Genomic Environment Impacts Color Vision Evolution in a Family with Visually Based Sexual Selection

Benjamin A. Sandkam[1,2,*], Jeffrey B. Joy[1,3], Corey T. Watson[1,4], and Felix Breden[1]

[1]Department of Biological Sciences, Simon Fraser University, Burnaby, British Columbia, Canada

[2]Department of Biology, University of Maryland, Maryland, USA

[3]Department of Medicine, University of British Columbia, Vancouver, British Columbia, Canada

[4]Department of Biochemistry and Molecular Genetics, University of Louisville School of Medicine, Kentucky, USA

*Corresponding author: E-mail: bsandkam@umd.edu.

Accepted: November 6, 2017

Data deposition: This project has been deposited at GenBank under the accession KX768552-KX768664.

Abstract

Many models of evolution by sexual selection predict a coevolution of sensory systems and mate preferences, but the genomic architecture (number and arrangement of contributing loci) underlying these characters could constrain this coevolution. Here, we examine how the genomic organization and evolution of the opsin genes (responsible for tuning color vision) can influence the evolutionary trajectory of sexually selected traits across 15 species in the family Poeciliidae, which includes classic systems for studies of color-mediated sexual selection such as guppies, swordtails, and mollies. Although male coloration patterns and the importance of this coloration in female mate choice vary widely within and among genera, sequencing revealed low variability at amino acid sites that tune Long Wavelength-Sensitive (LWS) opsins in this speciose family. Although most opsin genes in these species appear to have evolved along traditional mutation-selection dynamics, we identified high rates of gene conversion between two of the LWS loci (*LWS-1* and *LWS-3*), likely due to the inverted tandem repeat nature of these genes. Yet members of the subgenus Lebistes appear to resist LWS gene conversion. The LWS opsins are responsible for detecting and discriminating red and orange coloration—a key sexually selected trait in members of the subgenus Lebistes. Taken together these results suggest selection is acting against the homogenizing effects of gene conversion to maintain LWS-1/LWS-3 differences within this subgenus.

Key words: opsin, gene conversion, Poeciliidae, vision, LWS.

Introduction

The genomic repertoire (size and composition of gene families) underlying traits can strongly influence the evolution of behaviors including decision-making processes (Robinson et al. 2008; Renn et al. 2008). However, gene evolution is subject to its own processes and constraints that can facilitate or inhibit genomic repertoire diversity (Lynch 2007). Therefore, understanding the evolution of behaviors requires knowledge of the impacts of the genomic environment on the evolution of genes mediating such traits (Wilkinson et al. 2015).

Sexual selection can lead to behavioral variation, large-scale phenotypic evolution, and even speciation (Andersson 1994; Panhuis et al. 2001; Andersson and Simmons 2006). Variation in the genes tuning sensory systems has been shown to strongly influence the evolution of mate preference behaviors (reviewed in Horth 2007). Genomic predispositions toward divergence or stasis of relevant sensory genes could play a dramatic role in shaping the direction of sexually selected traits.

The Poeciliid family of freshwater fishes includes such classic models for understanding sexual selection as guppies (*Poecilia reticulata*), swordtails (*Xiphophorous helleri*), and sailfin mollies (*Poecilia latipinna*). Although nearly all sexually selected traits in this group are transmitted visually, the role of coloration differs strikingly across Poeciliid species (Pollux et al. 2014). Color vision is accomplished by comparing signals from cone cells with differences in wavelength sensitivity, which is largely determined by the tuning of their opsin proteins (Gegenfurtner and Sharpe 1999). Opsin tuning is primarily

determined by the protein's amino acid sequence, and thus opsin sequences can be used to estimate the functional repertoire of a species (Yokoyama and Yokoyama 1996).

Guppies and swordtails have an expanded repertoire 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 yellow, orange, and red end of the visible light spectrum (Yokoyama et al. 2008). It has been proposed that this expanded LWS repertoire has driven the evolution of strong female mate preferences for red and orange male coloration observed in the subgenus Lebistes (sampled here as *Poecilia parae, P. reticulata, P. wingei, P. picta*, and *P. bifurca*; *sensu* Rosen and Bailey 1963) (Archer and Lythgoe 1990; Hoffmann et al. 2007; Ward et al. 2008; Watson et al. 2011). Yet, despite the fact that the LWS duplication history is shared across the family (Rennison et al. 2012), not all Poeciliid species have mate preferences for long wavelength colors (Pollux et al. 2014). In addition, the evolution of nucleotide variation at key LWS coding sites does not appear to be uniform among these species (Ward et al. 2008; Watson et al. 2010, 2011).

Gene conversion is a process that typically homogenizes sequence between loci and can act to reduce multilocus gene repertoire diversity (Chen et al. 2007). Previously we found stronger effects of gene conversion homogenizing LWS loci in *Xiphophorus helleri* (Watson et al. 2010), which have no mate-preferences for long wavelength colors, compared with *Poecilia wingei* (Watson et al. 2011), which have strong mate-preferences for long wavelength colors. Here, we expand our previous analyses to ask whether patterns of gene conversion differ across the family. We did this by sequencing eight of the nine cone opsins and the rhodopsin gene in 15 Poeciliid species. We then inferred patterns of gene conversion on the LWS opsins. The differences observed led us to hypothesize that these differences could be related to variability in mate preference.

Materials and Methods

Sequencing

Using DNeasy blood and tissue kits (QIAGEN) DNA was extracted from tissue samples of single specimens of: *Heterandria formosa*, *Xiphophorus helleri*, *Poecilia caymanensis*, *P. vittata*, *P. nigrofasciata*, *P. latipinna*, *P. velifera*, *P. petenensis*, *P. mexicana*, *P. minor*, *P. reticulata*, *P. bifurca*, *P. picta*, *P. parae*, and *P. wingei.* (Note that *Heterandria formosa* is not to be confused with *Poecilia formosa*—two distinct members of the family Poeciliidae from different genera). Also, *P. wingei* was formerly considered a strain of *P. reticulata* but has been shown to be a distinct species (Pollux et al. 2014). Primers specific to 5′ and 3′ UTR regions were designed using genomic data: LWS and SWS2 loci—*Poecilia wingei* (Watson et al. 2011) and *Xiphophorous helleri* (Watson et al. 2010); *SWS1* and *RH1* loci—*Xiphophorus maculatus* (GenBank accessions: AGAJ01036758.1 and AGAJ01019341.1, respectively). UTR primers of *RH2-1* were taken from Sandkam et al. (2013). For primer sequences see supplementary table S1, Supplementary Material online. We follow Sandkam et al. (2013) and refer to LWS loci by their location relative to one another, with *LWS-R* being a retrotransposed gene in a separate linkage group compared with *LWS-1*, *LWS-2*, and *LWS-3*. PCR products of *LWS-2* were generated in two overlapping segments, each with one UTR and one internal primer. The LWS opsin genes share a duplication history that predates the family Poeciliidae (Watson et al. 2010, 2011; Rennison et al. 2012). By generating sequencing products with UTR primers, we ensure that phylogenetic clustering of opsins within species are gene conversion events rather than independent duplications (Watson et al. 2010). 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) when sites were ambiguous we compared overlapping reads and called a base by the highest peak.

A phylogeny of the species used in this study was inferred from *ND2* mitochondrial sequences retrieved from GenBank (supplementary table S2, Supplementary Material online). 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 amplified for the *RH2-2* locus in the majority of the species so this opsin gene was left out of all analyses. All sequences generated are available under GenBank Accessions KX768552 - KX768664.

Phylogenetic Inference

The divergence of the opsin classes occurred prior to the emergence of the family Poeciliidae (Rennison et al. 2012), making alignment of introns and UTRs across classes difficult. Therefore, we inferred a series of trees based on several sets of sequences: 1) mitochondrial ND2, 2) all opsins: exon sequence only, 3) *SWS1*: full sequence (UTR, introns and exons), 4) SWS2: full sequences, 5) *RH1*: full sequence, 6) *RH2-1*: full sequence, 7) LWS: exons and introns, and 8) LWS: UTR only sequence (see supplementary table S4 in the online Supplementary Material for all UTR lengths).

For each set, sequences were aligned using a command line implementation of Mafft v7.221 (Katoh and Standley 2013) and edited manually using AliView v1.17.1 (Larsson 2014) to ensure that intron–exon boundaries were consistent.

Best-fit models of molecular evolution were determined using MrModelTest 3.04 (Nylander 2004). Maximum likelihood phylogenetic trees were reconstructed using RAxML v8.2.9 (Stamatakis 2006). We performed Bayesian analysis of each aligned data partition using BEAST v 1.8.2 (Drummond et al. 2012) under a log normal relaxed molecular clock model (Drummond et al. 2006). We ran four independent chains for 50 million generations each; we assessed convergence of the runs graphically using Tracer v1.5 (Rambaut and Drummond 2007) and through evaluation of the effective sample size estimates (ESS) for each parameter (Drummond et al. 2006). ESS values above 200 were taken as evidence of parameter stability. We generated maximum clade credibility (MCC) phylogenetic trees from the resulting distributions of trees using TreeAnnotator v1.8.2 (Drummond et al. 2012). ML bootstrap values and Bayesian posterior probabilities were employed to assess support for inferred topologies.

Gene Conversion Analyses

We tested whether the *LWS-1/LWS-3* clustering within species and clades was due to gene conversion [as observed in some members of Poeciliidae (Watson et al. 2010, 2011)] using the program GARD on an alignment of all *LWS-1* and *LWS-3* sequences (Kosakovsky Pond et al. 2006). To identify within species gene conversion tracks between *LWS-1* and *LWS-3* we ran GENECONV (Sawyer 1989) on the same alignment of all *LWS-1* and *LWS-3* sequences and identified significant within-species pairwise tracks. GENECONV significance was set to $P < 0.05$ with gscale = 0. The length of within species conversion tracks, percent of *LWS-1/LWS-3* converted within species, intron length and within species percent amino acid similarity was compared using *t* tests in R v3.3.2 (R Core Team 2016).

Results and Discussion

In the absence of gene conversion and homogenization, we would expect to recover species relationships from phylogenies constructed from each opsin gene sequence. However, we have previously reported that phylogenies constructed using LWS coding sequence from *X. helleri* and *P. wingei* contained ambiguities in the evolutionary relationships between gene loci and species, which we attributed to the sequence homogenizing effects of gene conversion (Watson et al. 2010, 2011). To examine this expectation across poeciliids, we compared independent phylogenies constructed from each of the eight opsin genes sequenced here to a control phylogeny inferred from the mitochondrial *NADH dehydrogenase 2* (*ND2*) gene from each species. As expected, species relationships seen in each of the non-LWS opsin phylogenies mirrored those observed in the *ND2* tree (fig. 1, supplementary figs. S1–S5, Supplementary Material online) [these also matched previous Poeciliid phylogenetic studies

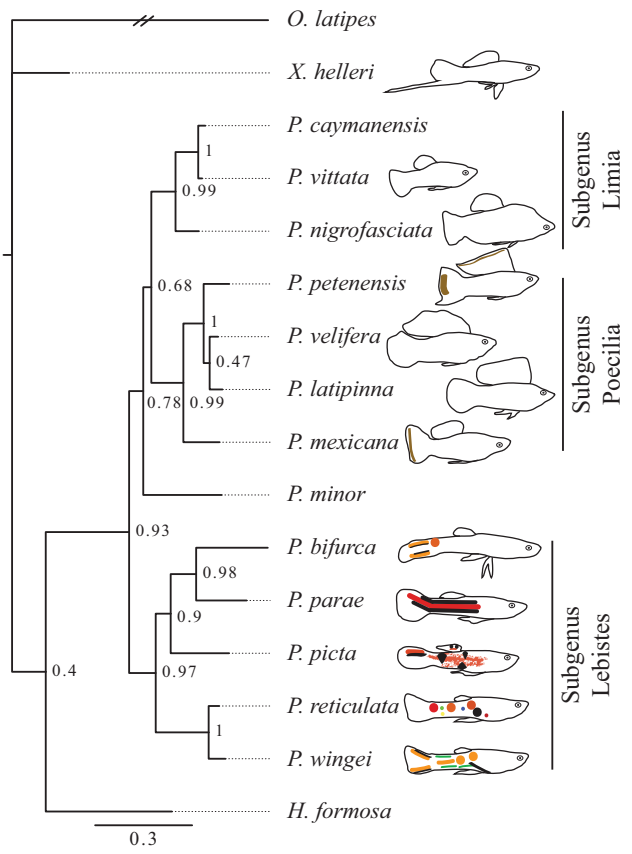

Fig. 1.—*ND2* mitochondrial Bayesian consensus tree with posterior probabilities. To show differences in body shape and coloration across species, sketches were made by outlining photographs. Traits involved in mate choice were included on sketches. Note the importance of red/orange coloration in subgenus Lebistes. Outgroup is *Oryzias latipes*.

(Breden et al. 1999; Pollux et al. 2014)]. Such expected species relationships were also recovered for *LWS-2* and *LWS-R*, demonstrating that these loci are evolving through traditional mutation-selection dynamics. In contrast, as we found previously in *X. helleri* and *P. wingei*, the introns and exons of *LWS-1* and *LWS-3* loci frequently clustered within species and clades (fig. 2*A*) indicating that these loci have likely undergone gene conversion (Watson et al. 2010, 2011; Rennison et al. 2012). This is in contrast to the LWS phylogeny built using untranslated region (UTR) (fig. 2*B*), which again recovered the expected species relationships seen in the *ND2* tree. The within species/clade clustering observed in LWS intron/exon trees likely reflects signatures of gene conversion rather than independent duplication (Watson et al. 2010, 2011).

We next identified the regions of LWS sequence that have undergone conversion. The program GARD (Genetic Algorithm Recombination Detection) (Kosakovsky Pond et al. 2006) revealed prevalent gene conversion between *LWS-1* and *LWS-3*, but only four shared significant breakpoints ($P < 0.05$) (fig. 3). This reaffirms the prevalence of gene conversion and suggests the tracks of sequence

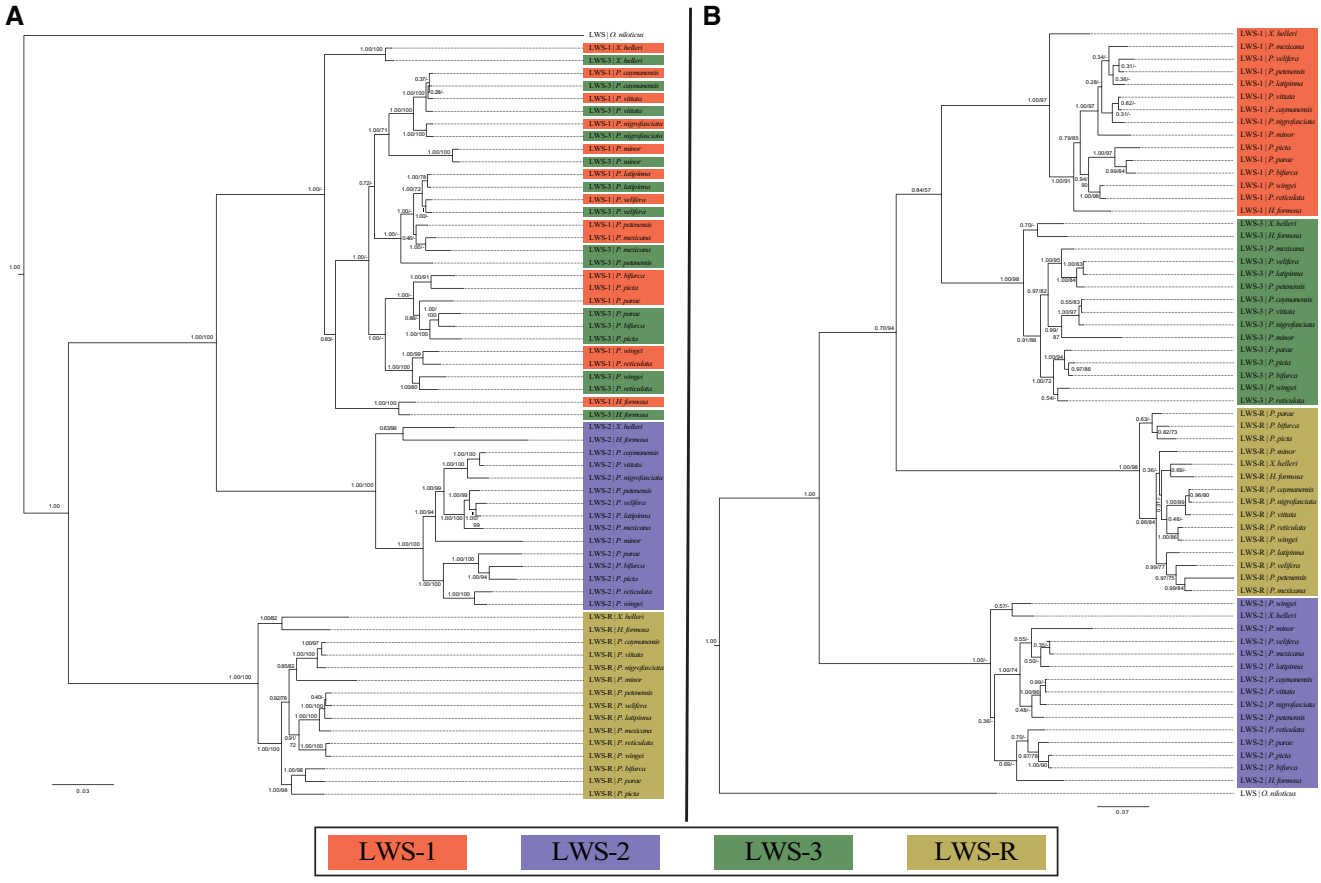

Fig. 2.—Bayesian consensus trees with posterior probabilities followed by ML bootstrap values >0.6 of (*A*) LWS intron/exon sequences, and (*B*) LWS UTR sequence. Outgroup is *Oreochromis niloticus*. Note the expected duplication history recovered in UTR sequence and general species relationships (based on ND2 tree) recovered within clades except for *LWS-1* and *LWS-3* of the intron/exon LWS tree.

experiencing gene conversion differ across the family. We also used the program GENECONV (Sawyer 1989) to identify tracks of gene conversion between *LWS-1* and *LWS-3* that have occurred at the species level (fig. 3). The five members of the subgenus Lebistes tested (*Poecilia parae, P. picta, P. bifurca, P. reticulata*, and *P. wingei*) have shorter tracks of within species *LWS-1/LWS-3* gene conversion (mean = 203.8 bp, SD = 79.4) compared with the ten species of non-Poeciliids (mean = 513.8 bp, SD = 220.6) ($t(10.97) = -3.80$, $P < 0.01$). Lebistes species also had a lower percentage of their *LWS-1/LWS-3* sequence converted (mean- = 13.14%, SD = 7.34) compared with other Poeciliids (mean- = 41.98%, SD = 20.69) ($t(10.92) = -3.77$, $P < 0.01$).

Such differences in conversion tracks likely have functional implications as a pair-wise BLAST analysis revealed that within species *LWS-1/LWS-3* percent amino acid sequence similarity was lower in Lebistes (mean = 97.92%, SD = 0.43) than non-Lebistes species (mean = 98.75%, sd = 1.01) ($t(16.46) = -2.56$, $P = 0.02$) (table 1).

These results raise two important questions: Why is gene conversion so prevalent between *LWS-1* and *LWS-3*? And what are the implications of gene conversion to color vision? Below we propose hypotheses for both questions and discuss their support.

Why Is Gene Conversion So Prevalent between *LWS-1* and *LWS-3*?

Gene conversion frequently occurs during double strand break (DSB) repair; the broken ends find a template strand matching intact sequence adjacent to the break and 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 on LG5 of the *P. reticulata* genome (GenBank ID 23338) in a tandem array with <6 kb between *LWS-1* and *LWS-2*, and <4 kb between *LWS-2* and *LWS-3* (Watson et al. 2010, 2011). The LWS duplication that resulted in this conformation predates the emergence of family Poeciliidae and has resulted in a shared genomic architecture of these genes (Watson et al. 2010, 2011; Rennison et al. 2012). The close proximity of *LWS-1*, *LWS-2*, and *LWS-3* greatly increases the likelihood of gene conversion occurring between these loci.

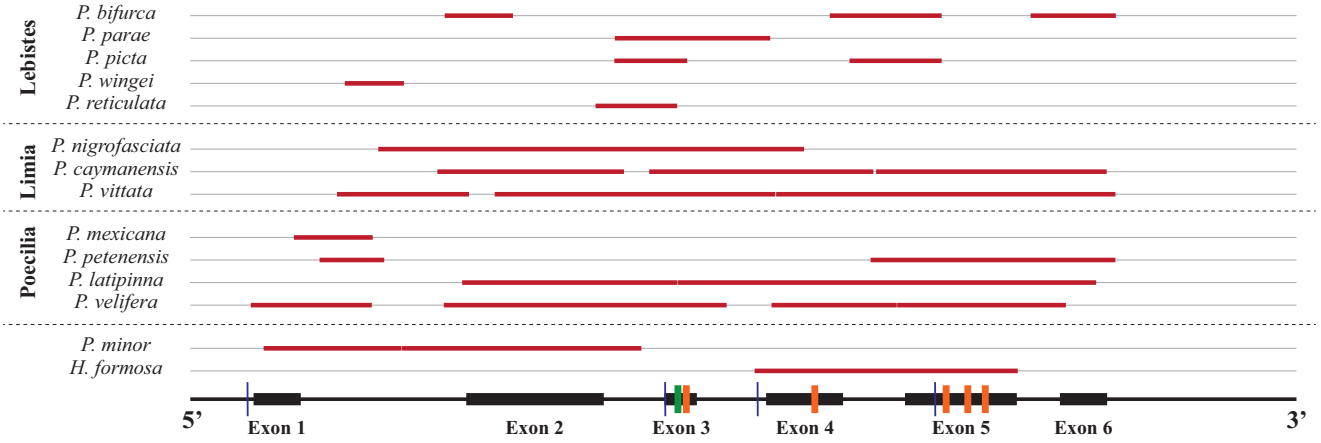

Fig. 3.—Overview of gene conversion between *LWS-1* and *LWS-3*. Within-species conversion tracks are shown in red. The intron/exon structure is noted below. The thin blue lines denote the shared breakpoints identified by GARD. Thick orange vertical lines denote the five "key sites" (Yokoyama and Radlwimmer 2001) and the thick green vertical line denotes the additional key site proposed for Poeciliid LWS by Kawamura et al. (2016). Note: GeneConv did not detect gene conversion between *LWS-1* and *LWS-3* within *X. helleri* because conversion has resulted in these loci being 91% identical throughout (1945/2138 bp) (Watson et al. 2010), making them too similar for the program to distinguish tracks without conversion (Mansai and Innan 2010).

Table 1

The "Five Key-Site" (Yokoyama and Raddlewimmer 2001) Haplotype (Using the Human Numbering System) of the Tandem LWS Loci Plus Site 194 Identified by Kawamura et al. (2016) as a Potentially Important Tuning Site

	LWS-1						*LWS-2*						*LWS-3*						# LWS	*LWS-1/3* Differences (%)
	180	194	197	277	285	308	180	194	197	277	285	308	180	194	197	277	285	308		
X. helleri	S	Y	H	Y	T	A	P	Y	H	F	A	A	S	Y	H	Y	T	A	2	0.28
H. formosa	S	Y	H	Y	T	A	P	Y	H	F	A	A	S	Y	H	Y	T	A	2	0.28
P. minor	S	Y	H	Y	T	A	S	Y	H	F	A	A	S	Y	H	Y	T	A	2	0.02
P. caymanensis	S	Y	H	Y	T	A	P	Y	H	F	A	A	S	Y	H	Y	T	A	2	0.00
P. nigrofasciata	S	Y	H	Y	T	A	P	Y	H	F	A	A	S	Y	H	Y	T	A	2	0.01
P. vittata	S	Y	H	Y	T	A	P	Y	H	F	A	A	S	Y	H	Y	T	A	2	0.00
P. latipinna	S	Y	H	Y	T	A	P	Y	H	F	A	A	S	Y	H	Y	T	A	2	0.00
P. mexicana	S	Y	H	Y	T	A	P	Y	H	F	A	A	S	F	H	Y	T	A	**3**	0.02
P. velifera	S	Y	H	Y	T	A	P	Y	H	F	A	A	S	Y	H	Y	T	A	2	0.00
P. petenensis	S	Y	H	Y	T	A	P	Y	H	F	A	A	S	F	H	Y	T	A	**3**	0.02
P. bifurca	A	Y	H	Y	T	A	P	Y	H	F	A	A	S	F	H	Y	T	A	**3**	**0.02**
P. picta	A	Y	H	Y	T	A	P	Y	H	F	A	A	S	F	H	Y	T	A	**3**	**0.03**
P. wingei	A	Y	H	Y	T	A	P	Y	H	F	A	A	S	F	H	Y	T	A	**3**	**0.02**
P. parae	S/A	Y	H	Y	T	A	P	Y	H	F	A	A	S	F	H	Y	T	A	**3**	**0.03**
P. reticulata	S/A	Y	H	Y	T	A	P	Y	H	F	A	A	S	F	H	Y	T	A	**3**	**0.02**

Note.—The number of LWS "influential-site" haplotypes within species is summarized as "# LWS." The within species amino acid differences between *LWS-1* and *LWS-3* is given as "*LWS-1*/3 Differences." The Lebistes species are highlighted in orange. Note: *P. parae* and *P. reticulata* are polymorphic for alleles with an A or S at site 180 in *LWS-1* (Tezuka et al. 2014; Sandkam et al. 2015a, 2015b).

Surprisingly, while most studies find that conversion occurs more frequently between adjacent loci (Katju and Bergthorsson 2010; Cortesi et al. 2015), we observed gene conversion only between *LWS-1* and *LWS-3*. The differences in intron–exon structure and sequence length among the LWS duplicates may be one factor affecting the differential gene conversion rates. Within species, *LWS-1* and *LWS-3* introns differ in length by an average of only 0.8 base pairs wheresas *LWS-1* and *LWS-2* differ by an average of 378.3 base pairs (supplementary table S5, Supplementary Material online). This is especially pronounced in the expanded length of intron 1 of *LWS-2* which is up to 7.2 times longer than intron 1 of *LWS-1* (*P. reticulata*). The difference in size of intron 1 between *LWS-2* and *LWS-1*/*LWS-3* may make it less likely that *LWS-1* or *LWS-3* will use *LWS-2* as a template following a DSB due to the dramatic size differences making alignment difficult. Other systems with opsin gene conversion generally have similar length introns between converting genes (Verrelli and Tishkoff 2004; Cortesi et al. 2015). The close proximity of *LWS-1* to *LWS-3* and matching length of all

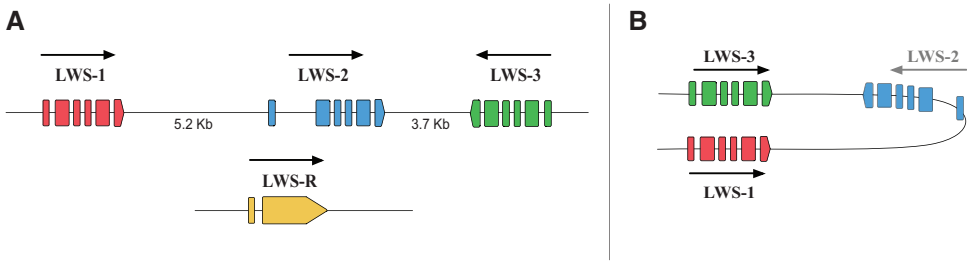

Fig 4.—(*A*) The genomic organization of the LWS opsin loci in *P. wingei* [adopted from (Watson et al. 2011)]. (*B*) The proposed conformation leading to high rates of gene conversion between *LWS-1* and *LWS-3* in the event of a double strand break. Arrows denote 5′ to 3′ direction. Colored boxes denote exons whereas spaces between boxes denote introns.

introns/exons make it likely they will experience gene conversion in the event of a DSB. This is not to say that gene conversion is strictly the result or cause of the similar sizes of *LWS-1* and *LWS-3*, as conversion will further homogenize intron lengths. However, as conversion homogenizes intron length, it will become more likely that conversion will occur in the future; leading to a self-reinforcing propensity toward gene conversion between these loci. Indeed, the length of each *LWS-1* and *LWS-3* intron is far more variable within a locus across species (average standard deviation 3.58) than across loci within species (average standard deviation 0.42) (supplementary table S5, Supplementary Material online).

We propose that the orientation of *LWS-1* and *LWS-3* makes it even more likely these two genes will experience gene conversion, because they are inverted relative to one another. When a DSB occurs and the DNA doubles back on itself *LWS-1* and *LWS-3* are in the same orientation, facilitating the use of the opposite locus as a template, resulting in a conversion event (fig. 4). 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 explains the high prevalence of this process within the family Poeciliidae.

Implications of Gene Conversion for Color Vision

Observations of gene conversion between opsin genes has now been made across many species and opsin subfamilies, including SWS opsin loci in percomorph fishes, RH2 opsin loci in Amazonian cichlids, and L/M opsins in humans and non-human primates, suggesting a key role for gene conversion in opsin evolution across taxa (Hiwatashi et al. 2011; Cortesi et al. 2015; Escobar-Camacho et al. 2017). Importantly, in many cases opsin gene conversion has been shown to have direct impacts on opsin function and phenotypic variation, where it converges opsin tuning and decreases the potential for color discrimination (Verrelli and Tishkoff 2004; Cortesi et al. 2015). However, in other gene families (such as MHC) gene conversion has been shown to be capable of increasing allelic diversity when it occurs in small sections; effectively "shuffling the deck" (reviewed in Ohta 2010).

Our results suggest that gene conversion tracts are shorter in Lebistes compared with the species in the subgenera Poecilia or Limia. There are two potential explanations for this; conversion occurs as shorter segments in Lebistes or there is stronger selection for recombination breaking up introgressed sequence. Either way this has allowed *LWS-1*/*LWS-3* amino acid sequences to be more different in Lebistes species, which likely facilitates color discrimination [e.g. behavioral differences were found in P. reticulata with the Ala versus Ser alleles of LWS-1 (Sakai et al 2016)]. Indeed, studies using microspectrophotometry (MSP) to determine the wavelength to which cone cells are most sensitive have found that members of Lebistes have more cone cell types with maximum sensitivity in the range of the LWS opsins. In Lebistes; *P. reticulata* and *P. wingei* both have three cone cell types in the LWS range (*P. reticulata*: 525, 540, and 560 nm; *P. wingei*: 533, 548, and 572 nm) (Archer and Lythgoe 1990; Watson et al. 2011) whereas *P. parae* has four cone types in the LWS range (526, 533, 543, and 553 nm) (Hurtado-Gonzales et al. 2014). Meanwhile *P. latipinna* and *P. mexicana*, both members of the subgenera Poecilia, each have only two cone types in the LWS range (*P. latipinna*: 551 and 576 nm; *P. mexicana:* 536 and 563 nm) (Korner et al. 2006) as does *X. helleri* (534 and 568 nm) (Watson et al. 2010).

Since gene conversion is a function of genomic environment, conversion rates are likely to be the same across all species sharing the same genomic environment [e.g., the shared genomic structure of LWS opsins throughout Poeciliidae (Watson et al. 2010, 2011)]. We propose that strong sexual selection for red/orange discrimination in the subgenus Lebistes (Liley 1965; Houde 1997; Lindholm et al. 2004) has resulted in selection against homogenizing *LWS-1*/*LWS-3* gene conversion which has resulted in more differentially tuned LWS opsins in this clade. Normally, such selection would be examined using tools that identify molecular signatures of selection, such as PAML (Yang 2007; Hofmann et al 2012). However, gene conversion violates the assumptions of such approaches. It would be interesting to conduct large population studies of Lebistes species to determine if there is a low frequency of individuals with long conversion tracks.

Previously, questions of LWS opsin tuning have focused on five key amino acid sites (180, 197, 277, 285, and 308 relative

to human opsin) that have been shown to have dramatic effects on LWS tuning across a broad range of taxa (Yokoyama and Radlwimmer 1998; Yokoyama et al. 2008). Our sequence results for these 15 species reveal relatively few differences in the five-sites across Poeciliids (table 1). However, recent findings show the five-sites likely do not provide the full story of Poeciliid LWS tuning: using in vitro expression of guppy LWS genes, Kawamura et al. (2016) showed that the tuning of the *LWS-1* Serine (180) allele and *LWS-3* differ dramatically despite having the same amino acids at the five-sites. They identified the amino acid site corresponding to human 194 (reported as 178 by bovine numbering) as a potentially influential site. We found Lebistes species differ at this site between *LWS-1* and *LWS-3* (table 1). We also found variation at other sites in the transmembrane domain (which generally contains the most influential sites for opsin tuning) (see Supplementary Material online for annotated amino acid alignments). Resolving the phenotypic impacts of such conversion patterns will require large comparisons of in vitro expression to determine the effects of tuning on the opsin genes.

Interestingly, while the *P. reticulata* and *P. parae* individuals sequenced here had a Serine at the 180 site (one of the "five key sites") (table 1), both species have been shown to also possess an allele with an Alanine at the 180 site, and the frequency of that allele varies across populations (*P. reticulata*—Tezuka et al. 2014; Sandkam et al. 2015a, 2015b; *P. parae*—Sandkam et al. 2015b). The presence of multiple *LWS-1* alleles in Lebistes raises the possibility that other species, including species used in this study, may also possess multiple alleles at opsin loci. However, within species the alleles differ by only 1 base-pair. Therefore, even if other species also possess multiple alleles and follow a similar pattern, such small differences are unlikely to impact our analyses of gene conversion.

Conclusion

We characterized the phylogenetic relationships for nine of the ten visual opsin genes in 15 species throughout the family Poeciliidae. We showed the *LWS-1* and *LWS-3* loci have undergone gene conversion in this family, acting to homogenize these loci within species, albeit to a different extent across species. Members of the subgenus Lebistes experience gene conversion in smaller segments leading to more differences between *LWS-1* and *LWS-3* compared with non-Lebistes species. The LWS opsins are responsible for detecting the reds, oranges, and yellows upon which Lebistes predominantly base mating decisions. The important role *LWS-1*/*LWS-3* plays in sexual selection for Lebistes may result in selection against homogenizing gene conversion in these species.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

We wish to thank members of the Breden lab: Laura Chow, Ian McNeil, and Luke Rawle for assistance in the lab. We also thank Drs. Karen Carleton, Kimberly Hughes, Belinda Chang, Soojin Yi, and four anonymous reviewers for providing comments on early drafts of the manuscript. This work was supported by the National Science and Engineering Council of Canada Discovery Grant #138178.

Literature Cited

Andersson MB. 1994. Sexual selection. Princeton (NJ): Princeton University Press.

Andersson MB, Simmons LW. 2006. Sexual selection and mate choice. Trends Ecol Evol. 21(6):296–302.

Archer SN, Lythgoe JN. 1990. The visual pigment basis for cone polymorphism in the guppy, *Poecilia reticulata*. Vision Res. 30(2):225–233.

Breden F, Ptacek MB, Rashed M, Taphorn D, Figueiredo CA. 1999. Molecular phylogeny of the live-bearing fish genus Poecilia (Cyprinodontiformes: Poeciliidae). Mol Phylogenet Evol. 12(2):95–104.

Chen J, Cooper D, Chuzhanova N, Férec C, Patrinos G. 2007. Gene conversion: mechanisms, evolution and human disease. Nat Rev Genet. 8(10):762–775.

Cortesi F, et al. 2015. Ancestral duplications and highly dynamic opsin gene evolution in percomorph fishes. Proc Natl Acad Sci U S A. 112(5):1493–1498.

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4(5):e88.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol 29(8):1969–1973.

Escobar-Camacho D, Ramos E, Martins C, Carleton KL. 2017. The opsin genes of amazonian cichlids. Mol Ecol 26(5):1343–1356.

Gegenfurtner KR, Sharpe LT. 1999. Color vision: from genes to perception. Cambridge, (UK): Cambridge University Press.

Hiwatashi T, et al. 2011. Gene conversion and purifying selection shape nucleotide variation in gibbon L/M opsin genes. BMC Evol Biol. 11(1):312.

Hoffmann M, et al. 2007. Opsin gene duplication and diversification in the guppy, a model for sexual selection. Proc Biol Sci. 274(1606):33–42.

Hofmann CM, et al. 2012. Opsin evolution in Damselfish: convergence, reversal, and parallel evolution across tuning sites. J Mol Evol. 75(3–4):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(2):159–175.

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

Hurtado-Gonzales JL, Loew ER, Uy JAC. 2014. Variation in visual habitat may mediate the maintenance of color polymorphism in a Poeciliid fish. PLoS One 9(7):e101497.

Katju V, Bergthorsson U. 2010. Genomic and population-level effects of gene conversion in *Caenorhabditis* paralogs. Genes 1(3):452–468.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780.

Kawamura S, et al. 2016. Spectral sensitivity of guppy visual pigments reconstituted in vitro to resolve association of opsins with cone cell types. Vis Res. 127:67–73.

Korner KE, Schlupp I, Plath M, Loew ER. 2006. Spectral sensitivity of mollies: comparing surface- and cave-dweeling Atlantic mollies, *Poecilia mexicana*. J Fish Biol. 69(1):54–65.

Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SDW. 2006. GARD: a genetic algorithm for recombination detection. Bioinformatics 22(24):3096–3098.
Larsson A. 2014. AliView: a fast and lightweight alignment viewer and editor for large data sets. Bioinformatics 30(22):3276–3278.
Liley NR. 1965. Ethological isolating mechanisms in four sympatric species of poeciliid fishes. Behaviour 13:1–197.
Lindholm AK, Brooks R, Breden F. 2004. Extreme polymorphism in a Y-linked sexually selected trait. Heredity 92(3):156–162.
Lynch M. 2007. The origins of genome architecture. Sunderland (MA): Sinauer Associates.
Mansai SP, Innan H. 2010. The power of the methods for detecting inter-locus gene conversion. Genetics 184(2):517–527.
Nylander JAA. 2004. MrModeltest v2. Program distributed by the author. Uppsala: Evolutionary Biology Center, Uppsala University.
Ohta T. 2010. Gene conversion and evolution of gene families: an overview. Genes 1(3):349–356.
Panhuis TM, Butlin R, Zuk M, Tregenza T. 2001. Sexual selection and speciation. Trends Ecol Evol. 16(7):364–371.
Pollux BJA, Meredith RW, Springer MS, Garland T, Reznick DN. 2014. The evolution of the placenta drives a shift in sexual selection in livebearing fish. Nature 513(7517):233–236.
R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from: https://www.R-project.org/, last accessed July 20, 2017.
Rambaut A, Drummond A. 2007. Tracer v1.4. doi: 10.1186/1471-2148-7-214.
Renn SCP, Aubin-Horth N, Hofmann HA. 2008. Fish and chips: functional genomics of social plasticity in an African cichlid fish. J Exp Biol. 211(18):3041–3056.
Rennison DJ, Owens GL, Taylor JS. 2012. Opsin gene duplication and divergence in ray-finned fish. Mol Phylogenet Evol. 62(3):986–1008.
Robinson GE, Fernald RD, Clayton DF. 2008. Genes and social behavior. Science 322(5903):896–900.
Rosen DE, Bailey RM. 1963. The Poeciliid fishes (Cyprinodontiformes), their structure, zoogeography, and systematics. Bull Am Mus Nat Hist. 126:1–176.
Sakai Y, Ohtsuki H, Kasagi S, Kawamura S, Kawata M. 2016. Effects of light environment during growth on the expression of cone opsin genes and behavioral spectral sensitivities in guppies (*Poecilia reticulata*). BMC Evol Biol. 16(1):1869.
Sandkam BA, Young CM, Breden F. 2015a. Beauty in the eyes of the beholders: colour vision is tuned to mate preference in the Trinidadian guppy (*Poecilia reticulata*). Mol Ecol. 24(3):596–609.
Sandkam BA, Young CM, Breden FMW, Bourne GR, Breden F. 2015b. Color vision varies more among populations than among species of live-bearing fish from South America. BMC Evol Biol. 15:225.
Sandkam BA, et al. 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(1):120–130.
Sawyer SA. 1989. Statistical tests for detecting gene conversion. Mol Biol Evol. 6(5):526–538.
Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22(21):2688–2690.
Tezuka A, et al. 2014. Divergent selection for opsin gene variation in guppy (*Poecilia reticulata*) populations of Trinidad and Tobago. Heredity 113(5):381–389.
Verrelli BC, Tishkoff SA. 2004. Signatures of selection and gene conversion associated with human color vision variation. Am J Hum Genet. 75(3):363–375.
Ward MN, 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 Evol Biol. 8(1):210.
Watson CT, Lubieniecki KP, Loew ER, Davidson WS, Breden F. 2010. Genomic organization of duplicated short wave-sensitive and long wave-sensitive opsin genes in the green swordtail, *Xiphophorus helleri*. BMC Evol Biol. 10(1):87.
Watson CT, et al. 2011. Gene duplication and divergence of long wavelength-sensitive opsin genes in the Guppy, *Poecilia reticulata*. J Mol Evol. 72(2):240–252.
Wilkinson GS, et al. 2015. The locus of sexual selection: moving sexual selection studies into the post-genomics era. J Evol Biol. 28:739–755.
Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 24(8):1586–1591.
Yokoyama S, Radlwimmer FB. 2001. The molecular genetics and evolution of red and green color vision in vertebrates. Genetics 158(4):1697–1710.
Yokoyama S, Radlwimmer FB. 1998. The "five-sites" rule and the evolution of red and green color vision in mammals. Mol Biol Evol. 15(5):560–567.
Yokoyama S, Yokoyama R. 1996. Adaptive evolution of photoreceptors and visual pigments in vertebrates. Annu Rev Ecol Syst. 27(1):543–567.
Yokoyama S, Yang H, Starmer WT. 2008. Molecular basis of spectral tuning in the red- and green-sensitive (M/LWS) pigments in vertebrates. Genetics 179(4):2037–2043.

Associate editor: Soojin Yi