Check for updates

RESEARCH ARTICLE

REVISED Phylogeny-wide analysis of G-protein coupled receptors in social amoebas and implications for the evolution of multicellularity

[version 2; peer review: 2 approved]

Grant Hall[1], Sarah Kelly[1,2], Pauline Schaap[1], Christina Schilde[1]

[1]School of Life Sciences, University of Dundee, Dundee, DD1 5EH, UK
[2]Centre for Inflammation Research, University of Edinburgh, Edinburgh, EH16 4TJ, UK

v2 **First published:** 06 Dec 2022, **2**:134
https://doi.org/10.12688/openreseurope.15250.1
Latest published: 15 Feb 2023, **2**:134
https://doi.org/10.12688/openreseurope.15250.2

Abstract

G-protein coupled receptors (GPCRs) are seven-transmembrane proteins and constitute the largest group of receptors within eukaryotes. The presence of a large set of GPCRs in the unicellular Amoebozoa was surprising and is indicative of the largely undiscovered environmental sensing capabilities in this group. Evolutionary transitions from unicellular to multicellular lifestyles, like we see in social amoebas, have occurred several times independently in the Amoebozoa, and GPCRs may have been co-opted for new functions in cell-cell communication.

Methods

We have analysed a set of GPCRs from fully sequenced Amoebozoan genomes by Bayesian inference, compared their phylogenetic distribution and domain composition, and analysed their temporal and spatial expression patterns in five species of dictyostelids.

Results

We found evidence that most GPCRs are conserved deeply in the Amoebozoa and are probably performing roles in general cell functions and complex environmental sensing. All families of GPCRs (apart from the family 4 fungal pheromone receptors) are present in dictyostelids with family 5 being the largest and family 2 the one with the fewest members. For the first time, we identify the presence of family 1 rhodopsin-like GPCRs in dictyostelids. Some GPCRs have been amplified in the dictyostelids and in specific lineages thereof and through changes in expression patterns may have been repurposed for signalling in multicellular development.

Discussion

Our phylogenetic analysis suggests that GPCR families 1, 2 and 6 already diverged early in the Amoebozoa, whereas families 3 and 5 expanded later within the dictyostelids. The family 6 cAMP receptors that have experimentally supported roles in multicellular development in dictyostelids (*carA-carD*; *tasA/B*) originated at the root of all

Open Peer Review

Approval Status ✓ ✓

	1	2
version 2 (revision) 15 Feb 2023	✓ view	✓ view
version 1 06 Dec 2022	? view	? view

1. **Jeffrey A. Hadwiger**, Oklahoma State University, Stillwater, USA

2. **Xuehua Xu**, National Institutes of Health (NIAID/NIH), Rockville, USA

Any reports and responses or comments on the article can be found at the end of the article.

dictyostelids and only have weakly associated homologs in *Physarum polycephalum.* Our analysis identified candidate GPCRs which have evolved in the dictyostelids and could have been co-opted for multicellular development.

Keywords
G-protein coupled receptor, dictyostelids, social amoebas, multicellularity, signalling, Amoebozoa

This article is included in the Cell, Molecular and Structural Biology gateway.

This article is included in the European Research Council (ERC) gateway.

This article is included in the Horizon 2020 gateway.

This article is included in the Evolution and Ecology gateway.

This article is included in the Evolutionary Biology collection.

Corresponding author: Christina Schilde (c.schilde@dundee.ac.uk)

Author roles: Hall G: Formal Analysis, Investigation, Visualization; **Kelly S**: Formal Analysis, Investigation, Visualization; **Schaap P**: Conceptualization, Funding Acquisition, Methodology, Writing – Review & Editing; **Schilde C**: Data Curation, Investigation, Supervision, Visualization, Writing – Original Draft Preparation

Competing interests: No competing interests were disclosed.

Grant information: This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 742288)
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2023 Hall G *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Hall G, Kelly S, Schaap P and Schilde C. **Phylogeny-wide analysis of G-protein coupled receptors in social amoebas and implications for the evolution of multicellularity [version 2; peer review: 2 approved]** Open Research Europe 2023, **2**:134 https://doi.org/10.12688/openreseurope.15250.2

First published: 06 Dec 2022, **2**:134 https://doi.org/10.12688/openreseurope.15250.1

REVISED **Amendments from Version 1**

A table with the abbreviations for all domains and gene names used has been added. The labelling in the heat maps of Figure 1–Figure 5 have been simplified for better readability and the legend to Figure 1 has been updated accordingly. The legend to Figure 3 has been corrected to reflect the fact that this phylogeny was obtained with RAxML and the values on the nodes represent bootstrap values. Figure 4 had been slighlty cut off at the bottom and a full Figure 4 has been included. A warning to the reader about inconsistency between RNAseq replicates has been inserted into the methods section. On suggestion of the reviewers, a brief discussion of the differences between the newly identified rhodpsin-like receptors of dictyostelids and animals as well as the bacteriorhodopsins has been inserted. The unknown binding affinity of Crls towards cAMP has been clarified.

Any further responses from the reviewers can be found at the end of the article

Introduction

G-protein coupled receptors, also known as seven-transmembrane receptors (7TMR) or serpentine receptors, comprise the largest class of conserved eukaryotic transmembrane proteins and are involved in sensing extracellular signals, such as odorants and signals used for cell-cell communication (Bockaert & Pin, 1999). They perform various environmental sensing and signalling functions using heterotrimeric G-proteins as mediators. GPCRs of vertebrates are important regulators during development and tissue homeostasis and the over 800 human members are the most researched drug targets (Insel *et al.*, 2019; Yang *et al.*, 2021). The seven transmembrane helices are the core unit of all GPCRs and they can have additional extracellular *N*-terminal sensing and *C*-terminal intracellular domains. Some GPCRs act as homo- or heterodimers and also associate with a variety of other membrane proteins (Bockaert & Pin, 1999). Ligand binding to the extracellular domain, or in case of the rhodopsin-like family within the transmembrane helices, results in intramolecular changes in the intracellular loops 2 and 3 which activates the GPCR's ability to exchange GDP for GTP and thereby activate the associated G proteins. The G-protein complex then dissociates into a G-alpha unit and a G-beta/gamma dimer, which can each activate downstream proteins, such as adenylyl cyclases, phospholipase C, ion channels, kinases, Rho-family guanine nucleotide exchange factors and other small G-proteins (McCudden *et al.*, 2005). Because the effects of many hormones, neurotransmitters and neuropeptides are mediated by GPCRs, they are important targets for drug development.

Based on sequence similarity, GPCRs have been subdivided into six families (Alexander *et al.*, 2019; Bockaert & Pin, 1999; Foord *et al.*, 2005; Hu *et al.*, 2017). A more recent phylogenetic study also recognises more atypical GPCRs like GPR108, which plays a role in the innate immune response and is activated by gambogic acid (Lyu *et al.*, 2022), and the intimal thickness-related receptor (ITR), an orphan receptor involved in vascular remodelling (Tsukada *et al.*, 2003), as universal eukaryote GPCRs based on their HMM (hidden Markov model) predictions (Nordstrom *et al.*, 2011). The Rhodopsin-like family 1 GPCRs is the largest group comprising several subfamilies of hormone, neuropeptide, neurotransmitter, nucleotide and light receptors. The family 2 secretin-like receptors are activated by glucagon-/secretin-like hormones and are related to the adhesion family GPCRs, which differ by containing a large cleavable *N*-terminus (Alexander *et al.*, 2019). Family 3 comprises the metabotropic glutamate receptors. GPCRs of family 4 are pheromone mating factor receptors, exclusively found in fungi. The family 5 frizzled/smoothened GPCRs have widespread signalling roles in animal development and adult tissue homeostasis. The dictyostelid cAMP receptors comprise family 6 and have no orthologues outside this group.

Dictyostelid social amoebas are multicellular members of the otherwise mostly unicellular Amoebozoa, a sister group to the Opisthokonta, which comprise fungi and and animals. Dictyostelia are phylogenetically divided into two main branches each consisting of two major groups and some minor group intermediates (Sheikh *et al.*, 2015). Branch I consists of group 1 (Cavenderiaceae) and group 2 (Acytosteliaceae) whereas branch II comprises groups 3 (Raperosteliaceae) and 4 (Dictyosteliaceae) (Sheikh *et al.*, 2018). Dictyostelids display a different type of multicellularity than animals, fungi and plants becoming multicellular by aggregation of individual cells. Dictyostelid amoebas feed on bacteria as single cells but survive starvation by aggregating and forming asexual multicellular fruiting bodies (sorocarps) that can consist of up to five different cell types. The two main cell types, spore and stalk cells, are present in all dictyostelid families except the Acytosteliaceae, which only produce spores. Whereas the unicellular spores can survive harsh environmental conditions for a long time and are the unit of dispersal for dictyostelids, stalk cells undergo altruistic cell death in the process of fruiting body formation. Differentiation into prespore and prestalk cells starts shortly after aggregation, but terminal encapsulation of spore and stalk cells occurs only during fruiting body formation. Like most Amoebozoa, dictyostelid amoebas live as unicellular bacteriophagic amoebas in soil. This requires a certain repertoire of environmental sensing receptors for finding food bacteria, avoiding predation and other stresses, and to find mating partners. Having inherited a repertoire of GPCRs from their last unicellular ancestor, it is conceivable that this was adapted and expanded during dictyostelid evolution to enable communication between cells during the multicellular phase. When Dictyostelid amoebas starve or are stressed, some species will encyst individually just like many other free-living amoebas. More commonly, deprivation of food will trigger aggregation and multicellular development, which all known dictyostelid species are capable of. The chemoattractant during aggregation is cAMP for most species in group 4 and glorin for some in groups 1 and 2 (Romeralo *et al.*, 2013), while some group 3 species use folate or pterin. cAMP is involved in morphogenesis of all dictyostelids and the cAMP receptors represent GPCR family 6. Genomes are available for *Dictyostelium* species that represent the four major

groups and for several related solitary Amoebozoa. For some Dictyostelia stage- and cell-type specific transcriptomes are also available. In this study we have used these data to assess conservation and change in the complete repertoire of GPCR genes across Dictyostelia and to investigate changes in developmental expression and functional domain architecture in the genes in the course of dictyostelid evolution.

Methods

Data retrieval

An InterProScan (RRID:SCR_005829) (Jones *et al.*, 2014) was carried out of all transmembrane proteins from the published proteomes of *Dictyostelium discoideum* (DDB) (Eichinger *et al.*, 2005), *Dictyostelium purpureum* (DPU) (Sucgang *et al.*, 2011), (Glöckner *et al.*, 2016), *Polysphondylium pallidum* (PPL), *Dictyostelium fasciculatum* (DFA) (Heidel *et al.*, 2011), *Acanthamoeba castellanii* (ACA) (Clarke *et al.*, 2013), *Physarum polycephalum* (Phypo) (Schaap *et al.*, 2015) and *Protostelium aurantium* var. *fungivorum* (PROFUN) (Hillmann *et al.*, 2018). G-protein coupled receptors were identified by their signature InterProScan GPCR domains. Comparison with earlier studies of GPCRs in some of the species revealed that they had not captured all family members. Representative members of each family were also used as bait in BLASTp (RRID:SCR_001010) searches of the eight Amoebozoan genomes to complete the dataset.

All GPCR sequences were aligned with Clustal Omega (RRID:SCR_001591) (Sievers *et al.*, 2011) and a pilot tree was inferred by RAxML (RRID:SCR_006086) (Stamatakis, 2014). This tree subdivided the full set into five major clades representing each GPCR family. New alignments and preliminary phylogenetic trees were generated for each clade with MrBayes 3.2.7 (RRID:SCR_012067) (Ronquist *et al.*, 2012). These trees were scrutinised for any missing members of otherwise orthologous sets, and further BLASTp or tBLASTn (RRID:SCR_011822) searches were performed with a member of the set as bait. This yielded in some cases incorrectly predicted gene models, which were repaired guided by orthologous sequences, before final phylogenetic trees were inferred.

Phylogenetic analysis and tree annotation

The sequences of the 7TM domains plus 20 amino acids of flanking sequence were aligned with Clustal Omega (Sievers *et al.*, 2011), large gaps and regions of poorly aligned sequence were deleted with Jalview 2.11.1.3 (RRID:SCR_006459) (Waterhouse *et al.*, 2009) or BioEdit 7.0.5.3 (RRID:SCR_007361) (Hall, 1999) and phylogenies were inferred using RAxML with 100 bootstrap replicates or using MrBayes with a mixed amino acid model and gamma distribution for rate variation between sites. Analyses were continued until the SD of split frequencies was <0.01 or up to 3,000,000 generations. The trees were mid-point rooted in FigTree v1.4.4 (RRID:SCR_008515) (A.Rambaut, University of Edinburgh, UK) and exported as .pdf files to Adobe Illustrator (RRID:SCR_010279) for further annotation.

Domain signatures were obtained using SMART (RRID:SCR_005026) (Schultz *et al.*, 1998). To avoid automatic retrieval of old models, we replaced the start methionine by another amino acid. In case of overlapping domains, the domains with the lowest e-value or largest coverage were selected. The domain architectures were saved as a .svg files and placed next to their corresponding locus tags onto the phylogenetic trees.

Developmental and cell-type specific expression data were retrieved from published RNAseq experiments of the five Dictyostelid species (Glöckner *et al.*, 2016; Kin *et al.*, 2018; Parikh *et al.*, 2010). Data were normalised as fraction of the maximum transcript read count for the developmental profiles (as in (Forbes *et al.*, 2019)) and as fraction of the summed read counts for the purified cell types. Expression value sums of 10 or less per experiment were highlighted by 30% opacity. Expression data was combined for each orthologous clade and shown as a heat map. Where no expression data could be found the row was left blank. This happened frequently when gene models had changed between different expression experiments. Where expression data was available, but the new gene model indicated the presence of an earlier gene fusion, the data was omitted because expression reads were not re-mapped and could not be attributed unequivocally to the either of the fused genes. In some instances, the duplicate gene expression experiments did vary considerably and caution should be taken in the interpretation of any gene expression trends.

Results and discussion

The proteomes of the dictyostelids *Dictyostelium discoideum* (DDB) (Eichinger *et al.*, 2005), *Dictyostelium purpureum* (DPU) (Sucgang *et al.*, 2011), *Dictyostelium lacteum* (DLA) (Glöckner *et al.*, 2016), *Polysphondylium pallidum* (PPL) and *Dictyostelium fasciculatum* (DFA) (Heidel *et al.*, 2011) and the solitary amoebozoans *Acanthamoeba castellanii* (ACA) (Clarke *et al.*, 2013), *Protostelium fungivorum* (PROFUN) (Hillmann *et al.*, 2018) and the related myxogastrid *Physarum polycephalum* (PHYPO) (Schaap *et al.*, 2015) were subjected to Interproscan (Jones *et al.*, 2014) to identify their repertoires of protein functional domains. Proteins in the six classes of GPCRs were first identified by the Interpro identifiers of each class. Comparison with earlier studies of GPCRs in some of the species (Eichinger *et al.*, 2005) (Heidel *et al.*, 2011) revealed that these data sets were incomplete since the earlier studies had not captured all family members. Further BLASTp searches with the most diverged members of each class were performed to identify any missing proteins. A total of 414 non-redundant GPCRs were identified (Table 1), which were assigned to the GPCR class for which their Interpro domains showed the lowest E-value and for cases where this was equivocal according to their phylogenetic affinity. No family 4 fungal pheromone GPCRs were found. Separate phylogenetic trees were prepared for each family, which were annotated with protein functional domains, and, for the dictyostelids, with heatmaps of developmental- and cell type- specific gene expression (Glöckner *et al.*, 2016; Kin *et al.*, 2018; Parikh *et al.*, 2010) (Figure 1–Figure 6).

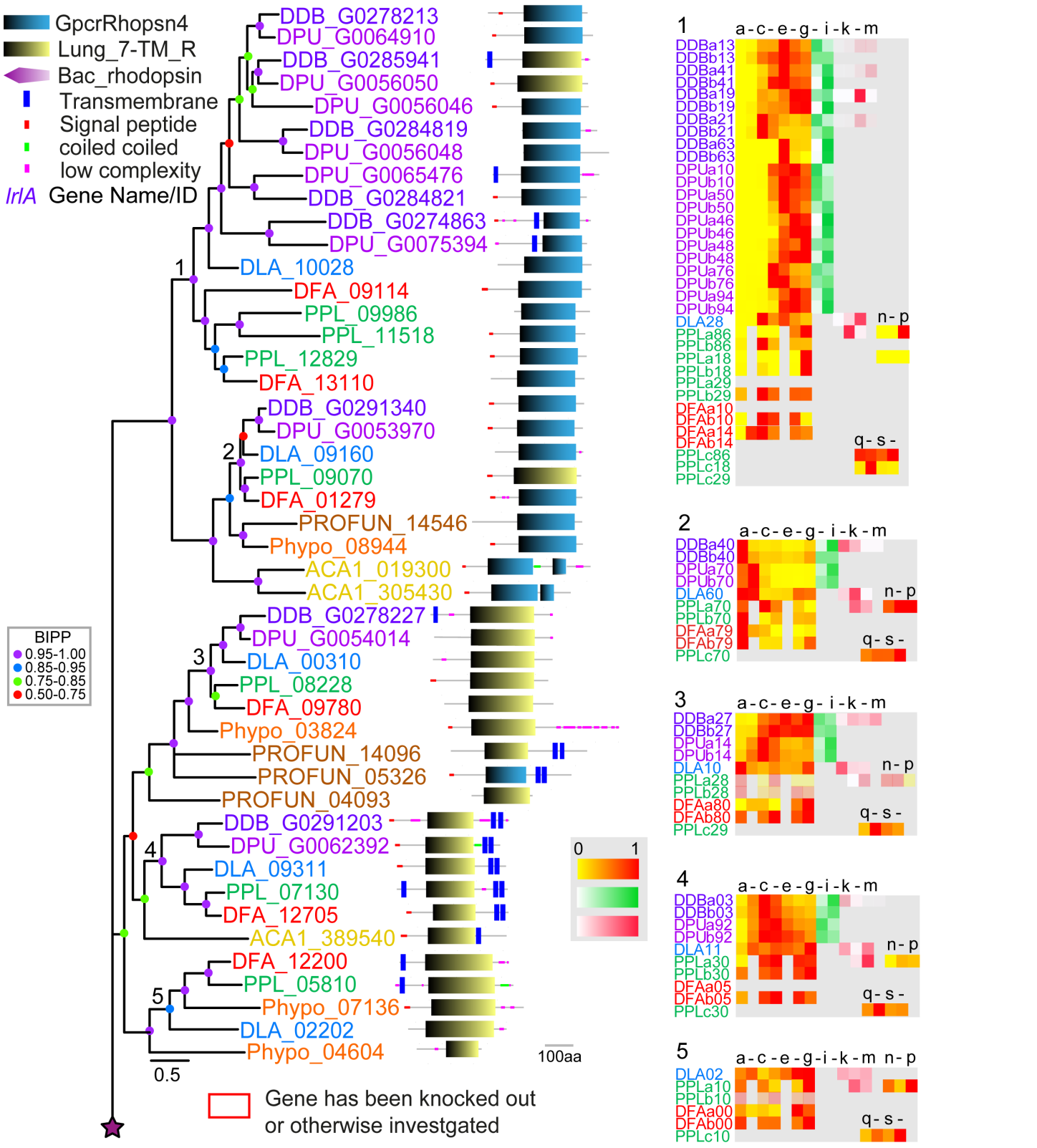
Table 1. List of the number of proteins found per family and per species. Note that family 4 contains only fungal pheromone receptors and is not present in Amoebozoa. Outliers in numbers compared to the whole set were identified and are indicated in bold italic numbers. * Clade numbers only refer to clades with Dictyostelid members. In brackets are previously reported gene numbers from *D.dis* (Prabhu & Eichinger, 2006)/(Nordstrom *et al.*, 2011) and *A.cas* (Clarke *et al.*, 2013).

species	Family 1	Family 2	Family 3	Family 4	Family 5	Family 6
A.cas	5 (8)	4 (5)	***0*** (0)	-	16 (17)	7 (6)
P.fun	8	1	***0***	-	0	5
Phy.pol	9	4	***37***	-	***35***	***39***
D.dis	10 (0/[3])	3 (2/1)	17 (17/15)	-	27 (27/25)	12 (13/12)
D.pur	***13***	3	11	-	15	9
D.lac	8	2	10	-	8	8
P.pal	9	3	15	-	8	12
D.fas	8	2	7	-	14	11
*clades**	7	2	11	-	8	10

Family 1: Rhodopsin-like GPCRs

19 subfamilies of rhodopsin-like GPCRs are recognised and grouped mainly by their type of ligands (Joost & Methner, 2002). They are often involved in autocrine, endocrine and paracrine functions but a number of family 1 receptors are still orphan receptors with unknown ligand. Most rhodopsin-type GPCRs only contain a short *N*-terminus and bind their ligands inside the 7TM region. They are additionally membrane-anchored by palmitoylation of a conserved cysteine at the *C*-terminus of an additional eighth helix (Hu *et al.*, 2017). No *C*-terminal palmitoylation motifs (apart from two low scoring ones in DFA_08662 and PHYPO_11504) were found in the Amoebozoan rhodopsin-like GPCRs using CSS-Palm palmitoylation site prediction (Ren *et al.*, 2008). Rhodopsin-like receptors were not noted in an earlier study (Prabhu & Eichinger, 2006), but six genes in *D.dis* were listed as "related to human transmembrane protein 145" and one as orphan receptor in a preceding comparative phylogenomic analysis (Heidel *et al.*, 2011). Those six genes are forming clades 1 and 2 of the phylogenetic tree of family 1 in the present analysis. Nordström *et al.* 2011 included atypical receptors and identified two GPR108-like and one ITR-like receptor in *D.dis* as separate branches. As demonstrated by the same authors, the *Dictyostelium* Rhodopsin-like GPCRs belong to a larger group of ancestral GPCRs from which the metazoan rhodopsins later evolved and diversified (Nordstrom *et al.*, 2011). They show no extensive sequence homology to either animal or bacterial rhodopsins and have not been acquired by recent horizontal gene transfer.

Phylogenetic inference subdivides the amoebozoan family 1 GPCRs into seven clades (Figure 1). Apart from clade 5, which is missing group 4 orthologues, all clades show a complete set of dictyostelid orthologues from all five species. This high level of gene conservation differs from the other GPCR families. Exhaustive Blast searches failed to identify orthologues of clade 5 in *D.dis* or *D.pur* so, we must assume those genes were lost in the group 4 last common ancestor (LCA). All clades apart from clade 1 also contain at least one orthologue from either *A.cas*, *Pro.fun* or *Phy.pol.*, suggesting that these clades have deep ancestral roots in the solitary Amoebozoa, a notion that is also supported by the presence of a small clade of associated *Phy.pol* sequences at the root. Clade 1 shows extensive amplification of receptor genes in groups 2 and 4, with the group 4 duplicated genes present in both *D.dis* and *D.pur*, suggesting that the duplication event occurred in the LCA to group 4. Compared to the other GPCR families (Figure 6D), the majority of dictyostelid family 1 GPCRs show a preference for expression in prestalk and stalk cells suggesting that these receptors might mediate effects of stalk-inducing signals. The receptors of clades 1 and 2 share their domain signature with animal rhodopsins whereas the other clades share domain similarity with the lung seven transmembrane receptors GPR107 and GPR108 (PF06814). GPR107 localises to the trans-Golgi network in human cells and is required for retrograde transport (Tafesse *et al.*, 2014). GPR108 is also a trans-Golgi protein and required for Adeno-associated virus entry into the nucleus (Dudek *et al.*, 2020). This subfamily also has homologues in fungi and plants and may have a conserved eukaryotic function in Golgi transport (Edgar, 2007). We find typical bacterial rhodopsin-like domains only in the small outgroup of *Phy.pol* proteins where the possibility of horizontal gene transfer from bacteria remains open. Unfortunately, no specific functions are known yet of rhodopsin family receptors in *Dictyostelium*.


GpcrRhopsn4
Lung_7-TM_R
Bac_rhodopsin
Transmembrane
Signal peptide
coiled coiled
low complexity
IrlA Gene Name/ID
DDB_G0278213
DPU_G0064910
DDB_G0285941
DPU_G0056050
DPU_G0056046
DDB_G0284819
DPU_G0056048
DPU_G0065476
DDB_G0284821
DDB_G0274863
DPU_G0075394
DLA_10028
DFA_09114
PPL_09986
PPL_11518
PPL_12829
DFA_13110
DDB_G0291340
DPU_G0053970
DLA_09160
PPL_09070
DFA_01279
PROFUN_14546
Phypo_08944
ACA1_019300
ACA1_305430
DDB_G0278227
DPU_G0054014
DLA_00310
PPL_08228
DFA_09780
Phypo_03824
PROFUN_14096
PROFUN_05326
PROFUN_04093
DDB_G0291203
DPU_G0062392
DLA_09311
PPL_07130
DFA_12705
ACA1_389540
DFA_12200
PPL_05810
Phypo_07136
DLA_02202
Phypo_04604
1
2
3
4
5
BIPP
0.95-1.00
0.85-0.95
0.75-0.85
0.50-0.75
0
1
0.5
100aa
Gene has been knocked out
or otherwise investgated
a-c-e-g-i-k-m
n-p
q-s-

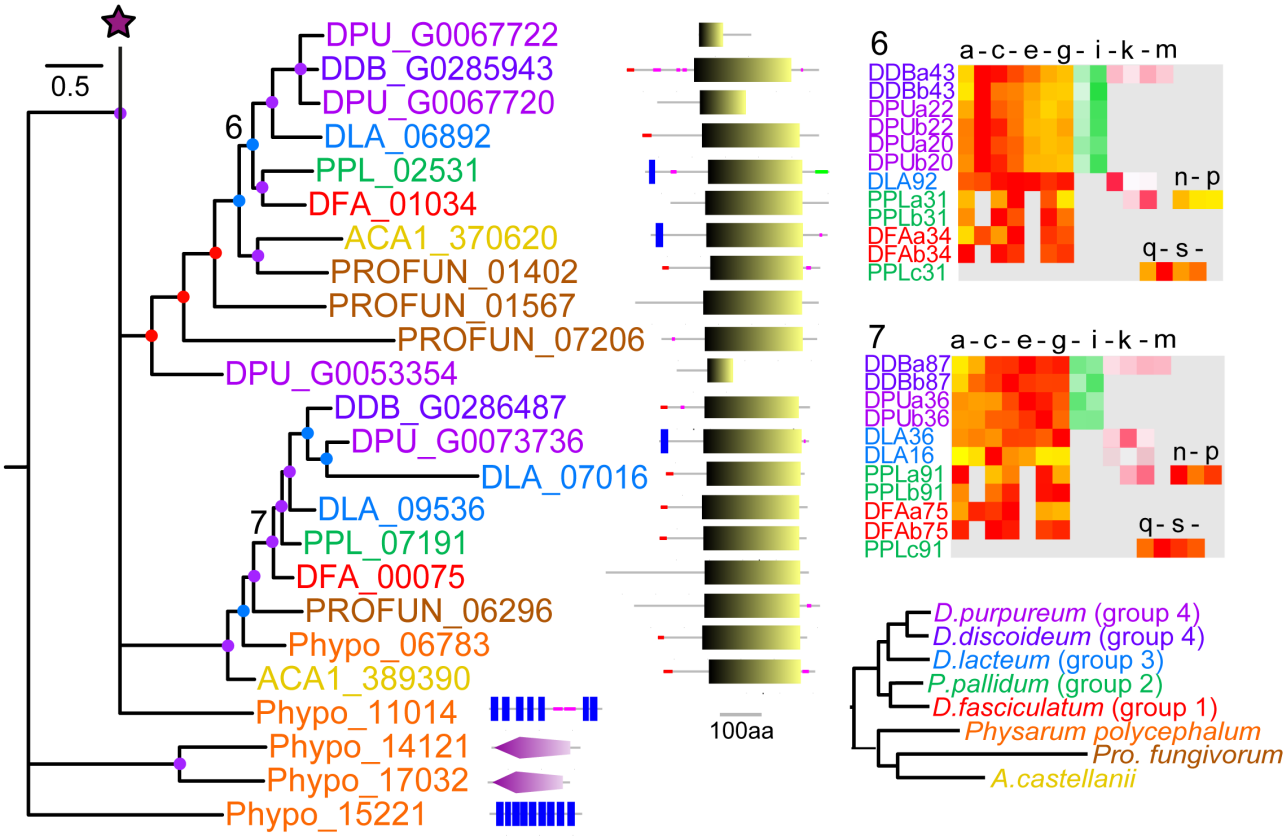


Figure 1. The family 1 rhodopsin-like GPCRs. Family 1 GPCRs were identified from InterproScan of the 8 Amoebozoan proteomes by the presence of InterproIDs IPR019336 or IPR009637 and further BLAST searches with rhodopsin-like receptors. Sequences were aligned and the phylogeny was inferred by Bayesian analysis. The scale bar indicates number of changes per site. Large trees have been split up and should be joined at the star signs. Locus tags are colour coded according to the species phylogenetic position (panel bottom right) and posterior probabilities of tree nodes (BIPP) are indicated by coloured dots. The locus tags are annotated with the SMART domain architecture of the protein shown next to it. Relative developmental stage and cell-type specific expression data for orthologous dictyostelid genes were presented as set of heatmaps for each clade, with transcriptional profiles for each gene identified by the first three and last two digits of the locus tag and 'a' and 'b' signifying replicate experiments. Legend for the heat maps (only alternative letters shown): (a-g) developmental stages, a: vegetative; b: lawn, starving cells, c: aggregates; d: tipped mounds; e: migrating slugs; f: early to mid-fruiting bodies; g: complete fruiting bodies; (h, i) precell types, h: prespore and i: prestalk cells; (j-m) mature cell types, j: vegetative, k: spore, l: stalk and m: cup cells; (n-p) and (q-t) encystation time courses of P.pal 0h to 48h into encystation, n: 0h, o: 24h, p: 48h, q: 0h, r: 8h, s: 16h, t: 24h. The colour legend beside the heat maps indicates the level of expression (yellow-red: (developmental expression): fraction of maximum value, white-green (pre-cell type expression) and white-red (cell-type specific expression): fraction of summed reads). For genes with summed maximal read counts <10, the heat map is shown at 30% opacity. The normalised transcript counts were obtained from published data (Glöckner *et al.*, 2016; Kin *et al.*, 2018; Parikh *et al.*, 2010). Note that not all stage and cell-type data is available for all species. Genes with known functions are boxed red and assigned names for *D.dis* genes are shown in purple italic font. *D.dis* locus tags for genes that result from the strain AX4 specific chromosome 2 duplication are indicated by 50% opacity. Table 2 contains a list of all abbreviations used including for domain names. Seven well-supported clades with either GpcrRhopsn4 or Lung_7-TM_R domains are recognised. In addition, two Phypo receptors forming an unresolved outgroup to the whole tree show bacteriorhodopsin domains.

Family 2: Secretin-like/adhesion GPCRs

Family 2 GPCRs combines secretin-like and adhesion-type GPCRs (Fredriksson *et al.*, 2003). The secretin-like GPCRs bind small peptides like secretin, calcitonin-gene related peptide, corticotropin-releasing factor, glucagon and glucagon-like peptides, growth hormone-releasing hormone, parathyroid hormone/parathyroid hormone-related peptides and vasoactive intestinal peptide which activate either the adenylyl cyclase pathway and/or the phosphatidyl-inositol-calcium pathway (Lin *et al.*, 1991; Takei, 2016; Yang *et al.*, 2021). The secretin-like GPCRs usually contain a large extracellular *N*-terminal hormone-binding domain (Yang *et al.*, 2021). Adhesion-type GPCRs mediate cell-cell and cell-extracellular matrix sensing and were identified as the ancestral type of this family (Schiöth *et al.*, 2010). The adhesion-type GPCRs of animals, too, contain a large extracellular domain which is cleaved from the C-terminal domain by an interspersed GPCR auto-proteolysis inducing (GAIN) domain (Vizurraga *et al.*, 2020). The GAIN domain encloses a tethered peptide-agonist stalk (also known as *Stachel* sequence) and dissociation of the *N*-terminal

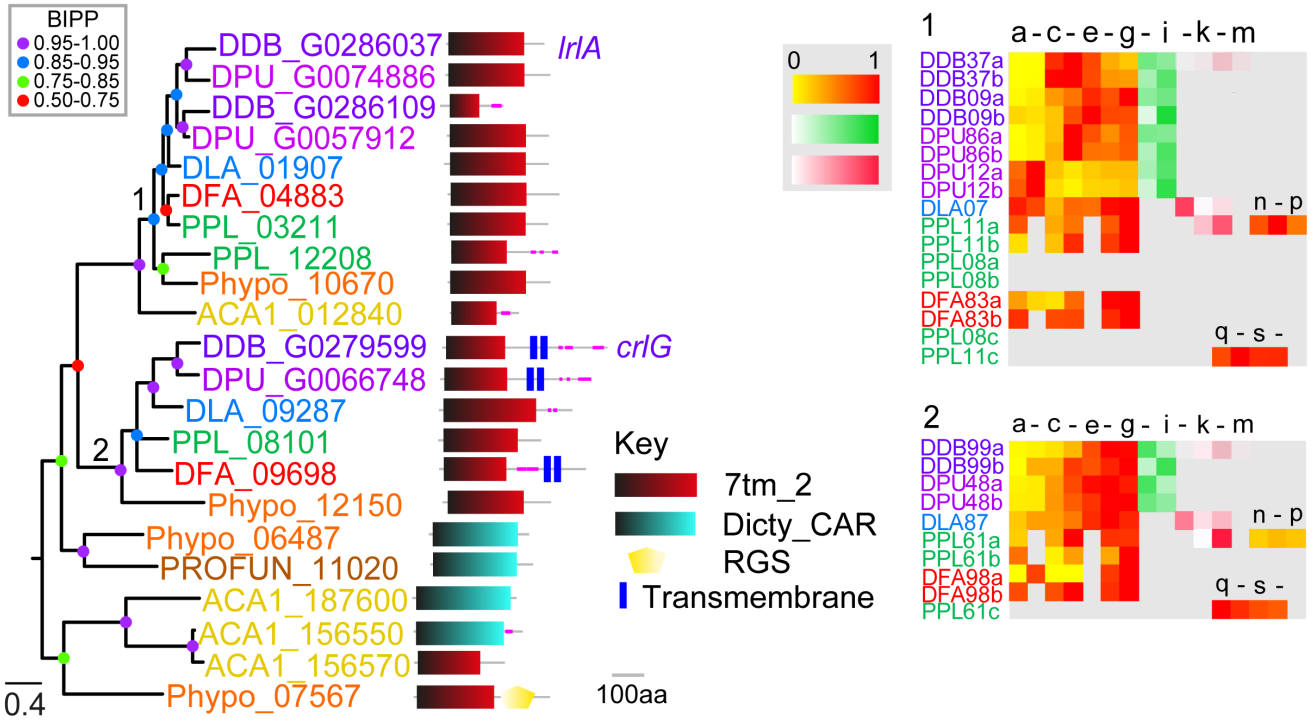
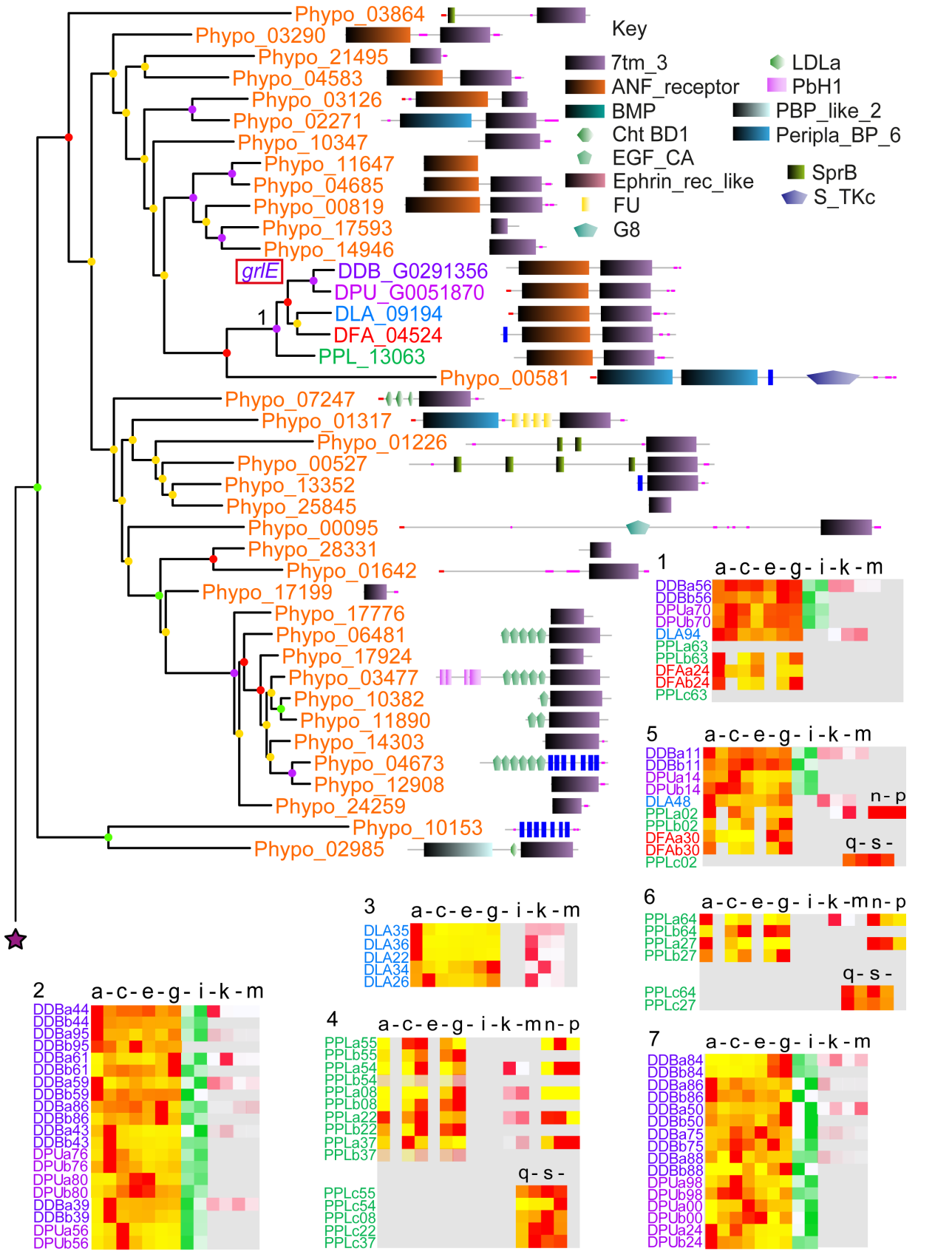


Figure 2. The family 2 secretin-like/Adhesion GPCRs. The family 2 GPCRs were identified by the presence of Interpro domain IPR000832 and by BLAST search and phylogenetic inference. A phylogeny was inferred from the final set of aligned proteins with MrBayes, which was annotated with the protein domain architectures and heatmaps of gene expression profiles as outlined in the legend to Figure 1. Note that clade 2 gene DDB_G0279599 was previously identified as a family 2 Car-like GPCR (Prabhu & Eichinger, 2006), but we found it showed stronger phylogenetic affinity to family 2 GPCRs. Legend and labels as in Figure 1.

domain upon ligand binding leads to activation of the GPCR by back-folding of the liberated agonist-peptide (Vizurraga *et al.*, 2020). A previous analysis revealed only a single *D.dis* secretin-like GPCR: Lathrophilin receptor-likeA (LrlA) (Prabhu & Eichinger, 2006), which was also placed into the adhesion-type family by Schiöth (Schiöth *et al.*, 2010). We found the Pfam GPCR family 2 domain signature (7tm_2) also in a clade comprising CrlG of *D.dis*. CrlG showed greater phylogenetic affinity to family 2 than to family 5 and was therefore included in this family here. Other dictyostelid GPCRs with the 7tm_2 domain showed a closer phylogenetic relationship to families 5 and 6 and were included there. No additional *N*-terminal domains or GAIN domains were detected in the family 2 GPCRs of Amoebozoa. BlastP searches with LrlA revealed the presence of such long N-terminal domains and the GAIN domain in Family 2 GPCRs of *Capsaspora owczarzaki*, a lineage within the opisthokonts closer to the origin of multicellular animals, but not in the latrophilin-like receptor A of *Hondaea fermentalgiana*, a member of the evolutionary distant stramenopiles (Dellero *et al.*, 2018). This suggests that in the lineages leading to multicellular animals an ancestral simple eukaryotic receptor type was modified for new functions in cell-cell and extracellular matrix sensing by the addition of *N*-terminal and autoproteolysis domains. LrlA has undergone a gene duplication in group 4 dictyostelids, which for *D.dis* resulted in a truncated gene (DDB_G0286109) containing only four transmembrane domains (Figure 2). We also detected simple adhesion-type receptors in the genomes of the three investigated solitary Amoebozoa. Nothing is known about the function of those receptors in *Dictyostelium* or other solitary Amoebozoa. Because of the small number of genes in this group, estimations of expression trends are difficult to interpret.

Family 3: Metabotropic glutamate receptors/GABA receptors

The Family 3 of metabotropic glutamate and gamma-amino butyric acid ($GABA_B$) receptors are only similar to other GPCRs on a structural level, but lack significant sequence similarity to them and probably represent an evolutionary separate group (Fredriksson *et al.*, 2003). The family also contains Ca^{2+}-sensing (CaS), pheromone, taste and smell receptors (Yang *et al.*, 2021). Ligand binding takes place in the *N*-terminus which forms a large extracellular domain. Family 3 GPCRs of vertebrates act as constitutive homo- or heterodimers (Yang *et al.*, 2021). In accordance with previous studies (Prabhu & Eichinger, 2006; Prabhu *et al.*, 2007b) we found 17 family 3 GPCRs in *D.dis* (Figure 3). Many of the *D.dis* receptors resulted from gene amplifications specific to *D.dis* or group 4, while independent amplifications also occurred in groups 2 and 3 (Figure 6A) or in groups 3 and 4 only (*grlQ* and *grlR*). Only GrlE, GrlL and GrlP are fully


Phypo_03864
Phypo_03290
Phypo_21495
Phypo_04583
Phypo_03126
Phypo_02271
Phypo_10347
Phypo_11647
Phypo_04685
Phypo_00819
Phypo_17593
Phypo_14946
grlE
DDB_G0291356
DPU_G0051870
DLA_09194
DFA_04524
PPL_13063
Phypo_00581
Phypo_07247
Phypo_01317
Phypo_01226
Phypo_00527
Phypo_13352
Phypo_25845
Phypo_00095
Phypo_28331
Phypo_01642
Phypo_17199
Phypo_17776
Phypo_06481
Phypo_17924
Phypo_03477
Phypo_10382
Phypo_11890
Phypo_14303
Phypo_04673
Phypo_12908
Phypo_24259
Phypo_10153
Phypo_02985
1
Key
7tm_3
ANF_receptor
BMP
Cht BD1
EGF_CA
Ephrin_rec_like
FU
G8
LDLa
PbH1
PBP_like_2
Peripla_BP_6
SprB
S_TKc
a-c-e-g-i-k-m
DDBa56
DDBb56
DPUa70
DPUb70
DLA94
PPLa63
PPLb63
DFAa24
DFAb24
PPLc63
5
DDBa11
DDBb11
DPUa14
DPUb14
DLA48
PPLa02
PPLb02
DFAa30
DFAb30
PPLc02
n-p
q-s-
3
DLA35
DLA36
DLA22
DLA34
DLA26
6
PPLa64
PPLb64
PPLa27
PPLb27
PPLc64
PPLc27
2
DDBa44
DDBb44
DDBa95
DDBb95
DDBa61
DDBb61
DDBa59
DDBb59
DDBa86
DDBb86
DDBa43
DDBb43
DPUa76
DPUb76
DPUa80
DPUb80
DDBa39
DDBb39
DPUa56
DPUb56
4
PPLa55
PPLb55
PPLa54
PPLb54
PPLa08
PPLb08
PPLa22
PPLb22
PPLa37
PPLb37
PPLc55
PPLc54
PPLc08
PPLc22
PPLc37
7
DDBa84
DDBb84
DDBa86
DDBb86
DDBa50
DDBb50
DDBa75
DDBb75
DDBa88
DDBb88
DPUa98
DPUb98
DPUa00
DPUb00
DPUa24
DPUb24

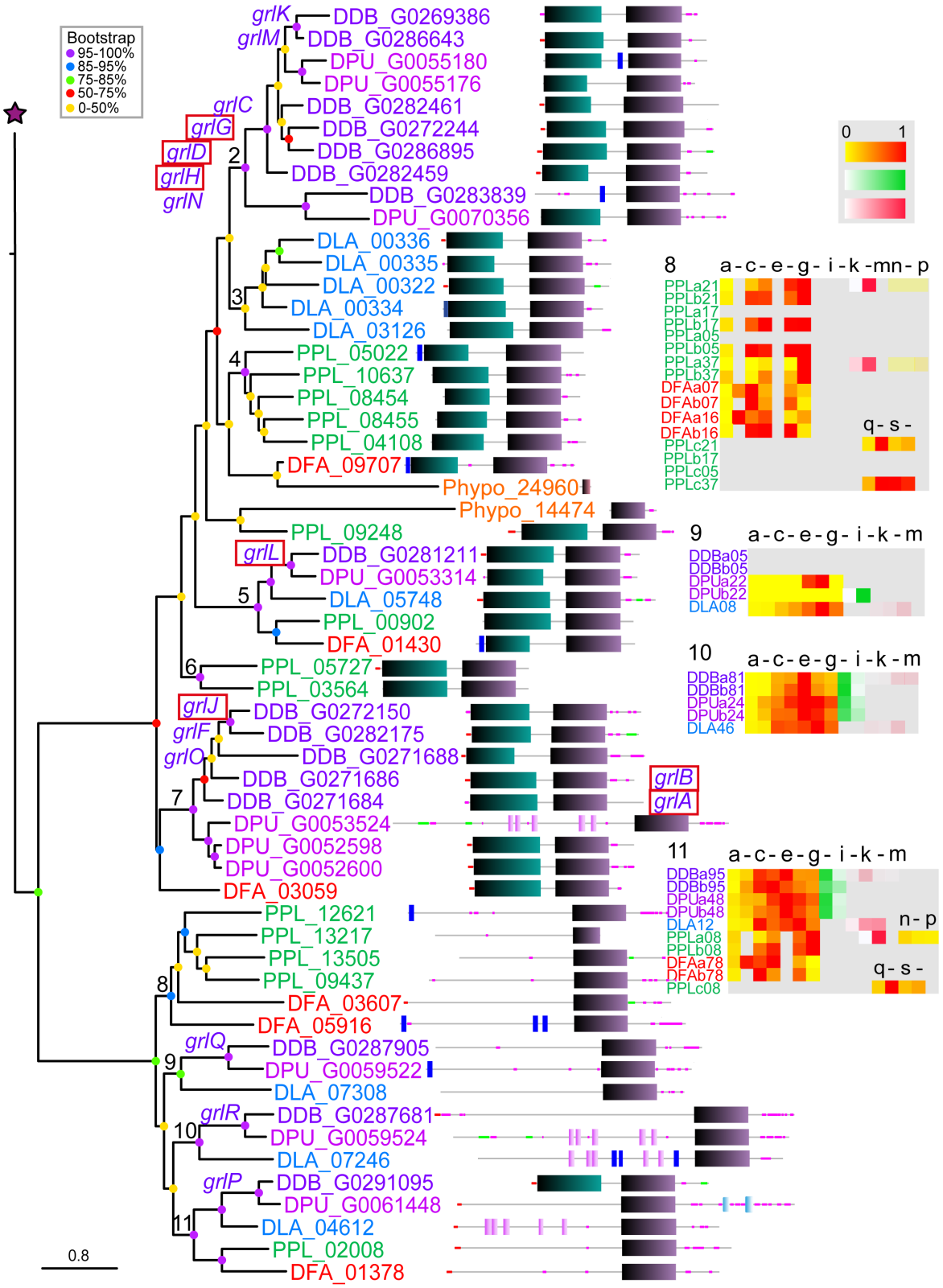
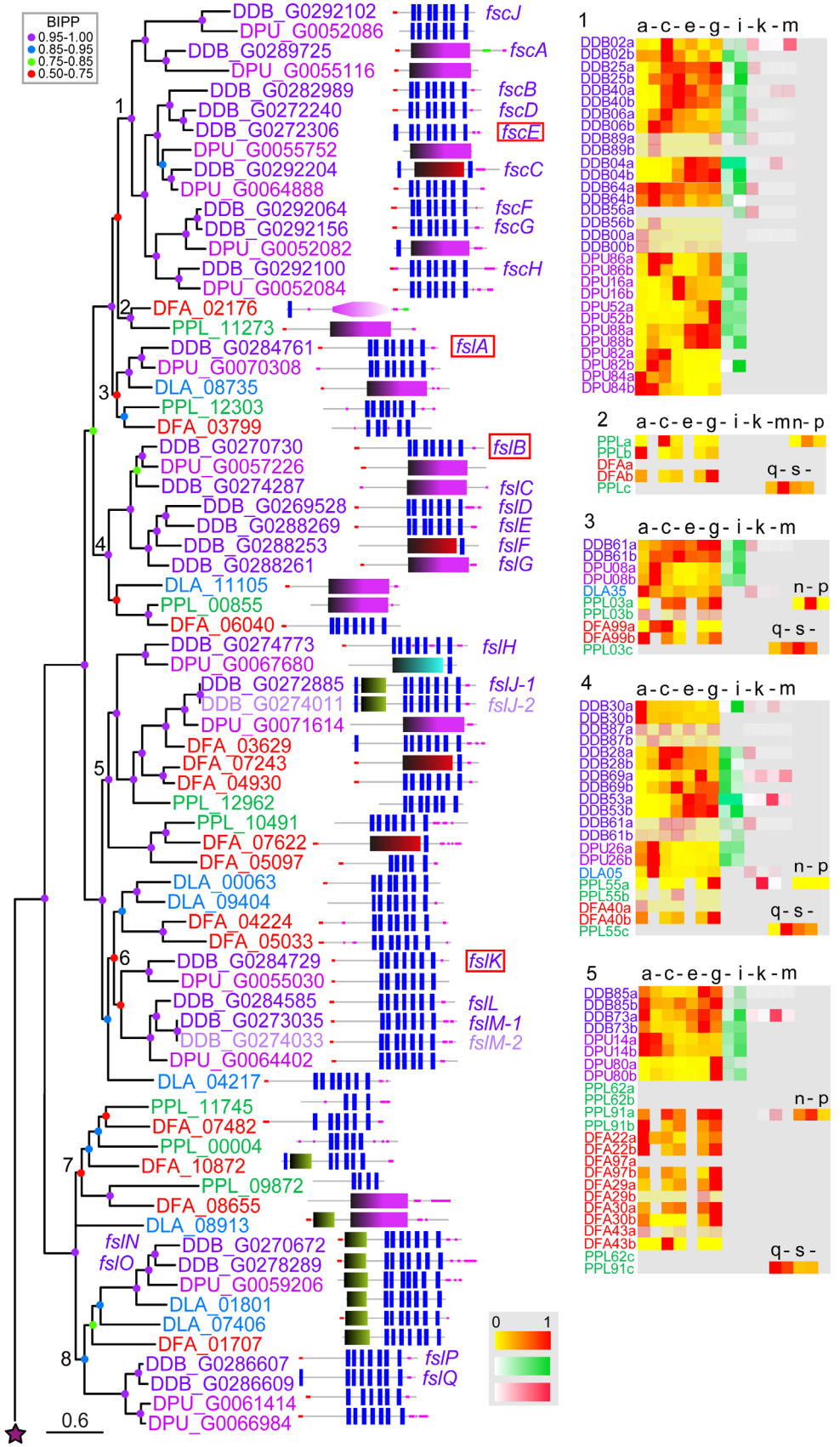


Figure 3. The family 3 metabotropic glutamate-like GPCRs. The family 3 GPCRs were identified by the presence of Interpro domain IPR017978 and by BLASTp search as outlined in the methods. A phylogeny was inferred from the final set of aligned proteins with RAxML, which was annotated with the protein domain architectures and heatmaps of gene expression profiles as outlined in the legend to Figure 1. Bootstrap support of nodes is shown colour coded according to the legend. Apart from clade 1, GrlE, all other family 3 GPCRs form a single cluster of almost exclusively dictyostelid members. Other receptors of this family are only found in *Physarum polycephalum,* but not in other Amoebozoa.


BIPP
0.95-1.00
0.85-0.95
0.75-0.85
0.50-0.75
DDB_G0292102
DPU_G0052086
DDB_G0289725
DPU_G0055116
DDB_G0282989
DDB_G0272240
DDB_G0272306
DPU_G0055752
DDB_G0292204
DPU_G0064888
DDB_G0292064
DDB_G0292156
DPU_G0052082
DDB_G0292100
DPU_G0052084
DFA_02176
PPL_11273
DDB_G0284761
DPU_G0070308
DLA_08735
PPL_12303
DFA_03799
DDB_G0270730
DPU_G0057226
DDB_G0274287
DDB_G0269528
DDB_G0288269
DDB_G0288253
DDB_G0288261
DLA_11105
PPL_00855
DFA_06040
DDB_G0274773
DPU_G0067680
DDB_G0272885
DDB_G0274011
DPU_G0071614
DFA_03629
DFA_07243
DFA_04930
PPL_12962
PPL_10491
DFA_07622
DFA_05097
DLA_00063
DLA_09404
DFA_04224
DFA_05033
DDB_G0284729
DPU_G0055030
DDB_G0284585
DDB_G0273035
DDB_G0274033
DPU_G0064402
DLA_04217
PPL_11745
DFA_07482
PPL_00004
DFA_10872
PPL_09872
DFA_08655
DLA_08913
DDB_G0270672
DDB_G0278289
DPU_G0059206
DLA_01801
DLA_07406
DFA_01707
DDB_G0286607
DDB_G0286609
DPU_G0061414
DPU_G0066984
fscJ
fscA
fscB
fscD
fscE
fscC
fscF
fscG
fscH
fslA
fslB
fslC
fslD
fslE
fslF
fslG
fslH
fslJ-1
fslJ-2
fslK
fslL
fslM-1
fslM-2
fslN
fslO
fslP
fslQ
1
2
3
4
5
6
7
8
0.6
0
1
1
a-c-e-g-i-k-m
2
a-c-e-g-i-k-mn-p
q-s-
3
a-c-e-g-i-k-m
n-p
q-s-
4
a-c-e-g-i-k-m
n-p
q-s-
5
a-c-e-g-i-k-m
n-p
q-s-

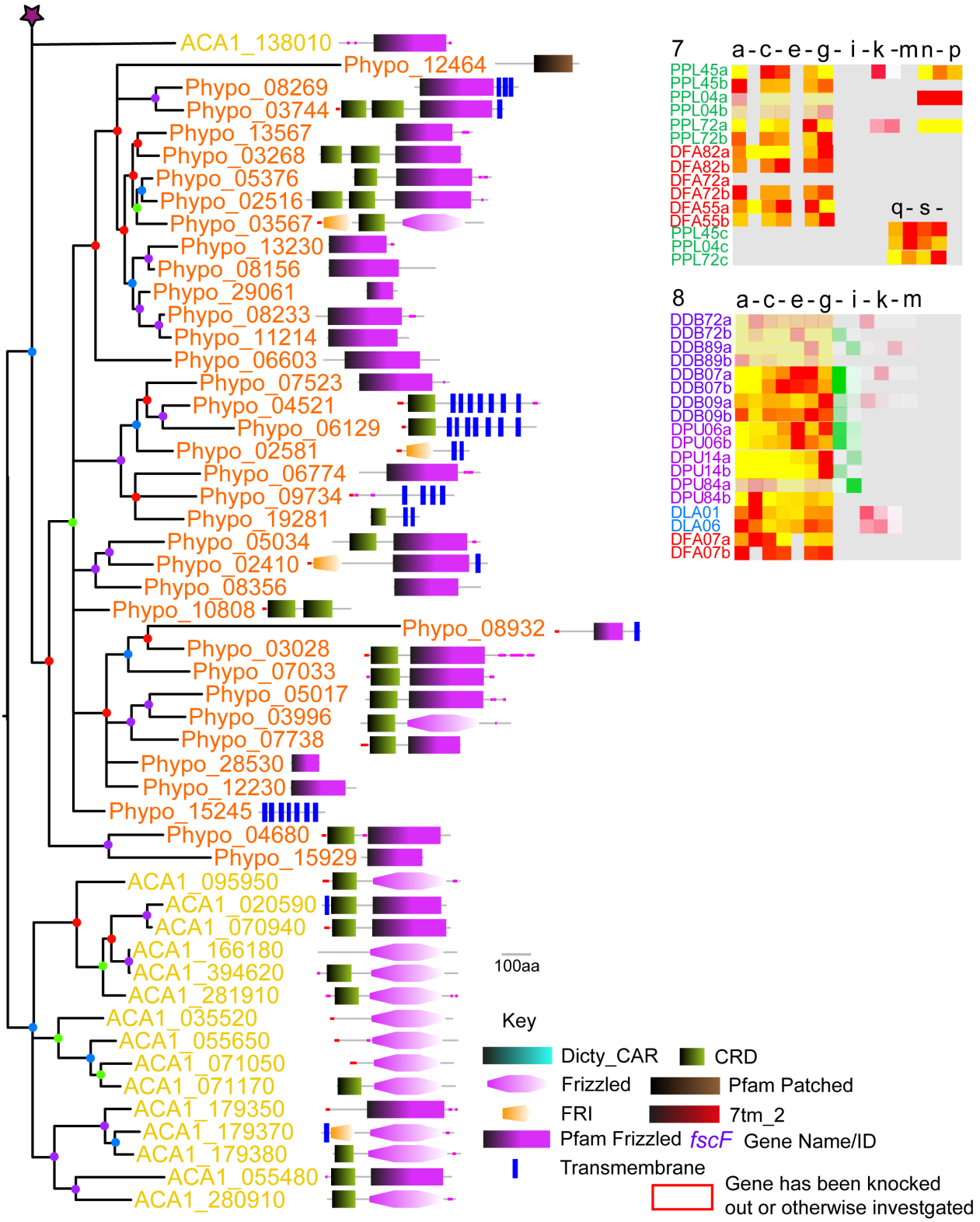
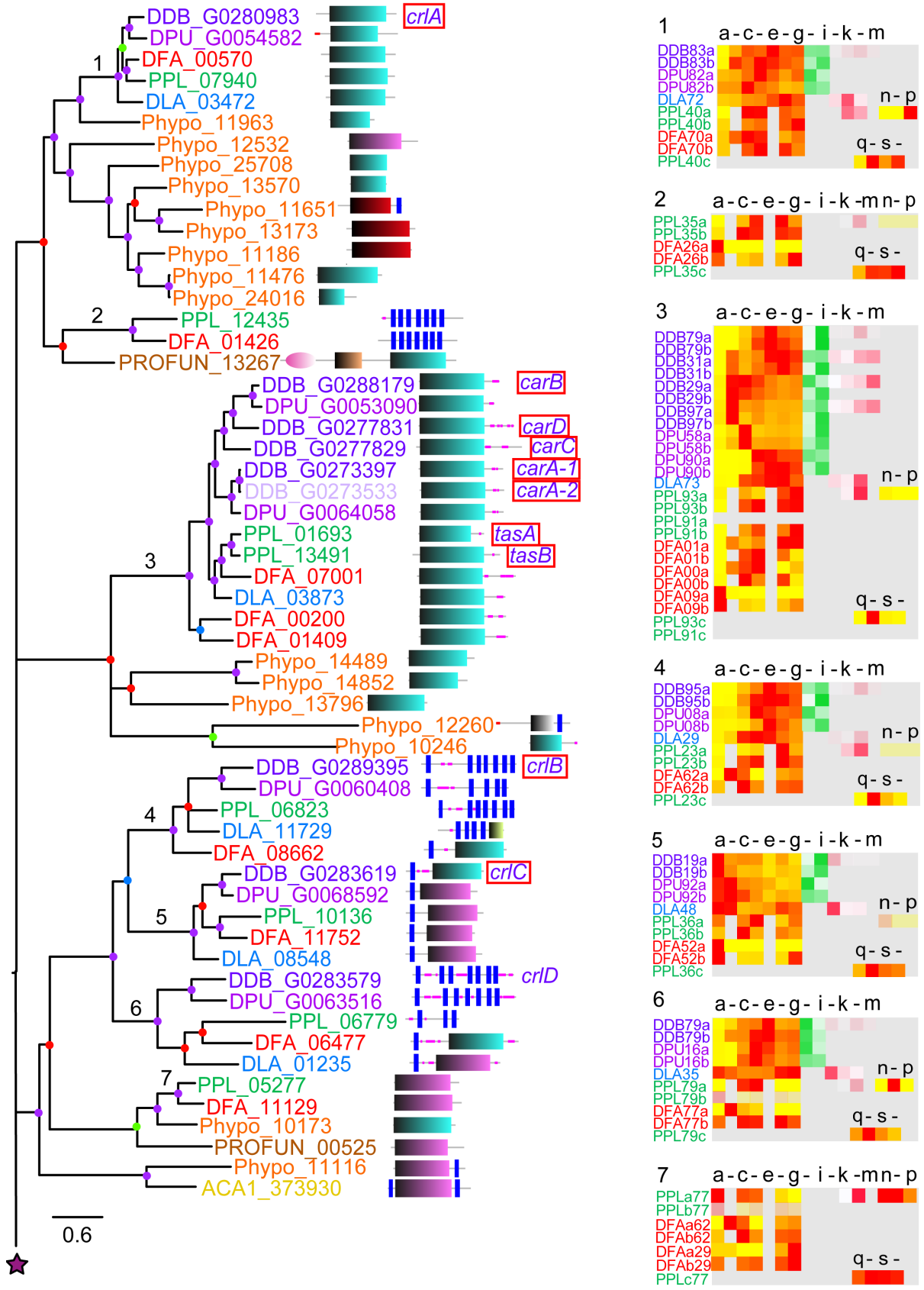


Figure 4. The family 5 frizzled-like GPCRs. The family 5 GPCRs were identified by the presence of Interpro domain IPR000539 and by BLASTp search as outlined in the methods. A phylogeny was inferred from the final set of aligned proteins with MrBayes, which was annotated with the protein domain architectures and heatmaps of gene expression profiles as outlined in the legend to Figure 1. The family 5 experienced independent amplifications in different lines of Amoebozoa and especially within group4 dictyostelids. The dictyostelid members all originate at the base of Dictyostelia and have no Amoebozoan orthologues.


DDB_G0280983
DPU_G0054582
DFA_00570
PPL_07940
DLA_03472
Phypo_11963
Phypo_12532
Phypo_25708
Phypo_13570
Phypo_11651
Phypo_13173
Phypo_11186
Phypo_11476
Phypo_24016
PPL_12435
DFA_01426
PROFUN_13267
DDB_G0288179
DPU_G0053090
DDB_G0277831
DDB_G0277829
DDB_G0273397
DDB_G0273533
DPU_G0064058
PPL_01693
PPL_13491
DFA_07001
DLA_03873
DFA_00200
DFA_01409
Phypo_14489
Phypo_14852
Phypo_13796
Phypo_12260
Phypo_10246
DDB_G0289395
DPU_G0060408
PPL_06823
DLA_11729
DFA_08662
DDB_G0283619
DPU_G0068592
PPL_10136
DFA_11752
DLA_08548
DDB_G0283579
DPU_G0063516
PPL_06779
DFA_06477
DLA_01235
PPL_05277
DFA_11129
Phypo_10173
PROFUN_00525
Phypo_11116
ACA1_373930
crlA
carB
carD
carC
carA-1
carA-2
tasA
tasB
crlB
crlC
crlD
1
2
3
4
5
6
7
0.6
a-c-e-g-i-k-m
n-p
q-s-
DDB83a
DDB83b
DPU82a
DPU82b
DLA72
PPL40a
PPL40b
DFA70a
DFA70b
PPL40c
a-c-e-g-i-k-mn-p
PPL35a
PPL35b
DFA26a
DFA26b
PPL35c
DDB79a
DDB79b
DDB31a
DDB31b
DDB29a
DDB29b
DDB97a
DDB97b
DPU58a
DPU58b
DPU90a
DPU90b
DLA73
PPL93a
PPL93b
PPL91a
PPL91b
DFA01a
DFA01b
DFA00a
DFA00b
DFA09a
DFA09b
PPL93c
PPL91c
DDB95a
DDB95b
DPU08a
DPU08b
DLA29
PPL23a
PPL23b
DFA62a
DFA62b
PPL23c
DDB19a
DDB19b
DPU92a
DPU92b
DLA48
PPL36a
PPL36b
DFA52a
DFA52b
PPL36c
DDB79a
DDB79b
DPU16a
DPU16b
DLA35
PPL79a
PPL79b
DFA77a
DFA77b
PPL79c
a-c-e-g-i-k-mn-p
PPLa77
PPLb77
DFAa62
DFAb62
DFAa29
DFAb29
PPLc77

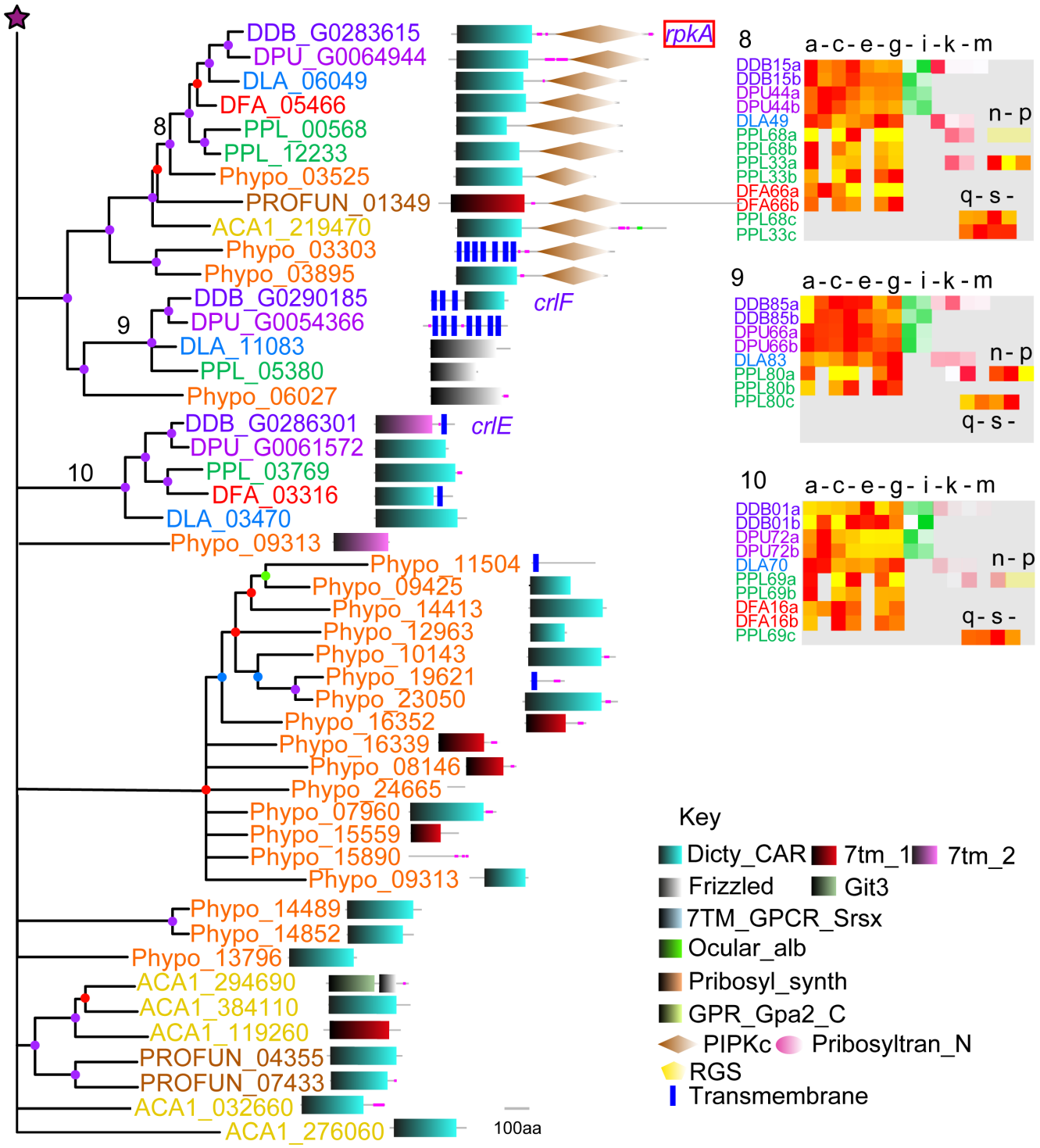


Figure 5. The family 6 cAR-like GPCRs. The family 6 GPCRs were identified by the presence of Interpro domain IPR000848 and by BLASTp search as outlined in the methods. A phylogeny was inferred from the final set of aligned proteins with MrBayes. Even after 3 million generations analysis did not converge and the deeper nodes of the tree are unresolved. Orthologous clades are, however, well defined. The tree was annotated with the protein domain architectures and heatmaps of gene expression profiles as outlined in the legend to Figure 1.

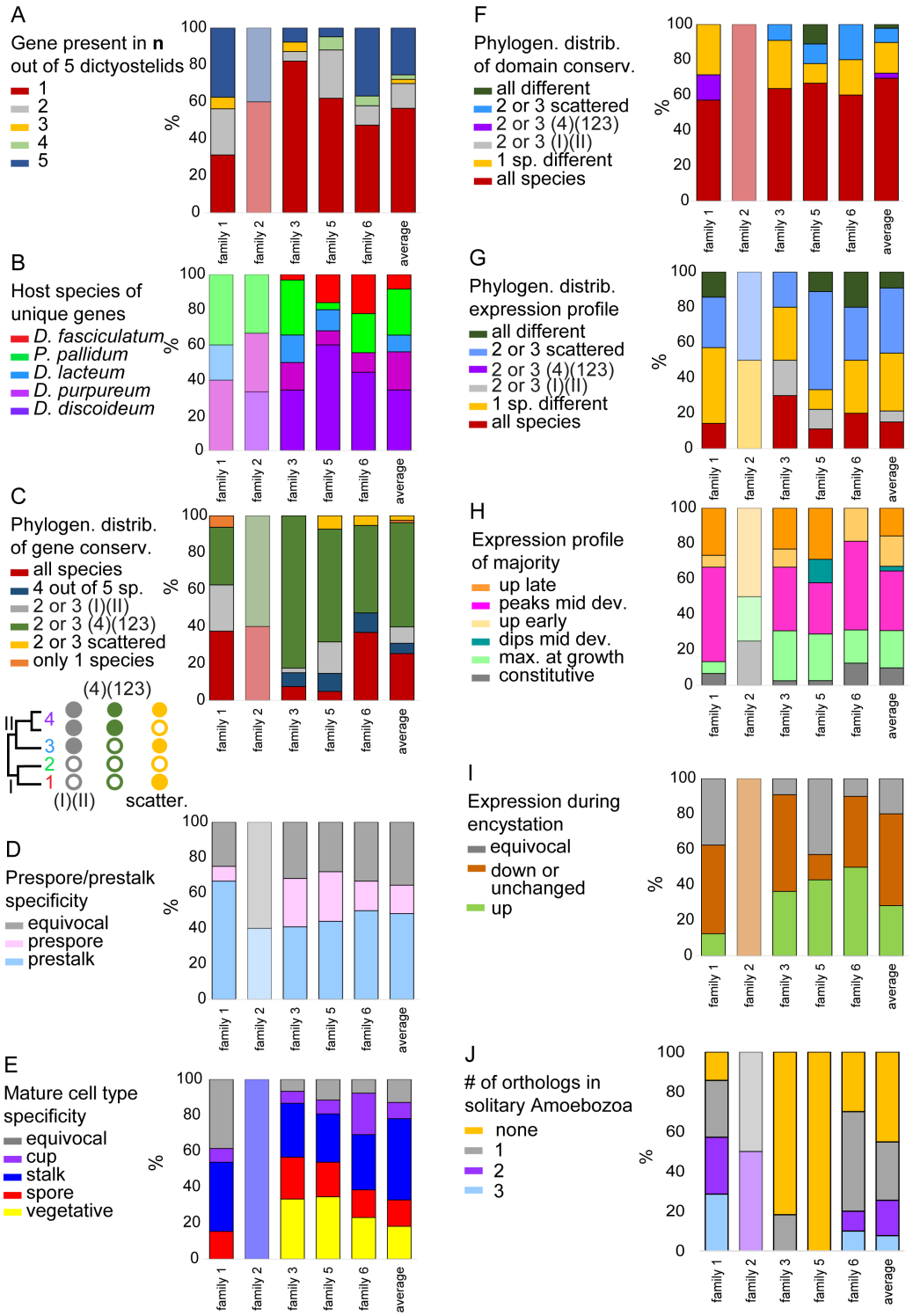


Figure 6. Summary of classed features across GPCR families. For each of the families we summarised and calculated the percentage of the different states of the following features: **A**. the total number of orthologs out of five species that were conserved for each gene. **B**. The host species of genes that were unique. **C**. The phylogenetic distribution of genes. **D**. The pre-cell (only for *D.dis* and *D.pur*) and **E**. mature cell type specificity (not for *D.fas*). The phylogenetic distribution of **F**. domain and **G**. expression profile conservation. **H**. The majority expression pattern during development. **I**. Expression tendency during encystation. (Data only available for *P.pal.*) **J**. Number of solitary Amoebozoan orthologues in clades. The name of each family is shown at the X-axis and the last bar represents the averages across all families. The figure is based on the data shown in Figure 1–Figure 5 that are compiled in supplemental file Compilation_GPCRs.xlsx and summarised in GPCR_traits.xlsx.

Table 2. List of Abbreviations.

7TMR	Seven transmembrane receptor
7tm_1/2/3	Seven transmembrane receptor class 1/2/3
7_TM	Seven transmembrane
7TM_GPCR_Srsx	Serpentine type 7TM GPCR chemoreceptor Srsx
ACA	*Acanthamoeba castellanii* locus tag
ANF_receptor	atrial natriuretic factor receptor
Bac_Rhodopsin	Bacterial rhodopsin
BLAST	Basic Local Alignment Search Tool
BMP	basic membrane lipoprotein
cAMP	Cyclic 3′.5′-adenosine monophosphate
CAR	Cyclic AMP-receptor
CaS	Ca^{2+}-sensing
Cht BD1	type 1 chitin binding domain
CRD	Cysteine-rich domain
CSS-Palm	Clustering and Scoring Strategy algorithm palmitoylation prediction
DDB	*Dictyostelium discoideum* locus tag
DFA	*Dictyostelium fasciculatum* locus tag
Dicty_CAR	Dictyostelium cAMP receptor
DLA	*Dictyostelium lacteum* locus tag
DPU	*Dictyostelium purpureum* locus tag
EGF_CA	Calcium-binding EGF-like domain
Ephrin_rec_like	Putative ephrin-receptor like
FRI	Frizzled
FU	Furin-like repeats
G8	Domain with 8 conserved glycines
GABA	gamma amino butyric acid
GAIN	GPCR auto-proteolysis inducing
Git3	Glucose receptor protein git3
GPR108	G-protein-coupled receptor 108
GPR_Gpa2_C	G-protein-coupled receptor that associates with GPA2 C-terminus
GPCR	G-protein coupled receptor
GpcrRhpsn4	Intimal thickness-related receptor-like, Rhodopsin-like 4
IPR	Interpro Database entry
ITR-like	Intimal-hyperplasia-related gene-like
LCA	last common ancestor
LDLa	Low-density lipoprotein receptor domain class A
Lung_7-TM_R	Mammalian Lung seven transmembrane receptor

HMM	Hidden Markov Model
Ocular_alb	ocular albinism type 1
PbH1	Parallel beta-helix repeats
PBP_like_2	Periplasmic binding protein-like 2
Peripla_BP_6	Periplasmic binding protein 6
PIPKc	Phosphatidylinositol phosphate kinase core
Phypo	*Physarum polycephalum* locus tag
PPL	*Polysphondylium pallidum* locus tag
Pribosyl_synth	phosphoribosyldiphosphate synthetase
Pribosyltran_N	N-terminal domain of ribose phosphate pyrophosphokinase
PROFUN	*Protostelium aurantium* var. *fungivorum* locus tag
RAxML	Randomized Axelerated Maximum Likelihood
RGS	Regulator of G protein signalling domain
SDF	spore differentiation factor
SMART	Simple Modular Architecture Research Tool
SprB	Surface protein B of *Flavobacterium johnsoniae*
S_TKc	Serine/Threonine protein kinases, catalytic domain

conserved throughout the Dictyostelia with single orthologs per clade. The majority of family 3 GPCRs evolved exclusively in the dictyostelids with no homologues in the unicellular Amoebozoa (Figure 3). Interestingly, GrlE has a number of related homologs in *Phy.pol* suggesting that this is the ancestral GABA receptor of the dictyostelids (Figure 3). Similar to other family 3 GPCRs, most Amoebozoan Grl receptors exhibit a large *N*-terminal extracellular domain that shares similarity with either atrial natriuretic factor receptor (ANF; clade 1), basic membrane lipoprotein (BMP. clades 2–7) or has no recognised domains (clades 8–11). ANF is a secreted peptide involved in regulation of arterial blood and electrolyte balance in humans but also has functions in the central nervous system (Stewart *et al.*, 1988). The phylogenetic tree of family 3 shows a significant dichotomy between GPCRs with and without the *N*-terminal ANF domain, with only GrlE belonging to the branch with ANF domains.

Biological roles for some *D.dis* Grls have been reported. GrlE mediates GABA activation of AcbA secretion from prespore cells (Wu & Janetopoulos, 2013). AcbA is then processed by the TagA protease on prestalk cells to yield spore differentiation-factor 2 (SDF-2), which in turn induces maturation of prespore cells into spores (Anjard & Loomis, 2006). This agrees with the observed cell-type specific expression of *grlE* in prespore and spore cells (Figure 3). GrlG (Far2) and GrlL (Far1) both bind folate, but only GrlL was found to mediate folate regulated processes like chemotaxis (Pan *et al.*, 2016). GrlH likely detects the chemorepellant chalone AprA, as *grlH*$^{-}$ cells phenocopy *aprA*$^{-}$ cells (Tang *et al.*, 2018). GrlD is the putative receptor for extracellular polyphosphate, an inhibitor of cell proliferation (Suess *et al.*, 2019). GrlJ is required for the correct timing of development and spore formation (Prabhu *et al.*, 2007b), while *grlB*$^{-}$ cells exhibit delayed aggregation (Wu & Janetopoulos, 2013). GrlA is required for proper late development and sporulation, probably mediating steroid-induced GABA release and subsequent SDF-2 production (Anjard *et al.*, 2009; Prabhu *et al.*, 2007a). Overall, the observed expression patterns follow the reported place of action for the genes with known effects. It is remarkable that many of the Grls with established roles, except GrlE and GrlL (Far1), are the result of gene amplifications that only occurred in group 4, suggesting that these roles uniquely evolved in group 4 or its LCA. The GPCRs of family 3 show varied expression patterns but the highest level of conserved expression across taxon groups (Figure 6G).

Family 5: Frizzled/smoothened 7TMRs

Roles for frizzled/smoothened type receptors were first identified in animal development, where they are involved in tissue morphogenesis and patterning. Frizzled/smoothened receptors are defined by an *N*-terminal ligand binding, cysteine-rich domain (CRD) and a core seven-transmembrane domain which is terminated by the motif KTXXXW. They bind to extracellular ligands like Hedgehog and Wnts and mediate G-protein dependent and -independent signalling. The canonical Wnt/Frizzled signalling pathway acts via beta-catenin but frizzled receptors also function in the non-canonical planar cell polarity and Wnt/Ca^{2+} pathways (Huang & Klein, 2004). In Amoebozoa, frizzled-smoothened

like (Fsl) receptors appear to have independently expanded from one or a few genes in *Phy.pol* and *A.cas*, but were not detected in *Pro.fun.* Dictyostelia also show extensive fsl gene amplification. Strikingly this particularly occurred in group 4, with most proteins found in *D.dis* (Figure 4). Only a few dictyostelid frizzled-like GPCRs in clades 5, 7 and 8 contain the CRD domain indicating the loss would have occurred multiple times in receptor expansion in the dictyostelids, even though only clade 1 of nine *D.dis* proteins is referred to as the Frizzled/Smoothened-like Sans CRD (FscA-J). These proteins also have shorter *N*-termini (Prabhu & Eichinger, 2006). Interestingly, Fscs only have homologues in *D.dis* and *D.pur* and most of the Fsc genes show developmental regulation. Of the remaining 14 *D.dis* Fsl receptors, several also lack a recognisable CRD domain even though longer *N*-terminal regions are present and only the *D.dis* Fsl receptors FslJ and FslK contain the conserved KTXXXW motif (Prabhu & Eichinger, 2006). Most *A.cas* and *Phy.pol* receptors contain CRD domains, with the highest similarity scores to animal frizzled receptors found in *A.cas* (Figure 4). Some dictyostelid Fsl GPCRs showed sequence similarity to the metazoan frizzled class receptors, but their SMART or PFAM frizzled domains were often not detected at E-values below the threshold of 0.001, indicating that they are considerably diverged.

Null mutants in *fscE*, *fslA*, *fslB* and *fslK* show defects in cytokinesis and responses to prestarvation signals, such as AprA (Suess *et al.*, 2017; Tang *et al.*, 2018). In agreement with these early roles *fscE* and *fslB* are downregulated after growth, but *fslA* and *fslK* persist throughout development indicating additional roles there (Figure 4). Like the family 1 rhodopsin-like receptors, the frizzled-like GPCRs (apart from clades 4 and 8) also show preferential expression in prestalk and stalk cells. Many family 5 GPCRs are also upregulated during encystation in *P.pal* and might perform signalling functions there.

Family 6: Dictyostelid cAMP receptors (Dicty_CAR)

The signature cAMP receptor (cAR) of this family was first purified and identified from *D.dis* as the GPCR mediating chemotaxis to cAMP (Klein *et al.*, 1987; Sun & Devreotes, 1991). Later, three lower affinity cAMP receptors and cAR-like receptors (Crl) were identified (Louis *et al.*, 1994; Raisley *et al.*, 2004; Saxe *et al.*, 1991). It is unknown if the latter set is able to bind cAMP and has homologues outside Dictyostelia. The present analysis identifies a large number of Crl proteins in *Phy.pol* and somewhat fewer in *A.cas* and *Pro.fun*. (Figure 5). The sequences that define the Dicty_CAR Pfam model overlap with those of other GPCR families (see *e.g.*, Figure 5, clades 7 and 9), but the overall sequence similarity of the proteins places them clearly within the Dicty_CAR family in phylogenetic analysis. While the deeper nodes of the phylogeny were not fully resolved in our analysis, we identified seven conserved clades that contain a full set of orthologs from dictyostelid species, with half of them also containing *Phy.pol* or *Pro.fun* orthologs or close homologs. Clade 1 contains a set of CrlA receptors including a *Phy.pol*. orthologue and a set of eight related *Phy.pol* proteins. CrlA was reported to detect the polyketide MPBD (4-methyl-5-pentylbenzene-1,3-diol), which stimulates both aggregation competence and release of the peptide SDF-1 (spore differentiation factor-1), which then promotes spore and stalk differentiation (Anjard *et al.*, 2011). However, other workers found that CrlA does not mediate MPBD induction of aggregation competence and is not needed for normal spore differentiation in AX2 wild type cells (Narita *et al.*, 2017). Clade 3 contains the four *D.dis* cAMP receptors (CarA, CarB, CarC and CarD). There are only two *D.pur* orthologues corresponding to CarA and CarB and separate duplication events have taken place in *P.pal* and *D.fas*. Cells in *D.dis* lacking the high affinity receptor CarA fail to aggregate and to express early aggregation genes (Sun *et al.*, 1990). CarB and CarD have lower cAMP affinity and are expressed later in development, with CarB directing tip-oriented migration of prestalk cells (Saxe *et al.*, 1991; Saxe *et al.*, 1993). The set of duplicated cARs in the group 2 species *P.pal* is required for post-aggregative morphogenesis and cAMP induction of prespore differentiation (Kawabe *et al.*, 2009). Clade 3 also contains five *Phy.pol* GPCRs, which form an outgroup to the dictyostelid cAMP receptors. It is not known whether any of these bind cAMP. It would be intriguing to investigate whether at least one organism outside Dictyostelia uses cAMP as a secreted signal, or alternatively what the ancestral signal for the cARs might have been.

The cAMP receptor-like proteins CrlB, CrlC and CrlD are conserved throughout Dictyostelia and together form a well-supported group (clades 4–6). The *D.dis crlB* and *crlD* genes are expressed with a peak in mid-development, however their knockouts had no discernible phenotypes (Raisley *et al.*, 2004). Whereas there were no solitary Amoebozoan orthologues of crlB-D, the larger grouping contains a clade (7) that consists of *P.pal* and *D.fas* (branch I), *Phy.pol* and *Pro.fun crl* genes. In addition, the grouping has a small basal clade of a *Phy.pol* and *A.cas crl* gene supporting an origin of the ancestral Crl of this clade in the Amoebozoa. It is worth noting that within this clade the domain sequences display a mixture of Dicty_CAR and 7TM_2 signatures, with the latter being more common for the secretin/adhesion GPCRs.

Clade 8 is made up of a full orthologous set of dictyostelid and amoebozoan Receptor Phosphatidylinositol Kinase A (*rpkA*) genes, a *P.pal rpkA* duplicate and two additional *Phy.pol rpkA*-like genes (Figure 5). The RpkA*s* consist of an *N*-terminal Dicty_CAR and a C-terminal phosphatidylinositol-4-phosphate 5-kinase (PIP5K) domain. The presence of clearly related homologues in all three solitary Amoebozoan species demonstrates that this domain arrangement was already present in the LCA of all Amoebozoa. In *D.dis,* RpkA is involved in phagocytosis and in cell density sensing (Bakthavatsalam *et al.*, 2006; Riyahi *et al.*, 2011). The related protein CrlF (clade 9) is present in most dictyostelids as well as *Phy.pol* but missing from *D.fas*. High BIPP support for a common origin of clades 8 and 9 points toward a split from an ancient receptor that happened

before the root of the Amoebozoa. The pfam Frizzled/Smoothened domain was recognised at a lower E-value than the Dicty_CAR domain for clade 9, but phylogenetic inference showed strong support for its inclusion in the cAR-like family. *D.dis* CrlE (clade 10) is also well conserved throughout the Dictyostelia even though no homologues for CrlE could be found in other Amoebozoa and the position of the clade remains unresolved. There are additional large groupings of *Phy.pol, A.cas* and *Pro.fun* cAR-like receptors with no dictyostelid homologues and unclear phylogenetic affinities. Overall, the expression patterns of dictyostelid Dicty_CARs varied with a preference for being expressed in prestalk over prespore cells and in stalk cells (Figure 6D, E).

Conclusions

Compared to an earlier analysis (Prabhu & Eichinger, 2006), some GPCRs have been assigned to different families and members of the Rhodopsin-like family have been identified for the first time in dictyostelids in this analysis. Receptor family diversification of original GPCRs is assumed to have taken place in the Amoebozoan ancestor (Nordstrom *et al.*, 2011) and it is possible that those GPCRs have not acquired sufficient family distinctive signatures yet to be classed into the traditional GPCR scheme derived from animals using conventional similarity methods. Previous studies recognised six (Heidel *et al.*, 2011) or three (Nordstrom *et al.*, 2011) additional GPCRs in *D.dis*, that lie outside the classical six families.

The data from all five GPCR families has been quantified and summarised in Figure 6. We observe the highest degree of gene conservation in families 1 and 5 and the lowest in family 3, where gene amplification is seen across all groups (Figure 6A). *D.dis* is overrepresented as host for unique genes indicating an overall higher degree of gene duplications in this species especially in Family 5 genes (Figure 6B). In family 1, *D.pur* appears to be the host for unique genes, but this is due to small numbers of amplified genes. There is a general trend of gene amplification with group 4 species in all families (Figure 6C).

The majority of GPCRs is more highly expressed in prestalk than prespore cells and this is particularly true for family 1 (Figure 6D). Also, expression in stalk cells is higher than in spores for all families (Figure 6E) The GPCRs individually show a range of different developmental expression profiles which is not very different between families. There is preference for peak expression in mid-development and only about ~20% of GPCRs are specifically expressed during growth. This seems to negate a notion that the many GPCRs of Dictyostelia may particularly be required for environmental sensing in the proliferative stage (Figure 6E).

Compared to the ~75% of GPCRs upregulated in multicellular development, less than ~30% are upregulated in encystation, with families 5 and 6 being most commonly upregulated. Combined, these observations suggest that the majority of *Dictyostelium* GPCRs contribute to the organisation of multicellular development.

GPCR families 1, 2 and 6 already diverged early in the Amoebozoa demonstrated by a larger number of orthologues in two or three of the solitary species, whereas families 3 and 5 mostly diverged within the dictyostelids. Most of the functionally analysed *D.dis* members of the Frizzled/Smoothened-like receptors have roles during growth. Our analysis identified homologues of this family which are developmentally upregulated and are worthy of further study. Likewise, the functions of the newly identified family 1 GPCRs are as yet unexplored. It could well be possible that this family harbours the yet unidentified dictyostelid photoreceptor. Whilst initially identified by their similarity to ITR and GPR108 proteins, the dictyostelid Rhodopsin-like GPCRs show GpcrRhopsn4-like and Lung_7-TM_R domains and are similar to metazoan rhodopsins in that they are lacking any other associated functional domains. It is likely that in both cases, ligand binding occurs within the transmembrane domains. However, whilst being derived from a set of ancestral 7TM receptors, the GPCRs in both metazoans and Amoebozoans have evolved independently for a long time, and it is not possible to directly impose a classification mainly based on mammalian GPCRs onto the set we have identified.

Data availability

Underlying data

All data files have been deposited within an OSF registry entitled "G protein coupled receptors of dictyostelids" under DOI 10.17605/OSF.IO/AP4NJ under license CCO 1.0 Universal. The Excel spreadsheet Social_Amoebas_Expression.xlsx is mostly identical to Additional file 2: Table S1 in (Forbes *et al.*, 2019) with addition of the *D.pur* genomic locus tags.

This project contains the following underlying data:

Social_Amoebas_Expression.xlsx (combined RNAseq expression data of social amoebas from several sources)

GPCRs_expression.xlsx (Raw and normalised expression data for GPCRs from social amoebas)

Compilation_GPCRs.xlsx (Family by family enumeration of traits by clades)

GPCRs_traits.xlsx (summarised trait data for all GPCR families from Compilation_GPCRs.xlsx used for generation of summary Figure 6)

GPCRfam1.fas (Fasta file of all GPCR family 1 protein sequences)

GPCRfam1.nex (edited alignment file for all GPCR family 1 sequences in nexus format)

GPCRfam2.txt (Fasta file of all GPCR family 2 protein sequences)

GPCRfam2.nex (edited alignment file for all GPCR family 2 sequences in nexus format)

GPCRfam3.fas (Fasta file of all GPCR family 3 protein sequences)

GPCRfam3.nex (edited alignment file for all GPCR family 3 sequences in nexus format)

GPCRfam5.txt (Fasta file of all GPCR family 5 protein sequences)

GPCRfam5.nex (edited alignment file for all family 5 sequences in nexus format)

GPCRfam6.txt (Fasta file of all GPCR family 6 protein sequences)

GPCRfam6.nex (edited alignment file for all GPCR family 3 sequences in nexus format)

Data are available under the terms of the Creative Commons Zero "No rights reserved" data waiver (CC0 1.0 Public domain dedication).

The raw RNA-Seq data of Dlac cell types and Ppal encystation time series are available from Arrayexpress https://www.ebi.ac.uk/biostudies/arrayexpress/studies/E-MTAB-7824?query=Dictyostelium%20lacteum under accession number E-MTAB-7824. The other RNAseq expression data were obtained from published sources as indicated.

Author contributions

G.H. and S.K. analysed all GPCRs and contributed equally to the work. C.S. supervised and recalculated family trees 3, 5 and 6, curated the data and wrote the original draft. P.S. obtained the funding, identified the GPCRs and conceived the project.

References

Alexander SPH, Christopoulos A, Davenport AP, *et al.*: **THE CONCISE GUIDE TO PHARMACOLOGY 2019/20: G protein-coupled receptors.** *Br J Pharmacol.* 2019; **176 Suppl 1**(Suppl 1): S21–S141.
PubMed Abstract | Publisher Full Text | Free Full Text

Anjard C, Loomis WF: **GABA induces terminal differentiation of *Dictyostelium* through a $GABA_B$ receptor.** *Development.* 2006; **133**(11): 2253–61.
PubMed Abstract | Publisher Full Text

Anjard C, Su Y, Loomis WF: **Steroids initiate a signaling cascade that triggers rapid sporulation in *Dictyostelium*.** *Development.* 2009; **136**(5): 803–12.
PubMed Abstract | Publisher Full Text | Free Full Text

Anjard C, Su Y, Loomis WF: **The polyketide MPBD initiates the SDF-1 signaling cascade that coordinates terminal differentiation in *Dictyostelium*.** *Eukaryot Cell.* 2011; **10**(7): 956–63.
PubMed Abstract | Publisher Full Text | Free Full Text

Bakthavatsalam D, Meijer HJ, Noegel AA, *et al.*: **Novel phosphatidylinositol phosphate kinases with a G-protein coupled receptor signature are shared by *Dictyostelium* and *Phytophthora*.** *Trends Microbiol.* 2006; **14**(9): 378–82.
PubMed Abstract | Publisher Full Text

Bockaert J, Pin JP: **Molecular tinkering of G protein-coupled receptors: an evolutionary success.** *EMBO J.* 1999; **18**(7): 1723–9.
PubMed Abstract | Publisher Full Text | Free Full Text

Clarke M, Lohan AJ, Liu B, *et al.*: **Genome of *Acanthamoeba castellanii* highlights extensive lateral gene transfer and early evolution of tyrosine kinase signaling.** *Genome Biol.* 2013; **14**(2): R11.
PubMed Abstract | Publisher Full Text | Free Full Text

Dellero Y, Cagnac O, Rose S, *et al.*: **Proposal of a new thraustochytrid genus *Hondaea* gen. nov. and comparison of its lipid dynamics with the closely related pseudo-cryptic genus *Aurantiochytrium*.** *Algal Res.* 2018; **35**: 125–141.
Publisher Full Text

Dudek AM, Zabaleta N, Zinn E, *et al.*: **GPR108 Is a Highly Conserved AAV Entry Factor.** *Mol Ther.* 2020; **28**(2): 367–381.
PubMed Abstract | Publisher Full Text | Free Full Text

Edgar AJ: **Human GPR107 and murine Gpr108 are members of the LUSTR family of proteins found in both plants and animals, having similar topology to G-protein coupled receptors.** *DNA Seq.* 2007; **18**(3): 235–41.
PubMed Abstract | Publisher Full Text

Eichinger L, Pachebat JA, Glöckner G, *et al.*: **The genome of the social amoeba *Dictyostelium discoideum*.** *Nature.* 2005; **435**(7038): 43–57.
PubMed Abstract | Publisher Full Text | Free Full Text

Foord SM, Bonner TI, Neubig RR, *et al.*: **International Union of Pharmacology. XLVI. G Protein-Coupled Receptor List.** *Pharmacol Rev.* 2005; **57**(2): 279–88.
PubMed Abstract | Publisher Full Text

Forbes G, Chen ZH, Kin K, *et al.*: **Phylogeny-wide conservation and change in developmental expression, cell-type specificity and functional domains of the transcriptional regulators of social amoebas.** *BMC Genomics.* 2019; **20**(1): 890.
PubMed Abstract | Publisher Full Text | Free Full Text

Fredriksson R, Lagerström MC, Lundin LG, *et al.*: **The G-Protein-Coupled Receptors in the Human Genome Form Five Main Families. Phylogenetic Analysis, Paralogon Groups, and Fingerprints.** *Mol Pharmacol.* 2003; **63**(6): 1256–72.
PubMed Abstract | Publisher Full Text

Glöckner G, Lawal HM, Felder M, *et al.*: **The multicellularity genes of dictyostelid social amoebas.** *Nat Commun.* 2016; **7**: 12085.
PubMed Abstract | Publisher Full Text | Free Full Text

Hall TA: **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.** *Nucl Acids Symp Ser.* 1999; **41**: 95–98.
Reference Source

Heidel AJ, Lawal HM, Felder M, *et al.*: **Phylogeny-wide analysis of social amoeba genomes highlights ancient origins for complex intercellular communication.** *Genome Res.* 2011; **21**(11): 1882–91.
PubMed Abstract | Publisher Full Text | Free Full Text

Hillmann F, Forbes G, Novohradská S, *et al.*: **Multiple Roots of Fruiting Body Formation in Amoebozoa.** *Genome Biol Evol.* 2018; **10**(2): 591–606.
PubMed Abstract | Publisher Full Text | Free Full Text

Hu GM, Mai TL, Chen CM: **Visualizing the GPCR Network: Classification and Evolution.** *Sci Rep.* 2017; **7**(1): 15495.
PubMed Abstract | Publisher Full Text | Free Full Text

Huang HC, Klein PS: **The Frizzled family: receptors for multiple signal transduction pathways.** *Genome Biol.* 2004; **5**(7): 234.
PubMed Abstract | Publisher Full Text | Free Full Text

Insel PA, Sriram K, Gorr MW, *et al.*: **GPCRomics: An Approach to Discover GPCR Drug Targets.** *Trends Pharmacol Sci.* 2019; **40**(6): 378–387.
PubMed Abstract | Publisher Full Text | Free Full Text

Jones P, Binns D, Chang HY, *et al.*: **InterProScan 5: genome-scale protein function classification.** *Bioinformatics.* 2014; **30**(9): 1236–40.
PubMed Abstract | Publisher Full Text | Free Full Text

Joost P, Methner A: **Phylogenetic analysis of 277 human G-protein-coupled receptors as a tool for the prediction of orphan receptor ligands.** *Genome Biol.* 2002; **3**(11): RESEARCH0063.
PubMed Abstract | Publisher Full Text | Free Full Text

Kawabe Y, Morio T, James JL, *et al.*: **Activated cAMP receptors switch encystation into sporulation.** *Proc Natl Acad Sci U S A.* 2009; **106**(17): 7089–7094.
PubMed Abstract | Publisher Full Text | Free Full Text

Kin K, Forbes G, Cassidy A, *et al.*: **Cell-type specific RNA-Seq reveals novel roles and regulatory programs for terminally differentiated *Dictyostelium* cells.** *BMC Genomics.* 2018; **19**(1): 764.
PubMed Abstract | Publisher Full Text | Free Full Text

Klein P, Vaughan R, Borleis J, *et al.*: **The surface cyclic AMP receptor in**

***Dictyostelium*. Levels of ligand-induced phosphorylation, solubilization, identification of primary transcript, and developmental regulation of expression.** *J Biol Chem.* 1987; **262**(1): 358–64.
PubMed Abstract | Publisher Full Text

Lin HY, Harris TL, Flannery MS, *et al.*: **Expression cloning of an adenylate cyclase-coupled calcitonin receptor.** *Science.* 1991; **254**(5034): 1022–4.
PubMed Abstract | Publisher Full Text

Louis JM, Ginsburg GT, Kimmel AR: **The cAMP receptor CAR4 regulates axial patterning and cellular differentiation during late development of Dictyostelium.** *Genes Dev.* 1994; **8**(17): 2086–96.
PubMed Abstract | Publisher Full Text

Lyu S, Zhang X, Tu Z, *et al.*: **GPR108 is required for gambogic acid inhibiting NF-kappaB signaling in cancer.** *Pharmacol Res.* 2022; **182**: 106279.
PubMed Abstract | Publisher Full Text

McCudden CR, Hains MD, Kimple RJ, *et al.*: **G-protein signaling: back to the future.** *Cell Mol Life Sci.* 2005; **62**(5): 551–77.
PubMed Abstract | Publisher Full Text | Free Full Text

Narita TB, Schaap P, Saito T: **Effects of deletion of the receptor CrlA on *Dictyostelium* aggregation and MPBD-mediated responses are strain dependent and not evident in strain Ax2.** *FEMS Microbiol Lett.* 2017; **364**(4).
PubMed Abstract | Publisher Full Text

Nordstrom KJV, Sällman Almén M, Edstam MM, *et al.*: **Independent HHsearch, Needleman--Wunsch-based, and motif analyses reveal the overall hierarchy for most of the G protein-coupled receptor families.** *Mol Biol Evol.* 2011; **28**(9): 2471–80.
PubMed Abstract | Publisher Full Text

Pan M, Xu X, Chen Y, *et al.*: **Identification of a Chemoattractant G-Protein-Coupled Receptor for Folic Acid that Controls Both Chemotaxis and Phagocytosis.** *Dev Cell.* 2016; **36**(4): 428–39.
PubMed Abstract | Publisher Full Text | Free Full Text

Parikh A, Miranda ER, Katoh-Kurasawa M, *et al.*: **Conserved developmental transcriptomes in evolutionarily divergent species.** *Genome Biol.* 2010; **11**(3): R35.
PubMed Abstract | Publisher Full Text | Free Full Text

Prabhu Y, Eichinger L: **The *Dictyostelium* repertoire of seven transmembrane domain receptors.** *Eur J Cell Biol.* 2006; **85**(9–10): 937–46.
PubMed Abstract | Publisher Full Text

Prabhu Y, Mondal S, Eichinger L, *et al.*: **A GPCR involved in post aggregation events in *Dictyostelium discoideum*.** *Dev Biol.* 2007a; **312**(1): 29–43.
PubMed Abstract | Publisher Full Text

Prabhu Y, Müller R, Anjard C, *et al.*: **GrlJ, a *Dictyostelium* $GABA_B$-like receptor with roles in post-aggregation development.** *BMC Dev Biol.* 2007b; **7**: 44.
PubMed Abstract | Publisher Full Text | Free Full Text

Raisley B, Zhang M, Hereld D, *et al.*: **A cAMP receptor-like G protein-coupled receptor with roles in growth regulation and development.** *Dev Biol.* 2004; **265**(2): 433–45.
PubMed Abstract | Publisher Full Text

Ren J, Wen L, Gao X, *et al.*: **CSS-Palm 2.0: an updated software for palmitoylation sites prediction.** *Protein Eng Des Sel.* 2008; **21**(11): 639–44.
PubMed Abstract | Publisher Full Text | Free Full Text

Riyahi TY, Frese F, Steinert M, *et al.*: **RpkA, a highly conserved GPCR with a lipid kinase domain, has a role in phagocytosis and anti-bacterial defense.** *PLoS One.* 2011; **6**(11): e27311.
PubMed Abstract | Publisher Full Text | Free Full Text

Romeralo M, Skiba A, Gonzalez-Voyer A, *et al.*: **Analysis of phenotypic evolution in Dictyostelia highlights developmental plasticity as a likely consequence of colonial multicellularity.** *Proc Biol Sci.* 2013; **280**(1764): 20130976.
PubMed Abstract | Publisher Full Text | Free Full Text

Ronquist F, Teslenko M, van der Mark P, *et al.*: **MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space.** *Syst Biol.* 2012; **61**(3): 539–542.
PubMed Abstract | Publisher Full Text | Free Full Text

Saxe CL 3rd, Ginsburg GT, Louis JM, *et al.*: **CAR2, a prestalk cAMP receptor required for normal tip formation and late development of Dictyostelium discoideum.** *Genes Dev.* 1993; **7**(2): 262–72.
PubMed Abstract | Publisher Full Text

Saxe CL 3rd, Johnson R, Devreotes PN, *et al.*: **Multiple genes for cell surface cAMP receptors in *Dictyostelium discoideum*.** *Dev Genet.* 1991; **12**(1–2): 6–13.
PubMed Abstract | Publisher Full Text

Schaap P, Barrantes I, Minx P, *et al.*: **The *Physarum polycephalum* Genome Reveals Extensive Use of Prokaryotic Two-Component and Metazoan-Type Tyrosine Kinase Signaling.** *Genome Biol Evol.* 2015; **8**(1): 109–125.
PubMed Abstract | Publisher Full Text | Free Full Text

Schiöth HB, Nordström KJV, Fredriksson R: **The *Adhesion* GPCRs; Gene Repertoire, Phylogeny and Evolution**. In: Yona, S. and Stacey, M., eds., *Adhesion-GPCRs: Structure to Function*. Boston, MA: Springer US, 2010; 1–13.
Publisher Full Text

Schultz J, Milpetz F, Bork P, *et al.*: **SMART, a simple modular architecture research tool: identification of signaling domains.** *Proc Natl Acad Sci U S A.* 1998; **95**(11): 5857–64.
PubMed Abstract | Publisher Full Text | Free Full Text

Sheikh S, Gloeckner G, Kuwayama H, *et al.*: **Root of Dictyostelia based on 213 universal proteins.** *Mol Phylogenet Evol.* 2015; **92**: 53–62.
PubMed Abstract | Publisher Full Text

Sheikh S, Thulin M, Cavender JC, *et al.*: **A New Classification of the Dictyostelids.** *Protist.* 2018; **169**(1): 1–28.
PubMed Abstract | Publisher Full Text

Sievers F, Wilm A, Dineen D, *et al.*: **Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega.** *Mol Syst Biol.* 2011; **7**: 539.
PubMed Abstract | Publisher Full Text | Free Full Text

Stamatakis A: **RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies.** *Bioinformatics.* 2014; **30**(9): 1312–3.
PubMed Abstract | Publisher Full Text | Free Full Text

Stewart RE, Swithers SE, Plunkett LM, *et al.*: **ANF receptors: distribution and regulation in central and peripheral tissues.** *Neurosci Biobehav Rev.* 1988; **12**(2): 151–68.
PubMed Abstract | Publisher Full Text

Sucgang R, Kuo A, Tian X, *et al.*: **Comparative genomics of the social amoebae *Dictyostelium discoideum* and *Dictyostelium purpureum*.** *Genome Biol.* 2011; **12**(2): R20.
PubMed Abstract | Publisher Full Text | Free Full Text

Suess PM, Tang Y, Gomer RH: **The putative G protein-coupled receptor GrlD mediates extracellular polyphosphate sensing in *Dictyostelium discoideum*.** *Mol Biol Cell.* 2019; **30**(9): 1118–1128.
PubMed Abstract | Publisher Full Text | Free Full Text

Suess PM, Watson J, Chen W, *et al.*: **Extracellular polyphosphate signals through Ras and Akt to prime *Dictyostelium discoideum* cells for development.** *J Cell Sci.* 2017; **130**(14): 2394–2404.
PubMed Abstract | Publisher Full Text | Free Full Text

Sun TJ, Devreotes PN: **Gene targeting of the aggregation stage cAMP receptor cAR1 in Dictyostelium.** *Genes Dev.* 1991; **5**(4): 572–82.
PubMed Abstract | Publisher Full Text

Sun TJ, Van Haastert PJ, Devreotes PN: **Surface cAMP receptors mediate multiple responses during development in Dictyostelium: evidenced by antisense mutagenesis.** *J Cell Biol.* 1990; **110**(5): 1549–54.
PubMed Abstract | Publisher Full Text | Free Full Text

Tafesse FG, Guimaraes CP, Maruyama T, *et al.*: **GPR107, a G-protein-coupled receptor essential for intoxication by *Pseudomonas aeruginosa* exotoxin A, localizes to the Golgi and is cleaved by furin.** *J Biol Chem.* 2014; **289**(35): 24005–18.
PubMed Abstract | Publisher Full Text | Free Full Text

Takei Y: **Chapter 18 - Secretin (Pituitary Adenylate Cyclase-Activating Polypeptide) Family**. In: Takei, Y., Ando, H. and Tsutsui, K., eds., *Handbook of Hormones*. San Diego: Academic Press, 2016; 140–141, e18-1–e18-2.
Publisher Full Text

Tang Y, Wu Y, Herlihy SE, *et al.*: **An Autocrine Proliferation Repressor Regulates *Dictyostelium discoideum* Proliferation and Chemorepulsion Using the G Protein-Coupled Receptor GrlH.** *mBio.* 2018; **9**(1): e02443–17.
PubMed Abstract | Publisher Full Text | Free Full Text

Tsukada S, Iwai M, Nishiu J, *et al.*: **Inhibition of experimental intimal thickening in mice lacking a novel G-protein-coupled receptor.** *Circulation.* 2003; **107**(2): 313–9.
PubMed Abstract | Publisher Full Text

Vizurraga A, Adhikari R, Yeung J, *et al.*: **Mechanisms of adhesion G protein-coupled receptor activation.** *J Biol Chem.* 2020; **295**(41): 14065–14083.
PubMed Abstract | Publisher Full Text | Free Full Text

Waterhouse AM, Procter JB, Martin DM, *et al.*: **Jalview Version 2--a multiple sequence alignment editor and analysis workbench.** *Bioinformatics.* 2009; **25**(9): 1189–91.
PubMed Abstract | Publisher Full Text | Free Full Text

Wu Y, Janetopoulos C: **Systematic analysis of γ-aminobutyric acid (GABA) metabolism and function in the social amoeba *Dictyostelium discoideum*.** *J Biol Chem.* 2013; **288**(21): 15280–90.
PubMed Abstract | Publisher Full Text | Free Full Text

Yang D, Zhou Q, Labroska V, *et al.*: **G protein-coupled receptors: structure- and function-based drug discovery.** *Signal Transduct Target Ther.* 2021; **6**(1): 7.
PubMed Abstract | Publisher Full Text | Free Full Text

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 2

Reviewer Report 16 February 2023

https://doi.org/10.21956/openreseurope.16865.r30850

© **2023 Hadwiger J.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Jeffrey A. Hadwiger**
Department of Microbiology and Molecular Genetics, Oklahoma State University, Stillwater, OK, USA

The authors have addressed all of my comments.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Signal transduction, chemotaxis, development, amoeba, G proteins, receptors, MAP kinases, and phosphodiesterases

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 15 February 2023

https://doi.org/10.21956/openreseurope.16865.r30851

© **2023 Xu X.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The author(s) is/are employees of the US Government and therefore domestic copyright protection in USA does not apply to this work. The work may be protected under the copyright laws of other jurisdictions when used in those jurisdictions.

✔ **Xuehua Xu**
Chemotaxis Signaling Section, Laboratory of Immunogenetics, National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIAID/NIH), Rockville, MD, USA

The manuscript is suitable for indexing.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: GPCR functions in diverse processes

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 03 February 2023

https://doi.org/10.21956/openreseurope.16492.r30522

© 2023 Xu X. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The author(s) is/are employees of the US Government and therefore domestic copyright protection in USA does not apply to this work. The work may be protected under the copyright laws of other jurisdictions when used in those jurisdictions.

?

Xuehua Xu
[1] Chemotaxis Signaling Section, Laboratory of Immunogenetics, National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIAID/NIH), Rockville, MD, USA
[2] Chemotaxis Signaling Section, Laboratory of Immunogenetics, National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIAID/NIH), Rockville, MD, USA

Authors analyzed phylogenetic distribution and domain composition, and their temporal/spatial expression of GPCRs in five species of dictyostelids. The work is informative and will provide essential information for future study.

Minor points:

1. It would be helpful for readers to briefly mention the known ligands and/or functions of GPCRs, such as GPR108 or IPR-like receptor when they were first mentioned the manuscript.

2. It would be very helpful to discuss a little more on the similarity and differences of the rhodopsin-like receptors between dictyostelids and their bacterial or animal rhodopsins.

3. Labels for figures are generally too small to read, especially those in the heat map of expression. Figure 2 is good.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and does the work have academic merit?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: GPCR functions in diverse processes

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 08 Mar 2023

Christina Schilde

Dear Reviewer 2, thank you for taking the time to review our article and for pointing out some omissions and errors. Your critical reading helped to improve the article. Minor points:

1. **"It would be helpful for readers to briefly mention the known ligands and/or functions of GPCRs, such as GPR108 or IPR-like receptor when they were first mentioned the manuscript."**

   We have added information about the function of GPR108 and IPR.

2. **"It would be very helpful to discuss a little more on the similarity and differences of the rhodopsin-like receptors between dictyostelids and their bacterial or animal rhodopsins."**

   We have added a brief discussion of the similarity to animal and bacterial rhodopsins. We have also inserted an additional section into the conclusions.

3. **"Labels for figures are generally too small to read, especially those in the heat map of expression. Figure 2 is good."**

   We appreciate the labels are too small and have tried to replace them by letter labels where only every other character is shown in order to improve readability because this was also an issue the other reviewer had raised. We hope that this has improved the figures.

Competing Interests: No competing interests were disclosed.

Reviewer Report 13 January 2023

https://doi.org/10.21956/openreseurope.16492.r30523

© 2023 Hadwiger J. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Jeffrey A. Hadwiger
[1] Department of Microbiology and Molecular Genetics, Oklahoma State University, Stillwater, OK, USA
[2] Department of Microbiology and Molecular Genetics, Oklahoma State University, Stillwater, OK, USA

The authors have conducted an extensive comparison of GPCRs in amoeba and produced family specific phylogenetic trees for each of the major GPCR families. The designation of new GPCRs and phylogenetic relationship of GPCRs should prove insightful for those studying GPCRs and amoeba. The analysis is based on existing genomic and expression data from several amoeba species including developmental and cell-type specific expression. Overall this report is well written and easy to understand but there are some clarifications that could be useful to the readers (see comments below).

- Several abbreviations are not defined.

For example GPR and IPR (InterPro terminology?).
A supplementary table with the full names of the abbreviated receptors and other designations. in the figure keys would help with understanding what is being referenced.
It is unclear what the c.0-c.24 refers to – counts? More explanation needed.

- The heat map labels on the X-axis are difficult to read – not sure of a remedy other than zooming in.

- The expression data in general appears to be weak and inconsistent in some cases between duplicate experiments. These inconsistencies are due to the original data source but the authors might caution the readers about the validity of this information.

- In the first paragraph describing family 6, the sentence “The latter set [referring to Crl receptors] did not bind cAMP and . . . “ The ability to bind cAMP was never tested for these receptors based on the recollection of this reviewer.

- The BIPP color coding key does not include yellow junctions. The significance of the yellow junctions should be noted.

- Fig. 4 appears to be cut off on the bottom

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and does the work have academic merit?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Signal transduction, chemotaxis, development, amoeba, G proteins, receptors, MAP kinases, and phosphodiesterases

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 06 Feb 2023

Christina Schilde

Dear reviewer 1, Thank you for taking the time to review our article and for pointing out some omissions and errors. Your critical reading helped to improve the article.

- ***"Several abbreviations are not defined."***

Thank you for the suggestion. We have inserted a table of abbreviations for all domains and receptors at the end. The names of the human genes GPR108 (G-protein-coupled receptor 108) and ITR-like (Intimal-hyperplasia-related gene-like) are derived from the HUGO (human gene nomenclature committee) terminology. IPR numbers are indeed Interpro domain identifiers. The c.0-c.48 are reads from encystation time courses of PPL and were explained in the figure legend to Figure 1. We have replaced those labels with small letters and have explained them in the figure legend to Figure 1.

- ***"The heat map labels on the X-axis are difficult to read – not sure of a remedy other than zooming in."***

We appreciate the labels are too small and have tried to replace them by letter labels where only every other character is shown in order to improve readability because this was also an issue the other reviewer had raised. We hope that this has improved the figures.

- ***"The expression data in general appears to be weak and inconsistent in some cases between duplicate experiments. These inconsistencies are due to the original data source but the authors might caution the readers about the validity of this information."***

Expression data with extremely weak read counts (<10) have already been indicated shaded in 30% opacity. Of course, we could raise the threshold, but we also sometimes observe problems between the duplicate datasets even with higher read counts. We have inserted a sentence at the end of the methods section cautioning the readers about incongruent duplicate datasets.

- ***"In the first paragraph describing family 6, the sentence "The latter set [referring to Crl receptors] did not bind cAMP and . . . " The ability to bind cAMP was never tested for these receptors based on the recollection of this reviewer."***

This was an overinterpretation of the fact that all Crl strains aggregate normally. We have reworded the sentence.

- ***"The BIPP color coding key does not include yellow junctions. The significance of the yellow junctions should be noted."***

Thanks for spotting this oversight. Yellow junctions only appear in Figure 3 as this was the only tree that could not be resolved with Bayesian phylogeny and hence was produced with RAxML. We have replaced the inset legend with a corrected legend and also corrected the figure legend.

- ***"Fig. 4 appears to be cut off on the bottom"***

Again, this was an oversight and you thankfully spotted this. A small number of clades was indeed cut off at the bottom of Figure 4 in the production of the PDF file. A complete Figure 4 is now attached.

Competing Interests: No competing interests were disclosed.