

**Conservation, duplication and divergence of
five opsin genes in insect evolution**

Roberto Feuda ^{1,2}, Ferdinand Marletaz ¹, Michael A. Bentley ¹ and Peter W H Holland ^{1*}

¹Department of Zoology, University of Oxford, South Parks Road, Oxford, OX1 3PS, UK

²Current address: Division of Biology and Biological Engineering, California Institute of
Technology. Pasadena, California, USA

* Correspondence: peter.holland@zoo.ox.ac.uk

Keywords: Lepidoptera, Arctiidae, butterfly, molecular evolution, adaptive evolution

Abstract

Opsin proteins covalently bind to small molecular chromophores and each protein-chromophore complex is sensitive to particular wavelengths of light. Multiple opsins with different wavelength absorbance peaks are required for colour vision. Comparing opsin responses is challenging at low light levels, explaining why colour vision is often lost in nocturnal species. Here we investigated opsin evolution in 27 phylogenetically diverse insect species including several transitions between photic niches (nocturnal, diurnal and crepuscular). We find widespread conservation of five distinct opsin genes, more than commonly considered. These comprise one c-opsin plus four r-opsins (long wavelength sensitive or LWS, blue sensitive, UV sensitive and the often overlooked Rh7 gene). Several recent opsin gene duplications are also detected. The diversity of opsin genes is consistent with colour vision in diurnal, crepuscular and nocturnal insects. Tests for positive selection in relation to photic niche reveal evidence for adaptive evolution in UV-sensitive opsins in day-flying insects in general, and in LWS opsins of day-flying Lepidoptera specifically.

Introduction

The ability of animals to respond to visual stimuli is essential for many aspects of life. In insects, and more generally in Eumetazoa, this response is mediated primarily by opsins: a set of proteins belonging to the G-protein-coupled receptor superclass. Opsin proteins covalently bind a small molecule chromophore, typically derived from vitamin A, and together the complex is able to react to light. The chromophore undergoes photo-isomerisation in response to a photon of light, inducing a conformational change in the opsin protein and activation of a downstream signalling cascade. Most opsin genes are expressed in photoreceptors, but there are opsins expressed in other tissues suggesting some non-visual functions (Oakley and Speiser 2015; Shichida and Matsuyama 2009; Terakita 2005).

Opsins originated early in metazoan evolution and duplicated to give three major gene families groups in bilaterians: c-opsins (ciliary opsins), r-opsins (rhabdomeric opsins) and RGR/Go opsins (including vertebrate peropsin) (Bellingham, et al. 2003; Fain, et al. 2010; Terakita 2005). The ciliary and rhabdomeric terminology reflects the structure of photoreceptor cells: in the former, the membrane of a cilium is folded to increase surface area for storage of opsin proteins, whereas in the latter the cell surface itself is deeply folded. Visual functions have rarely been found for RGR/Go opsins, although few have been studied (Gühmann, et al. 2015). In general, c-opsins are probably the main visual pigments of vertebrates (including all opsins expressed in ciliary rods and cones of the eye), and r-opsins are used as the principal visual pigments in the light-sensitive membranes (rhabdomeres) of arthropod compound eyes (Porter, et al. 2011). However, this long-assumed distinction between vertebrates and protostome invertebrates has been challenged in recent years and there is evidence for r-opsin expression in vertebrates and c-opsin use in protostomes (Arendt 2003; Arendt, et al. 2004; Passamanek, et al. 2011).

Within the visual r-opsin family of insects, there are three widely-accepted and well-studied paralogues: a long wavelength-sensitive opsin (LWS opsin, peak absorbance 500-600 nm), a blue sensitive opsin (Blue opsin, peak absorbance 400-500 nm) and an ultraviolet-sensitive opsin (UV opsin, peak absorbance 300-400 nm) (Briscoe 2008; Yuan, et al. 2010, Henze and Oakley 2015). In addition to this basic repertoire, an additional insect r-opsin paralogue has been identified in some species (Attardo, et al. 2014; Carulli, et al. 1994; Futahashi, et al. 2015), named Rh7, for which the phylogenetic distribution is unresolved.

These r-opsins are in addition to insect c-opsin, the evolutionary conservation of which is also unclear.

Opsin evolution cannot be understood without also considering the ecological interactions between animals and their environment. Colour vision requires the expression of opsins with different spectral sensitivity, since an animal must compare the responses of two or more opsins tuned to different wavelengths of light (Osorio and Vorobyev 2008). In vertebrates, this is achieved through possession of multiple c-opsins, while insects use the multiple r-opsins described above; the fact that these expanded gene sets evolved by different gene duplications suggests that colour vision evolved independently in vertebrates and insects. To some extent, the diversity of opsin genes present in a genome can give insight into the visual capability of the species. Possession of three visual opsins with different spectral sensitivity implies capacity for trichromatic vision, while a single opsin gene cannot provide colour vision (Kelber and Roth 2006). It might be assumed that ability to see colour would always be a trait favoured over monochrome vision, but there are situations when the selection pressure to retain colour vision might be relaxed. Nocturnal species or those inhabiting low-light environments such as caves tend to show reduced selective constraints on one or more opsin genes, or even gene loss (Tierney, et al. 2015). Light availability seems to impose strong selective pressures on opsin sequence evolution (Jacobs 2009; Kelber and Roth 2006; Osorio and Vorobyev 2008; Veilleux, et al. 2013; Zhao, et al. 2009); for example, the LW and UV-sensitive opsins each show differences in the amino acids subject to strong selection between nocturnal and diurnal fireflies (order Coleoptera (Sander and Hall 2015). Similarly, a change in the spectral tuning of the LW opsin protein was also observed in adaptation to dim-light environment in augochlorine bees (Tierney, et al. 2011).

In this study we took a genome scale approach to examine the diversity of opsin genes in insects. We identified all opsin genes in 27 insect genomes (spanning seven orders); these included lepidopteran species (Ferguson, et al. 2014) chosen to allow study of independent shifts from nocturnal to diurnal activity, such as the butterflies and the Arctiinae or 'tiger moths'. We find a dynamic pattern of gene duplication and gene loss, but overall conservation of five opsin types across insects (LWS, Blue, UV, Rh7, c-opsin). The gene family distribution is consistent with colour vision in diurnal, crepuscular and nocturnal insects. Tests for positive selection in relation to photic niche revealed evidence for adaptive evolution in UV-sensitive opsins in day-flying insects in general, and in LWS opsins of day-flying Lepidoptera specifically.

Material and Methods

Data mining

Genome sequences for 20 insects covering 7 orders were downloaded from Ensembl (Cunningham, et al. 2015), plus the genomes of *Manduca sexta* (<https://www.hgsc.bcm.edu/arthropods/tobacco-hornworm-genome-project>) and *Plutella xylostella* (You, et al. 2013). Each corresponding proteome was mined with blastp using the 449 opsin dataset of Feuda, et al. (2012). Consistent with previous work (Feuda, et al. 2012) sequences with e-value $<10^{-10}$ were retained, since this approach returns a diverse GPCR dataset including all opsins plus other genes, which can then be further analysed. To discriminate opsins from other GPCRs, we used a combination of a sequence similarity and motif analysis; to be retained as opsins, we required a top blastp hit with opsin in Uniprot and/or conservation of a recognizable retinal-binding domain. In addition, we used lower coverage genome sequences for *Polygonia c-album*, *Pararge aegeria*, *Callimorpha dominula*, *Cameraria ohridella* and *Glyptotaelius pellucidus* (Ferguson, et al. 2014) because these allow three independent shifts from nocturnal to diurnal lifestyle within a single order to be analysed; since these genomes are low coverage and not annotated a different strategy was necessary. Reads were assembled using Velvet with k-mer sizes 31 and 41, and searched with tblastn (e-value $<10^{-10}$) using opsins from *Apis mellifera*, *Drosophila melanogaster*, *Manduca sexta*, *Heliconius melpomene*, *Danaus plexippus* and *Tribolium castaneum*. We assembled a final data set of 166 opsin sequences for analysis. Opsin intron/exon structure predictions were generated using a homology-based method with the program Geca (Fawal, et al. 2012). Additionally, to gain a picture of the functional opsin number in each species, we predicted the number of transmembrane structure using Topcons (Tsirigos, et al. 2015); opsin sequences predicted to have either 6 or 7 transmembrane domains were retained as putative functional opsins.

Finally, photic niches (see below and table S1) were inferred integrating personal knowledge of most species and information from the literature (Gentile, et al. 2009; Kawada, et al. 2005; Marinotti, et al. 2013; Orr 1992; Porter and Tschinkel 1987; Rund, et al. 2011; Zimmerman 1992). However, it should be noted that in some species such as *Ae. aegypti*, *An. darlingi* and *At. cephalotes* there can be geographical or seasonal variation in behaviour.

Alignments and phylogeny

Protein sequence alignment was performed in PRANK (Loytynoja and Goldman 2008), which has been shown to outperform other alignments methods with similar data (Loytynoja and Goldman 2008). The alignment was manually curated and indel-rich regions of uncertain alignment removed, and is available from Oxford University Research Data Archive (ORA-Data), under DOI 10.5287/bodleian:st74cq83q. Phylogenetic reconstruction was performed using PhyloBayes 3.3e (Lartillot, et al. 2009) under the GTR- Γ model (the best fitting model for large opsin datasets (Feuda, et al. 2012; Feuda, et al. 2014). Trees were rooted using melatonin receptor (Feuda, et al. 2012; Feuda, et al. 2014; Fredriksson, et al. 2003; Plachetzki, et al. 2010). Unrooted trees were also constructed to exclude potential error caused by using a distant outgroup. For all analyses, two independent runs were performed and convergence monitored using the maxdiff statistics calculated using the bpcomp program for PhyloBayes. Analyses were considered to have converged when maxdiff dropped below 0.3. Results of the analyses of the opsin+outgroup data sets were further tested by performing Maximum Likelihood (ML) analyses under GTR- Γ (Stamatakis 2006).

Positive selection

To test for signatures of positive selection, the ratio of synonymous to non-synonymous substitution rates (d_N/d_S ratio or ω) was estimated in a maximum likelihood framework using the CodeML program of PAML (Yang 2007). Codon alignments for each of the five opsin gene families were generated using PRANK with the codon option (Loytynoja and Goldman 2008; Markova-Raina and Petrov 2011). Ambiguously aligned residues and positions with >60% gaps were removed. To test whether adaptation to diurnal or nocturnal lifestyle was accompanied by detectable adaptive sequence change, two data sets were analyzed. First, we included all insects; since inferring lifestyle of many ancestral nodes was not possible, we assigned ancestral states only when all descendants shared the character, with an undetermined state of character applied to the root. A branch-site model was applied allowing one d_N/d_S ratio for diurnal species, one for nocturnal species, one for crepuscular species or species active in day and night, and one ratio for undetermined ancestral nodes (NSsites=0 and mode=2, gamma with four categories). Second, we focused specifically on Lepidoptera, with an ancestral nocturnal state assumed for the species under study, accompanied by three independent shifts to diurnal activity (butterflies, Tiger moth, *Cameraria*). Different d_N/d_S ratios were allowed for nocturnal and diurnal. The Lepidoptera-

only LWS opsin analysis was performed both including and excluding intron-less duplicated loci. We also undertook analysis to sites under selection associated with the d_N/d_S increase in diurnal opsins by applying a branch-site model to the lepidopteran datasets (NSsites=2 and model=2, gamma with four categories). We considered the sites as candidates for positive selection according to the Bayes Empirical Bayes (BEB) criterion. Sites deduced to be under selection were plotted onto 2D transmembrane topology inferred using Protter (Omasits, et al. 2013).

Results and discussion

Conservation of five opsins in insect evolution

Phylogenetic analysis of all opsin sequences from 27 insect species confirms there are five major clades of insect opsin: one c-opsin and four r-opsins (LWS, UV, Blue and the enigmatic Rh7 opsin). All five opsins are very widespread across insect diversity (Figures 1, 2; Figure S1, S2 and S3; Table S1). Applying Bayesian phylogenetic analysis to an opsin plus outgroup data set, we find that UV and Blue opsins are sister clades (PP=1), which in turn form a monophyletic group with Rh7 opsins (PP=0.98) and finally LWS opsins are the sister group to other r-opsins (PP=1) (Figure 1, S1). An unrooted data set (Figure S2) and a maximum likelihood analysis (Figure S3) recover essentially the same topology. Most studies of opsin diversity in specific insect orders have focused on just three opsins (LWS, UV and Blue); our genome scale approach expands insect opsin diversity, as was also noted recently by Futahashi, et al. (2015). The few studies undertaken suggest that the c-opsins in insects are not primarily visual, but are expressed in the brain and probably involved in circadian rhythms (Velarde, et al. 2005). A dual role in vision and circadian entrainment has also been found for an insect r-opsin in an orthopteran insect (Komada et al. 2015). It is more difficult to assign a function to the Rh7 opsin, but we note that it is ancient in insects and very widespread; this conservation implies it certainly has functional relevance. In *Drosophila*, the gene is expressed in brain as well as retina (Chintapalli, et al. 2007; Graveley, et al. 2011; Kistenpfennig 2012; Papatsenko, et al. 2001), while there is low level expression in adult dragonfly eyes (Futahashi, et al. 2015). We suggest future studies of visual evolution in insects should not overlook this gene.

Combining the opsin phylogenetic analysis (Figure 1) with the species tree (Figure 2) implies that all four distinct r-opsin paralogues are ancient and date to early in insect evolution, or even prior to insect radiation (Misof, et al. 2014). Together with recent data from dragonflies, order Odonata (Futahashi, et al. 2015), this is consistent with the prediction of Briscoe and Chittka 2001) and Henze and Oakley (2015) that visual opsin diversity evolved very early in insect evolution and predates the origin of flowering plants ~150 million years ago. Such a conclusion would imply that the complex colour vision abilities of insects did not arise to adapt to the reflectance spectra of flowering plants, but rather the colours of fruit and flowers may have evolved to be more discernable to the pre-existing colour vision capabilities of insects (Briscoe and Chittka 2001; Marshall and Arikawa 2014; Osorio and Vorobyev 2008).

Gene duplication and gene loss in insect evolution

Recent work on dragonflies, order Odonata, a relatively basal clade within winged insects, identified two additional opsin genes that were either not found or not widespread in our study: RGR/Go and Arthropsin (Futahashi, et al. 2015). We found one potential RGR/Go gene in the Hemiptera *Acyrtosiphon pisum*, although the sequence is divergent and identification is equivocal (Table S1). RGR/Go genes are found elsewhere in the animal kingdom and in a basal insect clade (Futahashi, et al. 2015) so this clearly represents a gene loss in Holometabola. In most of the other cases where a particular opsin gene is missing from a species in our analysis, we cannot definitively conclude this is a gene loss, since few genome sequencing projects can claim completeness. The exception is absence of c-opsin in *Drosophila*; a particularly striking gene loss and one to which a high degree of confidence can be given.

We also found several intriguing cases of opsin gene duplication within the insects (Figure 2). Perhaps most interesting are the consistent duplications of LWS opsins in mosquitoes (with different numbers between species) and an extensive duplication of the LWS gene in a day-flying Scarlet Tiger moth *Callimorpha dominula* (Arctiidae; 4 genes). We find no cases as dramatic as the massive expansion of opsin gene numbers reported for dragonflies (Futahashi, et al. 2015). The conclusion that the multiple LWS sequences identified in the Scarlet Tiger moth draft genome represent independent genes is further supported by analysis of intron positions (Figure S4): three of the duplicate LWS genes of the *C. dominula* lack introns. It is not known if this reflects intron loss from tandemly duplicated loci or retroposition (integration of a DNA copy of an RNA transcript), although the latter mechanism would be intriguing since retroposed duplicates can only arise from germ-line expressed genes and would normally not copy regulatory sequences (Booth and Holland 2004). We note that intron-less opsin genes have evolved in other taxa, notably crustaceans, cnidarians and cephalopods, with the current understanding that they function as visual photoproteins (Liegertová, et al. 2015; Porter, et al. 2007).

Colour vision in nocturnal insects

The reduced amount of photons at night is thought to make the process of colour vision difficult (Warrant and Dacke 2011). It had long been assumed that nocturnal insects use olfaction rather than vision to find and recognize flowers (Brantjes 1978). However, recent work on nocturnal Lepidoptera, notably Sphingidae or hawkmoths (Kelber, et al. 2003), and on a species of nocturnal Hymenoptera (Warrant 2008), suggests that some

night-flying insects are able to distinguish colours. Based on the distribution of opsin paralogues determined in the present study (Figure 2), we suggest that the majority of nocturnal insects are capable, at least potentially, of discriminating colours. It seems likely from these data that nocturnal Lepidoptera (e.g. *Plutella xylostella* and *Manduca sexta*) have trichromatic vision (Figure 2). Additionally, a crepuscular coleopteran (*Tribolium castaneum*), a human ectoparasitic louse (*Pediculus humanus*), a nocturnal trichopteran (*Glyptotaelius pellucidus*), some nocturnal Diptera (e.g. *Anopheles gambiae*) and a nocturnal hemipteran (*Rhodnius prolixus*) have at least dichromatic colour vision.

These findings imply that, unlike many nocturnal mammals, nocturnal and crepuscular insects have generally not lost the opsins that are maximally sensitive to high light levels. Although further species sampling would be beneficial, it seems nocturnal and diurnal insects have similar opsin repertoires.

Adaptive evolution of opsins in day-flying Lepidoptera

We asked whether nocturnal or diurnal lifestyles had driven adaptive sequence change in the five opsin genes. We first performed an analysis across all the insect species in the dataset, making no assumptions about ancestral states where these are ambiguous. This revealed an increased global d_N/d_S ratio (ω value), suggestive of adaptive protein sequence change, in the UV opsin gene of diurnal insects compared to nocturnal insects (0.065 vs. 0.019; Figure 3a and tableS2). Comparison to a single d_N/d_S ratio model using a likelihood ratio test indicated that the multiple ratios have a better fit to the data ($p < 0.0001$). We repeated the same analysis focusing on Lepidoptera only, for which we defined the ancestral state as nocturnal. The majority of extant Lepidoptera are nocturnal, with shifts to diurnal lifestyle occurring in a few clades such as butterflies, tiger moths (Arctiinae) and some Leafminer moths (e.g. *Cameraria ohridella*). This analysis mirrored the pattern detected for all insects, with the UV opsin having a d_N/d_S ratio of 0.06 in diurnal lineages, compared to 0.03 in nocturnal lineages (Figure 3b tableS2). Comparison to a single d_N/d_S ratio model indicated that the two ratios have a better fit to the data ($p < 0.0005$). We speculate that the underlying adaptive reasons may be related to increased exploitation of UV-reflective patterns on flowers.

In addition, in Lepidoptera we detect a difference in d_N/d_S ratio in the LWS opsins, with nocturnal moths having a ratio of 0.03 and this value increasing to 0.07 in diurnal lineages (Figure 3b and tableS2). This difference is detected whether we include or exclude the duplicated intron-less copies of LWS opsin from *C. dominula*. In each case, using

different nocturnal and diurnal d_N/d_S ratios fits the data better than a single ratio ($p < 0.0001$) (Tierney, et al. 2011). Similar signatures of putative selection have been reported in LWS opsins of dim-light foraging sweat bees (Hymenoptera; Tierney, et al. 2011) and in LWS and UV opsins of fireflies (Coleoptera; Sander and Hall 2015).

To examine if changes to d_N/d_S ratio can be traced to positive selection in specific residues, we applied a branch-site model to the Lepidoptera UV and LWS opsin datasets, excluding the intronless *C. dominula* genes. We identified four and three amino acid sites, respectively, apparently under positive selection in the transitions from nocturnal and diurnal lifestyle (Figure S5); several of these sites are located in the loops between transmembrane domains, suggesting they may be involved in opsin properties other than direct spectral tuning, unless the latter is mediated through long-range effects (Sander and Hall 2015). The adaptive reasons in Lepidoptera are unclear, but could conceivably be related to the extensive deployment of the colour red in wing patterning of butterflies and diurnal tiger moths. In addition, if the absorption spectrum of LWS opsin extends into the very near-infrared, this may allow diurnal Lepidoptera to detect infrared scatter from healthy leaves (Gitelson, et al. 1996) or the direction of incident sunlight for increasing body temperature prior to flight. Further comparative and experimental work will be needed to test these and other hypotheses.

Conclusions

We find that most insects possess five distinct opsin gene families, which have been subject to a variety of gene loss and duplication events. The phylogenetic distribution of the different opsin genes suggests that colour vision may have evolved early in insects, before the origin of flowering plants. Furthermore, we suggest that the majority of nocturnal insects are able to discriminate colours. We identify a signature of adaptive sequence change in the UV opsin of diurnal insects, and the UV and LWS of Lepidoptera specifically. We speculate that the reasons for this positive selection may relate to wing colour pattern recognition, host food-plant detection or body temperature control.

Acknowledgements

We thank Stephen Richards, Michael Kanost, Gary Blissard and members of the Tobacco Hornworm Genome Project for use of *Manduca* data. This study was funded by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant [268513].

Figure 1. Phylogenetic tree of 166 opsins obtained using Phylobayes and a GTR-G model. On each node, the three support values shown are (left to right): PP of opsin+outgroup dataset, PP of opsin unrooted data set, Maximum likelihood bootstrap of the opsin+outgroup dataset. All phylogenetic trees were performed under GTR-G. The lack of support for the UV+Blue opsin clade in the maximum likelihood tree is caused by phylogenetic instability of a single opsin from the human ectoparasitic louse (compare Figure S1, S2 and S3). The orphan sequence not assigned to a named clade derives from pea aphid (*Acyrtosiphon pisum*). The same tree with species names is given in Supplementary Figure S1.

Figure 2. Opsin gene repertoire in genome sequences compared to a cladogram of insect evolution accordingly to Misof et al. (2014). When the number of genes identified is greater than one, this is indicated inside the rectangles. Additional genes are only accepted if the assembly predicts six or seven transmembrane domains; assignment deduced from phylogenetic analysis. White boxes indicate a gene is not found; this may reflect gene loss or genome incompleteness.

Figure 3. Pattern of positive selection inferred for the various opsin paralogues. (a) Comparison of ω for each paralogue across all insects, in relationship to lifestyle. (b) Comparison of ω for each paralogue across Lepidoptera, excluding intron-less *C. dominula* LWS sequences, assuming the common ancestor of the clade was nocturnal.

Table S1. Opsin distribution in the 27 genomes analyzed. The first number in each column indicates the total number of putative opsin loci detected which may include partial gene predictions or pseudogenes; numbers in parentheses indicate those genes predicted to have 6 or 7 transmembrane domains (shown in figure 2)

Table S2: Tests for positive selection on opsin paralogues in relation to lifestyle, for all insects (top) and Lepidoptera only (bottom). The Lepidoptera LWS-set 1 includes all loci; LWS-set 2 excludes intronless LWS loci. The *p*-values for ‘all insects’ are calculated with a likelihood ratio test (LRT) between a model with a single dN/dS value across all sequences and a model with four dN/dS values; this reveals a significant difference for UV opsins. The *p*-values for ‘Lepidoptera only’ are obtained with a LRT between a model with a single dN/dS ratio across all sequences and a model with two ratios corresponding to night and day values; this reveals a significant difference for UV and LWS opsins. The Rh7 dataset does not have a *Cameraria* sequence so for this gene we analyse two transitions to diurnal activity, compared to three transitions in the other lepidopteran opsin datasets. Asterisks (*) indicate significant *p*-value.

Figure S1. Phylogenetic tree of 166 opsins obtained using Phylobayes and a GTR-Γ model. The tree was rooted using the Melatonin receptor (see main text). This is the same tree as given in Figure 1, but showing species names and gene or protein identifier codes where available.

Figure S2. Phylogenetic tree of 166 opsins obtained using Phylobayes and a GTR-Γ model. This tree is calculated using only opsins without a non-opsin outgroup; the root position is then assumed to be between c-opsins and other opsin groups.

Figure S3. Maximum likelihood analysis performed on the opsin+outgroup data set. The tree was obtained using RAxML and a GTR-Γ model. With the exception of a few internal nodes, and the phylogenetic position of UV human ectoparasitic louse opsin (dashed lines), this tree is identical to Bayesian trees in figure S1 and S2.

Figure S4. Intron locations in the multiple predicted LWS sensitive opsins of the Scarlet Tiger moth *Callimorpha dominula*, compared to LWS genes from *Manduca sexta* and *Drosophila melanogaster*. Vertical coloured dashes denote homologous intron positions. The grey area indicates an alignment gap.

Figure S5. 2D predicted models of predicted LWS opsin proteins from (a) *Cameraria* and (b) *Polygonia*. The green circles indicate the sites inferred to be under positive selection in the Lepidoptera LW UV and sensitive opsin respectively (gene in A is *Cameraria* LWS; gene in B is *Polygonia_UVRh*).

References

- Arendt D 2003. Evolution of eyes and photoreceptor cell types. *International Journal of Developmental Biology* 47: 563-572.
- Arendt D, Tessmar-Raible K, Snyman H, Dorresteyn AW, Wittbrodt J 2004. Ciliary photoreceptors with a vertebrate-type opsin in an invertebrate brain. *Science* 306: 869-871.
- Attardo GM, et al. 2014. Genome sequence of the tsetse fly (*Glossina morsitans*): vector of African trypanosomiasis. *Science* 344: 380-386.
- Bellingham J, Tarttelin EE, Foster RG, Wells DJ 2003. Structure and evolution of the Teleost extraretinal rod - like opsin (erro) and ocular rod opsin (rho) genes: Is teleost rho a retrogene? *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* 297: 1-10.
- Booth HAF, Holland PW 2004. Eleven daughters of NANOG. *Genomics* 84: 229-238.
- Brantjes N 1978. Sensory responses to flowers in night-flying moths. The pollination of flowers by insects: 13-19.
- Briscoe AD 2008. Reconstructing the ancestral butterfly eye: focus on the opsins. *Journal of Experimental Biology* 211: 1805-1813.
- Briscoe AD, Chittka L 2001. The evolution of color vision in insects. *Annual review of entomology* 46: 471-510.
- Carulli JP, Chen D-M, Stark WS, Hartl DL 1994. Phylogeny and physiology of *Drosophila* opsins. *J Mol Evol* 38: 250-262.
- Chintapalli VR, Wang J, Dow JA 2007. Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. *Nat Genet* 39: 715-720.
- Cunningham F, et al. 2015. Ensembl 2015. *Nucleic Acids Res* 43: D662-D669.
- Fain GL, Hardie R, Laughlin SB 2010. Phototransduction and the evolution of photoreceptors. *Current Biology* 20: R114-124. doi: S0960-9822(09)02125-3 [pii] 10.1016/j.cub.2009.12.006
- Fawal N, Savelli B, Dunand C, Mathé C 2012. GECA: a fast tool for gene evolution and conservation analysis in eukaryotic protein families. *Bioinformatics* 28: 1398-1399.
- Ferguson L, et al. 2014. Ancient expansion of the Hox cluster in Lepidoptera generated four homeobox genes implicated in extra-embryonic tissue formation.

- Feuda R, Hamilton SC, McInerney JO, Pisani D 2012. Metazoan opsin evolution reveals a simple route to animal vision. *Proc Natl Acad Sci U S A* 109: 18868-18872. doi: 10.1073/pnas.1204609109
- Feuda R, Rota-Stabelli O, Oakley TH, Pisani D 2014. The comb jelly opsins and the origins of animal phototransduction. *Genome Biology and Evolution* 6: 1964-1971.
- Fredriksson R, Lagerstrom MC, Lundin LG, Schioth HB 2003. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol Pharmacol* 63: 1256-1272. doi: 10.1124/mol.63.6.1256
63/6/1256 [pii]
- Futahashi R, et al. 2015. Extraordinary diversity of visual opsin genes in dragonflies. *Proceedings of the National Academy of Sciences* 112: E1247-E1256.
- Gentile C, Rivas GB, Meireles-Filho AC, Lima JB, Peixoto AA 2009. Circadian expression of clock genes in two mosquito disease vectors: *cry2* is different. *Journal of biological rhythms* 24: 444-451.
- Gitelson AA, Merzlyak MN, Lichtenthaler HK 1996. Detection of red edge position and chlorophyll content by reflectance measurements near 700 nm. *Journal of Plant Physiology* 148: 501-508.
- Graveley BR, et al. 2011. The developmental transcriptome of *Drosophila melanogaster*. *Nature* 471: 473-479.
- Gühmann M, et al. 2015. Spectral tuning of phototaxis by a go-opsin in the rhabdomeric eyes of *Platynereis*. *Current Biology* 25: 2265-2271.
- Henze MJ, Oakley TH 2015. The dynamic evolutionary history of pancrustacean eyes and opsins. *Integr Comp Biol* 55: 830-842.
- Jacobs GH 2009. Evolution of colour vision in mammals. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364: 2957-2967.
- Kawada H, Takemura S-Y, Arikawa K, Takagi M 2005. Comparative study on nocturnal behavior of *Aedes aegypti* and *Aedes albopictus*. *Journal of medical entomology* 42: 312-318.
- Kelber A, Balkenius A, Warrant EJ 2003. Colour vision in diurnal and nocturnal hawkmoths. *Integr Comp Biol* 43: 571-579.
- Kelber A, Roth LS 2006. Nocturnal colour vision—not as rare as we might think. *Journal of Experimental Biology* 209: 781-788.
- Kistenpfennig CR 2012. Rhodopsin 7 and Cryptochrome—circadian photoreception in *Drosophila*. Dr rer nat thesis. Julius-Maximilians-Universität Würzburg.

- Lartillot N, Lepage T, Blanquart S 2009. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* 25: 2286-2288. doi: [btp368](https://doi.org/10.1093/bioinformatics/btp368) [pii] 10.1093/bioinformatics/btp368
- Liebertová M, et al. 2015. Cubozoan genome illuminates functional diversification of opsins and photoreceptor evolution. *Scientific reports* 5.
- Loytynoja A, Goldman N 2008. Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. *Science* 320: 1632-1635. doi: [10.1126/science.1158395](https://doi.org/10.1126/science.1158395)
- Marinotti O, et al. 2013. The genome of *Anopheles darlingi*, the main neotropical malaria vector. *Nucleic Acids Res* 41: 7387-7400.
- Markova-Raina P, Petrov D 2011. High sensitivity to aligner and high rate of false positives in the estimates of positive selection in the 12 *Drosophila* genomes. *Genome Res* 21: 863-874.
- Marshall J, Arikawa K 2014. Unconventional colour vision. *Current Biology* 24: R1150-R1154.
- Misof B, et al. 2014. Phylogenomics resolves the timing and pattern of insect evolution. *Science* 346: 763-767.
- Oakley TH, Speiser DI 2015. How complexity originates: The evolution of animal eyes. *bioRxiv*: 017129.
- Omasits U, Ahrens CH, Müller S, Wollscheid B 2013. Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics*: btt607.
- Orr MR 1992. Parasitic flies (Diptera: Phoridae) influence foraging rhythms and caste division of labor in the leaf-cutter ant, *Atta cephalotes* (Hymenoptera: Formicidae). *Behavioral Ecology and Sociobiology* 30: 395-402.
- Osorio D, Vorobyev M 2008. A review of the evolution of animal colour vision and visual communication signals. *Vision Res* 48: 2042-2051.
- Papatsenko D, Nazina A, Desplan C 2001. A conserved regulatory element present in all *Drosophila* rhodopsin genes mediates Pax6 functions and participates in the fine-tuning of cell-specific expression. *Mech Dev* 101: 143-153.
- Passamaneck YJ, Furchheim N, Hejnol A, Martindale MQ, Lütter C 2011. Ciliary photoreceptors in the cerebral eyes of a protostome larva. *EvoDevo* 2: 1-18.
- Plachetzki DC, Fong CR, Oakley TH 2010. The evolution of phototransduction from an ancestral cyclic nucleotide gated pathway. *Proceedings of the Royal Society B-Biological Sciences* 277: 1963-1969. doi: [Doi 10.1098/Rspb.2009.1797](https://doi.org/10.1098/Rspb.2009.1797)

- Porter ML, et al. 2011. Shedding new light on opsin evolution. *Proc Biol Sci*. doi: rspb.2011.1819 [pii]
10.1098/rspb.2011.1819
- Porter ML, Cronin TW, McClellan DA, Crandall KA 2007. Molecular characterization of crustacean visual pigments and the evolution of pancrustacean opsins. *Molecular Biology and Evolution* 24: 253-268.
- Porter SD, Tschinkel WR 1987. Foraging in *Solenopsis invicta* (Hymenoptera: Formicidae): effects of weather and season. *Environmental Entomology* 16: 802-808.
- Rund SS, Hou TY, Ward SM, Collins FH, Duffield GE 2011. Genome-wide profiling of diel and circadian gene expression in the malaria vector *Anopheles gambiae*. *Proceedings of the National Academy of Sciences* 108: E421-E430.
- Sander SE, Hall DW 2015. Variation in opsin genes correlates with signalling ecology in North American fireflies. *Mol Ecol* 24: 4679-4696.
- Shichida Y, Matsuyama T 2009. Evolution of opsins and phototransduction. *Philos Trans R Soc Lond B Biol Sci* 364: 2881-2895. doi: 10.1098/rstb.2009.0051
- Stamatakis A 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688-2690. doi: btl446 [pii]
10.1093/bioinformatics/btl446
- Terakita A 2005. The opsins. *Genome Biol* 6: 213. doi: gb-2005-6-3-213 [pii]
10.1186/gb-2005-6-3-213
- Tierney SM, et al. 2015. Opsin transcripts of predatory diving beetles: a comparison of surface and subterranean photic niches. *Royal Society Open Science* 2: 140386.
- Tierney SM, et al. 2011. Photic niche invasions: phylogenetic history of the dim-light foraging augochlorine bees (Halictidae). *Proceedings of the Royal Society of London B: Biological Sciences*: rspb20111355.
- Tsirigos KD, Peters C, Shu N, Käll L, Elofsson A 2015. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Res*: gkv485.
- Veilleux CC, Louis EE, Bolnick DA 2013. Nocturnal light environments influence color vision and signatures of selection on the OPN1SW opsin gene in nocturnal lemurs. *Molecular Biology and Evolution* 30: 1420-1437.
- Velarde RA, Sauer CD, Walden KK, Fahrbach SE, Robertson HM 2005. Pteropsin: a vertebrate-like non-visual opsin expressed in the honey bee brain. *Insect biochemistry and molecular biology* 35: 1367-1377.

Warrant E, Dacke M 2011. Vision and visual navigation in nocturnal insects. *Annual review of entomology* 56: 239-254.

Warrant EJ 2008. Seeing in the dark: vision and visual behaviour in nocturnal bees and wasps. *Journal of Experimental Biology* 211: 1737-1746.

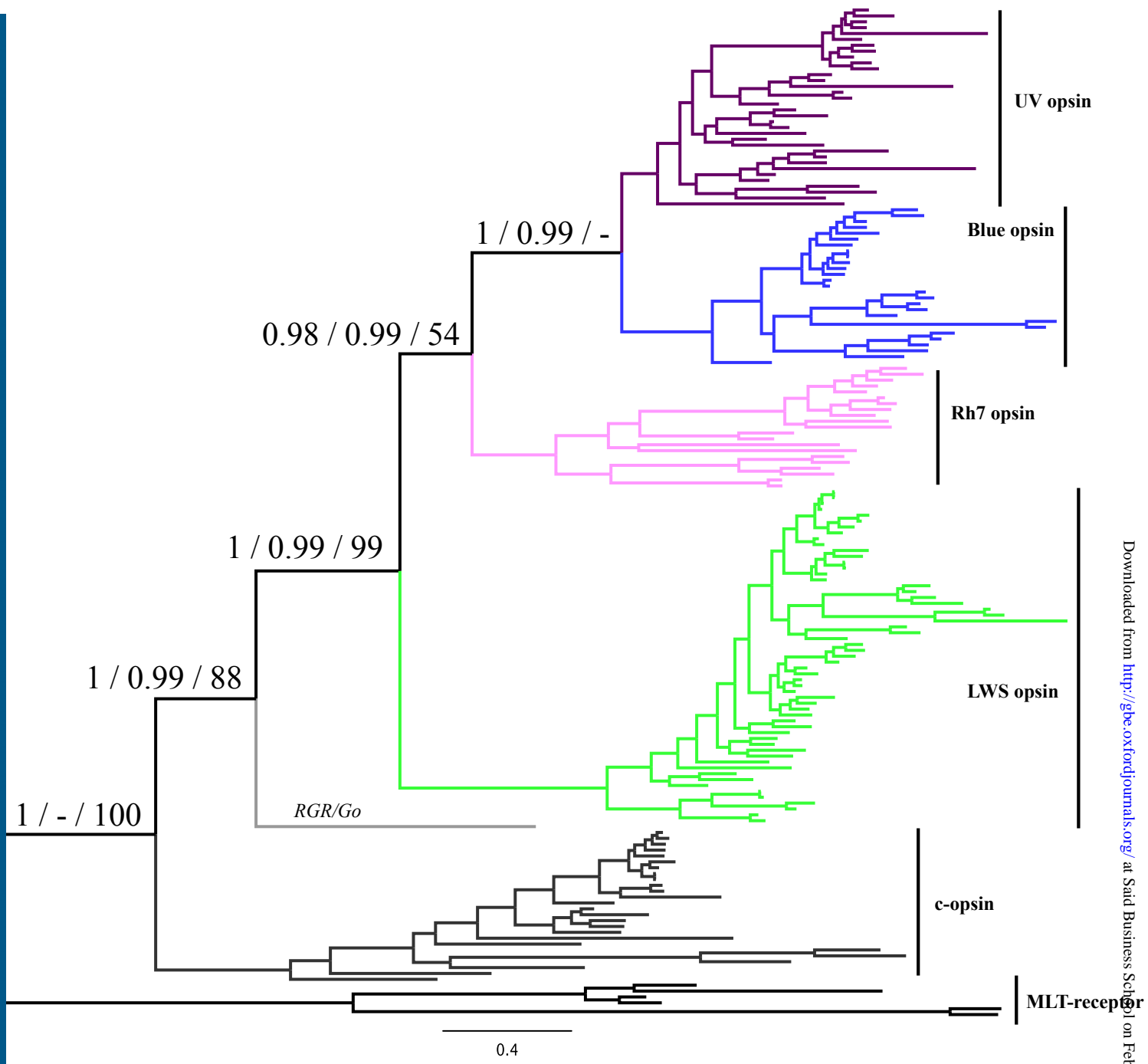
Yang Z 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* 24: 1586-1591.

You M, et al. 2013. A heterozygous moth genome provides insights into herbivory and detoxification. *Nat Genet* 45: 220-225.

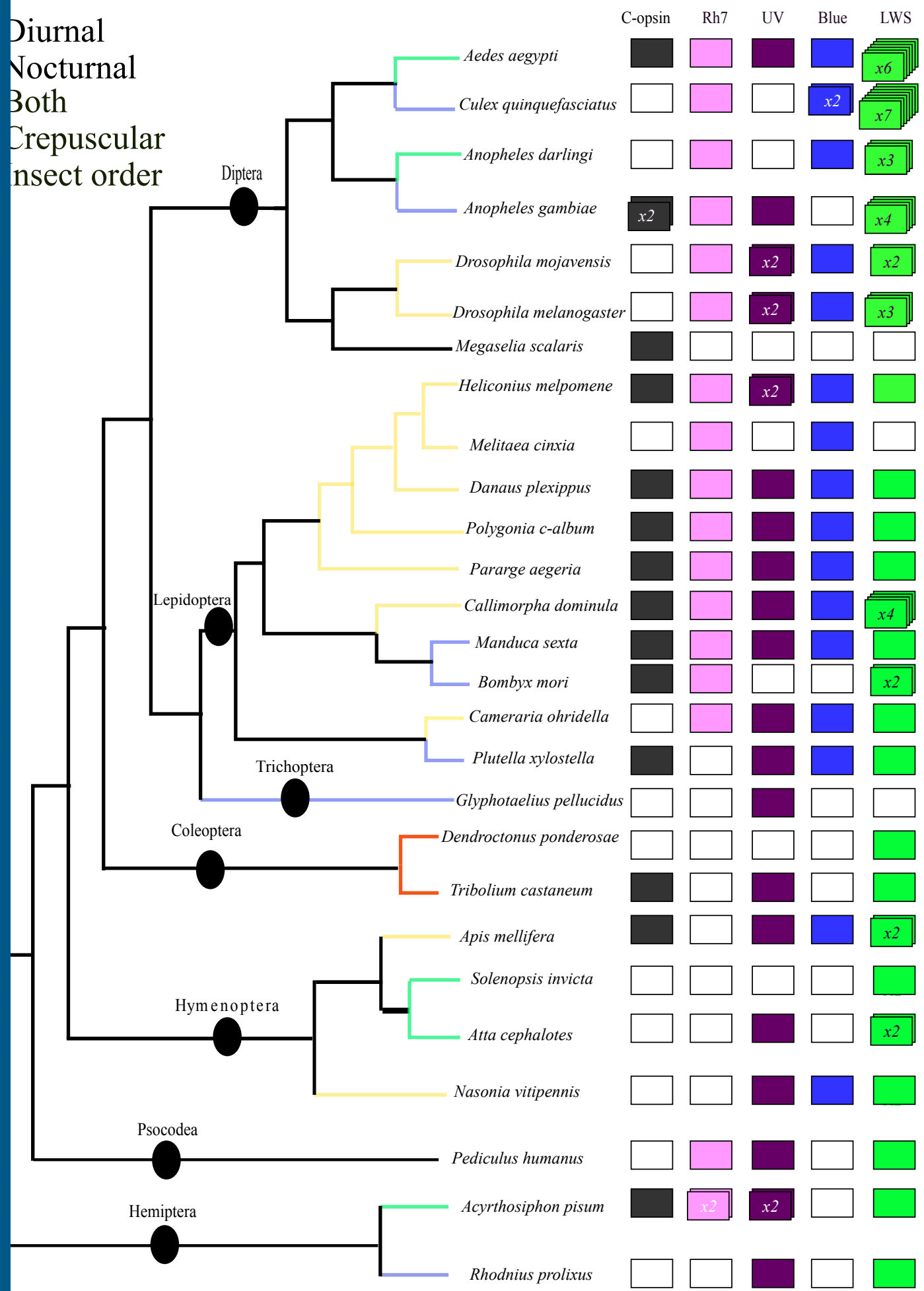
Yuan F, Bernard GD, Le J, Briscoe AD 2010. Contrasting modes of evolution of the visual pigments in *Heliconius* butterflies. *Molecular Biology and Evolution* 27: 2392-2405.

Zhao H, et al. 2009. The evolution of color vision in nocturnal mammals. *Proceedings of the National Academy of Sciences* 106: 8980-8985.

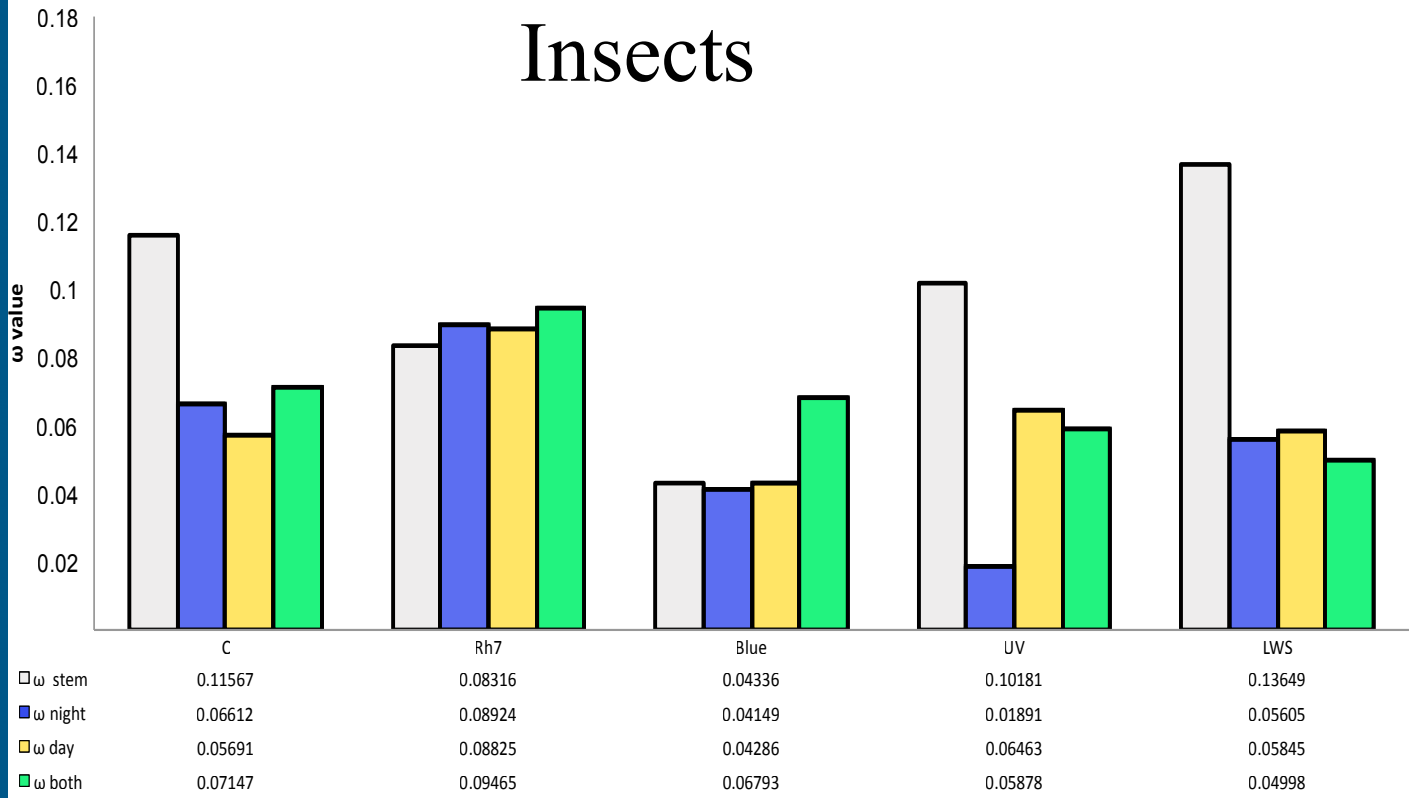
Zimmerman RH 1992. Ecology of malaria vectors in the Americas and future direction. *Memórias do Instituto Oswaldo Cruz* 87: 371-383.



Diurnal
Nocturnal
Both
Crepuscular
Insect order



Insects



Lepidoptera

