

Research article

Open Access

Candidate chemoreceptor subfamilies differentially expressed in the chemosensory organs of the mollusc *Aplysia*

Scott F Cummins*¹, Dirk Erpenbeck², Zhihua Zou³, Charles Claudianos⁴, Leonid L Moroz⁵, Gregg T Nagle³ and Bernard M Degnan¹

Address: ¹School of Biological Sciences, The University of Queensland, Brisbane, Queensland 4072, Australia, ²Department of Earth and Environmental Sciences, Ludwig-Maximilians-University, Munich, Germany, ³Department of Neuroscience and Cell Biology, University of Texas Medical Branch, Galveston, TX 77555, USA, ⁴Queensland Brain Institute, The University of Queensland, St. Lucia, Queensland, 4072, Australia and ⁵Whitney Laboratory for Marine Science and Department of Neuroscience, University of Florida, St Augustine, Florida 32080, USA

Email: Scott F Cummins* - s.cummins@uq.edu.au; Dirk Erpenbeck - derpenb@gwdg.de; Zhihua Zou - zhzhou@utmb.edu; Charles Claudianos - c.claudianos@uq.edu.au; Leonid L Moroz - moroz@whitney.ufl.edu; Gregg T Nagle - gtnagle@utmb.edu; Bernard M Degnan - b.degnan@uq.edu.au

* Corresponding author

Published: 4 June 2009

Received: 6 November 2008

BMC Biology 2009, 7:28 doi:10.1186/1741-7007-7-28

Accepted: 4 June 2009

This article is available from: <http://www.biomedcentral.com/1741-7007/7/28>

© 2009 Cummins et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Marine molluscs, as is the case with most aquatic animals, rely heavily on olfactory cues for survival. In the mollusc *Aplysia californica*, mate-attraction is mediated by a blend of water-borne protein pheromones that are detected by sensory structures called rhinophores. The expression of G protein and phospholipase C signaling molecules in this organ is consistent with chemosensory detection being via a G-protein-coupled signaling mechanism.

Results: Here we show that novel multi-transmembrane proteins with similarity to rhodopsin G-protein coupled receptors are expressed in sensory epithelia microdissected from the *Aplysia* rhinophore. Analysis of the *A. californica* genome reveals that these are part of larger multigene families that possess features found in metazoan chemosensory receptor families (that is, these families chiefly consist of single exon genes that are clustered in the genome). Phylogenetic analyses show that the novel *Aplysia* G-protein coupled receptor-like proteins represent three distinct monophyletic subfamilies. Representatives of each subfamily are restricted to or differentially expressed in the rhinophore and oral tentacles, suggesting that they encode functional chemoreceptors and that these olfactory organs sense different chemicals. Those expressed in rhinophores may sense water-borne pheromones. Secondary signaling component proteins $G\alpha_q$, $G\alpha_i$, and $G\alpha_o$ are also expressed in the rhinophore sensory epithelium.

Conclusion: The novel rhodopsin G-protein coupled receptor-like gene subfamilies identified here do not have closely related identifiable orthologs in other metazoans, suggesting that they arose by a lineage-specific expansion as has been observed in chemosensory receptor families in other bilaterians. These candidate chemosensory receptors are expressed and often restricted to rhinophores and oral tentacles, lending support to the notion that water-borne chemical detection in *Aplysia* involves species- or lineage-specific families of chemosensory receptors.

Background

All animals must recognize and respond to chemosensory information in their environment. Although the marine mollusc *Aplysia* has been a valuable model to investigate the molecular basis of behavior [1,2] and reproduction [3,4], our knowledge of how they recognize and respond to environmental signals is limited. In particular, it is unknown how they distinguish and bind water-soluble molecules and transfer exogenous information intracellularly. In contrast, the molecular components and mechanisms of chemical detection in a range of vertebrates and other invertebrates have been well studied.

Vertebrate chemoreception is made possible by six distinct classes of multi-transmembrane receptors: (i) olfactory receptors (ORs) [5], (ii) trace amine-associated receptors [6], vomeronasal receptors (iii) type 1 and (iv) type 2 [7,8] and taste receptors (v) type 1 and (vi) type 2 [9,10]. Besides binding chemical molecules, all share the common traits of seven transmembrane (7-TM) domains, G-protein signaling and precise sensory cell expression. In mammals, non-volatile pheromone perception is thought to act primarily through the vomeronasal organ sensory epithelium [11] and be mediated intracellularly via the interaction of chemical molecules with vomeronasal receptors located on the dendrites of vomeronasal sensory neurons [12]. However, in teleost fishes who do not have a vomeronasal organ, the vomeronasal receptors are found in the main olfactory epithelium [13]. It appears that genes involved in an animal's response to its environment are subject to extensive gene duplication, gene loss and lineage-specific expansion over time, leading to large gene families such as those observed in the OR and vomeronasal receptor repertoire. In fact, OR genes represent the largest mammalian gene family [14].

Chemoreception through 7-TM domain receptors appears to have evolved multiple times independently, as vertebrate chemoreceptors are not closely related to those known in insects and nematodes. Recognition of external chemicals in *Drosophila* is accomplished by families of 130 genes encoding 7-TM domain receptors [15,16], including OR (60) and gustatory receptors (70). Gustatory receptors are greatly reduced in the honeybee [17]. Insect chemoreceptors do not belong to the G-protein coupled receptor (GPCR) family due to a unique inverse membrane topology [18]. Rather, they use an alternative, non-G protein-based signaling pathway where receptors not only detect chemicals but can also act as ion channels [19]. In support of this, heterologous cells expressing silkworm, fruitfly or mosquito heteromeric OR complexes showed G-protein independent extracellular calcium influx and cation-non-selective ion conductance upon stimulation with odorant [19]. Nevertheless, chemical

detection is still mediated by a large and divergent family of 7-TM domain receptors.

A central issue that has not been adequately addressed is how water-borne chemicals are detected at the molecular level by the huge diversity of invertebrates that inhabit marine environments. In marine invertebrates, chemosensory abilities are essential for almost all aspects of their life, from feeding to predator avoidance and reproduction. A recent bioinformatic survey of the sea urchin genome resulted in the identification of a remarkable diversity of chemoreceptors, expressed specifically and differentially in adult sensory structures [20]. Meanwhile, there have been important findings forthcoming from research into the molluscan group. Olfactory studies of squid have shown that both phospholipase C (PLC) and cAMP-mediated pathways may be involved in olfactory sensory neurons activation [21]. In support of this, immunolocalization experiments revealed the presence of G proteins involved in both cAMP ($G\alpha_o$) and PLC ($G\alpha_q$) pathways which are clearly co-expressed in certain cell types.

Aplysia possesses many advantages necessary for chemical communication research, such as an extensive knowledge of its anatomy, a detailed understanding of the molecular and cellular basis of behavior, and now considerable genomic and expressed sequence tag (EST) resources. Moreover, we have found that in *Aplysia*, conspecific and congener attraction is mediated by a remarkable cocktail of water-borne protein pheromones [4,22]. In *Aplysia*, freshly laid egg cordons are considered to be a source of both water-borne and contact pheromones that attract conspecifics and closely related species to the area and induce them to mate and lay eggs. Egg laying results in the release of at least four proteinaceous attraction pheromones, including the 58-residue attractin [4,23,24]. T-maze bioassays have demonstrated that binary blends of attractin with either enticin, temptin or seductin are sufficient to attract potential mates [23].

At the anatomical level, *Aplysia* chemosensory detection is achieved by the rhinophore [25], specialized anterior sensory organs on the dorsal surface of the head. Rhinophores are retractile and primarily used for distance chemoreception and rheoreception (response to water current), whereas the oral tentacles, which are found more ventrally, are possibly involved in contact chemoreception and mechanoreception [26]. The neuroanatomical organization of rhinophores includes a rhinophore groove where most of the sensory cells appear to be concentrated. Its sensory epithelium contains sensory neurons that project axons back to rhinophore ganglia and dendrites that end in either a surface-exposed cilium or a small protuberance [26-28]. Consistent with a potential role in

chemical transduction, gene transcripts encoding G protein, PLC or inositol 1,4,5-trisphosphate receptor were found to be expressed in *Aplysia* rhinophore sensory epithelium [29]. The involvement of nitric oxide as a potential chemosensory processing component has also been implicated in molluscan chemoreception based on cytochemical NADH-diaphorase histochemistry of the *Aplysia* rhinophore [30]. Of significance, was the finding that nitric oxide synthase is present in epithelial sensory-like cells that had multiple apical ciliated processes exposed to the environment. This is consistent with findings demonstrating that inhibition of nitric oxide synthase disrupts slime trail following, suggesting a role for nitric oxide in neural processing of stimuli in snails [31].

The presence of G protein mRNAs in *Aplysia* sensory epithelium suggested that multi-transmembrane GPCR-like proteins could play an important role in chemosensory detection. With the availability of a 2× genome coverage for *Aplysia californica*, we expected that it would provide an excellent and first opportunity to investigate the molecular basis of chemical detection in a mollusc. Here, we performed iterative Basic Local Alignment Search Tool (BLAST) searches to identify genes similar to rhodopsin GPCR genes encoding 7-TM domains from the *A. californica* genome. We identified genes representing three unique monophyletic families that show rhinophore, oral tentacle and ovotestis expression. Based on their expression, these may encode chemosensory proteins, including pheromone and gustatory receptors. Antisera directed against a conserved region of a candidate chemosensory receptor, as well as *Aplysia* $G\alpha_q$, $G\alpha_i$, and $G\alpha_o$, confirmed their expression in sensory tissues, with localization to the outer sensory epithelium.

Results

Identification of genes encoding rhodopsin G-protein coupled receptor-like proteins in the *Aplysia californica* genome

We performed iterative tBLASTn for closely related novel genes and discovered a large number of genes encoding rhodopsin GPCR-like proteins. Using this approach, we successfully identified a total of 90 genes encoding proteins belonging to the GPCR superfamily. Of these, 72 were predicted to contain 7-TM domains. It was not possible to annotate the full-length sequence of all genes, especially in the 5'-regions, and 18 genes encoding six transmembrane domains were considered partial-length. Note that these numbers are the minimal estimates, because the genome sequencing of *A. californica* had not been fully completed. We expect that further rhodopsin GPCR-like genes will be found beyond the ones we have identified. Also, several of the multi-transmembrane gene models appeared to be pseudogenes with various defects, including the insertion of stop codons and frame-shifting

indels leading to premature termination of the coding region. These were not included in the final data set.

Phylogenetic construction and analysis of identified rhodopsin G-protein coupled receptor-like genes

We have performed phylogenetic analyses of the 90 selected rhodopsin GPCR-like genes (ingroup), together with four non-*Aplysia* GPCR genes (outgroup). The 94 sequences (including outgroup taxa) comprised 264 to 444 characters. For the two phylogenetic analyses (with and without outgroup, respectively) we had to restrict the character sets to 96 and 166 alignable positions, respectively, in order to maintain the conservative approach. Both phylogenetic reconstructions, with and without outgroup, display a congruent picture regarding the phylogenetic relationship of the *Aplysia* GPCRs (Figures 1a and 1b). Branch lengths of groups A (subfamily a) and B (subfamily b), especially in the former, are considerably shorter than of group C (subfamily c). GPCR-like gene group features are summarized in Additional file 1. The phylogenetic analyses support a monophyly of the three different subfamily groups, although sufficient phylogenetic signal for subfamily c monophyly is achievable only when non-*Aplysia* taxa are not present.

Besides structural predictions that place these genes within the 7-TM superfamily, which includes rhodopsin GPCRs, these genes have little amino acid identity (<10%) to any known genes, and do not appear in the published *Aplysia* EST neuronal transcriptome [32]. They show only distant similarity with known molluscan multi-transmembrane receptors, including well-characterized *Aplysia* neurotransmitter GPCRs. GenBank tBLASTn searches also reveal most amino acid identity with regions of orphan GPCRs of the sea urchin, *Strongylocentrotus purpuratus* (E value $1e-11$), various ghrelin receptors (for example, *Rattus norvegicus*, E value $2e-07$) and a candidate GPCR (*Caenorhabditis elegans*, E value $4e-04$) for subfamilies a, b and c, respectively. Based on the gene characteristics described below and observed tissue distribution (see Results, Tissue specificity of expression), we subsequently called them candidate *Aplysia californica* chemosensory receptors (AcCRs) subfamilies a to c.

Candidate *Aplysia californica* chemosensory receptors subfamily a

A total of 28 genes encoding rhodopsin GPCR-like proteins were identified within the *Aplysia* genome that grouped into the single monophyletic AcCRa. Of these, 10 appear to encode 7-TM domain proteins that range in size from 355 to 368 amino acids (40.1 to 41.5 kDa). The remaining 18 genes encoded 6-TM domain proteins; however, we were unable to identify an initiator methionine suggesting that these represent partial length genes. Figure 2a is a comparative representation of the 7-TM proteins, showing conservation of predicted amino acid sequences

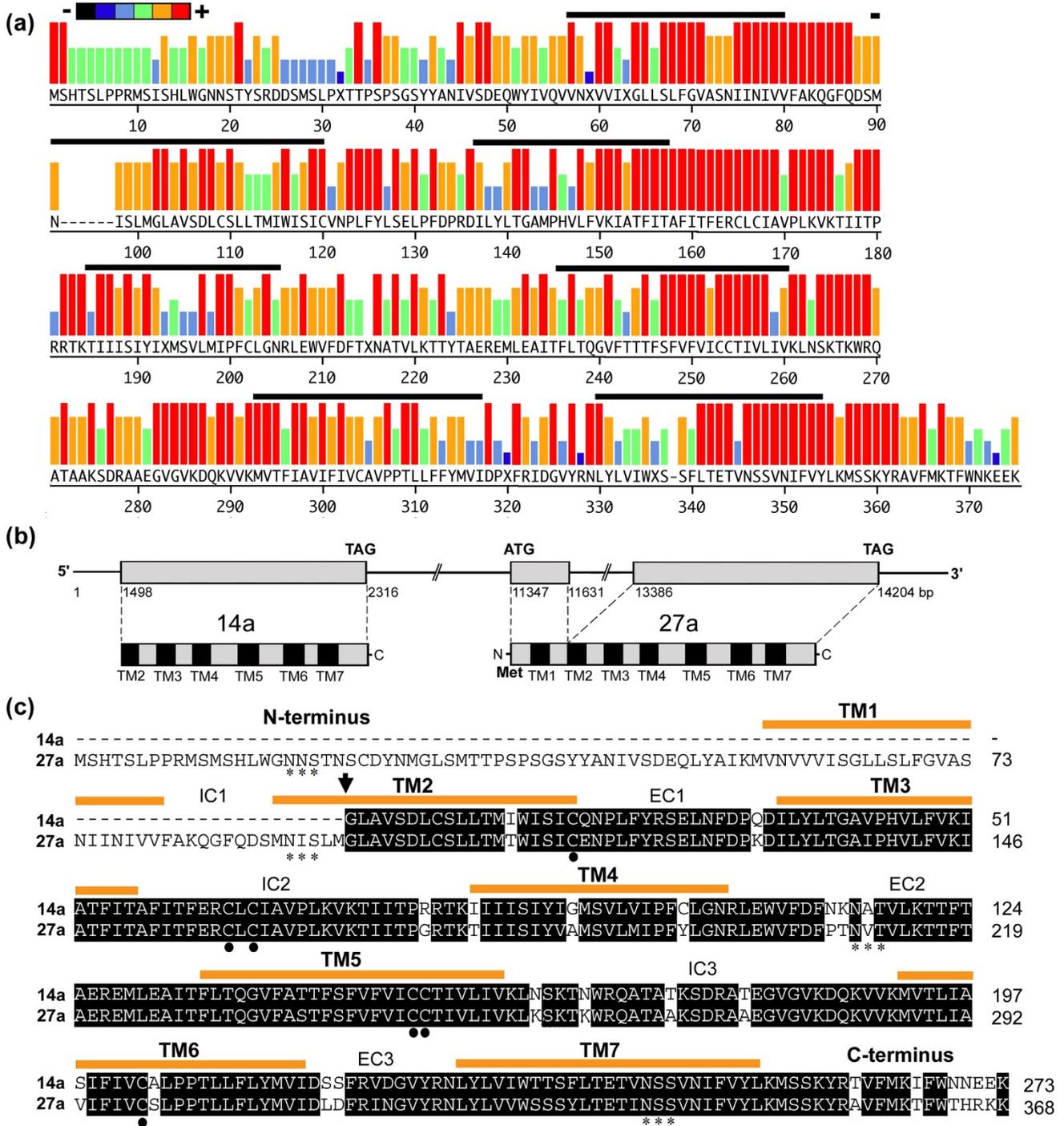


Figure 2 (see legend on next page)

Figure 2 (see previous page)

Analysis of candidate *Aplysia californica* chemosensory receptors subfamily a. (a) Conservation of predicted amino acid sequences for 10 full-length genes is displayed as a consensus strength as color-coded histogram. In this representation, the relative frequency with which an amino acid appears at a given position is reflected by the color, as depicted by the scale bar. The seven transmembrane domains (based on the HMMTOP version 2.0 program) are indicated by a solid bar above the sequence. (b) Schematic representation of the genome organization of clustered genes, partial *AcCR14a* and full-length *AcCR27a* (genomic contig AASC01105652), including predicted start (ATG) and stop codons (TAG) and intron/exon structure of *AcCR27a* leading to the mature multi-transmembrane protein. Met, methionine. (c) Comparative amino acid alignment of *AcCR14a* and *AcCR27a*. Identical amino acids are highlighted in black. Putative intracellular (IC), extracellular (EC), N-terminus and C-terminus domains are shown. Arrow indicates the intron/exon boundary; asterisks show potential N-linked glycosylation sites; black circles represent highly conserved cysteines.

for the gene repertoire. Overall, amino acid sequence identity among the subfamily ranges from 70% to 95%. These genes are most distinct from other subfamily sequences we identified due to the presence of a conserved intron between coding regions ISLM₉₅/GLAV (based on *AcCR27a*, Figures 2b and 2c). This was later verified by reverse transcription-polymerase chain reaction (RT-PCR) cloning and sequencing (See Results: laser capture microdissection (LCM)/RT-PCR gene identification). Also, conservation of GLA/SV₉₆₋₉₉, Y₁₂₀, ITAFITF₁₅₀₋₁₅₆, K₁₇₈, G₁₉₈, DRA/V₂₇₁₋₂₇₃, MVT₂₈₇₋₂₈₉, ET₃₃₆₋₃₃₇, and NSSVNI₃₃₉₋₃₄₄ are distinct this subfamily. Most variability is found within the N-terminal regions (also see alignments in the Additional file 2). Semi- and highly conserved cysteine residues are located at C₁₁₄, C₁₅₉, C₁₆₁, C₂₄₇, C₂₄₈ and C₂₉₈, while glycosylation sites can be found at N₁₈NS, N₉₁IS, N₂₁₀A/VT and N₃₃₉SS. Most share a signature motif with a FITAFITFERCLCIA amino acid sequence in the third transmembrane domain and second intracellular loop. We found that one genomic contig AASC01105652 contained two genes, *AcCR14a* and *AcCR27a* (Figure 2b). We predict that these are part of larger clusters that may become apparent upon completion of the full genomic assembly. Within the *Aplysia* genome (1.8 Gb), *AcCR14a* and *AcCR27a* are separated by 9031 bp, and are in the same transcriptional orientation. A comparative amino acid alignment of the partial *AcCR14a* and full-length *AcCR27a* proteins are shown in Figure 2c.

Candidate *Aplysia californica* chemosensory receptors subfamily b (*AcCRb*)

A total of 38 full-length intronless genes encoding predicted multi-transmembrane rhodopsin GPCR-like proteins were identified belonging to *AcCRb*. Sizes ranged from 319 to 364 amino acids in length (36.2 kDa to 40.8 kDa). Figure 3a is a comparative representation of these proteins, showing conservation of predicted amino acid sequence for the gene repertoire. Overall, amino acid sequence identity among the subfamily members ranges from 43% to 92%. Most variability is found at the N-terminal region and most conservation is within the pre-

dicted transmembrane 3 (also see alignments in Additional file 2). Most share a signature motif with a WITAFVTFERCLCIA amino acid sequence in the putative second intracellular loop. We found that at least some of the genes are clustered in the genome, including genes *AcCR11b* and *AcCR12b* (Figure 3b), as well as genes *AcCR13b* and *AcCR14b* (Figure 3c). Within the *Aplysia* genome, genes *AcCR11b* and *AcCR12b* are separated by 9173 bp while genes *AcCR13b* and *AcCR14b* are separated by 13297 bp, which also includes a putative non-long terminal repeat retrotransposon element in the reverse orientation (10532 to 9078 bp) (Figure S1 in Additional file 3). Amino acid identity between *AcCR11b* and *AcCR12b*, as well as *AcCR13b* and *AcCR14b* is high, 80.5% and 73.9% respectively (Figure S2 in Additional file 3). Conserved cysteines (based on *AcCR11b*) can be found at C₈₈, C₁₃₃, C₁₃₅, C₂₂₁ and C₂₆₉ and N-linked glycosylation sites include N₅ES, N₆₅IT, N₁₈₄KT, N₃₁₀ST, and N₃₂₀MS.

Candidate *Aplysia californica* chemosensory receptors subfamily c (*AcCRc*)

A total of 24 full-length genes containing uninterrupted coding regions were identified belonging to *AcCRc*. Sizes ranged from 324 to 433 amino acids in length (37.7 kDa to 49 kDa). Figure 4 is a comparative representation of these proteins, showing conservation of predicted amino acid sequence for the gene repertoire. Overall, amino acid sequence gene identity ranges from 19% to 91%. Most variability is found at the N-terminal region and within the proposed third intracellular domain, which carried length polymorphisms (also see alignments in Additional file 2). Conserved cysteines (based on *AcCR2c*) are located at C₁₃₆ and C₁₃₈, while semi- and highly conserved N-linked glycosylation sites are located at N₅ET, N₁₆IS, N₅₄IT, N₁₈₇TT, N₂₈₀IS, N₃₃₁TS.

Tissue specificity of expression

We next studied the expression of subfamily genes in adult tissues using degenerate primers designed to conserved codons specific to each of the *AcCR* subfamilies. This approach was designed to detect if any members within the three subfamilies were expressed in the target

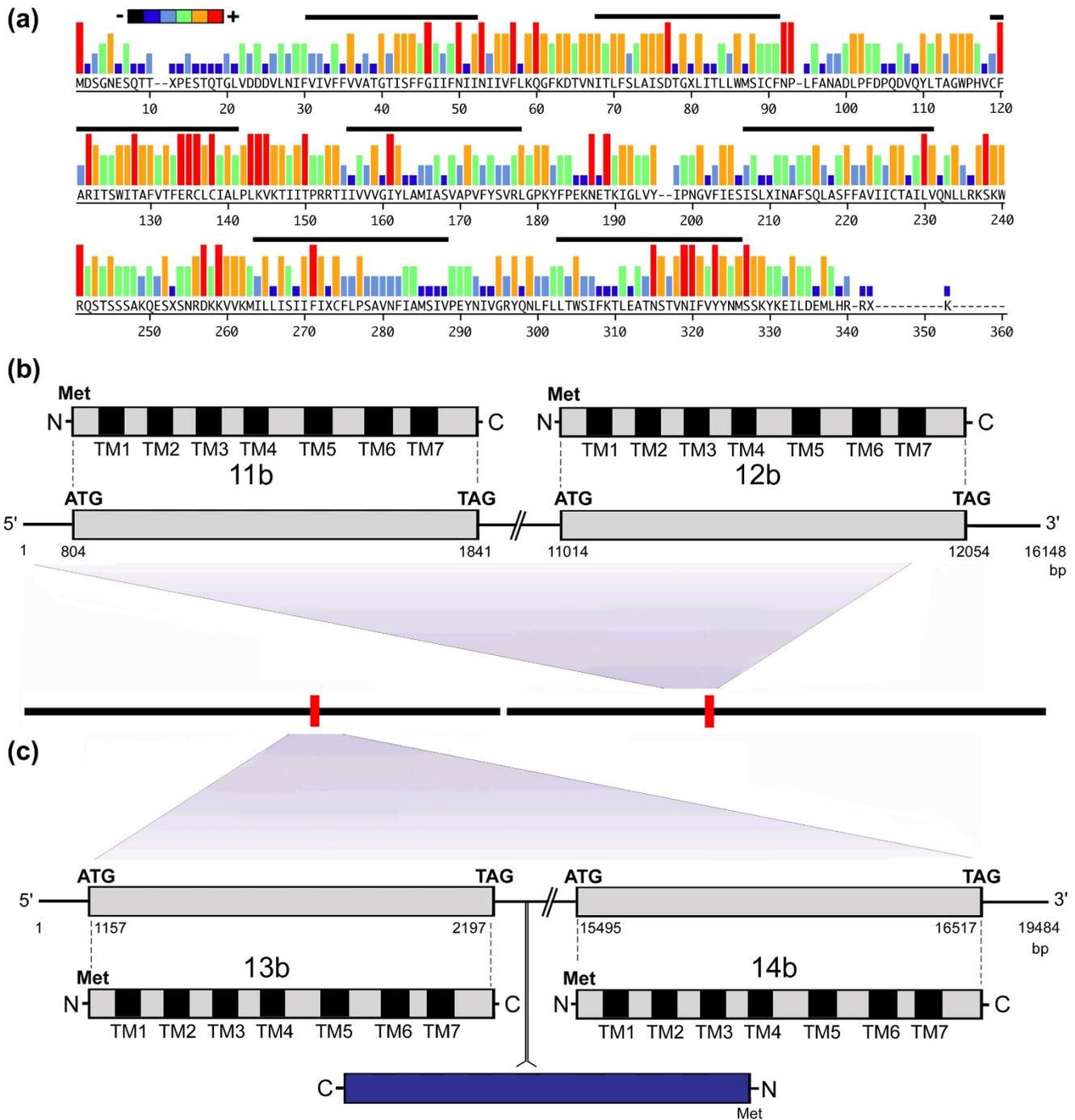


Figure 3
Analysis of candidate *Aplysia californica* chemosensory receptors subfamily b. (a) Conservation of predicted amino acid sequences for 38 full-length genes is displayed as a consensus strength as color-coded histogram. In this representation, the relative frequency with which an amino acid appears at a given position is reflected by the color, as depicted by the scale bar. The seven transmembrane domains (based on the HMMTOP version 2.0 program) are indicated by a solid bar above the sequence. (b) Schematic representation of the genome organization of clustered genes *AcCRI1b* and *AcCRI2b* (genomic contig AASC01159697), including predicted start (ATG) and stop codons (TAG). (c) Similar to (b) but clustered genes *AcCRI3b* and *AcCRI4b* (genomic contig AASC01064969). A predicted retrotransposon element is identified in the reverse orientation between the two gene sequences (see Figure S1 in Additional file 3). Met, methionine.

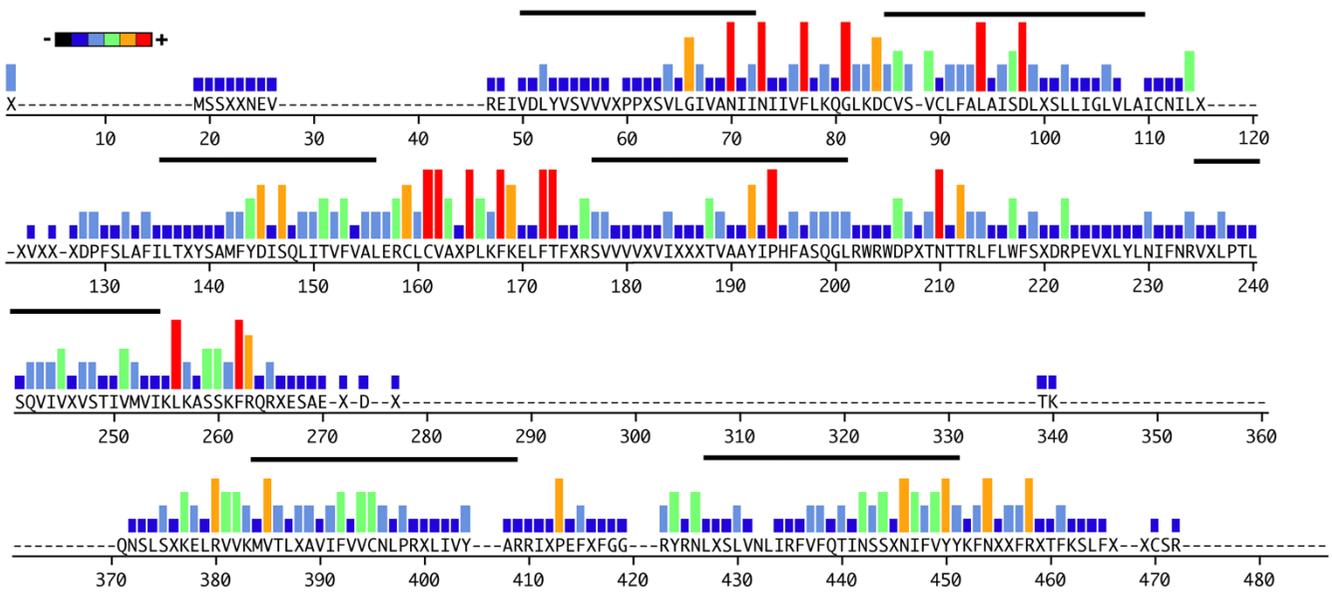


Figure 4
Analysis of candidate *Aplysia californica* chemosensory receptors subfamily c. Conservation of predicted amino acid sequences for 24 full-length genes is displayed as a consensus strength as color-coded histogram. In this representation, the relative frequency with which an amino acid appears at a given position is reflected by the color, as depicted by the scale bar. The seven transmembrane domains (based on the HMMTOP version 2.0 program) are indicated by a black bar above the sequence.

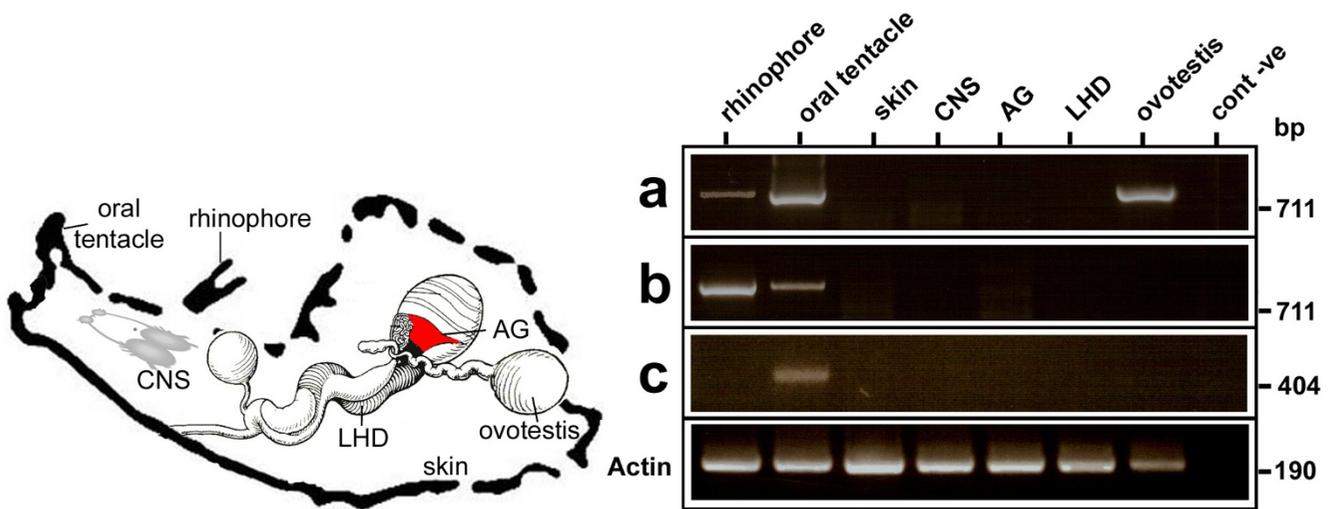


Figure 5
Tissue-specific expression of candidate *Aplysia* chemoreceptors. Schematic representation of *Aplysia californica* showing the location of selected sensory tissues (rhinophore, oral tentacle, skin), central nervous system (CNS) and reproductive organs (albumen gland (AG), ovotestis, large hermaphroditic duct (LHD)) used for RNA isolation and RT-PCR. Amplification products were 756 bp, 832 bp, 512 bp for AcCRa, AcCRb and AcCRc, respectively. The PCR control, using water instead of cDNA, was negative (cont -ve). Actin cDNA was amplified from each tissue preparation (224 bp).

tissues. Transcripts from AcCRa and AcCRb were identified in rhinophore, as well as the oral tentacle; AcCRa transcript was also present in the ovotestis (Figure 5). AcCRc transcripts were detected in the oral tentacle, indicating that each of the gene subfamilies are differentially expressed in the sensory tissues. No transcripts were detected in the skin, central nervous system, albumen gland or large hermaphroditic duct using the method described. We used the *Aplysia* housekeeping gene actin to demonstrate the integrity of each RT sample (224 bp). An alignment of deduced AcCRa amplicon sequences obtained from rhinophore, oral tentacle and ovotestis revealed that different members are present, which correspond most closely to genes *AcCR5a* (97%), *AcCR19a* (100%), and *AcCR7a* (90%), respectively (Figure S3 in Additional file 3). A comparative analysis of proteins encoded in AcCRb amplicons showed that the rhinophore and oral tentacle express a common candidate chemoreceptor gene, corresponding to the gene *AcCR17b*. Several point mutations, however, were present within the rhinophore transcript, which led to a premature stop codon (Figure S3 in Additional file 3). The AcCRc amplicon corresponded to the clone *AcCR2c*.

Scanning electron microscopy and molecular identification of candidate chemoreceptors

Based on results from tissue-specific expression experiments, we focused on the chemosensory organs. In reproductively mature *Aplysia* adults, the rhinophore (about 1 cm in length) is round and tapered from the base to the tip (Figure 6). Scanning electron microscopy (SEM) supports histological examination [27,28] and reveals that rhinophore grooves comprise folds of sensory epithelia bearing numerous cilia extending from a common pore (Figures 6a to 6c). The tip and outside surface of rhinophores are largely devoid of obvious cilia. This rhinophore groove epithelium was previously isolated by LCM and used to construct a cDNA library [12]. In a mature adult, the oral tentacle extends laterally and anteriorly from the ventral surface of the head, with an epithelium containing numerous bunched cilia (Figures 6d to 6f). Although ciliated regions were most common, the oral tentacle did contain regions of no obvious cilia. We next performed PCR with the aim to identify full-length candidate chemoreceptors from the rhinophore epithelium LCM library or prepared oral tentacle cDNA. Three clones were selected for further analysis. Subsequent protein sequence analyses using the Protein Families database of alignments and Interpro database [33] revealed that the deduced amino acid sequences have characteristics common to rhodopsin GPCRs (Figure S4 in Additional file 3).

AcCRa

PCR amplification of a rhinophore epithelium LCM library using degenerate primers that were selective for

members of AcCRa sequences (primer combinations A1 to A5) generated amplification products of 757 bp. Several amplicons were successfully cloned and sequenced, revealing multiple partial-length AcCRa genes. Subsequently, the full length of one gene sequence was identified by 5'- and 3'-RACE, containing 1269 bp and encoding a protein of 354 amino acids (Figure 7a – corresponding most closely, 91%, to partial gene *AcCR8a*). The sequence data has been submitted to the GenBank database under accession number [EU935862](#). It possesses three protein kinase C (PKC) phosphorylation sites (T₂₀₃YK, S₂₅₆DR, S₃₃₇SK,) and four N-linked glycosylation sites (N₁₈ET, N₇₆IS, N₁₉₆AT and N₃₂₅SS). The intron/exon boundary exists between coding regions ISLM₈₀/GLAV. Kyte-Doolittle hydrophathy profiles indicate the existence of seven hydrophobic transmembrane segments that were composed of between 20 and 25 residues, connected by extracellular and cytoplasmic loops.

AcCRb

PCR amplification of a rhinophore epithelium LCM library using degenerate primers that were selective for AcCRb sequences (primer combinations B1 to B8) generated amplification products of 744 bp. Several amplicons were successfully cloned and sequenced, revealing multiple partial-length AcCRb genes. Subsequently, the full length of one gene sequence was identified by 5'- and 3'-RACE, containing 1483 bp and encoding a protein of 354 amino acids (Figure 7b – corresponding most closely, 99.4%, to gene *AcCR29b*). The sequence data has been submitted to the GenBank database under accession number [EU808014](#). It possesses potentially six PKC phosphorylation sites (S₁₁SK, T₁₅QK, T₁₄₇PR, T₂₄₉NR, S₃₂₂SK, S₃₄₃ER) and four N-linked glycosylation sites (N₅NS, N₆₅IT, N₁₈₄VT and N₃₁₀ST) within the predicted N-terminal region and intracellular loop domains. Kyte-Doolittle hydrophathy profiles of the deduced amino acid sequence showed that it contained seven hydrophobic transmembrane segments that were each composed of 25 residues.

AcCRc

AcCRc genes could not be PCR-amplified from a rhinophore LCM library. However, transcripts could be obtained from oral tentacle cDNA preparations (Figure 7c – corresponding most closely, 98%, to identified gene *AcCR2c*). PCR amplification of oral tentacle cDNAs using degenerate primers selective for AcCRc (gene combination C1) generated an amplification product of 824 bp. The amplicon was successfully cloned and sequenced. Subsequently, the full-length gene sequence was identified by 5'- and 3'-RACE, containing 1752 bp and encoding a protein of 398 amino acids. The sequence data has been submitted to the GenBank database under accession number [EU808013](#). It possesses 10 PKC phosphorylation sites (T₇ER, T₂₈LR, T₁₅₀FK, T₁₈₈TR, S₁₉₅SK, S₂₇₅RR, S₂₉₅NK,

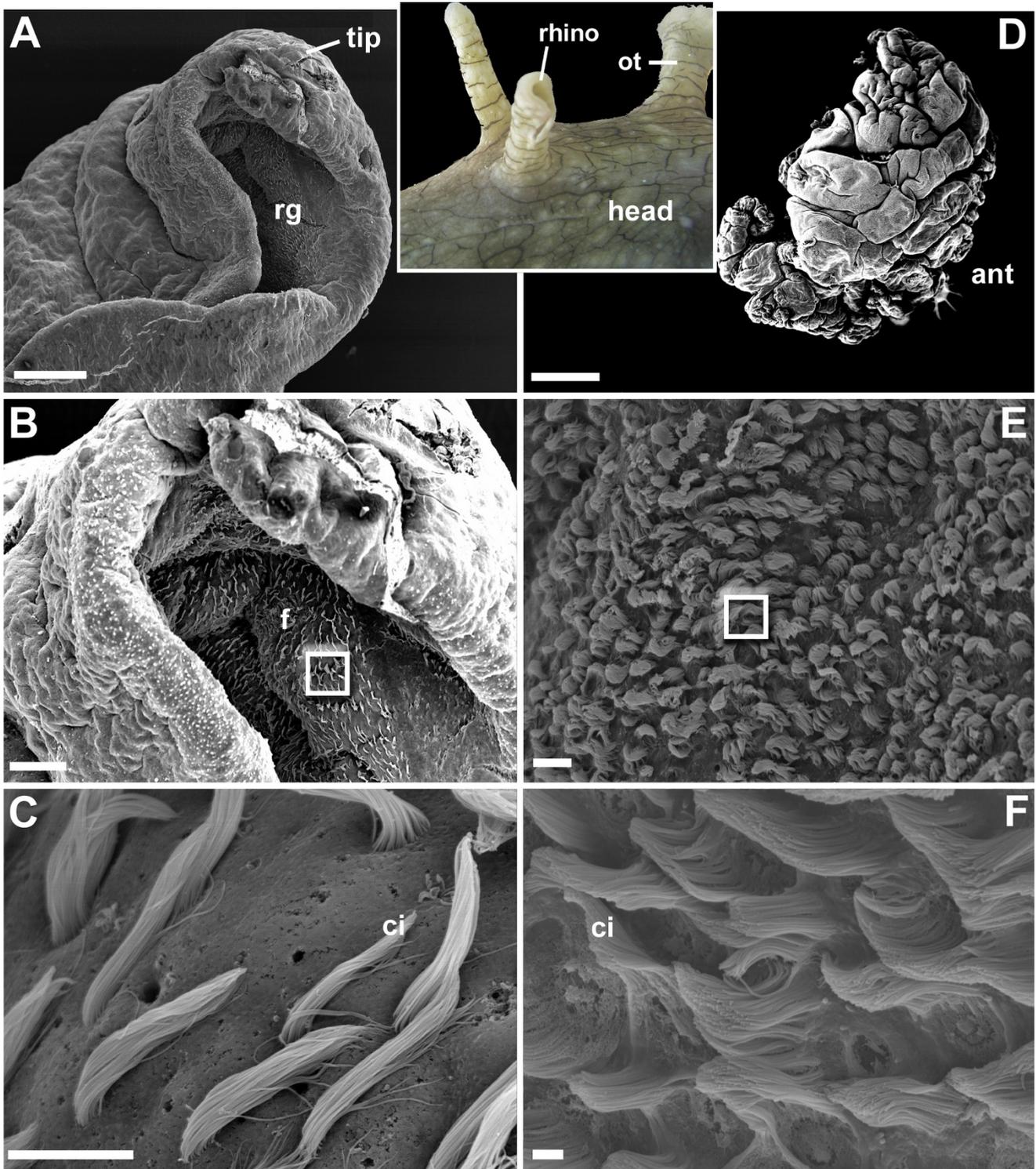


Figure 6 (see legend on next page)

Figure 6 (see previous page)

Scanning electron micrograph analysis of two *Aplysia* chemosensory organs. *Aplysia* species possess rhinophore (rhino) and oral tentacles (ot) to detect chemical stimuli in their marine environment. **(a to c)** Scanning electron microscopy (SEM) micrographs showing the surface of the *Aplysia* rhinophore. Scale bars: 300 μm , 100 μm and 10 μm , respectively. **(a)** A low-power SEM micrograph showing the rhinophore tip. rg, rhinophore groove; tip, rhinophore tip. **(b)** A medium-power SEM micrograph of the rhinophore tip showing the cilia-bearing epithelium within the rhinophore groove. f, folds. **(c)** Higher-power SEM micrograph of groove epithelium (boxed region in b) showing numerous bunched cilia extending from a common pore. Also evident are pores lacking obvious bunched cilia. ci, numerous long cilia. **(d to f)** SEM micrographs showing the surface of the *Aplysia* oral tentacle. Scale bars 1 mm, 10 μm and 1 μm , respectively. **(d)** A low-power SEM micrograph showing the oral tentacle. ant, anterior. **(e)** A medium-power SEM micrograph of the oral tentacle showing a mat of cilia-bearing epithelium. **(f)** Higher-power SEM micrograph of epithelium (boxed region in e) showing numerous bunched cilia extending from a common pore.

S₃₀₈AK, T₃₃₂SR, S₃₈₅YR, a cAMP phosphorylation site (K₂₃₄KSS) and six N-linked glycosylation sites (N₅ET, N₁₆IS, N₅₄IT, N₁₈₇TT, N₂₈₀IS, N₃₂₉TS) within the N-terminal region and intracellular loop domains. Kyte-Doolittle hydrophathy profiles of the deduced amino acid sequence showed that it contained seven hydrophobic transmembrane segments that were composed of between 20 and 25 residues.

Immunofluorescent localization of a candidate chemoreceptor

We complemented our RT-PCR gene expression study by analyzing the spatial distribution of AcCR29b protein within rhinophore and oral tentacle. Figures 8a and 8b shows representative sections of immunoreactivity within similar cell types located at the epithelial surface of rhinophore and oral tentacle, respectively. In both, antisera strongly label cell bodies and processes that extend to the surface, containing no apparent cilia. Rhinophore immunolabeling was most prominent in cells located in epithelia at the tip and outer surfaces, while not obvious within epithelium of the rhinophore groove at the magnification tested. Controls in which the primary antibody was preabsorbed against its antigenic peptide showed greatly reduced staining at the same exposure (Figure 8b, inset).

Distribution of $G\alpha_q$, $G\alpha_i$ and $G\alpha_o$ immunoreactivity in rhinophore

Sections were taken from the sites of pheromone detection, the rhinophore. *Aplysia* G proteins encoded by previously isolated transcripts from rhinophore sensory epithelium [29] are shown schematically in Figure 9a. Commercial antibodies used for this study were directed to the C-terminus which shares 100% identity with *Aplysia* $G\alpha$ proteins. Immunofluorescence studies confirmed the immunoblot expression results [29] and demonstrated localization of immunoreactive *Aplysia* $G\alpha_q$ in rhinophore sections. Numerous $G\alpha_q$ immunoreactive fibers were observed proximal to the rhinophore epithelium and within the distal layer (Figure 9b); $G\alpha_q$ immunoreactivity appeared to be present in the outer sensory surface, con-

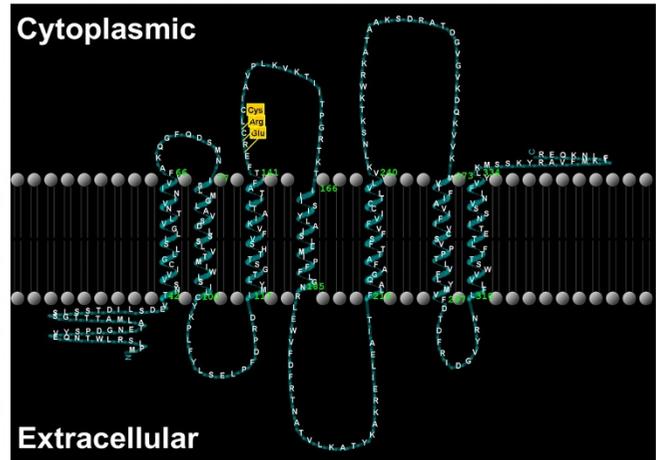
sistent with a potential role in pheromone signal transduction. $G\alpha_i$ -labeled cells were identified throughout rhinophore sections, where immunoreactivity was particularly concentrated in the distal regions of sensory epithelia. Inspection of these sections at high magnification revealed that the majority of cells in the epithelium were labeled, including cells presumed to be supporting cells and sensory neurons (Figure 9c). Immunoreactive fibers could also be found spreading into the cortex of the rhinophore, although this was less prominent. In contrast, immunoreactive *Aplysia* $G\alpha_o$ had a more restricted distribution, in that each section contained several immunoreactive fibers that were observed primarily within presumptive olfactory neuron dendrites (Figure 9d); fewer immunoreactive presumptive sensory neuron cell bodies were observed. No immunoreactivity was observed when primary antibody to $G\alpha_q$, $G\alpha_i$ or $G\alpha_o$ was omitted (data not shown).

Discussion

In this study, we provide an important step towards understanding the molecular and cellular basis of chemosensory recognition in molluscs. Mining of the incomplete *A. californica* genome revealed a novel and diverse set of genes encoding 90 (72 considered full-length intact) rhodopsin GPCR-like proteins that are likely to mediate chemosensory responses in *Aplysia*. To initially identify genes encoding multi-transmembrane GPCR-like proteins that may play a role in chemosensory detection, a number of assumptions were made that have proved important for the successful isolation of chemoreceptors in other metazoans. First, receptor genes would encode 7-TM domains and be clustered in the genome. Second, receptors would be relatively rapidly evolving and thus have limited amino acid sequence identity to members of known conserved GPCR families. Finally, receptors would be encoded by unique families of related genes, as has been observed in a range of other bilaterians [5,15,34]. Phylogenetic analysis of the identified *Aplysia* genes revealed the existence of three monophyletic subfamilies, which we have named candidate *Aplysia californica* chemosensory receptor

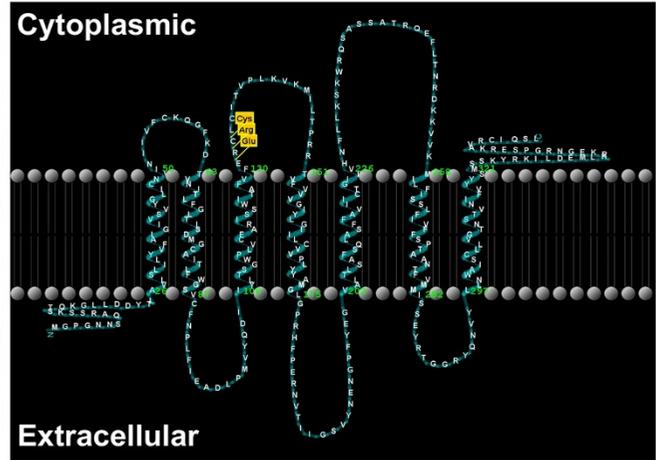
(a)

MPLSRLWLNQEVYSPDGNETALMATTGSSL 31
 SSTDILSDEV TNSVVICGLLSLLGVL TNVIN
 I V V F A K Q G F Q D S M N I S L M G L A V S D L S S L V T M 93
 I W L S I L C Y K P L F Y L S E L P F D P R D I M Y L T G S T
 T H S V F V K I A T L I T A F I T F E R C L C I A V P L K V K 155
 T I I T P G R T K T I I I S I Y L A I S L L M I P F F L G N R
 L E W V F D F R T N A T V L K A T Y K A K R E I L E A I T F L 217
 A Q G A F A T T F S F V F V I C C T I V L V V K L N S K T K W
 R K A T A A K S D R A T D G V G V K D Q K V V K M V T F I A V 279
 I F I V C S V P P T L V F L Y M V F D T D F R I D G V Y R N L
 F L V V W S T T F L T E T I N S S V N I F V Y L K M S S K Y R 341
 A V F M K T F L N K Q E R



(b)

M G P G N N S Q A R S S K S T Q K G L L D D Y T L A V L S L I 31
 L Y V F A I G I V S V Y G L I C N V I N I I V F C K Q G F K D
 T V N I T L F G L T I S D M G C A I T L F W G S V C F N P L F 93
 I E A D L P M V Y Q D I I Y L T S G W P L V C F A R I S S W I
 T A F V T F E R C L C I T V P L K V K M I L T P R R T V F V V 155
 V G I Y L G I I L C V V P L Y Y A M G L G P R H F P E R N V T
 I I G S V Y N E N G P F Y E G V A L T L S A F S Q L A S F F A 217
 V I I C T G I L V H N F L L K S K W R Q S A S S A T R Q E F L
 T N R D K K V V K M I L F I S S L F I V F F S P T A A N T F V 279
 M M I S S E Y R T G G R Y Q N V Y L L N W A I S C L L V G T N
 S T V N I F V Y Y S M S S K Y R K I L D E M L K R K E G N R G 341
 P S E R K A V R C I Q S L



(c)

M S L V N E T E R E L K G S H N I S E H G G L I D D Q T L R I 31
 F L L V F T F V T L G L S L L G S A F N C I N I T V F L T L G
 A K D C V S V C L L S L A V S D F T C L F L G A V C G V C D I 93
 L D A Y G S A D F Y V D P R G L Y Y Q V I F I S S M S Y D I S
 T Y I T A F I S L E R C L C V A L P F R F K E L F T F K R A V 155
 L A M A T I F C L T F C C Y L P H Y V T S G L R V Q W D P R T
 N T T R V L L W S S K D M P A I T A F L D L W N H L I L A V T 217
 S V V I V I V C T H I M V T G L K K S S Q F Q R R G A A R P S
 E P D G P N N F K T S N E V G E G E E N V L R D P D S R R D N 279
 N I S T P Y C P T N V E K E S S N K I K V E K S P Q T L S A K
 N R R R R V V K M V S T L A I V S I L C N T S R L L F V V A L 341
 R A E P D I N F G H R Y H N L Y M V I L V L A Y I F Q V I N A
 P V N I F I Y L K L N P S Y R K T F S Q I F G I G Q T K

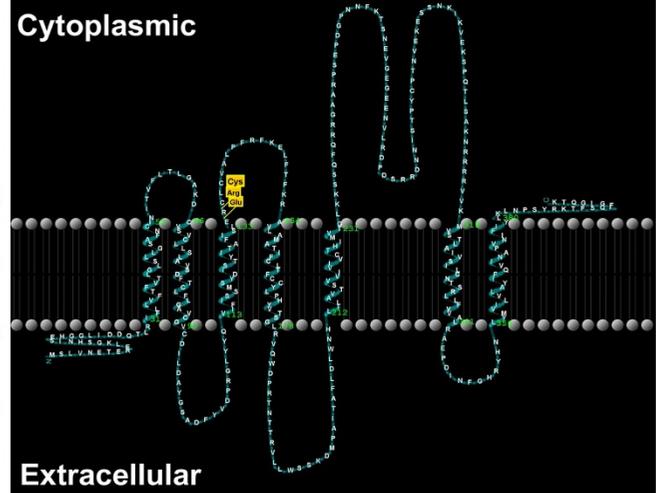


Figure 7

Molecular identification of candidate *Aplysia californica* chemoreceptor genes. Deduced amino acid sequence and schematic model of (a) AcCRa cDNA isolated from rhinophore laser capture microdissection (LCM) RNA (GenBank: [EU935862](#)), (b) AcCRb cDNA isolated from rhinophore LCM RNA (GenBank: [EU808014](#)) and, (c) AcCRc cDNA isolated from oral tentacle cDNA (GenBank: [EU808013](#)). Arrowhead, indicates position of intron/exon boundary and boxed grey areas indicate predicted transmembrane domains (based on HMMTOP version 2.0 program). Unfilled boxes show relative positions of the conserved motif Glu-Arg-Cys (ERC) at the proximal part of the cytoplasmic 2 domain.

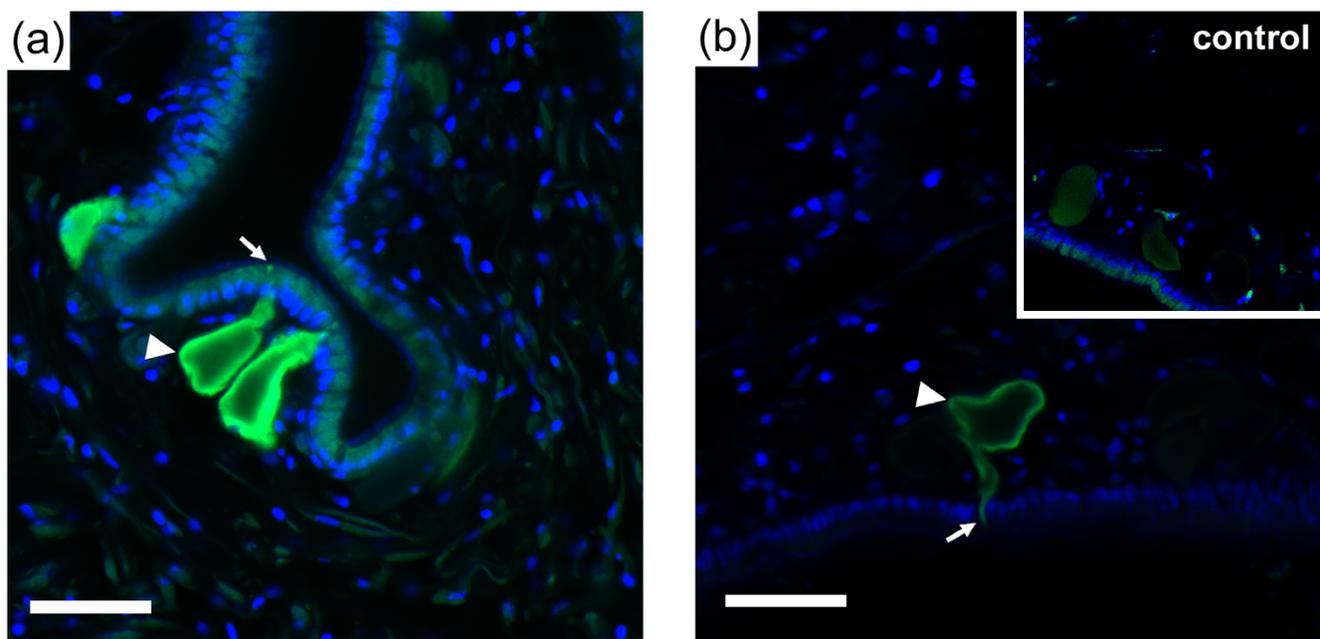


Figure 8
Immunofluorescence localization of AcCR29b. An antibody designed to the N-terminal region of AcCR29b was used to localize corresponding protein within epithelial cells of **(a)** rhinophore and **(b)** oral tentacle, shown in green. Arrowheads show immunopositive cell bodies and arrows point to processes exposed to the surface. Sections were counterstained with DAPI (blue) to show nuclear staining. A control in which the primary antibody was preabsorbed against its antigenic peptide showed greatly reduced staining at the same exposure (b, inset). Scale bars = 50 μ m.

(AcCR) subfamilies a to c; their primary features are summarized in Table 1. The gene expansion observed could provide the diversity of receptors required to enable the animal to recognize diverse water-soluble molecules, as well as complex pheromone blends during the coordination of attraction and reproduction.

Consistent with known chemosensory receptors (for example, insect, rodent, fish), all selected genes that were considered full length encoded 7-TM regions and semi-conserved glycosylation sites, as well as several common cysteine residues and amino acid sequence motifs. Conserved amino acids and post-translational modifications would likely contribute to the correct folding and functioning within the plasma membrane so that they may bind chemical stimuli and couple to appropriate secondary signaling molecules. For example, Katada and colleagues [35] demonstrated that in rodents, N-terminal glycosylation is critical for proper targeting of ORs to the plasma membrane. While proteins encoded by AcCRa and AcCRb genes share notably high sequence identity, comparative analysis within AcCRc shows that they share as little as 19% amino acid identity. As a consequence, there are very few defining sequence motifs which are retained throughout. However, it is not uncommon for chemoreceptors, and in particular gustatory receptors, to

be divergent; similarity between most insect and mammalian gustatory receptor pairs is only 15% and 25% or less at the amino acid sequence level, respectively [10,36].

Of the major GPCR superfamily groups, the identified *Aplysia* genes categorize most closely to the rhodopsin GPCR family (based on Interpro database [33]). As is the case with many other rhodopsin family GPCRs, these genes largely lack introns. Moreover, all encoded proteins have a short N-terminus and a highly conserved arginine (R) residue located at the cytosolic end of the third transmembrane domain. This residue is typically associated with the DRY (Asp-Arg-Tyr) motif, crucial for controlling agonist-dependent receptor activation. Of the three residues constituting DRY, arginine is the most conserved residue, and appears to be essential for forming intramolecular interactions that constrain receptors in either the inactive or activated conformation [37]. Consistent with this, receptors lacking the arginine side chain fail to activate G-protein signaling [38,39]. In the novel *Aplysia* proteins, however, this has been replaced by an ERC motif, a feature also observed in the human prostaglandin F2 α receptor and most other prostanoid receptors [40]. Studies show that substitution of the glutamic acid to a threonine residue leads to full constitutive activation and implicates the region in agonist-dependent G-protein

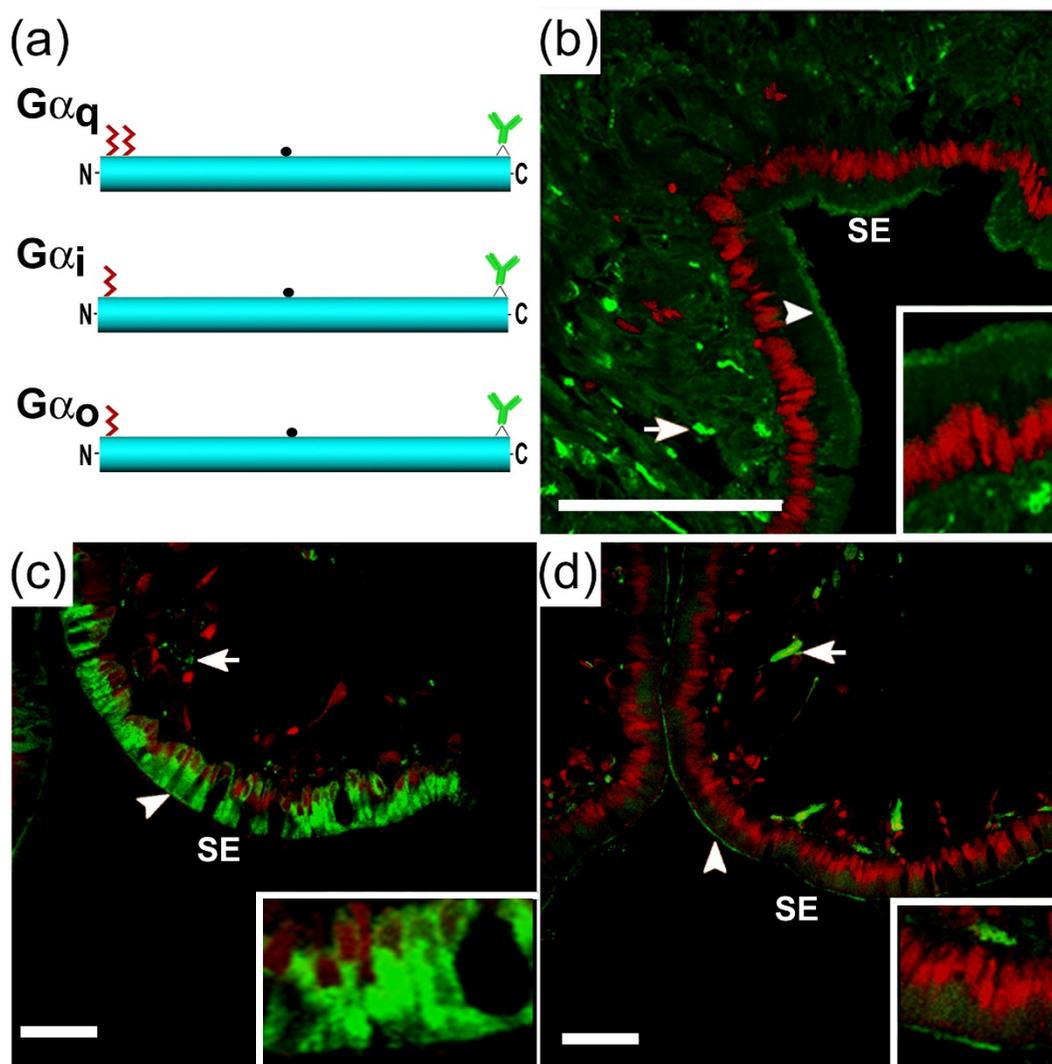


Figure 9
Analysis of *Aplysia* G α_q , G α_i , and G α_o proteins in rhinophore. (a) Schematic diagrams of G α_q (GenBank: [DQ397515](#)), G α_i (GenBank: [DQ656111](#)) and G α_o (GenBank: [DQ656112](#)) showing the location of N-terminal cysteines that may be sites for palmitoylation (wavy lines), a putative cholera toxin ADP-ribosylation site (●), and the binding site of G α -specific antibodies used for immunolocalization analyses. (b to d) Immunofluorescent (green) localization of G α_q (b), G α_i (c) and G α_o (d) in the rhinophore sensory epithelium (SE), including presumptive sensory neuron cell fibers (arrows) and outer epithelia (arrowheads). Sections were counterstained with propidium iodide (red). Scale bars = 100 μ m. Higher magnifications are shown in insets.

Table 1: Summary of candidate *Aplysia* chemoreceptor genes selected by genome mining

Subfamily	Number of intact genes	Size range (aa's)	Predicted TM domains	Intron number	Sites of expression**
a	10*	355–368	7	1	rhino, ot, ovo
a	38	319–364	7	0	rhino, ot
c	24	324–433	7	0	ot

* Not including those considered partial-length (non seven transmembrane domain).

** Based on degenerate reverse transcription-polymerase chain reaction strategy outlined in this study (Methods)
 rhino, rhinophore; ot, oral tentacle; ovo, ovotestis.

coupling control [40]. We predict that this may also be essential to receptor activation in the identified *Aplysia* receptors.

The canonical model of GPCR activation is via an interaction with intracellular heterotrimeric G-protein signaling components. The genes identified in this study show little amino acid identity to GPCRs found in the Metazoa and we have not shown that they directly interact with G-proteins. Despite this, our study indicates that $G\alpha$ proteins are present in the rhinophore sensory epithelium, possibly in close association with transmembrane GPCRs. The presence of sensory tissue G-protein immunoreactivity adds further support to studies of other marine invertebrate olfactory systems implicating G-proteins in sensory transduction [21,41]. Moreover, the existence of multiple G-type proteins in sensory epithelium suggests that multiple signal transduction pathways may be activated following ligand stimulus. In squid, for example, the pattern of immunolabeling implies that a G protein coupled to a PLC pathway ($G\alpha_q$ and $G\alpha_o$) may be present in similar cells as those coupled to a cAMP pathway ($G\alpha_i$). As suggested by Mobley and colleagues [21], overlapping G-protein pathways could facilitate discrimination between odorants detected by the same neuron. This contrasts rodent models where the role of G proteins in olfactory transmembrane signaling at the dendrites has been studied extensively. Researchers have demonstrated spatially restricted patterns of expression of respective G proteins [42-44]. $G\alpha_o$ and $G\alpha_i2$ are highly expressed by separate subsets of neurons that are located in different regions of the vomeronasal neuroepithelium [44].

Of particular relevance to *Aplysia* chemosensory studies is the rhinophore epithelium, where water-borne molecules such as pheromones presumably bind and initiate activation of pheromone-receptive neurons. In the rhinophore groove, receptor cells with a suspected chemosensory role have been described previously in molluscs [26-28] and their presence was further supported by our SEM analysis. The rhinophore groove ciliary aggregations are likely necessary in the separation and circulation of fluid throughout the groove, and may also be directly involved in detection of external chemical stimuli. It is from this precise location that we isolated cDNAs encoding identified novel GPCR-like proteins. Their presence raised the question as to whether their expression was specific to sensory tissues. Subsequently, representatives of each subfamily were found to be restricted to or differentially expressed in the rhinophore, oral tentacles and ovotestis, suggesting that they encode functional receptors and that these olfactory organs sense different chemicals. In the rhinophore, spatial expression of cells immunoreactive to candidate chemoreceptor AcCR29b was most prominent in the tip and outer epithelium, peripheral to the groove. Chemo-

sensory detection could likely benefit from this broad distribution, whereby stimulation may activate sensory fibers that extend to higher brain centers. Although this finding clearly indicates a sensory role, a more extensive study of *Aplysia* sensory organs at higher magnification is required to delineate the precise distribution of this receptor, as well as other receptor subfamily members.

Interestingly, we found gene expression within the ovotestis, and our preliminary analysis of various *Aplysia* neuronal EST databases indicate that a relatively small fraction of these genes may be expressed in the central nervous system. Deep sequencing of neuronal transcripts has resulted in identification of tags for 13 different genes in the central ganglia of *A. californica* (that is, *AcCR1a*, *AcCR5a*, *AcCR15a*, *AcCR32b*, *AcCR5c*, *AcCR9c*, *AcCR11c*, *AcCR15c*, *AcCR16c* and *AcCR20c*, see Additional file 1) as well as several orphan receptors similar to vertebrate ghrelin and histamine receptors (L Moroz, unpublished). Some of those are associated with centrally located sensory neurons that send neuronal processes to the periphery and therefore may be involved in chemoreception. Other neuronal cell types were previously described as motoneurons. Taken together, these findings imply that external chemical detection may not be their sole function, which is consistent with that described of chemoreceptors in mammalian olfactory bulb [45], cardiac muscle [46] and vertebrate germ tissues [47-49]. The functional significance of such expression is currently unknown. We speculate that there may be a role for *Aplysia* chemoreceptors in oocyte recognition, possibly because both are activated by identical or structurally similar hormones and pheromones. In addition to the exocrine albumen gland, the *Aplysia* pheromone seductin is known to be expressed in ovotestis tissue [24].

It will be of interest to perform a comparative gene analysis to determine whether *A. californica* candidate chemoreceptors are also present in other *Aplysia* species. This finding would suggest that these genes are highly similar throughout *Aplysia* species and strongly imply a selective pressure for conservation. We have already established that several *Aplysia* species share a comparable attraction pheromone blend [4,23], and therefore cognate receptor binding sites are likely to be similar. The identification of ligands for chemosensory receptors is often problematic and therefore it is an advantage that we have these attraction pheromones for use in future functional studies to definitively link identified genes to chemosensory detection. As these are the first such novel GPCRs identified in molluscs, it will also be of interest to see if analogous receptors are found in evolutionarily distant molluscs. However, apart from the highly conserved insect receptor Or83b [50], it is generally accepted that chemoreceptors

seem to be very divergent with little sequence conservation within and across orders [51,52].

Our analysis of the *Aplysia* genome noted that AcCRa genes as well as various AcCRb genes are clustered, a common feature of fast-evolving genes such as chemoreceptor genes [15,53-55]. We also found that, although the assembly of the genome used was incomplete, some genes contain mutations that introduce stop codons to encode truncated proteins, one of which appears to be expressed in the *A. californica* rhinophore. Hominoids, in particular, are known to possess a high pseudogene content (50%) among their ORs, whereas only 20% of OR genes are pseudogenes in the mouse [34] and less in the *Drosophila melanogaster* genome [56]. Upon genome completion, a more comprehensive analysis of GPCR gene families in *Aplysia* will be necessary to determine the precise pseudogene number. Indeed, some of the pseudogenes identified may in fact be 'flatliners', that is, genes whose functional versus pseudogene status is unclear [57]. As demonstrated in *C. elegans*, many of these genes have apparently functional alleles in one or more wild isolates and therefore are not pseudogenes. Evidence for this has also been shown for some *Drosophila* ORs [58] and gustatory receptors [59], as well as *Anopheles gambiae* gustatory receptors [51]. Although pseudogenes are generally accepted as nonfunctional and therefore not transcribed, occasionally it has been shown that such pseudogenes can be transcribed [60]; however, there is no evidence of the functional relevance.

Conclusion

Aplysia is an excellent model animal for studying the molecular mechanism of chemical communication in the marine environment. In this study we have isolated a novel group of genes encoding multi-transmembrane rhodopsin GPCR-like proteins that show expression in chemosensory tissues. This expression pattern and observed genomic clustering provide strong evidence that these have arisen via gene duplication and may be used to discriminate the large diversity of water-soluble molecules. The expression of some of these in rhinophore suggests that they are excellent candidates to be involved in pheromone detection. Further knowledge of the receptor gene genome organization, characterization of their developmental and spatial expression profile, secondary signaling and their evolutionary relationship to other molluscan species would be the next significant steps towards defining the logic behind how chemical communication in molluscs, and potentially other marine animals, operates.

Methods

Database mining for multi-transmembrane rhodopsin G-protein coupled receptor-like genes

A. californica genome contig sequences were procured from the NCBI trace database http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi?organ_ism=euk [61]. This was a preliminary genome assembly from 2× coverage of the genome by the Broad Institute at MIT <http://www.genome.gov/Pages/Research/Sequencing/SeqProposals/AplysiaSeq.pdf> [62]. An iterative tBLASTn strategy was adopted to identify multi-transmembrane rhodopsin GPCR-like genes in the *Aplysia* genome. Selection criteria included that receptors would be encoded by a family of related genes; at least some receptor genes would be clustered at the same genetic loci; receptors would have limited amino acid identity to members of known GPCR superfamilies; and a full-length coding region would encode multiple transmembrane domains. A search was initially performed using molluscan GPCR protein sequences already submitted to GenBank, as queries; a non-stringent expectation value cutoff of 1e-4 was employed. During this search we retrieved two sequences on contig AASC01105652 with genes encoding hydrophobic multi-transmembrane domains with no significant amino acid identity to other proteins. Putative *Aplysia* transmembrane receptors were in turn employed in searches to find more genes in an iterative tBLASTn process. A candidate rhodopsin GPCR-like gene having a complete open reading frame (methionine, 7-TM domains, three extracellular domains, three intracellular domains, and a stop codon) was considered intact and probably functional. To be conservative, genes that were >98% identical in amino acid sequence were considered allelic variants. As the *Aplysia* genome has not yet been fully assembled and consists of only contigs, this method does not identify splice-variants or provide a comprehensive analysis of gene clustering. Pseudogenes were identified by premature stop codons and we have arbitrarily chosen to disregard those genes that encode less than six transmembrane domains. Computer analyses of sequences were performed using BLAST and CLUSTALW for nucleotide alignment. Transmembrane helix domain and topology of predicted receptors was performed using HMMTOP version 2.0 program; <http://www.enzim.hu/hmmtop/index.html> [63]. Hydrophilicity plots were generated using the TMHMM program at <http://www.enzim.hu/hmmtop/html/submit.html> [64]. In order to categorize identified genes, we used <http://pfam.janelia.org/search> [65] and the Interpro database [33].

Phylogenetic and gene analysis

Selection of outgroup sequences was performed by GenBank tBLASTn searches. The approach provided homologous counterparts from *C. elegans* for the *Aplysia*

rhodopsin GPCR-like genes representing subfamilies a and b. We did not find any significant counterparts for subfamily c. The sequences were aligned with t-coffee [66] under default settings. Alignments have subsequently been improved by eye. Character positions, which could not be aligned unambiguously, were not considered for the phylogenetic analyses in order to avoid conflicting phylogenetic signal. The inclusion of outgroup sequences of *C. elegans* and *Strongylocentrotus purpuratus* resulted in a large proportion of only ambiguously alignable positions and consequently required deletion of a relatively high number of characters. Therefore we performed two separate analyses: first, we aligned and analyzed the ingroup sequences only in order to reconstruct their evolution under a maximal number of informative characters. Second, we aligned and analyzed the dataset with the non-*Aplysia* sequences included in order to infer the polarity to the phylogeny.

Phylogenetic analyses were conducted on a multi-processor Linux-Cluster under the likelihood criterion using RAxML v. 7.0 [67] for Maximum Likelihood and MrBayes v. 3.1.2 [68] for the Bayesian Inference. MrBayes analyses were performed in two runs of eight MCMCMC chains and under the GTR+G+I Model [69]. Chains ran for 10,000,000 generations or were stopped when the standard deviation of split frequencies between both runs fell below 0.01. RAxML bootstrap analyses on 1,000 replicates have been performed under the PROTMIX algorithm with the WAG amino acid substitution model [70]. In PROTMIX the tree inference is performed under the PROTCAT model followed by the final tree evaluation under the PROTGAMMA model in order to obtain stable likelihood values (see the RAxML manual for further details).

Animal and sample preparation

Adult *A. californica* (100 to 500 g) were obtained from Marine Research and Educational Products (Escondido, CA, USA). Animals were anesthetized in isotonic MgCl₂ (337 mM) equivalent to 50% of their weight, relevant tissues dissected out and either (1) embedded in optimal cutting temperature (OCT) compound for LCM or sectioning, (2) snap frozen in liquid nitrogen for RNA and protein isolation, or (3) prepared for SEM.

Reverse transcription-polymerase chain reaction

Total RNA was extracted from rhinophore, oral tentacle, skin, pooled central nervous system, albumen gland, large hermaphroditic duct, and ovotestis tissues of *A. californica* using a Tripure Isolation Reagent. Any contaminating genomic DNA was removed by treatment with DNase I followed by lithium chloride/ethanol precipitation. First-strand cDNA synthesis was performed in a 20 µl reverse transcription mixture containing oligo d(T)₁₂₋₁₈ and 200 U Superscript III RNase H⁻ reverse transcriptase, following

the manufacturer's instructions. PCR was performed using 1 µl of prepared cDNA using subfamily-specific primers (see Table S1 of Additional file 3: primer combinations A3, B3 and C1). Each reaction was performed in a final concentration of 1× PCR Buffer, 1.5 mM MgCl₂, 200 µM of each dNTP, 0.5 µM of sense and antisense primer, 1.25 units of Red Taq polymerase and ddH₂O. Negative controls contained no template cDNA. PCR using actin-specific primers (sense, 5'-GCTTCACCACCACTGCCGAGAG-3' and antisense, 5'-ACCAGCAGATTCCATACCCAGG-3') were used to ensure the integrity of each tissue cDNA sample. Reactions were heated at 94°C for 2 min and amplified for 36 cycles (94°C, 60 s; 50 to 55°C, 30 s; 72°C, 60 s). Following PCR, 15 µl of reaction mix was fractionated on 2% agarose gels and visualized by ethidium bromide staining. Based on the primer design, the expected amplification sizes were 756 bp (AcCRa), 832 bp (AcCRb), 512 bp (AcCRc) and 224 bp (actin). PCR products were cloned into pGEM-T vector and sequenced.

Scanning electron microscopy of rhinophore and oral tentacle

Adult *Aplysia* rhinophore and oral tentacle were fixed with 2.5% glutaraldehyde in phosphate buffer (pH 7.2 to 7.4) for 3 days at 4°C. Secondary fixation was in 1% OsO₄ (osmium tetroxide) in 0.1M sodium cacodylate. This material was dehydrated in a graded series of ethanol (20% to 100%). The samples were dried using hexamethyldisilazane and platinum-coated using Eiko IB-5 Sputter Coater. The specimens were viewed using a Jeol 6300 Field Emission Scanning Electron Microscope.

Laser capture microdissection and molecular identification of candidate A. californica chemosensory receptors

The location of the rhinophore groove, glomeruli underlying the sensory epithelium, and rhinophore ganglia in *Aplysia* have been described previously [26-28]. To examine whether selected genes are expressed in rhinophore sensory epithelial cells, a combination of LCM, total RNA isolation, library construction and RT-PCR were performed. Rhinophore tissue that had been embedded in OCT compound was sectioned (10 µm) onto slides and dehydrated. LCM was performed using a PixCell II laser capture microscope with an infrared diode laser (Arcturus Engineering Inc., Mountain View, CA, USA) and a laser spot size of 15 µm. Cells were marked, captured on Capture HS caps (Arcturus), and total RNA was isolated using the Picopure Isolation Kit (Arcturus) including DNase I incubation. RNA quality was determined by measurement of absorbance ratio at 260 nm/280 nm.

First-strand cDNA was synthesized according to the SMART cDNA Library Construction Kit protocol (Clontech, Palo Alto, CA, USA), with the minor modification of incorporated *Eco*RI restriction sites during cDNA synthesis

(5'-AAGCAGTGGTATCAACGCAGAGTGAAT-TCACGCGGG-3' and 5'-AAGCAGTGGTATCAACGCAGAGTGAATTCT₃₀VN-3'). Amplification of cDNA was performed by PCR using 5' and 3' PCR Primer mix (5'-CTAATACGACTCACTATAG-GGCAAGCAGTGGTATCAACGCAGAGT-3' and 5'-AAGCAGTGGTATCAACGCAGAGTGAATTCT₃₀VN-3'); samples were heated at 95°C for 1 min and amplified for 30 cycles (95°C, 1 min; 50°C, 30 s; 68°C, 4 min). Reaction volumes of 50 µl were treated with 2 µl of proteinase K (20 µg/µl) at 45°C for 20 min, and PCR products were precipitated. Dried samples were resuspended in 80 µl of deionized water. *EcoRI* digestion was performed and size fractionation was achieved using a CHROMA SPIN-400 column (Stratagene, La Jolla, CA, USA). Products were purified by precipitation and the dried pellet resuspended in 7 µl deionized water. *EcoRI* cDNA was cloned into the *EcoRI* sites of Lambda ZAP II vector, and the cDNA library (complexity 1 × 10⁶) amplified once.

The sequences of oligonucleotide primers (Sigma-Genosys, Australia) used for library PCR are located in Table S1 of Additional file 3. For AcCRa and AcCRb genes, PCR was performed using degenerate sense and antisense primer combinations A1 to A5 (AcCRa) and B1 to B8 (AcCRb). Each reaction was performed using Red *Taq* polymerase (Sigma) following the manufacturer's instructions. Samples were heated at 94°C for 3 min and amplified for 36 cycles (94°C, 60 s; 45°C, 30 s; 72°C, 60 s), followed by a 7-min extension at 72°C. PCR products were cloned into the TA vector pGEM-T (Promega) and sequenced as previously described [23]. To obtain 5' and 3' sequences, PCR was performed using gene-specific primers. For 3'-RACE, gene-specific sense primers (A3' and B3') were used in combination with vector primer T3. For 5'-RACE, gene-specific antisense primers (A5' and B5') were used in combination with vector primer T7. Samples were heated at 94°C for 2 min and amplified for 36 cycles (94°C, 60 s; 50°C, 30 s; 72°C, 60 s), followed by a 7-min extension at 72°C. PCR products were cloned into pGEM-T vector and sequenced.

Molecular identification of candidate *Aplysia* chemosensory receptors within subfamily c

Total RNA was extracted from oral tentacle tissue of *A. californica* using a Tripure Isolation Reagent (Roche), and any contaminating genomic DNA was removed by treatment with DNase I (Invitrogen). First strand cDNA synthesis was performed using 1 µg of total RNA in a 20 µl reverse transcription mixture containing oligo d(T)₁₂₋₁₈ and 200 U SuperscriptTM III RNase H⁻ reverse transcriptase, following the manufacturer's instructions. The sequences of oligonucleotide primers used for PCR are located in Table S1 of Additional file 3. PCR was performed using degenerate sense and antisense primer combinations C1-

C3. 3'- and 5'-RACE was performed with the SMART RACE amplification kit (BD Biosciences) and using gene-specific primers sense and antisense primers (C3' and C5'). Samples were heated at 94°C for 2 min and amplified for 36 cycles (94°C, 60 s; 50°C, 30 s; 72°C, 60 s), followed by a 7-min extension at 72°C. PCR products were cloned into a pGEM-T vector and sequenced.

Immunohistochemical localization

A rabbit polyclonal antibody was generated to the N-terminal region of candidate chemoreceptor 29b, corresponding to N₆SQARSSKSTQKGL (GenScript Corporation). This region was chosen due to its lack of significant amino acid identity to other receptors. Details of the immunohistochemical protocol have been described [23,24]. Briefly, tissue cryostat sections of rhinophore were incubated overnight at 4°C in either affinity-purified 29b antibody (0.6 mg/ml, 1:500), Gα_q, Gα_i or Gα_o antisera (Chemicon, 1:500 dilution), rinsed in phosphate buffered saline (PBS), incubated in fluorescein isothiocyanate (FITC)-conjugated goat anti-rabbit Ig (Sigma-Aldrich, St. Louis, MO, USA) for 1 h at 22°C, rinsed in PBS, and then mounted in FITC mounting media (90% glycerol/100 mM Tris pH 8.0). Preparations were examined using an Olympus FluoView confocal microscope (Leeds Precision Instruments, Inc., Minneapolis, USA), and the images captured on a spot-cooled charged coupled device camera. Sections were counterstained with 4',6-diamidino-2-phenylindole or propidium iodide at 1 µg/ml in water. As a control, the primary antiserum was replaced with no primary antibody. For 29b, a control also included using the primary antibody that had been preabsorbed against its antigenic peptide (20 µg/ml).

Abbreviations

7-TM: seven transmembrane; AcCR: *Aplysia californica* chemosensory receptor; BLAST: basic local alignment search tool; EST: expressed sequence tag; FITC: fluorescein isothiocyanate; GPCR: G-protein coupled receptor; LCM: laser capture microdissection; OCT: optimal cutting temperature; PBS: phosphate-buffered saline; PLC: phospholipase C; RT-PCR: reverse transcription-polymerase chain reaction; OR: olfactory receptor.

Authors' contributions

SFC, DE, GTN and BMD designed research. SFC and DE performed research. LLM and BMD analyzed data. SFC, ZZ, CC, GTN and BMD wrote the paper. All authors read and approved the final manuscript.

Additional material

Additional file 1

A. californica rhodopsin G-protein coupled receptor-like genes. This file summarizes details for all identified rhodopsin G-protein coupled receptor-like genes.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1741-7007-7-28-S1.xls>]

Additional file 2

Candidate A. californica chemosensory receptor alignments. The complete comparative amino acid alignments within AcCRa-c proteins.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1741-7007-7-28-S2.pdf>]

Additional file 3

Analysis of contig sequences and candidate chemosensory receptors.

Contains Figure S1, which shows the encoded protein sequence for a putative *Aplysia reverse transcriptase (RNA-dependent DNA polymerase)-like gene*; Figure S2, a comparative amino acid analysis for clustered AcCRb genes 11b/12b and 13b/14b; Figure S3, Comparative amino acid alignments of translated reverse transcription-polymerase chain reaction (RT-PCR) amplicons; Figure S4, Molecular identification of *Aplysia californica chemosensory genes*; and Table S1, a list of primers used for RT-PCR, PCR and gene cloning.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1741-7007-7-28-S3.pdf>]

Acknowledgements

We acknowledge the Broad Institute *Aplysia* Genome Initiative for making the partial genome sequence of *A. californica* available. We thank Dr Darren Boehning for comments on an earlier version of the manuscript. We acknowledge the assistance of the UTMB Protein Chemistry Lab, Steve LePage (MREP) and Erica Lovus (Institute of Molecular Biology). SFC is supported by a University of Queensland Fellowship. This research was supported by grants to BMD from the Australian Research Council, LLM from NIH and NSF, and to GTN from the NSF (grant IBN-0314377).

References

- Antonov I, Antonova I, Kandel ER, Hawkins RD: **Activity-dependent presynaptic facilitation and hebbian LTP are both required and interact during classical conditioning in *Aplysia*.** *Neuron* 2003, **37**:135-147.
- Balakrishnan R: **Learning from a sea snail: Eric kandel.** *Resonance* 2001, **6**:86-90.
- Nambu JR, Scheller RH: **Egg-laying hormone genes of *Aplysia*: evolution of the ELH gene family.** *J Neurosci* 1986, **6**:2026-2036.
- Painter SD, Cummins SF, Nichols AE, Akalal DB, Schein CH, Braun W, Smith JS, Susswein AJ, Levy M, de Boer PA, ter Maat A, Miller MW, Scanlan C, Milberg RM, Sweedler JV, Nagle GT: **Structural and functional analysis of *Aplysia* attractins, a family of water-borne protein pheromones with interspecific attractiveness.** *Proc Natl Acad Sci USA* 2004, **101**:6929-6933.
- Buck L, Axel R: **A novel multigene family may encode odorant receptors: a molecular basis for odor recognition.** *Cell* 1991, **65**:175-187.
- Liberles SD, Buck LB: **A second class of chemosensory receptors in the olfactory epithelium.** *Nature* 2006, **442**:645-650.
- Dulac C, Axel R: **A novel family of genes encoding putative pheromone receptors in mammals.** *Cell* 1995, **83**:195-206.
- Ryba NJ, Tirindelli R: **A new multigene family of putative pheromone receptors.** *Neuron* 1997, **19**:371-379.
- Li X, Staszewski L, Xu H, Durick K, Zoller M, Adler E: **Human receptors for sweet and umami taste.** *Proc Natl Acad Sci USA* 2002, **99**:4692-4696.
- Matsunami H, Montmayeur JP, Buck LB: **A family of candidate taste receptors in human and mouse.** *Nature* 2000, **404**:601-604.
- Touhara K: **Molecular biology of peptide pheromone production and reception in mice.** *Adv Genet* 2007, **59**:147-171.
- Zufall F, Leinders-Zufall T: **Mammalian pheromone sensing.** *Curr Opin Neurobiol* 2007, **17**:483-489.
- Cao Y, Oh BC, Stryer L: **Cloning and localization of two multigene receptor families in goldfish olfactory epithelium.** *Proc Natl Acad Sci USA* 1998, **95**:11987-11992.
- Zhang X, Firestein S: **The olfactory receptor genesuperfamily of the mouse.** *Nat Neurosci* 2002, **5**:124-133.
- Clyne PJ, Warr CG, Freeman MR, Lessing D, Kim J, Carlson JR: **A novel family of divergent seven-transmembrane proteins: candidate odorant receptors in *Drosophila*.** *Neuron* 1999, **22**:327-338.
- Vosshall LB, Amrein H, Morozov PS, Rzhetsky A, Axel R: **A spatial map of olfactory receptor expression in the *Drosophila* antenna.** *Cell* 1999, **96**:725-736.
- Robertson HM, Wanner KW: **The chemoreceptor superfamily in the honey bee, *Apis mellifera*: expansion of the odorant, but not gustatory, receptor family.** *Genome Res* 2006, **16**:1395-1403.
- Benton R, Sachse S, Michnick SW, Vosshall LB: **Atypical membrane topology and heteromeric function of *Drosophila* odorant receptors in vivo.** *PLoS Biol* 2006, **4**:e20.
- Sato K, Pellegrino M, Nakagawa T, Nakagawa T, Vosshall LB, Touhara K: **Insect olfactory receptors are heteromeric ligand-gated ion channels.** *Nature* 2008, **452**:1002-1006.
- Raible F, Tessmar-Raible K, Arboleda E, Kaller T, Bork P, Arendt D, Arnone MI: **Opsins and clusters of sensory G-protein-coupled receptors in the sea urchin genome.** *Dev Biol* 2006, **300**:461-475.
- Mobley AS, Mahendra G, Lucero MT: **Evidence for multiple signaling pathways in single squid olfactory receptor neurons.** *J Comp Neurol* 2007, **501**:231-242.
- Cummins SF, Schein CH, Xu Y, Braun W, Nagle GT: **Molluscan attractins, a family of water-borne protein pheromones with interspecific attractiveness.** *Peptides* 2005, **26**:121-129.
- Cummins SF, Nichols AE, Amare A, Hummon AB, Sweedler JV, Nagle GT: **Characterization of *Aplysia* antin and temptin, two novel water-borne protein pheromones that act in concert with attractin to stimulate mate attraction.** *J Biol Chem* 2004, **279**:25614-25622.
- Cummins SF, Nichols AE, Warso CJ, Nagle GT: ***Aplysia* seductin is a water-borne protein pheromone that acts in concert with attractin to stimulate mate attraction.** *Peptides* 2005, **26**:351-359.
- Levy M, Blumberg M, Susswein AJ: **The rhinophores sense pheromones regulating multiple behaviors in *Aplysia fasciata*.** *Neurosci Lett* 1997, **225**:113-116.
- Gobbeler K, Klussmann-Kolb A: **A comparative ultrastructural investigation of the cephalic sensory organs in Opisthobranchia (Mollusca, Gastropoda).** *Tissue Cell* 2007, **39**:399-414.
- Emery DG, Audesirk TE: **Sensory cells in *Aplysia*.** *J Neurobiol* 1978, **9**:173-179.
- Wertz A, Rossler W, Obermayer M, Bickmeyer U: **Functional neuroanatomy of the rhinophore of *Aplysia punctata*.** *Front Zool* 2006, **3**:6.
- Cummins SF, De Vries MR, Hill KS, Boehning D, Nagle GT: **Gene identification and evidence for expression of G protein alpha subunits, phospholipase C, and an inositol 1,4,5-trisphosphate receptor in *Aplysia californica* rhinophore.** *Genomics* 2007, **90**:110-120.
- Moroz LL: **Localization of putative nitrergic neurons in peripheral chemosensory areas and the central nervous system of *Aplysia californica*.** *J Comp Neurol* 2006, **495**:10-20.

31. Clifford KT, Gross L, Johnson K, Martin KJ, Shaheen N, Harrington MA: **Slime-trail tracking in the predatory snail, *Euglandina rosea*.** *Behav Neurosci* 2003, **117**:1086-1095.
32. Moroz LL, Edwards JR, Puthanveetil SV, Kohn AB, Ha T, Heyland A, Knudsen B, Sahni A, Yu F, Liu L, Jezzini S, Lovell P, Lannuccilli W, Chen M, Nguyen T, Sheng H, Shaw R, Kalachikov S, Panchin YV, Farmerie W, Russo JJ, Ju J, Kandel ER: **Neuronal transcriptome of *Aplysia*: neuronal compartments and circuitry.** *Cell* 2006, **127**:1453-1467.
33. Mulder NJ, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bork P, Bullard V, Cerutti L, Copley R, Courcelle E, Das U, Daugherty L, Dibley M, Finn R, Fleischmann W, Gough J, Haft D, Hulo N, Hunter S, Kahn D, Kanapin A, Kejariwal A, Labarga A, Langendijk-Genevaux PS, Lonsdale D, Lopez R, Letunic I, Madera M, Maslen J, et al: **New developments in the InterPro database.** *Nucleic Acids Res* 2007, **35**:D224-228.
34. Freitag J, Ludwig G, Andreini I, Rossler P, Breer H: **Olfactory receptors in aquatic and terrestrial vertebrates.** *J Comp Physiol [A]* 1998, **183**:635-650.
35. Katada S, Tanaka M, Touhara K: **Structural determinants for membrane trafficking and G protein selectivity of a mouse olfactory receptor.** *J Neurochem* 2004, **90**:1453-1463.
36. Scott K, Brady R Jr, Cravchik A, Morozov P, Rzhetsky A, Zuker C, Axel R: **A chemosensory gene family encoding candidate gustatory and olfactory receptors in *Drosophila*.** *Cell* 2001, **104**:661-673.
37. Rosenkilde MM, Kledal TN, Schwartz TW: **High constitutive activity of a virus-encoded seven transmembrane receptor in the absence of the conserved DRY motif (Asp-Arg-Tyr) in transmembrane helix 3.** *Mol Pharmacol* 2005, **68**:11-19.
38. Acharya S, Karnik SS: **Modulation of GDP release from transducin by the conserved Glu134-Arg135 sequence in rhodopsin.** *J Biol Chem* 1996, **271**:25406-25411.
39. Ballesteros J, Kitanovic S, Guarnieri F, Davies P, Fromme BJ, Konvicka K, Chi L, Millar RP, Davidson JS, Weinstein H, Sealfon SC: **Functional microdomains in G-protein-coupled receptors. The conserved arginine-cage motif in the gonadotropin-releasing hormone receptor.** *J Biol Chem* 1998, **273**:10445-10453.
40. Pathe-Neuschäfer-Rube A, Neuschäfer-Rube F, Püschel GP: **Role of the ERC motif in the proximal part of the second intracellular loop and the C-terminal domain of the human prostaglandin F2alpha receptor (hFP-R) in G-protein coupling control.** *Biochem J* 2005, **388**:317-324.
41. Wodicka LM, Morse DE: **cDNA sequences reveal mRNAs for two Gα signal transducing proteins from larval cilia.** *Biol Bull* 1991, **180**:318-327.
42. Shinohara H, Asano T, Kato K: **Differential localization of G-proteins Gi and Go in the accessory olfactory bulb of the rat.** *J Neurosci* 1992, **12**:1275-1279.
43. Jia C, Goldman G, Halpern M: **Development of vomeronasal receptor neuron subclasses and establishment of topographic projections to the accessory olfactory bulb.** *Brain Res Dev Brain Res* 1997, **102**:209-216.
44. Berghard A, Buck L: **Sensory transduction in vomeronasal neurons: evidence for Gα_o, Gα_{i2}, and adenylyl cyclase II as major components of a pheromone signaling cascade.** *J Neurosci* 1996, **16**:909-918.
45. Vassar R, Chao SK, Sitcheran R, Nunez JM, Vosshall LB, Axel R: **Topographic organization of sensory projections to the olfactory bulb.** *Cell* 1994, **79**:981-991.
46. Ferrand N, Pessah M, Frayon S, Marais J, Garel JM: **Olfactory receptors, Golf alpha and adenylyl cyclase mRNA expressions in the rat heart during ontogenic development.** *J Mol Cell Cardiol* 1999, **31**:1137-1142.
47. Parmentier M, Libert F, Schurmans S, Schiffmann S, Lefort A, Eggerickx D, Ledent C, Mollereau C, Gerard C, Perret J, Grootegoed A, Vassart G: **Expression of members of the putative olfactory receptor gene family in mammalian germ cells.** *Nature* 1992, **355**:453-455.
48. Fukuda N, Yomogida K, Okabe M, Touhara K: **Functional characterization of a mouse testicular olfactory receptor and its role in chemosensing and in regulation of sperm motility.** *J Cell Sci* 2004, **117**:5835-5845.
49. Branscomb A, Seger J, White RL: **Evolution of odorant receptors expressed in mammalian testes.** *Genetics* 2000, **156**:785-797.
50. Krieger J, Klink O, Mohl C, Raming K, Breer H: **A candidate olfactory receptor subtype highly conserved across different insect orders.** *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 2003, **189**:519-526.
51. Hill CA, Fox AN, Pitts RJ, Kent LB, Tan PL, Chrystal MA, Cravchik A, Collins FH, Robertson HM, Zwiebel LJ: **G protein-coupled receptors in *Anopheles gambiae*.** *Science* 2002, **298**:176-178.
52. Go Y, Niimura Y: **Similar numbers but different repertoires of olfactory receptor genes in humans and chimpanzees.** *Mol Biol Evol* 2008, **25**:1897-1907.
53. Bohbot J, Pitts RJ, Kwon HW, Rutzler M, Robertson HM, Zwiebel LJ: **Molecular characterization of the *Aedes aegypti* odorant receptor gene family.** *Insect Mol Biol* 2007, **16**:525-537.
54. Troemel ER, Chou JH, Dwyer ND, Colbert HA, Bargmann CI: **Divergent seven transmembrane receptors are candidate chemosensory receptors in *C. elegans*.** *Cell* 1995, **83**:207-218.
55. Xie SY, Feinstein P, Mombaerts P: **Characterization of a cluster comprising approximately 100 odorant receptor genes in mouse.** *Mamm Genome* 2000, **11**:1070-1078.
56. Nozawa N, Masatochi N: **Evolutionary dynamics of olfactory receptor genes in *Drosophila* species.** *Proc Natl Acad Sci USA* 2007, **104**:7122-7127.
57. Stewart MK, Clark NL, Merrihew G, Galloway EM, Thomas JH: **High genetic diversity in the chemoreceptor superfamily of *Caenorhabditis elegans*.** *Genetics* 2005, **169**:1985-1996.
58. McBride C: **Rapid evolution of smell and taste receptor genes during host specialization in *Drosophila sechellia*.** *Proc Natl Acad Sci USA* 2007, **104**:4996-5001.
59. Robertson HM, Warr CG, Carlson JR: **Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster*.** *Proc Natl Acad Sci USA* 2003, **100**(Suppl 2):14537-14542.
60. Liman ER, Innan H: **Relaxed selective pressure on an essential component of pheromone transduction in primate evolution.** *Proc Natl Acad Sci USA* 2003, **100**:3328-3332.
61. **NCBI trace database** [http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi?organism=euk]
62. ***Aplysia* genome sequencing, Broad Institute at MIT** [<http://www.genome.gov/Pages/Research/Sequencing/SeqProposals/AplysiaSeq.pdf>]
63. **HMMTOP v2** [<http://www.enzim.hu/hmmtop/index.html>]
64. **TMHMM program** [<http://www.enzim.hu/hmmtop/html/submit.html>]
65. **Pfam** [<http://pfam.janelia.org/search>]
66. Notredame C, Higgins DG, Heringa J: **T-Coffee: A novel method for fast and accurate multiple sequence alignment.** *J Mol Biol* 2000, **302**:205-217.
67. Stamatakis A: **RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models.** *Bioinformatics* 2006, **22**:2688-2690.
68. Ronquist F, Huelsenbeck JP: **MrBayes 3: Bayesian phylogenetic inference under mixed models.** *Bioinformatics* 2003, **19**:1572-1574.
69. Tavare S: **Some probabilistic and statistical problems on the analysis of DNA sequences.** *Lect Math Life Sci* 1986, **17**:57-86.
70. Whelan S, Goldman N: **A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach.** *Mol Biol Evol* 2001, **18**:691-699.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

