Metazoan opsin evolution reveals a simple route to animal vision

Roberto Feuda[a], Sinead C. Hamilton[a], James O. McInerney[a], and Davide Pisani[a,b,1]

[a]Department of Biology, National University of Ireland Maynooth, Kildare, Ireland; and [b]School of Biological Sciences and School of Earth Sciences, University of Bristol, Bristol BS8 1UG, United Kingdom

Edited by David M. Hillis, University of Texas at Austin, Austin, TX, and approved September 13, 2012 (received for review March 21, 2012)

All known visual pigments in Neuralia (Cnidaria, Ctenophora, and Bilateria) are composed of an opsin (a seven-transmembrane G protein-coupled receptor), and a light-sensitive chromophore, generally retinal. Accordingly, opsins play a key role in vision. There is no agreement on the relationships of the neuralian opsin subfamilies, and clarifying their phylogeny is key to elucidating the origin of this protein family and of vision. We used improved methods and data to resolve the opsin phylogeny and explain the evolution of animal vision. We found that the Placozoa have opsins, and that the opsins share a common ancestor with the melatonin receptors. Further to this, we found that all known neuralian opsins can be classified into the same three subfamilies into which the bilaterian opsins are classified: the ciliary (C), rhabdomeric (R), and go-coupled plus retinochrome, retinal G protein-coupled receptor (Go/RGR) opsins. Our results entail a simple scenario of opsin evolution. The first opsin originated from the duplication of the common ancestor of the melatonin and opsin genes in a eumetazoan (Placozoa plus Neuralia) ancestor, and an inference of its amino acid sequence suggests that this protein might not have been light-sensitive. Two more gene duplications in the ancestral neuralian lineage resulted in the origin of the R, C, and Go/RGR opsins. Accordingly, the first animal with at least a C, an R, and a Go/RGR opsin was a neuralian progenitor.

ancestral character state reconstruction | Metazoa | protein evolution

Understanding the origin and early evolution of vision at the molecular level has proven difficult (1–4). Both Protostomia (e.g., Mollusca and Arthropoda) and Deuterostomia (e.g., Vertebrata) have eyes, and it is plausible that the last common ancestor of the Bilateria possessed simple eyespots and some limited ability to detect light (5). In addition, eyes are known in jellyfishes (e.g., refs. 6, 7), and the common use of a Pax-6 regulated kernel [*sensu* Davidson and Erwin (8)] to control eye development in Cnidaria and Bilateria suggests a single origin of the neuralian eye (9). Furthermore, all neuralians for which data are available detect light by using visual pigments composed of an opsin and a chromophore, generally retinal (3), and their opsins link the chromophore through a Schiff base involving a lysine found at position 296 (K296) of the reference bovine rhodopsin sequence (10).

Opsins are seven-transmembrane proteins belonging to the G protein-coupled receptor (GPCR) superfamily (11). According to the glutamate, rhodopsin, adhesion, frizzled/taste2, and secretin (GRAFS) (12) classification system, opsins are members of the α-group of the rhodopsin-like receptors, and they are further classified in several subfamilies (11). Given that the opsins seem to be universally distributed within Neuralia (1, 2, 4, 7, 13), it is clear that, to understand the molecular foundations of vision, we must focus on the early branching metazoans: the Cnidaria, the Ctenophora, the Placozoa, and the sponges. Unfortunately, the phylogenetic relationships of the neuralian opsins are still debated (1–4), and, as a consequence, the early history of gene duplications and deletions within this family is still unknown (*SI Appendix*, Fig. S1). Should we wish to understand the origin of vision (in both its tempo and mode), the pattern of opsin duplications and deletions must be clarified first, and this can only be done by resolving the opsin phylogeny.

The current gap in our understanding of the evolution of vision is, at least in part, the consequence of an absence of genomic information for key, early branching metazoans. Data are still missing for two nonbilaterian lineages: the Ctenophora and the calcarean sponges. However, the genomes of four key taxa, the placoazoa *Trichoplax adhaerens* (14), the cnidarians *Hydra magnipapillata*(15) and *Nematostella vectensis* (16), and the demosponge *Amphimedon queenslandica* (17), have recently been released, improving data availability. Further to this, the genome of *Oscarella carmela*, a representative of a second sponge lineage (the Homoscleromorpha), has now been sequenced (18) and deposited in Compagen (http://compagen.zoologie.uni-kiel.de/).

The relationships among the sponges are still debated (19–23), and two competing hypotheses exist. The first suggests that the sponges are monophyletic (21, 22), whereas the second (19, 20, 23) suggests that they are paraphyletic. According to the sponge monophyly hypothesis, Porifera is the sister group of Eumetazoa, and both the Demospongiae and the Homoscleromorpha are valid outgroups to study the eumetazoan GPCRs (opsins included). According to the paraphyly hypothesis, the Homoscleromorpha is the sister group of the Eumetazoa, and proteins that are most closely related to the eumetazoan GPCRs should be found in this group only. Inclusion of the *Oscarella* genome is thus key to ensure that the closest sister group of the Eumetazoa is being considered when studying GPCR evolution, irrespective of what the relationships among the sponge classes are. Here, genomic information from all aforementioned taxa (*Oscarella* included) was used, together with a large sample of well-characterized neuralian opsins (*SI Appendix*, Table S1), to investigate the origin and evolution of the opsin family and of vision.

Bilaterian opsins have been classified in three major subfamilies (11): rhabdomeric (R) opsins, ciliary (C) opsins, and go-coupled plus retinochrome, retinal G protein-coupled receptor (Go/RGR) opsins. Usually there is an association between light receptors (i.e., the cells expressing these proteins) and specific opsin subfamilies, with the ciliary receptors expressing C and Go/RGR opsins, and the rhabdomeric receptors expressing R opsins (3, 24). A fourth opsin subfamily was suggested by Plachetzki et al. (1). These authors (*SI Appendix*, Fig. S1*A*) identified a large clan (*sensu* ref. 25) of cnidarian-specific opsins that they named Cnidopsins. In addition, they found that one cnidarian opsin in their data set clustered with the bilaterian C opsins, a result that is consistent with the observation that cnidarians have ciliary receptors (24).

Four studies (1–4) have addressed the relationships among the main opsin groups with a view of clarifying the gene duplication and deletion history within this family, but they reached contradictory

Author contributions: J.O.M. and D.P. designed research; R.F. performed research; S.C.H. contributed new reagents/analytic tools; R.F. and D.P. analyzed data; and R.F., S.C.H., J.O.M., and D.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: davide.pisani@nuim.ie.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1204609109/-/DCSupplemental.

results (*SI Appendix*, Fig. S1). A major source of uncertainty in these studies is that three of them (1–3) failed to include a representative sample of cnidarian opsins (*SI Appendix*, Fig. S1 *A*, *C*, and *D*). Accordingly, these studies did not have the power to test every possible hypothesis of opsin evolution. In addition, all four (1–4) used precomputed, empirical time reversible matrices to model amino acid substitutions. These matrices—WAG (1, 2), MtRev (3), and JTT (4)—are unlikely to fit an opsin dataset well because they were not derived from an opsin alignment. Further to this, all the aforementioned studies used uncritically selected outgroups. Plachetzki et al. (1) recognized that the use of problematic outgroups might negatively affect the opsin phylogeny, but failed to find a valid solution to this problem (*SI Appendix*). Consequently, all phylogenies in *SI Appendix*, Fig. S1, are questionable.

Here we performed detailed analyses to better understand opsin evolution. Unlike previous studies, we used modern, well-performing multiple sequence alignment software (26). We implemented better fitting evolutionary models, and considered all available genomic information for the deeply branching metazoans, including the newly sequenced genome of the homoscleromorph sponge *O. carmela*. We thoroughly tested a large sample of putative opsin outgroups and performed analyses by using only the less divergent ones. Most importantly, we used a comprehensive set of cnidarian opsins, including all sequences specific to two previous studies (1, 4). Accordingly, our data set has the power to test every proposed hypothesis of opsin relationships, and its analysis should allow the achievement of greater precision in pinpointing duplications and losses within the opsin family.

Results

Common problems with previous studies (1–4) were the use of under-sampled data sets, substitution models that did not fit the data [precomputed empirical time reversible (GTR) matrices], and inadequate outgroup selection (as detailed earlier). To avoid such problems, we assembled three GPCR and opsin alignments scoring hundreds of sequences (*Methods*), and estimated alignment-specific GTR matrices. Our matrices differ from available, precomputed GTR matrices (*SI Appendix*, Table S2 and Fig. S2), with the Akaike information criterion and Bayesian cross validation showing that they fit the data significantly better than any precomputed GTR matrix, and at least as well as any precomputed site-heterogeneous model (*SI Appendix*, Tables S3 and S4).

Fig. 1*A* represents the phylogeny derived from our all opsin master (AOM) alignment (*Methods*). AOM includes only neuralian opsins (no outgroups), and Fig. 1*A* is thus an unrooted phylogeny of our opsin data set (*SI Appendix*, Table S1). Fig. 1*A* (see also *SI Appendix*, Fig. S3) is consistent with the monophyly of the traditionally recognized bilaterian opsin subfamilies (C, R, and Go/RGR). In contrast, the Cnidarian opsins are split into three clans (hereafter referred to as groups A, B, and C). This is in agreement with the results of Suga et al. (4), but in disagreement with others (1–3). Group A includes only two sequences and sits on the branch separating the R opsins from all the other sequences in our dataset [posterior probability (PP) of 0.84]. The sequences in group A are from the study of Suga et al. (4), in which they were named group 3. These sequences were not included in the other three studies (1–3). Group B forms a relatively poorly supported clan with the Go/RGR opsins (PP = 0.69), whereas group C is found in a polytomy with the C opsins and the Go/RGR plus group B clans (Fig. 1*A*). Group C includes both the sequences that, in the study of Suga et al. (4), emerged as the sister group of the R opsins (their group 2 opsins) and the single sequence that Plachetzki et al. (1) classified as a C opsin. The phylogeny shown in Fig. 1*A* rejects the possibility that Suga et al.'s (4) group 2 opsins could be related to the R opsins. However, it could neither confirm nor reject the C opsin nature of Plachetzki et al.'s (1) putative C opsin. This is because Fig. 1*A* shows that all the aforementioned sequences belong to group C: a group that could not be placed with confidence with reference to the C and the Go/RGR plus group B opsins.

Posterior predictive analysis (*SI Appendix*, Table S5) showed that some of the sequences in AOM were compositionally heterogeneous. Because of their skewed amino acid composition, these sequences can mislead phylogenetic analyses (27). Heterogeneous sequences were included in AOM for the purposes of testing to which major opsin clan they belong. However, most of these sequences were excluded from further analyses (*Methods* and *SI Appendix*) to avoid their potentially biasing effect. Other sequences, such as short expressed sequence tags (ESTs) that, in Fig. 1*A*, were unequivocally identified as members of one of the opsin clans, were also excluded from further analyses.

We analyzed the GPCR and opsin master alignment (G&OM; *Methods*) to test what GPCR family is most closely related to the opsin family. These analyses (Fig. 1*B* and *SI Appendix*, Fig. S4) shown that the neuralian opsins form a monophyletic group. Importantly, the relationships among the neuralian opsins in Fig. 1*B* are consistent with those of Fig. 1*A*. That is, the tree in Fig. 1*B* is a rooted resolution of Fig. 1*A* in which the polytomy from which the C opsins, the Go/RGR plus group B opsins, and the group C opsins stem is resolved according to one of its possible resolutions. Fig. 1*B* also shows that the neuralian opsins are most closely related to a set of placozoan "opsin-like" sequences (PP = 0.98). By turn, the neuralian opsins and the placozoan opsin-like sequences are most closely related to the melatonin (MLT) receptors (PP = 0.89). Fig. 1*B* shows that both the placozoans and the cnidarians have MLT receptors, and, most importantly, that the placozoan opsin-like receptors are orthologues of the neuralian opsins. This implies that from an evolutionary point of view, the placozoan opsin-like receptors are members of the opsin family, even though they lack a retinal binding domain (RBD) with a K296 residue and might thus be unable to detect light. Neither an opsin nor an MLT receptor could be identified in *Oscarella* and *Amphimedon*, and we can thus conclude that both these protein families are eumetazoan specific. This confirms recent results showing that light sensitivity in *Amphimedon* is mediated by a cryptochrome, rather than an opsin (28). Fig. 1*B* shows that the MLT-plus-opsin clade is most closely related to a group including the lysosphingolipid and the orexin receptors (albeit with very low support; PP = 0.46; Fig. 1*B* and *SI Appendix*, Fig. S4). *Oscarella* and *Amphimedon* have sequences belonging to the latter (PP = 0.94; Fig. 1*B*), further confirming the eumetazoan nature of the opsin family.

We tested whether distant outgroups in the G&OM data set could have caused tree-reconstruction artifacts with reference to the opsin phylogeny. To do so, we analyzed the opsins and outgroups (O&O) alignment (*Methods*). The MLT receptors are the sole outgroups of O&O, which also include the placozoan opsin-like receptors. The Bayesian O&O phylogeny is reported in Fig. 1*C* (*SI Appendix*, Fig. S5), and the O&O maximum likelihood (ML) phylogeny is reported in *SI Appendix*, Fig. S6. Analyses of O&O confirmed the results obtained using G&OM (compare Fig. 1*B* vs. Fig. 1*C*). Both data sets show that the Cnidarian opsins can be classified in three groups (A, B, and C). These groups represent, respectively, the cnidarian orthologue of the bilaterian R opsins [group A; GTR PP = 0.89 and ML bootstrap proportion (BP) under an LG plus Γ model = 62%], the cnidarian orthologue of the bilaterian Go/RGR opsins (group B; PP = 0.81 and LG BP < 50), and the cnidarian orthologue of the bilaterian C opsins (group C; PP = 0.71 and LG BP < 50). ML bootstrap support values for the opsin internal relationships are low. Therefore, we used the approximately unbiased (AU) test (29) to evaluate whether the data, under the best-fitting GTR plus Γ model, can discriminate between alternative opsin phylogenies. The results of the AU test (Table 1) confirm that the data are informative and that the trees in Fig. 1*C* fit the O&O data set significantly better than the trees of the aforementioned previous publications (1–4).

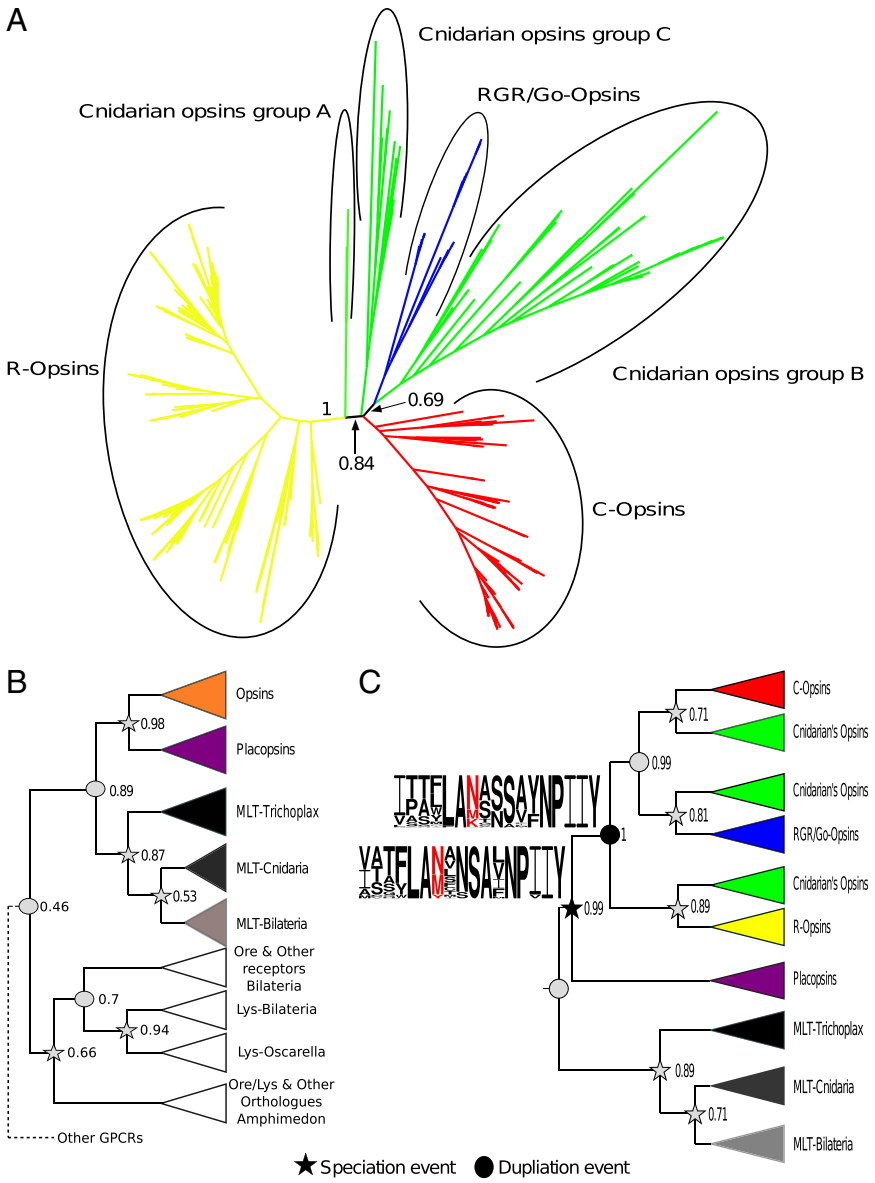

Fig. 1. Phylogeny of the opsin family. (*A*) Unrooted phylogeny of the neuralian opsins. (*B*) Rooted phylogeny of the neuralian opsins and of other GPCRs showing that the Placopsins are members of the opsin family (Ore, orexin; Lys, lysosphingolipid). (*C*) Opsin phylogeny rooted by using only the MLT receptors, and showing that cnidarians have orthologues of each bilaterian opsin subfamily: the C, R, and Go/RGR subfamilies. Support values (Bayesian PPs) are reported only for key nodes (*SI Appendix* shows all support values). The ancestral RBD of the LOCA and of the LOCNA are reported and are identified, respectively, by a black star and a black circle (*SI Appendix*, Fig. S7). The red position in the logos identifies position 296.

To provide further insights into opsin evolution, we carried out Bayesian and ML ancestral character state reconstruction of the RBD at key internal nodes. Results of the Bayesian reconstruction are reported as logos in Fig. 1*C* (*SI Appendix*, Fig. S7), and indicate that the last opsin common ancestor (LOCA) most likely did not have the key K296 residue (PP for K296 = 0.0034). Instead, position 296 was either occupied by an asparagine (PP for N296 = 0.51) or by a methionine (PP for M296 = 0.37). Absence of K296 in LOCA is confirmed by ML, which suggests with reasonable probability (P) that asparagine was the most likely amino acid in position 296 (P-N296 = 0.81 and P-K296 = 0.054). K296 is necessary to link the chromophore, and our results suggest that K296-mediated chromophore binding was not a feature of LOCA: it evolved within the opsin family. Indeed, even in the case of the last opsin common neuralian ancestor (LOCNA), the Bayesian reconstruction suggests that the RBD might not have had a K296 residue (PP for K296 = 0.15; Fig. 1*C*). However, ML contradict this result, as it finds a P-K296 value of 0.99. This incongruence leaves the question of occupancy of position 296 in LOCNA unresolved. No matter what the amino acid in LOCNA was, our results strongly suggest that a K296-based RBD was not a feature of LOCA.

Table 1. Results of AU tests

Hypothesis	Probability
Fig. 1C and *SI Appendix*, Fig. S5	0.7
Plachetzki et al. (1)	0.04
Porter et al. (3)	0.03
Plachetzki et al. (2)	0.008
Suga et al. (4)	5×10^{-18}

Discussion

Our results are markedly different from those of previous investigations. These differences reflect data completeness and methodological dissimilarities. We used a combination of recently developed multiple sequence alignment software that can better differentiate between insertions and deletions and performed extensive model selection analyses, resulting in the use of significantly better-fitting substitution models. We were careful to include the closest outgroups of the neuralian opsins (including sequences from the Placozoa and the Homoscleromorpha), and we used a very inclusive set of cnidarian opsins allowing for the simultaneous test of previous hypotheses (1–4).

Our results (summarized in Fig. 2) allow for a substantial clarification of the tempo and mode of opsin evolution. They confirm the results of Fredriksson et al. (12) that the sister group of the opsin family is represented by the MLT receptors, and they show that the opsin family originated from the duplication of the MLT and opsin ancestral gene in the stem eumetazoan lineage. Importantly, we were able to show that the placozoan genome contains sequences that are in an orthologous relationship with the neuralian opsins. From an evolutionary point of view, these sequences are members of the opsin family (Fig. 1*B* and Fig. 2) irrespective of whether they have the ability to detect light, and we propose to refer to these opsin-like receptors as placopsins. In addition, we show that cnidarians have R, Go/RGR, and C opsin orthologues. Accordingly, these opsin subfamilies must have evolved in the stem neuralian lineage, rather then in the stem bilaterian lineage, i.e., earlier than currently accepted.

Our results are largely phylogeny-independent. Nonetheless, uncertainty in the placement of the Placozoa still persists and deserves discussion. Consistently with our results, some of the most thorough analyses to date (20, 21) agree that the Placozoa are the sister group of Neuralia, even though some investigators (22, 30) found different results. However, Philippe et al. (31) have shown the results of Schierwater et al. (30) to be invalid. Differently, even though the study of Pick et al. (22) is sound, its conclusion that Placozoa is a member of Neuralia is questionable. This is because their dataset (22) is based on that of Dunn et al. (32), which has been shown to be unreliable (21, 22, 31). Importantly, even if Bilateria and Placozoa were confirmed to be sister groups (22), our results would still be valid, but our scenario would become less parsimonious as it would imply independent losses of the placopsin in Bilateria and Cnidaria, and of the C, R, and Go/RGR opsins in Placozoa.

Ancestral character state reconstruction suggests LOCA did not have a RBD containing a K296 residue. Accordingly, K296-mediated light detection most likely evolved in the stem Eumetazoan lineage, perhaps through autogenous evolution and neofunctionalization of a protein that was not light-sensitive. Neither of the two sponge taxa code for MLT or opsin receptors, yet their genomes include sequences clustering in the Lysosphingolipid plus orexin group (i.e., they code for proteins belonging to the sister group of the MLT-plus-opsins clade; Fig. 1*B*). These results confirm that the first opsin originated in the stem eumetazoan lineage, and imply that our conclusions are robust irrespective of whether sponges are monophyletic (21) or paraphyletic (20).

Identification of the duplication of the ancestral MLT plus opsin gene in the stem eumetazoan lineage lets us better constrain the timing of this event, as this lineage was dated to have existed between 755 and 711 Ma (19). In addition, the neuralian stem lineage was dated to have existed between 711 and 700 Ma (19). This relatively short time (11 million y) was a crucial period in opsin evolution. It was during this time that the K296-based RBD most likely evolved and the duplications separating the C plus Go/RGR opsin ancestor from the R opsins, and the C from the Go/RGR opsins, were fixed.

Our results suggest that the Go/RGR opsins represent the sister group of the C opsins. This is in disagreement with some previous findings (1–3), but is in agreement with others (4, 11). An additional line of evidence that seems to support our conclusion is that the Go/RGR opsins, exactly as the C opsins, are expressed in ciliary receptors (3, 24). Our results also predict that rhabdomeric receptors should exist in Cnidaria. This has not yet been proven, but cells with a strong resemblance to the bilaterian rhabdomeric receptors, which could be cnidarian rhabdomeric receptors, have been observed in cnidarian larvae (9, 24, 33).

Conclusions

We suggest an early and parsimonious explanation for the diversification of the opsin family (summarized in Fig. 2), and show that LOCA most likely did not have a K296-based RBD. Scarcity of signal for the deepest event in the history of the opsin family implies that some level of uncertainty in opsin evolution still remains, and might be unavoidable. However, results of the AU tests show that the topology uncovered in this study fits the data (under a GTR-plus-Γ model) significantly better than any previously proposed opsin phylogeny. Our results also indicate that a short 11-million-y period (711–700 Ma) was key in opsin evolution. During this time, two duplications in the stem neuralian lineage resulted in the evolution of the extant opsin paralogues. During this same time, the K296-based RBD most likely evolved, probably through a process of neofunctionalization. Our results are compatible with the view that the last common neuralian ancestor might have been more complex than generally assumed (34).

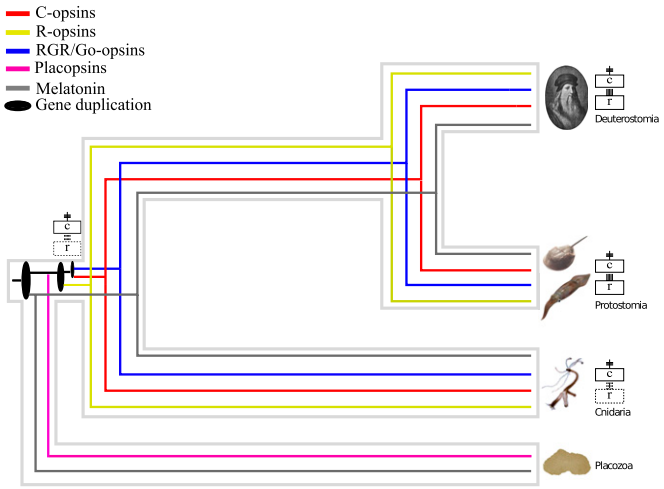

Fig. 2. Synopsis of opsin evolution. This figure represents a gene tree embedded within a species tree illustrating the evolutionary history of the opsins and MLT receptors in Metazoa. It shows that only three duplications and no deletions are necessary to explain opsin evolution.

Methods

Data Mining, Dataset Assembly, and Alignment. Taxonomic nomenclature follows the work of Nielsen (23). We assembled a large sample of well-characterized opsins from across Neuralia (*SI Appendix*, Table S1), including key sequences like the putative cnidarian C opsin of Plachetzki et al. (1) and the putative cnidarian R opsins of Suga et al. (4). To identify the closest outgroup(s) of the neuralian opsins, representatives of each monophyletic α-group of Rhodopsin-like receptors, and a set of sequences from the β-, γ-, and δ-groups (for a total of 139 sequences) were downloaded from GPCRDB (www.gpcrdb.org) and added to our dataset (*SI Appendix*, Table S1). Sequences in GPCRDB are of vertebrate origin. To enrich our data set of putative GPCRs from early branching metazoans, we mined the genomes of *H. magnipapillata*, *N. vectensis*, *T. adhaerens*, *A. queenslandica*, and *O. carmela* (*SI Appendix*). Our final dataset included 625 GPCRs (499 opsins and 176 putative opsin outgroups). From this data set, we generated two master alignments (26) (*SI Appendix*). The first alignment, the AOM alignment, included only the 499 neuralian

opsins. The second alignment, the G&OM, included all putative opsin outgroups (176 GPCRs in total) and a sample of 80 selected opsins (as detailed later; *SI Appendix*). The AOM and G&OM alignments were, respectively, 317 and 366 positions long. A third alignment was generated a posteriori after having inspected the results of the analyses of G&OM (as detailed later; Fig. 1*B*) to identify the closest sister group of the animal opsins. This third alignment, O&O, included the 80 opsins in G&OM plus the closest sister group of the animal opsins only (i.e., the MLT receptors; Fig. 1*B*). O&O included 104 sequences and was 366 positions long. All alignments are available upon request.

Phylogenetic Analyses and Ancestral Character State Reconstructions. In this section, we will focus on the logic of our analytical scheme. Technical details of the analyses performed are reported in *SI Appendix*. The AOM alignment was analyzed to recover an unrooted phylogeny including only well-characterized opsins from the three known bilaterian subfamilies (C, R, and Go/RGR) and an inclusive sample of cnidarian opsins. This analysis allowed the evaluation of the relative relationships among the cnidarian opsins in our data set, including those of Plachetzki et al. (1) and Suga et al (4). Results of the AOM analyses were used to select a subset of 80 opsins (20 C opsins, 20 R opsins, 20 Go/RGR opsins, and 20 cnidarian opsins) to be included in the G&OM and O&O data sets. Opsin subsampling was necessary to (*i*) reduce computational complexity and (*ii*) minimize the likelihood of tree reconstruction artifacts. Accordingly, fast-evolving, extremely short, and compositional heterogeneous sequences were not included in the G&OM and O&O alignments. However, a representative sample of sequences from every opsin clan identified in AOM was retained.

The G&OM alignment was analyzed to identify the closest outgroup of the opsin family. This alignment included the complete set of 176 putative opsin outgroups we identified. Because the closest opsin outgroup must belong to the α-group of Rhodopsin-like receptors, the G&OM phylogeny was rooted by using two γ-group receptors: two Galanin-like receptors (12).

To clarify the duplication and deletion history within the opsin family, we analyzed O&O, which we rooted by using the closest opsin outgroup (identified from the results of the G&OM analyses) only. Accordingly, O&O is simply a modification of G&OM from which distantly related opsin outgroups were excluded to minimize systematic artifacts (20–22, 31, 35).

The three alignments (AOM, G&OM, and O&O) were analyzed by using Bayesian tree reconstruction methods. O&O was also analyzed by using ML. The AU test was used to compare our O&O phylogeny against those from previous studies (1–4). Bayesian and ML-based ancestral character state reconstruction were performed to infer the sequence of the RBD at key internal nodes (LOCA and LOCNA).

ACKNOWLEDGMENTS. We thank Scott Nichols and Nicole King for providing access to the genome of *Oscarella carmela*, Stuart Longhorn for help in assembling our opsin data set, and Omar Rota Stabelli for discussion and suggestions. This work was supported by Irish Research Council for Science, Engineering, and Technology PhD fellowships (to R.F. and S.C.H.); Science Foundation Ireland Research Frontier Programme Grants 11/RFP/EOB/3106 and 09/RFP/EOB2510 (to D.P. and J.O.M.). All analyses were performed using the infrastructures provided by the Irish Centre for High End Computing and the National University of Ireland Maynooth supercomputing facility.

1. Plachetzki DC, Degnan BM, Oakley TH (2007) The origins of novel protein interactions during animal opsin evolution. *PLoS ONE* 2(10):e1054.
2. Plachetzki DC, Fong CR, Oakley TH (2010) The evolution of phototransduction from an ancestral cyclic nucleotide gated pathway. *Proc Biol Sci* 277(1690):1963–1969.
3. Porter ML, et al. (2011) Shedding new light on opsin evolution. *Proc Biol Sci* 279:3–14.
4. Suga H, Schmid V, Gehring WJ (2008) Evolution and functional diversity of jellyfish opsins. *Curr Biol* 18(1):51–55.
5. Land M, Nilsson DE (2002) *Animal Eyes* (Oxford Univ Press, Oxford, UK).
6. Nilsson DE, Gislén L, Coates MM, Skogh C, Garm A (2005) Advanced optics in a jellyfish eye. *Nature* 435(7039):201–205.
7. Kozmik Z, et al. (2008) Assembly of the cnidarian camera-type eye from vertebrate-like components. *Proc Natl Acad Sci USA* 105(26):8989–8993.
8. Davidson EH, Erwin DH (2006) Gene regulatory networks and the evolution of animal body plans. *Science* 311(5762):796–800.
9. Gehring WJ (2011) Chance and necessity in eye evolution. *Genome Biol Evol* 3: 1053–1066.
10. Nathans J, Hogness DS (1983) Isolation, sequence analysis, and intron-exon arrangement of the gene encoding bovine rhodopsin. *Cell* 34(3):807–814.
11. Terakita A (2005) The opsins. *Genome Biol* 6(3):213.
12. Fredriksson R, Lagerström MC, Lundin LG, Schiöth HB (2003) The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol Pharmacol* 63(6):1256–1272.
13. Koyanagi M, et al. (2008) Jellyfish vision starts with cAMP signaling mediated by opsin-G(s) cascade. *Proc Natl Acad Sci USA* 105(40):15576–15580.
14. Srivastava M, et al. (2008) The Trichoplax genome and the nature of placozoans. *Nature* 454(7207):955–960.
15. Chapman JA, et al. (2010) The dynamic genome of Hydra. *Nature* 464(7288):592–596.
16. Putnam NH, et al. (2007) Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science* 317(5834):86–94.
17. Srivastava M, et al. (2010) The Amphimedon queenslandica genome and the evolution of animal complexity. *Nature* 466(7307):720–726.
18. Nichols SA, Roberts BW, Richter DJ, Fairclough SR, King N (2012) Origin of metazoan cadherin diversity and the antiquity of the classical cadherin/β-catenin complex. *Proc Natl Acad Sci USA* 109(32):13046–13051.
19. Erwin DH, et al. (2011) The Cambrian conundrum: early divergence and later ecological success in the early history of animals. *Science* 334(6059):1091–1097.
20. Sperling EA, Peterson KJ, Pisani D (2009) Phylogenetic-signal dissection of nuclear housekeeping genes supports the paraphyly of sponges and the monophyly of Eumetazoa. *Mol Biol Evol* 26(10):2261–2274.
21. Philippe H, et al. (2009) Phylogenomics revives traditional views on deep animal relationships. *Curr Biol* 19(8):706–712.
22. Pick KS, et al. (2010) Improved phylogenomic taxon sampling noticeably affects nonbilaterian relationships. *Mol Biol Evol* 27(9):1983–1987.
23. Nielsen C (2012) *Animal Evolution* (Oxford Univ Press, London).
24. Fain GL, Hardie R, Laughlin SB (2010) Phototransduction and the evolution of photoreceptors. *Curr Biol* 20(3):R114–R124.
25. Wilkinson M, McInerney JO, Hirt RP, Foster PG, Embley TM (2007) Of clades and clans: Terms for phylogenetic relationships in unrooted trees. *Trends Ecol Evol* 22(3): 114–115.
26. Löytynoja A, Goldman N (2008) Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. *Science* 320(5883):1632–1635.
27. Foster PG (2004) Modeling compositional heterogeneity. *Syst Biol* 53(3):485–495.
28. Rivera AS, et al. (2012) Blue-light-receptive cryptochrome is expressed in a sponge eye lacking neurons and opsin. *J Exp Biol* 215(Pt 8):1278–1286.
29. Shimodaira H (2002) An approximately unbiased test of phylogenetic tree selection. *Syst Biol* 51(3):492–508.
30. Schierwater B, et al. (2009) Concatenated analysis sheds light on early metazoan evolution and fuels a modern "urmetazoon" hypothesis. *PLoS Biol* 7(1):e20.
31. Philippe H, et al. (2011) Resolving difficult phylogenetic questions: Why more sequences are not enough. *PLoS Biol* 9(3):e1000602.
32. Dunn CW, et al. (2008) Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* 452(7188):745–749.
33. Nordström K, Wallén R, Seymour J, Nilsson D (2003) A simple visual system without neurons in jellyfish larvae. *Proc Biol Sci* 270(1531):2349–2354.
34. Erwin DH (2009) Early origin of the bilaterian developmental toolkit. *Philos Trans R Soc Lond B Biol Sci* 364(1527):2253–2261.
35. Holton TA, Pisani D (2010) Deep genomic-scale analyses of the metazoa reject Coelomata: evidence from single- and multigene families analyzed under a supertree and supermatrix paradigm. *Genome Biol Evol* 2:310–324.