Available online at www.sciencedirect.com
ScienceDirect
Developmental Biology 300 (2006) 461–475
DEVELOPMENTAL BIOLOGY
www.elsevier.com/locate/ydbio

Opsins and clusters of sensory G-protein-coupled receptors in the sea urchin genome

Florian Raible [a,b,*,1], Kristin Tessmar-Raible [b,1], Enrique Arboleda [c,1], Tobias Kaller [b,d], Peer Bork [a], Detlev Arendt [b], Maria I. Arnone [c,*]

[a] *Computational Unit, EMBL, 69117 Heidelberg, Germany*
[b] *Developmental Unit, EMBL, 69117 Heidelberg, Germany*
[c] *Stazione Zoologica Anton Dohrn, 80121 Napoli, Italy*
[d] *FU Berlin, Institut für Biologie, Chemie, Pharmazie, Zoologie, Evolution und Systematik der Tiere, 14195 Berlin, Germany*

Received for publication 7 June 2006; revised 28 August 2006; accepted 30 August 2006
Available online 5 September 2006

Abstract

Rhodopsin-type G-protein-coupled receptors (GPCRs) contribute the majority of sensory receptors in vertebrates. With 979 members, they form the largest GPCR family in the sequenced sea urchin genome, constituting more than 3% of all predicted genes. The sea urchin genome encodes at least six Opsin proteins. Of these, one rhabdomeric, one ciliary and two G_o-type Opsins can be assigned to ancient bilaterian Opsin subfamilies. Moreover, we identified four greatly expanded subfamilies of rhodopsin-type GPCRs that we call sea urchin specific rapidly expanded lineages of GPCRs (*surreal-GPCRs*). Our analysis of two of these groups revealed genomic clustering and single-exon gene structures similar to the most expanded group of vertebrate rhodopsin-type GPCRs, the olfactory receptors. We hypothesize that these genes arose by rapid duplication in the echinoid lineage and act as chemosensory receptors of the animal. In support of this, group B *surreal-GPCRs* are most prominently expressed in distinct classes of pedicellariae and tube feet of the adult sea urchin, structures that have previously been shown to react to chemical stimuli and to harbor sensory neurons in echinoderms. Notably, these structures also express different *opsins*, indicating that sea urchins possess an intricate molecular set-up to sense their environment.

Keywords: Sea urchin; Rhodopsin-type G-protein-coupled receptor (GPCR); *Opsin*; *Surreal-GPCR*; Gene cluster; Chemosensation; Olfaction; Pedicellariae, Tube feet

Introduction

Echinoderms are slowly moving or even sessile animals with a strong dermal skeleton armament. At first sight, elaborate sensory organs seem to be missing in many echinoderms, including the sea urchin, which has led to the common perception that these animals have only poorly developed senses. For example, referring to chemoreceptive sensation, Aristotle noted that "of the walking or creeping species the urchin appears to have the least developed sense of smell" (*Historia Animalium*, Book IV.8, quoted from the translation of D'Arcy Wentworth Thompson). In contrast to this view, many lines of evidence now indicate that echinoderms react to a wide variety of environmental stimuli, such as light, touch, as well as to chemical cues released from predators or prey (see, e.g., Goldschmid, 1996; Millott, 1975; Pisut, 2004). For example, sea urchins have been described to be able to distinguish between an active, foraging and a passive predator located upstream in the water by adapting their responses accordingly (Phillips, 1978), indicating an intricate sensory circuitry. Moreover, although echinoids are usually viewed as not possessing elaborate sensory organs, their body wall has been shown to contain a variety of sensory neurons, as it is

* Corresponding authors. F. Raible is to be contacted at Computational and Developmental Unit, EMBL, 69117 Heidelberg, Germany. Fax: +49 6221 3878166. M.I. Arnone, fax: +39 081 5833285.
E-mail addresses: raible@embl.de (F. Raible), miarnone@szn.it (M.I. Arnone).
[1] Equal contribution.

0012-1606/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ydbio.2006.08.070

generally typical for echinoderms (Goldschmid, 1996). Up to 4000 sensory cells per square millimeter of skin surface have been reported in asteroids (Smith, 1937). In sea urchins, three main systems have been speculated to sense – and respond to – different cues based on behavioral and histological studies: the 'spine system', the tube feet and the pedicellariae (Campbell, 1973, 1974, 1983; Geis, 1936; Hamann, 1887; Millott, 1954; Peters and Campbell, 1987; Smith et al., 1985).

As far as photosensation is concerned, adult sea urchins exhibit a wide range of responses to light intensity, ranging from shelter seeking, covering reactions and daily migrations to light-dependent oriented movements, and spine defense reactions in response to predators shadowing an individual (reviewed in Millott, 1975; Smith, 1965, also see Barnes and Cook, 2001). In contrast to work on adult sea urchins, only scarce reports exist about the responses of sea urchin larvae to light. *Hemicentrotus* plutei prefer certain light intensities over others (Yoshida, 1966, summarized in Millott, 1975). Larvae of the echinoid *Dendraster* respond to direct sun illumination by avoiding the water surface. This descending behavior has been shown to depend on the ultraviolet light component of the sunlight and has been speculated to be the direct reason for the diel vertical migrations of these larvae (Pennington and Emlet, 1986). Whereas diffuse photoreception suffices for most of these behaviors, spatial vision, requiring more complex optical structures, has been described for the echinoid genus *Echinometra* (Blevins and Johnsen, 2004).

These reports suggest that photosensitivity is a common phenomenon in sea urchins. Some echinoderm classes, such as the Ophiuroidea, have been proposed to possess specialized photosensory organs by utilizing calcite ossicles of the dorsal arm plates as microlenses that bundle and project the beam of light to putative photoreceptors (Aizenberg et al., 2001; Döderlein, 1898; Hendler and Byrne, 1987). In principle, such photosensory structures could also exist in sea urchins but have remained unproven (Aizenberg et al., 2001). Whereas tube feet, pedicellariae, as well as spines have been shown to react in response to light, it has remained unclear if these structures are themselves photosensory or rather depend on photoreceptors located elsewhere on the animal's body (Millott, 1975).

Besides their photosensory responses, sea urchins display a rich chemosensory behavior. Chemical senses have been attributed to predator avoidance and defense, capture of prey, as well as homing (see e.g.Campbell, 1983; Phillips, 1978; Pisut, 2004). For example, the echinoid *Lytechinus variegatus* are able to detect and orient to chemicals emanating from potential food resources over a distance of 1 m, even under turbulent water flow conditions (Pisut, 2004). Similarly, *Strongylocentrotus* sp. is attracted by algae serving as its food over a distance of 1 m in a Y maze experiment (Vadas, 1977). Crushed urchins, tissue pieces of potential predators, as well as living predators located close to diverse sea urchins species trigger an activation of spines, tube feet, as well as pedicellariae (Campbell, 1973; Snyder and Snyder, 1970). Of those systems, the echinoid pedicellariae have been studied with regards to their morphology, sensation and responses, regeneration, fossil record and development (Burke, 1980; Campbell, 1983; Dubois and Ameye, 2001; Geis, 1936; O'Connell et al., 1974; Peters and Campbell, 1987). Four major types, (1) the globiferous, (2) the tridentate, (3) the ophiocephalous and (4) the trifoliate or triphyllous pedicellariae, are commonly distinguished in the literature, which can be further subdivided into additional subclasses (Agassiz and Clark, 1907; Campbell, 1983; Geis, 1936). *Strongylocentrotus purpuratus* possesses all four major types, plus the additional claviform type, which might have arisen from the globiferous type (Burke et al., 2006; Mortensen, 1943). The responses to chemical stimuli differ between different types of pedicellariae (Campbell, 1973, 1974, 1983; Smith, 1965), suggesting that they might harbor different chemosensory receptors types. Although receptor cells have been described to be located within the jaw epithelium of all four major pedicellarian types, real chemoreceptor cells have so far only been attributed to the globiferous type (Peters and Campbell, 1987), which release a poison from their venom glands upon direct chemical contact stimulation of their sensory hillock, a small thickening of tissue rich in chemosensory neurons close to the articulation points of the valves (Campbell, 1976, 1983).

Despite the rich amount of stimuli-dependent behaviors that have been described, and some basic knowledge of putative sensory structures in sea urchins that has been gathered, neither chemo- nor photosensation of the animal are understood at the molecular level. We therefore expected that the completion of the *S. purpuratus* genome might serve as a good opportunity to investigate the molecular basis of these senses and provide hints as to their diversification.

Chemoreception and light reception are – with the main exception of the Trp channels and cryptochrome molecules – mediated by members of the G-protein-coupled receptor superfamilies in vertebrates and invertebrates (Ache and Young, 2005). Among the five main superfamilies of GPCRs distinguished by the GRAFS classification scheme (Schiöth and Fredriksson, 2005), two encode photosensory and chemosensory functions of vertebrates, namely the glutamate-receptor superfamily (pheromone and taste) and the rhodopsin-type superfamily (light, olfaction and possibly pheromones) (Bjarnadóttir et al., 2005; Liberles and Buck, 2006).

The rhodopsin-type superfamily can be further subdivided into up to more than 70 subfamilies. Among these, olfactory receptors are unique because they show the largest differences in copy number and are still rapidly evolving. Whereas orthology between other GPCRs such as the Opsins can be determined across Bilateria, olfactory receptors of invertebrates, besides belonging to the GPCRs superfamily, constitute families on their own and are not clearly related to any other GPCR subfamily (see, e.g., Fredriksson and Schiöth, 2005). Olfactory receptors of vertebrates arose by multiple gene duplication events (reviewed, e.g., in Ache and Young, 2005). Gene duplications have long been implied as a major mechanism generating evolutionary innovations (Kimura, 1983; Ohno, 1970). Moreover, increasing organismal complexity has been

correlated with higher rates of duplicate retention (Yang et al., 2003). In the context of gene regulatory networks, gene duplication in the periphery of a network might be an important factor in generating cellular diversity (Davidson and Erwin, 2006), for instance by allowing new olfactory receptors to evolve that link to a common signal transduction machinery. If such duplications are accompanied by the mutual exclusion of the respective sister genes, gene duplication will facilitate the diversification of cell types.

In contrast to the uncertain evolutionary origin of the olfactory receptors, it is clear that several distinct Opsin families were present at the base of Bilateria, namely the rhabdomeric and the ciliary Opsins (Arendt et al., 2004), as well as the RGR Opsins, with only one clear protostome member so far (Hara-Nishimura et al., 1990; Terakita, 2005). A better understanding of Opsin evolution has proven to be of high value to understand the evolution of bilaterian photosensory structures, including the vertebrate eye (Arendt, 2003). A fourth possible ancient bilaterian family, G_o-Opsins, has been speculated upon based on sequences present in two species, *Patinopecten* and *Amphioxus* (Kojima et al., 1997; Koyanagi et al., 2002; Terakita, 2005), however, with varying support at the branch points. No *opsin*-gene has so far been described for any sea urchin species; however, experiments using antibodies directed against bovine rhodopsin indicated the presence of a ciliary-type Opsin in asteroids and ophiuroids (Johnsen, 1997). This finding, along with the evidence of light-dependent responses of sea urchins, prompted us to screen for *opsin* genes in the sequenced *S. purpuratus* genome. Given the phylogenetic position of the sea urchin at the base of deuterostomes (Delsuc et al., 2006), as well as its peculiar sensory appendages and responses (see above), we speculated that a genome-wide analysis of Opsin sub-families and potential chemosensory receptors would provide an interesting data set for the comparison of sensory systems across animals.

In this study, we use a bioinformatic survey of all rhodopsin-type GPCRs predicted from the sea urchin genome to identify putative light and chemoreceptors of the animal. We characterize the phylogenetic clustering and expression of six *S. purpuratus opsins* as well as the genomic organization, phylogenetic clustering, amino acid divergence and expression of representatives of two fast evolving rhodopsin-type GPCR families. We argue that these receptors are likely to function as chemoreceptors of the sea urchin. In agreement with their presumed function, we find representatives of both *opsins* and the presumed chemosensory receptors specifically and differentially expressed in prominent sensory structures of the adult sea urchin, the tube feet and pedicellariae, but not in other tissues, such as testes, ovaries, axial gland, coelomocytes or gut. Based on our additional analysis of the genomic organization and amino acid divergence of a second fast evolving rhodopsin-type GPCR subgroup, we suggest that the sea urchin has more than one family of chemoreceptors.

We thus provide evidence that the sea urchin uses a complex sensory receptor repertoire localized in decentralized appendages all over its body to analyze its environment.

Materials and methods

Animals, larvae and tissues

Adult *S. purpuratus* were obtained from Pat Leahy (Kerchoff Marine Laboratory, California Institute of Technology, USA). Spawning was induced by vigorous shaking of animals or by intracoelomic injection of 0.5 M KCl, and embryos were cultured at 15°C in Millipore filtered Mediterranean seawater (MFSW) diluted 9:1 in deionized water. Larvae were collected at one and two weeks of development, and either fixed for whole-mount in situ hybridization or used for RNA extraction addressing a time course analysis.

Tissue samples included ampulla, axial gland, coelomocyte, pedicellariae (tridentate, globiferous and trifoliate), gut, neural ring, neural tube, ovary, testis and tube feet. Dissected samples were immediately immersed in the respective reaction buffer for RNA extraction and kept at −80°C. Pedicellariae nomenclature follows that of Campbell, 1983.

RNA extraction

Larvae and tissue samples, kept in the cell-disruption buffer of the "RNeasy micro-kit" (Roche), were quickly frozen (with liquid nitrogen) and thawed three times, with a 1-min vortex step at maximum speed in between. Tissue samples had an additional step in a TissueLyser-Qiagen (Retsch). With a 1-min interval, the samples went trough a 2-min disruption at a frequency of 20 cycles per second. When the amount of solid organic matter still present demanded it, the samples were subsequently centrifuged at 13.000 rpm for 4 min and the supernatant taken. Total RNA from tissues and larvae were obtained following instructions of the "RNeasy micro-kit" (Roche), which include a DNase step. The concentration of RNA was estimated using a NanoDrop ND-1000 spectrophotometer.

Real-time quantitative PCR (QPCR)

First-strand cDNA was synthesized in a 50-µl reaction from 1 µg of total RNA using random hexamers and the TaqMan Reverse transcription Kit (Applied Biosystems) and later diluted to an RNA concentration of 10 ng/µl. Specific primer sets for each gene were designed using the program Primer3 (Rozen and Skaletsky, 2000; http://www.broad.mit.edu/cgi-bin/primer/primer3_www.cgi). Primer sets were chosen to amplify products 100–200 bp in length (see Tables S1 and S2 for details). Amplified PCR products were purified, cloned into pCRII-TOPO (Invitrogen) and fully sequenced on an Automated Capillary Electrophoresis Sequencer 3730 DNA Analyzer (Applied Biosystems, Foster City, CA) to determine their identity.

Reactions were performed in duplicate on Chromo 4 real-time detector (BioRad, Hercules, CA) using SYBR Green chemistry (Applied Biosystems) following the protocol described by Rast et al., 2000. Ubiquitin and 18S ribosomal RNA, which are known to be expressed at constant levels during development (Nemer et al., 1991; Ransick et al., 2002), were used as internal controls, and 18S Ribosomal RNA was used to normalize all data. Reactions typically used 2 µl of synthesized first-strand cDNA template further diluted 1:20, except for 18S RNA that was used at a 1:100 dilution. Primer efficiencies (i.e., the amplification factor for each cycle) were found to exceed 1.9. QPCR results reported in Figs. 2A and 4A are presented as percent of the maximum value. Percent of maximum indicates the percentage of expression of each sample as compared to the maximum level of expression found, among the analyzed samples (larval stages and tissues), for each primer set specific for a given gene. Calculations from QPCR raw data used the formula $1.9^{\Delta Ct}$, where 1.9 is the multiplier for amplification per PCR cycle, and ΔCt is the cycle threshold difference with 18S ribosomal RNA found for that sample.

Whole-mount in situ hybridization (WMISH)

Fragments of *Sp-opsin4* and *Sp-surreal-GPCRB27* were amplified from genomic DNA by PCR using specific primers (Tables S1 and S2). PCR products were purified and cloned into pCRII-TOPO (Invitrogen) according

to the manufacturer's instructions, and the identity of the inserts was determined by sequencing. Whole-mount in situ hybridizations were performed as described by Minokawa et al. (2004). The accuracy of the whole-mount in situ hybridization results was confirmed by control experiments using sense probes. Both antisense- and sense-digoxigenin-labeled RNA probes were obtained using DIG-RNA labeling kit (Roche, Indianapolis, IN), following the manufacturer's instructions by using 1 μg of linearized plasmids. RNA probes were purified using the Mini Quick Spin RNA Columns (Roche). After staining, embryos were mounted in glycerol and analyzed on a Zeiss Axio Imager M1 microscope operating in DIC mode.

Genes, sequence and domain analyses

The sea urchin proteome set that we used was based on the gene predictions of the first sea urchin genome assembly as provided by the sea urchin genome project (Sea Urchin Sequencing Consortium, 2006; http://www.hgsc.bcm.tmc.edu/projects/seaurchin/). This set was supplemented by the changes present in a manually curated data set (freeze 3) from the sea urchin research community. This set comprised a total of 29071 proteins. Proteins containing putative rhodopsin-type GPCR domains were systematically identified by profile hidden Markov model searches using the HMMER package by Sean Eddy (http://www.hmmer.wustl.edu/) and the PFAM model PF00001 (7tm_1). Candidates were compared against the full PFAM database (Finn et al., 2006), and only proteins containing full, non-overlapping domains (e-value cut-off: 0.001) were included in the final list of 979 peptides. For overall cross-comparisons, 7tm_1 domains were extracted from these peptides and compared using all-against-all Smith/Waterman alignments. Similarities between any pair of sequences were recorded as ratios of the sums of hit scores and self-hit scores. To obtain insight into the self-similarity of the sea urchin GPCR pool, this matrix was subjected to hierarchical clustering and visualized using the Cluster3/Treeview software (de Hoon et al., 2004; Saldanha, 2004).

To look for evidence for the transcription of the identified genes, systematic BLASTN searches (Altschul et al., 1997) were carried out against 130746 publicly available EST sequences (downloaded from the NCBI dbEST archive) as well as 15,275 additional EST sequences generated by the sea urchin sequencing consortium.

To test for distant relationships among GPCR domains, we aligned core sequences of each subgroup using the MUSCLE software (Edgar, 2004) and used this alignment to generate profile hidden Markov models (HMMs), using hmmbuild in hmmls multiple domain alignment mode. Profile HMMs were compared against the swissprot/uniprot/trembl reference database using hmmsearch. Assignments to distinct rhodopsin-type GPCR subfamilies were retrieved from the Interpro database (Mulder et al., 2005).

Phylogenetic analyses

Multiple sequences alignments were generated using MUSCLE (Edgar, 2004) and analyzed using neighbor-joining and maximum likelihood algorithms included in the CLUSTAL and TREE-PUZZLE 5.2 software packages, respectively (Schmidt et al., 2002; Thompson et al., 1994). For the TREE-PUZZLE analysis, we used the BLOSUM62 substitution model and 8 gamma rate categories.

GPCR amino acid divergence analyses

Core sequences of the group B and group C *surreal-GPCRs* were aligned using CLUSTAL, and consensus positions for the seven transmembrane regions was inferred from separate predictions for four representative sequences of each group (SPU_017092, SPU_013023, SPU_013027 and SPU_010318 for group B; SPU_001950, SPU_004502, SPU_019376 and SPU_008790 for group C) using the TMpred and TMHMM algorithms (Hofmann and Stoffel, 1993; http://www.ch.embnet.org/software/TMPRED_form.html; http://www.cbs.dtu.dk/services/TMHMM/). Similarities and divergence between one of the sequences (SPU_013027 for group B; SPU_001950 for group C) and the other sequences were scored for each position of the alignment.

Results

Rhodopsin-type GPCRs constitute the largest GPCR superfamily in the sea urchin genome

We started our analysis with a computational search for sea urchin proteins carrying a rhodopsin-type GPCR domain. By our approach, we could identify 979 rhodopsin-type GPCRs in the *S. purpuratus* proteome predicted from the first genome assembly. These are slightly more rhodopsin-type GPCRs than predicted by the independent analysis of Cameron et al. (2006). Rhodopsin-type GPCRs are by far the largest GPCR superfamily in the sea urchin genome, followed second by the secretin-type GPCR superfamily with only one sixth of the number (Cameron et al., 2006). The absolute number and the frequency of sea urchin rhodopsin-type GPCRs is thus comparable to several mammalian genomes, consistent with a rich molecular repertoire of sensory molecules in this species.

The sea urchin genome encodes six Opsins and supports the presence of an ancestral bilaterian G_o-Opsin family

So far, two Opsin families, the rhabdomeric Opsins (r-Opsins) and the ciliary Opsins (c-Opsins), have been determined to be ancestral for bilaterians (Arendt et al., 2004; Hill et al., 2002). Evidence for two additional ancient, pan-bilaterian Opsin families (the retinochrome or RGR subfamily and the G_o subfamily) exists from studies in Lophotrochozoa and Amphioxus (Hara-Nishimura et al., 1990; Kojima et al., 1997; Koyanagi et al., 2002; Terakita, 2005). However, whereas the RGR-Opsin subfamily, comprising vertebrate members and a squid molecule appeared highly supported, the existence of an additional G_o-Opsin subfamily has so far been lacking support because the *Amphioxus* as well as the *Patinopecten* putative G_o-Opsins have strongly drifted in phylogenetic trees, resulting in often unsupported branch points (data not shown).

Among the predicted sea urchin genes, six encode bona fide Opsins (Fig. 1A). Four of these are also reported in the independent publication by Burke et al. (2006). These rhodopsin-type GPCRs are Opsins based on three characteristics: (1) the presence of the key lysine residue necessary for the Schiff base formation with retinal (Fig. 1B); (2) amino acids specifically shared among Opsin (Fig. 1B and data not shown); and (3) their well-supported clustering in phylogenetic trees (Fig. 1A and data not shown). Phylogenetic analysis using neighbor joining and maximum likelihood methods, combined with an examination of the diagnostic amino acid fingerprints (Arendt et al., 2004), reveals that at least three pan-bilaterian Opsin families are present in the sea urchin (Fig. 1A). *Sp*-Opsin1 belongs to the c-Opsin family, whereas *Sp*-Opsin4 is a clear member of the r-Opsin family (Figs. 1A, B), confirming that both Opsin subfamilies were present at the beginning of Bilateria. Strikingly, two Opsins, *Sp*-Opsin3.1 and *Sp*-Opsin3.2, form one well-supported group with the *Patinopecten* G_o-Opsin (Kojima et al., 1997) and two Opsins from *Amphioxus* (Koyanagi et al., 2002) (Fig. 1A), thus supporting the existence of an ancestral bilaterian G_o-Opsin family.

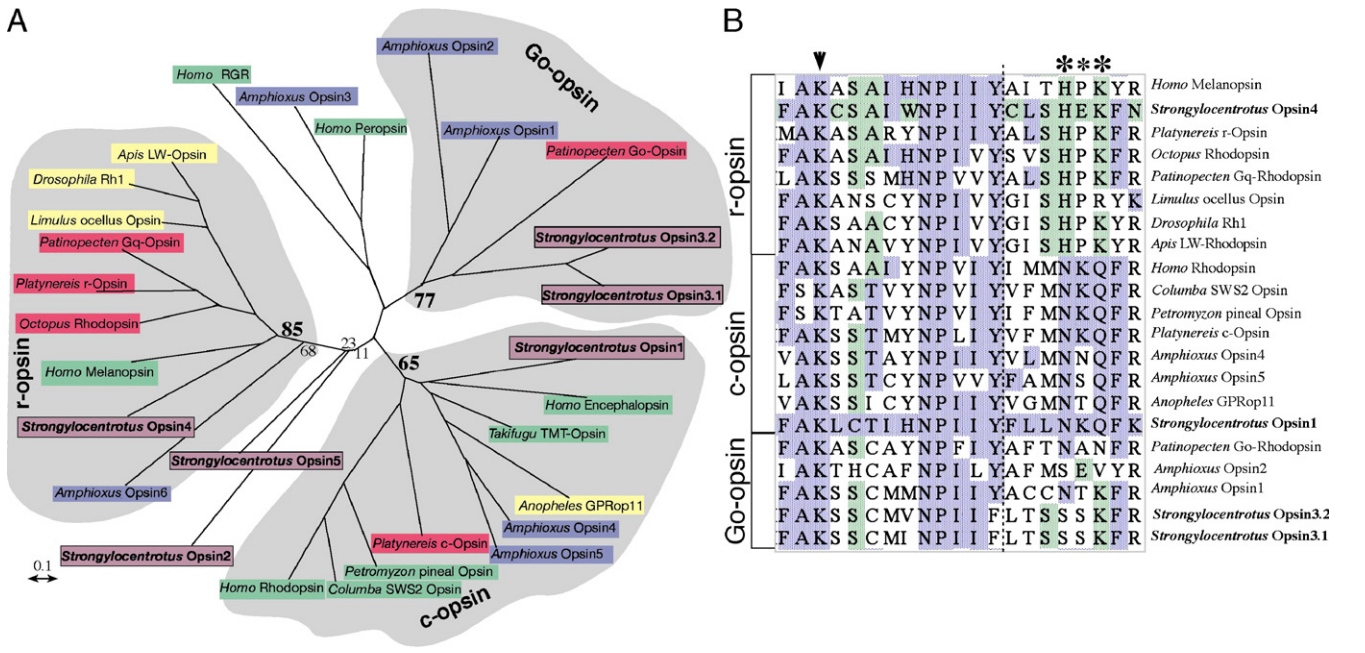


Fig. 1. Phylogenetic analysis and partial alignment of sea urchin Opsins. (A) Representative bilaterian Opsin members cluster into three significantly supported subfamilies in a maximum likelihood analysis. Branch support values are indicated next to crucial branching points. Branch length scale bar indicates relative amount of amino acid changes. Sequences are color-coded to indicate their phyletic association to Vertebrata (green), Lophotrochozoa (red), Ecdysozoa (yellow), Cephalochordata (blue) and Echinodermata (purple), respectively. (B) Amino acid alignment of members of the three Opsin subfamilies shown in panel A. The depicted alignment localizes to the border (dotted vertical line) between the 7th transmembrane domain and the C-terminal tail. Amino acid residues shared with the putative sea urchin c-Opsin (*Sp*-Opsin1) and r-Opsin (*Sp*-Opsin4) members are highlighted in blue and green, respectively. Asterisks demarcate a stretch of three amino acids in the C-terminus highly indicative of the c-Opsin and r-Opsin families. Note that no such fingerprint is detectable for the Go family at this position. Arrowhead demarcates the position of the Lysine residue critical for Schiff base formation.

The *S. purpuratus* genome also contains a G_o G-protein (SPU_007553), consistent with the possibility that the link between G_o-Opsin and the G_o G-protein is evolutionarily conserved. Remarkably, no vertebrate, insect or nematode G_o-Opsin family member is present in the sequence databases to date, indicating that the G_o-Opsin family might be connected to cell types present in marine species with primary ciliated larvae. The existence of two G_o-Opsin family members in the sea urchin is likely to be due to a recent duplication. This is suggested by the close similarity of the *Sp*-Opsin3.1 and *Sp*-Opsin3.2 sequences and is further supported by the adjacent position of these genes in the sea urchin genome. The two remaining sea urchin *opsins*, *Sp*-Opsin2 and *Sp*-Opsin5, show no clear affiliation to any of the larger Opsin subfamilies in phylogenetic and sequence analyses (Figs. 1A, B). Therefore, these proteins are likely to be highly derived members of one of the known groups.

We next studied the expression of the sea urchin *opsin* genes in larvae and adult tissues. Expression was assayed by QPCR with primers for *opsins 1*, *2*, *4*, *5* and *3.1*. Our analysis revealed expression for all the tested *opsins*. Except for *Sp-opsin5*, transcription of all *opsins* could already be detected in one week old larvae, the earliest developmental stage investigated (Fig. 2A). In adult sea urchins, all five tested *opsins are* expressed. Notably, the sampling of different tissues revealed that all *opsins are* most prominently transcribed in pedicellariae (Fig. 2A). To test if different pedicellaria classes showed different expression profiles, we sampled three out of the four major pedicellaria types present in sea urchin. This analysis revealed that all three classes express different *opsin* subsets at different levels (Fig. 2A). Besides their prominent expression in pedicellariae, transcripts of *Sp-opsin1*, *Sp-opsin2*, *Sp-opsin3.1* and *Sp-opsin4* were also present at high levels in the tube feet (Fig. 2A), structures that in asteroids have been described to bear the highest density of sensory cells at their tip (Smith, 1937). In addition, neural ring and tube express low levels of *Sp-opsin1*, whereas no *opsin* transcripts were detected in testis, ovary, ampulla, axial gland, coelomocyte and gut (Fig. 2A).

As a representative of the rhabdomeric *opsin* family, *Sp-opsin4 is* orthologous to the GPCR type present in the eyes of primary ciliated larvae (Arendt et al., 2002). Because the sea urchin pluteus larva represents such a primary ciliated larva, but has not been described to possess eyes, we were intrigued by the prominent larval expression of *Sp-opsin4* in our QPCR analysis. Therefore, we complemented our QPCR experiments by analyzing the spatial distribution of *Sp-opsin4* in the larva using whole-mount *in situ* hybridization (WMISH). This analysis revealed that *Sp-opsin4 is* specifically expressed in 1–2 cells close to the ciliated band at the tip of the postoral arms of two week old larvae (Figs. 2B3, 4, 5). Further analysis will show if these *Sp-opsin4* expressing cells resemble rhabdomeric photoreceptor cells by cellular morphology, and if any pigmented structures enable them to detect the direction of light.

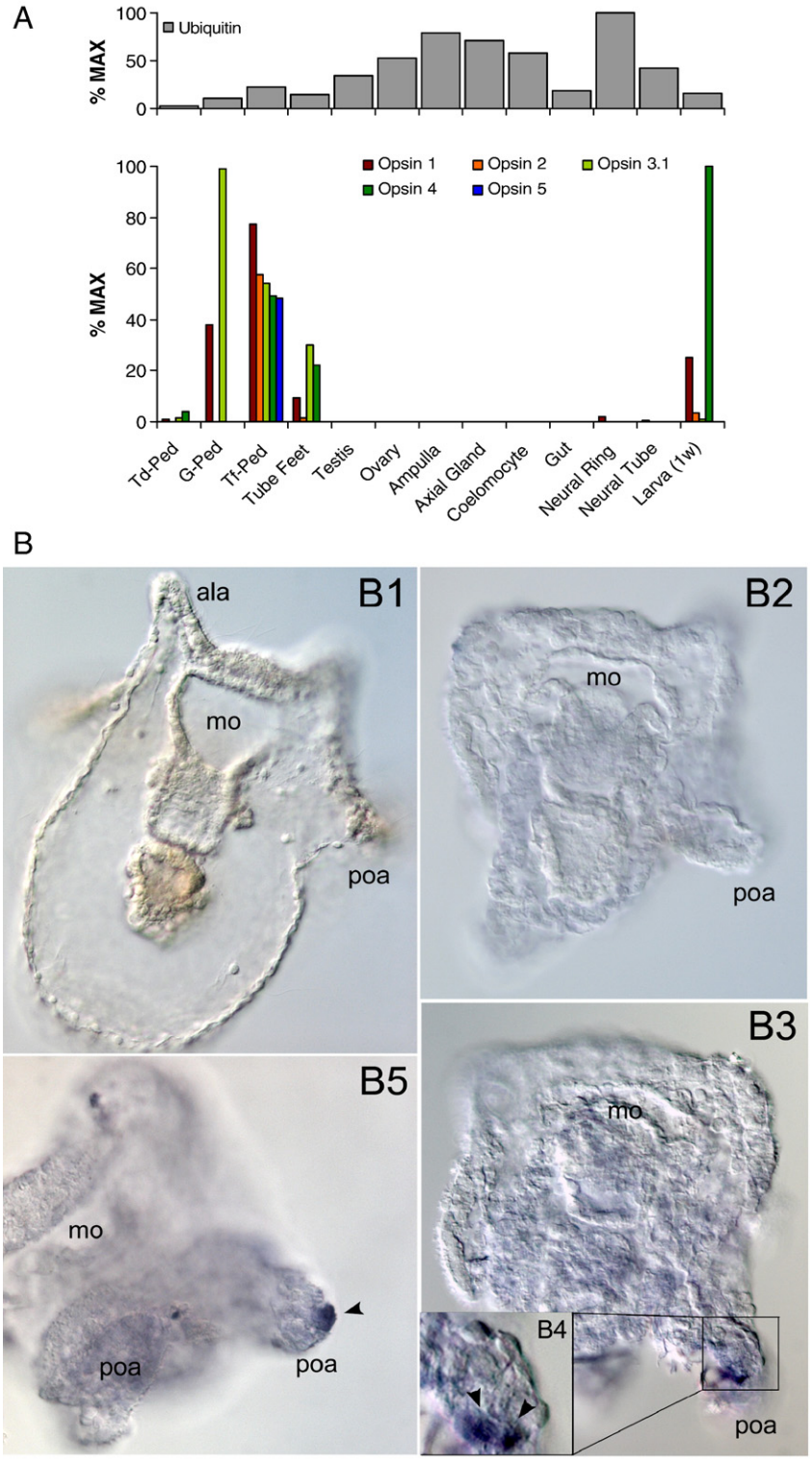


Fig. 2. Expression of *opsins* in the sea urchin larva and adult. (A) Expression of *opsins* in one-week-old larvae and adult tissues as detected by QPCR. QPCR data were normalized versus 18S ribosomal RNA C_Ts and results were expressed as percent of the maximum (% of MAX) as described in Materials and methods. The values of relative expression of *ubiquitin* normalized versus *18S ribosomal RNA* are reported as reference for each sample. Td-Ped, G-Ped and Tf-Ped are tridentate, globiferous and trifoliate pedicellariae, respectively. (B) Expression of *Sp-opsin4* as detected by whole-mount in situ hybridization in two-week-old larvae. (B1) Control fixed specimen viewed along the digestive tract, mouth on top, to display morphology before the hybridization procedure. (B2) Larva hybridized with sense probe shown as a control in the same orientation as in panel B1. (B3–5) Larvae stained with antisense probe. Positive cells are marked with arrowheads. Ala, anterolateral arm; mo, mouth; poa, postoral arm.

The sea urchin genome shows large specific expansions of GPCR lineages

Besides the Opsins, the rhodopsin-type sea urchin GPCRs contain other clear counterparts of functionally characterized GPCRs (Burke et al., 2006; Whittaker et al., 2006, and see below). However, we noted that many of the predicted molecules showed only distant similarity to known bilaterian GPCRs. To obtain a better insight into the interrelations between all rhodopsin-type GPCRs in the sea urchin, we performed a systematic cross-comparison of these receptors, both within the sea urchin genome and with the receptors present in the swissprot/uniprot/trembl database. In order to reduce the influence of neighboring protein domains and artifacts caused by incomplete or wrong gene models, we focused this analysis only on the protein regions identified by the Pfam model for rhodopsin-type GPCRs.

In the first analysis, we assigned each of the 979 rhodopsin-type GPCRs to its best hit in the swissprot/uniprot/trembl sequence database. 5 GPCRs did not match any known protein within the cut-off range (e-value of 0.001). Of those that had counterparts in the reference database, only 55 were also best reciprocal hits. This already indicated that a major proportion of sea urchin rhodopsin-type GPCRs was not part of 1:1 orthologous relationships. To obtain a further classification of the sea urchin rhodopsin-type GPCR repertoire, we also recorded, where possible, the specific rhodopsin-type GPCR subfamily that each hit belonged to. For this, we used the assignment of hits to the 77 subfamilies currently distinguished in the Interpro database (Mulder et al., 2005). This analysis showed that sea urchin GPCRs bear similarity to a broad range of diverse receptor subfamilies (Table S3). However, regarding the remote similarity that many sea urchin GPCRs display to their cognates in other animals, it appeared likely that they have also lost or modified important characteristics of these groups.

In our second analysis, we therefore addressed the characteristics and mutual similarities of sea urchin GPCRs more directly. Using hit score/self-hit score ratios as a measure of the mutual similarity of each GPCR pair, we performed all-against-all comparisons of the sea urchin rhodopsin-type GPCR domains. This analysis revealed several groups of sea urchin GPCR domains with high mutual similarity (Fig. 3A). Moreover, these GPCR domains are more closely related to other sea urchin GPCRs than to those of other organisms, indicating that they belong to paralogous groups. Together, these findings show that the sea urchin rhodopsin-type GPCR repertoire contains several largely expanded subfamilies that appear to be specific to the sea urchin lineage. We thus tentatively call these receptors "sea urchin specific rapidly expanded lineages of *GPCRs*", or *surreal-GPCRs* (Fig. 3). Because of the fast sequence evolution of these groups, we can presently not decide if all of these groups are mono- or polyphyletic. We suggest the common name '*surreal-GPCRs*' thus to emphasize the fact that all groups are fast evolving, without inferring any additional relationships between the groups.

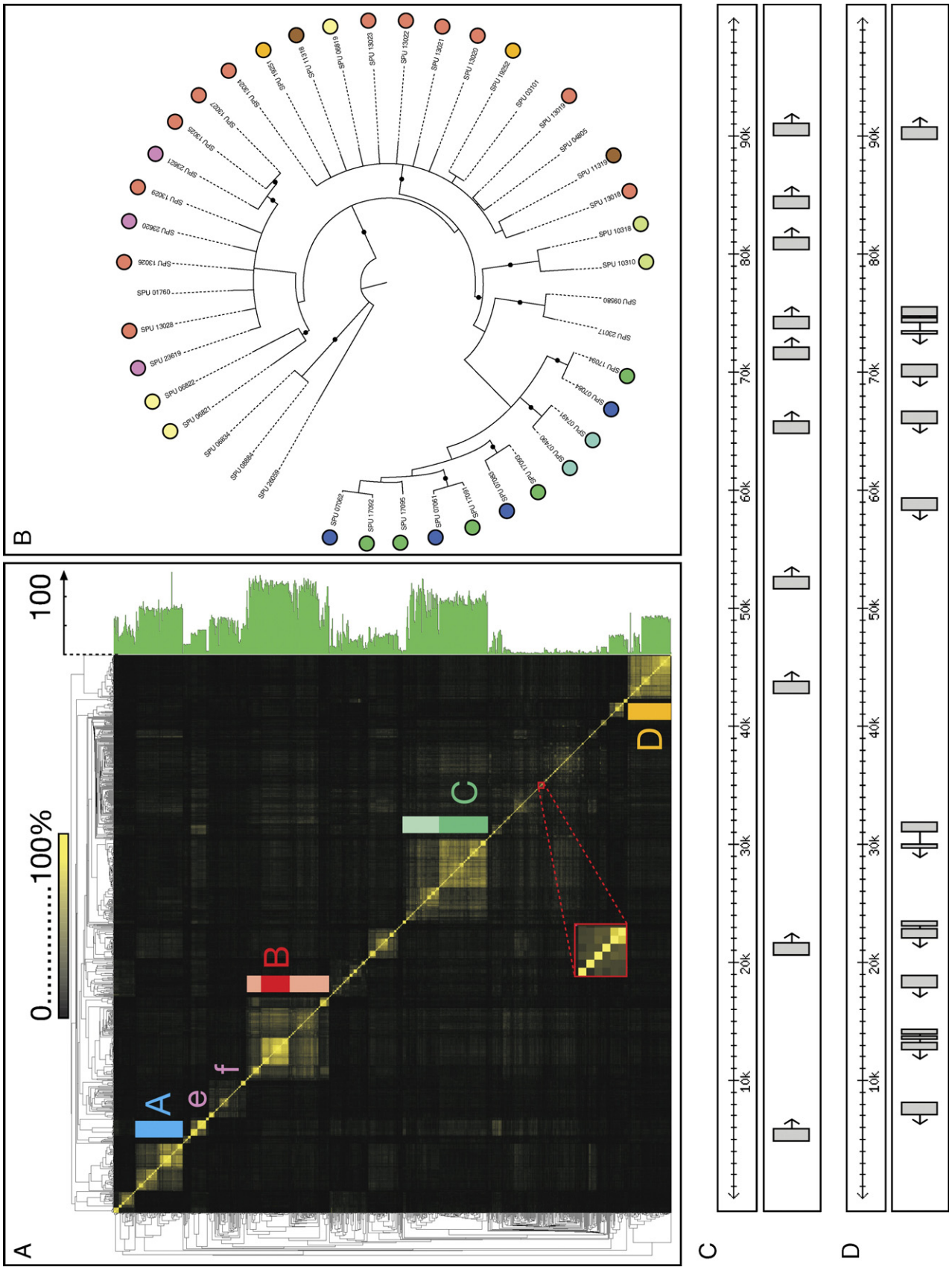
Based on their reciprocal similarity profiles, we can distinguish four major groups of *surreal-GPCRs* in the sea urchin genome (labeled A–D in Fig. 3A), and multiple small ones (exemplified as e and f). Consistent with a fast evolution of these receptor groups, the BLAST significance values for many protein hits in the reference database are comparably weak.

A more sensitive search, using profile hidden Markov models (profile HMMs), provided evidence that group C is most similar to the opioid receptor subfamily (IPR001418), suggesting that the ancestor of this group lay within this subfamily. The model derived from group D *surreal-GPCRs* showed highest similarity to a pond snail protein assigned as glycoprotein hormone receptor (GR101_LYMST), and, although with lower significance, to various vertebrate receptors of the relaxin receptor group (IPR008112). For group A *surreal-GPCRs*, the profile HMM search produced weak, yet similarly distant relationships to both muscarinic acetylcholine receptors and serotonin receptors. Finally, the profile HMM for group B *surreal-GPCRs* showed the least significant assignments. The first non-sea urchin GPCR picked up by the profile HMM was a human receptor, ADA1B_HUMAN, belonging to the adrenergic receptor GPCR subfamily (IPR002233), closely followed by representatives of other groups such as Opsins (IPR001760), dopamine receptors (IPR000929) or melatonin receptors (IPR000025).

These data suggest that the four main *surreal-GPCR* groups have arisen independently and from different members of the ancestral bilaterian GPCR repertoire. However, similar to the olfactory receptors of vertebrates, the ancestral molecules from which the different *surreal-GPCRs* arose are difficult to determine, resulting in distinct GPCR variants with little similarity to any existing classification scheme.

Surreal GPCRs are highly variable and found in genomic clusters

To further study the diversification of *surreal-GPCRs*, we examined the distribution of variable sites with respect to the putative structures of these receptors, focusing on the group B and group C *surreal-GPCRs*. We inferred a consensus structure for the 43 core sequences of the group B *surreal-GPCRs* and compared sequence variability in the different domains. This comparison revealed that variable positions (defined as residues not shared among 60% of the sequences) were distributed across all seven-transmembrane domains (Fig. 4A). Additional strong variability was seen in the N- and C-terminal regions as well as in the third intracellular loop between TM V and VI, all of which also carried length polymorphisms. Group C *surreal-GPCRs* exhibited a different variability pattern. This group shows an overall higher divergence than the group B *surreal-GPCRs* (Fig. 4B). Moreover, the rapidly changing amino acid residues are not distributed equally across the seven transmembrane domains (as is the case for the group B *surreal-GPCRs*), but preferentially cluster in the third to sixth transmembrane helix, with the TM IV and TM V being the most variable helices (Fig. 4B).


A
0.....100%
100
A
B
C
D
e
f
B
SPU 13023
SPU 13022
SPU 13021
SPU 13029
SPU 13032
SPU 03101
SPU 13019
SPU 04605
SPU 11316
SPU 13018
SPU 10318
SPU 10310
SPU 09580
SPU 23017
SPU 17094
SPU 07084
SPU 07481
SPU 07480
SPU 17093
SPU 17091
SPU 01061
SPU 17099
SPU 17092
SPU 07082
SPU 26659
SPU 08884
SPU 08834
SPU 08821
SPU 08822
SPU 23619
SPU 13028
SPU 01760
SPU 13026
SPU 23620
SPU 13029
SPU 23621
SPU 13025
SPU 13027
SPU 13024
SPU 13031
SPU 11319
SPU 08813
C
10k
20k
30k
40k
50k
60k
70k
80k
90k
D
10k
20k
30k
40k
50k
60k
70k
80k
90k

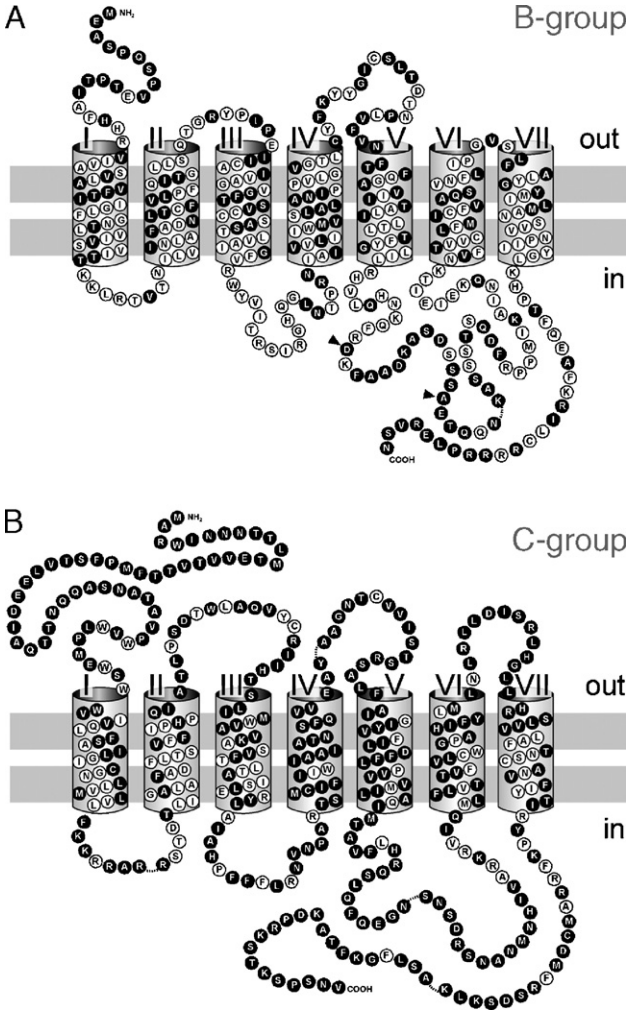


Fig. 4. Variability of amino acid residues in group B and group C *surreal-GPCRs*. (A) Diversity of group B *surreal-GPCRs*; reference: SPU_013027 (*surreal-GPCR-B27*). (B) Diversity of group C *surreal-GPCRs*; reference: SPU_001950. Both schemes show the amino acid sequence of one representative member of each group projected onto a consensus structure. The seven predicted transmembrane regions are displayed as vertical cylinders (I–VII). Amino acids shared with 60% or more of the other group B or group C *surreal-GPCRs*, respectively, are indicated as white balls. Less conserved positions are highlighted in black. Length polymorphisms are also indicated: Hatched lines mark positions at which other members of the group differ by the insertion of 5 or more amino acids; arrowheads in panel A demarcate an amino acid stretch missing in several other receptors of the same group.

In both group B and group C *surreal-GPCRs*, the third intracellular loop (ic3) as well as the C-terminal tail show a high divergence, as common for GPCRs in general (Mathi et al., 1997). The ic3 has been proposed to participate in G-protein binding, but also, like the C-terminus, to interact with G-protein interacting proteins (GIPs), which are implicated in the fine-tuning of receptor signaling (Bockaert et al., 2004). In line with such a role, we find a dopamine receptor interacting protein 78 (DRIP78) binding sequence (FxxxFxxxF) (Bermak et al., 2001) conserved in the C-termini of a subset of sequences (SPU_006834, SPU_008884, SPU_010318, SPU_010318; SPU_000576, SPU_015119, SPU_006431). In the case of the mammalian dopamine D1 receptor, the binding of DRIP78 has been suggested to influence receptor glycosylation and ligand binding kinetics (Bermak et al., 2001), providing an interesting possibility how the sensitivity of sea urchin group B and group C *surreal-GPCRs* might in part be regulated.

Next, we investigated if these genes appeared in single copies or in genomic clusters. Out of 43 core group B *surreal-GPCR* sequences predicted from the first genome assembly, all but 8 were found in direct neighborhood to at least one other *surreal-GPCR* gene (Fig. 3B). The largest cluster in the first genome assembly was found to contain 12 genes in tandem orientation on scaffold 548 (highlighted by red circles in Fig. 3B). Without exception, the 8 remaining, non-clustered genes were either located on small genomic scaffolds, or they were lying at the border of medium-sized scaffolds containing a maximum of 3 other genes. This suggested that some additional genomic clustering of *surreal-GPCRs* was not detectable due to incomplete assembly of the genome.

In the second genome assembly, similar fragmentation was observed. One scaffold (Scaffold_v2_21524) was found to contain 10 group B *surreal-GPCR*genes in a tandem array of single-exon genes, covering a locus of around 100 kb (Fig. 3C). Similarly, another scaffold (Scaffold_v2_66648) harbored 9 genes related to group C *surreal-GPCRs* in tandem duplication, in addition to one reversed gene (Fig. 3D). In both cases, these clusters are located at one end of the scaffold. We hypothesize that these clusters might be fragmented parts of larger genomic clusters that contain more of *surreal-GPCR* genes, a notion that might be testable after future polishing of the genomic assembly.

A characteristic feature of group B *surreal-GPCRs* that is possibly linked to their expansion is that many of them are encoded by single exons, including the members in the genomic cluster mentioned above. The intronless gene structure is also found in most group C *surreal-GPCRs* (see Table S3). Group A *surreal-GPCRs* display a lower proportion of single-exon genes, whereas group D has many genes with a high number of exons. We noted that group D *surreal-GPCRs* also possess the highest

Fig. 3. The *S. purpuratus* genome contains *s*ea *ur*chin specific *r*apidly *e*xpanded *l*ineages of *GPCRs* (*surreal-GPCRs*). (A) Clustered matrix showing the pairwise similarities between all 979 investigated rhodopsin-type sea urchin GPCR domains, expressed as hit/self-hit score ratios. For each GPCR domain, a green bar at the right side represents the number of sea urchin proteins that display a higher similarity than the best protein hit in the swissprot/uniprot/trembl reference database, thus providing an estimate of the size of paralog groups. Four major expanded lineages (group A–D *surreal-GPCRs*) are indicated with colored bars, and two smaller ones (e, f) are marked with small letters. For *surreal* groups B and C, core domains used for the generation of profile hidden Markov models are shaded in darker color. The magnified insert shows the six sea urchin *opsins*, with the two *opsin3* paralogs visible as bright quadruplet in the right lower corner. (B) Relationships among the core members of the *surreal-GPCR* group B as revealed by phylogenetic analysis. Nodes supported with bootstrap supports greater than 80% are marked in black. Genes that locate to clusters on the same genomic scaffold are coded in identical colors (C, D) Genomic clustering of *surreal-GPCRs*. (C) Ten representatives of group B *surreal-GPCRs* are shown in their genomic location on Scaffold_v2_21524. Coordinates (in kbp) of the scaffold are indicated on top of the annotated *surreal-GPCRs* (grey). Note that all genes are oriented in tandem orientation. (D) 10 genes related to group C *surreal-GPCRs* on Scaffold_v2_66648.

similarity to GPCRs of other organisms. One explanation for this apparent correlation is that class D receptors are under more evolutionary constraints. Alternatively, they might have diversified only in recent evolutionary time.

In turn, intronless sea urchin genes might be enriched for fast-evolving genes. A survey among all rhodopsin-type GPCRs revealed that more than 60% of all sea urchin rhodopsin-GPCRs are single-exon genes, many of which are only remotely related to receptors of other animals (Table S3). This proportion is in stark contrast to the genomic average of single-exon genes that we calculate to be 15% of all predicted gene models. The fast-evolving intronless genes are also found outside of the major clusters, for instance in the minor classes e and f (Fig. 3A). In both groups, we also detected genomic clustering, in one case (e) including a total of twelve closely related sequences. Notably, the vertebrate olfactory receptors are also characteristically single-exon genes that locate to large genomic clusters (reviewed in Dryer, 2000).

Surreal-GPCRs are expressed in larval and adult stages

The presence of largely expanded groups of rhodopsin-type GPCRs and the remarkable parallels to expanded GPCR

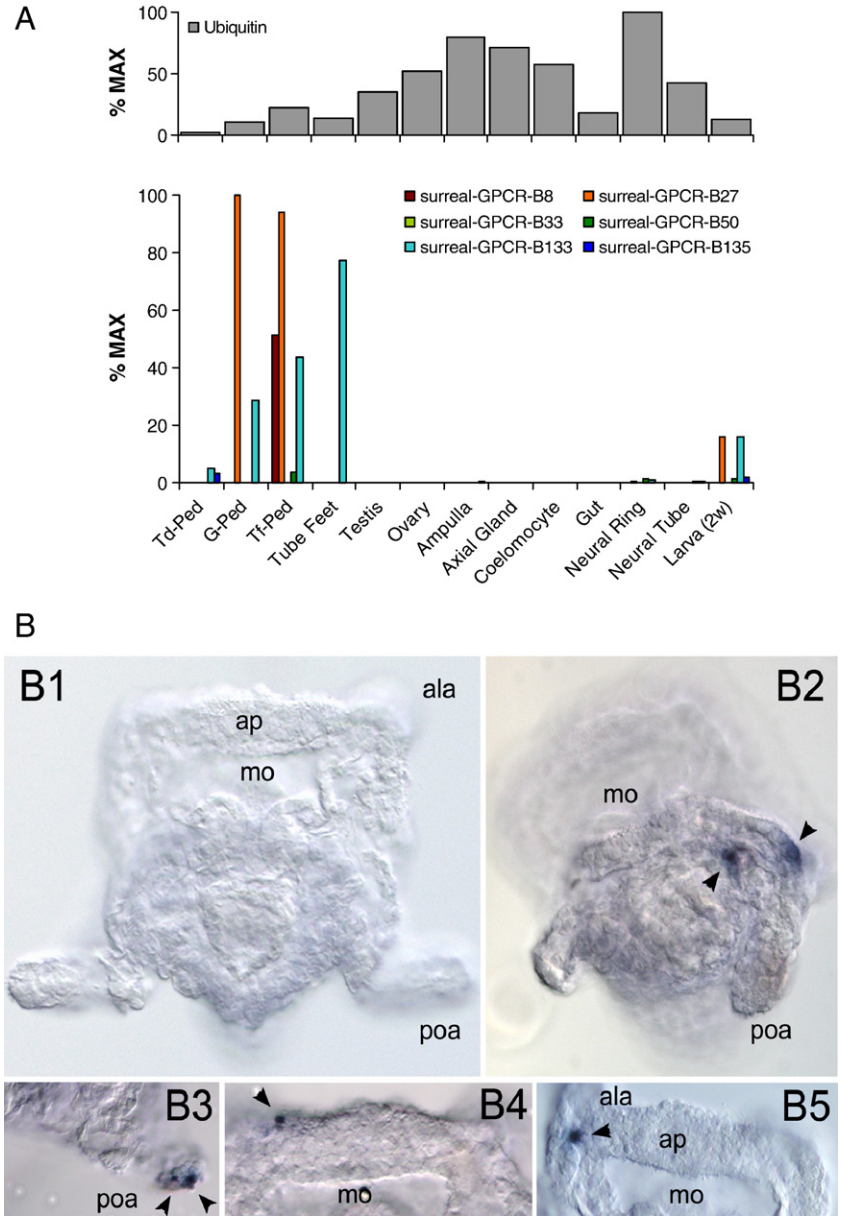


Fig. 5. Expression of representative *surreal-GPCRs* in the sea urchin larva and adult. (A) Expression of *surreal-GPCRs* in two weeks old larvae and adult tissues as detected by QPCR. Measurements and abbreviations are as in Fig. 2A. (B) Expression of *surreal-GPCR-B27* as detected by whole-mount in situ hybridization in two-week-old larvae. (B1) Larva hybridized with sense probe viewed from the anal surface. (B2–5) Larvae stained with antisense probe. In panel B2 a larva viewed as in panel B1 is observed in a different focal plane to show positive cells (arrowheads) along the postoral arm. In panels B3–5 details of different larvae are displayed to show staining in scattered cells (arrowheads). Ap, apical plate; other abbreviations are as in Fig. 2B.

families with chemosensory function in invertebrates and vertebrates prompted us to test if and where *surreal-GPCRs* were expressed. For this, we focused on representatives of the group B *surreal-GPCRs*. We used three different approaches: First, we systematically assessed transcriptional evidence from the available EST resources that cover various larval and adult cDNA libraries (see Materials and methods). Moreover, we selected oligonucleotide primers for a representative set of six group B *surreal-GPCRs* to monitor their expression using quantitative PCR. Finally, we performed in situ hybridizations with candidates that we expected to be expressed based on the QPCR results. All three approaches independently revealed that different members of the group B *surreal-GPCR* family are present at larval and adult stages (Figs. 5A, B and Table S4). In contrast, only limited expression was found for *surreal-GPCR* genes in the transcriptome map of Samanta et al. (2006). This probably reflects the fact that these genes are expressed only at late larval stages, whereas the tiling expression data were generated based on poly-A RNA purified from four embryonic stages (from egg to prism, 65 hpf) (Samanta et al., 2006).

QPCR analysis revealed that within adult urchins the clearly strongest level of the expression is present in pedicellariae and tube feet, which – similar to the *opsins* – express different sets of *surreal-GPCRs* at different levels (Fig. 5A). Possibly, this reflects a considerable sensory divergence between tube feet and different types of pedicellariae. Low additional expression (at least 10 times lower than in tube feet and globiferous and trifoliate pedicellariae) can only be detected in three additional adults tissues; the neural tube and ring, as well as in the ampulla (Fig. 5A). No expression is detected in testis, ovary, coelomocyte, axial gland and gut tissue, indicating that at least the members of our current group B *surreal-GPCR* sampling have a clear tendency to be almost exclusively present in sensory appendages, as well as in the nervous system of the sea urchin. The high expression of group B *surreal-GPCR* members in globiferous pedicellariae (Fig. 5A) is particularly intriguing, given the fact that these pedicellariae have been described to bear chemoreceptor cells on their sensory hillock (Campbell, 1976).

In order to gain further insight where these *surreal-GPCRs* might be expressed, we investigated their expression in larvae. Based on QPCR data, four out of six tested genes are clearly expressed at this stage (Fig. 5A). We further analyzed the expression of *surreal-GPCR-B27* by WMISH in two weeks old larvae (Fig. 5B). We find that *surreal-GPCR-B27 is* present in isolated single cells in the ectoderm at different positions (Figs. 5B2–5). Interestingly, one group of stained cells locates at the tip of the postoral arm (Fig. 5B3), in a position comparable to the *Sp-opsin4*-positive cells (Figs. 2B3–5), indicating that a sensory cell cluster might be located at the tip of these arms. Additional stained cells were observable between the anterior lateral arm and the apical plate (Figs. 2B2, 4, 5). Embryos stained with sense control RNA were never exhibiting any such single cell staining (Fig. 5B1).

Discussion

Sea urchins possess a surprisingly large molecular sensory repertoire

The rhodopsin-type GPCR superfamily appears to be a highly versatile substrate for evolution, as it has been adapted to functions as diverse as light reception, cell–cell signaling and chemoreception (Bockaert and Pin, 1999). Due to this versatility, the abundance of rhodopsin-type GPCRs in a given genome might be a good indicator of the overall versatility of cell types and functions within a species. From this perspective, it is remarkable to note that the number of rhodopsin-type GPCRs in *Strongylocentrotus* is similar to that found in mammals and exceeds that of many other invertebrates. The absolute number of rhodopsin-type GPCRs in the sea urchin genome reported in this study (979) might still change upon completion of the genome: Full separation of sequenced haplotypes will eliminate some of them as allelic variants, whereas further finishing of the sequence is expected to reveal more GPCR loci. Despite these changes, it is clear that rhodopsin-type GPCRs are the largest GPCR superfamily encoded by the sequenced sea urchin genome, constituting 3.4% of all the currently predicted genes. The latter fraction is well comparable to the abundance of rhodopsin-type GPCRs reported for the mouse and human genomes (3.4% and 2.8%, respectively) and exceeds that observed in the pufferfish (1.6%), nematode (2%) or fruit fly genomes (0.6%) (Waterston et al., 2002). This unexpectedly high abundance of rhodopsin-type GPCRs in the sea urchin is consistent with a rich molecular repertoire of sensory molecules in this species.

From our analyses of both *opsins* and *surreal-GPCRs*, it appears that the size of the sea urchin rhodopsin-type GPCR repertoire reflects two different trends. First, there are cases of ancestral rhodopsin-type GPCR ortholog genes that have been retained in the sea urchin genome, whereas they have been lost from other animals, indicating that the sea urchin has partially retained ancestral genomic complexity that has been secondarily reduced in other evolutionary lineages. This is in our analysis exemplified by the ciliary, rhabdomeric and G_o-Opsin family but is a reoccurring theme for other sea urchin gene families such as the *ets* (Rizzo et al., 2006), *fkh* (Tu et al., 2006), *homeobox* (Howard-Ashby et al., 2006) and *zinc finger* (Materna et al., 2006) transcription factor families, the circadian regulators *cry* and *tim* (Rubin et al., 2006), as well as the *vasotocin-like* preprohormone family (Burke et al., 2006). Second, as our discovery of *surreal-GPCRs* indicates, sea urchin rhodopsin-type GPCRs have undergone multiple independent expansions, causing a considerable gain of receptors in the sea urchin lineage. The four main *surreal-GPCR* expansions together account for a total of around 450 rhodopsin-type GPCRs in the sea urchin genome (Table S3). In addition, our systematic analysis indicated multiple small expansions in distinct rhodopsin GPCR subfamilies.

Group B and C surreal-GPCRs encode putative chemoreceptors of the sea urchin

Across the animal kingdom, expansions of rhodopsin-type GPCRs have generated ample substrates for evolutionary modification and structural adaptation of these receptors to diverse functions. Inherent to this phenomenon, it is difficult to make functional predictions for the expanded receptor families based on the roles of their closest relatives in other animals, and olfactory receptors are notoriously different in primary sequence, even among vertebrates (Dryer, 2000). Our argument that the members of the group B/C *surreal-GPCRs* that we investigated more closely are likely to be chemosensory receptors therefore does not build on sequence similarities to other known chemosensory receptor families. Rather, we infer the likely role of these *surreal-GPCRs* as sea urchin chemosensory receptors from two molecular features, (i) the striking similarities in the genes' single exon structure and clustered genomic organization to vertebrate olfactory/chemosensory receptors and (ii) the rapid and divergent expansion of these molecules within the sea urchin lineage. Specific rapid and large expansions of transmembrane domain genes are common to molecules that are utilized by the animal to detect a large variety of divergent molecules, as it is the case in the immune system and in chemoreception. In fact, the argument that the significant diversity is characteristic and essential for receptors functioning in 'olfactory'-type chemosensation was part of the reasoning that lead to the successful cloning strategy for the first olfactory receptors from the rhodopsin-type GPCR family (Buck and Axel, 1991). This reasoning has since been shown to be valid for 'olfactory' type chemosensory receptors of other animals, ranging from fruit fly and mosquito to human (e.g., Clyne et al., 1999, reviewed in Ache and Young, 2005).

Interestingly, the amino acid variance of the sea urchin group B *surreal-GPCRs* is distributed relatively evenly among the seven transmembrane domains. This is different from both mouse and *Drosophila* olfactory GPCRs. In mouse, the transmembrane domains 3–5 contain most of the variable residues, with TM IV and TMV being most divergent (Buck and Axel, 1991; Pilpel and Lancet, 1999), whereas in *Drosophila* the most divergent regions are the first extracellular loop, the third transmembrane helix, as well as the first and second transmembrane domains (Clyne et al., 1999). The dissimilarity between the distribution of divergent residues in vertebrates, sea urchin and insects suggests that the chemosensory receptors in these groups evolved independently. In contrast, the unequal distribution of variable amino acids in the group C *surreal-GPCRs*, and in particular their strongest variability in the fourth and fifth TM helix, is reminiscent of the classical vertebrate olfactory receptors (Buck and Axel, 1991; Pilpel and Lancet, 1999). It remains to be determined if this reflects convergent functional evolution or common ancestry between these vertebrate and sea urchin receptors.

With respect to the potential sequence divergence of the *S. purpuratus surreal-GPCRs*, it is noteworthy that preliminary sequence analysis of expressed group B *surreal-GPCR* sequences isolated from single individuals reveals several cases where more than two variants exist for a single gene present in the genome, indicating that the genome predictions rather reflect a lower bound estimate of the real *surreal-GPCR* diversity (data not shown).

Another strong line of evidence is provided by the fact that the few representatives of the group B *surreal-GPCRs* that we have tested so far show distinct expression profiles in tube feet and different pedicellariae classes of the adult, that have long been proposed or, in the case of globiferous pedicellariae, even demonstrated to contain chemosensory neurons. We expect that the knowledge gained from classical studies on the responses of tube feet and pedicellariae will lead us to determine the exact cues that each receptor is adapted to. This should, in turn, also be informative for the role of *surreal-GPCRs* in the arms of the sea urchin larva, and to determine if they are possibly involved in sensing or reacting to particular cues.

The S. purpuratus genome provides additional insights in Opsin and photoreceptor cell evolution

Our analysis of *opsin* genes has revealed that three ancestral bilaterian classes, the rhabdomeric, ciliary and G_o class are present in the sea urchin. The presence of G_o-Opsin orthologs is particularly revealing because the sea urchin sequences have helped to confirm G_o-Opsins as an ancestral group, improving bootstrap support for the monophyletic origin of G_o-Opsins prior to bilaterian radiation. Interestingly, G_o-Opsins appear to be absent from the other deuterostome groups, including vertebrates, contrasting with the general trend that vertebrates have retained many ancestral proteins (reviewed in Raible and Arendt, 2004).

In turn, the sea urchin might also have lost ancestral genes involved in photosensation. For instance, the currently predicted *S. purpuratus* proteome does not contain any clear retinochrome (RGR Opsins), a group of Opsins that has been found in vertebrates and squid, and possibly (with low branch point support) in the tunicate *Ciona intestinalis* (Hara-Nishimura et al., 1990; Nakashima et al., 2003; Terakita, 2005). This finding, together with the observed loss of G_o-type Opsins in all sequenced vertebrates, insects and nematodes emphasizes that at least the molecular set-up of the ancient bilaterian light sensory system was highly complex and that gene loss has frequently occurred during the evolution of the these systems in Bilateria. This interpretation is in line with recent findings on circadian rhythm regulatory genes and their presence in different Bilateria (Rubin et al., 2006).

As mentioned, dedicated complex photosensory organs have not yet been described in the sea urchin. Therefore, the comparative expression analysis of *opsins* provides an ideal entry point to detect putative photosensory structures in the sea urchin larva and adult. As indicated by our phylogenetic analysis, *Sp-opsin4* is orthologous to the *r-opsin* expressed in the rhabdomeric photoreceptor cells (PRCs) of the larval and adult eyes of primary ciliated protostome larvae, such as the trochophora larvae of mollusks and polychaetes. Our whole-mount *in situ* hybridization analysis revealed that *Sp-opsin4* is expressed in a few cells close to the ciliated band on the tip of

the postoral arms of two week old sea urchin larvae. It will be interesting to further investigate how these *Sp-opsin4*-expressing cells relate to the larval eyes of protostome larvae in respect to their molecular fingerprint, as well as to their position and functional relationship to the main ciliated locomotory band.

A molecular basis to dissect the sea urchin's sensory behavior

Classical reports on responses of larval and adult sea urchin have already provided rich details on the histological and ultrastructural level, but our understanding of the underlying molecular machineries is still very limited. On a first level, the identification of putative photo- and chemosensory receptors in the *S. purpuratus* genome provides candidates to characterize the sensory input of the sea urchin, for instance by mapping the light ranges and chemical cues each receptor responds to. Moreover, the possibility to address specific sensory cells by virtue of their molecular repertoire will help to dissect sensory-output circuits in the animal. In the adult urchin, such integrative circuits are expected to govern local responses of the pedicellariae, because at least some types of pedicellariae have been shown to retain their responsive behavior after isolation from the animal (Campbell, 1974). In this context, the expression of distinct *opsin* classes in pedicellariae and tube feet is particularly remarkable. So far, it had not been found that pedicellariae are themselves light-sensitive, but our data suggest that this is likely. Because pedicellariae have been proposed to be evolutionary derivations of spines (Haude, 1998), this also prompts the question if spines themselves are photosensory, implying that light-dependent responses of spines such as the shadow response could be directly evoked (Millott, 1975), thus representing an autonomous behavior.

Finally, the co-expression of multiple *opsins* in the same tissue also raises the question how different light stimuli are integrated within a tissue, and how they relate to the responsive output. It is noteworthy that the sea urchin does not only possess three ancestral Opsin types, as we describe in our study, but has also been reported to contain three ancient bilaterian families of Cryptochromes (Rubin et al., 2006). Cryptochromes are molecules derived from photolyases, and usually light sensitive in the UV range (Gehring and Rosbash, 2003). With two apparently non-reduced sets of ancestral light-sensitive molecules, the sea urchin will be a well-suited model to study the cross-talk between these systems.

In summary, it appears as if the sea urchin, despite its apparent lack of central sensory organs, possesses a rich repertoire of sensory molecules and sensory structures, defying Aristotle's view of the urchin sensory system as one of the "least developed" ones. The challenge will now be to understand the role of the single components, and how the various sensory units are integrated to a coordinate the behavior of the animal.

Acknowledgments

This work was supported by grants from the Marine Genomics Europe Network of Excellence (GOCE-04-505403, GARNET project; M.I.A, D.A., F.R.), fellowships of the Marie Curie RTN ZOONET (MRTN-CT-2004-005624; E.A., K.T.-R.) and the Deutsche Forschungsgemeinschaft (Deep Metazoan Phylogeny"; T.K.). We would also like to thank Francesca Rizzo of the Stazione Zoologica of Napoli (SZN) for her help and Marco Borra and Elvira Mauriello (SZN Molecular Biology Service) for their assistance with QPCR and sequencing; David Westley and Regina Herhoff for excellent librarian service, Ivica Letunic for help with the visualization of our phylogenetic analysis, and Guy Bloch for communicating results prior to publication. We furthermore thank the members of the laboratories of Ina Arnone, Detlev Arendt and Peer Bork for critical discussion and support, and two anonymous referees for their helpful and constructive criticism.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ydbio.2006.08.070.

References

Ache, B.W., Young, J.M., 2005. Olfaction: diverse species, conserved principles. Neuron 48, 417–430.

Agassiz, A., Clark, H.L., 1907. Hawaiian and other Pacific Echini. Mus. Comp. Zool. 34.

Aizenberg, J., Tkachenko, A., Weiner, S., Addadi, L., Hendler, G., 2001. Calcitic microlenses as part of the photoreceptor system in brittlestars. Nature 412, 819–822.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Arendt, D., 2003. Evolution of eyes and photoreceptor cell types. Int. J. Dev. Biol. 47, 563–571.

Arendt, D., Tessmar, K., de Campos-Baptista, M.I., Dorresteijn, A., Wittbrodt, J., 2002. Development of pigment-cup eyes in the polychaete *Platynereis dumerilii* and evolutionary conservation of larval eyes in Bilateria. Development 129, 1143–1154.

Arendt, D., Tessmar-Raible, K., Snyman, H., Dorresteijn, A.W., Wittbrodt, J., 2004. Ciliary photoreceptors with a vertebrate-type opsin in an invertebrate brain. Science 306, 869–871.

Barnes, D.K.A., Cook, A.C., 2001. Quantifying behavioural determinants of the coastal European sea-urchin *Paracentrotus lividus*. Mar. Biol. 138, 1205–1212.

Bermak, J.C., Li, M., Bullock, C., Zhou, Q.-Y., 2001. Regulation of transport of the dopamine D1 receptor by a new membrane-associated ER protein. Nat. Cell Biol. 3, 492–498.

Bjarnadóttir, T.K., Fredriksson, R., Schiöth, H.B., 2005. The gene repertoire and the common evolutionary history of glutamate, pheromone (V2R), taste and other related G protein-coupled receptors. Gene 362, 70–84.

Blevins, E., Johnsen, S., 2004. Spatial vision in the echinoid genus *Echinometra*. J. Exp. Biol. 207, 4249–4253.

Bockaert, J., Fagni, L., Dumuis, A., Marin, P., 2004. GPCR interacting proteins (GIP). Pharmacol. Ther. 103, 203–221.

Bockaert, J., Pin, J.P., 1999. Molecular tinkering of G protein-coupled receptors: an evolutionary success. EMBO J. 18, 1723–1729.

Buck, L., Axel, R., 1991. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. Cell 65, 175–187.

Burke, R.D., 1980. Development of pedicellariae in the pluteus larva of *Lytechinus pictus* (Echinodermata: Echinoidea). Can. J. Zool. 58, 1674–1682.

Burke, R.D., Angerer, L.M., Elphick, M., Humphrey, R., Yaguchi, G.W., Liang, S., Kiyama, S., Klein, T., Brandhorst, W.H., Rowe, B.P., Wilson, M., Mu, K., Murray, X., Mellott, G., Hallböök, D., Olinski, F., Thorndyke, R., 2006.

A genomic view of the sea urchin nervous system. Dev. Biol. 300, 434–460.
Cameron, R.A., Materna, S.C., Berney, K., 2006. The *Strongylocentrotus purpuratus* genome: a comparative perspective. Dev. Biol. 300, 485–495.
Campbell, A.C., 1973. Observations on the activity of echinoid pedicellariae: I. Stem responses and their significance. Mar. Behav. Physiol. 2, 33–61.
Campbell, A.C., 1974. Observations on the activity of echinoid pedicellariae: II. Jaw responses of tridentate and ophiocephalous pedicellariae. Mar. Behav. Physiol. 3, 17–34.
Campbell, A.C., 1976. Observations on the activity of echinoid pedicellariae: III. Jaw responses of globiferous pedicellariae and their significance. Mar. Behav. Physiol. 4, 25–39.
Campbell, A.C., 1983. Form and function of pedicellariae. Echinoderm Stud. 1, 139–167.
Clyne, P.J., Warr, C.G., Freeman, M.R., Lessing, D., Kim, J., Carlson, J.R., 1999. A novel family of divergent seven-transmembrane proteins: candidate odorant receptors in *Drosophila*. Neuron 22, 327–338.
Davidson, E.H., Erwin, D.H., 2006. Gene regulatory networks and the evolution of animal body plans. Science 311, 796–800.
de Hoon, M.J., Imoto, S., Nolan, J., Miyano, S., 2004. Open source clustering software. Bioinformatics 20, 1453–1454.
Delsuc, F., Brinkmann, H., Chourrout, D., Philippe, H., 2006. Tunicates and not cephalochordates are the closest living relatives of vertebrates. Nature 439, 965–968.
Döderlein, L., 1898. Ueber "Krystallkörper" bei Seesternen. Denkschr. Med. Nat. Ges. Jena 8, 491–494.
Dryer, L., 2000. Evolution of odorant receptors. Bioessays 22, 803–810.
Dubois, P., Ameye, L., 2001. Regeneration of spines and pedicellariae in echinoderms: a review. Microsc. Res. Tech. 55, 427–437.
Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.
Finn, R.D., Mistry, J., Schuster-Böckler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., Moxon, S., Marshall, M., Khanna, A., Durbin, R., Eddy, S.R., Sonnhammer, E.L., Bateman, A., 2006. Pfam: clans, Web tools and services. Nucleic Acids Res. 34, D247–D251.
Fredriksson, R., Schiöth, H.B., 2005. The repertoire of G-protein-coupled receptors in fully sequenced genomes. Mol. Pharmacol. 67, 1414–1425.
Gehring, W., Rosbash, M., 2003. The coevolution of blue-light photoreception and circadian rhythms. J. Mol. Evol. 57 Suppl. 1, S286–S289.
Geis, H.L., 1936. Recent and fossil pedicellariae. J. Paleontol. 10, 427–448.
Goldschmid, A., 1996. Echinodermata. In: Westheide, W., Rieger, R. (Eds.), Spezielle Zoologie-Teil 1: Einzeller und Wirbellose Tiere. Gustav Fischer Verlag, Stuttgart.
Hamann, O., 1887. Beiträge zur Histologie der Echinodermen. III. Anatomie und Histologie der Echiniden und Spatangiden. Gustav Fischer, Jena.
Hara-Nishimura, I., Matsumoto, T., Mori, H., Nishimura, M., Hara, R., Hara, T., 1990. Cloning and nucleotide sequence of cDNA for retinochrome, retinal photoisomerase from the squid retina. FEBS Lett. 271, 106–110.
Haude, R., 1998. Evolutionary Reconstruction of Primitive (Spinate) Echinoid Pedicellariae. Balkema, Rotterdam, London, UK.
Hendler, G., Byrne, M., 1987. Fine structure of the dorsal arm plate of *Ophiocoma wendti*: evidence for a photoreceptor system (Echinodermata Ophiuroidea). Zoomorphology 107, 261–272.
Hill, C.A., Fox, A.N., Pitts, R.J., Kent, L.B., Tan, P.L., Chrystal, M.A., Cravchik, A., Collins, F.H., Robertson, H.M., Zwiebel, L.J., 2002. G protein-coupled receptors in *Anopheles gambiae*. Science 298, 176–178.
Hofmann, K., Stoffel, W., 1993. TMbase—A database of membrane spanning protein segments. Biol. Chem. Hoppe Seyler 347, 166.
Howard-Ashby, M., Materna, S.C., Brown, C.T., Chen, L., Cameron, R.A., Davidson, E.H., 2006. Identification and characterization of homeobox transcription factor genes in *S. purpuratus*, and their expression in embryonic development. Dev. Biol. 300, 74–89.
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.
Kimura, M., 1983. The Neutral Theory of Molecular Evolution. Cambridge Univ. Press, Cambridge, U.K.
Kojima, D., Terakita, A., Ishikawa, T., Tsukahara, Y., Maeda, A., Shichida, Y., 1997. A novel Go-mediated phototransduction cascade in scallop visual cells. J. Biol. Chem. 272, 22979–22982.
Koyanagi, M., Terakita, A., Kubokawa, K., Shichida, Y., 2002. Amphioxus homologs of Go-coupled rhodopsin and peropsin having 11-*cis*- and all-*trans*-retinals as their chromophores. FEBS Lett. 531, 525–528.
Liberles, S.D., Buck, L.B., 2006. A second class of chemosensory receptors in the olfactory epithelium. Nature 442, 645–650.
Materna, S.C., Howard-Ashby, M., Gray, R.F., Davidson, E.H., 2006. The C2H2 zinc finger genes of *Strongylocentrotus purpuratus* and their expression in embryonic development. Dev. Biol. 300, 108–120.
Mathi, S.K., Chan, Y., Li, X., Wheeler, M.B., 1997. Scanning of the glucagon-like peptide-1 receptor localizes G protein-activating determinants primarily to the N terminus of the third intracellular loop. Mol. Endocrinol. 11, 424–432.
Millott, N., 1954. Sensitivity to light and the reactions to changes in light intensity of the echinoid *Diadema antillarum* Philippi. Philos. Trans. R. Soc. B238, 187–220.
Millott, N., 1975. The photosensitivity of echinoids. Adv. Mar. Biol. 13, 1–52.
Minokawa, T., Rast, J.P., Arenas-Mena, C., Franco, C.B., Davidson, E.H., 2004. Expression patterns of four different regulatory genes that function during sea urchin development. Gene Expr. Patterns 4, 449–456.
Mortensen, T., 1943. A Monograph of the Echinoidea. C.A. Reitzel, Copenhagen.
Mulder, N.J., Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Binns, D., Bradley, P., Bork, P., Bucher, P., Cerutti, L., Copley, R., Courcelle, E., Das, U., Durbin, R., Fleischmann, W., Gough, J., Haft, D., Harte, N., Hulo, N., Kahn, D., Kanapin, A., Krestyaninova, M., Lonsdale, D., Lopez, R., Letunic, I., Madera, M., Maslen, J., McDowall, J., Mitchell, A., Nikolskaya, A.N., Orchard, S., Pagni, M., Ponting, C.P., Quevillon, E., Selengut, J., Sigrist, C., Silventoinen, J., Studholme, V., Vaughan, D.J., Wu, R., 2005. InterPro, progress and status in 2005. Nucleic Acids Res. 33, D201–D205.
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, 180–190.
Nemer, M., Rondinelli, E., Infante, D., Infante, A.A., 1991. Polyubiquitin RNA characteristics and conditional induction in sea urchin embryos. Dev. Biol. 145, 255–265.
O'Connell, M.G., Alender, C.B., Wood, E.M., 1974. A Fine Structure Study of Venom Gland Cells in Globiferous pedicellariae from *Strongylocentrotus purpuratus* (Stimpson). J. Morph. 142, 411–432.
Ohno, S., 1970. Evolution by Gene Duplication. Springer-Verlag, Heidelberg.
Pennington, J.T., Emlet, R.B., 1986. Ontogenetic and diel vertical migration of a planktonic echinoid larva, *Dendraster excentricus* (Eschscholtz): occurrence, causes, and probable consequences. J. Exp. Mar. Biol. Ecol. 104, 69–95.
Peters, B.H., Campbell, A.C., 1987. Morphology of the nervous and muscular systems in the heads of pedicellariae from the sea urchin *Echinus esculentus* L. J. Morphol. 193, 35–51.
Phillips, D.W., 1978. Chemical mediation of invertebrate defensive behaviors and the ability to distinguish between foraging and inactive predators. Marine Biology 49, 237–243.
Pilpel, Y., Lancet, D., 1999. The variable and conserved interfaces of modeled olfactory receptor proteins. Protein Sci. 8, 969–977.
Pisut, D.P., 2004. The Distance Chemosensory Behavior of the Sea Urchin *Lytechinus variegatus*. Georgia Institute of Technology, Georgia.
Raible, F., Arendt, D., 2004. Metazoan evolution: some animals are more equal than others. Curr. Biol. 14, R106–R108.
Ransick, A., Rast, J.P., Minokawa, T., Calestani, C., Davidson, E.H., 2002. New early zygotic regulators expressed in endomesoderm of sea urchin embryos discovered by differential array hybridization. Dev. Biol. 246, 132–147.
Rast, J.P., Amore, G., Calestani, C., Livi, C.B., Ransick, A., Davidson, E.H., 2000. Recovery of developmentally defined gene sets from high-density cDNA macroarrays. Dev. Biol. 228, 270–286.
Rizzo, F., Fernandez-Serra, M., Squarzoni, P., Archimandritis, A., Arnone, M.I.,

2006. Identification and developmental expression of the ets gene family in the sea urchin (*Strongylocentrotus purpuratus*). Dev. Biol. 300, 35–48.
Rozen, S., Skaletsky, H., 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. 132, 365–386.
Rubin, E., Shemesh, Y., Cohen, M., Elgavish, S., Robertson, H.M., Bloch, G., 2006. Molecular and phylogenetic analyses reveal mammalian-like clockwork in the honey bee (*Apis mellifera*) and shed new light on the molecular evolution of the circadian clock. Genome Res. 16, 1352–1365.
Saldanha, A.J., 2004. Java Treeview-extensible visualization of microarray data. Bioinformatics 20, 3246–3248.
Samanta, M.P., Tongprasit, W., Istrail, S., Cameron, A., Tu, Q., Davidson, E.H., Stolc, V., 2006. A high-resolution transcriptome map of the sea urchin embryo. Science, in press.
Schiöth, H.B., Fredriksson, R., 2005. The GRAFS classification system of G-protein coupled receptors in comparative perspective. Gen. Comp. Endocrinol. 142, 94–101.
Schmidt, H.A., Strimmer, K., Vingron, M., von Haeseler, A., 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics 18, 502–504.
Sea Urchin Sequencing Consortium, 2006. The sequence of the sea urchin genome. Science, in press.
Smith, J.E., 1937. On the nervous system of the starfish *Marthasterias glacialis* (L.). Philos. Trans., B 227, 111–173.
Smith, J.E., 1965. Echinodermata. In: Bullock, H., Horridge, A. (Eds.), Structure and Function in the Nervous Systems of Invertebrates, vol. 2. W. H. Freeman, San Francisco, pp. 1519–1558.
Smith, D.S., Brink, D., del Castillo, J., 1985. Nerves in the spine of a sea urchin: a neglected division of the echinoderm nervous system. Proc. Natl. Acad. Sci. U. S. A. 82, 1555–1557.
Snyder, N., Snyder, H., 1970. Alarm response of *Diadema antillarum*. Science 168, 276–278.
Terakita, A., 2005. The opsins. Genome Biol. 6, 213.
Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.
Tu, Q.C., Brown, T., Davidson, E.H., Oliveri, P., 2006. Sea urchin forkhead gene family: phylogeny and embryo expression. Dev. Biol. 300, 49–62.
Vadas, R.L., 1977. Preferential feeding: an optimization strategy in sea urchins. Ecol. Monogr. 47, 337–371.
Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., Antonarakis, S. E., Attwood, J., Baertsch, R., Bailey, J., Barlow, K., Beck, S., Berry, E., Birren, B., Bloom, T., Bork, P., Botcherby, M., Bray, N., Brent, M.R., Brown, D.G., Brown, S.D., Bult, C., Burton, J., Butler, J., Campbell, R.D., Carninci, P., Cawley, S., Chiaromonte, F., Chinwalla, A.T., Church, D.M., Clamp, M., Clee, C., Collins, F.S., Cook, L.L., Copley, R.R., Coulson, A., Couronne, O., Cuff, J., Curwen, V., Cutts, T., Daly, M., David, R., Davies, J., Delehaunty, K.D., Deri, J., Dermitzakis, E.T., Dewey, C., Dickens, N.J., Diekhans, M., Dodge, S., Dubchak, I., Dunn, D.M., Eddy, S.R., Elnitski, L., Emes, R.D., Eswara, P., Eyras, E., Felsenfeld, A., Fewell, G.A., Flicek, P., Foley, K., Frankel, W.N., Fulton, L.A., Fulton, R.S., Furey, T.S., Gage, D., Gibbs, R.A., Glusman, G., Gnerre, S., Goldman, N., Goodstadt, L., Grafham, D., Graves, T.A., Green, E.D., Gregory, S., Guigo, R., Guyer, M., Hardison, R.C., Haussler, D., Hayashizaki, Y., Hillier, L.W., Hinrichs, A., Hlavina, W., Holzer, T., Hsu, F., Hua, A., Hubbard, T., Hunt, A., Jackson, I., Jaffe, D.B., Johnson, L.S., Jones, M., Jones, T.A., Joy, A., Kamal, M., Karlsson, E.K., et al., 2002. Initial sequencing and comparative analysis of the mouse genome. Nature 420, 520–562.
Whittaker, C.A., Bergeron, K.F., Whittle, J., Brandhorst, B.P., Bourke, R.D., Hynes, R.O., 2006. The echinoderm adhesome. Dev. Biol. 300, 252–266.
Yang, J., Lusk, R., Li, W.H., 2003. Organismal complexity, protein complexity, and gene duplicability. Proc. Natl. Acad. Sci. U. S. A. 100, 15661–15665.
Yoshida, M., 1966. Photosensitivity. Interscience, New York.