

Opsin Expression in the Central Nervous System of the Mantis Shrimp *Neogonodactylus oerstedii*

MARY W. DONOHUE^{1,*}, KAREN L. CARLETON², AND THOMAS W. CRONIN¹

¹*Department of Biological Sciences, University of Maryland, Baltimore County, 1000 Hilltop Circle, Baltimore, Maryland 21250; and* ²*Department of Biology, University of Maryland, College Park, Maryland 20742*

Abstract. Visual pigments, each composed of an opsin protein covalently bound to a chromophore molecule, confer light sensitivity for vision. The eyes of some species of stomatopod crustaceans, or mantis shrimp, can express dozens of different opsin genes. The opsin diversity, along with spectral filters and unique tripartite eye structure, bestow upon stomatopods unusually complex visual systems. Although opsins are found in tissues outside typical image-forming eyes in other animals, extraocular opsin expression in stomatopods, animals well known for their diversity of opsins, was unknown. Caudal photoreception in the central nervous system of decapod crustaceans, a group closely related to stomatopod crustaceans, is thought to be opsin based. However, electrophysiological data suggest that stomatopods do not have caudal photoreceptors. In this study, we identified mRNAs that could encode four different opsins and several components of a potential G_q-mediated phototransduction pathway in the central nervous system of the Caribbean mantis shrimp *Neogonodactylus oerstedii*. The four opsins are abundantly expressed in the cerebral ganglion, or brain, with little or no expression in the remainder of the ventral nerve cord. Our data suggest that

there are previously undiscovered cerebral photoreceptors in stomatopods.

Introduction

Vision scientists have studied the extraordinarily complex visual systems of stomatopod crustaceans, or mantis shrimp, for decades. The eyes of some species have more than a dozen physiologically distinct classes of receptors, many specialized for ultraviolet (UV), color, or polarization vision. Vision in stomatopods and other animals requires light-sensitive visual pigments, composed of a chromophore bound to an opsin protein, to initiate the phototransduction cascade.

There is a growing body of evidence that extraretinal opsin expression is commonplace across many diverse animal phyla, although the physiological functions associated with these opsins remain elusive in many cases. Opsin proteins are commonly found in central nervous systems (*e.g.*, crayfishes) and dermal tissues (*e.g.*, cephalopods), even in animals capable of image-forming vision (Kingston *et al.*, 2015; Kingston and Cronin, 2015; reviewed by Henze and Oakley, 2015; Kingston and Cronin, 2016). Although *Limulus* horseshoe crabs have several sets of eyes, the photosensitive segmental ganglia and tail also express several opsins (Battelle, 2016; Battelle *et al.*, 2016). The functional relevance of opsin coexpression in extraocular tissues is unclear, but it is thought that a single photopigment could detect irradiance, whereas multiple photopigments with different spectral sensitivities and located in different photoreceptor cells could reliably photoentrain circadian rhythms to dawn and dusk *via* wavelength discrimination (Roenneberg and Foster, 1997; Bertolucci and Foa, 2004). For example, the UV-, blue-, and long-wavelength-sensitive (LWS) opsins expressed in the optic lobes of hawkmoths are thought to entrain photoperiodic rhythms, since they are co-

Received 15 January 2017; Accepted 11 July 2017; Published online 25 October 2017.

* To whom correspondence should be addressed. E-mail: willard3@umbc.edu.

Abbreviations: A1–6, first through sixth abdominal ganglia; *Arr*, arrestin; CG, cerebral ganglion; *DGK*, diacylglycerol kinase; *Gprk*, G-protein-receptor kinases; *Gα_q*, G_q protein alpha subunit; *Gβ*, G protein beta subunit; *Gγ*, G protein gamma subunit; LWS, long-wavelength-sensitive; MWS, medium-wavelength-sensitive; *PKC*, protein kinase C; *PLC*, phospholipase C; *rdgB*, phosphatidylinositol transfer protein; SEG, subesophageal ganglion; T7–9, thoracic ganglia; *TRP*, transient receptor potential channel.

expressed with melatonin and arrestin (Lampel *et al.*, 2005). The presence and location of opsin expression in extraocular tissues provide circumstantial evidence for extraocular photosensitivity and, in some cases, supplement behavioral and physiological data for functional extraocular photoreceptors.

Decades before molecular techniques advanced the search for opsins, photosensitivity in the sixth abdominal ganglion (also called the caudal photoreceptor) had been detected in several decapod crustaceans (an order that includes lobsters and crayfishes), which could evoke a tail flexion escape response (Wilkens and Larimer, 1976; Edwards, 1984). More recently, Kingston and Cronin (2015) detected two opsins throughout the CNS of the crayfish *Procambarus clarkii*, including the location of the caudal photoreceptor. In this crayfish, opsin protein is expressed both in the eyes and in two clusters of cell bodies located in the cerebral ganglion (CG), or brain (Kingston and Cronin, 2015). The cerebral clusters of cell bodies connect to fibers extending to the subesophageal ganglion (SEG) and along the entire length of the nerve cord (Kingston and Cronin, 2015). The data by Kingston and Cronin (2015) suggest that opsins in decapod crustaceans confer photosensitivity throughout the entire CNS, not just to the caudal photoreceptor.

Both decapods (order Decapoda) and stomatopods (order Stomatopoda) belong to the same class of crustaceans (class Malacostraca). The decapod caudal photoreceptor, identified by molecular, behavioral, and physiological techniques, is not thought to exist in other malacostracan crustaceans. For example, it was not detected in the only stomatopod that has been tested, *Squilla empusa* (Wilkens and Larimer, 1976). Nevertheless, stomatopods have visual systems unparalleled in complexity and a sixth abdominal ganglion that is larger than the other five abdominal ganglia (shown in the scale drawing of the CNS in Fig. 1). Thus, stomatopods are promising models for probing the roles and evolutionary histories of extraocular photoreceptors.

The visual systems of stomatopods are noteworthy for their abundance of expressed opsins, a result of recent gene duplication (Porter *et al.*, 2009, 2013). A seminal study by Porter *et al.* (2013) identified a surprisingly large number of opsin transcripts in the eyes of stomatopods—up to 33 transcripts—that far outnumber physiologically distinct photoreceptor classes determined by microspectrophotometry or electrophysiology. Many of these opsins appear to have diverged at key spectral tuning and signal phototransduction sites (Porter *et al.*, 2013), but it is unclear whether all of the 33 opsins are responsible for the physiologically distinct receptor classes or whether the co-expression of several opsins tunes the photoreceptors' spectral sensitivities. These unusual crustaceans often occur in shallow tropical water, where light throughout the spectrum is abundant and can potentially penetrate the outer carapace to illuminate internal tissues. This, taken together with evidence for extraocular photoreception in the nerve cords of

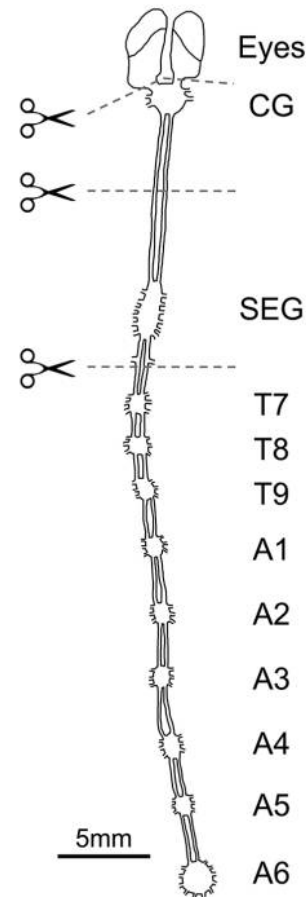


Figure 1. The central nervous system (CNS) of *Neogonodactylus oerstedii* includes the cerebral ganglion (CG), subesophageal ganglion (SEG), three thoracic ganglia (T7–9), and six abdominal ganglia (A1–6). The CNS was immediately cut into three regions upon dissection, and the regions were separated for RNA extraction and subsequent transcriptome analyses. RNA was isolated from the CG only, from the SEG only, and from the pooled thoracic and abdominal ganglia (T7–9, A1–6).

decapods, warrants further investigation into whether the abundance of stomatopod opsin transcripts found in retinas extends to extraretinal tissues and whether they there form functional photopigments that contribute to a significant physiological role.

This is the first study to search for genes that may encode photopigments in extraretinal tissues of stomatopods. We utilized data generated by RNA sequencing (RNA-Seq) to search for opsins and several potential phototransduction components in CNS tissue of *Neogonodactylus oerstedii* (Hansen, 1895). We found transcripts for one middle-wavelength-sensitive (MWS) opsin and three LWS opsins expressed in the CG, or brain, of *N. oerstedii*. Additionally, we found a dozen genes that putatively encode key G_q-mediated phototransduction components in the CNS of *N. oerstedii*. It is possible that opsins expressed in the CG confer light sensitivity for nonvisual functions. Further investigation will be needed to determine this.

Materials and Methods

Animals

Caribbean mantis shrimp (*Neogonodactylus oerstedii*) collected from the Florida Keys, Florida, were shipped to the University of Maryland, Baltimore County. The stomatopods were housed in plastic aquaria with approximately 800 ml artificial seawater, supplied with polyvinyl chloride tube burrows, and exposed to a light/dark cycle of 12 h:12 h daily. The animals were fed frozen shrimp weekly, and the seawater was changed after feedings.

RNA extraction

The CNSs from 10 *N. oerstedii* individuals were removed using Roboz dissecting tools that were treated with RNase Zap (ThermoFisher Scientific, Waltham, MA) and 0.1% diethylpyrocarbonate (DEPC) water to prevent RNase contamination. Each CNS was cut into 3 regions: (1) the CG only, (2) the SEG only, and (3) the pooled thoracic and abdominal ganglia (T7–9 and A1–6) and associated ventral nerve cord (Fig. 1). Tissue from all 10 individuals was pooled and immediately placed in TRIzol (ThermoFisher Scientific). RNA was isolated from each of the three tissue samples using a protocol slightly different from the protocol provided for TRIzol extractions by ThermoFisher Scientific. Our adjusted protocol is as follows. First, tissue was lysed and homogenized in 1 ml TRIzol using an RNase-free pestle and Eppendorf tube. Samples were centrifuged at $11,600 \times g$ for 10 min at 4 °C. The supernatant was transferred to a fresh tube and kept at room temperature for 5 min to allow the nucleoprotein complex to dissociate. In order to separate the mixture into a lower red phenol-chloroform, an interphase, and an upper RNA-containing aqueous phase, 250 μ l chloroform was added to each sample; and the tubes were shaken by hand for 15 s. Samples were kept at room temperature for 3 min, then centrifuged at $11,600 \times g$ for 15 min at 4 °C. The entire volume of the colorless top layer of supernatant (which contains the extracted RNA) was transferred to a fresh tube. After adding 500 μ l of isopropanol to the supernatant to precipitate the RNA, the solutions were incubated at room temperature for 10 min. The tubes were then centrifuged at $11,600 \times g$ for 10 min at 4 °C, and the isopropanol was immediately aspirated from the white RNA pellet at the bottom of the tube. The RNA pellet was washed with 700 μ l ethyl alcohol in 0.1% DEPC water by gently flicking the tube. After centrifuging at $7800 \times g$ for 5 min at 4 °C, the ethanol was aspirated off the RNA pellet. The RNA pellet was left to air-dry on a heat block at 42 °C, then dissolved in 0.1% DEPC water and stored at –80 °C.

Transcriptome sequencing, assembly, and analysis

The RNA libraries were prepared, multiplexed, and sequenced by Genewiz (South Plainfield, NJ) in a 2×100 -bp

paired-end configuration on the Illumina HiSeq2500 platform (Illumina, San Diego, CA) in high-output mode. The quality of the resulting data was checked using FastQC (Andrews, 2010), and only fragments sequenced from both the forward end and the reverse end were used for further analysis. Low-quality data (Phred score < 30) were removed using Trimmomatic (Bolger *et al.*, 2014); then the remaining data were assembled by the Trinity program (Grabherr *et al.*, 2011; Haas *et al.*, 2013) and run on the University of Maryland College Park computer “Deepthought.” To analyze the data, the sequence list was converted to a searchable database using the BLAST program available on the National Center for Biotechnology Information (NCBI) website (Camacho *et al.*, 2009). Sequences similar to those of selected arthropod opsins were identified by querying the searchable database. The transcript abundance in each tissue sample was determined using RNA-Seq by Expectation Maximization (RSEM) software (Li and Dewey, 2011). To assign the recovered opsin transcripts to the UV, short-wavelength-sensitive (SWS), blue, MWS, or LWS arthropod opsin clades, the phylogenetic relationships of the recovered transcripts and several other arthropod opsin sequences (from Porter *et al.*, 2007) were reconstructed by aligning the coding region of the predicted amino acid sequences using a ClustalW alignment (Larkin *et al.*, 2007; point accepted mutation [PAM] cost matrix: gap open cost, 22.2; gap extend cost, 0.1) run on the Geneious 6.0.6 platform (Kearse *et al.*, 2012). The alignment was then used to build a maximum-likelihood phylogeny using PhyML (Guindon and Gascuel, 2003). The *Sepia officinalis* cuttlefish opsin was used as an outgroup for the arthropod opsins (see Porter *et al.*, 2007). Phylogenetic relationships among the recovered sequences and several other stomatopod opsins (Porter *et al.*, 2009, 2013) were also reconstructed to determine whether the cerebral LWS opsins could be assigned to one or more of the stomatopod LWS opsin groups A–F. Previously identified *N. oerstedii* opsins were used to represent stomatopod opsin groups whenever possible. Two *Odontodactylus scyllarus* opsins (Porter *et al.*, 2009) were used to represent stomatopod LWS group E opsins, and an opsin transcript recovered from *Pseudosquilla ciliata* (Porter *et al.*, 2013) was used to represent stomatopod MWS opsins (there are no known *N. oerstedii* retinal opsins in either of these groups). Again, *S. officinalis* cuttlefish opsin was used as an outgroup for the reconstruction. References and NCBI GenBank accession numbers for opsins used to resolve phylogenetic relationships are available in Tables A1 and A2. Conserved regions of predicted amino acid sequences between NoT20086 and *N. oerstedii* group F LWS opsin No14 were identified using a Multiple Sequence Comparison by Log-Expectation (MUSCLE; Edgar, 2004) pairwise alignment run on the Geneious 6.0.6 platform, and conserved regions of predicted amino acid sequences among NoT27696g2 and *N. oerstedii* group B LWS opsins were identified using a MUSCLE multiple alignment.

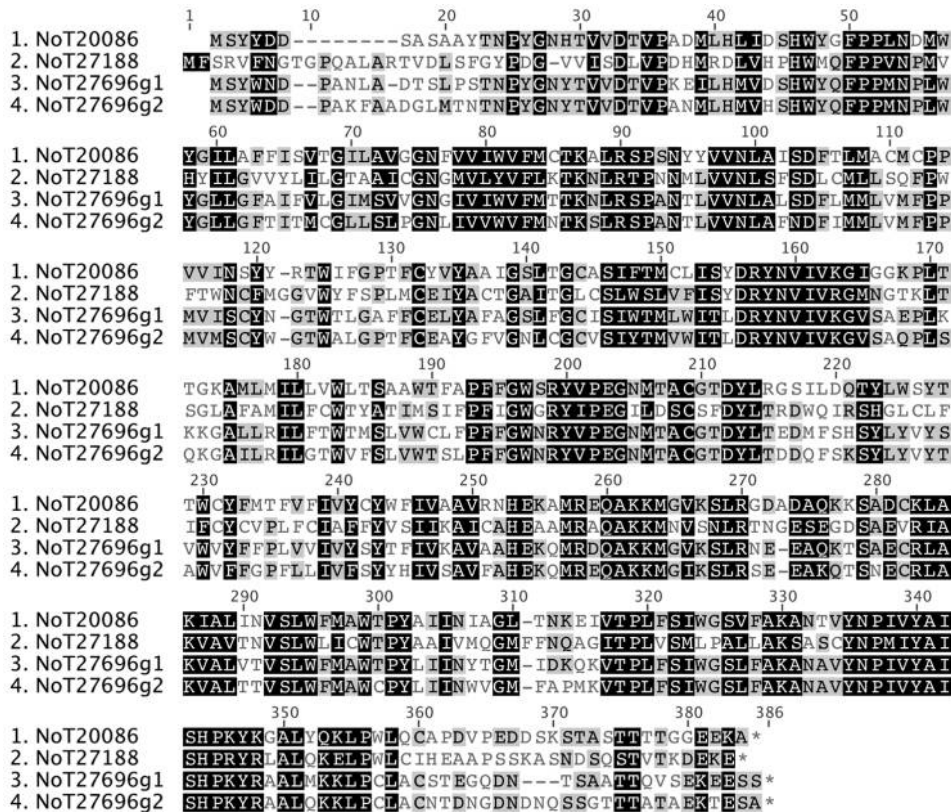


Figure 2. The predicted amino acid sequences of the four full-length opsins identified in the transcriptomic analyses. Conserved regions are indicated by a black or gray background.

Transcriptomes were also queried with the amino acid sequences of several putative arthropod phototransduction proteins that together are hypothesized to form the machinery required for typical G_q -mediated invertebrate phototransduction. Porter *et al.* (2013) identified several mRNAs encoding proteins commonly involved in G_q -mediated phototransduction in retinal transcriptomes from the stomatopods *Hemisquilla californiensis* and *Pseudosquilla ciliata*. Based on the transcripts identified by Porter *et al.* (2013), 12 genes of interest were investigated, including those encoding arrestins (*Arr1* and *Arr2*), diacylglycerol kinase (*DGK*), G-protein-coupled receptor (GPCR) kinases 1 and 2 (*Gprk1* and *Gprk2*), a G_q protein alpha subunit ($G\alpha_q$), a G protein beta subunit ($G\beta$), a G protein gamma subunit ($G\gamma$), protein kinase C (*PKC*), phospholipase C (*PLC*), phosphatidylinositol transfer protein (*rdgB*), and transient receptor potential (*TRP*) channel. Once candidate transcripts were identified, the coding region of the predicted amino acid sequence was aligned with orthologs of the respective gene using a ClustalW alignment (PAM cost matrix: gap open cost, 22.2; gap extend cost, 0.1) run on the Geneious 6.0.6 platform, and a maximum-likelihood phylogeny was built using PhyML (Guindon and Gascuel, 2003). Gene orthologs were selected by identifying the appropriate *Drosophila* gene on Ensembl Metazoa (Yates *et al.*, 2016) and downloading all orthologous invertebrate sequences, including those identified

in *Drosophila*. Scripts used to trim low-quality transcripts, assemble the transcripts, and estimate transcript abundance are reported in Table A3.

Results

Transcriptome analysis

After low-quality RNA-Seq data were removed using Trimmomatic (Bolger *et al.*, 2014), we retained 46–48 million reads,

Table 1

Fragments per kilobase per million reads (FPKM) values reveal that opsin transcripts are abundant in the cerebral ganglion (CG) and are lowly expressed or absent in the subesophageal ganglion (SEG), thoracic ganglia (T7–9), and abdominal ganglia (A1–6)

| Opsin sequence | CG | SEG | T7–9, A1–6 |
|---------------------|--------|------|------------|
| NoTranscript20086 | 95.02 | 0.07 | 0 |
| NoTranscript27188 | 36.66 | 0.05 | 0.02 |
| NoTranscript27696g1 | 20.75 | 0 | 0 |
| NoTranscript27696g2 | 121.74 | 0.05 | 0 |

FPKM values were determined from the transcriptome RNA-Seq by Expectation Maximization (RSEM) analysis (Li and Dewey, 2011), as described in the text.

representing 91.7%–92.4% per tissue sample. Four candidate transcripts were identified in our database based upon similarity to opsin sequences: NoT20086, NoT27188, NoT27696g1, and NoT27696g2 (Fig. 2). The accession numbers for the four opsin sequences are MG020521 (NoLWS1), MG020522 (NoMWS1), MG020523 (NoLWS2), and MG020524 (NoLWS3), respectively. In the SEG and T7–9/A1–6 transcriptomes, the fragments per kilobase per million reads (FPKM) values for all 4 sequences were low, ranging from 0 to 0.07 (Table 1). Conversely, the FPKM values for all 4 opsins were much higher in the CG transcriptome, ranging from 20.75 to 121.74 (Table 1). These data suggest that the four transcripts were al-

most exclusively expressed in the CG, with little or no expression in the SEG and the remainder of the nerve cord.

Opsins

The phylogenetic relationships of our four recovered sequences within arthropod opsins were reconstructed using the maximum-likelihood method on predicted amino acid sequences. Three of the derived amino acid sequences fall within the crustacean LWS clade, and one sequence falls within the crustacean MWS clade (Fig. 3). Then, a maximum-likelihood

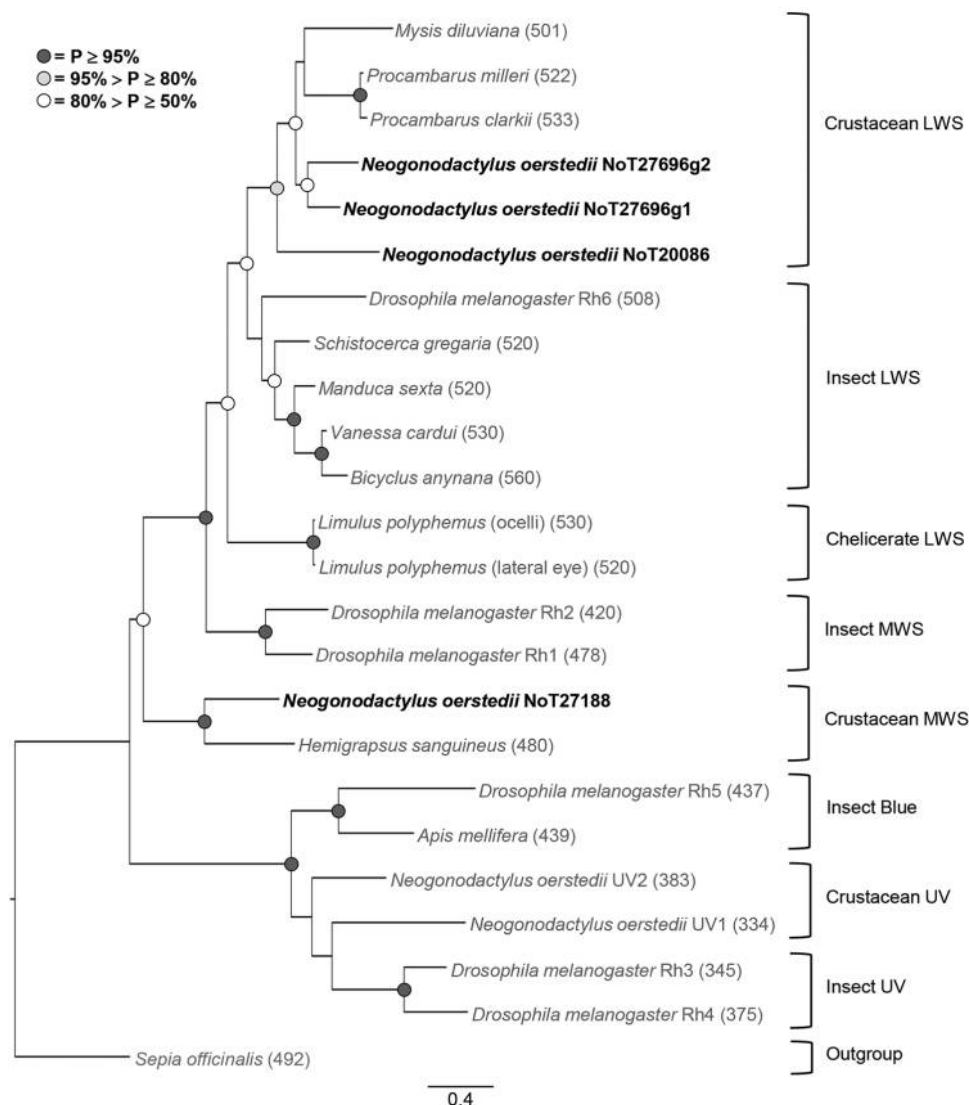


Figure 3. Maximum-likelihood phylogeny of invertebrate opsins based on predicted amino acid sequences. Opsins recovered by our transcriptomic analyses are shown in bold. T27696g1, T27696g2, and T20086 group with crustacean long-wavelength-sensitive (LWS) opsins, and opsin T27188 groups with crustacean middle-wavelength-sensitive (MWS) opsins. The known λ_{\max} values of opsins are noted in parentheses (data previously compiled by Porter *et al.*, 2007). The *Sepia officinalis* cuttlefish opsin (gray) was used as the outgroup. Dark gray circles on the nodes represent bootstrap confidences of $P \geq 95\%$, the light gray circle represents a bootstrap confidence of $95\% > P \geq 80\%$, and white circles represent bootstrap confidences of $80\% > P \geq 50\%$. UV, ultraviolet sensitive.

phylogeny was produced using the predicted amino acid sequences of our four recovered sequences and those of several stomatopod opsins. Porter *et al.* (2009) found that stomatopod opsin sequences form six monophyletic clusters, called LWS groups A–F. Thus far, LWS groups A–D contain only stomatopod opsins (Porter *et al.*, 2009). As described below, we found that the three crustacean LWS opsins expressed in the stomatopod CG are members of crustacean LWS opsin groups B, E, and F (Fig. 4).

Recovered crustacean LWS opsin transcripts T20086, T27696g1, and T27696g2 belong to stomatopod LWS opsin groups F, E, and B, respectively (Fig. 4). Two of these opsins, T20086 and T27696g2, are nearly identical to opsins previ-

ously identified in *N. oerstedii* eyes. The recovered LWS group F sequence, T20086, is similar to *N. oerstedii* opsin No14 (Fig. 4), differing at only amino acid positions 83, 205, and 206 (Fig. 5). The predicted amino acid sequences of T20086 and No14 may also differ from each other at positions 204 and 207 through 211, for which the No14 sequence is unknown (Fig. 5). Notably, the 2 sequences are identical in the 3' untranslated region (UTR) following the stop codon at position 373 (Fig. 5). The recovered LWS group B sequence, T27696g2, differs from the group B stomatopod opsins at only the 3' end of the amino acid sequence, which is highly conserved in the group B LWS opsins found in *N. oerstedii* eyes: No3, No4, No5, No10, and No11 (Fig. 6).

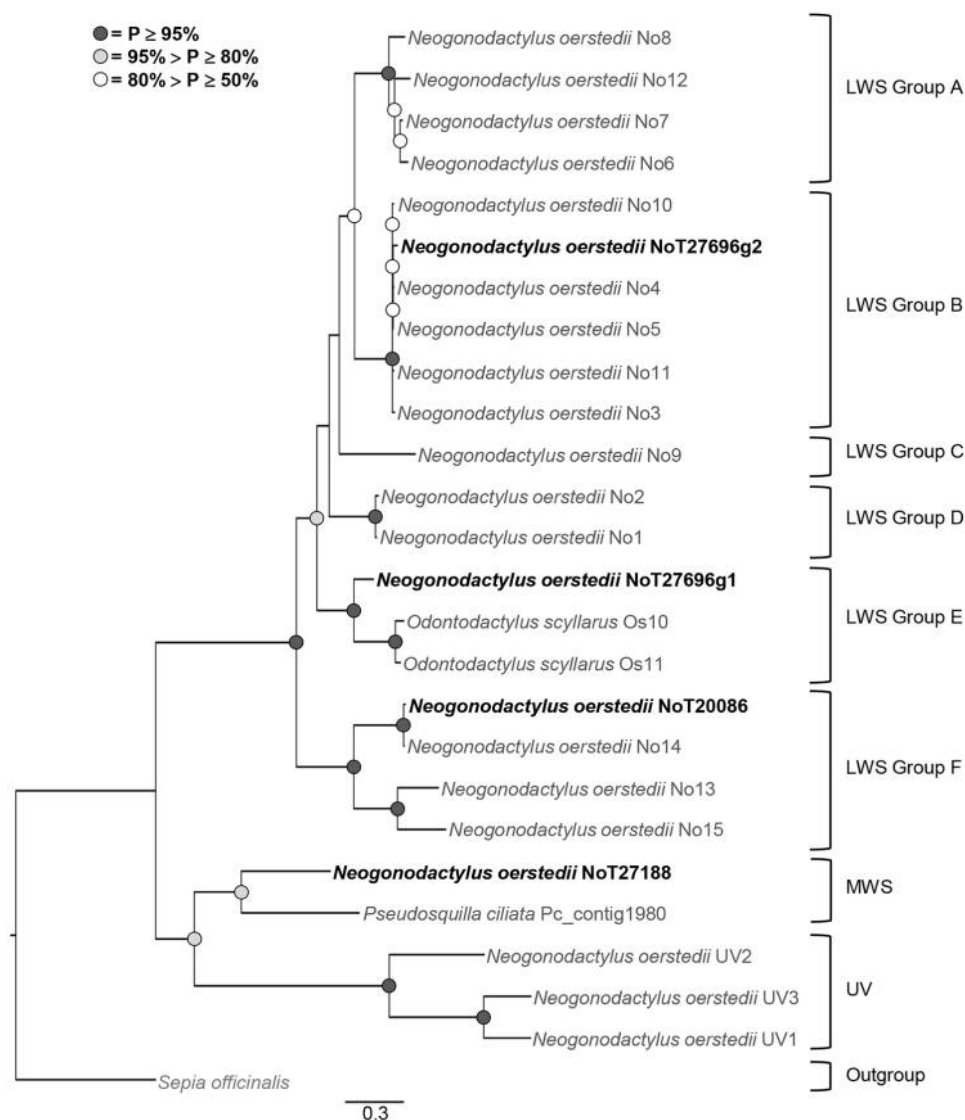


Figure 4. Maximum-likelihood phylogeny of stomatopod opsins based on predicted amino acid sequences. Opsins recovered by our transcriptome analyses are shown in bold. Opsins T27696g2, T27696g1, and T20086 group with group B, E, and F long-wavelength-sensitive (LWS) opsins, respectively. *Sepia officinalis* fly Rhodopsin 5 (gray) was used as the outgroup. Dark gray circles on the nodes represent bootstrap confidences of $P \geq 95\%$, light gray circles represent bootstrap confidences of $95\% > P \geq 80\%$, and white circles represent bootstrap confidences of $80\% > P \geq 50\%$. MWS opsins, middle-wavelength sensitive; UV, ultraviolet sensitive.

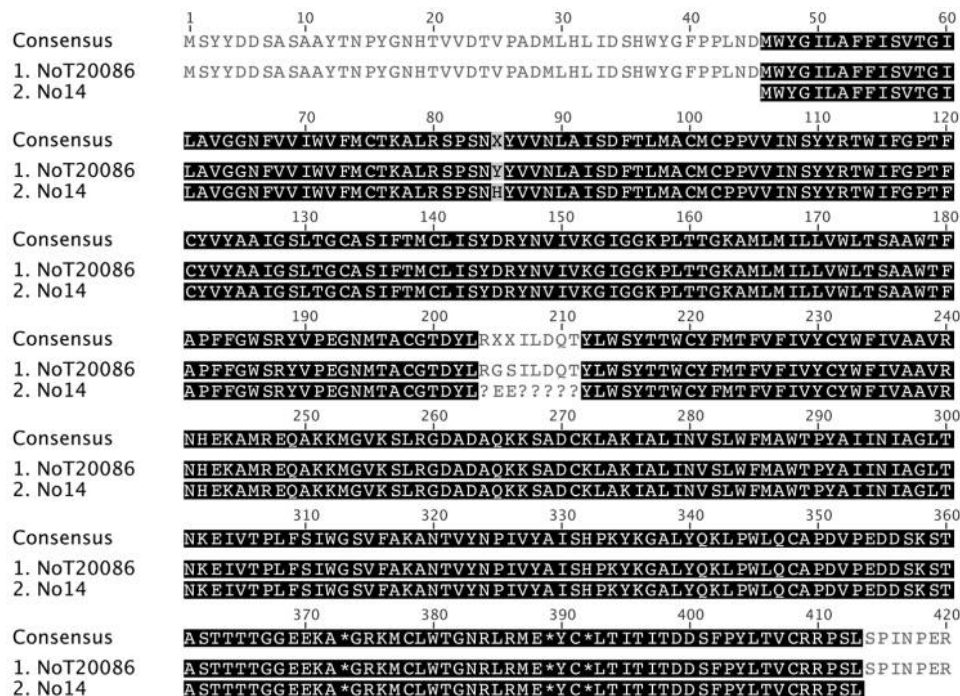


Figure 5. Multiple Sequence Comparison by Log-Expectation (MUSCLE) alignment of the predicted amino acid sequences for the group F long-wavelength-sensitive (LWS) cerebral transcript, NoT20086, and published *Neogonodactylus oerstedii* group F LWS opsin, No14. The 2 sequences differ at positions 85, 205, and 206. The sequences may differ from each other at position 204 and at positions between 207 and 211, for which the visual No14 sequence is unknown. Notably, the 2 sequences are identical in the untranslated region following the stop codon at position 373. Conserved regions are indicated by a black or gray background, and asterisks indicate stop codons.

Phototransduction components

We searched our transcriptomes for several genes that are required for a G_q -mediated phototransduction pathway. We identified putative sequences for all 12 genes of interest (Table 2) and confirmed their orthologies using maximum-likelihood phylogenies (data not shown).

Discussion

Our data suggest that stomatopods possess opsin-based extraocular photoreceptors that may utilize a G_q -mediated phototransduction pathway. Although the transcriptome analyses reveal that opsin transcript abundance is low in noncerebral nervous tissue, replicates would be required to confirm these results. The 46–48 million reads per RNA sample isolated from the nervous tissue from 10 individuals indicate that there are opsin transcripts expressed in the CG. Our phylogeny of transcriptome sequences and visual stomatopod sequences places the crustacean MWS opsin, NoT27188, in a group with the *Pseudosquilla ciliata* MWS opsin. Crustacean MWS opsins have been identified in the eyes of three stomatopod species—*Hemisquilla californiensis*, *P. ciliata*, and *Neogonodactylus oerstedii* (Bok, 2013; Porter *et al.*, 2013)—and MWS opsins have been found in several other arthropods (*e.g.*, the bioluminescent ostracod *Vargula hilgendorffii*, the copepod *Calanus fin-*

marchicus, and the springtail *Sminthurus viridis*) (reviewed by Henze and Oakley, 2015). The cerebral MWS opsin that we identified in *N. oerstedii* is unique compared to previously identified visual MWS opsins.

The three identified LWS opsin transcripts are assigned to stomatopod LWS groups B, E, and F. The group E cerebral opsin differs from previously published stomatopod group E opsins. Group E opsins have been detected in the eyes of the stomatopod species *P. ciliata*, *Odontodactylus scyllarus*, and *Squilla empusa* (Porter *et al.*, 2009, 2013) and have recently been identified in *N. oerstedii* eyes (M. Porter, University of Hawaii at Manoa, pers. comm.). The group B and F cerebral opsins are nearly identical to published stomatopod opsins. The group B cerebral transcript has a predicted amino acid sequence identical to one of the group B opsins in *N. oerstedii* eyes with the exception of the 3' end. The 3' end is highly conserved among other retinal group B opsins, so this disparity between the 3' ends of retinally and extraretinally expressed opsins may reflect differences between their visual and nonvisual functions. The predicted amino acid sequence of the group F cerebral opsin was nearly identical to published *N. oerstedii* opsin No14 in both the coding and 3' untranslated regions, suggesting that the same gene may be expressed in retinal and extraretinal tissues. Group F opsins represent the most basal group of stomatopod opsins (Porter *et al.*, 2009).

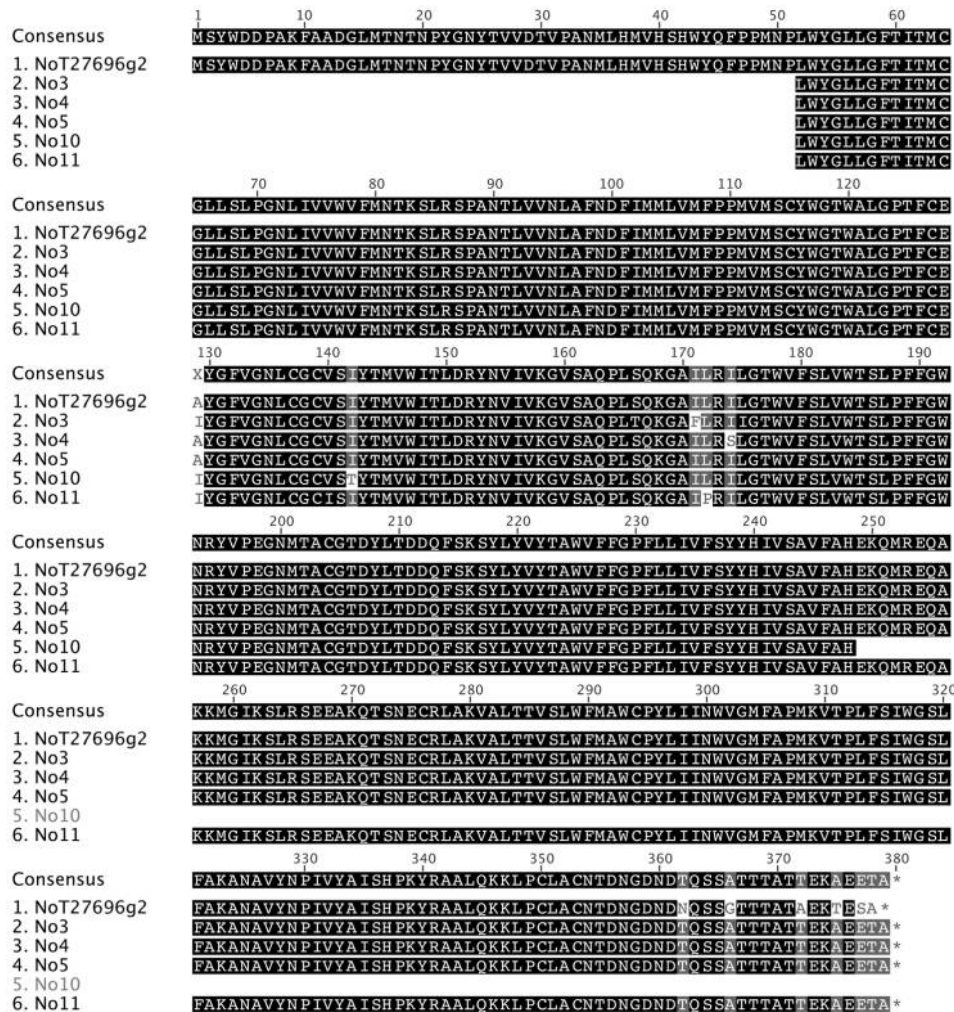


Figure 6. Multiple Sequence Comparison by Log-Expectation (MUSCLE) alignment of the predicted amino acid sequences for the group B long-wavelength-sensitive (LWS) cerebral transcript, NoT27696g2, and published *Neogonodactylus oerstedii* group B LWS opsins, No3, No4, No5, No10, and No11. Starting at amino acid position 52, where the predicted amino acid sequences begin for the available group B LWS opsin sequences, the assigned amino acid for the coding region of our transcript, NoT27696g2, is identical to at least 1 opsin sequence until position 375 onward at the 3' end. Conserved regions are indicated by a black or gray background, and asterisks indicate stop codons.

Our data raise new questions concerning the evolutionary history of stomatopod opsins. It is thought that the extremely similar visual stomatopod opsins indicate recent gene duplication events. Given the similarities among *N. oerstedii* opsins expressed in the eyes and extraocular tissues, it is possible that a subset of crustacean opsins used in vision was co-opted for nonvisual tasks in the CG. Conversely, the four extraretinally expressed opsins may represent the subset of crustacean opsins that gave rise to the other mantis shrimp opsins through gene duplication. Our data do not necessarily support one evolutionary progression over the other, but the question of which came first—opsins used for visual tasks or those used for nonvisual tasks—is one that certainly warrants further study in stomatopods, which are well known for their eyes that include a diversity of visual pigments.

Although we identify several transcriptomic sequences that putatively encode key components for a typical G_q -mediated phototransduction pathway, more data are required to confirm that the recovered opsin transcripts produce functional visual pigments of different spectral sensitivities or that they use a G_q -mediated phototransduction pathway. We also do not know whether the stomatopod extraretinal photopigments may be involved in circadian photoentrainment, modulation of neural processing, or some other function. Since two or more spectrally distinct visual pigments are required for color discrimination, it is plausible that the four opsins expressed in the CG of *N. oerstedii* confer color-discriminating abilities. Although the opsin groups were determined only by amino acid sequence similarities, crustacean MWS opsins typically produce visual pigments with wavelengths of maximum absorbance (λ_{\max})

Table 2

There are at least 12 putative G_q -mediated phototransduction components expressed in the central nervous systems of *Neogonodactylus oerstedii*

| Gene of interest | Accession number for query sequence | Reference | Transcript | Accession number for predicted transcript sequence |
|------------------|-------------------------------------|---------------------------------|--------------|--|
| <i>Arr1</i> | P15372 | Smith <i>et al.</i> , 1990 | T16832_g1_i1 | MG020525 |
| <i>Arr2</i> | P19107 | Yamada <i>et al.</i> , 1990 | T21375_g1_i1 | MG020526 |
| <i>Gprk1</i> | P32865 | Cassill <i>et al.</i> , 1991 | T25623_g1_i1 | MG020527 |
| <i>Gprk2</i> | AAF57152 | Misra <i>et al.</i> , 2002 | T20751_g2_i1 | MG020528 |
| $G\alpha_q$ | P91950 | McClintock <i>et al.</i> , 1997 | T33701_g1_i1 | MG020529 |
| $G\beta$ | NP525090 | Hoskins <i>et al.</i> , 2007 | T28638_g1_i3 | MG020530 |
| $G\gamma$ | XM002408603 | Nene, 2008 | T1758_g1_i1 | MG020531 |
| <i>PKC</i> | NP476682 | Hoskins <i>et al.</i> , 2007 | T19726_g1_i1 | MG020532 |
| <i>PLC</i> | P25455 | Shortridge <i>et al.</i> , 1991 | T24765_g1_i2 | MG020533 |
| <i>rdgB</i> | NP727733 | Hoskins <i>et al.</i> , 2007 | T24446_g1_i3 | MG020534 |
| <i>TRP</i> | P48994 | Phillips <i>et al.</i> , 1992 | T27206_g1_i1 | MG020535 |
| <i>DGK</i> | BAA04135 | Masai <i>et al.</i> , 1993 | T26120_g2_i2 | MG020536 |

Arr1 and *Arr2*, visual arrestins; *DGK*, diacylglycerol kinase; *Gprk1* and *Gprk2*, G-protein-coupled receptor kinases 1 and 2; $G\alpha_q$, G protein alpha subunit; $G\beta$, G protein beta subunit; $G\gamma$, G protein gamma subunit; *PKC*, protein kinase C; *PLC*, phospholipase C; *rdgB*, phosphatidylinositol transfer protein; *TRP*, transient receptor potential channels.

values between 400 and 490 nm, while LWS opsins produce visual pigments absorbing at peak wavelengths longer than 490 nm. If the four extraretinal stomatopod opsins do produce visual pigments with varying λ_{\max} values, it is possible that the putative brain photoreceptors could be used not only to detect the presence or absence of light but also to signal the spectral content and the spectral changes that occur with depth or at dawn and dusk. Alternatively, each spectrally unique opsin may serve wavelength-specific functions. Although these opsin transcripts could form visual pigments maximally sensitive to different wavelengths of light, more data would be required to reveal the spectral sensitivities of cerebral photoreceptors that express these opsins.

The cerebral opsins may also contribute to multisensory integration, providing a variety of sensory systems whenever the animal's head is protruding from its burrow. The CG of stomatopods includes several brain regions involved in the interpretation of sensory stimuli, including the olfactory lobes and antennal lobes. Another CG region called the central complex is similar in function to vertebrate basal ganglia, being crucial for motor function, memory formation, and attention (Strausfeld and Hirth, 2013). Light sensitivity in any of these brain regions may alter multisensory integration during critical interactions with the environment when the animals are vulnerable outside their burrows. For example, the extraocular photoreceptors in the tails of *Limulus* horseshoe crabs affect both circadian photoentrainment and the sensitivity of the visual system; the circadian clock can be entrained by illumination of the eyes or extraocular photoreceptors in the tail to affect visual system sensitivity *via* clock-driven efferent neurons originating in the brain (reviewed by Battelle, 2006, 2013). Future work in stomatopods could show that phototransduction pathways exist in opsin-expressing neurons and could investigate the function of the cerebral photoreceptors. Stomatopods may prove to be important model organisms in

the field of extraretinal photoreception given their exceptionally complex ocular vision, diversity of visual pigments, and variable light environments.

Literature Cited

- Andrews, S. 2010. FastQC: a quality control tool for high throughput sequence data. [Online]. Babraham Institute. Available: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> [2015, July 15].
- Battelle, B.-A. 2006. The eyes of *Limulus polyphemus* (Xiphosura, Chelicerata) and their afferent and efferent projections. *Arthropod Struct. Dev.* 35: 261–274.
- Battelle, B.-A. 2013. What the clock tells the eye: lessons from an ancient arthropod. *Integr. Comp. Biol.* 53: 144–153.
- Battelle, B.-A. 2016. Extraocular, non-visual, and simple photoreceptors. *Integr. Comp. Biol.* 56: 809–819.
- Battelle, B.-A., J. F. Ryan, K. E. Kempler, S. R. Saraf, C. E. Marten, W. C. Warren, P. Minx, M. J. Montague, P. J. Green, S. A. Schmidt *et al.* 2016. Opsin repertoire and expression patterns in horseshoe crabs: evidence from the genome of *Limulus polyphemus* (Arthropoda: Chelicerata). *Genome Biol. Evol.* 8: 1571–1589.
- Bertolucci, C., and A. Foa. 2004. Extraocular photoreception and circadian entrainment in nonmammalian vertebrates. *Chronobiol. Int.* 21: 501–519.
- Bok, M. J. 2013. The physiological, ecological, and evolutionary basis of polychromatic ultraviolet sensitivity in stomatopod crustaceans. Ph.D. dissertation, University of Maryland, Baltimore.
- Bok, M. J., M. L. Porter, A. R. Place, and T. W. Cronin. 2014. Biological sunscreens tune polychromatic ultraviolet vision in mantis shrimp. *Curr. Biol.* 24: 1636–1642.
- Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120.
- Briscoe, A. D., G. D. Bernard, A. S. Szeto, L. M. Nagy, and R. H. White. 2003. Not all butterfly eyes are created equal: rhodopsin absorption spectra, molecular identification, and localization of ultraviolet-, blue-, and green-sensitive rhodopsin-encoding mRNAs in the retina of *Vanessa cardui*. *J. Comp. Neurol.* 458: 334–349.
- Brown, P. K., and P. S. Brown. 1958. Visual pigments of the octopus and cuttlefish. *Nature* 182: 1288–1290.
- Camacho, C., G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, and T. L. Madden. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10: 421.

- Cassill, J. A., M. Whitney, C. A. Joazeiro, A. Becker, and C. S. Zuker. 1991. Isolation of *Drosophila* genes encoding G protein-coupled receptor kinases. *Proc. Natl. Acad. Sci. U.S.A.* **88**: 11067–11070.
- Crandall, K. A., and T. W. Cronin. 1997. The molecular evolution of visual pigments of freshwater crayfishes (Decapoda: Cambaridae). *J. Mol. Evol.* **45**: 524–534.
- Cronin, T. W., and T. H. Goldsmith. 1982. Photosensitivity spectrum of crayfish rhodopsin measured using fluorescence of metarhodopsin. *J. Gen. Physiol.* **79**: 313–332.
- Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**: 1792–1797.
- Edwards, D. H. 1984. Crayfish extraretinal photoreception. I. Behavioral and motoneuronal responses to abdominal illumination. *J. Exp. Biol.* **109**: 291–306.
- Feiler, R., W. A. Harris, K. Kirschfeld, C. Wehrhahn, and C. S. Zuker. 1988. Targeted misexpression of a *Drosophila* opsin gene leads to altered visual function. *Nature* **333**: 737–741.
- Feiler, R., R. Bjornson, K. Kirschfeld, D. Mismar, G. M. Rubin, D. P. Smith, M. Socolich, and C. S. Zuker. 1992. Ectopic expression of ultraviolet-rhodopsins in the blue photoreceptor cells of *Drosophila*: visual physiology and photochemistry of transgenic animals. *J. Neurosci.* **12**: 3862–3868.
- Gartner, W., and P. Towner. 1995. Invertebrate visual pigments. *Photochem. Photobiol.* **62**: 1–16.
- Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng *et al.* 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**: 644–652.
- Guindon, S., and O. Gascuel. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**: 696–704.
- Haas, B. J., A. Papanicolaou, M. Yassour, M. Grabherr, P. D. Blood, J. Bowden, M. B. Couger, D. Eccles, B. Li, M. Lieber *et al.* 2013. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**: 1494–1512.
- Henze, M. J., and T. H. Oakley. 2015. The dynamic evolutionary history of pancrustacean eyes and opsins. *Integr. Comp. Biol.* **55**: 830–842.
- Hoskins, R. A., J. W. Carlson, C. Kennedy, D. Acevedo, M. Evans-Holm, E. Frise, K. H. Wan, S. Park, M. Mendez-Lago, F. Rossi *et al.* 2007. Sequence finishing and mapping of *Drosophila melanogaster* heterochromatin. *Science* **316**: 1625–1628.
- Hubbard, R., and G. Wald. 1960. Visual pigment of the horseshoe crab, *Limulus polyphemus*. *Nature* **186**: 212–215.
- Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran *et al.* 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**: 1647–1649.
- Kingston, A. C. N., and T. W. Cronin. 2015. Short- and long-wavelength-sensitive opsins are involved in photoreception both in the retina and throughout the central nervous system of crayfish. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **201**: 1137–1145.
- Kingston, A. C. N., and T. W. Cronin. 2016. Diverse distributions of extraocular opsins in crustaceans, cephalopods, and fish. *Integr. Comp. Biol.* **56**: 820–833.
- Kingston, A. C., A. M. Kuzirian, R. T. Hanlon, and T. W. Cronin. 2015. Visual phototransduction components in cephalopod chromatophores suggest dermal photoreception. *J. Exp. Biol.* **218**: 1596–1602.
- Lampel, J., A. D. Briscoe, and L. T. Wasserthal. 2005. Expression of UV-, blue-, long-wavelength-sensitive opsins and melatonin in extra-retinal photoreceptors of the optic lobes of hawkmoths. *Cell Tissue Res.* **321**: 443–458.
- Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez *et al.* 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**: 2947–2948.
- Li, B., and C. N. Dewey. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**: 323.
- Masai, I., A. Okazaki, T. Hosoya, and Y. Hotta. 1993. *Drosophila* retinal degeneration A gene encodes an eye-specific diacylglycerol kinase with cysteine-rich zinc-finger motifs and ankyrin repeats. *Proc. Natl. Acad. Sci. U.S.A.* **90**: 11157–11161.
- McClintock, T. S., F. Xu, J. Quintero, A. M. Gress, and T. M. Landers. 1997. Molecular cloning of a lobster $G\alpha_q$ protein expressed in neurons of olfactory organ and brain. *J. Neurochem.* **68**: 2248–2254.
- Misra, S., M. A. Crosby, C. J. Mungall, B. B. Matthews, K. S. Campbell, P. Hradecky, Y. Huang, J. S. Kaminker, G. H. Millburn, S. E. Prochnik *et al.* 2002. Annotation of the *Drosophila melanogaster* euchromatic genome: a systematic review. *Genome Biol.* **3**: research0083.1.
- Nene, V. 2008. *Ixodes scapularis* G-protein gamma subunit, putative, mRNA. [Online]. GenBank. Available: https://www.ncbi.nlm.nih.gov/nuccore/XM_002408603 [2017, October 12].
- Nolte, J., and J. E. Brown. 1972. Electrophysiological properties of cells in the median ocellus of *Limulus*. *J. Gen. Physiol.* **59**: 167–185.
- Phillips, A. M., A. Bull, and L. E. Kelly. 1992. Identification of a *Drosophila* gene encoding a calmodulin-binding protein with homology to the trp phototransduction gene. *Neuron* **8**: 631–642.
- Porter, M. L., T. W. Cronin, D. A. McClellan, and K. A. Crandall. 2007. Molecular characterization of crustacean visual pigments and the evolution of pancrustacean opsins. *Mol. Biol. Evol.* **24**: 253–268.
- Porter, M. L., M. J. Bok, P. R. Robinson, and T. W. Cronin. 2009. Molecular diversity of visual pigments in Stomatopoda (Crustacea). *Vis. Neurosci.* **26**: 255–265.
- Porter, M. L., D. I. Speiser, A. K. Zaharoff, R. L. Caldwell, T. W. Cronin, and T. H. Oakley. 2013. The evolution of complexity in the visual systems of stomatopods: insights from transcriptomics. *Integr. Comp. Biol.* **53**: 39–49.
- Roenneberg, T., and R. G. Foster. 1997. Twilight times: light and the circadian system. *Photochem. Photobiol.* **66**: 549–561.
- Sakamoto, K., O. Hisatomi, F. Tokunaga, and E. Eguchi. 1996. Two opsins from the compound eye of the crab *Hemigrapsus sanguineus*. *J. Exp. Biol.* **199**: 441–450.
- Salcedo, E., A. Huber, S. Henrich, L. V. Chadwell, W. H. Chou, R. Paulsen, and S. G. Britt. 1999. Blue- and green-absorbing visual pigments of *Drosophila*: ectopic expression and physiological characterization of the R8 photoreceptor cell-specific Rh5 and Rh6 rhodopsins. *J. Neurosci.* **19**: 10716–10726.
- Shortridge, R. D., J. Yoon, C. R. Lending, B. T. Bloomquist, M. H. Perdew, and W. L. Pak. 1991. A *Drosophila* phospholipase C gene that is expressed in the central nervous system. *J. Biol. Chem.* **266**: 12474–12480.
- Smith, D. P., B. H. Shieh, and C. S. Zuker. 1990. Isolation and structure of an arrestin gene from *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* **87**: 1003–1007.
- Smith, W. C., D. M. Ayers, M. P. Popp, and P. A. Hargrave. 1997. Short wavelength-sensitive opsins from the Saharan silver and carpenter ants. *Invertebr. Neurosci.* **3**: 49–56.
- Strausfeld, N. J., and F. Hirth. 2013. Deep homology of arthropod central complex and vertebrate basal ganglia. *Science* **340**: 157–161.
- Townson, S. M., B. S. W. Chang, E. Salcedo, L. V. Chadwell, N. E. Pierce, and S. G. Britt. 1998. Honeybee blue- and ultraviolet-sensitive opsins: cloning, heterologous expression in *Drosophila*, and physiological characterization. *J. Neurosci.* **18**: 2412–2422.
- Vanhoutte, K. J. A., B. J. L. Eggen, J. J. M. Janssen, and D. G. Stavenga. 2002. Opsin cDNA sequences of a UV and green rhodopsin of the satyrine butterfly *Bicyclus anynana*. *Insect Biochem. Mol. Biol.* **32**: 1383–1390.
- White, R. H., P. K. Brown, A. K. Hurley, and R. R. Bennett. 1983. Rhodopsins, retinula cell ultrastructure, and receptor potentials in the developing pupal eye of the moth *Manduca sexta*. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **150**: 153–163.

- Wilkins, L. A., and J. L. Larimer. 1976. Photosensitivity in the sixth abdominal ganglion of decapod crustaceans: a comparative study. *J. Comp. Physiol.* **106**: 69–75.
- Yamada, T., Y. Takeuchi, N. Komori, H. Kobayashi, Y. Sakai, Y. Hotta, and H. Matsumoto. 1990. A 49-kilodalton phosphoprotein in the *Drosophila* photoreceptor is an arrestin homolog. *Science* **248**: 483–486.
- Yates, A., W. Akanni, M. R. Amode, D. Barrell, K. Billis, D. Carvalho-Silva, C. Cummins, P. Clapham, S. Fitzgerald, L. Gil *et al.* 2016. Ensembl 2016. *Nucleic Acids Res.* **44**: D710–D716.
- Zeiger, J., and T. H. Goldsmith. 1994. Behavior of crayfish rhodopsin and metarhodopsin in digitonin: The 510 and 562 nm “visual pigments” are artifacts. *Vision Res.* **34**: 2679–2688.

Appendix

Table A1

National Center for Biotechnology Information (NCBI) GenBank accession numbers and references for arthropod opsin sequences used in Figure 3

| Species | Accession number for opsin sequence | λ_{\max} (nm) | λ_{\max} reference |
|---|-------------------------------------|-----------------------|---|
| <i>Apis mellifera</i> | AF004168 | 439 | Townson <i>et al.</i> , 1998 |
| <i>Bicyclus anynana</i> | AF484249 | 560 | Vanhoutte <i>et al.</i> , 2002 |
| <i>Camponotus abdominalis</i> | AF042788 | 360 | Smith <i>et al.</i> , 1997 |
| <i>Sepia officinalis</i> | AF000947 | 492 | Brown and Brown, 1958 |
| <i>Drosophila melanogaster</i> Rh1 | AH001026 | 478 | Feiler <i>et al.</i> , 1988 |
| <i>Drosophila melanogaster</i> Rh2 | M12896 | 420 | Feiler <i>et al.</i> , 1988 |
| <i>Drosophila melanogaster</i> Rh3 | M17718 | 345 | Feiler <i>et al.</i> , 1992 |
| <i>Drosophila melanogaster</i> Rh4 | AH001040 | 375 | Feiler <i>et al.</i> , 1992 |
| <i>Drosophila melanogaster</i> Rh5 | U67905 | 437 | Salcedo <i>et al.</i> , 1999 |
| <i>Drosophila melanogaster</i> Rh6 | Z86118 | 508 | Salcedo <i>et al.</i> , 1999 |
| <i>Hemigrapsus sanguineus</i> | D50583 | 480 | Sakamoto <i>et al.</i> , 1996 |
| <i>Limulus polyphemus</i> (lateral eye) | L03781 | 520 | Hubbard and Wald, 1960 |
| <i>Limulus polyphemus</i> (ocelli) | L03782 | 530 | Nolte and Brown, 1972 |
| <i>Manduca sexta</i> | L78080 | 520 | White <i>et al.</i> , 1983 |
| <i>Mysis diluviana</i> | DQ852591 | 501 | Porter <i>et al.</i> , 2007 |
| <i>Neogonodactylus oerstedii</i> UV1 | KJ127520.1 | 334 | Bok <i>et al.</i> , 2014 |
| <i>Neogonodactylus oerstedii</i> UV2 | KJ127521.1 | 383 | Bok <i>et al.</i> , 2014 |
| <i>Procambarus clarkii</i> | S53494 | 533 | Zeiger and Goldsmith, 1994 |
| <i>Procambarus milleri</i> | AF003546 | 522 | Cronin and Goldsmith, 1982; Crandall and Cronin, 1997 |
| <i>Schistocerca gregaria</i> | X80071 | 520 | Gartner and Towner, 1995 |
| <i>Vanessa cardui</i> | AF385333 | 530 | Briscoe <i>et al.</i> , 2003 |

Porter *et al.* (2007) used these and additional data to investigate pancrustacean opsin evolution. λ_{\max} , wavelength of maximum absorbance.

Table A2

National Center for Biotechnology Information (NCBI) GenBank accession numbers and references for stomatopod opsin sequences used in Figure 4

| Species | Opsin | Accession number for opsin sequence | Reference |
|----------------------------------|-------|-------------------------------------|-----------------------------|
| <i>Neogonodactylus oerstedii</i> | No1 | GQ221725.1 | Porter <i>et al.</i> , 2009 |
| <i>Neogonodactylus oerstedii</i> | No2 | GQ221726.1 | Porter <i>et al.</i> , 2009 |
| <i>Neogonodactylus oerstedii</i> | No3 | GQ221727.1 | Porter <i>et al.</i> , 2009 |
| <i>Neogonodactylus oerstedii</i> | No4 | GQ221728.1 | Porter <i>et al.</i> , 2009 |
| <i>Neogonodactylus oerstedii</i> | No5 | GQ221729.1 | Porter <i>et al.</i> , 2009 |
| <i>Neogonodactylus oerstedii</i> | No6 | GQ221730.1 | Porter <i>et al.</i> , 2009 |
| <i>Neogonodactylus oerstedii</i> | No7 | GQ221731.1 | Porter <i>et al.</i> , 2009 |
| <i>Neogonodactylus oerstedii</i> | No8 | GQ221732.1 | Porter <i>et al.</i> , 2009 |
| <i>Neogonodactylus oerstedii</i> | No9 | GQ221733.1 | Porter <i>et al.</i> , 2009 |
| <i>Neogonodactylus oerstedii</i> | No10 | GQ221734.1 | Porter <i>et al.</i> , 2009 |
| <i>Neogonodactylus oerstedii</i> | No11 | GQ221735.1 | Porter <i>et al.</i> , 2009 |
| <i>Neogonodactylus oerstedii</i> | No12 | GQ221736.1 | Porter <i>et al.</i> , 2009 |
| <i>Neogonodactylus oerstedii</i> | No13 | GQ221737.1 | Porter <i>et al.</i> , 2009 |

Table A2 (Continued)

| Species | Opsin | Accession number for opsin sequence | Reference |
|----------------------------------|----------------|-------------------------------------|-----------------------------|
| <i>Neogonodactylus oerstedii</i> | No14 | GQ221738.1 | Porter <i>et al.</i> , 2009 |
| <i>Neogonodactylus oerstedii</i> | No15 | GQ221739.1 | Porter <i>et al.</i> , 2009 |
| <i>Neogonodactylus oerstedii</i> | NoUV1 | KJ127520.1 | Bok <i>et al.</i> , 2014 |
| <i>Neogonodactylus oerstedii</i> | NoUV2 | KJ127521.1 | Bok <i>et al.</i> , 2014 |
| <i>Neogonodactylus oerstedii</i> | NoUV3 | KJ127522.1 | Bok <i>et al.</i> , 2014 |
| <i>Odontodactylus scyllarus</i> | Os10 | GQ221749.1 | Porter <i>et al.</i> , 2009 |
| <i>Odontodactylus scyllarus</i> | Os11 | GQ221750.1 | Porter <i>et al.</i> , 2009 |
| <i>Pseudosquilla ciliata</i> | Pc_contig 1980 | NA | Porter <i>et al.</i> , 2013 |

NA, not applicable.

Table A3

Scripts used to trim low-quality transcripts, assemble the transcripts, and estimate transcript abundance from RSEM

| Purpose | Script |
|------------------|--|
| Trimming | java -jar /Users/Owner/Programs/Trimmomatic-0-4.33/trimmomatic-0.33.jar PE -phred33/Users/Owner/Desktop/Transcriptomes/Sample_NoerstediiCG/NoerstediiCG_GTG GCC_L002_R1_001.fastq.gz /Users/Owner/Desktop/Transcriptomes/Sample_NoerstediiCG/NoerstediiCG_GTG GCC_L002_R2_001.fastq.gz NoCG_trim_R1.fq.gz NoCG_trim_orphans_R1.fq.gz NoCG_trim_R2.fq.gz NoCG_trim_orphans_R2.fq.gz ILLUMINACLIP:/Users/Owner/Programs/Trimmomatic-0-4.33/adapters/Adapt.fa.txt:2:40:15 HEADCROP:5 SLIDINGWINDOW:4:20 LEADING:20 TRAILING:20MINLEN:80 |
| Trinity assembly | Trinity --seqType fq --max_memory 99G --normalize_reads \ --left /lustre/kcarleto/Stom/data/No_R1.fq \ --right /lustre/kcarleto/Stom/data/No_R2.fq \ --min_kmer_cov 2 --CPU 20 --output \$OUTDIR |
| RSEM | align_and_estimate_abundance.pl --transcripts /lustre/kcarleto/Stom/Noer.5851426/Trinity.fasta \ --thread_count 30 \ --seqType fq \ --left /lustre/kcarleto/Stom/data/NoCG_trim_R1.fq.gz \ --right /lustre/kcarleto/Stom/data/NoCG_trim_R2.fq.gz \ --est_method RSEM \ --aln_method bowtie \ --trinity_mode \ --prep_reference \ --output_dir \$OUTDIR |

RSEM, RNA-Seq by Expectation Maximization software (Li and Dewey, 2011).