Beauty in the Eyes of the Beholders: Colour Vision and Mate Choice in the Family Poeciliidae

by

Benjamin Alexander Sandkam

B.Sc. (Integrative Biology), University of Illinois at Urbana-Champaign, 2009

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
Department of Biological Sciences
Faculty of Science

© Benjamin Alexander Sandkam 2015

SIMON FRASER UNIVERSITY

Fall 2015

All rights reserved.
However, in accordance with the *Copyright Act of Canada*, this work may be reproduced, without authorization, under the conditions for "Fair Dealing." Therefore, limited reproduction of this work for the purposes of private study, research, criticism, review and news reporting is likely to be in accordance with the law, particularly if cited appropriately.

Approval

Name: **Benjamin Alexander Sandkam**

Degree: **Doctor of Philosophy (Biological Sciences)**

Title: ***Beauty in the Eyes of the Beholders: Colour Vision and Mate Choice in the Family Poeciliidae***

Examining Committee: **Chair:** Gerhard Gries
Professor

Felix Breden
Senior Supervisor
Professor

Wendy Palen
Supervisor
Associate Professor

Kimberly Hughes
Supervisor
Professor
Department of Biological Science
Florida State University

Michael Hart
Internal Examiner
Professor
Department of Biological Sciences

Karen Carleton
External Examiner
Associate Professor
Department of Biology
University of Maryland, College Park

Date Defended/Approved: September 21, 2015

Ethics Statement

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

c. as a co-investigator, collaborator or research assistant in a research project approved in advance,

or

d. as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, British Columbia, Canada

update Spring 2010

Abstract

Sexual selection plays a major role in numerous aspects of evolution. Many models have attempted to explain how mate preferences evolve both across populations within a species and across species. 'Sensory bias' predicts that the traits involved in mate choice will co-evolve with the tuning of the sensory systems responsible for detecting such traits. The family Poeciliidae is a classic system for studies of mate choice and provides an excellent opportunity to examine the co-evolution of preference for colour traits and the sensory system detecting such traits: colour vision. In this dissertation, I present a body of work investigating how colour vision differs across species and populations, thus exploring the potential role sensory systems have in shaping mate preferences. To do this, I focus on the opsin genes, which play a predominant role in tuning the wavelength sensitivity of cone cells – the detectors for colour vision. I found the Long Wavelength Sensitive opsins (detecting red/orange colours) experience high rates of gene conversion due to their genomic architecture. The effects of conversion may be influenced by the importance of red/orange in mate choice decisions. While traditional models of duplication and divergence suggest sensory repertoire expansion occurs slowly, I found hybridization can expand sensory repertoires in one generation. I have termed this process: Hybrid Sensory Expansion. I then focus on one species to show that differences in visual tuning (gene expression and allele frequency) co-vary with mate preferences across populations in a manner that is consistent with the Sensory Exploitation (SE) model for the evolution of female mate preferences. However, I go on to find that closely related, highly sympatric species differ in colour vision more across populations than across species within populations on mainland South America. This suggests that while SE could explain differences in mate preference across populations, it may not scale up to explain species level differences as generally assumed. Taken together, these results show that the evolution of visual tuning may not always evolve through traditional mutation-selection models and that visual systems are far more variable across populations within species than generally assumed.

Keywords: Opsin; Sensory System; Ecological Genetics; Sensory Bias; Sensory Exploitation; Sexual Selection

For family, friends, fish and fun

Acknowledgements

As this thesis is a result of lifelong learning both in the classroom and out, I owe a great deal to a great many people that have touched my life over the years. Specifically, I am tremendously thankful for the wonderful guidance and support of my advisor Dr. Felix Breden who always went along with my crazy plans for projects in both the lab and field- even when it meant exhausting fieldwork. Felix has been not only an excellent mentor but also a great friend and role model.

It is impossible to express my gratitude to my wife, Dr. Nora Prior, for her boundless support, insightful intellectual contributions and her uncanny ability to keep us moving forward. I thank our daughter, Ardea Sandkam, for keeping me active and sharing a love of the outdoors. I am also incredibly thankful for the amazing support of my in-laws, Dr. Julie Hengst and Dr. Paul Prior for their amazing patience, support and guidance into the world of academia. Thank you also to my close friend Matt Taves, for excellent scientific discussions while exploring nature on and under the water. I thank my parents, Tracy and Jeff Sandkam, for their support and kindling a lifelong love of the natural world.

I owe a great deal to all members of the Breden and FAB* labs past and present, specifically: Drs. Mike Hart, Arne Mooers, Bernie Crespi, Corey Watson, Jeff Joy, Will Stein, Mika Mokkonen, Kristen Gorman, and Iva Popovic, for insightful feedback, training, and excellent camaraderie. Megan Young and Kristina Pohl were invaluable to my ability to work off campus and I am very grateful for their patience and assistance.

A big thank you to my committee members, Drs. Kim Hughes and Wendy Palen, for the great feedback and understanding while my thesis morphed into what it has become. Thank you also to Dr. Karen Carleton for graciously taking the time to provide outside perspective as my external examiner.

While not directly involved with the work of this thesis, I am incredibly grateful to Dr. Isabelle Côté for training in scientific SCUBA diving, and inspirational enthusiasm for all things aquatic. Thank you also, to my undergraduate advisor, Dr. Becky Fuller, for the opportunities, advice and support that has continued even long after leaving her lab.

Table of Contents

Approval ii
Ethics Statement iii
Abstract iv
Dedication v
Acknowledgements vi
Table of Contents vii
List of Tables x
List of Figures xii
Introductory Image xiv

Chapter 1. Introduction 1
The Family Poeciliidae 2
Mate Choice Evolution 3
Colour Vision 5
Objectives 7
References 9

Chapter 2. The impact of genomic environment on the evolution of colour vision in the family Poeciliidae, a model for visually based sexual selection 16
Publication and Contributions 16
2.1. Abstract 16
2.2. Introduction 17
2.3. Materials and Methods 19
2.3.1. Sequencing 19
2.3.2. Tree Building 20
2.3.3. Gene Conversion Analyses in LWS Opsins 21
2.3.4. Identifying Sites Under Selection in Non-LWS Genes 22
2.4. Results 22
2.4.1. Data Set and Alignment 22
2.4.2. Phylogenetic Analyses 23
2.4.3. Gene Conversion Analyses 23
2.4.4. Identifying Sites Under Selection in Non-LWS Genes 23
2.5. Discussion 24
2.5.1. Gene Conversion, Not Independent Duplications 25
2.5.2. Why is Gene Conversion so Prevalent Between LWS-1 and LWS-3? 25
2.5.3. Implications of Gene Conversion to Colour Vision 26
2.5.4. Selection in Non-LWS Genes 28
2.6. Conclusion 28
Acknowledgements 29
Data Accessibility 29
References 30
Tables and Figures 35

Chapter 3. Hybridization leads to sensory repertoire expansion in a gynogenetic fish, the Amazon molly (*Poecilia formosa*): a test of the hybrid-sensory expansion hypothesis 55
Publication and Contributions 55
3.1. Abstract 55
3.2. Introduction 56
3.3. Materials and Methods 58
3.3.1. Sample Preparation, PCR, Cloning and Sequencing 58
3.3.2. Phylogenetic Analyses 59
3.3.3. Determination of Alleles Expressed in a Hybrid Species 60
3.3.4. Behavioural Response to Coloured Disks 61
3.4. Results 62
3.4.1. Dataset 62
3.4.2. Phylogenetic Analysis 63
3.4.3. Determination of Alleles Expressed in a Hybrid Species 63
3.4.4. Behavioural Response to Coloured Disks 64
3.5. Discussion 64
3.5.1. Strong Evidence of HSE in *Poecilia formosa* 64
3.5.2. What Does HSE Mean For *P. formosa* Vision? 66
3.5.3. Unanticipated Pecking Behaviour 67
3.5.4. Fitness Implications of HSE 68
3.6. Conclusion 68
Acknowledgements 69
Data Accessibility 69
References 70
Boxes, Figures and Tables 75
Supplementary Tables 81

Chapter 4. Beauty in the eyes of the beholders: Colour vision is tuned to mate preference in the Trinidadian guppy (*Poecilia reticulata*) 84
Publication and Contributions 84
4.1. Abstract 84
4.2. Introduction 85
4.3. Materials and Methods 88
4.3.1. Sample Collection and Environmental Parameters 88
4.3.2. qPCR Assay Design 89
4.3.3. Sample Processing and Analyses 91
4.3.4. LWS-1 A/S Allele Frequency 93
4.3.5. Statistical Analyses of Opsin Expression 93
4.4. Results 94
4.4.1. Guppy Opsin Expression Differs Across Populations 94
4.4.2. Low Predation Populations Express Higher Levels of LWS Opsins 95
4.4.3. Gene Frequencies of the 180 Ala Versus Ser Allele of LWS-1 96
4.4.4. Environmental Parameters 96
4.5. Discussion 97
4.5.1. Guppy Colour Vision Differs Across Populations 97
4.5.2. Low Predation Populations Express Higher Levels of LWS Opsins 99
4.5.3. Implications for Evolutionary Theory 101

4.6. Conclusion 103
Acknowledgements 103
Data Accessibility 103
References 104
Tables and Figures 110
Supplementary Tables and Figures 116

Chapter 5. Colour vision varies more among populations than species of live-bearing fish from South America 127
Publication and Contributions 127
5.1. Abstract 127
5.2. Introduction 128
5.3. Materials and Methods 130
5.3.1. Sample Collection 130
5.3.2. qPCR Assay Design 131
5.3.3. Sample Processing and Analyses for Opsin Expression 133
5.3.4. LWS-1 A/S Allele Frequency 135
5.3.5. Statistical Analyses of Opsin Expression 136
Do species and/or populations differ in colour vision? 136
Does opsin expression differ more by species or locations? 136
How do species differ in colour vision? 137
5.4. Results 137
5.4.1. *Do species and/or populations differ in color vision?* 137
5.4.2. *Does opsin expression differ more by species or location?* 137
5.4.3. *How do species differ in color vision?* 138
5.4.4. Gene Frequencies of the 180 Ala vs. Ser Allele Vary Across Species 138
5.5. Discussion 138
5.5.1. Colour Vision Differs Across Populations Within Species 139
5.5.2. Sensory variation and mate choice divergence 141
5.6. Conclusion 143
Acknowledgements 143
References 143
Tables and Figures 149
Supplementary Tables 155

Chapter 6. Concluding remarks and syntheses: Can beauty be found in the eyes of the beholder? 159
6.1. Colour Vision Can Vary Considerably Within Species 160
6.2. The Value of Sensory Exploitation as a Theoretical Framework for Understanding Mate Preferences 162
6.3. Conclusions 164
References 165

List of Tables

Table 2.1. The five key-site haplotype of the four LWS opsin loci with λ_{max} predicted following Yokoyama and Raddlewimmer (1998). The number of different LWS λ_{max} within species is summarized as '# LWS'. 35

Table 2.2. Selection tests on non-LWS opsins using models M0, M1, M2, M7 and M8 of the codeml program in PAML. Likelihood ratio tests (LRTs) were performed between models with selection and models with neural evolution to identify genes under positive selection. 36

Table 2.3. Sites identified as under positive selection using the M2 and M8 models in codeml of PAML, and the REL, FUBAR, and FEL programs of HYPHY. Grey highlight indicates the site is in the transmembrane domain. / indicates site was not identified to be under selection using that method. NA indicates run was not available. Numbers given for M2, M8, REL, and FUBAR are posterior probabilities while numbers for FEL are *P* values. 37

Table 3.1. Comparison of the two *P. formosa* alleles and parental species for each LWS locus and RH2-1. Percent similarity is given above the diagonal. Number of gaps is given below the diagonal. Dark grey boxes show the most similar parental sequence for each allele from *P. formosa*. 78

Table 3.2. Amino acid differences in LWS loci across species. Amino acid numbering follows that of Yokoyama (2000). Transmembrane designations indicate whether the amino acid site is located in the transmembrane domain (Y-yes; N-no), where changes in amino acids are more likely to shift spectral absorbances. '-' indicates there was no variation in amino acids at that site for that locus. The amount of within species variation is unknown. Shading is provided to facilitate visual comparison. 79

Table 3.3. Tukey's HSD comparing the hybrid *P. formosa* and parental species, *P. mexicana* and *P. latipinna*, in the number of pecks and approaches made toward coloured disks. Bold values indicate significance at $p < 0.05$. 80

Table 4.1. Results of MANOVA on opsin expression profiles. W- Watershed, T- Time, P(W)- Predation nested in watershed. 110

Table 4.2. Results of MANOVA on opsin expression profiles by watershed. 110

Table 4.3. Results of LWS ANOVA. W- Watershed, T- Time, P(W)- Predation nested in watershed. 111

Table 5.1. Results of MANOVA on opsin expression profiles as proportional and relative$_{(hk)}$ measures. *Loc*, Location; *Sp(Loc)*, Species nested in Location; *Sx(Sp(Loc))*, Sex nested in Species, nested in Location. Bold indicates $P < 0.05$. 149

Table 5.2. AIC values for the ability of models based on Location, Species, or Location*Species to explain variation in proportional measures of opsin expression for each opsin gene. Bold indicates $P < 0.05$. 149

Table 5.3. Results of opsin expression ANOVAs using proportional and relative$_{(hk)}$ measures by location. Bold indicates $P < 0.05$. 150

List of Figures

Figure 1.1. Male guppies (*Poecilia reticulata*) are highly variable in colour and number of spots..... 14

Figure 1.2. Males of the subgenus Lebistes are very colourful. While *Poecilia reticulata* males are highly variable in colouration, there are only two male morphs of *Poecilia picta* and five discrete male morphs of *Poecilia parae*..... 15

Figure 2.7. ND2 mitochondrial bayesian consensus tree with posterior probabilities. Outgroup is *Oryzias latipes*. 45

Figure 2.8. Bayesian consensus tree with posterior probabilities of SWS1. Outgroup is *Oryzias latipes*. 46

Figure 2.9. Bayesian consensus tree with posterior probabilities of SWS2A and SWS2B. Outgroup is *Oryzias latipes*..... 47

Figure 2.10. Bayesian consensus tree with posterior probabilities of RH1. Outgroup is *Oryzias latipes*. 48

Figure 2.11. Bayesian consensus tree with posterior probabilities of RH2-1. Outgroup is *Oryzias latipes*. 49

Figure 2.12. Bayesian consensus trees with posterior probabilities of (A) LWS UTR sequence, and (B) LWS full sequences. Outgroup is *Oryzias latipes*. Note the expected duplication history recovered in UTR sequence and general species relationships (based on ND2 tree) recovered within clades except for LWS-1 and LWS-3 of the full LWS tree. 50

Figure 2.13. (A) The genomic organization of the LWS opsin loci in *P. wingei* (adopted from Watson *et al.* (2011)). (B) The proposed conformation leading to high rates of gene conversion between LWS-1 and LWS-3 in the event of a double strand break. Arrows denote directionality. Coloured boxes denote exons while spaces between boxes denote introns. 51

Figure 3.1. Bayesian consensus phylogenies of (A) long wavelength sensitive (LWS) opsin loci, and (B) RH2-1 opsin locus in *Poecilia mexicana*, *P. latipinna*, and the two *P. formosa* alleles. Maximum likelihood bootstraps and Bayesian posterior probability values (respectively) are reported for each node..... 76

Figure 3.2. Violin plots (Hintze and Nelson 1998) showing the proportion of pecks and approaches made toward each coloured disk for the hybrid *P. formosa* and the parental species *P. mexicana*, and *P. latipinna*. Violin plots contain a box plot (white circles denote mean) surrounded by a grey kernel density plot which shows the distribution of the data followed by smoothing to facilitate visual comparison across plots. Letters denote groups that differ significantly (Tukey's HSD, p<0.05). 77

Figure 4.1. Long wavelength-sensitive (LWS) opsin expression by population for proportional measures. Stars denote significant differences between high and low predation within the same watershed (**P* < 0.05, ***P* <0.01, ****P* < 0.001); error bars are ±1 SE (see Table S8, Supporting information). Aripo low predation at 13:30 had *N* = 17; all other population time points had *N* = 18. 112

Figure 4.2. Long wavelength-sensitive (LWS) opsin expression by population for relative$_{(hk)}$ measures. Stars denote significant differences between high and low predation within the same watershed (**P* < 0.05, ***P* < 0.01); error bars are ±1 SE (see Table S9, Supporting information). Aripo low predation at 13:30 had *N* = 17; all other population time points had *N* = 18. 113

Figure 4.3. Proportional expression at time point 2. Stars denote significant differences within watershed between high- and low-predation populations (**P* < 0.05, ***P* < 0.01, ****P* < 0.001) (see Table S8, Supporting information). Letters denote significant differences between watersheds (see Table S6, Supporting information)..... 114

Figure 4.4. Genotype and gene frequencies of the 180 Ala (A) vs. 180 Ser (S) allele of LWS-1. LWS, long wavelength-sensitive. 115

Figure 5.1. Proportional measures of opsin expression across species and locations. Letters denote significant differences in expression of individual opsins across species within a location. Note: West Watuka has only *P. bifurca* and was not used in statistical analyses. 152

Figure 5.2. Relative$_{(hk)}$ measures of opsin expression across species and locations. Letters denote significant differences in expression of individual opsins across species within a location. Note: West Watuka has only *P. bifurca* and was not used in statistical analyses. 153

Figure 5.3. Genotype and gene frequencies of the 180 Ala (A) versus 180 Ser (S) allele of LWS-1. LWS, long-wavelength sensitive. 154

“Hooray for eyes! Hooray, hooray, hooray… for eyes!”

The Eye Book by Dr. Seuss

Image created by Todd Adamson, Integrated DNA Technologies, for a story featuring the published version of Chapter 5.

Chapter 1.

Introduction

Sexual selection is capable of playing a key role in the evolution and divergence of populations (Carleton *et al.* 2005; Seehausen *et al.* 2008), maintaining genetic polymorphisms within populations (Sinervo and Lively 1996; Hughes *et al.* 2013), and can even lead to speciation (Panhuis *et al.* 2001; Coyne and Orr 2004). A great many models have been put forth to explain the role of sexual selection on the evolution of a species (reviewed in Kuijper *et al.* 2012). The genes underlying differences in sexual selection are largely assumed to follow traditional 'mutation-selection' models of evolution (Jennions and Petrie 1997; Kirkpatrick and Hall 2004). However, even relatively small deviations from such models could result in significant effects on the direction, strength and role of sexual selection. When sexual selection takes the form of inter-sexual mate choice, both the characters preferred and the corresponding preferences are expected to co-evolve (Pomiankowski and Sheridan 1994; Breden *et al.* 1994; Andersson and Simmons 2006). As such, one can use differences in sexual characters across species as a proxy for corresponding differences in mate choice. For traits to play a role in mate choice they first must be detected by the assessor's relevant sensory system (Endler 1990; Horth 2007). Since sensory systems play multiple roles as organisms interact with their environment (such as foraging, navigating, predator avoidance, circadian regulation, etc.) they frequently adapt to local conditions through either evolution or plasticity (Loew and Lythgoe 1978; Fuller *et al.* 2004; Endler *et al.* 2005). Such variation in a crucial aspect of mate choice could offer a key factor influencing the strength, direction and role that sexual selection plays in the evolution of a species. It is therefore important to know the extent to which variation in peripheral sensory systems can explain differences in sexually selected traits.

In this thesis I use phylogenetic, laboratory and field based methods to investigate the variation and evolution of colour vision in the extensively studied family of freshwater, live-bearing fishes, the Poeciliidae. This family is one of the premier models for sexual selection as it includes such species as guppies (*Poecilia reticulata*), swordtails (*Xiphophorous helleri*), and sailfin mollies (*Poecilia latipinna*); since 1949, more than 1000 papers have been published on this group solely in the area of inter-sexual mate choice. Nearly all of the sexually selected traits described in this diverse group are transmitted visually and occur as either differences in morphology or colouration (Pollux *et al.* 2014). Poeciliid fishes have excellent colour vision, which plays a major role when making mating decisions based on colouration.

The Family Poeciliidae

The family Poeciliidae is comprised of over 250 species spanning 22-28 genera of live-bearing freshwater fishes (Stockwell and Henkanaththegedara 2011). Poeciliid fishes are native to areas from lower North America to the Amazon basin in South America but are now found world wide as a result of human introductions, frequently as pest control or aquarium releases (Stockwell and Henkanaththegedara 2011).

Sexual dimorphism is common across the family such that males are more ornamented than females, demonstrating the strong role that sexual selection has played in the evolution of such a diverse family (Farr 1989; Bisazza 1993). In several species males have exaggerated fins, and females base mating decisions on the size of the males' fin. This includes species of the genus *Xiphophorus* where females prefer to mate with males with elongated caudal fins (Basolo 1990a,b; Robinson *et al.* 2011) and *Poecilia latipinna* where females prefer males with enlarged dorsal fins (Ptacek and Travis 1997).

In other species sexual dimorphism occurs where males have bright colouration while females remain relatively drab grey or silver, in these species females make mating decisions based on male colouration. This pattern includes classic species such as guppies (*Poecilia reticulata*) where males vary substantially in the size and colour of spots (Figure 1.1) and females generally prefer more colourful males (Haskins and Haskins 1949; Haskins *et al.* 1961; Endler 1978; Breden and Stoner 1987; Magurran

2005). Several species that are closely related to guppies also make mating decisions based on male colouration; including *P. parae* where males occur in five discrete morphs (one large bodied with low colouration, three medium sized bodied with colour variation, and one small bodied female mimic) (Lindholm *et al.* 2004) (Figure 1.2). *Poecilia picta* is a sister species to *P. reticulata* and *P. parae*, which shows some male colouration but the role of colouration in mate choice is unknown (Breden and Bertrand 1999) (Figure 1.2). Surprisingly, *P. reticulata*, *P. parae*, and *P. picta* occur in extreme sympatry on mainland South America such that they are frequently found schooling together in small drainage ditches. Despite the close proximity and similar morphology they do not hybridize and cross species mating is rare (Liley 1965; Magurran and Ramnarine 2004; 2005). Such strong prezygotic isolation relies on correctly identifying and evaluating potential mates.

Mate Choice Evolution

Many models have been put forth in an effort to explain the evolution of mate preferences. Such models fall into the general categories of: direct benefits, indirect benefits, and sensory by-product (reviewed in Kuijper *et al.* 2012). Direct benefit models can be used to explain the evolution of female preferences in systems where males contribute more than just sperm - such as access to territory, protection from predators or parental investment (Kuijper *et al.* 2012). Direct benefit models are unlikely to explain the evolution of female preferences in the Poeciliidae family, as males are not known to make any direct contributions to females or offspring.

Indirect models can be further divided into three categories: 'Fisherian processes', 'good genes' and 'compatible genes'. 'Fisherian processes' (frequently referred to as 'sexy sons') follow the logic originally put forth by RA Fisher (Fisher 1915). This line of reason posits that by choosing attractive mates, females will have more attractive sons, which other females are more likely to choose as mates and thereby securing more grand-offspring. However, some researchers have questioned the ability of 'Fisherian processes' to act alone in driving the evolution of sexual selection as they are expected to break down when preferences are costly (Hall *et al.* 2000). Furthermore, while 'Fisherian processes' can explain increases in the strength of pre-existing female preference, they do little to explain the forces that led to the origin of such preferences.

‘Good genes’ and ‘compatible genes’ models of indirect benefits suggest that female preferences arise for traits that signal the genetic quality of a male. There has been some support for ‘good genes’ models to explain female preferences for male colour in Poeciliid fishes, particularly when the colours preferred by females cannot be synthesized by the male and must be obtained by foraging, such as with carotenoids (Grether 2000; Grether *et al.* 2005). However, ‘good genes’ models are generally thought to lead to a decrease in genetic diversity of a population or species (commonly referred to as the Lek paradox) (reviewed in Kotiaho *et al.* 2008) yet there is substantial genetic diversity in most Poeciliid species that have been examined (reviewed in Breden and Lindholm 2011) (although high diversity may be the result of genic capture at work (Rowe and Houle 1996)). The loss of genetic diversity is overcome in ‘compatible genes’ models, which suggest mate preferences could be the result of epistatic interactions, leading to variation in female mate preferences (Kuijper *et al.* 2012). However, neither of these models predict why specific traits arise.

‘Sensory by-product’ models suggest female preferences arise as a by-product of detection by the sensory system. For male signals to play a role in mate choice they must be displayed by an individual, transmitted through the environment, and detected by a female’s sensory system (Endler and Basolo 1998; Endler *et al.* 2005). While changes to male display traits and environmental transmission processes are both likely to impact female mating decisions, it is detection by the female’s sensory system that is the first stage of the process that is genetically inherited and is thereby capable of impacting the evolution of female mate preferences. There are many ‘sensory by-product’ models (primarily varying in the order or strength of selection), including: sensory traps, sensory drive, pre-existing bias, and sensory exploitation (reviewed in Basolo and Endler 1995; Endler and Basolo 1998). While many of these models were originally brought forth to explain the evolution of male display traits, the implications can be extended to explain the evolution of female mate preferences as well. The ‘sensory bias’ model has previously been implicated in the evolution of female preference for orange colouration in guppies (Rodd *et al.* 2002). ‘Sensory bias’ predicts that females have evolved a preference toward specific sensory stimuli outside the context of mate choice, and then males evolve to display matching stimuli (Basolo and Endler 1995; Endler and Basolo 1998). This model predicts that all populations should evolve the

same female mate preferences, yet there is considerable variation in mate preference for colouration across populations of guppies (Houde 1997; Brooks and Endler 2001). One alternative model that may better explain differences in mate preferences across populations is that of 'sensory exploitation'. Sensory exploitation predicts that males evolve sexual signals to maximize stimulation of females' sensory systems (Endler and Basolo 1998). Environments can vary considerably across populations of freshwater fishes in both quantity and quality of the available light spectrum (Lythgoe 1966; Loew and Lythgoe 1978; Levine and MacNichol 1979; Bowmaker 1995; Palen and Schindler 2010). Such variation would be expected to result in variation of the visual system across populations as organisms adapt to local conditions. Therefore the 'sensory exploitation' model is a prime candidate to explain the evolution of female mate preferences when there is variation both across species and across populations within species, such as seen in the family Poeciliidae. Many Poeciliid species base female mate preferences on male colouration, making variation in colour vision a prime candidate to test for the role of 'sensory exploitation' in the evolution of female mate preferences. Identifying variation in colour vision requires a detailed understanding of the mechanisms underlying colour vision.

Colour Vision

Colour vision is the discrimination of different wavelengths of light and is accomplished by the brain comparing signals from differentially tuned cone cells in the retina (Conway 2009). Cone cells detect light using transmembrane G protein-coupled receptors called opsins, which when bound to a chromophore are called visual pigments. In vertebrates opsins are either bound to an 11-*cis*-retinal (A1) or 11-*cis*-3,4-didehydroretinal (A2) chromophore (Terakita 2005). When a photon is absorbed there is a change in the conformation of the chromophore from 11-*cis* to all-*trans*, causing the visual pigment to couple with a G-protein and start a signal transduction cascade resulting in the perception of light. The wavelength of the photon to which a visual pigment is most sensitive is called its λ_{max}. Generally, each cone cell has only one visual pigment type, and it is the comparison of signals from cone cells with visual pigments that have different λ_{max} that allows colour vision. The λ_{max} of visual pigments are tuned by several methods across taxa including: selectively using either A1 or A2 chromophores

(e.g. crayfish (Zeiger and Goldsmith 1989)), differences in the amino acid sequence of the opsin protein (Yokoyama 2000; 2002), or differences in the expression of the opsin proteins (Hofmann and Carleton 2009). Of these methods for tuning visual receptors, differences in opsin sequence are by far the most common (Terakita 2005).

Opsin amino acid sequences are directly coded in the DNA and do not require post-translational modification (Yokoyama and Yokoyama 1996). This means the λ_{max} of a visual pigment made with a particular opsin can be directly estimated from the amino acid sequence of that opsin (hereafter referred to as the opsin's λ_{max}) (Yokoyama and Radlwimmer 1998). Therefore studies of opsin DNA sequence can be highly informative to investigations on the colour vision of a particular organism. This also means the evolution of colour vision is subject to genomic processes and constraints underlying the evolution of a multi-locus gene family.

Opsin genes are highly similar across taxa and are broken in classes depending on the range of light, and thereby colours, they detect: short wavelength sensitive 1 (SWS1) detect ultraviolet; short wavelength sensitive 2 (SWS2) detect blues and purples; rhodopsin like (RH2) detect greens; and long wavelength sensitive (LWS) detect reds and oranges. While humans with normal colour vision have only three different opsin proteins, Poeciliid fishes have an astounding nine cone opsin proteins: one SWS1, two SWS2 (SWS2A and SWS2B), two RH2 (RH2-1 and RH2-2), and four LWS (LWS-1, LWS-2, LWS-3, and LWS-R). Previously the LWS opsins have been named for the amino acid they possess at the location that corresponds to the 180th amino acid in humans (A180, S180, P180, and S180r) (Ward *et al.* 2008) but Chapter 3 proposes a new naming scheme for these genes, which has been taken up by the field (Tezuka *et al.* 2014). The LWS opsins are now named for the genomic location of their locus in the tandem array in which they are situated (for LWS-1, LWS-2 and LWS-3), while LWS-R is a retrotransposed gene (Sandkam *et al.* 2013). The expansion of the LWS opsin gene family in Poeciliid fishes, variation in sequence, and the role of red/orange in mate choice (detected by LWS opsins) make opsins a likely candidate in which to test for 'sensory exploitation' in the evolution of female mate preferences.

Objectives

My thesis is comprised of a series of studies examining the evolution of colour vision and mate choice in Poeciliid fishes. These studies used a combination of laboratory and field techniques to assess variation in opsin gene sequence and expression across species and populations. The goals of this body of work has been to examine: (1) the patterns underlying opsin evolution, (2) the relationship between opsin evolution and sexual selection, (3) the propensity for variation in the visual system to lead to differences in mate preferences across populations of the same species, and (4) whether visual systems vary more across species in the same location or across populations within species. Finally, whether phenotypic adaptation is primarily driven by changes in coding region or regulation of such sequences remains an ongoing controversy in evolutionary biology (King and Wilson 1975; Carroll 2005; Hoekstra and Coyne 2007). The experiments in this thesis directly estimate the relative contributions of these factors to variation in colour vision within and between species.

Chapter 2 explores the evolution of the opsin genes across the family Poeciliidae by sequencing and building phylogenies of fifteen species. Phylogenies built on LWS opsins revealed a high incidence of gene conversion – a phenomenon resulting in one locus overwriting another. Gene conversion of the LWS opsins is likely the result of their genomic architecture, which may be impacting the rate at which these genes are evolving. The potential role of gene conversion in shaping sexually selected characters is explored using phylogenetic techniques.

The evolution of colour vision is generally thought to rely on traditional models of mutation and selection where small, functional, changes occur by chance and accumulate gradually over many generations (Terakita 2005). In time, species are expected to evolve opsin genes that are both functional, and tuned to their local conditions. This pattern results in species having unique, functional opsin genes. Expanding the number of sensory genes is thought to rely on gene duplication, which then must go through the traditional slow process of divergence. Hybridization is a process that brings together genetic material from two species. Chapter 3 proposes and tests the Hybrid Sensory Expansion hypothesis. This hypothesis predicts that when two

species with distinct sensory repertoires hybridize, the resulting offspring will have a functional sensory repertoire that is larger than either of the parent species.

While wavelength sensitivity is largely tuned by opsin gene sequence, colour vision can also be modulated through differential expression of the opsin genes (Hofmann and Carleton 2009). Chapter 4 explores how colour vision differs across populations of guppies in independently colonized watersheds of Trinidad through opsin gene expression and allele frequency variation. Populations known to have stronger female preferences for more red/orange males were found to express higher levels of LWS opsins (responsible for detecting those colours) and differ in the frequency of an allele known to affect tuning of the LWS-1 opsin. These results support the sensory exploitation model for the evolution of female mate preferences and demonstrate how dramatically populations can differ in colour vision.

Populations divergence in mate preference through sensory bias models that results in speciation relies on variation in sensory systems within and across species. Chapter 5 evaluates whether there is more variation within or across closely related sympatric species on mainland South America. Surprisingly, opsin expression and allele frequencies of *P. reticulata*, *P. picta*, and *P. parae* differed as much or more across populations within species as they did between species within populations. These results suggest sensory exploitation may explain differences in mate choice across populations.

In general, this body of work furthers our understanding of how sensory system evolution can be promoted or constrained by the genomic architecture- a process frequently assumed to only follow traditional mutation selection models. This work also highlights how highly variable sensory systems can be across populations within species- a phenomenon largely unconsidered when exploring the visual ecology of an organism.

References

Andersson, M. B., and L. W. Simmons. 2006. Sexual selection and mate choice. *Trends Ecol Evol* 21:296–302.

Basolo, A. L. 1990a. Female preference for male sword length in the green swordtail, *Xiphophorus helleri* (Pisces, Poeciliidae). *Anim Behav* 40:332–338.

Basolo, A. L. 1990b. Female preference predates the evolution of the sword in swordtail fish. *Science* 250: 808-810.

Basolo, A. L., and J. A. Endler. 1995. Sensory biases and the evolution of sensory systems. *Trends Ecol Evol* 10:489–489.

Bisazza, A. 1993. Male competition, female mate choice and sexual size dimorphism in poeciliid fishes. *Marine Behaviour and Physiology* 23:257–286.

Bowmaker, J. K. 1995. The visual pigments of fish. *Progress in Retinal and Eye Research* 15:1–31.

Breden, F., and A. K. Lindholm. 2011 Genetic variation in natural populations. in *Ecology and Evolution of Poeciliid Fishes*. University of Chicago Press, Chicago IL.

Breden, F., and G. Stoner. 1987. Male predation risk determines female preference in the Trinidad guppy. *Nature* 329:831–833.

Breden, F., and M. Bertrand. 1999. A test for female attraction to male orange coloration in *Poecilia picta*. *Environ Biol Fish* 55:449–453.

Breden, F., H. C. Gerhardt, and R. K. Butlin. 1994. Female choice and genetic correlations. *Trends Ecol Evol* 9:343.

Brooks, R., and J. A. Endler. 2001. Female guppies agree to differ: phenotypic and genetic variation in mate-choice behavior and the consequences for sexual selection. *Evolution* 55:1644–1655.

Carleton, K. L., J. W. L. Parry, J. K. Bowmaker, D. M. Hunt, and O. Seehausen. 2005. Colour vision and speciation in Lake Victoria cichlids of the genus Pundamilia. *Mol Ecol* 14:4341–4353.

Carroll, S. B. 2005. Evolution at two levels: on genes and form. *Plos Biol* 3:1159–1166.

Conway, B. R. 2009. Color vision, cones, and color-coding in the cortex. *The Neuroscientist* 15:274–290.

Coyne, J. A., and H. A. Orr. 2004. *Speciation*. Sinauer Associates, Sunderland, MA.

Endler, J. A. 1978. A predator's view of animal color patterns. Pp. 319–364 in M. K. Hecht, W. C. Steere, and B. Wallace, eds. *Evolutionary Biology*. Plenum Press, Boston, MA.

Endler, J. A. 1990. On the measurement and classification of colour in studies of animal colour patterns. *Biological Journal of the Linnean Society* 41:315–352.

Endler, J. A., and A. L. Basolo. 1998. Sensory ecology, receiver biases and sexual selection. *Trends Ecol Evol* 13:415–420.

Endler, J. A., D. A. Westcott, J. R. Madden, and T. Robson. 2005. Animal visual systems and the evolution of color patterns: Sensory processing illuminates signal evolution. *Evolution* 59:1795–1818.

Farr, J. A. 1989. Sexual Selection and Secondary Sexual Differentiation in Poeciliids: Determinants of Male Mating Success and the Evolution of Female Choice in G. K. Meffe and F. F. J. Snelson, eds. *Ecology & Evolution of Livebearing Fishes*. Prentice Hall, New Jersey.

Fisher, R. A. 1915. The evolution of sexual preference. *Eugen Rev* 7:184–192.

Fuller, R. C., K. L. Carleton, J. M. Fadool, T. C. Spady, and J. Travis. 2004. Population variation in opsin expression in the bluefin killifish, *Lucania goodei*: a real-time PCR study. *J Comp Phys A* 190:147–154.

Grether, G. F. 2000. Carotenoid limitation and mate preference evolution: a test of the indicator hypothesis in guppies (*Poecilia reticulata*). *Evolution* 54:1712–1724.

Grether, G. F., G. R. Kolluru, F. H. Rodd, J. de la Cerda, and K. Shimazaki. 2005. Carotenoid availability affects the development of a colour-based mate preference and the sensory bias to which it is genetically linked. *Proc. R. Soc. B* 272:2181–2188.

Hall, D. W., M. Kirkpatrick, and B. West. 2000. Runaway sexual selection when female preferences are directly selected. *Evolution* 54:1862–1869.

Haskins, C. P., and E. F. Haskins. 1949. The role of sexual selection as an isolating mechanism in three species of poeciliid fishes. *Evolution* 3:160–169.

Haskins, C. P., E. F. Haskins, J. J. A. McLaughlin, and R. E. Hewitt. 1961. Polymorphism and population structure in *Lebistes reticulatus*, an ecological study in W. F. Blair, ed. Vertebrate Speciation.

Hoekstra, H. E., and J. A. Coyne. 2007. The locus of evolution: evo devo and the genetics of adaptation. *Evolution* 61:995–1016.

Hofmann, C. M., and K. L. Carleton. 2009. Gene duplication and differential gene expression play an important role in the diversification of visual pigments in fish. *Integr Comp Biol* 49:630–643.

Horth, L. 2007. Sensory genes and mate choice: Evidence that duplications, mutations, and adaptive evolution alter variation in mating cue genes and their receptors. *Genomics* 90:159–175.

Houde, A. E. 1997. *Sex, Color, and Mate Choice in Guppies*. Princeton University Press, Princeton NJ.

Hughes, K. A., A. E. Houde, A. C. Price, and F. H. Rodd. 2013. Mating advantage for rare males in wild guppy populations. *Nature* 503:108–110.

Jennions, M. D., and M. Petrie. 1997. Variation in mate choice and mating preferences: A review of causes and consequences. Biol Rev 72:283–327.

King, M. C., and A. C. Wilson. 1975. Evolution at two levels in humans and chimpanzees. *Science* 188:107–116.

Kirkpatrick, M., and D. W. Hall. 2004. Sexual selection and sex linkage. *Evolution* 58:683–691.

Kotiaho, J. S., N. R. LeBas, M. Puurtinen, and J. L. Tomkins. 2008. On the resolution of the lek paradox. *Trends Ecol Evol* 23:1–3.

Kuijper, B., I. Pen, and F. J. Weissing. 2012. A Guide to Sexual Selection Theory. *Annu Rev Ecol Evol S* 43:287–311.

Levine, J. S., and E. F. MacNichol. 1979. Visual pigments in teleost fishes: Effects of habitat, microhabitat, and behavior on visual system evolution. *Sensory Processes* 3: 95-131.

Liley, N. R. 1965. Ethological isolating mechanisms in four sympatric species of poeciliid fishes. *Behaviour. Supplement* 13:1–197.

Lindholm, A. K., R. Brooks, and F. Breden. 2004. Extreme polymorphism in a Y-linked sexually selected trait. *Heredity* 92:156–162.

Loew, E. R., and J. N. Lythgoe. 1978. The ecology of cone pigments in teleost fishes. *Vision Res* 18: 715-722.

Lythgoe, J. N. 1966. Visual pigments and underwater vision. Pp. 375–391 in *Light as an Ecological Factor*. Blackwell Scientific Publishers. Oxford, UK.

Magurran, A. E. 2005. *Evolutionary Ecology: The Trinidadian Guppy*. Oxford University Press, New York, NY.

Magurran, A. E., and I. Ramnarine. 2005. Evolution of mate discrimination in a fish. *Curr Biol* 15:R867.

Magurran, A. E., and I. W. Ramnarine. 2004. Learned mate recognition and reproductive isolation in guppies. *Anim Behav* 67:1077–1082.

Palen, W. J., and D. E. Schindler. 2010. Water clarity, maternal behavior, and physiology combine to eliminate UV radiation risk to amphibians in a montane landscape. *Proc Natl Acad Sci U S A* 107:9701–9706.

Panhuis, T. M., R. Butlin, M. Zuk, and T. Tregenza. 2001. Sexual selection and speciation. *Trends Ecol Evol* 16:364–371.

Pollux, B. J. A., R. W. Meredith, M. S. Springer, T. Garland, and D. N. Reznick. 2014. The evolution of the placenta drives a shift in sexual selection in livebearing fish. *Nature* 513:233–236.

Pomiankowski, A., and L. Sheridan. 1994. Linked sexiness and choosiness. *Trends Ecol Evol* 9:242–244.

Ptacek, M. B., and J. Travis. 1997. Mate choice in the sailfin molly, *Poecilia latipinna*. *Evolution* 51:1217–1231.

Robinson, D. M., M. S. Tudor, and M. R. Morris. 2011. Female preference and the evolution of an exaggerated male ornament: the shape of the preference function matters. *Anim Behav* 81:1015–1021.

Rodd, F. H., K. A. Hughes, G. F. Grether, and C. T. Baril. 2002. A possible non-sexual origin of mate preference: are male guppies mimicking fruit? *Proc. R. Soc. B* 269:475–481.

Rowe, L., and D. Houle. 1996. The lek paradox and the capture of genetic variance by condition dependent traits. *Proc. R. Soc. B* 263:1415–1421.

Sandkam, B. A., J. B. Joy, C. T. Watson, P. Gonzalez-Bendiksen, C. R. Gabor, and F. Breden. 2013. Hybridization leads to sensory repertoire expansion in a gynogenetic fish, the Amazon molly (*Poecilia formosa*): a test of the hybrid-sensory expansion hypothesis. *Evolution* 67:120–130.

Seehausen, O., Y. Terai, I. S. Magalhaes, K. L. Carleton, H. D. J. Mrosso, R. Miyagi, I. van der Sluijs, M. V. Schneider, M. E. Maan, H. Tachida, H. Imai, and N. Okada. 2008. Speciation through sensory drive in cichlid fish. *Nature* 455:620–626.

Sinervo, B., and C. M. Lively. 1996. The rock-paper-scissors game and the evolution of alternative male strategies. *Nature* 380:240–243.

Stockwell, C. A., and S. M. Henkanaththegedara. 2011. Evolutionary conservation biology. Pp. 128–142 in *Ecology and Evolution of Poeciliid Fishes*. University of Chicago Press, Chicago, IL.

Terakita, A. 2005. The opsins. *Genome Biol* 6:213.

Tezuka, A., S. Kasagi, C. Van Oosterhout, M. McMullan, W. M. Iwasaki, D. Kasai, M. Yamamichi, H. Innan, S. Kawamura, and M. Kawata. 2014. Divergent selection for opsin gene variation in guppy (*Poecilia reticulata*) populations of Trinidad and Tobago. *Heredity* 113:381–389.

Ward, M. N., A. M. Churcher, K. J. Dick, C. R. Laver, G. L. Owens, M. D. Polack, P. R. Ward, F. Breden, and J. S. Taylor. 2008. The molecular basis of color vision in colorful fish: Four Long Wave-Sensitive (LWS) opsins in guppies (*Poecilia reticulata*) are defined by amino acid substitutions at key functional sites. *BMC Evol Biol* 8.

Yokoyama, S. 2002. Molecular evolution of color vision in vertebrates. *Gene* 300:69–78.

Yokoyama, S. 2000. Molecular evolution of vertebrate visual pigments. *Progress in Retinal and Eye Research* 19:385–420.

Yokoyama, S., and F. B. Radlwimmer. 1998. The “five-sites” rule and the evolution of red and green color vision in mammals. *Mol Biol Evol* 15:560–567.

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

Zeiger, J., and T. H. Goldsmith. 1989. Spectral properties of porphyropsin from an invertebrate. *Vision Res* 29:519–527.

Figure 1.1. **Male guppies (*Poecilia reticulata*) are highly variable in colour and number of spots.**

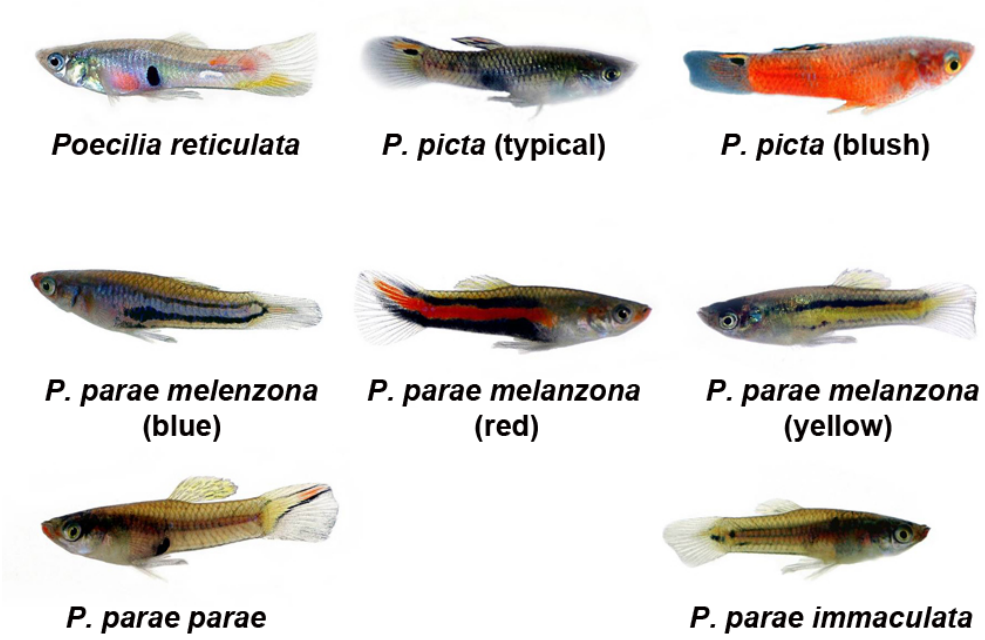


Figure 1.2. **Males of the subgenus Lebistes are very colourful. While *Poecilia reticulata* males are highly variable in colouration, there are only two male morphs of *Poecilia picta* and five discrete male morphs of *Poecilia parae*.**

Chapter 2.

The impact of genomic environment on the evolution of colour vision in the family Poeciliidae, a model for visually based sexual selection

Publication and Contributions

A version of this chapter is currently being prepared for submission: Sandkam, BA, JB Joy, CT Watson, and F Breden. The impact of genomic environment on the evolution of colour vision in the family Poeciliidae, a model for visually based sexual selection.

Contributions: BAS and CTW designed primers. BAS performed sequencing and ran site selection analyses. BAS and JBJ made alignments and generated phylogenies. BAS, JBJ, CTW and FB designed the project and wrote the manuscript.

2.1. Abstract

Sexual selection is predicted to lead to a co-evolution of sensory systems and mate preferences, but sensory system evolution can be constrained by genomic architecture. Here we examine the evolutionary interplay between sexual selection and the opsin genes (responsible for tuning colour vision) of 15 species across the family Poeciliidae, which includes such classic systems for studies of sexual selection as guppies, swordtails and mollies. Male colouration patterns vary widely within and among genera, yet only one of the known amino acid sites that tune Long Wavelength Opsins was found to be variable in this specious family that is characterized by colour-mediated sexual selection. While most opsin genes appear to evolve without genomic constraint, we identified high rates of gene conversion between two of the LWS loci (LWS-1 and

LWS-3), likely due to the genomic architecture of the LWS genes. The LWS opsins are responsible for detecting and discriminating red and orange colouration- a key sexually selected trait in members of the subgenus Lebistes. This is the only subgenus with differences in wavelength sensitivity between LWS-1 and LWS-3. Taken together these results suggest selection is acting against the homogenizing effects of gene conversion to maintain LWS diversity. Additionally, while we found mixed results for selection acting on opsin tuning, across the family there was strong selection acting on the rhodopsin gene, which acts in scotopic rather than photopic vision.

2.2. Introduction

Behaviours and decision-making processes are understood to be partially determined by the genomic repertoire available to an organism (Robinson *et al.* 2008; Renn *et al.* 2008). However, gene repertoire evolution is subject to its own set of processes and constraints (Lynch 2007). Therefore it is of great importance for studies investigating the evolution of behaviours to explore the associated changes in the genomic repertoire impacting the behaviours in question (Wilkinson *et al.* 2015).

Sexual selection is a suite of behaviours that has received a great deal of attention with respect to evolution in terms of both origins and changes (Andersson 1994; Panhuis *et al.* 2001; Andersson and Simmons 2006). Sexual selection is capable of driving large-scale evolutionary changes through mechanisms such as behavioural isolation (Ryan and Rand 1993) and the elaboration of sexually selected traits (Andersson 1994). Several evolutionary models predict sensory systems responsible for detecting traits under sexual selection and mate preferences are tightly linked, including Sensory Bias (Basolo and Endler 1995; Endler and Basolo 1998; Boughman 2002). Indeed, the genes underlying relevant sensory systems have been shown to strongly influence the strength and direction of trait evolution (reviewed in Horth 2007). It is well known that gene duplication is the driving force underlying the expansion of sensory repertoires, specifically those repertoires that are then co-opted for use in mate choice (Horth 2007). Gene duplication can occur through multiple mechanisms, resulting in differences in the genomic architecture, which can impose constraints and selection upon gene evolution (Zhang 2003).

The family Poeciliidae is a group of freshwater fishes that provides an excellent model for studies of sexual selection as it includes such classic model systems in sexual selection as; guppies (*Poecilia reticulata*), swordtails (*Xiphophorous helleri*), and sailfin mollies (*Poecilia latipinna*). Whereas sexual selection via female mate choice is common within this family, the dominant criteria females are selecting differ across species, for example: guppies base mating decisions on male colouration (Houde 1997; Magurran 2005), swordtails base mating decisions on the length of the male caudal fin (Basolo 1990; Robinson *et al.* 2011), and sailfin mollies base mating decisions on the size of the male dorsal fin (Ptacek and Travis 1997).

While nearly all of the sexually selected traits described in this diverse group are transmitted visually, the role of colouration differs strikingly across species. Colour vision (the discrimination of different wavelengths of light) is accomplished by having cone cells in the retina that are maximally sensitive to different wavelengths of light. The brain interprets the strength of signals from the different types of cone cells as colour (Conway 2009), thereby having a greater repertoire of cone cells facilitates better colour discrimination (Gegenfurtner & Sharpe 1999). The most influential factor determining the wavelength to which a cone cell is maximally sensitive (λ_{max}) is the transmembrane protein expressed by that cell, called an opsin (Yokoyama 2002). The λ_{max} of an opsin protein is directly determined by its amino acid sequence, hence the number of functionally different opsins possessed by an organism can be directly determined by the genomic repertoire and selection analyses on opsin genes can reveal insights to visual tuning (Yokoyama and Yokoyama 1996). Guppies and swordtails have recently been shown to have an expanded repertoire of opsin genes compared to other fish families resulting in a total of nine cone opsin genes, among the largest known opsin repertoire of any vertebrate (Ward *et al.* 2008; Watson *et al.* 2010; 2011). This expanded repertoire is especially pronounced in the long wavelength-sensitive (LWS) class of opsins, which detect wavelengths in the red, yellow, and orange end of the visible light spectrum (Yokoyama *et al.* 2008). This expanded LWS opsin repertoire has been proposed to be involved in the origin and prevalence of sexual selection in this group (Hoffmann *et al.* 2007; Ward *et al.* 2008; Watson *et al.* 2011), but has not been tested in a phylogenetic framework. A phylogenetic approach requires a detailed understanding of the evolution of colour vision in this family.

To investigate the evolution of colour vision in the family Poeciliidae we sequenced eight of the nine cone opsins as well as the rod opsin (rhodopsin) and built gene phylogenies for 15 species across the family. We proposed a targeted hypotheses regarding LWS opsin evolution and mate choice. Specifically, if differences across species in mate preference are driven by differences in the sensory system, then we would expect to see differences in the tuning of LWS opsins in the species for which red/orange plays a predominant role in mate choice. Additionally, we describe patterns of evolution in four of the five non-LWS cone opsins (SWS1, SWS2A, SWS2B, RH2-1) as well as rhodopsin (RH1). Since the lighting environment varies considerably across these species we tested for sites under selection that could describe differences in visual tuning. While most opsin genes appear to be under neutral evolution, we identified high rates of gene conversion between two of the LWS loci (LWS-1 and LWS-3) that occurs throughout the family Poeciliidae. We propose that the evolution of the LWS opsin genes may be influenced by an interplay between gene conversion due to genomic architecture and the role of red/orange colouration in sexual selection.

2.3. Materials and Methods

2.3.1. Sequencing

Using DNeasy blood and tissue kits (QIAGEN), DNA was extracted from tissue samples of single specimens of the following species: *Heterandria formosa*, *Xiphophorus helleri*, *Poecilia caymanensis*, *P. vittata*, *P. nigrofasciata*, *P. latipinna*, *P. velifera*, *P. petenensis*, *P. mexicana*, *P. minor*, *P. reticulata*, *P. wingei*, *P. bifurca*, *P. picta*, and *P. parae*. Primers specific to 5' and 3' UTR regions of each of the four LWS loci and the two SWS2 loci were designed using genomic data from *Poecilia wingei* (Watson *et al.* 2011) and *Xiphophorous helleri* (Watson *et al.* 2010). Primers specific to 5' and 3' UTR of RH2-1 were taken from Sandkam *et al.* (2013). SWS1 and RH1 UTR primers were designed using genomic data from whole genome shotgun sequence of *Xiphophorus maculatus* (GenBank accessions: AGAJ01036758.1 and AGAJ01019341.1 respectively). PCR primers were designed to be specific to the UTR regions upstream and downstream of each locus due to high exon sequence similarity in the LWS opsins; in such cases it has

been shown that UTR primers are critical to ensure locus specificity of the PCR products (Watson *et al.* 2011) (for primer sequences see Supplementary Table 2.1). We follow Sandkam *et al.* (2013) and refer to LWS loci by their location in the genome relative to one another. Due to the length of LWS-2, PCR products for this locus were generated in two overlapping segments. Locus specificity was maintained for each of these segments by PCR amplifying with one UTR primer and one internal primer.

All sequencing was performed by Molecular Cloning Laboratories (MCLAB; San Francisco, CA, USA). Sequence chromatograms were viewed and analyzed using SeqMan Pro (Lasergene 8.0; DNASTAR). All sequences are available on GenBank (accession numbers pending submission_).

A firm understanding of the phylogenetic relationships of the species used in this study was obtained from ND2 mitochondrial sequences retrieved from GenBank (see Supplementary Table 2.2 for accession numbers). LWS and SWS2 sequences for *P. wingei* were taken from Watson *et al.* (2011) (GenBank Accession: HM540108 and HM540107) and *X. helleri* from Watson *et al.* (2010) (GenBank Accession: GQ999832 and GQ999833). The LWS and RH2-1 sequences of *P. mexicana* and *P. latipinna* were taken from Sandkam *et al.* (2013) (GenBank Accessions: JF823552 – JF823560). PCR products could not be reliably made for the RH2-2 locus in the majority of the species so this opsin was left out of all analyses.

2.3.2. Tree Building

The divergence of the opsin classes occurred before the rise of the family Poeciliidae (Rennison *et al.* 2012), making alignment of introns and UTRs difficult, however, opsins are highly conserved structurally, allowing alignments of coding sequence to be made across classes. To best explore the evolutionary history of the opsins in the family Poeciliidae we built a series of trees based on several sets of sequences: (1) mitochondrial ND2, (2) all opsins: exon sequence only, (3) SWS1: full sequences (UTR, introns and exons), (4) SWS2: full sequences, (5) RH1: full sequences, (6) RH2-1: full sequences, and (7) LWS: full sequences. Upon seeing high

rates of gene conversion in the LWS opsins we built one additional tree: (8) LWS: UTR-only sequence.

For each set, sequences were aligned using Mafft v6.833b (Katoh *et al.* 2009) and edited manually using Se-Al v2.0a11 (Rambaut 1996) to ensure that intron-exon boundaries (where applicable) were consistent across all species. Best-fit models of molecular evolution were determined using MrModelTest 3.04 (Nylander 2004). Phylogenetic trees were reconstructed under Maximum Likelihood (ML) using PAUP* 4.0b10 (Swofford 2003) and Bayesian methods as implemented in MrBayes 3.1.2 (Ronquist and Huelsenbeck 2003). Two runs utilizing four Markov chains (three heated and one cold) were run for 10^7 generations, with trees sampled every 1000 generations. Convergence was assessed using the standard deviation of the split frequencies between runs, and graphically using the program Tracer (Rambaut and Drummond 2007) and AWTY (Nylander *et al.* 2008). ML bootstrap values and Bayesian posterior probabilities were employed to assess support.

2.3.3. Gene Conversion Analyses in LWS Opsins

The LWS duplication history is known to precede the emergence of the family Poeciliidae (Rennison *et al.* 2012). Yet when building a phylogeny of the LWS loci, LWS-1 and LWS-3 cluster within species and clades, indicative of high frequencies of gene conversion between these loci (see discussion for further justification) (Chen *et al.* 2007). To understand how gene conversion differed across the family Poeciliidae we identified gene conversion breakpoints using GARD (Kosakovsky Pond *et al.* 2006), a part of the HyPhy software package (Kosakovsky Pond *et al.* 2005), on the DataMonkey Server (Kosakovsky Pond and Frost 2005a). We ran GARD by subgenus: Lebistes (*Poecilia reticulata*, *P. parae*, *P. picta*, and *P. bifurca*), Poecilia (*P. latipinna*, *P. velifera*, *P. mexicana*, and *P. pentenensis*), and Limia (*P. caymenensis*, *P. vittata*, and *P. nigrofasciata*). Running GARD on subgenus rather than our full set of species allowed us to focus on how gene conversion has acted within these clades. The high incidence of gene conversion in the LWS opsins invalidates the assumptions required for traditional selection level analyses.

2.3.4. Identifying Sites Under Selection in Non-LWS Genes

GARD analyses found no evidence of recombination in any of the non-LWS gene classes, thereby meeting the assumption of no recombination required for selection analyses. To explore selection in the SWS1, SWS2A, SWS2B, RH2-1, and RH1 opsin classes; positively selected sites were identified using the random sites models (M0, M1, M2, M7, M8) in the codeml program from the PAML 4.0 software package (Yang 2007) run by the pamlX GUI (Xu and Yang 2013). The ND2 mitochondrial tree was used without branch lengths as the known species relationships for each of the opsin classes run through codeml. The highly robust Bayes' Empirical Bayes (BEB) analysis was used to identify sites under positive selection in the M2 and M8 models. Due to unknown reasons, we were unable to get the PAML analyses to run on the RH2-1 data set.

To increase confidence in the sites found using PAML, we also identified sites under selection using the REL, FEL, and FUBAR models (Kosakovsky Pond *et al.* 2006; Murrell *et al.* 2013) in the HYPHY software package on the Datamonkey server (Kosakovsky Pond and Frost 2005a).

2.4. Results

2.4.1. Data Set and Alignment

All sequences were generated and aligned successfully. Within each locus the size of the opsin genes did not differ more than 50 base pairs (bp) across species with the exception of intron 1 in LWS-2; this intron varied between 1808 bp in *P. minor* to 2410 bp in *P. reticulata*. The differences in length of intron 1 of LWS-2 are caused by expansions of micro-satellites in some species (Watson *et al.* 2011). For each opsin class, amino acid alignments are presented in Figures 2.1-2.6 following the style of (Hofmann *et al.* 2012), such that annotations are provided denoting the sites that correspond to known tuning sites, the transmembrane region, and the retinal binding pocket of bovine rhodopsin.

2.4.2. Phylogenetic Analyses

Phylogenetic trees inferred using ML (PAUP*) and Bayesian (MrBayes) methods converged upon similar topologies for all eight datasets. Support for recovered nodes was robust for all topologies inferred under both ML and Bayesian methods with no well-supported nodes (posteriori value > 0.8) differing among analyses. The ND2 mitochondrial tree provides expected species relationships within each opsin and is presented in Figure 2.7. Due to the large size of the phylogeny built on the exons of all opsins, it is presented as Supplementary Figure 2.1. The phylogenies of SWS1, SWS2, RH1, and RH2-1 are presented in Figures 2.8, 2.9, 2.10, and 2.11 respectively. The full LWS and the LWS UTR only trees are presented in Figure 2.12. In the phylogeny built on LWS-UTR sequences, each opsin locus (LWS-1, LWS-2, LWS-3 and LWS-R) formed a clear clade and generally followed the phylogenetic relationships between species in the ND2 mitochondrial tree. In the LWS-full sequences tree, LWS-2 and LWS-R formed clades with relationships between species generally matching those identified in the ND2 and LWS-UTR phylogenies. However, LWS-1 and LWS-3 clustered together (within species and clades) more frequently than expected based on the ND2 and LWS-UTR phylogenies. The frequent clustering of LWS-1 and LWS-3 is indicative of high rates of gene conversion between these two loci throughout the family Poeciliidae.

2.4.3. Gene Conversion Analyses

In the subgenus Lebistes, GARD identified 5 significant breakpoints ($p<0.05$) and 1 possible breakpoint ($p<0.10$). In the subgenus Limia, GARD identified 2 significant breakpoints ($p<0.05$). In the subgenus Poecilia, GARD identified 2 significant breakpoints ($p<0.05$) and 2 possible breakpoints ($p<0.10$). The number of breakpoints identified in the different subgenera could indicate that the sections of LWS-1 and LWS-3 loci undergoing gene conversion in Lebistes are shorter than the sections of LWS-1 and LWS-3 loci undergoing gene conversion in Limia or Poecilia.

2.4.4. Identifying Sites Under Selection in Non-LWS Genes

The different analyses (M2, M8, REL, FUBAR, and FEL) identified some of the same and some different sites under selection. Likelihood ratio tests between models

M2-M1 and M8-M7 in codeml of PAML revealed only RH1 to be under strong positive selection (Table 2.2) yet several sites identified to be under selection by M2 or M8 were also found to be under selection using either REL, FUBAR or FEL programs in the HyPhy package (Table 3). Sites identified using Bayes Empirical Bayes (BEB) analysis of M8 in PAML (P > 95%) revealed: 2 sites in SWS1, 1 site in SWS2B, and 5 sites in RH1 with an additional 9 sites at P > 0.99%. All of these sites concurred with at least one other analysis (M2, REL, FUBAR or FEL) and have been marked on the presented amino acid alignments (Figures 2.1–2.6). The REL, FUBAR and FEL analyses found no positively selected sites in either SWS1 or SWS2B.

2.5. Discussion

We sequenced and built phylogenies of rhodopsin and eight of the nine cone-opsin genes across 15 species in the family Poeciliidae. The lighting environments of the species examined here vary considerably both across and within species. Therefore we expected to see divergent visual systems but conserved species relationships. While expected species relationships were recovered for most of the opsins, we found the LWS-1 and LWS-3 loci cluster within species and clade, which would not be expected from their known duplication history. The unanticipated clustering of LWS-1 and LWS-3 is likely due to the process of gene conversion. Gene conversion is expected to homogenize gene sequence between genes, which would be expected in opsins to produce products with the same predicted tuning optimum. This has occurred for all species in this group except those in the subgenus Lebistes. Gene conversion is a process that results in one region of DNA “overwriting” another region of DNA; this makes the two regions identical and can look like independent duplications on a phylogeny (Chen *et al.* 2007). The maintenance of LWS-1 / LWS-3 diversity in the species in the subgenus Lebistes may be selected for because this facilitates the role of red/orange male colouration in mate choice throughout this clade. Below we outline our rationale for our conclusion of gene conversion between LWS-1 and LWS-3, discuss how the genomic architecture is likely promoting such gene conversion, and hypothesize how such patterns of conversion and life history could impact the evolution of these

genes. We go on to identify sites under selection in the non-LWS genes and hypothesize on the interplay between these changes and life history.

2.5.1. Gene Conversion, Not Independent Duplications

The genomic organization of the LWS opsin genes is shared between *Xiphophorus helleri* and *Poecilia reticulata* (Watson *et al.* 2010; 2011), suggesting that the duplication history of the opsin genes precedes the divergence of Poeciliidae. Rennison *et al.* (2012) also supported this with a wide range of species in the family. The work presented here further supports a shared duplication history across the family Poeciliidae, because primers were designed in the untranslated regions (UTR) upstream and downstream of the opsin genes to generate sequencing products. Having primers anneal in the UTR makes them specific to a location in the genome. Had there been independent duplications throughout the family Poeciliidae, we would expect the LWS loci to occur in different locations in the genome. The fact that UTR primers worked across all species suggests that all LWS loci occur at the same place and therefore share a duplication history that is older than the family Poeciliidae. Furthermore, the LWS phylogeny built on only UTR sequence recovered clear LWS-1, LWS-2, LWS-3, and LWS-R clades. Had the conserved UTR sequence been a result of independent duplications that included the UTR regions, we would expect to see LWS-1 and LWS-3 clustering within species (as observed in the phylogeny built on full LWS-1 and LWS-3 sequences).

The recovered full sequence LWS opsin gene phylogeny (Figure 2.12) revealed LWS-1 and LWS-3 cluster within species and clades more frequently than expected from the true duplication history, indicative of gene conversion.

2.5.2. Why is Gene Conversion so Prevalent Between LWS-1 and LWS-3?

Gene conversion frequently occurs following a double strand break (DSB). To repair the DNA the broken ends find a template strand that matches intact sequence adjacent to the break, and the ends are extended as complementary base-pairs to the template strand. The likelihood of using an incorrect template strand is a function of

proximity and similarity (Chen *et al.* 2007). LWS-1, LWS-2, and LWS-3 occur in a tandem array with less than 6 kb between LWS-1 and LWS-2, while there is less than 4 kb between LWS-2 and LWS-3 (Watson *et al.* 2010; 2011). The close proximity of LWS-1, LWS-2, and LWS-3 greatly increases the likelihood of gene conversion occurring between these loci, yet gene conversion only seems to occur between LWS-1 and LWS-3. The sequence length of LWS-2 is likely what decreases the chances of gene conversion involving this locus. While the introns and exons of LWS-1 and LWS-3 are nearly identical in length, LWS-2 has a greatly expanded intron 1 in all species throughout the family Poeciliidae. The difference in size of intron 1 between LWS-2 and LWS-1 / LWS-3 makes it less likely that LWS-1 or LWS-3 will use LWS-2 as a template in the event of a DSB due to the dramatic size differences that make alignment difficult. However, the close proximity of LWS-1 and LWS-3 in addition to the fact that all of their introns and exons match in length make it likely they will experience gene conversion with one another in the event of a DSB.

The orientation of LWS-1 and LWS-3 make it even more likely these two genes will experience gene conversion because they are inverted to one another. When a DSB occurs and the DNA doubles back on itself LWS-1 and LWS-3 are in the same orientation (Figure 2.13). The close proximity, identical sequence length and inverted orientation of LWS-1 relative to LWS-3 make these genes prime candidates for gene conversion and explain the high prevalence of such a process across the family Poeciliidae.

2.5.3. Implications of Gene Conversion to Colour Vision

The λ_{max} of the LWS opsins in vertebrates is predominantly determined by the amino acids at five key sites (Yokoyama and Radlwimmer 1998; Yokoyama *et al.* 2008). We found low allelic diversity across the family Poeciliidae at the key sites tuning the LWS loci (Table 2.1). Across the 15 species sequenced: LWS-1 has two alleles, LWS-2 has three alleles, LWS-3 has only one allele, and LWS-R has two alleles. The five key tuning sites correspond to the 180, 197, 277, 285 and 308 amino acid in the human opsin sequence (Yokoyama and Radlwimmer 1998) and were SHYTA for both LWS-1 and LWS-3 for all species except *P. wingei*, *P. bifurca*, and *P. picta* (which were all

AHYTA at LWS-1 and SHYTA at LWS-3). These three species, together with *P. reticulata* and *P. parae*, make up the subgenus Lebistes- a group well known to possess high levels of red/orange male colouration that are involved in female mate decisions (Liley 1965; Houde 1997; Lindholm *et al.* 2004). Interestingly, while the *P. reticulata* and *P. parae* individuals sequenced for this study had an S at the 180 site, both of these species have recently been shown to possess an allele with an A at the 180 site and vary in the frequency of this allele across populations (*P. reticulata*- (Tezuka *et al.* 2014; Sandkam *et al.* 2015); *P. parae*- Sandkam *et al.* in review). In guppies, the frequency of the A and the S has been observed to vary in a predictable manner between upstream and downstream populations (Sandkam *et al.* 2015). The most highly expressed LWS in these three species is LWS-1 (Sandkam *et al.* 2015 and Sandkam *et al.* in review), therefore the change from an S to an A at the 180 amino acid likely has important tuning implications in these species.

The presence of an A at the 180 site in LWS-1 relative to the fixed S at LWS-3 has previously been postulated to play a role in mate choice in the subgenus to which these species belong; Lebistes (*P. bifurca*, *P. parae*, *P. picta*, *P. reticulata*, and *P. wingei*) (Ward *et al.* 2008). Our data may offer some support to this hypothesis. When looking for breakpoints of conversion events between LWS-1 and LWS-3 within clades, GARD analyses reveal more breakpoints in the subgenus Lebistes than either the subgenus Poecilia (*P. latipinna*, *P. mexicana*, *P. petenensis*, and *P. velifera*) or the subgenus Limia (*P. caymanensis*, *P. nigrofasciata*, and *P. vittata*). The increased number of breakpoints in Lebistes shows that effective gene conversion is occurring in shorter segments compared to conversion events in Poecilia or Limia species. By occurring in smaller segments, gene conversion does not completely homogenize LWS-1 and LWS-3 in Lebistes species. The smaller segments of conversion in Lebistes may be the result of selection acting against the homogenizing effects of large gene conversion events in these species where red/orange discrimination plays a large role in mate choice. Normally differential selection in LWS-1 and LWS-3 of Lebistes versus non-Lebistes would be tested using methods such as dN/dS. However, the presence of gene conversion makes this impossible because it violates the assumptions of molecular evolution models underlying methods such as dN/dS (Kosakovsky Pond and Frost 2005b).

The presence of multiple alleles at LWS-1 in *P. parae*, *P. picta* and *P. reticulata* (Tezuka *et al.* 2014; Sandkam *et al.* 2015 and Sandkam *et al.* in review) raises the possibility that other species may also possess multiple alleles at opsin loci. The sequences presented here provide an excellent tool for future studies to investigate variation in colour vision across populations or individuals by examining the sequence and/or expression of the opsin genes.

2.5.4. Selection in Non-LWS Genes

Using dN/dS tests we identified sites under selection in several of the non-LWS genes. We found RH1 has many sites under positive selection, all of which occur either in the transmembrane domain or the retinal-binding pocket. Changes in both of these regions can be responsible for changing the tuning of the rod cells (Hargrave *et al.* 1983). RH1 has recently been shown to experience positive selection in cichlids that is driven by both ecology and phylogeny (Nagai *et al.* 2010; Schott *et al.* 2014; Torres-Dowdall *et al.* 2015). The positive selection we identified in RH1 is likely also indicative of differences in life history and presents an exciting system in which to test how differences in ecology can drive differences in visual tuning. It is noteworthy that in two colourful families of freshwater fishes (Cichlidae and now Poeciliidae) the RH1 gene appears to be so strongly under selection despite its use in non-colour based vision.

Surprisingly, most of the sites identified to be under selection by PAML and HyPhy were not located in the retinal binding pocket, which is thought to be the important sites influencing the tuning of the opsin λ_{max}. Therefore care should be taken when interpreting these selection level results until protein expression studies can be performed to show whether or not changes to these sites are indeed altering the phenotype.

2.6. Conclusion

We identified the phylogenetic relationships for nine of the ten colour opsin genes in 15 species throughout the family Poeciliidae. The LWS-1 and LWS-3 loci have undergone gene conversion at several points across this family, acting to homogenize

LWS sequence at these loci within species. Members of the subgenus Lebistes experience gene conversion in smaller segments leading to greater LWS allelic diversity compared to non-Lebistes species. The LWS opsins are responsible for detecting the reds, oranges, and yellows upon which Lebistes base mating decisions. It is possible that the LWS-1 / LWS-3 diversity observed in Lebistes species is a result of selection pressure on the sensory system for its role in sexual selection. We also found RH1 is under strong positive selection in the family Poeciliidae.

Acknowledgements

We wish to thank members of the Breden lab: Laura Chow, Ian McNeil, and Luke Rawle for assistance in the lab.

Data Accessibility

All sequences will be available on GenBank. Accession numbers pending submission.

References

Andersson, M. B. 1994. *Sexual selection*. Princeton University Press, Princeton, New Jersey.

Andersson, M. B., and L. W. Simmons. 2006. Sexual selection and mate choice. *Trends Ecol Evol* 21:296–302.

Basolo, A. L. 1990. Female preference for male sword length in the green swordtail, *Xiphophorus helleri* (Pisces, Poeciliidae). *Anim Behav* 40:332–338.

Basolo, A. L., and J. A. Endler. 1995. Sensory biases and the evolution of sensory systems. *Trends Ecol Evol* 10:489–489.

Boughman, J. W. 2002. How sensory drive can promote speciation. *Trends Ecol Evol* 17:571–577.

Chen, J., D. Cooper, N. Chuzhanova, C. Férec, and G. Patrinos. 2007. Gene conversion: mechanisms, evolution and human disease. *Nature Reviews Genetics* 8:762–775.

Conway, B. R. 2009. Color vision, cones, and color-coding in the cortex. *The Neuroscientist* 15:274–290.

Endler, J. A., and A. L. Basolo. 1998. Sensory ecology, receiver biases and sexual selection. *Trends Ecol Evol* 13:415–420.

Gegenfurtner KR, Sharpe LT (1999) *Color Vision: From genes to perception*. Cambridge University Press.

Hargrave, P. A., J. H. McDowell, D. R. Curtis, J. K. Wang, E. Juszczak, S. L. Fong, J. K. Rao, and P. Argos. 1983. The structure of bovine rhodopsin. *Biophys. Struct. Mech.* 9:235–244.

Hoffmann, M., N. Tripathi, S. R. Henz, A. K. Lindholm, D. Weigel, F. Breden, and C. Dreyer. 2007. Opsin gene duplication and diversification in the guppy, a model for sexual selection. *Proc. R. Soc. B* 274:33–42.

Hofmann, C. M., N. J. Marshall, K. Abdilleh, Z. Patel, U. E. Siebeck, and K. L. Carleton. 2012. Opsin evolution in damselfish: convergence, reversal, and parallel evolution across tuning sites. *J Mol Evol* 75:79–91.

Horth, L. 2007. Sensory genes and mate choice: Evidence that duplications, mutations, and adaptive evolution alter variation in mating cue genes and their receptors. *Genomics* 90:159–175.

Houde, A. E. 1997. *Sex, color, and mate choice in guppies*. Princeton University Press, Princeton, NJ.

Katoh, K., G. Asimenos, and H. Toh. 2009. Multiple alignment of DNA sequences with MAFFT. *Methods Mol Biol.* 537:39–64.

Kosakovsky Pond, S. E., and S. D. W. Frost. 2005a. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* 21:2531–2533.

Kosakovsky Pond, S. E., and S. D. W. Frost. 2005b. Not so different after all: A comparison of methods for detecting amino acid sites under selection. *Mol Biol Evol* 22:1208–1222.

Kosakovsky Pond, S. E., S. D. W. Frost, and S. V. Muse. 2005. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21:676–679.

Kosakovsky Pond, S. L., D. Posada, M. B. Gravenor, C. H. Woelk, and S. D. W. Frost. 2006. GARD: a genetic algorithm for recombination detection. *Bioinformatics* 22:3096–3098.

Liley, N. R. 1965. Ethological isolating mechanisms in four sympatric species of poeciliid fishes. *Behaviour. Supplement* 13:1–197.

Lindholm, A. K., R. Brooks, and F. Breden. 2004. Extreme polymorphism in a Y-linked sexually selected trait. *Heredity* 92:156–162.

Lynch, M. 2007. *The origins of genome architecture*. Sinauer Associates, Sunderland, MA.

Magurran, A. E. 2005. *Evolutionary Ecology: The Trinidadian Guppy*. Oxford University Press, New York, NY.

Murrell, B., S. Moola, A. Mabona, T. Weighill, D. Sheward, S. L. Kosakovsky Pond, and K. Scheffler. 2013. FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol Biol Evol* 30:1196–1205.

Nagai, H., Y. Terai, T. Sugawara, H. Imai, H. Nishihara, M. Hori, and N. Okada. 2010. Reverse evolution in RH1 for adaptation of cichlids to water depth in Lake Tanganyika. *Mol Biol Evol* 28:msq344–1776.

Nylander, J. A. A. 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Center, Uppsala University.

Nylander, J. A. A., J. C. Wilgenbusch, D. L. Warren, and D. L. Swofford. 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* 24:581–583.

Panhuis, T. M., R. Butlin, M. Zuk, and T. Tregenza. 2001. Sexual selection and speciation. *Trends Ecol Evol* 16:364–371.

Ptacek, M. B., and J. Travis. 1997. Mate choice in the sailfin molly, *Poecilia latipinna*. *Evolution* 51:1217.

Rambaut, A. 1996. Se-Al: Sequence Alignment Editor.

Rambaut, A., and A. Drummond. 2007. Tracer v1.4.

Renn, S. C. P., N. Aubin-Horth, and H. A. Hofmann. 2008. Fish and chips: functional genomics of social plasticity in an African cichlid fish. *J. Exp. Biol.* 211:3041–3056.

Rennison, D. J., G. L. Owens, and J. S. Taylor. 2012. Opsin gene duplication and divergence in ray-finned fish. *Mol Phylogenet Evol* 62:986–1008.

Robinson, D. M., M. S. Tudor, and M. R. Morris. 2011. Female preference and the evolution of an exaggerated male ornament: the shape of the preference function matters. *Anim Behav* 81:1015–1021.

Robinson, G. E., R. D. Fernald, and D. F. Clayton. 2008. Genes and social behavior. *Science* 322:896–900.

Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.

Ryan, M. J., and A. S. Rand. 1993. Species recognition and sexual selection as a unitary problem in animal communication. *Evolution* 47:647.

Sandkam, B. A., C. M. Young, and F. Breden. 2015. Beauty in the eyes of the beholders: colour vision is tuned to mate preference in the Trinidadian guppy (*Poecilia reticulata*). *Mol Ecol* 24:596–609.

Sandkam, B. A., J. B. Joy, C. T. Watson, P. Gonzalez-Bendiksen, C. R. Gabor, and F. Breden. 2013. Hybridization leads to sensory repertoire expansion in a gynogenetic fish, the Amazon molly (*Poecilia formosa*): a test of the hybrid-sensory expansion hypothesis. *Evolution* 67:120–130.

Schott, R. K., S. P. Refvik, F. E. Hauser, H. López-Fernández, and B. S. W. Chang. 2014. Divergent positive selection in rhodopsin from lake and riverine cichlid fishes. *Mol Biol Evol* 31:1149–1165.

Swofford, D. L. 2003. PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.

Tezuka, A., S. Kasagi, C. Van Oosterhout, M. McMullan, W. M. Iwasaki, D. Kasai, M. Yamamichi, H. Innan, S. Kawamura, and M. Kawata. 2014. Divergent selection for opsin gene variation in guppy (*Poecilia reticulata*) populations of Trinidad and Tobago. *Heredity* 113:381–389.

Torres-Dowdall, J., F. Henning, K. R. Elmer, and A. Meyer. 2015. Ecological and lineage specific factors drive the molecular evolution of rhodopsin in cichlid fishes. *Mol Biol Evol* msv159.

Ward, M. N., A. M. Churcher, K. J. Dick, C. R. Laver, G. L. Owens, M. D. Polack, P. R. Ward, F. Breden, and J. S. Taylor. 2008. The molecular basis of color vision in colorful fish: Four Long Wave-Sensitive (LWS) opsins in guppies (*Poecilia reticulata*) are defined by amino acid substitutions at key functional sites. *BMC Evol Biol* 8.

Watson, C. T., K. P. Lubieniecki, E. R. Loew, W. S. Davidson, and F. Breden. 2010. Genomic organization of duplicated short wave-sensitive and long wave-sensitive opsin genes in the green swordtail, *Xiphophorus helleri*. *BMC Evol Biol* 10:87.

Watson, C. T., S. M. Gray, M. Hoffmann, K. P. Lubieniecki, J. B. Joy, B. A. Sandkam, D. Weigel, E. R. Loew, C. Dreyer, W. S. Davidson, and F. Breden. 2011. Gene duplication and divergence of long wavelength-sensitive opsin genes in the Guppy, *Poecilia reticulata*. *J Mol Evol* 72:240–252.

Wilkinson, G. S., F. Breden, J. E. Mank, M. G. Ritchie, A. D. Higginson, J. Radwan, J. Jaquiery, W. Salzburger, E. Arriero, S. M. Barribeau, Phillips, P. C., S. C. P. Renn, and L. Rowe. 2015. The locus of sexual selection: moving sexual selection studies into the post-genomics era. *J Evolution Biol* 28:739–755.

Xu, B., and Z. Yang. 2013. PAMLX: a graphical user interface for PAML. *Mol Biol Evol* 30:2723–2724.

Yang, Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24:1586–1591.

Yokoyama, S. 2002. Molecular evolution of color vision in vertebrates. *Gene* 300:69–78.

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

Yokoyama, S., and F. B. Radlwimmer. 1998. The “five-sites” rule and the evolution of red and green color vision in mammals. *Mol Biol Evol* 15:560–567.

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

Yokoyama, S., H. Yang, and W. T. Starmer. 2008. Molecular basis of spectral tuning in the red- and green-sensitive (M/LWS) pigments in vertebrates. *Genetics* 179:2037–2043.

Zhang, J. 2003. Evolution by gene duplication: an update. *Trends Ecol Evol* 18:292–298.

Tables and Figures

Table 2.1. **The five key-site haplotype of the four LWS opsin loci with λ_{max} predicted following Yokoyama and Raddlewimmer (1998). The number of different LWS λ_{max} within species is summarized as '# LWS'.**

	LWS-1						LWS-2						LWS-3						LWS-R						#
Amino Acid	180	197	277	285	308	λ_{max}	180	197	277	285	308	λ_{max}	180	197	277	285	308	λ_{max}	180	197	277	285	308	λ_{max}	
X. helleri	S	H	Y	T	A	560	P	H	F	A	A	545	S	H	Y	T	A	560	S	H	Y	T	A	560	2
H. formosa	S	H	Y	T	A	560	P	H	F	A	A	545	S	H	Y	T	A	560	S	H	Y	T	A	560	2
P. minor	S	H	Y	T	A	560	S	H	F	A	A	534	S	H	Y	T	A	560	S	H	Y	T	A	560	2
P. caymanensis	S	H	Y	T	A	560	P	H	F	A	A	545	S	H	Y	T	A	560	S	H	F	A	A	534	3
P. nigrofasciatus	S	H	Y	T	A	560	P	H	F	A	A	545	S	H	Y	T	A	560	S	H	F	A	A	534	3
P. vittata	S	H	Y	T	A	560	P	H	F	A	A	545	S	H	Y	T	A	560	S	H	F	A	A	534	3
P. latipinna	S	H	Y	T	A	560	P	H	F	A	A	545	S	H	Y	T	A	560	S	H	F	A	A	534	3
P. mexicana	S	H	Y	T	A	560	P	H	F	A	A	545	S	H	Y	T	A	560	S	H	F	A	A	534	3
P. velifera	S	H	Y	T	A	560	P	H	F	A	A	545	S	H	Y	T	A	560	S	H	F	A	A	534	3
P. petenensis	S	H	Y	T	A	560	P	H	F	A	A	545	S	H	Y	T	A	560	S	H	F	A	A	534	3
P. bifurca	A	H	Y	T	A	553	P	H	F	A	A	545	S	H	Y	T	A	560	S	H	Y	T	A	560	3
P. picta	A	H	Y	T	A	553	P	H	F	A	A	545	S	H	Y	T	A	560	S	H	Y	T	A	560	3
P. wingei	A	H	Y	T	A	553	P	H	F	A	A	545	S	H	Y	T	A	560	S	H	Y	T	A	560	3
P. parae	S	H	Y	T	A	560	P	H	F	A	A	545	S	H	Y	T	A	560	S	H	Y	T	A	560	2
P. reticulata	S	H	Y	T	A	560	P	H	F	A	A	545	S	H	Y	T	A	560	S	H	Y	T	A	560	2

Table 2.2. Selection tests on non-LWS opsins using models M0, M1, M2, M7 and M8 of the codeml program in PAML. Likelihood ratio tests (LRTs) were performed between models with selection and models with neural evolution to identify genes under positive selection.

Gene	Model	np	*ln*L	*K*	w0/p	w1/q	w2/wp	Null	LRT	*P*
	M0	30	-2212.50	2.09	0.08			n/a		
	M1	31	-2185.55	2.03	0.00	1.00		n/a		
SWS1	M2	33	-2183.62	2.10	0.02	1.00	2.39	M1	3.86	> 0.05
	M7	31	-2186.57	2.07	0.01	0.05		n/a		
	M8	33	-2184.79	2.10	1.98	99.00	2.38	M7	3.55	> 0.05
	M0	30	-2617.56	4.87	0.19			n/a		
	M1	31	-2599.21	4.77	0.05	1.00		n/a		
SWS2A	M2	33	-2597.28	4.88	0.10	1.00	2.27	M1	3.86	> 0.05
	M7	31	-2599.92	4.80	0.06	0.27		n/a		
	M8	33	-2597.26	4.88	1.18	9.56	2.45	M7	5.32	> 0.05
	M0	30	-2365.32	2.74	0.10			n/a		
	M1	31	-2348.02	2.73	0.00	1.00		n/a		
SWS2B	M2	33	-2347.33	2.77	0.01	1.00	5.53	M1	1.36	> 0.05
	M7	31	-2348.02	2.73	0.01	0.07		n/a		
	M8	33	-2347.24	2.76	0.03	0.30	5.20	M7	1.54	> 0.05
	M0	30	-2385.05	3.30	0.34			n/a		
	M1	31	-2321.15	2.77	0.00	1.00		n/a		
RH1	M2	33	-2291.25	3.32	0.01	1.00	5.41	M1	59.80	**<0.0001**
	M7	31	-2321.82	2.72	0.01	0.05		n/a		
	M8	33	-2291.25	3.31	1.50	99.00	5.46	M7	61.14	**<0.0001**

Table 2.3. Sites identified as under positive selection using the M2 and M8 models in codeml of PAML, and the REL, FUBAR, and FEL programs of HYPHY. Grey highlight indicates the site is in the transmembrane domain. / indicates site was not identified to be under selection using that method. NA indicates run was not available. Numbers given for M2, M8, REL, and FUBAR are posterior probabilities while numbers for FEL are *P* values.

Gene	Site	M2 (Pr)	M8 (Pr)	Rel (Pr)	FUBAR (Pr)	FEL (*P*)
SWS1	62	/	/	0.712	/	/
	129	/	0.640	/	/	/
	144	0.780	0.912	0.718	/	/
	155	0.757	0.903	0.679	/	/
	163	0.541	0.682	0.761	/	/
	193	0.910	0.976	0.707	/	/
	274	0.908	0.975	/	/	/
SWS2A	5	0.721	0.908	0.989	0.939	0.036
	7	/	0.580	0.734	/	/
	31	/	0.682	0.825	/	/
	66	/	0.589	/	/	/
	100	/	0.598	/	/	/
	160	0.747	0.936	0.981	18.427	/
	161	/	0.617	0.713	/	/
	164	0.766	0.944	0.905	/	/
	234	0.750	0.938	0.990	24.249	/
	288	/	0.595	0.738	/	/
	324	0.637	0.830	0.928	/	/
	325	/	0.589	0.760	/	/
SWS2B	11	0.896	0.972	/	/	/
	54	0.549	0.666	/	/	/
	64	0.772	0.898	/	/	/
	126	0.585	0.698	/	/	/
	170	0.653	0.777	/	/	/
	171	0.539	0.656	/	/	/
	274	/	0.576	/	/	/
RH2-1	32	NA	NA	0.793	/	/
	166	NA	NA	0.998	/	0.027
	187	NA	NA	0.975	/	/
	217	NA	NA	0.961	/	/
	274	NA	NA	0.827	/	/

	327	NA	NA	0.998	0.990	0.044
	343	NA	NA	0.966	/	/
	11	/	/	0.963	/	/
	19	0.627	0.759	0.986	/	/
	38	0.689	0.807	0.847	/	/
	63	/	/	0.952	/	/
	95	/	/	0.952	/	/
	112	/	/	0.980	/	/
	133	0.990	0.998	0.987	0.934	0.072
	149	0.657	0.782	0.974	/	/
	166	0.993	0.999	0.975	0.919	/
	173	0.719	0.829	0.964	/	/
	189	0.937	0.980	/	/	/
RH1	205	0.934	0.979	/	/	/
	213	0.990	0.998	0.968	0.908	/
	217	0.948	0.984	/	/	/
	218	0.936	0.980	0.987	/	/
	255	0.991	0.998	0.883	/	/
	263	0.991	0.998	0.877	/	/
	266	0.998	1.000	/	/	/
	290	0.932	0.979	0.971	/	/
	297	0.999	1.000	0.832	/	/
	298	/	/	0.962	/	/
	304	0.984	0.997	/	/	/
	309	0.977	0.995	0.943	/	/

Figure 2.1. Amino acid alignment of SWS1.

Figure 2.2. Amino acid alignment of SWS2A.

Figure 2.3. Amino acid alignment of SWS2B.

Figure 2.4. Amino acid alignment of RH1.

Figure 2.5. Amino acid alignment of RH2-1.

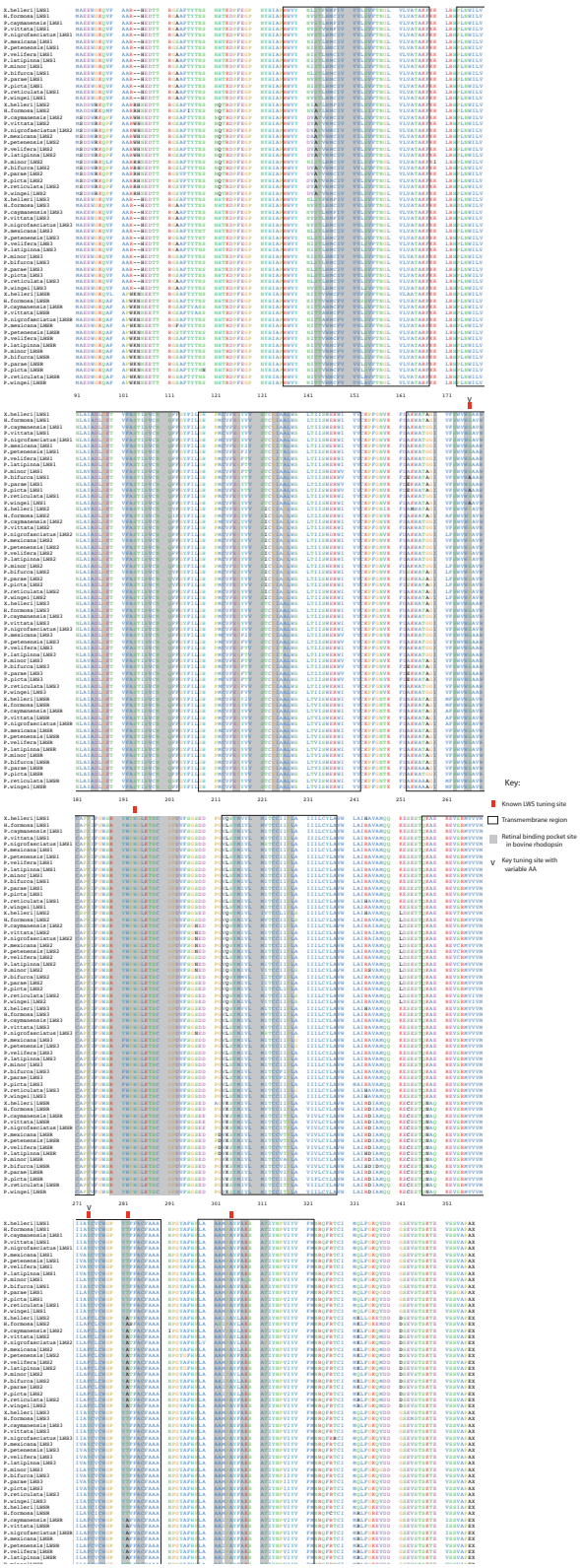


Figure 2.6. Amino acid alignment of LWS.

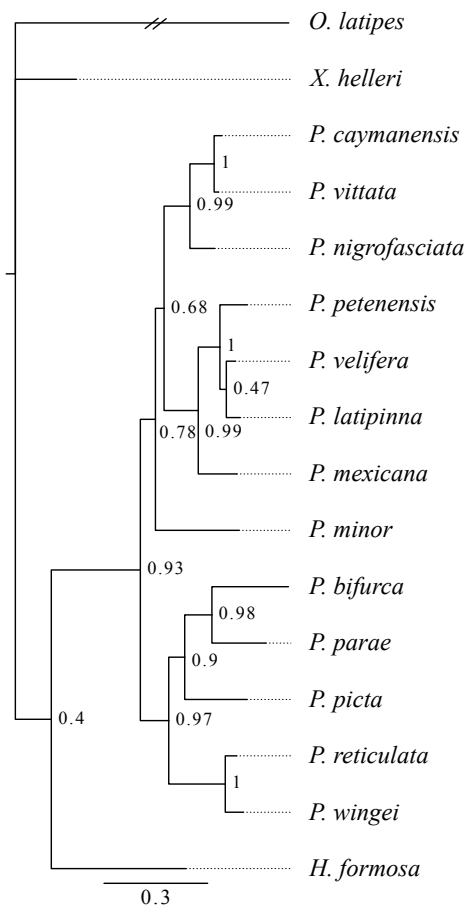


Figure 2.7. ND2 mitochondrial bayesian consensus tree with posterior probabilities. Outgroup is *Oryzias latipes*.

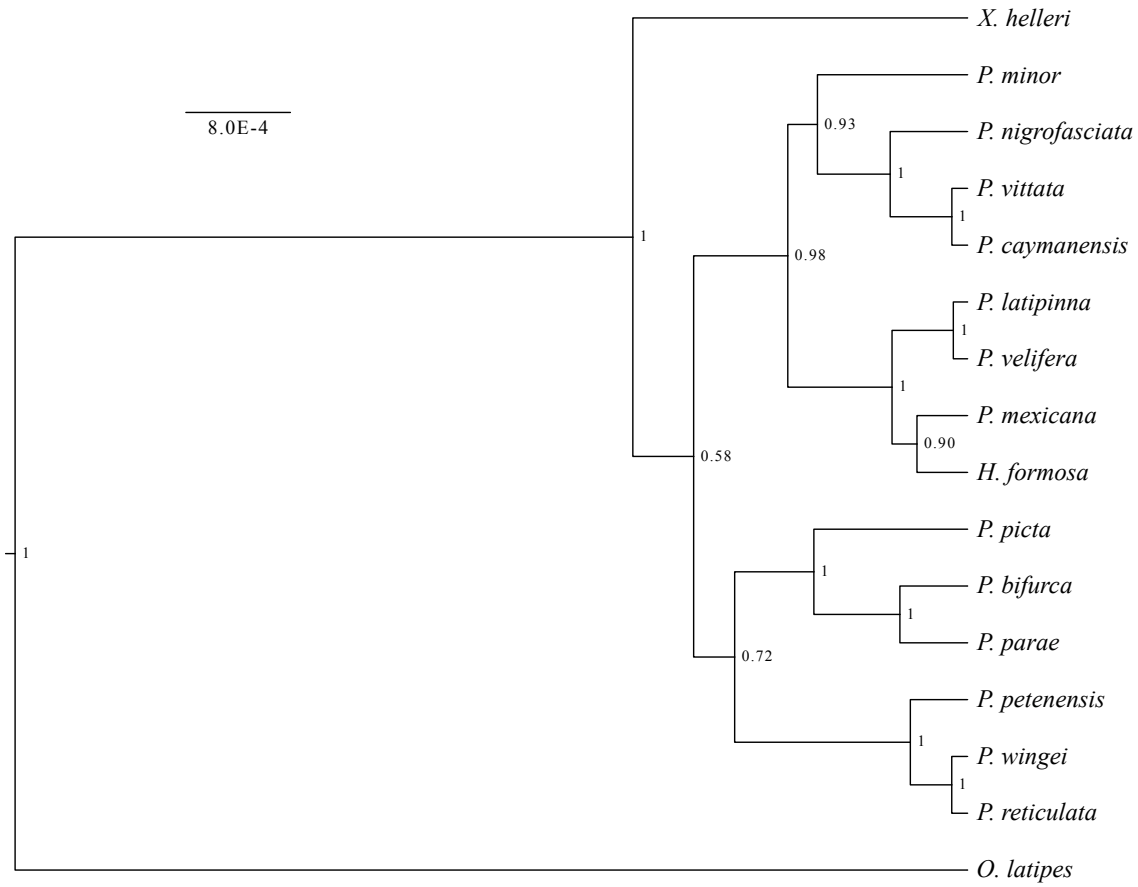


Figure 2.8. Bayesian consensus tree with posterior probabilities of SWS1. Outgroup is *Oryzias latipes*.

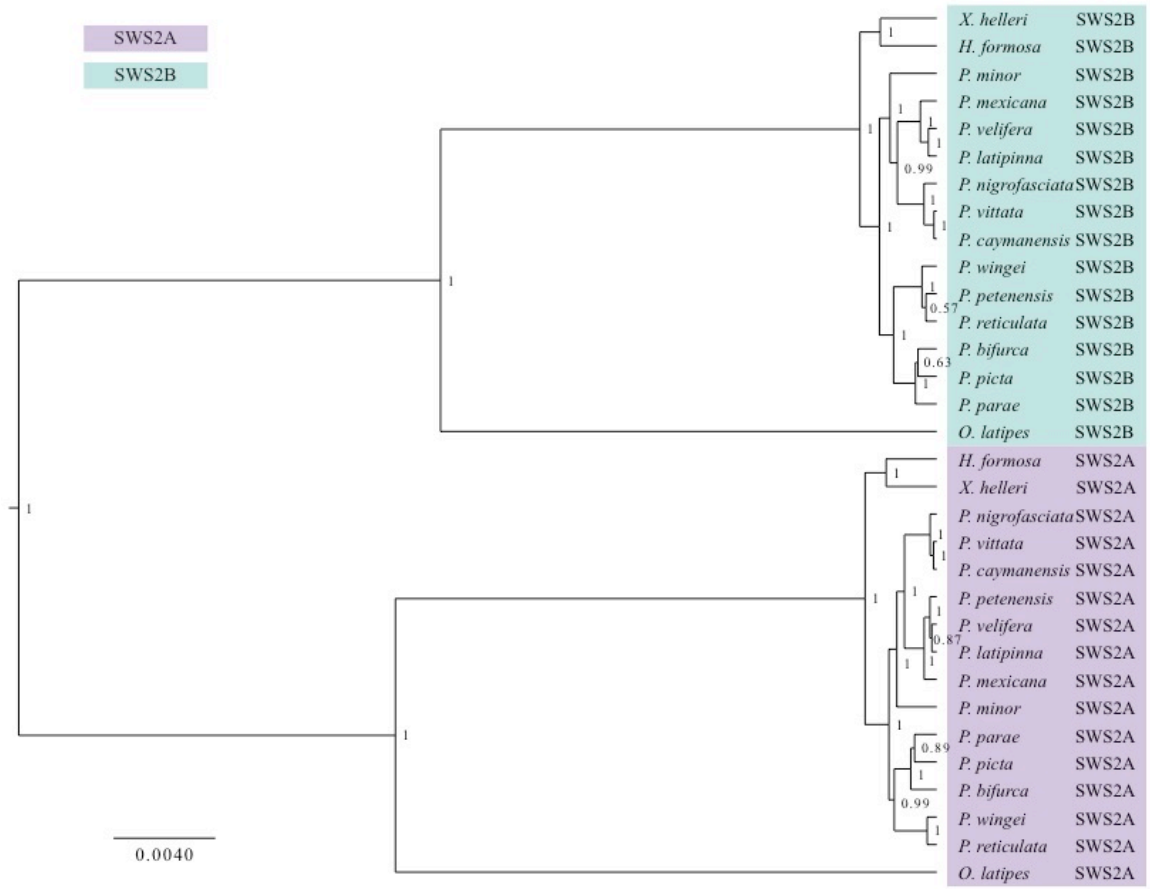


Figure 2.9. Bayesian consensus tree with posterior probabilities of SWS2A and SWS2B. Outgroup is *Oryzias latipes*.

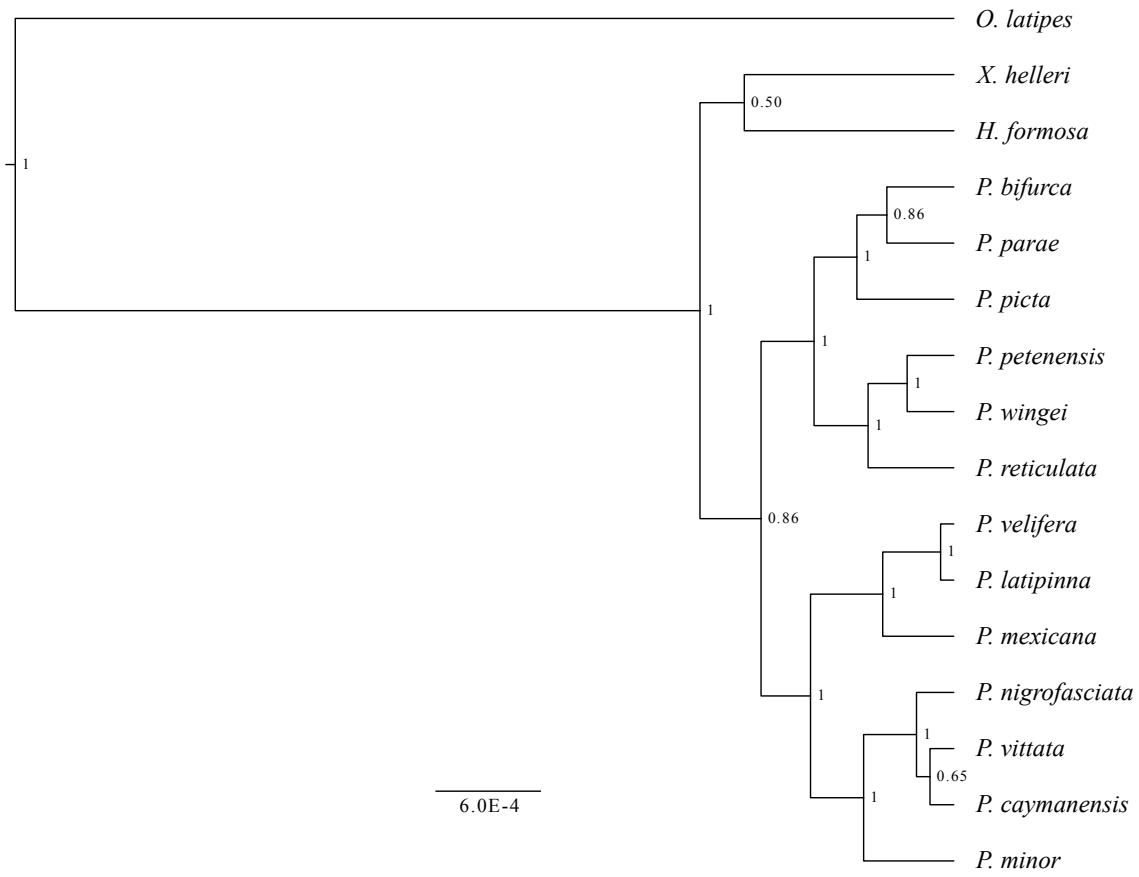


Figure 2.10. Bayesian consensus tree with posterior probabilities of RH1. Outgroup is *Oryzias latipes*.

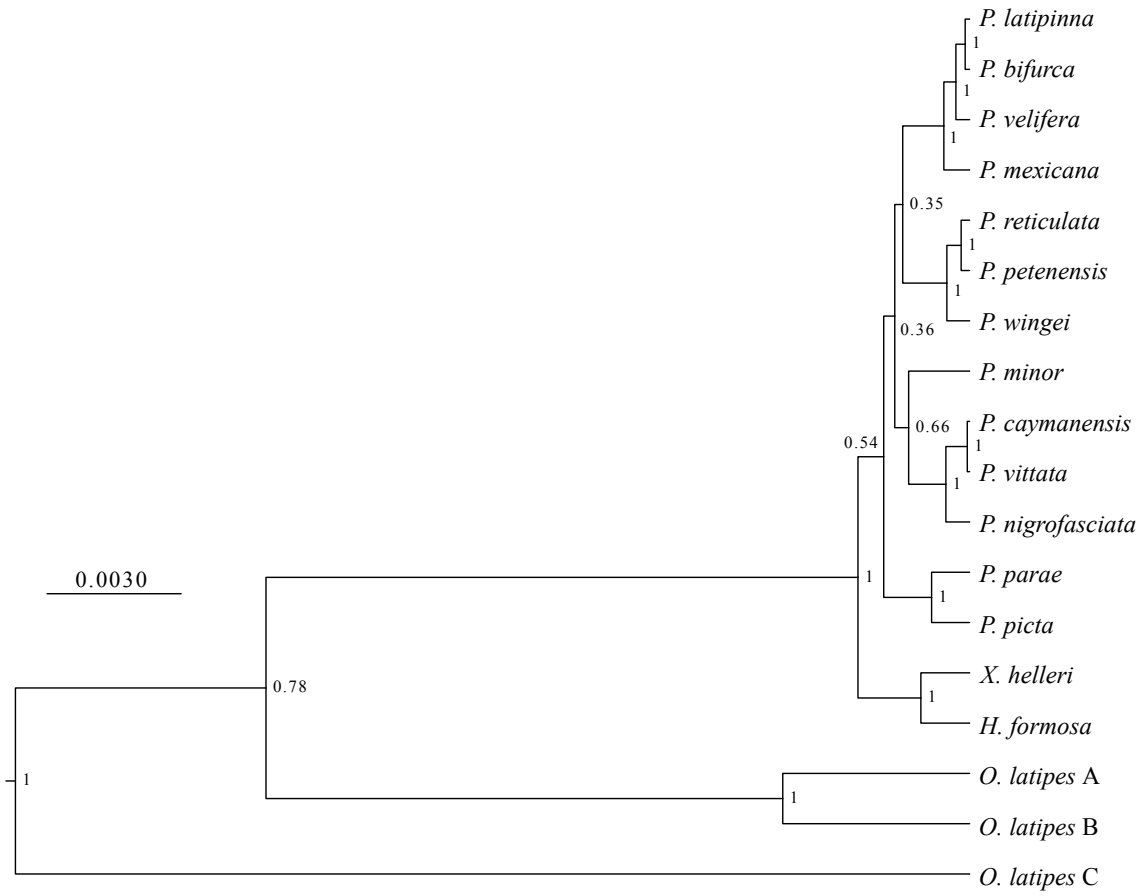


Figure 2.11. Bayesian consensus tree with posterior probabilities of RH2-1. Outgroup is *Oryzias latipes*.

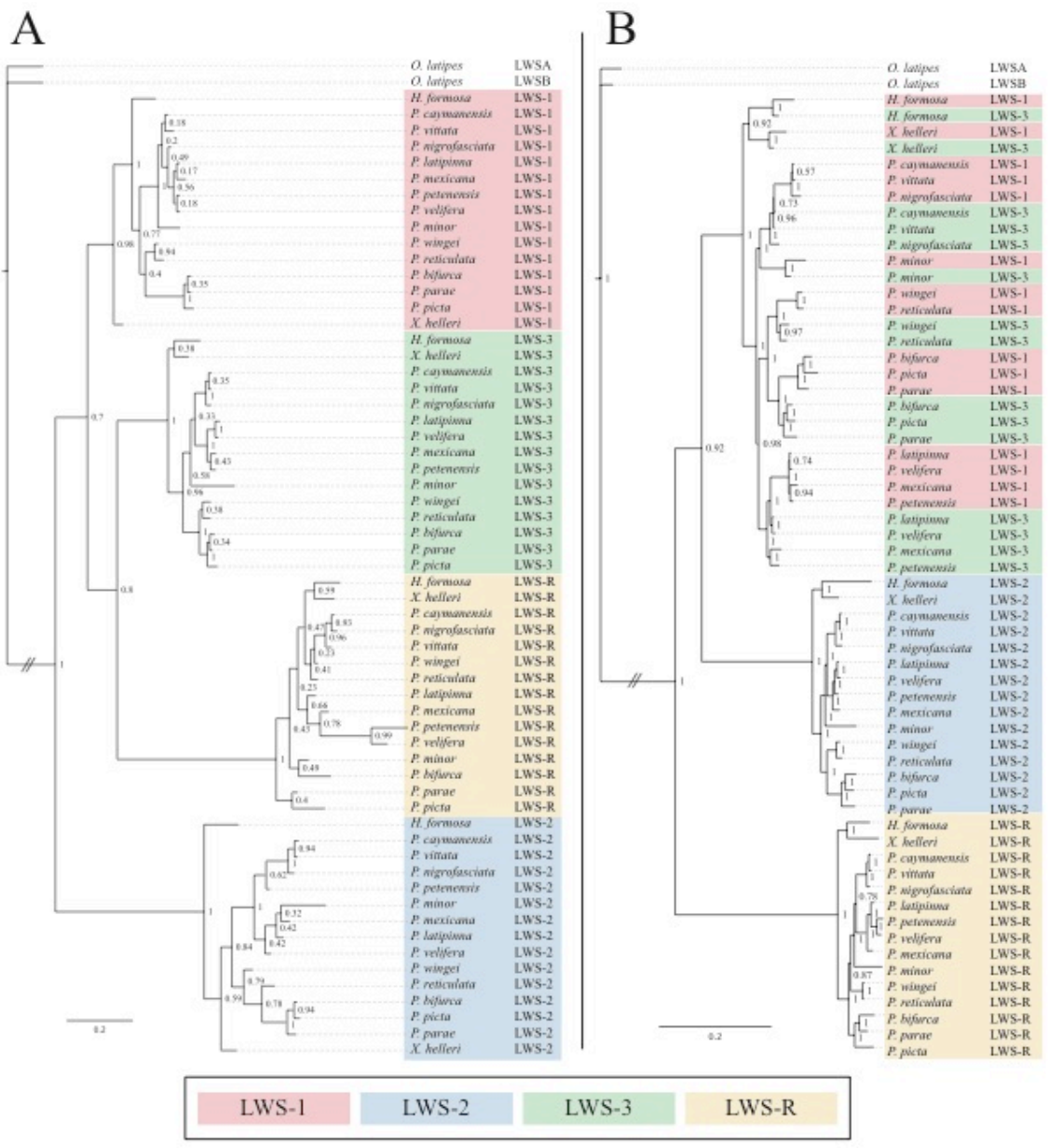


Figure 2.12. Bayesian consensus trees with posterior probabilities of (A) LWS UTR sequence, and (B) LWS full sequences. Outgroup is *Oryzias latipes*. Note the expected duplication history recovered in UTR sequence and general species relationships (based on ND2 tree) recovered within clades except for LWS-1 and LWS-3 of the full LWS tree.

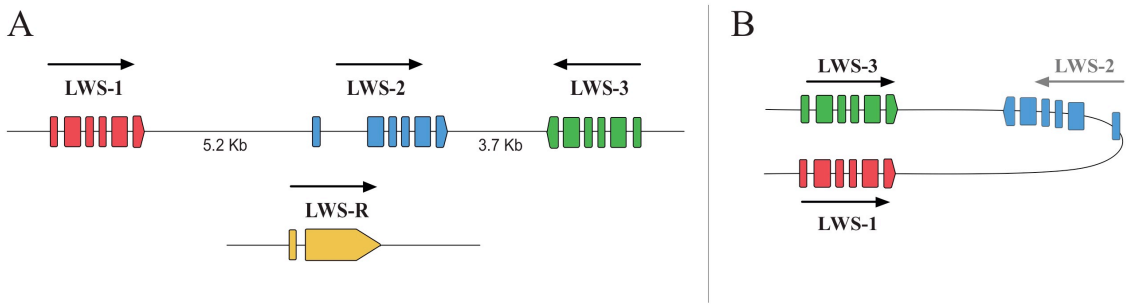
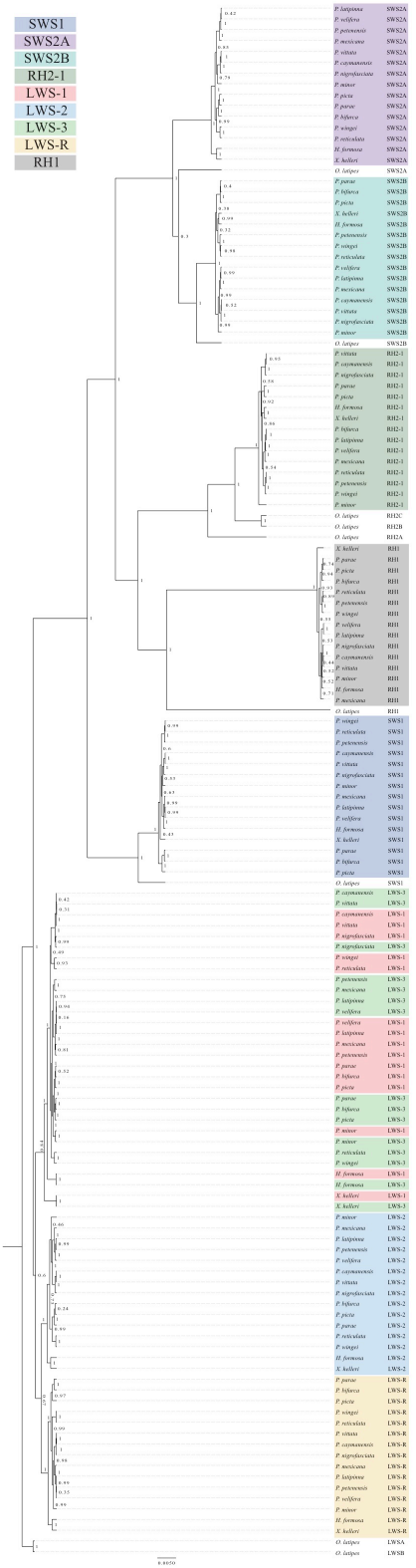


Figure 2.13. **(A) The genomic organization of the LWS opsin loci in *P. wingei* (adopted from Watson *et al.* (2011)). (B) The proposed conformation leading to high rates of gene conversion between LWS-1 and LWS-3 in the event of a double strand break. Arrows denote directionality. Coloured boxes denote exons while spaces between boxes denote introns.**



Supplementary Figure 2.1. Bayesian consensus tree with posterior probabilities of exon only sequence from all opsins. Outgroup is *Oryzias latipes*.

Supplementary Table 2.1. PCR primer sequences.

Species	LWS1		LWS2 (Piece 1)		LWS2 (Piece 2)		LWS3		LWSR		SWS1		SWS2A		SWS2B		RH1		RH2-1	
	For	Rev	For	Rev	For	Rev	For	Rev	For	Rev	For	Rev	For	Rev	For	Rev	For	Rev	For	Rev
Poecilia nigrofasciatis	1	2	6	7	8	9	11	12	16	17	18	19	20	21	22	23	24	25	28	29
Poecilia vittata	1	3	6	7	8	9	11	12	16	17	18	19	20	21	22	23	26	27	28	29
Pamphorichthys minor	1	4	6	7	8	10	13	12	16	17	18	19	20	21	22	23	26	27	28	29
Poecilia caymanensis	1	3	6	7	8	9	14	15	16	17	18	19	20	21	22	23	26	27	28	29
Heterandrai formosa	1	3	6	7	8	9	11	12	16	17	18	19	20	21	22	23	26	27	28	29
Peocilia parae	1	5	6	7	8	9	13	12	16	17	18	19	20	21	22	23	26	27	28	29
Poecilia picta	1	3	6	7	8	9	13	12	16	17	18	19	20	21	22	23	26	27	28	29
Poecilia bifurca	1	3	6	7	8	9	13	12	16	17	18	19	20	21	22	23	26	27	28	29
Poecilia latipinna	1	3	6	7	8	9	11	12	16	17	18	19	20	21	22	23	26	27	28	29
Poecilia reticulata	1	5	6	7	8	9	13	12	16	17	18	19	20	21	22	23	26	27	28	29
Poecilia petenensis	1	3	6	7	8	9	11	12	16	17	18	19	20	21	22	23	26	27	28	29
Poecilia mexicana	1	3	6	7	8	10	11	12	16	17	18	19	20	21	22	23	26	27	28	29
Poecilia velifera	1	3	6	7	8	9	13	12	16	17	18	19	20	21	22	23	26	27	28	29
Poecilia wingei	-	-	-	-	-	-	-	-	-	-	18	19	-	-	-	-	26	27	28	29
Xiphophorous helleri	-	-	-	-	-	-	-	-	-	-	18	19	-	-	-	-	26	27	28	29

Primer	Sequence
1	TGTGAAGTGCAGATCACCTAG
2	CTTCATAAACAGAGAACATGAAGG
3	AGAATTTGTTTTCTCCCAGCC
4	GTTACACATTCATGCATGATGC
5	ACACATTCATGCATGATGCAG
6	CCCAGCAAAACTCTCAAGGT
7	GCAAAGACTGTCTCTCCAAGATC
8	GATCCCTTTGAAGGACCAAACT
9	CAGTCCCGGCAGTAATAACAAAC
10	GGTAATACAATGCAAAAGTTCCATC
11	CCCAGCAAACGCAGAAGGTTATAAG
12	GTAGCACTCAAAATACTTTCAGTAC
13	GATAAGCGTTGGTATATAAGGAG
14	GGTGAAGCACCCAGCAAACG
15	CATTTGCCATAAAGTTTCCGTTTATC
16	CCACAAGGTTACAAGCCAGAG
17	TGCATTGACACATCCATGAGTTC
18	GTCCATCATCATCTTCTCCT
19	AGTTGGCGTACCAGAGG
20	ATGTTTTAGATTTGGACTTGGTG
21	GCAGAATGTTGCTAAGTTGTTC
22	CGTATTTAATCATCCAGTCTAG
23	GCAGATGTGCTACATGCATT
24	CAACCATGAATGGCACAGAG
25	CTCGGTCTTGGAAGCAGTAG
26	CTGTGGCCTAAATCAATTTGTGC
27	AAGCCAGTATCATATAGGATCC
28	GAGTGTAGGATTGGCATTTACCA
29	TTGCACGCGTAACCAGATGA

Supplementary Table 2.2. GenBank accession numbers of ND2 sequences used to identify species relationships.

Species	ND2 GeneBank Accession
Poecilia nigrofasciatis	AF031391.1
Poecilia vittata	AF353201.1
Poecilia minor	GU179242.1
Poecilia caymanensis	AF353192.1
Heterandria formosa	AF084973.1
Poecilia parae	GU179235.1
Poecilia picta	GU179237.1
Poecilia bifurca	GU179232.1
Poecilia latipinna	AF031389.1
Poecilia reticulata	AF031394.1
Poecilia petenensis	AF031401.1
Poecilia mexicana	AF080488.1
Poecilia velifera	AF031402.1
Poecilia wingei	AY249318.1
Xiphophorous helleri	FJ226476.1
Oryzias latipes	AP004421.1

Chapter 3.

Hybridization leads to sensory repertoire expansion in a gynogenetic fish, the Amazon molly (*Poecilia formosa*): a test of the hybrid-sensory expansion hypothesis

Publication and Contributions

A version of this chapter is published as: Sandkam, B. A., Joy, J. B., Watson, C. T., Gonzalez-Bendiksen, P., Gabor, C. R. and Breden, F. (2013), Hybridization leads to sensory repertoire expansion in a gynogenetic fish, the Amazon Molly (*Poecilia formosa*): a test of the hybrid sensory expansion hypothesis. Evolution, 67: 120–130. doi: 10.1111/j.1558-5646.2012.01779.x

Contributions: BAS, JBJ, CTW and FB conceived and designed the project. BAS performed all sequencing and analyses. BAS and JBJ built phylogenies and drafted early versions of the manuscript. GBP and CRG ran all behavioural trials.

3.1. Abstract

Expansions in sensory systems usually require processes such as gene duplication and divergence, and thus evolve slowly. We evaluate a novel mechanism leading to rapid sensory repertoire expansion: hybrid-sensory expansion (HSE). HSE occurs when two species with differently tuned sensory systems form a hybrid, bringing together alleles from each of the parental species. In one generation a sensory repertoire is created that is the sum of the variance between parental species. The Amazon molly presents a unique opportunity to test the HSE hypothesis in a “frozen” hybrid. We compared opsin sequences of the Amazon molly, *Poecilia formosa*, to those of the parental species. Both parental species are homozygous at the RH2-1 locus and

each of the four LWS loci, while *P. formosa* possess two different alleles at these loci; one matching each parental allele. Gene expression analysis showed *P. formosa* use the expanded opsin repertoire that was the result of HSE. Additionally, behavioural tests revealed *P. formosa* respond to coloured stimuli in a manner similar or intermediate to the parental species *P. mexicana* and *P. latipinna*. Together these results strongly support the HSE hypothesis. Hybrid sensory repertoire expansion is likely important in other hybrid species and in other sensory systems.

3.2. Introduction

The evolution of sensory systems is generally driven by selection acting upon variation arising through gradual intragenomic processes such as mutation, gene duplication, and recombination (Horth 2007). More rarely, novel variation in sensory systems could also arise through more rapid intergenomic processes such as hybridization, although empirical evidence for such a process has not yet been demonstrated.

We suggest that hybridization, a previously unexplored mechanism behind sensory repertoire expansion, could act to rapidly combine variation accumulated in two different genomes, often from different environments, into a single organism. By bringing together variation from two organisms in one mating event, hybridization expands the sensory repertoire much more quickly than more traditional models of duplication and divergence. We term this process the hybrid-sensory expansion (HSE) hypothesis (outlined in Box 3.1). A diploid F1 hybrid receives one allele of each sensory locus from both parental species (Dowling and Secor 1997; Lampert and Schartl 2008). The HSE hypothesis predicts that if the parental species' sensory repertoires are differentially tuned, then the sensory repertoire of the hybrid will be different than that of either parental species. When a hybrid species possesses a sensory repertoire which is a combination of the variation present in the two parental species, selection may act on that variation. When a hybrid lives in a similar selective environment to the parental species, selection is likely to maintain the sensory repertoire to be more similar to either or both of the parental species, which could impact hybrid fitness.

The sensory repertoires of species capable of colour vision are often tuned to take advantage of their specific spectral environments (Carleton and Kocher 2001; Fuller *et al.* 2005; Endler *et al.* 2005). Colour vision is mediated via cone cells in the retina of the eye (Loew and Lythgoe 1978). The peak spectral sensitivity (λ_{max}) of a cone cell is determined by its visual pigment, consisting of a chromophore coupled with a transmembrane protein called an opsin, which is maximally sensitive to a specific wavelength of light (Yokoyama 2000, 2002). The fishes in the family Poeciliidae possess excellent colour vision (Schwanzara 1967; Houde 1987; Anstis *et al.* 1998; Grether *et al.* 2005) and have undergone extensive gene duplication and subsequent differentiation in their long wavelength sensitive (LWS) opsin genes. Thus, some poeciliid species possess the largest known opsin repertoire among vertebrates (Hoffmann *et al.* 2007; Windsor and Owens 2009; Watson *et al.* 2011).

The Amazon molly (*Poecilia formosa*) is a member of the Poeciliidae family and is the result of a hybridization event between a female *P. mexicana* and male *P. latipinna* (Hubbs and Hubbs 1932; Avise *et al.* 1991; Schartl *et al.* 1995) that likely occurred once about 120,000 years ago (Stöck *et al.* 2010). *Poecilia formosa* reproduce through gynogenesis (Kallman 1962; Schlupp 2005), a form of clonal reproduction in which a diploid egg is created through mitosis. During this process the development of the egg is triggered by the presence of sperm from a sexual species. However, the sperm's genetic material is almost never incorporated into the egg (Beukeboom and Vrijenhoek 1998; Schlupp and Riesch 2011). This unusual form of reproduction has led some people to call *P. formosa* "frozen F1s" (Vrijenhoek 1979), because they are predicted to still possess full haploid genomes from each parental species. Previous work on the cone cells of the parental species *P. mexicana* and *P. latipinna* show that they are tuned to different wavelengths of light in ranges characteristic of the RH2-1/LWS class of opsins, while the cone cells of *P. formosa* span the range of the parental species (Körner *et al.* 2006). Both of these classes of opsins function in the cone cells to provide photopic (colour) vision. While RH2-1 typically detects greens and yellows, the LWS opsins detect greens, yellows, reds, and oranges (Yokoyama *et al.* 2008). The maintenance of the "F1 genotype" together with the potential of an expanded opsin repertoire makes specific predictions about the sensory repertoire of *P. formosa* allowing a formal evaluation of the HSE hypothesis.

We tested the HSE hypothesis by comparing RH2-1 and LWS opsin sequences of *P. formosa* to those of its parental species, *P. mexicana* and *P. latipinna*. The HSE hypothesis predicts that *P. formosa* RH2-1 and LWS opsin alleles should be more similar to one of the parental species (*P. mexicana* or *P. latipinna*) than they would be to each other, and that this combination of diverged parentally-derived opsins would result in a unique, differentially tuned sensory system. We compared genomic sequence of RH2-1 and the four LWS opsins to evaluate our hypothesis that there is an expanded RH2-1/LWS opsin repertoire in *P. formosa* due to one chromosome coming from the maternal ancestor, *P. mexicana* and the other chromosome coming from the paternal ancestor, *P. latipinna*. To verify that the hybrid species utilizes its expanded opsin repertoire we identified which of the opsin alleles were being expressed in the eyes of *P. formosa*. In addition, we identified behavioural differences in response to colour stimuli between *P. formosa* across these three species. We conclude by interpreting our findings of expanded opsin repertoire, expression profiles, and behavioural differences as support for the HSE hypothesis in light of experimental data on spectral tuning of cone cells in *P. formosa* and its parental species (Körner *et al.* 2006).

3.3. Materials and Methods

3.3.1. Sample Preparation, PCR, Cloning and Sequencing

DNA was extracted from tissue samples of single specimens of *P. mexicana*, *P. latipinna* and *P. formosa* from natural populations using a DNeasy blood and tissue kit (QIAGEN). Primers specific to 5' and 3' UTR regions of each of the four LWS loci were designed using genomic data from *Poecilia reticulata* (Watson *et al.* 2011; GenBank Accession: HM540108 and HM540107) and *Xiphophorous hellerii* (Watson *et al.* 2010; GenBank Accession: GQ999832 and GQ999833). Primers were also designed in the 5' and 3' UTR regions of an RH2-1 transcript from *P. reticulata* (Hoffmann *et al.* 2007; GenBank Accession: DQ234859). Given the high exon sequence similarity between LWS loci we designed all primers to be locus specific (for a list of primers used and their sequences see Supplemental Table 3.1). The LWS opsins in the Poeciliidae family have previously been identified by amino acids at position 180 of the protein, because this site

is known to dramatically influence the λ_{max} of LWS opsins (Yokoyama and Radlwimmer 1998), and has been shown to differ between LWS loci in *Poecilia reticulata* and the sister taxa in the subgenus, Micropoecilia (Ward *et al.* 2008). However, this nomenclature becomes confusing when describing poeciliids outside of this clade, as several species have multiple loci with the 'S180' genotype (Watson *et al.* 2010; this study; Sandkam unpublished data). To avoid confusion between species, we refer to LWS loci by their location in the genome relative to one another: LWS-1 (previously A180 or S180-1), LWS-2 (previously P180), LWS-3 (previously S180 or S180-2), and LWS-R (because it is the result of a retrotransposition event) (previously S180r) (Ward *et al.* 2008; Watson *et al.* 2010, 2011; Laver and Taylor 2011). Due to the length of LWS-2, PCR products for this gene were generated in two overlapping segments. To maintain locus specificity for these segments we amplified each with internal and external primers. LWS-3 and the second segment of LWS-2 each had two primer sets designed for them; one specific to *P. latipinna* and one specific to *P. mexicana*.

PCR products generated from *P. formosa* DNA were cloned prior to sequencing to ensure *P. mexicana* and *P. latipinna* derived alleles were sequenced and analyzed separately. PCR products were cloned into One Shot® chemically competent TOP10 *E. coli* cells using a TOPO TA Cloning® kit with pCR®2.1-TOPO® vector (Invitrogen®). Plasmid DNA was isolated from overnight cultures using the QIAprep® miniprep kit (QIAGEN). All sequencing was performed by Molecular Cloning Laboratories (McLab; San Francisco, CA, USA). Sequence chromatograms were viewed and analyzed using SeqMan Pro (Lasergene 8.0; DNASTAR). All sequences are available on GenBank; accession numbers JF823551 - JF823570.

3.3.2. Phylogenetic Analyses

All phylogenetic analyses were performed separately on the RH2-1 and LWS classes of genes, because the divergence of these opsin subfamilies pre-dates the Poeciliidae; as a result, many gene regions within the RH2 and LWS genes are highly diverged and cannot be aligned (Rennison *et al.* 2011). Sequences for each group of loci were aligned using Mafft v6.833b (Katoh *et al.* 2009) and edited using Se-Al v2.0a11 (Rambaut 1996) to ensure intron-exon boundaries were consistent across LWS or RH2-

1 loci for each of the three species. Best-fit models of molecular evolution were determined using MrModelTest 3.04 (Nylander 2004). Phylogenetic trees were reconstructed under Maximum Likelihood (ML) using PAUP* 4.0b10 (Swofford 2003) and Bayesian methods as implemented in MrBayes 3.1.2 (Ronquist and Huelsenbeck 2003). Two runs utilizing four Markov chains (three heated and one cold) were run for 10^7 generations, with trees sampled every 1000 generations for the LWS loci and RH2-1. Convergence was assessed using the standard deviation of the split frequencies between runs, and graphically using the program Tracer (Rambaut and Drummond 2007) and AWTY (Nylander *et al.* 2008). ML bootstrap values and Bayesian posterior probabilities were employed to assess support. Pair-wise nucleotide similarities within and between species were calculated for each locus using BLASTn (Zhang *et al.* 2000).

3.3.3. Determination of Alleles Expressed in a Hybrid Species

Both eyes from one individual of *P. formosa* were removed and placed into RNAlater (Life Technologies) immediately after being sacrificed in an overdose of MS-222. Eyes were ground with a disposable pestle (VWR®) in 600 µL of TRIzol® reagent (Life Technologies) and homogenized in a Homogenizer Cartridge (Life Technologies). RNA was isolated using the TRIzol® Plus Purification System (Life Technologies) and subjected to an on-column PureLink® DNase treatment (Life Technologies) to ensure no DNA was present before generating cDNA. Reverse transcription was carried out using a MultiScribe Reverse Transcriptase™ kit (Life Technologies) with the addition of an RNase inhibitor (Life Technologies).

Using genomic data generated above, one PCR primer set was designed to be specific to each parental species for each locus (Supplemental Table 3.2). The primer sets for LWS-R did not work and so the same primer set used for genomic PCR was used. PCR products of locus RH2-1 and LWS-1 were sequenced directly, resulting in 'clean' single consensus sequences, without the presence of multiple peaks per base pair position in the chromatograms. A direct sequencing approach did not generate 'clean' sequence for LWS-2, LWS-3, and LWS-R therefore PCR products of these loci were cloned and sequenced following the methods described above. This approach did

yield ‘clean’ single consensus sequences. Sequences were visually compared to genomic sequences generated for each locus to confirm allele type.

3.3.4. Behavioural Response to Coloured Disks

Adult female *P. latipinna* and *P. formosa* were collected in 2011 from an introduced population (*P. latipinna* from Louisiana and Florida, (Brown 1953); and *P. formosa* from Brownsville, Texas (Hubbs *et al.* 1991)) in Spring Lake at the headwaters of the San Marcos River in Hays County, Texas (29.89˚N, 97.82˚W) and maintained in the lab for several months before testing. Allopatric adult female *P. mexicana* were collected in Campeche, Mexico in 2002 (19.14˚N, 90.50˚W). Testing individuals were maintained at 25˚C on a 14:10 h light:dark cycle with full spectrum UV lights (Rayon Lighting Group) and fed daily with Purina AquaMax 200 pellets supplemented with live brine shrimp. The spectrum and intensity of the lighting environment was not measured because all holding tanks were exposed to the same lighting.

To determine level of attraction to different colours, simultaneous choice tests were performed on individual fish by measuring the number of approaches toward and pecks at coloured discs following the methodologies of Rodd *et al.* (2002). The testing arena used was 61x51cm, filled to a depth of 15cm and lit overhead with full spectrum fluorescent lighting (Rayon Lighting Group). Black plastic covered three sides of the arena, the fourth was covered with one-way glass to minimize disturbance from the observer. Female *P. mexicana* (n = 17), *P. formosa* (n = 17), and *P. latipinna* (n = 15) were tested individually. All behavioural tests were performed at Texas State University.

Four clear plastic petri dishes (diameter 9.7cm) were randomly placed on the bottom of the tank and buried in gravel so that the lip of the dish formed a ring that could be seen by the observer and acted as an association zone with a coloured disk placed in the center. Coloured disks were made by painting pennies using Acrylic Colors (by Liquitex) either orange (Cadmium Orange Hue), yellow (Cadmium Yellow Hue), red (Naphthol Crimson), or green (Light Green). These were the same paints used by Rodd *et al.* 2002, which include reflectance spectra for each of the coloured disks. Preliminary trials revealed that individuals were reluctant to begin foraging unless food was present;

therefore a brown food pellet (AquaMax 200) was placed on top of each coloured disk to elicit foraging behaviour. The order in which the coloured disks were arranged was randomized for each trial. A focal individual was introduced to the aquarium and allowed 10 min to begin foraging; if they did not begin foraging they were returned to the holding tank to be tested at another time. A trial began when an individual began foraging and lasted 10 min, during which time the number of approaches (snout passes into the association zone) and number of pecks toward each coloured disk were recorded. To compare the raw number of pecks and approaches toward the different coloured disks we ran repeated measures ANOVAs followed by Tukey's HSD in R 2.14.0 (R Core Development Team).

The presence of food on a disk at the start of a trial could impact the measured 'preference' (pecks at and approaches toward) the coloured disks. However, this is unlikely to be the case because it would be expected to lead to one of two scenarios: (a) an individual could be attracted to the coloured disk only to consume the food pellet- this would result in pecks/approaches toward all colours to be roughly equal, (b) an individual could be attracted to the first colour with food it comes across- this would result in the individual associating with/pecking at the first colour it comes across and would be random across individuals. Neither of these two scenarios was observed; individuals had clear preferences and species were similar in preference across individuals.

3.4. Results

3.4.1. Dataset

At each of the five loci (four LWS and one RH2-1) sequenced in *P. formosa*, we identified two alleles. For each locus, one allele was more similar to *P. mexicana* and thus labeled the '-M' allele, and the other was more similar to *P. latipinna* and labeled the '-L' allele (for RH2-1, '-L' and *P. latipinna* are 100% identical). At every locus the *P. formosa* '-M' allele and the *P. formosa* '-L' allele differ from one another roughly as much as the *P. mexicana* allele differs from the *P. latipinna* allele (Table 3.1). Assuming equal rates of molecular change, this suggests that the *P. formosa* '-M' allele and the *P.*

formosa '-L' allele diverged at the same time as the *P. mexicana* allele and the *P. latipinna* allele, which is supported by our phylogenetic findings described below.

To assess functional opsin diversity we compared amino acid sequences between species for each locus. The RH2-1 amino acid sequences were identical between *P. formosa*, *P. latipinna*, and *P. mexicana*. However, the LWS loci demonstrated considerable amino acid diversity (Table 3.2). As expected, all variable parental opsin haplotypes were observed in the *P. formosa* LWS repertoire with the exception of novel amino acids at one position in LWS-1, 2 and 3, and two in LWS-R. The λ_{max} of the RH2/LWS classes of opsins are primarily determined by five specific amino acid changes that occur in the region of the protein spanning the membrane (Yokoyama and Radlwimmer 1998, 2001; Yokoyama *et al.* 2008). Of the novel amino acids identified in *P. formosa*, none were located at any of the five sites that determine the λ_{max} of an opsin; therefore, the alleles gained from each of the parental species likely behave identically to the alleles in the parental species.

3.4.2. Phylogenetic Analysis

Phylogenetic trees inferred using ML and Bayesian methods converged upon identical topologies for both the RH2-1 and LWS opsin datasets. Support for recovered nodes was robust for all topologies inferred under both ML and Bayesian methods with no well-supported nodes differing among analyses. The fully resolved phylogenies of RH2-1 and LWS opsins are presented in Figure 3.1. For each locus the '-M' and '-L' alleles of *P. formosa* group with the respective alleles from *P. mexicana* and *P. latipinna*.

3.4.3. Determination of Alleles Expressed in a Hybrid Species

For four of the loci (LWS-1, LWS-3, LWS-R, and RH2-1) alleles from both parental species were expressed in the eyes of *P. formosa*. We were unable to detect the presence of LWS-2 transcripts in the eyes of *P. formosa*. This is not surprising given that recent work has shown LWS-2 to be expressed at very low levels in the closely related *P. reticulata* (Ward *et al.* 2008; Laver and Taylor 2011; Sandkam and Breden unpublished data). The LWS genomic organization of *P. reticulata* is identical to *P.*

formosa, which likely leads to similar patterns of relative LWS expression. If *P. formosa* relative LWS expression is similar to *P. reticulata* then it is possible that LWS-2 alleles from both parents are being expressed at such low levels that we were unable to detect either of them. Based on the absence of introns, we were able to verify that each sequence originated from cDNA.

3.4.4. Behavioural Response to Coloured Disks

The proportions of pecks at and approaches toward the different coloured disks are presented in Figure 3.2. The number of pecks made towards the green disks significantly differed between *P. mexicana* and *P. latipinna* (p=0.0331), but did not differ between *P. formosa* and *P. latipinna* (p=0.1279), or *P. mexicana* and *P. formosa* (p=0.8023). These data indicate that *P. formosa* exhibits a level of attraction to green disks that is intermediate between that observed for the two parental species. The intermediate attraction of *P. formosa* was also observed in the number of approaches toward green. In contrast, *P. formosa* had a greater number of pecks toward yellow disks than *P. mexicana* (p=0.0334) but not *P. latipinna* (p=0.2595), and the parental species also did not differ (p=0.6310). There were no significant differences between any of the species in approaches or pecks toward orange or red.

3.5. Discussion

3.5.1. Strong Evidence of HSE in *Poecilia formosa*

The phylogenetic analyses and pairwise sequence comparisons presented here clearly demonstrate that *P. formosa* possess one allele for every RH2-1 and LWS locus from each parental species, and these alleles differ from one another in amino acid sequence at several sites for the LWS opsins. Our expression analysis revealed that *P. formosa* simultaneously express alleles from both parental species at four out of the five loci we investigated (with no expression of either parental allele identified for locus LWS-2).

Microspectrophotometry (MSP), a technique used to identify the peak spectral sensitivities (λ_{max}) of a cone cell, further supports our evidence that *P. formosa* have experienced HSE in the RH2-1 and LWS opsins. Körner and colleagues used MSP in *P. formosa*, *P. mexicana*, and *P. latipinna* to show that the λ_{max} values of the cone cells were virtually identical across all three species except for cones tuned to long wavelength light (*i.e.* RH2-1/LWS). The two cone types that differed in λ_{max} absorbance between *P. mexicana* and *P. latipinna* fell in the range of the RH2-1 and LWS opsin classes (*P. mexicana*- 536.6 ± 4.7 nm and 563.0 ± 3.0 nm; *P. latipinna* 551.2 ± 4.9 and 576.7 ± 5.9 nm) (Körner *et al.* 2006). Distinct spectral peaks were not able to be determined in this range for *P. formosa*, although an average peak was reported at 560.3 ± 16.1 nm. MSP is performed by directing a monochromatic light source at an individual cone cell and changing the wavelength while measuring the amount of light absorbed (Loew and Lythgoe 1978). Specific absorbance peaks for cone cells are determined by averaging groups of similar absorbances; this procedure makes assigning cone types to groups difficult when cones have similar λ_{max} values. Given that *P. formosa* have opsins that closely match the amino acid sequence of each of their parental species' opsins (Table 3.2), and that our expression data show they are all expressed, the inability to determine a spectral peak in *P. formosa* is most likely the result of producing all of the parental opsin types in their cone cells. Because the parental species have cones with λ_{max} values at 536.6 ± 4.7 and 563.0 ± 3.0 nm (*P. mexicana*) and 551.2 ± 5.9 and 576.7 ± 5.9 (*P. latipinna*) it is possible that the 560.3 ± 16.1 nm range in *P. formosa* is actually four cone types with λ_{max} values at 536, 551, 563, and 576 nm. If one is to consider all *P. mexicana* and *P. latipinna* RH2-1/LWS cone types at once, the distance between the λ_{max} plus the standard deviation of one cone and λ_{max} minus standard deviation of the next highest cone is smaller than the standard deviations themselves; 4.0, 3.1, 4.8 nm between 536.6, 551.2, 563.0, and 576.7 nm respectively. Therefore, if all cone types from the parental species were present in the same eye the resulting λ_{max} values would form a near-continuous distribution, which would be indistinguishable from one cone cell type with a λ_{max} that has a high standard deviation, such as that reported for *P. formosa*. While it is clear that the additional opsin alleles are being used, it is also possible that *P. formosa* are expressing two opsin alleles per cone cell (functional implications will be discussed below).

Another possibility is that alleles of either *P. latipinna* or *P. mexicana* are acting in a dominant/recessive manner. Having dominant alleles from one parental species would result in *P. formosa* possessing a sensory system that matches one of the parental species. The data generated from the MSP study by Körner *et al.* (2006) clearly show that *P. latipinna*, *P. mexicana*, and *P. formosa* have dramatically different visual systems in the range of the LWS opsins, and thus that the LWS opsins are not acting in a dominant/recessive manner. Therefore, the LWS opsin alleles from one locus are either expressed in different cone cells (as discussed above), or in the same cone cell, and are acting in a co-dominant fashion. Either way, the RH2-1/LWS opsin repertoire is clearly expanded in *P. formosa* as a result of the hybridization of *P. mexicana* and *P. latipinna*. Thus, our results provide strong support for the hybrid-sensory expansion (HSE) hypothesis.

3.5.2. What Does HSE Mean For *P. formosa* Vision?

Vision is the detection of different wavelengths of light, whereas colour vision is the discrimination of different wavelengths of light. Adding photoreceptor pigments can expand the detectable spectral window and/or expand the ability to discriminate wavelengths (*i.e.* colour discrimination) (Jacobs *et al.* 1999). Colour discrimination is accomplished by comparing the signals from cone cells with different λ_{max}. In normal colour vision each cone cell expresses only one opsin (Hagstrom *et al.* 2000), thereby maximally detecting only one wavelength of light. Because *P. formosa* are expressing two alleles at the same locus, it is possible that their cone cells have two wavelengths that they 'maximally' detect. Some New World primates are heterozygous for MWS/LWS opsin alleles but do not have this "problem" because their opsin locus is on the X chromosome, therefore only one allele is expressed per cell due to X-inactivation (Jacobs 1998). Mapping data and synteny comparisons of LWS opsin gene regions between the *P. reticulata* and *Oryzias latipes* genomes suggest that *P. reticulata* LWS opsin genes are not sex-linked (Tripathi *et al.* 2009; Watson *et al.* 2011). Given the close similarity between species in Poeciliidae (Breden *et al.* 1999; Hamilton 2001), it is likely that this is also true for *P. formosa*. However, it is not known whether multiple opsin genes are expressed in individual cone cells in fish retinas, or if some form of allelic exclusion exists.

Two studies by Jacobs *et al.* (1999, 2007) highlight possible implications of an expanded opsin repertoire. Jacobs *et al.* (2007) created a line of transgenic mice in which the cone cells either expressed an MWS opsin or an LWS opsin. The mice had an expanded range of wavelengths they could detect and were able to differentiate additional wavelengths of light. However, Jacobs *et al.* (1999) also created a line of transgenic mice with MWS and LWS opsins but had to co-express them in individual cone cells. The mice could detect an expanded range of wavelengths but were unable to differentiate wavelengths of light (*i.e.* see colour). It is unknown if hybrid sensory expansion in *P. formosa* has resulted in an increased ability to differentiate wavelengths or only an expanded visual range. Our behaviour experiments suggest that *P. formosa* still have the ability to differentiate wavelengths of light. The extent that *P. formosa* pecked at/approached painted disks differed by colour in a manner similar to the parental species which are homozygous at opsin loci (Figure 3.2). This pattern held except for attraction to green (*P. formosa* is intermediate to the parental species) and attraction to yellow (*P. formosa* is more attracted than either parental species). Thus, the combination of parental opsin loci through hybridization has led to functional differences in the tuning of the sensory system of *P. formosa*.

3.5.3. Unanticipated Pecking Behaviour

Using MSP, both *P. mexicana* and *P. latipinna* have been shown to possess two classes of cone cells that maximally detect light in the range of green, yellow, red, and orange. While *P. mexicana* cones detect wavelengths around 536 and 563 nm, cones of *P. latipinna* detect wavelengths around 551 and 576 nm (Körner *et al.* 2006). Both cone classes in *P. mexicana* maximally detect shorter wavelengths of light than those of *P. latipinna*. Because *P. mexicana* are better tuned to detect shorter wavelengths than *P. latipinna*, we hypothesized that *P. mexicana* would peck at coloured disks that reflect shorter wavelengths of light more than *P. latipinna.* The yellow and green coloured disks used in our behavioural trials reflect shorter wavelengths than the red and orange coloured disks (Rodd *et al.* 2002). Therefore, we predicted that *P. mexicana* would peck at green and yellow disks more than *P. latipinna,* while *P. formosa* would be intermediate to the two parental species because it has opsins from both. Surprisingly, we found that *P. mexicana* pecked at the green disks less than *P. latipinna,* while *P. formosa* behaved

as predicted (intermediate to the parental species). The unexpected switch in pecking behaviour between *P. mexicana* and *P. latipinna* is interesting and may be an outcome of a female preference for the green-blue colouring of large male sailfin mollies but further exploration by the field of behavioural ecology is needed to test this hypothesis. Despite the unanticipated behaviour of the parental species, it should be noted that *P. formosa* did behave as predicted; pecking the green disk at a proportion that was intermediate to the parental species. Clearly *P. formosa* has inherited opsin alleles from both *P. mexicana* and *P. latipinna,* which has influenced the sensory system and thus behaviour of this hybrid species.

3.5.4. Fitness Implications of HSE

Our data show that *P. formosa* have received the RH2-1 and LWS opsin repertoire of both parental species. This raises the question of whether or not this expanded repertoire has impacts on *P. formosa* fitness. There have been at least 840,000 generations since the hybridization event that lead to the formation of *P. formosa* (Stöck *et al.* 2010) and yet opsin amino acid sequences remain almost identical to their parental sequences (Table 3.2); this is possibly due to selection for the maintenance of functional opsins with the same λ_{max} as their parental species (although slow neutral evolution is still a possibility (Schartl *et al.* 1991)). We therefore suggest that tests of the HSE in systems where sensory stimuli influence fitness through foraging or mating decisions may prove useful for understanding the importance of hybridization in expanding sensory system repertoires (such as crickets (Shaw and Herlihy 2000), flies (Schwarz *et al.* 2005; Mallet 2007), cichlids (Carleton and Kocher 2001), and butterflies (Awata *et al.* 2009)).

3.6. Conclusion

Taken together our results clearly show that hybrid species can have a different sensory repertoire than either of their parental species, because they share sensory genetic variation with both parents. Bringing together sensory allelic variation from both parental species results in an expanded sensory system in one generation that is much faster than the more traditional models of duplication and divergence. We have termed

this phenomenon 'hybrid-sensory expansion' (HSE) and propose that it could help explain many differences between hybrids and their parental species. Providing hybrids with large sensory repertoires made up of alleles that have experienced selection in two different species could result in different outcomes depending upon the selective environment of the hybrid. Given a selective environment that is similar to one or both of the parental species, the hybrid may exhibit the same tuning of the sensory repertoire to either or both parental species. However, it is also possible that divergent selection could act on the allelic diversity within the hybrid to favor differentiation of the sensory system, resulting in different sensory alleles from the parental species.

In some cases, hybridization may thus be an important driver of the diversification of sensory systems via repertoire expansion and this process likely catalyzes sensory diversification among other organisms and in the context of other sensory systems.

Acknowledgements

We thank Dr. Margaret Ptacek for *Poecilia mexicana* specimens. We also thank Dr. Catherine Peichel, Dr. Ingo Schlupp and one anonymous reviewer for insightful comments. This work was funded by the Natural Sciences and Engineering Research Council (NSERC) of Canada and approved by the UACC of Simon Fraser University (protocol: 982B-06). We declare no conflicts of interest.

Data Accessibility

DNA sequences: GenBank Accessions JF823551 - JF823570

References

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

Avise, J. C., J. C. Trexler, J. Travis, and W. S. Nelson. 1991. *Poecilia mexicana* is the recent female parent of the unisexual fish *P. formosa*. *Evolution* 45:1530–1533.

Awata, H., M. Wakakuwa, and K. Arikawa. 2009. Evolution of color vision in pierid butterflies: blue opsin duplication, ommatidial heterogeneity and eye regionalization in *Colias erate*. *J. Comp. Physiol. A* 195:401–408.

Beukeboom, L. W., and R. C. Vrijenhoek. 1998. Evolutionary genetics and ecology of sperm-dependent parthenogenesis*. J. Evol. Biol*. 11:755–782.

Breden, F., M. Ptacek, M. Rashed, D. Taphorn, and C. Figueiredo. 1999. Molecular phylogeny of the live-bearing fish genus Poecilia (Cyprinodontiformes: Poeciliidae). *Mol. Phylogenet. Evol.* 12:95–104.

Brown, W.H. 1953. Introduced fish species of Guadalupe River Basin. *Texas Journal of Science* 5:245-251.

Carleton, K., and T. Kocher. 2001. Cone opsin genes of African cichlid fishes: Tuning spectral sensitivity by differential gene expression. *Mol. Biol. Evol.* 18:1540–1550.

Dowling, T., and C. Secor. 1997. The role of hybridization and introgression in the diversification of animals. *Annu. Rev. Ecol. Syst.* 28:593–619.

Endler, J. A., D. A. Westcott, J. R. Madden, and T. Robson. 2005. Animal visual systems and the evolution of color patterns: Sensory processing illuminates signal evolution. *Evolution* 59:1795–1818.

Fuller, R., K. Carleton, J. Fadool, T. C. Spady, and J. Travis. 2005. Genetic and environmental variation in the visual properties of bluefin killifish, *Lucania goodei*. *J. Evol. Biol*. 18:516–523.

Grether, G. F., G. R. Kolluru, F. H. Rodd, J. de la Cerda, and K. Shimazaki. 2005. Carotenoid availability affects the development of a colour-based mate preference and the sensory bias to which it is genetically linked. *Proc. R. Soc. B.* 272:2181–2188.

Hagstrom, S. A., M. Neitz, and J. Neitz. 2000. Cone pigment gene expression in individual photoreceptors and the chromatic topography of the retina. *J. Opt. Soc. Am. A* 17:527–537.

Hamilton, A. 2001. Phylogeny of Limia (Teleostei: Poeciliidae) based on NADH dehydrogenase subunit 2 sequences. *Mol. Phylogenet. Evol*. 19:277–289.

Hintze, J., and R. Nelson. 1998. Violin plots: A box plot-density trace synergism. *American Statistician* 52:181–184.

Hoffmann, M., N. Tripathi, S. R. Henz, A. K. Lindholm, D. Weigel, F. Breden, and C. Dreyer. 2007. Opsin gene duplication and diversification in the guppy, a model for sexual selection. *Proc. R. Soc*. B. 274:33–42.

Horth, L. 2007. Sensory genes and mate choice: Evidence that duplications, mutations, and adaptive evolution alter variation in mating cue genes and their receptors. *Genomics* 90:159–175.

Houde, A. E. 1987. Mate choice based upon naturally occurring color-pattern variation in a guppy population. *Evolution* 41:1–10.

Hubbs, C. L., and L. C. Hubbs. 1932. Apparent parthenogenesis in nature, in a form of fish of hybrid origin. *Science* 76:628–630.

Hubbs, C., R.J. Edwards, and G.P. Garrett. 1991. An annotated checklist of the freshwater fishes of Texas, with keys to identification of species. *Texas Journal of Science* 43:4 (Suppl.)

Jacobs, G. H. 1998. A perspective on color vision in platyrrhine monkeys. *Vision Res*. 38:3307–3313.

Jacobs, G. H., G. A. Williams, H. Cahill, and J. Nathans. 2007. Emergence of novel color vision in mice engineered to express a human cone photopigment. *Science* 315:1723–1725.

Jacobs, G. H., J. C. Fenwick, J. B. Calderone, and S. S. Deeb. 1999. Human cone pigment expressed in transgenic mice yields altered vision. *Journal Neurosci.* 19:3258–3265.

Kallman, K. D. 1962. Gynogenesis in the teleost, *Mollienesia formosa* (Girard), with a discussion of the detection of parthenogenesis in vertebrates by tissue transplantation. *J. Genetics* 58:7–24.

Katoh, K., G. Asimenos, and H. Toh. 2009. Multiple alignment of DNA sequences with MAFFT. *Methods in molecular biology* 537:39–64.

Körner, K., I. Schlupp, and M. Plath. 2006. Spectral sensitivity of mollies: comparing surface–and cave–dwelling Atlantic mollies, *Poecilia mexicana*. *J Fish Biol*. 69:54–65.

Lampert, K. P., and M. Schartl. 2008. The origin and evolution of a unisexual hybrid: *Poecilia formosa. Phil. Trans. R. Soc B* 363:2901–2909.

Laver, C. R., and J. S. Taylor. 2011. RT-qPCR reveals opsin gene upregulation associated with age and sex in guppies (*Poecilia reticulata*) - a species with color-based sexual selection and 11 visual-opsin genes. *BMC Evolutionary Biology* 11:81.

Loew, E. R., and J. N. Lythgoe. 1978. The ecology of cone pigments in teleost fishes. *Vision Res.* 18:715–722.

Mallet, J. 2007. Hybrid speciation. *Nature* 446:279–283.

Nylander, J. 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Center, Uppsala University.

Nylander, J., J. C. Wilgenbusch, D. L. Warren, and D. L. Swofford. 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* 24:581–583.

Rambaut, A. 1996. Se-Al: Sequence Alignment Editor http://evolve.zoo.ox.ac.uk.

Rambaut, A., and A. Drummond. 2007. Tracer v1. 4, Available from http://beast.bio.ed.ac.uk/Tracer.

Rennison, D. J., G. L. Owens, and J. S. Taylor. 2011. Opsin gene duplication and divergence in ray-finned fish. *Mol. Phylogenet. Evol.* 1–23.

Rodd, F. H., K. A. Hughes, G. F. Grether, and C. T. Baril. 2002. A possible non-sexual origin of mate preference: are male guppies mimicking fruit? *Proc. R. Soc. B.* 269:475–481.

Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.

Schartl, M., B. Wilde, I. Schlupp, and J. Parzefall. 1995. Evolutionary origin of a parthenoform, the Amazon molly *Poecilia formosa*, on the basis of a molecular genealogy. *Evolution* 49:827–835.

Schartl, M., I. Schlupp, A. Schartl, M. Meyer, I. Nanda, M. Schmid, J. Epplen, and J. Parzefall. 1991. On the stability of dispensable constituents of the eukaryotic genome - stability of coding sequences versus truly hypervariable sequences in a clonal vertebrate, the Amazon Molly, *Poecilia formosa*. *Proc. Natl. Acad. Sci. USA* 88:8759–8763.

Schlupp, I. 2005. The evolutionary ecology of gynogenesis. *Ann. Rev. of Ecol. Evol. Syst.* 36:399–417.

Schlupp, I., and R. Riesch. 2011. Evolution of unisexual reproduction. Pages 50–58 *in*. *Ecology and evolution of Poeciliid Fishes*. University of Chicago Press, Chicago, IL.

Schwanzara, S. 1967. The visual pigments of freshwater fishes. *Vision Res*. 7:121–148.

Schwarz, D., B. M. Matta, N. L. Shakir-Botteri, and B. A. McPheron. 2005. Host shift to an invasive plant triggers rapid animal hybrid speciation. *Nature* 436:546–549.

Shaw, K. L., and D. P. Herlihy. 2000. Acoustic preference functions and song variability in the Hawaiian cricket *Laupala cerasina*. *Proc. R. Soc. B.* 267:577–584.

Stöck, M., K. P. Lampert, D. Möller, I. Schlupp, and M. Schartl. 2010. Monophyletic origin of multiple clonal lineages in an asexual fish (*Poecilia formosa*). *Mol. Ecol.* 19:5204–5215.

Swofford, D. L. 2003. PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Team, R. D. C. (n.d.). R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Tripathi, N., M. Hoffmann, E.-M. Willing, C. Lanz, D. Weigel, and C. Dreyer. 2009. Genetic linkage map of the guppy, *Poecilia reticulata*, and quantitative trait loci analysis of male size and colour variation. *Proc. R. Soc. B*. 276:2195–2208.

Vrijenhoek, R. 1979. Factors affecting clonal diversity and coexistence. *American Zoologist.*

Ward, M. N., A. M. Churcher, K. J. Dick, C. R. J. Laver, G. L. Owens, M. D. Polack, P. R. Ward, F. Breden, and J. S. Taylor. 2008. The molecular basis of color vision in colorful fish: four long wave-sensitive (LWS) opsins in guppies (*Poecilia reticulata*) are defined by amino acid substitutions at key functional sites. *BMC Evolutionary Biology* 8:210.

Watson, C. T., K. P. Lubieniecki, E. Loew, W. S. Davidson, and F. Breden. 2010. Genomic organization of duplicated short wave-sensitive and long wave-sensitive opsin genes in the green swordtail, *Xiphophorus helleri*. *BMC Evolutionary Biology* 10:87.

Watson, C. T., S. M. Gray, M. Hoffmann, K. P. Lubieniecki, J. B. Joy, B. A. Sandkam, D. Weigel, E. Loew, C. Dreyer, W. S. Davidson, and F. Breden. 2011. Gene Duplication and Divergence of Long Wavelength-Sensitive Opsin Genes in the Guppy, *Poecilia reticulata*. *Journal Of Molecular Evolution* 72:240–252.

Windsor, D. J., and G. L. Owens. 2009. The opsin repertoire of *Jenynsia onca*: a new perspective on gene duplication and divergence in livebearers. *BMC research notes* 2:159.

Yokoyama, S. 2000. Molecular evolution of vertebrate visual pigments. *Progress in Retinal and Eye Research* 19:385–420.

Yokoyama, S. 2002. Molecular evolution of color vision in vertebrates. *Gene* 300:69–78.

Yokoyama, S., and F. B. Radlwimmer. 1998. The “five-sites” rule and the evolution of red and green color vision in mammals. *Mol. Biol. Evol*. 15:560–567.

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

Yokoyama, S., H. Yang, and W. T. Starmer. 2008. Molecular basis of spectral tuning in the red- and green-sensitive (M/LWS) pigments in vertebrates. *Genetics* 179:2037–2043.

Zhang, Z., S. Schwartz, L. Wagner, and W. Miller. 2000. A greedy algorithm for aligning DNA sequences. *J. Comp. Bio*. 7:203–214.

Boxes, Figures and Tables

Hybrid Sensory Expansion (HSE) Hypothesis

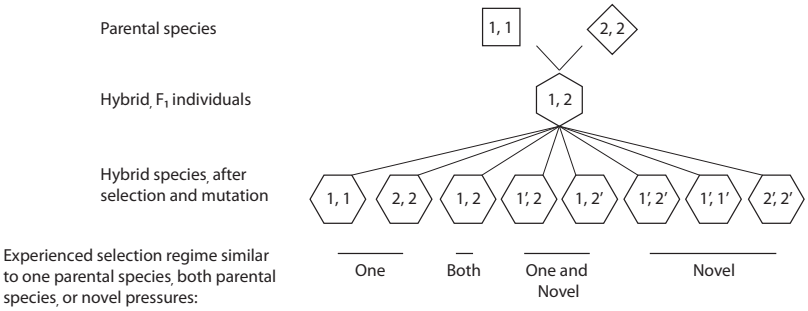
We propose that when a hybrid is formed between two parental species (P1 and P2) possessing functionally diverged sensory alleles (1 and 2) at the same autosomal locus; then the hybrid will be heterozygous, possessing the sensory alleles of both P1 and P2. Bringing together sensory alleles from P1 and P2 provides the hybrid with a combined sensory repertoire of both P1 and P2 (tested in this manuscript). While sensory expansion normally depends upon relatively slow evolutionary processes such as gene duplication and divergence, HSE occurs in a single generation. The expanded sensory repertoire may then be shaped by selection.

$$P1_{1,1} \times P2_{2,2}$$

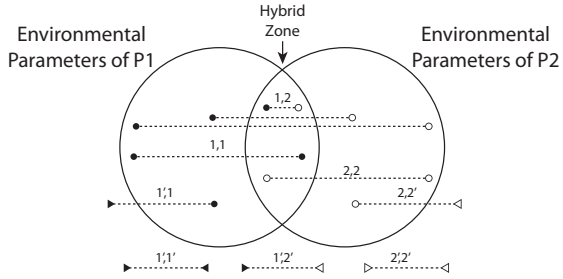
$$H1_{1,2}$$

Hybrid Sensory Expansion (HSE) Predictions

The resulting sensory repertoire of the hybrid can be shaped by selection from various aspects of life history including: sexual selection, prey detection, predator detection, etc. The resultant sensory repertoire could resemble the sensory system of one, both, or neither parental species if the selective regime experienced by the hybrid resembles those experienced by either or both parental species or is entirely or partially novel.



Sensory systems are the tools by which organisms perceive and interact with their environment. Changes to the sensory system can change the environmental parameters under which an organism can survive. Sensory alleles inherited from parental species P1 and P2 have been tuned to facilitate interaction with the environmental parameters of P1 and P2 respectively. First generation hybrids have sensory alleles from both parental species suggesting they could live in the environments of P1 and P2. Depending upon the direction of selection the set of environmental parameters available to the hybrid species could decrease, or shift to a new set of parameters. Selection working to maintain the full set of P1 and P2 alleles in the hybrid species would be observed as heterozygote advantage and could explain behavioral differences within populations of hybrid species as some individuals would be homozygous for each of the parental sensory alleles. Similarly, different populations of hybrids could experience selection pressures in different directions; matching P1 or P2.



Box 3.1. Hybrid Sensory Expansion (HSE): hypotheses and predictions.

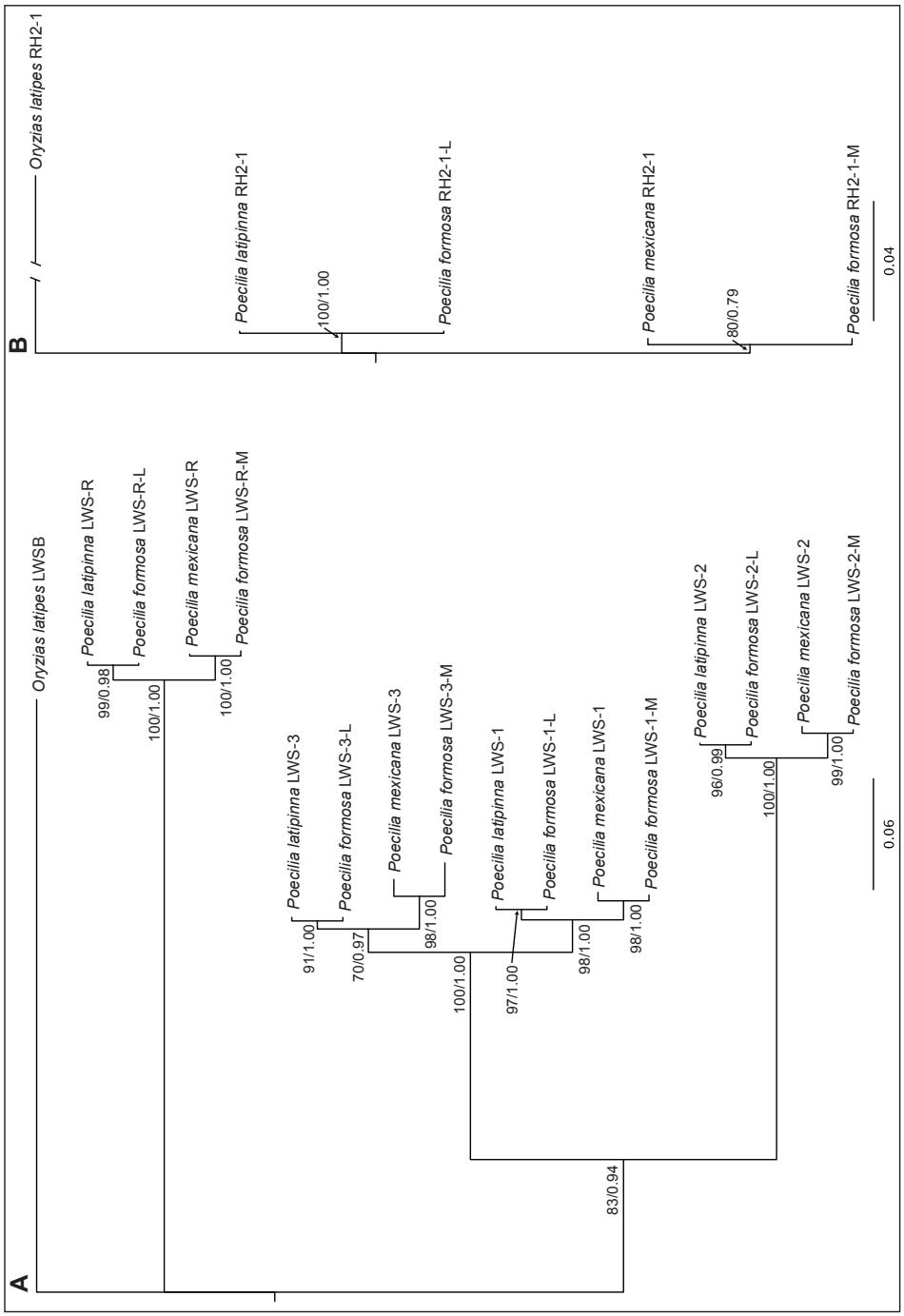


Figure 3.1. Bayesian consensus phylogenies of (A) long wavelength sensitive (LWS) opsin loci, and (B) RH2-1 opsin locus in *Poecilia mexicana*, *P. latipinna*, and the two *P. formosa* alleles. Maximum likelihood bootstraps and Bayesian posterior probability values (respectively) are reported for each node.

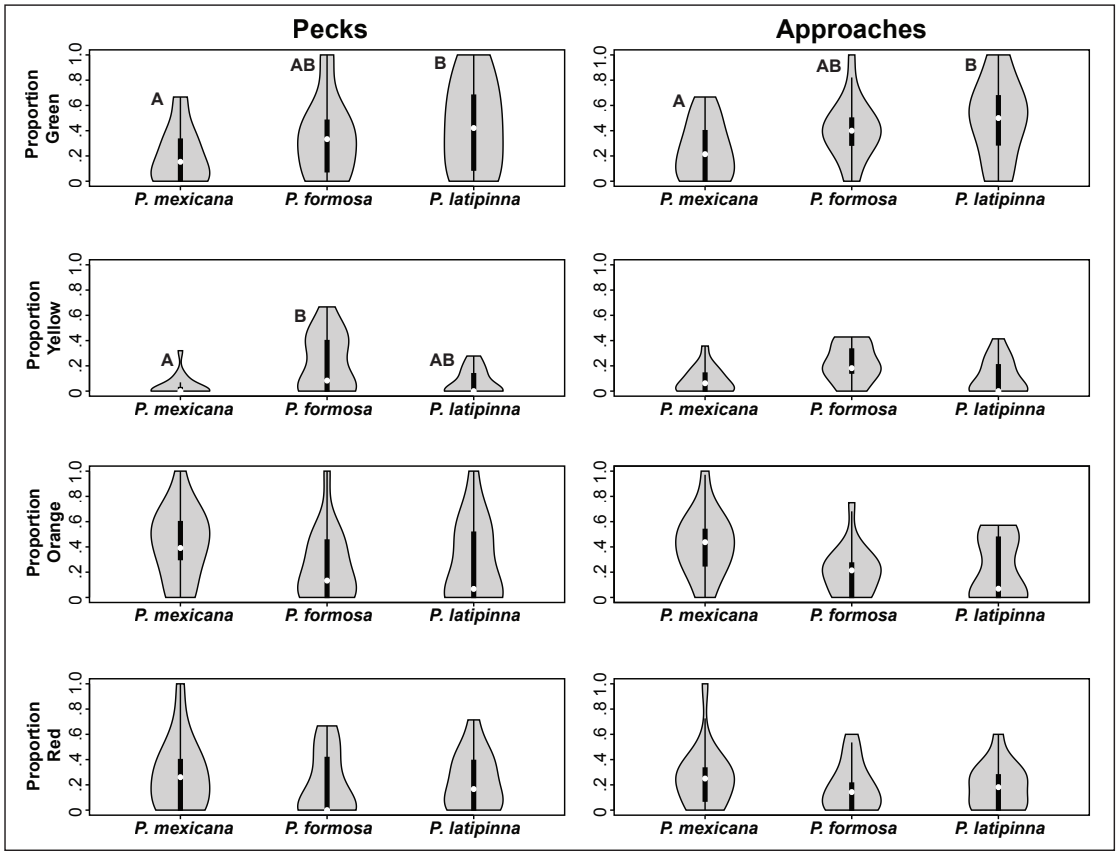


Figure 3.2. **Violin plots (Hintze and Nelson 1998) showing the proportion of pecks and approaches made toward each coloured disk for the hybrid *P. formosa* and the parental species *P. mexicana*, and *P. latipinna*. Violin plots contain a box plot (white circles denote mean) surrounded by a grey kernel density plot which shows the distribution of the data followed by smoothing to facilitate visual comparison across plots. Letters denote groups that differ significantly (Tukey's HSD, $p<0.05$).**

Table 3.1. **Comparison of the two *P. formosa* alleles and parental species for each LWS locus and RH2-1. Percent similarity is given above the diagonal. Number of gaps is given below the diagonal. Dark grey boxes show the most similar parental sequence for each allele from *P. formosa*.**

		Formosa-M	Formosa-L	Mexicana	Latipinna
LWS-1	Formosa-M	-	98.647	99.621	98.701
	Formosa-L	3	-	98.593	99.837
	Mexicana	0	3	-	98.647
	Latipinna	3	0	3	-
LWS-2	Formosa-M	-	97.772	98.859	97.799
	Formosa-L	25	-	98.138	99.532
	Mexicana	27	4	-	96.905
	Latipinna	25	12	50	-
LWS-3	Formosa-M	-	96.346	97.310	96.346
	Formosa-L	21	-	97.676	99.783
	Mexicana	16	5	-	97.676
	Latipinna	21	0	5	-
LWS-R	Formosa-M	-	98.685	99.862	98.616
	Formosa-L	2	-	98.824	99.931
	Mexicana	0	2	-	98.754
	Latipinna	2	0	2	-
RH2-1	Formosa-M	-	99.057	99.730	99.057
	Formosa-L	3	-	99.125	100.000
	Mexicana	1	4	-	99.125
	Latipinna	3	0	4	-

Table 3.2. **Amino acid differences in LWS loci across species. Amino acid numbering follows that of Yokoyama (2000). Transmembrane designations indicate whether the amino acid site is located in the transmembrane domain (Y-yes; N-no), where changes in amino acids are more likely to shift spectral absorbances. '-' indicates there was no variation in amino acids at that site for that locus. The amount of within species variation is unknown. Shading is provided to facilitate visual comparison.**

		Human amino acid number															
	Locus	11	20	26	33	55	65	131	132	139	153	188	194	233	247	298	310
P. mexicana	1	-	-	S	-	-	-	Y	I	-	-	V	-	-	-	-	-
P. formosa-M	1	-	-	S	-	-	-	F	I	-	-	V	-	-	-	-	-
P. formosa-L	1	-	-	A	-	-	-	F	T	-	-	I	-	-	-	-	-
P. latipinna	1	-	-	A	-	-	-	F	T	-	-	I	-	-	-	-	-
P. mexicana	2	-	-	-	-	A	I	-	-	-	-	-	-	-	-	A	L
P. formosa-M	2	-	-	-	-	A	I	-	-	-	-	-	-	-	-	G	L
P. formosa-L	2	-	-	-	-	V	V	-	-	-	-	-	-	-	-	G	F
P. latipinna	2	-	-	-	-	V	V	-	-	-	-	-	-	-	-	A	F
P. mexicana	3	Q	-	S	T	*L*	-	-	I	A	V	-	F	G	-	-	-
P. formosa-M	3	P	-	S	T	*L*	-	-	I	A	V	-	F	G	-	-	-
P. formosa-L	3	Q	-	A	S	I	-	-	T	T	I	-	Y	A	-	-	-
P. latipinna	3	Q	-	A	S	I	-	-	T	T	I	-	Y	A	-	-	-
P. mexicana	R	-	E	F	-	-	-	-	-	-	-	-	-	-	R	A	-
P. formosa-M	R	-	K	F	-	-	-	-	-	-	-	-	-	-	R	T	-
P. formosa-L	R	-	E	S	-	-	-	-	-	-	-	-	-	-	H	A	-
P. latipinna	R	-	E	S	-	-	-	-	-	-	-	-	-	-	H	A	-
Transmembrane		N	N	N	N	Y	Y	Y	Y	Y	N	Y	N	Y	N	N	Y

Table 3.3. Tukey's HSD comparing the hybrid *P. formosa* and parental species, *P. mexicana* and *P. latipinna*, in the number of pecks and approaches made toward coloured disks. Bold values indicate significance at p < 0.05.

				Pecks			Approaches		
				P-value	*F*-value	df	*P*-value	*F*-value	df
	P. latipinna	x	*P. formosa*	0.1279	3.625	2	0.8792	3.920	2
Green	*P. mexicana*	x	*P. formosa*	0.8023	3.625	2	0.0856	3.920	2
	P. mexicana	x	*P. latipinna*	**0.0331**	**3.625**	**2**	**0.0334**	**3.920**	**2**
	P. latipinna	x	*P. formosa*	0.2595	3.436	2	0.4062	1.860	2
Yellow	*P. mexicana*	x	*P. formosa*	**0.0334**	**3.436**	**2**	0.1565	1.860	2
	P. mexicana	x	*P. latipinna*	0.6310	3.436	2	0.8594	1.860	2
	P. latipinna	x	*P. formosa*	0.9969	1.323	2	0.5459	1.815	2
Orange	*P. mexicana*	x	*P. formosa*	0.3567	1.323	2	0.6578	1.815	2
	P. mexicana	x	*P. latipinna*	0.3411	1.323	2	0.1488	1.815	2
	P. latipinna	x	*P. formosa*	0.4646	0.757	2	0.9766	0.521	2
Red	*P. mexicana*	x	*P. formosa*	0.6756	0.757	2	0.7205	0.521	2
	P. mexicana	x	*P. latipinna*	0.9274	0.757	2	0.6075	0.521	2

Supplementary Tables

Supplementary Table 3.1. Primers used for creating PCR products and sequencing

Species	PCR Product	PCR Primers	Sequencing Primers
Poecilia formosa-L	LWS-1	01, 06	22, 19, 33, 35
Poecilia formosa-L	LWS-2A	26, 13	12, 16, 33, 35
Poecilia formosa-L	LWS-2B	13, 40	02, 03, 33, 35
Poecilia formosa-L	LWS-3	35, 01	17, 18, 33, 35
Poecilia formosa-L	LWS-R	37, 36	22, 33, 35
Poecilia formosa-L	RH2-1	31, 32	33, 34
Poecilia formosa-M	LWS-1	01, 06	22, 19, 33, 35
Poecilia formosa-M	LWS-2A	26, 13	12, 16, 33, 35
Poecilia formosa-M	LWS-2B	27, 28	02, 03, 33, 35
Poecilia formosa-M	LWS-3	29, 30	17, 18, 33, 35
Poecilia formosa-M	LWS-R	37, 36	22, 33, 35
Poecilia formosa-M	RH2-1	31, 32	33, 34
Poecilia latipinna	LWS-1	01, 06	09, 10, 13, 18, 19, 20, 22
Poecilia latipinna	LWS-2A	13, 26	05, 12, 13, 15, 26
Poecilia latipinna	LWS-2B	13, 40	02, 03, 08, 13, 20
Poecilia latipinna	LWS-3	01, 35	04, 07, 13, 17, 18, 20
Poecilia latipinna	LWS-R	36, 37	21, 22, 24, 25
Poecilia latipinna	RH2-1	31, 32	31, 32
Poecilia mexicana	LWS-1	01, 06	07, 09, 10, 11, 13, 20, 22
Poecilia mexicana	LWS-2A	13, 26	05, 10, 12, 13, 14, 15, 23, 26
Poecilia mexicana	LWS-2B	13, 32	02, 03, 08, 13, 20
Poecilia mexicana	LWS-3	01, 35	07, 10, 17, 18, 20
Poecilia mexicana	LWS-R	36, 37	21, 22, 24, 25
Poecilia mexicana	RH2-1	31, 32	31, 32

Primer #	Primer Sequence
01	TGTGAAGTGCAGATCACCTAG
02	GTGAAGACCCTGGAGTCCAG
03	TGGGTTGATTCACTATCTAGC
04	ACATCCAGAGTGTGGAGACG

05	AGGAAGATGACTCTGGCATTG
06	AGAATTTGTTTTCTCCCAGCC
07	TTATGCAGGAGCCACAGAGG
08	TCTTATCAGTCTTCACCAACGG
09	TTACTGTGTCTGCTGGGGACCCTAC
10	ATGTGGTAGTTTGGTCCTTC
11	AGGACGTTTTCAGTCCATGAGGC
12	ATACAGCTTGTCAGCAGAAG
13	GCAAAGACTGTCTCTCCAAGATC
14	AGATGTCCAGAAGACAGTCATG
15	GTCTCTATGCTATGCCACA
16	GGAACAACAGCCTTGAGAGGATTGTC
17	GACCCAGGAGAAAACTATTCCAGC
18	CATGACTACAACCATCCTGG
19	GCCCACCTGTCGGTTCATGAAG
20	CTTCATGAACCGACAGGTGGGC
21	CCCAGAAGGAAGTATCTAGGATG
22	ATCGCTCCTCGATGGGTTTAC
23	GTAAACCCATCGAGGAGCGAT
24	GGCAAGGTTGACTAAGATCCAGT
25	GGACACTTCAGAGCCGTCAT
26	CCCAGCAAAACTCTCAAGGT
27	CCTCGATGGGTTTATGATGC
28	CTAACAAGTTACAGTTACAGTCCC
29	CAGTTTCTGTACAGGTCACAGA
30	GAAGGTTATAAGCCAGAGGAAG
31	GAGTGTAGGATTGGCATTTACCA
32	TTGCACGCGTAACCAGATGA
33	TAATACGACTCACTATAGGG
34	CGCTGATTGTTTATTCAGGTGC
35	ATTAACCCTCACTAAAGGGA
36	TGCATTGACACATCCATGAGTTC
37	CCACAAGGTTACAAGCCAGAG

Supplementary Table 3.2. Primers used for creating cDNA products

Species	Allele	Primers
Poecilia formosa	LWS-1-M	38, 39
Poecilia formosa	LWS-2-M	40, 41
Poecilia formosa	LWS-3-M	42, 43
Poecilia formosa	LWS-R-M	44, 45
Poecilia formosa	RH2-1-M	46, 47
Poecilia formosa	LWS-1-L	48, 49
Poecilia formosa	LWS-2-L	50, 51
Poecilia formosa	LWS-3-L	52, 53
Poecilia formosa	LWS-R-L	54, 55
Poecilia formosa	RH2-1-L	56, 57

Primer #	Primer Sequence
38	GGACAATCATGTAGGACAGGACA
39	CTGAGGAACATTATTTCTAG
40	CCAAAACCGAGAACACTGACA
41	CCTCGATGGGTTTATGATGC
42	GTTTCCTCCTAAACAGAACATGAAGG
43	CGAAGACACAACAAGAGGTT
44	TTCCCAAAAAGCCGCATGATG
45	GAAGATCAAGCAGCTCAGATC
46	GCAGTTACGGAACTGTTTA
47	CTCCGCAGCTTTTGGCACTTT
48	GGACAATCATGTAGGACAGGACC
49	CTGAGGAACATTATTTCCAG
50	CCAAAACCGAGAACATTGACT
51	CCTCGATGGGTTTATGATGT
52	GTTTCTTCCTAAACAGAATATGAAGG
53	CGAAGATACAACAAGAGGCG
54	TTCCCAAAAAGCCGCATGATA
55	GAAGATCAAGCAGCTCAGATT
56	GCAGTTACGGAACTGTTTG
57	CTCCGCAGCTTTTGGCACTTC

Chapter 4.

Beauty in the eyes of the beholders: Colour vision is tuned to mate preference in the Trinidadian guppy (*Poecilia reticulata*)

Publication and Contributions

A version of this chapter is published as: Sandkam, B., Young, C. M. and Breden, F. (2015), Beauty in the eyes of the beholders: colour vision is tuned to mate preference in the Trinidadian guppy (*Poecilia reticulata*). Molecular Ecology, 24: 596–609. doi: 10.1111/mec.13058

Contributions: BAS and FB conceived and designed the project, conducted fieldwork, and wrote the manuscript; BAS designed assays and conducted all qPCR work; BAS and MY sequenced LWS-1 loci.

4.1. Abstract

A broad range of animals use visual signals to assess potential mates, and the theory of sensory exploitation suggests variation in visual systems drives mate preference variation due to sensory bias. Trinidadian guppies (*Poecilia reticulata*), a classic system for studies of the evolution of female mate choice, provide a unique opportunity to test this theory by looking for co-variation in visual tuning, light environment, and mate preferences. Female preference co-evolves with male colouration, such that guppy females from 'low predation' environments have stronger preferences for males with more orange/red colouration than do females from 'high predation' environments. Here we show that colour vision also varies across populations, with 'low' predation guppies investing more of their colour vision to detect red/orange

colouration. In independently colonized watersheds, guppies expressed higher levels of both LWS-1 and LWS-3 (the most abundant LWS opsins) in 'low predation' populations than 'high predation' populations at a time that corresponds to differences in cone cell abundance. We also observed that the frequency of a coding polymorphism differed between high and low predation populations. Together this shows that the variation underlying preference could be explained by simple changes in expression and coding of opsins, providing important candidate genes to investigate the genetic basis of variation in this model system.

4.2. Introduction

Understanding the forces driving population divergence is critical to understanding phenotypic evolution and ultimately speciation (Coyne & Orr 2004). Population divergence via sexual selection can occur rapidly when variation in mate preferences change across populations (Panhuis *et al.* 2001). Many hypotheses have been brought forth to explain the rapid evolution of female preference and male traits, including indirect models (such as Fisher's runaway and Good Genes), and direct models (such as sensory bias) (reviewed in Kuijper *et al.* 2012). Female preferences for male colouration occurs in a broad range of taxa from insects to mammals making the question of what shapes evolution of mating preferences fundamental to a wide range of fields including evolution, ecology and animal behaviour. One problem with evaluating the models for the evolution of female preference is that there are few systems in which variation in female preference exists across populations in a consistent replicated manner.

In the Trinidadian guppy (*Poecilia reticulata*), female mate preferences vary both within and between populations in response to environmental factors, making this system a premier model for evaluating the relative importance of the above models (Haskins *et al.* 1961; Endler 1980; Houde & Endler 1990; Brooks & Endler 2001; Hughes *et al.* 2013). High predation populations undergo strong forces of natural selection, which favor the evolution of males with drab colouration and weak (to non-existent) female preferences for colourful males. Alternatively, females from low predation populations have strong preferences for colourful males (usually reds/oranges) resulting in more

colourful males (Haskins *et al.* 1961; Endler 1978; Breden & Stoner 1987; Houde & Endler 1990; Endler & Houde 1995; Houde 1997; Magurran 2005). The watersheds of Trinidad were colonized independently, in each case from high to low predation regimes, thereby acting as naturally occurring experimental replicates (Reznick *et al.* 1996).

It has been strongly argued that female preferences for orange and red have evolved as a pleiotropic effect due to a pre-existing bias in guppies (following the Sensory Bias model of sexual selection) (Rodd *et al.* 2002). As highlighted by Endler and Basolo (1998) there are two models included in sensory bias that differ in their focus: in the sensory exploitation (SE) model, variation in the peripheral sensory system leads to variation in mate preferences; while in the sensory trap (ST) model mate preferences are due to neural responses that have evolved in a context outside of sexual selection. According to the SE model, variation in the tuning of peripheral sensory systems alone should be capable of driving changes in preferences for mate characteristics (Endler & Basolo 1998).

Mating decisions based on colour requires colour vision, which is largely accomplished by comparing the signals from cone cells in the retina, which are sensitive to different wavelengths of light (Gegenfurtner & Sharpe 1999). The wavelengths at which photopigments absorb light is largely determined by the transmembrane opsin protein expressed by the cone cell (Yokoyama 2000; 2002). While humans have 3 cone opsin proteins (and thus 3 cone cell types) guppies have an astounding 9 cone opsin proteins (Hoffmann *et al.* 2007; Ward *et al.* 2008; Watson *et al.* 2011). Each cone opsin protein is coded by a single gene which are grouped and named for the range of light they detect: short-wavelength sensitive (SWS1) detects ultraviolet; short-wavelength sensitive 2 (SWS2A and SWS2B) detect blues and purples; rhodopsin-like (RH2-1 and RH2-2) detect greens; and long-wavelength sensitive (LWS-1, LWS-2, LWS-3 and LWS-R) detect reds and oranges (Ward *et al.* 2008; Watson *et al.* 2011; Sandkam *et al.* 2012).

Differences in sensory tuning are traditionally assumed to occur via changes in gene sequences. However, differences in gene expression can also alter the tuning of sensory systems both across species (Cichlids- Carleton & Kocher 2001) and even

across populations within a species (Bluefin killifish- Fuller *et al.* 2004) (Reviewed in Horth 2007). Therefore both factors may be variable at the level that female preference changes in guppies. A large population genetics study of guppy opsin sequences in Trinidad by Tezuka *et al.* (2014) found variation in the frequency of polymorphisms known to impact colour vision only in LWS-1, where there are two alleles named for the amino acid at the position that corresponds to residue 180 in human M/LWS opsins: 180 Ala and 180 Ser. This same polymorphism occurs multiple times in the evolution of vertebrate opsins, and can result in a significant change in the tuning of an opsin pigment up to 15nm (Ward *et al.* 2008). Cone cell proportions in the retina are frequently estimated using opsin expression profiles, thereby opsin expression provides an excellent measure of the allocation of an individual's colour vision repertoire to a certain region of the spectrum (Fuller *et al.* 2003; 2004; Cheng & Flamarique 2004; 2007; Parry *et al.* 2005; Shand *et al.* 2008).

Here we assess the potential for the sensory exploitation model to explain female preferences in this model system by examining whether variation in visual tuning is associated with variation in mate choice at the population level. The parallel colonizations of watersheds each with high and low predation populations (with known differences in mate choice for reds and oranges) allows us to directly test if variation in the peripheral nervous system exhibits parallel variation in mate preferences by assessing the differences in allele frequency and expression of LWS opsins (responsible for detecting red/orange colouration). Using this framework we asked 'Do guppy populations vary in colour vision?' and pose the following hypothesis: If there is a causal relationship between visual tuning and mate choice, then low predation populations in independently colonized watersheds should show higher LWS expression and consistent changes in the frequency of genetic polymorphisms known to affect opsin sensitivity compared to high predation populations.

4.3. Materials and Methods

4.3.1. Sample Collection and Environmental Parameters

Poecilia reticulata were sampled from one high and one low predation population in the Aripo watershed, as well as one high and one low predation population in the Marianne watershed. These are all populations in the Northern Range Mountains of Trinidad where the Marianne watershed lies on the north slope while the Aripo watershed lies on the south slope. These watersheds were colonized independently and are thus a prime example of parallel evolution (Reznick & Bryga 1996). High and low predation populations were chosen from previously established study sites (see Supplementary Table 4.1 for GPS coordinates of each population)(Crispo *et al.* 2006). Previous work has shown female preference for males with more red/orange colouration in low predation populations of both the Aripo (Kemp *et al.* 2009) and Marianne (Endler & Houde 1995) watersheds. Each population was sampled at four time points: 07:30, 10:30, 13:30, and 16:30 (here-after referred to as "Time") on three days between July 3-14, 2011. Sampling consisted of catching 3 males and 3 females, which were rapidly sacrificed in an overdose of MS-222, measured, and photographed (totaling 288 individuals). Eyes were immediately removed and a small puncture was made to facilitate complete penetration of RNA*later*® Stabilization Solution (Life Technologies™). Both eyes from an individual were placed into a vial of RNA*later*® and kept at 10°C for 24 hours, to allow tissue to be saturated per manufacturer's recommendation. After 24 hours, the vial was transferred to liquid nitrogen. The vials were removed from liquid nitrogen just prior to being placed in checked baggage and flown to Simon Fraser University where they were placed in a -20°C freezer until RNA extraction. Time spent at room temperature totaled less than 24 hours and fell well under the 1-week maximum suggested by manufacturer. The bodies of individuals sampled were placed in tubes of 95% EtOH buffered with EDTA and kept at -20ºC until DNA extraction.

During the 10:30 collection the absolute irradiance of the light environment was measured using a Jaz Spectrophotometer fitted with a 600 um fiber cable and cosine corrector running the JazIrrad Program (Ocean Optics). This configuration measured the intensity of light arriving at the detector over a 180º field of view at 0.33 nm intervals

from 250-800 nm. The detector was placed 15cm below the surface and faced up to measure the absolute irradiance of the down-welling light. Guppies are usually found less than 15cm deep and therefore these measures of light environment are representative of the light experienced by guppies. To compare light environments across populations the λp50 was determined for each measure taken. The λp50 is the median wavelength of photons between 300-800 nm (McFarland & Munz 1975; Hortado-Gonzales *et al.* 2014). A larger λp50 is indicative of a light environment that is shifted toward longer wavelengths. Absolute irradiance measures are subject to variation in environmental parameters across days, such as cloud cover, therefore measures from the identical spot were taken on all 3 days of sampling at the same time to incorporate natural variation. On each day of sampling, environmental parameters were measured of the population using a DO700 Waterproof Dissolved Oxygen kit (ExTech®). Parameters measured included dissolved oxygen (DO), water temperature, conductivity, pH, Total Dissolved Solids (TDS), and salinity. Measures of λp50, DO, water temperature, conductivity, pH, TDS and salinity were compared between high and low predation populations within watersheds using independent Welch Two Sample t-tests in R v3.0.2 (R Core Team 2014) with the 3 days acting as replicates for each population.

4.3.2. qPCR Assay Design

Opsin expression profiles are a well established methodology by which to measure an organism's visual system, as differences in opsin expression profiles in the mid-morning correlate with differences in the ratio of cone cells expressing those opsins in fish and humans (Hagstrom *et al.* 2000; Fuller *et al.* 2004; Cheng & Flamarique 2004). Probe based PrimeTime® qPCR assays (IDT® Technologies) were designed to be specific for each of the 9 opsins, 1 rhodopsin (RH1), and 3 housekeeping genes (beta actin (B-actin); cytochrome c oxidase subunit I (COI); myosin heavy chain (Myosin HC)) such that a primer or probe spanned intron-exon boundaries whenever possible. Assays consisted of a forward primer, reverse primer and 5' FAM labeled probe with both 3' Iowa Black® and internal ZEN™ quenchers (IDT® Technologies). Assay specificity was verified by the presence of a single band when running PCR products on an agarose gel. LWS-1 and LWS-R assays resulted in products of the same size, while LWS-1 and LWS-3 loci are similar in sequence. To ensure LWS-1, LWS-3, and LWS-R were truly

locus specific assays, we measured the pairwise covariance of these 3 assays on the final relative$_{(hk)}$ data set (described below) using R v3.0.2 (R Core Team 2014). If assays were binding to non-specific targets we would expect to see positive covariance between assays. We found no positive covariance between either LWS-1 and LWS-R (-0.0284) or LWS-1 and LWS-3 (-6.5021), demonstrating locus specificity of the LWS assays.

Three gBlocks® Gene Fragments (synthetic double stranded, sequence-verified genomic blocks made by IDT® Technologies) were built to determine PCR efficiency (E_i) for each assay. The gene fragments contained sequence for each of the genes being assayed from 10 bp upstream of the forward primer to 10 bp downstream of the reverse primer. The order of genes was randomized such that gBlock® 1 contained LWS-1, RH2-1, SWS2A; gBlock® 2 contained SWS1, LWS-2, RH2-2, SWS2B; and gBlock® 3 contained RH1, LWS-3, LWS-R. The 3 gBlocks® were mixed in equal proportions and brought to a concentration of 0.001 ng/µl resulting in a control with equal ratios of all the opsin and housekeeping genes. Six replicates of each assay were run using 4.5 µl of the control. The relative primer efficiencies (E_i) were then calculated following Carleton and Kocher (2001) using the equation:

$$\frac{(1+1)^{Ct_{High}}}{(1+E_i)^{Ct_i}} = 1$$

The critical threshold for the opsin with the highest expression (lowest *Ct* value) is Ct_{High} and Ct_i is the critical threshold for opsin *i*. Running six replicates resulted in a mean relative efficiency and standard error for each of the assays (Supplementary Table 4.2).

As COI had the highest relative efficiency it was used to measure absolute efficiency. A ten fold serial dilution was made of a randomly chosen fish from those sampled in Trinidad. qPCR was performed on five replicates of each concentration using the COI assay. The absolute efficiency of COI was found using the slope of *ln*(concentration) plotted on *Ct* such that $\text{absolute } E_{COI} = e^{-slope} - 1$. The absolute efficiencies of the other primer/probes were determined following Fuller *et al.* (2004) using the equation:

$$\text{absolute } E_i = \left(\text{relative } E_i \text{ x absolute } E_{COI}\right) / \text{relative } E_{COI}$$

4.3.3. Sample Processing and Analyses

Both eyes from one individual were placed in 600 uL of TRIzol® reagent (Life Technologies™) and ground in a 1.5 mL RNase-free Kontes® Pellet Pestle Grinder (Kimble Chase). Solution was then run through an Ambion® Homogenizer (Life Technologies™) on a Beckman Coulter Microfuge® 18 microcentrifuge for 8 minutes at 12,000 rpm to reduce viscosity. RNA was extracted following the manufacturer's instructions using PureLink® RNA Mini Kit with the addition of an on column treatment with PureLink® DNase (Life Technologies™) to eliminate any potential genomic contamination during qPCR.

RNA yields were quantified by measuring each sample 3 times using a NanoDrop 2000c spectrophotometer and averaging across the 3 measures. Concentrations of RNA were adjusted to 50 ng/uL using UltraPure™ DNase/RNase-Free Distilled Water (Life Technologies™). For each sample 500 ng RNA was reverse transcribed using a High Capacity cDNA Reverse Transcription Kit with RNase Inhibitor (Life Technologies™) following the manufacturer's instructions. cDNA yields were quantified by measuring each sample 3 times using a NanoDrop 2000c spectrophotometer and averaging across the 3 measures. Concentrations of cDNA were adjusted to 11 ng/uL using UltraPure™ DNase/RNase-Free Distilled Water (Life Technologies™) for use in qPCR reactions.

For each individual the 9 opsins, 1 rhodopsin, and 3 housekeeping genes were quantified using the qPCR probe based assays described above, each reaction was run in triplicate. All 39 reactions for each individual were run simultaneously on the same 384 well plate in addition to negative controls (UltraPure™ water) for each assay. Each 10 uL reaction consisted of: 5 uL Brilliant III Ultra-Fast qPCR Master Mix (Agilent Technologies), 0.5 uL FAM labeled assay (described above) and 4.5 uL sample. All reactions were set up on ice and the plates were briefly spun down before being run on an Applied Biosystems® 7900HT qPCR machine (Life Technologies™). PCR conditions were as follows: 95°C for 3:00 followed by 40 cycles of 95°C for 0:05, 60°C for 0:15.

Since each assay was done in triplicate, the standard deviation was taken and when >2, outliers were removed. One of the 288 individuals had all poor data and was removed from further analyses. To assess differences in colour vision, the proportion of total opsin expression (T_{all}) made up of each opsin (T_i) was calculated following Fuller *et al.* (2004) and Carleton and Kocher (2001) with the following equation:

$$\frac{T_i}{T_{all}} = \frac{\left(1 \Big/ \left((1+E_i)^{Ct_i}\right)\right)}{\sum\left(1 \Big/ \left((1+E_i)^{Ct_i}\right)\right)}$$

E_i is the mean primer/probe efficiency for assay *i* (described above) and Ct_i is the critical cycle number for gene *i* (such that expression of the 9 opsins add to 1 for each individual).

Each of the 9 opsins as well as the 1 rhodopsin were also compared to the average of the 3 housekeeping genes (T_{House}) to determine differential regulation of the genes. Using the mean of 3 housekeeping genes decreases the potential for random variation in housekeeping gene expression to lead to errors (Vandesompele *et al.* 2002). The level of expression for each opsin and rhodopsin (T_i) was calculated based on the following equation modified from Fuller and Claricoates (2011):

$$\frac{T_i}{\overline{T}_{House}} = \frac{\left(1 \Big/ (1+E_i)^{Ct_i}\right)}{\left(\left(\sum\left(1 \Big/ (1+E_{House})^{Ct_{House}}\right)\right) \Big/ 3\right)}$$

E_{House} is the primer/probe efficiency for a housekeeping gene and Ct_{House} is the critical cycle number for that gene.

The above two methods resulted in each individual having two measures for each opsin: proportional opsin expression (proportional), and relative to housekeeping genes ($relative_{(hk)}$). Proportional measures of opsin expression provide a measure of colour vision, as colour vision is accomplished by comparing the signal from different cone cell types and opsins are the major differentiating character of cone cell types. Proportional opsin expression acts as a measure of the ratio of the different cone cell

types in the back of the retina (Fuller *et al.* 2004; Fuller & Claricoates 2011). Meanwhile relative$_{(hk)}$ measures of opsin expression reveal which opsins are differentially regulated (Fuller & Claricoates 2011), compared to a measure of overall gene activity, as evidenced by the mean of 3 housekeeping genes.

4.3.4. LWS-1 A/S Allele Frequency

A DNeasy blood and tissue kit (Qiagen) was used to extract genomic DNA from tail tissue of the same individuals used for expression analyses. Primers specific to the 5' and 3' UTR region of LWS-1 were used to generate PCR products of the LWS-1 locus. A 438 bp region including part of exon 2 and all of exon 3 was directly sequenced using internal sequencing primers by Molecular Cloning Laboratories (McLab, San Francisco, CA) (Supplemental Table 4.3). Sequence chromatograms were viewed and analyzed using SeqMan Pro (Lasergene 8.0; DNASTAR, Madison, WI). Individuals were identified to have either a Serine, Alanine or to be heterozygous at the 180 amino acid residue (no other variable sites were identified that would change wavelength sensitivity of LWS-1 (Tezuka *et al.* 2014)). The frequency of the three genotypes was determined in each population (sample sizes: Aripo watershed- high predation N= 54, low predation N=48; Marianne watershed- high predation N=49, low predation N= 46). F_{ST} based on the frequency of alleles with a Serine at amino acid position 180 were calculated between high and low predation for each watershed using:

$$F_{ST} = \frac{\mathrm{var}(S)}{\overline{S} * (1 - \overline{S})}$$

Where $\mathrm{var}(S)$ is the variance of the frequency of the Serine allele across high and low predation populations within the same watershed and $\overline{S}$ is the frequency of the Serine allele in the watershed.

4.3.5. Statistical Analyses of Opsin Expression

All expression data were analyzed using R v3.0.2 (R Core Team 2014). Time, sex, predation, and watershed were treated as fixed effects in all models with predation

nested in watershed. In all analyses time was treated as a categorical rather than continuous variable.

Do populations differ in colour vision? -

The overall effects of time, predation, watershed and sex on both proportional and relative$_{(hk)}$ opsin expression profiles was assessed using MANOVAs, with the levels for each of the 9 cone opsins log transformed and treated as dependent variables. Sex had no effect and was not involved in any significant interactions so was removed from the model. Significant interactions were followed up on individual watersheds where we examined the effects of time and predation.

Do low predation populations express higher levels of LWS opsins? -

Individual ANOVAs were conducted for each LWS opsin (both proportional and relative$_{(hk)}$ measures) to describe the effects of time, predation, sex, and watershed, because of the known importance of male orange and red colour patterns in these populations. Sex again had no effect nor was it involved in any significant interactions for any opsin and was removed from the models. Significant interactions were followed up with ANOVAs on data subsets by time, examining the effects of watershed and predation. Any significant watershed*predation interaction was further examined on subset data by watershed to examine the effects of predation.

4.4. Results

4.4.1. Guppy Opsin Expression Differs Across Populations

Across all populations SWS2A, LWS-2 and LWS-R showed lower levels of expression compared to other opsin genes (Figures 4.1, 4.2, and Supplementary Figures 4.1, 4.2). The most abundant opsin expressed varied across populations with RH2-1 being most highly expressed in Marianne low predation and Aripo high predation while SWS2B was the highest expressed in Marianne high predation and RH2-2 was highest expressed in Aripo low predation. MANOVAs of full opsin expression profiles

showed significant effects of predation (nested in watershed), watershed, time and every interaction therein for both proportional and relative$_{(hk)}$ measures (Table 4.1). Sex did not impact either proportional ($F_{1,271}$ = 0.7565, P = 0.6568) or relative$_{(hk)}$ ($F_{1,271}$ = 0.6815, P = 0.7255) measures of opsin expression profiles, nor was sex involved in any significant interactions with predation (nested in watershed), watershed or time. There were highly significant differences in opsin expression by predation, time and predation*time within both the Aripo and Marianne watersheds (Table 4.2). This shows that guppy colour vision, as measured by differences in opsin expression, does vary across populations.

4.4.2. Low Predation Populations Express Higher Levels of LWS Opsins

The ANOVAs of proportional and relative$_{(hk)}$ measures for LWS opsin expression revealed significant interactions between predation nested in watershed and time (Table 4.3; all other opsins presented in Supplementary Tables 4.4 and 4.5). To tease apart the effects of predation at each time point both datasets were subset by time point, and ANOVAs revealed significant effect of predation nested within watershed for many time points (Supplementary Tables 4.6 and 4.7). To tease apart the effect of predation in each watershed at the four time points, the time point data were subset by watershed and ANOVAs were run by predation; results for LWS are shown in Figures 4.1 and 4.2 (ANOVA results for all opsins are in presented in Supplementary Tables 4.8 and 4.9). When proportional LWS opsin expression differed within watershed by predation (Figure 4.1 and Supplementary Table 4.8), expression levels were higher in the low predation population, except for LWS-1 and LWS-2 at the end of the day in the Marianne watershed (Figure 4.1). This is likely due to the dramatic increase in LWS-3 expression experienced by this population at this timepoint (as can be seen in relative$_{(hk)}$ measures, Figure 4.2) which would result in relative expression of other opsins making up a smaller proportion of total opsin expression. When relative$_{(hk)}$ LWS opsin expression differed by predation within the Marianne watershed the low predation population had higher expression except for LWS-2 at time point 1 (Figure 4.2 and Supplementary Table 4.9); however the low level of LWS-2 expression in the low predation population at this time point may not be biologically relevant as it is indicative of very few cone cells expressing this gene and could have very minor contributions to colour detection. The higher

expression in both proportional and relative$_{(hk)}$ measures of LWS-1 and LWS-3 expression of low predation populations was accompanied by a decrease in expression of SWS2B (Figures 4.1, 4.2, 4.3; Supplementary Figure 4.2). The relative opsin expression of all opsins at time point 2 is shown in Figure 4.3.

4.4.3. Gene Frequencies of the 180 Ala Versus Ser Allele of LWS-1

Gene frequencies differed between high and low predation populations in both watersheds but all populations were found to follow Hardy Weinberg Equilibrium following independent chi-squared tests. The allele frequencies of the high predation populations were S- 62.37% A- 36.73% in the Marianne and S- 66.67% A- 33.33% in the Aripo, while low predation populations were S- 100% A- 0% in the Marianne and S- 97.92% A- 2.08% in the Aripo (Figure 4.4). F_{ST} between high and low predation populations was 0.3221 in the Aripo and 0.4393 in the Marianne watersheds. Sequences are available on GenBank under accession numbers: KP218062 – KP218254.

4.4.4. Environmental Parameters

We did not detect any differences in environmental parameters that likely drove differences in LWS opsin expression between high and low predation populations within watersheds (for raw data see Supplementary Table 4.10). The λp50 did not significantly differ between high and low predation populations in either the Marianne (t = 1.1373, P = 0.3542) or the Aripo (t = -2.6963, P = 0.0767) watersheds. Since increases to the λp50 is a measure of how far the available light is shifted toward longer wavelengths (Hurtado-Gonzales *et al.* 2014) we would expect to see significant within watershed differences between high and low predation populations if differences in LWS opsin expression were being driven by plastic responses to environmental differences in available light. Dissolved oxygen has been used as a proxy for differences in lighting environment elsewhere (Tezuka *et al.* 2014) yet for this environmental variable we still found no differences between high and low predation populations in either the Aripo (t = -0.0354, P = 0.9736) or the Marianne (t = -1.7165, P = 0.1954) watersheds. We also found no significant differences between high and low predation populations within watershed for temperature (Aripo – t = -1.2522, P = 0.2817 ; Marianne – t = 0.3288, P = 0.7629),

conductivity (Aripo – t = 4.0328, *P* = 0.0529 ; Marianne – t = 1.4242, *P* = 0.2332), pH (Aripo – t = -0.5048, *P* = 0.6593 ; Marianne – t = 1.8954, *P* = 0.1311), or total dissolved solids (Aripo – t = -0.3474, *P* = 0.7602 ; Marianne – t = 1.4661, *P* = 0.2218). However, there was a significant difference in salinity between the high and low predation populations of the Aripo watershed (t = 3.8891, *P* = 0.0399) but not in the Marianne watershed (t = 1.4444, *P* = 0.2302). As such no environmental parameters corresponded with differences in LWS opsin expression observed between high and low predation populations within both watersheds.

4.5. Discussion

Using guppies, a classic model system for studies of evolution by sexual selection, we found support for the sensory exploitation model to explain population divergence in mate preferences. This is the first study of opsin expression in natural populations of guppies and coupled with the allele frequency data we demonstrate that guppy vision differs across populations in a consistent manner. Proportionally, guppies in low predation populations express more long wavelength sensitive (LWS) opsins than do guppies from high predation populations, and low predation populations are almost fixed for an alternative amino acid known to affect retinal tuning. This suggests that much of the variation in female preferences in guppies could be explained by variation in the visual system, supporting the sensory exploitation model. This model for the evolution of mate preference posits that variation in the peripheral sensory systems can lead to variation in mate preferences (Endler and Basolo 1998). Thus, more broadly, support of the sensory exploitation model furthers our understanding of the mechanisms of mate choice that have the potential to lead to speciation (Panhuis *et al.* 2001; Coyne & Orr 2004).

4.5.1. Guppy Colour Vision Differs Across Populations

Variation in colour vision can be due to differences in the amino acid sequences of the opsin proteins, ratio of cone cells expressing different opsins, or neural processing of signals from the eyes. Here we show guppy populations differ both in opsin expression profiles and the frequency of alleles at the LWS-1 locus. Importantly, both

relative$_{(hk)}$ (measuring gene regulation) and proportional (measuring colour vision) guppy opsin expression profiles differ across watersheds but also by predation within each watershed (Table 4.2). Differences in relative$_{(hk)}$ expression suggest the regulation of opsin genes differ across populations, whereas differences in proportional expression suggest colour vision in guppies is not a fixed phenotype as has been assumed when modeling population differences in mate choice and colouration (such as Endler 1991 and Kemp *et al.* 2009). Specifically, the differences we found in proportional expression profiles at midmorning demonstrate that the ratio of cone cell types in the retina differ across populations. While implications of the diurnal variation in opsin expression across populations are less clear, it supports the conclusion that visual tuning differs dramatically across populations.

Further evidence of differences in guppy visual systems comes from the variation in allele frequencies seen across populations. The wavelength of light to which an opsin is maximally tuned, λ_{max}, provides a good estimate for how sequence level changes in the opsin genes do or do not impact the visual system (Yokoyama 2008). Determination of λ_{max} for the LWS opsins in guppies using protein expression systems has proven difficult (B. Chang (University of Toronto, Canada) personal communication). However, Tezuka *et al.* (2014) used the 'five site rule' (Yokoyama & Radlwimmer 2001) to identify two alleles of LWS-1 that differ in λ_{max} by 7nm (LWS-1 (180 Ala) ~553nm, LWS-1 (180 Ser) ~560nm), and estimated the following λ_{max} values for the other loci: LWS-2 ~518nm, LWS-3 ~560nm, and LWS-R ~560nm.

Tezuka *et al.* showed differences in frequency of the 180 Ala and 180 Ser alleles of LWS-1 across populations, such that most of the geographic variation occurred between streams in northern and southern parts of the island. However, Tezuka *et al.* also measured one Upper Aripo (equivalent to low predation Aripo in this study) and one Lower Aripo (equivalent to high predation Aripo in this study) population and the frequencies of the LWS-1 180 Ala/Ser alleles were very similar to those presented here. The frequency of the 180 Ser allele in the Lower Aripo (Tezuka) - 0.82 (N=45) compared to high predation (this study) - 0.73 (N=54), while the Upper Aripo (Tezuka) - 1.00 (N=33) compared to low predation (this study) - 0.96 (N=48). This demonstrates that low predation populations have a higher frequency of the longer wavelength LWS allele. The

F_{ST} based on LWS-1 between high and low predation populations in the Aripo was similar between Tezuka *et al.* (0.492) and this study (0.3221). Both these values are greater than the average F_{ST} estimate based on 7 non-opsin genes by Tezuka *et al.* for upstream/downstream pairs in the Aripo (0.070) and Guanapo (0.101) watersheds. The discrepancy in estimates of population divergence between those based on LWS-1 and those based on non-opsin genes is in line with the high levels of selection acting on LWS-1 reported by Tezuka *et al.* which corresponded with DO levels in their study. We found no differences in either DO or λp50, however this could be a seasonal effect of wet season driving high levels of DO across populations. It would be interesting to see if the other differences in allele frequency detected across watersheds by Tezuka *et al.* correspond with differences in predation level and/or mate preferences.

Classically, visual systems are assumed to be the same across different populations of the same species. However, our work adds guppies to the growing number of species that have recently been shown to vary in visual systems across populations due to differences in either opsin expression (*e.g.* Bluefin killifish (Fuller *et al.* 2004), cichlids (Smith *et al.* 2010), stickleback (Flamarique *et al.* 2013)) or sequence variation (*e.g.* Sand goby (Larmuseau *et al.* 2009), pied flycatchers (Lehtonen *et al.* 2012)). In consideration of this growing body of work we suggest care be taken when making assumptions that dismiss intra-species sensory system variation, such as when modeling quantum catch of available light.

4.5.2. Low Predation Populations Express Higher Levels of LWS Opsins

In most guppy populations, there is a trade off between natural selection against colourful males and sexual selection for colourful males, especially with respect to red/orange colouration (Endler 1980; Breden & Stoner 1987; Endler & Houde 1995; Lindholm *et al.* 2014). Guppy populations under low predation regimes evolve to invest more in female choice with increased male displays and stronger female preferences (Breden & Stoner 1987; Houde & Endler 1990; Endler & Houde 1995; Reznick *et al.* 2001; Lindholm *et al.* 2014). The increase in female preference for colourful males and decrease in predation pressure results in greater red/orange male colouration in low

than high predation populations (Haskins *et al.* 1961; Endler 1980). Furthermore transplant experiments have shown the evolution of female choice and male colouration can occur rapidly with changes to predation environments (Endler 1980; Reznick & Bryga 1987; O'Steen *et al.* 2002).

LWS opsins are the proteins responsible for detecting and discriminating red/orange colouration. We found greater LWS expression in low compared to high predation populations for both Aripo and Marianne watersheds. These watersheds were colonized independently and are thus a prime example of parallel evolution (Reznick & Bryga 1996). Therefore the greater expression of LWS opsins by low predation populations likely plays a major role in determining level of female preferences for the red/orange colouration of potential mates.

The differences in opsin expression profiles across populations could be due to either plastic response to differences in environmental conditions or heritable differences in opsin regulation. If expression differences in guppies between low and high predation are indeed due to plastic responses to lighting environment one would expect to see differences in the λp50 of the populations since the λp50 is a measure of whether a light environment is shifted to shorter or longer wavelengths. However, we did not find any differences in this parameter. The lack of lighting differences is not unprecedented as Smith *et al.* (2010) found significant differences in opsin expression across intra-specific populations of cichlids with no differences in lighting environment. Additionally Endler *et al.* (2001) found visual sensitivity to be heritable in guppies in response to artificial selection, and many other species have a strong heritable component to opsin expression (e.g. Bluefin killifish (Fuller *et al.* 2005), three spined stickleback (Flamarique *et al.* 2013), cichlids (Carleton & Kocher 2001; Hofmann & Carleton 2009)). This emphasizes the possibility that visual systems are far more genetically variable across populations within species than is commonly assumed. Since colour vision is constrained by the number of cone cells in the retina, even minor differences in selection pressures for the various roles vision plays (e.g. detecting and evaluating potential mates, competitors, predators, food, etc.) across populations could lead to differences in visual systems and may help explain population divergence in vision across a wide range of taxa.

Interestingly, we saw no difference in colour vision between males and females. The lack of sexual dimorphism in guppy opsin expression is not surprising since sexual dimorphism in opsin expression is rarely observed in other species for which it has been tested (such as killifish (Johnson *et al.* 2013), some cichlid species (Sabbah *et al.* 2010), or three-spined stickleback (Boulcott & Braithwaite 2006)), nor was there any sexual dimorphism found in Trinidadian guppy visual systems using microspectrophotometry to measure cone cell absorbencies (Archer & Lythgoe 1990). It should be noted that Laver and Taylor (2011) did report a sexual dimorphism in opsin expression in a laboratory population of guppies from mainland South America. We measured opsin expression profiles on individuals, with a large sample size (287 individuals), using highly specific probe based qPCR assays; therefore we feel it extremely unlikely that we failed to detect a true sexual dimorphism in guppy opsin expression for these Trinidadian populations. Further work is needed on opsin expression in guppies from mainland South America and laboratory populations of guppies before we can completely rule out sex differences in guppy opsin expression.

Differences in LWS opsin expression need not be the result of selection specifically for its role in mate choice for it to impact mate choice decisions. The sensory drive hypothesis states that sensory systems adapt to local environments and predicts that divergence in signals (such as the colours involved in mate choice) occurs with differences in those environments (Endler 1992, 1993; Endler & Basolo 1998; Boughman 2002). While this study found no significant differences in lighting environments it is possible that differences do occur at other times of the year (as potentially seen in DO levels by Tezuka *et al.* 2014) or measured in ways other than λp50. Further experiments are needed to test how much variation in opsin expression profiles are due to genetic versus environmental factors, and to what extent these differences in expression and allele frequency can affect female preferences.

4.5.3. Implications for Evolutionary Theory

According to the sensory bias model of sexual selection, variation in peripheral sensory systems can lead to differences in female mate preferences (Basolo & Endler 1995; Endler & Basolo 1998; Boughman 2002). Indeed changes to sensory systems

responsible for detecting the signals involved in mate choice are well known to influence both the direction and strength of preference for those signals (reviewed in Horth 2007 and Ryan & Cummings 2013). Traditional models of sensory system evolution through protein changes rely on the slow accumulation of sequence level changes due to mutation, gene duplication, and recombination that alter the tuning of sensory proteins (Horth 2007; Sandkam *et al.* 2012). These relatively slow models of sequence change do little to explain what happens to the sensory systems of populations that are capable of rapidly diverging with respect to mate choice, such as guppies. Meanwhile, regulatory differences are thought to be able to evolve more quickly since changes to regulatory regions are often less deleterious (Carroll 2005); these differences in opsin regulation could then persist in a population and be amenable to selection. As seen in Endler *et al.* (2001) artificial selection on guppies can result in different visual sensitivities in less than 9 generations. Clearly visual sensitivities have the capacity to keep pace with the rapidly evolving male colouration and female preferences seen in transplant experiments.

In 2002 Rodd *et al.* demonstrated that both male and female guppies prefer red and orange coloured objects outside of a mating context. Their data provided compelling support for their hypothesis that female mate choice has evolved following the sensory bias model of the evolution of sexual selection (Rodd *et al.* 2002). There are two models underlying Sensory Bias (Endler & Basolo 1998): the Sensory Trap (ST) model posits that female mate preferences toward a specific trait are due to selection historically leading to a preference for that trait outside of a mating context, allowing males to exploit that preference by displaying traits that resemble that preference (Christy 1995). The Sensory Exploitation (SE) model posits that female preferences for male traits are a result of male traits tuned to maximize detection by a female's sensory system (Ryan 1990; Ryan & Rand 1990, 1993; Ryan *et al.* 1990). The ST and SE models largely vary in the emphasis placed on what physiological changes occur to drive changes to female preference; the ST predicts changes at the level of higher order processing, while the SE model predicts changes at the peripheral sensory system should be capable of driving differences in female preferences. However, we saw no differences in opsin expression between males and females; this raises the possibility that red/orange attractions in both mating and non-mating contexts are the result of differences in visual tuning, and sexual selection in guppies is actually following the SE model.

4.6. Conclusion

Here we present support for the sensory exploitation model of the evolution of sexual selection. We found well-established differences in female mating preferences in guppies are associated with variation in the visual system. Populations with stronger female preferences for red/orange colouration invest a strikingly greater proportion of their colour vision in detecting those colours and are strongly differentiated for sequence variants at LWS-1 known to change spectral sensitivity. Furthermore, this pattern holds for two independently colonized watersheds. The intra-species differences in sensory systems across populations presented here and in other taxa suggest care should be taken when modeling sensory input based on measures of sensory systems taken from a different population.

Acknowledgements

We wish to thank Dr. Nadia Aubin-Horth, Dr. Becky Fuller and 2 anonymous reviewers with Axios Review for insightful comments and critiques of the manuscript; Frances Margaret Walker Breden for assistance in the field; Dr. Andrew Hendry and Kiyoko Gotanda for identifying field sites; Dr. Nora Hengst Prior, Dr. Kim Hughes, and Dr. Natasha Bloch for comments on early versions of the manuscript; and Dr. William Davidson and Dr. Krzysztof Lubieniecki for assistance and use of their qPCR machine at Simon Fraser University. Animals were collected under SFU animal protocol 982B-06 and permit # 001453 from The Trinidad Ministry of Housing and the Environment. This work was supported by NSERC discovery grant #138178 to FB, a Graduate International Research Award from Simon Fraser University to BS and an Undergraduate Student Research Award from NSERC to MY.

Data Accessibility

LWS-1 DNA sequences: GenBank Accessions KP218062-KP218254. All qPCR data, LWS-1 alignment, LWS-1 frequency, environmental parameter data, and commented R code used for statistical analyses are available on the Dryad Digital Repository at http://dx.doi.org/10.5061/dryad.sk6d5

References

Archer SN, Lythgoe JN (1990) The visual pigment basis for cone polymorphism in the guppy, *Poecilia reticulata*. **30**, 225–233.

Basolo AL, Endler JA (1995) Sensory biases and the evolution of sensory systems. *Trends In Ecology & Evolution*, **10**, 489.

Boughman JW (2002) How sensory drive can promote speciation. *Trends In Ecology & Evolution*, **17**, 571–577.

Boulcott P, Braithwaite VA (2006) Colour perception in three-spined sticklebacks: sexes are not so different after all. *Evolutionary Ecology*, **21**, 601–611.

Breden F, Stoner G (1987) Male predation risk determines female preference in the Trinidad guppy. *Nature*, **329**, 831–833.

Brooks R, Endler JA (2001) Female guppies agree to differ: phenotypic and genetic variation in mate-choice behavior and the consequences for sexual selection. *Evolution*, **55**, 1644–1655.

Carleton KL, Kocher TD (2001) Cone opsin genes of African cichlid fishes: tuning spectral sensitivity by differential gene expression. *Molecular Biology And Evolution*, **18**, 1540–1550.

Carroll SB (2005) Evolution at two levels: on genes and form. *PLoS Biology*, **3**, 1159–1166.

Cheng CL, Flamarique IN (2004) Opsin expression: new mechanism for modulating colour vision. *Nature*, **428**, 279–279.

Cheng C, Flamarique IN (2007) Chromatic organization of cone photoreceptors in the retina of rainbow trout: single cones irreversibly switch from UV (SWS1) to blue (SWS2) light sensitive opsin during natural development. *The Journal of Experimental Biology*, **210**, 4123-4135.

Christy JH (1995) Mimicry, mate choice, and the sensory trap hypothesis. *American Naturalist*, **146**, 171–181.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Inc.

Crispo E, Bentzen P, Reznick DN, Kinnison MT, Hendry AP (2006) The relative influence of natural selection and geography on gene flow in guppies. *Molecular Ecology*, **15**, 49–62.

Endler JA (1978) A predator's view of animal color patterns. In: *Evolutionary Biology* (eds Hecht MK, Steere WC, Wallace B), pp. 319–364. Plenum Press, Boston, MA.

Endler JA (1980) Natural selection on color patterns in *Poecilia reticulata*. *Evolution*, **34**, 76–91.

Endler JA (1991) Variation in the appearance of guppy color patterns to guppies and their predators under different visual conditions. *Vision Research*, **31**, 587–608.

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

Endler JA (1993) The color of light in forests and its implications. *Ecological Monographs*, **63**, 1–27.

Endler JA, Basolo AL (1998) Sensory ecology, receiver biases and sexual selection. *Trends In Ecology & Evolution*, **13**, 415–420.

Endler JA, Houde AE (1995) Geographic variation in female preferences for male traits in *Poecilia reticulata*. *Evolution*, **49**, 456–468.

Endler JA, Basolo AL, Glowacki S, Zerr J (2001) Variation in response to artificial selection for light sensitivity in guppies (*Poecilia reticulata*). *American Naturalist*, **158**, 36–48.

Flamarique IN, Cheng CL, Bergstrom C, Reimchen TE (2013) Pronounced heritable variation and limited phenotypic plasticity in visual pigments and opsin expression of threespine stickleback photoreceptors. *Journal of Experimental Biology*, **216**, 656–667.

Fuller RC, Claricoates KM (2011) Rapid light-induced shifts in opsin expression: finding new opsins, discerning mechanisms of change, and implications for visual sensitivity. *Molecular Ecology*, **20**, 3321–3335.

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: Sensory, Neural, and Behavioral Physiology*, **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, Fleishman LJ, Leal M, Travis J, Loew ER (2003) Intraspecific variation in retinal cone distribution in the bluefin killifish, *Lucania goodei*. *Journal Of Comparative Physiology A-Neuroethology Sensory Neural And Behavioral Physiology*, **189**, 609–616.

Gegenfurtner KR, Sharpe LT (1999) *Color Vision: From genes to perception*. Cambridge University Press.

Hagstrom SA, Neitz M, Neitz J (2000) Cone pigment gene expression in individual photoreceptors and the chromatic topography of the retina. *Journal of the Optical Society of America A*, **17**, 527–537.

Haskins CP, Haskins EF, McLaughlin JJA, Hewitt RE (1961) Polymorphism and population structure in *Lebistes reticulatus*, an ecological study. In: *Vertebrate Speciation* (ed Blair WF), pp. 320–395.

Hoffmann M, Tripathi N, Henz SR *et al.* (2007) Opsin gene duplication and diversification in the guppy, a model for sexual selection. *Proceedings Of The Royal Society Of London Series B-Biological Sciences*, **274**, 33–42.

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

Horth L (2007) Sensory genes and mate choice: Evidence that duplications, mutations, and adaptive evolution alter variation in mating cue genes and their receptors. *Genomics*, **90**, 159–175.

Houde AE (1997) *Sex, color, and mate choice in guppies*. Princeton University Press.

Houde AE, Endler JA (1990) Correlated evolution of female mating preferences and male color patterns in the guppy *Poecilia reticulata*. *Science*, **248**, 1405–1408.

Hughes KA, Houde AE, Price AC, Rodd FH (2013) Mating advantage for rare males in wild guppy populations. *Nature*, **503**, 108–110.

Hurtado-Gonzales JL, Loew ER, Uy JAC (2014) Variation in the visual habitat may mediate the maintenance of color polymorphism in a poeciliid fish. *Plos One*, **9**, e101497.

Johnson AM, Fuller RC, Stanis S (2013) Diurnal lighting patterns and habitat alter opsin expression and colour preferences in a killifish. *Proceedings Of The Royal Society Of London Series B-Biological Sciences*, **280**, 20130796.

Kemp DJ, Reznick DN, Grether GF, Endler JA (2009) Predicting the direction of ornament evolution in Trinidadian guppies (*Poecilia reticulata*). *Proceedings of the Royal Society of London Series B- Biological Sciences*, **276**, 4335–4343.

Kuijper B, Pen I, Weissing FJ (2012) A guide to sexual selection theory. *Annual Review of Ecology Evolution and Systematics*, **43**, 287–311.

Larmuseau MHD, Raeymaekers JAM, Ruddick KG, Van Houdt JKJ, Volckaert FAM (2009) To see in different seas: spatial variation in the rhodopsin gene of the sand goby (*Pomatoschistus minutus*). *Molecular Ecology*, **18**, 4227–4239.

Laver CRJ, Taylor JS (2011) RT-qPCR reveals opsin gene upregulation is associated with age and sex in guppies (*Poecilia reticulata*)-a species with color-based sexual selection and 11 visual-opsin genes. *BMC Evolutionary Biology*, **11**, 81.

Lehtonen PK, Laaksonen T, Artemyev AV *et al.* (2011) Candidate genes for colour and vision exhibit signals of selection across the pied flycatcher (*Ficedula hypoleuca*) breeding range. *Heredity*, **108**, 431–440.

Lindholm AK, Head ML, Brooks RC *et al.* (2014) Causes of male sexual trait divergence in introduced populations of guppies. *Journal Of Evolutionary Biology*, **27**, 437–448.

Magurran AE (2005) *Evolutionary Ecology: The Trinidadian Guppy*. Oxford University Press, New York.

McFarland WN, Munz FW (1975) Part II: The photic environment of clear tropical seas during the day. *Vision Research*, **15**, 1063-1070.

O'Steen S, Cullum AJ, Bennett AF (2002) Rapid evolution of escape ability in Trinidadian Guppies (*Poecilia reticulata*). *Evolution*, **56**, 776–784.

Panhuis TM, Butlin R, Zuk M, Tregenza T (2001) Sexual selection and speciation. *Trends In Ecology & Evolution*, **16**, 364–371.

Parry JWL, Carleton KL, Spady TC *et al.* (2005) Mix and match color vision: tuning spectral sensitivity by differential opsin gene expression in Lake Malawi cichlids. *Current biology : CB*, **15**, 1734–1739.

R Core Team (2014). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

Reznick DN, Bryga H (1987) Life-history evolution in guppies (*Poecilia reticulata*): 1. Phenotypic and genetic changes in an introduction experiment. *Evolution*, **41**, 1370-1385.

Reznick DN, Bryga HA (1996) Life-history evolution in guppies (*Poecilia reticulata*: Poeciliidae). V. Genetic basis of parallelism in life histories. *American Naturalist*, **147**, 339-359.

Reznick DN, Butler M, Rodd FH (2001) Life-history evolution in guppies. VII. The comparative ecology of high- and low-predation environments. *American Naturalist*, **157**, 126–140.

Reznick DN, Rodd FH, Cardenas M (1996) Life-history evolution in guppies (*Poecilia reticulata*: Poeciliidae). IV. Parallelism in life-history phenotypes. *American Naturalist*, **147**, 319-338.

Rodd FH, Hughes KA, Grether GF, Baril CT (2002) A possible non-sexual origin of mate preference: are male guppies mimicking fruit? *Proceedings Of The Royal Society Of London Series B-Biological Sciences*, **269**, 475–481.

Ryan MJ (1990) Sexual selection, sensory systems and sensory exploitation. *Oxford Surveys in Evolutionary Biology*, **7**, 157–195.

Ryan MJ, Fox JH, Wilczynski W, Rand AS (1990) Sexual selection for sensory exploitation in the frog *Physalaemus pustulosus*. *Nature*, **343**, 66–67.

Ryan MJ, Rand AS (1990) The sensory basis of sexual selection for complex calls in the Tungara Frog, *Physalaemus pustulosus* (Sexual Selection for Sensory Exploitation). *Evolution*, **44**, 305-314.

Ryan MJ, Rand AS (1993) Species recognition and sexual selection as a unitary problem in animal communication. *Evolution*, **47**, 647-657.

Ryan MJ, Cummings ME (2013) Perceptual Biases and Mate Choice. *Annual Review of Ecology, Evolution, and Systematics*, **44**, 437–459.

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

Sandkam BA, Joy JB, Watson CT *et al.* (2012) Hybridization leads to sensory repertoire expansion in a gynogenetic fish, the Amazon Molly (*Poecilia formosa*): a test of the Hybrid-Sensory Expansion hypothesis. *Evolution*, **67**, 120–130.

Shand J, Davies WL, Thomas N *et al.* (2008) The influence of ontogeny and light environment on the expression of visual pigment opsins in the retina of the black bream, *Acanthopagrus butcheri*. *The Journal of Experimental Biology*, **211**, 1495–1503.

Smith AR, D'Annunzio L, Smith AE *et al.* (2010) Intraspecific cone opsin expression variation in the cichlids of Lake Malawi. *Molecular Ecology*, **20**, 299–310.

Tezuka A, Kasagi S, Van Oosterhout C *et al.* (2014) Divergent selection for opsin gene variation in guppy (*Poecilia reticulata*) populations of Trinidad and Tobago. *Heredity*, **113**, 381–389.

Vandesompele J, De Preter K, Pattyn F *et al.* (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*, **3**, 0034.

Ward MN, Churcher AM, Dick KJ *et al.* (2008) The molecular basis of color vision in colorful fish: four Long Wave-Sensitive (LWS) opsins in guppies (*Poecilia reticulata*) are defined by amino acid substitutions at key functional sites. *BMC Evolutionary Biology*, **8**, 210.

Watson CT, Gray SM, Hoffmann M *et al.* (2011) Gene duplication and divergence of Long Wavelength-Sensitive opsin genes in the Guppy, *Poecilia reticulata*. *Journal Of Molecular Evolution*, **72**, 240–252.

Yokoyama S (2000) Molecular evolution of vertebrate visual pigments. *Progress in Retinal and Eye Research*, **19**, 385–420.

Yokoyama S (2002) Molecular evolution of color vision in vertebrates. *Gene*, **300**, 69–78.

Yokoyama S (2008) Evolution of dim-light and color vision pigments. *Annual Review of Genomics and Human 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.

Tables and Figures

Table 4.1. Results of MANOVA on opsin expression profiles. W- Watershed, T- Time, P(W)- Predation nested in watershed.

		Proportional Expression		Relative$_{(hk)}$ Expression	
	df	*F*	*P*	*F*	*P*
W	**1, 279**	**6.82**	**7.75E-09**	**7.16**	**2.53E-09**
T	**1, 279**	**20.56**	**< 2.2e-16**	**20.51**	**< 2.2e-16**
P(W)	**2, 279**	**9.33**	**< 2.2e-16**	**9.81**	**< 2.2e-16**
W*T	**1, 279**	**22.74**	**< 2.2e-16**	**22.68**	**< 2.2e-16**
P(W)*T	**2, 279**	**10.20**	**< 2.2e-16**	**11.67**	**< 2.2e-16**

Table 4.2. Results of MANOVA on opsin expression profiles by watershed.

	Aripo Watershed				Marianne Watershed			
	Proportional Expression		Relative$_{(hk)}$ Expression		Proportional Expression		Relative$_{(hk)}$ Expression	
	$F_{(1,139)}$	*P*	$F_{(1,139)}$	*P*	$F_{(1,140)}$	*P*	$F_{(1,140)}$	*P*
Predation (P)	**7.56**	**< 0.0001**	**7.78**	**< 0.0001**	**16.06**	**< 0.0001**	**15.84**	**< 0.0001**
Time (T)	**31.01**	**< 0.0001**	**27.33**	**< 0.0001**	**17.38**	**< 0.0001**	**17.79**	**< 0.0001**
P*T	**10.05**	**< 0.0001**	**9.12**	**< 0.0001**	**13.36**	**< 0.0001**	**14.78**	**< 0.0001**

Table 4.3. Results of LWS ANOVA. W- Watershed, T- Time, P(W)- Predation nested in watershed.

		Proportional Opsin				**Relative$_{(hk)}$**			
		LWS-1	LWS-2	LWS-3	LWS-R	LWS-1	LWS-2	LWS-3	LWS-R
W	$F_{(1,279)}$	3.70	**5.02**	0.16	1.34	**9.21**	**5.48**	1.41	**7.74**
	P	0.0553	**0.0258**	0.6836	0.2469	**0.0026**	**0.0198**	0.2347	**0.0057**
T	$F_{(1,279)}$	0.14	**2.82**	**15.87**	**16.53**	0.19	**5.83**	0.08	3.87
	P	0.7009	**0.0940**	**<0.0001**	**<0.0001**	0.6605	**0.0164**	0.7720	0.0500
P(W)	$F_{(1,279)}$	**9.81**	**7.90**	**7.44**	**11.74**	**8.33**	1.14	0.21	**9.21**
	P	**<0.0001**	**0.0004**	**0.0007**	**<0.0001**	**0.0003**	0.3205	0.8065	**0.0001**
W*T	$F_{(1,279)}$	0.94	**28.16**	2.53	**9.82**	0.50	**7.65**	0.22	**5.20**
	P	0.3324	**<0.0001**	0.1127	**0.0019**	0.4777	**0.0060**	0.6328	**0.0232**
P(W)*T	$F_{(1,279)}$	**13.00**	**5.69**	**8.20**	**22.82**	0.19	2.52	**4.28**	**5.00**
	P	**<0.0001**	**0.0037**	**0.0003**	**<0.0001**	0.8270	0.08217	**0.0147**	**0.0073**

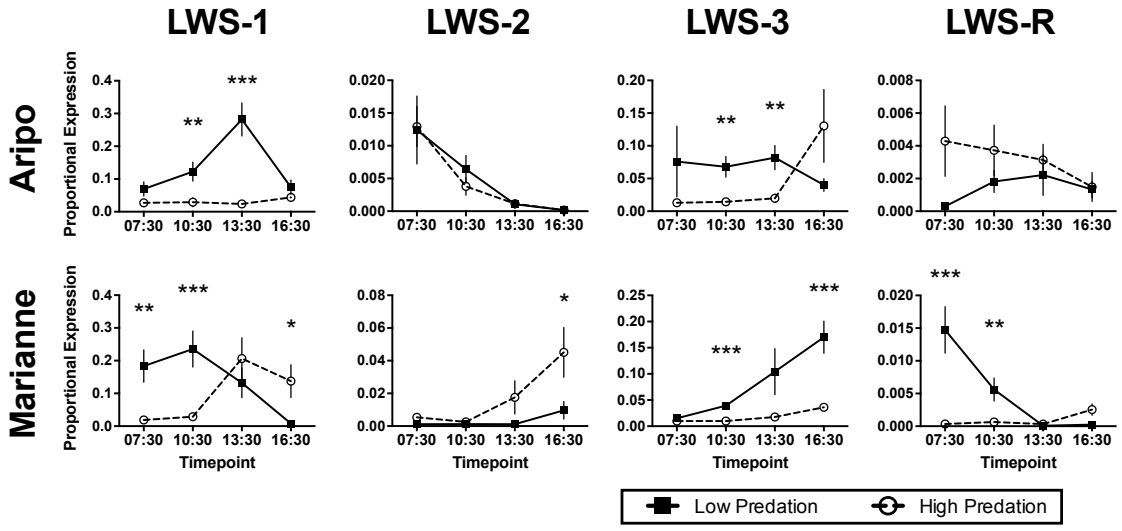


Figure 4.1. **Long wavelength-sensitive (LWS) opsin expression by population for proportional measures. Stars denote significant differences between high and low predation within the same watershed (**P* < 0.05, ***P* <0.01, ****P* < 0.001); error bars are ±1 SE (see Table S8, Supporting information). Aripo low predation at 13:30 had *N* = 17; all other population time points had *N* = 18.**

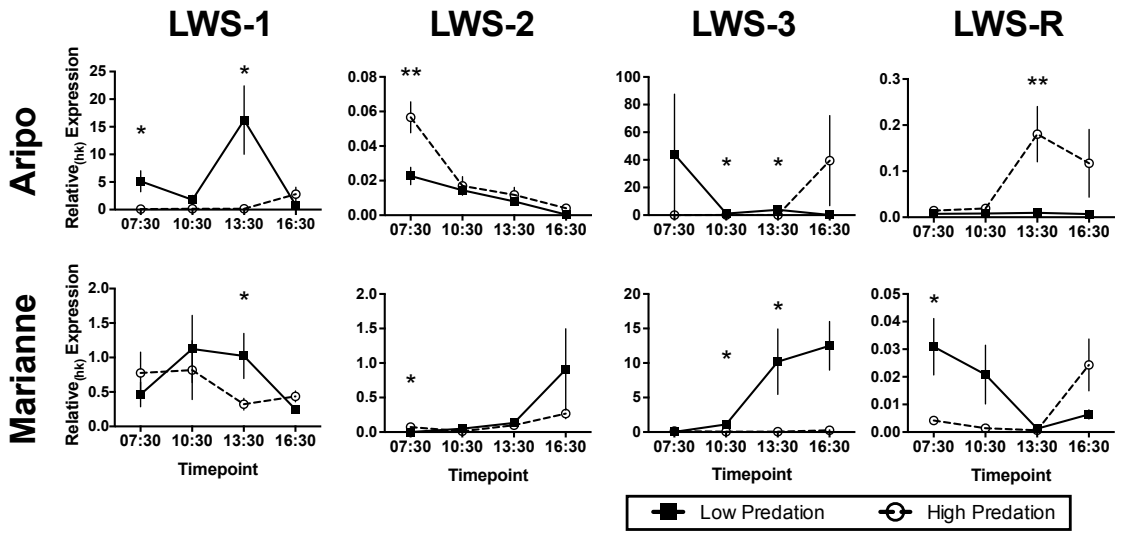


Figure 4.2. **Long wavelength-sensitive (LWS) opsin expression by population for relative$_{(hk)}$ measures. Stars denote significant differences between high and low predation within the same watershed (**P* < 0.05, ***P* < 0.01); error bars are ±1 SE (see Table S9, Supporting information). Aripo low predation at 13:30 had *N* = 17; all other population time points had *N* = 18.**

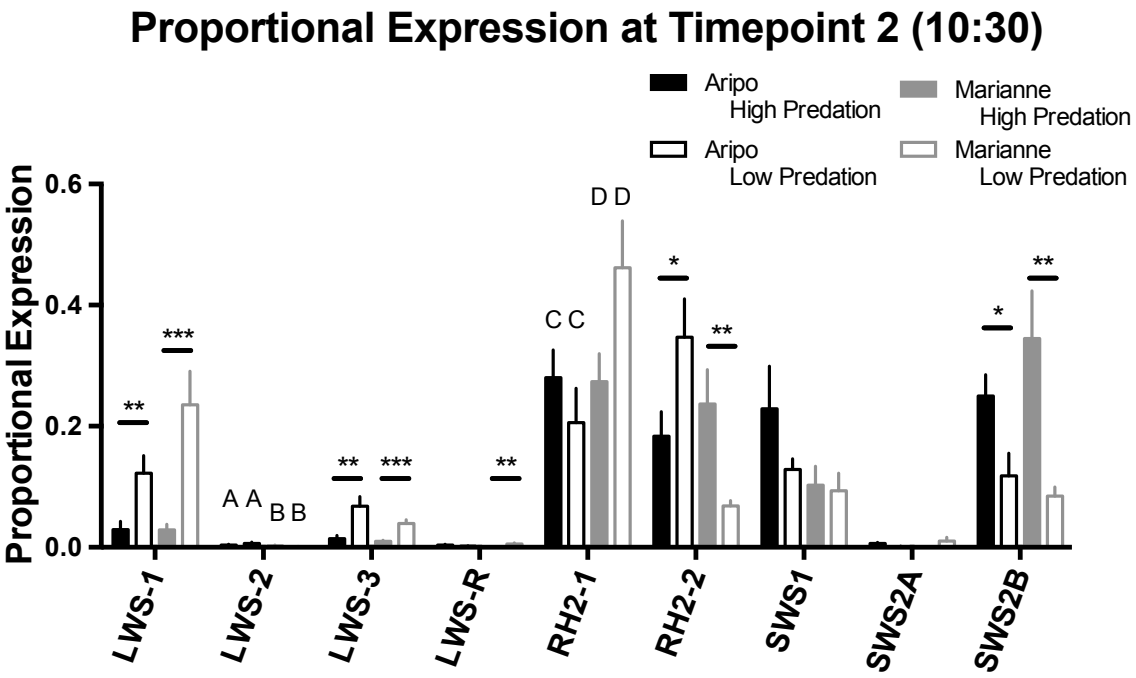


Figure 4.3. **Proportional expression at time point 2. Stars denote significant differences within watershed between high- and low-predation populations (**P* < 0.05, ***P* < 0.01, ****P* < 0.001) (see Table S8, Supporting information). Letters denote significant differences between watersheds (see Table S6, Supporting information).

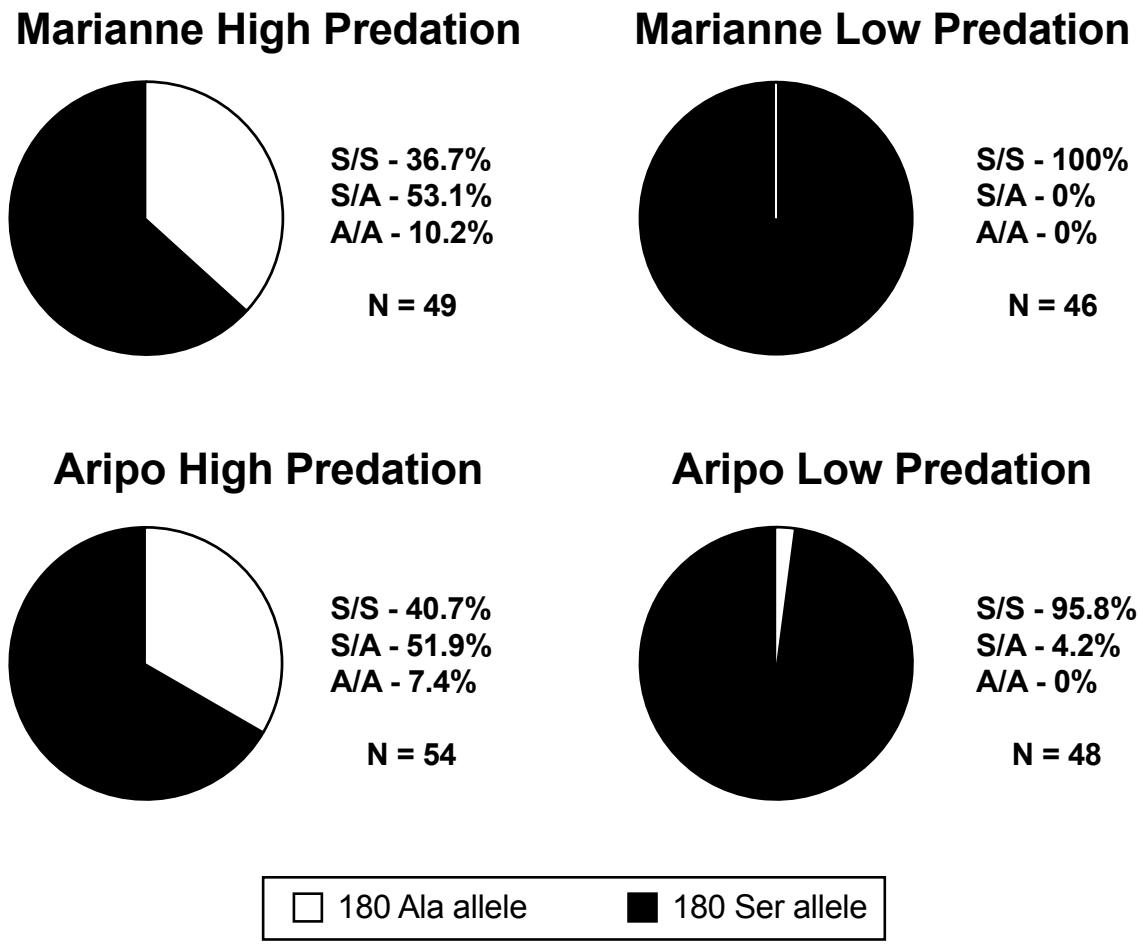


Figure 4.4. **Genotype and gene frequencies of the 180 Ala (A) vs. 180 Ser (S) allele of LWS-1. LWS, long wavelength-sensitive.**

Supplementary Tables and Figures

Supplementary Table 4.1. Coordinates of populations sampled.

	Latitude - N	Longitude - W
Aripo High Predation	10◦ 39.036	061◦ 13.387
Aripo Low Predation	10◦ 40.955	061◦ 13.846
Marianne High Predation	10◦ 45.894	061◦ 18.279
Marianne Low Predation	10◦ 45.015	061◦ 17.130

Supplementary Table 4.2. Primer/probe sequence, length of product (bp), GenBank accession number for sequence on which assay was designed, relative (± SE) and absolute qPCR efficiencies of each gene assay.

Gene	Forward Primer (5'-3')	FAM labeled Probe with quencher (5'-3')	Reverse Primer (5'-3')	Product Length	GenBank Accession	Relative efficiency	Absolute efficiency
B-actin	GCTTGTGCGGGATATCATTTG	CATGGCTGAACTTTGGGTGTGGG	GAATCCGGCTTTGCACATAC	137	EU143772	0.9623 ± 0.0287	0.8258
COI	GCCCACTTCCACTATGTTCTC	ATAGGAGCCGTCTTTGCCATTGTTGC	TGTTCAGGTGCTGTGTAGTG	108	HQ391989	0.9686 ± 0.0314	0.8313 ± 0.0457
Myosin HC	TCTCTGCTCACCTTCAACTTC	CCAGTTCTCGGACCCTCTGCT	CGAGGTTACAGTCAAGGATCTG	136	HQ260688	0.9156 ± 0.0340	0.7858
LWS-1	ATACAACAAGAGGCGCCG	AAGATCCCTTTGAAGGACCAACTACCACA	CACGATACACATCCAGAGAGT G	128	HM540108	0.9633 ± 0.0299	0.8267
LWS-2	GCGGCACAATGAAGATACAAC	AGAGATCCCTTTGAAGGACCAAACTACCAC	GACCACGATACACATCCAGAC	145	HM540108	0.9031 ± 0.0276	0.7751
LWS-3	AAGATACAACAAGAGGCGCT	AAGATCCCTTTGAAGGACCAACTACCACA	CCACGATACACATCCAGAGT	132	HM540108	0.9142 ± 0.0271	0.7856
LWS-R	AATGAAGAAACTACAAGGGG CT	AGAGATCCTTTTGAGGGACCAAACTACCAT	CACATCCAGACTGTCGTGAT	128	HM540107	0.9154 ± 0.0282	0.7856
RH1	TGGTCGTTATCATGGTCATCG	TTCCTGGTTTGCTGGATACCTTACGC	TTGTAGATGGAGGCTGACTTG	148	DQ912023	0.9066 ± 0.0277	0.7781
RH2-1	TCACCATCACCATCACATCTG	AGGGATTCATGGCAACACTTGGAGG	GAACAACCAGAGACCAGAGA G	120	DQ234859	0.9140 ± 0.0271	0.7844
RH2-2	GGACCCTTAGGCTGTGATATG	TGACCTCCAACCGTAGCAAAGAAACC	GCTAAGACCACAAGAGACCAT AG	80	DQ234858	0.9065 ± 0.0284	0.7780
SWS1	GTCTTCACCTGGTTCATGGG	CTGGAGCAGGTACATCCCGGAAG	AGCCGTACTCTTCATTGTGG	127	DQ912025	0.9083 ± 0.0265	0.7795
SWS2A	CACCGTCTACAATCCTGTCA	AGTCCTCAATAAGCAGTTTCGAACATGCAT	GACCAACTTTGGAGATTTCAG T	137	HM540108	0.8866 ± 0.0245	0.7609
SWS2B	TCCTGTGTCTCAAAAGCCTC	ACGTTCTCCTCAATAAGCAGTTCCGC	CACTCATCCCCAGCATCTTC	100	HM540108	0.9115 ± 0.0284	0.7823

Supplementary Table 4.3. PCR and sequencing primers used for LWS-1 180 Ala/Ser allele identification.

	Forward	Reverse
PCR	TGTGAAGTGCAGATCACCTAG	ACACATTCATGCATGATGCAG
Sequencing	GATCCCTTTGAAGGACCAAACT	GGACAATCATGTAGGACAGGACC

Supplementary Table 4.4. Results of ANOVAs for proportional measures of expression for each opsin (bold denotes *P* < 0.05). W- Watershed, T-Time, P(W)- Predation nested in watershed.

		LWS-1	LWS-2	LWS-3	LWS-R	SWS1	SWS2A	SWS2B	RH21	RH22
W	$F_{(1,279)}$	3.71	**5.02**	0.17	1.35	**5.41**	2.73	1.32	0.34	**21.17**
	P	0.0553	**0.0258**	0.6836	0.2469	**0.0207**	0.0997	0.2520	0.5600	**<0.0001**
T	$F_{(1,279)}$	0.15	**2.82**	**15.88**	**16.53**	3.32	**11.59**	0.28	0.17	3.10
	P	0.7009	**0.0940**	**<0.0001**	**<0.0001**	0.0695	**0.0008**	0.6010	0.6840	0.0794
P(W)	$F_{(1,279)}$	**9.81**	**7.90**	**7.44**	**11.75**	1.88	**5.30**	1.89	1.03	**13.03**
	P	**<0.0001**	**0.0004**	**0.0007**	**<0.0001**	0.1541	**0.0055**	0.1540	0.3600	**<0.0001**
W*T	$F_{(1,279)}$	0.94	**28.17**	2.53	**9.82**	**31.44**	**16.92**	0.09	**19.41**	2.72
	P	0.3324	**<0.0001**	0.1128	**0.0019**	**<0.0001**	**<0.0001**	0.762	**<0.0001**	0.1006
P(W)*T	$F_{(1,279)}$	**13.00**	**5.70**	**8.21**	**22.82**	**3.29**	**14.23**	**13.66**	0.07	**19.75**
	P	**<0.0001**	**0.0038**	**0.0003**	**<0.0001**	**0.0385**	**<0.0001**	**<0.0001**	0.931	**<0.0001**

Supplementary Table 4.5. Results of ANOVAs for relative$_{hk}$ measures of expression for each opsin (bold denotes *P* < 0.05). W- Watershed, T-Time, P(W)- Predation nested in watershed.

		LWS-1	LWS-2	LWS-3	LWS-R	SWS1	SWS2A	SWS2B	RH2-1	RH2-2	RH1
W	$F_{(1,279)}$	**9.22**	**5.49**	1.42	**7.74**	0.35	0.16	0.64	0.56	**28.49**	**17.14**
	P	**0.0026**	**0.0199**	0.2347	**0.0058**	0.5546	0.6926	0.4249	0.4543	**<0.0001**	**<0.0001**
T	$F_{(1,279)}$	0.19	**5.83**	0.08	3.87	2.95	**6.41**	**5.05**	2.30	**7.52**	**3.93**
	P	0.6605	**0.0164**	0.7720	0.0500	0.0870	**0.0119**	**0.0254**	0.1308	**0.0065**	**0.0483**
P(W)	$F_{(1,279)}$	**8.33**	1.14	0.22	**9.22**	**8.33**	0.37	1.62	**4.74**	**5.00**	**6.05**
	P	**0.0003**	0.3206	0.8065	**0.0001**	**0.0003**	0.6914	0.1999	**0.0095**	**0.0074**	**0.0027**
W*T	$F_{(1,279)}$	0.51	**7.66**	0.23	**5.21**	**6.09**	3.48	0.46	2.19	3.68	**4.36**
	P	0.4777	**0.0060**	0.6328	**0.0232**	**0.01420**	0.0631	0.4993	0.1400	0.0560	**0.0377**
P(W)*T	$F_{(1,279)}$	0.19	2.52	**4.28**	**5.00**	**6.55**	1.84	**4.27**	2.30	**27.76**	**14.45**
	P	0.8270	0.0822	**0.0147**	**0.0073**	**0.0017**	0.16100	**0.0149**	0.1019	**<0.0001**	**<0.0001**

Supplementary Table 4.6. Results of ANOVAs for proportional measures of expression for each opsin by time (bold denotes *P* < 0.05). T1- 07:30, T2- 10:30, T3- 13:30, T4- 16:30, W- Watershed, T- Time, P(W)- Predation nested in watershed.

			LWS-1	LWS-2	LWS-3	LWS-R	SWS1	SWS2A	SWS2B	RH21	RH22
T1	W	$F_{(1,68)}$	3.67	**9.00**	1.37	**6.34**	2.12	**9.33**	0.60	**7.22**	**15.13**
		P	0.0593	**0.0037**	0.2450	**0.0142**	0.1500	**0.0032**	0.4410	**0.0091**	**0.0002**
	P(W)	$F_{(2,68)}$	**9.48**	0.44	1.35	**12.88**	**3.35**	**7.63**	1.90	2.83	**4.05**
		P	**0.0002**	0.6427	0.2640	**<0.0001**	**0.0410**	**0.0010**	0.1570	0.0661	**0.0217**
T2	W	$F_{(2,67)}$	3.17	**5.71**	3.39	0.07	3.77	0.22	0.37	**4.51**	**5.58**
		P	0.0792	**0.0196**	0.0697	0.7906	0.0565	0.6381	0.5480	**0.0373**	**0.0211**
	P(W)	$F_{(2,67)}$	**12.43**	1.238	**12.42**	**4.74**	1.39	2.78	**9.32**	3.06	**6.31**
		P	**<0.0001**	0.2964	**<0.0001**	**0.0118**	0.2565	0.0694	**0.0003**	0.0533	**0.0031**
T3	W	$F_{(1,68)}$	0.12	2.61	0.17	**9.93**	**35.31**	0.11	1.27	**9.27**	1.37
		P	0.7241	0.1106	0.6765	**0.0024**	**<0.0001**	0.7470	0.2645	**0.0033**	0.246
	P(W)	$F_{(2,68)}$	**8.34**	2.57	**4.88**	0.36	**6.74**	1.52	**4.07**	**6.07**	**19.92**
		P	**0.0005**	0.0833	**0.0104**	0.6927	**0.0021**	0.2270	**0.0214**	**0.0038**	**<0.0001**
T4	W	$F_{(1,68)}$	0.22	**11.33**	0.30	0.00	**20.41**	1.95	0.01	**5.24**	**4.83**
		P	0.6343	**0.0012**	0.5816	0.9798	**<0.0001**	0.1674	0.9108	**0.0252**	**0.0314**
	P(W)	$F_{(2,68)}$	**5.63**	**4.81**	**6.25**	2.98	0.331	**3.62**	**12.62**	1.24	**16.53**
		P	**0.0054**	**0.0110**	**0.0032**	0.0574	0.7200	**0.0321**	**<0.0001**	0.2974	**<0.0001**

Supplementary Table 4.7. Results of ANOVAs for relative$_{(hk)}$ measures of expression for each opsin by time (bold denotes $P < 0.05$). T1- 07:30, T2- 10:30, T3- 13:30, T4- 16:30, W- Watershed, T- Time, P(W)- Predation nested in watershed.

			LWS-1	LWS-2	LWS-3	LWS-R	SWS1	SWS2A	SWS2B	RH21	RH22	RH1
T1	W	$F_{(1,68)}$	**4.24**	0.03	1.02	1.19	0.01	1.51	2.07	**15.37**	**18.50**	**6.04**
		P	**0.0433**	0.8566	0.3160	0.2801	0.9416	0.2240	0.1550	**0.0002**	**<0.0001**	**0.0165**
	P(W)	$F_{(2,68)}$	**6.80**	**5.52**	1.02	**5.01**	**6.24**	2.16	**4.75**	**29.85**	**11.37**	**5.71**
		P	**0.0020**	**0.0060**	0.3650	**0.0093**	**0.0033**	0.1230	**0.0117**	**<0.0001**	**<0.0001**	**0.0051**
T2	W	$F_{(2,67)}$	0.00	1.70	0.00	0.16	**6.65**	0.08	0.34	**4.57**	**15.03**	**11.10**
		P	0.9649	0.1963	0.9847	0.6919	**0.0121**	0.7730	0.5606	**0.0362**	**0.0002**	**0.0014**
	P(W)	$F_{(2,67)}$	2.68	2.49	**5.43**	2.84	**5.86**	2.44	**5.25**	3.12	**11.42**	**6.58**
		P	0.0759	0.0905	**0.0065**	0.0655	**0.0045**	0.0948	**0.0076**	0.0505	**<0.0001**	**0.0025**
T3	W	$F_{(1,68)}$	**5.89**	**6.83**	1.61	**9.75**	**12.93**	**8.15**	0.13	0.02	**12.49**	**5.89**
		P	**0.0178**	**0.0110**	0.2089	**0.0026**	**0.0006**	**0.0057**	0.7170	0.8836	**0.0007**	**0.0179**
	P(W)	$F_{(2,68)}$	**6.75**	0.17	**4.80**	**8.02**	**21.94**	**3.18**	**4.14**	**4.01**	**3.83**	**6.39**
		P	**0.0021**	0.8450	**0.0112**	**0.0007**	**<0.0001**	**0.0479**	**0.0200**	**0.0226**	**0.0264**	**0.0029**
T4	W	$F_{(1,68)}$	**4.84**	3.91	0.67	1.57	**11.88**	0.46	0.46	0.57	**7.11**	2.01
		P	**0.0313**	0.0522	0.4170	0.2140	**0.0009**	0.4990	0.5020	0.4520	**0.0096**	0.1606
	P(W)	$F_{(2,68)}$	2.36	1.17	1.58	2.26	1.59	0.74	2.03	1.96	**9.61**	2.96
		P	0.1017	0.3175	0.2140	0.1120	0.2119	0.4820	0.1390	0.1480	**0.0002**	0.0587

Supplementary Table 4.8. ANOVA table of proportional measures of opsin expression subset by watershed and run by predation for gene-time point combinations that had a significant watershed*predation interaction (bold denotes $P < 0.05$). Note: Aripo watershed had an $F_{1,33}$ at timepoint 2 due to one sample being discarded.

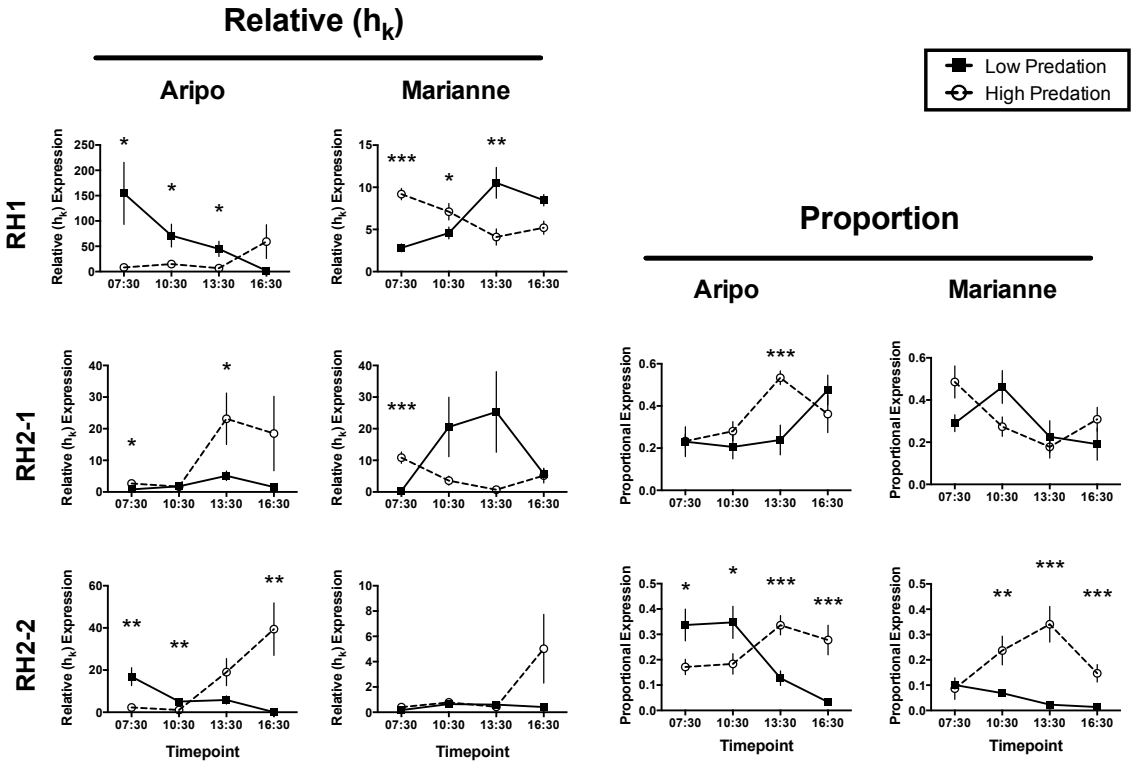
		Timepoint 1 (7:30)		Timepoint 2 (10:30)		Timepoint 3 (13:30)		Timepoint 4 (13:30)	
		$F_{(1,34)}$	P	$F_{(1,34)}$	P	$F_{(1,34)}$	P	$F_{(1,34)}$	P
LWS-1	Aripo	3.25	0.0801	**8.82 ($F_{1,33}$)**	**0.0055**	**25.78**	**<0.0001**	1.49	0.2300
	Marianne	**10.87**	**0.0023**	**13.67**	**0.0007**	0.92	0.3450	**6.69**	**0.0142**
LWS-2	Aripo	-	-	-	-	-	-	0.00	0.9650
	Marianne	-	-	-	-	-	-	**4.82**	**0.0351**
LWS-3	Aripo	-	-	**10.93 ($F_{1,33}$)**	**0.0023**	**10.16**	**0.0031**	2.55	0.1190
	Marianne	-	-	**20.55**	**<0.0001**	3.85	0.0579	**18.31**	**0.0001**
LWS-R	Aripo	3.43	0.0728	1.23 ($F_{1,33}$)	0.2750	-	-	-	-
	Marianne	**16.37**	**0.0003**	**8.03**	**0.0077**	-	-	-	-
SWS1	Aripo	3.72	0.0620	-	-	1.80	0.1880	-	-
	Marianne	2.41	0.1300	-	-	**6.96**	**0.0125**	-	-
SWS2A	Aripo	**10.11**	**0.0031**	-	-	-	-	3.65	0.0645
	Marianne	**7.47**	**0.0099**	-	-	-	-	2.692	0.1100
SWS2B	Aripo	-	-	**6.56 ($F_{1,33}$)**	**0.0152**	**5.11**	**0.0303**	3.059	0.0893
	Marianne	-	-	**10.50**	**0.0027**	3.24	0.0805	**32.47**	**<0.0001**
RH2-1	Aripo	-	-	-	-	**14.34**	**0.0006**	-	-
	Marianne	-	-	-	-	0.263	0.6110	-	-
RH2-2	Aripo	**5.68**	**0.0229**	**4.85 ($F_{1,33}$)**	**0.0348**	**19.14**	**0.0001**	**17.13**	**0.0002**
	Marianne	0.10	0.7510	**8.65**	**0.0059**	**20.27**	**<0.0001**	**14.8**	**0.0005**

Supplementary Table 4.9. ANOVA table of relative$_{(hk)}$ measures of opsin expression subset by watershed and run by predation for gene-time point combinations that had a significant watershed*predation interaction (bold denotes $P < 0.05$). Note: Aripo watershed had an $F_{1,33}$ at timepoint 2 due to one sample being discarded.

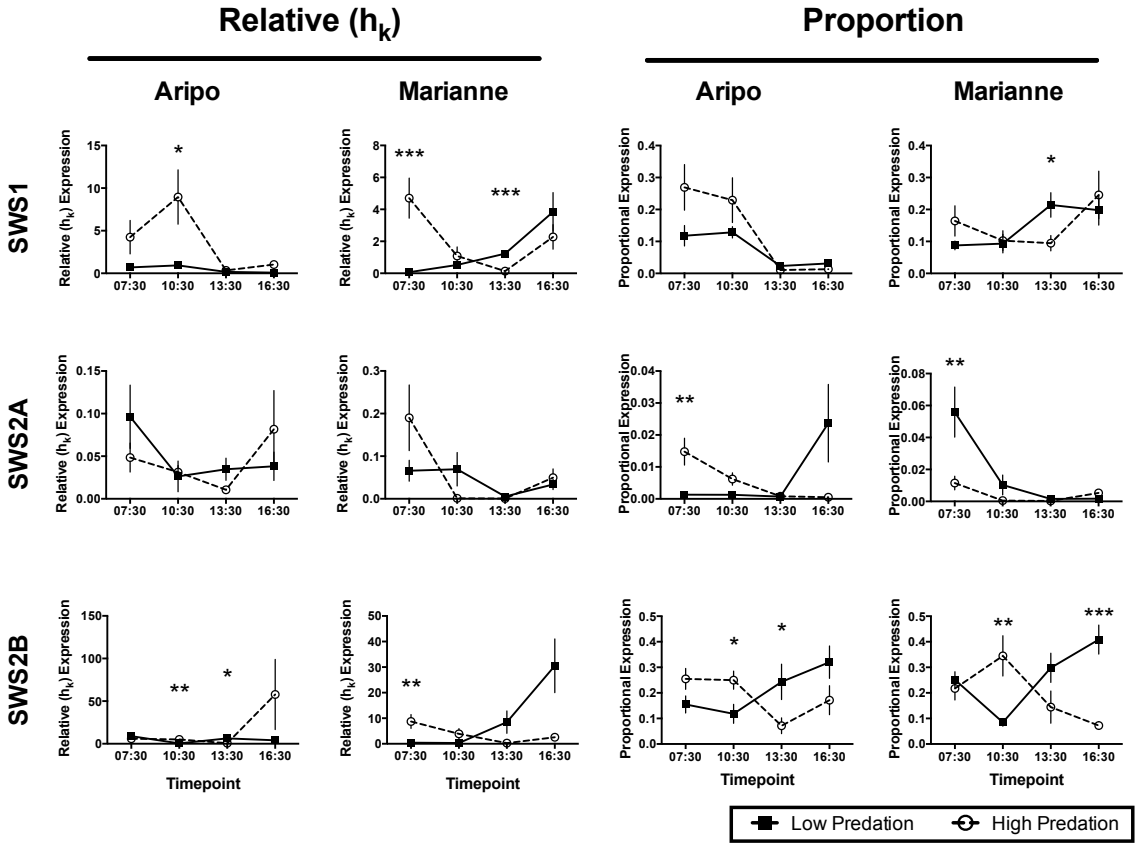
		Timepoint 1 (7:30)		Timepoint 2 (10:30)		Timepoint 3 (13:30)		Timepoint 4 (13:30)	
		$F_{(1,34)}$	P	$F_{(1,34)}$	P	$F_{(1,34)}$	P	$F_{(1,34)}$	P
LWS-1	Aripo	**7.01**	**0.0122**	-	-	**6.75**	**0.0137**	-	-
	Marianne	0.79	0.3810	-	-	**4.32**	**0.0454**	-	-
LWS-2	Aripo	**10.99**	**0.0022**	-	-	-	-	-	-
	Marianne	**4.95**	**0.0329**	-	-	-	-	-	-
LWS-3	Aripo	-	-	**4.23 ($F_{1,33}$)**	**0.0477**	**7.38**	**0.0103**	-	-
	Marianne	-	-	**7.17**	**0.0113**	**4.59**	**0.0394**	-	-
LWS-R	Aripo	1.04	0.3160	-	-	**8.02**	**0.0077**	-	-
	Marianne	**6.75**	**0.0138**	-	-	3.64	0.0648	-	-
SWS1	Aripo	3.24	0.0808	**5.93 ($F_{1,33}$)**	**0.0205**	2.64	0.1140	-	-
	Marianne	**13.62**	**0.0008**	0.94	0.3400	**26.61**	**<0.0001**	-	-
SWS2A	Aripo	-	-	-	-	3.15	0.0847	-	-
	Marianne	-	-	-	-	4.01	0.0534	-	-
SWS2B	Aripo	1.02	0.3210	**9.26 ($F_{1,33}$)**	**0.0046**	**6.71**	**0.0140**	-	-
	Marianne	**9.23**	**0.0046**	3.10	0.0872	3.52	0.0693	-	-
RH2-1	Aripo	**6.29**	**0.0171**	-	-	**4.72**	**0.0370**	-	-
	Marianne	**33.78**	**<0.0001**	-	-	3.71	0.0624	-	-
RH2-2	Aripo	**11.38**	**0.0019**	**11.86 ($F_{1,33}$)**	**0.0016**	3.83	0.0584	**9.93**	**0.0034**
	Marianne	4.00	0.0535	0.27	0.6070	0.64	0.4300	2.88	0.0988
RH1	Aripo	**5.70**	**0.0226**	**6.49 ($F_{1,33}$)**	**0.0157**	**6.33**	**0.0167**	-	-
	Marianne	**59.58**	**<0.0001**	**4.36**	**0.0443**	**9.57**	**0.0039**	-	-

Supplementary Table 4.10. Raw measures of environmental parameters.

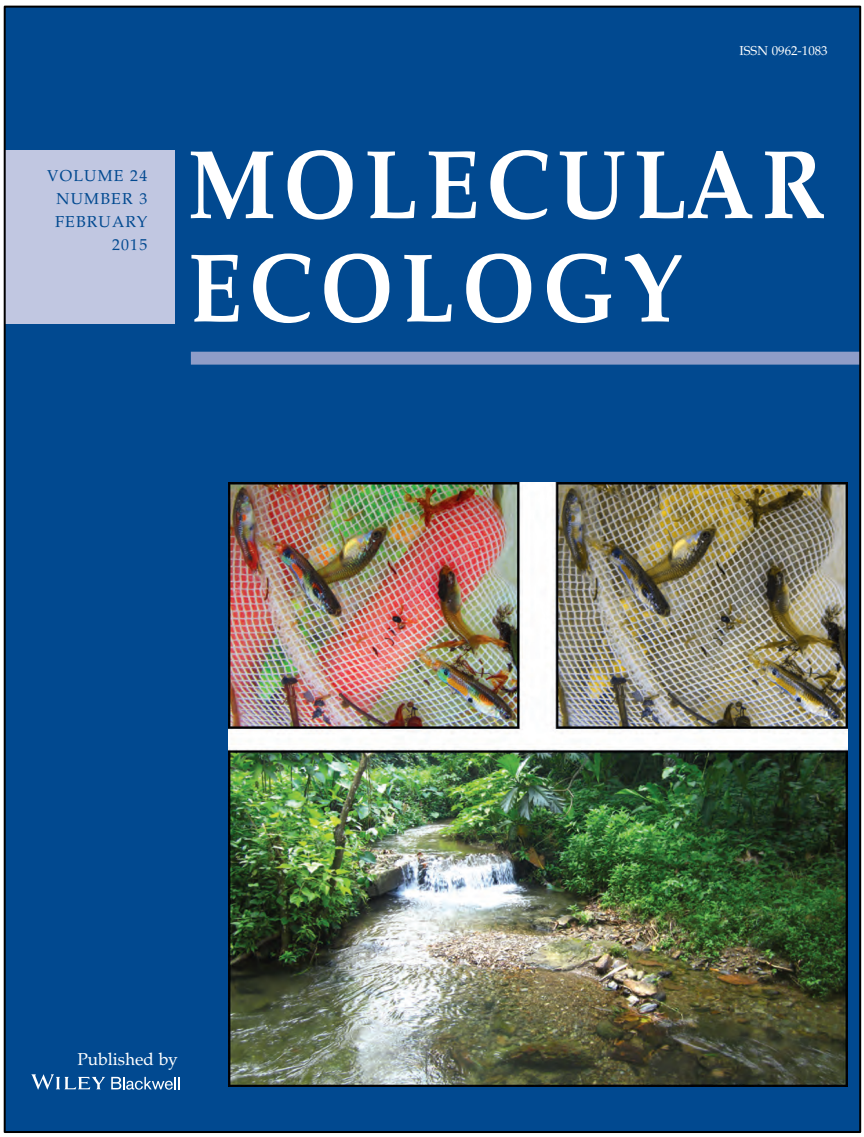
Watershed	Predation	Day	λp50	DO (ppm)	Temperature (°C)	Conductivity (µS/m)	pH	Total Dissolved Solids (mg/L)	Salinity (ppt)
Aripo	High	1	611	6.62	25.4	272	7.87	192	0.14
		2	571	7.52	24.9	330	8.08	221	0.16
		3	573	6.74	25.7	338	7.84	233	0.17
	Low	1	614	6.9	24.8	403	8.03	273	0.2
		2	704	7.01	24.4	393	7.14	264	0.19
		3	677	7.25	25.4	398	8.13	22.6	0.19
Marianne	High	1	651	8.03	24.4	115.8	7.63	77.9	0.05
		2	567	7.29	24.1	217	7.84	144	0.1
		3	585	7.62	24.4	247	7.82	164	0.12
	Low	1	555	7.3	24.7	306	8.04	205	0.15
		2	585	7.41	24.3	207	7.81	139	0.1
		3	571	7.34	24.1	278	7.99	185	0.13



Supplementary Figure 4.1. RH opsin expression by population. Stars denote significant differences between high and low predation within the same watershed (see Supplementary Tables S7 and S8).



Supplementary Figure 4.2. SWS opsin expression by population. Stars denote significant differences between high and low predation within the same watershed (see Supplementary tables S7 and S8).



Supplementary Figure 4.3. Cover of *Molecular Ecology* featuring the published version of chapter 5.

Caption: Barrier waterfalls in Trinidadian streams (bottom photo) restrict movement of predators upstream. Upstream guppies are typically very colorful, with high amounts of red and orange male coloration, and strong female preferences to mate with more red and orange males. Sandkam *et al.* show upstream guppies also express high levels of LWS opsins. The two photos show the dramatic effect of the presence (top left) or absence (top right) of such LWS opsins on the ability to perceive red and orange coloration. Photos by Ben Sandkam, Dr. Felix Breden and Frances Breden. Color blind (protanopia) simulation performed by Etre Limited.

Chapter 5.

Colour vision varies more among populations than species of live-bearing fish from South America

Publication and Contributions

A version of this chapter is in review at *BMC Evolutionary Biology*: Sandkam, BA, CM Young, FMW Breden, GR Bourne, and F Breden. Colour vision varies more among populations than species of live-bearing fishes from South America.

Contributions: BAS and FB conceived and designed the project and wrote the manuscript. BAS, GRB, FMWB and FB conducted fieldwork. BAS designed assays and conducted all qPCR work. CMY sequenced all LWS-1 loci.

5.1. Abstract

Sensory Bias models for the evolution of mate preference place a great emphasis on the role of sensory system variation in mate preferences. However, the extent to which sensory systems vary across- versus within-species remains largely unknown. Here we assessed whether color vision varies in natural locations where guppies and their two closest relatives, *Poecilia parae* and *Poecilia picta*, occur in extreme sympatry and school together. All three species base mate preferences on male coloration but differ in the colors preferred. Measuring opsin gene expression, we found that within sympatric locations these species have similar color vision and that color vision differed more across populations of conspecifics. In addition, all three species differ across populations in the frequency of the same opsin coding polymorphism that influences visual tuning. Together, this shows sensory systems vary considerably across populations and supports the possibility that sensory system variation is involved in population divergence of mate preference.

5.2. Introduction

Population divergence is widely accepted as a precursor to speciation (Coyne and Orr 2004), and can occur rapidly due to sexual selection via changes in mate preference (Panhuis et al. 2001). Many hypotheses have attempted to explain the drivers behind changes in mate preference, which fall into two general categories—indirect models, such as Fisher's runaway and good genes, or direct models such as Sensory Bias (reviewed in Andersson and Simmons 2006; Kuijper et al. 2012). Sensory Bias models emphasize the role of sensory system variation in driving divergence in mate preferences (Ryan and Rand 1993; Basolo and Endler 1995; Endler and Basolo 1998; Boughman 2002; Ryan and Cummings 2013). However, the extent to which sensory systems vary across- versus within-species remains largely unknown. Describing where the variation in sensory systems is partitioned is important for research aimed at directly testing such models of population divergence in mate choice.

Guppies (*Poecilia reticulata*) have been a valuable model for studies of the evolution of female mate preferences based on visual signals for nearly 100 years (reviewed in Houde 1997; Magurran 2005). Recently, *P. reticulata* has been shown to vary in the tuning of color vision across populations in a manner that correlates with female mate preferences on the island of Trinidad (Sandkam et al. 2015). In contrast to Trinidad, populations of guppies from mainland South America frequently occur in extreme sympatry and commonly school with two of their closest relatives, *Poecilia picta* and *P. parae* (Liley 1965; Breden et al. 1999). These species occupy similar ecological niches and have highly similar morphometrics, with the largest differences between species being male coloration (Liley 1965; Lindholm et al. 2004; Magurran and Ramnarine 2005; Lindholm et al. in review). In all three species, female mate preferences largely rely on male visual cues, yet the male traits preferred by females differ across species (Haskins and Haskins 1949; Liley 1965; Lindholm et al. 2004; Hurtado-Gonzales et al. 2014). Males of *P. reticulata* have highly variable numbers and colors of spots on their body (Haskins et al. 1961; Liley 1965; Luyten and Liley 1985). In *P. parae*, males occur in one of five discreet Y-linked morphs: three uni-color morphs (with color pattern dominated by a horizontal stripe that is either: red, yellow, or blue), one female mimic morph (colored like the female and engaging in sneak copulations),

and one large, aggressive morph (with vertical dark bars and most of its coloration on the caudal fin; Lindholm et al. 2004; Hurtado-Gonzales and Uy 2009). Males of *P. picta* all have an orange stripe on their caudal fin and yellow bands on their dorsal fin (Liley 1965). While most males of *P. picta* have no additional color, some males occur as a morph with red running diffusely throughout the entire body, although this red coloration doesn't strongly influence mating (Lindholm et al. in review) (see Supplementary Figure 5.1 for pictures of the male morphs for all three species). While males occasionally perform courtship displays for heterospecific females, female *P. picta* do not accept heterospecific males as mates (Liley 1965; Magurran and Ramnarine 2004; 2005). The variable role of color in female mate choice (Rodd et al. 2002; Bourne and Watson 2009), high similarity in niche and morphology, and occurrence in sympatry make these species, *P. reticulata*, *P. picta*, *P. parae*, an excellent system with which to examine whether color vision varies more across species or populations.

Color vision is accomplished by comparing the signals from different cone cells in the retina, which are maximally tuned to different wavelengths of light (Gegenfurtner and Sharpe 1999). The wavelength at which a cone cell maximally detects light is primarily determined by the transmembrane protein expressed, called an opsin (Yokoyama 2000; 2002). *P. reticulata*, *P. picta*, and *P. parae* have an astounding nine cone opsin proteins, among the highest for vertebrates (Hoffmann et al. 2007; Ward et al. 2008; Watson et al. 2011). Each cone opsin is coded by a single gene which is grouped and named for the range of light they detect: SWS1 (SWS1, short wavelength-sensitive) detects ultra-violet; SWS2A and SWS2B (SWS2, short wavelength sensitive 2) detect blues and purples; RH2-1 and RH2-2 (RH2, rhodopsin-like) detect greens; and LWS-1, LWS-2, LWS-3, and LWS-R (LWS, long wavelength-sensitive) detect reds and oranges (Ward et al. 2008; Watson et al. 2011; Sandkam et al. 2012; Tezuka et al. 2014).

Differences in tuning of color vision can occur through changes in either gene sequence or expression (reviewed in Horth 2007). Guppies have been shown to vary in tuning of color vision across populations through both differences in the frequency of an allele known to affect tuning of LWS-1 (Tezuka et al. 2014; Sandkam et al. 2015) and also differences in opsin expression profiles (Sandkam et al. 2015). Opsin expression profiles provide an estimate of cone cell proportions in the retina and thereby offer an

excellent measure of the allocation of an individual's color vision repertoire to different cone cell types (Fuller et al. 2003; Fuller and Travis 2004; Cheng and Flamarique 2004; 2007; Parry et al. 2005; Shand et al. 2008; Sandkam et al. 2015). Guppy populations with stronger female preferences for males with more red/orange coloration have higher expression of LWS opsins (Sandkam et al. 2015).

Here, we examined whether there is more variation in visual tuning within species or across species. The occurrence of multiple sympatric locations of the three closely related species *P. parae*, *P. picta*, and *P. reticulata* on mainland South America allows us to examine differences in visual tuning of all three species from the same environment in a replicated manner across populations. By examining visual tuning of each species across multiple locations we are able to assess the variability of opsin expression across populations and compare this to species differences.

5.3. Materials and Methods

5.3.1. Sample Collection

We sampled *Poecilia parae*, *P. picta*, and *P. reticulata* from four sympatric locations within Guyana between 10:00 and 16:00 in June-July 2010 (see Supplementary Table 5.1 for GPS coordinates and Supplementary Figure 5.2 for map). One sympatric location (Seawall Trench) was sampled on two days. Efforts were made to collect five adult males and five adult females of each species at every location; however not all species were present at the same density within locations resulting in smaller sample sizes for some collections (see Supplementary Table 5.1 for sample sizes of each species and population). *P. bifurca* is a close relative of *P. reticulata*, *P. picta*, and *P. parae* but does not occur sympatrically. Opsin expression for one population of *P. bifurca* (~79 km from closest sympatric location sampled) is reported here only as a qualitative comparison and is not included in statistical analyses. All four species occur in similar environments; small drainage ditches of Guyana that are usually only a few meters wide and less than a meter deep (Liley 1965). Adult males show pronounced species specific coloration, while females are all grey with minor differences in black patterning around the urogenital opening; these traits allow us to rapidly perform

visual identification of species and sex (Rosen and Bailey 1963). Sampling followed the protocols of Sandkam et al. (2015). Briefly, adult fish were caught with dip nets—individuals were rapidly sacrificed in an overdose of MS-222, measured and photographed. We immediately removed eyes and made a small puncture to facilitate complete penetration of RNAlater® Stabilization Solution (Life Technologies™). Both eyes from an individual were placed into a vial of RNAlater® and kept on ice for 24 hours, to allow tissue to be saturated per manufacturer's recommendation. After 24 hours, we transferred the vials to liquid nitrogen. The vials were removed from liquid nitrogen just prior to being placed in checked baggage and flown to Simon Fraser University where we placed them in a -20°C freezer until RNA extraction. Time spent at room temperature totalled less than 24 hours and fell well under the one-week maximum suggested by manufacturer. The bodies of individuals sampled were placed in tubes of 95% EtOH buffered with EDTA and kept at -20ºC until DNA extraction.

5.3.2. qPCR Assay Design

To measure opsin expression, we designed qPCR assays following the methods of Sandkam et al. (2015). We modified the primers reported in Sandkam et al. for *P. reticulata*, such that one set of qPCR assays could be used across all four species. Sequences from *P. reticulata*, *P. picta*, *P. parae*, and *P. bifurca* were aligned and viewed using SeqMan Pro (Lasergene 8.0; DNASTAR, Madison, WI). We designed probe based PrimeTime® qPCR assays (IDT® Technologies) in regions of conserved sequence such that there were no SNPs between any of the four species or primer/probe combinations. Assays were designed to be specific for each of the nine opsins, one rhodopsin (RH1), and three housekeeping genes (beta actin (B-actin); cytochrome c oxidase subunit I (COI); myosin heavy chain (Myosin HC)). Whenever possible, primers spanned intron-exon boundaries. Each assay consisted of a forward primer, reverse primer and 5' FAM labeled probe with both 3' Iowa Black® and internal ZEN™ quenchers (IDT® Technologies) (for primer/probe sequences and product length see Supplementary Table 5.2). Assay specificity for each species was verified by the presence of a single band when running PCR products on an agarose gel. Within each species, LWS-1 and LWS-R assays resulted in products of the same size, while LWS-1 and LWS-3 loci are similar in sequence. To ensure that LWS-1, LWS-3, and LWS-R were truly locus specific

assays, we measured the pairwise covariance of these three assays on the final relative$_{(hk)}$ data set (described below) using R v3.0.2 (R Development Core Team 2014). If assays were binding to non-specific targets we would expect to see large positive covariances between assays. We found no substantial covariance in any of the four species between either LWS-1 and LWS-R (covariance: *P. reticulata*, -9*10-6; *P. parae*, 0.0011; *P. picta*, -0.001; *P. bifurca*, -0.0028) or LWS-1 and LWS-3 (covariance: *P. reticulata*, -1*10-6; *P. parae*, 0.0066; *P. picta*, 0.0004; *P. bifurca*, 0.0059), demonstrating locus specificity of the LWS assays.

We determined the relative PCR efficiency (E_i) for each assay as in Sandkam et al. (2015), using four gBlocks® Gene Fragments (synthetic double stranded, sequence-verified genomic blocks made by IDT® Technologies). gBlocks® were designed on sequence from *P. reticulata* and contained sequence for each of the genes being assayed from 20 bp upstream of the forward primer to 20 bp downstream of the reverse primer. To ensure equal proportions of each gene when calculating relative efficiency, we adjusted the length of each opsin gBlocks® to 728 bp by adding upstream and downstream sequence from the first and last opsin. There were less than seven SNPs per gene across the four species in regions spanned by the assays and none of these differences occurred in primer/probes sites. Relative primer efficiencies calculated using these constructs were used for all four species. Gene order in the gBlocks® was randomized: gBlock® 1 contained LWS-1, RH1, SWS1, LWS-R; gBlock® 2 contained SWS2B, LWS-3, RH2-2; gBlock® 3 contained LWS-2, SWS2A, RH2-1; and gBlock® 4 contained B-actin, COI, Myosin-HC. The 4 gBlocks® were mixed in equal proportions and brought to a concentration of 0.001 ng/µl resulting in a control with equal ratios of all the opsin and housekeeping genes. We ran six replicates of each assay using 4.5 µl of the control. The relative primer efficiencies (E_i) were then calculated following Carleton and Kocher (2001) using the equation:

$$\frac{(1+1)^{Ct_{High}}}{(1+E_i)^{Ct_i}} = 1$$

such that Ct_{High} is the critical threshold of the opsin with the highest expression (lowest *Ct* value) and Ct_i is the critical threshold for opsin *i*. The mean relative efficiency and standard error was calculated across the six replicates (Supplementary Table 5.2).

The SWS2A assay had the highest relative efficiency and was used to measure absolute efficiency. A thousand fold serial dilution was made of a random sample of *P. picta*. Three replicates of qPCR were performed on each concentration using the SWS2A assay. The absolute efficiency of COI was found using the slope of ln(concentration) plotted on *Ct* such that $E_{SWS2A} = e^{-slope} - 1$. The absolute efficiencies of the other primer/probes were determined following Fuller et al. (2004) using the equation:

$$\text{absolute } E_i = \left(\text{relative } E_i \times \text{absolute } E_{SWS2A}\right) \Big/ \text{relative } E_{SWS2A}$$

5.3.3. Sample Processing and Analyses for Opsin Expression

Sample processing followed Sandkam et al. (2015). We placed both eyes from one individual in 600 uL of TRIzol® reagent (Life Technologies™) and ground them with a 1.5 mL RNase-free Kontes® Pellet Pestle Grinder (Kimble Chase). Solution was then run through an Ambion® Homogenizer (Life Technologies™) to reduce viscosity. We extracted RNA following the manufacturer's instructions using PureLink® RNA Mini Kits with the addition of on column treatments with PureLink® DNase (Life Technologies™) to eliminate any potential genomic contamination during qPCR. To verify quality of extracted RNA we ran a subset of samples on an Experion Bioanalyzer. RNA concentrations were adjusted to 50 ng/uL using UltraPure™ DNase/RNase-Free Distilled Water (Life Technologies™). For each sample 500 ng RNA was reverse transcribed using a High Capacity cDNA Reverse Transcription Kit with RNase Inhibitor (Life Technologies™) following manufacturer's instructions. cDNA samples were diluted roughly 20-fold using UltraPure™ DNase/RNase-Free Distilled Water (Life Technologies™) for use in qPCR reactions. Triplicate qPCR reactions were run on each individual for the nine opsins, one rhodopsin, and three housekeeping genes using the qPCR probe based assays described above. We ran all 39 reactions for each individual

simultaneously on the same 384 well plate in addition to negative controls (UltraPure™ water) for each assay. Each 10 uL reaction consisted of: 5 uL Brilliant III Ultra-Fast qPCR Master Mix (Agilent Technologies), 0.5 uL FAM labeled assay (described above) and 4.5 uL sample. We set up all reactions on ice and the plates were briefly spun down before being run on an Applied Biosystems® 7900HT qPCR machine (Life Technologies™). PCR conditions were as follows: 95°C for 3:00 followed by 40 cycles of 95°C for 0:05, 60°C for 0:15. The standard deviation of the triplicate reactions was taken and when >2, outliers were removed (comprising only 5% of the 5538 reactions for this study).

We assessed differences in color vision by calculating the proportion of total opsin expression (T_{all}) made up of each opsin (T_i) following Fuller et al. (2004) and Carleton and Kocher (2001) with the following equation:

$$\frac{T_i}{T_{all}} = \frac{\left(1 \Big/ \left(\left(1+E_i\right)^{Ct_i}\right)\right)}{\sum\left(1 \Big/ \left(\left(1+E_i\right)^{Ct_i}\right)\right)}$$

where E_i is the mean primer/probe efficiency of assay i and Ct_i is the mean critical cycle number for gene i (expression of the nine opsins adds to one for each individual).

We assessed differences in regulation of the nine opsin and one rhodopsin genes by comparing expression relative to housekeeping genes (T_{House}). To control for random variation in housekeeping gene expression, we took the average of three housekeeping genes (Vandesompele et al. 2002). We calculated measures following Sandkam et al. (2015) using the equation:

$$\frac{T_i}{\overline{T}_{House}} = \frac{\left(1 \Big/ \left(1+E_i\right)^{Ct_i}\right)}{\left(\left(\sum\left(1 \Big/ \left(1+E_{House}\right)^{Ct_{House}}\right)\right) \Big/ 3\right)}$$

where E_{House} is the primer/probe efficiency for a housekeeping gene and Ct_{House} is the critical cycle number for that gene.

This resulted in each individual having two measures for each opsin: the proportion of total opsin expression made up by that opsin (proportional), and relative to housekeeping genes ($relative_{(hk)}$). As opsins are the major differentiating character of cone cell types, and color vision is accomplished by comparing the signal from different cone cell types, proportional measures of opsin expression provide a measure of color vision (Fuller et al. 2004; Fuller and Claricoates 2011). Differences in regulation of individual opsins compared to overall gene transcription are revealed by $relative_{(hk)}$ measures of opsin expression, whereas overall gene activity is measured as the mean of the three housekeeping genes (Fuller and Claricoates 2011; Sandkam et al. 2015).

5.3.4. LWS-1 A/S Allele Frequency

Variation in visual systems across species or populations can also occur through differences in the frequency of polymorphisms known to alter tuning of the visual pigments. Only LWS-1 is known to possess a polymorphism that alters spectral tuning in *P. parae*, *P. picta*, or *P. reticulata* (Ward et al. 2008; Tezuka et al. 2014; Sandkam et al. 2015). The key difference between these alleles is the presence of either a serine or alanine as the amino acid at the position that corresponds to residue 180 in human M/LWS opsins (termed 'LWS-1 (180 Ser)' and 'LWS-1 (180 Ala)' respectively). This polymorphism can result in a change of tuning of the LWS-1 opsin protein by up to 15nm (Ward et al. 2008). To determine species and population frequencies of the LWS-1 (180 Ala) and LWS-1 (180 Ser) polymorphism we extracted genomic DNA from tail tissue of the same individuals we used for expression analyses using a DNeasy blood and tissue kit (Qiagen). We generated PCR products using 5' and 3' UTR-specific primers of the LWS-1 locus. Internal sequencing primers were used directly on PCR products to generate chromatograms, spanning part of exon 2 and all of exon 3, at Molecular Cloning Laboratories (McLab, San Francisco, CA) (for primer sequences see Supplementary Table 5.3). We viewed and analyzed the sequencing chromatograms using SeqMan Pro (LASERGENE 8.0; DNASTAR, Madison, WI). For each individual, we determined if the 180 amino acid residue was either a serine, alanine or heterozygous. We calculated genotype and allele frequencies for each species in each location. F_{ST} values based on the frequency of alleles with a serine at amino acid position 180 were

calculated two ways; between species within the same location, and within species across populations. F_{ST} values were calculated as in Sandkam et al. (2015), using:

$$F_{ST} = \frac{\mathrm{var}(S)}{\overline{S} \times (1 - \overline{S})}$$

where var(*S*) is the variance of the frequency of the serine allele, either across species (within locations) or across populations (within species), and $\overline{S}$ is the frequency of the serine allele in either the location or species (respectively).

5.3.5. Statistical Analyses of Opsin Expression

Do species and/or populations differ in colour vision?

To test whether closely related sympatric species differ in either proportional or relative$_{(hk)}$ measures of opsin expression profiles, we ran MANOVAs with the levels for each of the nine cone opsins log transformed and treated as dependent variables. Sex was nested within Species and since Locations were sampled at different times of the day, which can impact opsin expression (Sandkam et al. 2015), we further nested Species within Location. All expression data were analyzed using R v3.0.2 (R Development Core Team 2014).

Sex had no main or interaction effect (Table 1). There are no differences between sexes in opsin expression of natural guppy populations on Trinidad (Sandkam et al. 2015), nor for opsin expression in most other species in which it has been examined (such as stickleback (*Gasterosteus aculeatus*)- Boulcott and Braithwaite 2006; and killifish (*Lucania goodei*)- Johnson et al. 2013). Therefore we pooled sexes for all of the following analyses.

Does opsin expression differ more by species or locations?

Akaike Information Criterion (AIC) scores were calculated by Species, Location, or a Species*Location interaction to determine the ability of models built on these factors to explain variation in proportional expression for each of the opsin genes.

How do species differ in colour vision?

To better understand differences in colour vision across species within the same population, independent ANOVAs on data subset by population were run by Species. Opsins that significantly differed in ANOVAs were followed up with a Tukey test to determine which of the species contributed to the significant effects.

5.4. Results

5.4.1. *Do species and/or populations differ in color vision?*

Across all populations and species SWS2A, RH2-2, LWS-2, LWS-3, and LWS-R showed rather low levels of expression compared to other opsin genes (Figures 5.1, 5.2). The most abundant opsin expressed was either SWS1 or LWS-1 but varied across populations and species (Figure 5.1). MANOVAs revealed that colour vision, as measured by differences in opsin expression, varied significantly within species across populations and within populations across species (Table 5.1). However, expression profiles did not vary by sex (nested in species-in populations) for either proportional ($F_{12,109}$ = 1.1275, P = 0.1872) or relative$_{(hk)}$ ($F_{12,109}$ = 0.977, P = 0.5490) measures.

5.4.2. *Does opsin expression differ more by species or location?*

We found opsin expression to vary significantly by Location and Species nested within Location. To determine if Species, Location or a Location*Species interaction best explained variation in proportional expression we calculated AIC scores. Location alone best explained variation in SWS1, SWS2B, RH2-1, RH2-2, and LWS-1 (Table 5.2). Location and Location*Species explained the variation equally well for SWS2A and LWS-3, as seen by AIC scores differing by less than 2 for these two models in both genes. Location and Species explained variation of LWS-R equally well with AIC scores within 1 for these two models. Only for LWS-2 did Species best explain variation in opsin expression (Table 5.2). Overall, we found that variation in color vision, as measured by opsin expression, was better explained by Location than by Species.

5.4.3. *How do species differ in color vision?*

When relative$_{(hk)}$ measures of opsin expression differed by species, *P. parae* generally had highest expression (Figure 5.2). Despite differences in the abundance of opsin expression, colour vision was highly similar across species, as indicated by few occurrences in which locations differed in proportional measures of opsin expression across species (10 of 45 possible opsin by location combinations tested) (Table 5.3).

5.4.4. Gene Frequencies of the 180 Ala vs. Ser Allele Vary Across Species

We found within-species gene frequencies were similar across locations, but across-species gene frequencies differed within locations such that *P. reticulata* showed higher frequencies of alleles with LWS-1 (180 Ala) than either *P. parae* or *P. picta* (Figure 5.3). The F_{ST} scores we observed confirmed low differences within species in the frequency of the serine allele across locations, with the most variation occurring in *P. parae* (F_{ST} = 0.1304) (Table 5.4). The within location differences in frequency of the serine allele observed across species were moderate, as demonstrated by F_{ST} scores between 0.336 – 0.608 (Table 5.4). This is a result of the high frequency of LWS-1 (180 Ala) in *P. reticulata* while it is nearly absent in *P. parae* and *P. picta*. However, while qualitative differences in frequency and F_{ST} scores are clear, one should take care when interpreting the results quantitatively as the samples sizes were low (ranging from 2-20 individuals) (Figure 5.3).

5.5. Discussion

Intersexual mate choice was first proposed as a mechanism for speciation through divergence in the traits preferred (Darwin 1871). Over time, divergence of populations has widely been accepted as the first step toward speciation (Coyne and Orr 2004). Therefore, understanding the factors involved in population divergence of mate choice has far reaching implications to many fields including animal behaviour and evolutionary biology. Sensory Bias models for the evolution of mate preference place a strong emphasis on the role of sensory system variation in mate preferences (Ryan and

Rand 1993; Basolo and Endler 1995; Endler and Basolo 1998; Boughman 2002; Ryan and Cummings 2013). However, the extent to which sensory systems vary across, versus within, species remains largely unknown. Here we describe variation in color vision across three closely related sympatrically occurring species. We found that color vision varies across populations but does not vary consistently between species across locations. Below we describe the variation we found and discuss our findings in the framework of mate choice evolution.

5.5.1. Colour Vision Differs Across Populations Within Species

Visual systems are classically thought to show low to no variation across individuals within a species, especially when color vision is modeled in the context of mate choice (such as Endler 1991; Kemp et al. 2009; Hurtado-Gonzales et al. 2014; Cole and Endler 2015). Yet color vision can vary through differences in the abundance of different cone cell types, amino acid sequences of the opsin proteins, or neural processing of signals from the eyes (Horth 2007). Opsin expression profiles provide an estimate of the ratio of cone cells expressing different opsins in the eyes, which can impact color vision (Carleton and Kocher 2001; Fuller et al. 2004; Horth 2007; Sandkam et al. 2015). We present the first study of opsin expression in *P. bifurca*, *P. parae* and *P. picta*, and one of the few to carefully compare within versus between species color vision in similar environments. We demonstrate that color vision in these species, and sympatric *P. reticulata*, likely differs across populations through differences in opsin expression, and could differ through frequency of an alternative amino acid sequence.

Color vision of *P. parae* from nearby Guyanan populations has recently been explored using microspectrophotometry (MSP) to identify the peak wavelength sensitivity (λ_{max}) of the cone cells (Hurtado-Gonzales et al. 2014). While *P. parae* possess nine opsin genes, Hurtado-Gonzales and colleagues only found seven cone cell types across individuals. This can be explained by the low-to-no expression of both SWS2A and LWS-2 we found, emphasizing the important role of expression in tuning the visual system. While our measures of opsin gene expression provide insight to differences in color vision, it is possible that the exact cone cell proportions could differ from measures

of proportional RNA expression through differences in non-transcriptional control such as translation rates.

The long wavelength-sensitive-1 (LWS-1) opsin is the only opsin in Poeciliidae known to vary in tuning due to differences in amino acid sequence. The LWS-1 opsin can either have an Alanine or Serine at the site corresponding to the 180th amino acid in the human opsin sequence and is therefore termed: LWS-1 (180 Ala) and LWS-1 (180 Ser) respectively. The frequency of these two alleles varies across populations in *P. reticulata* (Tezuka et al. 2014; Sandkam et al. 2015). This same variation is also thought to be the only sequence difference in important tuning sites across the species *P. reticulata*, *P. parae*, *P. picta*, and *P. bifurca* where as the later three have, until now, been thought to be fixed for either LWS-1 (180 Ala) or (180 Ser) (Ward et al. 2008). Using the 'five site rule' (Yokoyama and Radlwimmer 1998), Tezuka et al. (2014) estimated the λ_{max} of the LWS-1 (180 Ala) to be ~553 nm and LWS-1 (180 Ser) to be ~560nm. Previously *P. parae* was thought to possess only one LWS-1 allele (180 Ser), however we show that *P. parae* also possesses the alternate LWS-1 (180 Ala), albeit at a low frequency. Interestingly, only 5 out of the 17 individuals (29.4%) examined by Hurtado-Gonzales et al. (2014) possessed cone cells with a λ_{max} of 553 ± 1.9 nm, this matches the predicted λ_{max} of LWS-1 (180 Ala) (Tezuka et al. 2014). The closest location to the populations sampled by Hurtado-Gonzales et al. we sampled for this study was West Patentia, where 22.2% of individuals had at least one copy of LWS-1 (180 Ala). The close proximity and lack of geographic barrier between these two populations makes it likely that allele frequencies are similar in these populations. The close frequencies of individuals with λ_{max} of 553 (Hurtado-Gonzales et al. 2014) and those with the LWS-1 (180 Ala) (this study) suggest variation in allele frequency across populations likely explains the presence/absence of cone cell types across individuals and further demonstrates that color vision varies across individuals in *P. parae*. We found differences across populations in the frequency of LWS-1 (180 Ala) in all three species (except for the allopatric species, *P. bifurca* which appears fixed for LWS-1 (180 Ala)) and demonstrate that color vision differs across populations not only in opsin expression, but also through differences in tuning of the opsins.

The frequency of the LWS-1 (180 Ala) allele varies across *P. reticulata* populations in Trinidad such that low predation populations are nearly fixed for the LWS-1 (180 Ser) allele while high predation populations had a greater frequency of the LWS-1 (180 Ala) allele. All the populations we use in this study are in the lowland region of Guyana where predation is extremely high (predators include birds, snakes, caiman and many fishes). Interestingly, for all of our populations of *P. reticulata*, the frequency of the LWS-1 (Ala) allele was as high or higher than the high predation populations of Trinidad (Sandkam et al. 2015) highlighting the potential role of predation in determining LWS-1 allele frequencies.

We also found the closely related but not sympatric species, *P. bifurca*, has a similar color vision system to *P. picta*, *P. parae* and *P. reticulata*. *P. bifurca* lives in environments that are generally tannin stained and occur further inland, thereby further demonstrating the surprisingly conserved nature of color vision across species in this group.

It should be noted that sites were not all sampled at the same time of day, therefore differences across populations could be confounded with diurnal variation in opsin expression, which has been shown to occur in Trinidadian guppy populations (Sandkam et al. 2015). However, the diurnal differences across locations are expected to occur in all three of the sympatric species; therefore we feel our cross- versus within-species comparisons are justified.

5.5.2. Sensory variation and mate choice divergence

For sensory systems to play a role in mate choice divergence there needs to be variation in sensory systems (Ryan 1990; Ryan and Cummings 2013). We recently found support for the Sensory Exploitation model to explain population divergence in female mate preferences of Trinidadian guppies (*Poecilia reticulata*) (Sandkam et al. 2015) such that populations with stronger female preferences for orange male coloration expressed higher levels of LWS opsins. These results raised two important questions regarding the feasibility of sensory variation to explain mate choice differences: (1) was population variation in sensory systems present before the split of the guppy from the

clade of *P. picta* and *P. parae*, and (2) if there is variation in other species, do visual systems vary more across species or populations?

We found visual systems do vary in both opsin gene expression and LWS-1 allele frequency across populations of *P. parae*, *P. picta*, and *P. reticulata* on mainland South America. Although these three species occur in such extreme sympatry that they frequently school together, they exhibit female preferences for males with different color traits (Haskins and Haskins 1949; Liley 1965; Bourne et al. 2003). All three of these species are known to vary in the frequency of their respective different male morphs across populations (*P. parae*- Hurtado-Gonzales et al. 2014; *P. picta*- Lindholm et al. in review; *P. reticulata*- Alexander et al. 2004). It will be especially interesting for future work to determine if the variation we found in visual systems across populations correlates with differences in mate choice.

As more species are shown to vary in sensory systems across populations [e.g. Bluefin killifish (Fuller et al. 2004), cichlids (Smith et al. 2010), stickleback (Flamarique et al. 2013), sand goby (Larmuseau et al. 2009), pied flycatchers (Lehtonen et al. 2011), guppies (Sandkam et al. 2015)], Sensory Bias models become a more likely candidate to explain divergence in mate preferences across species. This raises the question of whether differences in mate preference across species are maintained through consistent differences in their peripheral sensory systems. We found variation in color vision, as measured by opsin expression, is better explained by location than by species for most of the opsins, showing no consistent differences across species. Our data suggest mate preferences are likely maintained at another level of preference (such as higher order processing). However, we did find that the frequency of the LWS-1 (180 Ala) allele not only differs across populations, but also across species. This raises the possibility that the minor tuning differences between the LWS-1 (180 Ala) and LWS-1 (180 Ser) alleles could also play a role in mate choice differences across species. It will be interesting for future work in these species to investigate the correlation between color preference and genotype to further our understanding of visual tuning and mate preference.

5.6. Conclusion

We show that color vision, as measured by opsin expression, differs more across populations of the same species than across species in the same location. These differences provide support for Sensory Bias models to explain population divergence in mate preference since these models rely on sensory system differences across populations. Opsin expression did not differ consistently between species across locations, suggesting species level differences in mate preference are likely maintained at levels of higher order processing.

Acknowledgements

We wish to thank Orrin Clarke, Selwin France, and Joyce Wade for assistance in the field; Dr. William Davidson and Dr. Krzysztof Lubieniecki for use of their qPCR machine at Simon Fraser University; and CEIBA Biological Center for hosting us during fieldwork. We acknowledge the Guyana EPA for issuing research permit # 120710 BR 135.

References

Alexander, H. J., J. S. Taylor, S. S. T. Wu, and F. Breden. 2006. Parallel evolution and vicariance in the guppy (*Poecilia reticulata*) over multiple spatial and temporal scales. *Evolution* 60:2352–2369.

Andersson, M. B., and L. W. Simmons. 2006. Sexual selection and mate choice. *Trends Ecol Evol* 21:296–302.

Basolo, A. L., and J. A. Endler. 1995. Sensory Biases and the evolution of sensory systems. *Trends Ecol Evol* 10:489–489.

Boughman, J. W. 2002. How sensory drive can promote speciation. *Trends Ecol Evol* 17:571–577.

Boulcott, P., and V. A. Braithwaite. 2006. Colour perception in three-spined sticklebacks: sexes are not so different after all. *Evolutionary Ecology* 21:601–611.

Bourne, G. R., F. Breden, and T. C. Allen. 2003. Females prefer carotenoid colored males as mates in the pentamorphic livebearing fish, *Poecilia parae*. *Naturwissenschaften* 90:402–405.

Bourne, G. R., and L. C. Watson. 2009. Receiver-bias implicated in the nonsexual origin of female mate choice in the pentamorphic fish *Poecilia parae* Eigenmann, 1894. *AACL Bioflux* 2:299–317.

Breden, F., M. B. Ptacek, M. Rashed, D. Taphorn, and C. A. Figueiredo. 1999. Molecular phylogeny of the live-bearing fish genus Poecilia (Cyprinodontiformes: Poeciliidae). *Mol Phylogenet Evol* 12:95–104.

Carleton, K. L., and T. D. Kocher. 2001. Cone opsin genes of African cichlid fishes: tuning spectral sensitivity by differential gene expression. *Mol Biol Evol* 18:1540–1550.

Cheng, C. L., and I. N. Flamarique. 2004. Opsin expression: new mechanism for modulating colour vision. *Nature* 428:279–279.

Cheng, C. L., and I. N. Flamarique. 2007. Chromatic organization of cone photoreceptors in the retina of rainbow trout: single cones irreversibly switch from UV (SWS1) to blue (SWS2) light sensitive opsin during natural development. *J. Exp. Biol.* 210:4123–4135.

Cole, G. L., and J. A. Endler. 2015. Variable environmental effects on a multicomponent sexually selected trait. *The American Naturalist* 185:452–468.

Coyne, J. A., and H. A. Orr. 2004. *Speciation*. Sinauer Associates, Sunderland, MA.

Darwin, C. 1871. *The Descent of Man and Selection in Relation to Sex*. John Murray, London, UK.

Endler, J. A. 1991. Variation in the appearance of guppy color patterns to guppies and their predators under different visual conditions. *Vision Res* 31:587–608.

Endler, J. A., and A. L. Basolo. 1998. Sensory ecology, receiver biases and sexual selection. *Trends Ecol Evol* 13:415–420.

Flamarique, I. N., C. L. Cheng, C. Bergstrom, and T. E. Reimchen. 2013. Pronounced heritable variation and limited phenotypic plasticity in visual pigments and opsin expression of threespine stickleback photoreceptors. *Journal of Experimental Biology* 216:656–667.

Fuller, R. C., and J. Travis. 2004. Genetics, lighting environment, and heritable responses to lighting environment affect male color morph expression in bluefin killifish, *Lucania goodei*. *Evolution* 58:1086–1098.

Fuller, R. C., and K. M. Claricoates. 2011. Rapid light-induced shifts in opsin expression: finding new opsins, discerning mechanisms of change, and implications for visual sensitivity. *Mol Ecol* 20:3321–3335.

Fuller, R. C., K. L. Carleton, J. M. Fadool, T. C. Spady, and J. Travis. 2004. Population variation in opsin expression in the bluefin killifish, *Lucania goodei*: a real-time PCR study. *J Comp Physiol A* 190:147–154.

Fuller, R. C., L. J. Fleishman, M. Leal, J. Travis, and E. R. Loew. 2003. Intraspecific variation in retinal cone distribution in the bluefin killifish, *Lucania goodei*. *J Comp Physiol A* 189:609–616.

Gegenfurtner, K. R., and L.T. Sharpe. 1999. *Color Vision: From Genes to Perception*. Cambridge University Press, Cambridge, UK.

Haskins, C. P., and E. F. Haskins. 1949. The role of sexual selection as an isolating mechanism in three species of poeciliid fishes. *Evolution* 3:160–169.

Haskins, C. P., E. F. Haskins, J. J. A. McLaughlin, and R. E. Hewitt. 1961. Polymorphism and population structure in *Lebistes reticulatus*, an ecological study. Pp. 320–395 in W. F. Blair, ed. *Vertebrate Speciation*.

Hoffmann, M., N. Tripathi, S. R. Henz, A. K. Lindholm, D. Weigel, F. Breden, and C. Dreyer. 2007. Opsin gene duplication and diversification in the guppy, a model for sexual selection. *Proc. R. Soc. B* 274:33–42.

Horth, L. 2007. Sensory genes and mate choice: Evidence that duplications, mutations, and adaptive evolution alter variation in mating cue genes and their receptors. *Genomics* 90:159–175.

Houde, A. E. 1997. *Sex, color, and mate choice in guppies*. Princeton University Press. Princeton, NJ.

Hurtado-Gonzales, J. L., and J. A. C. Uy. 2009. Alternative mating strategies may favour the persistence of a genetically based colour polymorphism in a pentamorphic fish. *Anim Behav* 77:1187–1194.

Hurtado-Gonzales, J. L., E. R. Loew, and J. A. C. Uy. 2014. Variation in the visual habitat may mediate the maintenance of color polymorphism in a poeciliid fish. *PLoS ONE* 9:e101497.

Johnson, A. M., S. Stanis, and R. C. Fuller. 2013. Diurnal lighting patterns and habitat alter opsin expression and colour preferences in a killifish. *Proc. Biol. Sci.* 280:20130796.

Kemp, D. J., D. N. Reznick, G. F. Grether, and J. A. Endler. 2009. Predicting the direction of ornament evolution in Trinidadian guppies (*Poecilia reticulata*). *Proc. Biol. Sci*. 276:4335–4343.

Kuijper, B., I. Pen, and F. J. Weissing. 2012. A guide to sexual selection theory. *Annu Rev Ecol Evol S* 43:287–311.

Larmuseau, M. H. D., J. A. M. Raeymaekers, K. G. Ruddick, J. K. J. Van Houdt, and F. A. M. Volckaert. 2009. To see in different seas: spatial variation in the rhodopsin gene of the sand goby (*Pomatoschistus minutus*). *Mol Ecol* 18:4227–4239.

Lehtonen, P. K., T. Laaksonen, A. V. Artemyev, E. Belskii, P. R. Berg, C. Both, L. Buggiotti, S. Bureš, M. D. Burgess, A. V. Bushuev, I. Krams, J. Moreno, M. Mägi, A. Nord, J. Potti, P.-A. Ravussin, P. M. Sirkiä, G.-P. Sætre, W. Winkel, and C. R. Primmer. 2011. Candidate genes for colour and vision exhibit signals of selection across the pied flycatcher (*Ficedula hypoleuca*) breeding range. *Heredity* 108:431–440.

Liley, N. R. 1965. Ethological isolating mechanisms in four sympatric species of poeciliid fishes. *Behaviour. Supplement* 13:1–197.

Lindholm, A. K., R. Brooks, and F. Breden. 2004. Extreme polymorphism in a Y-linked sexually selected trait. *Heredity* 92:156–162.

Lindholm, A. K., B. A. Sandkam, K. Pohl, and F. Breden. *Poecilia picta*, a close relative to the guppy, exhibits red male coloration polymnot orphism: a system for phylogenetic comparisons. *PLoS ONE* - In review.

Luyten, P. H., and N. R. Liley. 1985. Geographic variation in the sexual behaviour of the guppy, *Poecilia reticulata* (Peters). *Behaviour* 95:164-179.

Magurran, A. E. 2005. *Evolutionary Ecology: The Trinidadian Guppy*. Oxford University Press, New York.

Magurran, A. E., and I. Ramnarine. 2005. Evolution of mate discrimination in a fish. *Curr Biol* 15:R867.

Magurran, A. E., and I. W. Ramnarine. 2004. Learned mate recognition and reproductive isolation in guppies. *Anim Behav* 67:1077–1082.

Panhuis, T. M., R. Butlin, M. Zuk, and T. Tregenza. 2001. Sexual selection and speciation. *Trends Ecol Evol* 16:364–371.

Parry, J. W. L., K. L. Carleton, T. C. Spady, A. Carboo, D. M. Hunt, and J. K. Bowmaker. 2005. Mix and match color vision: tuning spectral sensitivity by differential opsin gene expression in Lake Malawi cichlids. *Curr Biol* 15:1734–1739.

R Development Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rodd, F. H., K. A. Hughes, G. F. Grether, and C. T. Baril. 2002. A possible non-sexual origin of mate preference: are male guppies mimicking fruit? *Proc. R. Soc. B* 269:475–481.

Rosen, D. E., and R. M. Bailey. 1963. The Poeciliid fishes (Cyprinodontiformes), their structure, zoogeography, and systematics. *Bulletin of the America Museum of Natural History* 126:1–176.

Ryan, M. J. 1990. Sexual selection, sensory systems and sensory exploitation. *Oxford Surveys in Evolutionary Biology* 7:157–195.

Ryan, M. J., and A. S. Rand. 1993. Sexual selection and signal evolution: The ghost of biases past. *Philos. Trans. R. Soc. Lond., B* 340:187–195.

Ryan, M. J., and M. E. Cummings. 2013. Perceptual biases and mate choice. *Annu Rev Ecol Evol S* 44:437–459.

Sandkam, B. A., J. B. Joy, C. T. Watson, P. Gonzalez-Bendiksen, C. R. Gabor, and F. Breden. 2012. Hybridization leads to sensory repertoire expansion in a gynogenetic fish, the Amazon Molly (*Poecilia formosa*): a Test of the hybrid-sensory expansion hypothesis. *Evolution* 67:120–130.

Sandkam, B., C. M. Young, and F. Breden. 2015. Beauty in the eyes of the beholders: colour vision is tuned to mate preference in the Trinidadian guppy (*Poecilia reticulata*). *Mol Ecol* 24:596–609.

Shand, J., W. L. Davies, N. Thomas, L. Balmer, J. A. Cowing, M. Pointer, L. S. Carvalho, A. E. O. Trezise, S. P. Collin, L. D. Beazley, and D. M. Hunt. 2008. The influence of ontogeny and light environment on the expression of visual pigment opsins in the retina of the black bream, *Acanthopagrus butcheri*. *J. Exp. Biol.* 211:1495–1503.

Smith, A. R., L. D'Annunzio, A. E. Smith, A. Sharma, C. M. Hofmann, N. J. Marshall, and K. L. Carleton. 2010. Intraspecific cone opsin expression variation in the cichlids of Lake Malawi. *Mol Ecol* 20:299–310.

Tezuka, A., S. Kasagi, C. Van Oosterhout, M. McMullan, W. M. Iwasaki, D. Kasai, M. Yamamichi, H. Innan, S. Kawamura, and M. Kawata. 2014. Divergent selection for opsin gene variation in guppy (*Poecilia reticulata*) populations of Trinidad and Tobago. *Heredity* 113:381–389.

Vandesompele, J., K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe, and F. Speleman. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3:RESEARCH0034.

Ward, M. N., A. M. Churcher, K. J. Dick, C. R. Laver, G. L. Owens, M. D. Polack, P. R. Ward, F. Breden, and J. S. Taylor. 2008. The molecular basis of color vision in colorful fish: Four Long Wave-Sensitive (LWS) opsins in guppies (*Poecilia reticulata*) are defined by amino acid substitutions at key functional sites. *BMC Evol Biol* 8.

Watson, C. T., S. M. Gray, M. Hoffmann, K. P. Lubieniecki, J. B. Joy, B. A. Sandkam, D. Weigel, E. R. Loew, C. Dreyer, W. S. Davidson, and F. Breden. 2011. Gene duplication and divergence of long wavelength-sensitive opsin genes in the Guppy, *Poecilia reticulata*. *J Mol Evol* 72:240–252.

Yokoyama, S. 2002. Molecular evolution of color vision in vertebrates. *Gene* 300:69–78.

Yokoyama, S. 2000. Molecular evolution of vertebrate visual pigments. *Progress in Retinal and Eye Research* 19:385–420.

Yokoyama, S., and F. B. Radlwimmer. 1998. The “five-sites” rule and the evolution of red and green color vision in mammals. *Mol Biol Evol* 15:560–567.

Tables and Figures

Table 5.1. Results of MANOVA on opsin expression profiles as proportional and relative$_{(hk)}$ measures. *Loc*, Location; *Sp(Loc)*, Species nested in Location; *Sx(Sp(Loc))*, Sex nested in Species, nested in Location. Bold indicates $P < 0.05$.

		Proportional	Relative$_{(hk)}$
Population	$F_{(3,109)}$	**7.13**	**5.87**
	P	**<0.0001**	**<0.0001**
Species(Population)	$F_{(8,109)}$	**1.61**	**2.04**
	P	**0.0014**	**<0.0001**
Sex(Species(Population))	$F_{(12,109)}$	1.13	0.98
	P	0.1872	0.5490

Table 5.2. AIC values for the ability of models based on Location, Species, or Location*Species to explain variation in proportional measures of opsin expression for each opsin gene. Bold indicates $P < 0.05$.

	df	SWS1	SWS2A	SWS2B	RH2-1	RH2-2	LWS-1	LWS-2	LWS-3	LWS-R
Location*Species	13	-156.12	**-2015.89**	-147.09	-244.66	-847.86	-71.63	-1022.10	**-678.85**	-626.69
Location	5	**-160.52**	**-2014.14**	**-154.32**	**-255.37**	**-855.18**	**-78.62**	-1024.31	**-677.00**	**-630.76**
Species	4	-152.08	-2008.59	-134.21	-250.90	-839.24	-70.27	**-1027.32**	-669.96	**-631.74**

Table 5.3. Results of opsin expression ANOVAs using proportional and relative$_{(hk)}$ measures by location. Bold indicates $P < 0.05$.

		Princess Cemetery		Seawall Trench Day 1		Seawall Trench Day 2		Turkeyen		West Patentia	
Proportional Measures		$F_{(2,27)}$	P	$F_{(2,21)}$	P	$F_{(2,28)}$	P	$F_{(2,23)}$	P	$F_{(2,19)}$	P
	SWS1	1.97	0.1590	**10.11**	**0.0008**	2.15	0.1350	0.21	0.8100	3.03	0.0722
	SWS2A	3.22	0.0556	2.64	0.0951	0.25	0.7810	0.83	0.4500	**23.86**	**<0.0001**
	SWS2B	0.06	0.9460	1.37	0.2750	1.24	0.3060	0.98	0.3911	2.40	0.1170
	RH2-1	0.02	0.9840	**9.94**	**0.0009**	2.85	0.0747	2.09	0.1470	0.28	0.7590
	RH2-2	0.94	0.4030	**4.44**	**0.0247**	0.79	0.4660	1.89	0.1730	**8.56**	**0.0022**
	LWS-1	**3.45**	**0.0464**	**10.03**	**0.0009**	**5.17**	**0.0123**	1.26	0.3020	0.62	0.5470
	LWS-2	1.71	0.1990	**4.88**	**0.0182**	0.86	0.4350	1.20	0.3200	**4.01**	**0.0354**
	LWS-3	0.38	0.6880	2.58	0.0999	2.41	0.1090	0.12	0.8880	0.57	0.5740
	LWS-R	0.60	0.5580	1.65	0.2160	0.96	0.3960	1.03	0.3730	1.59	0.2290
Relative$_{(hk)}$ Measures	SWS1	**8.84**	**0.0011**	**19.99**	**<0.0001**	**4.35**	**0.0227**	1.14	0.3370	**5.30**	**0.0149**
	SWS2A	**4.22**	**0.0253**	0.90	0.4220	0.76	0.4790	0.75	0.4820	**4.47**	**0.0256**
	SWS2B	2.71	0.0847	**9.98**	**0.0009**	**3.85**	**0.0332**	0.73	0.4920	**5.55**	**0.0126**
	RH2-1	**3.46**	**0.0429**	**18.74**	**<0.0001**	0.76	0.4760	2.65	0.0922	**7.49**	**0.0040**
	RH2-2	**4.19**	**0.0259**	2.04	0.1550	1.52	0.2370	0.80	0.4630	0.82	0.4550
	LWS-1	**15.04**	**<0.0001**	2.94	0.0747	2.91	0.0711	**3.99**	**0.0323**	**13.52**	**0.0002**
	LWS-2	**10.05**	**0.0005**	**5.67**	**0.0108**	0.16	0.8560	0.56	0.5800	1.32	0.2030
	LWS-3	1.75	0.1930	**11.42**	**0.0004**	1.65	0.2110	0.68	0.5160	0.44	0.6520
	LWS-R	1.22	0.3110	2.34	0.1210	1.15	0.3310	1.00	0.3820	0.72	0.5000
	RH1	**5.55**	**0.0096**	**5.73**	**0.0103**	2.66	0.0876	**6.19**	**0.0045**	**12.64**	**0.0003**

Table 5.4. **F_{ST} based on the frequency of the 180 Ser allele of LWS-1. Calculated across populations of the same species or across species in the same population.**

Within Species – Across Populations				
	P. parae	*P. picta*		*P. reticulata*
F_{st}	0.130	0.077		0.033
Within Locations – Across Species				
	Princess Cemetery	Seawall Trench	Turkeyen	West Patentia
F_{st}	0.609	0.353	0.385	0.336

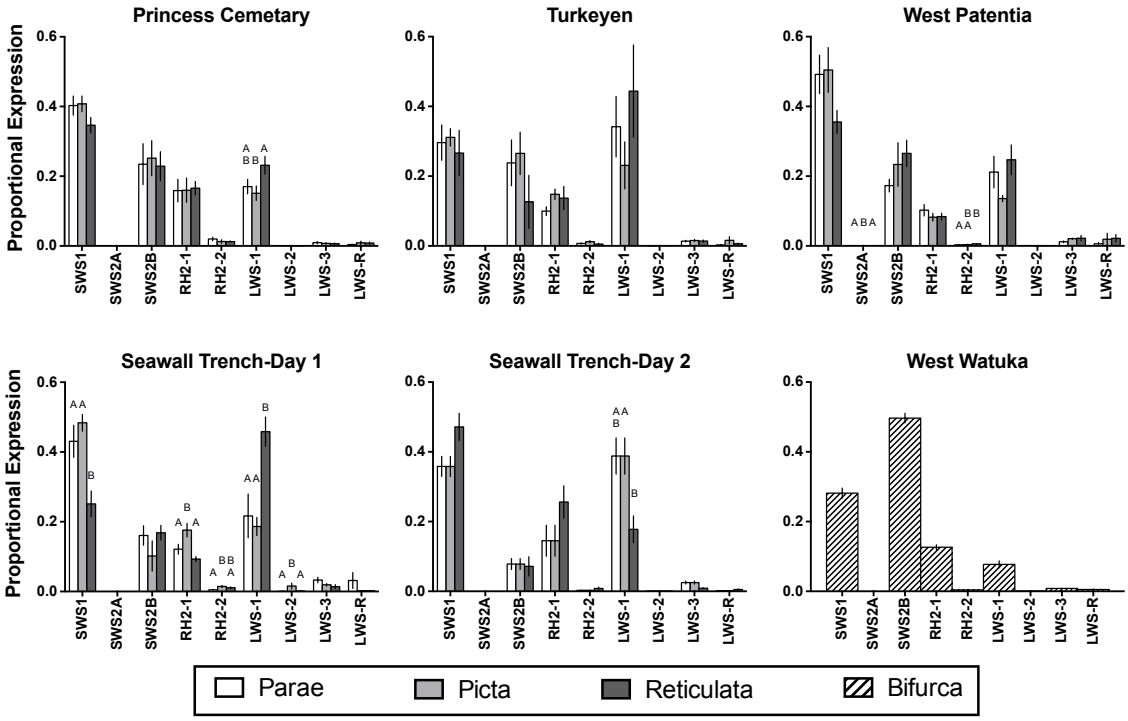


Figure 5.1. Proportional measures of opsin expression across species and locations. Letters denote significant differences in expression of individual opsins across species within a location. Note: West Watuka has only *P. bifurca* and was not used in statistical analyses.

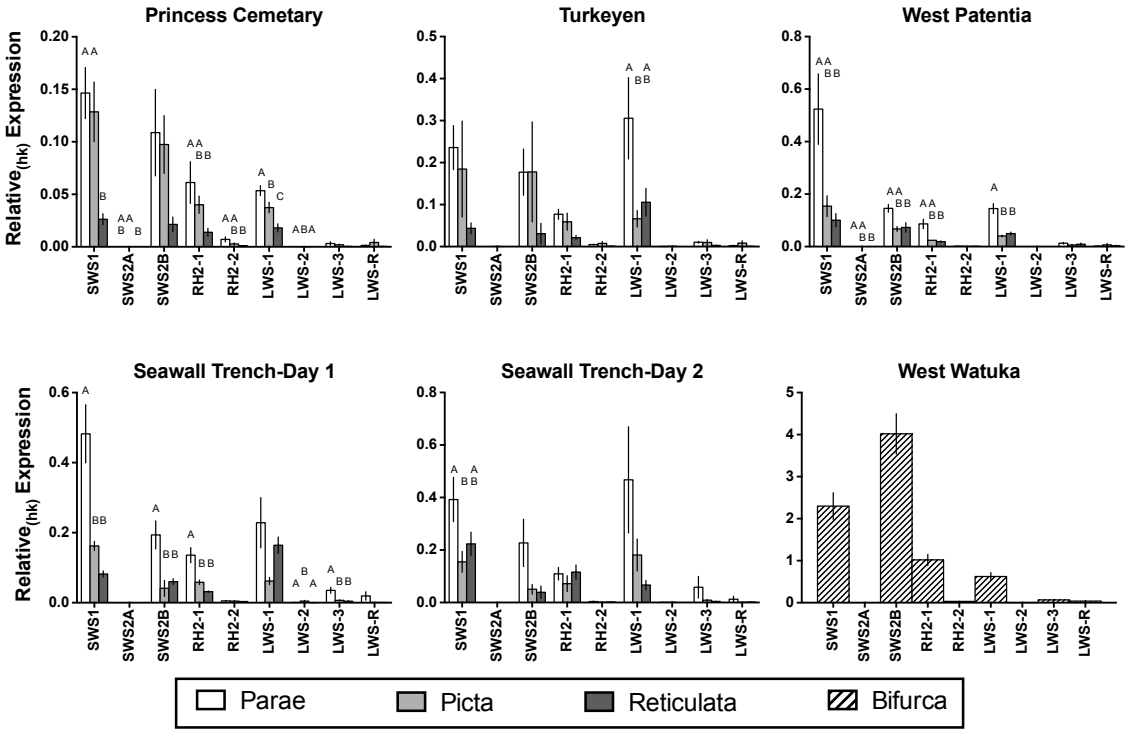


Figure 5.2. **Relative$_{(hk)}$ measures of opsin expression across species and locations. Letters denote significant differences in expression of individual opsins across species within a location. Note: West Watuka has only *P. bifurca* and was not used in statistical analyses.**

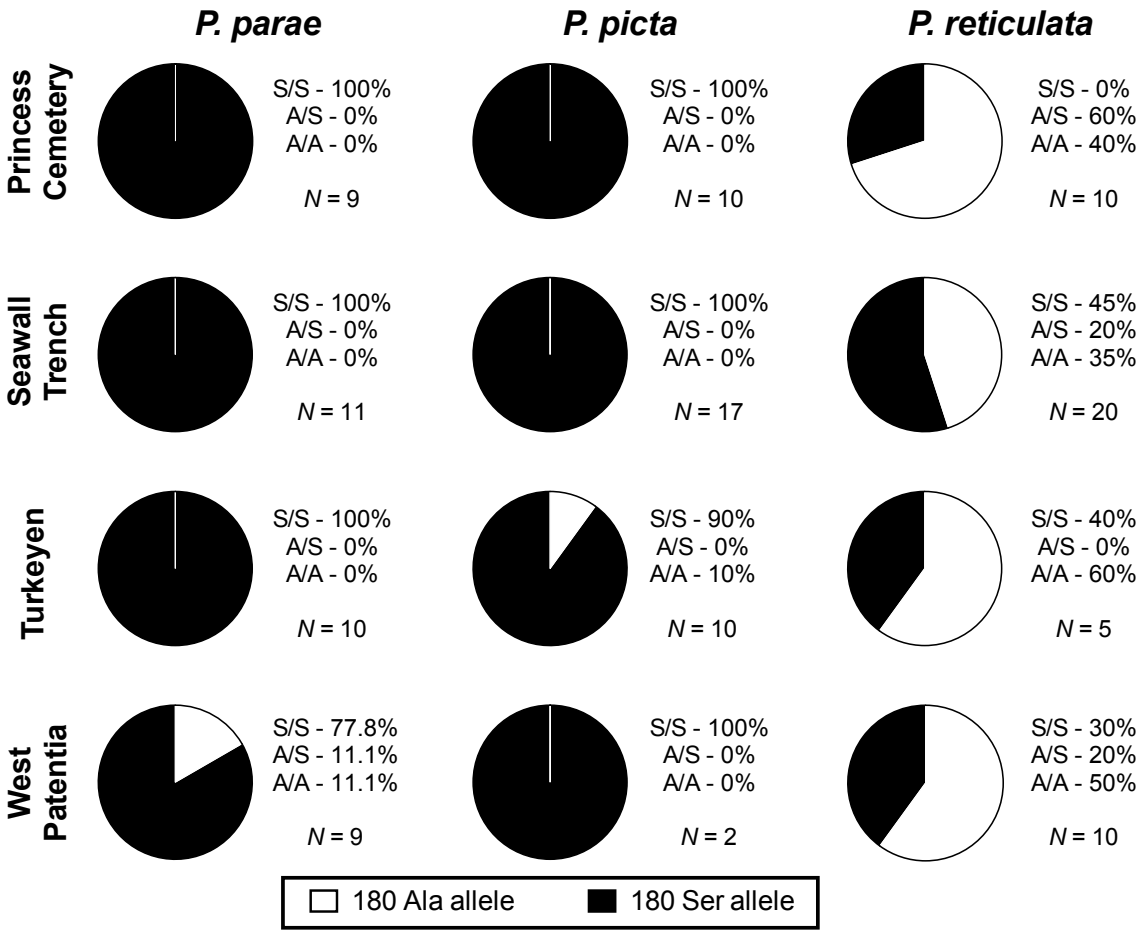


Figure 5.3. Genotype and gene frequencies of the 180 Ala (A) versus 180 Ser (S) allele of LWS-1. LWS, long-wavelength sensitive.

Supplementary Tables

Supplementary Table 5.1. GPS location and sample size for each location. Note: West Watuka has only *P. bifurca* and was not used in statistical analyses.

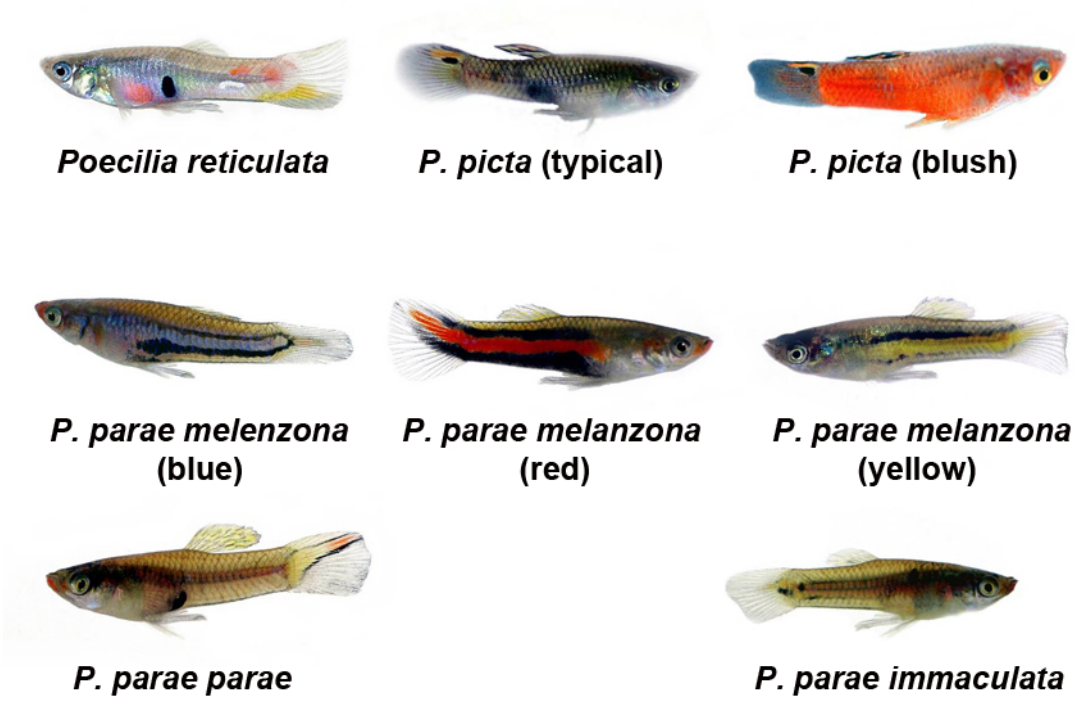
	GPS-N	GPS-W	Date	Time Sampled	Salinity (d20/20)	*P. parae*		*P. picta*		*P. reticulata*		*P. bifurca*	
						M	F	M	F	M	F	M	F
Princess Cemetary	06º 48.045'	058º 09.086'	7/14/10	10:00	1.002	5	5	5	5	5	5	-	-
Seawall Trench – Day1	06º 49.520'	058º 08.637'	7/12/10	15:30	1.003	5	3	5	1	5	5	-	-
Seawall Trench – Day2	06º 49.520'	058º 08.637'	7/13/10	10:00	1.002	5	5	6	5	5	5	-	-
Turkeyen	06º 49.067'	058º 06.764'	7/8/10	13:00	1.000	5	5	5	5	4	2	-	-
West Patentia	06º 41.752'	058º 12.066'	7/11/10	12:00	1.003	5	5	1	1	4	6	-	-
West Watuka	05º 59.509'	058º 18.282'	7/9/10	12:30	1.000	-	-	-	-	-	-	4	5

Supplementary Table 5.2. Relative and absolute efficiency (% efficiency) of qPCR assays.

	Forward Primer	Probe	Reverse Primer	Product Length	Relative Efficiency	Standard Error	Absolute Efficiency
B-actin	GCTTGTGCGGGATATCATTTG	CCCACACCCAAAGTTCAGCCATG	GTATGTGCAAAGCCGGATTC	137	0.7593	0.0114	1.3258
COI	GCCCACTTCCACTATGTTCTC	ATAGGAGCCGTCTTTGCCATTGTTGC	CACTACACAGCACCTGAACA	108	0.5462	0.0126	0.9537
LWS1	GAGCTGAGAAACCTTCTTT	ATGGCAGAGGAATGGGGAAAACA	GCTGTTTGTGTATGTGAATGC	116	0.7056	0.0173	1.2322
LWS2	CAATGTGTGTCTTTGAAGGCTA	GTTGTCTCAATTTGTGGAATTGCTGGGC	CAAACTTGACATTTCCAAAGGG	123	0.8685	0.0144	1.5165
LWS3	CAGAGGAAGGTGTGACAG	ATGGCAGAGGAATGGGGAAAACA	GCTGTTTGTGTATGTGAATGC	131	0.9694	0.0146	1.6927
LWSR	GGGCTCTGCTTTCACATAC	CATACAAGAGATCCTTTTGAGGGACCAAACTAC	CCAAGACCAGACCATTTGTG	155	0.6412	0.0117	1.1196
Myosin HC	TCTCTGCTCACCTTCAACTTC	CCAGTTCTCGGACCCTCTGCT	CAGATCCTTGACTGTAACCTCG	136	0.5303	0.0136	0.9260
RH1	CTGGTTTGCTGGATACCTTAC	CCAGTGTGGCCTGGTACATCTTCAC	GCTTGTTCATGCAGATGTAGATC	151	0.6168	0.0131	1.0770
RH2-1	ACCATCACATCTGCTCTTAA	GCTATTGAGGGATTCATGGCAACACT	AACAACCAGAGACCAGAGAG	111	0.6637	0.0139	1.1589
RH2-2	TAGGCTGTGATATGGAAGGT	GGAGGTCAAGTATCACTATGGTCTCTTGT	CCACTATGTATCTCTCAATAGCTAAG	93	0.8852	0.0258	1.5457
SWS1	GCGTTCTTCTCCAAGAGC	ACAAACAGTTCAATGCCTGCATCATGG	GTGGACACCTCAGTCTTTGA	146	0.6268	0.0164	1.0945
SWS2 A	CCATCCTGCCTGTCGAAG	AGTCCTCAATAAGCAGTTTCGAACATGCATG	GTCACTGACTGAGTTGTAGAGAC	140	0.9737	0.0263	1.7003
SWS2 B	TCCTGTGTCTCAAAAGCCTC	ACGTTCTCCTCAATAAGCAGTTCCGC	CACTCATCCCCAGCATCTTC	100	0.7110	0.0167	1.2414

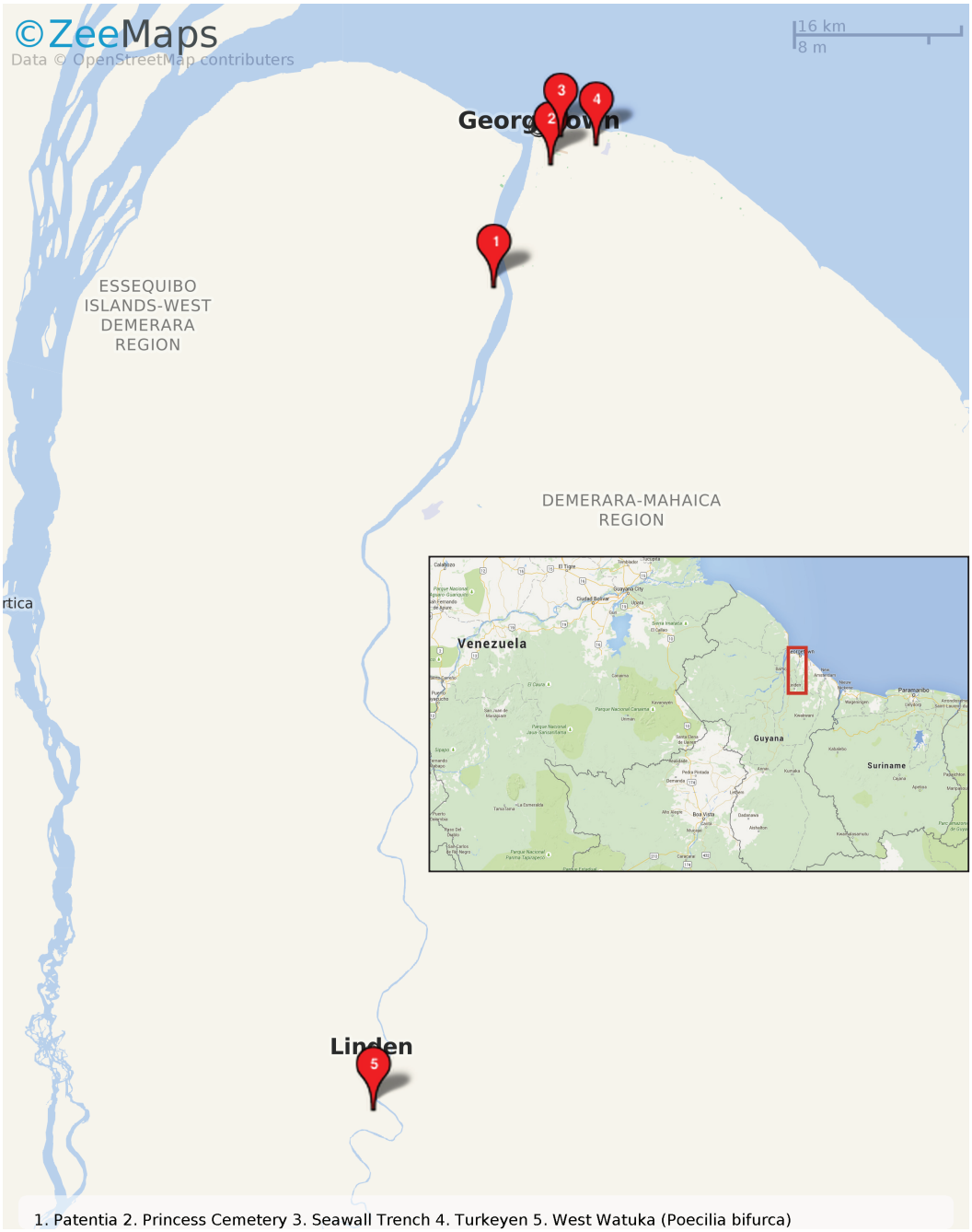
Supplementary Table 5.3. PCR and sequencing primers used for LWS-1 180 Ala/Ser allele identification.

	Forward	Reverse
PCR	TGTGAAGTGCAGATCACCTAG	ACACATTCATGCATGATGCAG
Sequencing	GATCCCTTTGAAGGACCAAACT	GGACAATCATGTAGGACAGGACC



Supplementary Figure 5.1. Male morphs of the three sympatric species.

Field Sites



Supplementary Figure 5.2. Map of field sites in Guyana.

Chapter 6.

Concluding remarks and syntheses: Can beauty be found in the eyes of the beholder?

Sexual selection has been a focus of evolutionary biology since being proposed by Darwin (Darwin 1871). While at first controversial, in time we have come to truly appreciate the far-reaching effects of sexual selection on the evolution of species, and many models have been put forth to explain the evolution of sexual selection (reviewed in Kuijper *et al.* 2012). One form of sexual selection is intersexual mate choice, which is expected to result in a coevolution of characters preferred and preferences for those characters (Pomiankowski & Sheridan 1994; Breden *et al.* 1994; Andersson & Simmons 2006). Making mating decisions based on the characters of a conspecific requires a sensory system capable of detecting and discriminating such traits (Endler *et al.* 2005). The Sensory Exploitation model predicts that variation in the sensory systems involved in evaluating potential mates will lead to differences in mate preferences (Basolo & Endler 1995; Endler & Basolo 1998). The family Poeciliidae provides an astounding model system for studies of the evolution of mate preference as intersexual selection is rampant. Many of the species in this family make mating decisions based on male colouration, which requires colour vision, but there is variation in the colours preferred. If Sensory Exploitation plays a role in mate preference evolution in this family then we expect to also see variation in colour vision.

In this dissertation I have presented a series of studies examining the co-variation in colour vision and mate choice in Poeciliid fishes. These studies used a combination of laboratory and field techniques to assess variation in opsin gene sequence and expression across species and populations. In brief, I report that, (a) the LWS-1 and LWS-3 opsins experience a high incidence of gene conversion in the family Poeciliidae but diversity is maintained in the subgenus Lebistes (Chapter 2), (b) the evolution of sensory repertoires does not need to rely on the slow processes of

duplication and divergence (Chapter 3), (c) differences in colour vision correlate with differences in mate preference across natural guppy populations (Chapter 4), and (d) colour vision differs more across populations than across species in the same location (Chapter 5). More broadly, these findings demonstrate; (1) colour vision can vary considerably within species and (2) the value of sensory exploitation as a theoretical framework for understanding mate preferences. Below I discuss how my work has supported these conclusions and suggest further lines of work that could build off the results of this dissertation.

6.1. Colour Vision Can Vary Considerably Within Species

Sensory systems are classically assumed to be constant within species, especially when modeling visual reception (*eg*. Endler 1991; Kemp *et al.* 2009; Hurtado-Gonzales *et al.* 2014). The work presented here clearly shows there can be high variability in sensory systems across populations within species, arising from differences in both allele frequencies and gene expression (Chapters 4 and 5). Such differences could be the result of adaptations to local light environments or differences in the selection pressures from the various roles of such sensory systems. Differences in selection pressures experienced by the visual system are likely as the visual system plays a key role in an organism's ability to both detect and evaluate: potential mates, conspecific and heterospecific competitors, predators, food location and quality, as well as obstacles in their environment. The relative importance of each of these functions can vary dramatically across time and space depending on the strength of selection pressures. Since colour vision is constrained by the number of cone cells in the back of the retina it is a trade off to find a visual system that is best suited to the relative selection pressures of the current environment. Thereby even minor differences in selection pressures across populations could lead to differences in visual systems.

Chapter 4 found natural populations of guppies (*Poecilia reticulata*) in Trinidad differ in colour vision through differences in the frequency of the LWS-1 (180 Ala) allele (Sandkam *et al.* 2015). The LWS-1 (180 Ala) is the only known allele variant of the LWS opsins that is thought to alter the tuning of cone cells in guppies (Tezuka *et al.* 2014) and is thought to decrease λ_{max} by 7nm relative to the LWS-1 (180 Ser) allele. The frequency

of the LWS-1 (180 Ala) allele differed between high and low predation populations in a very similar manner across two independently colonized watersheds (Aripo watershed: high predation- 0.33, low predation- 0.02; Marianne watershed: high predation- 0.36, low predation- 0.00). Such replicated allele frequency differences suggest populations with similar predation regimes experience similar selective pressures, which is in line with a large body of previous work on guppies and strengthens the conclusion that the colour vision of populations adapt to local conditions (Reznick *et al.* 2001; Tezuka *et al.* 2014).

Differences in the frequency of the LWS-1 (180 Ala) allele across populations were not restricted to guppies (as previously thought); Chapter 5 found two closely related species (*Poecilia picta* and *Poecilia parae*) also possess the LWS-1 (180 Ala) allele at a varying frequency across populations. This highlights the possibility that possessing multiple sensory alleles at a given locus may be much more widespread than is frequently assumed. Classically, studies that examine variation in the sequence of sensory genes examine only one or a few individuals per species, which is then assumed to be the same across all individuals of that species (reviewed in Rennison *et al.* 2012). However, more and more recent studies are finding allelic variation in sensory systems within species (*eg.* Smith & Carleton 2010; Tezuka *et al.* 2014; Sandkam *et al.* 2015). Such variation has important implications to sensory ecology and may help explain individual variation in behaviour, as these individuals would be expected to differ slightly in the information detected or discriminated in their environment. Future work is needed to assess how differences in sensory alleles influence how an organism interacts with their environment.

The process of Hybrid Sensory Expansion (HSE), proposed in Chapter 3, also supports variation in sensory alleles within species. HSE posits that when two species with differentially tuned sensory repertories form a hybrid, that hybrid will have an expanded sensory repertoire by being a heterozygote for the loci that differ in the parental species. Chapter 3 uses a unique system to test this part of the process, the species *Poecilia formosa*, which reproduces gynogenetically and is thereby a ‘frozen F1’ (Vrijenhoek 1979; Rogers & Vamosi 2010). However, when most hybrid species are formed they reproduce sexually, and according to traditional mendelian genetics, when two heterozygotes reproduce they are expected to form offspring of three genotypes at

each locus with different alleles. Clearly HSE results in differentially tuned sensory systems within species through allelic variation.

While changes to the wavelength at which an opsin absorbs light largely occurs at the sequence level, differences in the expression of the different opsin genes can also dramatically alter colour vision (reviewed in Horth 2007). Variation in colour vision due to differences in expression has been observed across species (cichlids- Carleton & Kocher 2001) and even across populations within species (Bluefin killifish- Fuller *et al.* 2004; Cichlids- Smith *et al.* 2010; Stickleback- Flamarique *et al.* 2013). Chapter 4 shows opsin gene expression varies across natural populations of guppies. As with the allele frequency differences the differences in opsin expression between high and low predation populations are similar in two independently colonized watersheds. Similarly, in Chapter 5 I showed opsin gene expression differs across populations in Guyana, not only in guppies, but also in the closely related *P. picta* and *P. parae*. Despite variation in opsin expression across populations of these three species, there was little to no difference in opsin gene expression across species from the same location.

The high variation in colour vision across populations due to allelic and expression level differences demonstrated in this body of work brings up an important but frequently overlooked aspect of visual ecology- that sensory systems vary within species. Models of visual ecology generally use one set of sensory receptors and assume equal proportions of all receptors then ask how changes in the light environment drive differences in life history (*eg.* Endler 1991; Kemp *et al.* 2009; Hurtado-Gonzales *et al.* 2014). As more and more technology and resources are available for field measures of sensory systems, it is becoming more feasible to measure sensory systems of many individuals in both sequence and expression to get a much higher resolution understanding of how sensory systems can drive differences in life history traits.

6.2. The Value of Sensory Exploitation as a Theoretical Framework for Understanding Mate Preferences

The Sensory Exploitation (SE) model for the evolution of mate preferences predicts that variation in peripheral sensory systems can lead to variation in the traits

preferred (Basolo & Endler 1995; Endler & Basolo 1998). Support for the SE model has largely come from comparison of preferences and sensory systems across species (reviewed in Ryan & Cummings 2013). Chapter 2 adds qualitative support to this pattern, such that a group of fishes with mate preferences based on red/orange colouration (the subgenus Lebistes) possess a unique allele of LWS-1 (responsible for detecting and discriminating those colours). Chapters 4 and 5 show sensory systems can vary dramatically across populations within species, raising the possibility that SE may be able to explain variation in mate preferences, not only across, but also within species. Indeed, Chapter 4 found that sensory systems co-vary with mate preferences across populations in guppies, supporting the idea that SE may play a role in population differences in mate preference. Chapter 5 further demonstrated colour vision varies across populations within species. Unfortunately we did not have measures of mate preference for these populations, it would be interesting to go back and measure the mate preferences for these populations to determine if the differences in colour vision could predict differences in mate preference.

The work of this dissertation offers qualitative support for SE and paves the way for future work to develop a more nuanced sense of the role that SE plays in leading to mate preferences. Particularly, directly determining colour preferences of individuals and populations, then measuring their opsin expression and allele frequencies could provide a stronger sense of how variation in sensory systems can lead to differences across populations in mate preferences. Additionally, manipulating colour vision and looking for differences in colour preferences would provide a key examination of whether changing sensory systems can in fact drive changes to mate preference, as is predicted from SE. New tools for genomic editing, such as the CRIPR/cas9 system, may provide a way to directly test how changes in opsin tuning influence mate preferences. Additionally, modification of the putative Locus Control Region (LCR) upstream of the LWS opsin array in guppies (Tam *et al.* 2011) may help us understand how changes in gene expression can influence colour preferences.

6.3. Conclusions

In this body of work I have examined variation in colour vision throughout the family Poeciliidae. I have found opsin amino acid sequence varies relatively little across species, which may be due to the genomic environment resulting in gene conversion homogenizing loci. However, I found there to be relatively high diversity within species in colour vision across populations in nature. The Sensory Exploitation hypothesis for the evolution of female preferences posits that variation in peripheral sensory systems can lead to differences in mate preference. As such, this work highlights a potential driver to population divergence and opens an interesting line of inquiry for future work to explore.

References

Andersson MB, Simmons LW (2006) Sexual selection and mate choice. *Trends In Ecology & Evolution*, 21, 296–302.

Basolo AL, Endler JA (1995) Sensory biases and the evolution of sensory systems. *Trends In Ecology & Evolution*, 10, 489–489.

Breden F, Gerhardt HC, Butlin RK (1994) Female choice and genetic correlations. *Trends In Ecology & Evolution*, 9, 343.

Carleton KL, Kocher TD (2001) Cone opsin genes of African cichlid fishes: tuning spectral sensitivity by differential gene expression. *Molecular Biology And Evolution*, 18, 1540–1550.

Darwin C (1871) *The Descent of Man and Selection in Relation to Sex*. John Murray, London, UK.

Endler JA (1991) Variation in the appearance of guppy color patterns to guppies and their predators under different visual conditions. *Vision Research*, 31, 587–608.

Endler JA, Basolo AL (1998) Sensory ecology, receiver biases and sexual selection. *Trends In Ecology & Evolution*, 13, 415–420.

Endler JA, Westcott DA, Madden JR, Robson T (2005) Animal visual systems and the evolution of color patterns: Sensory processing illuminates signal evolution. *Evolution*, 59, 1795–1818.

Flamarique IN, Cheng CL, Bergstrom C, Reimchen TE (2013) Pronounced heritable variation and limited phenotypic plasticity in visual pigments and opsin expression of threespine stickleback photoreceptors. *Journal of Experimental Biology*, 216, 656–667.

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: Sensory, Neural, and Behavioral Physiology*, 190, 147–154.

Horth L (2007) Sensory genes and mate choice: Evidence that duplications, mutations, and adaptive evolution alter variation in mating cue genes and their receptors. *Genomics*, 90, 159–175.

Hurtado-Gonzales JL, Loew ER, Uy JAC (2014) Variation in the visual habitat may mediate the maintenance of color polymorphism in a poeciliid fish. PLoS One, 9, e101497.

Kemp DJ, Reznick DN, Grether GF, Endler JA (2009) Predicting the direction of ornament evolution in Trinidadian guppies (*Poecilia reticulata*). *Proceedings of the Royal Society of London- B*, 276, 4335–4343.

Kuijper B, Pen I, Weissing FJ (2012) A guide to sexual selection theory. *Annual Review of Ecology Evolution and Systematics*, 43, 287–311.

Pomiankowski A, Sheridan L (1994) Linked sexiness and choosiness. *Trends In Ecology & Evolution*, 9, 242–244.

Rennison DJ, Owens GL, Taylor JS (2012) Opsin gene duplication and divergence in ray-finned fish. *Molecular Phylogenetics And Evolution*, 62, 986–1008.

Reznick DN, Butler M, Rodd FH (2001) Life-history evolution in guppies. VII. The comparative ecology of high- and low-predation environments. *American Naturalist*, 157, 126–140.

Rogers SM, Vamosi SM (2010) Frozen F1's amidst a masterpiece of nature: new insights into the rare hybrid origin of gynogenesis in the Amazon molly (*Poecilia formosa*). *Molecular Ecology*, 19, 5086–5089.

Ryan MJ, Cummings ME (2013) Perceptual biases and mate choice. *Annual Review of Ecology Evolution and Systematics*, 44, 437–459.

Sandkam BA, Young CM, Breden F (2015) Beauty in the eyes of the beholders: colour vision is tuned to mate preference in the Trinidadian guppy (*Poecilia reticulata*). *Molecular Ecology*, 24, 596–609.

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 *et al.* (2010) Intraspecific cone opsin expression variation in the cichlids of Lake Malawi. *Molecular Ecology*, 20, 299–310.

Tam KJ, Watson CT, Massah S *et al.* (2011) Regulatory function of conserved sequences upstream of the long-wave sensitive opsin genes in teleost fishes. *Vision Research*, 51, 2295–2303.

Tezuka A, Kasagi S, Van Oosterhout C *et al.* (2014) Divergent selection for opsin gene variation in guppy (*Poecilia reticulata*) populations of Trinidad and Tobago. *Heredity*, 113, 381–389.

Vrijenhoek RC (1979) Factors affecting clonal diversity and coexistence. *American Zoologist*, 19, 787–797.