

## **An extended family of novel vertebrate photopigments is widely expressed and displays a diversity of function**

Wayne I. L. Davies<sup>1,2§</sup>, T. Katherine Tamai<sup>3§</sup>, Lei Zheng<sup>2</sup>, Josephine K. Fu<sup>2</sup>, Jason Rihel<sup>4</sup>, Russell G. Foster<sup>2</sup>, David Whitmore<sup>3#</sup> and Mark W. Hankins<sup>2#</sup>

<sup>1</sup>School of Animal Biology and University of Western Australia Oceans Institute, University of Western Australia, Perth, Western Australia, Australia.

<sup>2</sup>Nuffield Laboratory of Ophthalmology, Nuffield Department of Clinical Neurosciences, John Radcliffe Hospital, University of Oxford, Oxford, OX3 9DU, UK.

<sup>3</sup>Centre for Cell and Molecular Dynamics, Department of Cell and Developmental Biology, University College London, London, WC1E 6DE, UK.

<sup>4</sup>Department of Cell and Developmental Biology, University College London, London, WC1E 6BT, UK.

<sup>§</sup>These authors contributed equally to this study.

#Correspondence: mark.hankins@eye.ox.ac.uk (M.W.H.) and d.whitmore@ucl.ac.uk (D.W.):

Prof. Mark W. Hankins, Nuffield Laboratory of Ophthalmology, Nuffield Department of Clinical Neurosciences, John Radcliffe Hospital, University of Oxford, Oxford, OX3 9DU, UK. Tel.: +44 (0)1865 223380/234782. Fax: +44 (0)1865 234795; Prof. David Whitmore, Centre for Cell and Molecular Dynamics, Department of Cell and Developmental Biology, University College London, London, WC1E 6DE, UK. Tel.: +44 (0)20 7679 6132. Fax: +44 (0)20 7679 7349.

**Running title:** Novel opsins in a model vertebrate

**Keywords:** vertebrate; zebrafish; expression; opsin; diversity; photosensitivity; photopigment

## SUPPLEMENTARY METHODS

### Fish care

Adult wild type zebrafish (AB, AB/TL, and the *per3*-luciferase transgenic line (Kaneko and Cahill 2005; Kaneko et al. 2006)) were raised according to standard protocols used by the University College London (UCL) Fish Facility at 28.5°C on a 14 h light: 10 h dark (14L:10D) cycle. All procedures were conducted in accordance with the United Kingdom Animals (Scientific Procedures) Act 1986 and local Animal Ethics Committee guidelines (UCL).

### Bioluminescence assay

Multiple tissues were dissected from adult *per3*-luciferase transgenic zebrafish (Kaneko and Cahill 2005; Kaneko et al. 2006). Small tissue pieces of brain, fin, gill, gut, liver, muscle, skin, and testis were placed into 96-well dishes, while eye, heart, and pineal and pituitary glands were placed in whole organ culture, and incubated in Leibovitz (L15) media, supplemented with 15% fetal calf serum, penicillin/streptomycin (50 U/mL), gentamicin (50 mg/mL), and 0.5 mM beetle luciferin (Promega). Subsequently, plates were exposed for up to six d to two different lighting regimes: a forward light cycle consisting of 12 h of light, followed by 12 h of dark (LD), or a reverse light cycle comprising 12 h of dark, followed by 12 h of light (DL). Plates were then transferred into constant darkness (DD) for up to seven days. Bioluminescence reporter gene activity was measured during the entire two weeks of the experiment (LD or DL into DD) on a Packard TopCount NXT scintillation counter at 28°C.

### Functional genomics and gene discovery

Database repositories were used that contain nucleotide and amino acid sequence data generated from the sequencing of whole genomes from a variety of different vertebrate organisms: (1) Ensembl (release 76), a joint project between EMBL-EBI and the Wellcome Trust Sanger Institute (<http://www.ensembl.org/index.html>), and (2) NCBI, run by the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). Ensembl is further divided into e!Ensembl that contains post-build genomes that have been fully annotated and Pre!Ensembl (<http://pre.ensembl.org/index.html>) for those pre-build genomes that are in the process of being annotated. Initial gene searches were performed on representative vertebrate species using the e!Ensembl and Pre!Ensembl databases, many of which incorporate sequences submitted to the NCBI database. Initially, the five zebrafish *melanopsin* (*opn4*) nucleotide coding sequences (*opn4m1* [ZFIN: *opn4a*], *opn4m2* [ZFIN: *opn4.1*], *opn4m3* [ZFIN:

*opn4b*], *opn4x1* [ZFIN: *opn4xa*], and *opn4x2* [ZFIN: *opn4b*]) (Davies et al. 2011) and 10 visual opsin genes (*lws1* [ZFIN: *opn1lw1*], *lws2* [ZFIN: *opn1lw2*], *sws1* [ZFIN: *opn1sw1*], *sws2* [ZFIN: *opn1sw2*], *rh2.1* [ZFIN: *opn1mw1*], *rh2.2* [ZFIN: *opn1mw2*], *rh2.3* [ZFIN: *opn1mw3*], *rh2.4* [ZFIN: *opn1mw4*], *rh1.1* [ZFIN: *rho*], and *rh1.2* [ZFIN: *rhoI*]) (Davies et al. 2012), being at the extreme ends of the opsin superfamily phylogenetic tree (Figures 2 and S1), were used as bait to mine the zebrafish genome (Zv9 assembly) ([http://www.ensembl.org/Danio\\_rerio/Info/Index](http://www.ensembl.org/Danio_rerio/Info/Index)) for opsin-like genes. As well as identifying the original query sequence (100% identity), each subsequent hit (less than 100% identity) was in turn subjected to Basic Local Alignment Search Tool (BLAST) analyses. By utilising two main algorithms under default conditions, namely the nucleotide blast (nblast) program ([https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE\\_TYPE=BlastSearch&LINK\\_LOC=blasthome](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome)) and the translated blast (tblastx) program ([https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=tblastx&PAGE\\_TYPE=BlastSearch&LINK\\_LOC=blasthome](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=tblastx&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome)), each BLAST hit was putatively assigned to a particular opsin class. This genomics approach was repeated a number of times using the BLAST results from one gene mining experiment as bait for subsequent bioinformatics searches. Thus, with every BLAST search, the resultant opsin sequences became more closely related in evolutionary terms. In addition, this convergent pipeline was able to identify a full complement of 42 opsin genes present in the zebrafish genome (see Figures 2 and S1 for GenBank Accession Numbers). To evaluate the presence or absence of pigment gene orthologues of the opsin genes discovered in the zebrafish genome, each opsin sequence was used, by using a nucleotide blast pipeline, as bait to mine the genomes of a number of vertebrate genomes that are currently listed in e!Ensembl, Pre!Ensembl, and NCBI databases. In addition to zebrafish, other key species were chosen as representatives of each major gnathostome class. These include bony fishes (teleosts) (e.g. *Takifugu rubripes* (Japanese pufferfish) (current assembly FUGU version 4) ([http://asia.ensembl.org/Takifugu\\_rubripes/Info/Index](http://asia.ensembl.org/Takifugu_rubripes/Info/Index))); amphibians (e.g. *Xenopus tropicalis* (African clawed frog) (current assembly JGI version 4.2) ([http://asia.ensembl.org/Xenopus\\_tropicalis/Info/Index](http://asia.ensembl.org/Xenopus_tropicalis/Info/Index))); reptiles (e.g. *Anolis carolinensis* (green anole) (current assembly AnoCar version 2) ([http://asia.ensembl.org/Anolis\\_carolinensis/Info/Index](http://asia.ensembl.org/Anolis_carolinensis/Info/Index))); birds (e.g. *Gallus gallus* (chicken) (current assembly Galgal version 4) ([http://asia.ensembl.org/Gallus\\_gallus/Info/Index](http://asia.ensembl.org/Gallus_gallus/Info/Index))); monotremes (e.g. *Ornithorhynchus anatinus* (duck-billed platypus) (current assembly OANA version 5) ([http://asia.ensembl.org/Ornithorhynchus\\_anatinus/Info/Index](http://asia.ensembl.org/Ornithorhynchus_anatinus/Info/Index))); marsupials (e.g. *Monodelphis domestica* (opossum) (current assembly MonDom version 5) ([http://asia.ensembl.org/Monodelphis\\_domestica/Info/Index](http://asia.ensembl.org/Monodelphis_domestica/Info/Index))), and eutherian mammals (e.g. *Homo sapiens* (human) (current assembly Genome Reference Consortium Human

version 38 (GRCh38)) ([http://asia.ensembl.org/Homo\\_sapiens/Info/Index](http://asia.ensembl.org/Homo_sapiens/Info/Index)) (see Figure S1 for GenBank Accession Numbers).

### **Phylogenetic analysis**

A codon-matched nucleotide sequence alignment of all 42 zebrafish opsin coding regions was generated by ClustalW (Higgins et al. 1996) and manually manipulated to refine the accuracy of cross-species comparison. These sequences were compared to visual and non-visual pigment gene sequences derived from a variety of gnathostome vertebrates ranging from teleosts to primates. Human *GPR21* (GenBank Accession Number NM005294) and *GPR52* (GenBank Accession Number NM005684) were included as outgroups given their close phylogenetic relationships to the opsin superfamily. A Neighbor-Joining phylogenetic tree (Saitou and Nei 1987) was generated by using the MEGA Version 5 computer package (Tamura et al. 2011) and applying a Maximum Composite Likelihood (MCL) model algorithm (Tamura and Nei 1993) with a homogenous pattern of nucleotide substitution among lineages, uniform rates, pairwise deletion and bootstrapping with 1000 replicates to assess the degree of support for internal branching (expressed as a percentage). A reduced phylogenetic tree was also generated in an identical way except that only coding sequences derived from zebrafish and human were used.

### **NanoString nCounter assay**

The same tissues (outlined above) were dissected from adult AB and AB/TL zebrafish strains between ZT3 and ZT6 (where ZT0 is lights on). Tissue samples were isolated ( $n \geq 3$ , where each sample ( $n$ ) contained 1 or more whole or tissue pieces from multiple animals), homogenized in TRIzol Reagent (Life Technologies) and stored at  $-20^{\circ}\text{C}$ . Total RNA was extracted following the manufacturer's instructions (Life Technologies) and gene-specific probes were designed to the coding region of all 42 opsins and five clock components by the manufacturer (NanoString Technologies). In addition, probes were designed to three reference genes, namely *eukaryotic translation elongation factor 1 alpha 1, like 1* (*eef1a1l1*), *ribosomal protein L13a* (*rpl13a*), and *ubiquitin B* (*ubb*). RNA samples were diluted to the appropriate concentration, and the NanoString nCounter assay was performed at Scientific Support Services (UCL Genomics). The data was first normalized to internal positive and negative controls for each lane using a medium level of stringency. Subsequently, opsin and clock gene expression levels were normalized again to the geometric mean (Vandesompele et al. 2002) of the reference genes for each tissue studied. The hierarchical cluster analyses were performed in MATLAB (Release 2012a), and the bar graphs were produced in Microsoft Excel 2010.

### **RNA *in situ* hybridization**

Whole adult zebrafish eyes and brains were collected at ZT3 and immediately fixed in 4% paraformaldehyde (PFA) (Pierce) dissolved in phosphate-buffered saline (PBS) containing NaCl (140 mM), KCl (3 mM), and PO<sub>4</sub> buffer (10 mM) (pH 7.0) (Invitrogen) for 12-24 h. Fixed tissues were cryoprotected in sucrose (30%) for 48 h before embedding in optimum cutting temperature (OCT) medium (VWR) and freezing at -80°C. Cryostat sections of zebrafish tissues were generated at a thickness of 18 µm and carefully placed on Superfrost Plus microscope slides (VWR). RNA *in situ* hybridization on adult zebrafish eye and brain sections was performed as previously described (Davies et al. 2011). Briefly, antisense RNA probes were generated by cloning full-length novel opsin coding sequences into pBluescript II SK plasmid (Stratagene) via *EcoRI* and *SaII* restriction enzyme sites, linearized and used as template in the SP6/T7 DIG RNA Labelling Kit (Roche), as described (Davies et al. 2011). Slides holding the zebrafish tissues were incubated at room temperature (RT) and dried, before permeabilization with proteinase K (1 µg/ml) (Sigma) for 5 min and acetylation for 10 min, both treatments being performed at RT. Treated slides were subjected to prehybridization with ULTRAhyb® Ultrasensitive Hybridization Buffer (Ambion) at RT for at least 1 h. Riboprobes labelled with digoxigenin (DIG) were added to preheated ULTRAhyb® Ultrasensitive Hybridization Buffer (Ambion), prior to addition to the prehybridized slides. All slides supplemented with riboprobes were incubated in a humidified chamber overnight at 68°C. After at least 12 h, riboprobes were removed by washing the slides twice at 72°C in a saline-sodium citrate (SSC) buffer (0.2X) (Invitrogen). Subsequent, the slides were blocked for 1 h in PBS supplemented with 10% heat-inactivated goat serum (Sigma) at RT. Post-blocking, the slides were incubated overnight at 4°C with an anti-DIG antibody conjugated to alkaline phosphatase (AP) (Roche). The detection of hybridized probes was subsequently performed using the BCIP/NBT AP Substrate Kit IV (Vector Laboratories) according to the manufacturer's instructions. Glycergel Mounting Medium (Dako) was used to mount the slides, prior to imaging with an automated SCN400 Slide Scanner (Leica) at the IQPath Facility (UCL Institute of Neurology).

### **UV-visible spectroscopy of recombinant novel opsin proteins**

Each of the new opsin full-length coding sequences was cloned into the pMT4 mammalian expression vector via *EcoRI* and *SaII* restriction enzyme sites as previously described (Davies et al. 2011). Each zebrafish novel opsin plasmid (210 µg DNA in total) was mixed with GeneJuice transfection reagent at a DNA to GeneJuice ratio of 3 to 1, according to manufacturer's instructions (Novagen), and used to transfect 12 large tissue culture plates (140 mm)

that were preseeded with a nearly confluent monolayer of human embryonic kidney (HEK293T) cells. The transfected cells were harvested after 48 h, pelleted, and washed with PBS four times to remove all dead cells and any remaining traces of cell culture media. The new pigments were repeatedly regenerated with excess 11-*cis* retinal or all-*trans* retinal (Sigma) under dark conditions. Cells were lysed with dodecyl-maltoside (DDM) detergent (1%), in the presence of the protease inhibitor phenylmethanesulphonylfluoride (PMSF) (20 mg/ml), to release membrane fractions containing the novel opsin proteins. Lysate was then subjected to affinity chromatography over a CNBr-activated Sepharose-binding column coupled to an anti-1D4 monoclonal antibody (Molday and MacKenzie 1983) as previously described (Davies et al. 2011). Triplicate absorbance spectra were recorded in complete darkness using a Shimadzu UV-visible spectrophotometer (UV-2550) (Shimadzu), before treatment with either exposure to light for 1 h (to bleach UV-sensitive pigments with a peak sensitivity value, the lambda max ( $\lambda_{\max}$ ), greater than 400 nm) or HCl acid (27.3 mM) for 30 min (to denature pigments with a  $\lambda_{\max}$  less than 400 nm). A difference spectrum was calculated by subtracting the treated spectral trace from that measured in the dark, which was then fitted to a modified Govardovskii rhodopsin A<sub>1</sub>-template (Govardovskii et al. 2000) using Microsoft Excel to determine the  $\lambda_{\max}$  as previously shown (Davies et al. 2007; Davies et al. 2009; Davies et al. 2011). For the biphasic difference spectrum of the Opn7d pigment regenerated with all-*trans* retinal, the  $\lambda_{\max}$  value for each peak was determined separately and followed the same analytical procedure as for all monophasic pigments. The two partial traces were then merged together to generate the overall biphasic spectrum.

### **Expression of novel opsins for electrophysiology**

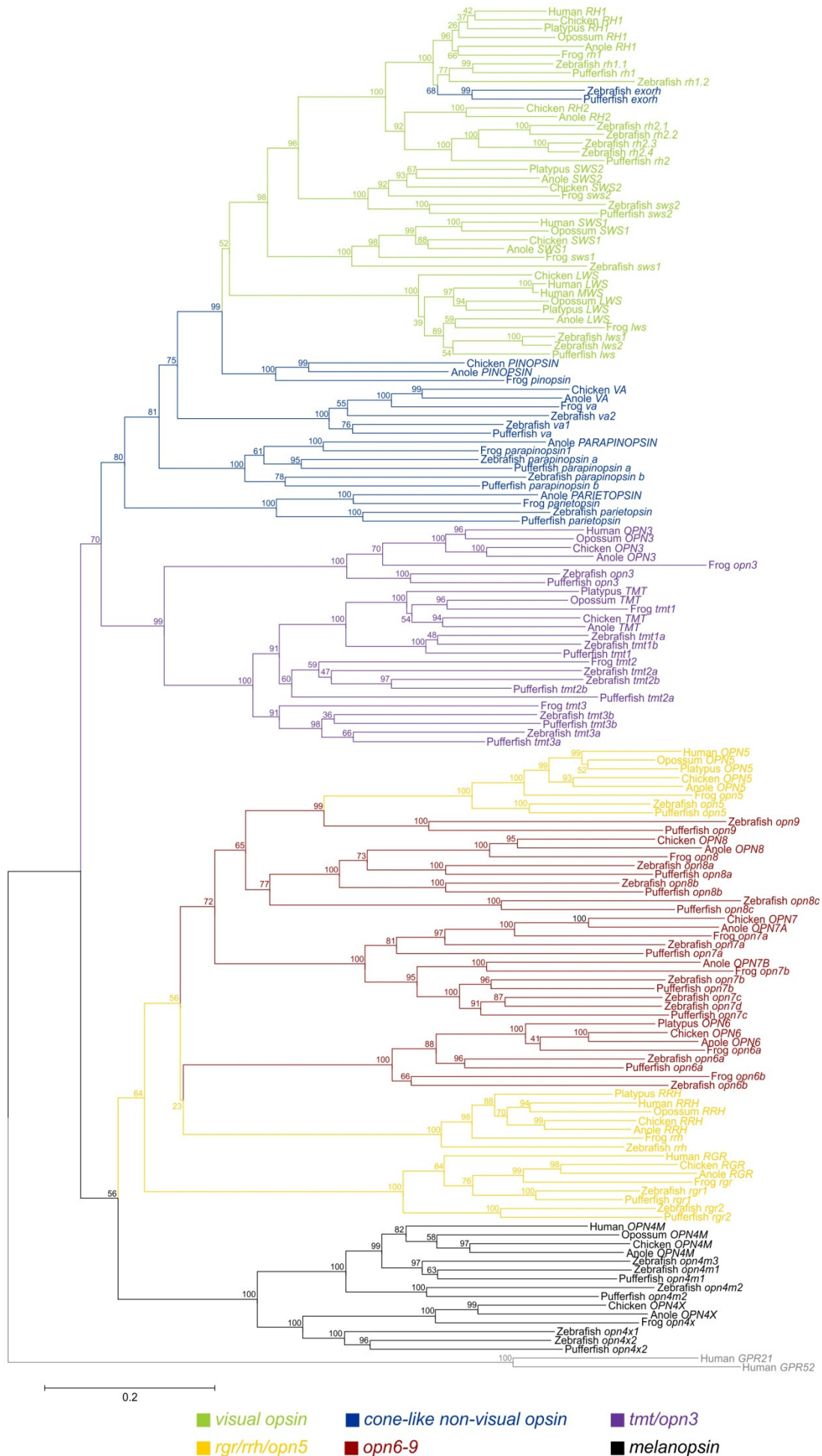
Neuro-2a cells, a neuroblastoma cell line, were grown in 30 mm tissue culture dishes (VWR) containing Dulbecco's Modified Eagle's Medium (DMEM) (Sigma) with the addition of 10% foetal bovine serum (FBS) (Life Technologies), 2 mM L-glutamine (Sigma), and 1% (v/v) penicillin/streptomycin (Sigma). Cells were incubated at 37°C in a humidified chamber, with a constant flow of 5% CO<sub>2</sub>. Fresh media was added every 48–60 h and cells reaching confluence were passaged. In preparation for transfection, cells were trypsinized, diluted, and cultured at  $1 \times 10^5$  cells per ml. Transient transfections were performed 24 h after passaging, where new opsin pMT4 constructs (see above) were expressed in Neuro-2a cells by using GeneJuice reagent (Novagen) according to manufacturer's instructions. For each transfection, a control pSIREN-DNR-DsRed-Express vector (Clontech) was co-transfected with each novel opsin plasmid to allow for identification of cells expressing a particular pigment gene to be analyzed by electrophysiology. Cell differentiation was induced with the addition of retinoic acid (20  $\mu$ M) (Tocris) 24 h after transfection.

Transfected differentiated cells were maintained in complete darkness for 48–60 h before electrophysiological analysis was conducted.

### **Whole-cell electrophysiology**

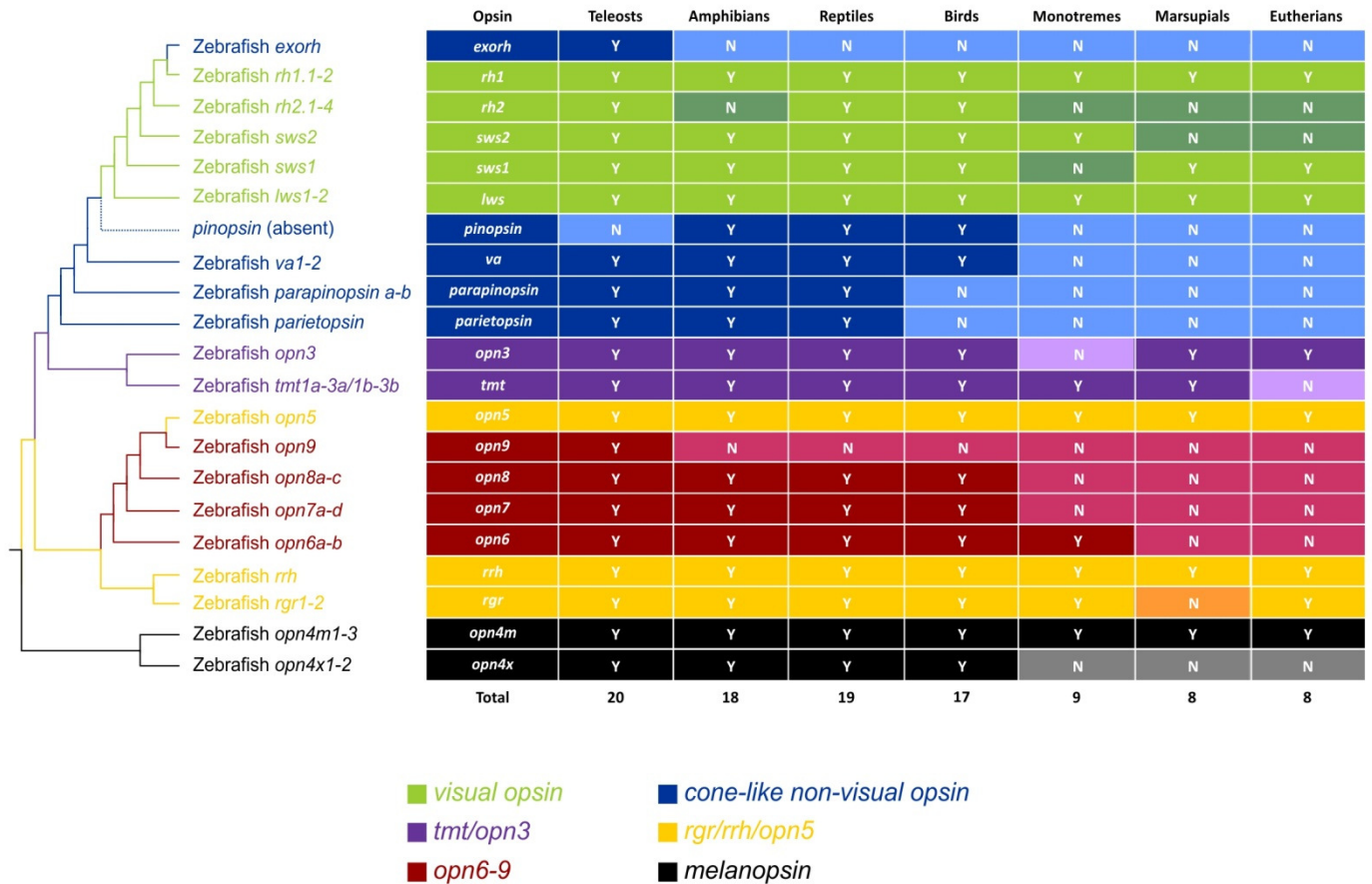
Transfected cells were identified by using a Zeiss Axioskop FS2 microscope to monitor DsRed fluorescence. The perfusion solution, containing NaCl (140 mM), KCl (4 mM), MgCl<sub>2</sub> (1 mM), CaCl<sub>2</sub> (2 mM), glucose (5 mM), and HEPES (10 mM) (pH 7.3–7.4, 24°C), was supplemented with 9-*cis* retinal (Sigma, UK) or all-*trans* retinal (Sigma, UK), both at 20 μM, or no retinal as a control. Cells were maintained in darkness for at least 1 h prior to recording. It should be noted that 11-*cis* retinal (the native chromophore most commonly used) and 9-*cis* retinal (an analogue that is commercially available) isomers are functionally similar in their ability to activate photopigments (Melyan et al. 2005). However, when interchanged, the photopigments bound to 9-*cis* retinal present with a  $\lambda_{\max}$  that is blue-wavelength shifted compared to photopigments that interact with 11-*cis* retinal. As large amounts of chromophore are required in the electrophysiology assays, 9-*cis* retinal was used, whereas 11-*cis* retinal was adopted for UV-vis spectrophotometric experiments. Whole-cell patch-clamp recordings were measured using pipettes containing a solution of KCl (140 mM), NaCl (10 mM), MgCl<sub>2</sub> (1 mM), HEPES (10 mM), and EGTA (10 mM), where the pH and osmolarity were adjusted to 7.3–7.4 (with KOH) and  $285 \pm 5$  mosmol.l<sup>-1</sup>, respectively. Open pipette resistance and access resistance during recordings were 3.5–5 MΩ and <20 MΩ, respectively. Neuro-2a cells were voltage-clamped at holding potentials of -50 mV and currents were recorded using Axopatch 200B (Axon Instruments), where recordings were subsequently filtered (1 kHz) and sampled (20 kHz). Specific wavelengths of light were produced by a slit monochromator using a Cairn Optoscan Xenon Arc lamp source. All light stimuli lasted 10 s with a half-bandwidth of 20 nm. Irradiance levels were determined using an optical power meter (Macam Photometrics), which were in turn transformed into photon flux values. Clampfit (Axon Instruments) was used to determine each response magnitude as defined by measuring the peak sustained. All statistical assessments were performed by applying multiple, independent Student t-tests (two-tailed distribution with two-sample unequal variance (heteroscedastic) parameters) between light-evoked currents generated in the presence of either 9-*cis* retinal or all-*trans* retinal compared to no retinal controls.

SUPPLEMENTARY FIGURES

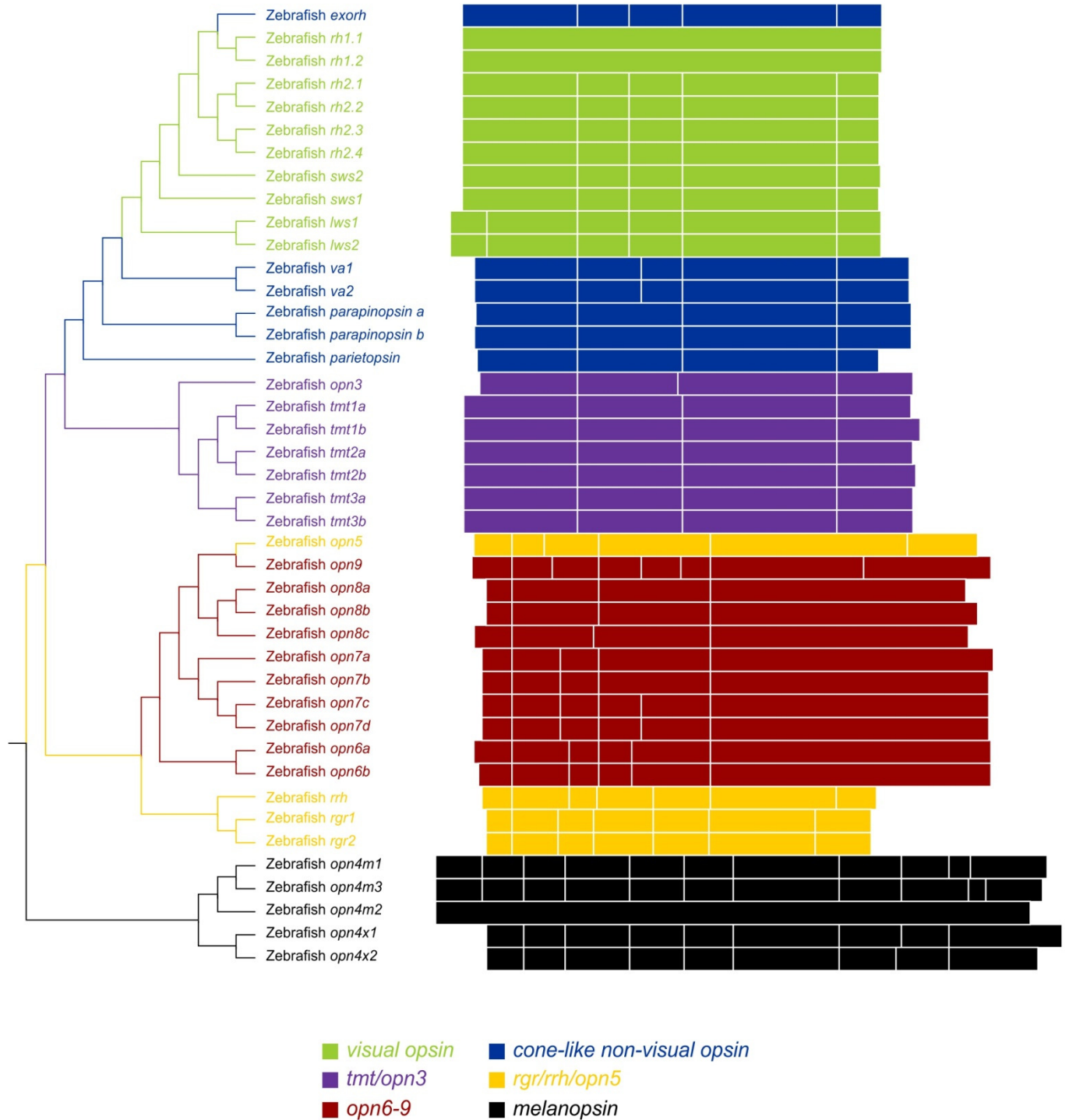




**Figure S1.** Expanded phylogenetic analysis of vertebrate pigment genes, showing the presence of 42 opsin sequences in the zebrafish (*Danio rerio*), in particular the identification of 10 novel opsins, and their conservation in other representative gnathostomes. Human *GPR21* (GenBank Accession Number NM005294) and *GPR52* (GenBank Accession Number NM005684) were included as outgroups (grey). The scale bar indicates the number of nucleotide substitutions per site and the degree of internal branching is expressed as a percentage at each node. The opsin sequences, including their Genbank Accession Numbers, used for generating the tree are as follows: KT008391–KT008432. Opsin classes are subdivided into 6 main groups that contain visual opsins (green), cone-like non-visual opsins (blue), *opn3/tmt* opsins (purple), *rgr/rrh/opn5* opsins (yellow), *opn4* opsins (black), and the new opsins (red). Specifically, these are subdivided into *extra-ocular rhodopsin (exorh)*, *rhodopsin-like-1 (rh1)* rod opsins (*rh1.1* [ZFIN: *rho*] and *rh1.2* [ZFIN: *rho*]), *rhodopsin-like-2 (rh2)* cone opsins (*rh2.1* [ZFIN: *opn1mw1*], *rh2.2* [ZFIN: *opn1mw2*], *rh2.3* [ZFIN: *opn1mw3*], and *rh2.4* [ZFIN: *opn1mw4*]), *short-wavelength-sensitive-2 (sws2)* cone opsin [ZFIN: *opn1sw2*], *short-wavelength-sensitive-1 (sws1)* cone opsin [ZFIN: *opn1sw1*], *long-wavelength-sensitive (lws)* cone opsins (*lws1* [ZFIN: *opn1lw1*] and *lws2* [ZFIN: *opn1lw2*]), *vertebrate ancient (va) long* opsins (*va1* [ZFIN: *valopa*] and *va2* [ZFIN: *valopb*]), *parapinopsins (parapinopsin a and parapinopsin b)*, *parietopsin*, *encephalopsin/panopsin (opn3)*, *teleost multiple tissue (tmt)* opsins (*tmt1a* [ZFIN: *tmtops2a*], *tmt1b* [ZFIN: *tmtops2b*], *tmt2a* [ZFIN: *tmtops3b*], *tmt2b* [ZFIN: *tmtops3a*], *tmt3a* [ZFIN: *tmtopsa*], and *tmt3b* [ZFIN: *tmtopsb*]), *neuropsin (opn5)*, novel opsins (*opn6*, *opn7*, *opn8*, and *opn9*), *retinal pigment epithelium-derived rhodopsin homolog (rrh)*; also known as *peropsin*), *retinal pigment epithelium (RPE) retinal G protein receptors (rgr)* (*rgr1* [ZFIN: *rgra*] and *rgr2* [ZFIN: *rgrb*]), *mammalian-like melanopsins (opn4m)* (*opn4m1* [ZFIN: *opn4a*], *opn4m2* [ZFIN: *opn4.1*], and *opn4m3* [ZFIN: *opn4b*]), and *Xenopus-like melanopsins (opn4x)* (*opn4x1* [ZFIN: *opn4xa*] and *opn4x2* [ZFIN: *opn4b*]).



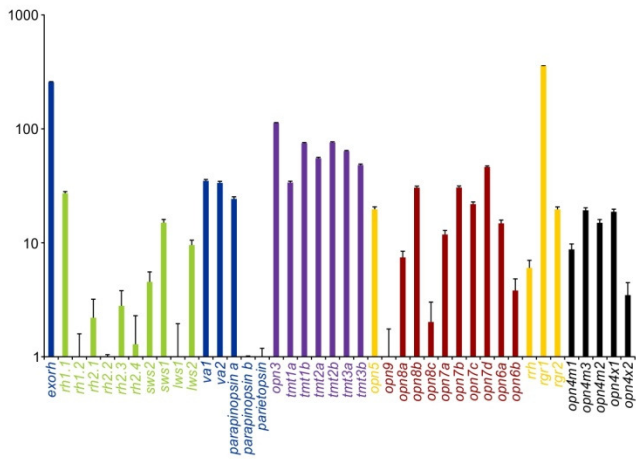
**Figure S2.** A summary table showing the evolution of all opsin classes in the gnathostomes (teleosts to eutherians). Megagenomic analyses, aligned to a reduced complement of zebrafish opsins. Neighbor-Joining (NJ) phylogenetic tree, demonstrate that pigments were significantly lost throughout vertebrate evolution, especially at the base of the mammalian lineage. See Figures 2 and S1 for colour coding and gene name abbreviations. The presence or absence of a particular gene class is indicated by a Yes (Y) or No (N) designation.



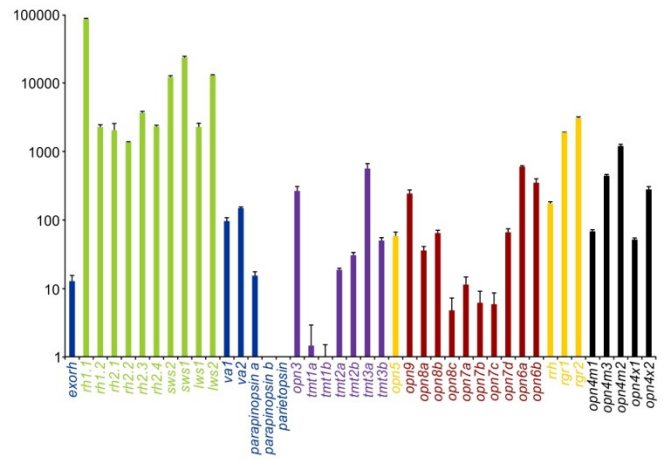
**Figure S3.** A comparison of genomic structures identified within the zebrafish opsin gene repertoire, aligned to an Neighbor-Joining (NJ) phylogenetic tree. Data demonstrate that some intron locations (white lines) are invariably conserved through a large number of opsin subclasses. For example, intron 1 (visual opsins (equivalent to intron 2 in the two *lws* genes and absent in the two intronless *rh1* genes), cone-like non-visual opsins (the two *va* genes, the two *parapinopsin* genes, and *parietopsin*), *opn3*, and all six *tmt* genes); intron 3 (visual opsins (equivalent to intron 4 in the two *lws* genes and absent in the two intronless *rh1* genes), cone-like non-visual opsins (the two *va* genes and the equivalent to intron 2 in both the *parapinopsin* (a and b) and *parietopsin* genes), and equivalent to intron 2 in all six *tmt* genes); and intron 4 (visual opsins (equivalent to intron 5 in the two *lws* genes and absent in the two

intronless *rh1* genes), cone-like non-visual opsins (the two *va* genes and the equivalent to intron 3 in the *parapinopsins* (a and b), *parietopsin*, *opn3*, and all six *tmt* genes). Within the novel opsin group, and compared to *opn5*, two introns are invariably conserved: intron 1 and intron 4 (equivalent to intron 5 in the *opn6* class, intron 4 (*opn7a* and *opn7b*), or intron 5 (*opn7c* and *opn7d*) in the *opn7* class, intron 3 in the *opn8* class, and intron 6 in the *opn9* class. Given the distinct pattern of exon-intron boundaries in all of the new opsin genes compared to their closest sister clades (i.e. *opn5*, *rrh*, and *rgr*), the new opsin genes (i.e. *opn6*, *opn7*, *opn8*, and *opn9*) are discrete opsin classes in their own right and not simply *opn5*-like. See Figures 2 and S1 for colour coding and gene name abbreviations.

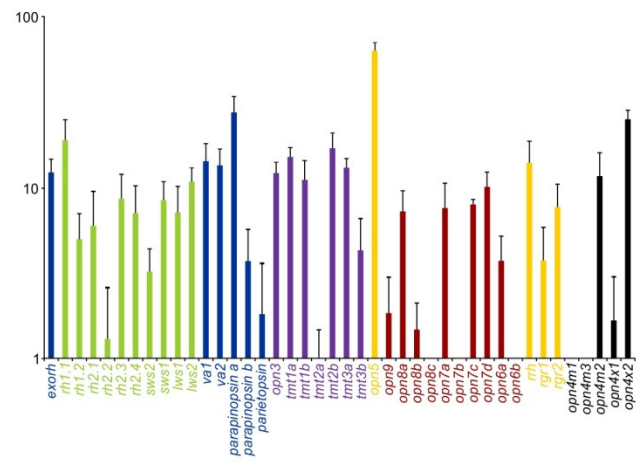
(A) Brain



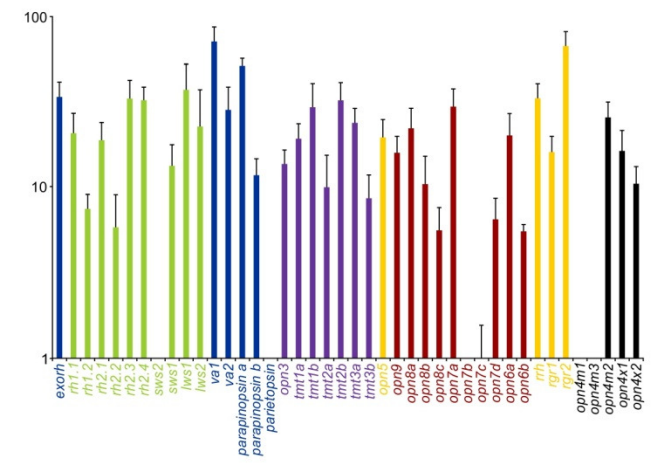
(B) Eye



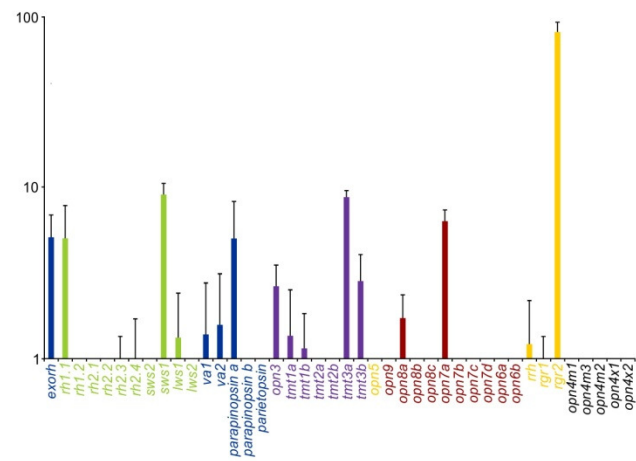
(C) Fin



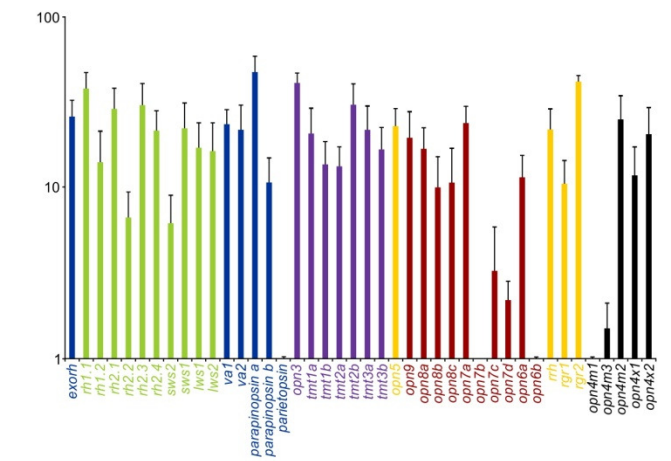
(D) Gill



(E) Gut

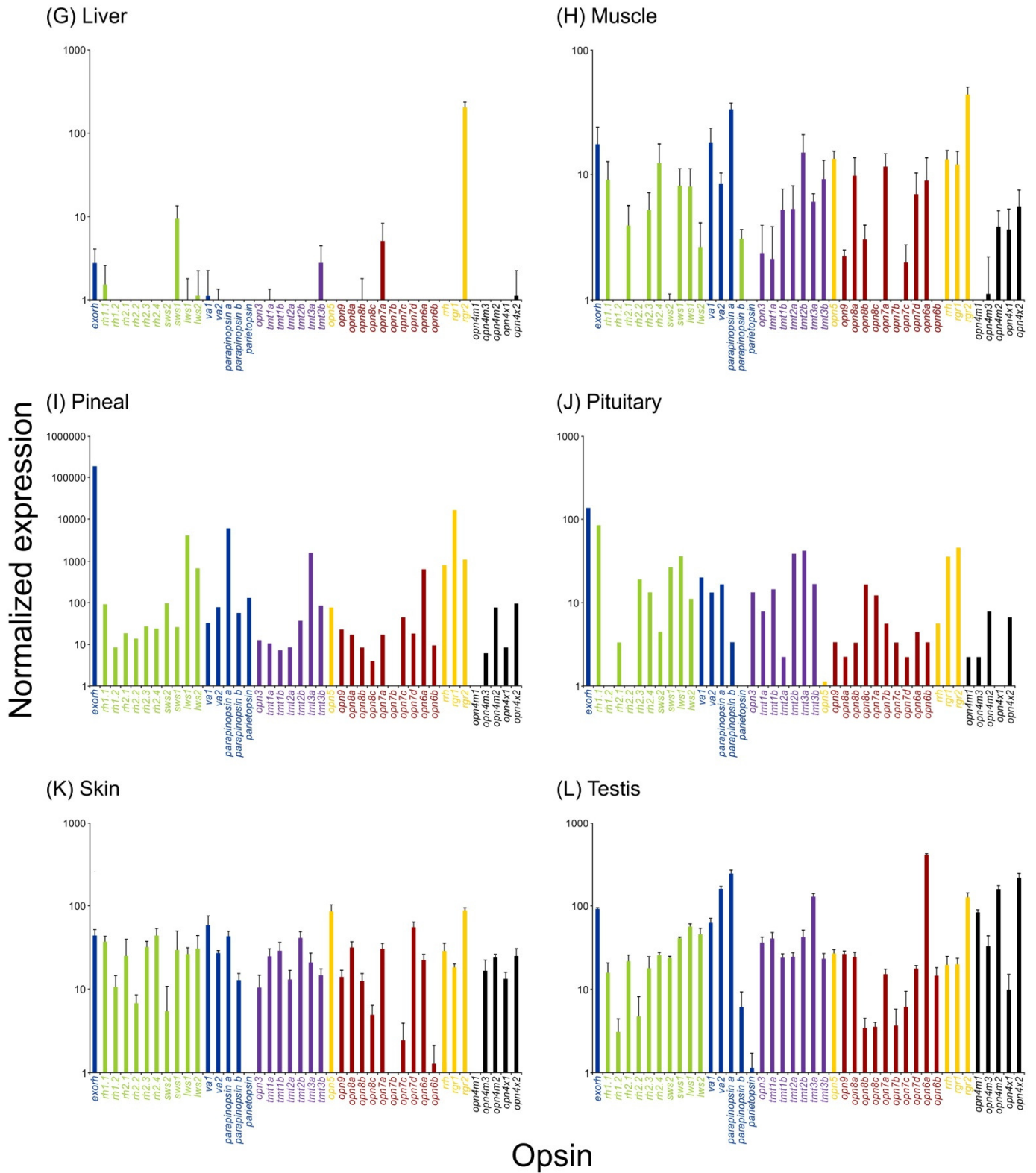


(F) Heart

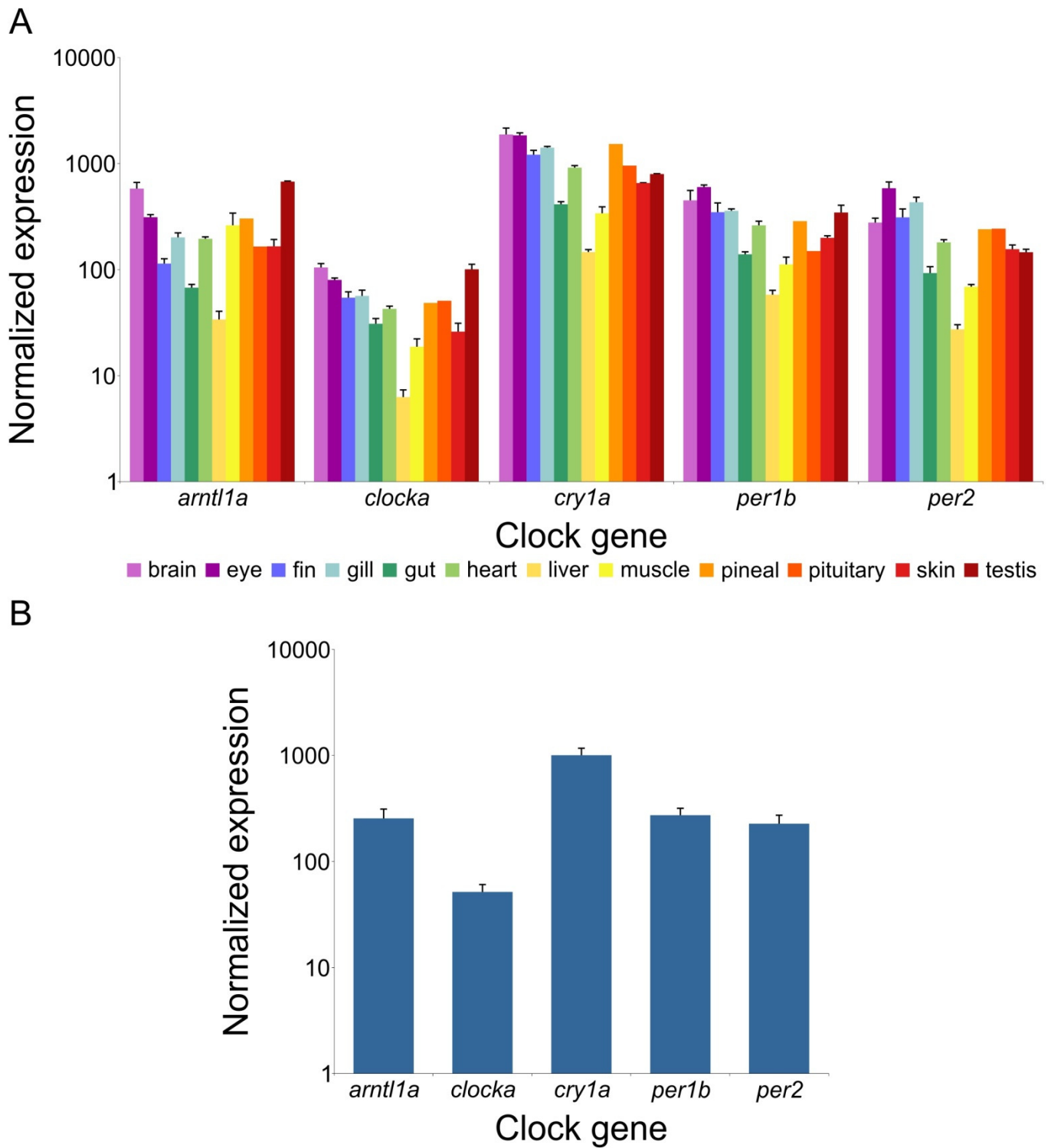


Normalized expression

Opsin



**Figure S4.** Gene expression profiling of zebrafish opsin transcripts using the NanoString nCounter assay. Adult zebrafish tissues ((A) brain; (B) eye; (C) fin; (D) gill; (E) gut; (F) heart; (G) liver; (H) muscle; (I) pineal; (J) pituitary; (K) skin, and (L) testis) were dissected and maintained as described in Methods. Relative opsin gene expression was measured as described in Methods and is presented as bar charts on a  $\log_{10}$  scale. Each column represents a different opsin, which is color-coded as in Figures 2 and S1.



**Figure S5.** Clock gene expression in adult zebrafish tissues. (A) Adult zebrafish tissues (brain, eye, fin, gill, gut, heart, liver, muscle, pineal, pituitary, skin, and testis) were dissected and maintained as described in Methods. The relative transcript levels of *arntl1a*, *clocka*, *cry1a*, *per1b*, and *per2* were measured using the NanoString nCounter system as described in Methods and is presented as bar charts on a log<sub>10</sub> scale. (B) The data show that *cry1a* is the most abundantly expressed and *clocka* is the least abundantly expressed clock genes across all tissues sampled at ZT3.

**SUPPLEMENTARY TABLE**

Structural or functional attribute	Novel opsin pigments										Typical visual opsin	Typical non-visual opsin
	Opn6a	Opn6b	Opn7a	Opn7b	Opn7c	Opn7d	Opn8a	Opn8b	Opn8c	Opn9	e.g. Rh1.1	e.g. Opn4m1
Seven transmembrane (TM) domains	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Disulphide bridge (Cys110)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Counterion (Glu113 or Tyr113)	Asp	Asp	Tyr	Tyr	Tyr	Tyr	Tyr	Tyr	Tyr	Tyr	Glu	Tyr
Glu/Asp-Arg-Tyr (E/DRY) motif (134-136)	TRF	IRY	VCC	VCW	VCC	VCC	VRY	VRF	IRY	DRC	ERW	DRY
Alternative counterion (Glu181)	Arg	Arg	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu
Disulphide bridge (Cys187)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Chromophore binding site (Lys296)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NPxxY(x) <sub>5,6</sub> F/Y motif (302-313)	Yes (NPLIYFGLSSKF)	Yes (NPFYIGMSSKF)	Yes (NPLIYLLKPNF)	Yes (NPMVYLVFKPSF)	Yes (NPMVYLLFKPNF)	Yes (NPVVYLVFKPNF)	Yes (NPLVYVFRKSF)	Yes (NPFYIYFSKTF)	Yes (NPFYFYFNKAF)	Yes (NPFYQIMNRDF)	Yes (NPCIYICMNKQF)	Yes (NPIIYAITHPKY)
Retinal chromophore (11- <i>cis</i> retinal or all- <i>trans</i> retinal)	Both	11- <i>cis</i> retinal	Both	Both	n.d.	all- <i>trans</i> retinal	Both	n.d.	Both	n.d.	11- <i>cis</i> retinal	Both
Chromophore valency	Bistable	Monostable	Bistable	Bistable	n.d.	Monostable	Bistable	n.d.	Bistable	n.d.	Monostable	Bistable
Spectral phase (in the dark)	Monophasic (510 nm)	n.d.	n.d.	n.d.	n.d.	Biphasic (369/508 nm)	n.d.	n.d.	Monophasic (375 nm)	n.d.	Monophasic	Monophasic

**Table S1.** Conservation of residues critical for structure and function of zebrafish photopigments.



## SUPPLEMENTARY REFERENCES

- Davies WI, Collin SP, Hunt DM. 2012. Molecular ecology and adaptation of visual photopigments in craniates. *Mol Ecol* 21: 3121-3158.
- Davies WI, Zheng L, Hughes S, Tamai TK, Turton M, Halford S, Foster RG, Whitmore D, Hankins MW. 2011. Functional diversity of melanopsins and their global expression in the teleost retina. *Cell Mol Life Sci* 68: 4115-4132.
- Davies WL, Carvalho LS, Tay BH, Brenner S, Hunt DM, Venkatesh B. 2009. Into the blue: gene duplication and loss underlie color vision adaptations in a deep-sea chimaera, the elephant shark *Callorhynchus milii*. *Genome Res* 19: 415-426.
- Davies WL, Cowing JA, Carvalho LS, Potter IC, Trezise AE, Hunt DM, Collin SP. 2007. Functional characterization, tuning, and regulation of visual pigment gene expression in an anadromous lamprey. *FASEB J* 21: 2713-2724.
- Govardovskii VI, Fyhrquist N, Reuter T, Kuzmin DG, Donner K. 2000. In search of the visual pigment template. *Vis Neurosci* 17: 509-528.
- Higgins DG, Thompson JD, Gibson TJ. 1996. Using CLUSTAL for multiple sequence alignments. *Methods Enzymol* 266: 383-402.
- Kaneko M, Cahill GM. 2005. Light-dependent development of circadian gene expression in transgenic zebrafish. *PLoS Biol* 3: e34.
- Kaneko M, Hernandez-Borsetti N, Cahill GM. 2006. Diversity of zebrafish peripheral oscillators revealed by luciferase reporting. *Proc Natl Acad Sci U S A* 103: 14614-14619.
- Melyan Z, Tarttelin EE, Bellingham J, Lucas RJ, Hankins MW. 2005. Addition of human melanopsin renders mammalian cells photoresponsive. *Nature* 433: 741-745.
- Molday RS, MacKenzie D. 1983. Monoclonal antibodies to rhodopsin: characterization, cross-reactivity, and application as structural probes. *Biochemistry* 22: 653-660.
- Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4: 406-425.
- Tamura K, Nei M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 10: 512-526.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28: 2731-2739.

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