Published in final edited form as:

Mol Ecol Resour. 2019 November ; 19(6): 1447–1460. doi:10.1111/1755-0998.13062.

Diurnal variation in opsin expression and common housekeeping genes necessitates comprehensive normalization methods for quantitative real-time PCR analyses

Miranda R Yourick^{1,+}, Benjamin A Sandkam^{1,2}, William J Gammerdinger^{1,3}, Daniel Escobar-Camacho¹, Sri Pratima Nandamuri^{1,4}, Frances E Clark¹, Brendan Joyce^{1,5}, Matthew A Conte¹, Thomas D Kocher¹, Karen L Carleton¹

¹Department of Biology, University of Maryland, College Park MD

²Present address: Department of Zoology, University of British Columbia, Vancouver, Canada

³Present address: Institute of Science and Technology (IST) Austria, Am Campus 1, 3400 Klosterneuburg, Austria

⁴Present address: Department of Human Genetics, University of Utah, Salt Lake City 84112

⁵Present address: Epic Systems Corporation, Verona, WI 53593

Abstract

To determine the visual sensitivities of an organism of interest, quantitative reverse-transcription polymerase chain reaction (qRT-PCR) is often used to quantify expression of the light-sensitive opsins in the retina. While qRT-PCR is an affordable, high-throughput method for measuring expression, it comes with inherent normalization issues that affect the interpretation of results, especially as opsin expression can vary greatly based on developmental stage, light environment, or diurnal cycles. We tested for diurnal cycles of opsin expression over a period of 24 hours at 1 hour increments and examined how normalization affects a dataset with fluctuating expression levels using qRT-PCR and transcriptome data from the retinae of the cichlid *Pelmatolapia mariae*. We compared five methods of normalizing opsin expression relative to: (1) the average of three stably expressed housekeeping genes (*Ube2z*, *EF1- α* , and *β -actin*), (2) total RNA concentration, (3) *GNAT2*, (the cone-specific subunit of transducin), (4) total opsin expression, and (5) only opsins expressed in the same cone-type. Normalizing by proportion of cone-type produced the least variation and would be best for removing time-of-day variation. In contrast, normalizing by housekeeping genes produced the highest daily variation in expression and demonstrated that the peak of cone opsin expression was in the late afternoon. A weighted correlation network analysis showed that the expression of different cone opsins follow a very similar daily cycle. With the

⁺ corresponding author Address of correspondence: Department of Biology, University of Maryland, College Park, Maryland, 20742, USA.

Author Contributions

M.R.Y, K.L.C, B.S, and W.J.G conceptualized and designed the experiment. M.R.Y, K.L.C, W.J.G, S.P.N, F.E.C, and D.E-C participated in data collection. M.R.Y, K.L.C, B.J, and M.A.C performed analyses. M.R.Y drafted the manuscript. All authors critically reviewed the article.

Data Accessibility Statement

Raw reads from transcriptomes have been deposited in the Sequence Read Archive (Accession: PRJNA545728).

All authors declare no conflicts of interest.

knowledge of how these normalization methods affect opsin expression data, we make recommendations for designing sampling approaches and quantification methods based upon the scientific question being examined.

Keywords

vision; cichlid; circadian; qRT-PCR; color vision; visual ecology

Introduction

For many organisms, visual systems are essential for tasks that directly impact survival and fitness. Whether it be for tracking down prey, avoiding predators, or interacting with conspecifics, strong selection shapes animal visual systems. In general, the visual systems of vertebrates are tuned to be sensitive to the most abundant wavelengths in the local light environment (Levine & Machnichol, 1979; Loew & Lythgoe, 1978; Lythgoe, 1984). The strong selection placed on the visual system has made visual tuning one of the classic models for studies of ecological and evolutionary questions (Yokoyama, 2008).

While traditional methods (eg. microspectrophotometry, and electrophysiology) allow for the direct measure of visual tuning, these techniques have significant drawbacks including being very low throughput and requiring highly controlled laboratory conditions. To overcome these limitations recent studies have used quantitative real-time PCR (qRT-PCR) techniques to measure the expression of the opsin genes that are primarily responsible for tuning visual systems. qRT-PCR techniques have the benefits of being relatively low cost, high throughput, repeatable, and allowing for samples to be easily stored and measured at a later time, which is particularly helpful when sampling in the field (Ginzinger, 2002). However, quantification of opsins by qRT-PCR is inherently a relative measurement. The raw data must be normalized, and the chosen normalization method could have a profound effect on the results (Bustin, 2002). As of yet there has been no consensus on how to best normalize opsin gene expression.

Opsins are G-protein coupled receptor proteins found in the outer segment of specialized retinal cells called photoreceptors. When the chromophore bound to the opsin absorbs light, it induces a conformational change in the opsin resulting in the activation of the phototransduction cascade (reviewed in Rodieck, 1998). Most vertebrates have two kinds of photoreceptors- the rods that are active during the night and under low-light conditions, and the cones which are utilized for daytime color vision. Rods and cones are packed with rhodopsin and cone opsins, respectively, which are sensitive to light of specific wavelengths (Bowmaker, 2008). Perception of color begins in the retina where light differentially stimulates cone cells that vary in sensitivity (Bowmaker & Hunt, 2006; Jacobs & Rowe, 2004; Kelber, Vorobyev, & Osorio, 2003). Because it is not practical to count the number of cone cells sensitive to different wavelengths in a retina, opsin gene expression is used as a substitute to approximate the visual sensitivity of an organism. Opsin expression measured by qRT-PCR provides a good approximation of visual sensitivity compared to microspectrophotometry, which only samples a subset of cone cells, and is more practical

than protein analysis, especially for non-model species (Carleton et al., 2008; Chinen, Hamaoka, Yamada, & Kawamura, 2003; Fuller, Carleton, Fadool, Spady, & Travis, 2004; Matsumoto, Fukamachi, Mitani, & Kawamura, 2006; Spady et al., 2006).

Photoreceptors are exposed to high light levels and high metabolic loads. As a result, the photoreceptor outer segments are replenished over time to replace opsins and other proteins. In rods, a small proportion of the outer segment is shed and replaced every day (Basinger, Hoffman, & Matthes, 1976; LaVail, 1980; Young, 1978). The rate of cone opsin replacement has not been investigated, but with the high intensity of daytime light it is possible that they are replaced at a similar or even faster rate. Interestingly, the amount of opsin expression is often not consistent throughout the day and varies in a cyclical manner. Daily variation in opsin expression has been described for several terrestrial vertebrates including chicken (Pierce et al., 1993), mouse (von Schantz, Lucas, & Foster, 1999), and toad (Korenbrodt & Fernald, 1989). However, far more is known about daily variation in opsin expression from aquatic species, primarily fishes. Opsins in the cichlid *Astatotilapia burtoni* exhibit marked diurnal variation (Halstenberg et al., 2005). In *A. burtoni*, expression of the day-light receptive cone opsins is highest around 18:00 and remains relatively high during the night. Similar patterns of cone opsin expression have been observed in zebrafish and bluefin killifish (Johnson, Stanis, & Fuller, 2013; Li et al., 2005). Korenbrot & Fernald (1989) showed that expression of the nocturnally active rhodopsin (*RH1*) peaks in the middle of the day, around 12:00, and decreases during the night. Therefore across taxa, it appears that opsins are being replenished at a time when that class of photoreceptor is less active; cone opsins tend to be highly expressed in the late afternoon into the night, while rod opsins are replenished during the day, when rods are retracted into the retinal pigment epithelium.

Not only can the amount of opsin expression change throughout the day, but also the relative composition of the total pool of opsins could change. It is possible that the cone opsins have differing diurnal cycles, especially if they are expressed in different cone cell types. While early work suggested that the cone opsins all have a similar diurnal pattern of expression (Halstenberg et al., 2005), more recent work did not show such a clear similarity (Johnson et al., 2013; Sandkam, Young, & Breden, 2015). An issue could arise when sampling at different times of the day as the relative makeup of the opsin pool could differ simply because of the diurnal variation in expression.

The particular methods used to normalize opsin expression are important because there is incredible variation in opsin expression across animals depending on the time of day (Halstenberg et al., 2005; Johnson et al., 2013; Korenbrot & Fernald, 1989; Li et al., 2005; Pierce et al., 1993; von Schantz et al., 1999), genetic variation and species differences (Fuller, Carleton, Fadool, Spady, & Travis, 2005; O'Quin et al., 2012), developmental stage (Carleton et al., 2008; Cheng & Flammarique, 2004; Roberts, Srinivas, Forrest, Escobar, & Reh, 2006), and plastic response to environmental changes (Everett, Tong, Briscoe, & Monteiro, 2012; Fuller & Claricoates, 2011; Hofmann, O'Quin, Smith, & Carleton, 2010; Luehrmann et al., 2018; Stieb, Carleton, Cortesi, Marshall, & Salzburger, 2016; Valen, Karlsen, & Helvik, 2018; Veen, Brock, Rennison, & Bolnick, 2017). There have been five commonly used methods, each with their own merits and drawbacks, to normalize opsin expression for questions of visual ecology and evolution: (1) relative to housekeeping gene

expression, (2) relative to total RNA concentration, (3) relative to expression of the cone-specific alpha subunit of transducin (*GNAT2*), (4) proportion of total pool of opsin expression, and (5) relative to pool of opsin expression from the same cone-type.

Normalizing expression to housekeeping genes is the most common method to observe changes in expression of a gene of interest. A housekeeping gene is assumed to be stably expressed because it is essential for normal cell function regardless of time, environment or cell-type. In theory, the transcript level of this gene in a sample will be linearly correlated with cell number and total transcript abundance. Genes such as β -actin, glyceraldehyde 3-phosphate dehydrogenase (*G3PDH*), elongation factor 1 alpha (*EF1- α*), and ubiquitin conjugating enzyme (*Ube2z*) have all been previously used as housekeeping genes to normalize opsin expression (Fuller & Claricoates, 2011; Halstenberg et al., 2005; Johnson et al., 2013; Sandkam et al., 2015; Shand et al., 2008; Valen et al., 2018). However, none of these studies have tested the validity of their housekeeping genes by showing that their expression is stable in retinal tissue. As the expression of some common housekeeping genes has been shown to vary by tissue, it is crucial to validate these genes in your tissue of interest before use (Bustin, 2002; Kozera & Rapacz, 2013; Radoni et al., 2004). To minimize potential effects of variable expression of housekeeping genes, some studies use an average of three housekeeping genes (Bustin, 2002; Sandkam et al., 2015). There is growing knowledge that housekeeping genes may not have the stability required for normalization. This method also requires identifying the best housekeeping genes, which can be difficult and require extensive testing across species and tissues. Normalizing to total RNA concentration has been used as an alternative (Bustin, 2000, 2002). In addition, a data driven method that assumes expression data is log normally distributed has been developed (Heckmann, Sørensen, Krogh, & Sørensen, 2011). With these methods in mind, normalizing by any reference which has relatively constant expression throughout the day would cause a gene that has cyclic expression to show higher values at one time of the day than another (Figure 1A). This would be useful if the aim of the study was to show diurnal variation in expression of a gene, but would be a source of error when samples taken at different times are used to characterize increases or decreases in opsin expression under different lighting conditions. In this case, some or all of the difference in expression between groups could be due to diurnal variation and not to the treatment.

Another method of opsin normalization compares opsin expression to that of the cone-specific alpha subunit of transducin, *GNAT2*, as it is an essential part of the cone signal transduction cascade and is thought to be common to all cone photoreceptors and highly expressed (Dalton, Lu, Leips, Cronin, & Carleton, 2015; Nandamuri, Dalton, & Carleton, 2017). This acts to normalize to the total number of cone cells rather than total retinal tissue. From our analysis, *GNAT2* has little variation in expression across 27 cichlid retinal transcriptomes (K.L. Carleton, unpublished data) but it is not known if it varies diurnally. Owing to its requirement for photopic vision, it is possible that *GNAT2* varies diurnally in conjunction with the cone opsins under a similar regulatory mechanism. In that case, it might provide a normalization method that acts to remove time of day variation (Figure 1B).

Measuring opsins relative to the total opsin pool, a proportional measure, is common in vision research and is used to give an estimation of which cones contribute to visual

sensitivity, as color vision is achieved through differential stimulation of cones (Carleton & Kocher, 2001; Fuller et al., 2004). However, use of proportions makes the opsin measurements non-independent. If expression of one opsin decreases, the others will appear to increase in abundance when in fact their expression may not have changed. This problem of interdependence has led some to normalize opsins to housekeeping genes, as this would allow detection of patterns of up and down-regulation (Fuller & Claricoates, 2011). Similarly, if it is known that certain opsins are expressed only in specific cone cell types, they can be normalized to the total amount of opsin within that subtype. For example, in cichlid fish, the short-wavelength sensitive opsins have been shown through *in-situ* analysis to only be expressed in the single cones (Dalton, de Busserolles, Marshall, & Carleton, 2017; Fernald, 1981; Nandamuri et al., 2017). Measuring proportional expression in one cell type decreases the noise introduced by variation in expression in other cell types when comparisons to all (total) opsins are made. This is especially useful when the cone mosaic is known and the ratio of cone cell types can be inferred. Thus, it becomes apparent that quantification of opsin expression can lead to vastly different results depending on the type of normalization procedure used. The optimal normalization procedure depends on the scientific question being asked. In this study, we used a large sample of qRT-PCR data coupled with RNAseq to examine diurnal variation in opsin expression in a cichlid. This dataset allows us to resolve four outstanding questions regarding diurnal patterns of opsin expression and qRT-PCR methodology: 1) What time of day is opsin expression the highest? 2) Do all cone opsins have similar diurnal expression cycles? 3) What is the pattern of expression of housekeeping genes and which are the most stable throughout the day? 4) What is the best method of normalization for quantifying or removing diurnal variation in expression?

Methods

Study system

Pelmatolapia mariae (Boulenger, 1899) is a sexually monomorphic fish native to the rivers of central West Africa, from Cote d'Ivoire to Cameroon, although it has been introduced to Australia and North America. It is a sister taxon to cichlids of the Great African lakes (Dunz & Schliewen, 2013; Klett & Meyer, 2002). African cichlids have seven opsin genes, *SWS1*, *SWS2B*, *SWS2A*, *RH2B*, *RH2Aα*, *RH2Aβ*, and *LWS*, although they only express a subset of those opsins (Hofmann & Carleton, 2009; O'Quin, Smith, Naseer, et al., 2011; O'Quin, Smith, Sharma, et al., 2011; Parry et al., 2005; Spady et al., 2006). Laboratory populations were established from the aquarium trade and bred in the Tropical Aquaculture Facility at the University of Maryland for several generations. All fish sampled were sexually mature adults (>6 months of age) from a single brood.

Sample Collection

Samples were collected in June 2016 in one-hour intervals over a continuous 24-hour period starting at 18:00, and including one hour of overlap (18:00) for a total of 25 sampling points. The room was lit by fluorescent lights (Phillips T8 32 watt bulbs) on a 14/10 light cycle with half lights from 06–07:00 and 20–21:00 to simulate dawn and dusk. Fish were immediately euthanized in an overdose of buffered MS-222 according to approved IACUC protocols

(Protocol#: R-AUG-18–41). Retinae were collected from two individuals of each sex per time-point and stored in RNAlater. Sex was determined by gonadal dissection.

Quantitative Real-Time PCR

Opsin expression was quantified using previously described methods (Carleton, 2011) and was averaged across all four individuals per time-point. Briefly, total RNA was extracted using QIAshredder and RNeasy Mini Kits (Qiagen, Valencia, CA, USA) followed by a DNase step to remove genomic DNA. Total RNA was quantified using the Quant-iT RiboGreen RNA assay kit (Thermo Fisher Scientific, Waltham, MA, USA) and compared to a set of RNA standards of known concentrations. cDNA was reverse transcribed from 0.5 µg of total RNA (Superscript III, Invitrogen, Carlsbad, CA, USA). Opsin expression from each cDNA sample was quantified by quantitative real-time PCR (qRT-PCR) using Taqman primers (Life Tech, Carlsbad, CA, USA) specific to each opsin on a Roche Lightcycler 480 (Roche, Bavaria, Germany). The *RH2Aα* and *RH2Aβ* genes are more than 99% identical in sequence and were combined to become the '*RH2A*' measurement for qRT-PCR. Each reaction was performed in duplicate. PCR efficiencies were determined by amplifying an opsin construct containing a single copy of each of the six opsin genes, which was made by amplifying relevant sections of each gene from *O. niloticus* and cloning them into a single vector (Spady et al., 2006). A second multigene gBlock® construct (Integrated DNA Technologies, Inc., Coralville, IA) contained a single copy of *RH1*, *β-actin*, *EF1-α*, *G3PDH*, *Ube2z*, and *GNAT2* and was mixed with the opsin construct in equal proportions; the sequence of this construct can be found in the Supplemental Information. Primer and probe sequences for the cone opsins and *GNAT2* have been published previously (Dalton et al., 2015); sequences for the housekeeping genes and *RH1* can be found in the Supplemental Information.

Housekeeping gene analysis

The ideal housekeeping gene would have no variance in expression throughout the day. Although it is unlikely that any gene could meet that criteria, we can determine which genes have the most stable expression to use for normalization. qRT-PCR Ct (critical threshold) values for each housekeeping gene were analyzed with the program RefFinder (Xie, Xiao, & Chen, 2012). RefFinder runs four top tools for housekeeping gene determination: BestKeeper, geNorm, Normfinder and the comparative delta-Ct method, where each method gives a ranking of the five genes, 1 being the most stable and 5 being the least. RefFinder then assigns a weight to an individual gene, then calculates the geometric mean of these weights for each gene to determine the final ranking score. The average for the top three genes was used from each sample for the three housekeeping (3HK) gene normalization.

Normalization methods for expression

Relative to Housekeeping Genes | 3HK expression was calculated as follows:

$$f_{gene, 3HK} = \frac{(\sum (1 + E_{house}))^{Ct_{house}/3}}{(1 + E_{gene})^{Ct_{gene}}}$$

Relative to total RNA | All cDNA samples analyzed with qRT-PCR were made with equal amounts of RNA, therefore, the expression of each gene can be calculated as:

$$f_{gene, totalRNA} = 1 / (1 + E_{gene})^{Ct_{gene}}$$

Relative to *GNAT2* | All genes were normalized to *GNAT2* as a proxy for cone cell abundance (Dalton et al., 2015):

$$f_{opsin, GNAT2} = \frac{(1 + E_{GNAT2})^{Ct_{GNAT2}}}{(1 + E_{opsin})^{Ct_{opsin}}}$$

Total Opsin Proportional Expression | For each opsin, expression was calculated as a proportion of the total opsin pool (TOT). Here T_{gene} is a value proportional to $1 / (1 + E_{gene})^{Ct_{gene}}$ and the fraction (f) of single cone opsins made up by a particular opsin is:

$$f_{opsin, TOT} = \frac{T_{opsin}}{T_{SWS1} + T_{SWS2B} + T_{SWS2A} + T_{LWS} + T_{RH2A} + T_{RH2B}}$$

Cone-specific Proportional Expression | Proportional normalization for the opsins was also employed using cone-type specific normalization for single cones (SC: *SWS1*, *SWS2A*, *SWS2B*) and double cones (DC: *RH2A*, *RH2B*, *LWS*), calculated as a percent of all opsins within that cell type. Although localization of opsin expression by cone type is not known specifically for this species, microspectrophotometry and fluorescent *in situ* hybridization done in other African cichlids have all shown this opsin expression pattern in the single and double cones (Carleton et al., 2008; Dalton et al., 2017; Dalton, Loew, Cronin, & Carleton, 2014). The fraction of single cone opsins made up by a particular opsin is:

$$f_{opsin, SC} = \frac{T_{opsin}}{T_{SWS1} + T_{SWS2B} + T_{SWS2A}}$$

Similarly, the fraction of double cone opsins made up by a particular opsin is:

$$f_{opsin, DC} = \frac{T_{opsin}}{T_{RH2B} + T_{RH2A} + T_{LWS}}$$

Overall daily variation in expression for a particular normalization method was determined as the coefficient of variance (CV) calculated from the standard deviation of the expression data divided by the mean expression. Expression values for each normalization method can be found in Supplementary File 1. In order to test for a pattern of *RH1* diurnal expression, sliding 6-hour windows of *RH1* expression were tested for significance against all other windows using a post hoc Tukey's multiple comparison of means with a 95% family-wise

confidence level in R version 3.4.3 (R Core Team, 2018). The results of this analysis can be found in Supplementary File 1 under the *RHI* tab.

Validation of qRT-PCR with transcriptomes

RNAseq was performed on one male individual at each of eight time-points (hours 02, 05, 08, 11, 14, 17, 20, and 23). The same RNA extraction was used for RNAseq and to generate the qRT-PCR data for those individuals. Samples were checked for quality on an Agilent Bioanalyzer and RNAseq libraries were made using the TruSeq RNA Sample Preparation Kit v.2 (Illumina, San Diego). Libraries were sequenced on an Illumina HiSeq1500 at the University of Maryland Institute for Bioscience and Biotechnology Research core facility. All samples were pooled and sequenced across two lanes to generate 100bp single-end reads. Overall sequencing quality was examined using FastQC (Andrews, 2010). TruSeq adapters and low-quality reads and bases were trimmed using Trimmomatic version 0.32 (Bolger, Lohse, & Usadel, 2014). Reads were mapped to the Nile tilapia (*Oreochromis niloticus*, version “O_niloticus_UMD_NMBU”) reference genome (Conte et al., 2019; Conte, Gammerdinger, Bartie, Penman, & Kocher, 2017) using Tophat2 version 2.0.13 (Kim et al., 2013) with the *Metriaclyma zebra* UMD2a gene annotations (Conte et al., 2018). Gene expression was quantified as FPKM (Fragments Per Kilobase of transcript per Million mapped reads) using Cufflinks version 2.2.1 (Trapnell et al., 2013). FPKM values for each gene in each sample can be found in Supplementary File 1. Gene expression was then regressed against the qRT-PCR values for the three housekeeping genes (3HK) and total RNA in the corresponding sample in R version 3.4.3 (R Core Team, 2018) to determine the strength of correlation (R^2) and significance. Since the *RH2A α* and *RH2A β* genes are distinct in the genome, they can be separated during read mapping. We were therefore able to get separate expression values for these two homologs. However, when comparing their expression to the qRT-PCR results, we add the values of the two genes because the *RH2A* PCR primers detect both genes together.

Weighted Gene Correlation Network Analysis (WGCNA)

To determine if the pattern of expression of the opsins is similar throughout the day we conducted Weighted Gene Correlation Network Analysis on the transcriptome datasets using the R package WGCNA (Langfelder & Horvath, 2008). WGCNA finds genes that are highly correlated and places them into groups called modules. First, FPKM values were log2-transformed and then filtered to remove genes with too many missing values. The function *pickSoftThreshold* was used to find the soft-thresholding power for which the scale-free topology index curve reaches 0.9. The one-step network construction function *blockwiseModules* was then used to create modules of interconnected genes. The *corType* “bicor” and a conservative value of 0.05 for *mergeCutHeight* was used in the function. The summary profile (eigengene) of each module was used to generate a heat-map of module-trait correlation and significance (*corPvalueStudent*) with the traits being the log2-transformed expression values for each gene from the transcriptome data. Correlations between each opsin or housekeeping gene and *LWS* were calculated using *cor* and the significance of the correlation was found using *corPvalueStudent*. *LWS* was used as the reference as it is the most highly expressed cone opsin in the *P. mariae* retina.

Results

Housekeeping gene selection

We first wanted to identify which of the five housekeeping genes that were previously used in studies of opsin expression was the most consistently expressed throughout the day. RefFinder determines a stability score with lower scores indicating the more stably expressed genes. The housekeeping genes in order from most to least stable were *Ube2z* (1.73), *β-actin* (1.86), *EF1-α* (2), *G3PDH* (3.13), and *GNAT2* (5) with stability scores shown in parentheses. As there was little difference between the scores of *Ube2z*, *β-actin*, and *EF1-α*, it was decided that the average of the top three genes would be used for normalizing the opsins.

Diurnal variation in expression

Quantification of opsin expression revealed that *P. mariae* expresses the cichlid long palette, *SWS2A*, *RH2A*, and *LWS* (Parry et al., 2005). We therefore refer to these three opsins as the palette opsins. As mentioned above, *RH2A* is actually the sum of two homologs, *RH2Aα* and *RH2Aβ* which are not distinguished in the qRT-PCR data. Variation in opsin expression observed throughout the day depended greatly on the normalization method used (Figure 2).

Three housekeeping (3HK) gene normalization—Using the 3HK (*Ube2z*, *β-actin*, and *EF1-α*) normalization for the qRT-PCR data the cone opsins show a variable pattern of expression throughout the day where expression is lowest during the night through early afternoon and increases in the late afternoon to evening. The highest expression is seen between 14:00 and 20:00 (Figure 2A). *LWS* is the most highly expressed cone opsin followed by *RH2A* and *SWS2A*. *SWS2B* is very lowly expressed and *SWS1* and *RH2B* are effectively not expressed. The palette cone opsins (*SWS2A*, *RH2A*, and *LWS*) and *GNAT2* appear to be in phase, except between the hours of 05:00 to 10:00 when *GNAT2* is much higher than the cone opsins. *G3PDH* expression is less variable throughout the day than the opsins or *GNAT2*, as expected given its stability score. The 3HK method produces the high variability in expression of the palette opsins as shown by the large average coefficient of variation of the palette opsins at 0.78 (Table 1).

RH1, the rod opsin, expression varies greatly from hour to hour (Figure 3). In general, *RH1* is highly expressed throughout the day at a level more than 10 times higher than any of the cone opsins. In order to find a pattern of *RH1* expression, sliding 6-hour windows were tested for significance against all other windows. This method demonstrated that only the time between 06:00 and 12:00 had significantly higher *RH1* expression than the other windows.

To check the reproducibility of opsin expression, we compared expression of each of the opsins between the first time-point (hour 18) and the overlap time-point from 24 hours later (also hour 18) using student's t-tests with multiple test correction (Table S1). The repeat sampling showed no significant difference in any of the opsins.

Normalization to total RNA—Interestingly, normalizing genes to total RNA produced a pattern of daily variation strikingly similar to the 3HK normalization (Figure S1), but with a

higher coefficient of variation than 3HK for the palette opsins (0.90). To test for a relationship between these two methods, a regression of the average expression for each gene at each time-point was done comparing 3HK to total RNA (Figure S2). A significant linear relationship was observed between the values obtained with 3HK and total RNA normalization with ($R^2=0.78$; $p < 2.2e-16$) and without *RHI* ($R^2=0.50$; $p < 2.2e-16$).

Normalization with cone-specific gene GNAT2—When normalized to *GNAT2*, the cone opsins show relatively high diurnal variation, especially when looking at *LWS* expression. The average coefficient of variation for the palette opsins was 0.55 (Table 1). For all palette opsins, there is very little expression in the morning between 06:00 and 09:00 (Figure 2B). For *LWS*, expression increases throughout the day and peaks at 20:00 (Figure 2B). While expression decreases after 20:00, it remains high during the night. *RH2A* shows a similar trend although it is expressed at a lower level than *LWS*. *SWS2A* is expressed at much lower levels and also shows less diurnal variation.

Normalization by total opsin pool—Normalizing to the total opsin pool shows a moderate level of diurnal variation; the coefficient of variation between the opsins was 0.30 (Table 1). *RH2A* remained constant relative to the other opsins (Figure 2C). *LWS* and *SWS2A* are relatively constant except in the morning hours between 04:00 and 10:00, when they slowly reach equal proportions (40% of each) at around 08:00. *SWS2B* expression is near zero for most of the day, but increases to as much as 6% of the total opsin pool between 07:00 and 12:00.

Normalization by cone-type—Normalizing by cone-type produces the least amount of diurnal variation in opsin expression with a coefficient of variation of the palette opsins at 0.14 (Table 1). The composition of opsin by cone-type remains quite constant, especially in the double cones which contain *LWS* and/or *RH2A* (Figure 2D). *LWS* makes up 70–90% of the double cone opsin pool throughout the day with *RH2A* making up the rest. At all time-points, the single cone opsin pool is comprised mostly of *SWS2A* with less than 25% *SWS2B* expression (Figure 2D).

Transcriptomes

After filtering, the average number of Illumina reads for each sample was 46.2 (± 4.9) million. For each sample, the FPKM for each gene was compared against its expression level from the 3HK normalization. Both the FPKM and 3HK measures show a similar diurnal expression pattern for all genes, except *G3PDH*. *G3PDH* is higher in the 3HK qRT-PCR results than in the RNAseq data compared to the other genes (Figure 4). A regression of all of these values shows that they are highly correlated ($R^2=0.97$; $p < 2 \times 10^{-16}$). While this is largely driven by *RHI*, there was still a strong correlation ($R^2=0.53$, $p=8.48 \times 10^{-12}$) after removing *RHI* (Figure 5). A second regression of the total RNA values against the FPKM also showed a positive correlation with ($R^2=0.84$; $p < 2 \times 10^{-16}$) and without *RHI* ($R^2=0.68$; $p < 2 \times 10^{-16}$; Figure S3).

Weighted gene co-expression network analysis (WGCNA)

A total of 29,510 genes passed the sparseness quality filtering step and were included in the WGCNA analysis. A soft-thresholding power of 20 was chosen from the scale-free topology (Figure S4). The analysis produced 145 modules with the number of genes ranging from 3,089 in the grey (uncorrelated) module to only 32 in the lavenderblush module (Figure S5). The palette opsins were placed in the same module, sienna4, and are strongly correlated (Table 2). The sienna4 module contains only 59 other genes. Because of the strong correlation between the palette opsins (*SWS2A*, *RH2A*, and *LWS*), correlation to *LWS* was used as a proxy for cone opsin expression for comparisons to all other genes. Interestingly, although *SWS2B* did not get placed in the same module as the palette opsins, it was still significantly correlated with *LWS* (cor=0.739, p=0.036). *GNAT2* was placed in a different module from the cone opsins but did not show a significant correlation with *LWS* (cor=0.697, p=0.054). *RH1* and *GNAT2* were placed in the same module. The other focal genes fell into separate modules except for two of the three normalization housekeeping genes, *EF1-α* and *β-actin*, which are both in the yellow module. The expression of all other housekeeping genes was uncorrelated with *LWS*. A heatmap of module correlations for each gene can be viewed in Figure S6.

Discussion

As opsin expression is tightly linked to its functional role in determining visual sensitivities, numerous examples have demonstrated a connection to evolutionary adaptation to local environments. With the advent of high throughput and affordable technologies such as qRT-PCR, quantifying opsin expression has become more affordable and less time consuming. While qRT-PCR provides a convenient tool to determine expression, the normalization method can have a strong impact on the interpretation of the results. Here, we show using qRT-PCR and transcriptome measures that opsin expression measurements vary significantly depending on the normalization method. We found that normalizing by the average of three housekeeping genes (3HK) performs the best for studies of diurnal variation in opsin expression. Meanwhile normalizing opsin expression by cone-type is the best approach when asking questions about visual tuning and when samples cannot be taken at the same time of day. Below we weigh each of the methods and outline their strengths and weaknesses, summarizing these recommendations in Table 3.

Normalization strategies for daily variation

The recommended method for observing and quantifying diurnal variation in expression is 3HK, which uses the average of the three housekeeping genes *B-actin*, *EF1-a*, and *Ube2z* for normalization (Table 3). The inclusion of three housekeeping genes instead of one provides more stability (Vanguilder, Vrana, & Freeman, 2008). In addition, the use of housekeeping genes allows actual changes in transcript number of the opsins to be quantified, whereas use of proportional measures does not. Normalizing to total RNA may also be a viable normalization method for observing diurnal variation, as it was highly correlated with 3HK expression. Total RNA normalization might serve as the better choice if a housekeeping gene has not been validated for use in a specific tissue or species, but this method is very sensitive to measurement error and can only be used on samples that were all collected,

processed, and measured via qRT-PCR at the same time. Additionally, extreme caution must be taken when quantifying total RNA, using dye-based assays such as Ribogreen, as absorbance-based assays such as nanodrop are too variable for such techniques, especially for low quantity samples (Bustin, 2002). Using these methods we have shown that cone opsins in *P. mariae* are more highly expressed in the late afternoon, reaching a peak around 18:00. This lines up well with observations in other fishes such as *A. burtoni* (Halstenberg et al., 2005), zebrafish (Li et al., 2005), and bluefin killifish (Johnson et al., 2013) which also have peak cone opsin expression in the late afternoon (Figure 6). This assemblage of diurnally active fishes represents about 250 million years of evolution (Kumar, Stecher, Suleski, & Hedges, 2017), which may indicate that diurnal patterns are conserved over a substantial part of fish diversity. This suggests that to maximize signal from the cone opsins, retinal samples should be taken between the hours of 14:00 and 20:00. *RHI* expression was highest between 06:00 and 12:00, which did follow the pattern seen in the related African cichlid, *A. burtoni*, where expression was highest between 09:00 and 14:00 (Halstenberg et al., 2005). Interestingly, the increase in *RHI* expression seems to correspond with the beginning of light onset in both of these studies, as lights were turned on at 06:00 in our study and 09:00 in the Halstenberg study. In our study, *RHI* was always highly expressed, but that amount seemed to vary drastically from hour to hour (Figure 3). FPKM from transcriptomes of 8 samples were compared to the 3HK qRT-PCR results of those same samples as a validation of the normalization method. The results from the two methods were highly correlated for the opsins, *GNAT2*, and *G3PDH*, but especially for *RHI* (Figure 5). This supports the idea that the variation we observed in *RHI* is real and likely not due to an issue with the 3HK normalization or qRT-PCR methodology.

GNAT2 is not stable throughout the day as indicated by the low score it received from RefFinder. Because it does not cycle in perfect synchrony with the opsins it is not a valid normalization method for observing or removing diurnal variation. However, it could still be used for comparing samples taken at the same time of day to look at changes in expression level (Table 3).

Of the two proportional methods, opsin expression relative to cone type produced less time of day variation than expression relative to the total pool. Therefore, normalizing opsin expression by cone-type is the best method for removing diurnal variation and for comparison of opsin expression across studies with the only drawback being that only proportional changes in expression will be observed and not up or down-regulation (Table 3). However, if cone cell localization has not been determined for a particular study species, then normalizing to total opsin is the best choice for removing diurnal variation.

Cone opsins show a similar diurnal pattern of expression

It is theoretically possible that one opsin could have its highest expression at 10:00 while another peaks at 19:00; this would mean that sampling time could greatly affect the composition of opsins sampled. However, we found that all the cone opsins had a similar diurnal rhythm. While the 3HK results suggested that the cone opsins have a very similar cycle of expression, we showed this empirically using weighted gene correlation network analysis (WGCNA). Using WGCNA, we found that the palette opsins, *SWS2A*, *RH2A*, and

LWS, are in the same module, sienna4, demonstrating that their expression is highly correlated. Therefore, sampling at different times of the day will give a similar composition of the most abundantly expressed opsins. This is also demonstrated by the higher stability of proportional normalization by total opsin and cone-type.

GNAT2 did not cluster in the same module as the palette opsins and but did have a nearly significant correlation with *LWS* expression throughout the day. *GNAT2* has a more similar diurnal expression pattern to the cone opsins than the other genes examined. This suggests that *GNAT2* might be regulated in a similar way to the cone opsins as it is expressed in the same cell type. Although *RHI* and *GNAT2* were placed into the same module, that module is large (1,530 genes) and less informative. There is no evidence to suggest that *RHI* and *GNAT2* should be correlated in their expression throughout the day as they are expressed in separate photoreceptor cell types (*RHI* in rods and *GNAT2* in cones). Other studies have shown that *RHI* is expressed highly during the first half of the day, then it decreases at the same time that cone opsins begin increasing to their maximum expression in the second half of the day (Halstenberg et al., 2005; Korenbrot & Fernald, 1989). It was expected that *RHI* and *LWS* expression might show a negative relationship, but we did not find a correlation. Two of the housekeeping genes used in the 3HK normalization, *β -actin* and *EF1- α* , were placed in the same module, although that module was again large at 1,126 genes. None of the potential housekeeping genes, *β -actin*, *EF1- α* , *Ube2z*, or *G3PDH*, were correlated with *LWS* expression. This was expected as RefFinder showed they are relatively stable throughout the day and therefore would not cycle with the highly variable cone opsins.

Best practices for measuring opsin expression with qRT-PCR

Understanding the limitations and specific questions of an opsin expression study are essential to obtaining meaningful results. For instance, measuring plasticity of opsin expression in response to changing light environments requires the utmost sensitivity to detect up or down-regulation of the opsins. For such an exacting study, it is essential to standardize the time of day at which fish are dissected in order to ensure that when normalizing to housekeeping gene(s), changes in expression are not simply due to time-of-day variation (Figure 1). As an example, Nandamuri et al. (2017) state that all retinae were dissected between 15:00 and 17:00. This specific window was needed as qRT-PCR normalization was done using *GNAT2*, which can be used to observe up or down-regulation between samples taken on different days only if the sampling time is strictly controlled. For other studies attempting to genetically map quantitative trait loci that underlie the molecular mechanisms of opsin expression, it may be sufficient to use a simpler method such as normalization to total opsin or cone-type (Nandamuri, Conte, & Carleton, 2018; O'Quin et al., 2012). This decreases sampling noise from diurnal variation and so improves phenotype accuracy to enhance mapping, but also fails to separate mechanisms that change absolute levels of opsin expression. While it is easy to stick to an exact sampling time in a lab setting, it may be difficult in the field. Situations also arise where collaborators share samples where sampling time is unknown or different from each other's datasets. In these instances, it may be better to use a more robust normalization method such as total opsin or cone-type to reduce time-of-day noise.

Measuring opsin expression with qRT-PCR is useful for determining the visual sensitivity of organisms, but comes with important caveats relating to normalization. While we have established some best practices for qRT-PCR measurements using diurnal patterns of opsin expression in the cichlid fish *P. mariae*, these concepts could certainly be applied more generally to fishes and other organisms. Very little is known about the exact daily rhythms of cone opsin expression in animals other than fish. A study in chickens demonstrated there is a peak of expression of the red-sensitive pigment, iodopsin, in the late afternoon similar to several fishes mentioned previously (Pierce et al., 1993). A study in mice found the peak of ultraviolet cone opsin expression earlier in the afternoon, around 12:00 (von Schantz et al., 1999). Some evidence shows that daily variation may affect visual sensitivity and even behavioral responses (Johnson et al., 2013; Li et al., 2005), but more work needs to be done to assess these methods for use in behavioral studies. Despite this dearth of knowledge, those studying opsin expression in systems where daily cycles are unknown can still follow these recommendations of strict sampling times and normalization methods to avoid time-of-day variation. With the understanding of how normalization affects diurnal opsin expression data gathered by qRT-PCR, we have a new hope to improve the accuracy and standardization across study systems.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We would like to thank Suwei Zhao of the Institute for Bioscience and Biotechnology, UMD for help preparing the transcriptome libraries. All animal procedures were approved by the IACUC at the University of Maryland, College Park (UMD# R-AUG-18-41). This work was supported with funds from the National Eye Institute at the National Institutes of Health (NIH R01EY024639 to K.L.C).

References

- Andrews S. (2010). FastQC: a quality control tool for high throughput sequence data.
- Basinger S., Hoffman R., & Matthes M. (1976). Photoreceptor shedding is initiated by light in the frog retina. *Science*, 194(4269), 1074–1076. 10.1126/science.1086510 [PubMed: 1086510]
- Bolger AM, Lohse M., & Usadel B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120. 10.1093/bioinformatics/btu170 [PubMed: 24695404]
- Bowmaker JK (2008). Evolution of vertebrate visual pigments. *Vision Research*, 48(20), 2022–2041. 10.1016/j.visres.2008.03.025 [PubMed: 18590925]
- Bowmaker JK, & Hunt DM (2006). Evolution of vertebrate visual pigments. *Current Biology*, 16(13), pR484–R489. 10.1016/j.cub.2006.06.016 [PubMed: 16824904]
- Bustin SA (2000). Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of Molecular Endocrinology*, 25, 169–193. [PubMed: 11013345]
- Bustin SA (2002). Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): Trends and problems. *Journal of Molecular Endocrinology*, 29(1), 23–39. 10.1677/jme.0.0290023 [PubMed: 12200227]
- Carleton KL (2011). Quantification of Transcript Levels with Quantitative RT-PCR In Orgogozo V. & Rockman M. (Eds.), *Methods in Molecular Biology* (pp. 279–295). Springer Science+Business Media, LLC 2011.

- Carleton KL, & Kocher TD (2001). Cone opsin genes of African cichlid fishes: Tuning spectral sensitivity by differential gene expression. *Molecular Biology and Evolution*, 18(8), 1540–1550. 10.1093/oxfordjournals.molbev.a003940 [PubMed: 11470845]
- Carleton KL, Spady TC, Streelman JT, Kidd MR, McFarland WN, & Loew ER (2008). Visual sensitivities tuned by heterochronic shifts in opsin gene expression. *BMC Biology*, 6(1), 22. 10.1186/1741-7007-6-22 [PubMed: 18500997]
- Cheng CL, & Flammarique IN (2004). A new mechanism for modulating color vision. *Nature*, 428(6980), 279.
- Chinen A., Hamaoka T., Yamada Y., & Kawamura S. (2003). Gene Duplication and Spectral Diversification of Cone Visual Pigments of Zebrafish, 675(February), 663–675.
- Conte MA, Gammerdinger WJ, Bartie KL, Penman DJ, & Kocher TD (2017). A high quality assembly of the Nile Tilapia (*Oreochromis niloticus*) genome reveals the structure of two sex determination regions. *BMC Genomics*, 18(1), 1–19. 10.1186/s12864-017-3723-5 [PubMed: 28049423]
- Conte MA, Joshi R., Moore EC, Nandamuri SP, Gammerdinger WJ, Roberts RB, ... Kocher TD (2018). Chromosome-scale assemblies reveal the structural evolution of African cichlid genomes. *BioRxiv*.
- Conte MA, Joshi R., Moore EC, Nandamuri SP, Gammerdinger WJ, Roberts RB, ... Kocher TD (2019). Chromosome-scale assemblies reveal the structural evolution of African cichlid genomes. *Gigascience*, 8(4). 10.1093/gigascience/giz030
- Dalton BE, de Busserolles F., Marshall NJ, & Carleton KL (2017). Retinal specialization through spatially varying cell densities and opsin coexpression in cichlid fish. *The Journal of Experimental Biology*, 220, 266–277. 10.1242/jeb.149211 [PubMed: 27811302]
- Dalton BE, Loew ER, Cronin TW, & Carleton KL (2014). Spectral tuning by opsin coexpression in retinal regions that view different parts of the visual field. *Proceedings of the Royal Society B: Biological Sciences*, 281(1797), 20141980–20141980. 10.1098/rspb.2014.1980
- Dalton BE, Lu J., Leips J., Cronin TW, & Carleton KL (2015). Variable light environments induce plastic spectral tuning by regional opsin coexpression in the African cichlid fish, *Metriacrima zebra*. *Molecular Ecology*, 24, 4193–4204. 10.1111/mec.13312 [PubMed: 26175094]
- Dunz AR, & Schlieffen UK (2013). Molecular phylogeny and revised classification of the haplotilapiine cichlid fishes formerly referred to as “Tilapia.” *Molecular Phylogenetics and Evolution*, 68(1), 64–80. 10.1016/j.ympev.2013.03.015 [PubMed: 23542002]
- Everett A., Tong X., Briscoe AD, & Monteiro A. (2012). Phenotypic plasticity in opsin expression in a butterfly compound eye complements sex role reversal. *BMC Evolutionary Biology*, 12(1), 232. 10.1186/1471-2148-12-232 [PubMed: 23194112]
- Fernald RD (1981). Chromatic organization of a cichlid fish retina. *Vision Research*, 21(12), 1749–1753. [PubMed: 7336611]
- Fuller RC, Carleton KL, Fadool JM, Spady TC, & Travis J. (2004). Population variation in opsin expression in the bluefin killifish, *Lucania goodei*: A real-time PCR study. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 190(2), 147–154. 10.1007/s00359-003-0478-z
- Fuller RC, Carleton KL, Fadool JM, Spady TC, & Travis J. (2005). Genetic and environmental variation in the visual properties of bluefin killifish, *Lucania goodei*. *Journal of Evolutionary Biology*, 18(3), 516–523. 10.1111/j.1420-9101.2005.00886.x [PubMed: 15842481]
- Fuller RC, & Claricoates KM (2011). Rapid light-induced shifts in opsin expression: Finding new opsins, discerning mechanisms of change, and implications for visual sensitivity. *Molecular Ecology*, 20(16), 3321–3335. 10.1111/j.1365-294X.2011.05180.x [PubMed: 21749514]
- Ginzinger DG (2002). Gene quantification using real-time quantitative PCR. *Experimental Hematology*, 30(6), 503–512. 10.1016/s0301-472x(02)00806-8 [PubMed: 12063017]
- Halstenberg S., Samagh SPS, Lindgren KM, Balt S., Fernald RD, & Nadal-Vicens M. (2005). Diurnal rhythm of cone opsin expression in the teleost fish *Haplochromis burtoni*. *Visual Neuroscience*, 22(02), 135–141. 10.1017/s0952523805222022 [PubMed: 15935106]
- Heckmann L., Sørensen PB, Krogh PH, & Sørensen JG (2011). NORMA-Gene: A simple and robust method for qPCR normalization based on target gene data. *BMC Bioinformatics*, 12(250).

- Hofmann CM, & Carleton KL (2009). Gene duplication and differential gene expression play an important role in the diversification of visual pigments in fish. *Integrative and Comparative Biology*, 49(6), 630–643. 10.1093/icb/icp079 [PubMed: 21665846]
- Hofmann CM, O'Quin KE, Smith AR, & Carleton KL (2010). Plasticity of opsin gene expression in cichlids from Lake Malawi. *Molecular Ecology*, 19(10), 2064–2074. 10.1111/j.1365-294X.2010.04621.x [PubMed: 20374487]
- Jacobs GH, & Rowe MP (2004). Evolution of vertebrate colour vision. *Clinical and Experimental Optometry*, 87(4–5), 206–216. [PubMed: 15312024]
- Johnson AM, Stanis S., & Fuller RC (2013). Diurnal lighting patterns and habitat alter opsin expression and colour preferences in a killifish. *Proceedings of the Royal Society B: Biological Sciences*, 280(1763). 10.1098/rspb.2013.0796
- Kelber A., Vorobyev M., & Osorio D. (2003). Animal colour vision – behavioural tests and physiological concepts. *Biological Reviews of the Cambridge Philosophical Society*, 78(1), S1464793102005985. 10.1017/S1464793102005985
- Kim D., Perteau G., Trapnell C., Pimentel H., Kelley R., & Salzberg SL (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology*, 14(R36), 1–13.
- Klett V., & Meyer A. (2002). What, if anything, is a Tilapia? - Mitochondrial ND2 phylogeny of tilapiines and the evolution of parental care systems in the African cichlid fishes. *Molecular Biology and Evolution*, 19(6), 865–883. 10.1093/oxfordjournals.molbev.a004144 [PubMed: 12032243]
- Korenbrot JJ, & Fernald RD (1989). Circadian rhythm and light regulate opsin mRNA in rod photoreceptors. *Nature*. 10.1038/337454a0
- Kozera B., & Rapacz M. (2013). Reference genes in real-time PCR. *Plant Genetics*, 54, 391–406. 10.1007/s13353-013-0173-x
- Kumar S., Stecher G., Suleski M., & Hedges SB (2017). TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Molecular Biology and Evolution*, 34(7), 1812–1819. 10.1093/molbev/msx116 [PubMed: 28387841]
- Langfelder P., & Horvath S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, 9(1), 559. 10.1186/1471-2105-9-559 [PubMed: 19114008]
- LaVail MM (1980). Circadian nature of rod outer segment disc. *Invest Ophthalmol Vis Sci*, 407–411. [PubMed: 7358492]
- Levine JS, & Machnichol E. (1979). Visual pigments in teleost fishes - effects of habitat, microhabitat, and behavior on visual-system evolution. *Sensory Processes*, 3, 95–131. [PubMed: 545702]
- Li P, Temple S., Gao Y, Haimberger T., Hawryshyn CW, & Li L. (2005). Circadian rhythms of behavioral cone sensitivity and long wavelength opsin mRNA expression: a correlation study in zebrafish. *Journal of Experimental Biology*, 208(3), 497–504. 10.1242/jeb.01424 [PubMed: 15671338]
- Loew ER, & Lythgoe J. (1978). The ecology of cone pigments in teleost fishes. *Vision Research*, 18, 715–722. [PubMed: 664357]
- Luehrmann M., Stieb SM, Carleton KL, Pietzker A., Cheney KL, & Marshall NJ (2018). Short-term colour vision plasticity on the reef: changes in opsin expression under varying light conditions differ between ecologically distinct fish species. *The Journal of Experimental Biology*, 221(22), jeb175281. 10.1242/jeb.175281
- Lythgoe J. (1984). Visual pigments and environmental light. *Vision Research*, 24, 1539–1550. [PubMed: 6398560]
- Matsumoto Y., Fukamachi S., Mitani H., & Kawamura S. (2006). Functional characterization of visual opsin repertoire in Medaka (*Oryzias latipes*). 371, 268–278. 10.1016/j.gene.2005.12.005
- Nandamuri SP, Conte MA, & Carleton KL (2018). Multiple trans QTL and one cis-regulatory deletion are associated with the differential expression of cone opsins in African cichlids. *BMC Genomics*, 19(1), 1–17. 10.1186/s12864-018-5328-z [PubMed: 29291715]
- Nandamuri SP, Dalton BE, & Carleton KL (2017). Determination of the genetic architecture underlying short wavelength sensitivity in Lake Malawi Cichlids. *Journal of Heredity*, 108(4), 379–390. 10.1093/jhered/esx020 [PubMed: 28498989]

- O'Quin KE, Schulte JE, Patel Z., Kahn, NadiaNaseer Z., Wang H., Conte MA, & Carleton KL (2012). Evolution of cichlid vision via trans-regulatory divergence. *BMC Evolutionary Biology*, 12(1), 251 10.1186/1471-2148-12-251 [PubMed: 23267665]
- O'Quin KE, Smith AAR, Sharma A., Carleton KL, O'Quin KE, & Smith AR (2011). New evidence for the role of heterochrony in the repeated evolution of cichlid opsin expression. *Evolution and Development*, 13(2), 193–203. 10.1111/j.1525-142X.2011.00469.x [PubMed: 21410875]
- O'Quin KE, Smith D., Naseer Z., Schulte JE, Engel SD, Loh Y-HE, ... Carleton KL (2011). Divergence in cis-regulatory sequences surrounding the opsin gene arrays of African cichlid fishes. *BMC Evolutionary Biology*, 11(1), 120 10.1186/1471-2148-11-120 [PubMed: 21554730]
- Parry JW, Carleton KL, Spady TC, Carboo A., Hunt DM, & Bowmaker JK (2005). Mix and match color vision: Tuning spectral sensitivity by differential opsin gene expression in Lake Malawi cichlids. *Current Biology*, 15(19), 1734–1739. 10.1016/j.cub.2005.08.010 [PubMed: 16213819]
- Pierce ME, Sheshberadaran H., Zhang Z., Fox LE, Applebury ML, & Takahashi JS (1993). Circadian regulation of iodopsin gene expression in embryonic photoreceptors. *Neuron*, 10, 579–584. [PubMed: 8476610]
- Radoni A., Thulke S., Mackay IM, Landt O., Siegert W., & Nitsche A. (2004). Guideline to reference gene selection for quantitative real-time PCR. *Biochemical and Biophysical Research Communications*, 313, 856–862. 10.1016/j.bbrc.2003.11.177 [PubMed: 14706621]
- Roberts MR, Srinivas M., Forrest D., Escobar G. M. de, & Reh TA (2006). Making the gradient: Thyroid hormone regulates cone opsin expression in the developing mouse retina. *Proceedings of the National Academy of Sciences*, 103(16), 6218–6223. 10.1073/PNAS.0509981103
- Rodieck RW (1998). *First Steps in Seeing*. Sunderland, MA: Sinauer Associates, Inc.
- Sandkam BA, Young CM, & Breden F. (2015). Beauty in the eyes of the beholders: colour vision is tuned to mate preference in the Trinidadian guppy (*Poecilia reticulata*). *Molecular Ecology*, 24, 596–609. 10.1111/mec.13058 [PubMed: 25556876]
- Shand J., Davies WL, Thomas N., Balmer L., Cowing JA, Pointer M., ... Hunt DM (2008). The influence of ontogeny and light environment on the expression of visual pigment opsins in the retina of the black bream, *Acanthopagrus butcheri*. *Journal of Experimental Biology*, 211(9), 1495–1503. 10.1242/jeb.012047 [PubMed: 18424684]
- Spady TC, Parry JW, Robinson PR, Hunt DM, Bowmaker JK, & Carleton KL (2006). Evolution of the Cichlid Visual Palette through Ontogenetic Subfunctionalization of the Opsin Gene Arrays. *Molecular Biology and Evolution*, 23(8), 1538–1547. 10.1093/molbev/msl014 [PubMed: 16720697]
- Stieb SM, Carleton KL, Cortesi F., Marshall NJ, & Salzburger W. (2016). Depth-dependent plasticity in opsin gene expression varies between damselfish (Pomacentridae) species. *Molecular Ecology*, 25(15), 3645–3661. 10.1111/mec.13712 [PubMed: 27262029]
- Trapnell C., Roberts A., Goff L., Pertea G., Kim D., Kelley DR, ... Pachter L. (2013). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols*, 7(3), 562–578. 10.1038/nprot.2012.016.Differential
- Valen R., Karlsen R., & Helvik JV (2018). Environmental, population and life-stage plasticity in the visual system of Atlantic cod. *The Journal of Experimental Biology*, 221(1), jeb165191. 10.1242/jeb.165191
- Vanguilder HD, Vrana KE, & Freeman WM (2008). Twenty-five years of quantitative PCR for gene expression analysis. *BioTechniques*, 44(5), 619–626. 10.2144/000112776 [PubMed: 18474036]
- Veen T., Brock C., Rennison D., & Bolnick D. (2017). Plasticity contributes to a fine-scale depth gradient in sticklebacks' visual system. *Molecular Ecology*, 26(16), 4339–4350. 10.1111/mec.14193 [PubMed: 28570029]
- von Schantz M. Von, Lucas RJ, & Foster RG (1999). Circadian oscillation of photopigment transcript levels in the mouse. *Molecular Brain Research*, 72, 108–114. [PubMed: 10521605]
- Xie F., Xiao P., & Chen D. (2012). miRDeepFinder: a miRNA analysis tool for deep sequencing of plant small RNAs. *Plant Molecular Biology*, 80, 75–84. 10.1007/s11103-012-9885-2
- Yokoyama S. (2008). Evolution of Dim-Light and Color Vision Pigments. *Annual Review of Genomics and Human Genetics*, 9(1), 259–282. 10.1146/annurev.genom.9.081307.164228

Young RW (1978). The daily rhythm of shedding and degradation of rod and cone outer segment membranes in the chick retina. *Investigative Ophthalmology & Visual Science*, 17(2), 105–116. [PubMed: 624604]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1. qRT-PCR normalization method significantly affects patterns of expression. a) When an opsin gene is normalized to a stable reference gene, the diurnal patterns of expression can be observed. b) When an opsin is normalized to a reference with a similar diurnal cycle, the diurnal patterns of expression disappear.

Figure 2.

Diurnal expression as measured by qRT-PCR with different normalization methods. Error bars show standard error of the mean (n=4). Bars under the hour of the day show night (black), twilight (grey), and day (gold). a) Cone opsin, *G3PDH*, and *GNAT2* expression normalized relative to the average of three housekeeping genes (*Ube2z*, EF1- α , β -actin). b) Cone opsin expression normalized relative to *GNAT2*. c) Cone opsin expression normalized relative to the total opsin pool. d) Cone opsin expression normalized relative to cone-type (single cones and double cones).

Figure 3. Diurnal expression of rhodopsin (*RH1*) as measured by qRT-PCR relative to the average of three housekeeping genes. Error bars show standard error of the mean (n=4). Bars under the hour of the day show night (black), dawn or dusk (grey), and day (gold).

Figure 4. A comparison of expression measured by a) qRT-PCR normalized to the average of three housekeeping genes and b) FPKM for the single individuals for which qRT-PCR and RNAseq was completed. *RH1* is presented on the secondary axis.

Figure 5.

A regression of the qRT-PCR results against FPKM from the transcriptomes of samples from 8 time-points by gene ($R^2=0.97$). Inset shows the regression excluding *RH1* ($R^2=0.53$).

Diurnal expression of *LWS* in fishes

Figure 6.

Expression of the long-wavelength sensitive opsin, *LWS*, throughout the day in four fish species, *Pelmatolapia mariae*, *Astatotilapia burtoni* (Halstenberg et al., 2005), *Lucania goodei* (Johnson et al., 2013), and *Danio rerio* (Li et al., 2005). Values are normalized to the highest expression value from the dataset

Table 1.

Coefficient of variation for the palette opsins by normalization method. The average coefficient of variation across the three palette opsins is shown in the last row. ⁺*RH2A* is a combination of *RH2A α* and *RH2A β* expression.

Palette cone opsin	3HK	Total RNA	<i>GNAT2</i>	Total opsin	Cone-type
<i>SWS2A</i>	0.68	0.86	0.43	0.52	0.07
<i>RH2A</i> ⁺	0.82	0.97	0.59	0.22	0.28
<i>LWS</i>	0.83	0.88	0.64	0.15	0.07
Average	0.78	0.90	0.55	0.30	0.14

Table 2.

WGCNA module membership of opsins and housekeeping genes. Significant correlations of a given gene with *LWS* are shown with an asterisk ($p < 0.05$). The palette opsins are highlighted in dark grey. The three housekeeping genes used in 3HK normalization are highlighted in light grey.

Short name	Module	Correlation with <i>LWS</i>	p-value	Full gene name
LOC101487452	sienna4	0.996	1.59E-07*	red-sensitive opsin (<i>LWS</i>)
LOC101482689	sienna4	0.986	6.61E-06*	green-sensitive opsin (<i>RH2Aα</i>)
LOC101483175	sienna4	0.985	7.33E-06*	green-sensitive opsin (<i>RH2Aβ</i>)
LOC101486588	sienna4	0.925	9.89E-04*	blue-sensitive opsin (<i>SWS2A</i>)
opn1sw	magenta	0.644	0.084	opsin 1, short wave sensitive (<i>SWS1</i>)
LOC101482402	white	0.428	0.289	green-sensitive opsin (<i>RH2B</i>)
LOC101487155	darkviolet	0.739	0.036*	blue-sensitive opsin (<i>SWS2B</i>)
LOC101473594	brown	0.583	0.128	rhodopsin (<i>RH1</i>)
LOC101481561	brown	0.697	0.054	guanine nucleotide-binding protein G(t) subunit alpha-2 (<i>GNAT2</i>)
actb	yellow	-0.305	0.462	actin beta (<i>β-actin</i>)
ube2z	red	0.322	0.436	ubiquitin conjugating enzyme E2 Z (<i>Ube2z</i>)
LOC101482309	yellow	-0.615	0.104	elongation factor 1-alpha (<i>EF1-α</i>)
gapdh	darkolivegreen1	-0.535	0.171	glyceraldehyde-3-phosphate dehydrogenase (<i>G3PDH</i>)

Table 3.

Sampling and normalization method recommendations based on the particular needs and limitations of a qRT-PCR study. ⁺Proportional gene expression is a non-independent measure of the opsins relative to each other.

Sampling time for highest cone opsin signal: 14:00 to 20:00						
Method	Remove diurnal variation	Observe diurnal variation	Up or down-regulation of gene expression	Proportional gene expression ⁺	Notes	Study Example
Three housekeeping (3HK) genes	No	Yes	Yes	No	Best for observing diurnal variation. Can use for up/down regulation if samples are taken at same time of day.	Determining the daily pattern of opsin expression for different populations. (Sandkam et al., 2015)
Total RNA concentration	No	Yes	Yes	No	Fine for same observations as 3HK if there is no validated HK gene. Sensitive to RNA quantification.	Present study
Cone-specific gene (<i>GNAT2</i>)	No	No	Yes	No	Not recommended unless all samples are taken at the same time of day.	Observing an increase or decrease in expression of opsins in response to changes in light. (Nandamuri et al., 2017)
Total opsin pool	Some	No	No	Yes	Best for removing variation only if cone-type localization is not known.	Determining opsin expression throughout development. (Carleton et al., 2008)
Cone-type	Yes	No	No	Yes	Best for removing diurnal variation if cone-type localization is known.	Mapping the genetic basis of opsin expression. (Nandamuri et al., 2018)