Seeing Nemo: molecular evolution of ultraviolet visual opsins and spectral tuning of photoreceptors in anemonefishes (Amphiprioninae)

Laurie J. Mitchell 1, Karen L. Cheney 1, Wen-Sung Chung 2, N. Justin Marshall 2, Kyle Michie 2,3, Fabio Cortesi 2

1 School of Biological Sciences, The University of Queensland, Brisbane, QLD 4072, Australia
2 Queensland Brain Institute, The University of Queensland, Brisbane, QLD 4072, Australia
3 King’s College, Cambridge, CB2 1ST, UK

*Corresponding authors: laurie.mitchell@uqconnect.edu.au, fabio.cortesi@uqconnect.edu.au

Keywords: Visual opsins, spectral tuning, reef fish vision, anemonefishes, ultraviolet, gene duplication

Abstract

Many animals can see short-wavelength or ultraviolet (UV) light (shorter than 400 nm) undetectable to human vision. UV vision may have functional importance in many taxa including for foraging and communication in birds, reptiles, insects and teleost fishes. Shallow coral reefs transmit a broad spectrum of light and are rich in UV; driving the evolution of diverse spectral sensitivities in teleost reef fishes, including UV-sensitivity. However, the identities and sites of the specific visual genes that underly vision in reef fishes remain elusive and are useful in determining how molecular evolution has tuned vision to meet the ecological demands of life on the reef. We investigated the visual systems of eleven anemonefish (Amphiprioninae) species, specifically probing for the molecular pathways that
facilitate UV-sensitivity. Searching the genomes of anemonefishes, we identified a total of seven functional visual genes from all five vertebrate opsin gene subfamilies. We found rare instances of UV-sensitive SWS1 opsin gene duplications, that produced two functional paralogs (SWS1α and SWS1β) and a pseudogene. We also found separate RH2A opsin gene duplicates which has not been previously reported in the family Pomacentridae. Finally, we report on the visual opsin gene expression levels measured in the adult retina of the false clown anemonefish (Amphiprion ocellaris), and their photoreceptor spectral sensitivities measured using microspectrophotometry.

**Introduction**

Ultraviolet (UV) vision is widespread across the animal kingdom (Jacobs, 1992; Tovée, 1995), and is relied upon for many essential behaviours including foraging (Church et al. 1998; Siitari et al. 2002; Boulcoup & Braithwaite, 2004; Novales Flamarique, 2013), mate selection (Bennett et al. 1997; Andersson & Amundsen, 1997; Smith et al. 2002; Rick et al. 2007) and detecting potential competitors (Rick & Bakker, 2008; Siebeck et al. 2010; Bohórquez-Alonso et al. 2018). Underlying vision in all animals are the opsins: G-protein-coupled receptors that bind via a Schiff base linkage to a carotenoid-derived (Vitamin A1 or A2) chromophore to form visual pigments mediating light absorbance from around 300-700nm (Collin et al. 2009). UV-sensitivity is mediated by visual pigments with peak sensitivities shorter than 400 nm and are found in many vertebrates (Bowmaker, 2008), including multiple teleost lineages (Lin et al. 2017). However, for many teleost fishes we still lack detailed information on the identities, sites and molecular evolution of specific genes and regulatory pathways that facilitate vision, including UV-sensitivity.
Visual opsins from all five vertebrate opsin subfamilies can be found in teleosts, including the UV-sensitive or very-short-wavelength-sensitive 1 (SWS1), short-wavelength-sensitive 2 (SWS2), medium-wavelength-sensitive (MWS), rhodopsin 1 (rod opsin, RH1) and rhodopsin-like 2 (RH2), and long-wavelength-sensitive (LWS) opsins (Bowmaker, 1990; Bowmaker, 2008). All of these families arose from an ancient vertebrate opsin that underwent multiple whole-genome and individual gene duplication events (Van de Peer et al. 2009), the latter of which facilitated the acquisition of novel visual opsins in multiple teleost lineages (Hofmann & Carleton, 2009; Rennison et al. 2012; Musilova et al. 2019). Opsin gene function can be preserved via gene conversion; a major homogeniser and repair mechanism of paralogous genes (Katju & Bergthorsson, 2010), that may also resurrect pseudogenes (Cortesi et al. 2016). Currently most of what is known on teleost opsin gene duplications pertains to the short-, medium- and long-wavelength sensitive genes (i.e. SWS2, RH1, RH2 and LWS) for which functional paralogs are found in many species (Rennison et al. 2012; Cortesi et al. 2016), whereas the duplication and retention of SWS1 paralogs is considerably rarer (Rennison et al. 2012; Lin et al. 2017; Musilova et al. 2019).

UV-sensitivity in the teleost retina is conveyed by the expression of SWS1 and/or SWS2 opsins in the outer-segments of single cone cells, while sensitivity to longer-wavelengths is facilitated by the expression of RH2A and/or LWS in double cone cells (i.e. two fused single cones) (Carleton et al. 2005, 2008; Parry et al. 2005; Spady et al. 2006; Dalton et al. 2015; Härer et al. 2018; Stieb et al. 2019). Aside from a change in opsin, the spectral tuning of visual pigments can also be altered by an A1 to A2 chromophore switch mediated by a cytochrome P450 family member, Cyp27c1 (Enright et al. 2015) that can induce a substantial shift upwards of 60 nm in the wavelength of maximum spectral absorbance ($\lambda_{\text{max}}$) of a visual pigment (Govardovskii et al. 2000); however, it is changes in the opsin protein that drive most of the variation in spectral tuning (Nickle & Robinson,
2007; Carleton et al. 2020), particularly in the UV-region (Yokoyama et al. 2016). More specifically, the interaction between the opsin protein and lysine residues on the surfaces of the chromophore binding-pocket determine the visual pigment’s wavelength of maximum absorption sensitivity (Yokoyama, 2000; Kelber et al. 2003). Changes in the polarity and/or charge of amino acid residues at ‘key-tuning’ sites near the binding pocket can induce a short- or long-wavelength shift in spectral sensitivity (Carleton, 2009; Carleton et al. 2005; Terai et al. 2006). While the cumulative known tuning effects of these sites may be used to infer the peak spectral absorbance of visual pigments (Yokoyama et al. 2008), including SWS1-based pigments (Shi & Yokoyama, 2003; Yokoyama et al. 2016), there may be further changes in the retina that spectrally tune vision including differential opsin expression (Fuller et al. 2004; Johnson et al. 2013; Cortesi et al. 2016; Shimmura et al. 2017) and/or coexpression of opsins in the retina (Cheng & Flamarique, 2007; Dalton et al. 2014, 2017; Cortesi et al. 2016; Lührmann et al. 2019; Stieb et al. 2019). The combined effect of these visual tuning mechanisms on photoreceptor spectral absorbance are measurable using microspectrophotometry (MSP); a technique that combines microscopy and spectrophotometry (Dartnall et al. 1983; Kelber et al. 2003).

Many small bodied teleosts that inhabit shallow environments, rich in UV-wavelengths such as coral reefs often possess SWS1 opsin genes contained in single cones (Stieb et al. 2016, 2017; Lin et al. 2017; Musilova et al. 2019) and have UV-transmissive lenses (Losey et al. 2003; Marshall et al. 2006). UV-vision in reef fishes may aid the detection of UV-absorbing zooplankton prey (Stieb et al. 2017) and/or facilitate a short-distance communication channel hidden from predators, most of which lack UV-sensitive photoreceptors (Losey et al. 1999; Siebeck, 2004; Marshall et al. 2006; Siebeck et al. 2010). However, despite the widespread nature of UV-sensitivity and its importance in reef fishes, its genetic basis remains largely uncharacterised, and therefore was the aim of this study.
To do this, we used anemonefishes (subfamily, Amphiprioninae), which are an iconic group of reef fishes that obligately associate with one or more species of sea anemone. They are also sequential hermaphrodites living in strict social hierarchies governed by body size (Buston, 2003). The visual system of *Amphiprion akindynos* has been previously characterised revealing UV-sensitive single cones containing SWS1 opsin, in addition to other spectral sensitivities (~400, 520, 498, 541 nm λmax/λmax) (Stieb et al. 2019). Now, the public availability of short- and long-read sequenced genomes for eleven species of anemonefishes has made it possible to study in detail the evolution of SWS1 opsin genes in this group of reef fishes.

We first identified the SWS1 opsin genes and other visual opsin genes found in the genomes of eleven anemonefish species and provide information from their synteny analysis and phylogenetic reconstruction. Second, we identified the visual transduction and shut-off pathway genes that regulate opsin activity to show their synteny in all examined species. Third, we measured the opsin gene expression levels and photoreceptor spectral sensitivities in the retina (using MSP) of the false-clown anemonefish (*Amphiprion ocellaris*). This fundamental information on anemonefish vision will enable further studies to take a whole-organism investigative approach to relate visual genes with the retinal neuroanatomy and behavioural ecology of a reef fish.

**Material and Methods**

*Anemonefish visual opsin genes: identification, phylogeny and molecular analysis*

All genetic sequence analyses were performed using Geneious Prime (v. 2019.2.1). Our *in-silico* searches of anemonefish visual opsin genes involved annotating the regions
containing the SWS1, SWS2, RH2, LWS and RH1 opsin genes, along with their immediate upstream and downstream (flanking) genes in the genomes of eleven species of anemonefish. The publicly available assembled genomes for anemonefishes and a sister species, the Lemon damselfish, Pomacentrus moluccensis, were accessed from various sources including for A. ocellaris (Tan et al. 2018; accession no. NXFZ00000000.1), A. percula (Lehmann et al. 2018; “Nemo Genome DB, 2018”; accession no. GCA_003047355.1), A. frenatus (Marcionetti et al. 2018; https://doi.org/10.5061/dryad.nv1sv), A. akallopisos, A. melanopus, A. perideraion, A. nigripes, A. bicinctus, A. polynmus, A. sebae and Premnas biaculeatus (Marcionetti et al. 2019). Additional opsin gene sequences from other unrelated species including Anolis carolinensis (accession no. NM_001293118), Oreochromis niloticus (accession no. AY775108; AF247128; JF262086; JF262087), Pseudochromis fuscus (accession no. KP004335; KP017247), Danio rerio (accession no. AB087811; HM367062; AB087803; NM_001002443; AF109369; KT008394; NM_182892; NM_131254; KT008398; KT008399) and Oryzias latipes (accession no. AB180742; XM_004069094; NM_001104694; AB223056; AB223057) were obtained from genbank (www.ncbi.nlm.nih.gov/genbank/). Reference gene sequences from O. niloticus were used to detect anemonefish visual genes including opsin genes, visual transduction pathway genes and shutoff genes by mapping individual exons using low specificity (<30% similarity) against anemonefish genomes. Annotated gene sequences were confirmed by BLAST search in ENSEMBL v. 97 (accession date: 15.08.2019) (Zerbino et al. 2018) against the genomes of D. rerio (assembly no. GRCz11), G. aculeatus (assembly no. GCA_006229165.1) and O. niloticus (assembly no. GCA_001858045.3). Visual transduction and shut-off pathway genes were also identified in the A. percula and A. ocellaris genomes using predicted gene sequences obtained from Ensembl (ensembl.org) and their coding sequences were confirmed by assembly against the transcriptome of the A. ocellaris retina.
Phylogenetic trees based on a reference dataset containing 149 visual opsin coding sequences were generated using Bayesian inference in MrBayes v.3.2.7a (Huelsenbeck & Ronquist, 2001) in a workflow run through CIPRES (Miller et al. 2010) and viewed using Figtree v. 1.4.4 (Rambaut, 2018). Two opsin trees were reconstructed: 1) overall opsin tree in a GTR+I+G model using default settings, and 2) an identical model, but with SWS1 sequences restricted to only the fourth exon and intronic region between the fourth and fifth exons; where the majority of sequence variation exists between the two coding SWS1 genes. Both trees used Anolis carolinensis vertebrate ancestral (VA) opsin as an outgroup and were selected based on the best-fit model AIC from Jmodeltest2 (Darriba et al. 2012; see supplementary materials for output). The Bayesian reconstructions included MCMC searches for 10 million generations with two independent runs and four chains each, a sampling frequency of 1000 generations and a burn-in of 25%.

Visual opsin gene duplicates were tested for gene conversion: a phenomenon commonly found in teleosts (Owens et al. 2009; Watson et al. 2010; Nakamura et al. 2013; Cortesi et al. 2015; Sandkam et al. 2017; Escobar-Camacho et al. 2017, 2020). This was analysed using the program GARD (Genetic Algorithm Recombination Detection) (Kosakovsky Pond et al. 2006) on the aligned whole sequences of both the RH2A and SWS1 duplicates in A. ocellaris and A. percula to reveal the presence/absence of recombination. Identified sites of recombination were corroborated using different phylogenetic tree topologies based on whole opsin gene sequences, and sequences limited to suspected regions of recombination.

*Animals and ethics statement*

For visual gene expression analysis, lens transmission measurements and microspectrophotometry (described below), we used 13 captive-bred *Amphiprion ocellaris*...
(supplier Gallery Aquatica, Wynnum, 4178 QLD, Australia), which were housed in recirculating aquaria at the Queensland Brain Institute at The University of Queensland, Australia. Experiments were conducted in accordance with The University of Queensland’s Animal Ethics Committee guidelines under the approval numbers: QBI/304/16 and SBS/077/17.

**Visual gene expression analysis of A. ocellaris**

Adult retinas from two female (mean standard-length = 45 mm) and two male (mean standard-length = 30 mm) *A. ocellaris* were sampled for opsin gene expression analysis. All fish were kept under standard aquarium lighting for a minimum of three weeks (Fig. S1). Isolated retinas were homogenized using a high-speed bench-top homogenizer and total RNA was extracted using the QiaGen RNeasy Mini Kit. RNA was purified from any possible DNA contamination by treating samples with DNase following the protocol outlined by the manufacturer (QiaGen). The integrity of the extract was subsequently determined using a Eukaryotic Total RNA Nanochip on an Agilent 2100 Bioanalyzer (Agilent Technologies). Total RNA was sent to Novogene (https://en.novogene.com/) for library preparation and strand-specific transcriptome sequencing on a Hiseq2500 (PE150, 250~300 bp insert). Retinal transcriptomes were then filtered and transcripts *de novo* assembled on a customised Galaxy (v2.4.0.2; usegalaxy.org) (Afghan et al. 2016) workflow following the protocol described in de Busserolles et al. (2017). Individual cone opsin expression levels for single and double cones separately were then calculated as per de Busserolles *et al.* (2017). Rod versus cone opsin expression was calculated as the total proportion out of all mapped opsin reads.

To determine whether the identity of the chromophore could have a role in the tuning of MWS/LWS cone spectral sensitivities, we calculated the gene expression levels of the
protein-encoding gene CYP27C1; a catalyst in the conversion of an A1 to A2 chromophore (Enright et al. 2015). We searched and annotated A. ocellaris CYP27C1 by mapping D. rerio CYP27C1 against the A. ocellaris genome. The coding regions of CYP27C1 were mapped against retinal transcriptome assemblies and relative expression was quantified as the number of mapped reads per one million reads, as per Härer et al. 2018.

**Lens transmission and photoreceptor spectral sensitivities of A. ocellaris**

For the measurement of lens transmission in A. ocellaris, the lenses (n=3 fish) were isolated and rinsed in PBS to remove any blood and vitreous. Spectral transmission (300-800 nm) was then measured by mounting the lens on a drilled metal plate between two fibres connected to an Ocean Optics USB2000 spectrometer and a pulsed PX2 xenon light source (Ocean Optics). Light spectra were normalised to the transmission value at 700 nm (Siebeck & Marshall, 2001), and lens transmission values were taken at the wavelength at which 50% of the maximal transmittance (T50) was attained (Douglas & McGuigan, 1989; Siebeck & Marshall, 2001).

The spectral absorbance of A. ocellaris photoreceptors were measured in a total of nine fish using single-beam wavelength scanning microspectrophotometry (MSP). This procedure followed that outlined elsewhere (Levine & MacNichol, 1979; Cheney et al. 2009; Chung & Marshall, 2016). In short, small pieces (~1 mm²) of tissue were excised from the eyes of two-hour dark-adapted fish, then immersed in a drop of 6% sucrose (1X) PBS solution and viewed on a cover slide (sealed with a coverslip) under a dissection microscope fitted with an infra-red (IR) image converter. A dark scan was first taken to control for inherent dark noise of the machine and a baseline scan measured light transmission in a vacant space free of retinal tissue. Pre-bleach absorbance measurements were then taken by aligning the outer segment of a photoreceptor with the path of an IR measuring beam that
scanned light transmittance over a wavelength range of 300-800 nm. Photoreceptors were then bleached using bright white light for one minute. Successful (pre-bleach) cell scans were then re-measured and compared to photo-bleached scans of the same cell to confirm the observed light absorption was by a visual pigment. Confirmed visual pigment spectral absorbance data was then fitted using bovine rhodopsin as a template pigment to estimate the maximum absorbance ($\lambda_{\text{max}}$) value of the visual pigment (Govardovskii et al. 2000; Partridge & De Grip, 1991). Individual scans were binned based on their grouping of similar ($\leq 10$ nm difference) $\lambda_{\text{max}}$ values and then averaged across individuals.

Comparisons between anemonefish opsin protein sequences and those of other fishes with known $\lambda_{\text{max}}$ values were made to infer the spectral tuning effects of individual amino acid sites (see supplementary material). Estimates of anemonefish opsin $\lambda_{\text{max}}$ values were calculated from the known spectral absorbances of *O. niloticus* (Parry et al. 2005), *Oryzias latipes* (Matsumoto et al. 2006), *Lucania goodei* (Yokoyama et al. 2007), *Maylandia zebra* (Spady et al. 2006), *Dascyllus trimaculatus* (Hofmann et al. 2012) and *Pomacentrus amboinensis* (Siebeck et al. 2010). These opsins have been thoroughly studied using either MSP or the *in-vitro* reconstitution of opsin proteins for a direct measurement of pure protein spectral absorbance (Carleton 2009; Matsumoto et al. 2006; Yokoyama & Jia, 2020). Our analysis involved identifying variable amino acid residues located at sites within the retinal binding pocket attributed to a shift in polarity and/or substitutions at previously reported tuning sites (Nakayama & Khorana, 1991; Fasick et al. 1999; Neitz et al. 1991; Jan & Farrens, 2001; Nagata et al. 2002; Shi & Yokoyama, 2003; Carleton et al. 2005; Parry et al. 2005; Spady et al. 2006; Yokoyama et al. 2007; Yokoyama, 2008; Yokoyama & Jia, 2020; Lührmann et al. 2019). Opsin protein sequences were aligned using MAFFT alignment (MAFFT v. 7.388; Katoh & Standley, 2013) with bovine (*Bos taurus*) rhodopsin as a template (PDB accession no. 1U19).
Visual pigment spectral absorbance curves were modelled and fitted against measured absorbance spectra for the (co)expression of opsins. The shape of the absorbance curve was calculated according to standard A1 and A2 chromophore templates (see Govardovskii et al. 2000) for each visual pigment based on estimated $\lambda_{\text{max}}$ values. For modelling coexpression, the individual absorbance curves were multiplied by varying input proportions to simulate different amounts contained in the cone cells until the summed absorbance curve had a $\lambda_{\text{max}}$ value close to that measured, while still fitting the shape of the cone absorbance spectra, with particularly close adherence in the long-wavelength limb.

Results

Anemonefish visual opsin genes: identification, phylogeny and molecular analysis

Our in-silico searches in the genomes of anemonefishes and one damselfish (P. moluccensis) yielded a total of eight fully-coding opsin genes belonging to five subfamilies including one rod opsin RH1, two ultraviolet-sensitive SWS1 opsins, a single violet-sensitive SWS2 opsin, three blue-green-sensitive RH2 opsins, and a single yellow-red-sensitive LWS opsin (Fig. 1; Tab. S1). All visual transduction pathway genes and shutoff genes present in other vertebrates/fish species were also identified in A. ocellaris and A. percula, with no extra duplicates for these genes found (Tab. S2). Phylogenetic analyses placed all the identified opsin genes into distinct homologous clusters with their predicted opsin subfamilies (Fig. 1A). Furthermore, the translated and aligned protein sequences for the identified genes (Fig. S2) exhibited typical opsin traits including the conserved chromophore binding site residue (K296) and in most cases intact seven transmembrane domains; however, the latter was only found for all opsin genes in the A. ocellaris and A. percula genomes, whereas in other species the full gene sequences for some opsins could not be found including for duplicate RH2A-2 and SWS1/$\beta$ opsin genes. The sister-species, A. ocellaris and A. percula were found to share
an identical palette of opsins with highly conserved coding sequences (≥99% similarity) between all homologous pairs and both possessed an SWS1 pseudogene not found in other species.

All anemonefishes were found to have three blue-green sensitive opsin genes; RH2A-1, RH2A-2 and RH2B, that closely matched sequence predictions on ENSEMBL (ensembl transcript id: ENSAOCT0000002771.1, ENSAPET0000018561.1, ENSAOCT0000002796.1, ENSAPET0000018598.1, ENSAOCT0000002830.1, ENSAPET0000018632.1). Like in other examined teleosts, the RH2 genes (Fig. 1B) in A. ocellaris and A. percula were found in-tandem, spanning a region of approximately 29kbp and flanked by identical genes immediately up- and down-stream of the syntenic region. The translated protein sequences of both RH2A and RH2B (Fig. S2) encoded 353- and 346-amino acids (aa), respectively. RH2A paralogs were found to be highly conserved within A. percula and A. ocellaris, sharing 95.2% (RH2A-1) and 98% (RH2A-2) similarity, respectively.

The chromosomal resolution of the A. percula genome revealed the locations of the RH2 and LWS-SWS2B syntenic regions (Fig. 1B) in-tandem on Chromosome 6, separated by approximately 8.6mbp. A conserved syntenic region shared by many other neo-teleost species (Musilova et al. 2019). A single red-sensitive, LWS opsin and violet-sensitive, SWS2B opsin was identified in both A. percula and A. ocellaris that matched sequence predictions on ensembl (ensembl transcript id: ENSAOCT00000024298.1, ENSAPET0000033644.1, ENSAOCT00000031935.1, ENSAPET00000033670.1). The translated protein sequences of LWS and SWS2B were found to encode 358- and 352- aa, respectively. Phylogenetic analysis of SWS2B, RH2B and LWS in anemonefishes showed a typical pattern of species-relatedness that for the most part resembled the inferred phylogenies reported elsewhere (Litsios et al. 2014; Rolland et al. 2018).
Two UV-sensitive opsins, SWS1α and SWS1β, were identified in all examined anemonefishes. Those found in A. percula and A. ocellaris, matched the sequence predictions on ensembl (ensembl transcript id: ENSAOCT00000007343.1, ENSAOCT00000031935.1, ENSAPET00000033679.1, ENSAPET00000035052.1). Protein sequence analysis revealed highly conserved SWS1 opsins, where both encoded 339 aa (Fig. S2) and shared 96% similarity. Closer examination of the protein sequences showed that 15% of the differences between paralogs occurred at known SWS1 tuning sites (Yokoyama, 2008; Shi & Yokoyama, 2003), and 90% were within the seven transmembrane domains. All SWS1 protein sequences had conserved F86 and S90 aa sites, an invariable feature of UV-sensitive opsins (Cowling et al. 2002). The A. percula genome showed that both SWS1 genes were located in-tandem on Chromosome 21, separated by 19kbp and spanning a total syntenic region of 29kbp. A third non-functional SWS1 pseudogene (Fig. 1B) was also found between the two paralogs (encoding 196 amino acids), with highest sequence homology to the initial two exons of SWS1β. The genes flanking the SWS1 region were identical across all species. Initial SWS1 phylogeny based on complete coding sequences revealed a partial clustering of two gene groups; the paralogs SWS1α and SWS1β (Fig. S3). Re-construction of the phylogeny used only Exon 4 and part of the intron between exons 4 and 5, where most of the informative sites (i.e. single nucleotide polymorphisms, ‘SNPs’) were located, and provided a clearer separation of the paralogs into two distinct clusters (Fig. 1A).

Like most teleost fishes (Musilova et al. 2019), anemonefishes were found to possess a single RH1, with a conserved RH1 syntenic region. Similar to most cone opsin genes, the RH1 phylogeny closely followed the species phylogeny (Fig. 1A).
Figure 1. (A) Anemonefish visual opsin gene phylogeny reconstructed using the fourth exon and intron of SWS1α and SWS1β. Shown are internodal Bayesian posterior probabilities (P) depicted in dark grey and light grey markers that indicate posterior probabilities greater than 0.95 and 0.90, respectively. Posterior probabilities less than 0.90 are given in full. For the purpose of presentation, all clustered outgroup branches were collapsed. To view the complete non-collapsed presentation, all clustered outgroup branches were collapsed. To view the complete non-collapsed phylogeny, see the supplementary material. (B) Synteny of anemonefish visual genes found in the genomes of including opsin genes and their flanking genes (yellow), wavelength-sensitive, SWS1* = Short-wavelength-sensitive 1/Short-wavelength-sensitive (rod opsin), RH2 = Rhodopsin 2, SWS2 = Short-wavelength-sensitive 2, LWS = Long-wavelength-sensitive.

**Gene acronyms**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNG</td>
<td>Cyclic nucleotide gated channel</td>
</tr>
<tr>
<td>GRK</td>
<td>G protein-coupled receptor kinase</td>
</tr>
<tr>
<td>GNAT</td>
<td>G protein subunit alpha transducin</td>
</tr>
<tr>
<td>GNB</td>
<td>G protein subunit beta</td>
</tr>
<tr>
<td>ARR</td>
<td>Arrestin</td>
</tr>
<tr>
<td>MAGI</td>
<td>MAGUK family protein</td>
</tr>
<tr>
<td>PDE</td>
<td>Phosphodiesterase</td>
</tr>
<tr>
<td>RH</td>
<td>Rhodopsin</td>
</tr>
<tr>
<td>SWS</td>
<td>Short-wavelength-sensitive</td>
</tr>
<tr>
<td>LWS</td>
<td>Long-wavelength-sensitive</td>
</tr>
<tr>
<td>SAG</td>
<td>S-antigen visual arrestin</td>
</tr>
<tr>
<td>GNAT2</td>
<td>G protein subunit alpha transducin</td>
</tr>
<tr>
<td>ARM</td>
<td>Arrestin</td>
</tr>
<tr>
<td>GRK</td>
<td>G protein-coupled receptor kinase</td>
</tr>
</tbody>
</table>
We found evidence of gene conversion in both the RH2A and SWS1 duplicates in *A. ocellaris* and *A. percula* (Fig. 2). GARD analysis revealed two major breakpoints in RH2A-I and RH2A-2 located at 634-892 bp (exons 3 and 4) and 893-1,059 bp (exons 4 and 5), that indicated recombination had occurred within the last few exons. Alternative tree topologies based on different RH2A breakpoint regions (Fig. 2) supported one of the two regions (892-1,059 bp) having recombination, as evident by the orthologous grouping of RH2A genes rather than paralogous grouping that was observed when tree topologies were based on either the 634-892 bp region or complete coding sequence. Analysis of the aligned RH2A opsin protein sequences identified only one known key tuning site (aa site 292; site number given according to bovine rhodopsin) (Yokoyama & Jia, 2020) located within the region of recombination.

One major breakpoint was found in SWS1α and SWS1β indicating that recombination occurred in the last two exons (839-1,017 bp). Alternative tree topologies supported the notion of gene conversion within this breakpoint region (Fig. 2), where opsin copies grouped as orthologs or as paralogs depending on whether trees were constructed using the breakpoint region or the complete coding sequence, respectively. Analysis of the aligned SWS1 opsin protein sequences did not yield any known tuning site changes within the recombined region.

We note that aside from *A. percula* and *A. ocellaris*, the translated protein sequences of RH2A-2 and SWS1 opsin genes were not examined in other anemonefishes due to the inability to extract complete gene sequences and in the case of SWS1β, the presence of premature stop codons. Whether these incomplete/pseudogenized genes are a true biological occurrence or sequencing errors as a result of short-read assembled draft genomes remains to be investigated.
Visual gene expression analysis of *A. ocellaris*

Analyses of retinal transcriptomes (Fig. 3) revealed that under aquarium lighting the *A. ocellaris* retina (n=4) expressed one rod opsin; *RHI* (mean ± s.d. = 70.63 ± 11.51%), and six cone opsins including four double cone opsins; *RH2A-1* (43.13 ± 6.68%), *RH2B* (35.34 ± 3.21% *LWS* (11.45 ± 6.77%) and *RH2A-2* (10.08 ± 11.66%), and two single cone opsins; *SWS1β* (59.09 ± 9.41%) and *SWS2B* (40.8 ± 8.70%) (Fig. 3). No trace of *SWS1α* expression was detected in three out of the four retinas and only a low amount (1.5%) was detected in
one retina. No difference in opsin expression was found between males (n=2) and females (n=2) (Pairwise ANOVA, p-value > 0.05).

The gene CYP27C1 was found to be expressed in very low amounts in all four A. ocellaris retinae, where relative expression levels ranged from 0.92 to 4.04 mapped reads per million reads.

**Figure 3.** Relative cone opsin expression levels in the retina of captive A. ocellaris (N=4; 2 females and 2 males) kept under aquarium lighting (see supplementary material for illumination spectra). SWS1 = Short-wavelength-sensitive 1, SWS2 = Short-wavelength-sensitive 2, RH2 = Rhodopsin 2, LWS = Long-wavelength-sensitive. Lines represent the median proportion of opsin expression relative to the total opsin expression levels of single cones opsins (SWSs) and double cone opsins (RH2s, LWS), respectively. Individual points represent a sampled fish.

The gene CYP27C1 was found to be expressed in very low amounts in all four A. ocellaris retinae, where relative expression levels ranged from 0.92 to 4.04 mapped reads per million reads.

**Lens transmission and photoreceptor spectral sensitivities of A. ocellaris**

Transmission measurements on the lenses of A. ocellaris (individual T50 values = 320/339/362 nm) were averaged to give a mean T50 wavelength value of 341 nm (Fig. 4A). This T50 value was used to correct the photoreceptor spectral absorbance curves for the filtered light reaching the retina.

MSP in A. ocellaris (Fig. 4A) found one medium-wavelength-sensitive (MWS) rod (mean $\lambda_{\text{max}}$ value = 502 ± 2.0 nm; n=10 cells, N = 4 fish; Fig. 4A), one short-wavelength-sensitive (SWS) single cone, three MWS double cones, and one LWS double cone. Unless otherwise mentioned, all measured absorbance spectra fitted an A1 visual pigment template.

Single cone $\lambda_{\text{max}}$ measurements (mean $\lambda_{\text{max}}$ value = 386 ± 4.6 nm; n=4, N = 4 fish; Fig. 4A) did not closely match any of the $\lambda_{\text{max}}$ estimates for SWS1 ($SWS1\alpha$ estimated $\lambda_{\text{max}}$...
value = 356 nm; SWS1β estimated $\lambda_{\text{max}}$ value = 370 nm; Fig. S2) or SWS2B (estimated $\lambda_{\text{max}}$ value = 407 nm); however, modelling the absorbance curve from the coexpression of SWS1β (40%) and SWS2B (60%) provided a close fit ($\lambda_{\text{max}}$ value = 384-386 nm; Fig. 4B). The single cone averaged absorbance spectra had an asymmetrical shape with a breadth more like an A2 visual pigment template.

One MWS double cone had $\lambda_{\text{max}}$ measurements (mean $\lambda_{\text{max}}$ value = 496.7 ± 0.58 nm; n=3, N = 3 fish; Fig. 4A) that closely matched the estimate for RH2B (estimated $\lambda_{\text{max}}$ value = 497 nm; Fig. S2), while another double cone had $\lambda_{\text{max}}$ measurements (mean $\lambda_{\text{max}}$ value = 515 ± 2.1 nm; n=6, N = 3 fish; Fig. 4A) that closely matched one of the estimates for RH2A-1/RH2A-2 (estimated $\lambda_{\text{max}}$ value = 516 nm; Fig. S2). A separate $\lambda_{\text{max}}$ estimate for RH2A-2 opsin could not be made, as the variable amino acid sites between the paralogs had unknown effects on spectral tuning and were absent in all reference sequences. One of these unknown variable sites, T266V exhibited a polarity shift that consistently alternated between the RH2A paralogs of A. percula and A. ocellaris.

Two measurements in one individual hinted at the presence of an intermediate MWS double cone (508/509 nm $\lambda_{\text{max}}$, N = 2 fish; Fig. 4A), with a peak spectral sensitivity that did not match any visual pigment estimates. Although it was reproducible by modelling the coexpression of RH2B (35%) and RH2A-I (65%) ($\lambda_{\text{max}}$ value = 507-509 nm; Fig. 4C), there was little difference to the modelled absorbance spectra of a lone visual pigment (Fig. 4C).

Similarly, the longer-wavelength sensitive double cone measurements (531/538 nm $\lambda_{\text{max}}$, N = 2 fish; Fig. 4A) provided a poor match to the LWS opsin estimate (estimated $\lambda_{\text{max}}$ value = 560 nm; Fig. S2) and could only be produced by modelling the coexpression of RH2A-I (55%) with LWS (45%) ($\lambda_{\text{max}}$ value = 529-531 nm; Fig. 4D), and RH2A-I (40%) with LWS (60%) ($\lambda_{\text{max}}$ value = 537-539 nm; Fig. S4), respectively. However, the latter provided an overall poor fit owing to the high amount of noise in the absorbance data that
combined with only having two measurements made it also difficult to compare the fit between A1 and A2 visual pigment templates.

![Graphs showing absorbance spectra](image)

**Figure.** (A) Normalised average lens transmission of *A. ocellaris*, and absorbance curves fitted to average rod and cone absorbance spectra measured in *A. ocellaris* (N=9). Modelled spectral absorbance curves for different cone opsin coexpression scenarios for (B) single cone ($\lambda_{max} = 385$ nm) coexpressing SWS1α (40%) and SWS2B (60%), (C) double cone ($\lambda_{max} = 508/509$ nm) coexpressing RH2B (35%) with RH2A-1 (65%), and (D) double cone ($\lambda_{max} = 531$ nm) coexpressing RH2A-1 (55%) with LWS (45%). Plotted alongside for comparison are modelled spectral absorbance curves for lone visual pigments and actual measured cone absorbance spectra.

**Discussion**

Anemonefish genomes revealed visual opsin orthologs belonging to all five ancestral vertebrate subfamilies (Davies et al. 2012) including SWS1, SWS2, RH1, RH2, and LWS, along with additional SWS1 and RH2A duplicates. We found rare instances of UV-sensitive SWS1 opsin gene duplication events that produced two functional paralogs (SWS1α and SWS1β) and a pseudogene. Moreover, the expressed cone opsin palette in the adult retinæ of *A. ocellaris* comprised of SWS1β and SWS2B single cone opsins, in addition to RH2A-1, RH2A-2, RH2B, and a small amount of LWS in double cone opsin genes. Lens transmission measurements and MSP strongly implicate UV-sensitivity in *A. ocellaris* including the
presence of UV-transmissive lenses and retinas that contain a single cone with a λ_max value of ~386 nm. Additional spectral sensitivities found by MSP in *A. ocellaris* included one rod with a λ_max value of 502 nm and four double cone spectral sensitivities with λ_max values of 497 nm, 515 nm, ~535 nm, and a possible fifth cone with a λ_max value of ~508 nm. We assume that *A. ocellaris* SWS opsins are only expressed in single cones, while the longer-wavelength sensitive RH2 and LWS genes are strictly-expressed in the double cones, as has been found in the anemonefish, *A. akindynos* (Stieb et al. 2019) and in cichlid fishes (Carleton et al. 2005, 2008; Parry et al. 2005; Spady et al. 2006; Dalton et al. 2015; Härer et al. 2019).

Visual opsin genes found in the genomes of anemonefishes are situated in conserved regions that share a common synteny with other teleost fishes (Lin et al. 2017). It also appears that anemonefish *SWS1* and *RH2* paralogs emerged through tandem duplication, as is common for opsin paralogs in many fish species (Cortesi et al. 2015; Lin et al. 2017; Musilova et al. 2019). It is extremely rare to find an *SWS1* opsin gene duplicate in teleosts (Rennison et al. 2012; Lin et al. 2017; Musilova et al. 2019). Here, we show that *SWS1* opsin duplicates are conserved across eleven members of Amphiprioninae and in a species of damselfish (*P. moluccensis*), as similarly reported in *Chromis chromis*; a member of the Pomacentrid subfamily, Chrominae (Musilova et al. 2019). Conversely, only one *SWS1* opsin gene was reported in the damselfish, *Stegastes partitus* (Musilova et al. 2019), and hence not all Pomacentridae possess *SWS1* opsin duplicates. Thus, it remains unclear whether the gene duplication event that produced *SWS1α* and *SWS1β* was lineage- or subfamily-specific. More genomic and/or transcriptomic data is needed to resolve whether this *SWS1* gene duplication occurred at the base of Pomacentridae or multiple times independently in its radiated lineages.
Interestingly, we found evidence of a second gene duplication event that produced a third pseudogenised SWS1 paralog in the genomes of *A. ocellaris* and *A. percula*. This SWS1 pseudogene was more alike *SWS1β* than *SWS1α*, suggesting it was either a duplicate of the former or gene conversion had occurred. The genomes of these species were published by two different lab-groups (Tan et al. 2018; Lehmann et al. 2018) and assembled differently, making it unlikely that this pseudogene is an assembly artefact. Due to the basal placement of these sister-species in the anemonefish phylogeny (Santini & Polacco, 2006), it remains possible this second duplication event occurred relatively early in anemonefish evolution. Alternatively, this duplication-event was specific to the clade containing *Premnas* and the sister-species, or even limited to only the sister-species. This pseudogene was not detected in the other eight anemonefish genomes; however, this was likely owing to these draft genomes being comparatively poorly resolved, that similarly meant only the fourth and fifth exons of *SWS1α* could be identified in those species. This incomplete or incorrect assembly of highly similar genomic regions is an inherent problem of short-read draft genomes.

Both SWS1 opsin protein sequences in *A. ocellaris* and *A. percula* had a conserved alanine at site 86 and serine at site 90; crucial for UV-sensitivity (Shi & Yokoyama, 2003; Tada et al. 2009). Furthermore, both anemonefish SWS1 duplicates have intact open reading frames but only *SWS1β* is expressed in the adult retina. This combined with a UV-transmissive lens and presence of a UV/violet sensitive single cone provide a strong indication of UV-vision mediated by SWS1 and SWS2B visual pigments in the single cones. Like other damselfishes, the examined anemonefishes all possess the violet sensitive SWS2B opsin (Hofmann et al. 2012; Stieb et al. 2016; Stieb et al. 2017). Moreover, no indication for the percomorph-specific SWS2Aα and SWS2Aβ was found, likely due to a double-gene loss in the Pomacentrid ancestor (Cortesi et al. 2015).
The large mismatch between spectral sensitivity measurements and estimates of *A. ocellaris* single cones and their opsins (SWS1 and SWS2B) suggest the measured spectral sensitivities are not produced by the expression of either opsin alone. Rather, it is more likely that *SWS1β* and *SWS2B* are coexpressed in the single cones to produce an intermediate absorbance curve; a phenomenon supported by fitting the absorbance spectra to the modelled equal coexpression of both opsins. Moreover, the asymmetrical and broad absorbance curve of the *A. ocellaris* single cone resembles that found in other species such as *A. akindynos*, where the single cones have a broad absorbance curve and \( \lambda_{\text{max}} \) value of 400 nm, despite SWS1 and SWS2B having estimated \( \lambda_{\text{max}} \) values of 370 nm and 407 nm, respectively (Stieb et al. 2019). Likewise, the coexpression of SWS visual pigments has been reported in the single cones of African cichlids (Dalton et al. 2017). Although not examined in this study, there can also be spatial patterns in the coexpression of SWS opsins across the retina, such as in *A. akindynos*, where coexpression is localised to a small, dorso-temporal (i.e. forward-looking) area of high acuity that may aid specific tasks (Stieb et al. 2019). Together, the close to equal ratio of *SWS1α* and *SWS2B* expression levels in the *A. ocellaris* retina, along with the failure to find a pure UV-sensitive and/or violet-sensitive cone strongly supports the notion of opsin coexpression in the single cones. Future in-situ mapping of cone opsin expression across the *A. ocellaris* retina will resolve the spatial patterns of opsin (co)expression.

The importance of UV-vision in anemonefishes is unclear but looking to other UV-sensitive teleosts could direct further investigation. Like the anemonefishes, the rainbow trout (*Oncorhynchus mykiss*) has retained two *SWS1* opsins that originated from a family-specific whole-genome duplication event, and their UV-sensitivity improves zooplanktivory efficacy by enhancing prey contrast (Flamarique & Hawryshyn, 1994; Flamarique, 2013). Similar foraging benefits conveyed by UV-sensitivity have been shown in perch (Loew et al. 1993), cichlids (Jordan et al. 2004), sticklebacks (Rick et al. 2012), and zebrafish (Flamarique,
Indeed, in damselfishes higher SWS1 expression correlates strongly with zooplanktivory (Stieb et al. 2017). Anemonefishes are life-long zooplanktivores (Daphne & Gerald, 1997) that could similarly benefit from the enhanced UV contrast of prey in the water column.

Fish also use UV-signals for communicating with rivals and mates, as reported in the guppy (Smith et al. 2002), stickleback (Rick et al. 2007, 2008) and a damselfish (Siebeck et al. 2010). In A. akindynos, the coexpression of SWS1 and SWS2 is believed to increase the chromatic contrast of its skin colour patterns that may improve the detection of conspecifics (Stieb et al. 2019). The use of UV-signalling in communication is worth further investigation in anemonefishes, as they possess UV-reflective skin patterns (Marshall et al. 2006; Stieb et al. 2019) that could conceivably be used to communicate their dominance status to other members within a family group and/or convey the occupied-status of their hosted anemone to the members of rival groups from nearby anemones.

Interestingly, the advantages of UV-vision could be conceivably conveyed by a single UV-sensitive opsin, and the functional significance of possessing SWS1 duplicates remains unknown. Another teleost species found to have two functionally coding SWS1 duplicates is the smelt, Plecoglossus altivelis and like A. ocellaris it only expresses one paralog (Minamoto & Shimizu, 2005). As similarly suggested in P. altivelis, it is possible that the second SWS1 paralog is expressed by A. ocellaris during specific seasons such as winter months on the reef, when there is higher ambient UV and lower turbidity promoting the more efficient transmission of shorter wavelengths of light. Increased SWS1 opsin expression has been found in the damselfish, P. nagasakiensis during winter months and may be a visual tuning response for taking advantage of the higher UV (Stieb et al. 2016). Seasonal changes in opsin gene expression levels have been reported to alter colour perception in widespread taxa (see review by Shimmura et al. 2018).
Analysing opsin expression levels in the larval and/or early-juvenile anemonefish retina may also reveal whether the shorter-wavelength sensitive $SWS1\alpha$ is expressed during earlier life stages, and whether an ontogenetic shift from $SWS1\alpha$ to $SWS1\beta$ occurs as a response to change(s) in light environment, particularly during the settlement stage when pelagic larvae return to the reef to seek a host anemone. Ontogenetic changes in cone opsin expression levels as a response to change in habitat have been reported in other reef fishes including the spotted unicornfish ($Naso brevirostris$) (Tettamanti et al. 2019) and dusky dottyback ($P. fuscus$) (Cortesi et al. 2016). Alternatively, the $SWS1\alpha$ opsin may be expressed in other tissues and/or organs, where animal opsins have been demonstrated to serve other sensory modalities including non-image forming vision, thermosensation, hearing (see review by Leung & Montell, 2017) and most recently in taste (Leung et al. 2020).

The anemonefish double cones expressed a greater variety of opsins including two $RH2A$ paralogs ($RH2A-1$ and $RH2A-2$), $RH2B$ and $LWS$, almost all of which have been found in other pomacentrids (Hoffmann et al. 2012; Stieb et al. 2016; Stieb et al. 2017). However, to our knowledge, this is the first report of an $RH2A$ duplication in a pomacentrid. Similar $RH2A$ duplications have been reported in most percomorphs (as reviewed by Musilova et al. 2019). It was previously difficult to separate $RH2A$ duplicates in pomacentrids due to their high degree of similarity, coupled with the low-resolution genomes available. Anemonefish $RH2A$ duplicates share a high level of similarity partly due to gene conversion detected in the fourth and fifth exons; a region that may be preserved to maintain the opsin-chromophore binding site (Lys296), as is likely the case in $SWS1$. However, unlike the region of gene conversion found in $SWS1$, that found in $RH2A$ also encompasses a known tuning site, bovine rhodopsin site number 292, where a large shift of -11 nm occurs when Ala is substituted with Ser (Yokoyama & Jia, 2020). Thus, preservation of this site by gene conversion may also have importance in maintaining the spectral tuning of $RH2A$ visual pigments.
Based on only the spectral sensitivities of the double cones, the spatial patterns of opsin expression in the anemonefish retina cannot be inferred. However, we speculate the most likely patterns of opsin expression in anemonefish double cones. In *A. ocellaris*, it appears that *RH2A-1* opsin (\(\lambda_{\text{max}}\) value = 516 nm) and *RH2B* opsin (\(\lambda_{\text{max}}\) value = 497 nm) are expressed in separate double cone members, as indicated by the close matches between measurements of cone spectral absorbance and estimates of visual pigment tuning. It is unclear whether the two visual pigments are also coexpressed in another double cone to produce an intermediate spectral sensitivity (\(\lambda_{\text{max}}\) value = 508/509 nm), as the possible expression of a short-wavelength shifted *RH2A-2* opsin cannot be out-ruled. Multiple variable amino acid sites were found between the *RH2A* duplicates that could not be accounted for when estimating tuning effects on the opsin. Moreover, many of these substitutions exhibit changes in polarity suggesting a likely effect on spectral tuning, for example site T266V that alternated between paralogs and may be a key tuning site in anemonefish RH2A tuning. In cichlids, the differential expression of duplicate *RH2A* opsin in the double cones is believed to improve blue-green sensitivity that may aid in viewing nuptial skin colours and/or colonising different depths (Weadick & Chang, 2012; Dalton et al. 2014). Moreover, the coexpression of *RH2A* paralogs with either *RH2B* or *LWS* opsin have also been shown to enhance luminance contrast (Dalton et al. 2014). Modelling the coexpression of anemonefish *LWS* and *RH2A-1/RH2A-2* opsins indicated it was likely responsible for producing the longer-wavelength sensitive double cones measured in the *A. ocellaris* retina. The in-situ mapping of cone opsin expression across the anemonefish retina is needed to resolve spatial patterns of opsin gene expression. Alternatively, it is feasible that a switch from an A1 to A2 chromophore could contribute towards the MWS/LWS tuning of the double cones; however, this seems unlikely when considering the very small amounts of *CYP27C1* expression.
detected in the *A. ocellaris* retina that were a couple orders of magnitude below levels typically detected in cichlids that express an A2 chromophore (Härer et al. 2018).

**Conclusions and future directions**

Here we have shown that anemonefishes possess eight visual opsin genes including duplications of the *SWS1* and *RH2* genes. Moreover, most of these opsins are expressed in the adult retinae of *A. ocellaris*. The presence of two functional *SWS1* opsins in all examined anemonefishes suggests an anemonefish specific *SWS1* gene duplication event, and a possibly clade-specific second duplication event that produced an *SWS1* pseudogene. Circumstantial evidence including the expression of *SWS1* opsin, a UV-transmissive lens and presence of UV/violet-sensitive cone cells in the *A. ocellaris* retina strongly implicate UV-vision. Finally, we hope that our characterisation of the anemonefish visual genes will encourage future experiments using them as a model for studying reef fish vision.

**References**


https://doi.org/10.1126/sciadv.aao4709


https://doi.org/10.1016/j.cub.2015.10.018


https://doi.org/10.1111/mec.13957


https://doi.org/10.1111/mec.15474


**Bioinformatics Applications Note, 17.** Retrieved from Yang and Rannala website:


https://doi.org/10.3390/genes1030452


