modeled pigment under that stimulus and the integral of I representing the total light intensity for either the green or other color stimulus. In order to test the Real-time expression corrected (RT-S) models, another predictive model was developed that weights the predicted inputs from each cone type by their gene expression fraction such that:

$$S_{RT-S} = f_{LWS}S_{LWS} + f_{RH2A}S_{RH2A} + f_{RH2B}S_{RH2B} \quad (4)$$

where $f_{gene}$ is the relative expression coefficient for the gene as part of the double-cone expression system (Smith *et al*. 2011). The efficacy of the models compared to actual behavioral measures was tested using ANOVAs and Tukey's HSD tests, while multivariate correlation analyses were used to define the correlation matrix between expression measures and behavioral thresholds.

<u>*Results*</u>

*Behavioral thresholds*

A significant difference in behavioral sensitivity thresholds across groups was found for two of the four color stimuli. For the red stimulus, the broad-spectrum *Metriaclima lombardoi* and the *Melanochromis auratus* both had significantly lower behavioral thresholds than the *M. lombardoi* raised under narrow-spectrum lighting ($F = 5.52$, $p = 0.008$, Fig. 4-2a). *M. auratus* performed significantly better for the blue stimulus than either of the *M. lombardoi* ($F = 4.78$, $p = 0.018$, Fig. 4-2b). No significant difference was found amongst the three groupings for either the green ($F = 2.06$, $p = 0.140$, Fig. 2c) or blue-green ($F = 1.28$, $p = 0.291$, Fig. 2d) stimuli.

*Cone opsin gene expression*

Significant differences in the expression of the LWS gene were found between the light rearing treatment groups, with narrow-spectrum *Metriaclima lombardoi* expressing more LWS than their counterparts reared under broad-spectrum lighting. *Melanochromis auratus* expressed LWS at the same level as the narrow-spectrum *M. lombardoi* ($F = 15.34$, $p < 0.001$, Fig. 4-3a). No significant difference was found in the expression of the RH2B pigment ($F = 1.29$, $p = 0.287$). For the SWS cone complement, there was a significant difference in the expression of the violet-sensitive SWS2B gene, with broad-spectrum *M. lombardoi* expressing more SWS2B than narrow-spectrum *M. lombardoi* ($F = 10.94$, $p = 0.003$, Fig. 4-3b). No significant difference was found in the expression of the SWS1 ($F = 0.08$, $p = 0.786$) or SWS2A ($F = 2.64$, $p = 0.115$) pigments.

*Predictive models and cone inputs*

We examined how well the different predictive models determined the threshold responses under blue, blue-green, and red stimuli as compared to the green stimulus (see equation 3 and 4). The RH2B single-pigment model produced very poor fits and will not be considered further. The performance of the predictive models based on RH2A or LWS single-pigments was strongly contingent on both the color stimulus and the cone inputs. Overall, only the LWS single-pigment model predicted the experimentally measured behavioral thresholds for the red light stimulus for all three subject groups (*M. lombardoi* broad, $F = 45.59$, $p < 0.001$, Fig. S4-4a; *M. lombardoi* narrow, $F = 222.87$, $p < 0.001$, Fig. S4-5a; *M. auratus*, $F = 315.\ 26$, $p < 0.001$, Fig. 4-4a). For this color stimulus, the RT-S model always underestimates the performance of the fish (*i.e.* the predicted behavioral thresholds are of higher intensity

than those observed). For the blue-green stimulus, the RT-S and RH2A models correctly predicted behavioral thresholds for all three groups, while the LWS-only model predicted behavior for the narrow-spectrum *Metriaclima lombardoi* (*M. lombardoi* broad, $F = 9.705$, $p < 0.001$, Fig. S4-4b; *M. lombardoi* narrow, $F = 4.05$, $p = 0.005$, Fig. S4-5b; *M. auratus*, $F = 46.96$, $p < 0.001$, Fig. 4-4b). For the blue stimulus, only the LWS-specific model correctly predicts behavior for the broad-spectrum *M. lombardoi* and *Melanochromis auratus*, while no models were able to accurately predict behavior for the narrow-spectrum *M. lombardoi*, with the LWS model being the closest (*M. lombardoi* broad, $F = 10.92$, $p < 0.001$, Fig. S4-4c; *M. lombardoi* narrow, $F = 30.79$, $p < 0.001$, Fig. S5c; *M. auratus*, $F = 63.72$, $p < 0.001$, Fig. 4-4c). In figure 4-5, we have translated the total mean values for all behavioral measures and the predictive models to sensitivity tuning curves to illustrate the conformity of different models to both the shape and mean of the behavioral measures.

*Correlation between gene expression and behavior*

To examine the relationship between gene expression and behavior, a multivariate correlation analysis was done on the complete pooled dataset rather than on the individual subject groupings, due to sample size requirements (data shown in Fig. S4-6). Similarly, multivariate analyses were necessary to determine the fits of single-pigment models, since the linear correlation coefficients for all single-pigment Govardovskii models are identical in univariate analyses despite conformity to the mean (Fig. S4-7). In accordance with the predominance of the LWS models in the predictive study and the developmental relationship between LWS expression and

behavior, a strong correlation was observed between LWS expression and the color stimuli thresholds. For the red, green, and blue-green stimuli, there was a roughly 20% correlation coefficient (calculated at $p < 0.05$) between quantitative variance in LWS expression and behavioral thresholds. This effect was weaker for the blue treatment, at only 8%. The full correlation matrix is available in Table S4-1.

*Discussion*

While prior work has tested the link between opsin gene expression and complex foraging and mate-choice preferences (Fuller *et al*. 2010; Fuller & Noa 2010), this is the first study to investigate the quantitative relationship between the relative expression of multiple sensory receptor genes and behavioral sensitivity thresholds. This work indicates a pronounced link between these levels of the visual processing cascade. Variation in the expression of the LWS cone opsin gene in particular is correlated with the performance of the optomotor response behavior. Fishes expressing relatively more LWS generally performed better, with lower absolute OMR thresholds than their counterparts. From a genetic standpoint, the correlation between LWS and the observed behavior is substantial. In the case of morphological traits with multiple quantitative genetic underpinnings, a 20% association would be considered quite large (Beavis 1998). The strength of this relationship is even more surprising given the numerous molecular, cellular, and neural processes that occur between receptor expression and behavior.

Prior work described plasticity in opsin gene expression patterns between wild fishes and lab-reared offspring (Hofmann *et al*. 2010), but our data indicate that this effect can be mediated by manipulating the ambient light environment fishes

experience during development in the laboratory. Compared to other opsins, LWS expression is particularly plastic. By rearing the fish in broader lighting which includes a UV component, opsin gene expression is closer to that found in the wild. For example, the LWS measures for broad-spectrum *Metriaclima lombardoi* are statistically indistinguishable from the wild-caught fishes measured by Hofmann *et al*. (2010; ANOVA, $F = 0.21$, $p = 0.648$). This expression plasticity is mirrored by behavioral plasticity, suggesting that both traits are partially mediated by the environmental conditions a cichlid experiences during development. Furthermore, the differential expression patterns tested in this study are similar to examples of intraspecific expression diversity found in Lake Malawi (Smith *et al*. 2011). Therefore, we can extrapolate our behavioral data to fishes in the lake. However, it is important to note that we were unable to fully recover the wild-type gene expression patterns with altered lighting in the lab. This may be the result of our lab conditions having intensities lower than those found in the wild (~5-50 lux in the lab depending on tank placement versus ~5000 lux at 10 m depth in the field). Alternatively, it may indicate additional mechanisms independent of ambient light environment are mediating expression plasticity in these species.

Entering this study we hypothesized that (i) the optomotor response would demonstrate LWS and MWS inputs contribute to the OMR in a random-wiring model, and (ii) differential opsin gene expression would result in lower thresholds for long-wavelength stimuli in fishes expressing more LWS while fishes expressing more MWS would have lower thresholds for shorter-wavelength stimuli as a result of this wiring pattern. However, our predictive models do not indicate that random wiring of

cone types controls the OMR, and our data similarly suggest that our latter hypothesis was consequently incorrect. Rather, it appears that the LWS cone mechanism predominates in the system, with increased LWS expression conferring increased sensitivity to both long- and short-wavelength dominant stimuli. Using the idealized LWS absorption template and a single behavioral measure, we were able to accurately predict the behavioral response of fishes at the long-wavelength end of the spectrum. This LWS-dominant model accurately predicted behavioral sensitivities under the shortest-wavelength stimulus in two of the three test groups. Meanwhile the RT-S model, which represents a random wiring pattern based on relative gene expression, generally performed quite poorly.

The exception to this pattern was found for the blue-green stimulus where the RT-S model performed better than the LWS absorption template. This may be explained by the spectral proximity of the green and blue-green filter treatments. The behavioral thresholds for these two filters were similar for most fishes, and the pigment models that would result in similar measures for these treatments would be those for the RH2A and RT-S. The decay of the LWS pigment absorption template predicts a sharper decline in behavioral sensitivities than is observed in the data. However, given natural variation in OMR performance, the detection of a sharp peak around the spectral region of greatest sensitivity seems unlikely. It is well known that behavioral thresholds are generally more sensitive than those observed in electrophysiological data for the auditory system (Heffner et *al.* 2008). The magnitude of this effect is not necessarily consistent across the full sensory range of an individual. Generally, this phenomenon is often attributed to accessory cues from

alternate sensory modalities. Given that the OMR threshold is partially mediated by alternative sensory cues (such as those from the lateral line), it could be hypothesized that sensitivity curves would be "flatter" than a pure Govardovskii curve at the region of highest sensitivity. This is indeed the effect we observe when we treat the mean predictions of the models and the behavioral measures as sensitivity tuning curves (Fig. 4-5), further supporting the idea that the LWS pigment dominates the OMR response.

The characteristics of the OMR behavior are consistent with the response characteristics of the magnocellular visual processing pathway. A wealth of previous work has shown that this pathway sums information from both LWS and MWS cones in primates (Sun *et al.* 2006) and zebrafish (Orger & Baier 2005). This finding translates to the description of the optomotor response as being a color-blind luminance pathway. However, Schaerer and Neumeyer (1996) found that the OMR in goldfish seemed to be controlled through a monochromatic mechanism resulting from the activity of only the LWS cone type. Our results in the cichlid fishes of Lake Malawi suggest that their OMR is best explained by an LWS cone mechanism similar to that in goldfish, but dissimilar to the results obtained from other model systems. Given this, it is possible that at some point in the evolutionary history of teleost fishes a change occurred in the basic wiring properties of a basal visual processing pathway. This hypothesis could have profound implications for understanding the visual ecology of teleosts, which are often used as model systems to study the evolution of sensory adaptation and signal recognition. Therefore, the comparative biology of the teleost magnocellular system may be a promising avenue of future research that needs

to be addressed with a wider range of molecular, physiological, and behavioral experimental techniques.

As a model for sexual selection based on male coloration, the visual biology of African cichlids has been important in the formulation of sexual selection hypotheses (Dominey 1984). A wide range of research investigating both the genetics of color vision (reviewed by Carleton 2009) and female mate choice based on visual cues (Couldridge & Alexander 2002; Jordan *et al*. 2003; Kidd *et al.* 2006) has been performed to date in the cichlids of Lake Malawi. This study is the first to assign a predictive behavioral value to a characteristic of the cone opsin expression palette. Moreover, we have demonstrated that the LWS cone opsin channels into a luminance-based visual pathway, and this finding has implications for the structure of male nuptial coloration. If red coloration is key to the male courtship signal due to the LWS component of the magnocellular system, we can hypothesize that this specific signal is relatively unconstrained by surrounding contrast patterns. Similarly, if the absolute sensitivity of the motion-sensitive magnocellular system is important in mate identification and selection, we would expect male nuptial patterns to highlight red/yellow coloration on the most mobile portions of the body, *i.e.* the fins. While these ideas are obviously preliminary, they represent testable hypotheses that may be addressed by broad phylogenetic association studies investigating LWS opsin expression and specific facets of male nuptial colorations. As such, the proposed predominance of the LWS pigment in the cichlid magnocellular pathway has profound implications for both the study of cichlid evolution and for the broader study of teleost visual pathway organization.

Figures

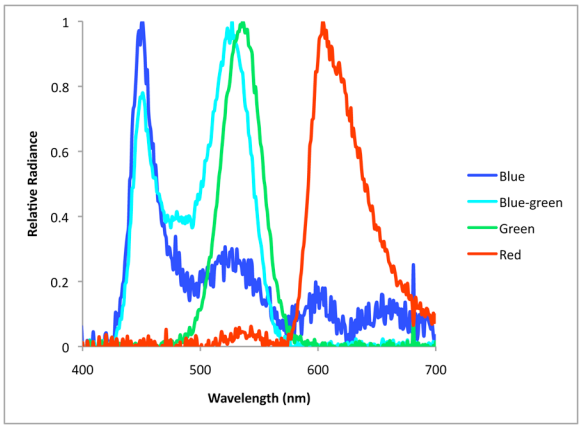



**Figure 4-1:** Normalized radiance spectra of the different experimental color treatments. Based on measures in photons/cm$^2$s.

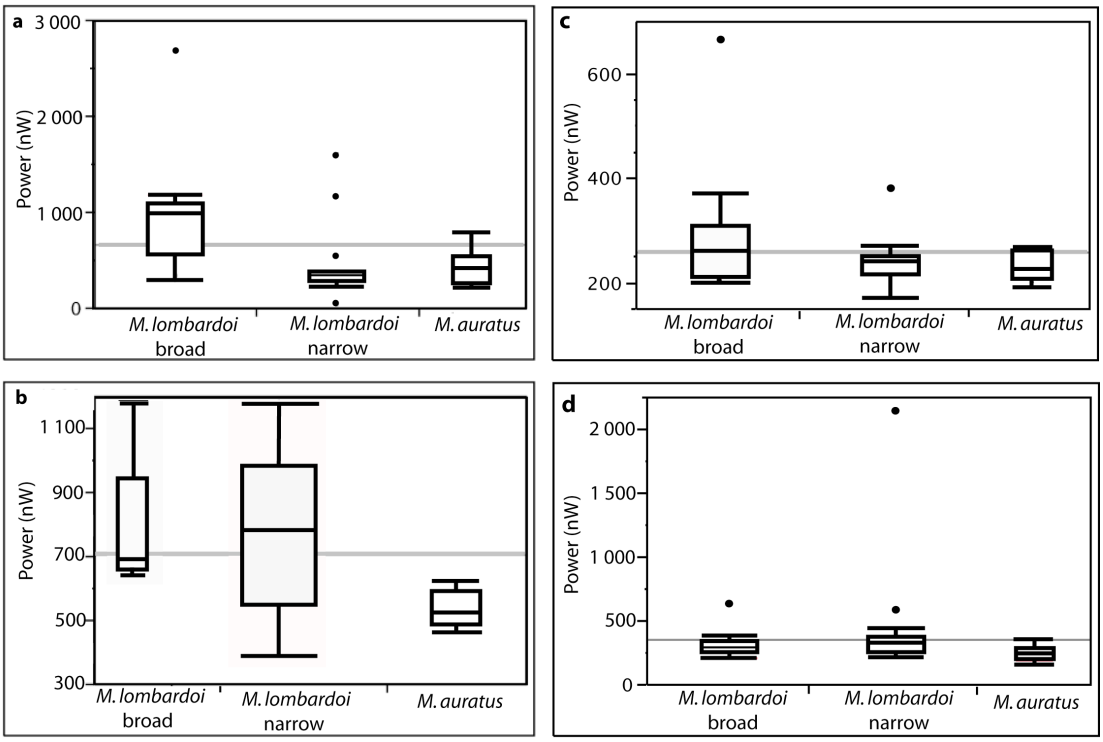



**Figure 4-2:** OMR sensitivities of different treatment groups for different chromatic stimuli: (a) red, (b) blue, (c) green, and (d) blue-green. The gray bar represents the grand mean of the ANOVA model.

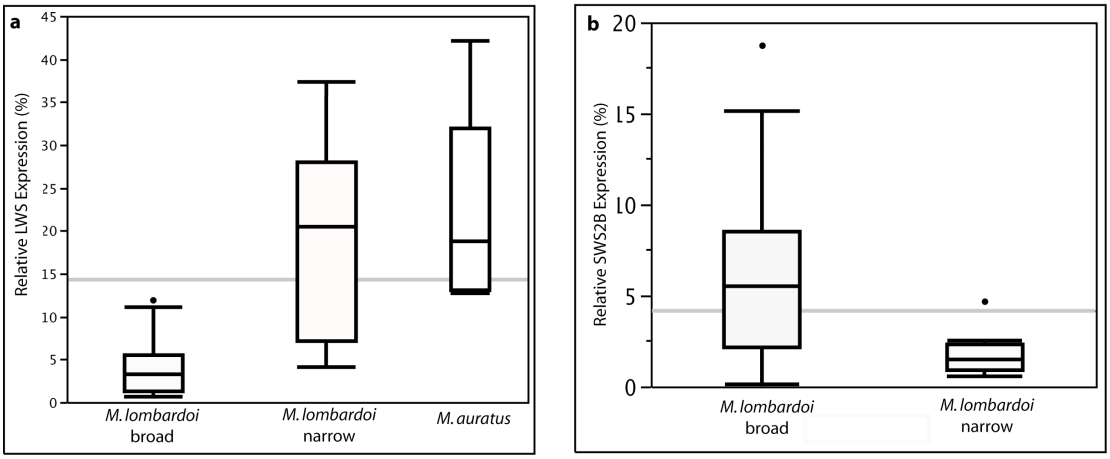



**Figure 4-3:** Relative gene expression differences across treatment groups for (a) the LWS gene and (b) the SWS2B gene. The gray bar represents the grand mean of the ANOVA model.

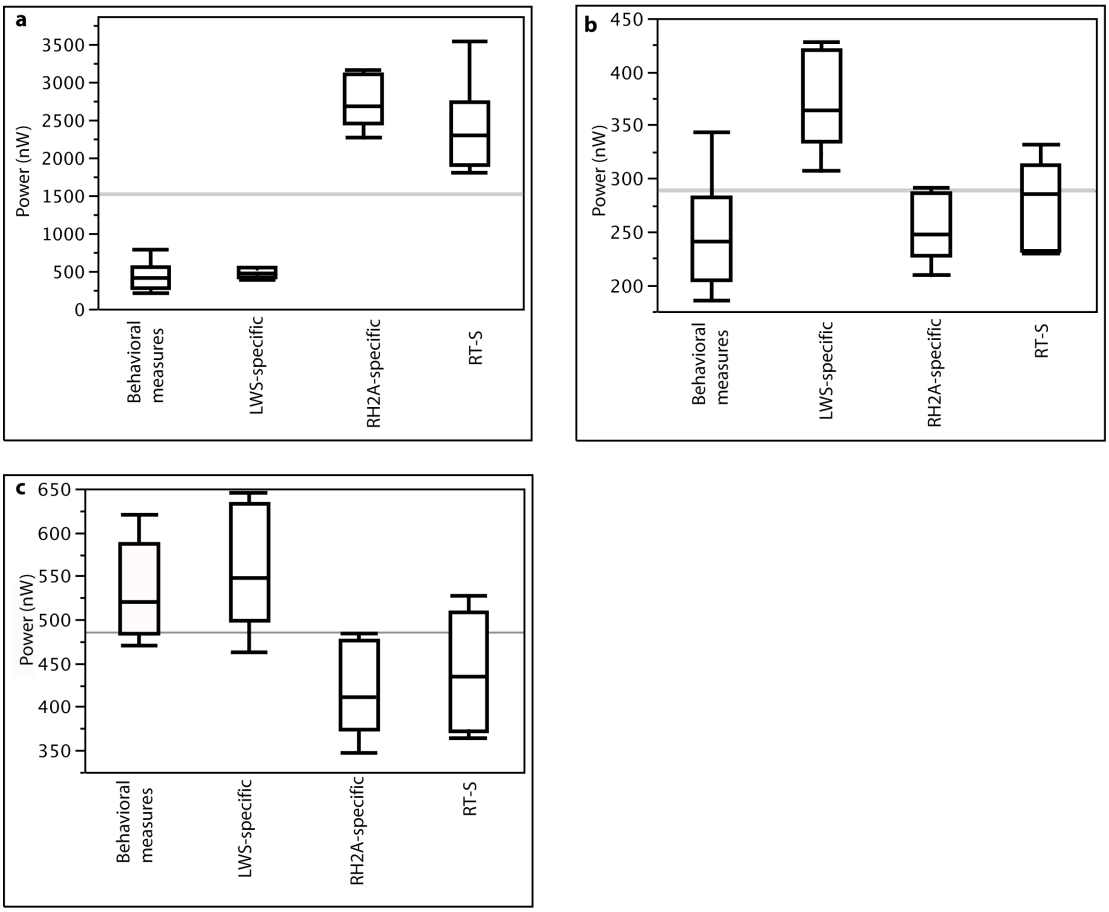



**Figure 4-4:** Behavioral predictive model performance for the *Melanochromis auratus* treatment for the (a) red, (b) blue-green, and (c) blue light environments. The gray bar represents the grand mean of the ANOVA model.

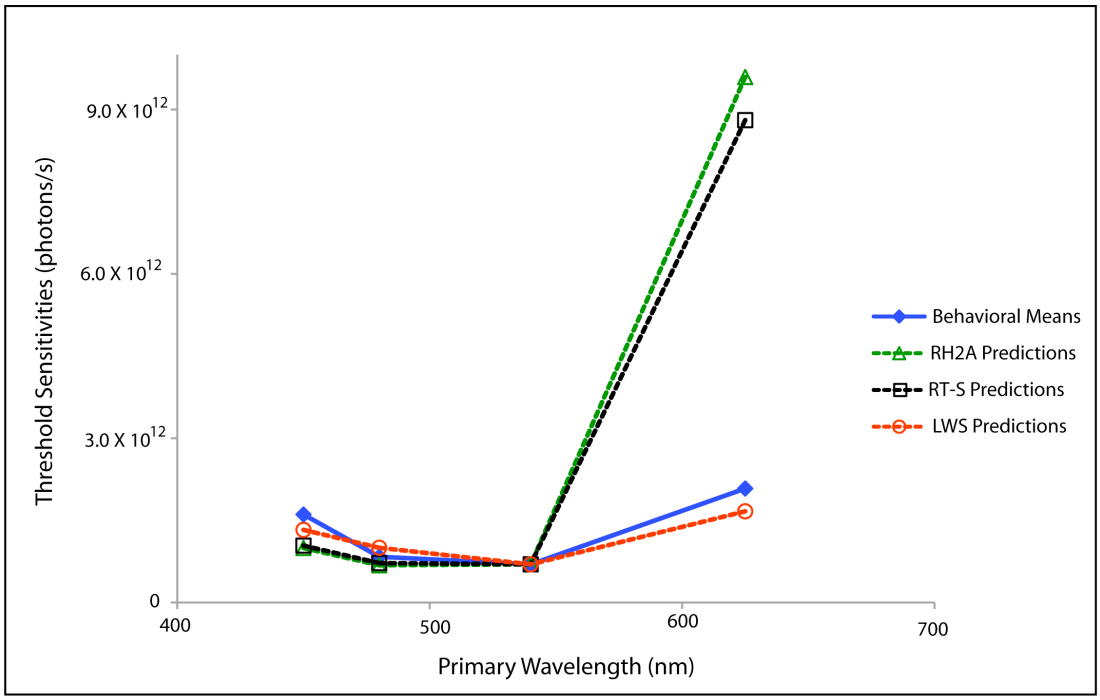



**Figure 4-5:** Sensitivity curves for the mean response of all fishes and the corresponding predicted sensitivities for each model type. Sensitivities in photons/s were estimated by treating each stimulus as monochromatic at its peak wavelength.

# Chapter 5:

# Building New Models: Conclusions and the Implications for the Study of Evolution

*Looking back: significant findings of previous chapters*

This dissertation was designed to address three fundamental hypotheses of cichlid visual biology: (i) changes in receptor gene sequence will be observed across species with varying ecologies and nuptial colorations, such that long- or short-wavelength shifts in spectral absorbance compared to a reference species will have long- or short-wavelength shifted nuptial signals, (ii) cone opsin expression will vary with changes in ambient light environment (depth), with fishes from deeper waters with a more constrained ambient light spectrum expressing less LWS opsin and with lower quantitative variance in gene expression than conspecifics from shallow waters, and (iii) shifts in gene expression will result in quantitative changes in behavioral thresholds, such that fishes expressing relatively more LWS will be more sensitive to long-wavelength stimuli while fishes expressing more RH2A and B will be more sensitive to shorter-wavelength stimuli. This suite of hypotheses was structured to investigate the molecular and behavioral facets of the sensory drive sexual selection model. While our work demonstrated significant differences at each of the three hypothesized levels, the nature of the findings suggests that selection models other than the sensory drive hypothesis are likely at play in Lake Malawi.

*Chapter 2: Opsin gene sequence variation*

Opsin sequence variation was characterized in two Malawi cichlid genera with different ecologies. Potentially functional gene sequence variation was found in four of the seven cone opsin genes in a panel of fourteen species from two genera. The SWS1 (ultraviolet) and RH2Aβ (green) were variable in the genus *Metriaclima*, while the SWS2B (violet) and LWS (red) genes were variable in the genus *Labidochromis*. Along with these functional variants, a potential pseudogenization was found in one gene for each genus (SWS2A in *Metriaclima* and RH2Aα in *Labidochromis*).

The divergent alleles found for each gene likely differ in function. The two SWS1 allelic variations, termed UV368 and UV378, are known to cause a 10nm shift in pigment absorbance (Parry *et al*. 2005). These alleles are separated by four fixed amino-acid sites, with no intermediate versions known in Lake Malawi. More specifically, it appears that the UV368 allele arose in Lake Malawi and has undergone rapid diversification. The single-site variations in both the SWS2B and LWS gene are known to cause 10nm shifts in cottoid fishes (Cowing *et al.* 2002) and humans respectively (Asenjo *et al*. 1994), but were previously undescribed in cichlids. Finally, the origin of the RH2Aβ variation is uncertain, with the derived variant seemingly resulting from an insertion event. This data supports the idea that opsin sequence tuning can contribute to fine scale differences (*i.e.* < 15nm in peak shift) in visual sensitivity, particularly in genes at the shortest and longest regions of the light spectrum.

*Chapter 3: Wild intraspecific variation in opsin expression*

Opsin expression variation was compared between two populations in three species. Qualitative variation in gene expression was found in *Copadichromis eucinostomus* collected at different depths from a single site, with shallow individuals (collected at 5m) expressing LWS while deep individual (collected at 15m) did not. Similarly, *Metriaclima zebra* collected at two geographically distinct locations exhibited qualitative variation in the expression of the SWS2B gene. Fish from Zimbawe rock expressed SWS2B while fish from Thumbi West Island did not. Finally, while *Tropheops gracilior* collected at two different depths (5m and 20m) displayed no qualitative variation in their opsin expression palette, the deeper fish had significantly more quantitative variation in gene expression than their shallow counterparts. This demonstrates that opsin gene expression can vary at a local scale and could contribute to visual sensitivity differences.

In order to test the effects of these gene expression differences, we developed a new model for analyzing visual sensitivities based on RT-qPCR expression data: the real-time corrected sensitivities model (henceforth referred to as RT-S). The RT-S uses the relative gene expression of each pigment as a scalar coefficient to determine the absolute magnitude of that pigment's predicted absorbance curve (Govardovskii *et al*. 2000). The scaled curves for all the pigments are then summed, and the integral of this complex curve can be used to calculate a predicted quantum catch for the visual system in a given light environment. Using these RT-S calculations, we determined that the broad down-welling ambient light spectrum in Lake Malawi likely does not impose significant constraint on gene expression palettes, particularly in the case of the qualitative LWS shift in *C. eucinostomus*. However, when we

calculated sensitivities using a red-shifted ambient light environment from Lake Victoria, the model predicted a significant sensitivity advantage for the individuals that expressed LWS. This finding further supported our conclusion that the ambient light environment in Lake Malawi is less likely to impose selection on differential gene expression patterns than other mechanisms such as drift or sexual selection.

*Chapter 4: The effects of developmental environment on LWS expression and behavior*

Behavioral and gene expression measures were taken from two species, with young fish from one of the species (*Metriaclima lombardoi*) being reared under two different developmental lighting treatments. Visual sensitivities for each subject were measured using the optomotor response (OMR) behavioral paradigm, with behavioral thresholds being measured for each of four different chromatic lighting stimuli. Relative opsin gene expression was then measured for each individual so that a specific correlation between expression and behavior could be tested. We found that developmental light environment had strong effects on (i) expression of the LWS gene, and (ii) OMR threshold measures for a long-wavelength dominant chromatic stimulus. The degree of developmental plasticity in *Metriaclima lombardoi* was such that individuals raised under narrow-spectrum (fluorescent) lighting expressed significantly more LWS than conspecifics raised under broad-spectrum lighting. Furthermore, the narrow-spectrum *M. lombardoi* had an LWS expression pattern similar to that observed in the other experimental species (*Melanochromis auratus*). This expression shift mirrored a behavioral shift, with narrow-spectrum *M. lombardoi* having behavioral thresholds similar to *M. auratus* under red light but dissimilar to

their conspecifics. Using these combined measures, we calculated that LWS gene expression explained roughly 20% of the behavioral variation in our data.

In order to test the practical applications of the RT-S model proposed in chapter three we constructed various models using expression measures and opsin absorbance templates to predict behavioral thresholds using a single behavioral measure (the OMR threshold under a green chromatic environment). In particular, we tested two types of models: (i) the RT-S model, which represents a random-wiring model scaled by expression measures, and (ii) cone-predominance models that include single pigments and don't use expression scaling factors. In general, the RT-S model performed poorly. The LWS single-pigment model, however, predicted behavior accurately for the long-wavelength stimulus, and even for the short-wavelength stimulus in the case of *Melanochromis auratus*. Therefore, we concluded that the LWS-predominance model was the most parsimonious pigment mechanism to define the optomotor response in the species we studied.

*Tying it all together: variation at multiple sensory levels*

The previous chapters highlighted three primary facets of cichlid visual biology: (i) opsin sequence diversity occurs between genera with different ecological traits and within genera with diverse male nuptial colorations, (ii) gene expression differentiation is found within species in the wild, and (iii) expression shifts such as those observed in the wild have measurable behavioral consequences. The degree of genetic and behavioral variation observed in Malawi cichlids is astounding, particularly when considering the small phylogenetic and temporal scales which our target species represent. Furthermore, this work combined with prior work in our lab

has demonstrated that the potential for environmental plasticity in expression is also variable across species (Hofmann *et al*. 2010). So visual diversification may take place along two axes: (i) the characteristics of the visual system, and (ii) differential developmental lability of the visual system within or between species.

Perhaps most importantly, this work has implicated the LWS pigment as a viable candidate for further work in cichlid speciation studies. The LWS gene was found to be polymorphic in the species *Labidochromis caeruleus,* showed plastic intraspecific expression variation in *Copadichromis eucinostomus*, and is a key determiner of performance on the optomotor behavioral response. Variation in relative LWS expression accounted for roughly 20% of the total observed behavioral variation. This effect is immense, particularly considering all of the neural processing steps that occur between the receptor cells and final behavioral output.

*The spaces in-between: what does this data say about the brain?*

One of the primary advantages of using the OMR to measure behavioral sensitivities relates to the cellular underpinnings that mediate the response. The sensitivity to motion is thought to be driven by cells in the retinal periphery that are tied to the magnocellular pathway. The general characteristics of the magnocellular pathway are well understood, as are some of the signal-recognition tasks that it performs (Lee & Sun 2004). Our work suggests that the LWS predominant wiring mechanism that is present in cichlids is similar to that previously posited for goldfish (Schaerer & Neumeyer 1996). However, exclusion of MWS cone input into this pathway disagrees with results from other systems, such as zebrafish (Orger & Baier 2005) and macaques (Sun *et al.* 2006). This suggests two things: (i) the magnocellular

MWS exclusion mechanism may be unique to certain teleosts, and (ii) this derived trait may exert specific and previously overlooked selective pressures in teleost model systems.

The magnocellular pathway is particularly sensitive to two types of visual stimuli: (i) chromatic motion of stimuli in the luminance channel of the long wavelength cone system, and (ii) pronounced contrast boundaries.
In addition, this neural circuit appears to contain a red/green opponency mechanism which is likely mediated by the H1 horizontal cell types (Lee & Sun 2004). If the MWS cone is eliminated from the pathway, this opponency system would likely be lost, resulting in a purely luminance-driven information stream that is truly "color-blind". Furthermore, the spectral range of the magnocellular response should long-wavelength shift and change the absolute response profile of the neural circuit.

*The implications for sexual selection*

We have found significant variation in visual sensitivity at multiple levels in the visual system. This could have important implications for how cichlids recognize conspecifics, how they choose mates and potentially whether cichlid species diverge. To illustrate this idea, we use the example of the wild-caught *Copadichromis eucinostomus* which were described in chapter three. The variation in LWS expression of *C. eucinostomus* was quite large, with shallow individuals expressing significant amounts of LWS and deeper living individuals generally expressing very little or none. This difference was of even greater magnitude than the environmentally induced plasticity in LWS expression found for *Metriaclima lombardoi* reared under different laboratory light environments as part of chapter four. Given the effect we

demonstrated between LWS expression and the optomotor response, the shallow *C. eucinostromus* individuals should exhibit a stronger response to stimuli that target the magnocellular pathway than their conspecifics from deeper waters. Given the traits of this circuit, we would hypothesize that male traits exhibiting yellow/red coloration or hard black/white contrast edges would stimulate this pathway more strongly. Moreover, since the magnocellular pathway is highly motion-sensitive, we posit that these signals should be elaborated on the fins of the fish, which tend to be highly-mobile during the courtship dance. Therefore, we can put forth the hypothesis that females from shallow waters should prefer males with red coloration or pronounced contrast bars on their fins, while deeper conspecifics might use separate cues to assess males.

*How does this affect current evolutionary models?*

To the best of our knowledge, this is the first complete work that posits an explicit, quantitative link between sensory receptor gene expression and corresponding behavioral sensitivities that could result in differential mate choice. Unlike the standing sensory drive hypothesis, our model requires no evolutionary constraint on sensory capabilities by environmental transmission characteristics. Rather, our model can accommodate greater sensory diversification due to a lack of constraint via stabilizing selection. Although the environment may influence the expression phenotype in plastic species (as we have demonstrated for species raised under different lighting environments in the lab), this may not be the only mechanism that induces such plasticity. For example, random drift of genetic factors controlling opsin expression may also contribute differential visual palettes. These visual palettes

can then be linked evolutionarily to specific male display traits, resulting in a sexual selection mechanism similar to the Fisherian Runaway model (Fisher 1915). The incorporation of different cone pigments into the sexual-selection mechanism would generate diverse hypotheses regarding male signal evolution. For example, if the SWS pigments became the dominant mate-choice mechanism, male traits should evolve to exploit the parvocellular neural pathway (which contains most of the information from SWS cone types). This neural circuit primarily processes color opponent mechanisms (particularly between single and double cones) and contrast patterning in high-acuity regions of the visual field, and therefore male nuptial displays would evolve on a different selective axis than in species that primarily use the magnocellular system. In the meantime, we would like to suggest that divergent mate choice and speciation through sexual selection in Lake Malawi (and possibly alternative systems as well) functions via "magnocellular drive". While the specific neural predictions of this model still need to be tested, we believe it is a promising avenue of speciation research in this and other model systems.

*What's next? Suggestions for future research*

There are numerous techniques for separately investigating mate choice, chromatic perception, and the function of the magnocellular pathway. However, performing research at the interface of these topics will require careful attention to experimental detail. Experimental paradigms from the molecular, physiological, and behavioral disciplines generally include inherent biases or assumptions that can be easily overlooked. Some previous work using the OMR have failed to account for the known magnocellular underpinnings of the behavior and have offered up what are

likely erroneous conclusions regarding cone inputs to the behavior. For example, Kroger *et al*. (2003) attributed their behavioral findings using the OMR to variation in SWS cones. Since the OMR is known to function through neural pathways that exclude SWS cone inputs, it is likely that their data are best explained using a mechanism that is mutually exclusive to the one they proposed. Similarly, Sabbah *et al*. (2010) collected both molecular and electrophysiological data, but did not utilize either measure as an information criterion to posit a single parsimonious unifying hypothesis. Sabbah *et al*. (2010) interpreted their electroretinogram (ERG) data using model fits that included very few mechanistic restrictions. Essentially they allowed their fits to include contributions from all known cone visual pigments. However, we known that opsin genes are differentially expressed with some genes not expressed in some species. Combined with the broad spectral bandwidth that was required for them to collect sufficient electrophysiological data, the likelihood that their optimal models would fit their expression data seems quite low. Rather, they should have used their data on relative opsin expression as a restrictive factor on their electrophysiology model to determine the difference in the fit of the constrained and unconstrained models. This would have provided a realistic assessment of the relationship between gene expression and the ERG response. Rather Sabbah *et al*. (2010) dismissed the two datasets as unrelated without rigorous statistical testing which would have addressed the relationship in an appropriate fashion. In the following section I will propose some basic experiments that could be used to further investigate the conclusions drawn in this dissertation, while avoiding the downfalls of some studies currently present in the literature.

*Is the magnocellular pathway really monochromatic? Investigating LWS-RH2 opponency through physiology and behavior*

One of the characteristic functions of the mammalian magnocellular pathway is a weak red/green color opponency mechanism (Lee & Sun 2004). The techniques used to measure this function tend to be fairly complex, and involve using *in vivo* electrophysiology of a gaze-fixated macaque attending to a color-modulated stimulus. Generally, such techniques are intractable in fish, although optokinetic eye-tracking techniques have been used for zebrafish and medaka (Mueller & Neuhass 2010). However, a variation of the OMR stimulus that employs a red/green isoluminance grid may be applicable here. By changing the standard black/white sinusoid stimulus to a red/green sinusoid, theoretical luminance gratings can be constructed using pigment model predictions. For example, if the LWS pigment wires into the pathway without any MWS output, a grid with red and green components that would stimulate LWS cones equally would cause no optomotor response (*i.e.* the subject would see the grid as a uniform background. By measuring the way different red/green intensity ratios affects behavior, it may be possible to determine not only what cone types channel into the magnocellular pathway, but also if red/green opponency is present in the system. Also, pharmaceutical knockouts could be used to ablate an opponency response via decreased activity of the H1 horizontal cells. If the disruption of horizontal cell activity causes no behavioral change, then the premise of a monochromatic magnocellular pathway would be supported.

*SWS expression variation and behavior: choice tests for discrimination*

One of the plastic gene expression effects documented by Hofmann *et al*. (2010) and confirmed in chapter four is a "switch" from an SWS1-only system to an SWS1+SWS2B system in the laboratory. However, a similar gene expression shift was observed in *Metriaclima zebra* from two separate populations in the wild (chapter 3). Similar to our *a priori* hypothesis of a random-wiring magnocellular model for the OMR behavior, we hypothesize that a dichromatic SWS system would function as an expression-weighted random wiring model. The expression of SWS1 is thought to be linked to planktivory (Hofmann *et al*. 2009), as planktonic organisms should appear as dark spots against a bright-field background to a UV-sensitive visual system. For this reason, a two-choice discrimination task could be used to test how the SWS1/SWS2B shift affects the perception of a bright-field stimulus by a fish. Test subjects could be trained to retrieve a food reward at the target with greater overall intensity. Trained fishes could then be offered two test fields with the same total power, with one being composed solely of UV-emissive LED lights and the other field being a mixture of UV- and violet-emitting LEDs. If the cone pigments are summed into a single channel as proposed, the subject should choose the target with light sources matching its gene expression palette. In essence, this would test the single-cone component of the RT-S model. However, there are two primary caveats for this experiment. First, training experiments are time intensive and typically require small sample sizes to be completed in a timely fashion. These small sample sizes are not conducive to model tests if there is a large degree of variance in the behavioral measures. While the two-choice experiment would help to constrain this variation, it would not alleviate statistically problematic variance *per se*. Furthermore,

to date we have not been able to reproducibly vary the SWS1/SWS2B shift in a species in the laboratory as we have with the LWS gene. Therefore, wild-caught fishes or heterospecific comparisons would likely be needed for this study.

*Variable LWS expression and mate choice: can we predict relevant nuptial signals?*

As previously mentioned, if a red-dominant magnocellular pathway is an important component of mate choice, we can hypothesize that females expressing more LWS should display a greater preference for males with yellow or red coloration of the fins. While a basic correlation analysis could be used to look at broad male coloration patterns vs. LWS expression for different species, direct behavioral tests must be more carefully structured. As of yet, the only species that we know has clearly tangible and reproducible LWS variation in the lab is *Metriaclima lombardoi*. Since mature, dominant males of this species are uniformly yellow, intraspecific mate choice tests will not necessarily be informative. However, given the propensity of *M. lombardoi* to hybridize with congeners, it may be possible to use other *Metriaclima* species pairs for this purpose. Specifically, we can utilize congeneric choice tests in the mate choice apparatus of Kidd *et al*. (2006). By placing female *M. lombardoi* with a species pair consisting of *M. zebra* and a similarly-patterned "red top" congener (*M. emmiltos* that has a red dorsal fin for example), it may be possible to assess the preference of *M. lombardoi* females for the red-finned phenotype. The number of eggs females lay with either blue or red-finned fishes can then be plotted against relative LWS expression. Much like the proposed training experiment, to perform this study will likely require a large female sample size. Given the amount of time it may take individual females to lay eggs with a

heterospecific male, total time for the completion of the experiment may be prohibitive.

*In conclusion: the intellectual value of this work*

The central theme of this dissertation can be reduced to one primary goal: the development of interdisciplinary predictive models explicitly relating genes and behavior. To date, despite the wealth of research involving the genetics of vision and the role of visual cues in mate choice, quantitative models describing the association between these two intellectual areas have been exceedingly rare. Since sexual selection and evolution as a whole can be thought of as an exercise in probabilities, this quantitative gap has been a point of contention when interpreting results on both the molecular and behavioral ends of the experimental spectrum. However, our studies demonstrate that well-described, straightforward techniques regularly used by both geneticists and ethologists can be applied to this problem. By combining these techniques in the appropriate fashion for a given study system, it is possible to perform insightful, nuanced interdisciplinary research. In essence, we hope this body of work can be viewed as a novel experimental scaffold on which to formulate new interdisciplinary hypotheses for the study of adaptation, speciation, and evolution.

# Appendix I: Supplemental Material for Chapter 2

**Table S2-1**: Collection data for *Labidochromis* samples used in this study.

| Species | Finclip Ref # | Retina Ref # | Location | Date Collected | Sex |
|---|---|---|---|---|---|
| *L. flavigulus* | N/A | 1400 | Thumbi West "Aquarium" | 7/23/08 | F |
| *L. gigas* | 2162 | 1322 | Thumbi West | 7/18/08 | M |
| | 2163 | 1323 | Thumbi West | 7/18/08 | M |
| *L. ianthinus* | 2286 | 2286 | Mbenji Island | 7/23/08 | M |
| | 2287 | 2287 | Mbenji Island | 7/23/08 | M |
| *L. vellicans* | 2164 | 1324 | Thumbi West | 7/18/08 | M |
| | 2165 | 1325 | Thumbi West | 7/18/08 | M |
| | 2198 | 1344 | Thumbi West | 7/19/08 | M |

**Table S2-2**: Novel opsin gene primers developed for this study.

| Gene | Primer Direction | Primer Sequence |
|---|---|---|
| SWS1 | Forward | TCCCAGCAGCATAATCTTTG |
| SWS1 | Reverse | AGCTCAGTCACGCCCTYTTA |
| SWS2B | Forward | AGAGGAAATCGTCCCACTGA |
| SWS2B | Forward | AGGAAGCCCGATGATCTTTT |
| SWS2B | Forward | AGAGGAAATCGTCCCACTGA |
| SWS2B | Reverse | GACAGCAAAGCAGAAGCAGA |
| SWS2B | Reverse | CAGACCCAAGTCAATCCACA |
| SWS2A | Forward | AGCGTTCATGTCTGCTCCTT |
| SWS2A | Forward | CTGGTTCCACAGGACCACTT |
| SWS2A | Forward | GCTGTTCCTTTGACCACGAT |
| SWS2A | Forward | ACATTAACAGGCAGCCAAGG |
| SWS2A | Reverse | CTGGTCACTTCCCTCTCTGC |
| SWS2A | Reverse | AGCACTGTAGGCCTTCTGGA |
| SWS2A | Reverse | CAGAAATCAGCGAGCATTGA |
| SWS2A | Reverse | GGAGAAGACAACCCATTTGC |
| RH2B | Forward | CTCGTGAGGAGTCCTTTTGAA |
| RH2B | Forward | TGCTGTTGAGAGATACATTGTGG |
| RH2B | Forward | ACACTCTGGCCCCAGGATAC |
| RH2B | Reverse | ACCTTGACCAGCCAAACAAT |
| RH2B | Reverse | GACACAGAAGTGGCAGGTGA |
| RH2B | Reverse | TGAACAATGCTGAGCTCTTTG |
| RH2A | Forward | TGGCCACACTTGGAGGTAAT |
| RH2A | Reverse | TCAGGACCACAGGAACACTG |
| LWS | Forward | GTCTTCACCAACGGTCTCGT |
| LWS | Reverse | CACAGCAAGGTAGCACAGGA |

**Table S2-3:** Accession numbers for the experimental and reference genome sequences used in this study.

| Genus | Species | SWS1 | SWS2b | SWS2a | RH2B-1 | RH2B-2 | RH2Aα | RH2Aβ | LWS |
|---|---|---|---|---|---|---|---|---|---|
| Labidochromis | caeruleus | HM04194 | HM049232 | HM049270 | HM049307 | HM049345 | HM049416 | HM049419 | HM049455 |
| Labidochromis | caeruleus | HM04195 | HM049233 | HM049271 | HM049308 | HM049346 | HM049417 | HM049420 | HM049456 |
| Labidochromis | caeruleus | HM04196 | HM049234 | HM049272 | HM049309 | HM049347 | HM049418 | HM049421 | HM049457 |
| Labidochromis | chisumulae | HM04197 | HM049235 | HM049273 | HM049310 | HM049348 | HM049383 | HM049422 | HM049458 |
| Labidochromis | chisumulae | HM04198 | HM049236 | HM049274 | HM049311 | HM049349 | HM049384 | HM049423 | HM049459 |
| Labidochromis | chisumulae | HM04199 | HM049237 | HM049275 | HM049312 | HM049350 | HM049385 | HM049424 | HM049460 |
| Labidochromis | flavigulus | HM04200 | HM049238 | HM049276 | HM049313 | HM049351 | HM049386 | NA | HM049461 |
| Labidochromis | gigas | HM04201 | HM049239 | HM049277 | HM049314 | HM049352 | HM049387 | HM049425 | HM049490 |
| Labidochromis | gigas | HM04202 | HM049240 | HM049278 | HM049315 | HM049353 | HM049407 | HM049426 | HM049462 |
| Labidochromis | ianthinus | HM04203 | HM049241 | HM049279 | HM049316 | HM049354 | HM049388 | NA | HM049463 |
| Labidochromis | ianthinus | HM04204 | HM049242 | HM049280 | HM049317 | HM049355 | HM049389 | HM049427 | HM060335 |
| Labidochromis | vellicans | HM04205 | HM049243 | HM049281 | HM049318 | HM049356 | NA | HM049428 | HM049464 |
| Labidochromis | vellicans | HM04206 | HM049244 | HM049282 | HM049319 | HM049357 | HM049390 | HM049429 | NA |
| Labidochromis | vellicans | HM04207 | HM049245 | HM049283 | HM049320 | HM049358 | HM049391 | HM049430 | HM049465 |
| Metriaclima | barlowi | HM04208 | HM049246 | HM049284 | HM049321 | HM049359 | HM049392 | HM049431 | HM049466 |
| Metriaclima | barlowi | HM04209 | HM049247 | HM049285 | HM049322 | HM049360 | HM049393 | HM049432 | HM049467 |
| Metriaclima | barlowi | HM04210 | HM049248 | HM049286 | HM049323 | HM049361 | HM049394 | HM049433 | HM049468 |
| Metriaclima | benetos | HM04211 | HM049249 | HM049287 | HM049324 | HM049362 | NA | HM049434 | HM049469 |
| Metriaclima | benetos | HM04212 | HM049250 | HM049288 | HM049325 | HM049363 | HM049395 | HM049435 | HM049470 |
| Metriaclima | benetos | HM04213 | HM049251 | NA | HM049326 | HM049364 | HM049396 | HM049436 | HM049471 |
| Metriaclima | callainos | HM04214 | HM049252 | HM049289 | HM049327 | HM049365 | HM049408 | HM049437 | HM049472 |
| Metriaclima | callainos | HM04215 | HM049253 | HM049290 | HM049328 | HM049366 | HM049409 | HM049438 | HM049473 |
| Metriaclima | callainos | HM04216 | HM049254 | HM049291 | HM049329 | HM049367 | HM049410 | HM049439 | HM049474 |
| Metriaclima | fainzilberi | HM04217 | HM049255 | HM049292 | HM049330 | HM049368 | HM049411 | HM049440 | HM049475 |
| Metriaclima | fainzilberi | HM04218 | HM049256 | HM049293 | HM049331 | HM049369 | HM049397 | HM049441 | HM049476 |
| Metriaclima | fainzilberi | HM04219 | HM049257 | HM049294 | HM049332 | HM049370 | HM049412 | HM049442 | HM049477 |
| Metriaclima | lombardoi | HM04220 | HM049258 | HM049295 | HM049333 | HM049371 | HM049398 | HM049443 | HM049478 |
| Metriaclima | lombardoi | HM04221 | HM049259 | HM049296 | HM049334 | HM049372 | HM049399 | HM049444 | HM049479 |
| Metriaclima | lombardoi | HM04222 | HM049260 | HM049297 | HM049335 | HM049373 | HM049400 | HM049445 | HM049480 |
| Metriaclima | mbenji | HM04223 | HM049261 | HM049298 | HM049336 | HM049374 | HM049401 | HM049446 | HM049481 |
| Metriaclima | mbenji | HM04224 | HM049262 | HM049299 | HM049337 | HM049375 | HM049402 | HM049447 | HM049482 |
| Metriaclima | mbenji | HM04225 | HM049263 | HM049300 | HM049338 | HM049376 | HM049403 | HM049448 | HM049483 |
| Metriaclima | phaeos | HM04226 | HM049264 | HM049301 | HM049339 | HM049377 | HM049404 | HM049449 | HM049484 |
| Metriaclima | phaeos | HM04227 | HM049265 | HM049302 | HM049340 | HM049378 | HM049405 | HM049450 | HM049485 |
| Metriaclima | phaeos | HM04228 | HM049266 | HM049303 | HM049341 | HM049379 | HM049406 | HM049451 | HM049486 |
| Metriaclima | pyrsonotus | HM04229 | HM049267 | HM049304 | HM049342 | HM049380 | HM049413 | HM049452 | HM049487 |
| Metriaclima | pyrsonotus | HM04230 | HM049268 | HM049305 | HM049343 | HM049381 | HM049414 | HM049453 | HM049488 |
| Metriaclima | pyrsonotus | HM04231 | HM049269 | HM049306 | HM049344 | HM049382 | HM049415 | HM049454 | HM049489 |

| Genus | Species | Source | SWS1 | RH2Aβ | LWS |
|---|---|---|---|---|---|
| Metriaclima | zebra | Malawi | AF191222 | AF247122 | AF247126 |
| Dimidiochromis | compressiceps | Malawi | AF191220 | AF247121 | AF247125 |
| Labeotropheus | fuelleborni | Malawi | AF191223 | AF247123 | AF247127 |
| Aulonocara | heuseri | Malawi | AY775100 | AY775090 | AY780517 |
| Melanochromis | auratus | Malawi | AY775101 | AY775091 | AY780518 |
| Tyrannochromis | maculatus | Malawi | AY775103 | AY775093 | AY780520 |
| Cynotilapia | afra | Malawi | AY775104 | AY775094 | AY780521 |
| Stigmatochromis | modestus | Malawi | AY775107 | AY775070 | AY780523 |
| Pseudotropheus | acei | Malawi | DQ088642 | DQ088633 | DQ088627 |
| Tramitichromis | intermedius | Malawi | DQ088644 | DQ088635 | DQ088629 |
| Pundamilia | nyererei | Victoria | AY673728 | AY673698 | AY673688 |
| Pundamilia | pundamilia | Victoria | NA | NA | AY673689 |
| Lipochromis | melanopterus | Victoria | AY673733 | AY673703 | AY673693 |
| Lithochromis | rubripinnis | Victoria | NA | NA | AY673781 |
| Neochromis | greenwoodi | Victoria | AY673734 | AY673704 | AY673694 |
| Paralabidochromis | chilotes | Victoria | AY673736 | AY673706 | AY673696 |
| Opthalmotilapia | ventralis | Tanganyika | AY775097 | AY775067 | AY780512 |
| Neolamprologous | brichardi | Tanganyika | AY775096 | AY775068 | AY780513 |
| Tropheus | duboisi | Tanganyika | AY775099 | AY775089 | AY780516 |
| Oreochromis | niloticus | River | AF191221 | AF247124 | AF247128 |

# Appendix II: Supplemental Material for Chapter 3

**Table S3-1:** Raw relative expression data for all of the fishes used in this study.

| Species | Source | Depth | SWS1 | SWS2B | SWS2A | RH2B | RH2Aa | LWS |
|---|---|---|---|---|---|---|---|---|
| Copadichromis eucinostomus | Otter Point | 20m | 16.56 | 0.12 | 0.03 | 18.80 | 64.09 | 0.39 |
| Copadichromis eucinostomus | Otter Point | 20m | 16.74 | 0.12 | 0.00 | 21.64 | 61.46 | 0.03 |
| Copadichromis eucinostomus | Otter Point | 20m | 20.68 | 0.57 | 0.02 | 16.82 | 61.23 | 0.67 |
| Copadichromis eucinostomus | Otter Point | 20m | 1.59 | 25.23 | 0.07 | 17.10 | 55.74 | 0.26 |
| Copadichromis eucinostomus | Otter Point | 20m | 11.11 | 1.22 | 0.01 | 24.40 | 63.17 | 0.09 |
| Copadichromis eucinostomus | Otter Point | 20m | 20.98 | 0.03 | 0.04 | 20.98 | 56.57 | 1.39 |
| Copadichromis eucinostomus | Otter Point | 20m | 25.28 | 0.00 | 0.05 | 18.82 | 55.84 | 0.00 |
| Copadichromis eucinostomus | Otter Point | 20m | 19.03 | 0.00 | 0.04 | 24.62 | 56.32 | 0.00 |
| Copadichromis eucinostomus | Otter Point | 20m | 16.32 | 0.50 | 0.05 | 26.65 | 56.48 | 0.00 |
| Copadichromis eucinostomus | Otter Point | 20m | 21.49 | 1.42 | 0.05 | 22.01 | 55.03 | 0.00 |
| Copadichromis eucinostomus | Otter Point | 20m | 17.93 | 3.81 | 0.20 | 22.69 | 53.60 | 1.78 |
| Copadichromis eucinostomus | Otter Point | 20m | 11.71 | 0.55 | 0.82 | 2.73 | 16.86 | 67.32 |
| Copadichromis eucinostomus | Otter Point | 20m | 18.27 | 0.49 | 0.23 | 15.23 | 44.40 | 21.39 |
| Copadichromis eucinostomus | Otter Point | 5m | 25.17 | 0.10 | 0.04 | 13.84 | 25.19 | 35.66 |
| Copadichromis eucinostomus | Otter Point | 5m | 23.23 | 0.00 | 0.00 | 26.10 | 15.48 | 35.19 |
| Copadichromis eucinostomus | Otter Point | 5m | 16.65 | 0.03 | 0.00 | 16.70 | 49.52 | 17.09 |
| Copadichromis eucinostomus | Otter Point | 5m | 14.08 | 0.08 | 0.00 | 18.80 | 47.89 | 19.16 |
| Copadichromis eucinostomus | Otter Point | 5m | 17.58 | 0.26 | 0.02 | 20.03 | 29.66 | 32.44 |
| Copadichromis eucinostomus | Otter Point | 5m | 14.46 | 1.08 | 0.04 | 18.20 | 34.96 | 31.26 |
| Copadichromis eucinostomus | Otter Point | 5m | 13.51 | 0.74 | 0.02 | 20.83 | 30.57 | 34.33 |
| Copadichromis eucinostomus | Otter Point | 5m | 16.13 | 0.26 | 0.00 | 11.74 | 38.28 | 33.59 |
| Copadichromis eucinostomus | Otter Point | 5m | 14.59 | 0.05 | 0.01 | 17.65 | 31.92 | 35.79 |
| Copadichromis eucinostomus | Otter Point | 5m | 19.11 | 0.55 | 0.04 | 16.17 | 37.24 | 26.89 |
| Copadichromis eucinostomus | Otter Point | 5m | 20.26 | 0.23 | 0.04 | 10.90 | 32.90 | 35.67 |
| Copadichromis eucinostomus | Otter Point | 5m | 21.84 | 0.12 | 0.01 | 13.48 | 44.76 | 19.78 |
| Copadichromis eucinostomus | Otter Point | 5m | 20.92 | 1.00 | 0.05 | 15.49 | 32.65 | 29.88 |
| Copadichromis eucinostomus | Otter Point | 5m | 18.72 | 0.23 | 0.01 | 15.57 | 40.10 | 25.37 |
| Copadichromis eucinostomus | Otter Point | 5m | 19.78 | 0.32 | 0.02 | 15.25 | 41.56 | 23.07 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Metriaclima zebra | Thumbi West | 10m | 23.08 | 0.03 | 0.01 | 13.96 | 62.53 | 0.29 |
| Metriaclima zebra | Thumbi West | 10m | 17.30 | 0.52 | 4.97 | 23.08 | 53.73 | 0.20 |
| Metriaclima zebra | Thumbi West | 10m | 25.25 | 0.14 | 0.04 | 20.05 | 54.31 | 0.00 |
| Metriaclima zebra | Thumbi West | 10m | 36.95 | 0.95 | 0.12 | 6.79 | 54.86 | 0.28 |
| Metriaclima zebra | Thumbi West | 10m | 20.51 | 0.07 | 0.07 | 14.35 | 64.46 | 0.47 |
| Metriaclima zebra | Thumbi West | 10m | 22.60 | 0.75 | 0.14 | 34.18 | 41.09 | 0.99 |
| Metriaclima zebra | Thumbi West | 10m | 26.41 | 3.15 | 0.08 | 8.19 | 50.13 | 11.39 |
| Metriaclima zebra | Thumbi West | 10m | 17.28 | 1.40 | 2.34 | 22.54 | 51.98 | 4.46 |
| Metriaclima zebra | Thumbi West | 10m | 21.00 | 0.56 | 0.16 | 8.98 | 68.53 | 0.76 |
| Metriaclima zebra | Thumbi West | 10m | 13.67 | 0.84 | 1.04 | 23.55 | 59.14 | 1.77 |
| Metriaclima zebra | Thumbi West | 10m | 24.60 | 0.37 | 0.49 | 21.85 | 51.27 | 1.43 |
| Metriaclima zebra | Zimbawe Rock | 10m | 35.10 | 5.14 | 0.08 | 9.23 | 47.14 | 0.41 |
| Metriaclima zebra | Zimbawe Rock | 10m | 27.32 | 16.18 | 0.74 | 1.04 | 38.02 | 8.06 |
| Metriaclima zebra | Zimbawe Rock | 10m | 9.87 | 10.66 | 0.07 | 20.77 | 46.64 | 0.64 |
| Metriaclima zebra | Zimbawe Rock | 10m | 24.79 | 3.71 | 0.00 | 10.33 | 56.03 | 1.27 |
| Metriaclima zebra | Zimbawe Rock | 10m | 20.53 | 20.66 | 0.05 | 27.07 | 28.97 | 0.11 |
| Metriaclima zebra | Zimbawe Rock | 10m | 11.69 | 13.42 | 0.01 | 33.67 | 26.71 | 0.00 |
| Metriaclima zebra | Zimbawe Rock | 10m | 6.77 | 20.80 | 0.02 | 30.79 | 34.50 | 0.05 |
| Metriaclima zebra | Zimbawe Rock | 10m | 12.17 | 26.79 | 0.04 | 25.86 | 32.31 | 0.20 |
| Metriaclima zebra | Zimbawe Rock | 10m | 17.69 | 5.22 | 0.81 | 4.04 | 69.74 | 1.63 |
| Metriaclima zebra | Zimbawe Rock | 10m | 12.43 | 13.93 | 0.24 | 19.84 | 47.20 | 0.76 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Tropheops gracilior | Otter Island | 20m | 52.49 | 0.00 | 4.16 | 2.16 | 41.20 | 0.00 |
| Tropheops gracilior | Otter Island | 20m | 27.23 | 0.00 | 0.89 | 17.59 | 54.29 | 0.00 |
| Tropheops gracilior | Otter Island | 20m | 16.81 | 0.44 | 0.09 | 20.51 | 62.17 | 0.00 |
| Tropheops gracilior | Otter Island | 20m | 36.77 | 0.92 | 0.31 | 11.86 | 50.13 | 0.00 |
| Tropheops gracilior | Otter Island | 20m | 36.74 | 1.72 | 1.19 | 2.98 | 47.62 | 9.75 |
| Tropheops gracilior | Otter Island | 20m | 39.33 | 0.98 | 0.63 | 4.64 | 47.23 | 7.20 |
| Tropheops gracilior | Otter Island | 20m | 21.65 | 1.77 | 0.41 | 29.06 | 42.98 | 4.13 |
| Tropheops gracilior | Otter Island | 20m | 28.05 | 6.26 | 3.54 | 15.14 | 36.84 | 10.17 |
| Tropheops gracilior | Otter Island | 20m | 11.68 | 1.10 | 0.03 | 34.56 | 48.22 | 4.42 |
| Tropheops gracilior | Otter Island | 20m | 27.39 | 7.66 | 1.68 | 3.88 | 44.25 | 15.14 |
| Tropheops gracilior | Otter Island | 20m | 29.39 | 3.42 | 0.47 | 35.57 | 29.63 | 1.53 |
| Tropheops gracilior | Otter Island | 20m | 18.32 | 1.32 | 0.09 | 43.22 | 35.69 | 1.36 |
| Tropheops gracilior | Otter Island | 20m | 18.16 | 0.19 | 0.00 | 33.13 | 47.71 | 0.80 |
| Tropheops gracilior | Otter Island | 20m | 30.25 | 6.23 | 1.97 | 1.48 | 42.36 | 17.71 |
| Tropheops gracilior | Otter Island | 20m | 42.84 | 4.27 | 7.53 | 6.29 | 35.51 | 3.56 |
| Tropheops gracilior | Otter Island | 20m | 17.28 | 0.51 | 0.01 | 26.48 | 51.09 | 4.64 |
| Tropheops gracilior | Otter Island | 5m | 28.54 | 1.52 | 0.90 | 27.55 | 41.49 | 0.00 |
| Tropheops gracilior | Otter Island | 5m | 29.95 | 0.00 | 0.00 | 21.19 | 48.86 | 0.00 |
| Tropheops gracilior | Otter Island | 5m | 30.45 | 6.13 | 3.82 | 19.68 | 39.92 | 0.00 |
| Tropheops gracilior | Otter Island | 5m | 26.44 | 0.76 | 0.25 | 27.73 | 39.78 | 5.03 |
| Tropheops gracilior | Otter Island | 5m | 24.55 | 1.15 | 0.93 | 30.19 | 40.23 | 2.95 |
| Tropheops gracilior | Otter Island | 5m | 22.58 | 0.58 | 0.13 | 19.24 | 56.64 | 0.82 |
| Tropheops gracilior | Otter Island | 5m | 26.65 | 4.66 | 4.09 | 23.17 | 41.44 | 0.00 |
| Tropheops gracilior | Otter Island | 5m | 22.58 | 3.04 | 1.48 | 19.63 | 49.34 | 3.94 |
| Tropheops gracilior | Otter Island | 5m | 23.02 | 3.53 | 2.25 | 19.37 | 46.43 | 5.40 |
| Tropheops gracilior | Otter Island | 5m | 20.07 | 1.25 | 0.52 | 28.94 | 47.65 | 1.57 |
| Tropheops gracilior | Otter Island | 5m | 22.71 | 1.71 | 0.81 | 23.33 | 50.02 | 1.41 |
| Tropheops gracilior | Otter Island | 5m | 20.52 | 0.93 | 0.24 | 22.29 | 54.93 | 1.09 |
| Tropheops gracilior | Otter Island | 5m | 10.58 | 0.22 | 0.09 | 30.80 | 57.70 | 0.61 |
| Tropheops gracilior | Otter Island | 5m | 21.97 | 0.65 | 0.38 | 24.54 | 50.80 | 1.65 |
| Tropheops gracilior | Otter Island | 5m | 18.32 | 0.33 | 0.05 | 30.63 | 49.11 | 1.56 |
| Tropheops gracilior | Otter Island | 5m | 21.15 | 1.68 | 1.13 | 17.27 | 54.22 | 4.56 |
| Tropheops gracilior | Otter Island | 5m | 21.38 | 4.53 | 1.97 | 16.78 | 50.77 | 4.58 |
| Tropheops gracilior | Otter Island | 5m | 20.22 | 2.59 | 1.41 | 22.07 | 51.02 | 2.71 |
| Tropheops gracilior | Otter Island | 5m | 19.94 | 8.61 | 8.48 | 17.20 | 36.98 | 8.78 |
| Tropheops gracilior | Otter Island | 5m | 20.68 | 2.04 | 1.21 | 19.31 | 54.38 | 2.39 |

**Table S3-2:** Mean relative expression values for each experimental group +/- standard deviations.

| Species | Source | Depth | SWS1 | SWS2B | SWS2A | RH2B | RH2Aa | LWS |
|---|---|---|---|---|---|---|---|---|
| Metriaclima zebra | Thumbi West | 10m | 22.60 +/- 6.12 | 0.80 +/- 0.88 | 0.86 +/- 1.53 | 17.96 +/- 8.31 | 55.64 +/- 7.64 | 2.00 +/- 3.36 |
| Metriaclima zebra | Zimbawe Rock | 10m | 17.84 +/- 8.90 | 13.65 +/- 7.68 | 0.21 +/- 0.31 | 18.26 +/- 11.46 | 42.73 +/- 13.36 | 1.31 +/- 2.43 |
| Copadichromis eucinostomus | Otter Point | 5m | 18.40 +/- 3.53 | 0.34 +/- 0.35 | 0.02 +/- 0.02 | 16.72 +/- 3.84 | 35.51 +/- 8.86 | 29.01 +/- 6.64 |
| Copadichromis eucinostomus | Otter Point | 20m | 16.74 +/- 5.95 | 2.62 +/- 6.87 | 0.12 +/- 0.22 | 19.42 +/- 6.04 | 53.90 +/- 12.21 | 7.18 +/- 18.99 |
| Tropheops gracilior | Otter Island | 5m | 22.62 +/- 4.47 | 2.29 +/- 2.23 | 1.51 +/- 2.01 | 23.04 +/- 4.73 | 48.09 +/- 6.20 | 2.45 +/- 2.32 |
| Tropheops gracilior | Otter Island | 20m | 28.40 +/- 11.09 | 2.30 +/- 2.49 | 1.44 +/- 2.06 | 18.03 +/- 14.04 | 44.81 +/- 8.08 | 5.02 +/- 5.57 |

# Appendix III: Supplemental Material for Chapter 4

**Table S4-1:** Full multivariate correlation matrix for relative gene expression and OMR tresholds in different light environments. All values are determined using a minimum *p* value of 0.05. Gene expression-behavior correlations are highlighted in red text, while the other correlations are in black text.

| | Red Stimulus | Green Stimulus | Blue-Green Stimulus | Blue Stimulus | Relative RH2B | Relative RH2A | Relative LWS |
|---|---|---|---|---|---|---|---|
| Red Stimulus | 1 | | | | | | |
| Green Stimulus | 0.6262 | 1 | | | | | |
| Blue-Green Stimulus | 0.2854 | 0.034 | 1 | | | | |
| Blue Stimulus | 0.5109 | 0.3059 | 0.8733 | 1 | | | |
| Relative RH2B | -0.1236 | -0.039 | 0.0597 | -0.0024 | 1 | | |
| Relative RH2A | 0.2736 | 0.2484 | -0.204 | -0.073 | -0.4432 | 1 | |
| Relative LWS | -0.2256 | -0.2504 | 0.1885 | 0.0823 | -0.1323 | -0.8299 | 1 |

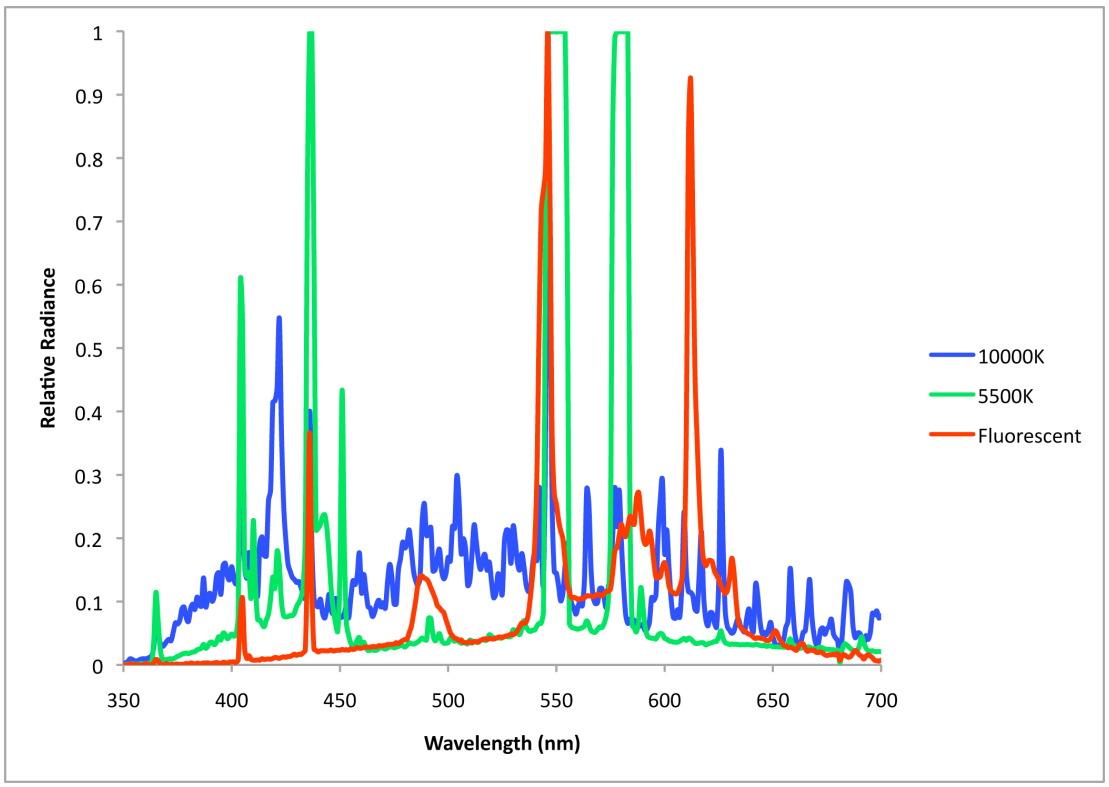



**Figure S4-1:** Relative radiance spectra for the rearing treatment groups. The fluorescent (red) line represents the narrow rearing treatment. The 5500K (green) and 10000K (blue) lines represent the two bulb types used for the broad rearing treatment.

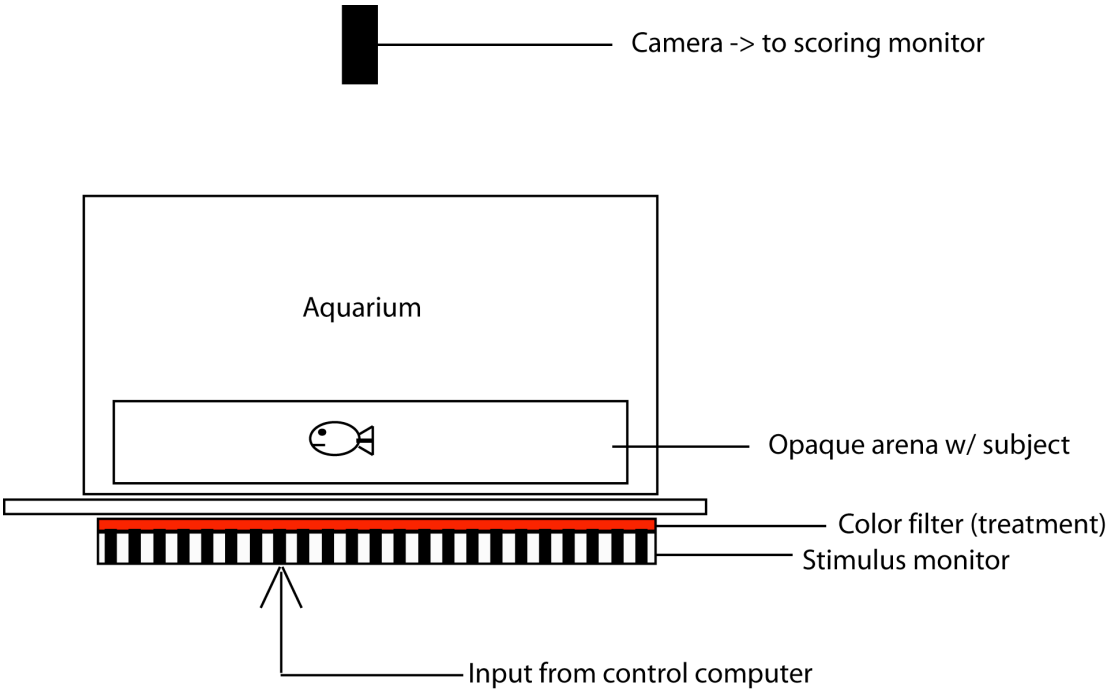
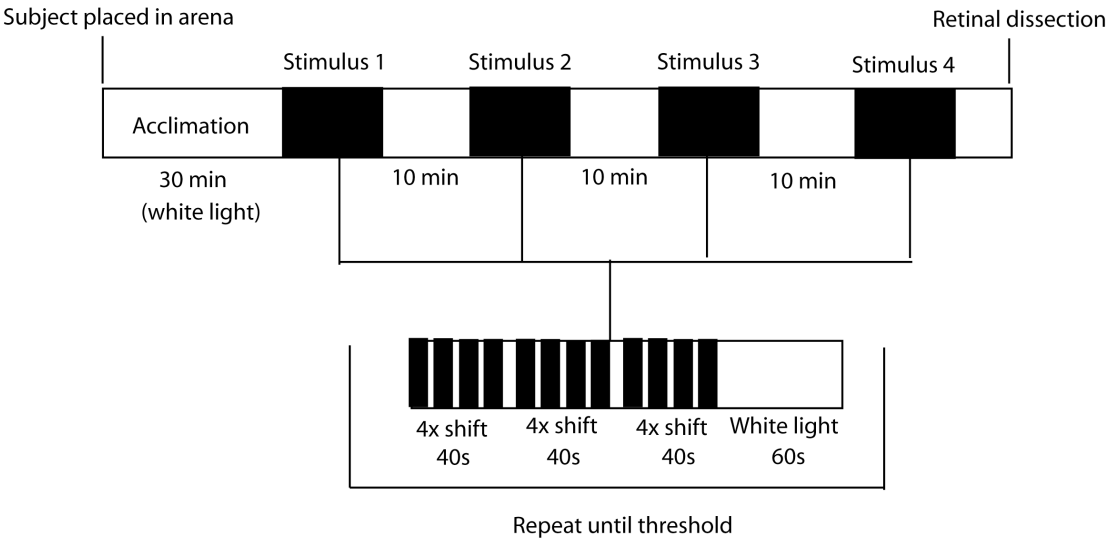



**Figure S4-2:** Schematic of the experimental apparatus.




Hypothetical sequence of keys (lower key number = higher intensity gradient)

Sequence 1: Keys 4, 5, 6

If yes on 4 only - Sequence 2: Keys 3, 4, 5

If yes on 5 - Sequence 2: Keys 4, 5, 6

If yes on 6 - Sequence s: Keys 6, 7, 8

**Figure S4-3:** Basic schematic of the experimental stimulus and presentation method.

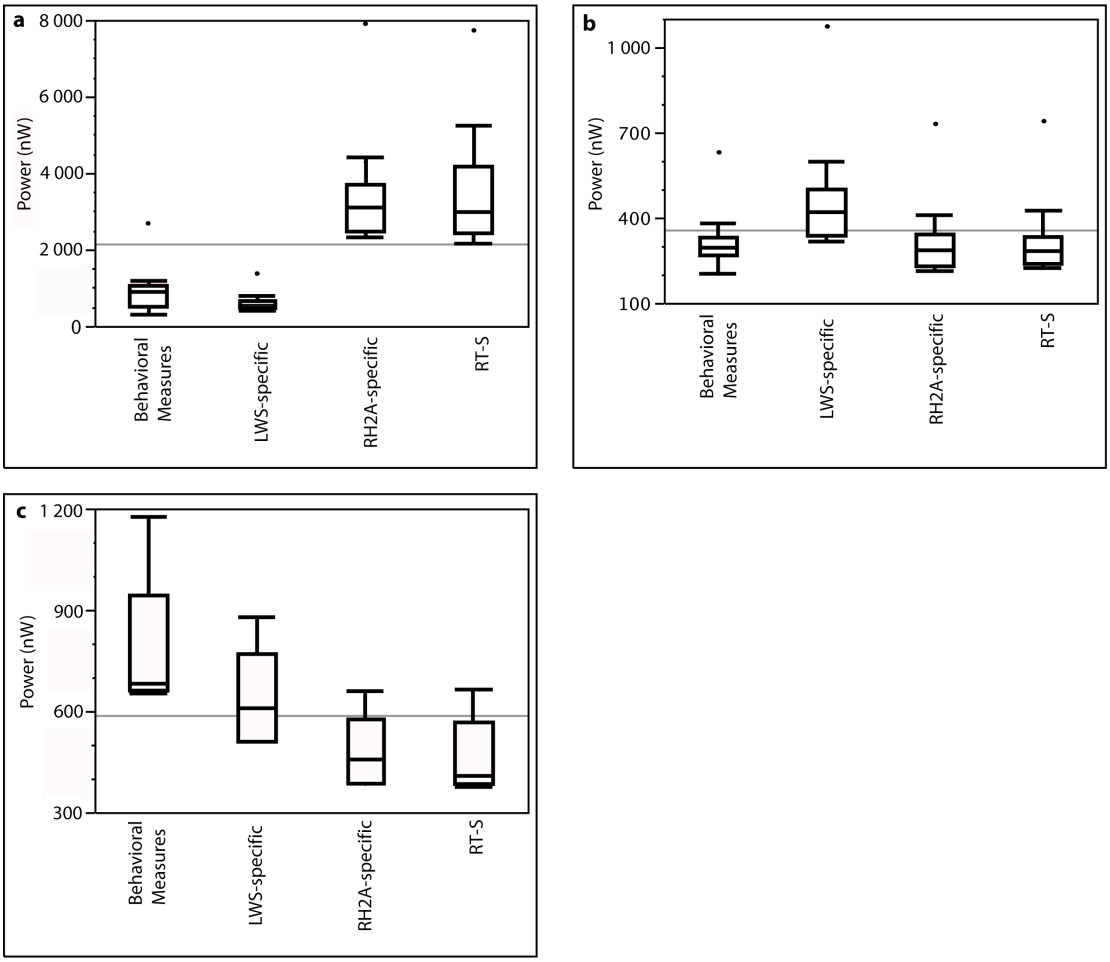



**Figure S4-4:** Behavioral predictive model performance for the broad-spectrum treatment *Metriaclima lombardoi* treatment for the (a) red, (b) blue-green, and (c) blue light environments. The gray bar represents the grand mean of the ANOVA model.

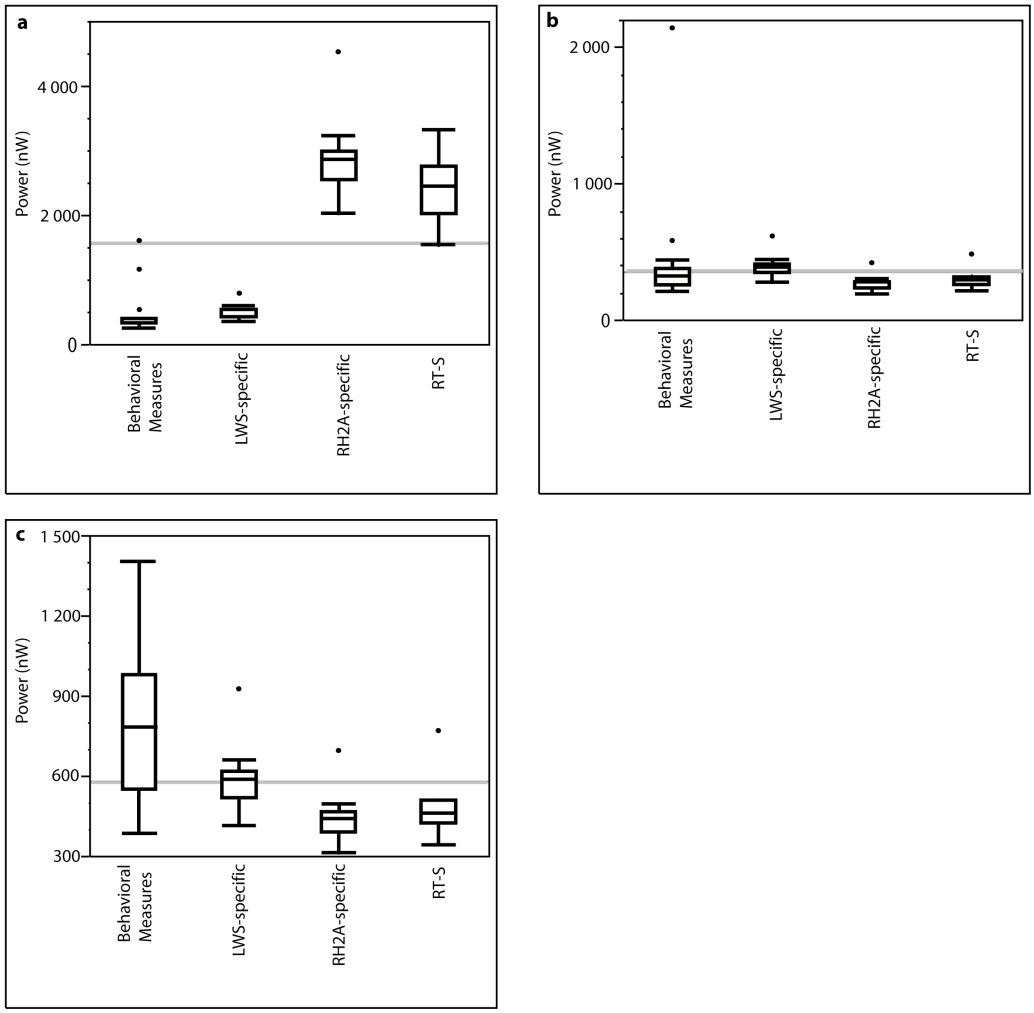



**Figure S4-5:** Behavioral predictive model performance for the narrow-spectrum treatment *Metriaclima lombardoi* treatment for the (a) red, (b) blue-green, and (c) blue light environments. The gray bar represents the grand mean of the ANOVA model.

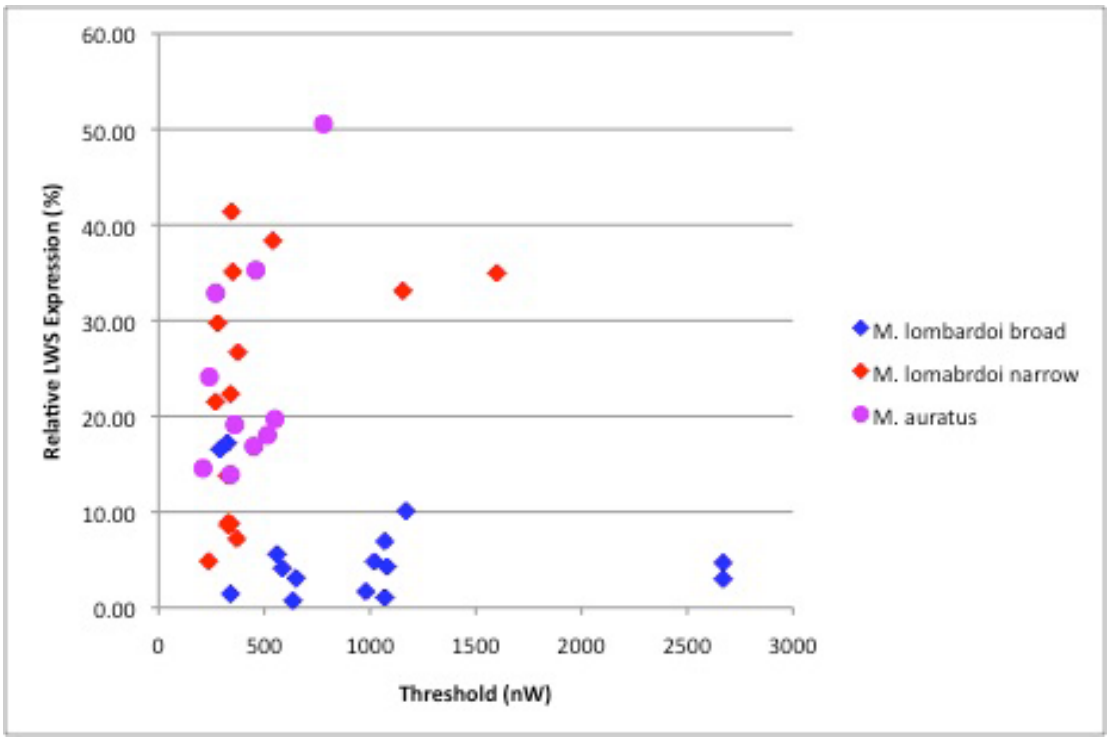



**Figure S4-6:** Scatterplot of all data points for the optomotor threshold under the red filter (nW) against relative LWS expression. Pink circles represent *Melanochromis auratus*, red diamonds represent narrow-spectrum treatment *Metriaclima lombardoi*, and blue diamonds represent broad-spectrum treatment *M. lombardoi*.

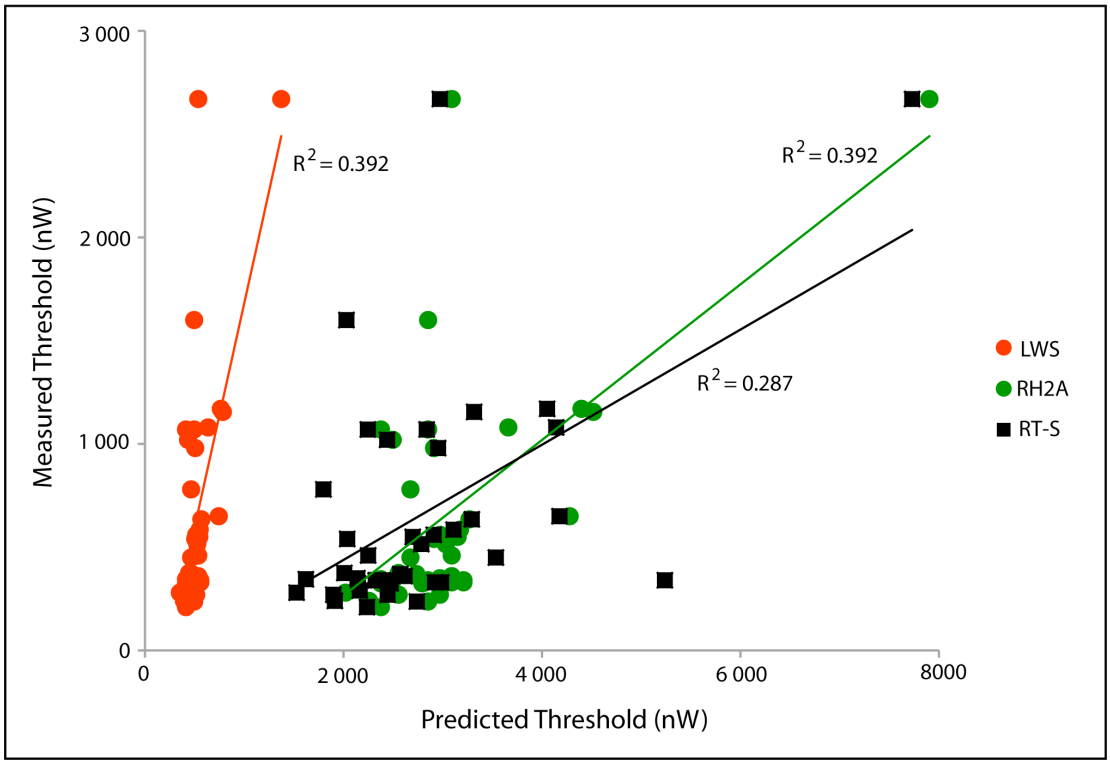



**Figure S4-7:** Linear regressions for model predictions vs. observed thresholds for all individuals for the red stimulus. The single pigment (LWS and RH2A) models predict the spread of data better than the RT-S model. The linear fit of the LWS and RH2A models is equivalent because they are both fit to a Govardovskii curve of the same shape and magnitude. This results in a constant r-squared value for single pigment models but differing prediction means.

# Bibliography


Anstis S, Hutahajan P, Cavanagh P (1998) Optomotor test for wavelength sensitivity in guppyfish (*Poecilia reticulata*). *Vision Research,* **38**, 45-53.

Asenjo AB, Rim J, Oprian DO (1994) Molecular determinants of human red/green color discrimination. *Neuron*, **12**, 1131-1138.

Azuma Y, Kumuzawa Y, Miya M, Mabuchi K, Nishida M (2008) Mitogenomic evaluation of the historical biogeography of cichlids toward reliable dating of teleostean divergences. *BMC Evolutionary Biology*, **8**, 215.

Baerends GP, Baerends-van Roon JM (1950) *An introduction to the study of the ethology of cichlid fishes*. Leiden E.J. Brill

Beavis WD (1998) QTL analyses: power, precision, and accuracy. In *Molecular dissection of complex traits*. Puberson, AH (ed) pp. 145-162, CRC Press.

Borgia G (1986) Sexual selection in bowerbirds. *Scientific American*, **254**, 92-101.

Boughman JW (2002) How sensory drive can promote speciation. *Trends in Ecology and Evolution*, **17**, 571-577.

Brown MB, Forsythe AB (1974) Robust tests for equality of variances. *Journal of the American Statistical Association*, **69**, 364-367.

Carleton KL (2009) Cichlid visual systems: mechanisms of spectral tuning. *Integrative Zoology*, **4**, 75-86.

Carleton KL, Harosi FI, Kocher TD (2000) Visual pigments of African cichlids: evidence for ultraviolet vision from microspectrophotometry and DNA sequences. *Vision Research*, **40**, 879-890.

Carleton KL, Hofmann CM, Klisz C, Patel Z, Chircus LM, Simenauer LH, Soodoo N, Albertson RC, Ser JR (2010) Genetic basis of differential opsin gene expression in cichlid fishes. *Journal of Evolutionary Biology*, **23**, 840-853.

Carleton KL, Parry JWL, Bowmaker JK, Hunt DM, Seehausen O (2005) Colour vision and speciation in Lake Victoria cichlids of the genus *Pundamilia*. *Molecular Ecology*, **14**, 4341-4353.

Carleton KL, Spady T, Cote RH (2005) Rod and cone opsin families differ in spectral tuning domains but not signal transducing domains as judged by saturated evolutionary trace analysis. *Journal of Molecular Evolution*, **61**, 75-89.

Carleton KL, Spady TC, Kocher TD (2006) Visual communication in East African cichlid fishes: diversity in a phylogenetic context. In: Ladich F, Collin SP, Moller P, Kapoor BG (eds) Communication in Fishes. Science Publishers Inc, Enfield NH, pp. 485-515

Chang BSW, Crandall KA, Carulli JP, Hartl D (1995) Opsin phylogeny and evolution: a model for blue shifts in wavelength regulation. *Molecular Phylogenetics and Evolution*, **4**, 31-43.

Couldridge VCK, Alexander GJ (2002) Color patterns and species recognition in four closely related species of Lake Malawi cichlid. *Behavioral Ecology*, **13**, 59-64.

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

Dalton BE, Cronin TW, Marshall NJ, Carleton KL (2010) The fish eye view: are

cichlids conspicuous? *Journals of Experimental Biology*, **213**, 2243-2255.

Dangles O, Irschick D, Chittka L, Casas J (2009) Variability in sensory ecology: expanding the bridge between physiology and evolutionary biology. *The Quarterly Review of Biology*, **84**, 51-74.

Danley PD, Kocher TD (2001) Speciation in rapidly diverging systems: lessons from Lake Malawi. *Molecular Ecology*, **10**, 1075-1086.

Dobzhansky T (1973) Nothing in biology makes sense except in the light of evolution. *American Biology Teacher*, **35**, 125-129.

Dominey WJ (1984) Effects of sexual selection and life history on speciation: species flocks in African cichlids and Hawaiian *Drosophila*. In *Evolution of fish species flocks* (ed. Echelle AA, Kornfield I) pp231-249: University of Maine at Orono Press.

Ebrey T, Koutalos Y (2001) Vertebrate photoreceptors. *Progress in Retinal and Eye Research*, **20**, 49-94.

Endler JA (1992) Signals, signal conditions and the direction of evolution. *American Naturalist,* **139**, S125-153.

Endler JA (1993) Some general comments on the evolution and design of animal communication systems. *Philosophical Transactions of the Royal Society of London B*, **340**, 215-225.

Ewert J-P (1984) Tectal mechanisms that underlie prey-catching and avoidance behaviors in toads. In *Comparative neurology of the optic tectum* (ed Vanages H) pp247-416: Plenum Publishing Corp.

Fisher RA (1930) *The genetical theory of evolution*. Clarendon Press, Oxford.

Fuller RC, Carleton KL, Fadool JM, Spady TC, Travis J (2004) Population variation in opsin expression in the bluefin killifish, *Lucania goodei*: a real-time PCR study. *Journal of Comparative Physiology A*, **190**, 147-154.

Fuller RC, Carleton KL, Fadool JM, Spady TC, Travis J (2005) Genetic and environmental variation in the visual properties of bluefin killifish *Lucania goodei*. *Journal of Evolutionary Biology*, **18**, 516-523.

Fuller RC, Noa LA (2010) Female mating preferences, lighting environment, and a test of the sensory bias hypothesis in the bluefin killifish. *Animal Behavior,* **80**, 23-35.

Fuller RC, Noa LA, Strellner RS (2010) Teasing apart the many effects of lighting environment on opsin expression and foraging preference in bluefin killifish. *American Naturalist,* **176**, 1-13.

Genner MJ, Seehausen O, Lunt DH, Joyce DA, Shaw PW, Carvalho GR, Turner GF (2007) Age of cichlids: new dates for ancient lake fish radiations. *Molecular Biology and Evolution*, **24**, 1269-1282.

Genner MJ, Turner GF (2005) The mbuna cichlids of Lake Malawi: a model for rapid speciation and adaptive radiation. *Fish and Fisheries*, **6**, 1-34.

Govardovskii V, Fyhrquist N, Reuter T, Kuzmin D, Donner K (2000) In search of the visual pigment template. *Visual Neuroscience,* **17**, 509-528.

Graham DJ, Midgley NG (2000) Graphical presentation of particle shape using triangular diagrams: an Excel spreadsheet method. *Earth Surface Processes and Landforms,* **25**, 1473-1477.

Hanlon RT, Naud M.-J., Saw PW, Havenhand JN (2005) Transient sexual mimicry

leads to fertilization. *Nature*, **430**, 212.

Heffner HE, Koay G, Heffner RS (2008) Comparison of behavioral and auditory brainstem response measures of threshold shift in rats exposed to loud sound. *Journal of the Acoustic Society of America,* **124**, 1093-1104.

Hofmann CM, Carleton KL (2009) Gene duplication and differential gene expression play an important role in the diversification of visual pigments in fish. *Integrative Comparative Biology*, **49**, 630-643.

Hofmann CM, O'Quin KE, Marshall NJ, Cronin TW, Carleton KL (2009) The eyes have it: regulatory and structural changes both underlie cichlid visual pigment diversity. *PLoS Biology*, **7**, e1000266.

Hofmann CM, O'Quin KE, Smith A, Carleton KL (2010) Plasticity of opsin gene expression in cichlids from Lake Malawi. *Molecular Ecology,* **19**, 2064-2074.

Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *Journal of Physiology,* **160**, 106-154.

Jordan R, Kellogg K, Howe D, Juanes F, Stauffer JRJ, Loew E (2006) Photopigment spectral absorbance of Lake Malawi cichlids. *Journal of Fish Biology*, **68**, 1291-1299.

Jordan R, Kellogg K, Juanes F, Stauffer JRJ (2003) Evaluation of female mate choice cues in a group of Lake Malawi *mbuna* (cichlidae). *Copeia,* **1**, 181-186.

Kawata M, Shoji A, Kawamura S, Seehausen O (2007) A genetically explicit model of speciation by sensory drive within a continuous population in aquatic environments. *BMC Evolutionary Biology*, 7, 99.

Kidd MR, Danley PD, Kocher TD (2006) A direct assay of female choice in cichlids:

all the eggs in one basket. *Journal of Fish Biology*, **68**, 373-384

Kirkpatrick M, Ravigne V (2002) Speciation by natural and sexual selection: models and experiments. *American Naturalist*, **159**, S22-S35.

Kocher TD (2004) Adaptive evolution and explosive speciation: the cichlid fish model. *Nature Reviews Genetics,* **5**, 288-298.

Kornfield I, Smith PF (2000) African cichlid fishes: model systems for evolutionary biology. *Annual Review of Ecology and Systematics*, **31**, 163-196.

Kroger RHH, Knoblauch B, Wagner HJ (2003) Rearing in different photic and spectral environments changes the optomotor response to chromatic stimuli in the cichlid fish *Aequidens pulcher*. *Journal of Experimental Biology*, **206**, 1643-1648.

Lande R (1981) Models of speciation by sexual selection on polygenic traits. *Proceedings of the National Academy of Science,* **78**, 3721-3725.

Lee BB, Sun H (2004) Chromatic input to cells of the magnocellular pathway: mean chromaticity and the relative phase of modulated lights. *Visual Neuroscience*, **21**, 309-314.

Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451-1452.

Lisney TJ, Studd E, Hawryshyn CW (2010) Electrophysiological assessment of spectral sensitivity in adult Nile tilapia *Oreochromis niloticus*:evidence for violet sensitivity. *Journal of Experimental Biology*, **213**, 1453-1463.

Loh Y-HE, Katz LS, Mims MC, Kocher TD, Yi SV, Streelman JT (2008)

Comparative analysis reveals signatures of differentiation and genomic polymorphisms in Lake Malawi cichlids. *Genome Biology*, **9**, R113.

Maan ME, Hofker KD, van Alphen JJM, Seehausen O (2006) Sensory drive in cichlid speciation. *American Naturalist*, **167**, 947-954.

Maan ME, Seehausen O, Soderberg L, Johnson L, Ripmeester EAP, Mrosso HDJ, Taylor MI, van Dooren TJM, van Alphen JJM (2004) Intraspecific sexual selection on a speciation trait, male coloration, in the Lake Victoria cichlid *Pundamilia nyererei*. *Proceedings of the Royral Society of London B Biology*, **271**, 2445-2452.

Mayr E. (1982) Speciation and macroevolution. *Evolution,* **36**, 1119-1132.

Mueller KP, Neuhass SCF (2010) Quantitative measurements of the optokinetic response in adult fish. *Journal of Neuroscience Methods*, **186**, 29-34.

Neumeyer C (2003) Wavelength dependence of visual acuity in goldfish. *Journal of Comparative Physiology A*, **189**, 811-821.

Orger MB, Baier H (2005) Channeling of red and green cone inputs to the zebrafish optomotor response. *Visual Neuroscience*, **22**, 275-281.

Orger MB, Smear MC, Anstis S, Baier H (2000) Perception of fourier and non-fourier motion by larval zebrafish. *Nature Neuroscience,* **3**, 1128-1133.

Palczewski K, Kumusaka 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.

Parry JWL, Carleton KL, Spady T, Carboo A, Hunt DM, Bowmaker JK (2005) Mix

and match color vision: tuning spectral sensitivity by differential opsin gene expression in Lake Malawi cichlids. *Current Biology,* **15**, 1734-1739.

Porter ML, Bok MJ, Robinson PR, Cronin TW (2009) Molecular diversity of visual pigments in Stomatopoda (Crustacea). *Visual Neuroscience*, **26**, 255-265.

Ramon y Cajal S (c1900) *Structure of the mammalian retina*. Madrid.

Ribbink AJ, Marsh BA, March AC, Ribbink AC, Sharp BJ (1983) A preliminary survey of the cichlid fishes of rocky habitats in Lake Malawi. *South African Journal of Zoology*, **18**, 149-310.

Rovner JS, Barth FG (1981) Vibratory communication through living plants by a tropical wandering spider. *Science*, **214**, 464-466.

Sabbah S, Lamela R, Gray SM, Hawryshyn CW (2010) Functional diversity in the color vision of cichlid fishes. *BMC Biology*, **8**, 133.

Schaerer S, Neumeyer C (1996) Motion detection in goldfish investigated with the optomotor response is "color blind". *Vision Research*, **36**, 275-281.

Schluter D (2001) Ecology and the origin of species. *TRENDS in Ecology and Evolution*, **16**, 372-380.

Seehausen O (2004) Hybridization and adaptive radiation. *TRENDS in Ecology and Evolution*, **19**, 198-207.

Seehausen O, Terai Y, Magalhaes IS, Carleton KL, Mrosso HDJ, Miyagi R, van der Sluijs I, Schneider MV, Maan ME, Tachida H, Imai H (2008) Speciation through sensory drive in cichlid fish. *Nature*, **455**, 620-627.

Seehausen O, van Alphen JJM (1998) The effect of male coloration on female mate

choice in closely related Lake Victoria cichlids *(Haplochromis nyererei* complex). *Behavioral Ecology and Sociobiology*, **42**, 1-8.

Seehausen O, van Alphen JJM (1999) Can sympatric speciation by disruptive sexual selection explain rapid evolution of cichlid diversity in Lake Victoria? *Ecology Letters*, **2**, 262-271.

Seehausen O, van Alphen JJM, Witte F (1997) Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science*, **277**, 1808-1811.

Shand J, Davies WL, Thomas N, Balmer L, Cowing JA, Pointer M, Carvalho LS, Trezise AEO, Colling SP, Beazley LD, Hunt DM (2008) The influence of ontogeny and light environment on the expression of visual pigmet opsins in the retina of the black bream, *Acanthopagrus butcheri*. *Journal of Experimental Biology*, **211**, 1495-1503.

Smith AR, Carleton KL (2010) Allelic variation in Malawi cichlid opsins: a tale of two genera. *Journal of Molecular Evolution,* **70**, 593-604.

Smith AR, D'Annunzio L, Smith AE, Sharma A, Hofmann CM, Marshall NJ, Carleton KL (2011) Intraspecific cone opsin expression variation in the cichlids of Lake Malawi. *Molecular Ecology,* **20**, 299-310.

Spady TC, Parry JWL, Robinson PR, Hunt DM, Bowmaker JK, Carleton KL (2006) Evolution of the cichlid visual palette through ontogenetic subfunctionalization of the opsin gene arrays. *Molecular Biology and Evolution*, **23**, 1538-1547.

Spady TC, Seehausen O, Loew ER, Jordan RC, Kocher TD, Carleton KL (2005)

Adaptive molecular evolution in the opsin genes of rapidly speciating cichlid species. *Molecular Biology and Evolution*, **22**, 1412-1422.

Straw AD (2008) Vision Egg: an open-source library for realtime visual stimulus generations. *Frontiers in Neuroinformatics,* **2**, 4.

Streelman JT, Albertson RC (2006) Evolution of novelty in the cichlid dentition. J *Experimental Biology*, **306B**, 216-226.

Sugawara T, Terai Y, Imai H, Turner GG, Kolbmuller S, Sturmbauer C, Shichida Y, Okada N (2005) Parallelism of amino acid changes of the RH1 affecting spectral sensitivity among deep-water cichlids from Lakes Tanganyika and Malawi. *PNAS*, **102**, 5448-5453.

Sun H, Smithson HE, Zaidi Q, Lee BB (2006) Do magnocellular and parvocellular ganglion cells avoid short-wavelength cone input? *Visual Neuroscience*, **23**, 441-446.

Swofford DL (2002) PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sunderland, MA: Sinauer Associates.

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

Terai Y, Mayer WE, Klein J, Tichy H, Okada N (2002) The effect of selection on a long wavelength sensitive (LWS) opsin gene of Lake Victoria cichlid fishes. *Proceedings of the National Academy of Sciences*, **99**, 15501-15506.

Terai Y, Seehausen O, Sasaki T, Takahashi K, Mizoiri S, Sugawara T, Sato T,

Watanabe M, Konijnendijk N, Mrosso HDJ, Tachinda H, Imai H, Shichida Y, Okada N (2006) Divergent selection on opsins drives incipient speciation in Lake Victoria cichlids. *PLoS Biology*, **4**, 2244-2251.

Weatherhead PJ, Robertson RJ (1979) Offspring quality and the polygyny threshold: 'The sexy son hypothesis'. *American Naturalist*, **113**, 201-208.

Wright S (1932) The roles of mutation, inbreeing, crossbreeding and selection in evolution. In *Proceedings of the Sixth International Congress on Genetics*: 355-366.

Yokoyama S (2008) Evolution of dim-light and color vision pigments. *Annual Review Genomics and Humam Genetics*, **9**, 259-282.

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, Zhang H, Britt L (2008) Elucidation of phenotypic adaptations: molecular analyses of dim-light vision proteins in vertebrates. *Proceedings of the National Academy of Sciences USA*, **105**, 13480-13485.

Yokoyama S, Yokoyama R (1996) Adaptive evolution of photoreceptors and visual pigments in vertebrates. *Annual Review of Ecology Systematics*, **27**, 543-567.

Zahavi A (1975) Mate selection – a selection for a handicap. *Journal of Theoretical Biology*, **53**, 205-214.

Zana Y, Ventura DF, de Souza JM, DeVoe RD (2001) Tetrachromatic input to turtle horizontal cells. *Visual Neuroscience*, **18**, 759-765.

Zwickl DJ (2006) Genetic algorithm approaches for the phylogenetic analysis of large

biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation, The University of Texas at Austin.