

Opsin evolution in the Ambulacraria.

D'Aniello, S; Delroisse, J; Valero-Gracia, A; Lowe, EK; Byrne, M; Cannon, JT; Halanych, KM; Elphick, MR; Mallefet, J; Kaul-Strehlow, S; Lowe, CJ; Flammang, P; Ullrich-Lüter, E; Wanninger, A; Arnone, MI

© 2015 The Authors.

CC-BY-NC-ND

For additional information about this publication click this link.

<http://qmro.qmul.ac.uk/xmlui/handle/123456789/13563>

Information about this research object was correct at the time of download; we occasionally make corrections to records, please therefore check the published record when citing. For more information contact scholarlycommunications@qmul.ac.uk

Opsin evolution in the Ambulacraria

1
2
3 D’Aniello S^{1,§,*}, Delroisse J^{2,3,§}, Valero-Gracia A¹, Lowe EK^{1,4}, Byrne M⁵, Cannon
4
5
6 JT^{6,7}, Halanych KM⁶, Elphick MR³, Mallefet J⁸, Kaul-Strehlow S⁹, Lowe CJ¹⁰,
7
8
9 Flammang P², Ullrich-Lüter E¹¹, Wanninger A¹² and Arnone MI¹
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

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 pentamerous Echinodermata comprising five classes (Crinoidea, Ophiuroidea, Asteroidea, Holothuroidea 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

1
2 + Ophiuroidea), which is now well supported by recent molecular phylogenies (Cannon
3 *et al.*, 2014; Telford *et al.*, 2014; O'Hara *et al.*, 2014).

4 Other than a few structural investigations of eye-like structures in some asteroid species
5 (e.g. the starfish optic cushion) and in enteropneust larvae (Brandenburger *et al.*, 1973;
6 Nezhlin and Yushin, 2004; Braun *et al.*, 2015), the molecular mechanisms of echinoderm
7 and hemichordate photoreception remained enigmatic until recently.
8 Immunohistochemical studies indicated the presence of a putative rhodopsin in the
9 asteroid *Asterias forbesi* and in the ophiuroid *Ophioderma brevispinum* (Johnsen,
10 1997). Subsequently, Raible *et al.* (2006) analyzed the 'rhodopsin-type' G-protein-
11 coupled receptors family in an echinoid genome (*Strongylocentrotus purpuratus*). They
12 predicted six *bona fide* opsin sequences, four of which were reported independently by
13 Burke *et al.* (2006). Later, Ooka *et al.* (2010) cloned an "encephalopsin" orthologue in
14 the sea urchin *Hemicentrotus pulcherrimus*. Recently, more opsin sequences have been
15 found in sea urchins (*S. purpuratus*; *Paracentrotus lividus*), starfish (*Asterias rubens*),
16 and brittle stars (*Ophiocomina nigra*, *Amphiura filiformis*) (Delroisse *et al.*, 2013, 2014,
17 2015 a,b ; Ullrich-Lüter *et al.*, 2011, 2013). These studies highlighted the expression of
18 ciliary and rhabdomeric opsins in various echinoderm tissues. Also, a large opsin gene
19 repertoire was identified in the brittle star *A. filiformis*, pinpointing notable differences
20 with findings from the previously published sea urchin genome (Delroisse *et al.*, 2014).
21 However, a comprehensive description of opsin diversity in echinoderms is still lacking
22 and almost nothing is known about hemichordate opsins.

23
24 Therefore, to characterize and describe the diversity of the opsin family in the
25 Ambulacraria, we conducted a detailed analysis of 6 genomic and 24 transcriptomic
26 sequence databases. This work represents the first attempt to describe and characterize
27 the evolution of the opsin "toolkit" in the ambulacrarian lineage. We performed a
28 phylogenetic study using the largest dataset of ambulacrarian opsin sequences to date,
29 including representatives of a previously neglected group, Hemichordata.

30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 **Materials and methods**

57 58 59 **Data mining**

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Strongylocentrotus purpuratus opsins belonging to all the paralogous classes (Suppl. 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*, *Astrotonna 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*, *Parastichopus 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

1 best-fit model analysis was performed using MEGA6 (following the AIC criteria)
2 (Tamura *et al.*, 2007; Kumar *et al.*, 2008) and WAG+G+F amino acid substitution
3 model was found to be the best suited (Whelan and Goldman, 2001). Three melatonin
4 receptor sequences from *S. purpuratus* (Echinodermata) and three from *Saccoglossus*
5 *kowalevskii* (Hemichordata) were chosen as the best outgroup for the opsin phylogeny,
6 as previously proposed by Plachetzki *et al.* (2010) and Feuda *et al.* (2014).
7
8
9
10

11 **Consensus fingerprint of ambulacrarian opsin groups**

12 Ambulacraria opsins were clustered according their estimated position within opsin
13 subfamilies and a multiple alignment of a 35 amino-acid long peptide region, including
14 the 7th transmembrane domain with the opsin-specific lysine (K296), was performed
15 with SeaView v4.2.12 for each opsin group supported in our phylogenetic tree. The
16 selected region spanned residues 286 to 320 of the *Rattus norvegicus* rhodopsin
17 sequence used as a reference (Palczewski *et al.*, 2000). The consensus sequence was
18 generated on the basis of the alignment for each class of ambulacrarian opsin using
19 Geneious®8.1.5.
20
21
22
23
24
25
26
27
28
29
30
31

32 **Results**

33 **Phylogeny and opsin distribution within ambulacrarian groups**

34 Using a collection of both genomic and transcriptomic data (see Materials and Methods
35 and Suppl. File 2 for details), a final set of 119 protein sequences, representing 31
36 ambulacrarian species, was generated for our phylogenetic reconstruction, which
37 included 6 outgroup sequences and 6 human reference opsin sequences (Suppl. File 1
38 and Suppl. File 3 for raw predicted protein sequences). The trimmed opsin alignment is
39 shown in the Suppl. File 5 (see Suppl. File 4 for the alignment phylip file). We
40 employed maximum likelihood using the WAG+G+F model with melatonin receptors
41 as an outgroup. Canonical opsin groups are well supported in our analysis (Figure 2),
42 demonstrating the presence of a complex opsin toolkit in Ambulacraria.
43
44
45
46
47
48
49
50
51
52
53

54 Interestingly, according to our data, two novel groups of opsins were found, which we
55 have named echinopsin-A and echinopsin-B groups. *Ad hoc* BLAST searches against
56 several metazoan genomes clearly indicated the absence of these two opsin types
57
58
59
60
61
62
63
64
65

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

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

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

37 **Alignment of the transmembrane domain and opsin fingerprint**

38 In order to build a consensus fingerprint to distinguish the various ambulacrarian opsin
39 groups, the 7th transmembrane domain and C-terminal tail region of our sequence
40 dataset were aligned and a graphical representation was generated (Fig. 3). All
41 sequences were characterized by the general structure of G protein-coupled receptors
42 (GPCRs) comprising seven transmembrane (TM) domains. Numerous residues
43 characteristic of opsins are present in the opsin sequences of *A. filiformis*. However, as
44 several sequences are partial, not all characteristic residues could be detected in all
45 sequences. Most of the opsin sequences also contained the highly conserved lysine
46 residue (equivalent to K296 of the *R. norvegicus* rhodopsin) critical for Schiff base
47 linkage formed with retinal, except three sea-urchin peropsins (Sp-opsin 6, Pl-opsin 6,
48 Lv-opsin 6) in which it is substituted by a glutamate (E). The dipeptide NP (position
49 302-303 of the *R. norvegicus* rhodopsin sequence) is also highly conserved among all
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 the subfamilies except in peropsins (N/HP) and RGR-opsins, which show divergence in
2 these residues (also rhabdomeric opsins to a lesser extent). Amino-acid conservation for
3 each opsin group from our phylogenetic analysis is shown in Figure 3. Ambulacrarian
4 c-opsins, r-opsins and echinopsins-A displayed a highly conserved tyrosine (Y306).
5 Conversely, the histidine (H310) appears distinctive of the ambulacrarian r-opsins
6 (Figure 3) and r-opsins in general (human melanopsin, octopus rhodopsin and
7 *Drosophila* Rh1-opsin). In our dataset the tripeptide SSS, positioned at residues 309-
8 402 of the reference protein, is a distinctive feature of ambulacrarian Go-opsins.

9 These representations will be particularly useful in future studies in support of
10 phylogenetic analysis to assign novel, unknown sequences to lineage-specific opsin
11 groups.
12
13
14
15
16
17
18
19
20
21
22

23 Discussion

24
25
26
27 Our phylogenetic analyses showed ambulacrarian opsin sequences to be represented in
28 all six prototypical bilaterian opsin groups: ciliary opsins, rhabdomeric opsins,
29 neuropsins, Go-opsins, peropsins and RGR-opsins (Fig. 4). In addition we confirmed
30 the presence of two novel echinoderm-specific opsin groups, which we have named
31 echinopsins (echinopsin-A and echinopsin-B). These novel groups of opsins, which
32 were found only in Echinoidea, Ophiuroidea and Asteroidea, respectively cluster as a
33 sister group of all other opsins and as a sister group of all opsins except Echinopsins-A
34 and ciliary opsins (Fig.4). A deeper analysis of these groups of proteins, including more
35 hemichordate opsin sequences, is needed in order to determine if they represent an
36 echinoderm or ambulacrarian novelty.
37
38
39

40 Our analysis failed to reveal a rhabdomeric opsin (r-opsin) in hemichordates. The
41 absence of such an opsin type is surprising because many enteropneust tornaria larvae
42 possess eyespots that bear photoreceptors with clear microvillar surface enlargement
43 (Brandenburger *et al.*, 1973; Nezhlin and Yushin, 2004; Braun *et al.*, 2015). So far,
44 photoreception in microvillar photoreceptor cell types has been demonstrated to
45 generally deploy opsins of the so-called rhabdomeric type (r-opsins), although co-
46 expression of other opsin types in microvillar/rhabdomeric photoreceptors has been
47 shown in recent studies (Randel *et al.*, 2013). However, although our analysis reveals no
48 such opsin in any of the examined enteropneust species, it should be noted that genomic
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 information is only available from the direct developer *S. kowalevskii*, which does not
2 have a larval (tornarian) stage in its life cycle. Moreover, most of hemichordate
3 transcriptomes in our study were generated using adult tissues; it is therefore possible
4 that the absence of r-opsin in this group of animals is due to a limitation of data
5 availability from this understudied group of animals.
6
7

8
9 In contrast to the lack of r-opsins in enteropneusts, our analyses showed several cases of
10 opsin gene duplication. Obviously in some instances the locus of duplication prompted
11 a large expansion of the gene family, as is the case of the five neuropsins found in *S.*
12 *kowalevskii*, and the six rhabdomeric opsins in *A. filiformis*, with the latter previously
13 described by Delroisse *et al.* (2014). However, the fragmentary information about these
14 duplicates makes it difficult to predict the exact number of functional opsin proteins in
15 Ambulacraria. Whether or not these duplicated genes have sub-functionalized roles
16 should be experimentally investigated by knock-out or silencing experiments.
17
18

19
20 Until recently, under-representation of many taxonomic groups in comparative studies
21 of photoreceptor evolution has hidden the real extent of opsin diversity (Porter *et al.*,
22 2012; Feuda *et al.*, 2014). As more opsins have been characterized, these sequences
23 have been classified into narrow pre-defined groups (e.g. Group 4 opsins), implying
24 theoretical functional similarities that might not always be correct (Shichida and
25 Matsuyama, 2009). At present, however, the rapidly increasing availability of entire
26 genomes and transcriptomes provides a large number of sequences for investigating the
27 evolution and functional diversity of the opsin family in greater detail. Likewise, our
28 detailed phylogenetic analyses of ambulacrarian opsins not only provide a better
29 understanding of opsin evolution in general, but are also essential for future
30 photoreceptor studies elucidating the evolution of opsin functions.
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

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 pentamerous 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 found are not reported in the table (for additional

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

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 asterisks 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

1
2
3
4 Arendt, D., 2008. The evolution of cell types in animals: emerging principles from
5 molecular studies. *Nat. Rev. Genet.* 9, 868-82.
6

7
8
9 Angueyra, J.M., Pulido, C., Malagón, G., Nasi, E., Gomez, M.P., 2012. Melanopsin-
10 expressing amphioxus photoreceptors transduce light via a phospholipase C signaling
11 cascade. *PLoS ONE* 7, e29813.
12
13

14
15
16 Bortoff, A., 1964. Localization of slow potential responses in the *Necturus retina*.
17 *Vision Res.* 4, 627-635.
18
19

20
21
22 Brandenburger J.L., Woollacott R.M., Eakin R.M., 1973. Fine structure of eyespots in
23 tornarian larvae (phylum: Hemichordata). *Z Zellforsch Mikrosk Anat.* 142: 89-102.
24
25

26
27 Braun, K., Kaul-Strehlow, S., Ullrich-Lüter, E. and Stach, T., 2015. Structure and
28 ultrastructure of eyes of tornaria larvae of *Glossobalanus marginatus*. *Org. Divers.*
29 *Evol.* 15, 423-428.
30
31

32
33
34 Burke, R.D., Angerer, L.M., Elphick, M.R., Humphrey, G.W., Yaguchi, S., Kiyama, T.,
35 *et al.*, 2006. A genomic view of the sea urchin nervous system. *Developmental Biology*
36 300, 434-60.
37
38
39

40
41
42 Cannon, J.T., Kocot, K.M., Waits, D.S., Weese, D.A., Swalla, B.J., Santos, S.R., *et al.*,
43 2014. Phylogenomic resolution of the hemichordate and echinoderm clade. *Curr. Biol.*
44 24, 2827-32.
45
46
47

48
49 Delroisse, J., Lanterbecq, D., Eeckhaut, I., Mallefet, J., Flammang, P., 2013. Opsin
50 detection in the sea urchin *Paracentrotus lividus* and the sea star *Asterias rubens*.
51 *Cahiers de Biologie Marine* 54, 721-727.
52
53
54

55
56 Delroisse, J., Ortega-Martinez, O., Dupont, S., Mallefet, J., Flammang, P. 2015. *De*
57 *novo* transcriptome of the European brittle star *Amphiura filiformis* pluteus larvae,
58 *Marine Genomics* 23, 109-121.
59
60
61
62

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

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.

1 Hargrave, P.A., McDowell, J.H., Curtis, D.R., Wang, J.K., Juszczak, E., Fong, S.L., *et*
2 *al.*, 1983. The structure of bovine rhodopsin. *Biophysics of Structure and Mechanism* 9,
3 235-244.
4
5

6
7
8
9 Holland, L., *et al.*, 2008. The amphioxus genome illuminates vertebrate origins and
10 cephalochordate biology. *Genome Research* 18, 1100-11.
11
12

13
14 Johnsen, S., 1997. Identification and localization of a possible rhodopsin in the
15 echinoderms *Asterias forbesi* (Asteroidea) and *Ophioderma brevispinum* (Ophiuroidea).
16 *Biol. Bull.* 193, 97-105.
17
18
19

20
21 Koyanagi, M., Kubokawa, K., Tsukamoto, H., Shichida, Y., Terakita A., 2005.
22 Cephalochordate melanopsin: evolutionary linkage between invertebrate visual cells
23 and vertebrate photosensitive retinal ganglion cells. *Current Biology* 15, 1065-1069.
24
25
26
27

28
29 Kumar, S., Nei, M., Dudley, J., Tamura, K., 2008. MEGA: a biologist-centric software
30 for evolutionary analysis of DNA and protein sequences. *Brief Bioinform.* 9, 299-306.
31
32
33

34 Kusakabe, T., Kusakabe, R., Kawakami, I., Satou, Y., Satoh, N., Tsuda, M., 2001. Ci-
35 opsin1, a vertebrate-type opsin gene, expressed in the larval ocellus of the ascidian
36 *Ciona intestinalis*. *FEBS Letters* 506, 69-72.
37
38
39
40

41 Land, M.F. and Fernald, R.D., 1992. The evolution of eyes. *Annu Rev Neurosci.* 15, 1-
42 29.
43
44
45

46
47 Millott, N., 1975. The photosensitivity of echinoids. *Advances in Marine Biology* 13, 1-
48 52.
49
50
51

52
53 Nielsen, C., 2012. *Animal Evolution*. Oxford University Press.
54
55

56
57 Nezhlin, L.P., Yushin, V.V., 2004. Structure of the nervous system in the tornaria larva
58 of *Balanoglossus proterogonius* (Hemichordata: Enteropneusta) and its physiogenetic
59 implications. *Zoomorphology* 123, 1-13.
60
61
62

1 Nakashima, Y., Kusakabe, T., Kusakabe, R., Terakita, A., Shichida, Y., Tsuda, M.,
2 2003. Origin of the vertebrate visual cycle: genes encoding retinal photoisomerase and
3 two putative visual cycle proteins are expressed in whole brain of a primitive chordate.
4 J. Comp. Neurol. 460, 180-90.
5
6
7
8
9

10
11 Metschnikoff, V.E., 1881. Über die systematische Stellung von Balanoglossus. Zool.
12 Anz. 4, 139-157.
13
14
15

16
17 O'Hara, T.D., Hugall, A.f., Thuy, B., Moussalli, A., 2014. Phylogenomic resolution of
18 the class Ophiuroidea unlocks a global microfossil record. Curr. Biol. 24, 1874-79.
19
20
21

22
23 Ooka, S., Katow, T., Yaguchi, S., Yaguchi, J., Katow, H., 2010. Spatiotemporal
24 expression pattern of an encephalopsin orthologue of the sea urchin *Hemicentrotus*
25 *pulcherrimus* during early development, and its potential role in larval vertical
26 migration. Development, Growth & Differentiation 52, 195-207.
27
28
29
30

31
32 Palczewski, K., Kumasaka, T., Hori, T., Behnke, C.A., Motoshima, H., Fox, B.A., *et*
33 *al.*, 2000. Crystal structure of rhodopsin: A G protein-coupled receptor. Science 289,
34 739-45.
35
36
37

38
39 Plachetzki, DC, Fong, CR, Oakley, TH., 2010. The evolution of phototransduction from
40 an ancestral cyclic nucleotide gated pathway. Proc. Biol. Sci. 277, 1963-69.
41
42
43

44
45 Porter, M.L., Blasic, J.R., Bok, M.J., Cameron, E.G., Pringle, T., Cronin, T. W.,
46 Robinson, P.R., 2012. Shedding new light on opsin evolution. Proceedings of the Royal
47 Society B: Biological Sciences 279, 3-14.
48
49
50

51
52 Raible, F., Tessmar-Raible, K., Arboleda, E., Kaller, T., Bork, P., Arendt, D., *et al.*,
53 2006. Opsins and clusters of sensory G protein-coupled receptors in the sea urchin
54 genome. Developmental Biology 300, 461-75.
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

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., Filipowski, A., Kumar, S., 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725-2729.

Telford, M.J., Lowe, C.J., 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.

1 Terakita, A., Kawano-Yamashita, E., Koyanagi, M., 2012. Evolution and diversity of
2 opsins. *Wiley Interdisciplinary Reviews: Membrane Transport and Signaling* 1, 104-
3 111.
4

5
6
7 Ullrich-Lüter, E.M., Dupont, S., Arboleda, E., Hausen, H., Arnone, M.I., 2011. Unique
8 system of photoreceptors in sea urchin tube feet. *PNAS* 108, 8367-72.
9

10
11
12 Ullrich-Lüter, E.M., D'Aniello, S., Arnone, M.I., 2013. C-opsin expressing
13 photoreceptors in echinoderms. *Integrative and Comparative Biology* 53, 27-38.
14
15

16
17
18 Vopalensky, P., Pergner, J., Liegertova, M., Benito-Gutierrez, E., Arendt, D., Kozmik,
19 K., 2012. Molecular analysis of the amphioxus frontal eye unravels the evolutionary
20 origin of the retina and pigment cells of the vertebrate eye. *PNAS* 109, 15383-88.
21
22

23
24
25 Whelan, S., Goldman, N., 2001. A general empirical model of protein evolution derived
26 from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.*
27 18, 691-699.
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Figure 1

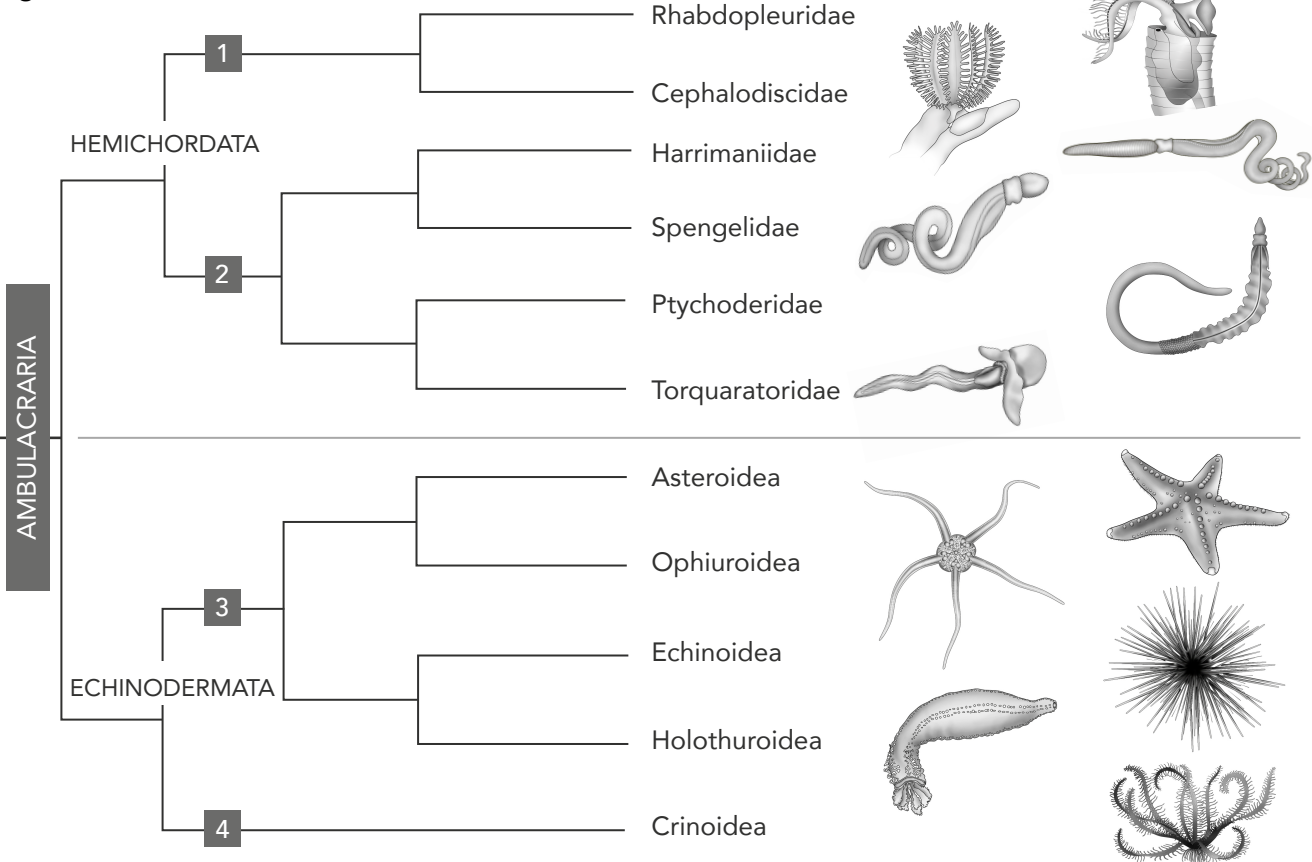


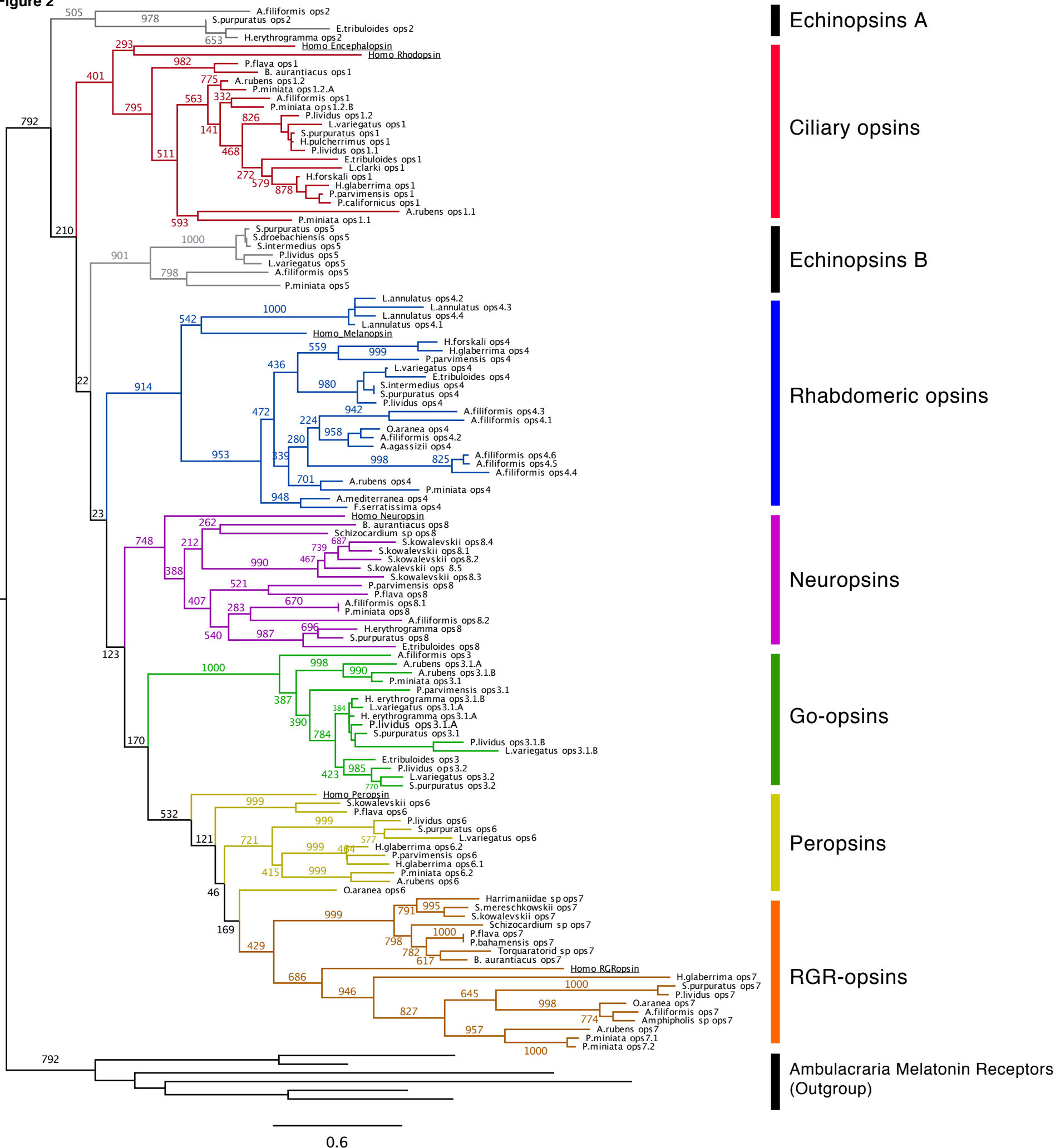
Figure 2

Figure 3

10

20

30

34

ECHINOPSINS A



CILIARY OPSINS



ECHINOPSINS B



RHABDOMERIC OPSINS



NEUROPSINS



Go-OPSINS



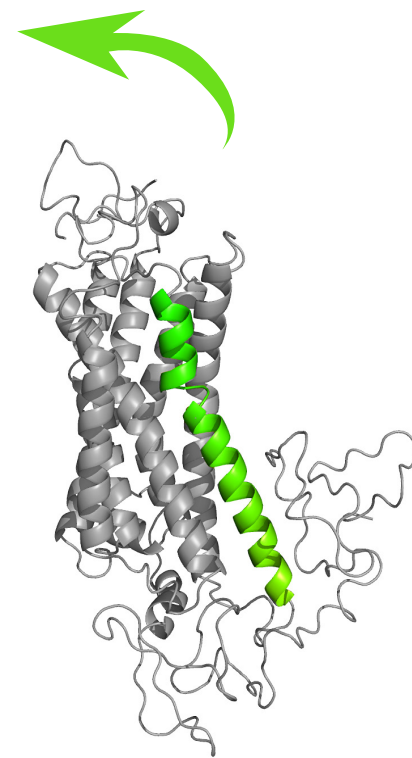
PEROPSINS



RGR OPSINS



OUTGROUP (Melatonin Receptors)



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](#)