Opsin evolution in the Ambulacraria

D'Aniello S[1,§,*], Delroisse J[2,3,§], Valero-Gracia A[1], Lowe EK[1,4], Byrne M[5], Cannon JT[6,7], Halanych KM[6], Elphick MR[3], Mallefet J[8], Kaul-Strehlow S[9], Lowe CJ[10], Flammang P[2], Ullrich-Lüter E[11], Wanninger A[12] and Arnone MI[1]

[1] Department of Biology and Evolution of Marine Organisms, Stazione Zoologica Anton Dohrn, Villa Comunale, 80121 Napoli, Italy.

[2] Biology of Marine Organisms and Biomimetics, Research Institute for Biosciences, University of Mons, Avenue du Champs de Mars 6, 7000 Mons, Belgium.

[3] School of Biological & Chemical Sciences, Queen Mary University of London, London, E1 4NS, UK.

[4] BEACON Center for the Study of Evolution in Action, Michigan State University, East Lansing, Michigan, USA

[5] Schools of Medical and Biological Sciences, The University of Sydney, Sydney, NSW, Australia.

[6] Department of Biological Sciences and Molette Biology Laboratory for Environmental and Climate Change Studies, Auburn University, Auburn, USA.

[7] Department of Zoology, Naturhistoriska Riksmuseet, Stockholm, Sweden; Friday Harbor Laboratories, University of Washington, Friday Harbor, WA 98250, USA.

[8] Laboratory of Marine Biology, Earth and Life Institute, Université Catholique de Louvain, Louvain-La-Neuve, Place Croix du Sud 3, bt L7.06.04, 1348 Louvain-la-Neuve, Belgium.

[9] Department of Molecular Evolution and Development, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria.

[10] Hopkins Marine Station of Stanford University, Pacific Grove, CA 93950, USA.

[11] Museum fuer Naturkunde Berlin, Invalidenstr 43, 10115 Berlin, Germany.

[12] Department of Integrative Zoology, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria.

§ These authors contributed equally to this work.

*Corresponding author: salvatore.daniello@szn.it

Abstract

Opsins - G-protein coupled receptors involved in photoreception - have been extensively studied in the animal kingdom. The present work provides new insights into opsin-based photoreception and photoreceptor cell evolution with a first analysis of opsin sequence data for a major deuterostome clade, the Ambulacraria. Systematic data analysis, including for the first time hemichordate opsin sequences and an expanded echinoderm dataset, led to a robust opsin phylogeny for this cornerstone superphylum. Multiple genomic and transcriptomic resources were surveyed to cover each class of Hemichordata and Echinodermata. In total, 119 ambulacrarian opsin sequences were found, 22 new sequences in hemichordates and 97 in echinoderms (including 67 new sequences). We framed the ambulacrarian opsin repertoire within eumetazoan diversity by including selected reference opsins from non-ambulacrarians. Our findings corroborate the presence of all major ancestral bilaterian opsin groups in Ambulacraria. Furthermore, we identified two opsin groups specific to echinoderms. In conclusion, a molecular phylogenetic framework for investigating light-perception and photobiological behaviours in marine deuterostomes has been obtained.

Keywords: Opsin; Photoreceptor Cell Evolution; Ambulacraria; Echinoderm; Hemichordate; Phylogeny; Echinopsin.

Manuscript type: Original data paper.

Introduction

In animals, the prototypical molecules involved in photoreception and vision are opsin proteins (Nilsson, 2004). Opsins are G-protein coupled receptors (GPCR) that consist of an apoprotein plus a covalently bound to a chromophore (11-retinal) (Terakita, 2005). The nitrogen atom of the amino group of residue K296, situated in helix VII, binds to the retinal molecule through a Schiff-base linkage, forming a double bond with the carbon atom at the end of this molecule (Hargrave *et al*., 1983). Residue K296 is, therefore, crucial for light absorption, and its presence or absence can be used as a molecular fingerprint to judge whether or not a GPCR is a *bona fide* opsin.

Recent investigations on opsin phylogeny resolved six distinct groups present in metazoans: ciliary opsins, rhabdomeric opsins, Go-opsins, neuropsins, peropsins, and RGR (RPE-retinal G protein-coupled receptor) opsins (Porter *et al*., 2012; Feuda *et al*. 2012; Terakita *et al*., 2012). A vast number of opsins are also expressed in non-ocular tissues (Porter *et al*., 2012; Plachetzki *et al*., 2005; Koyanagi *et al*., 2005; Terakita *et al*., 2012).

With regard to opsin evolution in the deuterostomes, genomic and transcriptomic data of a number of chordates have been used to identify and characterize their opsins (*e.g*. Holland *et al*., 2008; Kusakabe *et al*., 2001). However, little attention has been paid to Ambulacraria, the sister group to all extant chordates, (i.e. cephalochordates, urochordates, and vertebrates, Edgecombe *et al*., 2011), a key clade to reconstruct the opsin set of the common ancestor of extant deuterostomes.

The present study integrates opsin sequences from two ambulacrarian sub-lineages: enteropneust Hemichordata, (Harrimaniidae, Spengelidae, Ptychoderidae and Torquaratoridae), and the pentameral Echinodermata comprising five classes (Crinoidea, Ophiuroidea, Asteroidea, Holoturoidea and Echinoidea).

The phylogenetic relationship of echinoderms and hemichordates as sister groups within Ambulacraria, as shown in Figure 1, was already suggested by Metschnikoff (1881), and supported by Nielsen (2012). The monophyly of Ambulacraria is also well supported by molecular phylogenetic analyses (Cannon *et al*., 2014; Telford *et al*., 2014). Moreover, Cannon and colleagues showed that the six hemichordate subgroups cluster into two monophyletic taxa, Enteropneusta and Pterobranchia (Rhabdopleuridae and Cephalodiscidae). Finally, Figure 1 conforms to the Asterozoa hypothesis separating the Echinozoa (Echinoidea + Holothuroidea) and the Asterozoa (Asteroidea

+ Ophiuroidea), which is now well supported by recent molecular phylogenies (Cannon *et al*., 2014; Telford *et al*., 2014; O'Hara *et al*., 2014).

Other than a few structural investigations of eye-like structures in some asteroid species (e.g. the starfish optic cushion) and in enteropneust larvae (Brandenburger *et al*., 1973; Nezlin and Yushin, 2004; Braun *et al*., 2015), the molecular mechanisms of echinoderm and hemichordate photoreception remained enigmatic until recently. Immunohistochemical studies indicated the presence of a putative rhodopsin in the asteroid *Asterias forbesi* and in the ophiuroid *Ophioderma brevispinum* (Johnsen, 1997). Subsequently, Raible *et al*. (2006) analyzed the 'rhodopsin-type' G-protein-coupled receptors family in an echinoid genome (*Strongylocentrotus purpuratus*). They predicted six *bona fide* opsin sequences, four of which were reported independently by Burke *et al*. (2006). Later, Ooka *et al*. (2010) cloned an "encephalopsin" orthologue in the sea urchin *Hemicentrotus pulcherrimus*. Recently, more opsin sequences have been found in sea urchins (*S. purpuratus; Paracentrotus lividus*), starfish (*Asterias rubens*), and brittle stars (*Ophiocomina nigra*, *Amphiura filiformis*) (Delroisse *et al*., 2013, 2014, 2015 a,b ; Ullrich-Lüter *et al*., 2011, 2013). These studies highlighted the expression of ciliary and rhabdomeric opsins in various echinoderm tissues. Also, a large opsin gene repertoire was identified in the brittle star *A. filiformis*, pinpointing notable differences with findings from the previously published sea urchin genome (Delroisse *et al*., 2014). However, a comprehensive description of opsin diversity in echinoderms is still lacking and almost nothing is known about hemichordate opsins.

Therefore, to characterize and describe the diversity of the opsin family in the Ambulacraria, we conducted a detailed analysis of 6 genomic and 24 transcriptomic sequence databases. This work represents the first attempt to describe and characterize the evolution of the opsin "toolkit" in the ambulacrarian lineage. We performed a phylogenetic study using the largest dataset of ambulacrarian opsin sequences to date, including representatives of a previously neglected group, Hemichordata.

Materials and methods

Data mining

Strongylocentrotus purpuratus opsins belonging to all the paralogous classes (Supp. File 1) were used as starting query sequences for tBLASTx against transcriptomic and genomic databases including public databases (NCBI, JGI, Ensemble, Echinobase (www.echinobase.org/), BioInformatique CNRS-UPMC (http://octopus.obs-vlfr.fr/) and Genoscope (http://www.genoscope.cns.fr/spip/Generation-de-ressources.html). The parameters used across all our tBLASTx searches were the following: Matrix: Blosum62; gap penalties: existence: 11; extension: 1; neighboring words threshold: 13; window for multiple hits: 40. Additionally, our dataset was further enriched using various unpublished genomic and transcriptomic databases obtained from several independent research projects (Suppl. File 1 and Suppl. File 2). This includes transcriptomes from adult specimens' tissues, such as cuverian tubules and integument from *Holothuria forskali*, muscle of *Parastichopus californicus,* radial nerve from *Asterias rubens*, arms from *Labidiaster annulatus*, *Ophiopsila aranea*, *Astrotomma agassizii* and *Antedon mediterranea*, proboscis from *Saccoglossus mereschkowskii and Torquaratorid sp*, whole adult body of *Leptosynapta clarki* and anterior part of the body *from Harrimaniidae sp* and *Schizocardium braziliense*. Several other transcriptomes obtained from embryos or larvae from *Paracentrotus lividus*, *Heliocidaris erythrogramma*, *Eucidaris tribuloides*, *Parasticopus parvimensis*, *Saccoglossus kowalevskii* and *Ptychodera flava* (Suppl. File 1 and Suppl. File 2) were also screened.
The raw predicted opsin sequences used in this study are listed in the Suppl. File 3 in fasta format.

Alignment and phylogenetic analyses

Predicted protein alignments were performed with SeaView v4.2.12 (Galtier *et al*., 1996; Gouy *et al*., 2010) using the MUSCLE algorithm (Edgar, 2004). To improve phylogenetic reconstruction, N-terminal and C-terminal ends were trimmed and the alignment was manually corrected in order to minimize gaps and eliminate ambiguous and misaligned regions. Sequences that were shorter than 60 amino acids were removed to avoid bias. However, these could potentially correspond to true opsins and merit further study.

Maximum likelihood analyses (ML) of our dataset were conducted on Michigan State University's High Performance Computing Cluster using PhyML v3.0 (Guindon and Gascuel, 2003), and nodal support assessed with 1000 bootstrap replicates is indicated. The alignment is shown in Suppl. File 4 (phylip format) and Suppl. File 5 (image). A

best-fit model analysis was performed using MEGA6 (following the AIC criteria) (Tamura *et al*., 2007; Kumar *et al*., 2008) and WAG+G+F amino acid substitution model was found to be the best suited (Whelan and Goldman, 2001). Three melatonin receptor sequences from *S*. *purpuratus* (Echinodermata) and three from *Saccoglossus kowalevskii* (Hemichordata) were chosen as the best outgroup for the opsin phylogeny, as previously proposed by Plachetzki *et al*. (2010) and Feuda *et al*. (2014).

Consensus fingerprint of ambulacrarian opsin groups

Ambulacraria opsins were clustered according their estimated position within opsin subfamilies and a multiple alignment of a 35 amino-acid long peptide region, including the 7th transmembrane domain with the opsin-specific lysine (K296), was performed with SeaView v4.2.12 for each opsin group supported in our phylogenetic tree. The selected region spanned residues 286 to 320 of the *Rattus norvegicus* rhodopsin sequence used as a reference (Palczewski *et al*., 2000). The consensus sequence was generated on the basis of the alignment for each class of ambulacrarian opsin using Geneious®8.1.5.

Results

Phylogeny and opsin distribution within ambulacrarian groups

Using a collection of both genomic and transcriptomic data (see Materials and Methods and Suppl. File 2 for details), a final set of 119 protein sequences, representing 31 ambulacrarian species, was generated for our phylogenetic reconstruction, which included 6 outgroup sequences and 6 human reference opsin sequences (Suppl. File 1 and Suppl. File 3 for raw predicted protein sequences). The trimmed opsin alignment is shown in the Suppl. File 5 (see Suppl. File 4 for the alignment phylip file). We employed maximum likelihood using the WAG+G+F model with melatonin receptors as an outgroup. Canonical opsin groups are well supported in our analysis (Figure 2), demonstrating the presence of a complex opsin toolkit in Ambulacraria.

Interestingly, according to our data, two novel groups of opsins were found, which we have named echinopsin-A and echinopsin-B groups. *Ad hoc* BLAST searches against several metazoan genomes clearly indicated the absence of these two opsin types

outside the echinoderm lineage. The previously identified Sp-opsin2 and Sp-opsin5 belong to echinopsins-A and echinopsins-B, respectively (Raible *et al*., 2006).

A complete opsin profile including at least one representative of each prototypical opsin group (opsin 1-8) was detected in the sea urchin *S*. *purpuratus*, but not in *Lytechinus variegatus* or *P*. *lividus*. The genomes of the latter two species have not yet been comprehensively sequenced and annotated, and therefore some opsin genes may be missing due to incomplete sequence coverage. With the exception of echinopsin-B, a complete opsin profile was found in the genome sequence data of the starfish *Patiria miniata*. The starfish *A*. *rubens* radial nerve transcriptome also contained several opsins, including ciliary, Go-, RGR-opsins.

Surprisingly, rhabdomeric and Go-opsins do not seem to be present in hemichordates in our dataset. However, this requires confirmation through more extensive taxonomic sampling of hemichordate sequence data because, at present, only one hemichordate genome has been fully sequenced (*S*. *kowalevskii*). In several opsin groups we observed lineage-specific duplications: two opsins in *P*. *miniata* and *A*. *rubens*; five neuropsins in *S*. *kowalevskii*; four r-opsins in *L*. *annulatus* and six r-opsins in *A*. *filiformis*; two Go-opsins in the echinoids *L*. *variegatus*, *S*. *purpuratus* and *Heliocidaris erythrogramma*. Nevertheless, some of these molecules present a short overlapping sequence therefore we cannot exclude that they could be part of unique genes and therefore to overestimate their number.

Alignment of the transmembrane domain and opsin fingerprint

In order to build a consensus fingerprint to distinguish the various ambulacrarian opsin groups, the 7th transmembrane domain and C-terminal tail region of our sequence dataset were aligned and a graphical representation was generated (Fig. 3). All sequences were characterized by the general structure of G protein-coupled receptors (GPCRs) comprising seven transmembrane (TM) domains. Numerous residues characteristic of opsins are present in the opsin sequences of *A*. *filiformis*. However, as several sequences are partial, not all characteristic residues could be detected in all sequences. Most of the opsin sequences also contained the highly conserved lysine residue (equivalent to K296 of the *R*. *norvegicus* rhodopsin) critical for Schiff base linkage formed with retinal, except three sea-urchin peropsins (Sp-opsin 6, Pl-opsin 6, Lv-opsin 6) in which it is substituted by a glutamate (E). The dipeptide NP (position 302-303 of the *R*. *norvegicus* rhodopsin sequence) is also highly conserved among all

the subfamilies except in peropsins (N/HP) and RGR-opsins, which show divergence in these residues (also rhabdomeric opsins to a lesser extent). Amino-acid conservation for each opsin group from our phylogenetic analysis is shown in Figure 3. Ambulacrarian c-opsins, r-opsins and echinopsins-A displayed a highly conserved tyrosine (Y306). Conversely, the histidine (H310) appears distinctive of the ambulacrarian r-opsins (Figure 3) and r-opsins in general (human melanopsin, octopus rhodopsin and *Drosophila* Rh1-opsin). In our dataset the tripeptide SSS, positioned at residues 309-402 of the reference protein, is a distinctive feature of ambulacrarian Go-opsins.

These representations will be particularly useful in future studies in support of phylogenetic analysis to assign novel, unknown sequences to lineage-specific opsin groups.

Discussion

Our phylogenetic analyses showed ambulacrarian opsin sequences to be represented in all six prototypical bilaterian opsin groups: ciliary opsins, rhabdomeric opsins, neuropsins, Go-opsins, peropsins and RGR-opsins (Fig. 4). In addition we confirmed the presence of two novel echinoderm-specific opsin groups, which we have named echinopsins (echinopsin-A and echinopsin-B). These novel groups of opsins, which were found only in Echinoidea, Ophiuroidea and Asteroidea, respectively cluster as a sister group of all other opsins and as a sister group of all opsins except Echinopsins-A and ciliary opsins (Fig.4). A deeper analysis of these groups of proteins, including more hemichordate opsin sequences, is needed in order to determine if they represent an echinoderm or ambulacrarian novelty.

Our analysis failed to reveal a rhabdomeric opsin (r-opsin) in hemichordates. The absence of such an opsin type is surprising because many enteropneust tornaria larvae possess eyespots that bear photoreceptors with clear microvillar surface enlargement (Brandenburger *et al*., 1973; Nezlin and Yushin, 2004; Braun *et al*., 2015). So far, photoreception in microvillar photoreceptor cell types has been demonstrated to generally deploy opsins of the so-called rhabdomeric type (r-opsins), although co-expression of other opsin types in microvillar/rhabdomeric photoreceptors has been shown in recent studies (Randel *et al*., 2013). However, although our analysis reveals no such opsin in any of the examined enteropneust species, it should be noted that genomic

information is only available from the direct developer *S*. *kowalevskii*, which does not have a larval (tornarian) stage in its life cycle. Moreover, most of hemichordate transcriptomes in our study were generated using adult tissues; it is therefore possible that the absence of r-opsin in this group of animals is due to a limitation of data availability from this understudied group of animals.

In contrast to the lack of r-opsins in enteropneusts, our analyses showed several cases of opsin gene duplication. Obviously in some instances the locus of duplication prompted a large expansion of the gene family, as is the case of the five neuropsins found in *S*. *kowalevskii*, and the six rhabdomeric opsins in *A*. *filiformis*, with the latter previously described by Delroisse *et al*. (2014). However, the fragmentary information about these duplicates makes it difficult to predict the exact number of functional opsin proteins in Ambulacraria. Whether or not these duplicated genes have sub-functionalized roles should be experimentally investigated by knock-out or silencing experiments.

Until recently, under-representation of many taxonomic groups in comparative studies of photoreceptor evolution has hidden the real extent of opsin diversity (Porter *et al*., 2012; Feuda *et al*., 2014). As more opsins have been characterized, these sequences have been classified into narrow pre-defined groups (e.g. Group 4 opsins), implying theoretical functional similarities that might not always be correct (Shichida and Matsuyama, 2009). At present, however, the rapidly increasing availability of entire genomes and transcriptomes provides a large number of sequences for investigating the evolution and functional diversity of the opsin family in greater detail. Likewise, our detailed phylogenetic analyses of ambulacrarian opsins not only provide a better understanding of opsin evolution in general, but are also essential for future photoreceptor studies elucidating the evolution of opsin functions.

Figures

Figure 1. Ambulacrarian phylogenetic relationships and their adult forms.

The Ambulacraria consist of two groups: Hemichordata, bilateral animals subdivided in six clades: Cephalodiscidae, Rhabdopleuridae, Harrimaniidae, Spengelidae, Ptychoderidae and Torquaratoridae, and the pentameral Echinodermata, comprising: Crinoidea, Ophiuroidea, Asteroidea, Holoturoidea and Echinoidea. For each class there is a representation of the adult body plan. The numbers represented on the figure correspond to the two hemichordate subgroups: 1. Pterobranchia and 2. Enteropneusta, and the two echinoderm subgroups 3. Eleutherozoa and 4. Crinozoa.

Figure 2. Phylogenetic reconstruction of ambulacrarian opsins.

119 opsin from 31 different ambulacrarian species cluster in eight highly supported groups in this maximum likelihood (ML) based analysis. R-opsins in blue, c-opsins in red, Go-opsins in green, neuropsins in purple, peropsins in yellow and RGR-opsin in orange. Visualization was generated with Figtree.

Figure 3. Consensus sequences of different opsin groups.

Graphical representations of opsin amino acid patterns within the multiple alignments of the 7th transmembrane domain and the protein G linkage site. The 7th transmembrane domain is highlighted in green in the tridimensional representation of a typical opsin receptor. Alignment is limited to the highly conserved regions including the opsin-specific lysine residue and the "NPxxY(x)6F" pattern. The lysine residue involved in the Schiff base formation - equivalent to K296 of the *R. norvegicus* rhodopsin - is present in position 10. The pattern "NPxxY(x)6F" (position 302-313 of the *R. norvegicus* rhodopsin sequence) is present in position 17-28. The size of each amino acid indicates the probability to find this specific amino acid for the considered position. Amino acid patterns of Melatonin receptors used as an outgroup in the phylogenetic analysis is also presented.

Figure 4. Opsin distribution within the investigated ambulacraria species.

For each species the number of opsin belonging to classical groups were reported. Those species for which no opsins were find are not reported in the table (for additional

informations see Suppl. File 1). Species for which the genome data is available are in bold.

Supplementary Figure 1. List of all the ambulacraria species used in our analysis and the opsin gene content per species.

Supplementary Figure 2. List of the ambulacraria transcriptomes surveyed in our analysis. The available information regarding the transcriptomes used is provided and the RNA source material is detailed.

Supplementary Figure 3. Opsin fasta file. Predicted opsin proteins used in the phylogenetic analysis are listed in fasta format.

Supplementary Figure 4. The phylip version of the trimmed alignment used in the study.

Supplementary Figure 5. Trimmed alignment supporting the phylogenetic tree shown in Figure 2.

Trimmed alignment of all deduced amino acid sequences of Ambulacraria opsins. The alignment mainly contain the "TM cores" of the opsins. Non-conserved N-terminus and C-terminus ends were trimmed. The Schiff base equivalent to the lysine residue in the position 296 of the *R. norvegicus* rhodopsin – is indicated by an asteriscs in the alignment. Alignment performed in Seaview and edited in Geneious®8.1.5.

Aknowledgements

The authors would like to thank Santiago Valero-Medranda for the illustrations in Figure 1; Prof. Andy Cameron, Dr. Joel Smith, the staff of Echinobase for giving access to various echinoderm sequences, and Thierry Lepage and Genoscope *P. lividus* Project N°17. Paola Oliveri, UCL (UK) and Yi-Hsien Su, Academia Sinica of Taipei (Taiwan) for their suggestions in the early phase of the project. The authors would like to thank Elise Hennebert, Mélanie Demeuldre and Akihito Omori for providing access to the *A. rubens* tube foot transcriptome, the *H. forskali* cuverian tubule transcriptome and the Oxycomanthus japonicus embryo transcriptome respectively. S. D'Aniello is supported by a Marie Curie Career Integration Grant, FP7-PEOPLE-2011-CIG (grant no. PCIG09-GA-2011-293871). J. Delroisse is currently supported by BBSRC grant no. BB/M001644/1 awarded to M.R. Elphick (QMUL). A. Valero-Gracia is supported by the Marie Curie ITN 'Neptune' (grant no. 317172, PI: M.I. Arnone). This work is supported in part by the "National funds for research" (FNRS) of Belgium (FRFC grant no. 2.4590.11). K.M. Halanych and J.T Cannon hemichordate's transcriptomes were funded by the USA National Science Foundation (grant no. DEB-0816892). E. Ullrich-Lüter is supported by a grant of the German Research foundation (DFG) grant no: UL 428/2-1. M. Byrne is supported by that Australian Research Council (DP120102849). S. Kaul-Strehlow is supported by a grant of the Austrian Science Fund (FWF, M-1485).

References

Arendt, D., 2008. The evolution of cell types in animals: emerging principles from molecular studies. Nat. Rev. Genet. 9, 868-82.

Angueyra, J.M., Pulido, C., Malagón, G., Nasi, E., Gomez, M.P., 2012. Melanopsin-expressing amphioxus photoreceptors transduce light via a phospholipase C signaling cascade. PLoS ONE 7, e29813.

Bortoff, A., 1964. Localization of slow potential responses in the *Necturus retina*. Vision Res. 4, 627-635.

Brandenburger J.L., Woollacott R.M., Eakin R.M., 1973. Fine structure of eyespots in tornarian larvae (phylum: Hemichordata). Z Zellforsch Mikrosk Anat. 142: 89-102.

Braun, K., Kaul-Strehlow, S., Ullrich-Lüter, E. and Stach, T., 2015. Structure and ultrastructure of eyes of tornaria larvae of *Glossobalanus marginatus*. Org. Divers. Evol. 15, 423-428.

Burke, R.D., Angerer, L.M., Elphick, M.R., Humphrey, G.W., Yaguchi, S., Kiyama, T., *et al*., 2006. A genomic view of the sea urchin nervous system. Developmental Biology 300, 434-60.

Cannon, J.T., Kocot, K.M., Waits, D.S., Weese, D.A., Swalla, B.J., Santos, S.R., *et al*., 2014. Phylogenomic resolution of the hemichordate and echinoderm clade. Curr. Biol. 24, 2827-32.

Delroisse, J., Lanterbecq, D., Eeckhaut, I., Mallefet, J., Flammang, P., 2013. Opsin detection in the sea urchin *Paracentrotus lividus* and the sea star *Asterias rubens*. Cahiers de Biologie Marine 54, 721-727.

Delroisse, J., Ortega-Martinez, O., Dupont, S., Mallefet, J., Flammang, P. 2015. *De novo* transcriptome of the European brittle star *Amphiura filiformis* pluteus larvae, Marine Genomics 23, 109-121.

Delroisse, J., Ullrich-Lüter, E., Ortega-Martinez, O., Dupont, S., Arnone, M.I., Mallefet, J, *et al*., 2014. High opsin diversity in a non-visual infaunal brittle star. BMC Genomics 15, 1035.

Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32,1792-1797.

Edgecombe, G.D., Giribet, G., Dunn, C.W., Hejnol, A., Kristensen, R.M., Neves, R.C. *et al*., 2011. Higher-level metazoan relationships: recent progress and remaining questions. Organisms Diversity & Evolution 11, 151-172.

Feuda, R., Hamilton, S.C., McInerney, J.O., Pisani, D., 2012. Metazoan opsin evolution reveals a simple route to animal vision. Proceedings of the National Academy of Sciences 109, 18868-18872.

Feuda, R., Rota-Stabelli, O., Oakley, T.H., Pisani, D., 2014. The comb jelly opsins and the origins of animal phototransduction. Genome Biol. Evol. 6, 1964-71.

Galtier, N., Gouy, M., Gautier, C., 1996. SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. Comput. Applic. Biosci. 12, 543-548.

Gouy, M., Guindon, S., Gascuel, O., 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol. 27, 221-224.

Guindon, S., Delsuc, F., Dufayard, J.F., Gascuel, O., 2009. Estimating Maximum Likelihood Phylogenies with PhyML. Methods Mol. Biol. 537, 113-137.

Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696-704.

Hagins, A., 1962. Local membrane current in the outer segments of squid photoreceptors. Nature 194, 844-847.

Hargrave, P.A., McDowell, J.H., Curtis, D.R., Wang, J.K., Juszczak, E., Fong, S.L., *et al*., 1983. The structure of bovine rhodopsin. Biophysics of Structure and Mechanism 9, 235-244.

Holland, L., *et al*., 2008. The amphioxus genome illuminates vertebrate origins and cephalochordate biology. Genome Research 18, 1100-11.

Johnsen, S., 1997. Identification and localization of a possible rhodopsin in the echinoderms *Asterias forbesi* (Asteroidea) and *Ophioderma brevispinum* (Ophiuroidea). Biol. Bull. 193, 97-105.

Koyanagi, M., Kubokawa, K., Tsukamoto, H., Shichida, Y., Terakita A., 2005. Cephalochordate melanopsin: evolutionary linkage between invertebrate visual cells and vertebrate photosensitive retinal ganglion cells. Current Biology 15, 1065-1069.

Kumar, S., Nei, M., Dudley, J., Tamura, K., 2008. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinform. 9, 299-306.

Kusakabe, T., Kusakabe, R., Kawakami, I., Satou, Y., Satoh, N., Tsuda, M., 2001. Ci-opsin1, a vertebrate-type opsin gene, expressed in the larval ocellus of the ascidian *Ciona intestinalis*. FEBS Letters 506, 69-72.

Land, M.F. and Fernald, R.D., 1992. The evolution of eyes. Annu Rev Neurosci. 15, 1-29.

Millott, N., 1975. The photosensitivity of echinoids. Advances in Marine Biology 13, 1-52.

Nielsen, C., 2012. Animal Evolution. Oxford University Press.

Nezlin, L.P., Yushin, V.V., 2004. Structure of the nervous system in the tornaria larva of Balanoglossus proterogonius (Hemichordata: Enteropneusta) and its physiogenetic implications. Zoomorphology 123, 1-13.

Nakashima, Y., Kusakabe, T., Kusakabe, R., Terakita, A., Shichida, Y., Tsuda, M., 2003. Origin of the vertebrate visual cycle: genes encoding retinal photoisomerase and two putative visual cycle proteins are expressed in whole brain of a primitive chordate. J. Comp. Neurol. 460, 180-90.

Metschnikoff, V.E., 1881. Über die systematische Stellung von Balanoglossus. Zool. Anz. 4, 139-157.

O'Hara, T.D., Hugall, A.f., Thuy, B., Moussalli, A., 2014. Phylogenomic resolution of the class Ophiuroidea unlocks a global microfossil record. Curr. Biol. 24, 1874-79.

Ooka, S., Katow, T., Yaguchi, S., Yaguchi, J., Katow, H., 2010. Spatiotemporal expression pattern of an encephalopsin orthologue of the sea urchin *Hemicentrotus pulcherrimus* during early development, and its potential role in larval vertical migration. Development, Growth & Differentiation 52, 195-207.

Palczewski, K., Kumasaka, T., Hori, T., Behnke, C.A., Motoshima, H., Fox, B.A., *et al*., 2000. Crystal structure of rhodopsin: A G protein-coupled receptor. Science 289, 739-45.

Plachetzki, DC, Fong, CR, Oakley, TH., 2010. The evolution of phototransduction from an ancestral cyclic nucleotide gated pathway. Proc. Biol. Sci. 277, 1963-69.

Porter, M.L., Blasic, J.R., Bok, M.J., Cameron, E.G., Pringle, T., Cronin, T. W., Robinson, P.R., 2012. Shedding new light on opsin evolution. Proceedings of the Royal Society B: Biological Sciences 279, 3-14.

Raible, F., Tessmar-Raible, K., Arboleda, E., Kaller, T., Bork, P., Arendt, D., *et al*., 2006. Opsins and clusters of sensory G protein-coupled receptors in the sea urchin genome. Developmental Biology 300, 461-75.

Randel, N., Bezares-Calderón, L.A., Gühmann, M., Shahidi, R., Jékely, G., 2013. Expression dynamics and protein localization of rhabdomeric opsins in Platynereis larvae. Integr. Comp. Biol. 53, 7-16.

Ridge, K.D., Palczewski, K., 2007. Visual rhodopsin sees the light: structure and mechanism of G protein signaling. Journal of Biological Chemistry 282, 9297-301.

Schertler, G.F.X., 2008. Signal transduction: The rhodopsin story continued. Nature 453, 292-3.

Scheerer, P., Park, J.H., Hildebrand, P.W., Kim, Y.J., Krauß, N., Choe, H.W. *et al*., 2008. Crystal structure of opsin in its G protein-interacting conformation. Nature 455, 497-502.

Schmidt-Rhaesa, A., 2007. The Evolution of Organ Systems. Oxford University Press, USA.

Schwartz, T.W., Hubbell, W.L., 2008. Structural biology: A moving story of receptors. Nature 455, 473-4.

Shichida, Y., Matsuyama, T., 2009. Evolution of opsins and phototransduction. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 364, 2881-95.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725-2729.

Telford, M.J., Lowe, CJ, Cameron C.B., Ortega-Martinez, O., Aronowicz, J., Oliveri, P., Copley, R.R., 2014. Phylogenomic analysis of echinoderm class relationships supports Asterozoa. *Proc. R. Soc. Lond. Ser. B* 281: 20140479.

Terakita, A., 2005. The opsins. Genome Biology 6, 213.

Terakita, A., Kawano-Yamashita, E., Koyanagi, M., 2012. Evolution and diversity of opsins. Wiley Interdisciplinary Reviews: Membrane Transport and Signaling 1, 104-111.

Ullrich-Lüter, E.M., Dupont, S., Arboleda, E., Hausen, H., Arnone, M.I., 2011. Unique system of photoreceptors in sea urchin tube feet. PNAS 108, 8367-72.

Ullrich-Lüter, E.M., D'Aniello, S., Arnone, M.I., 2013. C-opsin expressing photoreceptors in echinoderms. Integrative and Comparative Biology 53, 27-38.

Vopalensky, P., Pergner, J., Liegertova, M., Benito-Gutierrez, E., Arendt, D., Kozmik, K., 2012. Molecular analysis of the amphioxus frontal eye unravels the evolutionary origin of the retina and pigment cells of the vertebrate eye. PNAS 109, 15383-88.

Whelan, S., Goldman, N., 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol. Biol. Evol. 18, 691-699.

Figure 1

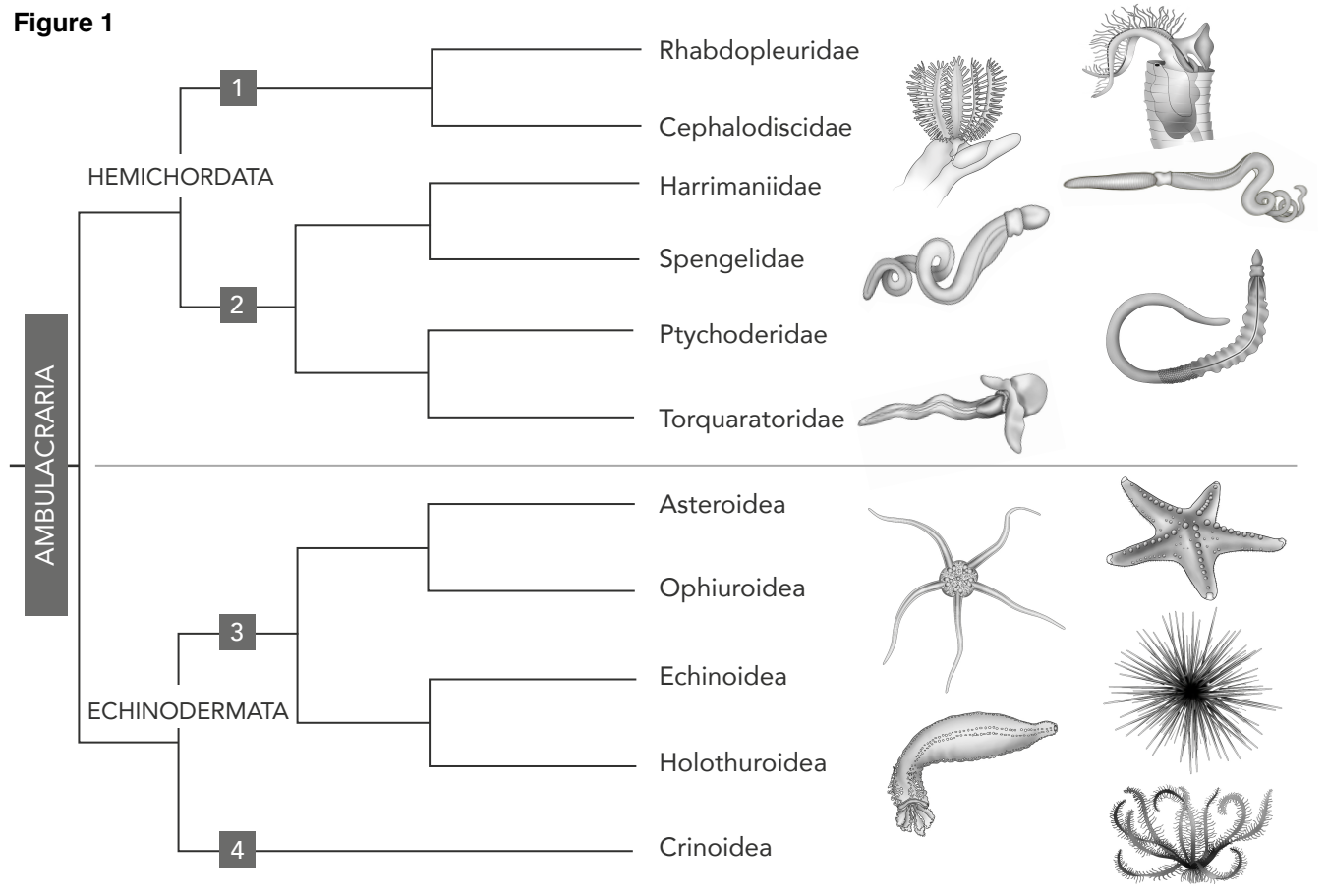
AMBULACRARIA
HEMICHORDATA
ECHINODERMATA
1
2
3
4
Rhabdopleuridae
Cephalodiscidae
Harrimaniidae
Spengelidae
Ptychoderidae
Torquaratoridae
Asteroidea
Ophiuroidea
Echinoidea
Holothuroidea
Crinoidea

Figure 2

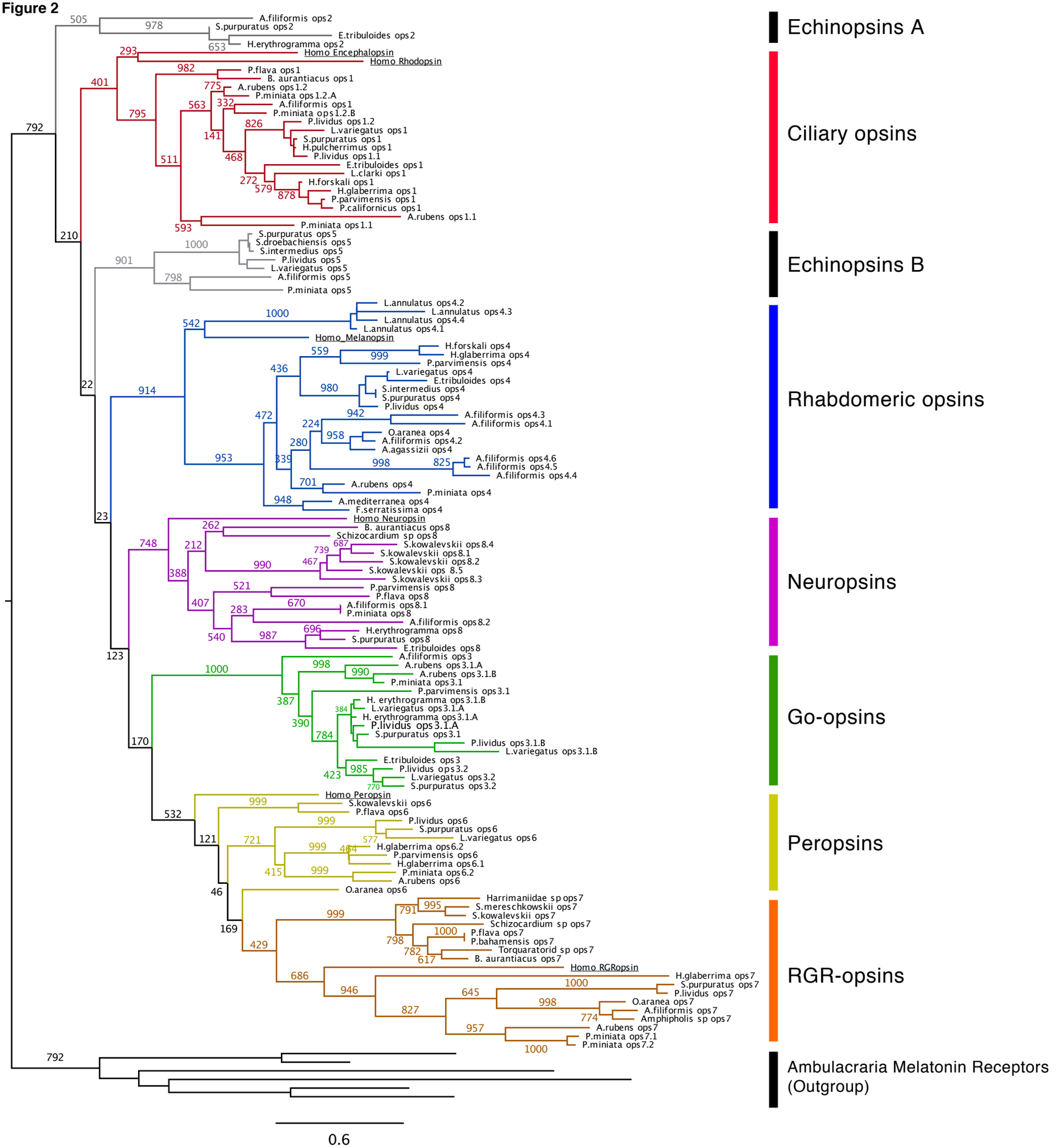
A.filiformis ops2
S.purpuratus ops2
E.tribuloides ops2
H.erythrogramma ops2
Homo Encephalopsin
Homo Rhodopsin
P.flava ops1
B. aurantiacus ops1
A.rubens ops1.2
P.miniata ops1.2.A
A.filiformis ops1
P.miniata ops1.2.B
P.lividus ops1.2
L.variegatus ops1
S.purpuratus ops1
H.pulcherrimus ops1
P.lividus ops1.1
E.tribuloides ops1
L.clarki ops1
H.forskali ops1
H.glaberrima ops1
P.parvimensis ops1
P.californicus ops1
A.rubens ops1.1
P.miniata ops1.1
S.purpuratus ops5
S.droebachiensis ops5
S.intermedius ops5
P.lividus ops5
L.variegatus ops5
A.filiformis ops5
P.miniata ops5
L.annulatus ops4.2
L.annulatus ops4.3
L.annulatus ops4.4
L.annulatus ops4.1
Homo_Melanopsin
H.forskali ops4
H.glaberrima ops4
P.parvimensis ops4
L.variegatus ops4
E.tribuloides ops4
S.intermedius ops4
S.purpuratus ops4
P.lividus ops4
A.filiformis ops4.3
A.filiformis ops4.1
O.aranea ops4
A.filiformis ops4.2
A.agassizii ops4
A.filiformis ops4.6
A.filiformis ops4.5
A.filiformis ops4.4
A.rubens ops4
P.miniata ops4
A.mediterranea ops4
F.serratissima ops4
Homo Neuropsin
B. aurantiacus ops8
Schizocardium sp ops8
S.kowalevskii ops8.4
S.kowalevskii ops8.1
S.kowalevskii ops8.2
S.kowalevskii ops 8.5
S.kowalevskii ops8.3
P.parvimensis ops8
P.flava ops8
A.filiformis ops8.1
P.miniata ops8
A.filiformis ops8.2
H.erythrogramma ops8
S.purpuratus ops8
E.tribuloides ops8
A.filiformis ops3
A.rubens ops3.1.A
A.rubens ops3.1.B
P.miniata ops3.1
P.parvimensis ops3.1
H. erythrogramma ops3.1.B
L.variegatus ops3.1.A
H. erythrogramma ops3.1.A
P.lividus ops3.1.A
S.purpuratus ops3.1
P.lividus ops3.1.B
L.variegatus ops3.1.B
E.tribuloides ops3
P.lividus ops3.2
L.variegatus ops3.2
S.purpuratus ops3.2
Homo Peropsin
S.kowalevskii ops6
P.flava ops6
P.lividus ops6
S.purpuratus ops6
L.variegatus ops6
H.glaberrima ops6.2
P.parvimensis ops6
H.glaberrima ops6.1
P.miniata ops6.2
A.rubens ops6
O.aranea ops6
Harrimaniidae sp ops7
S.mereschkowskii ops7
S.kowalevskii ops7
Schizocardium sp ops7
P.flava ops7
P.bahamensis ops7
Torquaratorid sp ops7
B. aurantiacus ops7
Homo RGRopsin
H.glaberrima ops7
S.purpuratus ops7
P.lividus ops7
O.aranea ops7
A.filiformis ops7
Amphipholis sp ops7
A.rubens ops7
P.miniata ops7.1
P.miniata ops7.2
0.6
Echinopsins A
Ciliary opsins
Echinopsins B
Rhabdomeric opsins
Neuropsins
Go-opsins
Peropsins
RGR-opsins
Ambulacraria Melatonin Receptors (Outgroup)

Figure 3


1
10
20
30
34
ECHINOPSINS A
CILIARY OPSINS
ECHINOPSINS B
RHABDOMERIC OPSINS
NEUROPSINS
Go-OPSINS
PEROPSINS
RGR OPSINS
OUTGROUP (Melatonin Receptors)
Bits
0
2
4

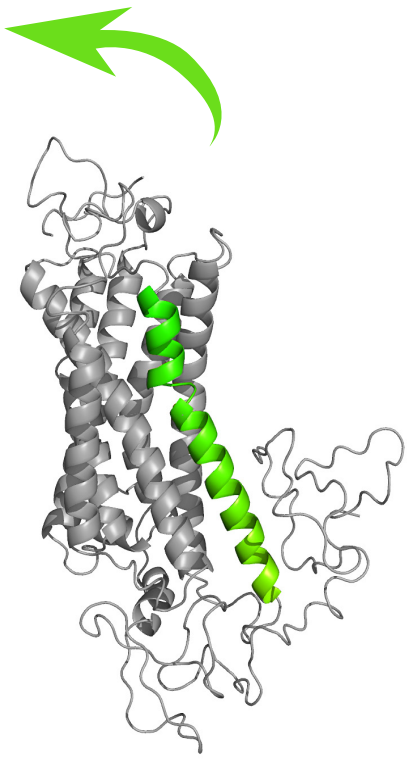

Figure 4

Phylum	Class	Species	TOTAL	R-opsins / opsins 4	C-opsins / opsins 1	Go opsins / opsins 3	Neuropsins / opsins 8	Peropsins / opsins 6	RGR-opsins / opsins 7	Echinopsins A / opsins 2	Echinopsins B / opsins 5
Echinodermata	*Echinoidea*	***Strongylocentrotus purpuratus***	**8**	**1**	**1**	**2**	**1**	**1**	**1**	**1**	**1**
		Strongylocentrotus droebachiensis	1								1
		Strongylocentrotus intermedius	2	1							1
		Lytechinus variegatus	**6**	**1**	**1**	**2**		**1**			**1**
		Paracentrotus lividus	**7**	**1**	**1**	**2**		**1**	**1**		**1**
		Hemicentrotus pulcherrimus	1		1						
		Eucidaris tribuloides	5	1	1	1	1			1	
		Heliocidaris erythrogramma	3			1	1			1	
	Holothuroidea	*Parastichopus californicus*	1		1						
		Parastichopus parvamensis	5	1	1	1	1	1			
		Leptosynapta clarki	1		1						
		Holothuria forskali	2	1	1						
		Holothuria glaberrima	4	1	1			1	1		
	Ophiuroidea	***Amphiura filiformis***	**13**	**6**	**1**	**1**	**2**		**1**	**1**	**1**
		Ophiopsila aranea	3	1				1	1		
		Astrotomma agassizii	1	1							
		Amphipholis sp	1						1		
	Asteroidea	*Asterias rubens*	6	1	2	1		1	1		
		Patiria miniata	**10**	**1**	**2**	**1**	**1**	**2**	**2**		**1**
		Labidiaster annulatus	4	4							
	Crinoidea	*Antedon mediterranea*	1	1							
		Florometra serratissima	1	1							
Hemichorda	*Harrimaniidae*	***Saccoglossus kowalevskii***	7				**5**	**1**	**1**		
		Saccoglossus mereschkowskii	1						1		
		Harrimaniidae sp. (Iceland)	1						1		
	Spengelidae	*Schizocardium c.f. braziliense*	2				1		1		
	Ptychoderidae	*Ptychodera flava*	4		1		1	1	1		
		Ptychodera bahamensis	1						1		
		Balanoglossus c.f. aurantiacus	3		1		1		1		
	Torquaratoridae	*Torquaratoid sp. (iceland)*	1						1		

Supplementary Material File 1

Click here to download Supplementary Material: Suppl. File 1 NEW 29 September.xlsx

Supplementary Material File 2

Click here to download Supplementary Material: Suppl. File 2 NEW Transcriptome table.xlsx

Supplementary Material File 3

Click here to download Supplementary Material: Suppl. File 3 NEW Opsin Fasta.docx

Supplementary Material File 4

Click here to download Supplementary Material: Suppl. File 4 NEW Opsin alignment.phy.txt

Supplementary Material File 5

Click here to download Supplementary Material: Suppl. File 5 NEW Opsin alignment image 29Sept.pdf